17th International Conference on Machine Vision Applications (MVA)
Fully Online, July 25–27, 2021.

An Optical Model for Show-through Cancellation in Ancient Document Imaging with Dark and Bright Mounts

Yuri Ueno[1] Kenichiro Tanaka[1,2] Takuya Funatomi[1] Yasuhiro Mukaigawa[1]

[1]Nara Institute of Science and Technology
8916–5 Takayama-cho, Ikoma, Nara 630–0192, JAPAN
{ueno.yuri.uq1, ktanaka, funatomi, mukaigawa}@is.naist.jp
[2]Ritsumeikan University
1–1–1 Noji-higashi, Kusatsu, Shiga 525–8577, JAPAN
ken-t@fc.ritsumei.ac.jp

Abstract

There is a need to make ancient paper documents from Asian countries such as Japan and China more readable. Many of them have writing or illustrations on both sides for reuse, which can show through the paper. In such documents, it is often difficult to distinguish just the front image because of the ink show-through from the back side. Our aim is to obtain clearer front images by removing the show-through from the back side. Hence, we propose an imaging method that uses dark and bright mounts as well as a prediction step based on a simple optical model. We successfully remove the show-through from the back and extract the images of the front of the document.

1 Introduction

Ancient documents are important in the field of literature for understanding history, culture, and people. Many researchers have developed methods for recording these documents without physical damage so that this precious literature can be conserved [1, 2].

Many such documents are inked on both sides because paper was extremely valuable in the past. Therefore, when the images of these documents are taken, the front and back sides become mixed. While it is not too difficult for humans to read such documents, a distinct image of one side of the paper without visible artifacts from the other side is more suitable for the archive data in for example, a digital database.

Some methods for contemporary printed documents have been proposed; for instance, there are methods based on an optical linear filtering scheme [3], blind source separation [4], and neural networks [5]. Here, we propose a method to remove the show-through from an image of one piece of an ancient paper document that has writing on both sides.

(a) Stitched binding (b) Ink show-through

Figure 1: Examples of ancient documents.

2 Ancient paper document optical model

2.1 Ancient paper document

An ancient paper document is composed of a large amount of paper folded and bound with a thread, which is called a "stitched binding" as shown in Fig. 1 (a). This type of document was once widespread in Asia mainly in the past time. The ancient paper used in these documents differs from the parchment. It is made of just the botanical materials, for example, in Japan, paper mulberry (*Broussonetia papyrifera*), mitsumata (*Edgeworthia chrysantha*), and Gampi (*Diplomorpha sikokiana*) have been used.

Ancient paper documents sometimes shows through because in the past, paper craftsman worked hard to make the paper thin, as Fig. 1 (b) shows. This was feasible because most of the components were botanical materials. Such paper was extremely expensive, and most people used the back side of the paper as much as possible. As a result, a huge amount of ancient literature is written on both sides of the paper.

2.2 Optical model

An image of an ancient paper document is a mixture of the light reflected from the surface and light transmitted through the multiple paper layers of the stitched binding structure. The reflected light rebounds from the front of the paper. In contrast, the transmitted

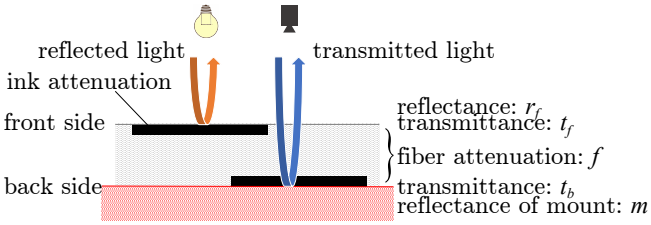


Figure 2: Optical model. Orange and blue rays indicate the reflected and transmitted rays, respectively. Paper and mount are indicated by the gray and red regions. Small black sections indicate the ink attenuation.

light passes through the paper once but reflects on the object underneath, which we call a "mount." This light transmits through the paper once more and reaches the camera.

