Chapter 2
Survey of Pen-and-Paper Computing

Over several decades, a large body of research has been established that focuses on Pen-and-Paper Computing. This chapter reviews previous work of the field – both from a technological and interface perspective – and discusses future directions of research and development.

2.1 Technologies

Pen-and-Paper Interfaces require that the computer system be able to capture how a user is interacting with physical documents in the real world. This capturing should be robust and perform in real-time without requiring a complicated technological setup or adding interactional overhead. Ideally, paper documents and interactions on paper should be tracked without the need to modify them in any form. In this section, we first discuss technologies for realizing input to a computer system: this includes capturing contents of paper documents, identifying documents and tracking their locations. We further present technologies that allow for capturing input that users make on paper using their hands, fingers or pens. Finally, we briefly discuss how computer output can be provided on paper, by using projection or by augmenting paper with electronic components.

2.1.1 Digitizing Contents of Paper Documents

Even though increasingly more documents are available in an electronic form, there are still documents that only exist on paper. Visual scanners allow us to digitize their contents. A scanner captures a static image of a page's contents at a given point in time. If an image contains text, optical character recognition (OCR) [119] can be used to convert the graphical marks into a machine-readable symbolic representation.

J. Steimle, *Pen-and-Paper User Interfaces*, Human-Computer Interaction Series,
DOI 10.1007/978-3-642-20276-6_2, © Springer-Verlag Berlin Heidelberg 2012

Desktop scanners are well-suited for scanning large numbers of pages and offer a high resolution. Modern desktop scanners feature an automatic document feeder and scan up to 35–60 pages per minute with a resolution of 600 dots per inch or more.

In contrast to desktop scanners, handheld scanners are small and light, so they can be used in mobile settings. A first class of handheld scanners has a scanning unit integrated into the tip of a pen-like device. The interaction for scanning resembles to using a pen on a paper document. By moving the pen along the lines of text, the document is successively scanned. A computing unit built into the pen integrates the small fragments to one single image of the page. This approach is well-suited for scanning individual words or short passages but is too slow for scanning entire pages. An example is the Wizcom InfoScan2 Elite pen[1]. The pen is able to scan with a speed of about 15 cm/s and a resolution of 400 dpi. It features optical character recognition. This enables the pen to read the scanned text aloud the scanned text using voice synthetization, to translate scanned text and to provide definitions for scanned words. Users can transfer scanned data to a computer or a mobile phone via an infrared port or via USB. A second class of handheld scanners is not used like a pen, but is placed flat on the document, very much like a ruler. The scanning module therefore has a larger width, which significantly speeds up the scanning processes. This enables scanning an entire page in 4 to 8 seconds. An example is the Planon DocuPen RC 805[2] (Fig. 2.1).

Fig. 2.1 Planon DocuPen RC 805 handheld scanner (photo copyright Planon Ltd.)

[1] http://www.wizcomtech.com (all references to web pages contained within this book were retrieved on 2011-10-10)

[2] http://www.planon.com

The advantages of desktop and handheld scanners are that they can scan all types of paper documents without additional provisions made to the documents. However, only one static picture is taken at a given point in time. So, it is not possible to continuously track changes made to the document. Moreover, all visual contents are digitized in one single layer. For example, it is therefore a complex task to separate the underlying printed document from handwritings made on it, such as annotations and sketches. Hence, scanning is not optimal for interactive applications.

Another approach supports more interactive uses. The contents of documents are captured by one or several cameras which are mounted above the table or in front of an interactive wall. While contents can be continuously tracked, camera capturing provides lower resolutions than most desktop and handheld scanners. We will discuss this approach in more detail below.

2.1.2 Page Identification and Location Tracking

Many settings that integrate paper with computing require that a physical sheet of paper can be uniquely identified. For instance, a physical paper card could be used as a physical token to access a specific digital object. Alternatively, the system could help the user to find a paper document in the office by indicating the physical location of the document. In both cases, the system must identify the document. Two main approaches can be distinguished: marker-based and content-based identification. Marker-based approaches require that the objects that are to be tracked contain a machine-readable tag. This tag can be visible to the human eye, such as a printed barcode, or invisibly integrated into the object, such as an electronic RFID tag. Content-based techniques do not interfere with the visible artwork, but result in lower processing speed and can distinguish between only a smaller number of objects than marker-based techniques.

Visual Markers

Visual markers (also called fiducials) encode an identifier in an optical machine-readable representation. A marker is captured by one or several cameras, by a light sensor or by a laser scanner. The most widespread form of visual markers, contained on almost any product, is the linear barcode. It encodes a binary sequence by varying the width of black bars that are arranged in a linear sequence. The EAN 13 coding scheme [44] is used worldwide to identify products at cashpoints. Figure 2.2 (a) shows an example of this barcode.

Linear barcodes are not used by many Pen-and-Paper Interfaces. More common are two-dimensional fiducials, since they allow for storing more data and also for tracking the location of objects in 3d space. Several types of two-dimensional fiducials can be distinguished:

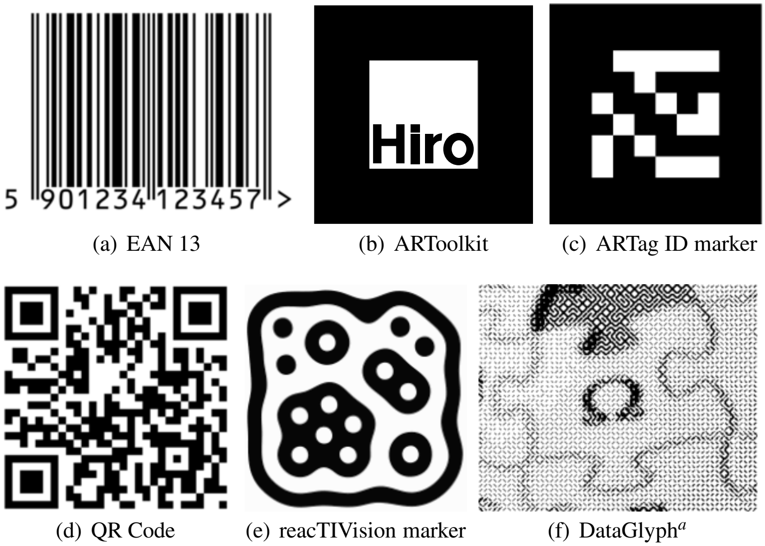

Fig. 2.2 Linear and two-dimensional barcodes ([a]DataGlyph example by Jeff Breidenbach)

Pattern-based fiducials are used for instance in the widely-used open source marker tracking library ARToolkit[3]. The fiducial has a black frame which encloses a black and white pattern (Fig. 2.2 (b)). The library identifies a barcode by comparing the pattern with a set of pre-registered templates using pattern matching techniques from computer vision. The advantage of pattern-based fiducials is that the application developer has some influence on how the fiducial looks like and can create meaningful markers. For instance, it can contain a symbol or some text. However, identification is more error-prone than with other techniques.

Matrix-based fiducials are also known as 2D barcodes. They encode binary data by a two-dimensional grid of black and white points. Each point encodes one bit of data, whereby some bits are usually reserved for error correction. Figure 2.2 (c) depicts an ID marker of ARTag [28]. 2D barcodes allow not only to encode an identifier but a relatively large amount of data. For instance, one single QR code [46] (see Fig. 2.2 d) can store more than 4,000 alphanumeric characters. The DataMatrix [45] code can store more than 2,000 characters. It is for example used for electronic stamps by the German postal service. The data density of 2D barcodes can be further increased by varying visual properties of the points, such as color or brightness, but at the price of a decrease in robustness.

In addition to *identifying* objects, matrix-based fiducials can also be used for tracking the *location* of objects. If the size and shape of the marker is known, it is possible to calculate the marker's relative position and orientation with full 6

[3] http://www.hitl.washington.edu/artoolkit/

degrees-of-freedom information from a 2D camera image. This principle is widely used in augmented reality and tangible interaction systems. Several toolkits offer out-of-the-box support for application developers. ARTag [28] supports up to 1024 barcode markers which have been optimized for fast and reliable detection. ARToolkitPlus [162] is an improved version of the original ARToolkit[4] that is inspired by ARTag's approach. It offers binary markers that can be more robustly detected than the pattern-based markers of the original ARToolkit.

Topological region adjacency is leveraged by a third class of fiducials. The basic idea of this approach is, instead of encoding a sequence of bits as a sequence of black and white points, to encode a hierarchical graph. The original bit sequence is decoded by identifying and then traversing this graph. The advantage of this approach is that it is fast and very robust against false positive detection [104]. However, only a small number of identifiers can be encoded. reacTIVision [53] is an open-source toolkit that uses this technique. Figure 2.2 (e) shows a reacTIVision fiducial. reacTIVision is used in many projects that require tracking the location and orientation of tangibles on interactive tabletops. While reacTIVision tracks the 2D position and orientation of objects on a flat surface, it cannot provide full 6 degrees-of-freedom information. Recent research showed how to obtain 6 degrees-of-freedom information using topological region adjacency markers [104].

The main advantage of using fiducial markers for identifying and tracking objects is that this is a relatively inexpensive tracking solution. However, this approach requires that the fiducial is in the line-of-sight of the camera. This requirement can severely restrict natural interactions, for instance, when objects are piled. Moreover, the fiducial interferes with the artwork of the object.

Content-embedded fiducials are visually less obtrusive. DataGlyphs [52] (Fig. 2.2 (e)) encodes binary data with a pattern of forward and backward slashes. They are flexible in size, shape and color. This makes it possible to emulate the look of a grayscale or even of a color image by an appropriate pattern of small dashes, similar to how offset printing emulates images by small raster dots. At 600 dpi printing resolution, DataGlyphs can encode up to 1,000 bytes of data per square inch. Another technology, Anoto digital pen and paper [4], performs tracking by decoding a pattern of tiny points that is printed onto paper documents and hardly visible to the human eye. The Anoto approach is presented in more detail in Section 2.1.4 below.

Some applications require real-time location tracking in three dimensions, even in cases when objects are moving very fast. For instance this is important in augmented reality applications that overlay physical objects with projected digital contents. Optical motion capture systems (e.g. Vicon[5], OptiTrack[6]) use several high-speed infrared cameras that observe a scene from different perspectives. Several small retro-reflective dots are attached to the object which is to be captured. These dots appear as white blobs in the camera images. If a dot is seen by at least two

[4] http://www.hitl.washington.edu/artoolkit/

[5] http://www.vicon.com

[6] http://www.naturalpoint.com/optitrack

Fig. 2.3 An infrared camera of the OptiTrack high-speed motion capture system

cameras, its 3D position can be calculated. If several dots form a known spatial arrangement, the object's 3D pose can also be identified. Figure 2.3 shows an OptiTrack camera.

Electronic Markers

Electronic markers avoid some of the major problems of visual markers. Radio-Frequency Identification (RFID) uses tags that are embedded into or applied to a physical object. The tag consists of an integrated circuit that stores a unique ID and manages the communication with an external reading device. Moreover, it includes an antenna for receiving and transmitting signals. If the tag is passive, it does not include an own battery but receives energy from the reading device via an electromagnetic field. When the RFID tag is within the range of a reading device, it transmits its unique ID.

No direct line of sight between the tag and the reading device is necessary. However, in contrast to fiducials, RFID tags produce significant costs. To date, an RFID tag still costs more than 0.1 USD. Depending on how paper is used in a Pen-and-Paper Application, these costs might be prohibitive. Moreover, in contrast to fiducials that can be printed directly with the paper document, an additional processing step is necessary for applying RFID tags to paper. Finally, while the technology enables to track whether a tag is within the range of a reader or not, it does not identify its precise location. A novel technology called Near Field Communication (NFC) enables mobile phones to act both as RFID tag and as RFID reading device. NFC is already available in several mobile phones. It is likely that RFID will gain more prominence in the near future and will supersede visual markers in some domains.

A range of electromagnetic solutions (e.g. Ubisense[7]) provide for capturing the physical location of objects. To date, however, they require large markers that cannot be used when working with thin sheets of paper.

Some research projects have developed individual solutions for identifying objects. Similarly to RFID, integrated circuits storing a unique identifier are applied to physical objects. The communication to a reading device is made via a wired connection (e.g. [130, 49]) or a wireless connection (e.g. [114]).

Content-based Identification and Tracking

Content-based approaches do not require markers. Instead, documents are identified and tracked solely by their visual appearance. A camera is observing the scene or documents are scanned, and image processing techniques are used for analyzing the images.

The most commonly used approach is SIFT features [83]. In order to identify a document page, it first has to be registered. The system captures specific optical features of the image and stores these features as a fingerprint of the document page. When a document page appears in the camera image, its features are captured and compared against the fingerprints in the database to identify the page. This technique performs quite well if the camera provides a good resolution and the number of different pages is not very high. For instance, Kim et al. [60] reached a recognition rate of more than 90 % if the document width in the camera image was at least 300 pixels. Liu et al. [79] present an improved set of features that was inspired by the SIFT feature set. They achieved recognition rates of more than 99 %, even if pages had to be identified from a very large set of pages. However, their evaluation excluded lighting effects and camera noise, so it must be assumed that in real setups performance will be lower.

