Modeling of Photoconductor Print Artifacts Using a Mixture of Gaussians

Ahmed H. Eid and Brian E. Cooper; Lexmark International, Inc.; Lexington, KY, USA

Abstract

Manufacturing imperfections of photoconductor (PC) drums in electrophotographic (EP) printers cause low-frequency artifacts that could produce objectionable non-uniformities in the final printouts. In this paper, we propose a technique to detect and quantify PC artifacts. A scanner-based system is utilized to scan printed pages of flat-field areas. A wavelet-based filtering technique is applied to the scanned images to isolate the PC-related artifacts from other printing artifacts, based on the frequency range and the direction of the PC defects. The prior knowledge of the PC circumference is utilized to determine the printed area at each revolution of the drum to be analyzed separately. The expectation maximization (EM) algorithm for probability density estimation is applied to the filtered images to model the PC defects as a mixture of three Gaussians. We use the estimated parameters of the Gaussians to measure the severity of the defect. The consistency of the PC artifacts, at subsequent revolutions of the drum, is studied by comparing the models from different revolutions. Results from experiments on different drums and print samples of different tone levels show a high correlation score between the proposed measure and the subjective evaluation of print quality experts.

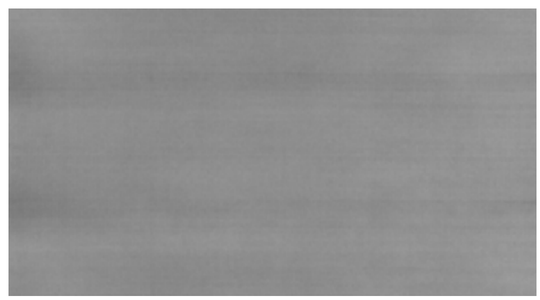
Introduction

Low frequency artifacts on the the photoconductor (PC) drum produce objectionable non-uniformities in the final printouts. This non-uniformity is most visible in low-density, large (constant) area, high-frequency halftoned regions. These artifacts can be caused by variation in the thickness of the charge layer of the photoconductor. This variation may occur during fabrication as well as during use (i.e., from wear). Environmental conditions can also influence the density variation.

During printing, successive revolutions of the cylindrical PC drum will form a repeating pattern of variation down the page (i.e., in the process direction). Vibration or inadequate photoconductor (PC) drum speed regulation can be the source of well-known PC banding artifacts, as shown in Figure 1 (top). Vibration and speed regulation issues can arise from gear manufacturing defects. Over-development (darkness) can occur if the PC drum slows down during rotation. Conversely, if the PC drum speeds up, areas of too little development (lightness) can be seen.

Due to the periodic nature of the PC artifacts, most of the methods developed to detect and quantify these artifacts have relied on generating the 1-D profile of the defect [1]- [5]. That assumption is valid for analyzing artifacts such as the one shown in Figure 1 (top). However, defects of a second type shown in Figure 1 (bottom) do not extend uniformly across the page. In such cases, the 1-D profile may not provide an accurate representation of the defect.

In this paper, we propose a technique to analyze the PC defects using the 2-D details of the input images assuming a prior knowledge of PC's circumference. A number of complete cycles of the PC drum from the input image is filtered using wavelet filtering. This preprocessing step is needed to remove the perpendicular defects from the image and to filter the high frequency defect parallel to the PC defects. The expectation maximization algorithm for density estimation is used to model the non-uniformities caused by the PC drum as a mixture of Gaussians. The model parameters are then used to provide a measure to quantify the PC defects.



