

IMPROVED CROSS-TALK CANCELLATION IN SCANNED IMAGES BY ADAPTIVE DECORRELATION

Boaz Ophir and David Malah

Technion IIT, Haifa 32000, Israel
 Department of Electrical Engineering

boazo@technix.technion.ac.il, malah@ee.technion.ac.il

ABSTRACT

Cross talk (or show-through) interference is a common occurrence when scanning duplex printed documents. The back-side printing shows through the paper thus contaminating the front side image. Previous work focused on modelling the non-linear process and offered a simple adaptive decorrelation process. We propose an improved cleaning process. We take into account local brightness variations, estimating local background brightness through a mean-shift process and adapting to it. The effects of cross interference are minimized by a post processing adaptive filtering stage. Further improvement is achieved by a cascaded multi-stage filtering scheme.

1. INTRODUCTION

Cross talk interference is a common occurrence when scanning duplex printed documents. The back-side printing shows through the paper thus contaminating the front side image. The same occurs when scanning the reverse side of the page. This is not a problem in low quality scans (as done in home/office scanners) where, up to a degree, image quality is not an issue. The matter becomes crucial when image quality is essential. Such a case is when creating a master copy, in the digital printing industry.

A previous work [1] focused on analyzing the process that causes the phenomena, tracking the passage of light in the scanner mechanism as it passes through the document, creating a physical model for it and trying to linearize it. The actual cleaning process, based on the developed model, is fairly simple, using basic signal processing and adaptive filtering techniques [2] to estimate the point spread function and clean the front-side image, using the back-side image as a reference noise signal.

Related work, such as [3] [4], dealing with the separation of transparent layers, assume linear image mixture models, which are not directly applicable in our scenario.

2. THE PHYSICAL MODEL

The model developed in [1] tracks the light passing through the printed page and the scanner mechanism. In a typical scanner, the light, originating from the scanner lamp, is reflected from the document back to a sensor, thus creating a reflectance profile of the document.

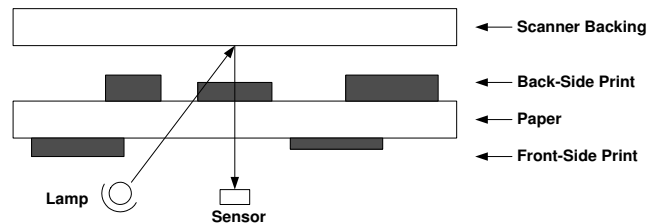


Fig. 1. Passage of light through duplex printed document.

The print and paper are set in a layered structure, each layer displaying reflectance, transmittance and absorbance properties. The print layers (front and back) are assumed to be non-reflecting surfaces, with the print itself representing the transmittance profile. The paper reflects most of the light that hits it and the scanner backing reflects light back toward the sensor. Intuitively, the generated image contains a component created by the light passing through the front side print and reflected from the paper back toward the sensor, but also another component created by light passing through front-side print, paper and back-side print and reflected back by the scanner backing (Fig.1). This component creates the show-through effect. This process is non-linear by nature [1].

Given R_f and R_b the clean front and back side reflectance images, respectively, and R_{wp} the reflectance of white (un-printed) paper, the normalized Optical Density images are defined by:

$$D_i(x, y) \equiv -\ln\left(\frac{R_i(x, y)}{R_{wp}}\right) \quad (1)$$

and the normalized Optical Absorbance images by :

$$A_i(x, y) \equiv 1 - \left(\frac{R_i(x, y)}{R_{wp}}\right) \quad (2)$$

where i =front/back.

The linearized model gives the scanned front side density image as a linear mixture of the front side density and the back side absorbance [1]:

$$D_f^s(x, y) = D_f(x, y) + h(x, y) * A_b(x, y) \quad (3)$$

$h(x, y)$ is a point spread function representing the (unknown) optical transmittance/absorbance properties of the paper (mostly attenuation and blurring).

