Shading Removal of Illustrated Documents

Daniel Marques Oliveira[1], Rafael Dueire Lins[1], and Gabriel França Pereira e Silva[1,2]

[1] Universidade Federal de Pernambuco, Recife, Brazil
[2] Universidade Federal Rural de Pernambuco, Garanhuns, Brazil
daniel.moliveira@ufpe.br, {rdl,gfps}@cin.ufpe.br

Abstract. Pictures of documents have non-uniform illumination causing shading which may yield to bad quality image for human visualization and unsuitable for some image processing algorithms. Most algorithms do not consider the scenario in which documents have large non-uniform regions such as photographs and illustrations. This paper proposes an algorithm to remove the shading of such documents. Once the background is identified, Natural Neighbor Interpolation estimates the shading for non-background pixels. The algorithm performed well on 33 synthetic images using SSIM and PSNR measures. The same quality of performance was confirmed in "real-world" images.

Keywords: Shading removal, Illumination normalization, Enhancement.

1 Introduction

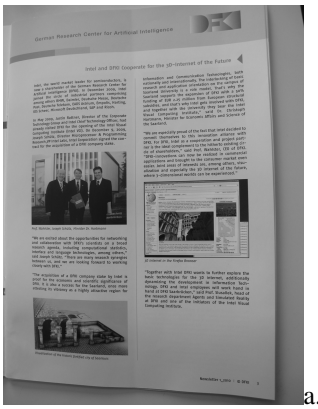
Portable digital cameras are widely used for the capture of document content due its low price, portability and image quality. Pictures of documents obtained by this device often have non-uniform illumination. Figure 1 presents two examples of photographed documents, in which one may observe shade differences along the images.

Due to its intrinsic uneven nature, shading removal is not a straightforward process. One strategy is to see the original document image as the result of the addition of two images: 1 – the background image, with the intensity exclusively related to the light color; 2 – the image with uniform illumination. Shading removal excludes the background image from the input image yielding only the uniformly illuminated one.

Most of times, algorithms identify pixels that belong to the background. The estimation of the unknown shading (foreground pixels) is done by interpolation or fitting of the known shading (background pixels). Once the shading of the whole image is obtained, it is removed from the input image. If the document contains only text, the estimation is straightforward. Otherwise, the estimation it must consider that the large foreground regions (e.g. pictures, charts, etc.), such the ones in Figure 1.

This paper presents an efficient algorithm to remove shading of color document images with pictures, drawings and photos such as the one in Figure 1. It extends the work presented by the authors of this paper in reference [12] by changing the shading estimation to Natural Neighbor Interpolation (NNI). This type of interpolation is claimed by Amidror [1] to be the best method when the distribution of known points varies highly, which is the case of the illustrations boundaries in the image of a document. The next sections present related works, the proposed algorithm and performance measurement of current work.

M. Kamel and A. Campilho (Eds.): ICIAR 2013, LNCS 7950, pp. 308–317, 2013.


a.

b.

Fig. 1. Camera documents with drawings and pictures

2 Related Works

The shading removal algorithms are classified by its shading estimation: interpolation and fitting. For the former approach, the resulting estimation passes exactly into the known points. For the shading fitting process, the final estimation is close to the known shading values. The next subsections overview these techniques.

2.1 Shading Interpolation

Gatos and his colleagues [6] interpolate the shading with equation (1), where: $I(x,y)$is the input image with low-pass Wiener filter, $N(x,y)$ is the Niblack's binarization of $I(x,y)$; $B(x,y)$is the interpolated background image. The window of size $dx \times dy$ should cover at least two characters [6]. In reference [7], Niblack's [9] algorithm was replaced with the one by Savoula [14].

$$B(x,y) = \begin{cases} I(x,y) & if N(x,y) = 0 \\ \dfrac{\sum_{i=x-dx}^{x+dx}\sum_{j=y-dy}^{y+dy} I(i,j)(1-N(i,j))}{\sum_{i=x-dx}^{x+dx}\sum_{j=y-dy}^{y+dy}(1-N(i,j)} & if N(x,y) = 1 \end{cases} \quad (1)$$

The approaches in references [12][13], proposed by the authors of this work, divide the image into blocks. The background colors of the blocks are obtained by the component modes on a window. The blocks that belong to the background are identified by the threshold region growing segmentation method. The shading of foreground values are estimated iteratively similar to the region growing algorithm.

2.2 Shading Fitting

Zhang and his colleagues [18] identify the background region by edge detector and morphological operations on the value component of the HSV model. The Radial Basis Function (RBF) fitting is used to estimate the shading of the entire image.