Therefore, as shown in the Fig. 2, the image of one piece of the ancient document I can be represented simply using two rays, one for reflected light from the front side reflectance r_f and one for transmitted light formed by the transmittance through the paper front side and back side and on the bottom, which corresponds to the show-through component. This reflectance and transmittance are determined by the attenuation of the ink or not. The transmission is composed of the transmission of the front side t_f , the fiber attenuation f, the transmission of the back side t_b, and the reflectance of the mount m. The entire formula is

$$I = r_f + t_f^2 f^2 t_b^2 m. \tag{1}$$

3 Show-through cancellation

3.1 Strategy

We propose an idea that uses traditional knowledge to reduce show-through artifacts in an image of an ancient paper document inked on both sides. A dark paper mount decreases the show-through artifacts. However, using the mount alone does not completely cancel the show-through. This is because it is not possible to realize an ideal black mount with a reflectance of zero. Hence, our show-through cancellation based on an optical model requires dark and bright mounts with different reflectance values.

3.2 Optical model using two types of mount

Ancient paper documents kept in a professional institute are unbound when they accept maintenance once every few decades. The whole images, inked on both sides of the pieces, are archived one by one during such rare events. In our approach we propose installing the dark and bright colored mounts as shown in Fig. 3.

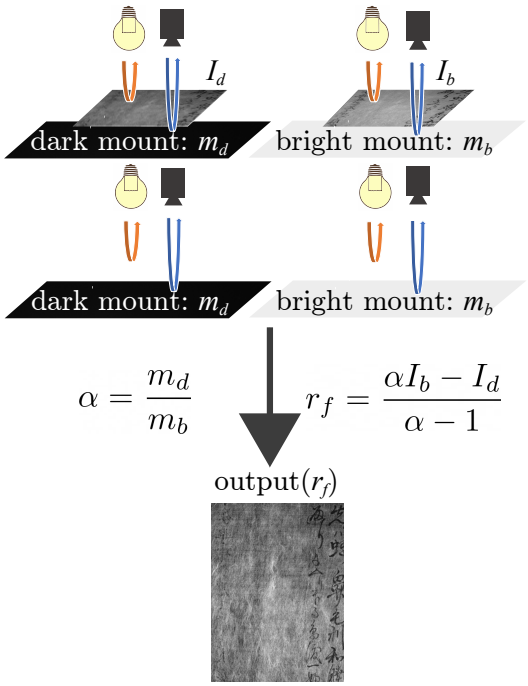


Figure 3: Flow of show-through cancellation.

These mounts must have different reflectance values, but they do not need to be completely dark and bright.

In this model, I_d and I_b denote the images taken with dark and bright mounts, respectively. The reflected ray from the front side reflectance r_f is not changed by the mount, whereas the transmittance $t_f^2 f^2 t_b^2 m$ in Eq. (1), depends on the color of each mount. Substituting $t_f^2 f^2 t_b^2$ into P, the transmission can be denoted by Pm_d and Pm_b for the dark mount m_d and bright mount m_b, respectively. Hence, the whole appearance can be expressed as

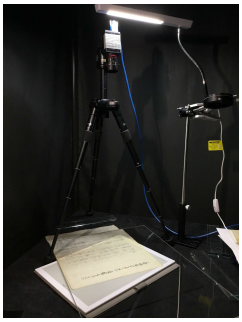
$$\begin{aligned} I_d &= r_f + Pm_d, \\ I_b &= r_f + Pm_b. \end{aligned} \tag{2}$$

3.3 Cancellation of the show-through artifacts

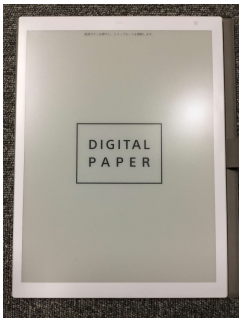
To cancel the show-through component, we estimate the front side reflectance r_f. The r_f is not affected by the back side. We denote the reflectance ratio of two mounts as

$$\alpha = \frac{m_d}{m_b}. \tag{3}$$

These ratios of the mounts are means of the entire mounts' images simply. The values of m_d and m_b can be calculated from images of the dark and bright mounts in the same environment used to record the documents. Then, Eq. (2) can be re-arranged to obtain the reflectance r_f, which suggests that the image



(a) Environment



(b) Electronic paper

Figure 4: Experimental setting.

