Techniques to Enhance Images for *Mokkan* Interpretation

Jun Takakura, Akihito Kitadai, Masaki Nakagawa
Tokyo University of Agri. & Tech.
Tokyo, Japan
j.takaqula@gmail.com

Hajime Baba, Akihiro Watanabe
National Research Institute for Cultural Properties, Nara.
Nara, Japan
hajime@nabunken.go.jp

***Abstract*— In this study, we investigate several methods for enhancing scanned *mokkan* images to aid archeologists and historians in the interpretation of *mokkans*. *Mokkans* are wooden tablets with handwritten characters used in 8th century Japan. Due to damages and natural deterioration, the interpretation of the text on *mokkans* is difficult even for archeologists and historians. The automatic interpretation of *mokkan* text is unrealistic due to drastic deterioration. Our aim is to realize an image processing software system for archeologists, specific to the processing of *mokkans*. This paper evaluates noise reduction and contrast enhancement techniques for inclusion in the system. Several operations such as thresholding and Gaussian blurring are shown to be effective, when used in conjunction with contrast enhancement in the HSV space. We also identified several techniques that allow convenient selection of character strokes in of *mokkan* images. The selection can then be used as input for character recognition.**

Keywords-historical document, image processing, mokkan,

I. Introduction

The analysis of historical documents is often complicated by the damage or deterioration of the documents. Image processing techniques have come in very useful for the analysis of these deteriorated texts. These methods typically perform binarization [1-8], noise removal/reduction [9], or contrast enhancement [10].

In this paper several methods for enhancing *mokkans* for their interpretation are considered. "*Mokkan*" is a Japanese generic name which refers to wooden tablets used as documents in ancient Japan. Texts are written on *mokkans* using a brush with Indian ink. There are about 320,000 old *mokkans* excavated at the ruins of ancient cities in Japan. Figure 1 shows some of the *mokkans* from the ruin of the Heijo palace site, the capital of Japan in the Nara period (from A.D. 710 to 794). Many *mokkans* were used for luggage tags of gifts, commodities, goods for tax, and so on. Therefore, the decoding of these *mokkans* is important to find the flow of materials, the relations among regions and the condition of economy at the period.

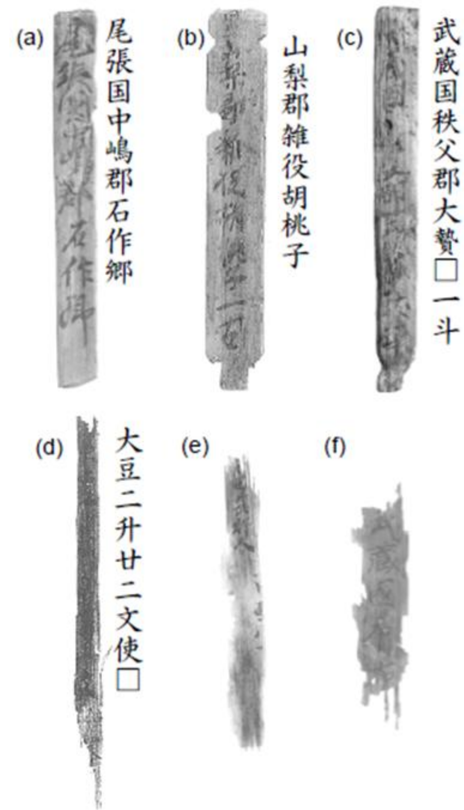


Figure 1. Mokkans from the ruin of the Heijo palace site.

II. The condition of *mokkans*

The interpretation of *mokkans* is made difficult by their extensive deterioration. Since they have been buried underground for more than 1300 years, the written ink has become worn out. Frequently, the ink has seeped through other parts of the wooden tablet, resulting in deformation of the written characters. The deformation is difficult to isolate since they compose of the same ink material as the actual written text.

Another problem with *mokkans* is that the wooden tablets may be partially damaged due to the conditions which buried them, or during excavation. This causes parts of the text to be missing and has to be filled in manually through the interpretation of experts.

Due to these problems, in spite of the archaeological value of the *mokkans*, only a small portion of them have been deciphered so far. As mentioned, many of the excavated *mokkans* have been stained and damaged, and frequently requires heavy guesswork for their decipherment.

