Joint Enhancement-Compression of Handwritten Document Images through DjVu Encoder

Mohamed Riad YAGOUBI, Amina SERIR and Azeddine BEGHDADI*
USTHB - Department of Telecommunications, LTIR, Algiers, Algeria
*Institut Galilée, L2TI, Paris, France

Abstract

Despite all the technological developments in image acquisition and processing, preserving old documents and other data of historical interest is still a very challenging issue. Indeed, these documents are often proned to several types of artifacts affecting their readability. Furthermore, due to the considerable information considered in such media, reducing the size of the digitized documents is another challenging problem. We believe that driving lost information onto artifacts could bring an elegant solution to this issue. In this paper, a novel approach joining compression-enhancement of single-side handwritten document is proposed. This approach presents a novel foreground/background segmentation algorithm, using both directional and contrast features to highlight the original information. This pre-treatment step is embedded into DjVu encoder, which is commonly used in national archives and libraries frameworks, to drive the compression rate. Both objective evaluation and perceptual judgment demonstrate the efficiency of the proposed scheme on the whole DIBCO datasets.

Keywords: Document image enhancement, image compression, image segmentation, joint enhancement-compression, Wave atoms, directional filter, modified DjVu

1. Introduction

The motivation of this work comes from the need of national archives to digitize a large amount of historical documents, including a considerable number of single-sided handwritten documents and requiring a particular attention regarding their specificity. In this special context, the design of an optimal data digitization and processing framework raises some challenging problems. The most important one, to address here, is how to restore and compress these degraded document

images without affecting their relevant content.

One of the most studied problems is the artifact removing from given Degraded HandWritten (DHW) documents. In this context, we define document restoration as the process by which the original side (the recto-side) is extracted from the degraded image. Moreover, the digitization of books or articles requires a huge capacity of storage. Therefore, using lossy compression approaches is unavoidable. This paper presents a new method of joint compression-enhancement of handwritten signal-sided document images. The main idea behind the consideration of restoration and compression is to guarantee an efficient compression while sustaining a high quality of the compressed images. In the sequel of this paper, as in the literature, "restoration" and "enhancement" terms are used interchangeably, since they converge literary toward the same purpose.

Handwritten document images are affected by several types of degradations. Among them, we consider: ink/coffee stains, illumination variance and the bleed-through, due to sipping of ink through pages. The result is that handwritten characters from the reverse side appear as noise on the front side and even interfere with the front side characters. These artifacts are mainly caused by the poor quality of the ink or scanning materials, external sources such as heat or humidity, or improper handling of the documents when scanning. Consequently, the readability of these documents is considerably affected. In addition, reducing the size of these digitized documents is another challenging problem. Our purpose consists then in removing these artifacts, to improve the document quality, while compressing the file size to be stored. The objective, here, is to achieve both a good compression rate and a high image quality as much as possible.

The proposed method is based on a segmentation algorithm taking into account both contrast and ~~directionaly~~ features in order to decorrelate the information to be preserved. Thereafter, we incorporate the proposed segmentation algorithm into DjVu [1] compression encoder, that has proved its efficiency for document image compression [2], to restore and compress simultaneously the DHW document images.

This study brings several important contributions to the image processing field. Indeed, we developed ~~the first technique~~ ~~that joint compression and enhancement~~ for single side DHW document images. In addition, in the restoration step, we have proposed the first automatic approach using both contrast and directional features of strokes to highlight the front side information.

The directional feature is analyzed by using an original multi-directional filter based on wave atoms transform ~~exclusively built in this study~~. This filter could be used in various image processing problems. Furthermore, this filter is simple to implement, fast and exhibits interesting ~~proprieties~~.

This paper starts with an overview of the prior research in document image restoration, document compression and the existing joint enhancement-compression approaches in section 2. In section 3, we introduce and describe in details the proposed approach. The performance evaluation of the proposed method through experiments conducted on all DIBCO-DHW document images databases, is presented in section 4. Section 5 is devoted to the conclusion and possible direction of future works.