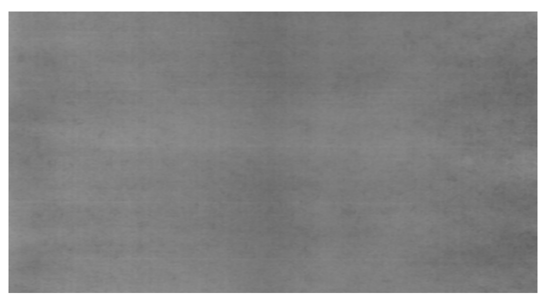


Figure 1. *Photoconductor (PC) artifacts: banding as a low frequency horizontal artifact that extends continuously across the page (top), a PC artifact that does not show continuity across the page but still shows the periodicity of the PC artifacts (bottom).*

System Setup

The system setup utilizes a scanner system with an automatic belt-fed document feeder. The feed mechanism of this scanner places the sample on the flatbed before scanning. As compared to more conventional paper-fed document feeders, this method introduces less distortion to the scanned image [5]. Test samples consist of ten flat field (constant coverage) pages, ranging from 10% to 100% coverage in steps of 10%. Each sample was halftoned and printed at 600 dpi on the ten printers included in the test. The print samples are then scanned at 600 ppi and calibrated to provide lightness (L^*) data. The scanned images are then sent to a workstation for imaging analysis. The proposed algorithms are implemented in MATLAB, and data are processed with a 3.4 GHz Pentium D dual-core HT processor.

Wavelet Analysis

Wavelet filtering is used to extract the low frequencies belonging to the PC drum in one direction while filtering the other defects on the perpendicular direction. The scaling property of wavelets, which is related to the frequency, helps to extract defects of different frequency bands even if they have the same extent and direction. In addition, the scale analysis of wavelet simulates the way the human visual system responds to image details [4].

Applying wavelet filtering to the scanned images needs three processing steps: decomposition, editing, and reconstruction. The 2D discrete wavelet representation is computed by applying a separable filter bank to the image $I(x,y)$ of size $M \times N$ as:

$$W_{\varphi}(j_0,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y)\varphi_{j_0,m,n}(x,y) \tag{1}$$

$$W_{\psi}^{i}(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y)\psi_{j,m,n}^{i}(x,y) \tag{2}$$

where j_0 is an arbitrary starting scale. $\varphi(x,y)$ and $\psi^i(x,y)$, $i \in \{H,V,D\}$ are 2D scaling and wavelet functions, respectively [6]:

$$\varphi(x,y) = \varphi(x)\varphi(y)$$

$$\psi^H(x,y) = \psi(x)\varphi(y)$$

$$\psi^V(x,y) = \varphi(x)\psi(y)$$

$$\psi^D(x,y) = \psi(x)\psi(y),$$

and the $W_{\varphi}(j_0,m,n)$ coefficients define an approximation of $I(x,y)$ at scale j_0. $W_{\psi}^{H}(j,m,n)$, $W_{\psi}^{V}(j,m,n)$, and $W_{\psi}^{D}(j,m,n)$ are the details of $I(x,y)$ at scales $j \geq j_0$.

The reconstructed image $I^R(x,y)$, up to L wavelet levels, that contains the details of PC artifacts isolated from other printing artifacts can be expressed as:

$$I^R(x,y) = \frac{1}{\sqrt{MN}} \sum_{j=j_0}^{L} \sum_{m,n} W_{\psi}^{H}(i,m,n)\psi_{j,m,n}^{H}, \tag{3}$$

which can be modeled by the expectation maximization (EM) algorithm.

Density Estimation using EM Algorithm

The expectation maximization algorithm is used to estimate the parameters of a mixture of Gaussians model. In this work, we assume the model is composed of three Gaussians representing dark, average, and light shades of the flat field printed samples showing the PC artifacts.

Assuming, in general, that we use a mixture of K Gaussians, then the Gaussian mixture distribution can be written as:

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k,\Sigma_k), \tag{4}$$

where x is the feature vector, π_k is the weight of the k^{th} mixture, and μ_k and Σ_k are the corresponding mean and covariance of the k^{th} Guassian (normal) distribution, respectively. To estimate the model parameters of a sample x of size N, the EM algorithm can be summarized as [7]:

1. Initialize the means μ_k, covariances Σ_k, and the weights π_k.
2. **E step**. Evaluate the responsibilities, γ_{kn}, where $1 \leq n \leq N$ using the current parameter values,

$$\gamma_{kn} = \frac{\pi_k \mathcal{N}(x_n|\mu_k,\Sigma_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(x_n|\mu_j,\Sigma_j)}. \tag{5}$$

