BLIND DOCUMENT IMAGE ENHANCEMENT BASED ON DIFFUSION PROCESS

*M. R. YAGOUBI, A. SERIR and A. BEGHDADI**

USTHB - Faculté d'Electronique et Informatique, LTIR, Alger, Algrie
*L2TI, Université Paris 13, Sorbonne Paris Cité, France
myagoubi@usthb.dz, aserir@usthb.dz, beghdadi@univ-paris13.fr

ABSTRACT

Document image enhancement is a well investigated field of research. Many approaches based on nonlinear anisotropic diffusion have been proved efficient for removing some common artifacts such as bleed-through. However, the proposed methods are somewhat ineffective in removing bleed-through artifacts while strengthening original side strokes. The purpose of this work is to propose a novel nonlinear anisotropic diffusion approach which enhances foreground edges coherence while smoothly removes background artifacts. The proposed method is based on a structure tensor using a new parameter which permits to better distinguish between foreground and artifacts characters. The diffusion scheme is expressed through a new set of equations using the tensor eigenvalues. This system allows to perform a second discrimination between foreground/background strokes. The performance of the proposed enhancement scheme is evaluated by means of objective measures and perceptual judgment of the obtained results on real images.

***Index Terms*—** Document image enhancement, Document image restoration, Diffusion process, PDE-based approach, Anisotropic diffusion

1. INTRODUCTION

Degraded document image enhancement represents an important step in document analysis process. Indeed, the documents may suffer from many artifacts and degradations such as noise, bleed-through and ink stains. The objective is to remove these artifacts and degradations without affecting the useful information on the original document. Many methods have been proposed in the literature [1, 2]. Some of the classical approaches are based on a binarization and enhancement of the degraded images. A good survey of such approaches can be found in [3].

One of the simplest and effective approaches used to enhance degraded document images is the one based on Partial Differential Equation (PDE) and especially the diffusion based methods. PDE approaches are based on a iterative process by which the input image is smoothed in a controlled manner [4]. They are roughly classified in [4] and [5] into two distinct categories, namely linear and nonlinear. Linear diffusion techniques consist of applying local operators uniformly to the degraded image. This corresponds to an isotropic smoothing process. For example, in [6], authors apply a Gaussian kernel in order to improve the geometrical coherence of the noisy image. Isotropic diffusion filter is expressed as follows:

$$\begin{cases} I_t = \operatorname{div}(\nabla(I)) \\ I_o = I_{\text{noisy}} \end{cases} \tag{1}$$

where I_o is the original noisy image and I_t is the considered image after t iterations.

However, linear models fail to discriminate edges from noise. This leads to a blurring effect that may be amplified during the iterative process. To cope with this problem and preserve the sharpness of edges, a nonlinear model was developed by Perona and Malik in [7]. In [7], authors have proposed to weight the linear diffusion function presented in equ. (1) by a conductance function that depends on the gradient magnitude. In this way, the conductance varies across the image according to the signal activity. The idea is then to stop the diffusion across edges and amplify the diffusion along contours. The threshold "stop diffusion" is controlled by the conductance parameter K and the process is expressed as follows:

$$\begin{cases} I_t = \operatorname{div}(c(\|\nabla(I)\|)\nabla(I)) \\ I_o = I_{\text{noisy}} \end{cases} \tag{2}$$

$$\text{where} \begin{cases} c(\|\nabla(I)\|) = \frac{1}{1+\frac{\|\nabla(I)\|^2}{K}} \\ \text{or} \\ c(\|\nabla(I)\|) = \exp^{-\frac{\|\nabla(I)\|}{K}^2} \end{cases}$$

In [8], authors have introduced a novel conductance function expressed in equ. (3). The proposed expression of C provides better results for some types of noisy image such as degraded document images. Furthermore, in [8], a simple method for estimating the conductance parameter at each iteration has been proposed.

$$c(\|\nabla(I)\|) = \frac{1}{2}[\tanh\left(\gamma(K - \|\nabla(I)\|)\right) + 1] \tag{3}$$

Since the pioneer work of Perona and Malik, many improvements have been proposed in the literature. The application of diffusion equation to document image enhancement has been proven efficient in some cases.

