Detection of Color Fading in Printed Customer Content [1]

Zuguang Xiao[a], Shaoyuan Xu[a], Eric Maggard[b], Katie Morse[b], Mark Shaw[b], Jan Allebach[a]
[a]School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906, U.S.A.
[b]Hewlett-Packard Inc., Boise, ID 83714, U.S.A

Abstract

The electrophotographic (EP) process widely used in modern laser printer is susceptible to various print defects. Among them, fading is one of the most important ones since it is easily detected by the user on his or her prints; and it occurs frequently i.e. whenever one or more cartridges is low on toner. In contrast, the identification of other types of defects usually requires a specially designed page to be printed and examined by a professionally trained expert. Since fading usually happens when one or more of the printer cartridges – cyan, magenta, yellow or black – is depleted, it is of interest to know which cartridge is depleted given a faded print, so that the user can replace that cartridge. To achieve this goal, we will propose an algorithm that detects color fading from users' prints. The print is scanned; and the resulting digital image is compared to the digital master that was sent to the marking engine to generate the print. This work is a continuation of our previous work, which proposed a text-fading detection algorithm. Here, we extend our work to predict fading in non-text regions as well.

Introduction

Page quality (PQ) is most important in the printing industry – it directly plays a role in the user's satisfaction about their products. Page quality is degraded when defects appear on the page, which could be caused by the Electrophotographic (EP) process and associated print mechanism. The traditional diagnosis of a page uses a diagnostic page. To find what the problem is, users need to print the diagnostic page and compare it against a user manual [1] [2]. This approach could effectively diagnose a page, however it's too costly for customers. Instead, we want to have an algorithm that automatically picks up all the defects from a customer page directly. Among these many defects, color fading can dramatically degrade the PQ. Fading usually occurs in a large area of a page and is easily noticed. An example can be seen from Figure 1, where Figure 1(a) is the raster (digital master) page, and Figure 1(b) is the customer printed page, which has severe fading on the whole page caused by low toner in black cartridge. Figure 2 shows a similar example for low toner in the yellow cartridge. So, we have put the focus of our work on detecting color fading first. Previously, we proposed an algorithm to detect text fading [3] [4] [5]. In the case of a page that has no text, our previous algorithm will be unable to detect fading if any exists. Here, we extend our previous algorithm to be able to detect color fading on non-text regions from printed customer contents, as well.

Our approach to detect fading is to simply compares the user's printed page against the raster page, which is the digital original. However, before we compare these two images, the printed page needs to be scanned by a scanning device, and the raster page needs to be calibrated into the printer color space. We also need to make sure that the scanning device is well calibrated. After these two pages are in the same color space, the scanned page may be spatially misaligned with respect to the raster page. The same image registration algorithm used in our previous work is also used here, which extracts feature points and features, estimates the

homography by using MLESAC or RANSAC based on the feature points, and transforms the scanned page with the homography [6] [7] [8] [9] [10] [11].

In this paper, we propose an algorithm that solves this misalignment problem and then detects color fading from the scanned page. Our algorithm will calculate the super pixel – average of the pixels within a block – of both the raster page, and the scanned page, then clustering all these super pixels based on their colors in a perceptually uniform color space. Each color cluster on the scanned page will be compared with the corresponding one on the raster page by calculating the color difference. When fading occurs, we will see that for some color clusters, the color difference between the scanned and the raster (master) will significantly increase. By comparing how the color cluster changes its color on the scanned page against the color cluster on the master page, we can predict which cartridge is depleted. For example, when the skin tone looks very cool on the scanned page compared with the raster page, as shown in Figure 2, we may infer that yellow cartridge is running low on toner. However, if we analyze the black color cluster, in which pixels are mainly from the hair of the woman, we may not see too much difference between them. Our proposed algorithm will formulate these analyses to predict from each color cluster which cartridge has low toner and sum all the prediction results. By using majority rule, we can confidently report the low toner cartridge that needs to be replaced.