Lu and Tan [16] estimate the shading by iteratively applying the Savitzky-Golay filter, which is a local polynomial fitting. Lee and his collaborators [19] remove shading of grayscale document images with pictures and illustrations obtained by scanners. First, they obtain the background and foreground regions using Söbel edge detector in several directions. Then, the background values for the foreground regions are estimated by vertical line segments. Finally, the shading is removed from the original image. This method is further improved by Athimethphat and Patanavijit [25]. Meng *et al* [26] defines a Convex Hull with uniform background and the shading are estimated by fitting of points that are outside the Convex Hull.

3 The Proposed Algorithm

The new algorithm proposed here splits the document image into blocks and encompasses: background identification; Foreground shading estimation; Shading removal.

3.1 Document Background Identification

The blocks are segmented by the region growing method [24]. Two neighboring blocks belong to the same area if the Euclidean distance of their background colors is smaller than a given threshold. The histograms of a sliding window for every RGB component are computed; the RGB modes define the background color. Once, defined all regions, the one with most elements are set to the document background. Herein, dimensions of 5x5 and 15x15 pixels were set to the block and sliding window sizes, respectively. One may observe that the sliding window of neighboring blocks share pixels, increasing the "continuity" of the background values.

3.2 Foreground Shading Estimation

The paper background blocks were obtained in the previous step. There is a high density variation of background blocks close to the illustrations. Document images are a few Mega pixel large. Such large size brings difficulties in estimating the unknown values. Natural Neighbor Interpolation is used in this step for the several reasons that are explained in the next section.

3.2.1 Natural Neighbor Interpolation

The estimation of values between non-uniform points is known as "scattered data interpolation". Amidror [1] provides a survey on this subject and compares Radial Basis Function (used by Zhang and his colleagues [18]), polynomial fitting (exploited by Athimethphat and Patanavijit [25]), natural neighbor interpolation and others. That study claims that the most efficient approach when the density of points has a high variation is the natural neighbor interpolation.

The first step of natural neighbor interpolation is to set a Voronoi diagram of the known points. A Voronoi diagram defines areas of the closest points. Figure 2.a shows one example with solid black lines as region limit between sites.

An unknown value is calculated by the weighted sum of neighboring sites. The weight is the area "borrowed" from each known point if the diagram had this point. Figure 2.a represents this by the blue region around the new point; the interpolation weights are represented by the green circles at each corresponding point.

The dual of a Voronoi Diagram is the Delaunay Triangulation [2] scheme, in which no point is inside the circumcircle of any triangle. The Voronoi diagram is obtained by connecting the circumcircle centers of the triangles. Delaunay triangulation can be used to compute the Natural Neighbor Interpolation; insertions and removals in it are easier to implement than using Voronoi Diagram directly [5].

The triangulation is built as points are inserted. If the points are inserted according to a z-order space filling curve (e.g. Figure 2.b), the triangulation creation has linear complexity in time [3]. If no specific order is used, the lower bound in time is log-linear [2]. Thus, the points are inserted in the z-order into the Delaunay triangulation.

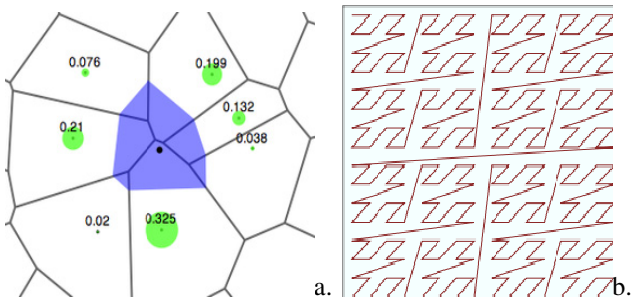


Fig. 2. Natural neighbor interpolation [22] (a); Z-order space filling curve [21] (b)

3.3 Shading Removal

Once the shading is obtained for the entire picture, it must be removed. Two assumptions can be made for document images [16]: the book surface has lambertian reflection (i.e. the specular index is too low) and it has non-uniform albedo distribution. Having these assumptions into account, the image of the paper background has a constant value of intensity (I_{us}), independently of the location of the viewer if illuminated by the same amount of light. As the image is the result of reflected light (I_o) at arbitrary levels, one can express the light variance by equation (2):

$$\frac{I_o(i,x,y)}{I_{us}(i,x,y)} = L(i,x,y);\ \ i \in \{r,g,b\} \qquad (2) \qquad\qquad I_{us}(i,x,y) = \frac{BG_{us}(i)}{BG_o(i,x,y)} I_o(i,x,y) \qquad (3)$$

Observing equation (2), $I_{us}(i,x,y)$ the values for the document background color should be constant for every RGB component, hence, $I_{us}(i,x,y) = I_{us}(i)$.

