AN EFFECTIVE BACKGROUND ESTIMATION METHOD FOR SHADOWS REMOVAL OF DOCUMENT IMAGES

Bingshu Wang, Student Member IEEE, C. L. Philip Chen, Fellow IEEE

Faculty of Science and Technology, University of Macau, Macau, China

ABSTRACT

Shadows of document images bring about difficulties for digitization application and uncomfortable perception in vision. This paper proposes an effective method to remove shadows from the single document images, which contains two stages: shadow detection and shadow removal. For the shadow detection, an iterative neighboring information-based approach is designed to estimate pixel-wise local background color image, which can be used to generate a shadow map. For the shadow removal, a global reference background color is obtained from the local background color image. The shadow scale, which is applied to relight shadowed regions, can be estimated by local and global background color. Moreover, a tone fine-tuning process is designed to make the output colors be normal values. Experiments on document images indicate that the proposed method can produce high-quality unshadowed document images with a comparable efficiency.

***Index Terms*—** Shadow removal, Document images, Neighboring information, Local background color

1. INTRODUCTION

The popularization of smart phones makes taking photos more convenient and diverse. With a smart phone in hand, people can share their pictures anytime through the Mobile Internet. A gradual trend is that more and more people choose mobile phone as a mainstream document capture device rather than scanner or camera. However, illumination changes in environments may result in shadows, which easily causes the degradation of document images [1, 2]. This motivates us to design algorithms to detect and remove shadows from document images.

Normally, shadows are generated when illumination source is partially occluded by objects [3, 4, 5, 6]. Most of shadow detection and removal algorithms focus on the outdoor scenes under the sunshine [6, 7, 8, 9]. The regions covered by shadows usually are objects such as ground, grass and wall that have background colors. Even for the fixed and static images captured by surveillance cameras [10, 11], the background is also colorful. However, dealing with document images is different [1, 12, 13]. In most cases, texts in document images are printed in black and the background is a constant color [14]. When the shadow regions and texts are both dark, the distinction becomes difficult. In this paper, our objective is to design a method to remove shadows of document images and preserve text details simultaneously.

In a document image with shadows, the background of shadow regions tends to be dark while the background of lit regions tends to be bright [13]. This inspires us to estimate the darkness degree of an image (i.e., shadow map). In return, a shadow map can assist shadow removal [14].

In this paper, we propose a new shadow removal method without any user interaction. Two main contributions are included. First, we design a strategy that takes advantage of neighboring information of each pixel to predict its background color, which generates shadow map and estimate a global reference background color. Second, we propose a technique that adopt shadow scale, which is computed by global and local background color, to generate unshadowed image. Experiments performed on a document dataset show the effectiveness of our method.

2. RELATED WORK

Previous work on shadow removal paid more attention to natural scenes [6, 9, 11, 15, 16, 17, 24]. The methods can be classified as two categories: automatic approaches [3, 6, 18, 19] and user interaction approaches [4, 15, 20]. The difference between these categories is the techniques used for shadow detection. The shadow removal part is always an automatic task. Guo et al [3] designed an automatic method using paired regions for single image shadow removal. Xiao et al [18] proposed a method using subregion matching illumination transfer for automatic shadow removal. Albu et al [24] proposed recursive filtering and contrast stretching techniques for shadow compensation. Gong et al [20] presented an interactive shadow method using two rough user inputs. Yu et al [15] suggested to use color-lines as interactions for shadow elimination. However, these methods have difficulties in removing document shadows.

This work was supported partly by the National Natural Science Foundation of China (61751202, 61751205, 61572540, U1801262, U1813203), Macau Science and Technology Development Fund (FDCT, 079/2017/A2, 024/2015/AMJ, 019/2015/A1), and Multiyear Research Grants of the University of Macau. (Corresponding author: philipchen@umac.mo).