978-0-7695-4221-8/10 $26.00 © 2010 IEEE
DOI 10.1109/ICFHR.2010.62

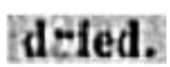
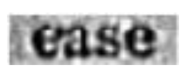
At this point, the automatic decipherment of *mokkans* is considered unfeasible.

III. Motivation

The mainstream method for enhancing historical documents is based on binarization, which is typically used as a pre-processing step for optical character recognition. In the present work we do not consider binarization techniques, for two reasons. One, binarization does not necessarily result in the most human-readable text. As an example consider the processed images in Table 1, taken from [3]. The second reason is that we consider the task of automatically deciphering *mokkan* text as being still far from achievable. For these reasons we do not consider the currently available automated methods for recognizing the text on historical documents [1-8].

Our approach to *mokkan* interpretation is to provide archaeologists with a set of methods for enhancing *mokkan* text. These methods may be effective with respect to only specific deteriorations. We consider it the choice of the archaeologists to decide which enhancement method is to be applied, based on what they perceive to be the condition of the *mokkan* under inspection. The output from the methods may also lead to an image that may be more suitable for character recognition.

TABLE I. Binarization may not result in more human readable text.

Method Name	Sample 1	Sample 2
Original Image		
Markov Model		
Global Fixed Threshold		

IV. Methods

The methods which we consider in this paper are mainly to do with processing image to enhance readability. We evaluate various commonly used methods for image processing and examine their effectiveness through consultation with archaeologists.

While we consider automated text recognition still unachievable, under the supervision of a specialist, it is conceivable for *mokkan* text to be manipulated in ways that may improve their chances to be correctly classified by a character recognizer. Hence, we also consider methods for identifying handwritten strokes in *mokkan*. The outputs of these methods are to be confirmed by archaeologists, and then possibly given as input to a character recognizer.

Two methods are considered to enhance the readability of *mokkans*, namely, noise reduction and contrast enhancement. A few methods that enable experts to prepare the character strokes for automatic character recognition are also considered.

V. Methods to Enhance Readability

A. Noise reduction

Two methods are attempted to reduce noise in this study. In the first method, Gaussian filters of different frequencies are applied to the *mokkan* image. In Figure 2, the image of a *mokkan* (interpreted as written with either the word “沸” or “佛”) is processed respectively using a Gaussian filters of 7 pixels and 23 pixels.

When Gaussian filters are applied, noise in the images becomes less noticeable. However, the strokes corresponding to the text in the images are also affected. The disability to distinguish character strokes from noise, due to their similarity, limits the effectiveness of the Gaussian filter approach to noise removal.

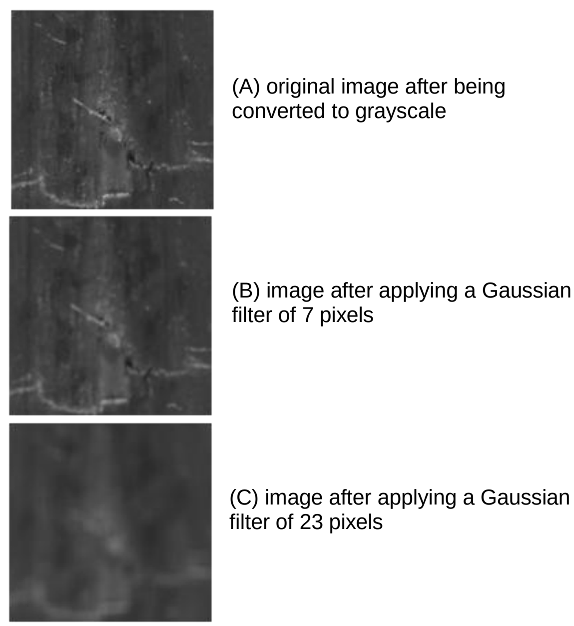


Figure 2. Images after applying noise removal method.

In the second method, we use Otsu’s method [11] to first obtain a threshold that corresponds to the background, and then remove the pixels below this threshold from the *mokkan* image. Figure 3 shows the same image used in Figure 2, before and after this processing.

Our experiments show that most images do not become more readable after the processing using this approach. Again, this is due to the process’ interference in the regions that correspond to character strokes.

In summary, both the methods for noise reduction appear to be ineffective in improving the readability of *mokkan*

(A) original image

(B) image after binarization using Ostu's method

(C) image after removing the pixels below the threshold obtained using Otsu's method

Figure 3. Color image masked by binary images.