For example, in [1], authors have used a diffusion process to model the physical degradation phenomena which may affect document images. They constructed through this model a reverse diffusion process to enhance degraded document images. However, this approach was not blind in the sense that it necessitates the availability of both recto and verso side of the document. Whereas, in the blind approach the verso side of the document is replaced by an estimated background from the recto side [9]. The most recent document enhancement approach based on diffusion process was proposed by Drira and LeBourgeois in [4]. The proposed approach is based on Weickert anisotropic diffusion model [10] using a new diffusion tensor formalism. Indeed, Drira and LeBourgeois have introduced a novel scheme based on eigenvalues functions instead of Weickert functions. This new method offers good performances for degraded document image enhancement. Other nonlinear anisotropic approaches have been proposed for image enhancement and particularly for image inpainting and image deblurring [11], where new anisotropic diffusion functions have been introduced. Instead of using the divergence, authors have proposed to use the trace of the Hessian matrix of image signal to perform an anisotropic diffusion.

The aim of the proposed approach is to enhance foreground coherence characters while removing artifacts in document images. In most cases, these artifacts are nothing else than the reverse side characters which appear on the front side of the document image. In this paper, a nonlinear anisotropic diffusion approach which smoothly removes reverse side strokes and further strengthens foreground side characters is proposed. This paper is organized into 4 sections. After this brief survey on the most recent methods based on diffusion process, the analytic expression of the nonlinear anisotropic diffusion filter based on structure tensor approach is provided in section 2. The proposed method is developed and discussed in details in section 3. The experimental results and the comparison with some approaches are given in section4. Finally, the last section is devoted to conclusion and some potential perspectives.

2. NONLINEAR ANISOTROPIC DIFFUSION

Before discussing the nonlinear anisotropic diffusion, let us first define Ω as an open discrete rectangle included in $\mathbf{R}$ and $I(x, y)$ as a degraded document image where $[x, y]^t \in \Omega$. I is defined as a weighted combination of both the original side (foreground side) I_o and the verso side (background side) I_v including the reverse side strokes and some background artifacts.

The nonlinear anisotropic diffusion based on image tensor introduced by Weickert is expressed through the diffusion process given by:

$$\begin{cases} I_t = \mathrm{div}(D(T)\nabla(I)) \\ I_o = I_{\text{noisy}} \end{cases} \tag{4}$$

Where $D(T)$ is analogous to the conductivity function in the isotropic diffusion and depends only on the structure tensor T which is expressed through eigenvalues $\lambda_\pm$ and eigenvectors $\Theta_\pm$ as follows:

$$T = \lambda_-.\Theta_-\Theta_-^t + \lambda_+.\Theta_+\Theta_+^t \tag{5}$$

The anisotropic diffusion tensor $D(T)$ is given by:

$$D(T) = f_+(\lambda_+, \lambda_-).\Theta_-\Theta_-^t + f_-(\lambda_+, \lambda_-).\Theta_+\Theta_+^t \tag{6}$$

where $f(\lambda_+, \lambda_-) \in [0, 1]$

The main characteristic of anisotropic diffusivity filter is to enhance edges smoothly along their orientations, which correspond to the eigenvector Θ_+ direction. Indeed, if $f_\pm(x, y) \gg 0$, an edge enhancement around (x, y) and along Θ_+ direction is obtained. Furthermore, an isotropic smoothing around (x, y) could appear where $f_\pm(x, y) \approx 0$. These properties are generally exploited in the design of anisotropic diffusion based filter by using various analytic definitions of $f_\pm$ eigenvalues functions. Table .1 summarizes the eigenvalues functions used by Drira-LeBourgois and Weickert approaches.

3. PROPOSED APPROACH

The proposed approach also exploits these observations in order to enhance foreground edges coherence while removing background artifacts smoothly. One could distinguish between foreground and background edges through both eigenvalues and a background estimation algorithms. However, all isotropic/anisotropic diffusion filters, except Moghaddam-Cheriet approach [9], which is based on isotropic diffusivity to model degradations, do not consider a degraded document images as a composition of foreground side edges, reverse side edges and some background artifacts. Therefore, all edges present in the document are considered as important; whereas, background strokes are identified as foreground edges. The proposed method is then motivated by the last observation.

