Digital Signal Processing 67 (2017) 61–75

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

A collaborative enhancement-compression approach for historical document images based on PDE-analysis

Mohamed Riad Yagoubi [a,*], Amina Serir [a], Azeddine Beghdadi [b]

[a] USTHB, Department of Telecommunications, LTIR, Algiers, Algeria
[b] Institut Galilée, L2TI, Paris, France

ARTICLE INFO

Article history:
Available online 27 April 2017

Keywords:
Historical document image
Compression
Enhancement
JPEG 2000
JPEG XR
Diffusion

ABSTRACT

Historical document images suffer from several kinds of degradation affecting their readability and increase their sizes when digitizing. Indeed, preserving the old document heritage requires an important storage infrastructures, while this space could be significantly reduced if one manages to safeguard these documents without their considerable amount of artifacts which increasing their entropy. We present in this work a new and more advantageous solution for the document image compression technology dedicated to National Archives and Libraries. The aim of this work is to propose a novel historical document compression scheme judiciously combined with a novel automatic enhancement scheme, both based on PDE-Analysis applied in an innovative way, to face this challenging issue without affecting the look and feel of the historical document image. Throughout experiment, the proposal is fairly evaluated on overall DIBCO datasets and some historical documents collected from the web by means of objective measures and perceptual judgment against the most used enhancement algorithms, the most robust compression standards such as JPEG-XR and the most used sequential compression–enhancement combinations techniques to proof the robustness of the proposed collaborative approach against classic sequential approaches.

1. Introduction

Actually, National Archives and libraries provide important efforts to digitize their large amount of books, articles and other historical manuscripts. In consequence, several virtual libraries, sharing a large number of digitized books in many languages and topics such as "Wikisource",[1] have appeared recently.

Archiving historical document images requires a huge capacity of storage. Indeed, the noise and distortions affecting the historical manuscripts contribute to increase the entropy which explicitly attains the compression efficiency. However, so far, due to the important amount of artifacts on which many ancient documents are suffering, existing image compression technologies fail to provide acceptable performances when facing historical document images. It is demonstrated that commonly used compression standards, such as JPEG, show their limit when applied on scanned document images [1,2]. The solution could be in applying a robust enhancement algorithm to reduce those artifacts before performing the compression process. However, we demonstrate in this study that applying consecutively a document image enhancement algorithm followed by a compression standard affects considerably the readability of the document if one tries to reach aggressive compression bitrates.

Mixed Raster Content (MRC) (ITU-T T.44) based standards such as DjVu [2,3] seem to be more effective when facing document images; nevertheless they do not perform on degraded historical documents as well as on natural documents regarding to some specificities characterizing historical documents as we will show throughout this study. MRC standard defined in [4] is the cornerstone of the most effective compression approaches specialized in document image. Roughly speaking, MRC strategy focuses on keeping the full structural information (text) while compressing massively outwards. Thus, the MRC strategy consists in preserving the text while reducing the compression rate in the rest of the document. MRC mode-one is the most used MRC approach. It is based on the separation of the document image into foreground, background, and mask layers. The latter is a binary image containing the text or/and drawing estimated from the document image. The foreground layer contains the color of the text or/and drawing present in the mask layer, while the background layer includes all

* Corresponding author.
E-mail address: myagoubi@usthb.dz (M.R. Yagoubi).

[1] https://fr.wikisource.org/wiki/Wikisource:Accueil.

http://dx.doi.org/10.1016/j.dsp.2017.04.009
1051-2004/

pictures present in the document. Then, with the aim to the preserve the structural information, a lossless or a high bitrate encoding is performed to compress the mask layer, while the background and the foreground layer are both compressed with a low bitrate.

One of the most used MRC commercial standards, which is based on a statistical analysis (HMM), is DjVu [2] (The whole Wikisource document images are compressed using DjVu technology). Within DjVu, the foreground and the background images are compressed at low resolution through a wavelet based algorithm, called IW44, where the binary mask layer is compressed with JB2, a variation of JBIG2 algorithm [5]. Unfortunately, while DjVu is the best document image compression algorithm [6], the embedded segmentation algorithm (HMM-based) in this standard does not seems to be suitable for historical document images. Fig. 1 illustrates results obtained when applying DjVu encoder on both natural and handwritten document images.

PDE-based approaches have also encrusted the image compression field [7,8]. Within these approaches, authors generally try to preserve only a well chosen subset of the image such as structural edges, then retrieve missing pixels by PDE-based interpolation technique. Compression using edge information and homogeneous diffusion seems to be efficient on cartoon images [8]. However, to our best knowledge, such compression techniques have never been applied on natural images or historical document images even if the latter present a great potential of structural information.

Indeed, historical document images present particular structural information and may suffer from many artifacts and degradations such as noise, bleed-through or ink smudges. By definition, document image enhancement consists of removing those degradation marks (e.g. artifact, noise) without affecting the original aspect of the document, so as preserving the look and the feel of the processed document image. Many methods have been proposed in the literature [9–11] and a good survey of such approaches can be found in [12].

Natural image enhancement could be the solution to reduce the artifacts on which the historical document images suffer from. Indeed, the natural image enhancement is a well investigated field and still a subject of ongoing research [13]. For example, recently, Gu et al. have proposed an effective approach [14], based on a histogram modification framework, which improves automatically the contrast of salient information in the image. However, the historical document images present particular structural information comparing to natural images since they suffer from singular kinds of degradations such as the bleed-through which is, often, hardly distinguishable from the foreground edges. In consequence, performing an image enhancement approach increases the readability alteration since the useful information and the bleed-though contrast are improved simultaneously. This underlines the importance to develop particular enhancement/restoration techniques dedicated to document images.

Within the state-of-the-art, document image quality enhancement is categorized into two distinct types, namely, double-side and single-side approaches. Double-side approaches deal with this task where considering that both the front and the reverse-side of the historical document could be available. Whereas, since the back-side face is rarely available, the single-side (also called blind) restoration problem is in consequence more investigated in the literature [9,15,16].

Among blind techniques, one can cite: Statistical-based methods, color-based [11] or PDE-based methods [9,15]. Due to their ability to preserve the aspect of the processed document, PDE-based approaches are considered as the most effective solutions for document image enhancement. The basic idea of PDE-based approaches is to consider the image as a physical medium where a controlled diffusion takes place. The diffusion process is then expressed through an iterative scheme starting from the initial state corresponding to the original degraded image. PDE-based image enhancement approaches could be categorized into three distinct segments, namely "linear", "nonlinear isotropic" and "nonlinear anisotropic" approaches.

Linear diffusion was originally used to describe heat diffusion process through a differential equation. Roughly speaking, linear diffusion techniques consist of applying local operators uniformly to the degraded image, which corresponds to an isotropic smoothing process. However, linear filtering operators fail to discriminate the salient features from noise. Therefore, they tend to remove noise while smoothing edges. This drawback has motivated Perona and Malik to develop the first nonlinear anisotropic process in their pioneer work in [17]. To avoid warping edges, Perona and Malik have proposed to weight the diffusion process by the gradient based conductance function so as to adapt the diffusion to the local activity of the signal. The diffusion is allowed inside uniform region and stopped across edges. However, to perform such discrimination, a threshold value of the conductance needs to be defined to separate edges from the noise. The whole process is then governed by the following equations:

$$\begin{cases} u_t = \mathrm{div}(c(\|\nabla(u)\|)\nabla(u)) \\ u_o = u_{\text{noisy}} \end{cases} \tag{1}$$

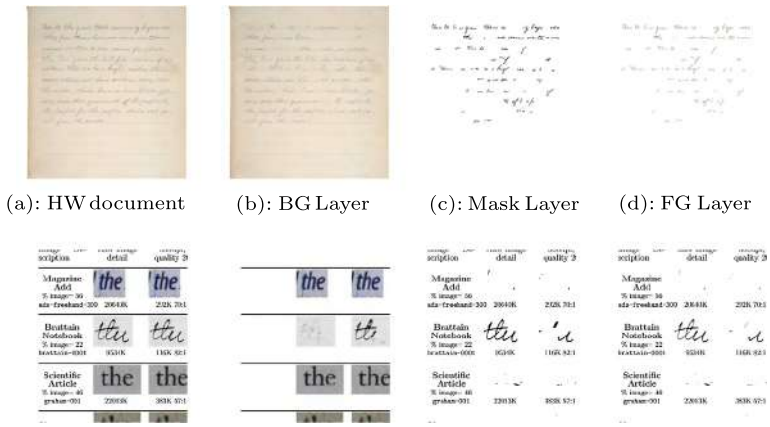
∇: divergence operator, $\nabla(u) = [\frac{du}{dx}, \frac{du}{dy}]^t$ where u_t is the output image at iteration t, u_o is the original degraded image and:

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

The efficiency of the Perona and Malik filter depends highly on the conductance parameter K. However, this constant is arduous to tune since its relies on several parameters such as the nature of the treated image (document image, medical image), the edges intensity and the distribution of the noise.

Weickert [18] introduced another framework based on structure tensor to overcome the limitations of classical PDE-based filtering methods. This framework has been further exploited and adapted to enhance document images in [9,15]. However, in Weickert formalism, the anisotropic diffusion requires to tune, not only one, but two parameters used to highlight eigenvalues representing foreground edges from those representing artifacts denoted $K_{\pm}$. This makes this task more challenging and one of the important purposes of this study is to develop an effective solution to overcome this drawback.

Besides, in this work, a diffusion approach for document image enhancement is developed in the first part. Since this process needs to tune some parameters to perform correctly, two solutions are proposed to automate this process. These solutions are incorporated thereafter within a PDE-based compression scheme developed in this work to deal with the historical document compression task. In fact, one could believe that a proper detection of foreground edges is the key to perform an acceptable compression when facing historical documents. Hence it is necessary to drive judiciously the encoder algorithm to compress massively around artifacts and perform a slight or lossless compression in edges areas. Otherwise, compressing historical document by using classical compression schemes introduces some distortions on the foreground edges and affects the readability of the document. Moreover, the compression performances would be less interesting if the encoding scheme is uniformly applied in the whole document image without any artifacts/edges differentiation as done in the widely used image compression standards JPEG [19] or JPEG2000 [20]. These observations are behind the main motivation of our work.



(a): HW document (b): BG Layer (c): Mask Layer (d): FG Layer

(e): Natural document (f): BG Layer (g): Mask Layer (h): FG Layer

Fig. 1. DjVu Encoder applied on both Handwritten document image (top) and natural document image (down).

