A Simplest Approach to Reduce the Yellowness and Foxing in Antique Book Pages

Anjali Chopra
M.tech Student
School of Engineering and
Technology, EECE Deptt.
ITM University

Amanpreet Kaur
Assistant Professor
School of Engineering and
Technology, EECE Deptt.
ITM University

Shaveta Arora
Assistant Professor
School of Engineering and
Technology, EECE Deptt.
ITM University

ABSTRACT

Digital data is gaining importance in communication domain. Thus, a simplest approach to correct the yellowness and foxing in pages of antique historical data is proposed here. This is the simplest approach to recover the data in digital form, is less expensive and doesn't require any trained personnel to perform the task. Digital reproductions are a powerful tool to share the information and to preserve the original materials. Large intellectual and financial resources have already been devoted to solve this purpose.

Keywords: Foxing, MRF, Thresholding

1. INTRODUCTION

In this technological era, the digital data holds a very important domain. This is mainly because it allows easy data handling, and transportation. Thus, the same data can be used by everyone and everywhere. Usually most of the data these days is stored in the digital form.

The data which is stored in the form of antique books and documents holds utmost importance in our culture and knowledge and should be preserved for the generations to come. This data is distributed in thousands of public libraries and museums and is therefore, not accessible to everyone. Thus, it is very important to convert this data in the digital form. This digital copy can be a better reproduction of the original document, improving its quality and recovering the damages on the original document. Some resources have already been devoted for this purpose. Some examples may include the Digital Libraries Initiative and the EU-funded Bricks Digital Library Project.

This conversion into the digital document requires an excellent scanning process and then the further processing. UV scanning can be used for this purpose but the book binding should not be harmed [3].

There are several approaches which can be used for the digital restoration of the antique document. In [4, 5, 6], the Markov-Random Field (MRF) method is used which describes the properties of the boundaries and the characters. It optimizes and produces an enhanced text but is a costly affair and uses approximations. To recover some text in the Archimedes Manuscripts, multispectral imaging techniques were also used. [7, 8]

In this paper, a very simple approach is proposed which enables restoration of data from antique documents with a similar kind of background, i.e., a similar kind of yellowness of paper.

The yellowness of the original document is mainly because of the aging and the local red coloration observed in the document is known as 'foxing'. This foxing can be because of certain metals used in the manufacturing of the paper or due to the action of the micro-organisms. The removal of foxing stains may involve the use of thresholding techniques [1].
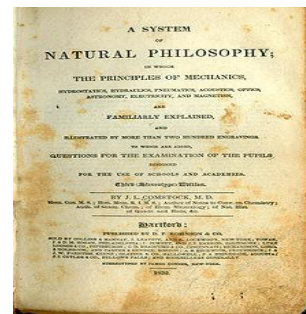

Fig 1: (a): An antique document with foxing stains received from Wikipedia.org.
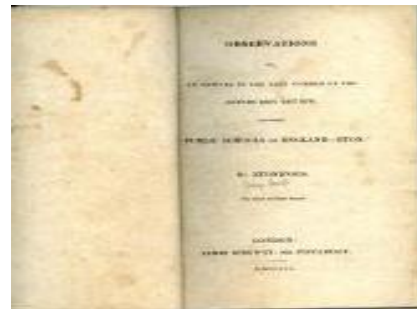

Fig1: (b): An antique document having yellowing and foxing from richardfordmanuscripts.co.uk

2. FOXING

Foxing is used to describe the local browning spots observed on the antique document. The causes of foxing may be the fungal growth over the paper because of aging. It may also be because of the oxidation of iron, copper over the paper. Humid conditions may also contribute to foxing.