Another approach uses optical character recognition of document text [24]. The recognized text is used for querying a database of documents and retrieving the digital version. As an alternative to using text recognition, the white spaces between words have proven to be quite unique for each text. Brick Wall Coding [26] detects the white gaps between words and takes them as features. It has lower recognition performance than SIFT or FIT features, but it can detect a document page even when the page is only partially visible to the camera.

A drawback of content-based approaches is that the visual appearance must contain enough information for unambiguous identification. For instance, it is not possible to uniquely identify one of several copies of the same document page. Moreover, content-based identification is challenging if the user modifies documents, for instance by making annotations.

[7] http://www.ubisense.net

2.1.3 Capturing Touch Input

A further important field for bridging paper and computers is to track touch input on physical sheets of paper. The most frequent approach for detecting touch input on physical paper consists of observing the scene with an overhead camera. Touch events are detected in the camera image using image processing techniques. An influential early example is Wellner's DigitalDesk [167]. The DigitalDesk recognizes the position of fingers by their shape in the video frames. A microphone attached to the bottom of the desk is used to detect the exact moment when the user's finger taps on the desk. However this does not allow the system to detect multiple touches that occur simultaneously. Another technique can detect multiple simultaneous touches. It tracks the color of the fingertips. When the color changes, it is assumed that the finger is pressed against a surface [93]. Further approaches analyze shadows cast by the fingers [171] or use depth information for detecting whether a finger is hovering over or is touching the surface. Wilson [172] recently presented a technique that uses a depth camera to detect touches not only on flat, but also on curved surfaces. Figure 2.4 shows how the approach works: The scene (left) is observed by a depth camera. Initially a model of the surface is set up that stores the depth of the surface for each pixel while no hands or fingers are visible in the camera image. The current input of the depth camera (center picture) is then compared against this model. Hands and fingers have lower depth values because they are situated above the surface. If the distance between the value stored in the model and the currently recorded value falls below a certain threshold, a contact is detected (right).

A drawback of camera-based capturing is that it significantly restricts the mobility of paper. Typically, the camera is mounted at a fix position. While there exist mobile solutions, such as the Docklamp [55] or Sixth Sense [96], these are still rather large and heavy and their use is not comparable to the flexibility of traditional pen and paper. This restriction can be alleviated by using multiple cameras that capture a large volume [173].

Other approaches for detecting touch use electromagnetic field sensing [6] or embed electronics into paper, e.g. push buttons [114, 9].

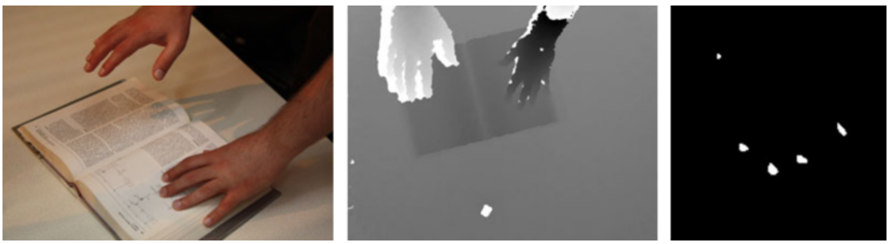

Fig. 2.4 Detecting touch using a depth camera. Left: Touching a book. Center: Depth image. Right: Detected contacts (photos courtesy of Andy Wilson)

2.1.4 Capturing Pen Input

Pen-input is an important style of interaction with paper documents. Users take notes, make annotations and sketches, or interact with printed user interfaces, e.g. by tapping with a pen on printed buttons. There is a wide range of technologies for capturing pen input. Technology for capturing pen input on real paper should offer high tracking performance while restricting the natural interaction as little as possible. A first class of approaches tracks the relative position of a pen with respect to a separate tracking device, e.g. a camera. A second class of capturing approaches is able to directly capture the absolute position on a page.

All approaches have in common to generate not only a two-dimensional representation of the pen traces (as does a photo or a scan of the document). In addition, they record temporal information of how the traces are made over time. Some technologies also track the force with which the pen tip is pressed onto the paper sheet. At regular intervals, e.g. 50 times per second, a so-called sample is captured. This sample contains the current position of the pen, the current time and optionally the pen tip force. A set of samples is called digital ink data. By interpolating curves through the sample coordinates, the pen traces can be visualized.

Camera-based Capturing

Above we have discussed how camera-based capturing can be used for detecting touch input. Similar approaches have been presented for tracking pen input. The DigitalDesk [167] analyzes images of the desk that an overhead camera is continuously capturing. It detects not the pen itself, but the traces that are made with a pen. While this allows the system to record traces, it is not possible to detect interactions that leave no visible ink traces, such as tapping with the pen.

Other approaches aim at tracking the pen itself, not only the visible traces the user has made with it. Typically one or several visual markers are attached to the pen. This allows the system to robustly identify the pen in the camera image. These markers can be active markers that emit light signals [70] or passive markers that reflect light of a given wavelength [40].

Ultrasonic Pens

A second technique relies on ultrasonic triangulation. The digital pen continuously emits an ultrasonic signal which is not audible by humans. An additional separate tracking device is attached to the paper sheet(s). It has two or more reference points that capture the ultrasonic signal. By calculating the time difference with which the signal reaches the reference points, the position of the pen with respect to these reference points can be calculated.

The temporal and spatial resolution of this technology is high enough for capturing handwriting. The tracking does not depend on the material the pen is used upon and it scales to large surfaces. So it can be used on arbitrary flat surfaces, such as tables, augmented walls or whiteboards. However, the position of the pen is tracked in relation to the external device and not in relation to the paper sheet. Moreover, this approach is not able to detect which sheet the user is writing on. This makes this approach hard to use in settings where users do not write on one single page but deal with many pages. Ultrasonic tracking is utilized in commercial solutions that target private end-users, for example the Pegasus Tablet NoteTaker[8]. It has a resolution of 100 dpi.

Digitizing Tablets

Digitizing tablets are devices that capture pen input for computer applications. The most common technology is patented by Wacom. A Wacom device has a flat surface and generates a magnetic field. Using induction, the position of a specific stylus can be detected on this surface. Figure 2.5 shows a Wacom Intuos4 graphics tablet[9]. In principle, digitizing tablets do not aim at supporting the use of a pen on paper. Instead, the tablet is used to directly interact with the digital system, e.g. for drawing in a graphics application or for positioning the mouse pointer. However, the induction principle still works if one or several sheets of paper are positioned between the graphics tablet and the stylus. For this reason, they can be used to track pen input on real paper. This approach enables very high resolutions (about 1000 to 5000 dpi). However, when used for capturing pen input on paper, it has the same drawbacks as ultrasonic tracking. The user must manually calibrate the paper sheet and must indicate page changes. Moreover, the interaction is restricted to the small surface of the tablet. Therefore this approach is typically used only in research prototypes (e.g. [88, 27, 150]).

Inductive Pens

The following techniques do not detect the position of the pen with respect to an external reading device, but they encode positional information directly on paper. Thereby local page coordinates can be detected without requiring the system to have knowledge about the absolute position of the sheet of paper. It is the digital pen that decodes its position, which makes external devices and calibration obsolete. As the position which is encoded on the physical sheets can also contain a page identifier, the pen is able to detect on which page it is used. Hence, it is not necessary to manually indicate on which page the pen is currently used. Users can therefore work very naturally with multiple sheets of paper.

Within the frame of the European Paper++ project, researchers have developed an approach that leverages conductive ink [134, p. 26]. Positional information is

[8] http://www.pegatech.com

[9] http://www.wacom.com

Fig. 2.5 Wacom Intuos 4 digitizing tablet (photo copyright Wacom)

encoded directly on paper by printing a grid of linear barcodes with conductive ink on paper documents. This ink is barely visible to the human eye. A digital pen measures the inductivity and decodes from the barcodes its two-dimensional position. The resolution of this technique is not high enough to capture handwritings but it allows for distinguishing larger hot-spot areas on a document.

Anoto Technology

Anoto digital pen and paper [4] is the currently most mature solution for capturing pen input on paper. An Anoto digital pen behaves like an ordinary ballpoint pen and leaves visible ink traces. Moreover, it has a built-in infrared camera and a processing unit to detect a specific pattern on the print products. By analyzing this pattern, the pen can decode its position and electronically capture all pen traces made. This is done at a frame rate of approximately 75 Hz and with a spatial resolution of about 850 dpi. In addition to the position in the two-dimensional coordinate space, current pens register pen tip force and timestamps. The technology further allows for detecting the rotary and tilt angles of the pen although this is not supported by the firmware of current pens. Anoto pens include a battery and can therefore be used in mobile conditions. An Anoto pen is depicted in Fig. 1.2 on p. 12.

Depending on the capabilities of the pen, data is temporally buffered on the pen until it is synchronized with a computer via USB or Bluetooth. Alternatively the data is streamed in real-time to a nearby computing device using a Bluetooth connection.

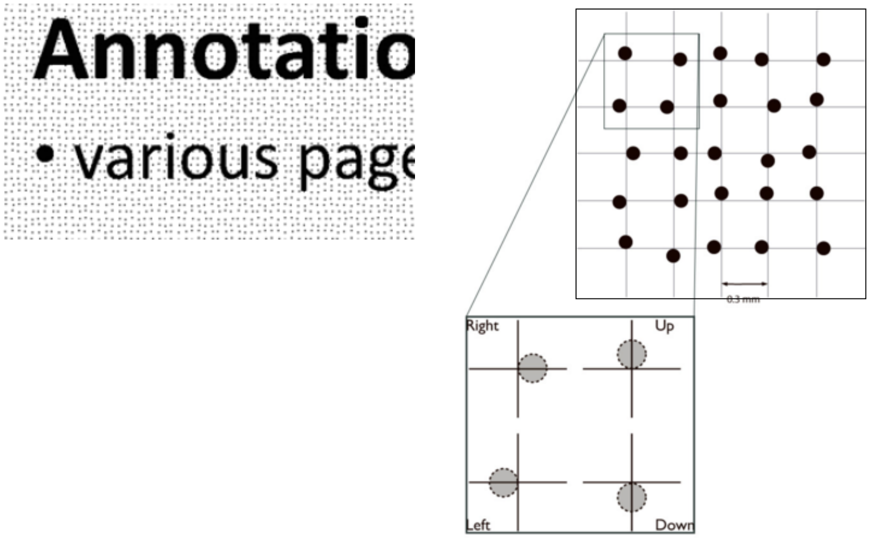

Fig. 2.6 The Anoto pattern. Left: Enlarged view of a document that contains the pattern. Right: Highly enlarged schematic view of a 6x6 grid of points of the Anoto pattern (illustration copyright Anoto)

Unfortunately, it is not end-users who can decide in a given situation whether the pen should temporally store the data or stream to a nearby device. Instead the designer of the application has to select beforehand which of the modes to use. This is necessary because two different types of Anoto pattern, which have to be licensed separately, enable either batch processing or streaming, but not both of them.[10]

Anoto Pattern The positional information is encoded directly on the paper sheets with a patented two-dimensional dot pattern (Fig. 2.6). The pattern is only slightly visible to the human eye, making the print product appear slightly gray. Each dot has a position on an imaginary grid which overlays the page. By slightly displacing each dot in one of the four directions, a single dot encodes 2 bit of data. Each 6x6 matrix of dots encodes a unique position in the Anoto coordinate space. This gives not only the position on a page, but also allows for distinguishing between different pages.

The Anoto pattern can be printed on paper products with various printing techniques, including offset, laser and inkjet printing. Printing is unproblematic if the dot pattern is to be printed on empty sheets of paper that do not contain any other printed contents. In this case, most desktop laser printers can be used, and also many inkjet printers yield acceptable results. Things get more complicated if the page contains printed contents on which the Anoto pattern is overlaid. The problem here is

[10] To our knowledge the only exception is the Nokia SU-1B pen, now discontinued, for which a firmware patch exists that enables switching between batch processing and streaming on non-streaming pattern. This pen is used in many research prototypes.

that the camera might not see the pattern dots at positions where other content is printed. To cope with this problem, the Anoto technology relies on a smart solution that leverages an optical characteristic of toner and ink. Some toners absorb infrared light and thus appear in the image of the infrared camera in black. These toners are used for printing the dot pattern. Other toners do not absorb infrared light. While these toners are visible to the human eye, they are not visible in the camera image. These toners can be used for printing contents other than the dot pattern. These contents do not interfere with the dot pattern in the camera image. The printer requires toner (or ink) that has these characteristics. It happens that the black toner (K) used in many laserjet printers typically absorbs infrared ink, while cyan (C), magenta (M) and yellow (Y) toners do not. This allows printing the dot pattern with K toner, while other contents are printed by using only C, M and Y colors. The Anoto pattern and other contents can be printed in two separate steps or in one single step. Anoto has tested printers of diverse manufacturers and has published a list of Anoto-certified printers. If the pattern is to be applied to larger paper products, it can be printed with some selected inkjet plotter models.[11]

Pen Models Anoto pens are produced by several manufacturers. The Anoto/Maxell DP-201 and the Logitech/Destiny io2 support USB batch processing and Bluetooth streaming. The Anoto ADP-301 supports only Bluetooth streaming, but features a lower latency than the DP-201. This makes it the product of choice for applications with hard real-time constraints on input. Finally, the Anoto ADP-501 features a thicker pen tip instead of the thin ballpoint tip. This provides for use on flipcharts and whiteboards. Several older pen models are discontinued, most notably Logitech io and io2, Nokia SU-1B and Nokia SU-27W.

