

A Variational Approach to Degraded Document Enhancement

Reza Farrahi Moghaddam and Mohamed Cheriet, *Senior Member, IEEE*

Abstract—The goal of this paper is to correct bleed-through in degraded documents using a variational approach. The variational model is adapted using an estimated background according to the availability of the verso side of the document image. Furthermore, for the latter case, a more advanced model based on a global control, the *flow field*, is introduced. The solution of each resulting model is obtained using *wavelet shrinkage* or a time-stepping scheme, depending on the complexity and nonlinearity of the models. When both sides of the document are available, the proposed model uses the reverse diffusion process for the enhancement of double-sided document images. The results of experiments with real and synthesized samples are promising. The proposed model, which is robust with respect to noise and complex background, can also be applied to other fields of image processing.

Index Terms—Variational framework, *PDE-based* image processing, document enhancement, bleed-through effect, **ACM Taxonomy:** I.4.3.a, I.7.5.a

I. INTRODUCTION

DOCUMENTS are usually archived in the form of digital images for reproduction, distribution, retrieval, indexing or analysis. Except for recently created electronic documents, it is usually the scanned version of a document that is archived. However, scanned images often suffer degradation, such as *noise, optical blur, CCD jitter*, and low resolution, among others. The acquired images can be preprocessed to correct these problems before archiving. This helps reduce file size, making archiving and retrieval easier. Also, preprocessed and clean documents are of more interest to the user. Some examples of processing tasks are Optical Character Recognition (OCR), feature extraction, and automatic retrieval. An example of a postprocessing task is *automatic translation* which requires digital content-available documents as the end-product. Moreover, the performance of processing systems is greatly improved if the input images are clean, which means that they are free from distortion, interference, and degradation.

Usually, the application of advanced imaging techniques and the use of high-quality materials in document scanning results in degradation-free output at the scanning stage. In addition, many degradation models have been developed for scanners which may be used in the development of ways to compensate for the degradation caused by the scanners themselves [1]. However, if the documents suffer from physical degradation, that degradation will pass through the system and appear on the output image. Proper handling of documents will reduce

the physical degradation caused by external sources, such as humidity, heat, etc. Aging also results in physical degradation, depending on the quality of the paper and ink used. In very old documents, this degradation is usually noticeable. Finally, physical degradation can be caused by poor quality printing materials even in recent documents.

Very old and ancient documents are useful for studying many of the cultural, typographical, and psychological aspects of the writer and his environment. Despite a great deal of interest in the analysis of old documents, access to these materials is limited to the digital images that have been made of them to extend their life. Because of the often complex nature of such documents, the whole of the scanned image is, in many cases, archived without any processing. Therefore, good document preprocessing is important for easy archiving and future processing. Also, a great deal of effort is being put into the development of huge searchable datasets of documents and books such as the Google Book Search dataset [2]. Aging and the poor quality of materials are two of the biggest challenges in processing and archiving large numbers of books.

One of the most difficult problems with very old documents of poor quality is the *bleed-through effect*. This effect is the result of the diffusion of ink from one side of a page through to the other side. There are some similar degradation phenomena, such as the *shadow-through effect*. In this case, because of the transparency of the paper, the information on one side of document interferes with the information contained on the other. The bleed-through and shadow-through effects are not found only in very old documents. Even in today's newspapers, bleed-through and other effects reduce legibility. From a physical point of view, the bleed-through phenomenon is a diffusion process of water/ink through paper [3]–[8]. There are many similar phenomena, such as the seepage of water or oil through soil, which are usually modeled *using diffusion-based models* [9]–[11].

In this work, we focus on degradation effects like bleed-through. The goal is to obtain a clear image of the scanned document. In this way, the archival process can be very cost-effective and fast. If needed, these enhanced versions are ready for input to a retrieval system, with the objective of obtaining a high-quality text-only version of a document.

The following terminology is used throughout the paper. As input, we have two images which are the scanned images of two sides of a document. The main side is identified as the “recto” image and the other side the “verso” image. It must be noted that neither the recto nor the verso image offers any advantage from a computational perspective, and both sides

Synchromedia Laboratory for Multimedia Communication in Telepresence, École de Technologie Supérieure, Montréal, (QC) H3C 1K3 Canada (e-mail: reza.farrahi-moghaddam.1@ens.etsmtl.ca and mohamed.cheriet@etsmtl.ca)

undergo equivalent processing. In other words, the calculations are completely symmetrical. It will be sufficient, because of that symmetry, to discuss only the computations for one side of the document. We refer here to the side that is currently under process as the recto side, the other being the verso side.

The structure of the paper is as follows. In section II, a summary of previous works on document enhancement, especially for double-sided documents, is presented. In section III, a variational model is introduced which enhances document images suffering from the bleed-through effect. The solution is presented in an analytic form in the wavelet domain. Then, in section IV, another variational model is introduced which uses estimated background information for the enhancement of degraded documents in the case when the verso side of the document is not available. For highly degraded documents containing very fine features and weak edges, a more sophisticated variational model is used, which is presented in section V. This model uses a global classifier, the flow field, for preserving fine features and weak edges. In the experimental section, the performance of the models is evaluated in both subjective and objective ways. Finally, in section VII, a summary of our work is presented and some future research directions suggested.

II. PREVIOUS WORKS

Several studies have been conducted on the separation of text from intense interference patterns. In [12], multistage thresholding was used for the extraction of text in degraded documents. They processed a single-sided document as the input. It was found that popular thresholding techniques such as Otsu, are not suitable for the multi-stage thresholding separation approach. The following are other examples of such studies: applying adaptive filters to the separation of forensic handwritten text from the background [13], analyzing ink traces in handwritten documents in the field of forensic science based on features such as start points and cross points [14], and applying multi-scale analysis and adaptive binarization for the correction of shadow-through in color images [15]. A noise-based thresholding technique has even been presented for the extraction of main text from background and noise [16]. In another approach, a statistical adjustment, based on entropy, is performed between the input gray-scale and output binary versions of an image in order to remove the interfering patterns [17]. In all these studies, separable gray levels or other distinctive features in the input image were assumed for the main text and the background.

If two images of a document are available, i.e. a scanned image of each side, the problem is usually formulated as a system of two undetermined and nonlinear equations with two variables (two source images). For example, in [18], a transmission model was developed to correct shadow-through patterns, on the assumption that the printed matter on the two sides of the document were two layers that influenced the scanned output. The model was then linearized. The equations were solved inversely to obtain the original layers. In other studies, the solution of the system of equations was expressed in terms of some correction functions or transformations.

In [19], the authors proposed a correction function, which actually substrates the two images to obtain undistorted images to correct the shadow-through problem in double-sided document scanning. Also, in [20], before applying their wavelet-based method, they used a transformation which weakens the suspected interference on the combination of double-sided inputs.

Transformation and filters are also used to extract other features from the source images, and those extracted features are used to reconstruct the text. For example, in [21], for double-sided scans of double-sided handwritten documents, a thresholding filter was applied for the extraction of foreground edges and the elimination of edges from the other side. Then, the edges were used to reconstruct the final output. The method was based on the fact that the edges of the interference patterns from the verso side are not very sharp. Their edge detector was adapted to ignore these types of interference pattern. Also, they used an orientation filter in the edge detector, as their target application was the restoration of historical documents in the National Archives of Singapore, in which the strokes are mainly oriented from lower-left to upper-right.

Multi-stage analysis of handwritten documents has also been proposed. In the first step of this type of analysis, an attempt is made to identify all parts of the text by operations like segmentation. Then, using the inpainting technique, the interference patterns are removed. For example, in [22] and [23], the correction involves the application of a series of operations, such as the registration of the two sides, the identification of bleed-through regions, and replacing the bleed-through pixels with background intensity.

Statistical methods have also been applied to double-sided scanned documents. Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are used in a Blind Source Separation (BSS) manner to extract the main text from the other interference patterns [24], [25]. Statistical methods are very powerful in separation tasks, but their computational cost is also very high. As well, to obtain a good human/machine readable image, other postprocessing tasks, such as removing residual information and connecting the gaps, are needed. Totally linear models have been proposed for the extraction of double-sided document text. In [26], the authors focused on documents which are degraded by the presence of bleed-through, proposing a linear model based on a BSS technique to separate recto from verso information. Linear models, despite their lower computational cost, are not very suitable for the analysis of nonlinear phenomena. Moreover, the presence of input noise may cause these models to fail.

Variational approaches based on the degradation model are also proposed for restoring a document with the shadow-through problem. In [27], a total variational regularization considering a location-dependent model for shadow-through degradation is used as a BSS technique to restore document images affected by shadow-through.

In conclusion, the above mentioned methods suffer from the following drawbacks: (i) In many approaches, several steps, such as segmentation, edge detection, identification, inpainting, etc, are required to obtain the desired result. It

is often assumed that the interference from the back side of the document does not have same the features as the front side. Also, some of them are based on specific features of the input script such as the orientation of the strokes. (ii) The statistical methods are point-wise and very sensitive to uniformity in the input images, and they do not use the two-dimensional spatial correlation between the input information. (iii) The model-based variational methods provide the degradation model along with the restoration facility. However, they are very sensitive to the selected degradation model. Therefore, like the statistical methods, they cannot be used in highly degraded cases. This limits their application to the shadow-through problem. (iv) Many of the methods are dependent on the presence of the verso side of the document image. (v) Usually, noise and other background artifacts are not removed. The above analysis is our motivation for proposing a unified variational-based framework for the bleed-through problem. In our approach, based on the availability of the verso side, different models are developed which can be applied to document images degraded by bleed-through or shadow-through. In this way, the interference patterns are completely removed from the document image. In addition, the computational cost will be addressed for different models.

III. RESTORATION OF DOUBLE-SIDED DOCUMENT IMAGES

A. Problem statement

A set of two images, resulting from the double-sided scanning of a document, is given (recto and verso side images): $u_{r0}(x)$ and $u_{v0}(x)$ where $x = [i, j]^T \in \Omega \subset \mathbb{R}^2$. The domain Ω is an open rectangle, and we call the set of all pixels on Ω “discrete Ω ”. The image values extend from zero to one, where a zero pixel value means that the pixel is white, and a value of one is equivalent to a black pixel. Because of the poor quality or the aging process of the document paper, ink from both sides has bled through the paper, causing interference patterns on the recto side. Also, external sources have resulted in some degree of noise in each image. The task is to obtain a clear, good quality image of the recto side of the document which is free of interference, noise, and other degradation, while all the document’s typographical and textual details, such as strokes and curves, are preserved.

