Handwritten Notes as a Visual Interface to Index, Edit and Publish Audio/Video Highlights

Anselm Spoerri

School of Communication, Information & Library Studies
Rutgers University, 4 Huntington Street, New Brunswick, NJ 08901, USA.
aspoerri@scils.rutgers.edu
http://scils.rutgers.edu/~aspoerri

Abstract. Digital libraries aim to make media-rich information accessible to "anyone, anywhere, anytime." However, digital audio and video are difficult to search and share. This paper describes Souvenir, a system which enables people to use their handwritten or text notes to retrieve and share specific media moments. Souvenir enables users to take time-stamped notes on a variety of devices, such as the paper-based CrossPad, the Palm Pilot and standard keyboard devices. Souvenir can segment unstructured handwriting into an effective media index without the need for handwriting recognition. People can use their notes to create hyperlinks to random-access media stored in digital libraries. Souvenir has also web publishing and email capabilities to enable anyone to access or email media moments directly from a web page. This paper presents an overview of Souvenir's functionality and describes its handwriting segmentation as well as note and media synchronization algorithms.

1 Introduction

Digital libraries aim to make media-rich information accessible to "*anyone, anywhere, anytime*." However, users need to find it easy to access, use and share this digital information. They have to be able to *personalize* the contents of a digital media library by creating *annotations* that can be used as a personal access index. Users want *flexibility* in how they can interact with digital libraries: 1) Users don't want to have to learn specific query languages. Instead, they want to be able to use familiar and "loosely structured" interaction modes to access highly structured digital libraries. 2) Users want to use diverse access and annotation devices, such as palmtop, laptop or desktop computers. 3) Users want to create and store their digital media annotations in a distributed fashion on different devices and use this meta-data in their workflow without being held captive by a digital media library.

This paper describes Souvenir which enables people to use what comes easy to them, creating handwritten notes, as an access interface and annotation mechanism to do something that is difficult for most, retrieving and sharing specific media moments in digital libraries. The goal of Souvenir is to offer a flexible and comprehensive way for anyone to share and integrate digital media moments into their workflow.

People are increasingly using handheld computers to record information. Tablets and devices combining paper and digitizing technology, such as the CrossPad or SmartPad, can be used to create handwritten notes that are captured as digital ink. Souvenir supports a variety of note-taking devices to enable people to use their notes to playback or share specific moments in digital media libraries.

1.1 User Need

Digital media content is growing exponentially and represents the next wave of the Internet. Corporations are adopting streaming media to save time, reduce costs and enrich their communications. Knowledge workers need to be able to integrate specific media moments into their workflow. Consumers are acquiring digital media recording equipment and starting to create personal digital media collections. However, audio/video is difficult to search and share in a personal way [1], [6], [8], [10]. General users do not possess the media editing skills and tools that are currently needed to pinpoint and organize specific media moments. What if people could use existing skills, such as taking notes and editing text, to pinpoint specific media moments and create personalized media presentations?

Souvenir offers users flexibility in how they can interact with digital media: 1) Users can use familiar skills and "loosely structured" interaction modes to access digital media libraries. 2) Users can personalize digital audio/video by creating annotations that can be used as a visual access index. They can use a variety of access and annotation devices, such as palmtop, laptop or desktop computers. 3) Users can create hyperlinks to random-access and integrate specific media moments into their workflow without being held captive by how and where the digital media is stored.

This paper is organized as follows: First, previous work is briefly reviewed. Second, Souvenir's functionality is described. Different scenarios are discussed of how users can take media-enabled notes on a variety of devices. We address the strengths and weaknesses of the different note-taking devices supported by Souvenir. A flowchart shows how Souvenir can convert unstructured handwritten notes into media-enabled web pages. A "segment-oriented" framework is introduced to support both handwritten and text notes as well as the playback of different media recordings in the same Souvenir document. Third, we discuss the special issues that arise when handwritten notes, captured as digital ink, are to be used to index digital media. An algorithm for segmenting digital ink is described that exploits spatial and temporal characteristics of how people write without having to recognize their handwriting. We also discuss the need for multiple and complementary ways to index digital media libraries. Fourth, a flexible mechanism for synchronizing notes and media is described. Finally, we provide an informal evaluation of Souvenir and discuss its effectiveness.

2 Previous Work

Previous research has investigated how to provide easy access to digital media collections [1], [4], [5], [8], [9], [10]. Xerox PARC has developed Tivoli [4] for a note-taker to summarize captured meetings. Classroom 2000 [1] automatically captures classes and integrates them with annotations created during class. Marquee [8], Filochat [10], Dynomite [11] and Audio Notebook [6] have clearly demonstrated how personal note-taking can make audio/video retrieval easier.

The use of digital ink to access digital media has been studied by several research groups [1], [3], [4], [6], [8], [10]. [11]. The Classroom 2000 system [2], [7] employs a simple temporal and spatial heuristic to link handwritten notes with audio recorded at the same time. Chiu and Wilcox [3] present a way to generalize simple heuristics into a more general algorithm using hierarchical clustering.

In terms of automatic methods for indexing digital media, there is the major Informedia initiative that combines advanced speech, language and image understanding technology to transcribe, segment and index large video media repositories [9].