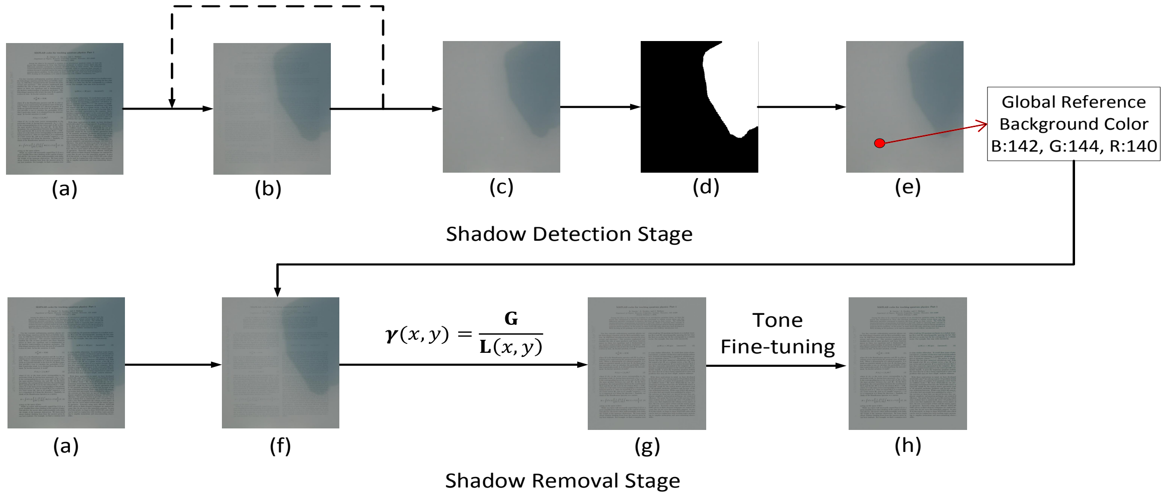

Fig. 1. The Flowchart of document shadows removal. The dash line represents an iterative process to estimate the local background color. (a) Input image, (b) Local background color image (neighbor:5×5), (c) Background color image after several iterations, (d) Shadow map, (e) Find global reference background color, (f) Local background color image (neighbor:3×3), (g) Coarse unshadowed result with MSE 30.73, (h) Final unshadowed result with MSE 22.26.

Some progresses have been made aiming to remove shadows from images of documents [1, 12, 13, 14, 21, 22]. Most of these methods are without user-aiding. Oliveira et al [12] used natural neighbor interpolation to estimate the shading image. However, it may fail when interpolation is excessive. Shah et al [13] proposed an iterative approach that regards the shadow removal as a problem of estimating the shading and reflectance images. Based on the observation that documents have a constant colored background, Bako et al [14] proposed a patch level clustering to estimate local background, and entire image clustering to estimate global reference background. Our method also estimates the local and global background, but it is with high efficiency. Experimental results demonstrate this point.

3. THE PROPOSED METHOD

The proposed method mainly includes two stages: shadow detection and shadow removal. The shadow detection stage consists of estimations of local background color image, shadow map and global reference background color. The shadow removal stage focuses on the estimation of the shadow scale $\gamma(x,y)$ which is applied to re-light shadow regions. Moreover, a tone fine-tuning compensation follows. The flowchart is shown in Fig. 1.

3.1. Local and Global Background Color Estimation

It is observed that texts in document images tend to be dark while the background tends to be bright. To find the local background color (bright), for each pixel in an image, its neighboring pixels provide valuable information for background color prediction. We design an effective selection policy to estimate the local background color as follows

$$\begin{aligned} \mathbf{I}_{max}(x,y) = max(I(i,j)), (i,j) \in \mathbf{W} \\ \mathbf{I}_{min}(x,y) = min(I(i,j)), (i,j) \in \mathbf{W} \end{aligned} \quad (1)$$

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

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

In Eq.1, $\mathbf{W}$ represents the neighboring window (like 5×5). The maximum value $\mathbf{I}_{max}(x,y)$ and the minimum value $\mathbf{I}_{min}(x,y)$ in the neighboring window are selected to compute a fusion factor α which is used to correct the extreme white noise. The local background color $\mathbf{L}_(x,y)$ is estimated by Eq.3 (shown in Fig. 1 (b)). For boldface type of texts, some residual components might persist in the local background image. To address this, an iteration process is carried out until there is no changes between two consecutive iteration results. In practice, three iterations can achieve good result with a neighboring window 5×5 employed.