(a) Raster (digital master) page　　*(b) Printed page with fading*

Figure 1. An example showing the effect of black cartridge running low on toner.

(a) Raster (digital master) page　　*(b) Printed page with fading.*

Figure 2.. An example showing the effect of yellow cartridge running low on toner.

IS&T International Symposium on Electronic Imaging 2018
Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications

297-1

Preprocessing

To compare the scanned page with the raster page, we first apply the image registration algorithm described previously to the scanned page so that the content of the scanned page can be roughly aligned with the raster page.

(a) Before image registration   *(b) After image registration*

Figure 3. Raster image overlapped with scanned image.

Figure 3(a) shows the two pages overlapped with each other. The blurring on the page indicates the area of the scanned page that is misaligned with the raster page. After image registration, we can clearly see that the two pages are now well aligned. The image registration is a global process, after which the scanned page is registered with the raster page. However, this algorithm, or any other current existing image registration algorithm cannot achieve pixel to pixel accuracy in alignment. If we zoom in to the left bottom corner of Figure 3(b), we can still see some blurring of those text characters.

Since our interest is to detect fading in non-text region, we could further locally align each text character of the scanned page with the corresponding text character on the raster page, after which we remove those text characters from our region of interest. This algorithm uses Otsu's thresholding [12] to binarize images, connected component analysis to extract text characters from each binary image, and template matching [13] to align them. The algorithm is explained in detail in our previous work [3]. Figure 4 shows the scanned page in Figure 1(b) after removing the text characters, represented by the black bounding boxes. The remaining areas will be our region of interest. Note even though some pixels in the graphic regions are masked out by mistake, we should still be able to detect fading since fading is a large area defect.

Figure 4. Scanned image after removing text characters.

Object Map

An object map of a raster page tells the printer how this page should be rendered. An example is provided in Figure 5, which is the object map of the raster image in Figure 1(a). Three types of objects are commonly seen in a customer page: symbol, raster, and vector. Symbol regions mainly contain text characters and symbols, which are not in our interest for this work. Raster are those rough regions that contain many details. Raster regions are represented by the dark gray in this object map. And light gray regions are vector objects, which are the smooth areas. Print quality can be improved by using object-oriented halftoning which requires this object map [14].

Without going into too much of the details of the object map, the main take away from here is that since the raster regions and the vector regions may be rendered with different frequency of halftone screens, we will also do our analysis on these two kinds of regions separately in the future.
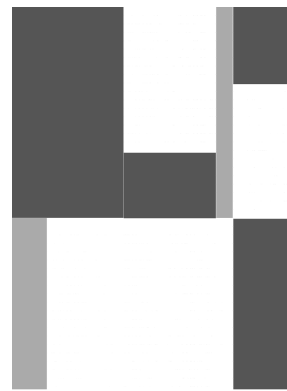

Figure 5. Object map of Figure 1(a). Dark gray is raster; light gray is vector.

Color Difference

To detect color fading on our region of interest, we cannot naively calculate the difference image. As mentioned earlier, the alignment is not perfect, and the misalignment is around 5 to 10 pixels. To compensate for this misalignment, we first convert each page from sRGB to CIE Lab color space, then average the pixels within each 100×100 size non-overlapping block, which we refer to as a *super pixel*. Next, we group all the super pixels in the raster page to different clusters by using the mean shift clustering technique [15]. The super pixels on the scanned page will be clustered the same way depending on their locations. In other words, a super pixel at a location on the scanned page will be classified to the same cluster as the super pixel at the same location on the raster page. At last, for each pair of color clusters from the raster page and the scanned page, we can calculate the color difference between them by using ΔE, which is the Euclidean distance between two points in CIE Lab color space. Note that this process is done separately for vector and raster regions.