2. Previous related works

2.1. Existing document images restoration techniques - a brief survey

Several document image restoration techniques have been proposed in the literature and several models were proposed to classify them [3]. Before starting a brief survey on document restoration techniques, it is important to stress that restoration have two distinct aims in the literature. Improving the visual quality of document images by preserving the look and the feel while reducing artifacts of these documents on the one hand, which is called enhancement. On the other hand, restoration is used to improve readability while associated with a local or a global thresholding, this is called binarization.

Binarization remains the most studied [4]. It consists of classifying pixels of the document into two classes, original text and artifacts. A detailed study focusing on the most effective binarization algorithms is provided in [3].

Enhancement process consists on reducing artifacts in the image [5, 6] while preserving, in some cases strengthening [7], the front side edges. In this context, enhancement approaches, as well as binarization approaches, could be roughly categorized into two major types: the front and back enhancement, also called enhancement with the back side and enhancement without the back side (or blind enhancement). The front and back restoration approaches is used if the back and the front side of the document are available. In this case, the problem is expressed through

a system of two nonlinear equations with two variables representing the front and the back sides of the document image. For example, in [8] authors developed a physical model in order to eliminate the bleed-trough effect. They assume that the scanned output is only involved by the front and back side, their model is then linearized, and the equations system could be solved inversely to derive the original cleaned front and back sides. Another front and back approach was proposed in [9]. In this work, an iterative process was applied to the double-sided document. In short, authors applied iteratively a wavelet transformation to reduce the interfering noise, and at the same time, strengthening foreground strokes gray-levels. One of the most effective approaches using both sides of the document was published by Modhaddam and Cheriet in [10]. In this study, these authors have used a diffusion process to model the physical degradation phenomena undergone by document images. Thereafter, through this study, they constructed a reverse diffusion model to enhance degraded document images.

However, all the above mentioned approaches suffer from major limitations. Indeed, the reverse sides of the scanned documents are not always available. Even if they are available, in all front and back restoration approaches, except Modhaddam-Cheriet one [10], in developing a double sided technique one has to make a hypothesis that the shadow-through of the front side was caused only by the back side. However, this is not always the case, because the quality of the paper sheet or other external sources as a hilt or humidity can also be involved in front side artifacts. For that, blind approaches seem more interesting to deal with restoration issue.

One of the most recent blind document image restoration approach is proposed in [11]. In this study, authors propose to enhance degraded document images by maximizing the amount of the contextual details, leading to improving foreground characters contrast. Nevertheless, the proposed approach fails to remove strong degradation as mentioned in [12].

Directional techniques could bring promising results in the document image enhancement field. An example of such an approach was proposed by Wang et al. in [13]. In this study, these authors use a directional wavelet transformation to exploit the difference between the foreground and the interfering strokes orientations. They make the assumption that foreground and interfering strokes are oriented on $\pi/4$ and $3\pi/4$, respectively. Then, they project the DHW document image onto a 2D wavelet basis oriented on $\pi/4$ and $3\pi/4$ instead of the common directions 0 and $\pi/2$. Since they assumed that foreground strokes are oriented on $\pi/4$, the reconstructed

image with the first basis provides an enhanced version of the degraded document. This process is used in the restoration project of The National Archives of Singapore. Nevertheless, this technique presents an important drawback. Indeed, foreground strokes are not all oriented on a single direction and each character could be constructed with several strokes, having different orientations. An example is shown in Figure 1. Thus, projecting these characters through an unidirectional filter could warp all edges oriented on other directions. Fortunately, this issue could be solved with multi-directional filters as we propose in this study.

Partial-Differential-Equations (PDE) based techniques, or especially, diffusion based techniques represent actually the most effective techniques which resolve degraded document images enhancement issue [14, 15]. In [16], authors improve the method proposed in [10] by replacing the needed reverse side with an estimated background of the degraded document. The most recent and effective approach dealing with degraded document image enhancement was proposed by Drira et al. for the first time in [7], and detailed afterward in [17]. Indeed, Drira et al. adapt the nonlinear anisotropic diffusion process, which has proven its efficiency in various image denoising methods [18], to document image enhancement by substituting analytic eigenvalues functions introduced in [19] by others, more appropriate to reduce documents artifacts while strengthening simultaneously foreground edges and corners. This method is tested on real images with providing promising results. Nevertheless, the proposed process is not automatic and needs to set, manually, two parameters denoted $K_{\pm}$ used for original/artifacts eigenvalues separation.

