

Contextual Note-Taking for Family History

Mitchell K. Harris, Dan R. Olsen Jr.

Brigham Young University
3361 TMCB PO Box 26576

Provo, UT 84602-6576

Tel: 1-801-422-7698

mitchell.k.harris@gmail.com, olsen@cs.byu.edu

ABSTRACT

Microfilms readers present a rich source of family history information. The paper and pencil method that most researchers employ to take notes is inadequate. Handwritten notes become difficult to modify and organize, are time consuming to record, and gives no context about the actual information found. In this paper we explore the flaws in the classical methods of note-taking and introduce a new method that involves image capturing and digital storage. This new solution allows the researcher to focus entirely on the task of information gathering rather than recording.

INTRODUCTION

Family History research is often laden with frustrating distractions that slow search. A few of these distractions are transcribing and organizing notes, error correcting, and rechecking facts previously collected. Information gathering is the goal of research and should be the central task; however, often the researcher spends their time painfully copying records from slides or rearranging their notebook. While these other tasks are currently necessary, they should be reduced or replaced with simpler methods. Then time spent researching would be undistracted and spent more productively.

Using computers to eliminate secondary tasks such as organizing and copying information can make research faster, easier, and more focused. However, the simple addition of a computer may not always help. We propose a new approach that uses a high resolution digital camera to capture and record images from microfilm readers and then stores them digitally. With this camera-computer system the researcher's time can be focused on research as the computer and camera handle many of the secondary tasks.

THE PROBLEM

While most researchers have varying techniques of gathering information, almost all use the same method of recording it: transcribing the information by hand into a

notebook. This method is plagued by several chronic problems:

1. Lack of organization, searchability, and modifiability
2. Wasted time copying data
3. Loss of context

The first chronic problem is that notes recorded in a notebook tend to become disorganized and messy. Microfilms are often organized in a different manner than the researcher wishes to organize his notes. As a result, while gathering information a researcher can do one of two things: either spend a great deal of time and effort trying to find an appropriate place to record the information he has discovered, or simply write the notes continuously in the order that they are found. While the first is more organized, it requires more effort and is likely unscaleable. The second gives no organization at all to the information and makes searching impossible. Both methods are inherently brittle to future additions. Revisions become confused and untidy and thus hard to comprehend and search. The recorder would have to remember where all of his information was stored; a daunting task considering the volumes of data gathered by most researchers. Notebooks become an inflexible and clumsy tool for organizing and collecting family history information.

In addition to the inadequacies of notebooks as an information organizing tool, notebooks are also insufficient as an information collection tool. Notebooks and other paper dependant mediums are expensive in time and effort to use. When a researcher finds a record that they believe may be worth recording he then copies down the entire set of information associated with the record. Since it is not known what information may prove useful in the future no information may be skipped or deemed unimportant. If information is skipped, the time cost of reloading the reel and scrolling to the record again is exorbitant. As a result, most researchers exhaustively copy down all data found in a record.

Often the researcher needs to continually look back and forth at the film and their notebook as they transcribe the information. It is not uncommon to see a frustrated researcher place a book or finger on the reader base board to keep their place while they are looking between the reader and his notebook. This causes fatigue on the researcher

This space has been kept available for the copyright notice

and a sense of futility. One feels that the vast majority of their efforts are probably unimportant and as a result is less inclined to record records while searching. If the researcher runs across a record that could be relevant, but possibly not, there is a brief battle of wills as the recorder internally debates if it is worth pausing for several minutes to record the entry. Much information that is useful is skipped because it is deemed too expensive for its value. Even more information that turns out to be entirely useless is recorded and therefore time is wasted. Writing information by hand turns out to be a demanding task that distracts from the desired goal: information gathering.

Another problem that interferes with information search is lack of context: information embedded in a frame that cannot be copied. All information that the user wishes to have must be copied down. However, it is nearly impossible to copy down the entire frame, and as a result information is invariably discarded. This lost information is a loss of context.

An excellent example of this loss of context is the way words are written. Most researchers have difficulty reading the script on microfilms from handwritten sources. As a result when they take notes they are often forced to make a guess about the spelling of a particular name or if they're seeing a 7 instead of a 1. Later, when the recorder is reviewing their notes, they cannot remember alternate possible spellings or compare the handwriting to other names that they have written. This lack of context can send the research into many dead ends, looking for names perpetuated by mistaken recordings or other transcription errors. Rarely, if ever, will the researcher return to the original slide to check for alternate spellings.