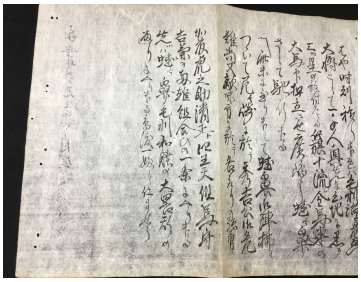


Figure 5: Unbounded ancient paper document.

of the front side is

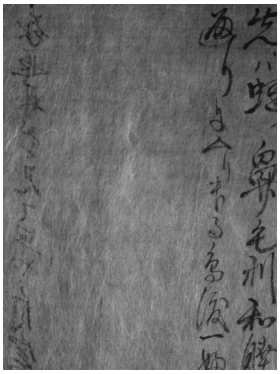
$$r_f = \frac{\alpha I_b - I_d}{\alpha - 1}. \tag{4}$$

4 Experiment

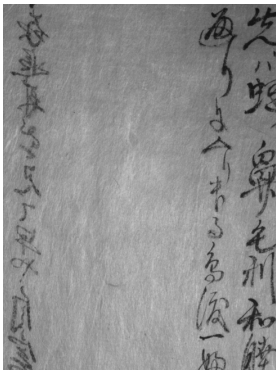
4.1 Experiment environment

We applied our proposed method with the optical model into an actual ancient paper document. This experiment was done inside a dark room to avoid the unexpected effects of ambient light. We used a monochrome camera (Grasshopper 3 GS3-U3-15S5M, Point Grey Inc.). The light source was a white LED. A glass plate was placed on the paper to flatten the surface, as shown in Fig. 4 (a). The target object was one piece of an ancient document, which was separated from the other pages in advance.

When the mount is physically changed, the paper document also moves. Therefore, an additional alignment process would be needed. In contrast, instead of switching the mounts physically, electronic paper can be used to change the reflectance, as shown in Fig. 4 (b). Hence, we utilized the electronic paper (DPT-RP1, Sony) as the mount to prepare dark and bright mounts effortlessly and avoid the alignment process. Note that this electronic paper display does not use light, so we need not be concerned about unexpected light emissions.



(a) On dark mount



(b) On bright mount

Figure 6: Images: The left part of the images is show-through from the back side. The show-through part of the left image is weaker than the right.

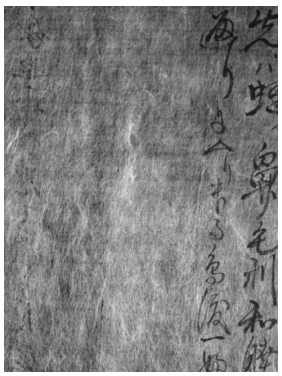


Figure 7: Result of show-through cancellation for one piece of paper inked on both-sides.

4.2 One piece of paper inked on both sides

An example of double-sided inked ancient paper is shown in Fig. 5. In this case, it is speculated that the writer made a mistake in the first part and reused the back side of this paper on the last page.

The image of the paper on the dark mount projected by the electronic paper is darker entirely than the image taken with the bright one. The left part of these images is inked on the back side, and the right part is inked on the front side, as shown in Fig. 6. The show-through characters from the back on dark mount disappeared somewhat because the overall intensity values went down and the contrast of the image also.

After processing the images with our proposed optical model solution, there is less show-through effect that appeared on the left part before about the expectation image Fig. 7 of its front side. The value of α used in this case was computed from images taken on the dark and bright mounts displayed by the electronic paper without the glass plate and the target object. This result is better than the result of the image

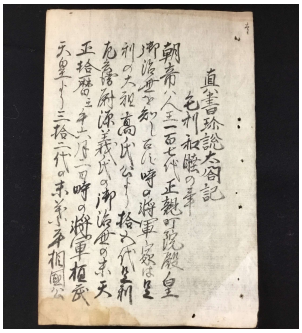
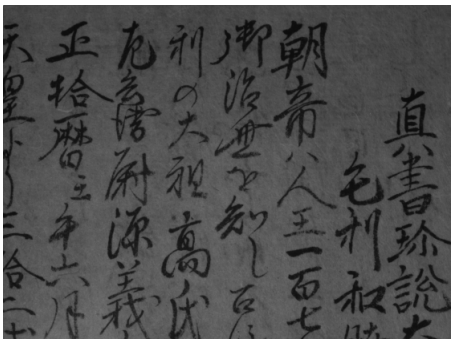
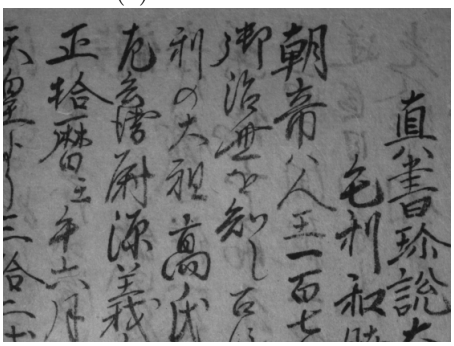


Figure 8: Usual appearance of one page of an ancient document. One side has written characters and it is folded once.