B. Notations

For better understanding of the key variables in the formulation of the work, their notations and a brief discussion are presented here.

- *Estimated background:*

This is an estimation of the background for each pixel, and it is used as *a priori* information. Even pixels that are on strokes have an estimated background value. We use the algorithm in [28] to calculate these values. In this algorithm, an *a priori* threshold value, δ_{bg} , is used to compute the estimated background. A brief description of the algorithm and the definition of δ_{bg} are provided in the Appendix. The estimated background is a function of the input image and will not change through the computations. In notations, for example, the estimated

background of the recto side u_r is denoted by u_{rb} using a subscript b .

- *Target background:*

The target background is a preselected value for each pixel. This value will be used as *a priori* information to fill the pixel, if that pixel is supposed to be part of an interfering pattern, noise, or other type of degradation. Usually, a fixed function for the target background of all pixels can be used to have a clean and uniform background at the output. However, any form of information can be used as the target background. In this work, a fixed value of 0.1 is used for target background. We use a subscript t to indicate the target background in the notations. For example, u_{rt} represents the target background for the recto side. Note that these two concepts, target background and estimated background, are completely different, and, despite their similar names, they have different roles in the formulation.

- *Wavelet representation:*

For wavelet representation, we use the Dual-Tree Complex Wavelet Transform (DTCWT)¹ [29]. This transform is shift-invariant and directionally selective in two or more dimensions. In terms of the DTCWT, a function $u(x)$ can be represented by its wavelet coefficients $d^{l,o_1,o_2}(x_l)$, where l is the level superscript, and o_1 and o_2 are the orientation superscripts. The level superscript varies from 1 to L . L is selected such that the block size of the coarse scale ($l = L$) is less than half the stroke width. The low-pass coefficients are also represented by $d^{L+1,o_1,o_2}(x_L)$. For simplicity’s sake, all the coefficients will be represented by $d^{l,o_1,o_2}(x_l)$ in the following sections. The DTCWT, by definition, requires that Ω be a square. The input images can be cropped or extended to fulfill this requirement. The coordinate x_l is in the reduced domain Ω_l ; If discrete Ω is a square of $2^N \times 2^N$ pixels, then Ω_l is a square of $2^{N-l} \times 2^{N-l}$ pixels. It is worth noting that the wavelet coefficients are complex numbers.

C. Formulation

Below, a variational model is presented which provides an enhanced version of the recto side. A similar formulation can be used for the verso side in a symmetrical way. Normally, image functions belong to $L_2(\Omega)$. However, this function space is not suitable for wavelet analysis. Therefore, the image functions are restricted to other function spaces, such as the Besov spaces $B_p^\alpha(L_p(\Omega))$ of minimal smoothness² [31]–[33]. It is worth noting that these function spaces are embedded in $L_2(\Omega)$ [34]. Therefore, image functions can be expressed using their equivalents from these spaces.

In terms of wavelet coefficients, we propose to use the following variational model for the double-sided bleed-through problem:

$$\tilde{d}_r = \arg \min_{d_r \in B_p^\alpha(l_p)} J(d_r) \quad (1)$$

¹<http://taco.poly.edu/WaveletSoftware/dt2D.html>

²Roughly speaking, a function in $B_p^\alpha(L_p(\Omega))$ has α bounded “derivatives” in $L_p(\Omega)$ [30]. Also, minimal smoothness requires that $\alpha = 2/p - 1$.

where

$$\begin{aligned}
 J(d_r) &= \|u_r - u_{r0}\|_{w_0}^2 + \|u_r - u_{rt}\|_{w_{rev}}^2 + 2\lambda \|d_r\|_{b_p^\alpha(l_p)}^p \\
 &= \int_{\Omega} (u_r(x) - u_{r0}(x))^2 w_0(u_{r0}(x)) d\Omega \\
 &+ \int_{\Omega} (u_r(x) - u_{rt}(x))^2 w_{rev}(u_{r0}(x)) d\Omega \\
 &+ 2\lambda \sum_l \sum_{x_l} \sum_{o_1, o_2} |d_r^{l, o_1, o_2}(x_l)|^p \\
 &= \sum_l \sum_{x_l} \sum_{o_1, o_2} \left\{ w_0^l(x_l) (d_r^{l, o_1, o_2}(x_l) - d_{r0}^{l, o_1, o_2}(x_l))^2 \right. \\
 &+ w_{rev}^l(x_l) (d_r^{l, o_1, o_2}(x_l) - d_{rt}^{l, o_1, o_2}(x_l))^2 \\
 &+ 2\lambda |d_r^{l, o_1, o_2}(x_l)|^p \left. \right\}, \tag{2}
 \end{aligned}$$

where d_r is the corresponding wavelet coefficient sequence of u_r , and $b_p^\alpha(l_p)$ and l_p are the corresponding Besov sequence space of $B_p^\alpha(L_p(\Omega))$ and sequence space of discrete Ω respectively [30], [35]. Here $\|u\|_w$ is the weighted norm of the function u with the weight function w in the L_2 sense, and $\|u_r - u_{rt}\|_{w_{rev}}$ represents the *reverse* diffusion based on the two functions u_{r0} and u_{v0} (discussed below). Also, the weight function of the fidelity term, w_0 , is selected as $1 - w_{rev}$. The reduced reverse coefficient, w_{rev}^l , is obtained by mapping³ w_{rev} to Ω_l and then multiplying it by 2^{-l} . Also, $w_0^l = 1 - w_{rev}^l$, and λ is the smoothing parameter (see subsection VI-C for more details). It is worth noting that $\|u\|_{B_p^\alpha(L_p(\Omega))}$, the Besov semi-norm in the Besov space $B_p^\alpha(L_p(\Omega))$ [32], is equivalent to the norm of the wavelet coefficients in the Besov sequence space $b_p^\alpha(l_p)$ [30], [36]:

$$\|u\|_{B_p^\alpha(L_p(\Omega))} \asymp \left(\sum_l \sum_{x_l} \sum_{o_1, o_2} |d^{l, o_1, o_2}(x_l)|^p \right)^{1/p} = \|d\|_{b_p^\alpha(l_p)}$$

in the sense that there exist two positive numbers c_1 and c_2 with $c_1 \leq c_2$ such that

$$c_1 \|u\|_{B_p^\alpha(L_p(\Omega))} < \|d\|_{b_p^\alpha(l_p)} < c_2 \|u\|_{B_p^\alpha(L_p(\Omega))}$$

In this way, $J(d_r)$ can also be connected to the following functional in the Besov space $B_p^\alpha(L_p(\Omega))$:

$$\|u_r - u_{r0}\|_{w_0}^2 + \|u_r - u_{rt}\|_{w_{rev}}^2 + 2\lambda \|u_r\|_{B_p^\alpha(L_p(\Omega))}^p$$

There are two special cases of interest in model (1): $p = 1$, which corresponds to soft wavelet shrinkage, and $p \rightarrow 0$, which represents hard wavelet shrinkage [37]. The performance of soft wavelet shrinkage is superior to that of hard wavelet shrinkage in the case of noise removal. Roughly speaking, it reduces all the coefficients by an amount depending on the noise variance [31]. This correction is reasonable for denoising applications, where all pixels are affected by noise. However, in the bleed-through problem, where interfering patterns appear mostly on the background regions, true data should not be affected by the shrinkage process. Therefore, hard wavelet shrinkage ($p \rightarrow 0$) is used from here on. Figure 1 schematically shows the difference in behavior of hard and soft wavelet shrinkages when dealing with degraded document images suffering the bleed-through effect.

Before presenting the solution of (1), the role of w_{rev} is discussed. As mentioned above, the functions u_{r0} and u_{rt} are the input recto image and target background respectively. The

³Using Gaussian pyramid reduction operation (*impyramid* function of Matlab).

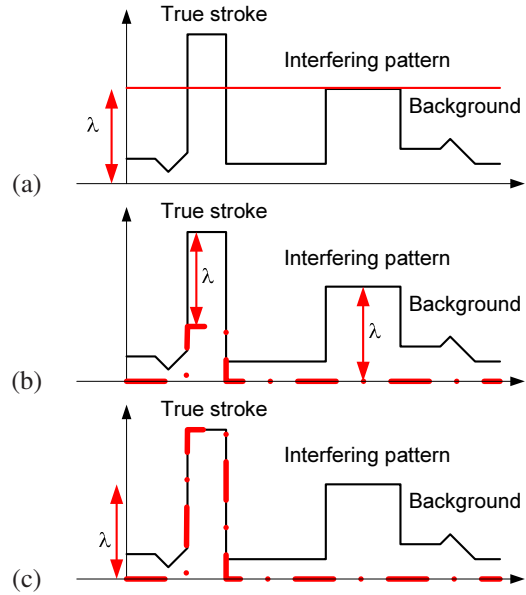


Fig. 1. (a) A typical schematic of the variation of a wavelet coefficient on a line of a document image. The estimated λ is shown by the red line. (b) Application of soft wavelet shrinkage. Although the interfering patterns have been removed, the true data are also affected. (c) Application of hard wavelet shrinkage. True data are preserved.

basic structure of J is very similar to that in the standard smoothing variational approaches: there are two terms, one is the fidelity term to measure the similarity between input and output, and one to compute the smoothness of the output. However, there is an additional term: the *reverse* norm is used to compute the closeness of the two sides of the document (recto and verso sides). In this term, the coefficient w_{rev} , which is a function of both u_{r0} and u_{v0} , measures the extent of bleed-through. For pixels with high values of w_{rev} , the reverse norm is significant and will help to introduce a separation between the text and the interfering patterns. The definition of w_{rev} is as follows:

$$w_{rev}(u_{r0}) = \frac{1}{2} \left(1 + \tanh \left(\frac{u_{v0} - u_{r0} - 2\sigma_{rev}}{\sigma_{rev}} \right) \right) \tag{3}$$

where σ_{rev} determines the degree of reverse diffusion. The rate of that diffusion is determined by the ratio σ_{rev} to the distance between the verso text intensity and its bleed-through pattern on the recto side. w_{rev} can also be considered as a nonlinear transform which produces an effective distance of size σ_{rev} between the main recto text and the interfering patterns. In addition to controlling the nonlinear transformation, σ_{rev} has another role, which is to provide enough room for preserving possible variations in text intensity. This parameter is set roughly to 0.1 for all experiments. It provides enough separation, and the shrinkage process can easily capture the weakened interfering patterns. As the smoothing parameter λ is chosen based on the estimated background (see subsection VI-C), λ implicitly depends on the interfering patterns intensity. Therefore, the smoothing process is not sensitive to σ_{rev} , and the rough value of σ_{rev} is sufficient.