3. **M step**. Re-estimate the parameters using the current responsibilities

$$\mu_k^{new} = \frac{\sum_{n=1}^{N} \gamma_{kn} x_n}{\sum_{n=1}^{N} \gamma_{kn}}, \tag{6}$$

$$\Sigma_k^{new} = \frac{\sum_{n=1}^{N} \gamma_{kn}(x_n - \mu_k^{new})(x_n - \mu_k^{new})^T}{\sum_{n=1}^{N} \gamma_{kn}}, \tag{7}$$

$$\pi_k^{new} = \frac{\sum_{n=1}^{N} \gamma_{kn}}{N}. \tag{8}$$

4. Repeat steps 2 and 3 until the relative difference of the subsequent values of Eqs. 6-8 are sufficiently small.

PC Artifacts Quantification

Periodic artifacts can be quantified by the contrast or a peak-to-peak measure of the signal representing the defect. For multimodal distribution like the mixture of Gaussians, the difference between the means of Gaussians can provide information about the defect severity. When the mixture of Gaussians is reduced to a single Gaussian, the variance, or the standard deviation, can be used to describe the variation caused by non-uniformities. In this paper, we propose a measure that uses the mixture of Gaussian parameters to quantify the PC defects as:

$$PC\ Score = \max\{\mu_{max} - \mu_{min},\ 2\sigma_{max}\}, \tag{9}$$

where μ_{max} and μ_{min} are the maximum and minimum means of the mixture, while σ_{max} is the maximum standard deviation of the mixture derived from the mixture covariances Σ_k. These model parameters used in the proposed measure should be selected from Gaussians that have weights of at least 0.1 to ensure a valid representation of the individual Gaussians in the mixture.

Experimental Results

In the first experiment, we show the consistency of PC models at different revolutions of the drum. Three consecutive samples of flat field areas of 50% coverage were printed using the

Table 1. PC artifact scores for 10 different printers, with their corresponding subjective and objective ranks.

Printer	Pr_1	Pr_2	Pr_3	Pr_4	Pr_5	Pr_6	Pr_7	Pr_8	Pr_9	Pr_{10}
PC Score	0.8892	0.9734	1.0594	1.1216	0.8181	1.2351	1.5674	0.7458	0.7879	1.0409
Objective Rank	4	5	7	8	3	9	10	1	2	6
Subjective Rank	3	7	8	6	5	9	10	2	1	4

same PC drum. Knowing the circumference of the drum lets us select an area of two cycles of the drum from a scanned image.

We applied the proposed technique to the two-cycle image. First, wavelet filtering is performed to extract the horizontal details with low frequencies and remove all vertical artifacts. Figure 2 shows the original image to the left and the corresponding wavelet filtering in the middle. Then, the EM algorithm is applied to the filtered image. The algorithm's estimated parameters and the mixture of Gaussians are shown in Figure 2 (right), along with the histogram of the image. The same procedure is repeated for the second and the third pages. Comparing model parameters of the three pages in Figure 2, we see the repeatability of the model of the PC artifact at different revolutions of the drum on consecutive printed pages.

In the second experiment we used ten printers to determine their quality based on the proposed approach. For each printer, we used ten printed samples of coverage ranging from 10% to 100% in steps of 10% for each printer. The PC artifact score of each print sample was computed. The results from each printer are compared at specific L^* values in Figure 3 (left). The PC scores are interpolated to provide scores at consistent L^* values. To compare the overall quality of each printer, we computed the mean of each printer's individual scores. The printers are ranked based on the overall PC scores in Figure 3 (right). Print quality experts ranked the printers from No. 1 (best quality) to No. 10 (worst quality), based on the objectionability of PC-related artifacts.

The results of the experiment are summarized in Table 1, comparing both the objective and the subjective rankings. The ordinal correlation between the objective rank derived from the PC score and the subjective rank provided by print quality experts is 0.8788. This shows a high correlation between the proposed technique and the subjective ranking.