Souvenir is innovative in that it offers a flexible and visual solution for anyone to personalize and retrieve audio/video and create hyperlinks to random-access media to share specific moments with others. Souvenir offers users a seamless way to structure their handwritten notes, combine them with text and to publish both digital ink and text as media-enabled web pages.

3 Souvenir

Souvenir is a digital media annotation tool with web publishing and email capabilities. It enables people to use their handwritten or text notes as a personal and visual audio/video index. Souvenir time stamps the user's note-taking activity and uses these time stamps to synchronize the notes with the timeline of a media recording (see Figure 1). Palmtop, laptop or desktop computers can be used to take handwritten or text notes while digital media is being recorded or played back. Users can take notes during (the playback of) a lecture, interview, meeting or movie. Users can be there in person or remotely via phone, teleconferencing or the Internet. At the same time, the audio/video recording or playback can occur on their palmtop, laptop, and desktop computer or remotely on an Internet server. Souvenir enables users to link their notes to the related media file wherever it is stored. Once linked, users can double click any digital ink or text to play the media at the time the note was taken.

Figure 2 provides a flowchart of how Souvenir can convert "unstructured" handwritten notes into structured data that (a) can be used to query and access a digital media, and (b) can be published as media-enabled web pages. Souvenir's functionality can be summarized as follows: 1) Users can take media-enabled notes in multiple ways: (a) the media is recorded or stored on the same device used for taking notes; thus making it easy to synchronize media and notes internally; (b) media and notes

Time-stamped Notes to be used as a Personal and Visual Index into a Digital Media Library
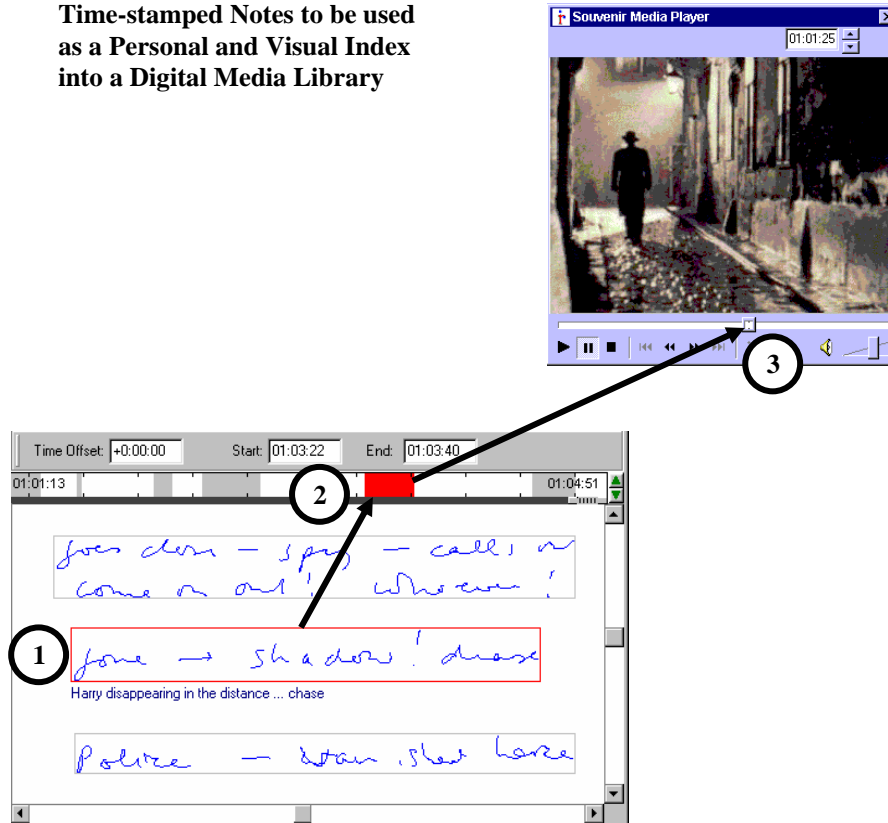

Fig 1: 1) Users take time-stamped notes, which are grouped into segments. 2) Once the time-line of the note activity has been linked to a media recording stored in digital library, a note segment references and indexes a specific media segment on the media timeline. 3) Double-clicking on a note plays the associated media segment. Handwritten notes can visualize in a personal way which parts of a media recording are of interest.

are stored on different devices, but they can be synchronized locally; (c) the notes can be synchronized with media hosted on a remote Internet server. 2) The referenced media can be uploaded to or stored in a digital media library. 3) The Souvenir Desk-top application offers users a flexible way to link and synchronize their handwritten or text notes via the network with the appropriate media file in a digital library. The Souvenir Desktop enables users to store, synchronize, annotate, edit and publish their notes created by a palmtop, laptop or desktop computer: (a) handwritten notes can be segmented and annotated with text; (b) only the text components of the segments can be displayed and shared with others. 4) Users can publish their Souvenir documents as a set of web pages. 5) Users can email hyperlinks to specific media moments di-rectly from a web page or using the Souvenir Desktop application.

Converting "Unstructured" Handwritten Notes into Media-Enabled Web Pages

1

The Third Man

Orson Welles is alive !