3.1. Proposed diffusion process

The proposed diffusion process introduces a novel parameter called $d \in [0; 1 - \epsilon]$, which represents a normalized difference between I and the estimated background matrix I_b. This parameter is weighted by the constant $\alpha \in \mathbf{R}$. Thus, the proposed process is expressed as:

Table 1. Eigenvalues functions expression of Weickert and Drira-LeBourgois filter

Method	$f_{\pm}$ functions
Weickert	$\begin{cases} f_- = \alpha + (1-\alpha)e^{\frac{-C}{(\lambda_+ + \lambda_-)^2}} & \text{if} \quad \lambda_+ \neq \lambda_- (C = 1^{-10}) \\ f_- = \alpha & \text{else} \\ f_+ = \alpha = 0.001 \end{cases}$
Drira-LeBourgois	$\begin{cases} f_- = e^{\frac{-\lambda_-}{K_-}} \\ f_+ = e^{\frac{-\lambda_+}{K_+}} \end{cases}$ or $\begin{cases} f_- = \frac{1}{1+\frac{\lambda_-}{K_-}} \\ f_+ = \frac{1}{1+\frac{\lambda_+}{K_+}} \end{cases}$

$$\begin{cases} I_t = \text{div}\big(\alpha(1-d(I, I_b)).D(T)\nabla(I)\big) \\ I_o = I_{\text{noisy}} \end{cases} \tag{7}$$

where $d(x, y) = A\,|I(x, y) - I_b(x, y)| + B$
with $A = \frac{1-\epsilon}{\max|I-I_b| - \min|I-I_b|}$ and $B = -A \min|I - I_b|$

It is worth noticing that weighting the divergence operator by $[1 - d]$ plays the same role as the conductance parameter in the case of isotropic diffusion. The matrix d is introduced to enhance discrimination between foreground edges and the rest of the degraded document image. Indeed, $d(x, y)$ is high around foreground edges and low in reverse side edges or in background artifacts. Consequently, $[1 - d(x, y)]$ is higher where $I(x, y)$ is considered as background, and smaller enough but not null where $I(x, y)$ is considered as foreground. In addition, it is worth pointing out that to allow I_t to survive the diffusion process, we do not permit to $[1 - d]$ to attain zero in foreground edges. Thus, $D(T)$ continues to increase foreground characters coherence on Θ_+. This is one of the advantages given by the proposed method. However, if $[1 - d]$ does not decrease enough where $I(x, y)$ is considered as foreground, this last will be blurred. For these reasons, the tuning of the parameter α is of great importance.

3.2. Eigenvalues functions construction

In our framework, we propose to estimate eigenvalues conductance functions as follows:

$$\begin{cases} f_+ = \frac{1}{2}[\tanh(\gamma - \frac{\gamma}{K_+}\lambda_+) + 1] \\ f_- = \frac{1}{2}[\tanh(\gamma - \frac{\gamma}{K_-}\lambda_-) + 1] \end{cases} \tag{8}$$

The construction of the proposed eigenvalues functions is inspired from Monteil and Beghdadi conduction function. Indeed, constructing eigenvalues diffusion functions through hyperbolic tangent allows:

1. Controlling the stop diffusion eigenvalues $\lambda^s_{\pm}$ which depends on both $K_{\pm}$ and γ. Theoretically, $\lambda^s_{\pm}$ should play a role of a separator between foreground strokes and artifacts eigenvalues. Indeed, this feature seems useful to stop the anisotropic diffusion when we deal with the bleed-through effect.

