Correcting Specular Noise in Multiple Images of Photographed Documents

Ednardo Mariano
Rafael Dueire Lins
Gabriel de França Pereira e Silva
Universidade Federal de Pernambuco
Recife, Brazil
{rdl, gabriel.psilva}@ufpe.br

Jian Fan
Peter Majewicz
HP Labs
Palo Alto, USA
{jian.fan, peter.majewicz}@hp.com

Marcelo Thielo
HP Labs.,
Porto Alegre, Brazil
marcelo.resende.thielo@hp.com

Abstract **—** Portable digital cameras have become omnipresent. Their low-lost, simplicity to use, flexibility, and good quality images have widened their applicability far beyond their original purpose of taking personal photos. Every day people discover new uses for them from photographing teaching boards to documents. One of the difficulties of using cameras is the occurrence of specular noise whenever the photographed object is glossy. This paper presents an efficient algorithm for removing the specular noise of photographed documents by taking multiple images with different illumination sources.

Keywords- *specular noise, photographed documents, multiple images.*

I. Introduction

Portable digital cameras may be considered as a pervasive good. The quality and resolution of the cameras embedded in cell phones today are as good as the ones of dedicated devices from a not distant past. Such omnipresence has widened its applicability into unforeseen domains. One of them is using portable digital cameras for digitizing documents. People now use those devices as a fast way to acquire document images, avoiding photocopying, take photos of teaching boards and bill boards instead of taking notes. Such application gave rise to new research area [3], which is evolving fast in many different directions and claims for new algorithms, tools and processing environments that are able to provide users in general with simple ways of visualizing, printing, transcribing, compressing, storing and transmitting through networks such images. Reference [6] points out some particular problems that arise in this document digitization process using portable digital cameras: the first of all is background removal. Very often the document photograph goes beyond the document size and incorporates parts of the area that served as mechanical support for taking the photo of the document. The second problem is due to the skew often found in the image in relation to the photograph axes, as documents have no fixed mechanical support very often there is some degree of inclination in the document image. The third problem is non-frontal perspective, due to the same reasons that give rise to skew. A fourth problem is caused by the distortion of the lens of the camera. This means that the perspective distortion is not a straight line but a convex line, depending on the quality of the lens and the relative position of the camera and the document. The fifth difficulty in processing document images acquired with portable cameras is due to non-uniform illumination. A even more complex situation is faced when the paper of the document is glossy: often the photo has areas in which the in-built strobe flash or intense lighting from the environment "erases" parts of the document presenting what is called the "specular noise" [1]. Figure 1 presents an example of a photo with specular noise. The photo from two bound pages of a magazine printed on glossy paper exhibits two areas of specular noise. On the right hand page one may observe an "erased" area to the left of the leg of the chair. On the opposite page, there is another damaged area on the pillows (which erases part of the stripes). Red arrows point at the specular noises in Figure 1. The photo was obtained with the built in strobe flash on.

This paper proposes an algorithm to remove the specular noise by taking three pictures of documents or 3D-objects with a fixed camera and varying the position of the light sources. results obtained yield to conclude that the presented scheme is valid and provided much better quality results than any other previous report in the technical literature [7][10].

Figure 1. Photo of a two page magazine printed on glossy paper with specular noises.

1520-5363/11 $26.00 © 2011 IEEE
DOI 10.1109/ICDAR.2011.186

II. MOTIVATION

Besides the aforementioned omnipresence of digital cameras, there are situations in which their use is more adequate and easier than flatbed scanners. Since scanners have been integrated with printers in "all-in-one" devices (scanner, printer and copier) the price for an A4 scanner became marginal. But larger-size scanners have kept their high prices unchanged and are not easily available to buy. Scanning documents larger than A4 size, such as maps, is a difficult task. The same happens with bound books (either hard or soft) for which one either damages the volume or has to address the difficult problem of geometrical warping and even blur in some areas. In the case of oversized old books or documents, using standard A4 flatbed scanners is unviable.

A simple digitalization platform for such documents based on portable digital cameras is shown in Figure 2.

Figure 2 – Digitization platform with two light sources and portable digital camera (*UFPE-Planetarium*).

The platform presented in Figure 2 was built at UFPE (Brazil) and it has shown to be a suitable test bed for document and book digitization for a low-cost and flexible platform, capable of digitizing oversized documents such as maps.

