Interactive Smoothing of Handwritten Text Images Using a Bilateral filter

Oliver A. Nina, Bryan S. Morse
oan3@cs.byu.edu, morse@cs.byu.edu
Brigham Young University

Abstract

The use of digital images of handwritten historical documents has become more popular in recent years. Volunteers around the world now read thousands of these images as part of their indexing process. Handwritten text images of old documents are sometimes difficult to read or noisy due to the preservation of the document and quality of the image.

In this paper, we present a technique that allows interactive smoothing and text reconstruction. This makes parts of the image where the text ink is hard to recognize clearer and visually better. Results show that this technique helps selectively reduce the noise of the image and makes the text more readable.

Introduction

Indexing is a process where people read text images to transcribe important information in them for fast retrieval of them later. This process has caught a lot of attention lately for members of the Church of Jesus Christ of Latter Day Saints for large volunteer efforts to index genealogical information. This is because current computer algorithms have not been able to process and read automatically images that contain handwritten text. For this reason we cannot yet rely on computers to read the text in these images.

With the help of computers and the Internet, the Church uses the efforts of volunteers around the world to read old handwritten documents from digitized microfilms.

Because of the nature of these documents, the text is sometimes unreadable due to the image's poor quality (noise) and the age and conservation of these documents

This paper presents a variation of a technique called Bilateral Filtering and applies it to this problem to first smooth non-uniform backgrounds in text images and then augment the soft pixels in the image to become more distinct to the reader.

Background

The bilateral filter was first introduced by Smith and Brady who called it “SUSAN” [1]. Later, Tomasi and Manduchi rediscovered it and called it the “bilateral filter” [2] which is the name that most people now know it by.

The bilateral filter replaces each pixel by a weighted sum of its neighbors based on both distance and brightness differences of the pixels in the image. The decreasing weighting function is a Gaussian function(G_σ), and considering a gray-level value *I*, we have

$$I'_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\|p-q\|)\ G_{\sigma_r}(|I_p - I_q|)\, I_q$$

where

$$W_p = \sum_{q \in S} G_{\sigma_s}(\|p-q\|)\ G_{\sigma_r}(|I_p - I_q|)$$

The parameter σ_s defines the extent of the spatial distribution around a pixel. The parameter σ_r controls the extent of the distribution of the difference of intensity between a pixel and its neighboring pixels.

One of the advantages of the bilateral filter is that it smooths areas where the pixels are similar. This allows us to leave relatively unaffected the edges in the image.

The bilateral filter has been used to smooth character images of handwritten text before as in Basavaraj et al. [5]. However, in this paper the parameters were selected manually with no information about them. There is also no mention about the properties and use of the filter. The paper was focused mainly on text feature extraction.

In contrast, in this paper we explore the potential of the bilateral filter as an interactive tool for image smoothing and reconstruction of handwritten text for manual transcription and indexing.

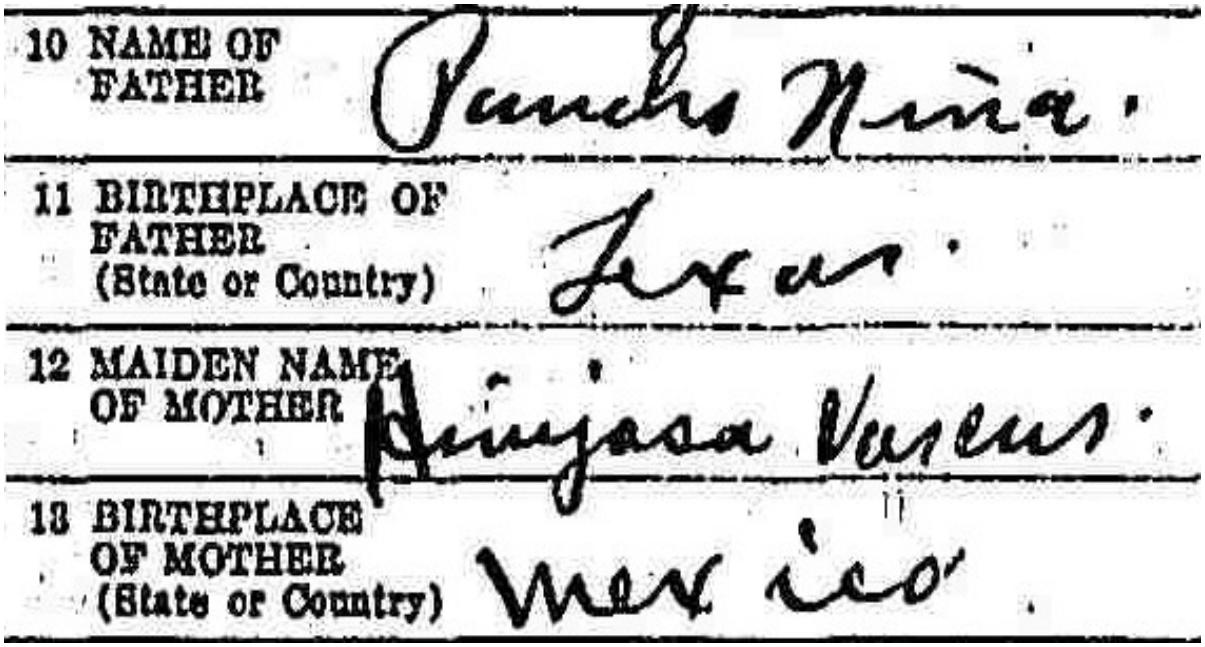
The bilateral filter is computational complex. It is difficult to use interactively due to computational constraints. However, in 2006, Paris and Durand [3] proposed a new approximation of the bilateral filter with convolution in a higher-dimensional space. This made the interactivity of the algorithm much faster and opened the door to new uses of the bilateral filter in different applications such as real-time edge aware processing [4].