This paper is organized into 7 sections. After presenting some efficient image compression standards followed by PDE-based image enhancement techniques, section 2 is devoted to recall the nonlinear anisotropic diffusion technique as formalized by Weickert [18]. Afterward, we present in section 3 the proposed diffusion formalism. In section 4 the automatic PDE-based document image enhancement technique is presented in details. Thereafter, the proposed compression scheme adapted to the scanned historical document is described and detailed in section 5. Section 6 is devoted to the experiment study while the conclusion and some potential future directions are presented in the last section.

2. Non-scalar anisotropic diffusion in the state-of-the-art

First, let us introduce some basic notations and definitions to make this paper self-contained. Let us denote u the degraded document image composed of the original side (commonly called foreground side) denoted u_f and the background side denoted u_b including all reverse-side edges and other background artifacts.

We introduce now the structure tensor T depending on the image eigenvalues $\lambda_\pm$ and their corresponding eigenvectors $\Theta_\pm$. T is expressed as follows:

$$T = \nabla(u).\nabla(u^t) = \begin{bmatrix} u_{x^2} & u_{x,y} \\ u_{y,x} & u_{y^2} \end{bmatrix} \tag{2}$$

or, through eigenvalues and eigenvectors:

$$T = \lambda_-.\Theta_-\Theta_-^t + \lambda_+.\Theta_+\Theta_+^t \tag{3}$$

In [18], Weickert has introduced a function denoted D, which depends on the structure tensor matrix T, to control the diffusion rate. In brief, the anisotropic diffusion tensor $D(T)$ introduces a local modification on the structure tensor T by re-calibrating the eigenvalues through functions $f_\pm(\lambda_+, \lambda_-)$ ($f_\pm \in [0, 1]$) as mentioned in Eq. (4). This modification is mainly due to some observed behaviors of the eigenvalues across image features (see Fig. 2). the structure tensor $D(T)$ is formalized as follows:

$$D(T) = f_+(\lambda_+, \lambda_-).\Theta_-\Theta_-^t + f_-(\lambda_+, \lambda_-).\Theta_+\Theta_+^t \tag{4}$$

The eigenvalues behavior across the image features could be summarized as follows:

- $\lambda_- = \lambda_+ \approx 0$ along smooth areas.
- Edges are characterized by $\lambda_- \gg \lambda_+ \approx 0$.
- Corners are characterized by $\lambda_- > \lambda_+ \gg 0$.

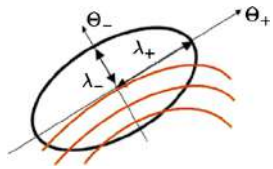


Fig. 2. Eigenvalues and eigenvectors behavior against edges.

Then, the first non-scalar anisotropic diffusion model, introduced by Weickert [18], is governed by the following model:

$$\begin{cases} u_t = \mathrm{div}(D(T)\nabla(u)) \\ u_o = u_{\text{noisy}} \end{cases} \tag{5}$$

Since the first modification given by Weickert, many attempts have been done to develop other analytic expressions of the eigenvalues functions $f_\pm(\lambda_-, \lambda_+)$. However, these solutions focus on highlighting edges and corners from noise, which is inappropriate when dealing with the bleed-through marks. The unique study which takes benefits from the non-scalar anisotropic diffusion to overcome the bleed-through issue within historical document images has been proposed by Drira et al. [21]. In [21], the authors propose to associate the anisotropic diffusion formalism as introduced by Weickert with new eigenvalues functions $f_\pm$ (see Table 1) inspired from conductance functions introduced by Perona and Malik [17]. Table 1 shows the eigenvalues functions used by Drira et al. and some commonly used diffusion formalism and eigenvalues functions in the state-of-the-art.

Models presented in Table 1 are each dedicated for a specific application, thus using these solutions for document image analysis could be benefic rise some serious disadvantages. The model based on the Hessian trace proposed by Tschumperle et al. [23] fails in enhancing less contrasted edges which are in coherence with their neighborhoods. The diffusion matrix introduced in the Beltrami model [22] decreases instantaneously even when facing smalls variations, which makes it inadequate for preserving foreground edges. The model proposed by Weickert [18] presents some interesting features, especially in enhancing the less contrasted foreground edges. However, an over soothing result is often observed since the edge enhancement process is not well governed. In consequence, the singularities are often affected when using this model as shown in Fig. 3. Advantages and disadvantages observed when applying these diffusion models for the document analysis field are summarized in Table 2.

Table 1
Diffusion models in the state-of-the-art.

Approach	Conductance function $f_{\pm}$	Diffusion equation
Weickert (EED) [18]	$\begin{cases} f_+ = \alpha + (1-\alpha)e^{\frac{-C}{(\lambda_+ + \lambda_-)^2}} & \text{if } \lambda_+ \neq \lambda_- \\ f_- = \alpha & \text{else} \\ f_- = \alpha = 0.001 \quad (C = 1^{-10}) \end{cases}$	$u_t = \text{div}(D(T)\nabla(u))$
Drira et al. [15]	$\begin{cases} f_- = e^{\frac{-\lambda_-}{K_-}} \\ f_+ = e^{\frac{-\lambda_+}{K_+}} \end{cases}$ or $\begin{cases} f_- = \frac{1}{1+\frac{\lambda_-}{K_-}} \\ f_+ = \frac{1}{1+\frac{\lambda_+}{K_+}} \end{cases}$	$u_t = \text{div}(D(T)\nabla(u))$
Beltrami et al. [22]	$\begin{cases} f_+ = \sqrt{\frac{1+\lambda_+}{1+\lambda_-}} \\ f_- = \sqrt{\frac{1+\lambda_+}{1+\lambda_-}} \end{cases}$	$u_t = \frac{1}{\sqrt{(1+\lambda_+)(1+\lambda_-)}}\text{div}(D(T)\nabla(u))$
Tschumperle et al. [23]	$\begin{cases} f_+ = \frac{1}{\sqrt{1+\lambda_+ + \lambda_-}} \\ f_- = \frac{1}{\sqrt{1+\lambda_+ + \lambda_-}} \end{cases}$	$u_t = trace(DH)$, H: Hessian matrix

Table 2
Common diffusion process – a brief comparison of their features in document analysis.

Diffusion process	Artifacts removing	Enhancement of less contrasted edges	Singularity preservation
Perona et Malik [17]	√	X	√
Beltrami et al. [22]	√	Moderately	X
Tschumperle et al. [23]	√	X	√
Weickert [18]	√	√	X

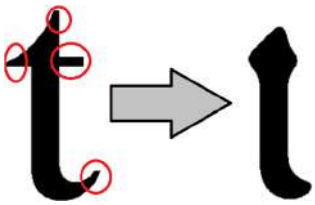


Fig. 3. Example of singularities deformations observed when performing Weickert model [18].

3. Proposed diffusion formalism

In this study, we propose to reformulate the basic scheme on which the non-scalar anisotropic process is based, by introducing a new formalism adapted for the specificity of document image distortions.

Indeed, we propose a new PDE-based method for removing distortion from historical document images without affecting the salient features. The proposed iterative system could remove the bleed-through artifacts and enhances gradually the foreground edges coherence simultaneously. The proposed scheme is governed by the following equations:

$$\begin{cases} u_t = \text{div}\big(\alpha(1 - d(u, u_{bkg})).D(T)\nabla(u)\big) \\ u_o = u_{\text{noisy}} \end{cases} \tag{6}$$

while d is the normalized difference between the degraded document u and its estimated background denoted u_{bkg}.

$d(x, y) = A\left|u(x, y) - u_{bkg}(x, y)\right| + B$

with $\begin{cases} A = \frac{1-\epsilon}{\max\left|u-u_{bkg}\right| - \min\left|u-u_{bkg}\right|} \\ B = -A\min\left|u - u_{bkg}\right| \end{cases}$

The key idea behind weighting the divergence parameter by $[1 - d]$ is to add "background artifacts" removal feature to the diffusion process. In other words, $[1 - d]$ plays the same role as the conductance parameter within the Perona and Malik formalism. However, in our case, d is the difference between the degraded document and its estimated background. Consequently, since the degraded document image and its estimated background image present similar values in background pixels, $[1 - d]$ is then high when facing artifacts areas and the process can diffuse massively. On the other hand, $[1 - d]$ is low outward (foreground pixels) which stop the diffusion process and preserve the foreground edges.

While the discrimination "edges"/"background artifacts" is performed through $[1 - d]$, the eigenvalues functions intervene to discriminate "foreground edges"/"reverse-side edges". Within the proposed diffusion process, we propose to estimate the eigenvalues functions through the hyperbolic tangent as follows:

$$\begin{cases} f_+ = \frac{1}{2}[\tanh(\gamma - \frac{\gamma}{K_+}\lambda_+) + 1] \\ f_- = \frac{1}{2}[\tanh(\gamma - \frac{\gamma}{K_-}\lambda_-) + 1] \end{cases} \tag{7}$$

The hyperbolic tangent allows to control the decreasing speed of the diffusion through γ and permits to stop diffusing instantly when facing foreground edges. This feature seems to be useful particularly when the reverse-side and foreground strokes contrasts are close and confused. In this special case, γ should be higher to increase the derivatives of $f_{\pm}$ in $K_{\pm}$. Fig. 4 illustrates the eigenvalues functions behaviors.

4. Automating the proposed diffusion formalism

Anisotropic diffusion is an innovative and wise solution to several image processing issues [7,9]. However, since this process relies highly on some crucial parameters, such as the diffusion shut-off and the conduction parameter, automating such an iterative process when facing image processing issues becomes extremely challenging.

4.1. A brief overview on automatic diffusion systems in the state-of-the-art

In the state-of-the-art, since the first automation proposed in [24], many attempts have been done to deal with this arduous task. The most recent study is proposed by Tsiotsios et al. in [25]. In this paper, the authors proposed to use a statistical study combined with the well-known thresholding Knee algorithm [26] to estimate the conductance parameter K. However, this approach deals with a Gaussian noise removal while the local structural information and the specificities of the images are not taken into

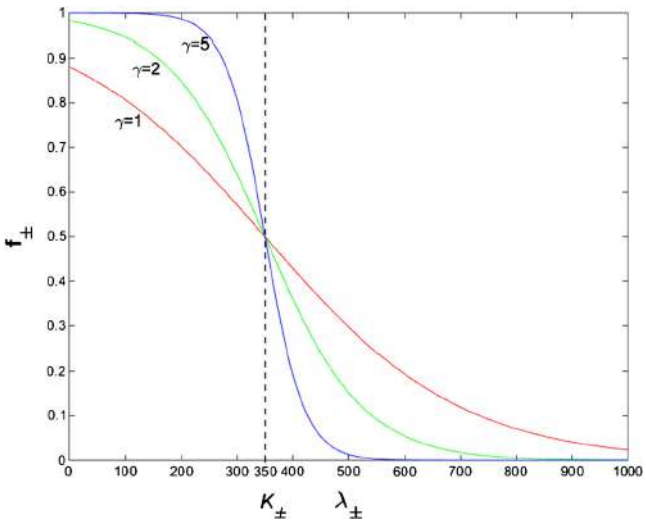


Fig. 4. The impact of γ on the decreasing speed of eigenvalues functions in $K_{\pm}$.

account. In addition, the proposed solutions are exclusive to the anisotropic diffusion filter as formalized by Prona and Malik [17], while the Weickert [18] formalism is more suitable to deal with document image artifacts [9].