An effective note-taking method would need to solve these issues. It would need to:

1. Allow for easy and searchable data that can be re-organized and modified at a later date.
2. Be inexpensive to take notes thus making searching the primary task of research, not recording.
3. Store contextual information about the frames, but be able to separate the data and the context.

A Partial Solution

The simple act of introducing a computer as a device for information gathering addresses several of these issues. Currently many users chose to forego the troubles of a notebook and bring a laptop computer along with them. They place it on the desk by the film reader and write their information in the application of their choice. Since the digital medium is easily modified the first chronic problem of microfilm research is solved. The researcher can record their information in the application of their choice and at a later date modify it, delete unwanted records, append more to already explored records, and search the records that they've created in the past. The simple addition of a personal computer makes the organization process of gathering information much more flexible.

However, the other issues remain unaddressed. The computer still stores no contextual information about the source frames. Computers also can be harder to take notes with than a notebook. Those that do not know how to touch-type find that entering data into a computer is much slower than simply writing it out. However, touch-typing may be even worse. Because the researcher uses both hands they cannot place a finger on the reader base board to keep their place as they transcribe. The researcher finds herself constantly glancing back and forth from the computer and the reader trying to see what they have written and what is left to write. Even worse, the computer can only be put in a small space by the reader. Since the reader is quite large the typist needs to twist at an awkward angle to get both of their hands on the keyboard and still be in front of the reader and have a good view of the frame. The result is an awkward exercise that deters many from using the computer at all.

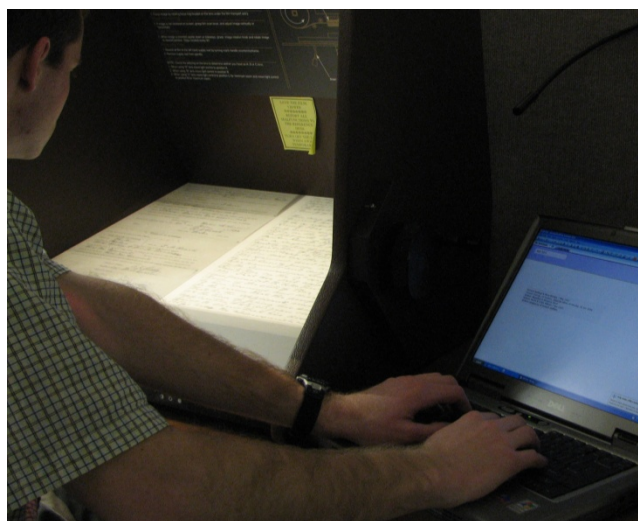


Figure 1: A researcher using a computer to record is forced to twist at an awkward angle to touch the keyboard with both hands.

A MORE COMPLETE SOLUTION

The solution we propose offers a unique way of making a computer far more useful by overcoming all three of the chronic problems of microfilm research. We propose that by using a high resolution digital camera in tandem with the computer, the user can experience the organization, scalability, and modifiability that digital medium offers and provide fast, easy, contextual information storage as well.

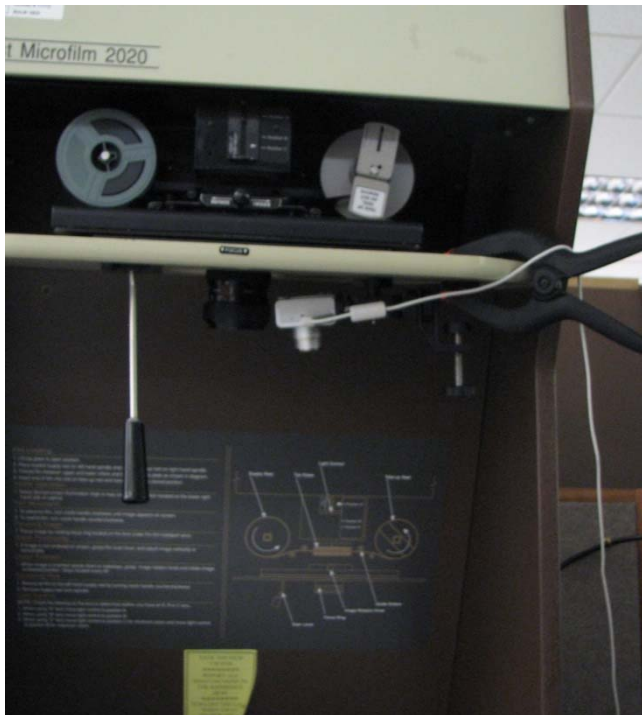


Figure 2: The camera-computer system attached to a reader

Attaching the camera to the reader and getting images to the computer is easy. As seen in Figure 2 an off-the-shelf digital camera is screwed onto a typical camera mount. That mount is attached to the top of the reader with a spring loaded grip (Figure 3). Ideally the camera should be placed directly adjacent to the reader's lens running length-wise parallel with the side of the reader. The camera is connected to the computer via a USB cable. The computer runs an application that is capable of controlling the camera. We have developed a note-taking application for the purpose of demonstration.