Our Method

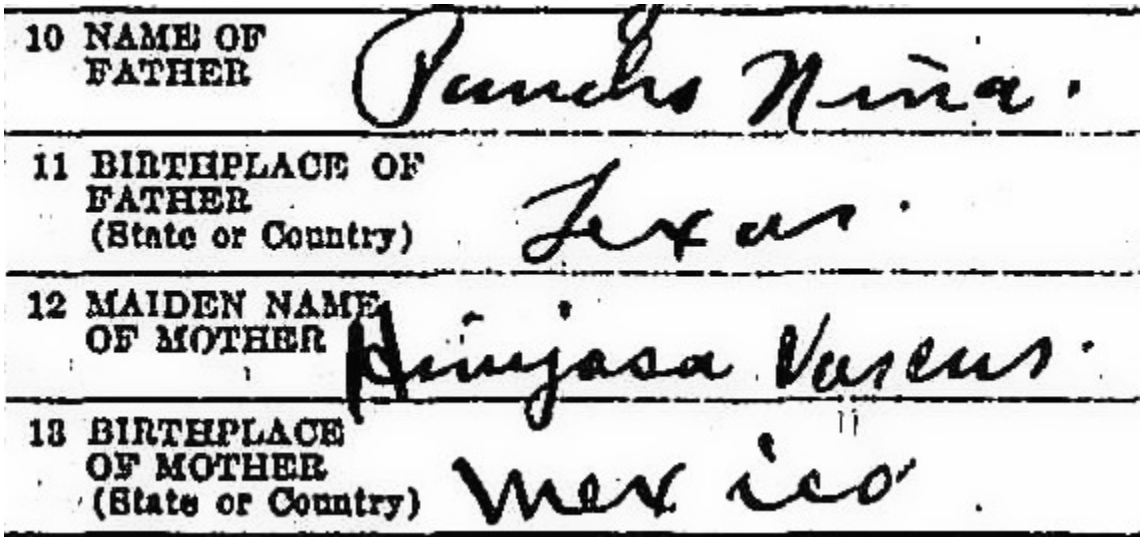
One of the reasons why indexers cannot read or recognize some of the handwritten text in the images is because the background often doesn't provide a good contrast with the text, and non-uniform backgrounds usually clutter the ink of the text.

The following picture taken from the labs.familysearch.org site, illustrate better how the background of poor quality images can limit the visibility and reading of the text. Notice the jagging around the text because of the lossy jpeg compression and the quality of the image.

The second image shows the resulting image after applying the bilateral filter on the original image.