Figure 1: DHW document images sample illustrating foreground strokes multi-directionality characteristic.

2.2. Existing document images compression standards - a brief overview

Document image compression field has been well investigated in the last two decades, especially after the massive development of the web which is the cradle of the numerical document. This leads to the emergence of specialized document images standards compression such as PDF or DjVu [1]. Document images are a combination of text, natural images, graphics and background.

Handwritten documents take an important part in several national archives amount and need special processing adapted to their specificity. Indeed, handwritten documents have well distinctive characteristics. They are often composed of foreground strokes, reverse side strokes and some background artifacts.

Existing natural images compression standards such as JPEG, JPEGXR or JPEG2000 seem not very effective when applied to document images [1], especially to the handwritten ones, due to their particular structural information. Otherwise, specialized natural documents compression standards such as LuraDocument [1] or DjVu are more effective for handwritten document images. However, they do not perform as well as on natural documents.

It is worth noticing that the most effective approaches and standards dealing with document image compression are based on Mixed Raster Content (MRC) (ITU-T T.44) standard defined in [20]. MRC logic 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 of 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 pictures present in the document. Then, depending on which MRC-based standard we use, a lossless or a high bitrate compression is performed to compress the mask layer, while the background and the foreground layer are both compressed with a low bitrate. An example of MRC-one decomposition is shown in Figure 2.

Image segmentation plays a key role in the MRC process, it consists in classifying pixels in the document (the pixel belongs to either the background layer or the mask layer). The compression bitrate and the quality of the processed document depend highly on the segmentation algorithm. Indeed, if the foreground text is improperly detected, the missed text could be blurred and deformed by the lossy background layer encoder. This impacts negatively the output document quality. On the other hand, if reverse strokes are erroneously detected as foreground strokes through the segmentation algorithm, the compression performance decreases, since the mask layer would contain more active pixels.

[1] http://www.luratech.com/en/

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

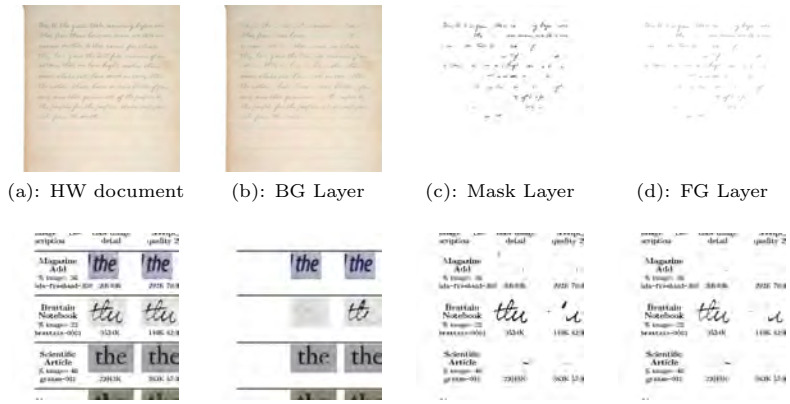
Figure 2: Example of MRC natural document image decomposition.

The most effective text segmentation techniques found in the literature are based on statistical models such as Hidden Markov Model (HMM), Markov Random Filed (MRF) model or Conditional Random Filed (CRF) model described respectively in [1, 21], [22] and [23]. A survey and performance evaluation study of some text segmentation techniques is presented in [24]

One of the most used MRC commercial standards which is, unsurprisingly, based on a statistical analysis (HMM) is DjVu [1]. 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 [25]. Unfortunately, even if DjVu is the best document images compression algorithm [2], the embedded segmentation algorithm (HMM-based) into this standard seems to be not totally suitable to handwritten document images. This is clearly shown in Figure 3 through results obtained when applying DjVu encoder on both natural and handwritten document images.

