Bleed-through removal in degraded documents

Róisín Rowley-Brooke and Anil Kokaram

Department of Electronic & Electrical Engineering, Sigmedia Group
Trinity College Dublin
Ireland

ABSTRACT

This paper presents a linear-based restoration method for bleed-through degraded document images and uses a Bayesian approach for bleed-through reduction. A variation of iterated conditional modes (ICM) optimisation is used whereby samples are drawn for the clean image estimates, whilst the remaining variables are estimated via the mode of their conditional probabilities. The proposed method is tested on various samples of scanned manuscript images with different degrees of degradation, and results visually compared with a recent user-assisted restoration method.

Keywords: Document Image Restoration, Bayesian estimation, Bleed-through removal

1. INTRODUCTION

Reduction in legibility due to progressive degradation is often encountered in the study of documents, particularly handwritten. Many libraries host large collections of manuscripts and documents which are especially vulnerable to such degradations due to the fragile nature of the media on which they were written. Physical restoration of degraded documents is a cost and time intensive process, and may affect the integrity of the original. Restoration methods using automatic image processing techniques therefore have become increasingly popular as they have the advantage of being able to make any number of alterations to the document appearance, whilst leaving the original intact.

Loss of textual information in degraded documents may be classified into four categories:

(i) Fading of text due to light exposure or flaking ink.

(ii) Obscured or missing text due to degradation of the writing medium. This may be caused by damp, mould, parasites, or inherent brittleness in the medium.

(iii) Bleed-through interference, where ink has seeped through from one side of a page to the other.

(iv) Digitisation of documents may introduce noise artifacts and degrade the textual information. For example non stationary noise due to variable illumination,[1] show-through caused by the scanning of double sided documents,[2] and images that appear warped as a result of inherent curvature in the document, for example due to binding.

This paper focuses on degradation caused by bleed-through interference, and proposes a method for bleed-through removal using a linear-based degradation model. A previous version of this work can be found in,[3] however the model has since been altered with the addition of several new priors and a change in the iteration scheme used.

Further author information:
R.R.B.: E-mail: rowleybr@tcd.ie
A.K.: E-mail: anil.kokaram@tcd.ie

2. RELATED WORK

The problem posed by bleed-through degradation has attracted an increasing amount of attention over the past decade. Different approaches may be categorised into one of two groups; blind and non-blind. Blind methods work with an image of one side of the document only, whereas non-blind methods work with registered images of both the recto and verso sides, on the assumption that images of both are available.

Due to the relative lack of image data, blind methods often involve an intensity based thresholding step, making the assumption that the bleed-through intensity is distinct from that of the foreground text, such as hysteresis thresholding in Ref. 4, and the recursive unsupervised classification method of Ref. 5. Other blind approaches include that of Tonazzini et al.,[6] where they separate the RGB colour channels of a single image into foreground, background, and bleed-through classes using independent component analysis (ICA). Tonazzini then, in Ref. 7, applies this method to different colour space representations of manuscript images for bleed-through removal and information content maximisation. More recently Wolf, in Ref. 1, addresses the problem via a dual-layer Markov Random Field (MRF), using two hidden label fields and one observation field.

Non-blind methods have the advantage of more available data, however they are essentially two-stage processes. Firstly recto and verso sides must be registered so that they are aligned and of the same resolution, then the second stage is the bleed-through removal. Many non-blind methods use the increased amount of data to improve the results of thresholding algorithms compared to blind methods. Castro et al.[8] use Sauvola's thresholding algorithm[9] combined with fuzzy classification, while Ref. 10 proposes extensions to two binarisation methods, namely symmetric/non-symmetric Kullback-Leibler (KL) thresholding algorithms and the binarisation algorithm of Gatos et al.;[11] adding in a second threshold level for the bleed-through interference. These methods both focus solely on early music documents. Tonazzini et al.[12] extended the ICA method to grayscale images of double-sided documents, using the flipped verso image as one of the sources. More recently Huang et al. in Refs. 13, 14, propose a method for bleed-through reduction on more severely degraded documents that takes a small set of user input training data for the background, foreground text, and bleed-through interference of both recto and verso pages, and uses this to locate and remove bleed-through interference via a dual-layer MRF model. Moghaddam et al.[15] use diffusion models for the recto and verso texts, and the background medium, and apply a reverse diffusion model to remove interference. They have since applied this diffusion method in a unified framework,[16] using variational models for blind, non-blind, and severe bleed-through removal.