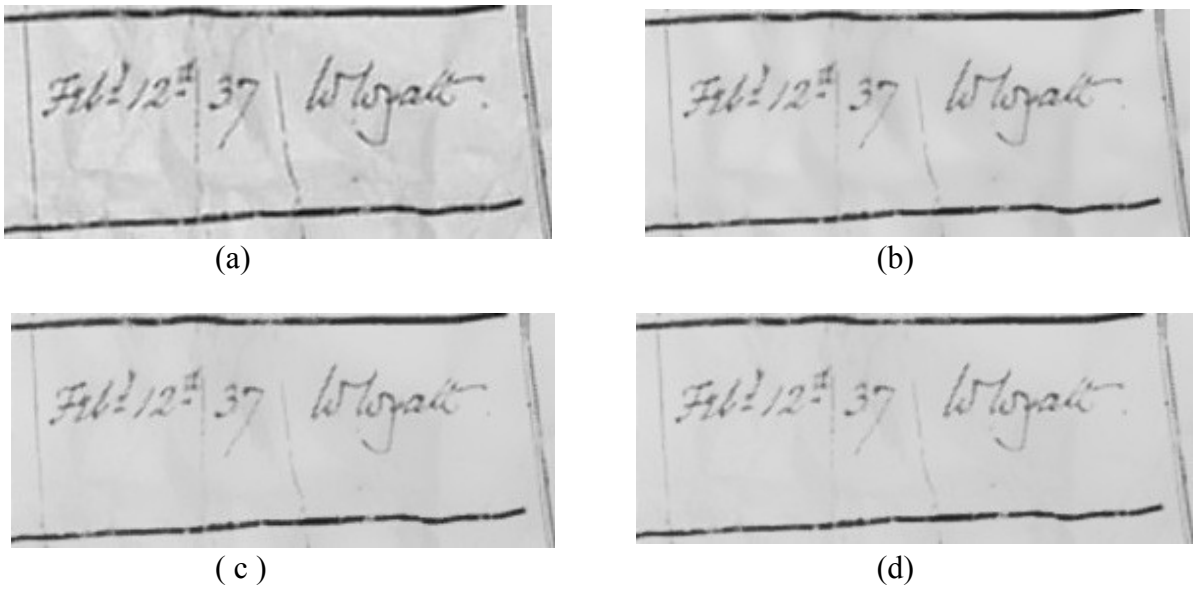
a) Original Image



b) Bilateral Filter with $\sigma_s = 10$ $\sigma_r = 4$

Figure 1: Before and after example of the bilateral filter

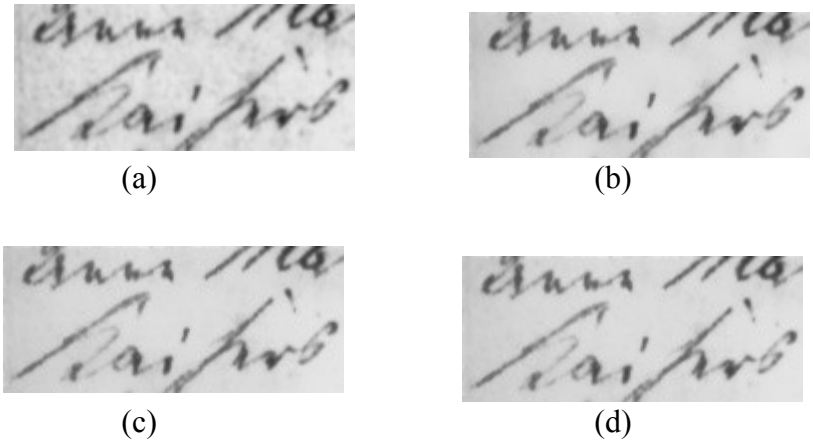
In order to fix the problem of uneven backgrounds we use the bilateral filter to "smooth" them. We can also do this interactively because in most cases the majority of the text can be read but there are only small parts of the image where it cannot be. We indicate with a mouse the areas where we want the filter to be applied. We do this by clicking and dragging the mouse where we want the algorithm to work. Figures 1 and 2 show the result after applying just the bilateral filter to smooth the background from the text images.