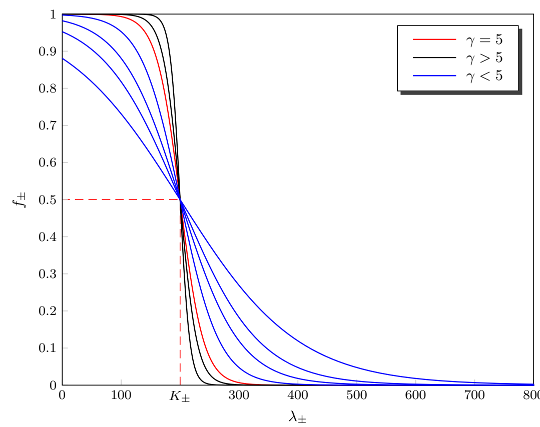

Fig. 1. The proposed eigenvalues function with many decreasing speed parameter. (red curve: the used decreasing speed parameter ($\gamma = 5$), blue curves: speed parameter inferior to 5, black curves: speed parameter superior to 5)

2. Controlling the decreasing speed of the diffusion through γ; this feature seems useful particularly when reverse side strokes are close to foreground strokes gray-level values. In this case, the coefficient γ should be higher to decrease $f_{\pm}$ immediately in background/foreground transition pixels.

The stop diffusion eigenvalues are computed with $f_{\pm}(\lambda^s_{\pm}) = p$, where$\lambda^s_{\pm}$ is given by:

$$\lambda^s_{\pm} = \frac{K_{\pm}}{\gamma}(\gamma - \tanh^{-1}(2p - 1)) \tag{9}$$

For the experiment p is set to $\frac{1}{2}$ and γ is set to 5. Therefore, the stop diffusion controller is nothing else than the coefficient $K_{\pm}$. Figure 1 illustrates this behavior through $f_{\pm}$.

3.3. Background estimation algorithm

Background estimation in still and moving picture has been well investigated during the past decades [12]. Here we introduce a simple background estimation algorithm based on an average operation. The proposed algorithm should be robust and computationally efficient since it is used at each iteration of the diffusion process.

To achieve this, we decompose the document image I_d into blocks $bl_{i,j}$ of size $z \times z$ and compute the average intensity in each block. This is considered as the first background estimation. Thereafter, we re-estimate the average of the block $bl_{i,j}$ computed in the last iteration by computing the average in the n order neighborhood blocks in order to smooth out local irregularities. This routine is iterated until the maximum number of iteration t_{max} fixed $a\ priori$ is reached.

With the aim to reduce the computation time, in the experiment, the parameters bloc size z and the maximum number of iteration t_{max} are set to 8 and 3, respectively. Likewise, only the 1^{st} order neighborhood is considered and the function AvgNeig(A, x, y) is used to compute the 1^{st} order average around (x, y) in the matrix A. Algorithm (1) summarizes the proposed background estimation process.

Algorithm 1 Background estimation

```
Require: I, t_max, z
  for iter = 1 to t_max do
    for each bl_{i,j} ∈ I do
      if iter = 1 then
        Avg[i, j] = E[bl_{i,j}]
      else
        Avg[i, j] = E[AvgNeig(Avg, i, j)]
      end if
    end for
  end for
```

4. EXPERIMENT RESULT

The performance of the proposed method is evaluated and compared subjectively and objectively with some methods of the state-of-the-art. The subjective comparison is based on subjective visual inspection of the obtained results. Whereas, the objective evaluation is performed on the binarized enhanced images. The original degraded images and the corresponding enhanced versions are binarized using some efficient gray-level thresholding algorithms.

4.1. Perceptual evaluation

For this purpose, we compare, visually, the proposed approach with the most effective document enhancement approaches of the state-of-the-art, on document images collected form DIBCO databases[1] and some documents collected from these previous works [1, 8].

For the purpose of ensuring a fair comparison between the considered methods, all the parameters are set to their optimal values in order to get the best perceptual quality of the outputs. Figure 2 depicts performances of Weickert, Monteil-Beghdadi, Drira-LeBourgois and the proposed method.