There are two main benefits for adopting this approach. First, grouping all the super pixels to color clusters will reduce the impact of misalignment. The super pixel represents the average color of a block, which is less affected by the minor populations of the misaligned pixels. Furthermore, after we have clustered all the super pixels, we do not care where each color is located on the page anymore. We only care how this color changes when fading occurs. The second benefit is that the computation cost is tremendously

lowered. The page samples we run for our algorithm are A4 pages printed and scanned by a scanner at 600 ppi resolution; the resulting digital scanned pages have size around 6400×4928×3 pixels. This approach avoids the cost of comparing pixel by pixel between the master and the scanned page over the whole page. Instead, we just analyze some dominant color clusters, which have relatively large populations of super pixels. For those small population clusters, we can safely skip them, since they do not contribute too much to color fading, which is a large area defect.
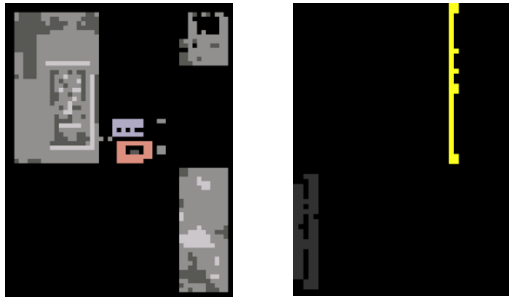
(a) Raster objects   (b) Vector objects

Figure 6. Color clusters and color map of them.

Figure 6 shows the color clusters generated by the mean shift algorithm for raster regions and vector regions for the image in Figure 4. To generate our test samples, we selected some customer raster pages and then alternatively printed these pages, which have about 16000 printed pages. The printed pages were automatically scanned. During this process, we had cartridges run out of toner at different page counts. We did not change the cartridge immediately when it ran out of toner, but rather waited until we had obtained a sufficient number of faded pages. Figure 1(a) is one of these customer pages; and we printed 189 of them. The dashed vertical lines in Figure 7 represent the ground truth that a page is faded, which is based on the examination by an expert. In our test pages, the cyan and magenta cartridge ran out of toner at about the same sampled page. This was followed by the yellow cartridge, and the black cartridge depleted last. In Figure 7, for each cluster, the plot shows the ΔE between the raster and the scanned images for the raster region and the vector region. On the right side of the plot, we have the RGB value, the percentage of super pixels in this cluster, and the CMYK value which comes from a printer LUT. Notice although there are many black color clusters in the raster region, they are actually composite black, which is a combination of C, M, Y instead of K and that is why all those color clusters have quite small ΔE when the black cartridge is depleted. Whereas in the vector regions, we have only two color clusters: black and yellow, and the ΔE significantly increases when K and Y cartridge is depleted. The ΔE drops immediately after we replace the cartridge. In Figure 8, we have also provided some scanned pages that have different types of fading caused by low toner in the selected cartridge.
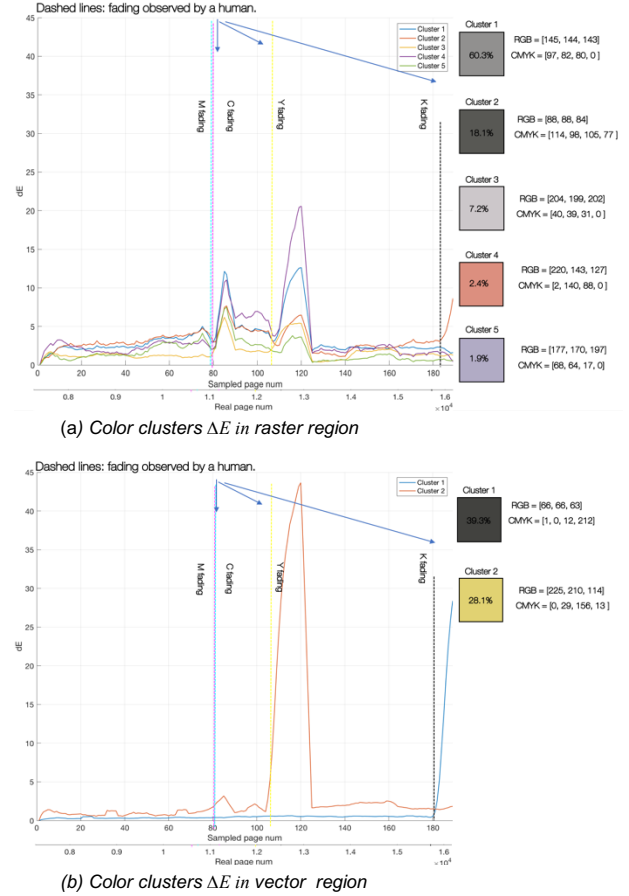