These standard models suffer of several shortcomings that will be discussed in the following: pen feedback, erasing of contents and use as a stylus on displays.

Pen Feedback The standard models do only capture and digitize pen traces to transfer them to a computer. They do not interpret them and therefore cannot provide system-specific feedback, even though all pens feature LEDs and some of them can vibrate. Moreover, the Bluetooth connection does not offer any back-channel to control the pen, so applications cannot generate feedback given by the pen. Liao et al. [76] extended one of these pen models with auditory, tactile and visual feedback and explored with that prototype several pen-based feedback mechanisms. Recently, several new pen models have been commercialized that include additional processing capabilities and output devices. These pens are able to interpret pen traces and react to user input. The Fly Fusion Pentop Computer[12] has a built-in speaker. It reads out the translations of handwritten words, can be used as a calculator and plays MP3 files. The more recent Fly Tag Reading System aims at helping kids learn to read. It features a digital pen that reads out the passages of a book the user points to. The Livescribe Echo Smartpen[13] (Fig. 2.7) has a built-in microphone, speaker and an

[11] For instance the plotters of the Canon imagePROGRAF series (iPF8000, iPF8300, iPF9000, iPF9100) provide quite good results on diverse types of paper and foils.

[12] http://www.flyworld.com

[13] http://www.livescribe.com

Fig. 2.7 Livescribe Echo SmartPen (photo copyright Livescribe)

OLED display. In addition to capturing notes, it can be used to record and playback audio and to perform calculations. An SDK provides for developing further applications (so-called penlets) which are executed directly on the pen. While these novel pens enable to develop more interactive interfaces including direct system feedback on the pen, they do not provide a wireless Bluetooth connection.

Erasing Pens A further shortcoming of current Anoto pens is that they do not allow users to erase physical pen traces. Olberding and Steimle [111] have demonstrated an easy and inexpensive method to add erasing capabilities to Anoto pens. Anoto plans to commercially deploy an eraser pen (ADE-501) in the near future that can be used to erase contents on whiteboards, but not on paper documents.[14]

Pen Input on Displays All these solutions aim at supporting pen input on paper. However, Anoto pens can also be used as a stylus for providing input on screens. Brandl et al. [10] presented an approach that enables using Anoto pens on rear-projection multi-touch screens. This type of screen is typically used in interactive tabletop and wall displays. Their approach relies on printing the Anoto pattern on a translucent foil. Like on paper, the pen decodes its position from the printed pattern. This position can be converted to screen coordinates. More recently, Hofer and Kunz [38] introduced a method that allows using Anoto pens also on LCD displays. Anoto is announcing their own solution for input on displays to be released soon. Theses approaches enable to produce very large pen-enabled displays that have a high input resolution at relatively low cost. Moreover, in contrast to competing technologies for pen input on displays, multiple pens can be simultaneously used.

[14] http://www.anoto.com/up-and-coming-products.aspx

Data-embedding Pens

A recent direction of research examines data-embedding pens [81]. In addition to leaving visible ink traces, a data-embedding pen visually encodes digital data onto the paper document. Liwicki et al. present a prototype device that features, in addition to an ordinary ink refill, a tiny ink-jet head. This ink-jet head can print a sequence of dots in a bright color close to the handwritten traces. This dotted line encodes a sequence of binary data, which can be subsequently decoded by using optical image processing. A data-embedding pen can be used for instance for encoding metadata of the handwriting, such as the time of writing, the ID of the author or the geo-coordinates of the current location. Alternatively, the pen can encode the temporal sequence in which the pen traces have been written. Such information about the writing process can improve recognition of scanned handwriting.

2.1.5 Digital Output on Paper

We have discussed a variety of approaches that allow systems to capture paper contents and input made on paper. To conclude the section on technologies, we will briefly discuss how systems can provide *output* directly on paper. While a straightforward way consists of (re-)printing a document, this is not very interactive, since each update requires a new printout. We will have a look at technologies that enable more dynamic output on paper.

Projection

The projection approach uses paper as a passive display onto which digital contents are projected. The position and orientation of the paper surface is tracked in real-time. The system then projects the display contents onto this surface. This enables not only to simulate screens, but also to overlay physical objects with additional digital information. If the performance of the tracking system is sufficently high and the projected image is correctly adjusted to fit the physical surface, the paper surface can be moved around freely and behaves similar as if it was an active display. However, even in state-of-the-art solutions, the projected image is slightly lagging behind the paper surface if the surface is moved quickly.

This approach is relatively straightforward if paper is used only on a flat surface. With the influential Augmented Surfaces system [123], Rekimoto and Saitoh demonstrated that a single camera and a standard projector are sufficient for tracking paper on a 2d surface. If the user is allowed to freely move and rotate paper surfaces in three dimensions, things get more complicated. In this case, the surface is not necessarily perpendicular to the orientation of the projector. So the projected image has to be distorted such that the projected image appears undistorted to an observer. Early examples are Dynamic Shader Lamps [7] and PaperWindows [40]. Most solutions use commercially available optical motion capture systems, such as Vicon or NaturalPoint OptiTrack. A set of several high-speed cameras detects the 3D positions of retro-reflective makers that are attached to the paper surface. This allows the system to identify the position, orientation, and possibly deformation of the surface. Recently Lee et al. [70] have demonstrated that even a low-cost tracking solution, using the PixArt camera within the Nintendo Wii Remote, can yield acceptable results for 3D tracking. The Microsoft Surface team presented a different approach for projecting contents onto passive displays [47]. Instead of top-projection, they use a sophisticated setup for projecting contents onto semi-transparent paper surfaces from the backside. For this purpose, a projector is integrated into an interactive tabletop display. The tabletop surface consists of an electronically switchable diffuser. This can alternate with very high frequency between a diffuse and a transparent state. During the diffuse time slots, imagery is projected onto the tabletop surface. During the transparent time slots, imagery is projected through the tabletop surface onto the paper displays.

Electronically Augmented Paper

Output on paper can be realized by attaching or embedding electronic components into paper. This enables a greater diversity of interactions than with passive paper, but electronically augmented paper documents are more complicated and more expensive to produce.

An example is *Pulp-based computing* [17]. The authors embed diverse components directly into paper during the papermaking process. These components comprise microphones, bend sensors, LEDs, speakers and vibrating motors. *Voodoo-Sketch* [9] is an ad-hoc physical interface toolkit. It allows plugging various electronic controls, such as push buttons, switches and sliders, onto specific paper palettes, which have embedded conductive layers. The *Computational Sketchbook* [11] follows a similar approach. The user can create interactive paintings by attaching electronic components onto ordinary paper. These components comprise speakers, motors, switches, LEDs and batteries. The user establishes the conductive connections by painting them with conductive ink. The electric circuit thus becomes a visible part of the artwork.

Electronic Paper

Upcoming novel display technologies have a large potential to significantly alter the way we are interacting with paper and with displays. Rapid advances in OLED and Electronic Paper technologies allow companies to develop displays that have similar characteristics as paper. Current commercialized displays, e.g. displays of recent e-book readers, are still rigid and much thicker than traditional paper. However, various research prototypes show promising advances. Japanese researchers have developed an electronic paper display which is as thin as 1 mm; Sony has developed an OLED TV of only 2 mm thickness [54]. Moreover, Sony [141] and Polymer Vision [120] have presented displays that can be rolled. Recent research has even used paper as the substrate of an electronic paper display. This is much thinner than traditional electronic paper and more flexible and soft [54]. Such displays will allow designers to develop applications that leverage the affordances of paper while offering full digital feedback comparable to a screen.

2.1.6 Pen-and-Paper Toolkits

As Anoto pens are the currently most mature and most widely used solution for realizing Pen-and-Paper Applications, we present some toolkits that ease developing applications featuring the Anoto technology. All of these toolkits offer support for establishing the low-level connection between the pen and a computing device and for accessing the pen data. Some toolkits additionally provide higher-level support for interpreting pen data.

Anoto SDK

Anoto offers commercial SDKs for developing Pen-and-Paper Applications with Anoto pens. Only little information about the SDKs is made publicly available. For guiding application developers, we outline core concepts of the Anoto SDKs.

Developers have to license some Anoto pattern space that can be used within the application. The entire Anoto pattern space is large enough to cover a surface equivalent to that of Europe and Asia combined. A license consists of a subregion of that space. For instance, a “book” license is equivalent to 256 letter-size or A4-size pages that each can be uniquely identified.

The Anoto pen delivers raw coordinates which describe a unique location in the entire pattern space. In typical applications, these huge coordinates are quite cumbersome to handle. One is rather interested in coordinates that are relative to a given document or a given page. The Anoto SDKs provide such higher-level abstractions of the raw pen coordinates. For instance, it is easily possible to retrieve all traces that were made on a specific page. The SDKs further allow application developers to define interactive regions on document pages. For instance, a developer might want to use a paper form in the application. This form might have some preprinted areas into which the user can indicate some information, e.g. for the user’s name and address. In addition it might contain some boxes that the user can optionally check, e.g. indicating if she wants to receive further information. All these areas can be defined as interactive regions. On a technical level these are rectangular areas with an associated identifier.

To define such interactive regions on paper products, Anoto offers the Form Design Toolkit, a plug-in for Adobe Acrobat. This allows for defining interactive regions and adding the Anoto pattern to any PDF document. However, the Form

Design Toolkit requires manually executing several commands within Adobe Acrobat. Manually generating these documents is not problematic if an application uses a limited number of paper documents that are identical for each user. However if users need personalized documents, possibly even creating and printing them during runtime, the manual approach is inadequate. For such cases, Anoto offers the Paper SDK. This SDK enables applications to add the Anoto pattern to documents during runtime.

For developing applications, Anoto offers an SDK for PC applications (including Java and for native Windows applications) and an SDK for network applications. Both SDKs support only non-streaming applications, offering classes and methods for accessing and parsing pen data that has been transferred via batch synchronization. The SDKs allow the developer to directly retrieve pen traces that were made on a specific interactive region. Furthermore, the SDKs offer basic additional functionality, such as for rendering and for saving pen data.

Access to the streaming functionality of the ADP-301 is supported by a further SDK, the Streaming Pen Connectivity Driver and its accompanying API. This provides basic methods for accessing the pen data arriving through the Bluetooth connection. While it gives access to the linear data stream, to the best of our knowledge it does not provide the abstractions offered by the non-streaming SDKs described above.

PaperToolkit

The PaperToolkit [178] is an open source toolkit developed by Ron Yeh at Stanford University. It builds on top of the Anoto SDK. PaperToolkit aims at an interactive use of digital pens in applications and provides more support and higher-level abstractions of digital pen data than the Anoto SDK. It is based on an event-based architecture, similar to frameworks for Graphical User Interfaces toolkits like Java Swing or Windows Forms. The application developer can compose a paper interface by composing predefined interface widgets, such as drawing areas, buttons and check boxes. The toolkit renders these widgets on top of the Anoto pattern to a PDF document and allows for printing it on paper.

For applications the framework provides event handlers that react to specific types of pen input. Event handlers can be added to paper interface widgets and trigger events on the software side. For instance, the application can register event handlers that are invoked when any pen activity has occurred within a specific widget, when a button widget has been tapped on or when a check box has been selected.

PaperToolkit unifies real-time and batched event handling. It is also possible to use several pens simultaneously in the same application. The toolkit contains plenty of further components that ease developing interactive Pen-and-Paper Applications. This includes a component that allows developers to define a fully functional paper interface simply by sketching it on paper, diverse components for rendering digital pen data as well as a tool for simulating pen data for debugging purposes.