2.3. Existing joint compression-restoration techniques

Joint compression-restoration is a less investigated filed of research. In literature, only two studies could be found. The first proposed approach [26] is very global and not-adapted to handwritten document images. Roughly speaking, authors propose a scheme for joint compression and restoration with a Non-Linear Interpolative Vector Quantification (NLIVQ) based on restoration techniques using wavelets. Unfortunately, using NLIVQ technique requires the availability of the original cleaned image, which is not the case in real application when dealing with document



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

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

Figure 3: DJVU Encoder applied on both Handwritten document image (top) and Naturel document image (down).

images.

The second approach proposed in [27] requires the availability of both front and back sides of the degraded document. Within this work, the authors propose to perform the restoration part into three steps: Registration, segmentation and inpainting step. In the first step, authors used an optimization method to align the reverse-side original information with the front side, so as each pixel from the reverse-side which, could be found in the front side, is regarded as front side bleed-through. In the second part, authors segment each side into four regions: "foreground", "background", "bleed-through" and "mixed bleed-through and foreground". Finally, in the last step, they use an inpainting technique to replace every pixel identified as "bleed-through" with "background" region. Concerning the compression part, authors proposed to compress the "foreground" areas with JPEG or JPEG-2000 standard and the "background" areas are compressed with one of lossless bilevel standards, such as JBIG2.

3. Proposed approach

The main purpose of this study consists of compressing DHW document images and reducing artifacts. This leads to deal with two different issues, having distinct purposes, namely restoration and compression. First, we propose to build a segmentation algorithm to highlight foreground and artifact areas in the degraded document. Thereafter, through the result of this segmenta-

tion, one can conduct the compression rate in the compression scheme to preserve the foreground information in terms of visual quality and readability, while eliminating the undesirable information.

To start, we define the observed DHW document image $I_d(x, y)$ as a 2D signal containing information (texts and/or patterns) of the recto side image $I_r(x, y)$ (also called the front-side) combined with the content of the reverse side image $I_v(x, y)$ (also called the back-side) and some artifacts which intervene in the background image $I_b(x, y)$, where $[x, y]^T \in \Omega$ and Ω is an open discrete rectangle included on $\mathbf{R}^2$. Some DHW document image samples extracted from DIBCO databases are depicted in Figure 4.

Now, we propose to construct the image called I_{bv} which contains both information of I_v and I_b. This is justified by the fact that the limit between these images is ill-defined. Moreover, to improve DHW document images quality, one does not need to extract I_v and I_b separately. In the sequel of this paper, I_{bv} is called background, reverse-side or artifact image. For each pixel (x, y), I_d is expressed as follows:

$$I_d(x, y) = p_r(x, y)I_r(x, y) + (1 - p_r(x, y))I_{vb}(x, y) \tag{1}$$

where $p_r(x, y)$ is the probability that (x, y) belongs to the recto side image I_r.

Figure 4: DHW document images samples.

As we deal with two distinct tasks, the proposed approach could be then, roughly, divided into two main steps:

The first step consists on performing a Foreground/Background (F/B) segmentation algorithm to detect pixels belonging to I_r according to two characteristics. Indeed, one can tackle this problem by considering both directional features and gray levels to detect which pixels are belonging to the recto-side in order to retrieve I_r. Indeed, foreground strokes gray levels are

more important, in terms of intensity, than artifacts gray levels in most degraded document areas. Furthermore, in several DHW document images, suffering especially from bleed-through, foreground strokes directions are different from interfering strokes ones. These characteristics could be observed in Figure 4.

Joining enhancement and compression processes in DjVu encoder is the next step of the proposed approach. In short, we substitute the segmentation algorithm embedded into DjVu with the proposed F/B segmentation algorithm performed through the first step, this provides enhanced-compressed images in DjVu process output.