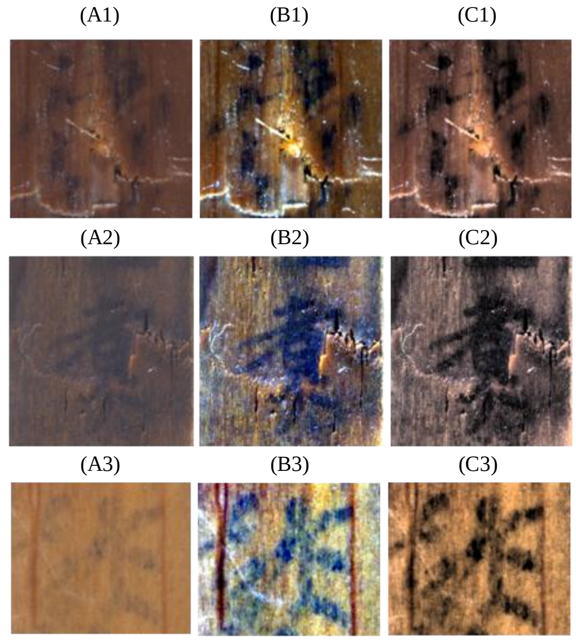
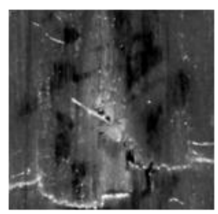
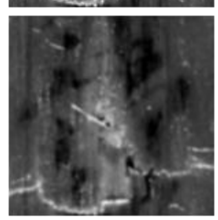


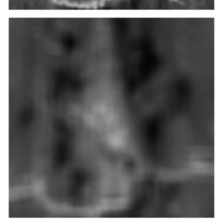
Figure 4. (A1-A3) original images (B1-B3) after contrast enhancement in RGB space (C1-C3) after contrast enhancement in HDV space.



(A) Figure 2(A) after HSV contrast enhancement



(B) Figure 5(A) after applying a Gaussian filter of 7 pixels



(C) Figure 5(A) after applying a Gaussian filter of 23 pixels

Figure 5. Effects of noise reduction on HSV contrast-enhanced image

images in our experiments. However, when combined with contrast enhancement, the effects from these techniques become more pronounce, as will be seen in the next section.

Wang et al. suggested another method for noise reduction [9]. We are currently extending our codes to experiment with this method.

B. Contrast enhancement

Contrast enhancement can be performed in several spaces of the RGB color model. In this study the original RGB space and the HSV space is used. When working in the RGB space, the values in each of R, G and B are linearly transformed so that they respectively take up the entire 0-255 range. For the HSV space, only the values of V are transformed to take up the entire 0-255 range. Values for each of R, G, and B typically form a single Gaussian distribution. When the transformation is performed in HSV space, the mean of each distribution is preserved.

Figure 4 shows three different *mokkans* when contrast enhanced in the RGB and HSV space respectively. Images enhanced in the HSV space typically appear more natural due to their preservation of the means of the Gaussian distributions.

When combined with contrast enhancement, the earlier methods for noise reduction produced significantly improvement in readability in the earlier *mokkan* images. Figure 5 shows the effects of noise reduction on HSV contrast-enhanced image for the *mokkan* in Figure 2(A).

The effect of Gaussian filters is further evaluated by overlapping filters of different frequencies. We use filters of 10 different frequencies from 3 pixels to 23 pixels. First, two images, H and L, are constructed by superimposing images from the 10 processed images, using two sets of multiplication factors that emphasizes the high frequency filters and low frequency filters, respectively. Then, H and L are combined, using various weights, to form the final image. The process is shown in Figure 7.