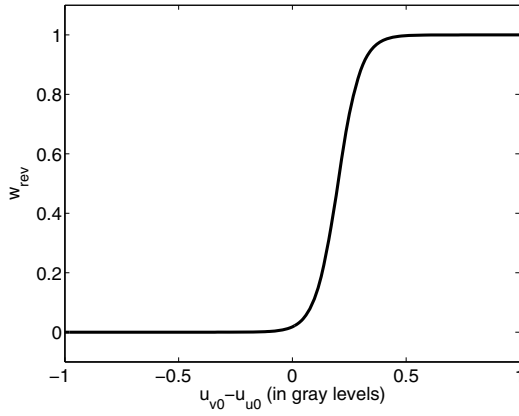


Fig. 2. A typical profile of w_{rev} with respect to $u_{v0} - u_{r0}$.

A plot of w_{rev} with respect to $u_{v0} - u_{r0}$ is shown in Figure 2. The reverse diffusion has its maximum value for the pixels that have black intensities on the verso side. On contrast, it is almost zero for stroke pixels on the recto side. There is a transition region of $2\sigma_{rev}$ width. The tanh function is selected because of its nonlinear and asymmetrical behavior. It provides a saturation effect to describe the nonlinear nature of the bleed-through effect.

The solution of the model (1) can be obtained using wavelet shrinkage. This will give the solution in one step, which will substantially improve the computational cost. Using hard wavelet shrinkage [38], the solution can easily be written as follows:

$$\begin{aligned} \tilde{d}_r^{l,o_1,o_2}(x_l) = & U \left(w_0^l(x_l) d_{r0}^{l,o_1,o_2}(x_l) + w_{rev}^l(x_l) d_{rt}^{l,o_1,o_2}(x_l) - \lambda \right) \times \\ & \left(w_0^l(x_l) d_{r0}^{l,o_1,o_2}(x_l) + w_{rev}^l(x_l) d_{rt}^{l,o_1,o_2}(x_l) \right) \end{aligned} \quad (4)$$

where $U(\cdot)$ is the unit-step function. Using the new wavelet coefficients, the enhanced document image can be constructed by applying the inverse wavelet transform [29]. Here after, we refer to the solution (4) as the result of the wavelet double-sided method.

IV. RESTORATION WITHOUT THE VERSO-SIDE IMAGE

In many cases, the verso side of the document is not available. Even if it is, because of physical and geometrical changes to the document surface, the relation between two sides may be very nonlinear, and therefore two images cannot be registered easily. In these cases, another model based on the estimated background can be used. The document's text can be considered as an overlay on the estimated background layer (as a substitute for the unknown true background). Therefore, the distance between the input image and the estimated background can be used as a measure to identify the text on the image from all other sources, as well as the bleed-through patterns. The variational form of the model becomes the following:

$$\tilde{d}_r = \arg \min_{d_r \in b_p^s(l_p)} J_b(d_r) \quad (5)$$

where

$$\begin{aligned} J_b(d_r) &= \|u_r - u_{r0}\|_{w_{0b}}^2 + \|u_r - u_{rt}\|_{w_{bkgd}}^2 + 2\lambda \|d_r\|_{b_p^s(l_p)}^p \\ &= \int_{\Omega} (u_r(x) - u_{r0}(x))^2 w_{0b}(u_{r0}(x)) d\Omega \\ &\quad + \int_{\Omega} (u_r(x) - u_{rt}(x))^2 w_{bkgd}(u_{r0}(x)) d\Omega \\ &\quad + 2\lambda \sum_l \sum_{x_l} \sum_{o_1,o_2} |d_r^{l,o_1,o_2}(x_l)|^p \\ &= \sum_l \sum_{x_l} \sum_{o_1,o_2} \left\{ w_{0b}^l(x_l) (d_r^{l,o_1,o_2}(x_l) - d_{r0}^{l,o_1,o_2}(x_l))^2 \right. \\ &\quad + w_{bkgd}^l(x_l) (d_r^{l,o_1,o_2}(x_l) - d_{rt}^{l,o_1,o_2}(x_l))^2 \\ &\quad \left. + 2\lambda |d_r^{l,o_1,o_2}(x_l)|^p \right\} \end{aligned} \quad (6)$$

where $w_{0b} = 1 - w_{bkgd}$, and w_{bkgd} is the weight function for the background diffusion:

$$w_{bkgd}(u_{r0}) = \frac{1}{2} \left(1 + \tanh \left(\frac{u_{rb} - u_{r0} + 2\sigma_{bkgd}}{\sigma_{bkgd}} \right) \right)$$

where u_{rb} is the estimated background, σ_{bkgd} is the extent of the gray values that are affected by the background diffusion. Like σ_{rev} , the model is not sensitive to σ_{bkgd} , and this parameter is set to 0.1 in all experiments. The estimated background is computed based on an algorithm in [28] (see Appendix). The w_{bkgd} formula is similar to (3), except for the sign of $2\sigma_{bkgd}$. This is because the estimated background of a pixel is always lower than its intensity. It provides a separation between the recto text and the interfering patterns based on the estimated background information. As in (4), the solution can be expressed as follows:

$$\begin{aligned} \tilde{d}_r^{l,o_1,o_2}(x_l) = & U \left(w_{0b}^l(x_l) d_{r0}^{l,o_1,o_2}(x_l) + w_{bkgd}^l(x_l) d_{rt}^{l,o_1,o_2}(x_l) - \lambda \right) \times \\ & \left(w_{0b}^l(x_l) d_{r0}^{l,o_1,o_2}(x_l) + w_{bkgd}^l(x_l) d_{rt}^{l,o_1,o_2}(x_l) \right) \end{aligned} \quad (7)$$

The inverse wavelet transform can then be used to reconstruct the solution. From here on, we refer to the solution (7) as the result of the single-sided wavelet method. In the experimental section, a comparison between the double-sided wavelet method and the single-sided wavelet method is presented. Obviously, in cases where the verso side of the document is not presented, the double-sided method cannot be applied.

V. INTRODUCTION OF THE FLOW FIELD

In highly degraded document images, the edges of the strokes are very indistinct. In this section, we introduce a global classifier which helps preserving these very weak edges, while at the same time achieving a high degree of smoothing and enhancement. To build a global classifier, we start with the correlation tensor. Consider the image u_r . For each pixel $x = [i, j]$ of u_r , the correlation tensor T_x is defined as follows:

$$T_{xk,l} = \exp \left(-(u_r(x) - u_r(x_{k,l}))^2 / h_f^2 \right), \quad k, l \in N_x$$

where N_x is a neighborhood of x , $x_{k,l} = [i + k, j + l]$ and h_f is the correlation parameter. In all experiments, a 15×15 pixel window is used as N_x . Many global classifiers can be constructed based on T_x .

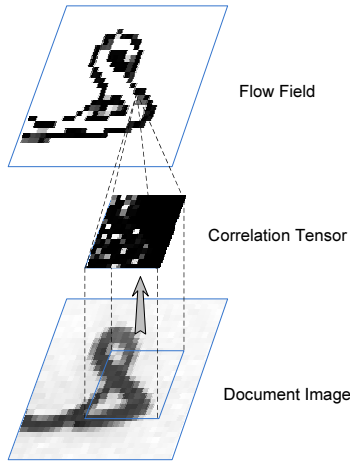


Fig. 3. An illustration of how the flow field is computed. A typical correlation tensor is also shown for a pixel near an edge.

The flow field f is defined as follows:

$$f(x) = \sum_{k,l,[k,l] \neq [0,0]} K_{k,l} T_{xk,l} \quad (8)$$

Where the kernel $K_{k,l} = s/(k^2 + l^2)$ and $\sum_{k,l,[k,l] \neq [0,0]} K_{k,l} = 1$. s is the normalization factor, and is equal to $1/\sum_{k,l,[k,l] \neq [0,0]} 1/(k^2 + l^2)$. $K_{k,l}$ is a spatial kernel to localize the information in the correlation tensor. The kernel is rotation invariant. As a consequence, the flow field f acts as an orientation-invariant global classifier for correlation. To make f robust with respect to the parameter h_f , the flow field is then normalized over the whole image. Here after, the constant value of 0.1 is used for h_f . Figure 3 shows a typical process of computation of the correlation tensor and the flow field value. The pixels of the uniform areas have different values of f compared to the pixels near the edges. For isolated pixels, all elements of the correlation tensor except the central element will be zero, and therefore, the flow field of those pixels will be zero. For pixels on the centerline of the edges, the number of non-zero elements is higher, but still is very small (less than the number of non-zero elements for off-edge pixels). For these pixels, the flow field value is small and near zero, and they appear as dark lines on the flow-field map. As we move far away from the edges, the number of similar pixels in the correlation tensor will increase rapidly. This results in almost white regions for off-edge pixels. Therefore, ∇f can be used in a very similar manner to ∇u_r in the smoothness term, which makes the model global.

Below, a variational model which uses the flow field and the estimated background to enhance the document images with the bleed-through problem is provided. As the nonlinearity of the model will not allow the solution to be found in a single step, an iterative approach must be applied. Also, for incorporating the flow field into the model, a square form for the smoothing term is selected [39]. Under these conditions, the function space of the bounded variations, $BV(\Omega)$ [40], is used. This larger space is also embedded in $L_2(\Omega)$ [31]. The

variational model (6) can be generalized by incorporating the flow field, as follows:

$$\tilde{u}_r = \arg \min_{u_r \in BV(\Omega)} J_{b,f}(u_r) \quad (9)$$

where

$$\begin{aligned} J_{b,f}(u_r) &= \|u_r - u_{r0}\|_{w_{0b}}^2 + \|u_r - u_{rt}\|_{w_{bkgd}}^2 \\ &+ 2\lambda \int_{\Omega} \phi(|\nabla u_r|, |\nabla f|) d\Omega \\ &= \int_{\Omega} (u_r(x) - u_{r0}(x))^2 w_{0b}(u_r(x)) d\Omega \\ &+ \int_{\Omega} (u_r(x) - u_{rt}(x))^2 w_{bkgd}(u_r(x)) d\Omega \\ &+ 2\lambda \int_{\Omega} \phi(|\nabla u_r|, |\nabla f|) d\Omega \end{aligned} \quad (10)$$

where w_{0b} is $1 - w_{bkgd}$ and ϕ is a regulation function in the square form [39]:

$$\phi = \sigma^2 \log(1 + |\nabla u_r|^2/\sigma^2) \frac{1}{1 + |\nabla f|^2/\sigma_f^2}$$

The parameter σ is the minimum-edge gradient that will be preserved. Usually, the value of σ is determined based on the required scale, and we use a morphological approach to obtain it [41], [42]. For the selection of the scale parameter in this approach, we require that the σ value be less than the minimum gradient value between the text and the background. As the variation range of the flow field is fixed between 0 and 1, a fixed value for the parameter σ_f is used: $\sigma_f = 0.1$. The regulation function ϕ is used for computing the degree of smoothness of the target function (based on the gradient of the function). The flow field regulates the behavior of the *regular* diffusion process on the recto side. It is worth noting that the interfering patterns also have some signatures on the flow field, and therefore are partially protected from the regular diffusion. However, at the same time, the *background* diffusion, which is the main degradation removal phenomenon in the model, is not affected by the flow field. As can be seen from the second term in $J_{b,f}$, the background diffusion does not depend on the flow field. Therefore, the model is capable of restoring the degradation while preserving fine structures and strokes.