The value without shading is $I_{us}(i,x,y)$ for every pixel in the image. It is calculated using equation (3) [4], where $BG_o(i,x,y)$and $BG_{us}(i)$, denote $I_o(i,x,y)$ and $I_{us}(i,x,y)$ of shading pattern, respectively. I_{us}, I_o, BG_{us} and $BG_o$are in the interval [0,255]. Figure 7 shows the result in (b) of the proposed algorithm applied to (a).

4 Results

The algorithm performance assessment is done using synthetic images. The test images were generated by three illumination models combined with 11 images without shading. For each output image, one with the same structure but without distortion was created.

The images without shading were obtained from the LiveMemory dataset [8] and two issues of the Negócios PE Magazine [10][11], which are scanned images. The shaded images were generated using Processing 1.2.1 [23] library. All images will be made publically available for non-commercial use after this paper publication. Figure 3 illustrates how these images were produced.

The quality of the proposed method is assessed by calculating the "similarity" between references and processed images as illustrated in Figure 4. This computation alone does not provide enough information regarding the algorithm performance, as the processed image can be less clear than its input. If this happens, the image information was "lost". Thus, the degree of "similarity" between input and reference image is obtained, then the gain of information can be calculated as in Equation (4).

$$\text{Gain} = \frac{\text{Similarity between } \textbf{processed} \text{ and } \textbf{ref}.}{\text{Similarity between } \textbf{input} \text{ and } \textbf{ref}.} - 1 \quad (4) \qquad \text{SNR} = 20 \times \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \quad (5)$$

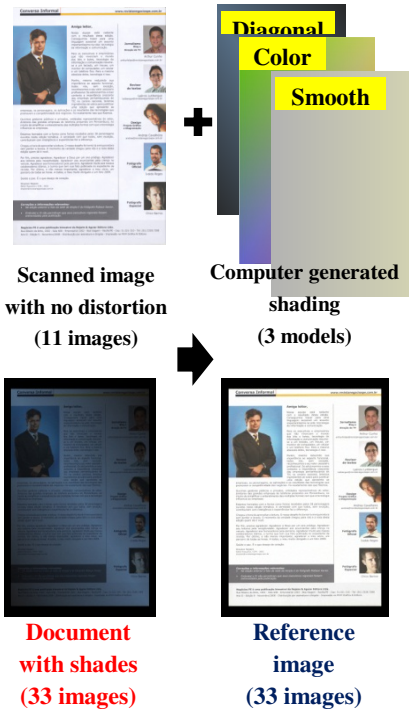


Fig. 3. Computer generated images

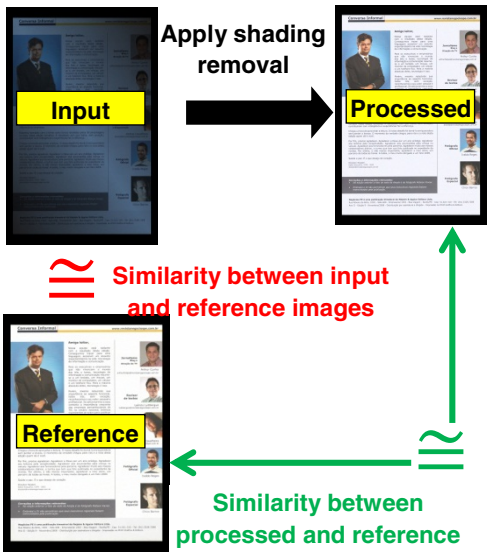


Fig. 4. Assessment of Image Quality

In this work two similarity measures were used: PSNR and SSIM [17]. The PSNR (Peak signal-to-noise ratio) is a classical similarity measure between two images which is calculated according to equation (5), where MAX_I are the maximum possible intensity and MSE the mean square error between the two images. In this context, the intensities are in the [0;1] scale, thus MAX_I is equal to 1.

The SSIM (Structural SIMilarity) [17] method was proposed by Wang and his colleagues. It is an objective measure that compares how two images are different from a human eye perceptual model. It is widely used for JPEG compression quality measurement. This measures ranges between 0 and 1, with 1 representing if two images are equivalent and 0 if they are totally different. This work used the Matlab™ code provided by the authors [17] with the default parameters.

