Capstone:
Gather Valuable Information Found on Headstones as an Automated Process

Cameron Christiansen
cam@byu.edu

William A. Barrett
barrett@cs.byu.edu

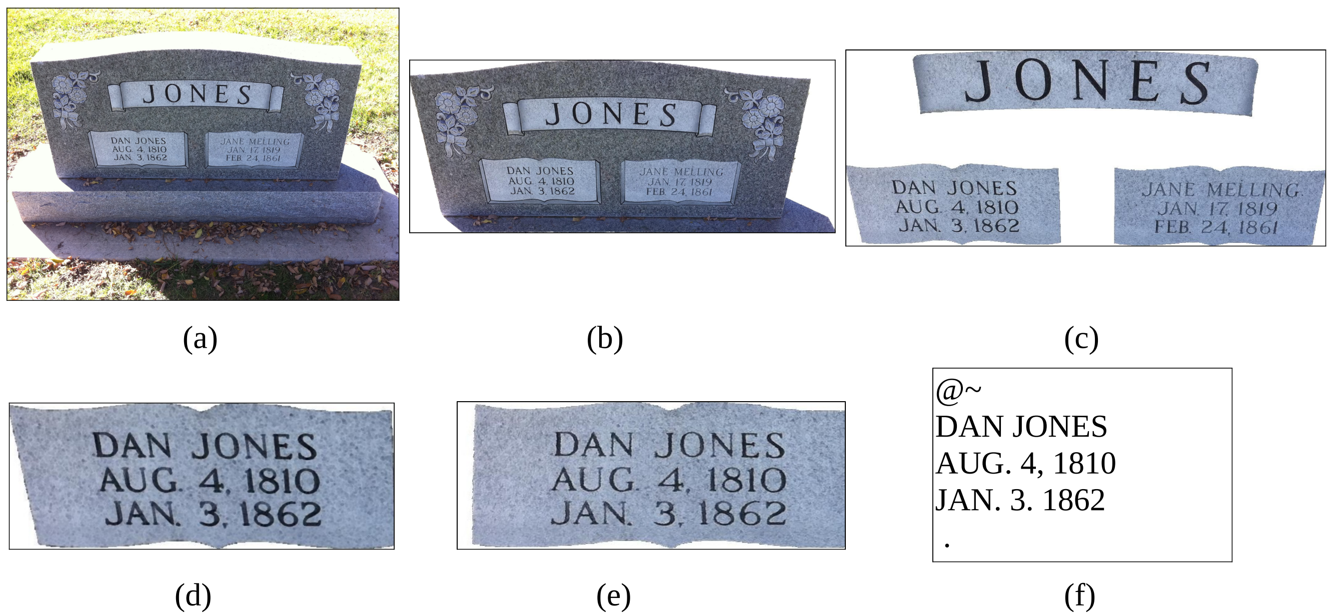


Figure 1. (a) Original image of headstone (b) headstone segmented from background using graph cut (c) segmentation of inscription areas (d) filtered image of text area (e) manual text rectification of text area (f) OCR output.

Abstract

Cemetery headstones contain historical and genealogical information that is largely untapped and is difficult to access. We introduce a solution through computer vision and OCR that allows the user to take a photo of a headstone with a GPS smart phone, review and correct, if necessary, the regions recognized as text, and have the text, images, and GPS location from the headstone extracted, recorded, and uploaded in digital form. A graphical link is then made to a cemetery map for browsing, searching and making the information accessible on the web. A segmentation of text is done through interactive graph cuts, as described in [1]. After image filtering and text rectification, the image of the text is then passed to an OCR engine for text recognition.

1. Introduction

Cemetery headstones contain a wealth of genealogical information that is largely untapped, unindexed and inaccessible. Such information could provide not only the vital birth and death information, but, in many cases, links to family relationships and clues as to where they lived. Significant efforts have been made to capture information from headstones and make it search-able and available on-line. For example, “Find A Grave” (http://www.findagrave.com/) has indexed over 50 million names worldwide, along with pictures of the headstone and/or cemetery, and sometimes the individual and their biographical information. However, time must be spent in capturing the photo, recording the cemetery and plot location, transcribing the information on the headstone, importing the photo and information onto a computer, then using a browser to upload the information to the web site. This is insufficient for the number of headstones that remain unindexed due to the time consuming process described and given the non-decreasing number of deaths (~ 2.5 million per year in the U.S.). In Capstone, the process will be reduced to capturing the photo with minimal correction by the user.