(a) On dark mount



(b) On bright mount

Figure 9: Images of the one folded piece of paper. Show-through from the back side is visible.

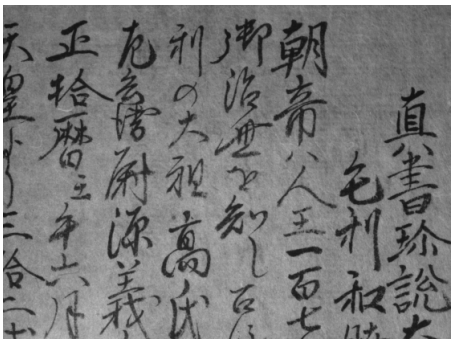


Figure 10: Result of the one piece of folded paper after show-through cancellation.

taken on the dark mount simply. However, the paper fiber has an impact on the whole image, which makes it somewhat difficult to read. This result is caused by the simple optical model used to process the factual ancient paper documents.

4.3 Folded one paper

Our proposed method can be applied on one piece of paper with a stitched binding structure, as shown in Fig. 1 (a). This piece of paper has characters written on only its front side, but the underlying image shows through slightly because of its thinness and structure. Thus, we performed an experiment in almost the same setting as that shown in Fig. 4. However, we used object shown in Fig. 8.

Fig. 9 shows the images captured on the bright and dark mounts. The expected front side image is better than in the previous experiment, as shown in Fig. 10. This suggests that our method matches this another structure although this result could be caused by the lower level of show-through. Furthermore, the overall effect from the paper fiber is less. Therefore, the readability is better.

5 Discussion

We succeeded in cancellation show-through from images of an ancient paper document inked on both sides and a piece of paper inked on one-side and folded. To do this, we used the images taken on dark and bright mounts. However, the results are not perfect because the proposed optical model ignores the effect of scattering. Therefore, we should improve the optical model by increasing its detail.

In addition, when it is not possible to open a stitched binding structure because of its physical damage and the need for conservation, the inside image of the ancient paper document may need to be restored instead of removed. For such cases, many methods such as [6] are too expensive and difficult to introduce. Hence, there is a need to develop simple approaches. Our study in this project would be supportive for the imaging inside the literature.

Acknowledgement

This work is partly supported by the National Institute of Japanese Literature and JST CREST JPMJCR1764. We are grateful to Prof. Kazuaki Yamamoto for his ancient paper document and helpful advice.

References

[1] J. Dambrogio, A. Ghassaei, D. S. Smith, H. Jackson, M. L. Demaine, G. Davis, D. Mills, R. Ahrendt, N. Akker-

man, D. van der Linden, and E. D. Demaine, "Unlocking history through automated virtual unfolding of sealed documents imaged by x-ray microtomography," *Nature Communications*, vol. 12, Mar 2021.

[2] L. Nodari and P. Ricciardi, "Non-invasive identification of paint binders in illuminated manuscripts by er-ftir spectroscopy: a systematic study of the influence of different pigments on the binders' characteristic spectral features," *Heritage Science*, vol. 7, Feb 2019.

[3] G. Sharma, "Show-through cancellation in scans of duplex printed documents," *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 736–754, 2001.

[4] B. Ophir and D. Malah, "Show-through cancellation in scanned images using blind source separation techniques," in *2007 IEEE International Conference on Image Processing*, IEEE, 2007.

[5] Y. Chen, "Scanning double-sided documents without incurring show-through by learning to fuse two complementary images using multilayer perceptron," *PLOS ONE*, vol. 12, p. e0176969, May 2017.

[6] A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, J. Romberg, and R. Raskar, "Terahertz time-gated spectral imaging for content extraction through layered structures," *Nature Communications*, vol. 7, Sep 2016.