2020 7th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS) | 978-1-7281-6246-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICCSS52145.2020.9336942

STRONG SHADOW REMOVAL OF TEXT DOCUMENT IMAGES BASED ON BACKGROUND ESTIMATION AND SHADING SCALE

Bingshu Wang[1], *Shuang Feng*[2], *C.L. Philip Chen* [3,4,*]

[1] School of Software, Northwestern Polytechnical University Taicang Campus, Suzhou, China.
[2] School of Applied Mathematics, Research Center for Applied Mathematics and Interdisciplinary Science, Beijing Normal University, Zhuhai, China
[3] School of Computer Science and Engineering, South China University of Technology, Guangzhou, China.
[4] Faculty of Science and Technology, University of Macau, Macau, China
Email: wangbingshu@nwpu.edu.cn, fengshuang@bnu.edu.cn, philip.chen@ieee.org

ABSTRACT

Shadows may bring uncomfortable perception when taking text document images. Previous work mainly focuses on weak shadow removal issue. However, strong shadow removal is still a challenging task. This paper proposes a method based on background estimation and shading scale to remove strong shadows from text document images. Firstly, background color estimation is designed by a number of iterations through neighboring pixels' propagation. Then, umbra and penumbra are separated by morphological operations and processed by a divide-and-conquer strategy. For umbra, shading scale strategy is exploited to obtain unshadowed result. For penumbra, background replace strategy is designed to remove the shadow regions. Finally, a reference background and a text binary image are combined to generate unshadowed image. Experiments conducted on some typical strong shadow images demonstrate the effectiveness of our method.

Index Terms— Document shadow removal, Background estimation, Umbra and penumbra, Shading scale

1. INTRODUCTION

Nowadays portable mobile devices such as smartphones and cameras have become very popular due to their various functions [1]. People are likely to use smartphones to take text document photos instead of scanners. When these photos are captured, the light is easily occluded by other objects such as the device self and the photographer's hands or arms, resulting in shadows [2, 3, 4]. In most cases, shadows are useless and even interference for document clarity [5, 6, 7]. For example, Fig. 1 (a) is a text document image with strong shadows, and (b) is the output image after removing shadows.

This work was funded in part by the National Key Research and Development Program of China under number 2019YFA0706200 and 2019YFB1703600, partly by the National Natural Science Foundation of China grant under number 61702195, 61751202, U1813203, U1801262, 61751205, in part by the Science and Technology Major Project of Guangzhou under number 202007030006 and Foundation for Young Talents in Higher Education of Guangdong, China under number 2018KQNCX324, in part by The Science and Technology Development Fund, Macau SAR (File no. 079/2017/A2, and 0119/2018/A3), in part by the Multi-year Research Grants of University of Macau. (Corresponding author: philip.chen@ieee.org).

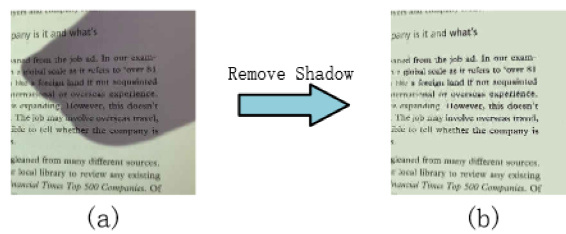

Fig. 1. An example for document image shadow removal. (a) input image, (b) output image after removing shadow.

In past decade, many publications are proposed to remove shadows or realize text binarization [8, 9, 10, 11, 12, 13]. Oliveira et al [9] utilized natural neighbor interpolation to estimate shading image. Bako et al [2] proposed to estimate background image by patch level clustering and global reference image by entire image clustering. It mainly concentrates on the scenes that document images have a uniform colored background. Kligler et al [11] used a 3D point cloud transformation technique to generate a new representation of image for shadow removal algorithm [2].

Shah et al [6] proposed to estimate shading map and reflectance images in an iterative manner. Jung et al [12] explored a water-filling approach to rectify illuminations by mapping image into topographic structure. Wang et al [13] designed an effective background estimation method to remove shadows. In addition, there are many other techniques [14, 15, 16, 17, 18, 19, 20, 21, 22]. Most of the mentioned methods can handle the weak shadows well. However, they face challenges in strong shadows of document images.