The toolkit is developed in Java. We have used it in our own projects and have experienced that it is very stable. However, it is not under active development any more. While older pen models (Logitech io2, Maxell DP-201, Nokia SU-1B and SU-27W) are supported, more recent pen models cannot be used with the toolkit unless extra code is written to access these pens.

iPaper

The iPaper framework [134] was developed by Beat Signer and his colleagues at ETH Zurich. It supports the development of interactive paper products by combining an authoring perspective with a service approach. The underlying concept is to decouple visual design, interaction design and development of the services. The graphic designer authors the document. A service developer implements a service that can be used with an interactive paper document. Finally, the interaction designer ties both the document and the services together by defining links between regions of the paper document and services.

iPaper is based on iServer, a generic framework that allows the definition of links between arbitary physical or digital information entities. For instance, links can be defined between any combination of images, videos, digital documents, paper sheets, RFID tags and many other types of resources. Each link can have multiple source and multiple target resources. The iPaper plug-in for iServer provides more specific support for Pen-and-Paper Interfaces. iPaper links are established between an active region on paper and an active component, which is the service associated with this region. In contrast to the Anoto SDK, active regions do not have to be rectangular. Other shapes, such as circles and polygons are supported. When the pen is used on an active region, the framework automatically executes the code of the active component that is linked to this region. While a number of predefined active components are available, the application developer can easily add own active components.

iPaper has a distributed architecture. For instance, an active component can be deployed on a server. Moreover, it includes a powerful printing component. This component allows printing documents that contain the Anoto pattern directly from within the end application, similar to the Anoto Paper SDK. It supports most of the current pen models (including Anoto ADP 201, Logitech io2), but not the most recent Anoto ADP-301.

Letras

Letras [36] is an open source toolkit that provides support for mobile Pen-and-Paper Interfaces in ubiquitous computing settings. Letras enables to use the same pen in a variety of different contexts, in combination with various computing devices. Perhaps most interesting is that it allows users to couple Anoto pens with Android phones and tablets. This provides for using interactive Pen-and-Paper Applications nearly everywhere. In this case, pen data is streamed to a personal mobile device that acts as an information hub and decides how to distribute pen data further to other computing devices over a network connection. Letras also supports different configurations; for instance the pen can be coupled with a PC.

Letras has a modular architecture. It relies on a distributed processing pipeline for pen data. Different processing stages are decoupled using generic interfaces. The first processing stage, the driver stage, connects to one or several Anoto pens. This includes simultaneous connection to different pen models. A next processing stage provides for abstracting from raw pen coordinates to higher-level interactive areas. This is similar to the interactive areas of the Anoto SDK, but with a distributed lookup of area definitions across multiple issuing organizations, allowing to distribute interactive paper material such as leaflets. Further processing stages support semantic processing, e.g. handwriting recognition, and application-level processing, e.g. rendering the ink traces. Application developers can develop additional processing stages and flexibly deploy processing stages to different computers and devices. While the first processing stage is typically deployed on a computing device that is used by end-users (e.g. the mobile phone, the tablet or the PC), the application developer can decide where to deploy other stages. For instance, handwriting recognition might be performed on a server. For rendering, the system might detect which computing devices in the room feature a large screen and dynamically deploy the rendering stage on one of these devices. The individual stages are interconnected with the MundoCore middleware [3]. This supports the dynamic configuration of the computing environment. A drawback of Letras is that it currently does not provide a print toolkit. For adding the Anoto pattern to documents, the developer must use one of the Anoto SDKs.

Letras is developed in Java and contains some native components for Mac OS X, Windows and Linux. It provides native support for streaming pen data to Android devices. It currently supports Anoto ADP-301, Anoto DP-201, Logitech/Destiny io2 and Nokia SU-1B.

Livescribe SDK

Livescribe offers an SDK for their Livescribe pens.[15] It can be downloaded and used free of charge. It follows a different approach than the frameworks presented above. In contrast to standard Anoto pens which do only capture pen data and transfer it to a computer, a Livescribe pen has a processing unit that can execute custom code. Livescribe applications consist of a paper product and a penlet. A penlet is a piece of software that is executed directly on the pen. This enables to develop applications that react on user input in real-time without requiring a streaming connection to a second computer device which handles the interpretation. Furthermore, a specific Desktop SDK offers support for developing desktop applications that access pen data (this is similar to the standard way of Anoto-based applications).

[15] http://www.livescribe.com

The Livescribe API is based on the Java Platform Micro Edition (Java ME). It comes with Windows and Mac versions of an Eclipse-based IDE and a software for emulating the Livescribe pen.

Livescribe has a management of paper regions that is different from the other toolkits. It distinguishes between fixed print and open paper. Fixed print areas are similar to the interactive regions of the other toolkits. They allow developers to define specific areas for widgets that control the application. In contrast, open regions are blank regions of paper that are not permanently associated with one specific application. Instead each application can claim an open region when the user is writing on it. This enables the flexible use of one single paper notebooks for many different applications.

2.2 Pen-and-Paper Interfaces

A wealth of applications and interaction techniques of Pen-and-Paper Interfaces have been presented in prior work. In this section, we review the large and ever-growing body of research that has been established during the past two decades.

Early, highly influential research was conducted at XeroxPARC and EuroPARC in the early 1990s. A number of seminal systems demonstrated how paper can be closely integrated with digital media on one single interactive tabletop surface. These systems already showed three central functions of augmented paper: 1) paper is used as a token for accessing and controlling a digital resource, 2) paper documents contain embedded hyperlinks to additional digital resources, and 3) the contents of a paper document are automatically synchronized with a digital version of the document. These systems laid the foundation for a wide range of paper-augmented tabletops and walls and also for a number of approaches that allow users to manage digital media on their PC by using paper tokens.

While in this first generation of works, paper was mostly restricted to be used solely on the interactive tabletop or on the user's desk, a subsequent generation focused on paper as a mobile medium. Augmented paper notebooks, form-filling applications and applications for annotating printed documents aimed at retaining the freedom to use paper at many different places. The advent of digital pens which can be used in mobile settings without a complicated technical setup certainly promoted this evolution. In contrast to the early desk systems, most of these mobile systems provide visual output not directly on paper, but on a separate computer screen. However, we will also review approaches that realize visual output directly within paper documents by using mobile projectors or tiny overlaid displays. Support for paper-based collaboration is an issue of more recent interest. This comprises not only co-located collaboration, but in particular how asynchronous remote sharing of paper-based contents can integrate paper with social networks and the Web 2.0.

Our review is structured following which paper media are augmented by digital functionality and following the main functions of augmented paper. Augmented paper cards and post-its clearly show how paper can be used as a token for accessing and controlling digital resources. Augmented books demonstrate paper-digital hyperlinking. Augmented paper notebooks and augmented printed documents present a variety of approaches to synchronize paper-based contents with a digital version. Finally, augmented tables, flipcharts and walls integrate all of these functionalities and provide a very seamless integration of physical and digital media on large surfaces. We will conclude this chapter with a discussion of future directions of research.

2.2.1 Augmented Paper Cards and Post-Its

Augmented paper cards and augmented post-its demonstrate the first main function of augmented paper: using paper as a physical token for accessing and managing digital resources. Each digital resource to access is represented by a physical object. This object has no functionality other than representing the digital resource; it does only contain a description of the resource, but not its actual contents. By manipulating this object (e.g. by holding it in front of a barcode reader or near to an RFID reader or by pressing a push button on that object), the associated digital resource is accessed. This provides an easy and intuitive way for selecting and opening documents or applications. Moreover, the physical tokens can be flexibly structured by arranging them in space and can be physically shared with co-workers.

An example is *WebStickers* [82] which uses post-it stickers that contain a barcode. The user can associate a post-it note with a single Web page or with a collection of Web pages. By holding the post-it under a barcode reader, the associated resource is displayed on a computer screen. *Palette* [101] follows a similar approach and uses barcode-enhanced paper cards for accessing individual slides of slide presentations.

PaperButtons [114] overcomes the limitations of barcodes that must be placed below a barcode reader or in front of a camera phone. It extends the Palette system by replacing the barcode on a paper card with a push button. A unique ID is transmitted to the system over a wireless connection when this button is pressed.

The previous systems allow easy access to the digital resource, but do not support the opposite direction: easily finding the physical token that belongs to a digital resource. To support this opposite direction, *Quickies* [95] augments physical post-it stickers with an RFID tag that is applied to the reverse side. The user can write a handwritten label on a sticker using a digital pen. The label is automatically digitized, which allows for searching through all stickers in a desktop application. the corresponding physical stickers can be found using an RFID reader. *Move-It* [121] further improved this approach. The system can electronically actuate individual Post-it stickers. This supports not only easy finding of a specific physical sticker, but also allows the system to trigger physical notifications. For instance, a post-it which contains a reminder for a meeting can be actuated shortly before the scheduled time of that meeting.

2.2.2 Augmented Books

Augmented books demonstrate the second main function of augmented paper: enriching printed contents by embedded hyperlinks which point to additional digital resources. Thereby classical reading of a physical book can seamlessly evolve into browsing a physical-digital hypertext. Augmented books are conceptually similar to paper tokens in that they link from paper to digital resources. However in contrast to a paper token, which merely acts as a physical representation of a digital resource, an augmented books has its own value, independently from the digital resources. Here, hyperlinks do not reference between two instantiations of the *same* resource, but between *different* resources.

We can classify these systems following the technology which is used to select hyperlinks on paper (and the resulting interactions), the type of the digitally linked media, and the device on which this digital media is made available. The underlying interaction metaphor is similar in all thee systems: selecting a link hot-spot with a deictic gesture for accessing the associated digital resource. Depending on the technology, the deictic gesture is realized by tapping with a stylus or the finger, clicking with the mouse, or scanning a marker with a camera or a barcode reader.

ActiveBook [137] consists of a paper book and a specific point-and-click selection device that contains a barcode reader and a mouse. Pages of the book contain active areas that serve as link hot-spots to Web resources. These areas can be of various shapes and sizes, allowing for very flexible link anchors. A hot-spot is selected by moving the selection device over it. The system does not automatically recognize which page of the book is open. For this purpose, each page contains a linear barcode that the user has to scan before selecting a link hot-spot.

The *Listen Reader* [6] is an augmented physical book with embedded hyperlinks to audio files. It conceptually improves over ActiveBook in two aspects. First, it does not require any interaction device but allows the user to select links by simply tapping on a link hot-spot. Second, the current page of the book is automatically detected using RFID technology. A passive RFID tag is embedded in each page of the book. An RFID reader in the back cover of the book detects which pages form the right hand side of the opened book and infers the current page. For detecting touch and hover interactions, the Listen Reader relies on inductive sensing.

Books with Voices [62] is an augmented paper book that provides links to video resources. Each link is represented by a linear barcode that is printed on the page margin besides an interlinked paragraph of text. When the user wants to follow this hyperlink, she scans the barcode with a PDA device that features a barcode reader (Fig. 2.8). The PDA identifies the target resource and plays back the video. While the interaction is less direct than with ListenReader and the large number of barcodes potentially interferes with the visual design, Books with Voices has the advantage to be easily deployed in real settings using standard hardware. Moreover, digital contents can be displayed in-situ using on the PDA. Today, a similar concept is widely deployed: QR codes [46] are printed on paper documents that encode the URL of a digital resource. The user can easily display the resource on a camera phone by reading the barcode with the camera.

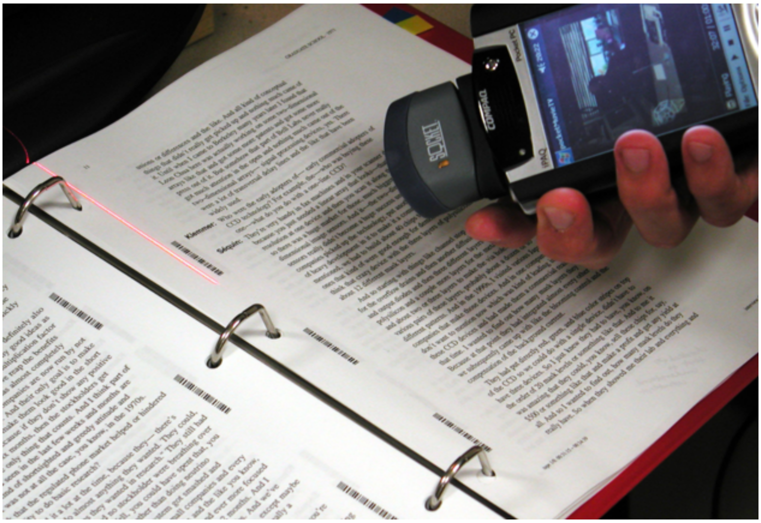

Fig. 2.8 Books with Voices (photo courtesy of Scott Klemmer)

The advent of Anoto pens enabled realizing augmented books that do not require visual markers nor an extensive hardware setup. *Print-n-Link* [108] supports easy retrieval of references in scientific articles by pointing with an Anoto pen on a printed reference. The *LeapFrog Tag Reading System*[16] (Fig. 2.9) is a commercial augmented book. The system turns a storybook interactive with the goal to help children learn to read. An Anoto pen, which is redesigned for kids, can be moved over the pages of the book. Audio information related to the selected content is then played back on the pen. For instance, words are read out. Moreover, the pen automatically logs information about how it is used. This information can be accessed by parents or teachers to examine the kid's learning progress.