(a) Color clusters ΔE in raster region

(b) Color clusters ΔE in vector region

Figure 7. ΔE of different color clusters between raster page and the scanned page.

(a) Non-faded page   (b) C&M faded page

(c) Y faded page   (b) K faded page.

Figure 8. Example of faded pages.

IS&T International Symposium on Electronic Imaging 2018
Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications

297-3
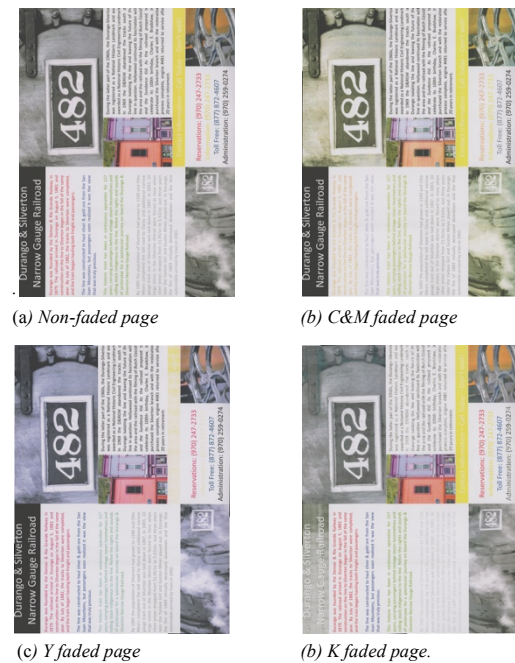
Local Analysis

The main drawback of our current algorithm is that it fails to pick up the color error when fading only occurs in some local regions. To be more specific, when a small population of pixels are faded and the remaining pixels in the same cluster are not faded, the average of all the pixels in this cluster will tend to be more like a non-faded color. Such an example is shown in Figure 9. In Figure 9, the faded pixels are magenta color, however, there are many more magenta pixels that are not faded. The average of these pixels is expected to have a very small color difference compared to the raster page. This can be hardly picked up by our algorithm.
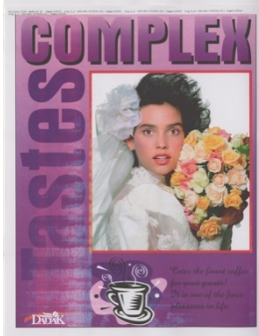

Figure 9. A locally faded page.

Figure 10. Partition of the raster image.

Fading can be better detected if the page is partitioned into many regions and we only look at one region of the page at a time so that those faded pixels will not be averaged out. Figure 10 shows how the raster image in Figure 1(a) is partitioned. The page is divided into a 12×8 array of blocks, and we apply the same algorithm for each block. In each block, we only look at the top three-color clusters ranked according to their populations. This allows us to detect local color fading from those blocks, if there are faded pixels in them.