Conclusions

In this paper, we presented a technique to model the photoconductor (PC) artifacts. The wavelet filtering is used to extract the horizontal low frequency defect representing the PC signature while filtering other printing artifacts in the vertical direction. The expectation maximization (EM) algorithm was used to model the PC artifacts using a mixture of three Gaussians. We introduced a measure of the PC quality based on the parameters estimated by the EM algorithm. Our experimental results show that the Gaussian mixture estimated from different prints of the same PC drum were consistent. Another experiment on 10 different printers show a high correlation between the objective ranking provided by the proposed technique and the subjective ranking provided by print quality experts.

References

[1] Bimal Mishra and D. Rene Rasmussen, "MicroUniformity: an image quality metric for measuring noise," *Proc. of IS&T's 2000 PICS Conference,* pp. 75-78, Portland, March 2000.

[2] D. Rene Rasmussen, Edul N. Dalal and Kristen M. Hoffman, "Measurement of macro-uniformity: streaks, bands, mottle and chromatic variation," *Proc. of IS&T's PICS*, pp. 90-95, Montréal, April 2001.

[3] Nathir A. Rawashdeh, Il-Won Shin, Kevin Donohue, and Shaun Love, "Printer banding estimation using the generalized spectrum," *Proc. of SPIE-IS&T*, Vol. 6059-15, San Jose, January 2006.

[4] Kevin Donohue, Chengwu Cui, and M. Vijay Venkatesh, "Wavelet analysis of print defects," *Proc. of IS&T's 2002 PICS Conference,* pp. 42-47, Portland, April 2002.

[5] A. Eid, B. Cooper, and E. Rippetoe, "A unified framework for physical print quality," *Proc. SPIE,* Vol. 6494, Jan. 2007.

[6] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, 2nd edition, Prentice Hall Inc., 2002.

[7] Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer Science+Busness Media, LCC, 2006.

Author Biography

Ahmed Eid received both his BS and MS in Electronics and Communications Engineering from Mansoura University (1994 and 1999 respectively). He received his PhD in Electrical Engineering from the University of Louisville (2004). He joined Lexmark International Inc. in 2005. His work focuses on image and print quality. His other research interests include 3-D data modeling, image registration, segmentation, and fusion. He is a member of IEEE and its computer and signal processing societies, IS&T, SPIE, and ACM.