From the obtained results, it could be noticed that the proposed approach as well as Drira-LeBourgois approach gives a good visual result compared to others as expected since these methods are elaborated especially to deal with document images. Moreover, the proposed method outperforms the Drira-LeBourgois one on all the considered document images (both handwritten and machine printed document). However, the proposed technique is more time consuming. This is mainly due to the estimation of the background at each diffusion iteration. This is the case even if all the background estimation algorithm parameters are optimized to minimize the computation time (see sec. 3.3). However, since this algorithm is used to compute the matrix d, which allows to blindly discriminate foreground strokes and background artifacts from the first iteration, all enhanced document images presented in figure 2 are obtained after 4 to 7 iterations. Furthermore, using both $d(I, I_b)$ and $D(T)$ to discriminate foreground strokes from document images gives a far better flexibility to set $K_{\pm}$ in a wide range of values. For example, the proposed scheme applied to the second document image shown in figure 2 gives good results in $K_{+} \in [150, 250]$ and $K_{-} \in [30, 100]$.

4.2. Quantitative evaluation

The quantitative evaluation of the proposed method is based on the binarization of the processed images. Both proposed approach and Drira-LeBourgois approach are applied to enhance all document images suffering from bleed-through in DIBCO 10-11-12 databases. Binarization algorithms are performed on the enhanced document images resulting from these filters. Thereafter, we use FMeasure (equ. (10)), Pseudo-FMeasure (this metric is computed as the same manner as FMeasure with substituting the recall by the pseudo-recall function defined on equ. (11)), DRD (this metric is used on DIBCO competition since 2011, and introduced in [13]) and the PSNR criterion to evaluate the proposed method and Drira-LeBourgois method. The images are binarized using the grey-level thresholding algorithms of Otsu [14], Tsai [15] and Howe [16], which is the winner of H-DIBCO 2012 competition.

As in the perceptual evaluation, parameters $K_{\pm}$ associated with the number of iterations of Drira-LeBourgois filter and parameters $K_{\pm}$ associated with the number of iterations of the proposed filter are all optimized to obtain the best results in terms of FMeasure criterion.

Table 2 shows the obtained results. Note that the proposed approach outperforms Drira-LeBourgois approach for all the considered binarization algorithms. However, even if the proposed method globally gives better results than the others in the case of Howe binarization, the obtained results do not differ significantly. This small gain is due especially to the coherence enhancement of foreground characters given

[1] http://utopia.duth.gr/ ipratika/DIBCO2013/resources.html

Fig. 2. From left to right: Original degraded document image, Weickert filter, Monteil-Beghdadi filter, Drira-LeBourgois filter, the proposed filter.

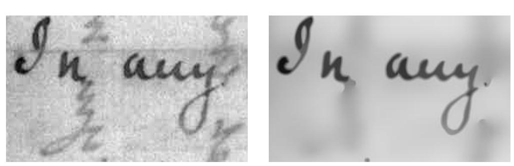

Fig. 3. Performances of the proposed approach in strengthening foreground coherence characters, left: degraded document image sample, right: enhanced document image with the proposed approach.

by the proposed approach and the fact that Howe binarization algorithm is well-adapted to eliminate bleed-through in degraded document images. Therefore, the performances of these methods when combined with the Howe binarization algorithm are somehow comparable. Figure 3 illustrates the strengthening foreground coherence characters performances given by the proposed approach.

$$\mathrm{F}_{\text{Measure}} = 100 * \frac{2 * \mathrm{PR} * \mathrm{RC}}{\mathrm{PR} + \mathrm{RC}} \tag{10}$$

where $\mathrm{RC} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$ and $\mathrm{PR} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$ and TP, FP and FN represent respectively, the true positive, false positive, and false negative values in the binarized image.

$$\text{Pseudo-RC} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{bin}(i,j) I_{sg}(i,j)}{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{sg}(i,j)} \tag{11}$$

where I_{bin} and I_{sg} represent the binarized document image, of size MxN, and its corresponding skeletonized ground truth version, respectively.