Recently, Drira et al. [21] have proposed an automatic diffusion filter, governed by the Weickert formalism, dedicated for removing bleed-through artifacts and fully adapted for historical documents. In this work, the authors have focused their efforts on tuning the conductance parameters when dealing with historical document. Through some observations, the authors have proposed to set $K_{\pm}$ as follows:

$$\begin{cases} K_{-} = 0.1 * \max(\lambda_{-}) \\ K_{+} = \frac{K_{-}}{2} \end{cases} \tag{8}$$

However, the diffusion shutoff parameter is the major drawback of this process. Indeed, this iterative process must be stopped manually when approaching acceptable results since it diverges quickly after a certain number of iterations. In addition, the Drira et al. filter seems to be not extendable on other document image databases. In fact, $K_{\pm}$ are simply estimated through some observations performed on a specific document image dataset, while this automation is not suitable if one tries to apply this automatic method to reduce bleed-through within numerical manuscripts collected from other databases.

4.2. Proposed automatic diffusion method

Automating the enhancement task when dealing with historical degraded documents provides new and more advantageous solutions for document image restoration technology. The automation targets two components: the diffusion shutoff and the conductance parameters. These latter seems to be crucial since overestimating $K_{\pm}$ leads certainly to an over-smoothed result, while an insufficient estimation of these parameters enhances the bleed-through edges contrast and may corrupt the document readability. Herein, two solutions are proposed to automate the proposed diffusion formalism.

Within the proposed processes, instead of providing an important computational effort to calibrate $K_{\pm}$ in each iteration as done in [25] or proposing a simple thresholding technique to separate foreground/background edges based on some observation performed *a priori* as done in [21], we propose to take into account the local characteristics of each pixel (e.g. coherence and the contrast relevance around the neighborhood) to determine its affiliation (e.g. foreground or artifacts).

The idea consists of estimating, in each pixel, a coefficient expressing the coherence degree of this pixel around its neighborhood in terms of intensity and orientation, simultaneously. Modeling those characteristics proves to be conceivable using the eigenvalues and eigenvectors; in addition, since the eigenvalues and eigenvectors are explicitly involved within the proposed anisotropic diffusion process, using those parameters avoid the proposals to provide a consequent computational effort.

We denote $\Phi(x, y)$, the coefficient which has a role to quantify the intensity and orientation coherence according to the eigenvalues and eigenvectors behaviors (see Sec. 2). $\Phi(x, y)$ produces important values across foreground edges and less enough values outwards to permit a proper discrimination. Aside to be simple-to-implement and provides a well foreground/background discrimination, the proposed iterative process is convergent and does not need shutoff parameter since the proposed filter provides a stable result after a few number of iteration.

To this end, two foreground/background discriminatory measures are proposed in the following:

4.2.1. Eigenvalues Coherence Measure – ECM

$\Phi(x, y)$ is estimated here with the "normalized coherence measure" and the "gradient norm", simultaneously. The coherence measure $\mathbf{c} \in [0; 1]$ is well known and widely used within inpainting approaches [27]. This measure is expressed generally as follows:

$$c = \left(\frac{\lambda_{-} - \lambda_{+}}{\lambda_{-} + \lambda_{+}}\right)^2 \tag{9}$$

Due to the eigenvalues behavior across relevant edges, corners and background (see sec. 2), the coherence measure c is high (≈ 1) in "smooth relevant edges" and "smooth background areas". Thereby, in order to develop a criterion that better represents smooth relevant edges without confuse them by any other features, we propose to weight the normalized coherence measure by the gradient norm $\|\nabla u\|$ of the degraded document u. Since the gradient is relevant on "reverse-side/foreground" edges and less into "smooth background pixels", ($c.\|\nabla u\|$) would be relevant only across "smooth relevant edges". Then, one could propose to express Φ, at element (x, y), as follows:

$$\Phi(x, y) = \left(\frac{\lambda_{-}(x, y) - \lambda_{+}(x, y)}{\lambda_{-}(x, y) + \lambda_{+}(x, y)}\right)^2 . \|\nabla u(x, y)\| \tag{10}$$

4.2.2. Eigenvalues–Eigenvectors Coherence Measure – EECM

Herein, the function $\Phi(x, y)$ estimates coherence degree, in terms of orientation, of the anisotropic tensor field across (x, y) and, at the same time, the relevance of the considered pixel. This is done by using the behavior of λ_{-} and eigenvectors couple $(\Theta_{+}; \Theta_{-})$ around edges (see Fig. 2). Then, we propose to estimate Φ at (x, y) as follows:

$$\Phi(x, y) = \lambda_{-}(x, y). \arctan \frac{\|\Theta_{-}(x, y)\|}{\|\Theta_{+}(x, y)\|} \tag{11}$$

The elliptical tensor is more stretched ($\lambda_{-} >> 0$) around relevant edges (see Fig. 2) and $\arctan \frac{\|\Theta_{-}\|}{\|\Theta_{+}\|}$ is somehow the degree to which the neighborhood directions of (x, y) are coherent according to (x, y) direction [18]. Then, $\Phi(x, y)$ exhibits high values when edges are contrasted and present a coherent direction around their neighborhood, which is used to discriminate between the different image components.

Table 3
ECM/EECM – brief comparative.

	Advantages	Inconvenient
ECM	Provides a proper discrimination where the whole foreground edges are highly contrasted	Could confuse foreground/background edges where they present close contrast values.
EECM	Greatly highlight less-contrasted foreground edges if they are coherent with their neighborhood.	Since pixels intensity vary gradually in the background, certain background pixels could be considered as foreground due to their high coherence among each other.

4.2.3. Thresholding process

Once the discriminatory measures have been introduced, it is important to point out that, even if these criteria are estimated through consistent basis to discriminate "foreground"/"reverse-side" strokes, both of the proposed expression of Φ present some limitations. In fact, since the structure tensor stems directly from the gradient measures as expressed in Eq. (2), characters with large widths may spread through the background and become unreadable when using the ECM measure. Moreover, the EECM measure could be high in some reverse-side strokes presenting high homogeneity even if they are low contrasted, which may affect the outputs. The main advantages and inconvenient observed when using each measure are summarized in Table 3.

In order to try to reduce these drawbacks, we propose to subdivide the matrix Φ into blocks of size $k * k$ called $s_i^k, i = \{1, ..., \frac{\text{card}(u)}{k^2}\}$. Afterward, we make use of two less restrictive assumptions:

1. Foreground strokes are usually more relevant than reverse-side one.
2. Blocks containing foreground edges and those containing background one do not exhibit similar structural and statistical features. From experiment, one observes that a simple standard deviation is sufficient to highlight this statistical dissimilarity.

Relying on the considered hypotheses, we first estimate the standard deviation std in each block. Then sort the given values in a decreasing order and compute the relative variation of the reordered coefficients as expressed in Eq. (12). According to the second assumption, the first maximum of ∂_i represents the frontier between "foreground edges"/"reverse-side strokes" while the second one represents the transition "reverse-side strokes"/"background".

$$\partial_j = \frac{\text{std}_j[s^k] - \text{std}_{j-1}[s^k]}{\text{std}_{j-1}[s^k]} \tag{12}$$

where j is the order of the coefficient stat$[s_i^k]$ when sorting in the decreasing order.

In other words, $\max_1(\partial_i)$ could be used to discriminate blocks containing foreground edges from the others in Φ. In addition, to make sure that s_i^k could fit on the whole foreground shapes regardless to their sizes, one can perform this process on $k = \{4, 8, 16\}$, and s_i^k is considered as foreground if and only if this block is considered as foreground in at least 2/3 of the considered block sizes.

The last step of the proposed automation algorithm consists of thresholding the eigenvalues $\lambda_\pm$. This thresholding process works as follows: if (x, y) is considered as foreground pixel through the last discrimination, the value 1 is attributed to $\lambda_\pm(x, y)$ to stop diffusion across the foreground edges. Whereas, the value of $\lambda_\pm(x, y)$ is kept to diffuse depending on its intensity since the pixel is considered as artifact. The proposed automatic document image enhancement algorithm is detailed bellow.

This algorithm is then incorporated within the proposed anisotropic diffusion process (see Eq. (6)) to produce a full automatic enhancement filter. However, it has been noticed that estimating the threshold matrix T_s after performing a number of iterations boosts significantly performances. The edges coherence enhancement along the structure tensor direction is the key to this improvement. Indeed, some less contrasted foreground edges are enhanced significantly after performing a number of iterations if these edges are in coherence with their neighborhood, an example is illustrated in Fig. 5. Thereby, we propose to perform the following strategy:

Step.01 Perform the diffusion process presented in Eq. (6) in few number of iteration (e.g. from experiment, 30 iterations seems to be quiet enough) with very low values of $K_\pm$ to improve the edge coherence.

Step.02 Estimate the thresholding matrix T_s through Algorithm 1.

Step.03 Perform the automatic diffusion process as presented in Algorithm 2.

Algorithm 1 Perform the threshold matrix T_s.

Require: $\lambda_\pm$,$\Theta_\pm$,$MeasureType$
  Declare vector K={4, 8, 16}
  if $MeasureType$=ECM **then**
    Compute Φ from Eq. (10).
  else if $MeasureType$=EECM **then**
    Compute Φ from Eq. (11).
  end if
  for each k on K **do**
    Declare structure $\tau^k : [\text{card}(u)/k^2]$
    for $i = \{1, ..., \text{card}(u)/k^2\}$ **do**
      Estimate $\text{std}_i[s^k]$.
      Sorting $\text{std}_i[s^k]$ on the decreasing order.
      Perform ∂_j from Eq. (12), then ∂_{max_1}.
      Extract **ord** the ancient order of max_1 before sorting.
      for $i = 1$ to $\text{card}(u)/k^2$ **do**
        if $\text{std}_i[s^k] < \text{std}_{ord}[s^k]$ **then**
          $\tau^k(i) \leftarrow 1$
        else
          $\tau^k(i) \leftarrow 0$
        end if
      end for
    end for
  end for
  for each pixel (x, y) in $\lambda_\pm$ **do**
    if $\sum_{k=1}^{card(K)} \tau^k(x, y) >= \lfloor \frac{card(K)}{2} + 1$ **then**
      $T_s(x, y) \leftarrow 1$
    else
      $T_s(x, y) \leftarrow 0$
    end if
  end for
  Output T_s

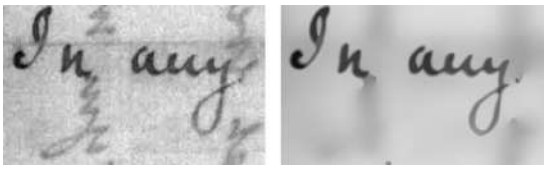


Fig. 5. Strengthening foreground coherence characters after 30 iterations.