The proposed method efficiently removes the foxing stains from the digital document using the chromatic properties of these foxing stains. The stains are removed and the area beneath them is filled with the color similar to that of the original background of the document.

$$Yr(i,j) = Y(i,j) + \frac{Y(i-1,j) + Y(i+1,j) - 2Y(i,j)}{k\left(Y(i-1,j) - Y(i+1,j)\right)2 + A} + \frac{Y(i,j-1) + Y(i,j+1) - 2Y(i,j)}{k\left(Y(i,j-1) - Y(i,j+1)\right)2 + A}$$

3. IMAGE ENHANCEMENT

Image enhancement is basically the modification of the pixels in the original image to some extent to make it suitable for

analysis as compared to the original document. It is the enhancement in the quality of the digital image without the consideration of the source of degradation. An image can be enhanced by increasing the contrast in the image. By increasing the contrast the readability of the document can be improved.

In the proposed approach, page enhancement is used to improve the quality of the restored document. It reduces the yellowness of the paper and provides uniformity in the background pixel values. The text is enhanced using contrast enhancement between the text and the background.

4. THRESHOLDING

Thresholding is the simplest approach for image segmentation. Otsu's method is used for histogram based thresholding. In the proposed approach, thresholding has been used to separate the image text with the background. It automatically computes an optimum threshold to segment the image into two classes. A binary image can thus be obtained using thresholding.

5. METHODOLOGY

The main aim is to remove the foxing stains on the document while retaining its original characteristics which include the yellowness of the document which is a mark of its age and its primary properties.

For this purpose the image text and background are separated, processed and then are added together so as to retain the original image.

Firstly, the colored image I represented by its R, G, B components is first converted into the Y, Cb, Cr color space and the gray color space.

The luminance coefficient is filtered for each pixel in the image so as to improve its edges, boundaries and provide a similar luminance in the areas between these boundaries.

This is done using the expression below:

where, A and k are constants whose values are defined later. [2]

Input the scanned image

↓

Convert into the red component image

↓

Select a threshold and convert into a binary image

↓
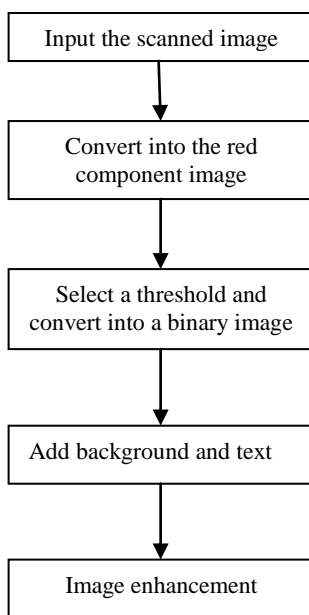
Add background and text

↓

Image enhancement

Fig : 3 Block Diagram

The red color component is extracted from the image and its maximum value is computed. This is because the foxing stains are red-brown in color. A threshold value of this image is computed based on which the image is converted into the binary image.

Now, the background image and the image with the data or text are added to obtain the desired results. This image is enhanced using the available functions on MATLAB.[9]

6. EXPERIMENTAL RESULTS

The proposed algorithm was used with some images collected from different websites and gives improved results for images or documents having a similar background luminance.

The values for the parameters used in these images is $A = 5$ and $k = 0.03$.