The solution of the model can be sought in the wavelet domain. Although, because of the nonlinear nature of the model, a series of wavelet shrinkage operations [38] must be applied, or a nonlinear equation must be solved in the wavelet domain [43]. The computational cost of this approach is equivalent to the computational cost of solving the model in spatial domain directly. Therefore, we will solve the problem in the spatial domain using a relaxation approach versus virtual time. As a consequence of introduction of virtual time, the definition of w_{bkgd} is changed; it now depends on the current value of u_r at each iteration instead of u_{r0} . As the virtual time passes, current value of u_r will provide a better estimation, and therefore more accurate value of weight function would be expected. Also, the signature of the interfering patterns on the flow field will diminish over time as better estimation of the text is available on u_r . The governing PDE takes the following form:

$$\begin{aligned} 0 &= (u_r - u_{r0})w_{0b,1} + (u_r - u_{rt})w_{bkgd,1} \\ &- \lambda \operatorname{div} \left(\frac{1}{1 + |\nabla u_r|^2/\sigma^2} \frac{1}{1 + |\nabla f|^2/\sigma_f^2} \nabla u_r \right) \end{aligned} \quad (11)$$

where

$$w_{0b,1} = w_{0b} + \frac{1}{2} \frac{\partial w_{0b}}{\partial u_r} (u_r - u_{r0})$$

$$w_{bkgd,1} = w_{bkgd} + \frac{1}{2} \frac{\partial w_{bkgd}}{\partial u_r} (u_r - u_{rt})$$

The governing equation can be rewritten in a relaxing form, the steady-state solution of which is the solution of the original variational problem:

$$\frac{\partial u_r}{\partial t} = -(u_r - u_{r0})w_{0b,1} - (u_r - u_{rt})w_{bkgd,1}$$

$$+ \lambda \operatorname{div} \left(\frac{\nabla u_r}{1 + \frac{|\nabla u_r|^2}{\sigma^2}} \frac{1}{1 + \frac{|\nabla f|^2}{\sigma_f^2}} \right), \quad (12)$$

$$u_r(t=0) = u_{r0} \quad (13)$$

The result can be obtained by solving governing equations (12) and (13). The governing PDE equations are discretized in the spatial domain on an eight-pixel neighborhood. The discrete form of the governing equation can be written as follows:

$$u_{r,i,j}^{n+1} = u_{r,i,j}^n - (u_{r,i,j}^n - u_{r0,i,j})w_{0b,1}^n \Delta t$$

$$- (u_{r,i,j}^n - u_{rt,i,j})w_{bkgd,1}^n \Delta t \quad (14)$$

$$+ \lambda' \sum_{k,l \in \{-1,0,1\}} d_{k,l} c_{k,l} \Delta_{k,l} u_{r,i,j}^n$$

where $u_{r,i,j}^n$ is the intensity of u_r at pixel $[i, j]$ at iteration n . $d_{k,l} = 1$ if one of k and l is zero, and $d_{k,l} = 1/2$ for other neighboring pixels. The distance to neighboring pixel (k, l) , $\Delta_{k,l} u_{r,i,j}^n$ is equal to $u_{r,i,j}^n - u_{r,i+k,j+l}^n$. And, $c_{k,l}$ is defined as follows:

$$c_{k,l} = \frac{1}{1 + |\Delta_{k,l} u_{r,i,j}^n|^2 / \sigma^2} \frac{1}{1 + |\Delta_{k,l} f_{i,j}^n|^2 / \sigma_f^2}.$$

For stability of numerical calculations, the parameter λ' must be less than $1/6$. It is assumed that the steady state is obtained if the variation in the solution norm is less than 0.001. The flow field is a dynamic feature, i.e. it depends on the current state of the image. In other words, in iterative schemes, the flow-field map is updated on each iteration. However, because of the global nature of the flow field, it is not necessary to update it at each iteration. In our experiments, the flow field is updated at every ten iterations. Here after, we refer to the solution of equations (12) and (13) as the result of the single-sided flow-based method.

The concept of the flow field can also be used in the case of double-sided document images. The variational model in section III can be then modified in the same way as in this section to arrive at a model which includes the flow-field classifier. Because of its similarity in terms of detail to the content of this section, the flow-based model for double-sided document images is not presented here. Also, in other fields of image processing, the flow-based model without background diffusion is capable of enhancing images which suffer from noise and at the same time help preserving weak and fine details.

In the next section, the performance of the models is studied using several real databases and a degradation model.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, performance of various methods will be evaluated in both subjective and objective evaluations. High resolution versions of all figures are available on Internet⁴.

A. Subjective evaluation

As was stated previously, the parameters of the weight functions are constant for all experiments: $\sigma_{rev} = 0.1$, $\sigma_{bkgd} = 0.1$ and $h_f = 0.1$. Also, the smoothing parameter λ is selected based on the estimated background range of variation. The detail of the λ calculations is presented in subsection VI-C. The performance of the models is evaluated using several datasets:

(i) Google Book Search dataset [2], which contains scanned, resized and flattened images of books. The DVD of the dataset contains data for 68 books, while the hard drive contains 1,000 books, making something like 300,000 pages available. Several samples from this dataset, which suffer from degradation, have been used in the experiments. For the registration of inputs from this dataset, we have used the standard spatial transform provided in the Matlab Image Processing Toolbox: cp2tform [44];

(ii) Manually generated degraded document dataset⁵ [23], which contains several pre-registered handwritten double-sided document images;

(iii) Latin and Arabic manuscript dataset [45], which contains a large number of ancient documents including Arabic manuscripts from Tombouctou;

(iv) Two old manuscripts, courtesy of the Juma Al Majid center for Cultural Heritage (Dubai).⁶ Together, these two books contain about 500 pages;

(v) Model-generated degraded datasets [3], [46]. These computer-generated datasets were obtained using a physical model of old document degradation. The degree of degradation has been chosen based on a number of parameters. A brief description of the degradation model [3], [46] is presented in below. These degraded document images are automatically registered.

The degradation model used for the last dataset (v) is capable to add degradation effects such as ink bleed-through and aging [3], [46]. The degradation model based on a physical diffusion produces the bleed-through effect on a set of two input images. The degree of bleed-through effect depends on the parameters of the degradation model. The basic process of the model consists of diffusion of the information from the various sources to the output image. The rate of each diffusion process is selected by specifying the diffusion coefficients “ c ”, of that process. The value of c also has a spatial dependency, in order to account for variations in the document paper. The governing equation of the degradation model can be written as follows:

$$\frac{\partial u}{\partial t} = \sum_i \operatorname{div}_{s_i} (c_i \nabla_{s_i} u) \quad (15)$$

⁴<http://www.synchromedia.ca/web/ets/expres33y9>

⁵<http://www.site.uottawa.ca/~edubois/documents/>

⁶<http://www.almajidcenter.org/English/Pages/default.aspx>

where i sums over all sources, and $\nabla_{s_i} u$ is the gradient of u with respect to source s_i . The time interval used represents the aging of the document. The longer the time interval, the higher the degree of degradation. Also, there is a plate for the noise and random variations in the background and text to represent external distortions and changes. In the case of bleed-through phenomena, there are three sources in the degradation model: the main text, the other (verso) side of document and the background. The output, which is created mainly from the main text, is also affected by the information on the verso side. Because of ink dispersion, dark regions in the main text are more likely to be affected by interference patterns originating on the verso side.

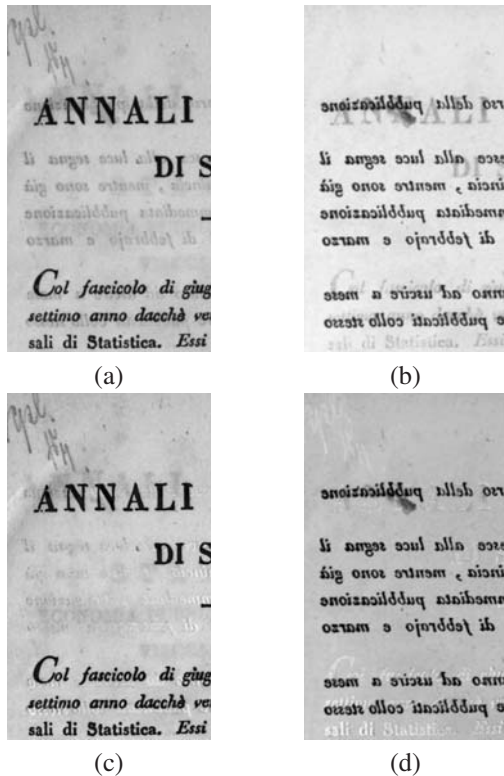


Fig. 4. (a) and (b) Input images used as the recto and verso sides of a degraded double-sided document image. (c) and (d) Restored version of the document using the ICA method.