An example of local background color image is shown in Fig. 1 (c). A shadow map (shown in Fig. 1 (d)) can be easily generated by applying Otsu [23] to Fig. 1 (c). Given the local background color image and shadow map, global reference background color can be obtained by an approximation means:

$$\mathbf{G} = \frac{1}{n}\sum \mathbf{L}(i,j), \quad (i,j) \in \textbf{Unshadowed Region} \quad (4)$$

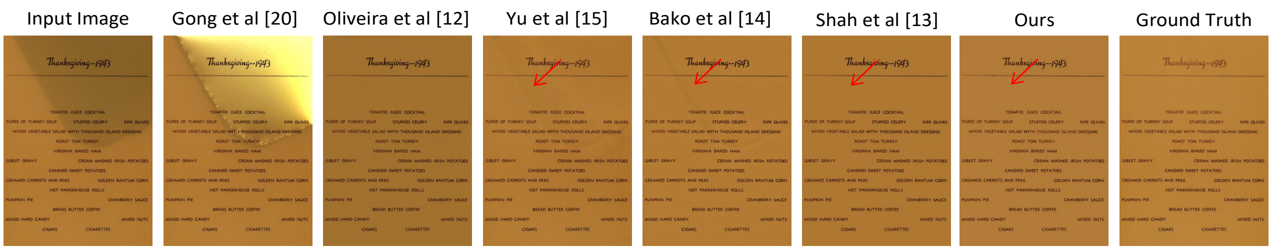

Fig. 2. Visual comparison of six methods on an image captured in a controlled environment.

3.2. Shadow Removal

Generally, the shadow regions are darker than the unshadowed regions. The shadow regions can be relighted by applying a shadow scale **r**(x,y)

$$\mathbf{r}(x, y) = \frac{\mathbf{G}}{\mathbf{L}(x, y)} \tag{5}$$

It should be noted that the **L**(x,y) in Eq.5 is derived from a 3×3 neighbor window, which ensures to acquire a relative accurate background color.

For each pixel, three channels in RGB color space are calculated respectively. For a document image with uniform illuminations, unshadowed image (Fig. 1 (g)) can be generated by a multiplication of shadow pixel's intensity and **r**(x,y). However, when the illumination is non-uniform, some artifacts may be produced and they should be corrected.

3.3. Tone Fine-tuning

This section aims at solving text artifacts problem. For a local background color image like Fig. 1 (f), we define tone scale as a ratio between lit regions and shadow regions.

$$\tau = \frac{lit\_bg}{shadow\_bg} \tag{6}$$

where lit_bg represents the average background color of unshadowed regions and $shadow\_bg$ represents the average background color of shadow regions. τ is calculated from a perspective of entire image, which is different from pixel's shadow scale **r**(x,y). Intrinsically, tone scale is a scale between bright regions and dark regions. For a pixel, when its shadow scale $\mathbf{r}(x, y)$ is abnormal, τ can provide a helpful reference.

For pixels belonging to text, if they are dark enough and the **r**(x,y) is much larger than the τ, the shadow scale should be reduced to keep the text details. After the tone fine-tuning process, a better shadow removal result is expected, which is illustrated in MSE comparison of Fig. 1 (g) and (h).

Table 1. Quantitative comparisons of the proposed method with some state-of-the-art approaches on two evaluation metrics: MSE and $ErrorRatio$.