Algorithm 2 Automatic $\lambda_{\pm}$ thresholding.

Require: $\lambda_{\pm}, T_s$
  for each pixel (x, y) in $\lambda_{\pm}$ **do**
    if $T_s(x, y) == 1$ **then**
      $\lambda_{\pm} \leftarrow 1$
    end if
  end for
  Output $\lambda_{\pm}$

5. Proposed historical document image compression scheme

Due to the artifacts from which they suffer, archiving the digitized document images without pretreatments needs a huge capacity of storage since each degradation mark, especially bleed-through edges, increases considerably the entropy. The motivation behind this study is then to produce an efficient compression scheme able to detect artifacts within historical manuscript to preserve foreground edges through slight or lossless compression and compress massively around degradation marks.

Roughly speaking, we propose to drive the compression scheme by the thresholding algorithm described previously. Indeed, before performing the encoding strategy, foreground text must be well localized to be slightly compressed and retrieved properly without losing on the readability of the document. In addition, knowing foreground edges positions, one can compress massively elsewhere without undue risk. When decoding, we perform a PDE-based interpolation to retrieve the missed information.

We believe that, due to the homogeneous aspect of the background observed on a great number of old historical documents, edges locations, their values and values of their adjacent pixels are widely sufficient to express old manuscripts. Thereby, through the encoding step within the proposed compression scheme, one proposes to safeguard and compress only the foreground strokes positions, their values and values of their adjacent pixels extracted from the enhanced document image given by Algorithm 1. When decompressing, we propose to perform in few iterations the Laplace heat equation (e.g. homogeneous diffusion), weighted by the binary version of the document, to preserve foreground edges from deformations and interpolate values of unsaved pixels within the encoding step. Then retrieving gradually the enhanced version of the historical document image.

5.1. Encoding scheme

The proposed automatic enhancement and compression approaches are linked. The proposed encoding technique starts by performing the Howe binarization algorithm [16] on the restored image through the proposed automatic enhancement algorithm. To our best knowledge, this binarization method remains the best historical document binarization algorithm as shown in several binarization survey such as in [28]. Furthermore, this algorithm has been crowned "best binarization algorithm" in DIBCO 2012 competition [29]. The resulted binary image through Howe algorithm is denoted u_{b_0}. Afterward, **and** logic operator is performed between T_s the unique output given by Algorithm 1 and u_{b_0} so as to obtain u_b, the best representative binary version of u (see Fig. 6).

The whole encoding process is summarized as below:

Step.01 Compressing u_b through the lossless JB2 algorithm, a variation of JBIG2 [5]. Our motivations behind using lossless encoder is to preserve the readability of the document whatever the compression rate.

Step.02 Extraction of active pixels colors intensities in u_b and their adjacent pixels in a well defined neighborhood. The whole extracted values are stored in their original position within the image u_v as shown in Fig. 6. This image is the key of the proposed decoding technique.

Step.03 Since the human visual system is much more sensitive to variations in brightness than color [30], one can transform u_v from RGB to YCrCb and downsamples the Cr and Cb components with rate 1:4.

Step.04 Sweeping and storing, in a one dimension vector, colors of active pixels along their appearing order in u_v. Since colors vary gradually along edges, safeguarding these values in the order of their occurrence optimizes certainly the entropy coding performance of the quantized values.

Step.05 Quantize the obtained vector using the histogram quantization technique. The quantization parameter is fixed by the user.

Step.06 Compressing the quantized vector using the entropy-based PaQ algorithm. PaQ is the winner of the unique loss-less compression competition called Hutter. In addition, this algorithm has been used in several effective compression schemes, such as in [31]. For our purpose, we perform PaQ8m,[2] which is an open source (GPL) file compressor designed for 32 bit architectures.

The obtained bitstream could be summarized as follows:

1. **Header:**
   (a) **Edges indexes:**
      i. Quantization dictionary length (1 Byte).
      ii. Dictionary (Variable).
      iii. Length of the Cr, Cb channels before 1:4 downsampling (1 Byte).
   (b) **Neighborhood indexes:**
      i. Quantization dictionary length (1 Byte).
      ii. Dictionary (Variable).
      iii. Length of the Cr, Cb channels before 1:4 downsampling (1 Byte).
      iv. Considered neighborhood scale (1 Byte).
2. **Main data:**
   (a) JB2 data: the binary version of the document.
   (b) PaQ data: the color nuances of the document.

5.2. Decoding scheme

Within National Archives laboratories, restoring damaged or incoherent portions of historical images is a common and daily practice. This practice is also known as inpainting or retouching [32, 33]. In the state-of-the-art, some authors has performed inpainting techniques to deal with several issues such as retrieving missed information within decompression process [8] or restoring degradation caused by the latter.

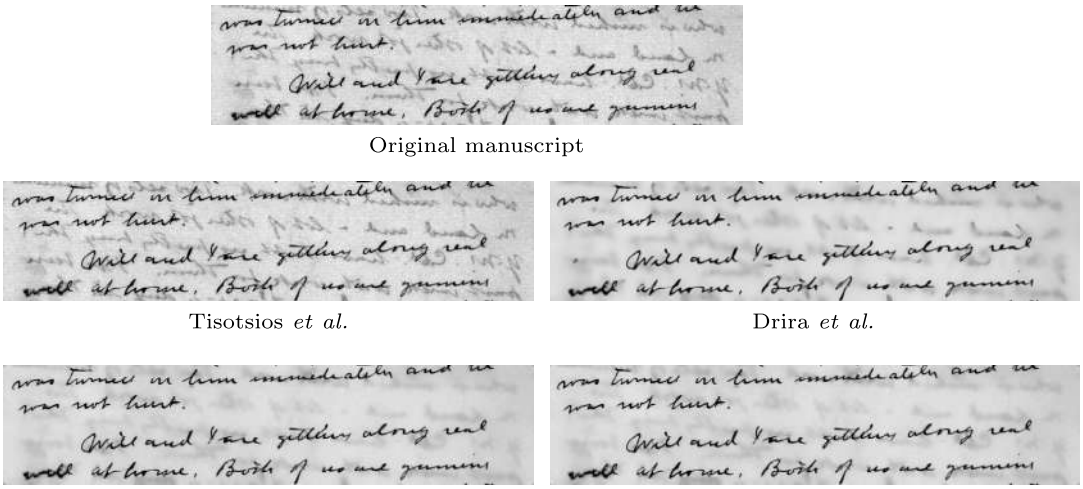
Obviously, when facing strong degradation, automatic inpainting algorithms could never reach manual restoration performance. However, if the missed regions are smooth and homogeneous, a PDE-based interpolation technique applied in a correct way could yield good results if the degradation did not affect foreground structural information.

To this purpose, we start by separating form the encoded file the bilevel image u_b and the quantized vector compressed with PaQ, then combining both of these information to retrieve u_v. So far, a PDE analysis has been applied to enhance the foreground edges coherence and eliminating bleed-through within historical document, so as to improve document readability by eliminating some embarrassing artifacts. However, herein, instead of using diffusion to eliminate useless information, one proposes to perform diffusion to restore missed information within the reconstructed image u_v. In other words, we propose to interpolate, iteration-

[2] https://github.com/JohannesBuchner/paq/tree/master/paq8m.

Original manuscript | u_b: Binarized manuscript | u_v: Color of edges and their adjacent pixels

Fig. 6. Encoding scheme – brief example.



Original manuscript

Tisotsios *et al.* | Drira *et al.*

Proposed ECM approach | Proposed EECM approach

Fig. 7. Historical manuscript extracted from DIBCO databases.

by-iteration, missed pixels in u_v using a diffusion filter. Indeed, performing such diffusion technique permits to preserve boundaries defined by u_b while diffusing the colors of their neighborhood gradually along the image to keep the smooth aspect characterizing historical documents. Thereby, we propose to drive the decoding scheme by the heat equation weighted by the boundary matrix u_b. The proposed decoding diffusion system is presented below:

$$\begin{cases} u_r = \mathrm{div}(u_b \nabla(u_r)) \\ u_0 = u_v \end{cases} \tag{13}$$

6. Experiments and discussion

In this section, the obtained results through the proposed document image enhancement and compression approaches are presented, assessed to illustrate performance against the most effective methods in the state-of-the-art and discussed. The considered dataset consists of the whole historical (e.g. machine-written and handwritten) document images suffering from the bleed-through artifact within DIBCO databases.[3] The advantage given by the DIBCO databases consists of the availability of the ground though images so as an objective assessment could be explicitly performed. In addition to DIBCO, the constructed dataset is consolidated with 8 historical documents collected from the web, such as some historical manuscripts extracted from the Word Digital Library database.[4]

Herein, we start by comparing subjectively the performance of the proposed manual-tuning and automatic enhancement algorithms against the most effective automatic and manual-tuning enhancement methods in the state-of-the-art. The fair objective evaluation is performed thereafter thanks to some classical and well known readability measures. Thereafter, the proposed compression scheme is compared with DjVu [2] the most accurate scanned document compression encoder, the well known JPEG and JPEG2000 standards to assess the proposal with the most used image compression techniques and the most recent compression standard namely JPEG-XR.[5] Finally, to proof the necessity to apply a join compression–enhancement technique instead of a simple sequential application of compression and enhancement methods to deal with historical document archiving issue, the proposed joint method is objectively evaluated with the best enhancement method followed by the best natural image and document image compression standards.

[3] http://users.iit.demokritos.gr/~kntir/HDIBCO2014.

[4] http://www.wdl.org.

[5] https://jpeg.org/jpegxr/index.html.

Table 4
Readability, edge sharpness and the visual quality evaluation.