Figure 3: A camera mounted adjacent to the reader lens

Our note-taking application has two purposes. The first is to act as an integration layer between the computer and the camera. The second is to allow the user to create, organize, modify, and annotate the notes and images taken with the camera in a form intuitive to family history researchers. Between the recording power added by the camera, and the

flexibility of the application we have developed, research becomes much less tedious and more search oriented.

The Software

The application allows the user to connect a variety of Canon digital cameras to the computer. Once the camera is connected the user can get the live viewfinder feed from the camera on the computer. This allows the user to see a preview of what images the camera will take and center the frame (Figure 4). The application is also capable of rotating the image if the camera is at an inconvenient angle as well as zooming and changing resolutions.



Figure 4: The viewfinder displays a low resolution preview of the image

After starting the application and configuring the camera, the researcher begins by creating a new task. Tasks are used to separate research into different areas. For example the researcher may create a new task titled "Find Daniel Heaton's children". After a task is added the researcher begins his search. Whenever the researcher changes reels he can place the new microfilm's box under the camera and press the "New Source..." button. The camera will snap a shot of the box and the researcher can continue scrolling through the film. By snapping the image the reel number and any other contextual information about the reel is stored in image format. The image will appear in a navigation bar to the right of the display area with an icon of a microfilm reel indicating that it is an image of a source.

Since working with the entire image can be clumsy, a way of making the image more useful has been developed. Once an image is snapped the researcher may add a "rubber band" to the image. A rubber band is a red bounding box around the portion of the image considered important. For instance a source image may have a rubber band surrounding the roll number on the box as seen in Figure 5. The rubber band is automatically zoomed on when resizing the window and draws the eye to the portion of data being recorded. If the researcher desires he may add other metadata

to the source image. However, none of this metadata need be added immediately at all, the researcher may continue on immediately to searching.

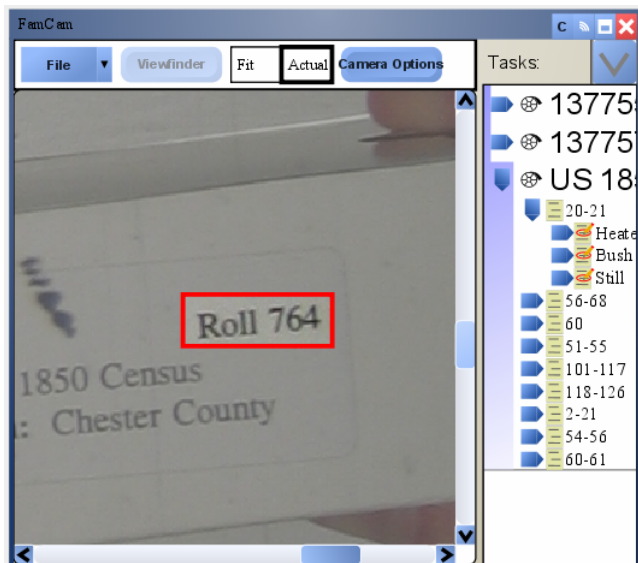


Figure 5: A source's roll number is emphasized by a rubber band

When a frame is found with information that the researcher wants to record, the researcher presses the “New Image” button to snap an image of the frame. That picture is automatically downloaded from the camera to the application. It is added to the navigation bar under the source that the image belongs along with an icon indicating it is an image. Again, if the researcher desires he can annotate the image then or choose to do so later and can immediately continue his research. Research becomes a cycle of snapping shots of frames of microfilm and continuing without having to pause and copy any information.

Later, when the researcher wishes to return to any of the images he has collected, he need only click on the icon of the image and it is shown in the display area. If any part of the image has been indicated with a rubber band, that part of the image will be centered on the display. In addition, to make the information on the image more accessible, for every interesting piece of information on the image the user can add a record. The record will appear underneath the image with another icon. For example if an image has 3 people of note on it, the research may add three records, one for each person. Each record also has a list of facts. The researcher can add a number of facts to the record including:

- Given Name
- Surname
- Date of Birth
- Date of Death
- Many other family history facts

Every fact can have a value that the researcher may enter.