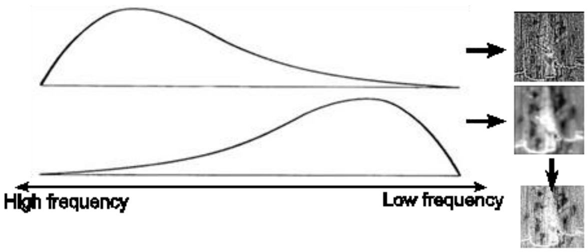
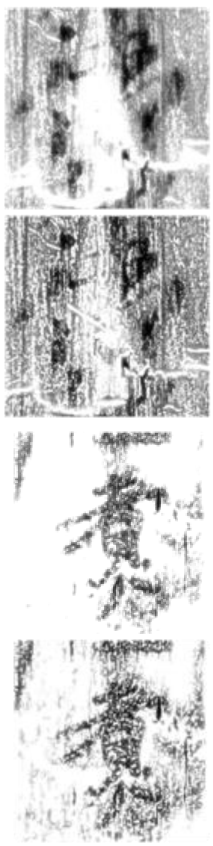


Figure 6. Noise reduction using a set of Gaussian filters.



(A1) 0.3H+1.55L on Figure 4(C1)

(A2) 0.6H+1.05L on Figure 4(C1)

(B1) 0.3H+2.13L on Figure 4(C2)

(B2) 0.6H+1.55L on Figure 4(C2)

Figure 7. Images obtained from overlapping multiple Gaussian filters

The results from application of these Gaussian filters are shown in Figure 7. Significant improvement is observed, when compared to the earlier images.

Contrast enhanced images tend to benefit more from binarization as well. A repetition of the processing performed on Figure 3 produced results better than that in the earlier case (see Figure 8).

We conclude in this section that contrast enhancement, particularly when performed in the HSV space, holds much promises in improving the readability of *mokkan* text. Agam et al. [10] suggested a method for contrast enhancement which may potentially improve the present approach.

(A) after removing pixels below threshold obtained using Otsu's method on Figure4(C1)

(B) same as (A) but on Figure 4(C2)

(C) same as (A) but on Figure 4(C3)

Figure 8. Removing noise pixels from improved contrast images

VI. Methods to Enhance Image for Character Recognition

A major problem with the automatic interpretation of *mokkan* is that the character strokes are not very significantly different from the noise, hence making the separation of these character strokes difficult. As a step towards the automatic interpretation of *mokkan*, our strategy is to simplify the process for a user to select regions on the image which the user believes to correspond to text.

A. Selection based on pixel similarity

An immediate approach to the selection of pixels is to select all the pixels that are similar a specified set of pixels. To specify this set of pixels, the user is presented with the original image and asked to select a few pixels which appear to correspond to character strokes. The selection of "similar" pixels may be performed in two different spaces: RGB, HSV and LAB. In the RGB space, the distance measure used is the value of R+G+B, whereas in the HSV space, the measure used is H+S+V. In the case of LAB, the value L+A+B is used. All the pixels within a distance of 10 from any pixel in the specified set are selected. Figure 9 shows the selection in each of RGB, HSV and LAB space.

As with noise reduction, we discovered that selection can be improved when the image is first contrast enhanced. Figure 10 shows the selection in each of RGB, HSV and LAB space, on *mokkan* images that are first contrast enhanced in the HSV space.

B. Border discovery through spectral clustering

A conceivable approach to pixel selection is to first organize the pixels into distinctive regions so that the user may select an entire region. We use spectral clustering [12] for this purpose. Clustering is performed using the software

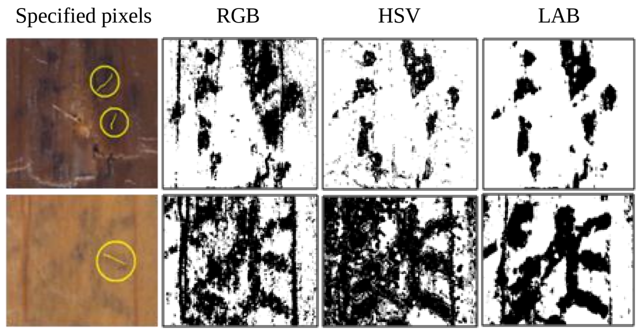


Figure 9. Selected pixels in RGB, HSV and LAB spaces corresponding to the specified pixels, on respectively the images from (A1) and (A3) in Figure 4.