The linear-based model presented here is similar to that of Tonazzini et al.[12] However, constant mixing parameters across the images are not assumed, nor is it assumed that the extent of bleed-through is the same on both sides.

3. REGISTRATION

In order that the bleed-through on each side is aligned with the originating text on the other, registration of the two images needs to be performed as a pre-processing step. The global affine warp method of Dubois and Pathak[17] is used here, however a slight modification is added whereby the warp matrix is initialised from a set of user defined points. The user selects corresponding feature locations in the recto and verso images, and a least squares fit is performed to the displacements between these points to obtain the initial affine warp. This simple step has helped to reduce computation time significantly, and improve the performance of the algorithm. For the rest of this paper *verso* will refer to the flipped, registered image of the original verso side.

4. DEGRADATION MODEL

In the proposed model, it is assumed that an observed degraded recto pixel intensity $I_r(h,k)$, at image location (h,k), is a linear combination of the corresponding clean recto pixel $Y_r(h,k)$ and some proportion of the clean verso pixel $Y_v(h,k)$. The proportion of bleed-through is controlled by $\alpha_v(h,k)$, a mixing parameter $(0 \leq \alpha_v < 1)$, and also binary masks defined on both sides $M_r(h,k), M_v(h,k)$. As the problem is symmetrical, the formation of the observed verso pixel $I_v(h,k)$ is similar. The model for each side therefore is as follows (where pixel coordinates are discarded for brevity).

$$\begin{aligned} I_r &= Y_r + M_r(1 - M_v)\alpha_v Y_v + \rho \\ I_v &= Y_v + M_v(1 - M_r)\alpha_r Y_r + \nu \end{aligned} \tag{1}$$

The binary mask terms here are defined to be value 0 where the corresponding image is foreground text, and 1 everywhere else. It is assumed, as in Ref. 1, that bleed-through degradation is not visible through foreground text, and the combination of the mask terms explicitly ensures that this is the case by limiting the presence of the clean verso image, Y_v in the observed recto image to regions where Y_r is background and Y_v is foreground. This assumption holds for all but the most heavily degraded documents, where bleed-through can be seen to darken foreground characters. It is also assumed that the document image suffers from added noise from other degradations and also the scanning process. This noise is represented in the terms ρ, ν, and is assumed to be zero mean and Gaussian: $\mathcal{N}(0, \sigma^2_{\rho\rho})$, $\mathcal{N}(0, \sigma^2_{\nu\nu})$.

5. BAYESIAN FRAMEWORK

The bleed-through removal method is formulated under a Bayesian maximum a posteriori (MAP) framework. From the degradation model it is clear that the unknown parameter vector for estimation contains six variables: $\boldsymbol{\theta} = [\alpha_v, \alpha_r, M_v, M_r, Y_r, Y_v]$. Applying Bayes' rule, the p.d.f. for the posterior probability of $\boldsymbol{\theta}$ given the observed data $I_r,\ I_v$, at a single location, is then

$$p(\boldsymbol{\theta}|I_r, I_v, \tilde{M}, \tilde{\alpha}) \propto p(I_r, I_v|\theta, \tilde{M}, \tilde{\alpha}) p(\theta|\tilde{M}, \tilde{\alpha}), \tag{2}$$

where $\tilde{M}, \tilde{\alpha}$ represent the existing state of the mask and linear mixing parameter (on the recto or verso side as appropriate) in the neighbourhood of the pixel site currently being considered. In what follows, the case of the observed recto image is considered only (due to the symmetrical nature of the model, the formulation is similar for the verso side).

5.1 Likelihood

Following the degradation model the likelihood combines the influence of both the recto and verso sides to yield a joint Gaussian distribution

$$p(I_r, I_v|\boldsymbol{\theta}, \tilde{M}, \tilde{\alpha}) \propto \exp -\left\{ \frac{1}{2\sigma^2_{\rho\rho}} [I_r - Y_r - M_r(1 - M_v)\alpha_v Y_v]^2 + \frac{1}{2\sigma^2_{\nu\nu}} [I_v - Y_v - M_v(1 - M_r)\alpha_r Y_r]^2 \right\}. \tag{3}$$

Here it is assumed that the noise generating processes of the two sides are independent.