Fig 4: The original document
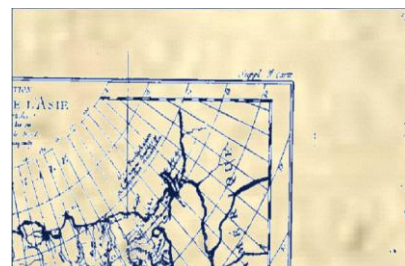

Fig 5: The red component image

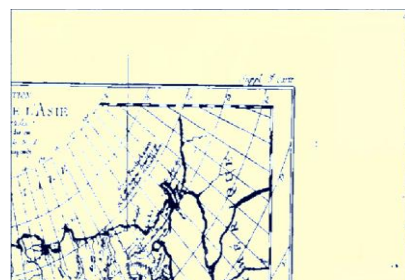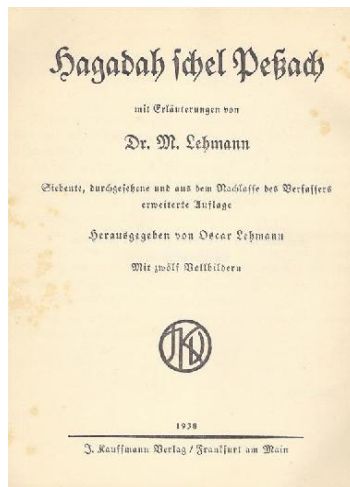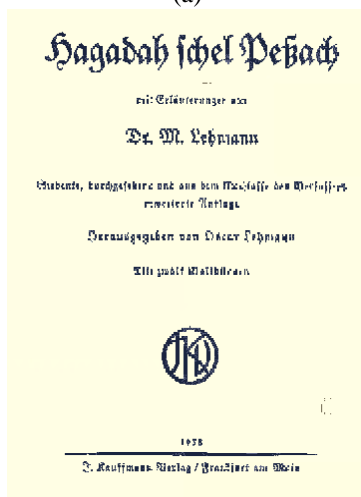**Fig 6: The recovered image having no stains.**

Fig 7: Image after enhancement.

The same algorithm was applied to another document and the results were recorded. This shows that the proposed algorithm is capable enough to be further exploited.

(a)

(b)

Fig 8: Example using the same algorithm
(a) The original document taken from danwymanbooks.com
(b) The resulting image with no stains and less yellowness.

7. COMPARISON

The other approach available to remove the affect of foxing and yellowness over a document uses the YCbCr color apace. In this approach, the luminance operator and the chrominance operators are used individually. The chrominance operators use the Cb, Cr and a map created using the luminance operators as its input. To remove the foxing stains the matrix related to the Cr component is operated upon. Rational Filtering is used to attenuate the small-image variations while preserving the edges. After removing the foxing stains, color filtering is done. The pages are merged now and enhanced further using page enhancing techniques. A comparison is drawn between this technique and the proposed approach.

	Proposed Approach	Other Approach
Color Space	RGB	YCbCr
Filtering Used	None	Rational Filtering and Color Filtering
Application	Can be used for documents with similar background luminance.	Can be used for documents with all kinds of background luminance.
Complexity	Less Complex	More complex
Page Enhancements	Used	Used

8. CONCLUSIONS

In this paper, a simple approach to recover the data from antique documents having similar background illumination has been introduced. This gives a better reproduction of the original document removing the foxing stains and reducing the yellowness of the paper. A comparison has been drawn between the existing and the proposed approach. It is also evident that the proposed method is very simple and economical and doesn't require any trained personnel. We can further extend this approach for all kinds of background illumination and text.

REFERENCES

[1]   Otsu, N.: ' A Threshold Selection Method from Gray-level Histograms'. IET Systems, man and Cybernet.,1979, vol. 1, pp. 62-66

[2]   Stanco, F., Tenze, L., Ramponi, G.:'Technique to correct yellowing and foxing in antique books'. IET Image Process., 2007, vol. 1, no. 2,pp. 123-133

[3]   Zhang, Z., and Tan, C.: 'Rstoration of images scanned from thick bound documents'. Proc. Int. Conf. on Image Process.,2001, vol. 1,pp. 1074-1077

[4]   Tonazzini, A., and Bedini, L.: 'Character segmentation in highly blurred ancient printed documents'. Proc. IEEE Int. Conf. on Image Anal. Process, 1999, pp. 836-841

[5]   Bedini, L., Bozzi, A., Tonazzini, A.: 'Digital technique for character recognition in old documents', ERCIM News, January 2004, http://www. ercim.org/

[6]   Bedini, L., Corso, G.D., Tonazzini, A.,: 'A preconditioning technique for edge-preserving image restoration'. Proc. IEEE Int. Conf. on Inf. Intell. Syst., 1999, pp. 519-526

[7]   Tonazzini, A., Bedini, L., Salerno, E.: 'Image analysis on the archimedes palimpsest', ERCIM News, July 2004, http://www.ercim.org/

[8]   Easton, R.L.J., Knox, K.T., and Christens-Barry, W.A.: 'Multispectral imaging of the archimedes palimpsest'. Proc. IEEE 32nd Appl. Imagery Pattern Recognit. Workshop (AIPR03), 2004

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