In addition to the increased difficulty and impracticality of continuing to enter and upload headstone data by hand, the amount of headstone information available on-line is only a fraction of what could be available. Careful estimates suggest that there are over 1 billion headstones in the U.S. alone. Furthermore, the data contained on these headstones continues to deteriorate each year and may be the only source of this information for 75% of the headstones. Clearly, a more automated way of capturing, transcribing and uploading this information would have much to offer.

The physical location of the headstone is also valuable information. The above mentioned "Find A Grave" and "Names In Stone" (http://www.namesinstone.com) provide GPS coordinates for a majority of their recorded cemeteries. However, due to inconsistent plot mapping among cemeteries, the GPS coordinates of the headstone itself becomes valuable.

To meet this need, we propose Capstone, a hand-held technology for capturing, OCR-ing and uploading the genealogical data, image, location and associated information found on and near a headstone. Current mobile phone offerings, such as the iPhone make this realizable due to increased resolution and focusing capability of the camera, processing power and memory on the phone, along with GPS and wireless access to the Internet. Design criteria anticipate that the Capstone technology will allow capture, recognition (OCR), recording and queuing the upload of headstone data as the operator moves from one headstone to another. This will allow the operator to verify the accuracy and sufficiency of what is captured and transcribed.

2. Related Research

We first briefly discuss research done in the area of noisy OCR. We then discuss two commercial applications of text recognition.

2.1 Noisy OCR

Noise found in documents has been a challenge for OCR systems. [2] discusses and analyzes the performance of four person-name recognizers on noisy data. Named entity recognition (NER) is applicable to our work and will likely influence our future recognition of names on the headstone. However, In addition to recognition of the names on the headstone, we must also recognize information such as dates.

The work done in [3] introduces an approach to improve recognition of noisy documents through specialized training of the OCR engine. OCR engines are typically trained on clean training data, which varies in nature greatly from the potentially noisy documents. The approach described is to train the OCR engine on noisy documents that have a high recognition rate despite the noise. This results in an OCR engine that is more appropriately modeled after its intended use. Although this approach is focused on recognizing text in a document, this research can help assist us in creating training an OCR engine adept for recognizing headstone information.

2.2 Google Goggles

Google has released the ability to search by sight through “Google Goggles.” This technology allows the user to take a photo of logos, labels, artwork, contact info, books, landmarks, and text to learn more about the subject. The use of contact information, books, and text are particularly relevant to our goal. The problem Google is attempting to solve is ambitious and difficult. We wish to simplify the problem and focus solely on headstones. In doing this, we will pay for more accuracy at the cost of less generality. Also, with Google Goggles, when an photo is taken by the user, a search is performed on the Internet for relevant information. What Google finds is what you get. The information available on headstones is not currently indexed and will not be found by a Google search. However, as Capstone is put into use and more headstones are indexed, this would potentially increase the success and end use of technologies such as Google Goggles in finding more biographical information regarding the deceased individual.

We believe there is potential for increased accuracy due to a limited scope of content (names, dates, etc.). However, in other ways, recognizing textual content in headstones poses a more difficult problem due due to the variation and complexity in the surrounding content and context and the inherent three-dimensionality of the problem.

2.3 Business Card Readers

There are currently many applications that can be purchased and downloaded for the iPhone with costs ranging from $0.99 to a few dollars. Several dozen of these contain OCR applications for business cards, documents, billboards, and the like. While some of these applications share similarities with capturing alphanumeric data from headstones (i.e. text with different fonts, found at different locations surrounded by graphical content), they usually have the advantage of a more predictable (light) background, with controllable lighting and less noise (from design, background, wearing, etc.). Headstones, on the other hand, come in a wide variety of shapes and colors with varying lighting conditions, shadowing, etc (see Figure 2). In addition, the text may be recessed or embossed, light or dark, faded or worn, and will usually be presented in a range of styles and fonts, even on the same headstone.