3. CLEANING ALGORITHM

The cleaning algorithm proposed by [1] is a 2-D adaptation of 1-D echo-cancellation techniques used in telephony. Adaptive linear filters estimate and track the show-through point spread function.

The Algorithm:

1. Manual Estimation of R_{wp} (averaging reflectance values of areas with no print on either side).
2. Convert front-side reflectance values to density.
3. Convert back-side reflectance values to absorbance.
4. For each pixel (progressing in a spatial contiguous order):
 - (a) Compute show-through corrected density.

$$\hat{D}_f(m, n) = D_f^s(m, n) - \sum_{k=-N}^{k=N} \sum_{l=-N}^{l=N} w(k, l) A_b^s(m - k, n - l). \quad (4)$$

- (b) If back side has activity but not front side, update filter coefficients by LMS method

$$w'(k, l) = w(k, l) + \mu \hat{D}_f(m, n) A_b^s(m - k, n - l). \quad (5)$$

$$w(k, l) = w'(k, l).$$

- (c) Convert density to reflectance.

3.1. White-paper reflectance

The original algorithm described in [1] computes the white paper reflectance R_{wp} by taking the average reflectance value of a manually selected area of the images, not containing print on either the front or back side. We propose an automatic method based on the mean shift algorithm proposed in [5] and [6].

Typically, the brightness histogram of a scanned document is multi-modal (Fig. 2). We define the white paper reflectance as the peak of the brightest (rightmost) mode in the image. Other choices (such as taking the brightest pixel value in the image), are also possible, though our choice seems intuitive and gives good results.

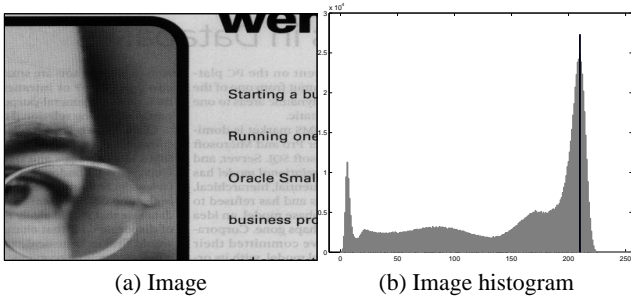


Fig. 2. Multi-modal PDF of image (R_{wp} marked by dark line).

The mean shift algorithm is a simple non-parametric technique for estimation of probability density gradient. An iterative steepest ascent algorithm is implemented, operating on the marginal probability density (the brightness histogram). The algorithm (with properly selected parameters) converges to a local probability density maximum (mode peak).

The Mean Shift Algorithm:

1. Choose the search window radius (number of brightness levels).
2. Choose initial window location (brightness level).
3. Compute mean value of pixels contained in the window and shift window location to it.
4. Repeat step 3 until convergence.

The point of convergence obviously depends on the starting point. Since we are seeking the peak of the brightest mode, the brightest pixel value in the image is chosen as the initial window position.

As a rule of thumb - the smaller the window radius - the better the separation resolution of the modes. However, if the radius is too small, the algorithm may get stuck on local maxima caused by discontinuities in the PDF estimate. Thus we want to find the maximal radius that gives us sufficient resolution, depending on the pixel distribution statistics of the image. In [5] the window radius chosen is proportional to the STD of the pixel brightness values. Values between 0.2σ and 0.4σ were chosen according to the application. We found that a radius of 0.1σ produced good results on the images tested.

We denote the global white paper reflectance values obtained by the above algorithm, R_{pg}^f and R_{pg}^b for the front and back-side images, respectively.

3.2. Local background normalization

The cleaning processes described above attempts to cancel or minimize in some sense (specifically least mean square) a difference, or error, value. According to the normalization proposed in [1], the zero value corresponds the global white paper reflectance. This means that when removing show-through, the algorithm attempts to bring the brightness value closer to R_{pg} , whether or not this is the "desired" value.