Below, some examples from the above datasets are presented. The first example is from the Google Book Search Dataset. The double-sided sample is shown in Figures 4(a) and 4(b). The result of applying the double-sided wavelet method (given by equation (4)) is shown in Figure 4(e) and 4(f). Despite the variable intensity of the bleed-through interference pattern in this example, the output background is very uniform, and all the main text has been recovered. It is worth noting that, in our model, the weight function w_{rev} actually presents the binarized result simultaneously in the same step. For the sake of comparison, the result of applying the ICA method to the same sample is also presented in Figure 4(c) and 4(d). The FastICA⁷ implementation is used for the calculations [47], [48]. There are some remaining patterns on the ICA output

⁷<http://www.cis.hut.fi/projects/ica/fastica/>

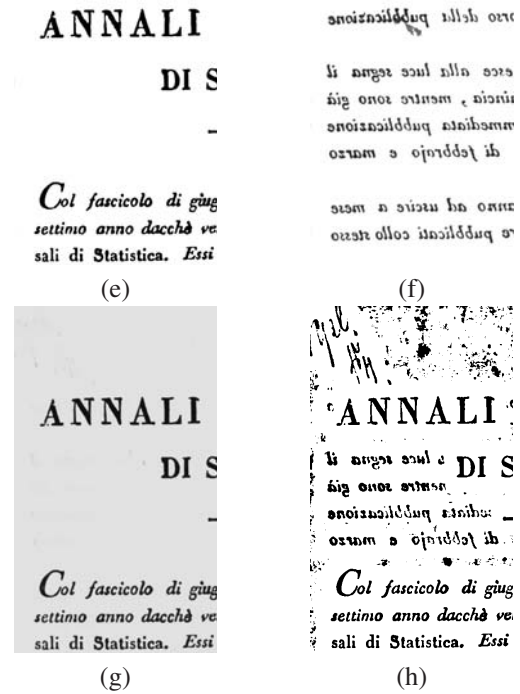


Fig. 4. (e) and (f) Restored version of the document using the double-sided wavelet method ($\delta_{bg} = 0.8$). (g) Restored version of the document using the single-sided flow-based method. (h) Restored version of the document using Niblack's method.

which are due to the nonlinear and nonuniform distribution of the bleed-through effect. This remaining interference forces the user to perform some post-binarization tasks which will introduce a loss of information, especially at fine edges and boundaries. The results of applying the single-sided flow-based method and Niblack's method are also provided in Figures 4(g) and 4(h) respectively. In Niblack's method, which is based on the local mean and local standard deviation, the threshold varies over the image. The threshold at pixel x , $T(x)$, is calculated as $T(x) = m(x) + ks(x)$, where $m(x)$ and $s(x)$ are the mean and standard deviation values respectively in a local neighborhood of x . The size of the neighborhood depends on the stroke width. We used a 35×35 pixel neighborhood for our high-resolution example. The value of k is used to adjust the total print object boundary taken as a part of the given object; it has been found that $k = -0.2$ gave well-separated print objects [49]. But, because of the highly degraded nature of the input, we performed an optimization for k . The best result of the Niblack's method for input image in Figure 4(a) has been obtained using the optimization value for the parameter $k = 1.7$. This optimal value is used in all examples. Because of similar behavior of interfering patterns and the main text, Niblack's method is not able to remove them. The performance of the single-sided flow-based method, given by equation (12), is the highest. A quantitative and detailed comparison of various methods will be provided at the end of this section.

In the second example, we use a real double-sided sample from [23], shown in Figures 5(a) and 5(b). Here as well, the

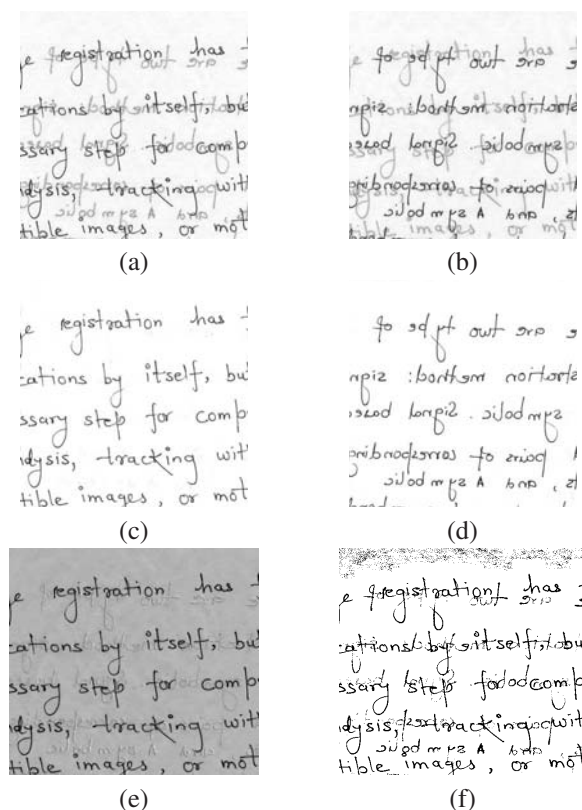


Fig. 5. (a) and (b) Input images used as recto and verso sides of a degraded double-sided document image. (c) and (d) Restored version of the document using the double-sided wavelet method. (e) Restored version of the document using the ICA method. (f) Restored version of the document using Niblack's method.

double-sided wavelet method is applied and the result is shown in Figure 5(c) and 5(d). The output of the ICA method (Figure 5(e)) again suffers from the remaining signatures, which were caused by the nonlinear nature of the ink seepage through the paper.

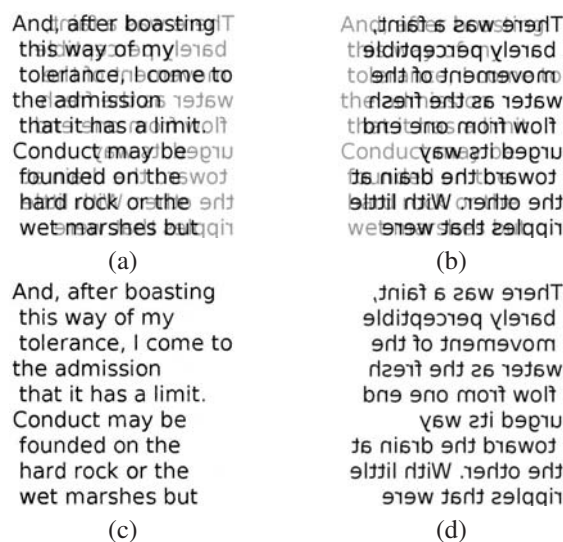


Fig. 6. (a) and (b) Input images used as recto and verso sides of a degraded double-sided document image. (c) and (d) Restored version of document using the double-sided wavelet method.

The third example is a double-sided model generated using

the degradation model [3] and suffers from the bleed-through problem (Figure 6). The output of the double-sided wavelet method is shown in the same figure. The main text has been successfully recovered.

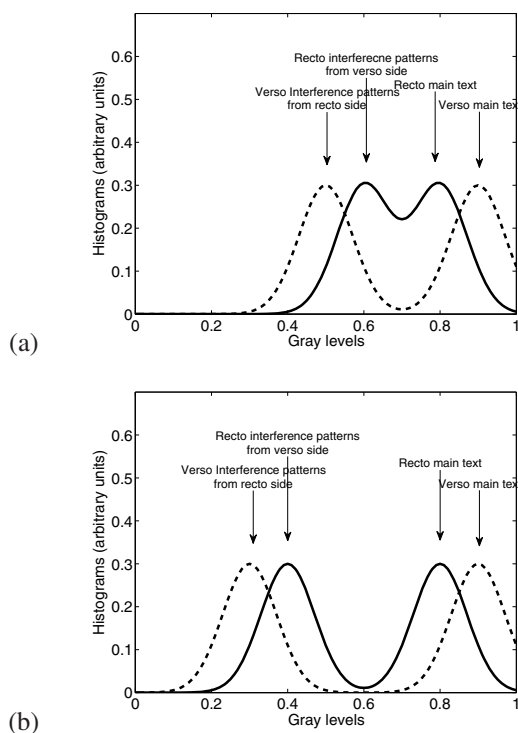


Fig. 7. A schematic histogram for the description of reverse diffusion. (a) Local histograms of two sides of a degraded double-sided document image. The continuous line represents a typical histogram of the recto side, and the verso side histogram is shown by the dashed line. (b) Local histograms following application of model (1) before application of the shrinkage. The interference patterns are pushed toward the background.

The mechanism governing the variational model (1) can be seen in Figure 7. In this figure, typical local histograms for the two sides of a document are presented. The interference patterns on the recto side may be as dark as the main strokes on that side. However, because of reverse diffusion from the main text on the verso side, this interference has been pushed toward the target background and removed.

For very degraded documents and those containing fine features and weak edges, such as the document image in Figure 8(a), the single-sided flow-based method, given by equation (12), can be applied. The output is shown in Figure 8(b). The very fine structures are preserved, while the background is completely uniform and all interference and noise has been removed. For the sake of comparison, the output of application of the variational model (9) is also presented in the same figure, which shows some disconnections and loss of information. Selection of the best model for removing the interference depends on the application, the type of strokes in the manuscript, and also the computational cost. The computational cost of the first two models (1) and (9) on a desktop PC for a 512×512 gray-scale double-sided document image is less than two seconds. For the flow-based model (12), the CPU time for every 10 iterations is about one second,

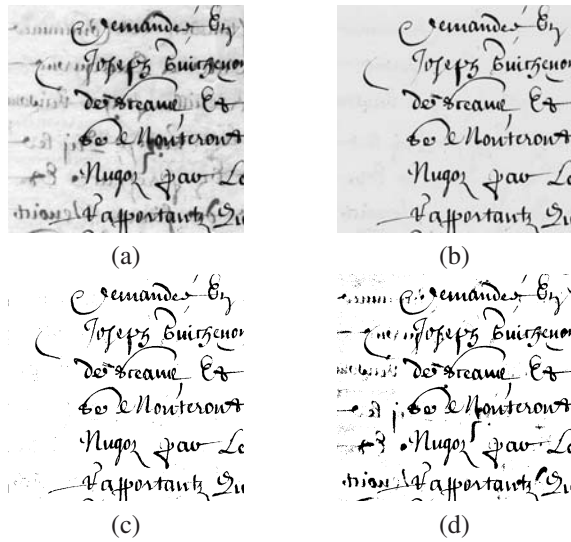


Fig. 8. (a) Input degraded single-sided document image. (b) Restored version of the document using the single-sided flow-based method. (c) Restored version of the document using the single-sided wavelet method. (d) Output of Niblack's method.

and usually the computations converge within one hundred iterations. If there are no fine features or structures in the input images, there is no need to use the flow-based method and the solution can be obtained very quickly. A typical situation before and after applying the single-sided flow-based model (9) or (12) to a document image is shown in Figure 9.