Evaluation Metric	*MSE*	*Error Ratio*
Gong et al [20]	661.02	1.245
Oliveira et al [12]	260.34	0.903
Yu et al [15]	72.24	0.497
Bako at al [14]	66.63	0.476
Shah at al [13]	73.97	0.518
Ours	**58.28**	**0.456**

4. EXPERIMENTAL RESULTS

In this section, we compare our method with some state-of-the-art approaches [20, 12, 15, 14, 13]. The experiments are performed on a controlled environment dataset (81 images) with ground truth and some real world images [14]. For fair comparison, these methods are conducted on 3.5GHz Xeon CPU machine with Windows OS and there is not special optimization techniques used. Bako's approach and ours [14] are implemented by C++, and Shah's approach[13] is implemented by Matlab.

4.1. Quantitative Comparison

We present a quantitative evaluation for six methods by Mean Squared Error (MSE) and $ErrorRatio$ [20] as follows:

$$MSE(R, GT) = \frac{1}{n} \sum (R(x, y) - GT(x, y))^2 \tag{7}$$

$$ErrorRatio = \frac{RMSE(R, GT)}{RMSE(I, GT)} \tag{8}$$

where $RMSE$ is the root MSE (i.e., $\sqrt{MSE}$) and R, GT, and I represents the result of unshadowed image, ground truth, and input image, respectively.

The comparison of methods is given in Table 1, which implies that the proposed method outperforms the compared

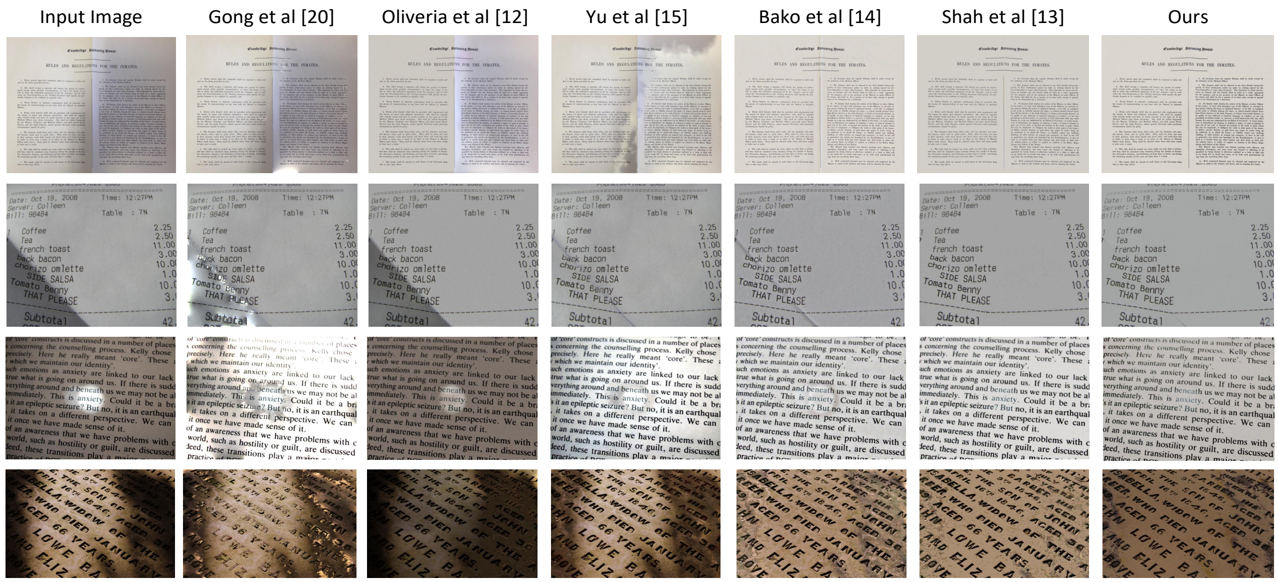

Fig. 3. Visual comparison of six methods on real world images.

approaches. Table 2 presents a running time comparison of our method and two state-of-the-art algorithms. Our method is almost three times faster than Bako's [14] and Shah's [13]. The clustering process in Bako's takes much time. In terms of complexity, Shah's approach and our method are both $O(n)$, where n represents the number of pixels in an image and each pixel needs constant operations. The iteration number of Shah's method is three times than ours, which leads to its lower efficiency than ours. Table 2 indicates the proposed method's high efficiency.