I believe in God.
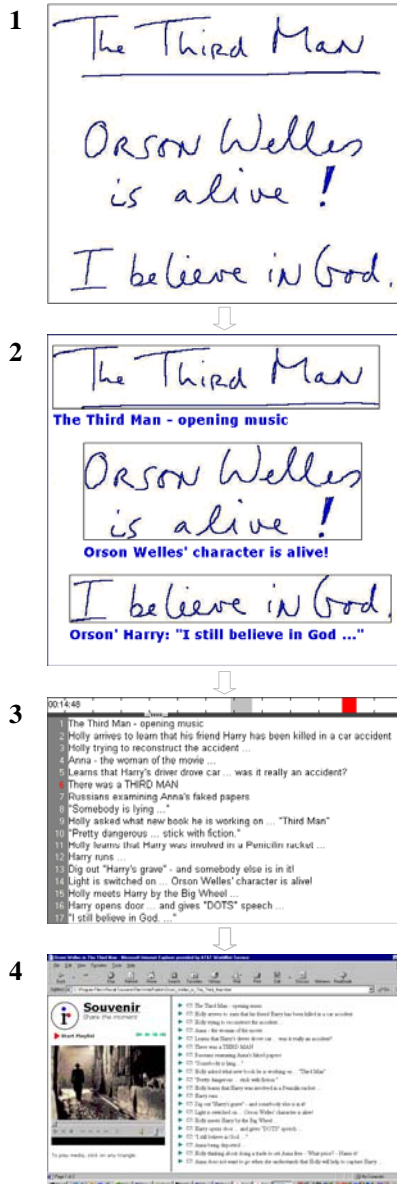
Unstructured Handwritten Notes

Handwritten notes are initially unstructured and need to be grouped to enable both digital ink and text in the same Souvenir document.

Media-enabled Notes

Souvenir time stamps the user's note-taking activity so that the notes can be synchronized with the timeline of an audio/video recording.

2

The Third Man
The Third Man - opening music

Orson Welles is alive !
Orson Welles' character is alive!

I believe in God.
Orson' Harry: "I still believe in God ..."

Segmentation of Handwritten Notes

Souvenir organizes digital ink into segments by exploiting spatial and temporal characteristics of how people write without the need to recognize the digital ink.

Text Annotation of Digital Ink

Users can annotate an ink segment, where the associated text is displayed directly below it.

Media Synchronization

Souvenir can automatically synchronize notes and media if (a) the audio/video is recorded on the same device used for taking notes, or (b) the media playback occurs in the Souvenir Media Player. In all other cases, users need to identify a specific media moment to be linked with a specific note. After that, Souvenir does the rest. Double clicking on the digital ink or text will play the associated media.

3

00:14:48

1 The Third Man - opening music
2 Holly arrives to learn that his friend Harry has been killed in a car accident
3 Holly trying to reconstruct the accident
4 Anna - the woman of the movie
5 Learns that Harry's driver drove car ... was it really an accident?
6 There was a THIRD MAN
7 Russians examining Anna's faked papers
8 "Somebody is lying ..."
9 Holly asked what new book he is working on ... "Third Man"
10 "Pretty dangerous ... stick with fiction."
11 Holly learns that Harry was involved in a Penicilin racket ...
12 Harry runs ...
13 Dig out "Harry's grave" - and somebody else is in it!
14 Light is switched on ... Orson Welles' character is alive!
15 Holly meets Harry by the Big Wheel ...
16 Harry opens door ... and gives "DOTS" speech ...
17 "I still believe in God. ..."

Text View of Notes

Users can use the "text only" view to expand their initial, quick notes to create a full report to be shared with others.

4

Souvenir
Share the moment

Web Publishing & Email Sharing

Souvenir documents can be published as a set of web pages so that others can access or share via email specific media moments directly from a web page. The media owners do not lose control over their content, because users are only creating and sharing pointers to specific media moments instead of actually copying the content.

Fig. 2. shows how Souvenir can be used to convert "unstructured" handwritten notes into structured data that (a) can be used to query and access a digital media library, and (b) can be published as media-enabled web pages.
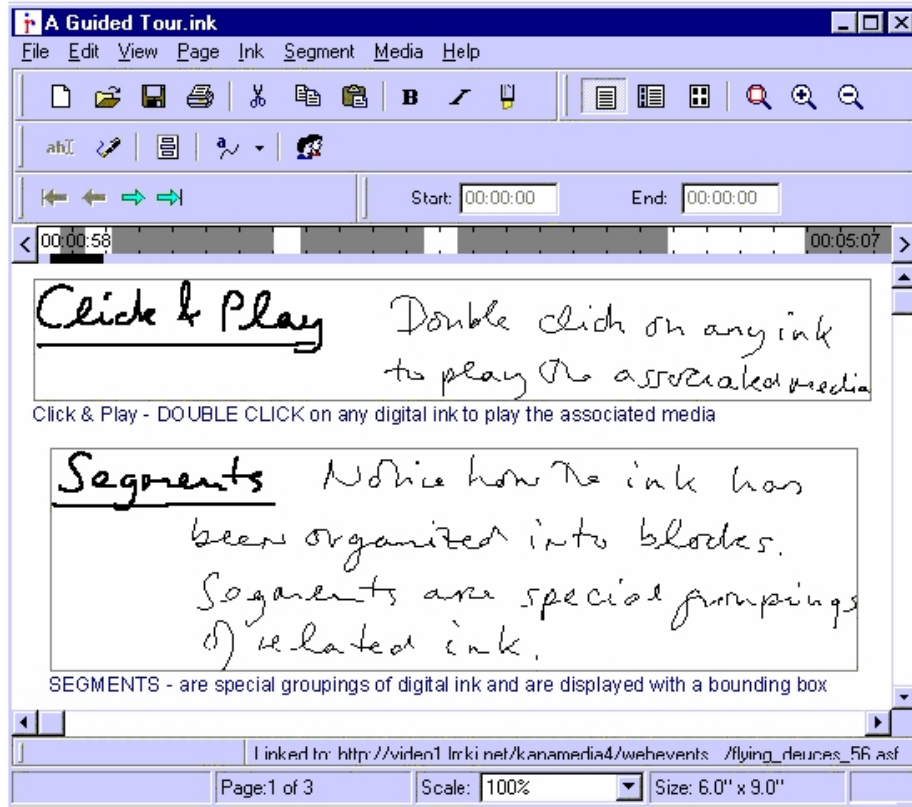
Fig. 3. The *Souvenir Document* window consists of a *media timeline* and a *document display*, which shows digital ink segments surrounded by a bounding box and the associated text directly below. Double-clicking on the digital ink or text plays the associated media. The media timeline shows how the segments, displayed in dark gray, are related to the media.