Figure 2. Headstones come in a variety of shapes, sizes, formats, text styles and fonts, with a mix of graphical content and descriptive text, making automated extraction of content a challenging problem.

Through a simple test of CamCard Lite (iPhone app) we found an increased accuracy of recognizing business cards over Google Goggles. This leads us to believe that the increase in accuracy is due to the limited scope discussed above. However, recognizing headstones pose more difficult problems than those of headstones due to the potential variations and variables from headstone to headstone. Thus, neither class of application is sufficient for recognizing headstones.

3. Methods

Using Capstone in its current state provides a desktop application in which the user can open an image and view segmentation and OCR results. This will evolve into a process in which the user will take a photo on a mobile device, which will trigger the start of the described Capstone process. The photo will need to be captured judiciously, meaning that the headstone be centered and leaving space on all sides (when possible) and that any debris such as mowed grass or fallen leaves be cleared. The user will also need to have location services turned on when capturing the photo. This will store geotagging information in the image header and provide us with the GPS coordinates, direction, and image orientation.

The methods used in this paper to capture the textual content found in headstones is summarized visually in Figure 1 and are discussed step by step below.

3.1 Segmentation

When the photo is captured, the first step in the process is to use watershed tobogganing as presented in [4] to create TRAPs. A TRAP (Tobogganed Regions of Accumulated Plateaus) is a grouped catchment basin of tens to hundreds of pixels found through the tobogganing process. We compute the TRAPs for the entire image, segmenting the image into small subregions. These TRAPs are invisible to the user, but will be used for future processing.

We then segment the image based on foreground and background seeds. Image Segmentation will be one of the most challenging aspects of this project. This is due to the potential variations shown above in Figure 2. We also cannot make assumptions on whether the inscription is darker than its stone, and whether the inscription will be recessed or embossed. The segmentation algorithms must therefore be quite general and will proceed in the following steps:

Figure 3. Capstone showing the segmentation of the foreground (green tint) from the background (red tint)

First, the headstone will be segmented from the surrounding background (Figure 1-b). Currently the user must manually identify the headstone through left-clicking and dragging the mouse over the headstone. To specify the background the user must right-click and drag over the surrounding areas. Upon each click of the mouse, a graph cut algorithm is performed to separate the foreground from the background, or in our case, the headstone from the rest of the image. The graph cut algorithm takes advantage of the TRAPs previously found. This provides for increased speed and efficiency in finding the graph cuts when compared to only using the image's pixels. This process is repeated if necessary until the headstone is properly defined from the background. When each graph cut is performed, the image is updated to show the graph cut based on the given seeds. The foreground is shown with a green tint, and the background with a red tint as shown in Figure 3. When the user is satisfied with the segmentation, the “Remove Background” option is chosen to trim the image and replace any leftover background with white pixels.

After the headstone is separated from the background we look to segment the areas of text from the rest of the headstone (Figure 1-c). This is done using the same approach described in the above paragraph, however, this segmentation poses a more difficult problem. A number of headstones have bound all of its text within a border, thus simplifying the search for text areas (as seen on the headstone in Figure 1-c). Unfortunately, the majority of headstones contain artwork or additional engravings alongside the text, or the text area simply has no bound other than the edge of the headstone. While we will preserve the option for user interaction in the placement of seed points, the goal is to zone the inscription areas containing the text automatically.

3.2 Image Filtering

Once the text has been isolated, we will apply image filtering techniques to make the text as recognizable as possible for the OCR engine (Figure 1-d). Images of headstones may contain shadows or glares and the stone of the headstone may vary in saturation/intensity and may also contain large amounts of speckled texture; all seen as noise to the OCR engine. We wish to reduce the noise as much as possible. Currently the only form of image filtering is the use of a median filter to reduce the noise of the stone and make the background closer to uniform.

When filteringing of the image is complete, we pass the image to the OCR engine and are given our result (Figure 1-f). We use the Tesseract-OCR open source OCR engine [6].

In addition to the text recognition process we also wish to capture the geotagging information. The geotagging information is extracted using the open source library “libexif” [7]. The GPS coordinates and direction are gathered.