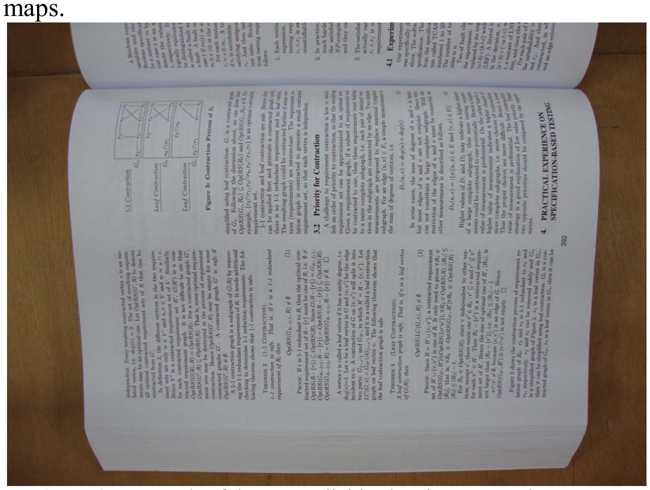

Figure 3 – Example of document digitized at the *UFPE-Planetarium* Both lamps on and strobe flash off.

The features of the *UFPE-Planetarium* are:

- Camera: Digital Sony Cyber-Shot W220, 12.1 Mpixels, in-built strobe flash.
- Camera height (lens) to support plan: 12.2 cm;
- Lamp: Compact fluorescent – 20 W;
- Right-hand side lamp height to support plan: 13.7 cm;
- Left-hand side lamp height to support plan: 13.5 cm;
- Distance between camera lens and right-lamp: 10.8 cm;
- Distance between camera lens and left-lamp: 11.2 cm.

A second platform was also built with the following features:

- Camera: HP 8 Mpixels with fixed resolution and zoom. Lens distortion is negligible;
- Camera height (lens) to support plan: 21 cm;
- Lamp: high-power phosphor-converted white LEDs;
- The center lamp is positioned as close to the lens as possible.
- Lamp offset left:7.5 cm
- Lamp offset right:7.5 cm

A total of 4 frames of the scene with different illumination are available. They are:

- ambient lighting only;
- left lamp on, right lamp off, center lamp off;
- left lamp off, right lamp on, center lamp off;
- left lamp off, right lamp off, center lamp on.

The correction of the geometrical warp caused by book binding was studied in a series of papers [4][9] with satisfactory results. The other problem faced is being addressed here: the removal of the specular noise.

III. METHODOLOGY

Removing the specular noise from images such as the one presented in Figure 1 is a difficult task. One possibility is to take several photos under different illumination set-ups to try to "re-assemble" a new image replacing the parts "erased" by the specular noise from one image with the "information" from another shot. This paper follows exactly such an alternative taking three photos of a given document under three different illumination patterns and generating a new image with "parts" of the three "tributaries". Figure 4 presents an example of three images of the same document. Although the solution proposed here seems to be a simple one, in reality to automatically extract the information from each image and yield a "natural looking" result is a complex task. After a large number of attempts that ranged from functor analysis of color variation to splitting the image into small blocks and using a neural classifier on each image to spot where the specular noise could be found in each of them [5], the scheme that is presented below was the one that provided the most consistent best results in terms of resulting image quality and also in processing time performance.

A. The algorithm.

The photos taken are true-color RGB 24-bits. Each document photo is split in their RGB components, which each of them is an 8-bit grayscale image (*Left-R*, *Left-G*, *left-B*; *Center-R*, *Center-G*, *Center-B*; *Right-R*, *Right-G*, *Right-B*).

Figure 4 – Document photographed with the second platform.
Top: Right lamp only. **Center:** Central lamp only. **Bottom:** Left lamp only.

The main idea of the algorithm proposed is to compare the intensity of the pixels of the image between components. The pixel with the lowest intensity of the same component of the three images is copied into the component of the final image. The image with the specular noise filtered out is formed by the RGB components, following the scheme depicted in Figure 5.

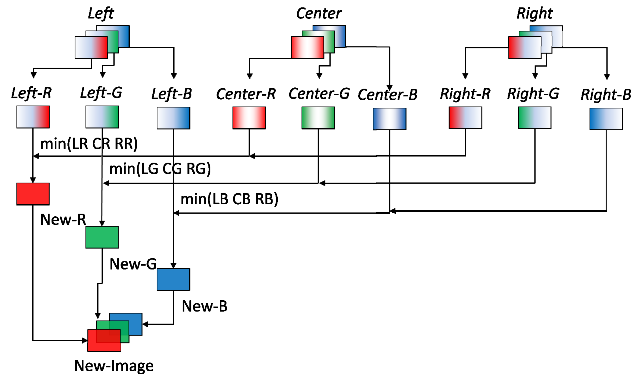

Figure 5 – Scheme adopted for removing the specular noise by analyzing three images simultaneously.