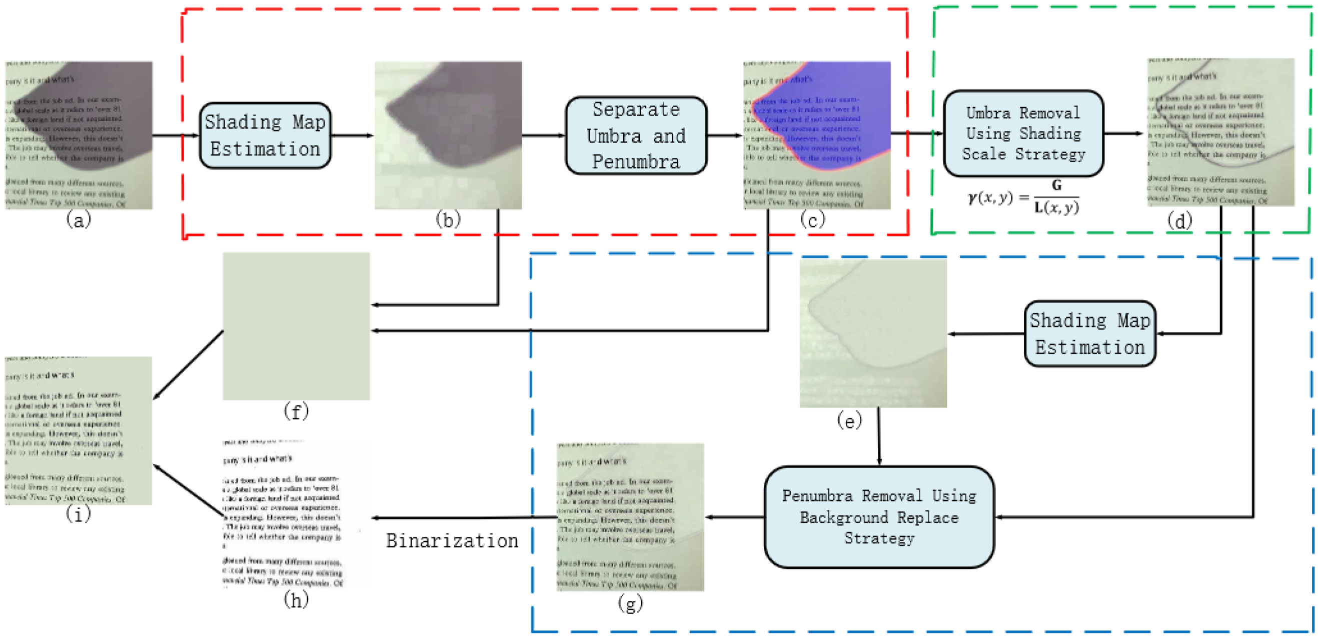

Fig. 2. The flowchart of shadow removal of text document image. The dash lines represent different modules. (a) input image, (b) estimated background image, (c) separated umbra and penumbra, (d) image with removed umbra, (e) estimated background based on (d) image, (f) reference background image, (g) image with removed penumbra, (h) binary image based on (g) image, (i) final output image.

In this paper, we focus on removing strong shadows from text document images. Assuming the images have a uniform color and only include texts. The contributions include several parts. Firstly, a background estimation strategy is used to get background color image and separate umbra and penumbra. Secondly, we propose penumbra removal using background replace strategy. Experiments on multiple images indicate the effectiveness of our method.

2. THE PROPOSED METHOD

The proposed method includes three parts as shown in Fig. 2. The red dotted line shows the process of background estimation (Fig. 2(b)) and the separation of umbra and penumbra (Fig. 2(c)). The green dotted line gives the umbra removal process (Fig. 2(d)). The blue dotted line presents the penumbra removal process, which is based on the result of umbra removal. Finally, the reference background image (Fig. 2 (f)) is combined with binary image (Fig. 2 (h)) to generate the final output (Fig. 2 (i)).

2.1. Background Estimation

From the observation of text document images, we found that texts are always black while the background is white. For one pixel belonging to text, its background color can be estimated by its neighboring information [13]. This can be regarded as a background propagation process. For an image with RGB channels, each channel is processed to obtain background color image as follows,

$$\begin{aligned} \mathbf{V}_{max}(x,y) &= max(V(i,j)), (i,j) \in \mathbf{S} \\ \mathbf{V}_{min}(x,y) &= min(V(i,j)), (i,j) \in \mathbf{S} \end{aligned} \tag{1}$$

$$\alpha = (\mathbf{V}_{max}(x,y) - \mathbf{V}_{min}(x,y))/\mathbf{V}_{max}(x,y) \tag{2}$$

$$\mathbf{L}(x,y) = \mathbf{V}_{max}(x,y) * (1-\alpha) + \mathbf{V}_{min}(x,y) * \alpha \tag{3}$$

In Eq.1, $\mathbf{S}$ denotes the neighboring size (e.g., 3×3). The $\mathbf{V}_{max}(x,y)$ is defined as local maximum value and the $\mathbf{V}_{min}(x,y)$ is local minimum value. Maximum and minimum values are used to compute a fusion factor α. It helps to avoid extreme white noise. $\mathbf{L}(x,y)$ is the estimated background color. In practice, the estimation process has a number of iterations to avoid texts as residual components. In practice, 3 iterations with a size of 3×3 are enough, for example, Fig. 2 (b). With this image obtained, we can derive a reference background image Fig. 2 (f) by computing the average background color of non-shadow regions.