5.2 Priors

Prior models for each of the unknown parameters are designed based on the a priori knowledge of what they represent and the effects that they embody.

5.2.1 Masks

The mask variables $M_r,\ M_v$ mark the regions of foreground text, and should therefore be smooth in local patches. The widely used Gibbs energy prior for spatial smoothness makes sense in this case. In regions where there is neither foreground text, nor bleed-through degradation, there is no ambiguity and it is clear that these variables should be 1. It is possible to estimate roughly the regions of text, bleed-through and background (using K-means clustering on the degraded image). Using this rough estimate it is sensible then to include a prior constraining the mask variables in the definite non-text, that is, background region. Therefore the prior for the masks is

$$p(M_r|\tilde{M}) \propto \exp -\left\{ \sum_{\tilde{M}} (M_r - \tilde{M})^2 \lambda_M + \beta_r (1 - M_r) \right\}. \tag{4}$$

Here $\tilde{M}$ represents the current state of M_r in the 8-connected neighbourhood of the current site, and λ_M is a smoothness hyperparameter to encourage the spatial smoothness (in essence an Ising model in this case due to the binary nature of the masks). In all experiments $\lambda_M = 1$ was set empirically. β_r represents a penalty for setting $M = 0$ in the regions of non-text, and is estimated in the initialisation step discussed below.

5.2.2 Mixing Parameters

Priors for the mixing parameters, $\alpha_v,\ \alpha_r$ follow similar logic. Firstly, smoothness is encouraged with the Gibbs energy prior. Secondly, it is necessary when the mixing parameters are present to constrain them to values that would yield material that is close in intensity to the background regions in the clean image. Without this constraint there is nothing limiting the mixing parameters to create useful images since the smoothness of itself does not constrain the absolute value. Finally to prevent the smoothness prior both from spreading the mixing parameters into foreground regions, and also from creating the impossible scenario of bleed-through on both sides in the same location, the mixing parameters are constrained to the initial mask estimates, and against each other. Hence the prior model is

$$p(\alpha_v|\tilde{\alpha}) \propto \exp -\left\{ \sum_{\tilde{\alpha}} (\alpha_v - \tilde{\alpha})^2 \lambda_\alpha + (\alpha_v - \bar{\alpha}_v)^2 \lambda_c + (\bar{M}_r \alpha_r \alpha_v^2) \lambda_c \right\}, \tag{5}$$

where $\tilde{\alpha}$ are values of the α_v in the 8-connected neighbourhood of the current site, $\bar{\alpha}_v$ is a *rough* estimate of α_v, obtained during initialisation, $\bar{M}_r$ is the initial estimate for the recto mask, and $\lambda_\alpha,\ \lambda_c$ are smoothness and constraining hyperparameters respectively, set empirically to $\lambda_\alpha = 55,\ \lambda_c = 5$. The smoothness parameter is much higher than for the mask prior so as to ensure that the final clean images retain as little trace of the bleed-through as possible, and to avoid harsh intensity cut offs at character edges; the purpose of the mixing parameters is to blend the bleed-through regions into the background.

5.2.3 Clean Images

A prior is used for the clean image data to encourage, as for the mixing parameters, that the average brightness of the restored document in the bleed-through regions matches the average brightness of the background. This is again a simple Gaussian prior with parameters derived at initialisation.

$$p(Y_r|\bar{Y}_b) \propto \exp -\left\{ \frac{1}{2\sigma_{\rho\rho}^2} M_r(1 - M_v)\alpha_v (Y_r - \bar{Y}_b)^2 \right\}. \tag{6}$$

Here, $\bar{Y}_b$ is an estimate of the average background intensity of the clean image, and $\sigma_{\rho\rho}^2$ is the variance of the background noise as defined above in Section 4. The multiplier $M_r(1 - M_v)\alpha_v$ ensures that this constraint is restricted to bleed-through restored regions only.

6. SOLUTION

To solve for all the variables, iterated conditional modes (ICM) optimisation[18] is used. A slight modification to the usual process however is made here in that *samples* are drawn for the underlying clean images $Y_r,\ Y_v$, while the mode of the conditionals are selected for the masks and mixing parameters. The solution for each of the variables is presented below in the case of the recto side only, as, again, the verso equations are similar due to the symmetry of the problem.