The images illuminated by the left and right hand side lamps (strobe flashes) tend to exhibit a circular halo at opposite sides to the lamp position causing a large variation to the pixel intensities and covering a large part of the image area. For better results, and such halo may have its contrast increased [1] by lowering the variation of the intensity of the pixels, through the equalization of the global histogram of each RGB component both of the left and right illuminated images.

The resulting image presents a significant reduction of the specular noise as may be observed in Figures 6 and 7.

Figure 6 – Image of Figure 4 with specular noise removed

Due to the choice of pixels with the lowest intensity in each component, the resulting image has its brightness reduced in comparison with the original ones. But all the information, colors and shapes are recovered, making possible the application of techniques of histogram correction to enhance the brightness and contrast of the resulting image [1].

Figure 7 – Image of Figure 1 with specular noise removed

In glossy paper, commonly used in magazines and color printed books the reflection of the light of the strobe flash may be so strong such as to cause saturation of the sensor of the camera, causing a "chroma noise" or "confetti" [1][2] at areas surrounding the specular noise. Figure 8 top zooms into part of Figure 1 and exhibits some of the chroma noise in the image. The algorithm proposed is not able to remove such noise as it may also be observed in the bottom part of the same figure.

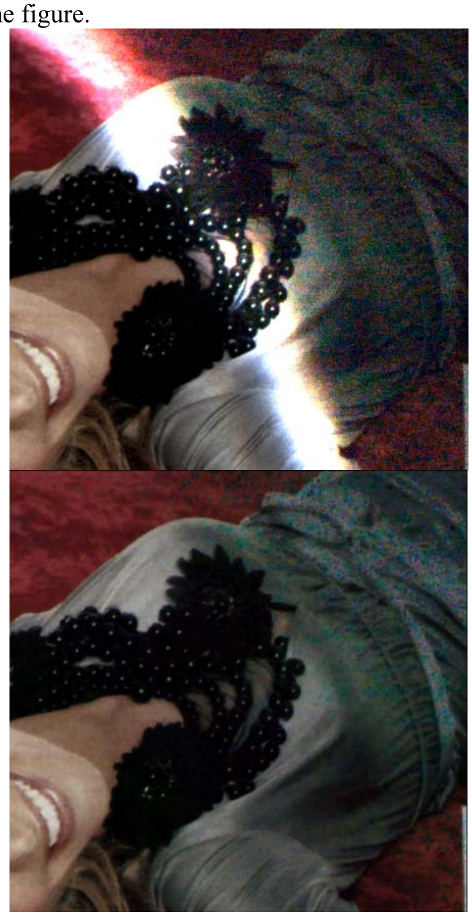

Figure 8 – Zoom into parts with chroma noises in Figures 1 and 7.

B. *Compensating Chroma Noises in Images*

The direct application of the algorithm presented in the last section does not filter out the confetti noise from the resulting image. A more careful study of the different images taken shows that such noise appears close to the specular halo and at the outskirts of images. Thus, it is possible to identify the regions where the chroma noise may appear and eliminate them before comparing the RGB components in the algorithm presented. This was done by creating three sub-images of the original image. The width of each image depends on the incidence angle of the lamp which determines the position of the halo and the chroma noise. For the angles of the current configuration of the second platform the images were split into 25%, 50%, and 25% of the original width. Regions will be labeled as "A", "B", and "C". Figure 9 illustrates such an image splitting in relation to the halo and specular noise. If one considers that each object is photographed three times under different illumination patterns and each has 3 RGB-components, then one has 27 sub-images in total.