The estimated background color image (Fig. 2 (b)) is also a shading map. The dark part belongs to the shadow region while the bright part belongs to non-shadow region. Therefore, the image can be used to separate umbra and penumbra.

2.2. Separate Umbra and Penumbra

A series of operations are designed to separate umbra and penumbra. Firstly, three pre-processing steps are carried out: convert color image into gray image, medium filtering and binarization. Then umbra region can be obtained, which is shown in blue in Fig. 2 (c). A number of dilation operations are carried out to produce a shadow mask. The shadow mask is subtracted by umbra mask to generate penumbra as shown in red in Fig. 2 (c). Next, umbra and penumbra will be removed, respectively.

2.3. Umbra Removal Using Shadow Scale Strategy

It can be easily seen that umbra is darker than the non-shadow regions. To relight the umbra, a shadow scale is exploited as follows, The umbra can be removed by applying a shading scale **r**(x,y)

$$\mathbf{r}(x,y) = \frac{\mathbf{G}}{\mathbf{L}(x,y)} \tag{4}$$

where $\mathbf{G} = \frac{1}{n}\sum \mathbf{L}(i,j), \ \ (i,j) \in$ **Non-shadowed Region**.

For each umbra point, it can be relighted by a multiplication of shadow pixel's intensity and shading scale. Three channels for RGB image are calculated, respectively. The follow equation shows the result, **Vu**(x,y) is defined as relighted umbra at point (x,y).

$$\mathbf{Vu}(x,y) = \mathbf{V}(x,y)\mathbf{r}(x,y) \tag{5}$$

2.4. Penumbra Removal Using Background Replace Strategy

Penumbra removal is based on the result of umbra removal (Fig. 2 (d)). The penumbra is the region that between non-shadow and umbra. The intensity become darker from non-shadow region to umbra side.

To realize penumbra removal, a background color image is estimated firstly, which is used to replace the non-text points in penumbra. Penumbra region includes texts and non-texts, which can be distinguished by a binarization operation. Then, we use estimated background color image (i.e.,Fig. 2 (e)) to replace the non-texts, and produce image without penumbra, for example, (Fig. 2 (g)). It can be seen that some borders might be left. To make the image be good perception, a binarization operation [8] is applied to generate a binary image. Finally, the binary image is combined with reference background image to generate unshadowed image.

3. EXPERIMENTAL RESULTS

In our experiments, the proposed method is used to compare with some state-of-the-art approaches [2, 12, 13]. All the methods are implemented by Visual C++ under Windows OS with a Core(TM) i5-5200U CPU @2.2GHz. To our knowledge, there are few datasets that include strong shadows in text document images. Thus, the experiments are performed on some images collected from internet. All the test images are typical cases that have strong shadows. Due to the lack of ground turth, herein, we choose to use visual perception as an intuitive evaluation metric.

Visual results are presented in Fig. 3 and quantitative comparison for running time are provided in Table 1. It can be seen from the Fig. 3 that our method can achieve better results than the compared approaches. The method [2] is able to produce good results for umbra removal. However, it fails to handle the shadow boundaries (penumbra). The approach proposed by [12] can locate the shadow regions accurately, but introduces other colors to the shadow regions, which seems differential. The strategy designed by [13] can remove umbra well, but have penumbra regions left.

It can be concluded from Table 1 that our method outperforms [2, 12] in the running speed. Our method runs slower than the method [13], because it takes more times of estimation background step for our method than that of [13]. It should be noted that although our method can remove the shadows well, there is text intensity difference between the shadow region and non-shadow region. This should be invested in future to keep all the texts with a similar intensity.

Table 1. The Running time (seconds) comparisons of our method, [2], [12] and [13] on two different sizes of images.

Size (pixels)	Bako et al [2]	Jung et al [12]	Wang et al [13]	Ours
640×480	1.172	0.927	0.115	0.707
822×616	1.952	1.614	0.181	1.173

4. CONCLUSION

In this paper, we propose a method to remove strong shadows from text document images. It is based on the background color estimation and Retinex theory. Umbra and penumbra are separated and processed by a divide-and-conquer manner. The umbra is enhanced by the implementation of shading scale estimation strategy. With the umbra removal result obtained, Penumbra is processed subsequently. Estimated background is exploited to replace the non-texts in penumbra. Finally, unshadowed image is generated by a combination of a global reference background image and a text binary image. Experimental results show the effectiveness of our method.