These systems offer appropriate support as long as the focus is on reading a document and it is sufficient to have access to some pre-programmed digital resources. However, if the user has a more active role and wants to create user-defined hyperlinks, pre-programmed links are too restricted. We will reviews approaches that allow users to define own hyperlinks on printed documents in Section 2.2.4.

2.2.3 Augmented Paper Notebooks

The third main function of augmented paper consists of synchronizing a paper-based with a digital version of the same resource. Augmented paper notebooks are an important class of interfaces that apply this principle. Contents of a traditional, paper-based notebook are automatically digitized and made available on a computer. This

[16] http:/www.leapfrog.com

Fig. 2.9 LeapFrog Tag Reading System (photo copyright Anoto)

retains many of the advantages of a paper notebook: the user can write and sketch freely with a pen, the notebook is mobile and it can be used in a variety of settings, including casual notetaking in cases when using a computer would be inappropriate, think for instance of a therapy session. Moreover, the paper notebook provides rich physical cues for navigating through the contents.

All pen traces are automatically captured in an electronic form and transferred to a computing device, either continuously in real-time or at specific points in time when the user synchronizes the pen with a computer. A software viewer allows the user to browse through a digital facsimile of the paper notebook. Subsequent updates made on the paper notebook are automatically integrated into the existing digital version. The digital version has the advantage to be searchable, e.g. by time. It is easily accessible even if the physical notebook is not available, for instance for the purpose of archiving. In addition, some augmented notebook systems automatically perform handwriting recognition of the pen traces. The recognized text is typically not displayed, but used in the background to enable full text search within the notebook.

Nowadays, this basic functionality of an augmented paper notebook is provided by many commercial solutions for end-users. Examples include the Livescribe Echo smart pen software[17] and the Oxford Easybook notebooks[18]. While these solutions

[17] http://www.livescribe.com

[18] http://www.oxfordeasybook.com

come with their own viewer software, the software Adapx Capturx for OneNote[19] automatically integrates handwritten notes and sketches into Microsoft OneNote. In addition, some augmented notebooks (e.g. Oxford EasyBook software) allow the user to select commands to be performed once the notes are transferred to the computer. For instance, a note can be automatically sent to a specific person by electronic mail or it can be added as a new task to Microsoft Outlook's task list. These commands are invoked on a note by writing a specific sign (e.g. an encircled letter) next to the note.

Integrating Additional Media

More advanced research prototypes demonstrate how the basic principle of synchronization can be combined with more advanced functionality, such as hyperlinks to additional digital resources. Influential in this respect was a series of research prototypes that aim at supporting biologists. Biologists make frequent use of notebooks for jotting down information when they are in the field, e.g. while observing species or collecting specimen, and also use their notebooks when they are in the lab. Many of these notes are closely related to other resources, such as articles, photos or physical specimen. While these typically remain separate from traditional paper notebooks, augmented notebooks allow biologists to integrate these different sources of information closely with their notebook. As a matter of course, such augmented notebooks do not only support the effective work of biologists, but equally apply to a wide range of professional occupations. For instance, also designers typically make heavy use of notebooks.

The *A-Book* [88] is an augmented lab notebook for biologists. It comprises the basic functionality for accessing paper-based notes via a computer interface that was introduced above. In addition, it contributed a set of important concepts for augmented notebooks. First, the A-Book introduced digital means for integrating the paper notebook with external resources. The user can create hyperlinks to Web pages and can paste physical objects, such as printouts and photos, into the paper notebook, provided that these are "known" to the computer. Second, the A-Book offers functionality for structuring the notebook. The user can easily define a digital table of contents by adding graphical snapshots of passages of the notebook which are listed in chronological order. Moreover, it is easy to connect different passages within the notebook by creating hyperlinks.

A-Book's probably most influential contribution is the Interaction Lens. The Interaction Lens makes digital information (such as hyperlinks) available in-situ, directly within the notebook, instead of on a separate computer screen. The Interaction Lens consists of a PDA, which can be placed onto the paper notebook, moved and rotated. The Interaction Lens appears to be transparent, visualizing on its display the physical contents of the notebook that it is occluding (Fig. 2.10). The position and orientation of the PDA is automatically tracked. This enables the system to maintain

[19] http://www.adapx.com

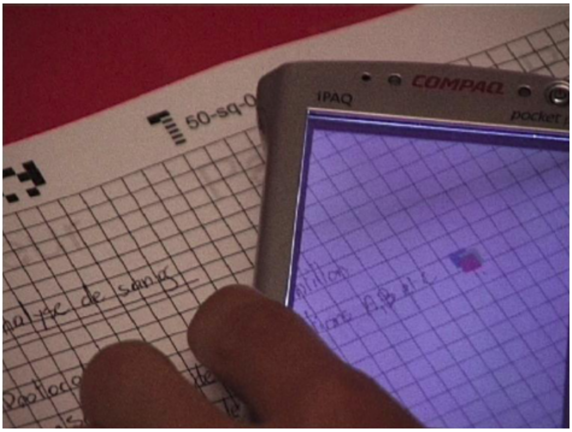

Fig. 2.10 Interaction Lens of the A-Book for accessing digital content directly within the physical notebook (photo courtesy of Wendy Mackay)

the illusion of a transparent PDA for any orientation. The Interaction Lens acts as a window between the physical and the digital documents. It overlays the physical contents with additional digital information, such as hyperlinks. It is not only used for visualizing contents, but also for interacting with the notebook, such as for creating hyperlinks or for adding passages to the table of contents. Hence, the Interactive Lens seamlessly integrates physical and digital information, but only on the small screen of the PDA which must be moved over the page to successively view the digital contents of the entire page. While the concept of the Interaction Lens *per se* is mobile, the A-Book prototype is not, since it relies on a stationary technical setup for capturing pen traces and the position of the PDA, using a digitizing tablet.

ButterflyNet [176] is an augmented notebook for field biologists. It improved upon the A-Book by contributing a set of interaction techniques that allow users to easily link from the notebook to digital resources without requiring a computer. For instance, photos taken with a digital photo camera are automatically added to the digital version of the notebook and aligned with the handwritten contents using their timestamps. Optionally the user can also perform a specific pen gesture to integrate a photo at a specific position within the digital version of the notebook. Moreover, it is possible to easily create links to physical objects, such as specimen collected in the field. The user has first to place the object into an envelope that a barcode is printed on. By taking a photograph of the barcode, the link is established. An important advantage here is that all activities can be done with the paper notebook – no manual post-processing of the digital version is necessary. The ButterflyNet software viewer then provides access to a digital version of the notebook that includes the photos (Fig. 2.11 left). *Memento* [168] is a hybrid physical and digital scrapbook that uses a similar concept for embedding pictures and videos.

Prism [149] is an augmented paper notebook that also targets biologists. The focus is here on the feature-rich Web-based digital version of the notebook into which handwritten notes are automatically integrated. In addition to a digital copy of the handwritten notes, the Web-based version can contain typewritten notes as well as hyperlinks to Web pages, e-mails and local documents or snapshot images

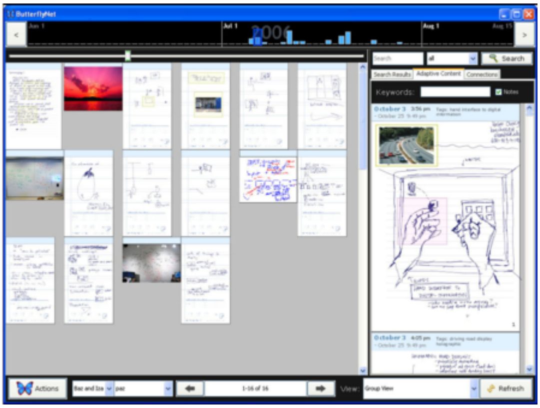

Fig. 2.11 The ButterflyNet/IDeas notebook [176, 69]. Left: Paper notebook. Right: Shared digital notebook including selected pages of multiple users (photo courtesy of Scott Klemmer)

of them. In contrast to ButterflyNet, additional media cannot be integrated using the paper notebook but only in the Web-based interface. The authors conducted one of the most extensive studies that examine how Pen-and-Paper User Interfaces are used. Prism was evaluated in a long-term study over a period of nine months with 5 participants. Amongst others, the authors found that the paper notebook was preferred to be used as a master notebook. This is an organized account of what the users did and planned to do – a central information hub that systematizes and integrates information from various sources.

Another stand of research examines how handwritten notes can be coupled with temporal data, such as audio or video recordings. The *Audio Notebook* [148] closely couples handwritten notes with an audio recorder. It comprises a digital audio recorder and a paper notebook that is placed on top of a digitizing tablet. Handwritten notes are automatically captured in a digital form and used to index the temporal audio stream. If notes are taken while audio is recorded, the notes are automatically linked to the temporal position within the audio stream that corresponds to their creation time. By tapping with the pen on a note, the audio recording is played back at that specific position in time. This is a lightweight, yet powerful indexing mechanism that lays the temporal audio stream out in space. This concept became commercially available with the Livescribe pen. The *ChronoViz* system [164] further refined this concept to support data collection and analysis in observational research. Handwritten field notes are automatically linked to audio and video recordings made during the observation and serve as bookmarks for subsequent data analysis.

The *PaperPDA* [35] is one of the early augmented paper notebook systems. It demonstrates how rich paper-based interactions can be realized even without digital pens, only by using a standard desktop scanner. PaperPDA combines a conventional paper notebook, calendar and organizer with electronic support. Using specific pre-printed paper forms, the user can for instance take notes and write e-mails (Fig. 2.12). This is particularly helpful in mobile settings. Back in the office, the

Fig. 2.12 The PaperPDA. Center: A form for writing e-mails on paper. Right: Physical stickers for creating hyperlinks (photo courtesy of Scott Hudson)

user scans all new or updated forms using a desktop scanner. Specific visual marks on each form allow the system to identify the orientation and the type of the form. PaperPDA then detects marks that the user has made on interface elements (such as printed check boxes) and uses optical character recognition for recognizing text. The PaperPDA system then automatically performs the electronic operations requested by the user, e.g. sending an e-mail. Hyperlinks between two paper documents can be created with physical stickers (Fig. 2.12 right). The user attaches two corresponding stickers to the locations of both link anchors. On each of these stickers, an ID is printed that enables the system to detect corresponding stickers when the forms are scanned. This automatically creates a digital version of the hyperlink.

Collaborative Notetaking

While earlier systems mostly addressed the integration of digital media into the augmented notebook, recent research focuses on supporting collaborative uses of notebooks.

A good example of co-located collaborative use of paper notebooks is *AirTransNote* [97], a system for classrooms. Students take handwritten notes on a paper notebook. The notes are digitally captured and automatically transmitted to the PC of the instructor. This enables the instructor to provide feedback and to discuss notes of students by projecting them onto an electronic whiteboard. Moreover, the system uses handwriting recognition and clusters the recognized notes of all students. This allows the instructor to quickly get an overview of the students' notes.

Remote sharing of notebook contents via a network connection is supported by a number of systems. In the most simple case, the entire notebook is made available to other users without the option to select individual contents to share. This is the case with Memento [168] which allows for accessing shared notebooks in a standard Web browser. The *iDeas* notebook [69] (Fig. 2.11), a collaborative version of ButterflyNet, offers similar functionality. However, it also lets users select specific contents of their personal notebooks to be added to a shared group notebook. Maldonaldo et al. [89] have conducted one of the very few long-term studies of Pen-and-Paper Interfaces. Over a period of 6 months, more than 50 design students used the iDeas notebook for their studies. Main findings were that approximately two third of the participants adopted the novel technology and regularly used the system. While the participants disliked the large size of the pen and its short battery lifespan, they did not create less notes than students using a traditional notebook. Main advantages were seen in sharing notes with other team members and in automatically having a digital copy of the notebook. Prism [149] (introduced above) enables users to share individual pages of their paper notebooks with collaborators using Atom feeds. Sharing is controlled via a Web interface, but not directly from paper. Recently, *UbiSketch* [20] demonstrated how paper notebooks can be integrated with social media. Users can easily share their paper-based sketches on Facebook, Twitter and via e-mail, even on-the-go by connecting the digital pen to a smart phone. Results from a four-week user study indicate that shared sketches stimulate a higher degree of social interaction than shared photos. *PaperSketch* [166] uses a similar concept for sharing sketches with remote users in real-time via a Skype connection. Simultaneously users can communicate over voice and video channels.

EdFest [134, p. 153 sqq.] is a particularly rich augmented notebook that integrates notetaking with printed information and audio feedback. EdFest is a mobile interactive festival guide in a notebook format. The guide contains information about events of a festival, including their title, time, location and a short textual description. With an Anoto pen, the user can write handwritten reviews of events. These notes are transmitted to a central sever. Another user requesting information on that specific event can then access this review via text-to-speech output on a headset. This supports true mobile use, not only for data input, but also for output. EdFest provides additional functionality for a variety of festival-related tasks, such as rating events and getting the directions to an event location.