6.1 Mask Estimate

As the masks are clearly present in both observations, the estimates for the masks are generated using the conditional

$$\begin{aligned} p(M_r|\alpha_r, \alpha_v, M_v, Y_r, Y_v, I_r, I_v) \propto \exp -\Bigg\{ & \frac{1}{2\sigma_{\rho\rho}^2} (I_r - Y_r - M_r(1 - M_v)\alpha_v Y_v)^2 \\ & + \frac{1}{2\sigma_{\nu\nu}^2} (I_v - Y_v - M_v(1 - M_r)\alpha_r Y_r)^2 + \sum_{\tilde{M}} (M_r - \tilde{M})^2 \lambda_M + \beta_r(1 - M_r) \Bigg\}. \end{aligned} \tag{7}$$

In this case the estimation is performed numerically since M_r is binary: both $M_r = 0$ and $M_r = 1$ are substituted in the expression above, and whichever yields the greater probability is selected.

6.2 Mixing Parameter Estimate

Each mixing parameter is present in only one of the observation terms, therefore the conditional probability at a site is

$$p(\alpha_v|M_v,M_r,Y_r,Y_v,I_r,I_v) \propto \exp -\Big\{ \frac{1}{2\sigma^2_{\rho\rho}}(I_r - Y_r - M_r(1-M_v)\alpha_v Y_v)^2 + \sum_{\tilde{\alpha}}(\alpha_v - \tilde{\alpha})^2\lambda_\alpha + (\alpha_v - \bar{\alpha}_v)^2\lambda_c + (\bar{M}_r\alpha_r\alpha_v^2)\lambda_c \Big\}. \quad (8)$$

The estimate of α_v at each site is obtained analytically since the expression is quadratic in α_v. Hence

$$\hat{\alpha}_v = \frac{2\sigma^2_{\rho\rho}\big[\sum_{\tilde{\alpha}}\tilde{\alpha}\lambda_\alpha + \bar{\alpha}_v\lambda_c\big] + (I_r - Y_r)M_r(1-M_v)Y_v}{2\sigma^2_{\rho\rho}\big[\sum_{\lambda_\alpha}\lambda_\alpha + \lambda_c(1+\bar{M}_r\alpha_r)\big] + (M_r(1-M_v)Y_v)^2}. \quad (9)$$

6.3 Clean Image Estimate

The estimates for the clean images are generated with the conditional

$$p(Y_r|\alpha_r,\alpha_v,M_r,M_v,Y_v,I_r,I_v) \propto \exp -\Big\{ \frac{1}{2\sigma^2_{\rho\rho}}(I_r - Y_r - M_r(1-M_v)\alpha_v Y_v)^2 + \frac{1}{2\sigma^2_{\nu\nu}}(I_v - Y_v - M_v(1-M_r)\alpha_r Y_r)^2 + \frac{1}{2\sigma^2_{\rho\rho}}M_r(1-M_v)\alpha_v(Y_r - \bar{Y}_b)^2 \Big\}. \quad (10)$$

This is clearly Gaussian, however instead of maximising the conditional following the ICM process, a sample is drawn from this distribution within $\pm T$ standard deviations of the mean (T is set to 0.5 in what follows). This strategy has been employed by other authors working in video and audio restoration.[19] The reasoning behind this is that the mean tends to generate over smooth images, while using an unconstrained random draw is visibly chaotic. Therefore by drawing samples within some distance of the mean, a textural component in the underlying signal is allowed for and the iterative process performs better. The required draw is therefore $Y_r \sim \mathcal{N}(\bar{Y}_r, \sigma^2_Y)$. By completing the square in the conditional above the mean and variance are extracted:

$$\bar{Y}_r = \frac{I_r + M_r(1-M_v)\alpha_v(\bar{Y}_b - Y_v) + \sigma^2(I_v - Y_v)M_v(1-M_r)\alpha_r}{1 + \sigma^2(M_v(1-M_r)\alpha_r)^2 + M_r(1-M_v)\alpha_v}, \quad (11)$$

where $\sigma^2 = \sigma^2_{\rho\rho}/\sigma^2_{\nu\nu}$

$$\sigma^2_Y = \frac{\sigma^2_{\rho\rho}\sigma^2_{\nu\nu}}{\sigma^2_{\nu\nu}(1 + M_r(1-M_v)) + M_v(1-M_r)\alpha_r\sigma^2_{\rho\rho}}. \quad (12)$$

The estimate for σ^2_Y depends on estimates for $\sigma^2_{\rho\rho}, \sigma^2_{\nu\nu}$, in the proposed method these estimates are obtained from the observed data using the initial mask estimates.