Color Fading Prediction

Even if the ΔE of color clusters increases when the page is faded, it does not give us any direct information about which cartridge is depleted. Here, we propose a prediction algorithm based on the local color clusters, as shown in Figure 11. For each color cluster pair, we first calculate the color difference, and if the color difference is smaller than a threshold then we should not process it. If it is greater than the threshold, this implies that this color cluster in the scanned page is faded. To predict the depleted cartridge, we first calculate the four most faded colors – C, M, Y, K – based on the raster color cluster. The calculations will be explained shortly. After we have one estimated most faded color for each cartridge, we calculate which one of these four colors is closest to the color cluster of the scanned page, and we will increment the counter of the corresponding cartridge. These counters, which count the number of faded clusters for all cartridges, will represent our confidence scores. We repeat this process until all the color cluster pairs have been processed. In total, we will process $12 \times 8 \times 3$ color cluster pairs at most, because some blocks may have less than 3 color clusters.
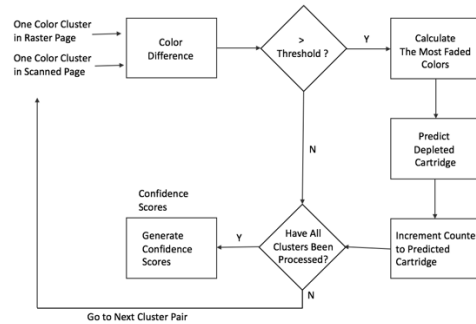

Figure 11. Color fading prediction algorithm flowchart.

The calculations of the most faded colors and prediction of the depleted cartridge are shown in Figure 12. Color clusters are initially in the CIE Lab color space. For each raster color cluster, we first convert it to sRGB color space, then CMY color space. The CMY is calculated by simply inverting the RGB values. The most faded color of each CMY can be treated as the color by setting each of the CMY channels to zero. For the black cartridge, we decrease all the CMY channels at the same time until one of the channels reaches zero. This is equivalent to subtracting the minimum of the three channels from all three channels. These calculated colors represent what the scanned color will look like when one cartridge is completely depleted. We then convert them back to CIE Lab color space and calculate the fading direction of each type of color fading. These directions, represented as normal vectors, are centered at the raster cluster color in CIE Lab color space. When fading occurs, we expect the scanned color cluster move away from its raster color, and towards one of these directions, depending on which cartridge is depleted.

Convert master cluster from RGB to CMY
$$CMY_{raster} = [255,255,255] - RGB_{raster}$$

Calculate the most faded color for each cartridge
$$CMY_{Cfade} = [0, M_{raster}, Y_{raster}]$$
$$CMY_{Mfade} = [C_{raster}, 0, Y_{raster}]$$
$$CMY_{Yfade} = [C_{raster}, M_{raster}, 0]$$
$$CMY_{Kfade} = CMY_{raster} - \min(CMY_{raster})$$

Convert all the most faded colors back to CIE Lab color space.
$$Lab_{CMYKfade} = RGB2Lab(255 - CMY_{CMYKfade})$$

Calculate the vectors of those most faded colors, and the vector of the scanned color.
$$Dir_{CMYK} = norm(Lab_{CMYKfade} - Lab_{raster})$$
$$Dir_{scanned} = norm(Lab_{scanned} - Lab_{raster})$$

Predict the depleted cartridge by finding the smallest angle between the vector of the faded color and the vector of the scanned color.
$$CartPredict = \min_{cart \in \{CMYK\}} \{Dir_{scanned} \cdot Dir_{cart}\}$$

Figure 12. Calculations of the most faded colors, and prediction of the depleted cartridge.

297-4

IS&T International Symposium on Electronic Imaging 2018
Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications

An example is shown in Figure 13. The top color box represents the color of one color cluster in a raster page. And the bottom four color boxes are the most faded colors for each channel. All of the fading directions are plotted in a 3D CIE Lab color space. As we can see from the plot, when C or M or Y is faded, the color will mainly change its hue; and when K is faded, the color remains relatively constant at its hue but increases its lightness. All these vectors are quite separated from each other, which is very important for our prediction problem, which is also a classification problem. After we have obtained the vector of the scanned cluster color by following the same calculations, we calculate its angles with respect to each of these four vectors and find the one with smallest angle. We then increment the counter of that cartridge. Eventually, we will have four counters; and we select the one with the maximum number of faded color clusters.
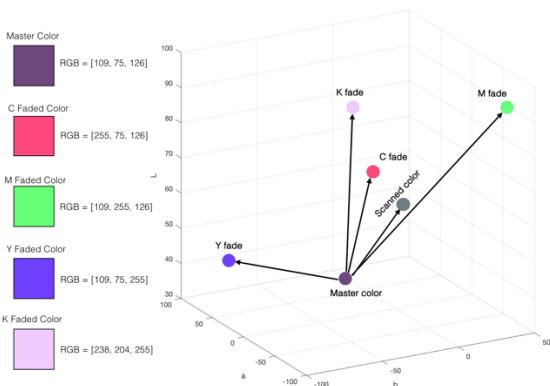