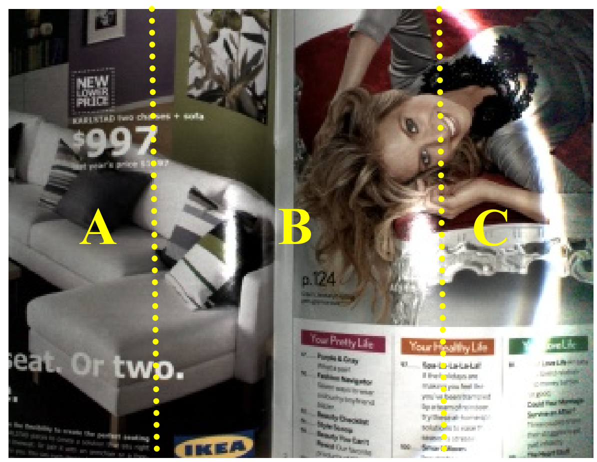

Figure 9 – Left illuminated image split into three regions

Instead of comparing the intensity of pixels directly of the three images following the scheme presented in Figure 5, one groups images in three sets:

- A-Left and A-Center;
- B-Left, B-Center, and B-Right;
- C-Right and C-Central;

Where "A-Left" means the region "A" of the Left-illuminated region, "A-Center" means the region "A" of the image with the centre lamp on and the other lamps off, etc. The sets above avoid that the chroma noise is taken into the final image. The images in each set have their intensity of the RGB-components compared in a similar fashion to the original algorithm, yielding one resulting image per region. The final image is obtained by merging the three images in their position. Figure 10 shows the resulting image of the images of Figure 4 with the specular and chroma noises reduced.

Figure 10 – Result of Figure 4 with specular and chroma noises removed

Figure 11 zooms into the image of Figure 1 with the specular and chroma noises reduced.

Figure 11–Zoom into part of Figure 1 with specular and chroma noises removed

One may still observe that the resulting image has lower brightness level than the original images.

The implementation of the algorithm was done using Matlab® of The MathWorks, running on Microsoft Windows Seven Professional®.

IV. CONCLUSIONS

Portable digital cameras are a low-cost option for digitizing large documents such as maps or even bound books. These are new applications that are far away from their original purpose that is to take souvenir photos of people and places. Whenever taking pictures of documents printed on glossy paper the specular noise may arise due to the short distance between the camera and the photographed object [7][10][8].

This paper presents an efficient method to remove the specular noise by taking multiple images under different illumination patterns. It was tested in images obtained in two different platforms and performed well in both of them. In the case of the second platform, besides the specular noise, the intensity of the strobe flash lamps also gave rise to chroma noises. Splitting the image in regions and analyzing them in a way not to take into account the noisy areas yielded good results.

The simple yet efficient method proposed here also provided suitable removal for the specular noise in grayscale images and also in images of 3-D objects.

REFERENCES

[1] R. Gonzalez, R. Woods. Digital Image Processing, 3° Edition, Prentice Hall, 2008, pp. 68-102,pp. 259-300.

[2] G. V. Landon, Y. Lin, and W. B. Seales. Towards Automatic Photometric Correction of Casually Illuminated Documents., IEEE Conference on Computer Vision and Pattern Recognition, 2007.

[3] J. Liang, D. Doermann and H. Li. Camera-Based Analysis of Text and Documents: A Survey. International Journal on Document Analysis and Recognition, 2005.

[4] J. Liang, D. DeMenthon, and D. Doermann. Geometric Rectification of Camera-Captured Document Images. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol 30 (4):591-605.

[5] R.D. Lins, G. F. P. e Silva, S. Banergee, S Kuchibhotla, and M. Thielo. Automatically Detecting and Classifying Noises in Document Images. ACM-SAC 2010. ACM Press, 2010. v.1. p.33 – 39.

[6] R. D. Lins, A. R. G e Silva, G. F. P. e Silva. Enhancing Document Images Acquired Using Portable Digital Cameras. ICIAR 2007. Springer Verlag, 2007. v.LNCS. p.1229 - 1241.

[7] C.-A. Saint-Pierre, J. Boisvert, G. Grimard, F. Cheriet, Detection and correction of specular reflections for automatic surgical tool segmentation in thoracoscopic images, Machine Vision and Applications. V22(1): 171-180. Springer-Verlag, 2007.

[8] D. M. Oliveira, R. D. Lins. Improving the Border Detection and Image Enhancement Algorithms in Tableau. ICIAR´2008. Springer Verlag, 2008. LNCS v.5112. p.1111 - 1121

[9] D. M. Oliveira, R. D. Lins, G. Torreão, J. Fan, M. Thielo. A New Method for Text-line Segmentation for Warped Documents ICIAR 2010. Springer Verlag, 2010. LNCS 6112. p.398 - 408.

[10] L. B. Wolff, L. B., "On the relative brightness of specular and diffuse reflection. In: Proceedings of CVPR" ,1994, pp. 369–376.