3.1 Souvenir Desktop

The Souvenir Desktop application enables users to store, synchronize, annotate, edit, publish and share their media-enabled notes that were created using a palmtop, laptop or desktop computer. The Souvenir Desktop lets users synchronize their notes via the network with the appropriate media wherever it is stored. Once synchronized, double-clicking on the digital ink or text plays the associated media. The Souvenir document window consists of a media timeline and a document display (see Figure 3). The media timeline shows how the digital ink and/or text segments (displayed in dark gray on the timeline) are related to the media. Users can interact with the media timeline to fine-tune the relationship between their notes and the associated media. The document display supports three display modes (ink only, text only, and ink & text), thumbnails and different magnifications views.

3.2 Supported Annotation Devices

Souvenir is flexible in terms of the devices that can be used to create time-stamped annotations. Souvenir currently has companion applications for the *Palm Pilot* and the *CrossPad*, a portable digital notepad that digitizes notes written on paper, to enable the capture of handwritten notes as time-coded digital ink. Notes taken with a Palm Pilot or CrossPad need to be uploaded to a PC.

The Souvenir Palm Application enables users to create time-stamped notes, where each note has three components: 1) "Digital Ink" – handwritten notes created by writing with a stylus on the Palm screen. 2) "Text" – created using Graffiti or a keyboard. 3) "Keywords" – up to six keywords can be assigned to a note.

Users can think of a Souvenir Palm note as a "post-it" note on which they can scribble short handwritten notes, add text notes on the backside and categorize it using up to six keywords at the same time. The notes created with the Souvenir Palm application are uploaded to a PC each time the user performs a HotSync and are automatically transferred to the Souvenir Desktop application.

The note-taking devices supported by Souvenir have different strengths and weaknesses in terms of creating media-enabled notes: (a) The CrossPad enables note-taking on paper, which is the preferred way for most people to take notes, and its writing area is large enough to take extensive handwritten notes. But, the CrossPad can not record or playback digital audio or video. CrossPad notes have to be uploaded to the Souvenir Desktop to be linked to the related media recording. Stifelman [6] has developed a digital notepad prototype that can record digital audio, but it could not be successfully commercialized. Further, the CrossPad and digital notepads like it have only been adopted by a small user population so far. (b) The Palm Pilot has been widely adopted because of its effective interface and small size, which in turn makes it difficult to take extensive digital ink notes. This is why the Souvenir Palm application also supports text notes. The Palm Pilot is ideal for taking short handwritten or text notes in a way that is non-intrusive during a meeting or class. (c) Keyboard-based devices are universally used. However, most people can not type as fast as they can write by hand. Users need to be skilled typists to take notes that do not "lag behind" the media moments they wish to pinpoint. (d) Tablet computers can be used to take media-enabled notes, but their current costs are prohibitive.

As the above discussion shows, there does not currently exist an "ideal device" for taking extensive media-enabled notes easily, quickly and affordably. Thus, Souvenir Desktop application has been designed to support notes created by a variety of devices, which are currently affordable and could be adopted or are being used by a large user population.

3.3 Segment-Oriented Framework

A Souvenir document is composed of *segments*, whose data structure stores: (a) digital ink and/or text; (b) the associated media file and its playback start and end times. Souvenir employs a "segment-oriented" framework to support: (a) both handwritten and text notes as well as (b) the playback of multiple media recordings in the

same document, (c) to enable users to interact with digital ink at a higher-level of organization instead of having to manage a large number of individual strokes, and (d) to create hyperlinks to random-access media wherever it is stored. These hyperlinks make it easy for anyone to share specific media moments without the need to transfer large media files. The media owners do not lose control over their content, because users only share pointers to specific media moments instead of copying them.