Acknowledgment

We really thank Seungjun Jung (KAIST, South Korea) and Steve Bako (University of Berkeley, Santa Barbara) for their code sharing.

Fig. 3. Visual comparison of four methods on real world images.

5. REFERENCES

[1] B. Jiang, S. Liu, S. Xia, X. Yu, and Y. Gao, "Video-based document image scanning using a mobile device," in *2015 International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 2016.

[2] S. Bako, S. Darabi, E. Shechtman, and J. Wang, "Removing shadows from images of documents," in *Asian Conference on Computer Vision*. Springer,Cham, 2016, pp. 173–183.

[3] M. Brown and Y. Tsoi, "Geometric and shading correction for images of printed materials using boundary," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1544–1554, 2006.

[4] L. Zhang, A. M. Yip, and C. Tan, "Removing shading distortions in camera-based document images using inpainting and surface fitting with radial basis functions," in *Ninth International Conference on Document Analysis and Recognition*. IEEE, 2007, vol. 2, pp. 984–988.

[5] J. Zhao, C. Shi, F. Jia, Y. Wang, and B. Xiao, "An effective binarization method for disturbed camera-captured document images," in *16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*. 2018, IEEE.

[6] V. Shah and V. Gandhi, "An iterative approach for shadow removal in document images," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2018, pp. 1892–1896.

[7] J. Wang and Y. Chuang, "Shadow removal of text document images by estimating local and global background colors," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 1534–1538.

[8] D. Bradley and G. Roth, "Adaptive thresholding using the integral image," *Journal of graphics tools*, vol. 12, no. 2, pp. 13–21, 2007.

[9] D. M. Oliveira, R. D. Lins, and G. Silva, "Shading removal of illustrated documents," in *International Conference Image Analysis and Recognition*. Springer, 2013, pp. 308–317.

[10] C. Xiao, D. Xiao, L. Zhang, and L. Chen, "Efficient shadow removal using subregion matching illumination transfer," in *Computer Graphics Forum*. Wiley Online Library, 2013, vol. 32, pp. 421–430.

[11] N. Kligler, S. Katz, and A. Tal, "Document enhancement using visibility detection," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2018, pp. 2374–2382.

[12] S. Jung, M. Hasan, and C. Kim, "Water-filling an efficient algorithm for digitized document shadow removal," in *Asia Conference on Computer Vision(ACCV)*. Springer, 2018, pp. 398–414.

[13] B. Wang and C.L. Chen, "An effective background estimation method for shadows removal of document images," in *Internal Conference on Image Processing*. IEEE, 2019, pp. 3611–3615.

[14] Q. Yang, K. Tan, and N. Ahuja, "Shadow removal using bilateral filtering," *IEEE Transactions on Image processing*, vol. 21, no. 10, pp. 4361–4368, 2012.

[15] H. Gong and D. Cosker, "Interactive shadow removal and ground truth for variable scene categories," in *2014-Proceedings of the British Machine Vision Conference*. University of Bath, 2014.

[16] M. Gryka, M. Terry, and G. Brostow, "Learning to remove soft shadows," *ACM Transactions on Graphics*, vol. 34, no. 5, pp. 153, 2015.

[17] L. Ma, J. Wang, E. Shechtman, K. Sunkavalli, and S. Hu, "Appearance harmonization for single image shadow removal," in *Computer Graphics Forum*. Wiley Online Library, 2016, vol. 35, pp. 189–197.

[18] X. Yu, G. Li, Z. Ying, and X. Guo, "A new shadow removal method using color-lines," in *International Conference on Computer Analysis of Images and Patterns*. Springer, 2017, pp. 307–319.

[19] B. Wang, C. L. Chen, Y. Li, and Y. Zhao, "Hard shadows removal using an approximate illumination invariant," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2018, pp. 1628–1632.

[20] X. Hu, L. Zhu, C. Fu, J. Qin, and P. Heng, "Direction-aware spatial context features for shadow detection," *Proceedings of the Computer Vision and Pattern Recognition*, 2018.

[21] B. Wang and C.L. Chen, "Moving cast shadows segmentation using illumination invariant feature," *IEEE Transactions on Multimedia*, vol. 22, pp. 2221–2233, 2019.

[22] G. Meng, S. Xiang, N. Zheng, and C. Pan, "Nonparametric illumination correction for scanned document images via convex hulls," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1730–1743, 2013.