Table 2. FM and PSNR results on applying both proposed and Drira-LeBourgois filter on degraded document images suffering from bleed-thorough in DIBCO 10-11-12 databases

		OTSU	TSAI	HOWE
FM	Drira-LeBourgois	76,54	76,80	94,11
	Proposed filter	**80,52**	**80,19**	94,14
P-FM	Drira-LeBourgois	79.03	77.92	95.81
	Proposed filter	**83.40**	**81.69**	**95.97**
DRD	Drira-LeBourgois	30.91	17.49	1.83
	Proposed filter	**28.26**	**15.12**	1.86
PSNR	Drira-LeBourgois	0,23	-0,55	13,71
	Proposed filter	**4,25**	**2,77**	**14,11**

5. CONCLUSION

In this work, we have introduced a novel blind approach for document image enhancement based on a nonlinear anisotropic structure tensor diffusion. The obtained results confirm its efficacy to eliminate bleed-through smoothly while strengthening foreground characters. The presented approach includes a novel anisotropic diffusion process system based on the background estimation of the processed document. Furthermore, the proposed anisotropic diffusion process introduces a new eigenvalues functions driving properly structure tensors. The subjective and objective evaluations performed on the obtained results on a set of images confirm the outperformance of the proposed method over the other considered methods. However, as Drira-LeBourgois the most recent and effective filter adapted to document images, the proposed approach also requires some efforts for tuning the parameters. Nevertheless, we have shown that the proposed approach offers a flexible way to tune the parameter $K_{\pm}$ in a wide range of values. One of the promising direction for future works is then to develop strategies for automatic setting of this parameter.

6. REFERENCES

[1] R.F.Moghaddam and M.Cheriet, "Low quality document image modeling and enhancement," *International Journal of Document Analysis and Recognition (IJDAR)*, vol. 11, no. 4, pp. 183–201, 2009.

[2] S.Chen, A.Beghdadi, and M.Cheriet, "Degraded color document image enhancement based on nrcir," in *Visual Information Processing (EUVIP), 2010 2nd European Workshop on*, July 2010, pp. 19–22.

[3] G.Leedham, Y.Chen, K.Takru, H.Joie, T.Nata, and M.Li, "Comparison of some thresholding algorithms for text/background segmentation in difficult document images," in *Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on*, Aug 2003, pp. 859–864.

[4] F.Drira, F.LeBourgeois, and H.Emptoz, "A new pde-based approach for singularity-preserving regularization: application to degraded characters restoration," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 15, no. 3, pp. 183–212, 2012.

[5] S.Morfu, "On some applications of diffusion processes for image processing," *Physics Letters A*, vol. 373, no. 29, pp. 2438–2444, 2009.

[6] A. Tikhonov and V. Arsenin, "Solution of ill-posed problems," *Winston and Wiley*, 1977.

[7] P.Perona and J.Malik, "Scale-space and edge detection using anisotropic diffusion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 7, pp. 629–639, Jul 1990.

[8] J.Monteil and A.Beghdadi, "A new interpretation and improvement of the nonlinear anisotropic diffusion for image enhancement," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 9, pp. 940–946, 1999.

[9] R.F.Moghaddam and M.Cheriet, "Rsldi: Restoration of single-sided low-quality document images," *Pattern Recognition*, vol. 42, no. 12, pp. 3355–3364, 2009.

[10] J.Weickert, "Coherence-enhancing diffusion of colour images," *Image and Vision Computing*, vol. 17, no. 3, pp. 201–212, 1999.

[11] D.Tschumperle and R.Deriche, "Vector-valued image regularization with pdes: A common framework for different applications," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 4, pp. 506–517, 2005.

[12] P.L.M.Bouttefroy, A.Bouzerdoum, S.L.Phung, and A.Beghdadi, "On the analysis of background subtraction techniques using gaussian mixture models," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, March 2010, pp. 4042–4045.

[13] H.Lu, A.Kot, C.Alex, and Y.Shi, "Distance-reciprocal distortion measure for binary document images," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 228–231, 2004.

[14] N.Otsu, "A threshold selection method from gray-level histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, no. 1, pp. 62–66, Jan 1979.

[15] C.Tsai and H.Lee, "Binarization of color document images via luminance and saturation color features," *Trans. Img. Proc.*, vol. 11, no. 4, pp. 434–451, Apr. 2002.

[16] N.Howe, "A Laplacian energy for document binarization," in *International Conference on Document Analysis and Recognition*, 2011, pp. 6–10.