As the input images are in color, the values of the PSNR and SSIM were computed for each RGB component and the grayscale version of the images. Table 1 shows the SSIM and PSNR values averages and standard deviation between the reference images and shading/processed images with the average gain. The results are grouped by the shading pattern described in Figure 3. One may see that the information gains are always positive, with the average of 68.8% in the diagonal scenario (worst case). Figure 6 shows the input images together with their shading estimation, corrected, reference images and corresponding SSIM and PSNR including their gains in parenthesis.

One important fact regarding proposed algorithm is that the illustrations do not change the quality of correction of textual areas as the interpolation always passes at the known points. This does not happen for the fitting methods, which are used by [18][16][19][25][26].

Figure 5 shows the graph that relates the processing time and the image size. One may see that the relation is almost linear for the images in test set. Further testing with more complex documents that may incorporate many different figures of different sizes is needed. Whenever there is a large variation in the density of the points in an image the performance of Delaunay triangulation may degrade [20].

Table 1. Reference SSIM related to input and processed

Model	SSIM to the reference image			PSNR to the reference image		
	Shading	Processed	Gain	Shading	Processed	Gain
Diag.	0.572 ± 0.048	0.952 ± 0.014	68.0% ± 13.5%	5.742 ± 1.504	32.4 ± 8.6	485.4% ± 270%
Color	0.814 ± 0.018	0.970 ± 0.011	19.5% ± 3.1%	9.320 ± 1.628	33.8 ± 11	268.7% ± 155%
Smooth	0.951 ± 0.006	0.981 ± 0.004	3.2% ± 0.6%	14.81 ± 1.653	38.5 ± 5.1	165.1% ± 47.6%

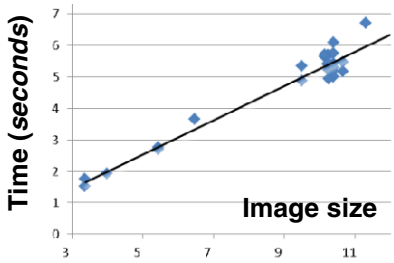


Fig. 5. Image size (in Mpixels) versus Processing Time

4.1 Processing of Real Images

Previous section presented the good performance measures towards computer-generated images dataset. This section analyses the behavior of the proposed algorithm applied to 14 Mpixels (with 3240x4320 pixels) real world images taken with Sony CyberShot W380.

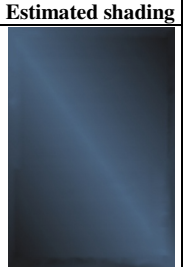
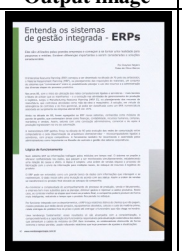
Input image	Estimated shading	Output image	Reference image	SSIM	PSNR
				R: 0.942 (+91.4%) G: 0.956 (+66.3%) B: 0.958 (+48.3%)	R: 37.1 (+800.2%) G: 39.0 (+711.0%) B: 39.4 (+632.5%)
				R: 0.967 (+16.8%) G: 0.974 (+17.3%) B: 0.951 (+33.1%)	R: 38.1 (+214.6%) G: 38.9 (+238.6%) B: 28.9 (+188.3%)

Fig. 6. Shading and output images of the proposed algorithm

Fig. 7. Processing of 3240x4320 pixels image: input (a); result (b); zoom into the left-upper part of processed image without shading (c)

Input image	Output image
	
	

Fig. 8. Processing of images of 3240x4320 pixels

Figure 7 shows one example of processing where the illumination became uniform in all areas of the image, even in region with a picture enclosed different shading as highlighted in Figure 7.c. Other examples of proposed algorithm processing are presented in Figure 8, with input images on the left and output on the right parts. The upper image has smooth illumination change and the proposed algorithm removed it turning the visualization more clear.

The lower image of Figure 8 has several pictures and two uniform regions, turning the illumination recovery harder. The shading of the output image was completely removed, confirmed the good performance of proposed algorithm.

5 Conclusions

This paper introduces a new algorithm that efficiently removes shading of documents acquired with portable digital cameras. The algorithm is robust enough to remove shading of documents with pictures, drawings and other illustrations. A widely used perceptual measure – the SSIM – shows that the proposed scheme yields high quality results. Besides that, the proposed algorithm performed well on real word 14 Mpixel images with 11s processing time on average.