Bin. algorithm	Enhancement technique	Readability		Sharpness	Visual-quality	
		Precision	$F_{Measure}$	$PF_{Measure}$	DRD	PSNR
TSAI	Direct Binarization (Without pretreatment)	0.68	75.73	77.61	25.20	15.06
	Perona and Malik **manual-tuning** app.	0.65	85.01	87.97	17.67	18.05
	Proposed diffusion scheme **manual-tuning** app.	0.84	85.19	87.83	17.77	18.01
	Weickert **automatic-tuning** app.	0.65	72.79	74.59	26.55	14.15
	Tsiotsios et al. **automatic-tuning** app.	0.80	81.81	84.77	20.11	16.98
	Drira et al. **automatic-tuning** app.	0.80	83.13	85.19	18.79	17.23
	Proposed ECM **automatic-tuning** app.	0.82	84.27	86.71	18.18	17.75
	Proposed EECM **automatic-tuning** app.	0.82	84.17	86.52	18.39	17.67
OTSU	Direct Binarization (Without pretreatment)	0.61	72.68	73.66	21.15	13.76
	Perona and Malik **manual-tuning** app.	0.71	79.29	80.95	14.63	15.56
	Proposed diffusion scheme **manual-tuning** app.	0.74	81.46	82.66	12.82	16.41
	Weickert **automatic-tuning** app.	0.59	70.64	71.67	21.60	13.26
	Tsiotsios et al. **automatic-tuning** app.	0.67	76.61	78.05	17.01	14.96
	Drira et al. **automatic-tuning** app.	0.70	79.05	79.89	14.39	15.56
	Proposed ECM **automatic-tuning** app.	0.74	81.48	82.56	12.18	16.31
	Proposed EECM **automatic-tuning** app.	0.74	81.24	82.28	12.33	16.16
HOWE	Direct Binarization (Without pretreatment)	0.90	92.66	94.13	3.39	21.12
	Perona and Malik **manual-tuning** app.	0.87	88.09	94.67	2.73	20.10
	Proposed diffusion scheme **manual-tuning** app.	0.94	94.67	96.28	1.72	21.67
	Weickert **automatic-tuning** app.	0.87	88.09	90.15	4.04	18.32
	Tsiotsios et al. **automatic-tuning** app.	0.92	93.03	94.53	2.73	20.98
	Drira et al. **automatic-tuning** app.	0.92	92.33	93.78	2.53	20.26
	Proposed ECM **automatic-tuning** app.	0.94	93.82	95.52	2.20	21.39
	Proposed EECM **automatic-tuning** app.	0.94	93.49	95.14	2.43	21.22

6.1. Enhancement evaluation

6.1.1. Objective evaluation criteria

The proposed manual-tuning formalism and its automated versions are evaluated herein by means of objectives judgments through various commonly used criteria for document image enhancement evaluation [29,34]. This objective evaluation is performed by means of some binarization measures to quantify the readability of the outputted documents. We propose to compare our proposed method against the most recent automatic and manual tuning schemes in the literature, characterized by their ability to preserve the look and the feel of old manuscripts suffering from bleed-through. The evaluated methods are:

- The Tsiotsios et al. automatic-tuning approach [25].
- The Drira et al. automatic-tuning approach [15,21].
- The Weickert automatic-tuning approach [18].
- The Perona and Malik manual-tuning approach [17].
- The proposed manual-tuning diffusion scheme.
- The proposed automatic-tuning technique based on ECM.
- The proposed automatic-tuning technique based on EECM.

As noticed earlier the proposed manual-tuning diffusion scheme and the Perona and Malik [17] approaches are not blind and require to perform an arduous tuning to set the conductance thresholds. In addition, the Drira et al. [15] technique is not fully-blind since the number of iterations plays a key role to attain acceptable performances.

Thus, to perform a fair evaluation study and optimize performances of the evaluated approaches, K and $K_{\pm}$ have been tuned and the number of iteration has been fixed at 150 in both the proposed manual diffusion scheme and the Perona and Malik approaches to optimize the $F_{Measure}$ within each document image treated with those manual-tuning algorithms; while we apply the Drira et al. technique with various iterations and the image obtaining the best $F_{Measure}$ is chosen for the evaluation.

Thereafter, three classical and well known thresholding algorithms are applied to binarize the enhanced documents by the considered enhancement methods. The thresholding algorithms considered in this study are:

1. Otsu method: The widely used binarization technique proposed in [35].
2. Tsai method: One of the most used binarization algorithm introduced in [36].
3. Howe method: To our best knowledge, this binarization algorithm is the best in document image binarization [16].

The comparison is performed by using a set of some measures used to evaluate binarization techniques in DIBCOs competitions. The objective tools are:

1. The $F_{Measure}$, which is most used measure to evaluate the document image binarization algorithms: This measure establish a score in the range of [0, 100] according to the number of true positive, true negative and false positive pixels. The larger is the $F_{Measure}$, the best is the readability of the document.

$$F_{Measure} = 100 * \frac{2 * PR * RC}{PR + RC} \tag{14}$$

were $RC = TP/(TP + FP)$, $PR = TP/(TP + FN)$ and TP, FP and FN represent respectively the number of, the true positive, the false positive, and the false negative pixels in the binarized image.
2. The Distance Reciprocal Distortion (DRD[6]) score [37]: This score correlates with the human visual perception and used to quantify the visual distortion in binary document images. Since 2012, this strong measure is performed to evaluate the binarization algorithms in DIBCO competitions. The low is the DRD, the best is the visual quality and the readability of the document.
3. The $PF_{Measure}$: This score expresses the readability of the document across its ground-through skeleton. The idea consists

[6] This measure reflects the readability.

Fig. 8. Old manuscript extracted from the WDL database – handwritten.

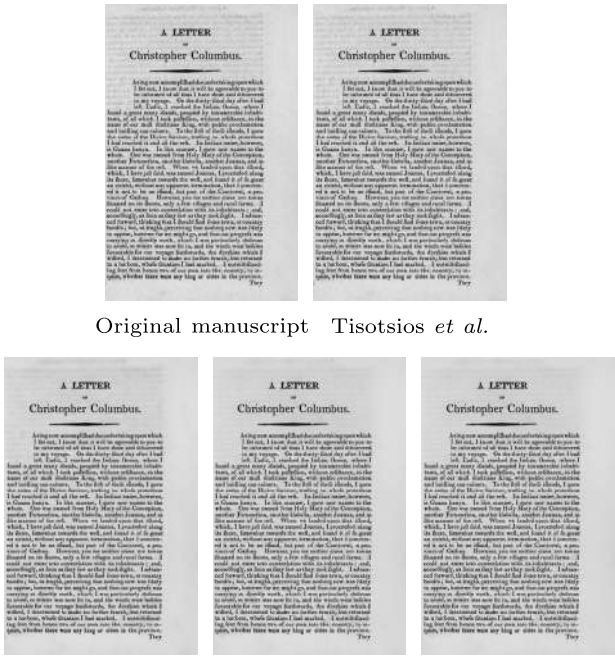


Fig. 9. Old manuscript extracted from the WDL database – machine written.

in estimating the sharpness by comparing the skeleton of the document with the skeleton ground-truth as explained in [38]. The $PF_{Measure}$ is computed as well as the $F_{measure}$ while substituting the RC score by the $Pseudo_{RC}$ as defined in Eq. (15).

$$Pseudo_{RC} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} u_{bin}(i,j)u_{sg}(i,j)}{\sum_{i=1}^{M}\sum_{j=1}^{N} u_{sg}(i,j)} \tag{15}$$

where u_{bin} is the binarized document image and u_{sg} is its skeletonized version.

6.1.2. Results and discussion

It is clear from the presented in Table 4 that the proposed automation techniques outperform the whole others automatic-tuning techniques in the state-of-the-art. However, even if the Drira et al. approach shows acceptable performances especially in the bleed-through removal task as seen in Fig. 7, 8 and 9; as noticed previously, the most embarrassing drawback on which this approach suffers consists in the shutoff criteria. In fact, when conducting this evaluation, the Drira et al. iterative process is stopped manually when the $F_{Measure}$ reaches its maximum. Fig. 10 shows performance given by the proposed ECM enhancement scheme against the Drira et al. scheme under a wide iteration range.

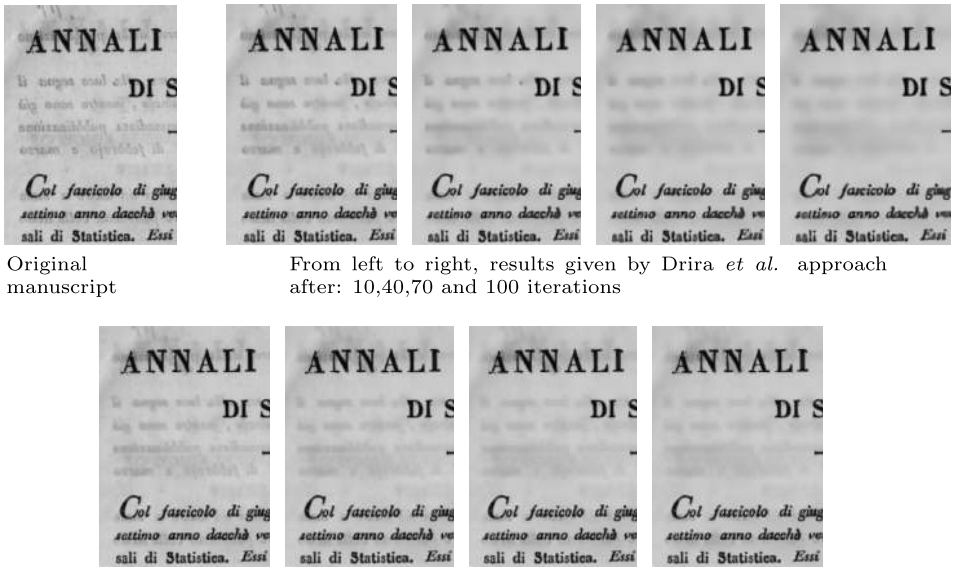
Regarding the visual aspect (see Figs. 7, 8 and 9), one could notice without providing any visual effort that the restored images through the proposed approaches present, for the majority, acceptable visual aspect compared to images obtained from the other methods. Furthermore, the feeling of the processed historical manuscript is literally kept. This is quiet important when dealing with this kind of images. The obtained Distance Reciprocal Distortion (DRD) [37] score could objectively confirm these observations (see Table 4). Indeed, since the proper correlation of the DRD with the human visual perception as explained in [29], the DRD score is usually used to quantify the visual distortion in binary document images as done in several DIBCOs competitions. The lower is the DRD, better are performances.