Souvenir is built on the premise that users' initial notes are "imperfect" or "noisy" in terms of how precisely they pinpoint the intended media moment and how completely they describe or annotate it. Souvenir aims to make it easy for users to "edit and polish" their initial Souvenir notes to create a report that can be published and shared with others via the Internet. Souvenir has been designed so that users can use their text editing skills to organize and refine their initial media-enabled notes. In particular, Souvenir organizes handwritten notes into segments so that users can interact with digital ink at a level of visual organization that is familiar and makes it easy to edit and refine their handwritten notes. Users can easily annotate digital ink segments with text or add new text and link it to a specific moment in a media recording. Users can edit and rearrange their notes as well as "copy & paste" notes from other Souvenir documents that reference different media files. Users can then publish their report as set of web pages so that others can access or share via email specific media moments directly from these pages (see Figure 2).

3.4 Media Editing

Initial notes can have a tendency to lag behind the moment in the media recording users wish to pinpoint. Souvenir makes it easy for users to correct for this "lag problem." On the one hand, users can specify a general "playback offset" that is added or subtracted from the "start" and "end" times of all segments in a Souvenir document. On the other hand, users can fine-tune the relationship between a specific segment and its associated media. Users can edit and change the "start" and "end" media times of a segment via edit controls in the document window or by interacting with the media timeline. The media timeline shows how the ink and text segments are related to the timeline of the associated media recording(s). The segments are displayed in dark gray on the media timeline (see Figures 1 and 3). Users can change the start and/or end time of a segment by changing its spatial area in the timeline.

4 Structuring Handwriting

Handwriting is ideal for taking quick notes. However, handwritten notes are "loosely structured" whereas text notes are highly structured. Hence, handwritten notes need to be structured to enable both digital ink and text in the same Souvenir document and to be an effective index for digital media. Furthermore, people need to be able to interact with digital ink at a level of visual organization that is easy and familiar. People most commonly interact with text at the word, sentence or paragraph level,

where the latter two units of organization are used for attaching annotations in text documents.

At the lowest level of organization, handwritten notes are a collection of strokes, where each stoke represents the pixels touched between successive "pen-down" and "pen-up" events. Figure 5(b) shows a collection of strokes that are surrounded by rectangular bounding boxes. Human handwriting is structured and governed by constraints. People (in the West) have a tendency: (a) to write consecutive or related words spatially close together; and (b) to create straight lines of text. Further, people have a tendency to write a few (lines of) words describing their current thought, to pause for a while and then to start a new thought or to add more to the current or to a previous thought. These constraints can be used to create a simple, yet robust digital ink segmentation algorithm, which assumes that users create lines of text and write top to bottom (see Figure 4). However, the algorithm makes no assumptions about the direction, orientation or line height of the hand-written text and can adapt to the user's current writing style. The algorithm segments digital inks into units of organization that are equivalent to text paragraphs (see Figure 5(c), 5(d), or 5(e)). Users can now interact with their handwritten notes in a familiar way, where the notes are organized in such a way to make it easier to annotate, edit or refine them.

However, digital ink in close spatial proximity is not always created close in time. This fact matters because we want to use the temporal properties of handwritten notes to index audio and video. Without access to the creation times of the digital ink, it is very difficult to detect that notes spatially close have not been created close in time. Souvenir uses note-taking devices, such as the CrossPad or Palm Pilot, that can automatically time-stamp handwritten notes. One of the goals of Souvenir is to make visible when digital ink was added later in time in the vicinity of already existing ink. Figures 5(c) and 5(d) show digital ink in close spatial proximity that has been detected by Souvenir as distinct segments because a certain amount of time elapsed without note-taking activity or some of the digital ink was added later in time.

The purpose or function of a note in relation to a media file can be: (a) to just comment on a specific moment in time; (b) to describe a limited time period in a media recording; (c) to fully capture the spoken content in a one-to-one relationship with the written notes, as is the case in closed captioning, but which can only be created by a specialist; or (d) to be an addition to an existing note that pinpoints a media moment other than the current media time. Now, it takes a certain amount of time to create a note and the note-taking timeline is linear and monotonic increasing. Additional information is needed to determine the specific purpose of a note in relation to the media timeline. For example, if the state of the media recording or playback device is in "pause" or "stop," then the written words are most likely to describe or comment on the current media time. Souvenir takes into account this piecewise linear relationship between media playback activity and simultaneous note-taking activity when synchronizing Souvenir notes and media (see Figure 6).

As the above discussion shows, structuring digital ink as well as synchronizing it to the appropriate media moments can be difficult. Souvenir offers users a seamless and flexible way to structure and synchronize their handwritten notes with digital media wherever it is stored.
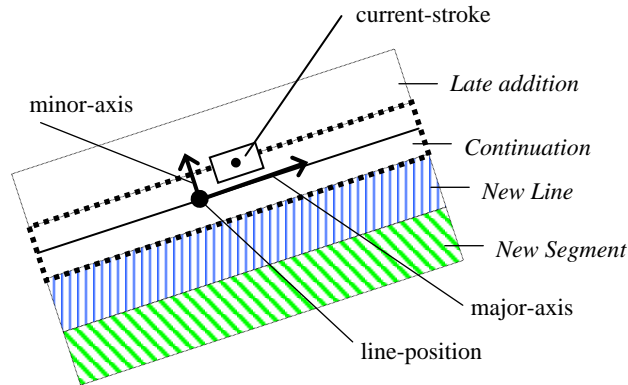
Fig. 4. The *Souvenir segmentation algorithm* computes the line-position, major-axis and minor-axis for the current-line to determine if the current-stroke lies in the "Late Addition," "Continuation," "New Line," or "New Segment" spatial band.