Figure 13. An example of calculations of color fade directions for a raster cluster color.

Experiment Results

Previously, we have shown the color difference results in Figure 7. Now, we apply our local analysis and color prediction algorithms to the same dataset, and the results are shown Figure 14. The x-axis of the plot in Figure 14 is the sampled page number, which has been explained for the previous results. The real printed page number is provided below as well. The y-axis represents the count of predictions for each cartridge. We use the same color for the plot to represent each cartridge. In the plot, at around 80th page, both the magenta and cyan cartridges are depleted. However, magenta fade is more dominant. That's why all the faded clusters are predicted as magenta fade. From the 107th page to the 123rd page, we have yellow color fade; and after 180th page, we have black fade. As one can see, the results coincide with the ground truth perfectly, which are represented by the dotted lines. And the results are also very consistent with the color difference results in Figure 7, where those peaks occur around the same pages. Figure 15 shows the color prediction result of another dataset. The fading events occur at almost at the same pages as with the previous dataset. We can see some mispredicted clusters under the magenta peak and the yellow peak. We can just use majority rule to predict one cartridge, since the population of those mispredicted clusters is very small.
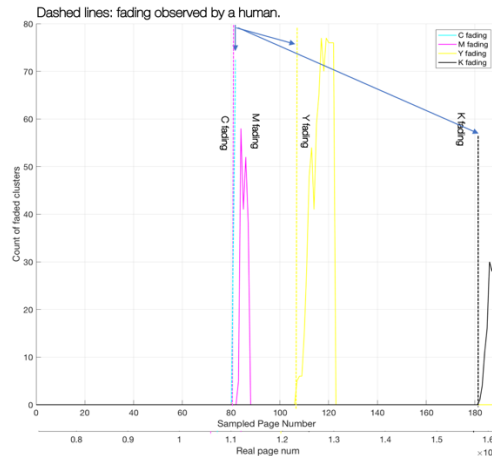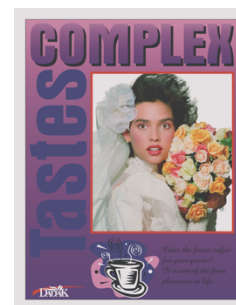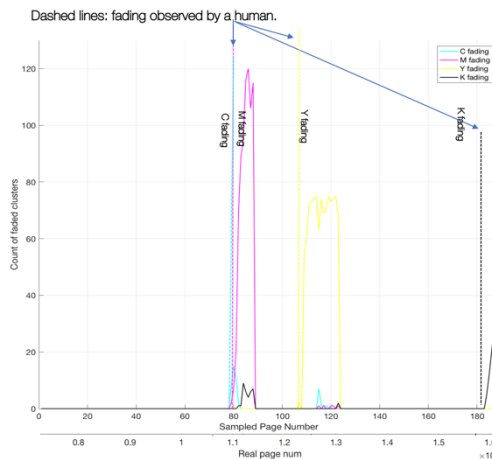

Figure 14. Color prediction result of the dataset whose raster page is in Figure 1.

(a) The raster page of a different dataset

(b) Color prediction result.

Figure 15. Results for a new dataset.