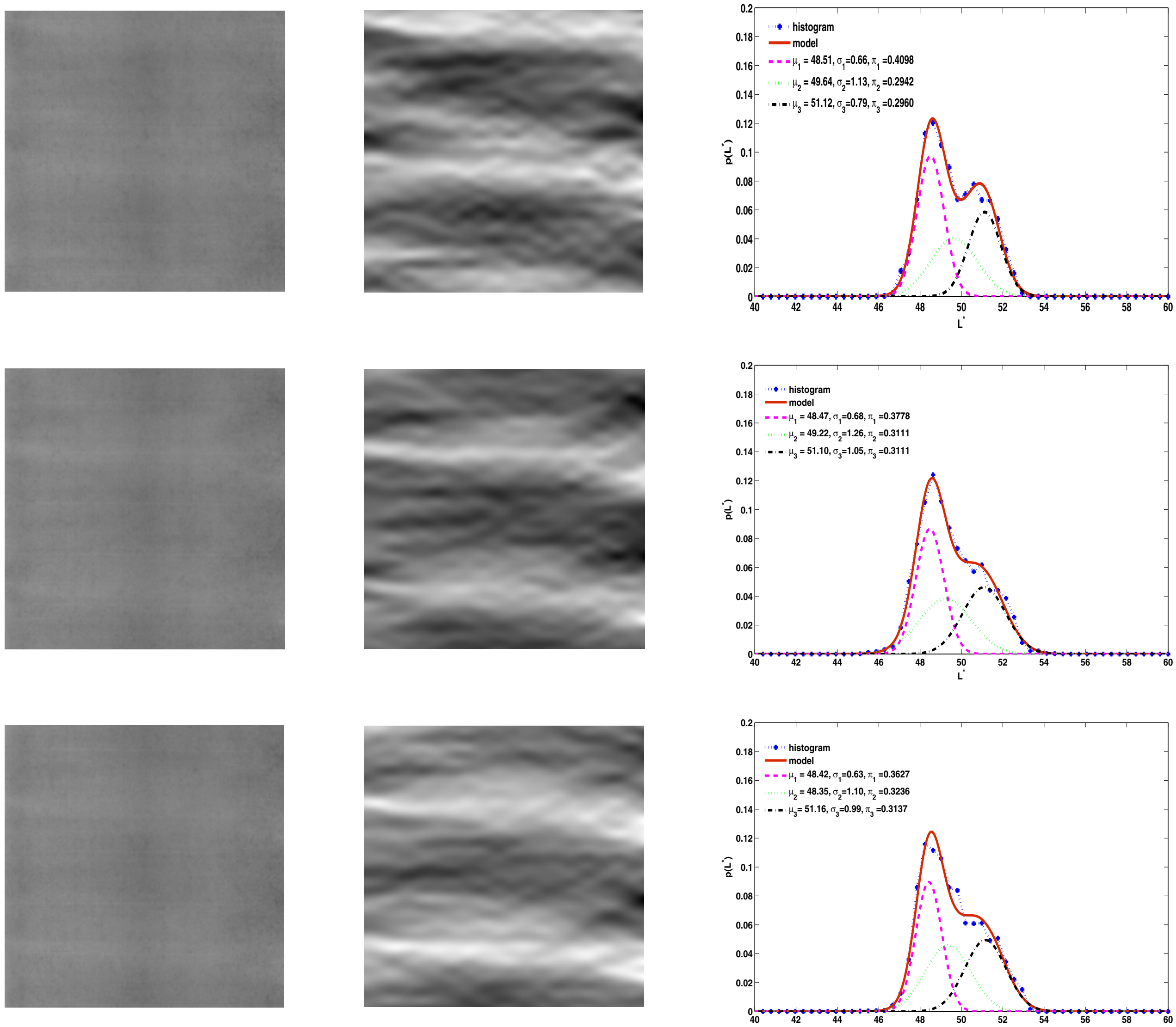


Figure 2. *PC artifact modeling: original images (left), wavelet filtered images (middle), and the corresponding mixture of Gaussians with the estimated model parameters (right) are shown. The three rows show scans of three consecutive printouts printed that used the same PC drum. The results show the consistency of model parameters from different cycles of the PC drum.*

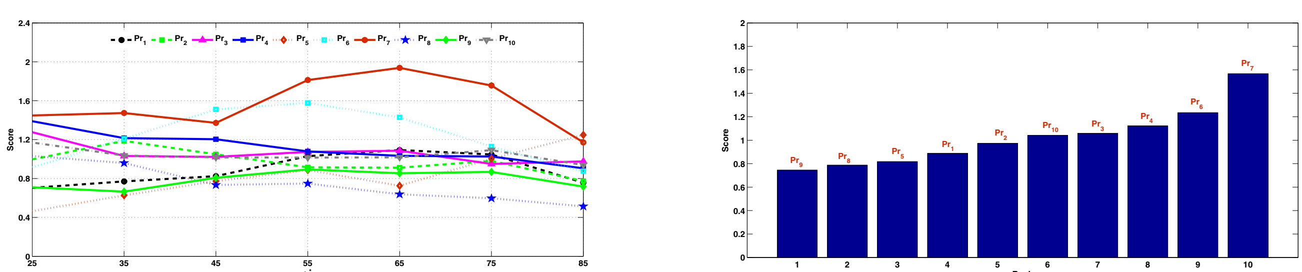


Figure 3. *Quantification of the PC artifacts: the PC artifact scores for ten printers using different print samples of different tones (left). The results are interpolated at specific L^* values to compare scores from each printer at consistent L^* values. The ten printers are ranked (right) by averaging together each printer's scores across the L^* range. A lower rank number means better quality.*