For example Figure 6 shows a record with two facts. The first is that the record's given name is Daniel, and the record's surname is Heaton. Both of these facts belong to the record above it indicated by the icon of a paper with a red mark and a pencil. If the Heaton icon were to be clicked the screen would change to look like Figure 7

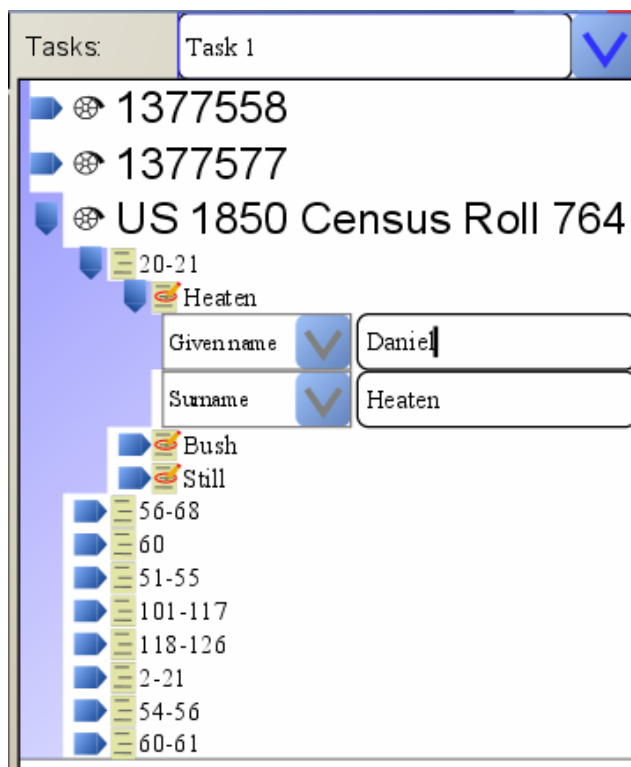


Figure 6: The navigation bar showing different parts of the image. In order from top to bottom are three separate images with the last having nine images. The first image is expanded to show three different records, the first having two facts listed.

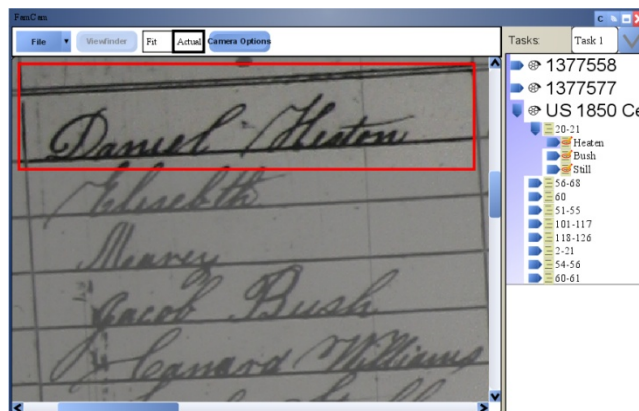


Figure 7: The rubber banded area is shown immediately when a record is clicked on.

Once metadata has been recorded it may be saved and exported. Although not currently implemented, it would be simple to add exporting to GEDCOM or other formats. The power of this application is that the images, and therefore the exact information, are preserved. However, their data can be easily denoted and extracted with records, facts, and rubber bands.

Overcoming the Chronic Problems of note-taking

The overall result is a system where taking notes becomes cheap. This system overcomes all of the chronic problems of note-taking with a notebook. The organizational, searchability, and modifiability problems have been addressed because the information is stored on a digital medium. However, adding the camera overcomes problems 2 and 3 (those associated with the time-cost of note-taking and of lack of contextual notes).

Previously in this paper we discussed as the second chronic problem of note-taking: copying notes via computer or by hand is an expensive operation. This new system is much easier. Taking notes becomes inexpensive. In order to get all of the information content from a frame the user need only press a single button and the frame is captured by the computer. One of the most useful advantages is that the image need not be annotated immediately. Instead of interrupting research to copy over all of the information, the researcher can merely snap a shot and move on. Later when the researcher is reviewing his notes he can annotate the images he feels are important. Because capturing frames takes mere seconds compared to the long process of transcribing all the information from a frame to paper by hand, the psychological barriers, that battle of wills over cost and benefit of pausing for taking notes, are eliminated. When a researcher sees a frame that may or may not be useful he can easily assume a, "may as well grab it" attitude because it is easy. Being a prolific researcher becomes natural. Frames that before would have been ignored can now be captured, and if those frames prove not to be useful they can be deleted without any significant amount of time wasted.