4. Results

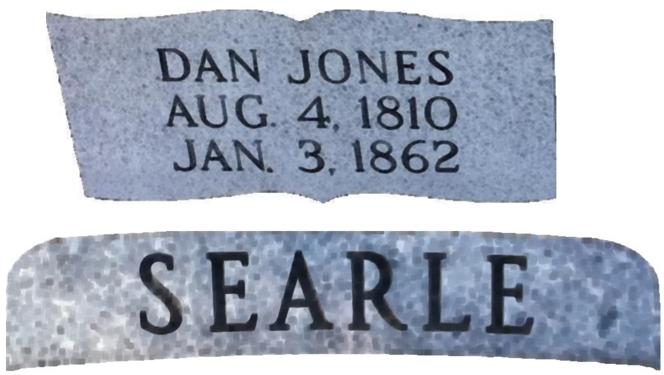
4.1 Segmentation

As mentioned previously, segmentation of an image is solely manual at this time. Headstones vary in the number of seeds needed to perform the segmentation. Figure 4 shows the segmentation on headstones from the same seed point lines. The seed lines are shown in green for the foreground and red for the background (upper left and lower right). These seeds were automatically placed with no regard to the image data. We can see that in some cases segmenting the image can be straight forward whereas in others, challenges occur due to glare, shape, or variation in color.

Figure 4. Graph cuts based on simple seeding lines

(a) (b)

Figure 5. (a) Images of headstone information inscribed on noisy stone. (b) Same images after passing through a median filter



DAN JONES 1.
Auc.. 4. \s\o 2
JAN. 3. \862

f SEARLE

Figure 6. Filtered images with their corresponding OCR'd text

4.2 Image Filtering

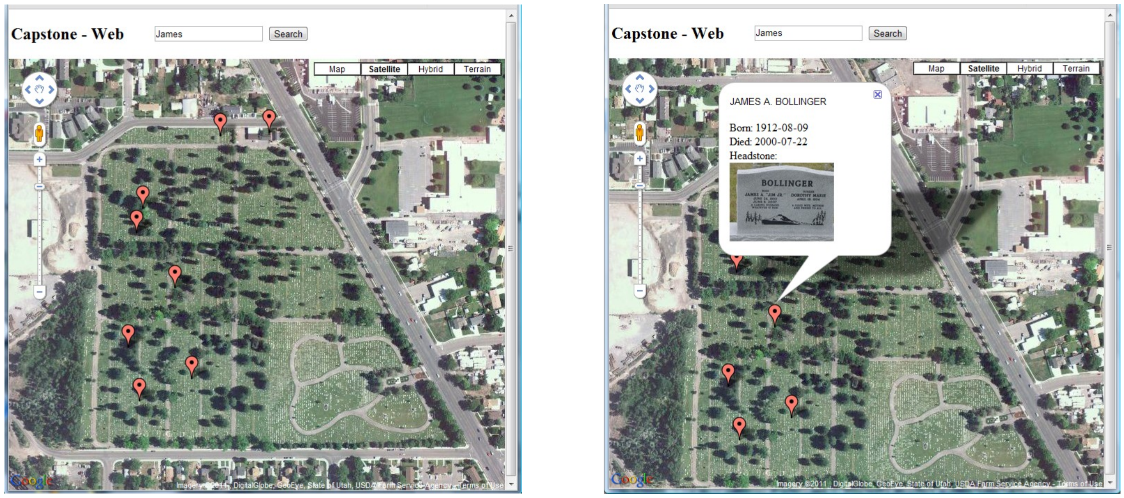
The filtering done to the image is simple as of yet, but as seen in Figure 5 the noise is reduced through the median filter, resulting in a much more usable image for the OCR engine.

4.3 OCR output

Figure 6 Displays two segmented and filtered images with their corresponding text. A clear difference can be seen between the OCR output of the rectified-text input (Figure 1-e, 1-f) and the above, non-rectified image. This reveals the next area to implement for improved accuracy. To give more meaningful results, a much larger test sets is required and is not available at this time.