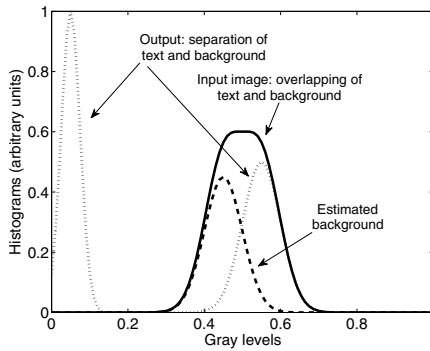


Fig. 9. Schematic histogram for the description of background diffusion. Based on the estimated background information, the main text and the background have been separated. The continuous line and dotted lines represent respectively the histogram before and after applying the background diffusion.

Finally, Figure 10 is an example of bleed-through restoration while preserving the paper's fiber texture. Also, a smaller δ_{bg} value is used along with the double-sided wavelet method ($\delta_{bg} = 0.5$, see Appendix for the definition of δ_{bg}) to preserve the handwriting. Figure 10(a) shows the output of the double-sided wavelet method. By combining this output and the result of a global thresholding, a mask for the interfering patterns is obtained (Figure 10(b)). Using this mask, any inpainting method can be applied to restore the document image while preserving its texture. In Figure 10(c), a possible restoration using the estimated background as the inpainting data is shown.

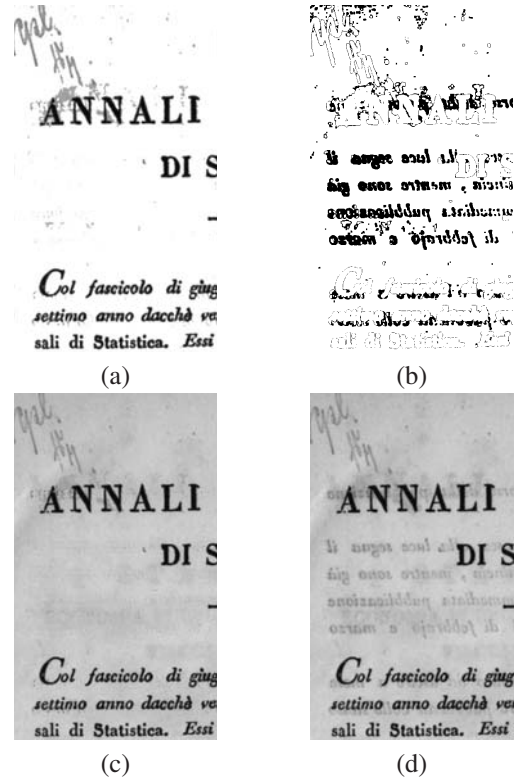


Fig. 10. (a) Output of the wavelet double-sided method for the inputs in Figures 4(a) and 4(b). (b) The inpainting mask obtained by combining (a) and the output of global thresholding. (c) Restored version of the document preserving the paper fiber texture. (d) The input recto side is presented for the sake of comparison.

B. Objective Evaluation

For quantitative evaluation, the goal-directed rules [49], [50] are used, in which an image understanding module (for example, an OCR system) is used which is based on the results of the target low-level image processing models. The objective evaluation is performed by comparing the recognition rate of a commercial OCR system⁸ on the outputs of different methods. A series of degraded document images generated by the degradation model in [3] are enhanced by applying different methods. The images and the recognized texts are available on Internet⁹. Five methods are compared: The double-sided wavelet method and the ICA method, which use both sides of the input image, and the single-sided wavelet method, the single-sided flow-based method and Niblack's method which use only the recto side of the input. Three samples of the series are presented in Figure 11. The impact of the each model in terms of the PSNR, the OCR recognition rate, and percentage of incorrect introduced characters are presented in Tables I, II, and III respectively. In terms of the PSNR, the single-sided flow-based method provides a robust improvement over whole the degradation range. However, for very less degraded cases, the ICA method has better performance. In terms of the recognition rate, the situation is more clear.

⁸FineReader version 9.0

⁹<http://www.synchromedia.ca/web/ets/expres33y9>



Fig. 11. Three samples of a series of degraded document images. (a) Input degraded double-sided document images (only the recto side is shown). The results of various methods: (b) Double-sided wavelet model; (c) ICA method; (d) Single-sided wavelet model; (e) Single-sided Flow-based model; (f) Niblack's method. The three columns correspond to the fourth, fifth and seventh columns in Tables I, II, and III respectively.

The single-sided flow-based method again outperforms all the methods. After the single-sided flow-based method, our two other methods, the double-sided wavelet and the single-sided wavelet methods, have acceptable performance taking into account their lower computational costs. In the last table, Table III, the percentage of incorrect introduced characters are provided. This percentage provides a measure of the extra characters that are presented in the output of OCR system because of reminding artifacts on the document image. Therefore, this measure is not complement of the recognition rate percentage. The performances in this measure are almost the same as previous two measure. We can conclude that the single-sided flow-based method is the best one in all three measures because of providing uniform background and the continuous strokes, while its computational cost is also higher than the others. The double-sided wavelet method outperforms the single-sided wavelet method in two of three measures. However, in cases where the verso side of the document is not presented, the single-sided wavelet method is the best choice. The ICA method, which is point-wise, failed especially in the highly degraded cases, where the boundaries of the interference patterns are highly deformed.

Method	Degradation is decreasing from left to right							
Input	8.71	9.81	11.37	13.41	15.85	18.63	21.64	24.84
Double-sided Wavelet	8.01	8.95	11.65	15.20	17.94	18.27	19.51	19.83
ICA	9.89	8.68	11.22	11.51	19.96	23.61	26.95	30.08
Single-sided Wavelet	9.83	10.20	12.75	14.60	16.20	15.70	17.58	17.84
Single-sided Flow-based	8.99	12.88	16.27	18.62	20.76	21.95	22.78	23.33
Niblack's	6.71	8.11	9.22	10.90	13.41	14.45	14.95	15.30

TABLE I

PSNR MEASURE OF VARIOUS METHODS FOR A SERIES OF EIGHT DEGRADED EXAMPLES. THE DEGRADATION INCREASES FROM THE RIGHT TO THE LEFT. THREE SAMPLES OF THE SERIES HAVE BEEN SHOWN IN FIGURE 11.

Method	Degradation is decreasing from left to right							
Input	2.55	1.82	6.20	18.25	100.0	100.00	100.0	100.0
Double-sided Wavelet	0.73	0.73	2.19	80.29	100.0	100.0	100.0	100.0
ICA	0.00	2.92	0.00	0.00	100.0	100.0	100.0	100.0
Single-sided Wavelet	0.00	0.00	57.30	97.81	99.64	100.0	100.0	100.0
Single-sided Flow-based	2.55	31.02	97.08	99.64	100.0	100.0	100.0	100.0
Niblack's	4.38	2.19	5.11	23.36	95.62	100.0	99.64	99.64

TABLE II

RECOGNITION RATES OF VARIOUS METHODS FOR A SERIES OF DEGRADED EXAMPLES. THE DETAILS ARE THE SAME AS IN TABLE I.

C. Model selection

The most important parameter in the Wavelet methods is λ . This parameter determines the level of hard wavelet threshold.

Method	Degradation is decreasing from left to right							
Input	48.18	31.75	93.07	109.49	0.36	0.36	0.00	0.00
Double-sided Wavelet	106.20	105.11	25.18	32.48	0.00	0.00	0.00	0.00
ICA	10.95	60.95	4.38	NaN	0.00	0.00	0.00	0.00
Single-sided Wavelet	144.89	NaN	67.15	6.20	0.36	0.00	0.00	0.00
Single-sided Flow-based	22.26	78.10	11.68	0.36	0.00	0.36	0.00	0.00
Niblack's	69.34	62.04	60.22	121.53	17.88	6.20	11.68	5.11

TABLE III

PERCENTAGE OF WRONG INTRODUCED CHARACTERS OF VARIOUS METHODS FOR A SERIES OF DEGRADED EXAMPLES. THE DETAILS ARE THE SAME AS IN TABLE I.

In order to estimate this parameter we use Algorithm 1. In this algorithm, a Gaussian estimation of the estimated background histogram is calculated, and based on it the value of λ is selected. In order to generate the sample data of the histogram, the RANDP¹⁰ function is used. Also, the Gaussian fit is obtained using a EM algorithm, GaussianMixture¹¹ function from the CLUSTER¹² library. Using this value of λ , the hard-shrinkage step removes the most of the background variation.

Algorithm 1: Estimation of λ

- 1 Calculate the estimated background u_{rb} ;
- 2 Calculate the histogram of u_{rb} ;
- 3 Use this histogram as a PDF and fit a Gaussian estimation with three clusters to it;
- 4 Set $\lambda = 1 - p_{min}$, where p_{min} is the minimum peak in the three clusters;

VII. CONCLUSION AND FUTURE PROSPECTS

For the enhancement of degraded double-sided document images, a new variational model is introduced. The model has an extra term for reverse diffusion between the two sides of a document. By transferring the model to the wavelet domain and using the hard wavelet shrinkage, the solution of the model is obtained in a single step. The transformation to the wavelet domain is performed using the dual-tree complex wavelet transform which is a shift-invariant discrete wavelet transform. For the cases where the verso side of the document is not available, a modified variational model is introduced which removes the interference patterns using the estimated background information. Owing to the fast wavelet shrinkage operator, both these variational models are very quick.

For document images containing very fine features and structures, the second model is modified to include a global classifier, the flow field. The flow field classifies the information on the image based on its relation to the edges and boundaries. The concept of the flow field helps the diffusion models to easily preserve fine details. The model

¹⁰<http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=8891>

¹¹<http://web.ics.purdue.edu/~wong17/gaussmix/gaussmix.html>

¹²<http://cobweb.ecn.purdue.edu/~bouman/software/cluster/>

is also applicable to many other image processing fields. The performance of the models is evaluated against other methods, such as the ICA method, in both subjective and objective ways using several databases of different types of document script and degradation. For objective evaluation, PSNR and the recognition rate of a commercial OCR system is used as the goal. The single-sided flow-based method outperforms the other methods in all cases.

Analysis of the contents of ancient documents, which usually contain various, and probably hidden, writings, is an advancing field in DIAR. One of the most powerful approaches for obtaining images of historical documents is multispectral acquisition [51]. This approach, in addition to three components of visible color scan, provides several other components in the non-visible spectral range, such as IR or UV. Processing the multispectral scans of ancient documents will provide very deep analysis and accurate restoration of the entire contents of these documents. In the future, we will use our variational framework, in conjunction with statistical methods such as ICA and MRF, to provide a powerful analysis and restoration tool for processing multi-spectral historical and ancient documents.