2.2.4 Augmented Printed Documents

Paper notebooks have the property to be initially empty. It is the user who, bit by bit, fills the notebook with handwritten contents. Yet, people do not only make handwritten notes on empty sheets of paper, but also write on paper documents that contain pre-printed contents. For instance, people use pen and paper to fill in questionnaires and forms, to make annotations on printed documents, to mark up contents or to create references. Such activities enable successful active reading, but also support effective presentation and discussion. The conceptual difference to augmented notebook systems is that handwritten contents must be matched with already existing printed contents. In this section we review a class of systems that allow users to create handwritten annotations on printed documents – annotations ranging from the highly structured answers written into a printed form to highly unstructured free-form comments.

Form Filling

Printed paper forms are still widely used today. They are inexpensive to produce and can be easily used at varying places. Paper forms are also the method of choice when computers are not appropriate for entering data, e.g. because they would create an interactional barrier, such as during interviews or medical consultations. However, most data collected in paper forms has eventually to be integrated into computer databases. Manually copying the data to a computer system or scanning paper forms requires additional effort and increases the time span until the data is electronically available.

This is where form-filling approaches have their benefits. They support structured data entry in mobile settings, retaining the advantages of paper forms. A paper form contains printed fields that act as placeholders where the user can fill in the requested data with a digital pen. This pen automatically captures all data and sends it to a computer system, e.g. via a mobile phone to the back-end system of a company, where the data is further processed. Figure 2.13 shows an example of a form-based interface.

The main commercial supplier of such solutions is Anoto, together with a number of partner companies. Form-based solutions are successfully used in a variety of settings.[20] For instance, field staff of several large companies use form-based solutions for filling-in order forms during their customer visits. Policemen use them

Fig. 2.13 Form-based interface (photo copyright Anoto)

[20] http://www.anoto.com/all-cases-4.aspx

for recording the details of traffic violations. Physicians and nurses take medical records in hospitals. The technology was even tested in the course of German elections [161], but eventually not used in subsequent elections due to security problems in the electronic voting system. More advanced solutions, that go beyond simply recording the data entered in a form, are currently investigated in research projects. For instance, one project examines how printed forms can support speech-language therapy [117].

Annotating Printed Documents

Capturing handwritten annotations on a printed document and synchronizing them with a digital version of the document is one of the most frequently addressed issues in Pen-and-Paper Computing. A very influential system is *Paper Augmented Digital Documents (PADD)* [32]. This system introduced the main principle of paper-based annotation: A digital document is printed on paper. This printout is used as a proxy to interact with the digital version of this document. Handwritten annotations made on the printout are automatically digitized and added to the digital document, which can be accessed on a computer. Moreover, the user can print updated versions of the document that include new annotations. This results in a digital-paper-digital annotation lifecycle. The PADD prototype uses an Anoto pen comes with a plug-in for Adobe Acrobat to annotate PDF documents. This concept was taken on by commercial solutions, including Anoto PenDocuments Pro[21] and Adapx Capturx Markup for PDF[22].

Proofrite [18] is a paper-augmented word processor that applies the PADD concept to the use case of proofreading. The user can annotate the printed version of a word processor document with proofreading marks that are synchronized with the digital version. In the digital view they reflow automatically when the surrounding text is modified. Neither PADD nor Proofrite do interpret the handwritings, but only capture and visualize them as an additional layer on top of the digital document. *PaperProof* [165] demonstrates how handwritten proofreading marks can be automatically interpreted to modify the underlying digital document. In contrast to the above systems, PaperProof creates a detailed semantic model of the printed document that captures what contents are printed at what locations on the paper document. This allows PaperProof to relate proofreading marks to document contents. For instance, if the user crosses out a word, PaperProof recognizes this gesture and identifies which word of the textual document is printed at the position of the cross-out mark. The word is then deleted in the word processor (see Fig. 2.14). Different gestures support inserting or moving text and making annotations. The prototype is integrated into OpenOffice.

The PADD principle was not only applied to word processing and standard office document formats. For instance, Pen-and-Paper Interfaces allow for annotations on

[21] http://www.anoto.com

[22] http://www.capturx.com

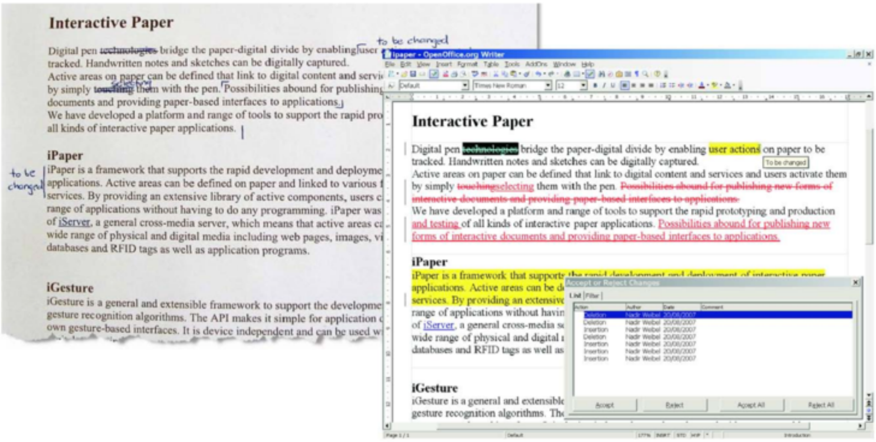

Fig. 2.14 PaperProof automatically interprets proofreading marks made on a printout (left) and applies them to the document in the word processor (right) (photo courtesy of Nadir Weibel)

large maps in geographic information systems [94]. Another example, *Musink*, supports annotations on musical partitions [154]. The system automatically analyzes them and integrates the printed partitions with OpenMusic, a computer-based music composition tool.

ModelCraft demonstrates that pen-and-paper annotations are not restricted to flat paper. The system supports handwritten annotations on physical 3D models that are made of paper [140]. Figure 2.15 shows a 3D model with some annotations. The annotations are automatically added to the underlying CAD model. However, in contrast to flat documents which can be easily re-printed, creating updated versions of physical models requires manual work. This limits the paper-digital annotation lifecycle.

One important benefit of digitized annotations is that they are not bound to the physical medium they have originally been made on. Hence, they can be more easily shared with other people, either in co-located settings or over distance:

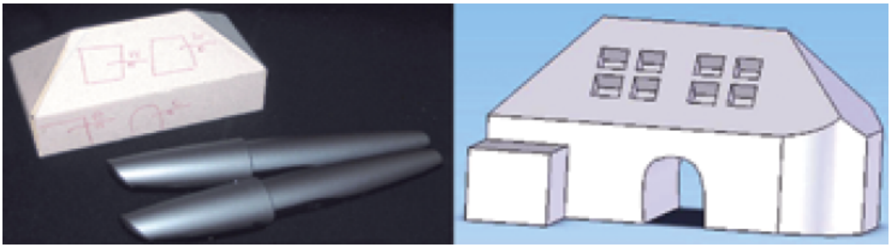

Fig. 2.15 ModelCraft provides for annotating three-dimensional paper models (photo courtesy of Hyunyoung Song)

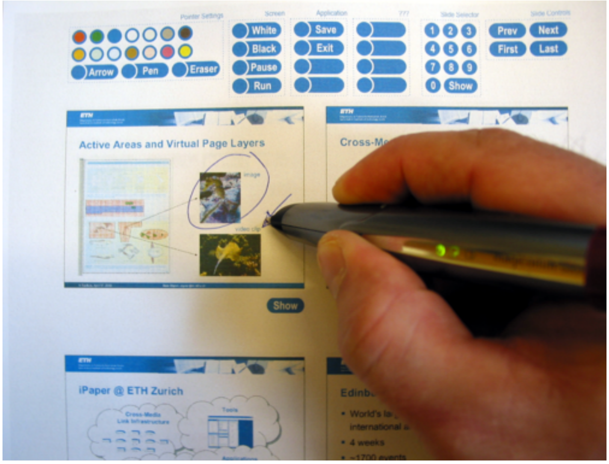

Fig. 2.16 PaperPoint integrates paper-based annotations into slide presentations (photo courtesy of Beat Signer)

PaperPoint [135] shows how a paper-based interface can support slide presentations. The presenter uses a printout of the presentation slides (Fig. 2.16). By making pen gestures on the printout, the presenter can control which slide is presented. In addition, handwritten annotations made on the printout are automatically integrated into the projected slide in real-time. The concept allows presenters to overcome the implicit linear character of slide presentations by developing a slide sequence on-the-fly that meets the demands of the audience. Anoto has commercialized this concept as a product called Anoto PenPresenter[23].

Similarly, *PaperCP* [74] supports in-classroom collaboration by slide annotations. Students can annotate printed handouts of lecture slides and electronically transmit their annotations to the instructor. The authors improve over PaperPoint by introducing a paper-based mechanism for defining which annotations are shared with the instructor and which annotations are kept private. This feature relies on two separate areas on each slide – one is used for making shared notes, the other for private notes.

CoScribe (cf. chapter 5 of this book and [144]) supports asynchronous sharing of annotations made on printed PDF documents and presentation slides. Different personal annotation styles are addressed by flexible print layouts, e.g. additional notetaking areas support extensive notes or many shared annotations. A paper-based sharing mechanism allows for sharing annotations with different groups of people. A novel visualization of shared handwritten annotations provides efficient access to annotations of multiple users in one single view.

[23] http://www.anoto.com/facilisis-a-consequat-quis-1.aspx

Handwritten Categorization of Contents

Tagging is a more formal type of annotation. A tag is an annotation made on a document that describes its contents in a very condensed form, e.g. by a keyword. The goal of tagging is to classify documents and passages of documents in order to retrieve them more easily in the future. On the Web, tagging of resources is nowadays very common. Browser bookmarks allow users to make their personal bookmarks. Collaborative bookmarking tools such as del.icio.us[24], digg[25] or cite-u-like[26] enable sharing bookmarks with other users. Here we review concepts for tagging contents that are printed on paper. Most previous work relies on pen gestures.

PapierCraft [75] extends the PADD system for paper-based annotations, which was introduced above. It is a gesture-based command system for paper documents, relying on the Anoto technology and containing a rich set of pen gestures. These gestures enable not only tagging documents, but also creating and following hyperlinks on documents, copy/paste within paper documents, performing Google searches using keywords from the printed document and mailing portions of a document to other persons. To perform a command, the user must first select the portion of the document that the command applies to: Text snippets can be selected by underlining or encircling; several lines of text can be selected with a vertical line drawn besides the lines in the margin of the document; also arbitrary rectangular areas can be selected. Then the command is specified with a specific gesture.

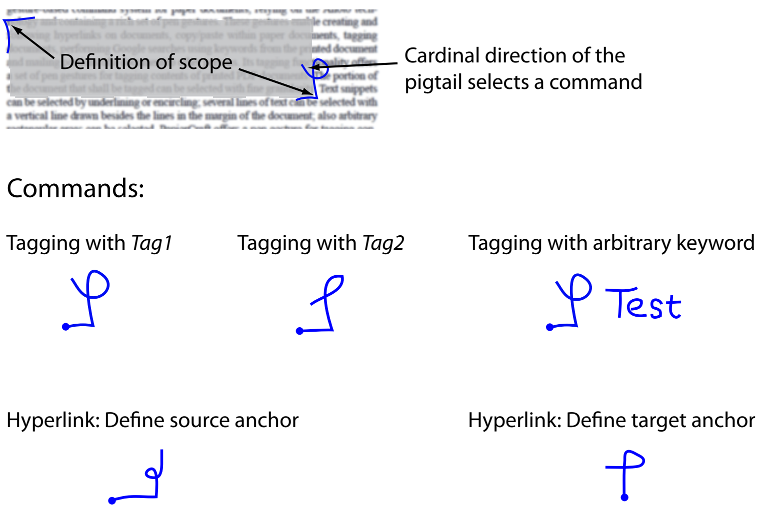

Fig. 2.17 PapierCraft's pen gestures for tagging document passages with predefined or freely-chosen keywords and for creating hyperlinks

24 http://del.icio.us

25 http://digg.com

26 http://www.citeulike.org

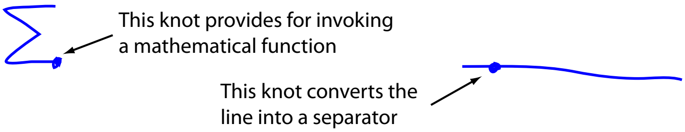

Fig. 2.18 Two examples of Knotty Gestures

As an example, Figure 2.17 depicts PapierCraft's gestures for tagging and linking contents. PapierCraft offers to select one of two predefined tagging categories. The cardinal direction of the ending of the pigtail gesture decides upon the category. This approach guarantees readability on paper and a reliable gesture recognition, but is restricted by the small number of categories that can be supported (up to the eight cardinal and secondary directions). Moreover, the abstract gestures have no natural connection to the category. As an alternative, the user can also create tags with an arbitrary handwritten keyword. To allow the system to reliably detect whether the user is writing or performing a gesture for invoking a command, the authors suggest using a secondary device, such as a push button.