4.2. Visual Results

Figure 2 shows visual results of these methods under controlled environments. The method proposed by [20] relights shadow regions excessively. Our method and [12, 13] produce results that are less affected by shadow boundaries while other methods produce some boundary vestiges.

Table 2. Running time (seconds) comparisons of our method, [14] and [13] on three different sizes of images.

Size (pixels)	***1287×1638***	***1632×1050***	***534×1194***
Bako et al [14]	3.366	2.844	1.062
Shah at al [13]	3.432	2.207	1.340
Ours	1.148	0.845	0.363

The comparisons of four real word images are given in Fig. 3. The method proposed by [20] fails to estimate shadow statistics even with some user-aiding cues provided. The method [12] depends too much on interpolation so that it is difficult to remove document shadows. Yu's algorithm [15] can remove the soft shadows, but face challenges on hard shadows. Bako's method [14] can provide good results but leave some artifacts along the boundaries. Shah's [13] results tend to be similar with ours on perception due to the adoption of background estimation. Experimental results show that our method can generate unshadowed images well.

5. CONCLUSION

In this paper, we proposed an automatic method to remove shadows of document images. It uses pixel's neighboring information to estimate the local background color image iteratively, which makes shadow map detection an easy work. By applying shadow scale to input image, unshadowed output can be generated with a tone fine-tuning followed. Quantitative evaluation and visual comparisons indicate that the proposed method outperforms some state-of-the-art approaches. In the future, our work will focus on the shadow removal of document images with figures, and improve efficiency further.

Acknowledgment

We really thank Steve Bako (University of Berkeley, Santa Barbara), Xiaoming Yu (Peking University) for interesting discussions and their code sharing.

6. REFERENCES

[1] 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.

[2] G. D. Finlayson, M. Drew, and C. Lu, "Entropy minimization for shadow removal," *International Journal of Computer Vision*, vol. 85, no. 1, pp. 35–57, 2009.

[3] R. Guo, Q. Dai, and D. Hoiem, "Paired regions for shadow detection and removal," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 12, pp. 2956–2967, 2013.

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

[5] Li. 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.

[6] 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.

[7] B. Wang, Y. Yuan, Y. Zhao, and W. Zou, "Adaptive moving shadows detection using local neighboring information," in *Asian Conference on Computer Vision*. Springer, 2016, pp. 521–535.

[8] L. Liu, C. L. Chen, Y. Zhou, and X. You, "A new weighted mean filter with a two-phase detector for removing impulse noise," *Information Sciences*, vol. 315, pp. 1–16, 2015.

[9] S. Bell, K. Bala, and N. Snavely, "Intrinsic images in the wild," *ACM Transactions on Graphics*, vol. 33, no. 4, pp. 159, 2014.

[10] M. Russell, J. Zou, G. Fang, and W. Cai, "Feature-based image patch classification for moving shadow detection," *IEEE Transactions on Circuits and Systems for Video Technology*, 2017.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] J. Zhou, L. Chen, C. L. Chen, Y. Zhang, and H. Li, "Fuzzy clustering with the entropy of attribute weights," *Neurocomputing*, vol. 198, pp. 125–134, 2016.

[17] 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.

[18] 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.

[19] T. Zhou, P. Krahenbuhl, and A. Efros, "Learning data-driven reflectance priors for intrinsic image decomposition," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 3469–3477.

[20] 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.

[21] 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.

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

[23] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[24] F. Albu, C. Vertan, C. Florea and A. Drimbarean, "One scan shadow compensation and visual enhancement of color images," in *2009-Proceedings of the 16th IEEE International Conference on Image Processing*. IEEE, 2009, pp. 3133–3136.