From the $PF_{Measure}$ column in Table 4 we can notice that the automation given in this study succeed in safeguarding, in certain cases improving, the strength points of the proposed manual-tuning technique, which consists on the Precision score and $PF_{Measure}$. The Precision measure increases when the degree of true positive pixels is important and false positive pixels tends to zero. Indeed, the proposed diffusion formalism, on which the proposed automatics techniques are based, improves the contrast while simultaneously strengthens the coherence of each pixel with its neighborhood. This refines each character shape iteration-by-iteration, consequently both precision and edge sharpness are improved.

The sharpness could be expressed through the $PF_{Measure}$. The $PF_{Measure}$ expresses the readability of the document across its ground-through skeleton. This is motivated by the fact that the silhouette is unique on each character within the skeleton image. Then, as explained in [38], comparing the skeleton version of the document with the skeleton ground-truth could be more accurate to estimate the edges sharpness instead of using a classic binarized-document/ground-truth-document comparison.

Nevertheless, tested with Howe algorithm, performance of the automatic-tuning schemes are quite interesting but does not attain same performance as the manual-tuning version, while the manual-tuning performances are beaten under Tsai and Otsu algorithms. In fact, it is important to point out that the proposed manual approach needs a significant effort for tuning the $K_{\pm}$ coefficients to reach such results, while this tuning is hard and tedious when dealing with a considerable number of degraded documents. Then, providing such results with a simple-to-implement and fully-automatic techniques could be so far considered as an achievement.

Finally, we notice from Table 4 that the proposed automatic-tuning approaches improve the performances of the used binarization algorithms. Indeed, even if Howe [16] algorithm is considered so far as the best thresholding algorithm for document image, joining this algorithm with one of the proposed enhancement methods improve performances as shown through the objective assessment.



Original manuscript

From left to right, results given by Drira *et al.* approach after: 10,40,70 and 100 iterations

From left to right, results given by the proposed ECM approach after: 10,40,70 and 100 iterations

Fig. 10. Proposed approach vs. Drira et al. approach – iteration-by-iteration comparison.

Furthermore, the performances of the classic methods Otsu and Tsai are significantly improved when applied on enhanced documents by the proposed techniques. Then associating the proposed enhancement techniques with these thresholding algorithms could significantly revitalize their use since their unequal rapidity and implementation simplicity compared to similar methods in the state-of-the-art [39].

6.2. Compression evaluation

Compression performances of the proposed enhancement-compression method is evaluated herein by means of some objective readability measures associated with perceptual subjective observations. Our proposal is fairly compared with four robust and well-known compression standards:

1. DjVu compression encoder [2]: Undoubtedly, DjVu is the most accurate compression scheme for scanned documents [6], which explains its used in the Wikipedia book digitization project Wikisource.[7] In this study, the DjVu encoder is performed thanks to the open source software DjVuLibre-3.5.[8]
2. JPEG compression standard [19]: So far, JPEG remains the most used image compression standard. Regardless of its efficiency, this is mainly due to its implementation simplicity.
3. JPEG2000 compression standard [20]: JPEG2000 is one of the most efficient lossy image compression standard.
4. JPEG-XR compression standard[9]: JPEG-XR is the most recent technology in image compression.

6.2.1. Proposal versus image compression standards

To start, the whole degraded historical documents included in the constructed dataset are compressed, in various bitrates, using DjVu, JPEG, JPEG2000, JPEG-XR standards and the proposed compression scheme. Thereafter, as done in the enhancement assessment, the obtained compressed documents are binarized through the Howe thresholding algorithm [16] and evaluated in their readability accuracy thanks to the most efficient readability metrics in the field. This process is done on each bitrate as shown in Fig. 11.

One can clearly notice from the obtained results (Fig. 11 and 12) that the proposed compression method greatly outperforms, under the whole readability measures, the whole compared standards. Indeed, the proposed compression technique focuses mainly on preserving the readability of the document whatever the compression rate. This is done by encoding its best binarized version using a lossless compression algorithm, which is objectively demonstrated through the obtained $F_{Measure}$ and $PF_{Measure}$. In addition to the readability improvement given by the proposed framework, the look and the feel of the compressed manuscripts are preserved comparing to the others standards. The obtained DRD scores are behind this observation (see Fig. 11).

It is clear from the obtained results that tackling the historical document compression issue with commons image compression standards seems to be inappropriate. This is due to the fact that JPEG and JPEG XR/2000 deal with pixels without taking into account their local characteristics, so as foreground pixels are compressed as well as artifacts; consequently, the readability is doubtless affected when one tries to reduce the compression bitrate. This supports the need to use a specific compression technique as we propose in this study.

Finally, it is important to notice that, under a logical context, the more we compress, the more the readability of the document is affected. However, paradoxically, a slight amelioration of the readability is observed when we are using the natural image standards (JPEG or JPEG XR/2000) at very high bitrates. Indeed, at this stage, the compression process targets more noise than useful edges, which produces a slight amelioration of the readability (see Fig. 13). On the other hand, since the proposed diffusion process reduces the degradations affecting the image, the amelioration seen in the high bitrates is not perceptible when applying the natural image compression standards into the enhanced document images with the proposal as we can notice in Fig. 13.

[7] http://en.wikipedia.org/wiki/Wikisource.
[8] http://djvu.sourceforge.net/.
[9] https://jpeg.org/jpegxr/index.html.

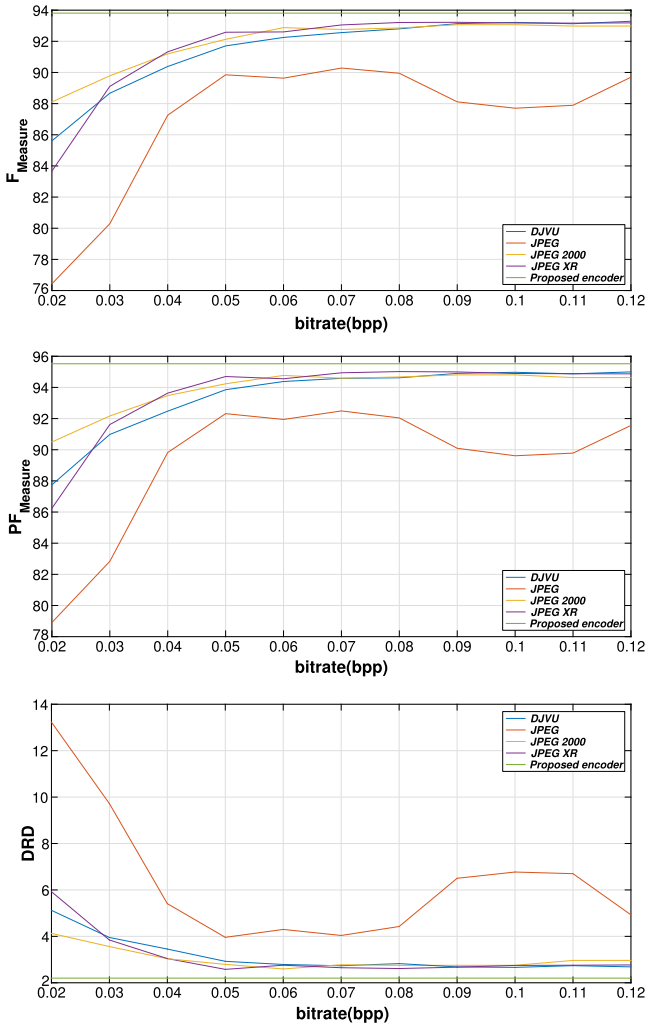


Fig. 11. Evolution of FM, PFM and DRD in DIBCO images compressed by the proposed (green), DjVu approach (blue), JPEG (red), JPEG2000 (yellow) and JPEG XR (violet) in various bitrates. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6.2.2. Join versus combination methods

It is worth to recall that, to improve compression performances, the proposed compression scheme is driven by a restoration technique able to target properly "artifacts"/"useful edges" within historical manuscript. Throughout this work, we assumed that a simple enhancement–compression algorithms combination could not reach same performance as full joint algorithm when it is about compressing historical scanned documents. Especially in low and very low bitrates.

In other words, applying consecutively a robust document image enhancement algorithm followed by a robust image compression standard alters certainly some foreground edges if we try to reach aggressive compression bitrate. In order to underpin or disprove this hypothesis, we propose to combine the proposed enhancement technique ECM with JPEG2000, DjVu and JPEG-XR to provide the most robust sequential enhancement-compression frameworks. Results given by these sequential frameworks are compared with the proposed join framework.

Fig. 13 and 14 shows scores of readability obtained at each bitrate using the evaluated schemes. Undoubtedly, performances of the whole evaluated standards are improved after being associated with an enhancement technique. However, it is clear from the obtained results that is more appropriate to use a join scheme instead of a consecutive scheme when dealing with historical document images especially in low bitarates.

Table 5
Encoding run-time comparison in 0.04 bpp.

Approach	Min. (s)	Max. (s)	Average (s)
Proposed encoder	1.3524	4.6088	2.5467
DjVu	0.6035	1.4192	2.6370
JPEG	0.3824	0.7958	0.5369
JPEG 2000	2.2529	17.5514	6.6045
JPEG–XR	0.6113	2.5691	1.3263

Indeed, even if artifacts are considerably reduced after performing an enhancement algorithm in the consecutive schemes, since the compression encoder does not discriminate foreground/background edges properly, foreground edges are still deformed when performing an aggressive compression. This could be supported objectively through the readability and visual aspect measures used in Fig. 13 and 14, and subjectively in Fig. 12.

Finally, it is worth to notice that before performing such experimentation it was expected that, since DjVu is the most efficient compression technique of scanned documents [6], DjVu would certainly outperform JPEG2000 when dealing with historical documents in the low bitrates. DjVu is based on HMM segmentation algorithm to differentiate edges from artifacts. However, as shown throughout this study, this segmentation has reached its limits against the specific historical document degradations. Moreover, it is well-known that the multi-resolution analysis given by the wavelet transform permits to highlight relevant edges in the image which makes JPEG2000 less sensitive to document degradation. All these features are behind the success of JPEG2000 against DjVu when facing historical document.

6.2.3. Compression complexity – discussion

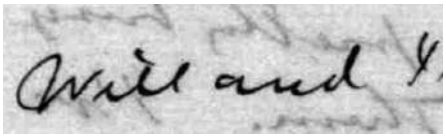
Besides compression efficiency, the encoding complexity and the run time takes an important role. While the complexity depends only on the constructed method, the run time relies on different factors such as the processor frequency, the memory bandwidth or the programming language.