An interesting observation can be made about the difference in work styles of those using notebooks and those using the camera-computer system. These observations show that users of the camera-computer system can focus more completely on searching the microfilm than those that use notebooks. Notebook users tend to position themselves between the reader and their notebook (see Figure 8). This is because it is the most comfortable position to be in while transcribing from the reader. However, those that research with the camera tend to place themselves entirely in front of the reader (as seen in Figure 9) and barely pay attention to the computer except to reach over from time to time and click a button.

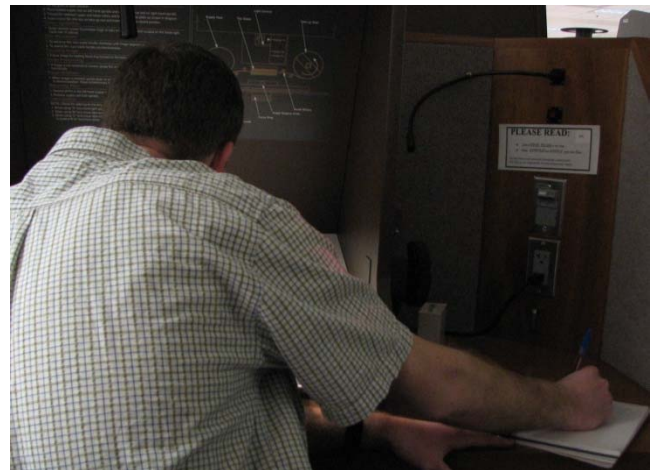


Figure 8: A researcher using a notebook must wedge himself between the reader and his notebook to be able to access both

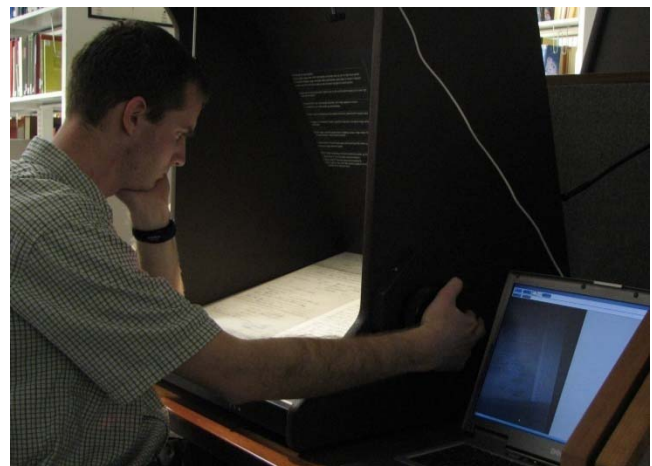


Figure 9: The user is able to comfortably place themselves fully in front of the reader.

It is easy to see that this researcher is able to devote himself entirely to the task of searching for information rather than transcribing it. The researcher in Figure 8, however, had to adopt a more awkward position because his primary task was not entirely search, but transcription. The camera-computer solution allows for the researcher to be dedicated to the activity of research and be un-interrupted by copying.

Another valuable enhancement is contextual note-taking. Where both the note book and the computer-only approach had notes that stripped away information, the camera-computer approach captures the context of the information. Images of the actual microfilm hold all of the contextual information that the user needs. Most frames have frame numbers, page numbers, or other locators that allows the user to store where the information came from. That information would be captured along with the entire frame for every note that the user makes with this system.

Earlier we described a situation where a researcher had to guess the spelling of a name as they recorded it in their notebook. With images the user can easily return and check if another spelling is possible or compare it with another name found in a different frame. Error checking and side by side comparison makes the retention of context useful. It is easy to imagine a situation where a researcher mistranscribes a record and spends hours in dead-end research looking for the wrong spelling of a name or an incorrect date. In contrast with images, researchers can quickly find a fact; they just click on the record and look at that part of the frame again. The researcher never needs to reload the reel and find the frame again.

Contextual storage of information is likewise important when the researcher wishes to collect more information from a record than they originally planned. Where the classical technique to note-taking forced the researcher to either record all information that might prove useful, or be forced to have to reload the reel at a later date, the camera-computer solution captures the entire frame and stores it very easily for later retrieval. The user does not need to be at a microfilm reader to check for more information about a frame. It is trivially easy to reload the image on the computer and gather any more information that the user may desire.

CONCLUSION

We introduced a new style of research that combines a digital camera with the microfilm research process. This style makes record capturing for personal use almost effortless. As a result researchers can collect more information than they need, and filter through it at a later time. Instead of hours spent in painstaking note-taking, users focus their time on collecting the information they are interested in quickly and continuing on. The camera-computer technique makes the researcher truly a researcher, not a scribe.