Schumacher et al. [127] demonstrate how handwritten keywords can be further automatically processed. Similar to PapierCraft, their concept allows the user to select a passage of a document and to define a handwritten keyword. The system automatically recognizes the handwritten keyword and also extracts the selected text. This information is stored in the user's personal information base, an RDF-based ontology, and can henceforth be used for queries.

Knotty Gestures [155] aims at overcoming some of the limitations of traditional pen gestures. Handwritten gestures, such as those of PapierCraft, clutter up the content and take valuable space on paper if they are used heavily. Knotty gestures are less obtrusive. A knotty gesture is a small knot that resides on top of any other pen trace (see Fig. 2.18). One or several knots can be added at arbitrary points on an existing trace. The knot defines the role of the trace on which is resides. For instance, a knot can convert a line, which per se has no clearly defined semantic meaning, into a link anchor, a separator or an interactive slider. As another example, knots can be used for controlling recording and playback of audio recordings. Moreover, knots are not only useful during their creation; they also define a point of interaction that can be revisited at a later point in time. For instance, tapping with the pen on an already existing knot triggers associated commands. Circling around the center of a knot gives access to a list of options and selects one of them. The system is prototypically implemented on a Livescibe pen. The authors show that gestures can be reliably recognized directly on the pen, without the need for an additional computing device.

CoScribe (cf. Chapter 7 of this book and [144]) introduces a set of techniques for tagging contents on paper documents that goes beyond pen gestures. Digital Paper Bookmarks are physical stickers that can be attached to pages of a physical document for tagging them. The position of the bookmark and its handwritten label are automatically captured and made available in several visualizations. Bookmarks can be shared with other users over a network connection and can be compared in a collaborative view. A second tagging technique uses separate paper cards containing an inventory of tagging categories that can be modified and extended by end-users.

Handwritten Referencing

Handwritten annotations often contain references to different passages of the same document or to a second documents. Such user-defined references are not only beneficial for quick retrieval of information but also help in integrating and structuring information from many sources. We will now review pen-and-paper systems that provide for creating and following hyperlinks on paper.

PaperLink [5] was an early system within this class. It uses a specific pen onto which a camera is attached (Fig. 2.19 shows a commercialized version of PaperLink). This camera serves for creating and detecting link hot-spots. The visual contents of the document serve as link anchors. If the pen is placed on a paper document, the camera captures an image of the document area around the pen tip. The image is processed on a computer using simple computer vision techniques. The pattern which appears at the center of the image (typically an individual word) is extracted. This pattern can be associated with a digital resource and serves from now on as a link hot-spot. If it is detected in the camera image, the target resource is opened on the computer.

The *Interactive Multimedia Textbook* [68] offers similar pen-based interactions for creating and following hyperlinks from printed documents to Web pages. It relies on an ultrasonic pen. Therefore link anchors do not have to be bound to specific

Fig. 2.19 The Hitachi LinkStick is a commercialized version of PaperLink. The pen has a built-in camera. From the camera image, a text pattern is extracted and acts as anchor for a hyperlink (photo courtesy of Toshifumi Arai)

visible marks on the document (e.g. an individual word). Instead hot-spot areas can be freely defined on the paper surface. Tapping with the pen on such an area displays the associated digital document. Both PaperLink and the Interactive Multimedia Textbook are limited to links from paper to digital media.

PapierCraft [75] (introduced above) allows the user to create hyperlinks between two paper documents. A hyperlink is created by drawing two specific pen gestures on both link anchors (see Fig. 2.17 on p. 53). To create links from or to digital documents, the same gestures can be used on a Tablet PC on which the digital document is displayed. In contrast to the other solutions discussed in this section, it is not possible to activate a hyperlink on paper, but only in the PapierCraft software viewer.

The hyperlinking functionality of *CoScribe* (cf. Chapter 6 of this book and [144]) focuses on tasks that require integrating and structuring information from a variety of documents, including printed documents and Web pages. Hyperlinks are created with pen gestures similar to PapierCraft's gestures. CoScribe introduces the following two novel aspects: On the one hand, it integrates the linking interactions; the same Anoto pen and the same gestures can be used both on paper and on displays. This avoids the need for switching between different devices and different interaction techniques, allowing seamless linking between both realms. A plug-in for Mozilla Firefox supports pen-based linking within Web pages. On the other hand, CoScribe automatically integrates linking activities of multiple users and multiple documents into one workgroup view. This view allows users to identify higher-level linking patterns that go beyond individual links.

In-place Visual Feedback

Almost all of the solutions discussed above strongly separate paper-based input from digital output. Digital versions of the documents are made available on a separate screen. We conclude this section by reviewing several publications that examine how digital information can be visualized directly *within* the paper document.

PenLight [138] introduced the concept of a digital pen that features a built-in projector. The pen projects a superimposed layer of digital information onto the paper document. This layer shows additional contents, such as annotations made by other users or additional images, but also provides visual guidance for invoking commands from pie menus. As current mobile projectors are still too large to be mounted to a pen, the authors simulat a projecting pen by tracking the position and orientation of the pen. Contents are then projected by a projector that is mounted at the ceiling above the table. *MouseLight* [139] is a follow-up work by the same authors. Here a real mobile projector is used, which however is detached from the pen and forms a separate device. This device, which resembles a mouse, can be placed onto paper documents and projects a superimposed layer of digital information. Figure 2.20 depicts MouseLight. The concept is similar to PenLight, but interaction is now bimanual. One hand is interacting with the pen while the other hand is manip-

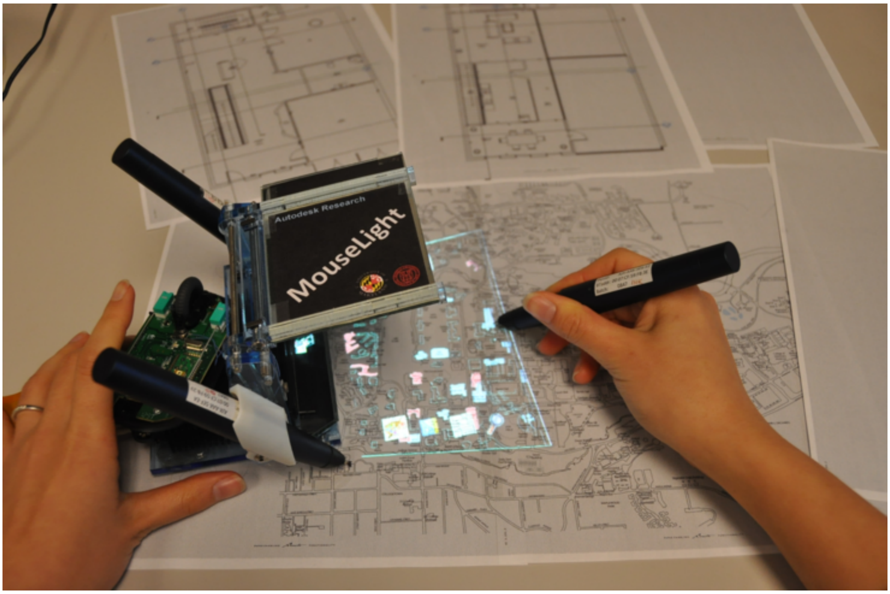

Fig. 2.20 MouseLight combines a digital pen with a separate mobile projector (photo courtesy of Hyunyoung Song)

ulating the projecting device. This leaves more interactional flexibility and ensures a stable projection even while the pen is used for writing.

Similar to MouseLight, *FACT* [78] uses a pen and a separate projection unit. In contrast to MouseLight, FACT leverages a camera and analyzes visual features to automatically identify paper documents and to track their position and orientation in 2D space. As a consequence, no Anoto pattern is required and documents can be placed and moved quite naturally on a table. The projector provides digital output on the paper documents in real time, enabling computer-like functionality on paper. For instance the user can select a word for performing a full-text search by tapping with the pen on a word that is printed on the document. All occurrences of this word are then highlighted by the projector directly within the printed document. Moreover, the system projects digital annotations onto the paper document, such as links to related Web pages. Further functionalities include Web search, copy/paste of paper contents and remote sharing. PACER [77], a predecessor system of FACT, enables similar interactions with paper and a mobile phone, however without in situ projection of digital contents.

2.2.5 Augmented Tables, Flipcharts and Whiteboards

Augmented tables, flipcharts and whiteboards combine paper-based media with interactive tabletop and wall displays. They provide a close integration of paper and digital media by combining them on one single surface. Paper documents can be used directly on top or in front of a large computer display. Alternatively, or in addition, digital information is displayed directly on paper.

Above we have already reviewed some concepts to display digital information in-situ, within a paper document, e.g. by leveraging mobile projectors or superimposed PDA displays. However, digital information was restricted to a small surface. In contrast, augmented tables, flipcharts and whiteboards realize very large digital display surfaces. These systems combine many of the augmented paper principles that we have introduced above.

Early works on augmented desks date back many decades. Already in 1945, Vannevar Bush envisioned *Memex* [12], an augmented desk that is able to display virtual documents on the table surface. Memex was thought of as an electro-mechanical machine that stores all books and other documents of a user on microfilm. Memex enables the user to read and annotate these documents as well as to create associations (we would call them hyperlinks today) between these documents. Memex never went beyond a theoretical state; however, it strongly influenced the development of hypertext systems and of augmented desks.

The foundations of digital augmented desks and tables, and of paper-based computing in general, were laid in a series of seminal works at EuroPARC in the early 1990s. A highly influential system is Wellner's *DigitalDesk* [167]. As shown in Fig. 2.21, the user can place printed documents on the table, very much like on a traditional desk. In addition, digital documents are projected onto the desk surface, and printed documents can be overlaid with projected information. This creates a very seamless integration of physical and digital contents.

The combined tracking and projection setup of the DigitalDesk inspired many follow-up works. It is depicted in Fig. 2.21 (left). A camera is mounted at a fix position above the desk. By analyzing the stream of images captured by this camera, the system identifies the position and the contents of printed documents. Moreover, the camera images are used for detecting pen and touch input on the desk – both on paper documents and on projected contents. Similar to the camera, the projector is mounted at a fix location above the desk.

Several example applications of the DigitalDesk introduced a set of novel interaction techniques. One interaction allows for physical-digital copy&paste of data. In a calculator application (Fig. 2.21 right), the user can copy numbers from printed documents by pointing on them. For recognizing numbers within the document snapshot captured by the camera, optical character recognition is used. Another interaction realizes physical-digital copy&paste of images. The user can draw a sketch on a sheet of paper and copy this sketch to a digital collage using a simple gesture. In the background, the system extracts an image of the sketch from the camera snapshot and projects this as a part of the collage. An extension of the initial system [124] in addition allows users to follow hyperlinks on printed versions of Web pages. By

Fig. 2.21 The DigitalDesk. Left: Tracking-projection setup of the prototype. Right: copy&paste between a paper document and a calculator application (photos courtesy of Pierre Wellner)

tapping with the pen on a printed hyperlink, the target Web page is displayed in in a browser window that is projected onto the desk besides the paper document.

The DigitalDesk was the conceptual basis for a number of subsequent systems developed at EuroPARC. The *Digital Drawing Board* [86] is both a conventional drawing board and a top-projection display. It allows designers to make paper-based construction sketches on the board and to easily digitize them. The digitized images can then serve for a variety of purposes. For instance, they can be automatically rendered as textures onto objects to ease comparison of alternative designs.

Video Mosaic [85] addresses video editing. It combines paper video storyboards with the capabilities of video editing software. Obviously, planning and creating the temporal sequence of scenes is at the heart of video editing. The work examines how to best support the user in seeing in one glance such dynamic data. The chosen solution is to lay out time in physical space. Paper-based storyboard elements act as proxies for video clips. By physically arranging storyboard elements in a sequence on the desk, the user can change the sequence of clips to be played. Each storyboard element can be annotated with handwritten notes and sketches and can be associated with additional digital resources. The system allows the user to easily play back the current sequence, to digitize and reprint the storyboard elements, as well as to share them with co-workers.

EnhancedDesk [64] is based on a similar setup as the DigitalDesk, but uses two cameras for tracking. It further improved hand tracking and gesture recognition.