A secondary estimation method for Y_r, Y_v that is used every N iterations ($N = 10$ in this case) is to maximise conditionals that ignore the mask terms and priors, hence

$$p(Y_r|\alpha_r,\alpha_v,Y_v,I_r,I_v) \propto exp - \Big\{ \frac{1}{2\sigma^2_{\rho\rho}}(I_r - Y_r - \alpha_v Y_v)^2 + \frac{1}{2\sigma^2_{\nu\nu}}(I_v - Y_v - \alpha_r Y_r)^2 \Big\}. \quad (13)$$

Due to the quadratic nature of Eq.13 in Y_r, an estimate of Y_r at a site is obtained analytically, giving

$$\hat{Y}_r = \frac{I_r - \alpha_v Y_v + \sigma^2(I_v - Y_v)\alpha_r}{1 + \sigma^2\alpha_r^2}, \quad (14)$$

where σ^2 is defined as above.

6.4 Initialisation

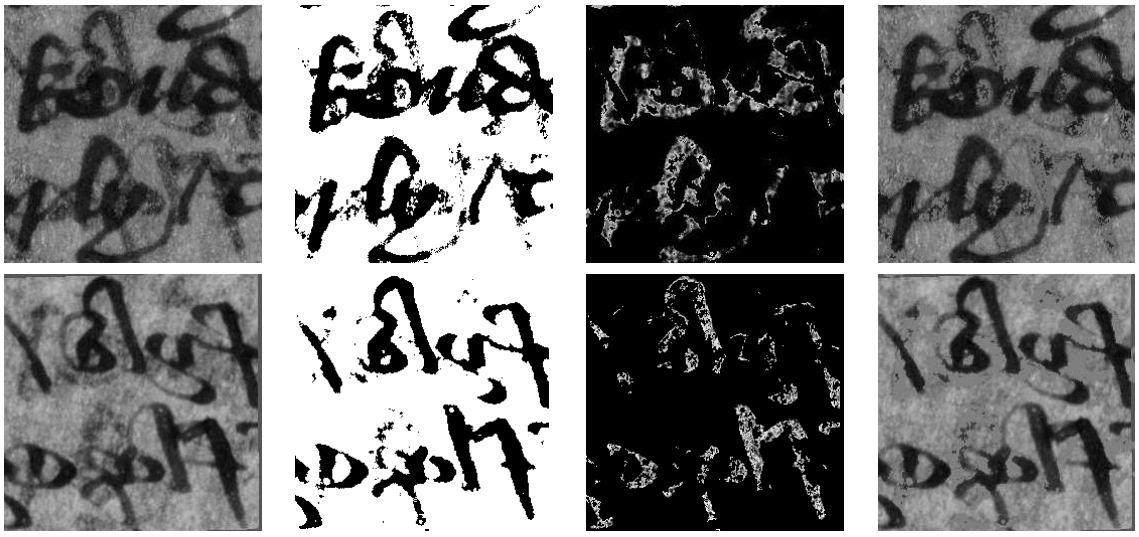
It is well known that good initial estimates are required for ICM to converge usefully, see Fig. 1. In this work the masks are initialised using k-means clustering on the intensity of the observed images, using 3 clusters. The lightest two clusters are considered to be background of the image and the darkest cluster is then the foreground text. The mask penalties β_r, β_v are calculated in the same step by creating further masks B_r, B_v to be to be coincident with the brightest cluster where there is no text. β_r is then configured as $100(B_r)$. The masks and observed images are then used to estimate the noise variance for both recto and verso sides, based on the variance of regions where $B_r and B_v = 0$, also the estimates for mean background intensity $\bar{Y}_b$ (for each side as appropriate) are taken to be the means of these regions. The initial mixing parameters are then obtained from the mask estimates and the observed images as follows (for α_v). Finally initial clean recto and verso estimates

---

At each pixel location
$\alpha_v = 0$
if recto is background ($M_r = 1$) **then**
  if verso is foreground ($M_v = 0$) **then**
    if $Y_r > \bar{Y}_b$ **then**
      $\alpha_v = \frac{Y_r - \bar{Y}_b}{Y_v}$
    end if
  end if
end if

---

are obtained by substituting the relevant initial estimates into Eq.1.