The desired value we are seeking is that of the background reflectance in the vicinity of the pixel being cleaned. We define this value as that of the brightest mode in a square area ($L \times L$ pixels) surrounding the pixel. This value can generally be found using a mean-shift process, with a few alterations, as explained below.

As before, the brightest pixel value in the area, is chosen as the initial window position. However, since the sample size is limited to a small area surrounding each pixel, the marginal probability may not give a good approximation of the density. Too small a mean-shift window radius may cause the mean-shift process to get stuck on an isolated value or local maxima that is not a true mode. In other cases the image segment analyzed contains only one narrow mode. This typically happens in image segments that are more or less uniform (such as segments with no print). In these cases, setting the window size as a small fraction of the STD, gives too small a value that may result in the algorithm getting stuck. A reasonable choice for the window radius (taking into account that the STD may be very low) is 0.5σ .

The solution to these problems is to allow the window radius to vary in size throughout the iterations [6]. The minimal radius is set in proportion to the STD, however if the window covers too few pixels (less than a certain percentage of the $L \times L$ area pixels), the

radius is increased until this condition is met. Increasing window size guarantees that the algorithm does not get stuck on isolated values, instead of an actual mode. We used 5% of the pixels as the minimal window coverage.

In some cases, where back-side activity is high, it is impossible to estimate the correct local background value, based on the immediate vicinity alone. In these cases the best we can do is use the global value. Back-side activity is measured by comparing average back-side and front-side reflectance values and by comparing the average back-side reflectance value to back-side R_{pg} . If the average back-side value is lower than 0.6 of the back-side R_{pg} , or if the local mean back-side value is lower than the local mean front-side value, the global front-side value is chosen.

The process described above is computationally heavy - a mean-shift process is run for most pixels. Several measures may be taken to reduce the computational load. Local background values need not be calculated for every pixel. Instead, we found it sufficient to re-calculate at horizontal and vertical intervals equal to half the dimensions of the pixel neighborhood. In order to avoid unwanted discontinuities, the local background images are smoothed by low-pass filters (gaussian, 15×15 pixels, $\text{std} = 2$). Additionally, when the STD is very low, i.e. the data is single modal, it is sufficient to use the mean pixel value in the neighborhood, instead of a mean-shift process.

A similar process is run for the back-side. We denote the local background reflectance values $R_{pl}^f(m, n)$ and $R_{pl}^b(m, n)$, for front and back-sides, respectively. These values replace the R_{wp} when calculating density and absorbance thus providing adaptation to local background levels.

3.3. Multi-stage filtering

One of the drawbacks of a MSE process, is that it has a tendency to concentrate it's effort on the larger filter coefficients, where most of the energy is concentrated. In doing so the process neglects the smaller coefficients which may not converge correctly.

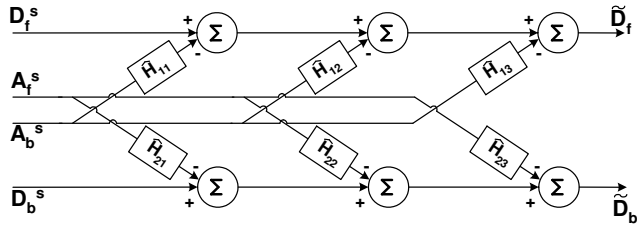


Fig. 3. Cascaded filter structure.

The solution to this problem is a cascaded filter structure (Fig. 3), replacing the single filtering stage, with large filter support, in the original algorithm. The filter support increases with each stage of the cascade. Each stage uses the previous stage's output as it's main input, but uses the same original reference (interference) signal.

Predictably, the later filtering stages contribute less with each stage. The first filtering stage produces the biggest improvement both visually and in similarity measures. Both mutual information and cross correlation are calculated between front and back side

images. We expect these values to decrease in magnitude with the reduction of show-through. The latter filtering stages provide visual improvement, however, improvement in the similarity measures is negligible. In some cases even a certain degradation occurs. We found that two or three filtering stages are sufficient.