4.1 Digital Ink Segmentation

Souvenir automatically groups the digital ink into segments to identify interesting entry points into a media recording. The digital ink segmentation algorithm exploits spatial and temporal characteristics of how people write without having to recognize their handwriting. The handwriting activity can be categorized as follows: the stroke just being written (a) is part of the "current segment" of strokes or words, or (b) forms the beginning of a "new segment," or (c) is a later addition to an "old segment." Souvenir assumes that users write "top to bottom" so that a later addition represents a "backwards motion" up the page. Souvenir makes no assumptions about the writing direction and computes a "current-line" estimate so that "dotting the i's, crossing the t's and underlining," which often represent a backward motion of the pen, do not get categorized as a later addition to an existing segment.

The Souvenir segmentation algorithm continuously calculates an estimate of the line currently being written to determine if the stroke next in time signals the start of new segment. In particular, it computes the line-position, major-axis and minor-axis for the current-line. The line-position is equal to the weighted averages of the x and y coordinates of all the strokes assigned so far to the current-line. The algorithm also computes major and minor axes based on the spatial distribution of the strokes so far assigned to the current-line so that the variance is maximal and minimal in the direction of the major-axis and minor-axis, respectively. The orientation of the current-line is equal to the orientation of the major-axis. The current line-height is proportional to the length of the minor-axis. The estimate of the current-line adjusts as more strokes are added. The handwriting segmentation algorithm loops through all the digital ink strokes in order of creation. The current-stroke is grouped with the current-line if it lies within a spatial band defined by the estimates of the major-axis and minor-axis for
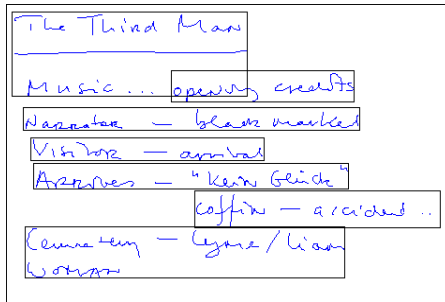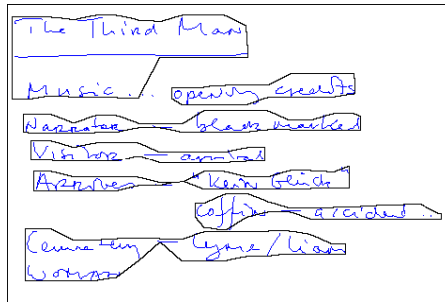
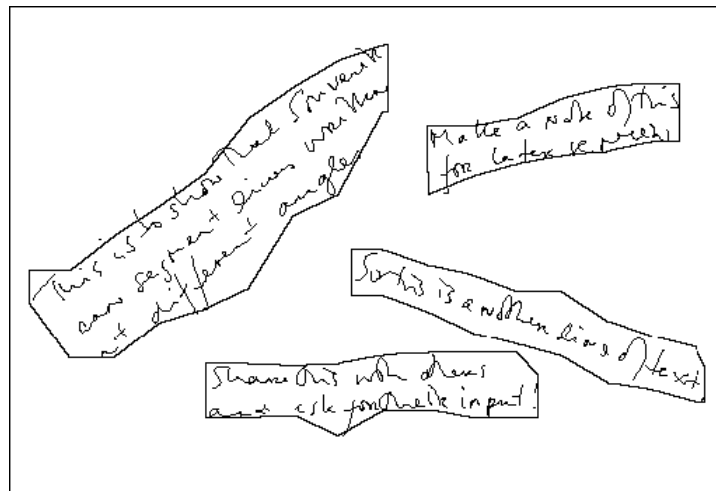# Digital Ink Segmentation
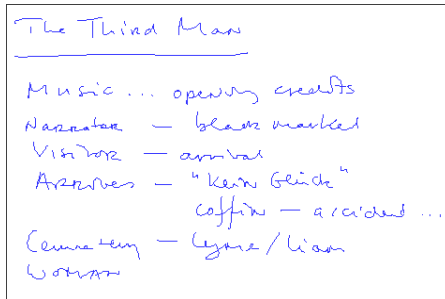
(a) Digital ink.
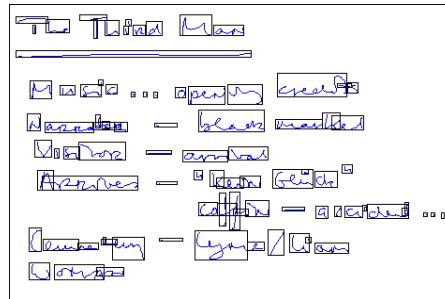
(b) Individual strokes surrounded by bounding boxes.

(c) Detected segments with rectangular bounding boxes.

(d) Detected segments with minimal area boxes.

(e) Detected segments for words written at different angles in the same document.

Fig. 5. shows how Souvenir can segment and structure digital ink that initially consists of a collection of strokes. In (c) and (d) "opening credits" is detected as a separate segment because of a significant time gap in note-taking activity.