3.1. Proposed Foreground/Background segmentation approach

This algorithm consists on detecting foreground strokes in the considered DHW document by analyzing directional and contrast features. To achieve this, firstly, the image called Shadow Foreground Image (SFI), which is an anisotropic estimation of the DHW document in well defined directions is computed. Thereafter, we use both the degraded document I_d and its anisotropic estimation to perform a robust F/B segmentation algorithm.

3.1.1. Directional estimation of the DHW document image

Herein, the anisotropic estimation of the DHW document is performed. Firstly, we develop the theoretical analysis on how to compute the SF image through a multi-scale multi-directional basis. Thereafter, we assume, through some observations, that the wave atoms transform is the most appropriate transformation in our context. Thus, a brief introduction to the wave atoms transformation is given; we propose thereafter a new multi-directional filter based on the wave atoms basis. Finally, the SF image is estimated through this filter.

Shadow Foreground Image (SFI) - Theoretical analysis

We propose to decompose the degraded image I_d into a multi-resolution (to take into account different strokes widths), multi-directional (to consider all orientations) basis $B = \{b_m\}_{(m \in N)}$, defined in $L^2(\mathbf{R})$. Basically, the decomposition of I_d into the basis B could be written as following:

$$I_d = \sum_{m \in N} < I_d, b_m > b_m \tag{2}$$

Firstly, let us consider $G = \{< I_d, b_m >\}_{m \in N}$ set containing all coefficients provided by the projection of I_d in the basis B. Thereafter, G is segmented on $K \times G_i$ subsets, then $G = \bigcup_{i=1}^{K} G_i$ with $\{G_i \cap G_j = \emptyset, i \neq j\}$. In addition, it is worth to point out that G_i should correspond on the I_d projection in one and well defined orientation. Then, We associate afterward to each G_i an orientation obtained through the function denoted $\vartheta : \mathbf{R}^{Card(G_i)} \mapsto [0, 2\pi]$, where $\vartheta(G_i)$ is used to extract the dominant direction of each subset G_i.

With the aim to estimate the relevance of each G_i direction's, one propose to introduce $\phi(\vartheta(G_i), \Omega_F)$, ϕ: $[0, 2\pi]^{k+1} \mapsto \mathbf{R}$ to estimate how $\vartheta(G_i)$ is closer to elements of the vector $\Omega_F = \{\Omega_1^F, \Omega_2^F, ..., \Omega_k^F\}$ containing all foreground strokes directions a priori fixed.

Thereafter, for each G_i, one can associate a global distance to elements of Ω_F by means of $\phi(\vartheta(G_i), \Omega_F)$. In other words, $\phi(\vartheta(G_i), \Omega_F)$ could be seen as a metric of the relevance of each G_i against the vector Ω_F. Then, we construct a novel space called N_Ω which contains the indexes m such as $< I_d, b_m >$ belongs to G_i and the value $\phi(\vartheta(G_i), \Omega_F)$ is higher than a predefined threshold T_d. Thus, N_Ω should contains only coefficients which represent strokes in I_d oriented on directions stated in Ω_F. N_Ω is written as follows:

$$N_\Omega = \{m \in N : \quad < I_d, b_m > \in G_i \quad and \quad \phi(\vartheta(G_i), \Omega_F) > T_d\} \tag{3}$$

The reconstructed image $\hat{I}_r$ with N_Ω is called Shadow Foreground Image (SFI). Which could be seen as an estimate of I_d regarding Ω_F, is given by:

$$\hat{I}_r = \sum_{m \in N_\Omega} < I_d, b_m > b_m \tag{4}$$

It is worth noticing that $\hat{I}_r$ is constructed by considering the directional characteristic only. Indeed, the contrast of I_d strokes oriented in one of a desired direction within Ω_F is preserved in $\hat{I}_r$. However, if some highly contrasted segments in I_d, which may belong to the foreground side, are not oriented on any of directions within Ω_F, the corresponding strokes in the constructed image $\hat{I}_r$ would have low contrast. Thus, I_d seems more appropriate than $\hat{I}_r$ to discriminate

strokes by their contrast, where $\hat{I}_r$ is adapted to discriminate the strokes by their orientation.