(a) (b) (c) (d)

Figure 1. (a): degraded recto (top) and verso images of the Piers example, (b): initial mask estimates using k-means clustering ($k = 3$), (c): initial estimates of α_r and α_v, (d): initial clean recto and verso estimates.

6.5 Algorithm

The overall algorithm may be enumerated as follows:

1. Register the recto and verso sides of the image using the process outlined in Section 3.
2. Initialise all variables as described above.

3. Using a checkerboard visitation pattern for sites repeat until convergence:

   (a) Generate $\hat{M}_v,\ \hat{M}_r,\ \hat{\alpha}_r,\ \hat{\alpha_v}$ using the expressions above, (across all sites in separate image passes) updating in place.

   (b) If mod(iteration no.,10) = 0 generate $Y_r,\ Y_v$ without the masks as in Eq.14
   Else draw samples for $Y_r,\ Y_v$ as above

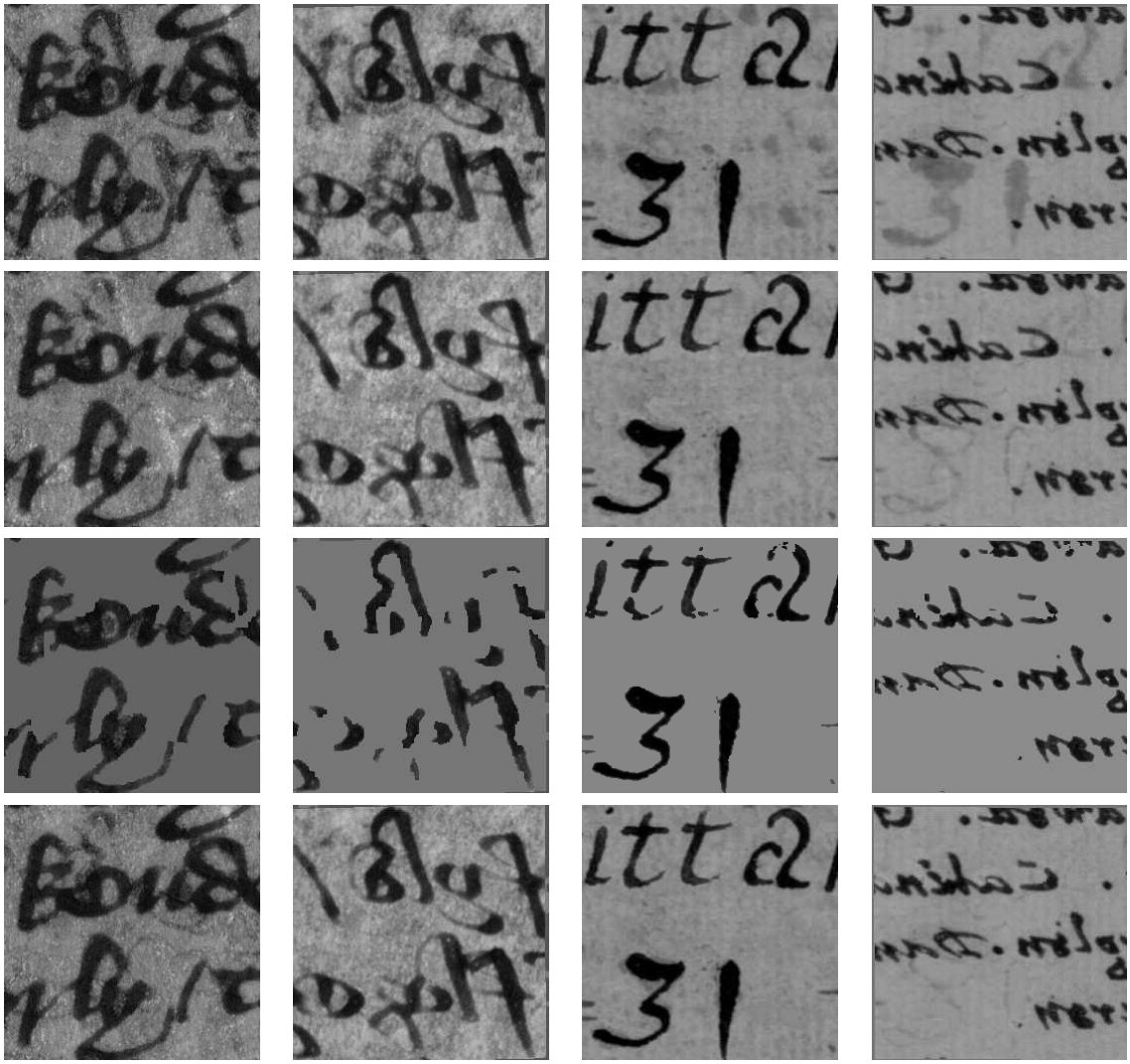


Figure 2. 1st row, left to right: degraded recto and verso of a sample from the Piers image and degraded recto and verso of a sample of the Welsh image. 2nd row: Results of the previous work. 3rd row: Results of the Huang et al. method. 4th row: Results of the proposed method.

7. RESULTS

The proposed algorithm was tested on sample extracts from larger high resolution (600dpi) images of from various sources, and with varying degrees of bleed-through degradation. The manuscript images used were from *Piers Plowman* 'B' text from the late 14th century, fol.22, MS.201, Corpus Christi Library, Oxford (*Piers*), a Welsh dictionary from 16th/17th Century, fol.1, MS.16 Jesus College Library, Oxford (*Welsh*), and a 16th Century

Irish Primer, fol.7,The Benjamin Iveagh Library, Farmleigh by kind permission of the Governors and Guardians of Marsh's Library, Ireland (*Irish*). The results obtained are compared here to results obtained from the related work,[3] and from the user-assisted method of Huang et al.[14] with a user markup of 9 strokes per image (3 each for the recto, verso, and bleed-through labels).