126

the current-line (see Figure 4). Otherwise, the current-stroke forms the seed for the estimate of a new current-line. If the vertical distance between the new current-line and previous current-line is greater than a multiple of the current estimate of the line-height, then a "new segment" is created. If this vertical distance is negative, then the new segment is categorized as a "later addition." A "new segment" is also automatically created if a certain amount of time has elapsed without note-taking activity.

Figure 5 shows how Souvenir can segment handwritten notes that initially consist of a collection of strokes: (a) displays the digital ink without segmentation; (b) shows the individual strokes, which represent the pixels touched between successive "pen-down" and "pen-up" events, surrounded by bounding boxes; (c) shows the segments detected by Souvenir and displays them with rectangular bounding boxes - "opening credits" is detected as a separate segment because of a significant time gap in note-taking activity; (d) shows the detected segments and displays them with minimal area bounding boxes to make the individual segments easier to see; (e) demonstrates that Souvenir can segment text written at different angles in the same document.

Once Souvenir has segmented the handwritten notes, users can easily edit the detected ink segments by merging, splitting, regrouping or annotating them manually. For example, users can split a segment by selecting it and moving the cursor inside of it - the segment is then split in two based on the cursor position in an interactive way.

4.2 Multiple & Complementary Indexes for Digital Media Libraries

Audio/video content can be described at multiple levels of organizations to facilitate the creation of a retrieval index. There are major research efforts to automatically identify the content of individual image frames or movie sequences [9]. The media index created by user annotations represents an additional and complementary way to index digital media. On the one hand, human media annotations are time consuming to create and have required so far specialized software tools to produce. On the other hand, the user annotations represent "value-added" information that can not be easily inferred by automatic techniques. Souvenir aims to "liberate" the creation of user media annotations: (a) a variety of inexpensive devices and tools can be used to create them; (b) the media-enabled annotations can be leveraged without being held captive by where and how the media is stored in a digital library; (c) users can use their time more productively, because they get a personalized media index "for free" when they take Souvenir notes while experiencing digital media; (d) the handwritten notes can visualize in a personal way which parts of a media recording are of interest.

5 Note & Media Synchronization

Souvenir time stamps the user's note-taking activity. It uses these time stamps to link and synchronize the notes with the timeline of a media recording. This way the notes can be used to access and playback specific moments in a media recording. The synchronization of the notes and media timelines is conceptually straightforward:
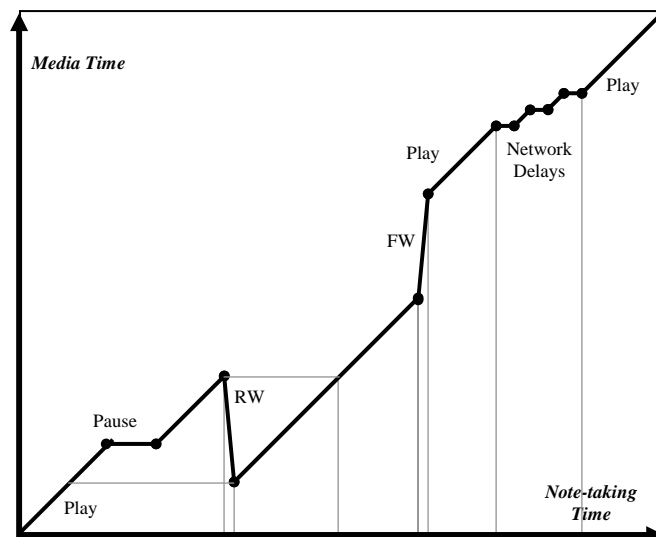
Fig 6: shows the piecewise linear relationship between media playback and simultaneous note-taking activity that needs to be considered when synchronizing notes and media.

identify a specific media moment to be linked with a specific note and the resulting relationship between their respective time-codes can be used to compute the media playback offsets for all the other notes. This presupposes that the recording or play-back of the media is a strictly monotonic, linear event. However, users can pause, play, rewind or fast forward the media while taking notes. If the media is delivered via a network, playback delays can also occur.

The Souvenir Media Player keeps track of when a user plays, pauses, rewinds or fast-forwards a media recording. This player, which can play local as well as Micro-soft or RealNetworks streaming media, also monitors the network performance if the media is streamed. Souvenir can use this database of media playback activity to link the user's note-taking activity with the appropriate media moments.

Souvenir automatically performs the synchronization of notes and media if (a) the audio/video is recorded or played back on the same device used for taking notes, or (b) the media playback occurs in the Souvenir Media Player. In all other cases, a wizard will guide users through a simple process where they need to identify a spe-cific media moment to be linked with a specific note. After that, Souvenir completes the synchronization of notes and media, taking into account the piecewise linear rela-tionship between media playback activity, network performance and simultaneous note-taking activity (see Figure 6).

Souvenir is built on the premise that users want the flexibility to use different de-vices to capture and playback digital media. However, users will want to be able to share their media-enabled notes with others. Hence, they will want to upload or store the referenced media in a digital library that accessible to others via the Internet. Souvenir makes it easy for users to update or change the location information of the media recordings referenced in a Souvenir document.