5. Application

An example application for use of this data is shown in Figure 7. The Google Maps API is used to drop a pin on the headstones we recorded in our visit to the Provo City Cemetery. We are also able to search by other information such as name. Figure 7-a shows the map after a search is performed for all those with the name of James buried in the cemetery. Figure 7-b shows detailed information about the headstone. This bubble potentially will include links to a biography, journal or other historical information. A link may also be made to this individual in genealogical web sites such as http://new.familysearch.org. The potential uses are encouraging and are left for the general public to find innovative ways to use this valuable data.



(a) (b)

Figure 7. Example application of data gathered from Capstone.

6. Future work

There is much work to be done to achieve the desired accuracy. However, the improvements discussed below are very realizable and cause a significant improvement in accuracy.

As discussed in section 3, we wish to automate segmentation of text. This will be attempted in several ways. One way is by analyzing the histogram of the colors/intensities found in the segmented headstone and splitting the histogram using the Otsu algorithm. Pixels corresponding to the intensity of the mode on each side of the split would them be used to automatically seed and segment these zones. Techniques such as using texture, connected components/lines and additional graph cuts will also be considered as means to find inscription areas.

We have not yet addressed issues such as if shadows or glares exist, causing the intensity and coloring of an area to be greatly distorted. To address we will apply techniques described in [5] to make these areas more visible. We will also look into further techniques on removing background noise, as the simple median filter will not provide the clarity we desire.

Another obstacle for the OCR engine is the text orientation. Many headstones contain text that is curved or slanted. Text may also become warped in the image simply by the angle from which the photo was taken. By artificially casting profiles at coarse to fine angles, we can determine the best horizontal orientation as that angle that minimizes the spread in the profile (the distance between leading and trailing edges). Once the text has been rectified, it can be passed to the OCR engine for recognition. In Figure 6 we can see the impact of not rectifying the text before passing it to the OCR engine. We see the '1' character confused as a '\', an understandable mistake given the slanted text. In running simple comparisons to rectified text versus slanted, accuracy was improved significantly.

Thus far Tesseract-OCR has proven sufficient, however we leave open the possibility of changing the OCR engine. Currently our build of the Tesseract-OCR has only the default training and default lexicon. This is also a prime area for future tuning. We will train the OCR engine directly on headstones and fine tune the lexicon available. When trained on headstones, the OCR engine will gain a better understanding of the size, fonts, and character usage given on headstones.

We will also look into using OCRopus, which builds upon the Tesseract-OCR engine, adding probabilistic language models, and layout analysis.

In addition to geotagging information, we will look into using additional image information that may assist in the filtering step above (image orientation, flash used indicator, etc.).

Currently no web services exist to upload the extracted headstone information. The data will be uploaded using the smart phone's wireless Internet capability to a server for storage and become accessible through the Internet in a medium such as that shown in section 5.

7. Conclusion

The current implementation of Capstone exists on a desktop machine. This will need be ported to a mobile device including integration with the device's camera and wireless services. At the present time, Apple's iPhone is our desired destination.

Capstone is still in its infancy and we expect to see great improvements. Capstone will make possible the gathering of valuable genealogical and biographical information found on headstones. The current processes of gathering this data is time consuming and laborious. We therefore seek to automate this process and make possible the indexing of this largely untapped resource.

8. References

[1] Y.Y. Boykov, M.-P. Jolly, 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *Computer Vision, 2001. ICCV 2001. Proceedings.* Vol. 1, 105 – 112. Interactive Graph Cuts – Boykov, Jolly
[2] T. Packer, J. Lutes, A. Stewart, D. Embley, E. Ringger, K. Seppi, 2010. Extracting Person Names from Diverse and Noisy OCR Text. Family History Technology Workshop, 2010.
[3] A. Kae, G. Huang, C. Doersch, and E. Learned-Miller, 2010. Improving State-of-the-Art OCR through High-Precision Document-Specific Modeling. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.*
[4] W. A. Barrett, A. S. Cheney, 2002. Object-based image editing. *In Proceedings of ACM SIGGRAPH 2002*
[5] L. Zhang, A. Yip, C. Tan, Removing shading distortions in camera-based document images using inpainting and surface fitting with radial basis functions. *International Conference on Document Analysis and Recognition, 2007,* pp. 984–988.
[6] http://code.google.com/p/tesseract-ocr/
[7] http://libexif.sourceforge.net/