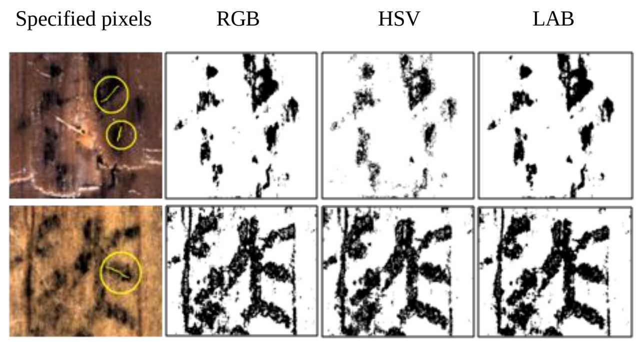


Figure 10. Selected pixels in RGB, HSV and LAB spaces corresponding to the specified pixels, on respectively the images from (C1) and (C3) in Figure 4.

R using its radial basis kernel function. Figure 11 shows the output from the algorithm on the input of a HSV contrast enhanced *mokkan* image. The method is able to identify major regions within the *mokkan* image.

C. Structure discovery after selection

After the pixels selection in Sections 6.1 and 6.2, secondary structures corresponding to character strokes may be constructed out of the selected pixels. This may be done by combining several regions into a single stroke, or by identifying line and stroke structures embedded within the selected pixels.

VII. Conclusion

In this paper we evaluated several methods for helping archaeologists interpret *mokkans*, wooden tablets with handwritten characters used in ancient Japan. Due to the deterioration suffered by these *mokkans*, it is difficult to automate their interpretation.

We identified several techniques which, when used in conjunction, may be very helpful in improving the readability of the *mokkan* images. As a step in allowing the semi-automated interpretation of *mokkans*, we evaluated several techniques for pixel selection, in the hope that eventually we may be able to develop technique to

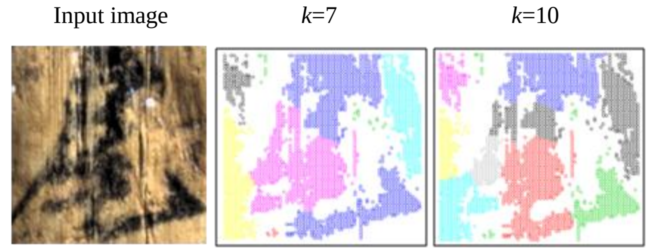


Figure 11. Clusters discovered using spectral clustering on the input image, when k=7 and k=10.

automatically identify character strokes from the selected pixels.

Constructing large database of *mokkan* images is necessary for evaluating above techniques. Also, adapting the techniques to other kinds of historical document images is our future work.

Acknowledgment

This work was supported by the Grant-in-Aid for Scientific Research (S)-20222002 and Young Scientists (B)-22720239.

References

[1] B. Gatos, I. Pratikakis, S. J. Perantonis, Adaptive degraded document image binarization, Pattern Recognition, 39(3):317-327, March, 2006

[2] B. Gatos, I. Pratikakis, S.J. Perantonis, An Adaptive Binarization Technique for Low Quality Historical Documents, DAS 2004, 3163:102-113 December 2004.

[3] M. R. Gupta , N. P. Jacobson , E. K. Garcia, OCR binarization and image pre-processing for searching historical documents, Pattern Recognition, 40(2):389-397, February 2007.

[4] J. He, Q. D. M. Do, A. C. Downton, J. H. Kim, A Comparison of Binarization Methods for Historical Archive Documents, in Proceedings of the eighth international conference on document analysis and recognition, ICDAR'05, 538-542, 2005.

[5] J. Sauvola, M. Pietikainen, Adaptive document image binarization, Pattern Recognition, 33:225-236, January, 1999.

[6] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging, 13(1):146-165, January, 2004.

[7] Ø. D. Trier, T. Taxt, Evaluation of Binarization Methods for Document Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 17 (3):312-315, March 1995.

[8] Y. Yang, H. Yan, An adaptive logical method for binarization of degraded document images, Pattern Recognition, 33:787-807, March 1999.

[9] Q. Wang, T. Xia, C. L. Tan, L. Li, Directional Wavelet Approach to Remove Document Image Interference, in Proceedings of the seventh international conference on document analysis and recognition, ICDAR '03, Volume 2, 736-740, 2003.

[10] G. Agam, G. Bal, G. Frieder, O. Frieder, Degraded document image enhancement, in Document Recognition and Retrieval XIV, Proc. SPIE 6500, 1-11, 2007.

[11] N. Otsu. A threshold selection method from gray-level histograms. IEEE Trans. Sys., Man., Cyber. 9:62–66. 1979.

[12] J. Shi, J. Malik, Normalized Cuts and Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8):888-905, August, 2000.