128

6 Evaluation & Discussion

Souvenir has been downloaded by quite a few people. Informal tests of Souvenir's digital ink segmentation algorithm are encouraging (see Figures 1, 2, 3 and 5). The algorithm can segment notes created by different note-takers without the need for any training or special instructions. However, a formal evaluation is needed and planned to test the effectiveness of the segmentation algorithm more rigorously. Considering the high variability of human handwriting, the algorithm tracks simple properties that are robust enough to detect segments that are useful and match how the note-takers would visually parse their notes. As mentioned, handwritten words in spatial proximity can be assigned to different segments when they were not all created at the same time. This can be "visually counter-intuitive" until the user takes into account that the timeline of how the notes were created is not always easy to see.

Souvenir offers users a flexible mechanism for linking notes and related media. In the "worst case," users have to identify a specific media moment to be linked with a specific note and then Souvenir does the rest. However, general users will adopt Souvenir only if the synchronization is automatic and transparent. So far users have had to use specialized tools to pinpoint and share specific media moments. Using Souvenir, users can create a media "edit list," which stores the start and end times of media clips, without having to learn new skills. Souvenir has been designed to enable users to use their note-taking and text editing skills to manipulate digital media without having to be concerned about where and how the referenced media has been stored.

Souvenir makes it easy for anybody to bookmark specific moments in streaming media hosted on the Internet. However, some media sites make it difficult to identify the URLs of their media content to prevent "deep linking." Having access to the URL makes it possible to play the streaming media in the Souvenir Media Player, which in turn makes it easy for Souvenir to automatically link the notes with the related media. The reluctance of media sites to make the media URLs easily available is a pity, because Souvenir encourages its users to only create and share pointers to specific media moments instead of actually copying the content and media owners losing control over their content. Actually, the Souvenir media annotations could be used for community building and data mining purposes. The Souvenir notes make it possible to access and share specific media moments without the need to view the entire recording, resulting in a more efficient use of server resources. Content providers can use the personal media annotations and the resulting targeted access of their content to visualize frequently accessed moments.

7 Conclusion

Souvenir is a versatile note-taking tool to pinpoint, edit and publish audio/video highlights stored in digital libraries. Souvenir is innovative in that it offers a comprehensive solution for users to use their handwritten or text notes to create a visual and effective audio/video index. Souvenir gives users flexibility in how they can access

and integrate digital media moments into their workflow. Users can take media-enabled notes on a variety of devices, such as the paper-based CrossPad, the Palm Pilot and standard keyboard devices. Souvenir can organize digital ink into segments, which are detected using a segmentation algorithm that exploits spatial and temporal characteristics of how people write without having to recognize their handwriting. Souvenir employs a "segment-oriented" framework to support both handwritten and text notes as well as the playback of multiple media recordings in the same document. Souvenir is flexible in terms of how to synchronize notes and media, taking into account the piecewise linear relationship between media playback and note-taking timelines. Souvenir also has web publishing and email capabilities to enable anyone to access or share specific media moments directly from a web page.

8 Acknowledgement

References

1. Abowd, G. D., Atkeson, C. G., Brotherton, J. A., Enqvist, T., Gulley, P., and LeMon, J. "Investigating the Capture, Integration and Access Problem of Ubiquitous Computing in an Educational Setting," Proc. of CHI '98 (April 1998), ACM Press, 440-447.
2. Brotherton, J. A., Bhalodia, J. R. and Abowd, G. D. "Automated Capture, Integration, and Visualization of Multiple Media Streams," Proc. of IEEE Multimedia '98, July, 1998.
3. Chiu, P. and Wilcox, L. "A Dynamic Grouping Technique for Ink and Audio," Symposium on User Interface Software and Technology '98, (San Francisco, Nov 1998), 195-202.
4. Cruz, G. and R. Hill. "Capturing and Playing Multimedia Events with STREAMS", Proc. of ACM Multimedia '94, October 1994, pp. 193-200.
5. Moran, T. P., Palen, L., Harrison, S., Chiu, P., Kimber, D., Minneman, S., van Melle, W., and Zelweger, P. "`I'll Get That Off the Audio,': A Case Study of Salvaging Multimedia Meeting Records," Proc. of CHI '97 (Atlanta, April 1997), ACM Press, 202-209.
6. Stifelman, L. J. "Augmenting Real-World Objects: A Paper-Based Audio Notebook," Proc. of CHI '96 (Vancouver, Canada, April 1996), ACM Press, 199-200.
7. Truong, K. N. and Abowd, G. D. "StuPad: Integrating Student Notes with Class Lectures," Proceedings of CHI '99, (Pittsburgh, May 1999), ACM Press, Volume 2, 208-209.
8. Weber, K., and Poon, A. "Marquee: A Tool for Real-Time Video Logging," Proc. of CHI '94 (Boston, April 1994), ACM Press, 58-64.
9. Wactlar, H., Kanade, T., Smith, M., and Stevens, S. "Intelligent Access to Digital Video: The Informedia Project." IEEE Computer, Vol. 29, No.5, May 1996.
10. Whittaker, S., Hyland, P., and Wiley, M. "Filochat: Handwritten Notes Provide Access to Recorded Conversations," Proc. of CHI '94 (Boston, April 1994), ACM Press, 271-277.
11. Wilcox, L., Schilit, B. and Sawhney, N. "Dynomite: A Dynamically Organized Ink and Audio Notebook," Proc. of CHI '97 (Atlanta, April 1997), ACM Press, 186-193.