APPENDIX

NUMERICAL APPROACHES TO VARIATIONAL PROBLEMS

A typical example of a variational problem is Total Variation (TV) denoising, which is one of the successful noise removal and image enhancement techniques [40]. It is a pioneering one in the rapidly advancing field of variational approaches to image processing. The basic idea of the TV model can be expressed as follows:

$$\tilde{u} = \arg \min_{u \in BV(\Omega)} J(u) = \int_{\Omega} (u - u_0)^2 dx + \lambda \int_{\Omega} |\nabla u| dx \quad (16)$$

where $BV(\Omega)$ is the Bounded Variant function space on Ω , u_0 is the input noisy image, λ is a parameter which balances the fidelity of the solution \tilde{u} to the source image u_0 and the smoothness of the solution. For denoising applications, a λ value near the estimated noise variance results in a good performance of the model [40].

The big challenge in variational problems is obtaining the numerical solution. The *Relaxation* technique is one of the direct numerical methods for solving variational problems, and it is also easy to implement. In this method, the corresponding Partial Differential Equations (PDEs) of the variational problem are transformed into a time-stepping scheme. The steady-state solution of the new PDE-based problem is actually the solution of the main variational problem. In this technique, the stability of the numerical implementation must be ensured. Usually, all the spatial derivatives are replaced with appropriate finite difference approximations. Then the resulting system of ordinary differential equations is solved along the virtual time variable [52]. This technique, despite its simplicity, usually requires more computational resources and requires more CPU time than the other techniques.

The direct solution to variational problems can be achieved using nonlinear solvers. There are many direct nonlinear

minimizers to reduce the cost function of these problems [53]. Usually, there is a discretization stage prior to the application of the minimizer. However, the problem can also be stated from the beginning in a proper discrete space [54], [55], which avoids discretization errors. The performance of the direct minimizers can be enhanced by adapting them to the specific types of images by the incorporation of Neural Networks (NN) [56], [57].

One of the most successful numerical methods for solving variational problems is the Finite Element Method (FEM) [58]. Although it is one of the oldest approaches, its generality makes it applicable to all kinds of problems and its performance is superior in terms of accuracy and convergence rate. In this approach, the *weak* form of the problem is usually used. Then, using a set of basis functions defined on the elements which partitions the domain of the problem, the problem is transformed into a system of linear (or nonlinear) algebraic equations. The primary advantage of the FEM is that, because it uses a measure of error, the convergence to the solution is very fast (a few nonlinear steps). Also, partitioning the domain with elements of variable size, enables the FEM to be applied to any complex and irregular domain. The third feature of the FEM is that it can be applied to any form of cost function. The main barrier to the application of the FEM to general problems is finding a way for automatic assembly of the nonlinear problem (or its linearized form). There are several projects, such as FEniCS [59], which provide a high-level FEM language interface for solving general variational problems.

Finally, transforming variational problems from the spatial domain to the wavelet domain may result in an analytical solution in many cases [32], [36]. These solutions, which are termed "wavelet shrinkage" solutions, can be applied very quickly [60], [61]. In this successful approach, the variational problem is restated in the wavelet domain. One of the most powerful discrete wavelet transforms is the Dual-Tree Complex Wavelet Transform (DTCWT) [29], [62]. The DTCWT has recently been the object of a great deal of interest in image processing because of its shift invariance and reduced sensitivity to direction [63], [64]. In this transform, there are six directions (according to two directional superscripts) in addition to the level superscript, and the coefficients are complex numbers.

ESTIMATED BACKGROUND

To compute the estimated background, we use the iterative method introduced in [28]. In the first step, an average background is computed based on an *a priori* threshold value, δ_{bg} . All intensities below δ_{bg} are assumed to be probable background data, and are considered in the first averaging process. For each pixel, the value of the computed average background is used as the new threshold of the background. Then, in an iterative manner, a new background value is computed, but now with a smaller sampling window. The adaptive and iterative nature of the algorithm makes it robust with respect to the parameter δ_{bg} . Therefore, even this threshold value can be set to 1. More accurate values for δ_{bg} can be

proposed considering threshold values of Otsu and Niblack methods, among others, as guidelines, or using a bootstrapping procedure. Throughout the paper, a value of 0.8 is used for δ_{bg} (except where stated otherwise). A factor of the stroke width is used as the smaller window size. The algorithm is shown as Algorithm 1. The function $BKGDEST(u, \delta, s)$ represents the averaging process on a sampling window of size s and using δ as the threshold value. The parameter ϵ is a small value used for the sake of the stability of the algorithm.

Algorithm 2: Computation of the estimated background of an image u

```

1  $n = 0$ ; Compute the average background:
    $u_{b,0} = BKGDEST(u, \delta_{bg}, \text{size}_{\text{large}})$ ;
2 repeat
3    $n = n + 1$ , compute new background:
    $u_{b,n} = BKGDEST(u, u_{b,n-1} - \epsilon, \text{size}_{\text{small}})$ ;
4 until steady state is obtained ;

```

An example of the estimated background is shown in Figure 12. The input image is the recto side of the input images in Figure 4. It is worth noting that, in computing the estimated background, only the information on the recto side has been used.

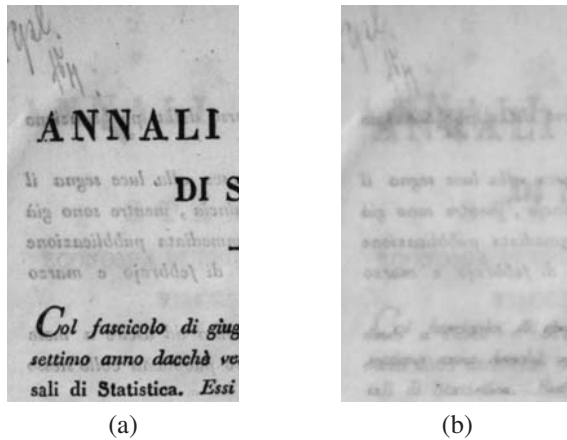


Fig. 12. An example of the estimated background. (a) Input image (the same as in Figure 4(a)). (b) Computed estimated background. $\delta_{bg} = 0.8$ has been used in this example.

ACKNOWLEDGMENTS

The authors would like to thank the NSERC of Canada and FQRNT for their financial support. They would also like to thank Mr. Nicola Nobile of CENPARMI at Concordia University for providing the OCR results.

REFERENCES

- [1] T. Kanungo, "Document degradation models and a methodology for degradation model validation," Ph.D. dissertation, University of Washington, 1996.
- [2] Google, *Book Search Dataset*, version v ed., 2007.
- [3] M. Cheriet and R. Farrahi Moghaddam, "Degradation modeling and enhancement of low quality documents," in *WOSPA'2008*, Sharjah, UAE, Invited paper, 2008.
- [4] S.-W. Huang, D.-L. Way, and Z.-C. Shih, "Physical-based model of ink diffusion in chinese ink paintings," *Journal of WSCG*, vol. 10, no. 3, pp. 520–527, 2003.

- [5] S. Mei-jun, S. Ji-zhou, and Y. Bin, "Physical modeling of "xuan" paper in the simulation of chinese ink-wash drawing," in *CGIV'05*, S. Ji-zhou, Ed., 2005, pp. 317–322.
- [6] W. Xiujin, J. Jingshan, and S. Jizhou, "Graphical simulator for chinese ink-wash drawing," *Transactions of Tianjin University*, vol. 8, no. 1, pp. 1–7, 2002.
- [7] S. Yongxin, S. Jizhou, and Z. Haijiang, "Graphical simulation algorithm for chinese ink wash drawing by particle system (chinese)," *Journal of Computer-Aided Design & Computer Graphics*, vol. 15, no. 6, pp. 667–672, 2003.
- [8] Q. Zhang, Y. Sato, J.-y. Takahashi, K. Muraoka, and N. Chiba, "Simple cellular automaton-based simulation of ink behaviour and its application to suibokuga-like 3d rendering of trees," *The Journal of Visualization and Computer Animation*, vol. 10, no. 1, pp. 27–37, 1999.
- [9] L. Chen, J. Zhu, M. Young, and R. Susfalk, "Modeling polyacrylamide transport in water delivery canals," in *ASA-CSSA-SSSA International Annual Meetings*, Indianapolis, IN, Nov. 12–16 2006, pp. 294–6.
- [10] K. Roth, "Scaling of water flow through porous media and soils," *European Journal of Soil Science*, vol. 59, no. 1, pp. 125–130, 2008.
- [11] H. H. Vaziri, Y. Xiao, R. Islam, and A. Nouri, "Numerical modeling of seepage-induced sand production in oil and gas reservoirs," *Journal of Petroleum Science and Engineering*, vol. 36, no. 1-2, pp. 71–86, Oct. 2002.
- [12] G. Leedham, S. Varma, A. Patankar, and V. Govindaraju, "Separating text and background in degraded document images - a comparison of global thresholding techniques for multi-stage thresholding," in *IWFHR'02*, 6–8 Aug. 2002, pp. 244–249.
- [13] K. Franke and M. Köppen, "A computer-based system to support forensic studies on handwritten documents," *IJDAR*, vol. 3, no. 4, pp. 218–231, May 2001.
- [14] K. Franke and S. Rose, "Ink-deposition analysis using temporal (online) data," in *IWFHR'06*, La Baule, France, 2006.
- [15] H. Nishida and T. Suzuki, "Correcting show-through effects on document images by multiscale analysis," in *ICPR'02*, T. Suzuki, Ed., vol. 3, 2002, pp. 65–68 vol.3.
- [16] H.-S. Don, "A noise attribute thresholding method for document image binarization," *IJDAR*, vol. 4, no. 2, pp. 131–138, Dec. 2001.
- [17] J. M. M. da Silva, R. D. Lins, F. M. J. Martins, and R. Wachenchauser, "A new and efficient algorithm to binarize document images removing back-to-front interference," *Journal of Universal Computer Science*, vol. 14, no. 2, pp. 299–313, Jan. 2008.
- [18] G. Sharma, "Show-through cancellation in scans of duplex printed documents," *Image Processing, IEEE Transactions on*, vol. 10, no. 5, pp. 736–754, 2001.
- [19] K. T. Knox, "Show-through correction for two-sided documents," United States Patent 5646744, July 1997.
- [20] C. L. Tan, R. Cao, and P. Shen, "Restoration of archival documents using a wavelet technique," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 10, pp. 1399–1404, 2002.
- [21] C. L. Tan, R. Cao, P. Shen, Q. Wang, J. Chee, and J. Chang, "Removal of interfering strokes in double-sided document images," in *WACV'00*, Palm Springs, CA, USA, 2000, pp. 16–21.
- [22] E. Dubois and A. Pathak, "Reduction of bleed-through in scanned manuscript documents," in *PICS'01*, Montreal, Canada, April 2001, pp. 177–180.
- [23] E. Dubois and P. Dano, "Joint compression and restoration of documents with bleed-through," in *Proc. IS&T Archiving 2005*, Washington DC, USA, April 2005, pp. 170–174.
- [24] I. Gerace, F. Cracco, and A. Tonazzini, *Lecture Notes in Computer Science: Independent Component Analysis and Blind Signal Separation*. Springer Berlin / Heidelberg, 2004, vol. 3195/2004, ch. An Extended Maximum Likelihood Approach for the Robust Blind Separation of Autocorrelated Images from Noisy Mixtures, pp. 954–961.
- [25] E. Salerno, A. Tonazzini, and L. Bedini, "Digital image analysis to enhance underwritten text in the archimedes palimpsest," *IJDAR*, vol. 9, no. 2, pp. 79–87, Apr. 2007.
- [26] A. Tonazzini, E. Salerno, and L. Bedini, "Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique," *IJDAR*, vol. 10, no. 1, pp. 17–25, Jun. 2007.
- [27] B. Ophir and D. Malah, "Show-through cancellation in scanned images using blind source separation techniques," in *ICIP'07*, D. Malah, Ed., vol. 3, 2007, pp. III–233–III–236.
- [28] R. Farrahi Moghaddam and M. Cheriet, "EFDM: Restoration of single-sided low-quality document images," in *ICFHR 2008*, Montréal, Québec, Canada, August 19–21 2008, pp. 204–209.