Conclusion

In this work, we proposed an algorithm to detect color fading from a scanned customer page by comparing it to its raster page. Our algorithm locally analyzes each color cluster to predict the depleted cartridge based on its raster cluster color. The final result consists of four confidence scores are the sum all the prediction results. We then use majority rule to select the most likely depleted cartridge. This algorithm along with some printer signal controls can be potentially implemented on an ASIC inside a printer that automatically pulls

IS&T International Symposium on Electronic Imaging 2018
Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications

297-5

out the raster page from the printer, and the scanned page from a scanning device to do the prediction analysis. This method has a clear advantage compared to the methods using a diagnostic page, which is too costly for customers.

Bibliography

[1] W. Jang, M.-C. Chen, J. Allebach and G. Chiu, "Print Quality Test Page," *Journal of Imaging Science and Technology,* vol. 48, no. 15, pp. 432-446, 1 Sept 2009.

[2] H. Santos, H. J. Park, R. Kumontoy, K. Low, M. Ortiz, C. W. Kim, P. Choe, S. Leman, K. Oldenburger, M. Lehto, X. Lehto and J. P. Allebach, "A web-based self-diagnosis tool to solve print quality issues," *Journal of Imaging Science and Technology,* vol. 54, no. 4, pp. 040503-1 - 040503-13, 2010.

[3] Z. Xiao, M. Nguyen, E. Maggard, M. Shaw, J. Allebach and A. Reibman, "Real-time print quality diagnostics," in *Image Quality and System Performance XIV*, Burlingame, CA, January 2017.

[4] N. Yan, E. Maggard, R. Fothergill, R. J. Jessome and J. P. Allebach, "Autonomous detection of iso fade point with color laser printers," in *Image Quality and System Performance XII*, San Fransisco, CA, February 2015.

[5] Y. Ju, E. Maggard, R. J. Jessome and J. P. Allebach, "Autonomous detection of text fade point with color laser printers," in *Image Quality and System Performance XII*, San Francisco, CA, February 2015.

[6] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," in *Fourth Alvey Vision Conference*, 1988.

[7] R. Hartley and A. Zisserman, "Chapter 4. Estimation – 2D Projective Transformations," in *Multiple View Geometry in Computer Vision*, Second ed., Cambridge University Press, 2004, pp. 87-131.

[8] J. Li and N. M. Allinson, "A Comprehensive Review of Current Local Features for Computer Vision," *Journal of Neurocomputing,* vol. 71, no. 10-12, pp. 1771-1787, Jun 2008.

[9] Y. Lu, Q. Feng, L. Qi and Q. Chen, "Sub-pixel Surface Fitting Algorithm in Digital Speckle Correlation Method," in *9th International Conference on Electronic Measurement & Instruments*, Beijing, 2009.

[10] P. Torr and A. Zisserman, "MLESAC: A New Robust Estimator with Application to Estimating Image Geometry," *Computer Vision and Image Understanding,* vol. 78, pp. 138-156, 2000.

[11] M. Fischler and R. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications To Image Analysis and Automated Cartography," *Communications of the ACM,* vol. 24, no. 6, pp. 381-395, Jun 1981.

[12] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 9, no. 1, pp. 62-66, Jan 1979.

[13] J. P. Lewis, "Fast Template Matching," in *Vision Interface*, Quebec City, 1995.

[14] S. J. Park, M. Q. Shaw, G. Kerby, T. Nelson, D.-Y. Tzeng, K. R. Bengtson and J. P. Allebach, "Halftone Blending Between Smooth and Detail Screens to Improve Print Quality With Electrophotographic Printers," *IEEE Transactions on Image Processing,* vol. 25, no. 2, pp. 601-614, 11 Nov 2015.

[15] Y. Cheng, "Mean Shift, Mode Seeking, and Clustering," *IEEE Transactions on Pattern Analysis and Macihne Intelligence,* vol. 17, no. 8, pp. 790-799, Aug 1995.

297-6

IS&T International Symposium on Electronic Imaging 2018
Color Imaging XXIII: Displaying, Processing, Hardcopy, and Applications