Figure 2 shows results from two 255x255 samples from the Piers and Welsh images respectively. Shown are the recto and verso of each, results from the previous, related work, results from the Huang et al. method, and results from the proposed linear-based method. Figure 3 similarly shows results from a 553x623 sample of the Irish image, and a 497x523 sample of the Piers image.

Figure 3. 1st row, left to right: degraded recto and verso of a sample from the Irish image and degraded recto and verso of a larger sample from the Piers image. 2nd row: Results of the previous work. 3rd row: Results of the Huang et al. method. 4th row: Results of the proposed method.

All the methods used perform well in the examples where the bleed-through is light, that is the Welsh sample (Fig. 2), and the Irish sample (Fig. 3). The Huang et al. result copes best with the lateral diffusion of the bleed-through text in the Welsh sample (see the handwritten number '3' in Fig. 2), whereas our previous method did not account for these regions, and there are still some artefacts in the results of the current method. However the Huang method does remove some foreground text most notably in the recto image. The proposed method performs best on the Irish sample, as our previous work leaves visible artefacts, and the Huang et. al method removes some foreground text.

In the more severely degraded samples from the Piers image (Figs. 2&3), our previous method successfully locates the bleed-through, but again leaves visible artefacts after restoration. In both these examples, the Huang et al. method performs better than the proposed method in regions where the bleed-through is equivalent in intensity to, or darker than foreground text. This is to be expected as a result of the initial user markup. However, overall the proposed method produces a visually more pleasing result, as the Huang et al. restoration is often too harsh such that many foreground regions are replaced with background. The proposed model allows for some ambiguity in the classification, since restoration and classification are essentially performed concurrently. This results in smoother restored images where the inherent character of the medium is preserved.