(a) (b)

(c) (d)

Figure 1: (a) Original image (b) Bilateral filter with $\sigma_s = 5$ $\sigma_r = 3$ (c)Bilateral filter with $\sigma_s = 10$ $\sigma_r = 3$ (d) Bilateral filter with $\sigma_s = 15$ $\sigma_r = 3$



(a) (b)

(c) (d)

Figure 2: (a) Original image (b) Bilateral filter with $\sigma_s = 5$ $\sigma_r = 3$ (c)Bilateral filter with $\sigma_s = 10$ $\sigma_r = 3$ (d) Bilateral filter with $\sigma_s = 15$ $\sigma_r = 3$

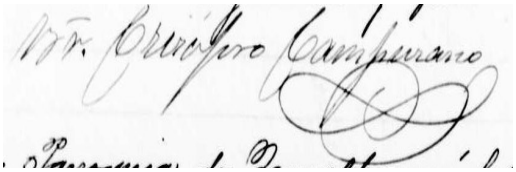
When we want to interactively allow a user to make soft strokes more pronounced, we can improve the method further. For this part of the algorithm we combine the bilateral filter with an edge detection algorithm that allows us to recognize when the mouse is over an edge. We use a Laplacian edge-detecting algorithm to identify the edges on the image.

We also introduce two parameters for the algorithm. One lets us control the threshold of the values in the Laplacian image where we want to differentiate between strokes and background, and the other lets us change the brightness of the soft strokes to a desired value.

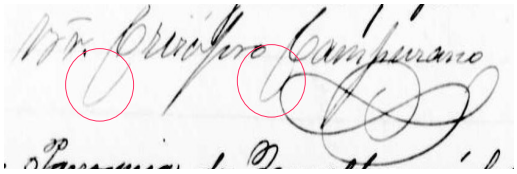
In Figure 3 we see how we can restore some of the softer strokes in the image and make it clearer. Figure 3a shows the original image. As it is shown, some of the strokes that belong to the two Cs of the name in the original image are very soft and not very clear. We apply our algorithm by clicking directly in these areas where the Cs appear very soft. By using the help of the Laplacian filter on Figure 3b we can have a more precise idea of where we want to accentuate the contrast of the strokes (make darker). Thus in Figure 3c we see the result after applying the algorithm in this soft areas.

As it is seen in Figure 3b, the Laplacian image cannot very accurately identify the exact edges where the ink is. However, by setting our threshold parameter we can visually adjust the parameter to allow more or less information from the Laplacian image according to what looks and reads better.

Hence, with the help of the bilateral filter, the user and the mouse, we can better accentuate the soft strokes, thus helping us read the character better.

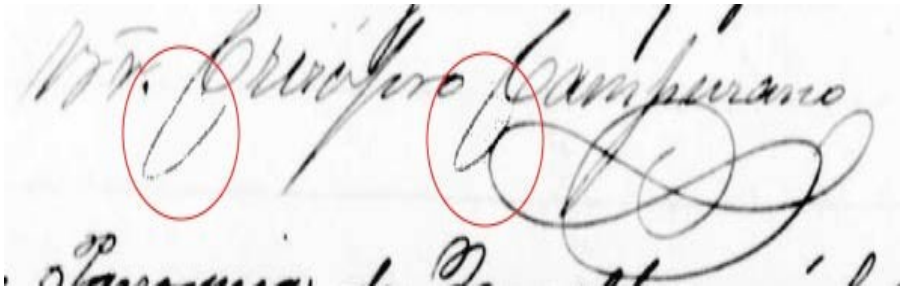


(a)



(b)

(c)



(d)

Figure 3: (a)(b) Original image (c) Laplacian Filter of the already smoothed image (d) Strokes recovered

Conclusion

In this paper we have presented a variation of a technique for Bilateral filtering and applied it to solve the problem of low quality in handwritten text images.

Our method shows promising results and opens a new alternative for more research in the area of interactive image processing and computer vision as they apply to family history documents.

Further work

One of the ways in which this technique can be improved is by identifying the start and end of the edges better. Also, a valuable alternative would be to be able to automatically generate the parameters necessary to work with the bilateral filter and the control of noise for the Laplacian filter.

References

1 Smith, S.M., Brady, J.M.: SUSAN – a new approach to low level image processing. International Journal of Computer Vision **23** (1997) 45-78

2 Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Proc. of International Conference on Computer Vision, IEEE (1998) 839-846

3 Paris, S., and Durand, F. 2006. A fast approximation of the bilateral filter using a signal processing approach. In Proceedings of the European Conference on Computer Vision.

4 Chen, J., Paris, S., Durand F.,:Real-time Edge-Aware Image Processing with the Bilateral Grid. In Transactions on Graphics, ACM (2007) Vol. 26, No. 3, Article 103.

5 Basavaraj, L., Samuel, S. 2007 : Offline Handwritten Character Detection using Image Components. In: International Conference on Computational Intelligence and Multimedia Applications.