3.4. Post-processing stage

Inherent in the LS approach (Fig. 4), as it is used in [1], is the assumption that there is no leakage of the primary input signal s_1 into the interfering signal s_2 . If both signals are coupled into each sensor (Fig. 5), the performance of a LS system may deteriorate. This is the case here where both front and back-side scanned images contain show-through. The LS approach, in this case, will cause a portion of the primary signal (the density image) to be cancelled out while removing the interfering signal (the absorbance image of the reverse side), thus distorting the recovered signal. A further post processing stage (Fig. 6) is required in the recovery system, to reduce this distortion [7].

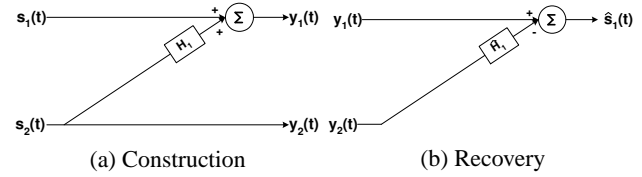


Fig. 4. Least-squares system.

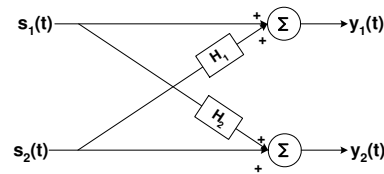


Fig. 5. Two channel construction system.

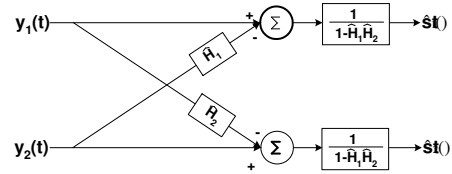


Fig. 6. Two channel recovery system.

Estimating the post-processing filter $\tilde{H} = \frac{1}{1-H_1H_2}$ is not in itself a trivial process. We propose using the fact that the scanned (contaminated) image contains an undistorted version of the clean image (except for noise). We estimate the post-processing filter and cancel the distortion through an LMS process (Fig. 7). In theory the error signal will contain only the uncorrelated show-through and the filter output the undistorted image. This process is run separately for the front and back side images.

3.5. Improved cleaning algorithm

To summarize, the proposed algorithm is:

1. Image registration.
2. Estimation of R_{pg} using the mean-shift process on the whole image for both front and back sides.

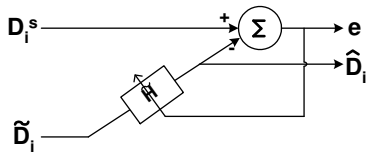


Fig. 7. Post processing filter estimation.

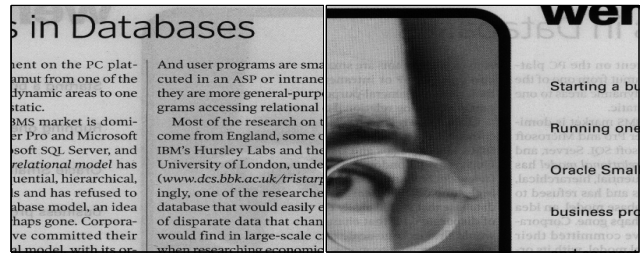
3. Estimation of R_{pl} image for both front and back: local mean-shift process coupled with activity estimate.
4. Convert front-side reflectance values to density using front-side R_{pl}^f .
5. For each pixel (progressing in a spatial contiguous order):
 - (a) Convert back-side reflectance values to absorbance around pixel location using front-side R_{pl}^f value at pixel location.
 - (b) Compute show-through corrected density.
 - (c) If back side has activity but not front side, update filter coefficients by LMS method.
6. Clip density values brighter than Density of R_{pg} .
7. Repeat stages 5 and 6 for larger filter supports (Fig. 3).
8. Post-processing stage - cancel distortion on front and back-side recovered images using their respective scanned versions in an LMS process (Fig. 7).
9. Convert density back to reflectance using R_{pl}^f .
10. Repeat stages 4-9 for back-side.