8. CONCLUSION

A relatively simple linear-based method for bleed-through removal has been proposed in this work. The algorithm relies on estimates generated from local pixel neighbourhoods in an ICM-based scheme. The performance of the proposed method has been evaluated on small samples images from various manuscripts with differing degrees of bleed-through degradation, and has been shown to perform well in comparison to one of the most recent new methods. However, as there is a lack of ground truth for restoration results, it is difficult to quantify and compare performance accurately, especially as the definition of a *restored document* can be subjective. The proposed restoration method resulted in document images where the degradation was significantly reduced, whilst allowing for some ambiguities in difficult regions, and maintaining the underlying texture of the original image. The significant downfall of this method, however, is apparent in regions where bleed-through is darker than foreground text and currently this issue is being addressed with the introduction of some user assistance.

ACKNOWLEDGMENTS

This research has been funded by the Irish Research Council for Science, Engineering, and Technology (IRCSET) 'Embark' Initiative.

REFERENCES

[1] Wolf, C., "Document ink bleed-through removal with two hidden markov random fields and a single observation field," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **32**(3), 431–447 (2010).

[2] Sharma, G., "Cancellation of show-through in duplex scanning," *Proc. Image Processing, International Conference on* **2**, 609–612 vol.2 (2000).

[3] Rowley-Brooke, R. and Kokaram, A., "Bleed-through removal in degraded manuscripts," *IET Irish Signals and Systems Conference* (2011).

[4] Estrada, R. and Tomasi, C., "Manuscript bleed-through removal via hysteresis thresholding," *Document Analysis and Recognition, ICDAR '09, 10th International Conference on* , 753–757 (2009).

[5] Drira, F., Le Bourgeois, F., and Emptoz, H., "Restoring ink bleed-through degraded document images using a recursive unsupervised classification technique," in [*Document Analysis Systems VII*], Bunke, H. and Spitz, A., eds., *Lecture Notes in Computer Science* **3872**, 38–49, Springer Berlin / Heidelberg (2006).

[6] Tonazzini, A., Bedini, L., and Salerno, E., "Independent component analysis for document restoration," *International Journal on Document Analysis and Recognition* **7**(1), 17–27 (2004).

[7] Tonazzini, A., "Color space transformations for analysis and enhancement of ancient degraded manuscripts," *Pattern Recognition and Image Analysis* **20**(3), 404–417 (2010).

[8] Castro, P., Almeida, R. J., and Pinto, J. R. C., "Restoration of double-sided ancient music documents with bleed-through," in [*Progress in Pattern Recognition, Image Analysis and Applications*], *Lecture Notes in Computer Science* **4756**, 940–949, Springer Berlin/Heidelberg (2007).

[9] Sauvola, J. and Pietikinen, M., "Adaptive document image binarization," *Pattern Recognition* **33**(2), 225–236 (2000).

[10] Burgoyne, J. A., Devaney, J., Pugin, L., and Fujinaga, I., "Enhanced bleedthrough correction for early music documents with recto-verso registration," *International Conference on Music Information Retrieval* , 407–412 (2008).

[11] Gatos, B., Pratikakis, I., and Perantonis, S. J., "Adaptive degraded document image binarization," *Pattern Recognition* **39**(3), 317–327 (2006).

[12] Tonazzini, A., Salerno, E., and Bedini, L., "Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique," *International Journal on Document Analysis and Recognition* **10**(1), 17–25 (2007).

[13] Huang, Y., Brown, M. S., and Xu, D., "A framework for reducing ink-bleed in old documents," *Computer Vision and Pattern Recognition, CVPR 2008. IEEE Conference on* , 1–7 (2008).

[14] Huang, Y., Brown, M. S., and Xu, D., "User-assisted ink-bleed reduction," *Image Processing, IEEE Transactions on* **19**(10), 2646–2658 (2010).

[15] Moghaddam, R. F. and Cheriet, M., "Low quality document image modeling and enhancement," *International Journal on Document Analysis and Recognition* **11**(4), 183–201 (2009).

[16] Moghaddam, R. F. and Cheriet, M., "A variational approach to degraded document enhancement," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **32**(8), 1347–1361 (2010).

[17] Dubois, E. and Pathak, A., "Reduction of bleed-through in scanned manuscript documents," *IS&T Image Processing, Image Quality, Image Capture Systems Conference (PICS2001)* **4**, 177–180 (2001).

[18] Besag, J., "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society. Series B (Methodological)* **48**(3), 259–302 (1986).

[19] O'Ruanaidh, J. and Fitzgerald, W., [*Numerical Bayesian Methods Applied to Signal Processing*], Springer (1996).