- [29] I. Selesnick, R. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform," *Signal Processing Magazine, IEEE*, vol. 22, no. 6, pp. 123–151, 2005.
- [30] K. Berkner, M. J. Gormish, and E. L. Schwartz, "Multiscale sharpening and smoothing in besov spaces with applications to image enhancement," *Applied and Computational Harmonic Analysis*, vol. 11, no. 1, pp. 2–31, Jul. 2001.
- [31] A. Chambolle, R. De Vore, N.-Y. Lee, and B. Lucier, "Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage," *Image Processing, IEEE Transactions on*, vol. 7, no. 3, pp. 319–335, 1998.
- [32] A. Chambolle and B. Lucier, "Interpreting translation-invariant wavelet shrinkage as a new image smoothing scale space," *Image Processing, IEEE Transactions on*, vol. 10, no. 7, pp. 993–1000, 2001.
- [33] H. Triebel, *Theory of Function Spaces II*, ser. Monographs in Mathematics. Basel: Birkhäuser Verlag, 1992, vol. 84.
- [34] Y. Meyer, *Wavelets and Operators*, ser. Cambridge Studies in Advanced Mathematics. Cambridge, U.K: Cambridge Univ. Press, 1995, vol. 37.
- [35] W. Rudin, *Real and Complex Analysis*, ser. International Series in Pure and Applied Mathematics. McGraw-Hill, 1986.
- [36] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, September 1994.
- [37] D. Lorenz, "Variational denoising in besov spaces and interpolation of hard and soft wavelet shrinkage," Univ. Bremen, Germany, Tech. Rep. Zentrum für Technomathematik, 2004. [Online]. Available: <http://www.math.uni-bremen.de/zetem/DFG-Schwerpunkt/preprints/orig/lorenz2003softhard.pdf>
- [38] P. Mrázek, J. Weickert, and G. Steidl, *Lecture Notes in Computer Science: Scale Space Methods in Computer Vision*. Springer Berlin / Heidelberg, 2003, ch. Correspondences between Wavelet Shrinkage and Nonlinear Diffusion, pp. 101–116.
- [39] J. Monteil and A. Beghdadi, "A new interpretation and improvement of the nonlinear anisotropic diffusion for image enhancement," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 940–946, Sept. 1999.
- [40] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [41] F. Voci, S. Eiho, N. Sugimoto, and H. Sekibuchi, "Estimating the gradient in the Perona-Malik equation," *Signal Processing Magazine, IEEE*, vol. 21, no. 3, pp. 39–65, 2004.
- [42] S. Mukhopadhyay and B. Chanda, "Multiscale morphological segmentation of gray-scale images," *Image Processing, IEEE Transactions on*, vol. 12, no. 5, pp. 533–549, 2003.
- [43] J. A. Dobrosotskaya and A. L. Bertozzi, "A wavelet-laplace variational technique for image deconvolution and inpainting," *Image Processing, IEEE Transactions on*, vol. 17, no. 5, pp. 657–663, 2008.
- [44] *MATLAB Version 7.5.0*, The Mathworks Inc., Natick, MA.
- [45] R. Deriche and O. Faugeras, "Les EDP en traitement des images et vision par ordinateur," INRIA, Tech. Rep. 2697, 1996.
- [46] R. Farrahi Moghaddam and M. Cheriet, "Low quality document image modeling and enhancement," *IJDAR*, vol. 11, no. 4, pp. 183–201, 2009.
- [47] E. Oja and Z. Yuan, "The FastICA algorithm revisited: Convergence analysis," *Neural Networks, IEEE Transactions on*, vol. 17, no. 6, pp. 1370–1381, 2006.
- [48] A. Cichocki, A. S. K. T. Tanaka, A. H. Phan, and R. Zdunek, "ICALAB – MATLAB toolbox ver. 3 for signal processing," 2007.
- [49] O. Trier and A. Jain, "Goal-directed evaluation of binarization methods," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, no. 12, pp. 1191–1201, 1995.
- [50] X. Ye, M. Cheriet, and C. Suen, "Stroke-model-based character extraction from gray-level document images," *Image Processing, IEEE Transactions on*, vol. 10, no. 8, pp. 1152–1161, 2001.
- [51] A. Tonazzini, E. Salerno, M. Mochi, and L. Bedini, "Blind source separation techniques for detecting hidden texts and textures in document images," *Image Analysis and Recognition*, pp. 241–248, 2004.
- [52] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, July 1990.
- [53] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1, pp. 89–97, Jan. 2004.
- [54] J. Bioucas-Dias, M. Figueiredo, and J. Oliveira, "Total variation-based image deconvolution: a majorization-minimization approach," in *ICASSP'06*, M. Figueiredo, Ed., vol. 2, 2006, pp. II-II.
- [55] Y. Boykov, V. Kolmogorov, D. Cremers, and A. Delong, "An integral solution to surface evolution PDEs via geo-cuts," in *ECCV'06*, 2006, pp. 409–422.
- [56] X. Hu and J. Wang, "Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network," *Neural Networks, IEEE Transactions on*, vol. 17, no. 6, pp. 1487–1499, 2006.
- [57] S. Effati and A. Nazemi, "Neural network models and its application for solving linear and quadratic programming problems," *Applied Mathematics and Computation*, vol. 172, no. 1, pp. 305–331, Jan. 2006.
- [58] O. C. Zienkiewicz, R. L. Taylor, and J. Z. Zhu, *The Finite Element Method: Its Basis and Fundamentals*, 6th ed. Amsterdam: Butterworth-Heinemann, 2005.
- [59] R. C. Kirby and A. Logg, "A compiler for variational forms," *ACM Trans. Math. Softw.*, vol. 32, no. 3, pp. 417–444, 2006.
- [60] F. Malgouyres, "Combining total variation and wavelet packet approaches for image deblurring," in *Variational and Level Set Methods in Computer Vision, 2001. Proceedings. IEEE Workshop on*, 2001, pp. 57–64.
- [61] —, "Minimizing the total variation under a general convex constraint for image restoration," *Image Processing, IEEE Transactions on*, vol. 11, no. 12, pp. 1450–1456, 2002.
- [62] N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.
- [63] I. Selesnick, "A new complex-directional wavelet transform and its application to image denoising," in *ICIP'02*, vol. 3, 2002, pp. 573–576 vol.3.
- [64] B. Dumitrescu, I. Bayram, and I. W. Selesnick, "Optimization of symmetric self-Hilbertian filters for the dual-tree complex wavelet transform," *Signal Processing Letters, IEEE*, vol. 15, pp. 146–149, 2008.

Reza Farrahi Moghaddam received his B.Sc. degree in Electrical Engineering and his Ph.D. degree in Physics from the Shahid Bahonar University of Kerman, Iran, in 1995 and 2003, respectively. Since 2003, he has been with Vali-Asr University of Rafsanjan, Iran. Since 2007, he has been a Postdoctoral Fellow with the Synchronmedia Laboratory for Multimedia Communication in Telepresence, École de Technologie Supérieure (University of Quebec), Montréal, QC, Canada. His research interests include mathematical modeling for image processing and



document image processing.



Mohamed Cheriet was born in Algiers (Algeria) in 1960. He received his B.Eng. from USTHB University (Algiers) in 1984 and his M.Sc. and Ph.D. degrees in Computer Science from the University of Pierre et Marie Curie (Paris VI) in 1985 and 1988 respectively. Since 1992, he has been a professor in the Automation Engineering department at the École de Technologie Supérieure (University of Quebec), Montréal, and was appointed full professor there in 1998. He co-founded the Laboratory for Imagery, Vision and Artificial Intelligence (LIVIA) at the University of Quebec, and was its director from 2000 to 2006. He also founded the SYNCHROMEDIA Consortium (Multimedia Communication in Telepresence) there, and has been its director since 1998. His interests include document image analysis, OCR, mathematical models for image processing, pattern classification models and learning algorithms, as well as perception in computer vision. Dr. Cheriet has published more than 200 technical papers in the field, and has served as chair or co-chair of the following international conferences: VI'1998, VI'2000, IWFHR'2002, and ICFHR'2008. He currently serves on the editorial board and is associate editor of several international journals: IJPRAI, IJDAR, and Pattern Recognition. He co-authored a book entitled, "Character Recognition Systems: A guide for Students and Practitioners," John Wiley and Sons, Spring 2007. Dr. Cheriet is a senior member of the IEEE and the chapter chair of IEEE Montreal Computational Intelligent Systems (CIS).