We present in Table 5 a run time evaluation involving the best image compression standards (JPEG, JPEG 2000, JPEG XR, DjVu) and the proposed encoder, the evaluation has been conducted in the Win 8.1 Intel® Core i7 machine at 2.6 GHz. We note that the smallest image in the considered dataset takes 982 × 665 pixels, while the largest image takes 3297 × 1097 pixels.

In terms of the computing time, it seemed obvious that the majority of the commercial standards perform slightly better than the proposal since they are highly optimized. Nevertheless, the proposed encoder shows acceptable run time performance even if this property is highly affected by the iterative filter (see Eq. (6)) and the diffusion automation algorithm (see Algorithm 1). For example, encoding a document image with the proposed process takes an average of 2.6 s and may go up to 4.6 s. Finally, it is important to recall that the digitization and the archiving processes are often performed offline, so as the encoding run time would not be seriously constraining. In addition, the decoding process is nearly instantaneous, which makes the proposed method efficient and places it among the top of archiving solutions.

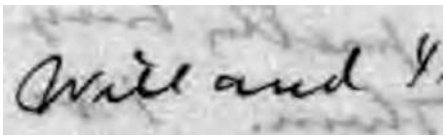
7. Conclusion

In this work, a new diffusion formalism dedicated for document image enhancement and restoration is proposed. Thereafter, two variants have been proposed to automate this diffusion formalism when dealing with the document enhancement task. The

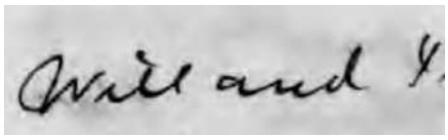


a: Original document

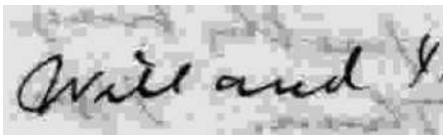
b: Joint enhanced-compressed document by the proposal



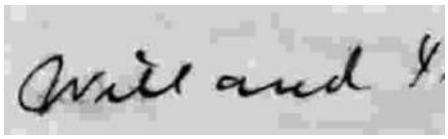
c: DjVu applied on original manuscript



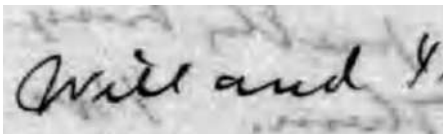
d: DjVu applied on the enhanced manuscript by the proposal



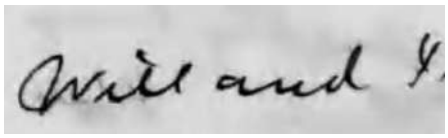
e: JPEG applied on original manuscript



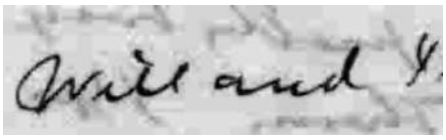
f: JPEG applied on the enhanced manuscript by the proposal



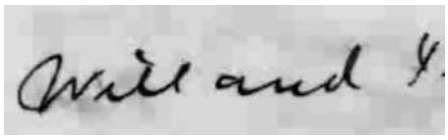
g: JPEG2000 applied on original manuscript



h: JPEG2000 applied on the enhanced manuscript by the proposal



i: JPEG-XR applied on original manuscript



j: JPEG-XR applied on the enhanced manuscript by the proposal

Fig. 12. Performance of the proposal against the most used compression standards at 0.04 bpp.

first method puts the emphasis on the edges contrast estimated through eigenvalues while the second method takes into account the edges coherence in terms of direction to discriminate foreground/artifacts within the degraded manuscript. Afterward, using judiciously one of these enhancement techniques, we develop a compression method totally suitable for historical document images. Indeed, we demonstrate throughout this study that current image compression standards, even those specialized in document images such as DjVu, fail to provide acceptable performances when applied on historical scanned manuscripts due to the important amount of artifacts on which those images are suffering. In addition, we demonstrate that applying an enhancement algorithm to reduce document artifacts before performing the encoding process is not the most suitable solution for archiving scanned document images.

Enhancement and compression performances of the proposed framework have been evaluated on overall DIBCO databases and some historical manuscripts collected from the Word Digital Library database. We have noticed, using some well known readability measures, that both proposed enhancement techniques outperform all the automatic-tuning methods of the state-of-the-art. In addition, we have shown that associating one of the proposed enhancement technique with a binarization algorithm improves considerably its performance. Concerning the compression, we have shown through some document image quality measurements that the proposed compression method greatly outperforms some well known standards such as JPEG, DjVu and JPEG2000/XR. Moreover, using the same document quality measurements, we have highlighted the advantages offered by the proposed joint method against some sequential compression–enhancement combination methods.

However, enhancement and compression performances could be both improved in future studies by incorporating other discriminant characteristics to distinguish foreground strokes in the image such as edges width could boost performances of the proposed scheme and extend the scope of its application.

Finally, it is important to stress that, except the Distance Reciprocal Distortion score [37] which correlates with the human visual perception, the document image enhancement field still does not have a solid quality score on which the readability and the contrast are quantified without dissociation. Furthermore, objective natural image measures could not be used to evaluate the quality of document images. Indeed, the natural image non-learning measures have not been designed to take into account the readability score [40] while the learning measures, such as [41], have not been trained on the degraded document images. In this study, we present interesting theoretical tools, such as the orientation of the structural tensor T and the intensity of the coherence measure ECM (see Eq. (10)), which could be used to develop new perceptual-based objective evaluation technologies dedicated to document images.

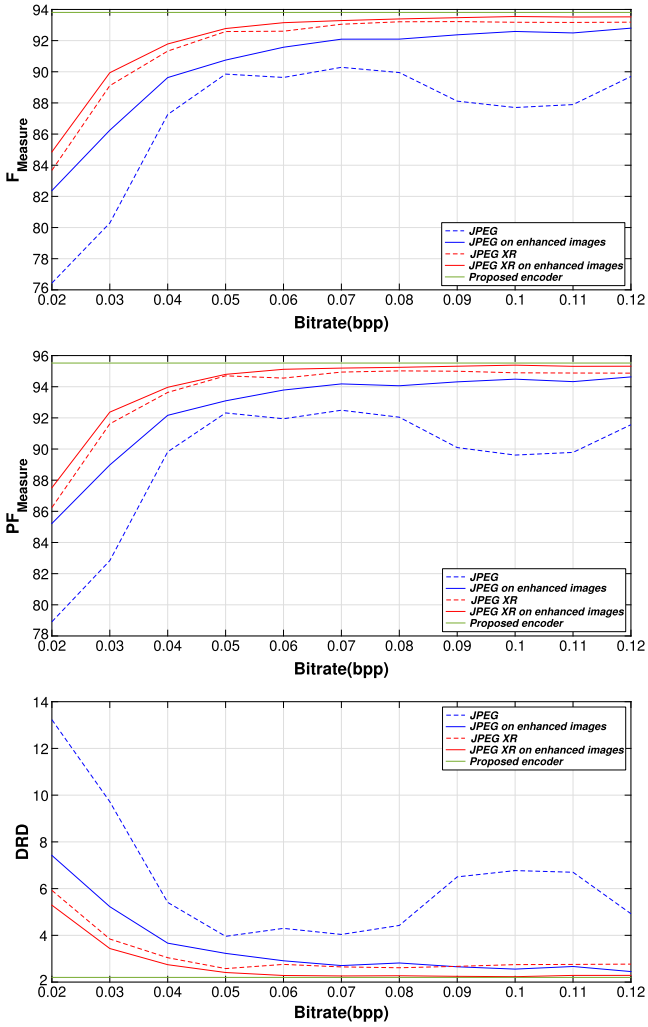


Fig. 13. Readability gain when performing JPEG and JPEG XR on enhanced document images with the ECM approach.

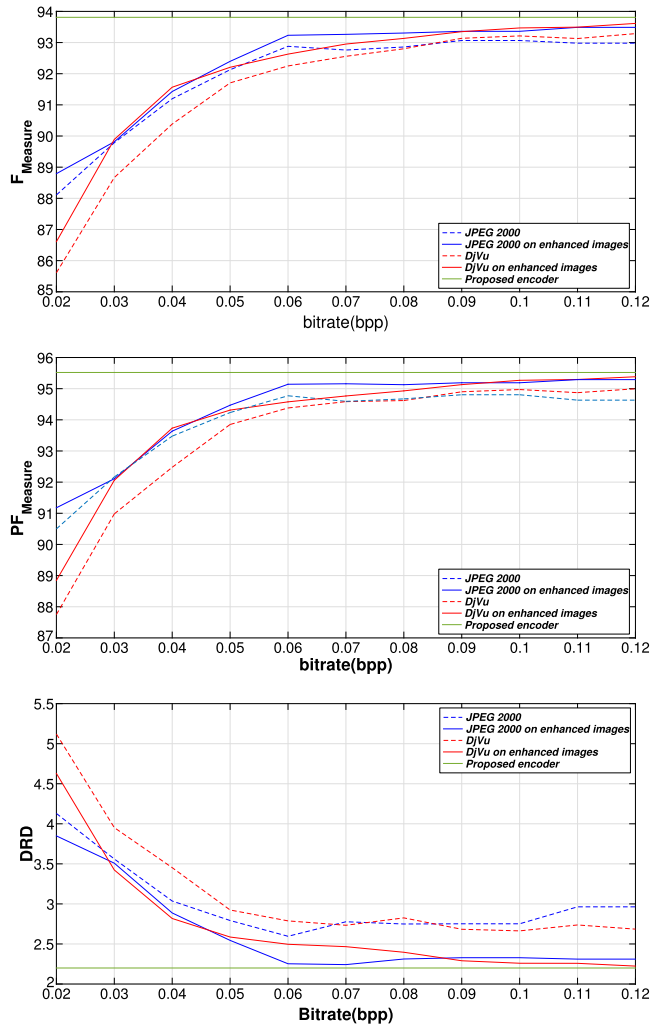


Fig. 14. Readability gain when performing DjVu and JPEG2000 on enhanced document images with the ECM approach.

References

[1] M. Yagoubi, A. Serir, A. Beghdadi, Joint enhancement–compression of handwritten document images through DjVu encoder, J. Vis. Commun. Image Represent. 41 (2016) 324–338, http://dx.doi.org/10.1016/j.jvcir.2016.10.012, http://www.sciencedirect.com/science/article/pii/S1047320316302188.

[2] P. Haffner, P. Howard, P. Simard, Y. Bengio, Y. Lecun, et al., High quality document image compression with DjVu, J. Electron. Imaging 7 (3) (1998) 410–425.

[3] P. Haffner, L. Bottou, Y. Lecun, L. Vincent, A general segmentation scheme for DjVu document compression, in: ISMM, vol. 2, 2002.