References

[1] Amidror, I.: Scattered data interpolation methods for electronic imaging systems: A survey. J. Electron. Imaging 11(2), 157–176 (2002)
[2] de Berg, M., et al.: Computational Geometry: Algorithms and Applications. Springer (2008)
[3] Buchin, K., Mulzer, W.: Delaunay Triangulations in O(sort(n)) Time and More. In: 50th Annual IEEE Symposium on Foundations of Computer Science, pp. 139–148 (2009)
[4] Fan, J.: Robust Color Image Enhancement of Digitized Books. In: Proceedings of 10th International Conference on Document Analysis and Recognition, pp. 561–565 (2009)
[5] Flötotto, J.: 2D and Surface Function Interpolation. CGAL User and Reference Manual, CGAL Editorial Board, 3.8th edn. (2011)
[6] Gatos, B., Pratikakis, I., Perantonis, S.J.: An Adaptive Binarization Technique for Low Quality Historical Documents. In: Marinai, S., Dengel, A.R. (eds.) DAS 2004. LNCS, vol. 3163, pp. 102–113. Springer, Heidelberg (2004)
[7] Gatos, B., Pratikakis, I., Perantonis, S.J.: Adaptive degraded document image binarization. Pattern Recognition 39(3), 317–327 (2006)
[8] Lins, R.D., Torreão, G., Pereira e Silva, G.: Content Recognition and Indexing in the liveMemory Platform. In: Ogier, J.-M., Liu, W., Lladós, J. (eds.) GREC 2009. LNCS, vol. 6020, pp. 220–230. Springer, Heidelberg (2010)
[9] Niblack, W.: An Introduction to Digital Image Processing. Englewood Cliffs, New Jersey (1986)
[10] Najaim & Aguiar Ltd. Negócios PE. 19th edn. Najaim & Aguiar Ltd., Recife (2011)
[11] Najaim & Aguiar Ltd. Negócios PE, 18th edn. Najaim & Aguiar Ltd., Recife (2011)
[12] Oliveira, D.M., Lins, R.D.: A New Method for Shading Removal and Binarization of Documents Acquired with Portable Digital Cameras. In: Third International Workshop on Camera-Based Document Analysis and Recognition, Barcelona, Spain, pp. 61–65 (2009)

[13] Oliveira, D.M., Lins, R.D.: Generalizing Tableau to Any Color of Teaching Boards. In: 20th International Conference on Pattern Recognition, Istanbul, Turkey, pp. 2411–2414 (2010)
[14] Sauvola, J., Pietikainen, M.: Adaptive document image binarization. Pattern Recognition 33(2) (2000)
[15] Bukhari, S.S., Shafait, F., Breuel, T.M.: The IUPR Dataset of Camera-Captured Document Images. In: 4th Int. Workshop on Camera-Based Document Analysis and Recognition, Beijing, China (2011)
[16] Lu, S., Tan, C.L.: Thresholding of badly illuminated document images through photometric correction. In: Proc. 2007 ACM Symp. Document Eng., Manitoba, Canada, pp. 3–8 (2007)
[17] Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004), https://ece.uwaterloo.ca/~z70wang/research/ssim/
[18] Zhang, L., Yip, A.M., Tan, C.L.: Photometric and geometric restoration of document images using inpainting and shape-from-shading. In: 22nd Conference on Artificial Intelligence, Vancouver, Canada, pp. 1121–1126 (2007)
[19] Lee, J.-S., Chen, C.-H., Chang, C.-C.: A novel illumination-balance technique for improving the quality of degraded text-photo images. IEEE Trans. Cir. and Sys. for Video Technol. 19, 6 (2009)
[20] Har-Peled, S.: Data structures for geometric approximation. American Mathematical Society (2011)
[21] Eppstein, D.: Four levels of the Z curve, showing the square that is eventually filled by the curve, http://en.wikipedia.org/wiki/File:Four-level_Z.svg (last visited on March 20, 2013)
[22] File, M.: Natural-neighbors-coefficients-example.png, http://en.wikipedia.org/wiki/File:Natural-neighbors-coefficients-example.png (last visited on March 20, 2013)
[23] Processing.org, http://processing.org (last visited on March 20, 2013)
[24] Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Prentice Hall, New Jersey (2008)
[25] Athimethphat, M., Patanavijit, V.: A non-linear illuminations balancing for reconstructed degraded scanned text-photo image. In: ISCIT 2010, pp. 1158–1163 (2010), doi:10.1109/ISCIT.2010.5665163
[26] Meng, G., Xiang, S., Zheng, N., Pan, C.: Non-parametric Illumination Correction for Scanned Document Images via Convex Hulls. IEEE Trans. on Pattern Ana. and Machine Intelligence (99)