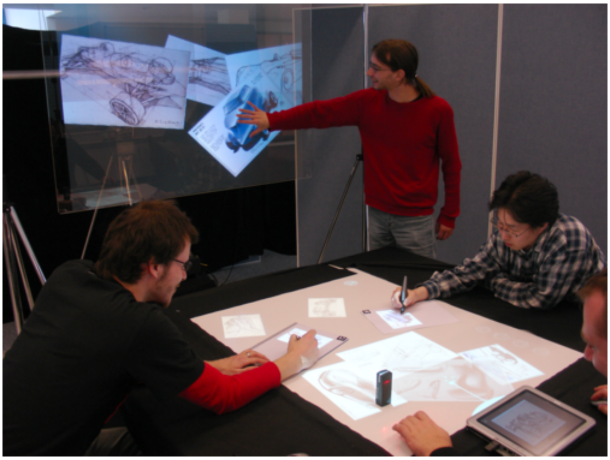

Fig. 2.22 The Shared Design Space integrates paper with an interactive tabletop and an interactive wall in a collaborative environment (photo courtesy of Michael Haller)

Digital documents are identified by fiducials. In an example application, a physical textbook is augmented by additional digital information that is projected onto the desk surface besides the textbook.

The *Augmented Surfaces* system [123] leverages both tables and walls as shared information surfaces. Using a mouse cursor, information can be seamlessly moved between these surfaces and ordinary computer displays. Moreover, the system leverages fiducials for identifying physical objects that are placed on the table. While the DigitalDesk provides for copy&paste from paper to digital, Augmented Surfaces support the reverse direction: digital contents can be virtually attached to a physical object. These digital contents are projected either directly onto the object or onto the table around the object, enabling compound collections of physical and digital objects. When the user moves or rotates the object, the projection is automatically updated to create the experience that the projected contents are physically attached to the object. A number of further seminal tabletop systems, such as *metaDESK* [156], were introduced in the late 1990s. These are not discussed here because they do not offer Pen-and-Paper Interaction.

The *Shared Design Space* [33] (Fig. 2.22) integrates paper with tabletop and wall displays. Similar to Augmented Surfaces, the user can move digital contents (e.g. videos and images) onto pages of paper notebooks. Moreover, the system automatically captures annotations and sketches that are made in the notebook with an Anoto pen. A clone functionality synchronizes notebook pages between multiple users in real time, blurring the boundary between the private notebook and the shared space. While the user who is making an annotation is leaving physical ink traces on his or her notebook, the collaborators simultaneously get a virtual copy projected on their own paper. Similar in approach, *Pictionaire* [34] is an augmented tabletop for collaborative design brainstorming. It coherently integrates the hybrid copy&paste

operations introduced by the DigitalDesk and Augmented Surfaces. A drag-off gesture creates a digital copy of sketches that were made on physical paper. In the reverse direction, the user can snap digital images to paper surfaces.

Some time after interactive desk and table systems allowed for combining paper and displays in a horizontal configuration, interactive wall systems addressed the use of paper directly on vertical displays. The *Designers' Outpost* [63] focuses on creative planning tasks with post-it stickers (Fig. 2.23). It features a rear-projection whiteboard on which users can create hybrid physical-digital collages. These collages consist of physical post-it stickers that are attached onto the whiteboard and of virtual pen traces that are made on the whiteboard. A set of interactions allows for attaching a post-it at an arbitrary location on the whiteboard, moving or removing it, as well as for adding virtual pen traces. The locations and contents of the post-it stickers are automatically captured by a camera. This allows the user to save all contents of the board with one simple click. Saved contents can be accessed either on the whiteboard or via a computer interface. *DigiPost* [50] uses a similar principle on a horizontal tabletop display.

Similar to Designers' Outpost, *DocuDesk* [27] captures the physical arrangement of paper items. Its aim is to support users in creating multi-way links between paper documents and digital documents that are displayed on a horizontal tabletop. The system automatically captures contents of paper documents as well as hyperlinks. This allows the user to quickly re-establish the state of open windows and documents when resuming a task. To do so, it suffices to place one of the documents onto the screen. All other documents which are linked to that one are then displayed on the screen.

Placing paper documents onto a display potentially occludes digital contents. Only recently, research started addressing paper-based occlusion of screen contents. An empirical study [146] analyzed spatial patterns of how printed and digital doc-

Fig. 2.23 The Designers' Outpost (photo courtesy of Scott Klemmer)

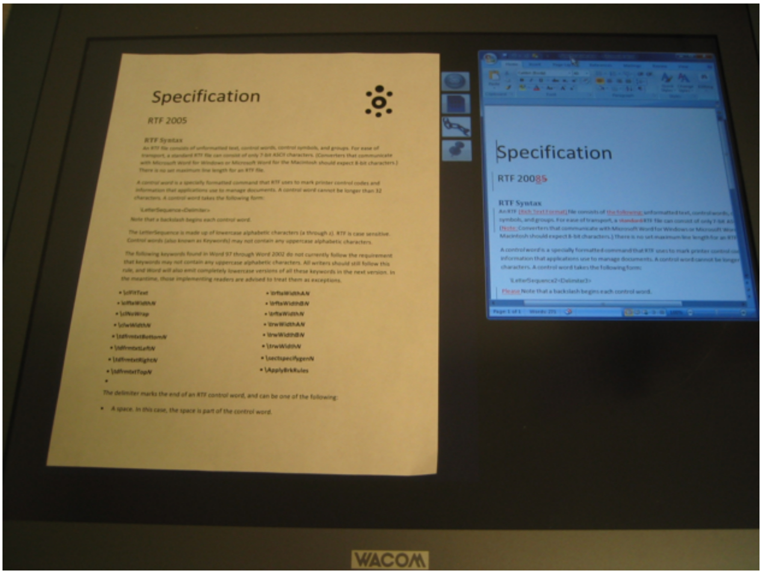

Fig. 2.24 DocuDesk (photo courtesy of Katherine Everitt)

uments are used simultaneously on the same surface and how users cope with occlusion. Another publication [57] introduced interaction techniques for managing hybrid paper-based and digital piles on tabletops that take into account paper-based occlusion.

Another strand of research examines how physical paper itself can become a large interactive surface, rather than combining (small) paper documents with (large) displays. The underlying premise of these works is that paper can be considered a display with a very slow update rate. A paper poster is therefore an inexpensive way of realizing a large, high-resolution display. *Gigapixel Prints* [177] presented a set of innovative applications of interactive paper posters. While input is realized with one or several Anoto pens, contents of the display can be updated either by reprinting the poster (slow, but high-resolution update) or by projecting additional contents onto the poster. The concept can play to its strengths when large data sets have to be visualized that are relatively stable over time. *PLink* [147] leverages large paper deskpads for integrating the physical desk and the computer desktop of office workers. The deskpad acts as a large linking area on which the user can easily and quickly create and access links to resources of the digital desktop. A four-week field study showed that PLink enables users to quickly access digital resources and supports flexible organization of digital information by laying out links in physical space.

The systems reviewed in the this section require that paper documents be kept flat on a two-dimensional surface. This restricts the natural interaction with paper, which to a large extent is manipulated above the table [151, 146]. Recent research opens up the 3D space above the table for paper-digital interaction:

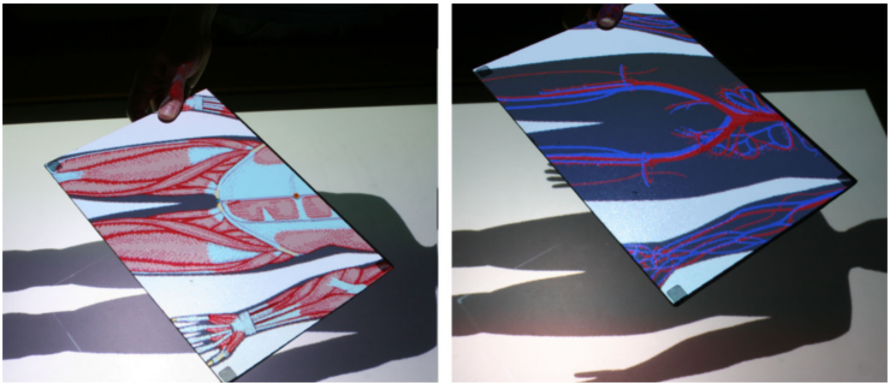

Fig. 2.25 PaperLens offers tangible views for information visualization above interactive tabletops (photo courtesy of Raimund Dachselt)

A very influential system is *PaperWindows* [40]. It introduces thin and lightweight paper displays as an interface for the windows of the operating system. Using a camera-projection unit, PaperWindows displays the contents of each GUI window onto one sheet of paper that the user can freely move on the desk. The authors present a set of interactions for copying windows from the computer to paper and back, for navigating through contents on the paper displays, for copy&paste of contents, and for annotating contents. These gestures leverage not only pen and touch input, but introduce paper manipulations, such as flipping or stacking pages, as a means for controlling the system.

PaperLens [142] (Fig. 2.25) is a tangible view for information visualization. A passive paper display can be used as an interactive lens for data that is displayed on an interactive tabletop. This is similar to the Interaction Lens of the A-Book (introduced above), but the PaperLens can be freely moved and rotated in three dimensions above the tabletop display. This allows for a set of novel interactions and visualizations that take into account not only the lens' position on the tabletop, but also its distance from the table surface and its 3D orientation.

The concepts discussed above use paper sheets of a fix size. One advantage of paper is that it can be modified in shape and size, for instance by folding. Lee et al. [70] examine different form factors of paper displays that allow for resizing the display. They present a large rectangular display that can be folded, a scroll that can be rolled in an out, as well as form factors for resizable displays which are inspired by fans and umbrellas. Inspired by this work, *Xpaaand* [56] addresses interaction techniques for such resizable displays. The authors present a hardware prototype of a passive rollable display and introduce a set of interaction techniques for manipulating digital information by resizing the display.

2.3 Directions of Future Research

This state of the art survey has shown that Pen-and-Paper Interfaces are technically mature and have found their way into commercially successful applications. A wide variety of technologies allow for capturing paper-based contents and for realizing digital output on paper. A large number of interface concepts cover three main functions of augmented paper with various paper media, ranging from tiny stickers over books and printed documents to large paper posters. The survey has also shown that a growing body of research aims at integrating paper more closely with real-time visual output, resulting in augmented digital pens and in paper-digital interactive surfaces.

To conclude this chapter, we briefly outline directions of future research on Pen-and-Paper Computing. A more comprehensive discussion can be found at the end of this book in Chapter 8.2. While most previous work has focused on data input on paper (output being provided on a separate computer screen), future work will aim at further *enhancing real-time feedback on paper*. Livescribe has shown the direction that future pens are likely to follow further: include more powerful processors to run complex applications directly on the pen and provide more real-time feedback on the pen. Future pens might feature larger displays, a built-in projector or even a built-in inkjet printer to leave permanent marks on paper. Moreover, mobile tracking-projection solutions are promising for transforming any paper document into an interactive surface. The increasing processing power of mobile phones and the advent of very small mobile projection units let us expect that in the near future, every smart phone comprises the components that are required for paper-digital interactive surfaces. A further, highly promising direction are flexible displays. These combine many affordances of paper with the powerful capabilities of displays. It is very likely that such displays will open up new and previously unforeseen ways of interacting with digital information.

On the level of applications, we see five major challenges. The first challenge is to fully leverage the *mobile character of paper*. Most prior applications either cannot be used at all in mobile settings or only parts of their functionality are available in a mobile setting. We expect more applications that couple paper and a digital pen with a mobile phone. This coupling results in a powerful device federation that requires only standard hardware which is already available today. Future work should examine how the user interface can be repartitioned between pen-and-paper and the mobile phone. A second challenge is related to how we manipulate paper in a Pen-and-Paper User Interface. Powerful new tracking technologies allow for capturing *manipulations that deform paper*, such as bending, folding and rolling. Future work should examine these interactions more deeply.

A third challenge consists of improving *large-scale collaboration*. Most current applications focus on a single user. In particular, it is still not fully understood how to process, integrate and visualize paper-based contents that are created by a very large community of users. This point is related to another challenge, the *interpretation* of contents. Almost all current systems interpret pen-and-paper interactions only to a very limited extent. The systems typically display only a facsimile of the handwritten contents, sometimes performing handwriting recognition in the background to allow for full-text search. It will be interesting to see how contents that are created on pen-and-paper can be more directly integrated with tagging platforms, blogs and social networks, known as the Web 2.0.

This survey has shown that the hardware for realizing Pen-and-Paper Interfaces is readily available. However, the field lacks *standards and interoperability* of solutions. This concerns not only interfaces for abstracting from the pen hardware and an effective standard for digital ink data, but also standardizations in authoring and publishing.

Finally, we have seen that, even though a large number of Pen-and-Paper Interfaces was developed, there is only very little research examining *how people use these interfaces*. Most publications do either not report on user feedback at all or provide only limited insights. Only a small number of of studies examine how Pen-and-Paper Interfaces are used over a longer period of time and how they are integrated into existing information ecologies. Future work should definitely deepen our understanding and conduct more long-term studies of Pen-and-Paper Interfaces.

http://www.springer.com/978-3-642-20275-9

Pen-and-Paper User Interfaces
Integrating Printed and Digital Documents
Steimle, J.
2012, XII, 192 p., Hardcover
ISBN: 978-3-642-20275-9