[4] R.D. Queiroz, R. Buckley, M. Xu, Mixed raster content (MRC) model for compound image compression, in: Electronic Imaging'99, International Society for Optics and Photonics, 1998, pp. 1106–1117.

[5] P. Howard, F. Kossentini, B. Martins, S. Forchhammer, W. Rucklidge, The emerging JBIG2 standard, IEEE Trans. Circuits Syst. Video Technol. 8 (7) (1998) 838–848.

[6] E. Haneda, C. Bouman, Text segmentation for MRC document compression, IEEE Trans. Image Process. 20 (6) (2011) 1611–1626, http://dx.doi.org/10.1109/TIP.2010.2101611.

[7] I. Galić, J. Weickert, M. Welk, A. Bruhn, A. Belyaev, H.-P. Seidel, Image compression with anisotropic diffusion, J. Math. Imaging Vis. 31 (2–3) (2008) 255–269, http://dx.doi.org/10.1007/s10851-008-0087-0.

[8] M. Mainberger, J. Weickert, Edge-based image compression with homogeneous diffusion, in: Computer Analysis of Images and Patterns, in: Lecture Notes in Computer Science, vol. 5702, Springer, Berlin, Heidelberg, 2009, pp. 476–483.

[9] M. Yagoubi, A. Serir, A. Beghdadi, Blind document image enhancement based on diffusion process, in: 5th European Workshop on Visual Information Processing, EUVIP, 2014, pp. 1–6.

[10] M. Yagoubi, A. Serir, A. Beghdadi, A new automatic framework for document image enhancement process based on anisotropic diffusion, in: 2015 13th International Conference on Document Analysis and Recognition (ICDAR), 2015, pp. 1126–1130.

[11] A. Tonazzini, Color space transformations for analysis and enhancement of ancient degraded manuscripts, Pattern Recognit. Image Anal. 20 (3) (2010) 404–417, http://dx.doi.org/10.1134/S105466181003017X.

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

[13] K. Gu, G. Zhai, W. Lin, M. Liu, The analysis of image contrast: from quality assessment to automatic enhancement, IEEE Trans. Cybern. 46 (1) (2016) 284–297.

[14] K. Gu, G. Zhai, X. Yang, W. Zhang, C.W. Chen, Automatic contrast enhancement technology with saliency preservation, IEEE Trans. Circuits Syst. Video Technol. 25 (9) (2015) 1480–1494.

[15] F. Drira, F. LeBourgeois, H. Emptoz, Document images restoration by a new tensor based diffusion process: application to the recognition of old printed documents, in: 10th International Conference on Document Analysis and Recognition, 2009, ICDAR'09, IEEE, 2009, pp. 321–325.

[16] N.R. Howe, A laplacian energy for document binarization, in: International Conference on Document Analysis and Recognition, ICDAR, IEEE, 2011, pp. 6–10.

[17] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Mach. Intell. 12 (7) (1990) 629–639, http://dx.doi.org/10.1109/34.56205.

[18] J. Weickert, Coherence-enhancing diffusion of colour images, Image Vis. Comput. 17 (3) (1999) 201–212.

[19] J. Mitchell, Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines, ITU-T Recommendation T 81.

[20] M.W. Marcellin, M.J. Gormish, A. Bilgin, M.P. Boliek, An overview of JPEG-2000, in: Proceedings of Data Compression Conference, DCC, IEEE, 2000, pp. 523–541.

[21] F. Drira, F. LeBourgeois, H. Emptoz, A new PDE-based approach for singularity-preserving regularization: application to degraded characters restoration, Int. J. Doc. Anal. Recognit. 15 (3) (2012) 183–212.
[22] N. Sochen, R. Kimmel, R. Malladi, A general framework for low level vision, IEEE Trans. Image Process. 7 (3) (1998) 310–318, http://dx.doi.org/10.1109/83.661181.
[23] D. Tschumperle, R. Deriche, Vector-valued image regularization with PDEs: a common framework for different applications, IEEE Trans. Pattern Anal. Mach. Intell. 27 (4) (2005) 506–517.
[24] J. Monteil, A. Beghdadi, A new interpretation and improvement of the non-linear anisotropic diffusion for image enhancement, IEEE Trans. Pattern Anal. Mach. Intell. 21 (9) (1999) 940–946.
[25] C. Tsiotsios, M. Petrou, On the choice of the parameters for anisotropic diffusion in image processing, Pattern Recognit. 46 (5) (2013) 1369–1381.
[26] M. Petrou, C. Petrou, Image Processing: The Fundamentals, John Wiley & Sons, 2010.
[27] O. Le Meur, J. Gautier, C. Guillemot, Examplar-based inpainting based on local geometry, in: 18th IEEE International Conference on Image Processing, ICIP, 2011, pp. 3401–3404.
[28] B. Su, S. Lu, C.L. Tan, Robust document image binarization technique for degraded document images, IEEE Trans. Image Process. 22 (4) (2013) 1408–1417, http://dx.doi.org/10.1109/TIP.2012.2231089.
[29] I. Pratikakis, B. Gatos, K. Ntirogiannis, ICFHR 2012 competition on handwritten document image binarization (H-DIBCO 2012), in: 2012 International Conference on Frontiers in Handwriting Recognition, CFHR, IEEE, 2012, pp. 817–822.
[30] S. Winkler, M. Kunt, C. van den Branden Lambrecht, Vision and video: models and applications, in: C. van den Branden Lambrecht (Ed.), Vision Models and Applications to Image and Video Processing, Springer, US, 2001, pp. 201–229.
[31] M. Mainberger, J. Weickert, Edge-based image compression with homogeneous diffusion, in: Computer Analysis of Images and Patterns, Springer, 2009, pp. 476–483.
[32] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, ACM Press/Addison-Wesley Publishing Co., 2000, pp. 417–424.
[33] T. Dang, A. Beghdadi, M.-C. Larabi, A perceptual image completion approach based on a hierarchical optimization scheme, in: Image Restoration and Enhancement: Recent Advances and Applications, Signal Process. 103 (2014) 127–141, http://dx.doi.org/10.1016/j.sigpro.2013.11.036, http://www.sciencedirect.com/science/article/pii/S0165168413004726.
[34] I. Pratikakis, B. Gatos, K. Ntirogiannis, ICDAR 2013 document image binarization contest, DIBCO 2013, in: 12th International Conference on Document Analysis and Recognition, ICDAR, IEEE, 2013, pp. 1471–1476.
[35] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66, http://dx.doi.org/10.1109/TSMC.1979.4310076.
[36] C. Tsai, H. Lee, Binarization of color document images via luminance and saturation color features, IEEE Trans. Image Process. 11 (4) (2002) 434–451, http://dx.doi.org/10.1109/TIP.2002.999677.
[37] H. Lu, A. Kot, C. Alex, Y. Shi, Distance-reciprocal distortion measure for binary document images, IEEE Signal Process. Lett. 11 (2) (2004) 228–231.
[38] K. Ntirogiannis, B. Gatos, I. Pratikakis, An objective evaluation methodology for document image binarization techniques, in: The Eighth IAPR International Workshop on Document Analysis Systems, DAS'08, IEEE, 2008, pp. 217–224.
[39] T.R. Singh, S. Roy, O.I. Singh, T. Sinam, K. Singh, et al., A new local adaptive thresholding technique in binarization, arXiv preprint, arXiv:1201.5227.
[40] K. Gu, G. Zhai, W. Lin, X. Yang, W. Zhang, No-reference image sharpness assessment in autoregressive parameter space, IEEE Trans. Image Process. 24 (10) (2015) 3218–3231.
[41] K. Gu, G. Zhai, W. Lin, X. Yang, W. Zhang, Learning a blind quality evaluation engine of screen content images, Neurocomputing 196 (2016) 140–149.

Mohamed Riad Yagoubi received the B.S. degree in 2008, the M.S. degree in Intelligent Systems in 2010 and the Ph.D. degree in Information Processing in 2017, all from the Telecommunication department in the EE & CS faculty of USTHB. From 2010 to 2014, he worked as a developer–consultant for many software development companies. He is now involved in several National R&D project in software development and image processing. His research interests include: image compression, image enhancement, PDE analysis for image processing and video analysis.

Amina Serir is a Professor at the University of Sciences and Technology, USTHB and the Director of the Image Processing and Radiation Lab, LTIR, where she does all her research in image and video processing. She received her Engineering degree in Electrical Engineering from State High School, Polytechnic School, EMP, Algiers in 1985. She had worked in the design office of Algeria Airlines. Then, she joined the University of Sciences and Technology and received in 2002, her Ph.D. degree in Image Processing. She contributed to the creation of a High School of Technology (2010–2011). She leads several national research projects through the integration of biometrics in smart cards for bank payment or in intelligent video. Her research interests include information processing systems in particular compression and information representation and analysis.

Azeddine Beghdadi is Full Professor at the University of Paris 13 (Institut Galilée) Sorbonne Paris Cite since 2000 and the director of L2TI laboratory. He received Maitrise in Physics and Diplôme d'Etudes Approfondies in Optics and Signal Processing from University Orsay-Paris XI (Equivalent: Masters of Sciences) in June 1982 and June 1983 respectively and the Ph.D. in Physics (Specialism: Optics and Signal Processing) from University Paris 6 in June 1986. Dr. Beghdadi worked at different places during his Ph.D. thesis, including "Laboratoire d'Optique des Solides" (University of Paris 6) and "Groupe d'Analyse d'Images Biomédicales" (CNAM Paris). From 1987 to 1989, he has been "Assistant Associé" (Assistant Professor) at University Paris 13. During the period 1987–1998, he was with LPMTM CNRS Laboratory working on Scanning Electron Microscope (SEM) materials image analysis. He published over than 220 international refereed scientific papers. He is a founding member of the L2TI laboratory. His research interests include image quality enhancement and assessment, image and video compression, bio-inspired models for image analysis and processing, and physics-based image analysis. Dr. Beghdadi is the founder and Co-Chair of the European Workshop on Visual Information Processing (EUVIP). He served as conference chair of ISSPA 2003, technical chair of ISSPA 2005 and Track chair of ISSPA 2007. He is founding member of the International Summer School on Signal Processing and its Applications. He also served as session organizer and a member of the organizing and technical committees for many IEEE conferences. Dr. Beghdadi is a member of the editorial board of "Signal Processing: Image Communication" journal Elsevier and member of the editorial board of the Scientific World Journal "Signal Processing", Hindawi Publisher. Dr. Beghdadi is a Senior member of IEEE, EURASIP member and member of the IEEE-MMTC.