4. RESULTS

The original scanned images are shown in Fig. 8. The similarity measurements for these images are Mutual Information $MI = 0.105$, and the Normalized Cross Correlation $XC = 0.248$. The original algorithm's results (Fig. 9) using a single 31×31 pixel filter, have $MI = 0.027$ and $XC = 0.052$. While most of the show-through is removed, artifacts, such as the bright patches on the man's forehead in the back-side image, are common. The improved algorithm's results are shown in Fig. 10. A 31×31 ($L=31$) pixel window was used for the local background computation. The cleaning process used a three stage cascade with filter supports of 5×5 , 9×9 and 15×15 pixels. The post processing stage used 5×5 pixel filters. After cleaning, the similarity measures are $MI = 0.0295$ and $XC = 0.013$. While the change in the similarity measures is not great, visual improvement is significant. A larger amount of the show-through is removed with significantly less distortion to the image.

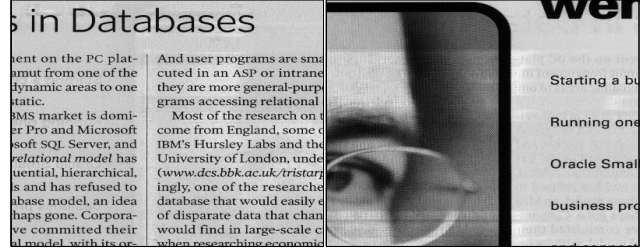
5. CONCLUSION

In this article we present an improved cleaning algorithm. The algorithm achieved superior results on the images tested. The results improved most noticeably on scans of complex documents containing not just text but also images. Future work on the problem may include employing more complex decorrelation algorithms (RLS) and methods relying on higher order statistics (ICA) for blind signal separation.



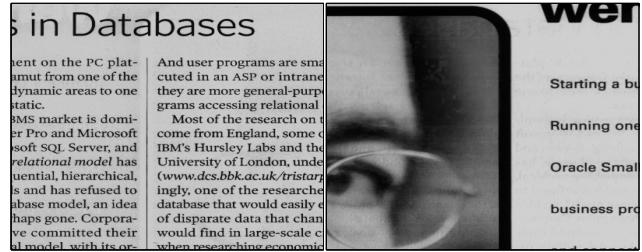
(a) Front (b) Back

Fig. 8. Scanned images.



(a) Front (b) Back

Fig. 9. Cleaned images based on [1].



(a) Front (b) Back

Fig. 10. Improved algorithm cleaned images.

6. ACKNOWLEDGEMENTS

We would like to thank Dr. Ehud Karnin and Dr. Eugene Walach of the IBM Research Laboratory in Haifa, for presenting us with the problem and providing the data used in the experiments.

7. REFERENCES

- [1] Gaurav Sharma, "Show-through cancellation in scans of duplex printed documents," *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 736–754, May 2001.
- [2] Simon Haykin, *Adaptive Filter Theory*, Prentice Hall, NJ, second edition, 1991.
- [3] Y.Y. Schechner, N. Kiryati, R. Basri, "Separation of transparent layers using focus," *International Journal of Computer Vision*, vol. 39, no. 1, pp. 25–39, Aug 2000.
- [4] B. Sarel, M. Irani, "Separating transparent layers through layer information exchange," in *ECCV 2004*.
- [5] D. Comaniciu, P. Meer, "Robust analysis of feature spaces: Color image segmentation," in *IEEE 1997 Conference on Computer Vision and Pattern Recognition*, pp. 750–755.
- [6] D. Comaniciu, V. Ramesh, P. Meer, "The variable bandwidth mean shift and data-driven scale selection," in *8th International Conference on Computer Vision*, vol. 1, pp. 438–445, 2001.
- [7] E. Weinstein, M. Feder, A.V. Oppenheim, "Multi-channel signal separation by de-correlation," *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 4, pp. 405–413, Oct 1993.