Doubtless, the most important step is to pick on the optimal basis $B = \{b_m\}_{(m \in N)}$ since performance of the proposed F/B segmentation algorithm depends mostly on the chosen basis efficiency. Indeed, the most appropriate transformation should emphasize the directional feature through a multi-resolution analysis. Among the existing directional methods, the Radon transform does not allow the multi-resolution analysis. Whereas, the Ridgelet transform [28] is more suitable for the representation of line segments. Curvelet [29] is another intersting analysis tool for document images. This transform is specified for representing curvatures and may be useful for our application. However, given the specificity of the handwritten documents which makes little apparent curvature, we believe that it may be more suitable to consider the information in the document as oriented oscillating textures. Indeed, we seek to separate through resolutions, textures differentiated by their orientations. For that, wave atoms transform seems to be more appropriate for document image analysis.

Wave Atoms (WA) transform

Wave atoms transform, developed in 2006 by Demanet [30], is a multi-scales transform characterized by its capacity of representing the oscillating structure and emphasizing multi-directional propriety. Consequently, this transformation is increasingly used in various image processing fields [31, 32].

In order to make this paper self-sufficient, let us recall some basis notions on WA transform following the description given in [30]. Herein, we briefly introduce the wave atoms frame followed by the 1D/2D wave atoms projection formulas.

Wave atoms frame:

Wave atoms transform is constructed from tensor products of adequately chosen 1D wavelets packets. The wave atom $\zeta^J_{\eta,n}(x)$ and its Fourier transform $\hat{\zeta}^J_\eta(\omega)$, centered in space domain in $x_{j,n} = 2^{-J}n$ and in frequency around $\pm\omega_{(j,\eta)} = \pm\pi 2^J \eta$ [30] with $\mu = (j, \eta, n)$, where $j \in \mathbf{N}$, $\eta \in \mathbf{N}$ and $n \in \mathbf{Z}$.

$\zeta^J_{\eta,n}(x)$ and its Fourier transfrom are obtained by combining dyadic dilates and translates of ζ^0_η, which is the mother wave packet satisfying well defined conditions expressed in [30], in space domain:

$$\zeta^J_{\eta,n}(x) = \zeta^J_\eta(x - 2^{-J}n) = 2^{J/2}\zeta^0_\eta(2^J x - n) \tag{5}$$

and $\hat{\zeta}^0_\eta(\omega)$ in frequency domain:

$$\hat{\zeta}^J_\eta(\omega) = 2^{\frac{-J}{2}}\hat{\zeta}^0_\eta(\omega 2^{-J}) \tag{6}$$

Both $\zeta^0_\eta(x)$ and its Fourier transform $\hat{\zeta}^0_\eta(\omega)$ are defined in [33].

Wave atoms projection

In 1-D, wave atoms coefficients $C_{j,\eta,n}$ of $u(x)$ ($x \in \mathbf{R}$), which is an 1D signal, can be computed as a decimated convolution at scale 2^{-J}:

$$C_{j,\eta,n} = \int \zeta^J_\eta(x - 2^{-J}n)u(x)dx \tag{7}$$

By Plancherel theorem:

$$C_{j,\eta,n} = \frac{1}{2\pi}\int e^{-i2^{-J}\omega n}\bar{\hat{\zeta}}^J_\eta(\omega)\hat{u}(\omega)dx \tag{8}$$

In 2D, the orthonotmal basis is formed by individually taking products of 1-D wavelet packets (equ. (5)). 2D wave atoms indexes are now: $\boldsymbol{\eta} = (\eta_1, \eta_2) \in \mathbf{N}^2$ and $n = (n_1, n_2) \in \mathbf{Z}^2$. Elements of 2D wave atoms frame are given in [30] as:

$$\varphi^1_\mu = \frac{\varphi^+_\mu + \varphi^-_\mu}{2}; \qquad \varphi^2_\mu = \frac{\varphi^+_\mu - \phi^-_\mu}{2} \tag{9}$$

with:

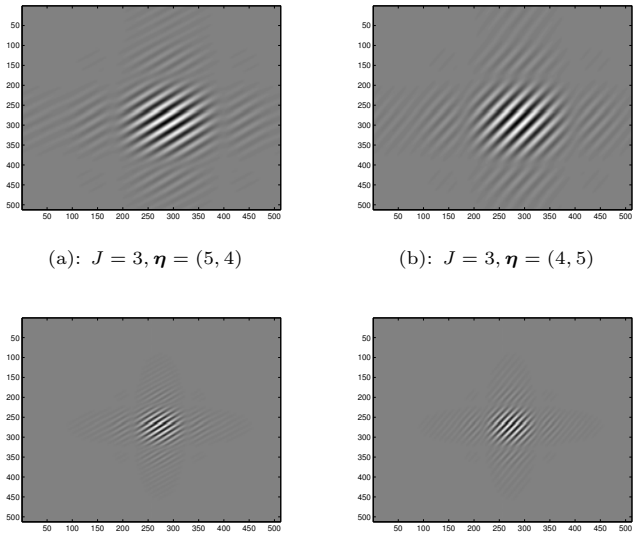
$$\varphi^+_\mu = \varphi^+_\mu(x_1, x_2) = \zeta^J_{\eta_1}(x_1 - 2^{-J}n_1)\zeta^J_{\eta_2}(x_2 - 2^{-J}n_2) \tag{10}$$

The second orthonormal basis φ^-_μ is built through the Hilbert transformation:

$$\varphi^-_\mu = \varphi^-_\mu(x_1, x_2) = H\zeta^J_{\eta_1}(x_1 - 2^{-J}n_1)H\zeta^J_{\eta_2}(x_2 - 2^{-J}n_2) \tag{11}$$

Shadow Foreground Image through the Wave Atoms Transform

Let us consider I_d image of size $[M \times N]$. We note $\{C(j, \boldsymbol{\eta})\}_{j\in\mathbf{N},\boldsymbol{\eta}\in\mathbf{N^2}}$ the resulting WA coefficients obtained through the projection of I_d on all 2D Wavelet packets $\varphi^\pm_{j,\boldsymbol{\eta},n}$. The projection of I_d into the specific Wavelet Packet (WP) $\varphi^\pm_{J,\boldsymbol{\eta},n}$, at a the scale J, provides the cell $C(J, \boldsymbol{\eta})$ of size $[2^J \times 2^J]$ holding 2^{2J} wave atoms coefficients. In [34], authors have noticed that each 2D WP $\varphi^\pm_{j,\boldsymbol{\eta},n}$ oscillates in a specific orientation denoted: θ.



(a): $J = 3, \boldsymbol{\eta} = (5, 4)$ (b): $J = 3, \boldsymbol{\eta} = (4, 5)$

(c): $J = 4, \boldsymbol{\eta} = (5, 4)$ (d): $J = 4, \boldsymbol{\eta} = (4, 5)$

Figure 5: 2D WP basis sample.

We consider $\theta(j, \boldsymbol{\eta}, n)$ as the oscillation angle of the 2D WP $\varphi^{\pm}_{j,\boldsymbol{\eta},n}$ used to compute $C(j, \boldsymbol{\eta})$. Due to the oscillatory behavior of the WA basis, if some strokes in the DHW document I_d are oriented in $\theta(j, \boldsymbol{\eta}, n)$, the considered strokes would be well represented in $C(j, \boldsymbol{\eta})$ compared to others. In other words, $C(j, \boldsymbol{\eta})$ could be seen as projection, at the scale j, of all strokes in I_d oriented in $\theta(j, \boldsymbol{\eta}, n)$. Thus, this feature seems to be interesting to perform a directionality discrimination. Some samples of WP at different scales and frequency indexes are shown in Figure 5.

The oscillation angle $\theta(j, \boldsymbol{\eta}, n)$ of the WP $\varphi^{\pm}_{j,\boldsymbol{\eta},n}$ depends only on the frequency index $\boldsymbol{\eta}$ (this could be observed in Figure 5). The orientation of the vector perpendicular to $\varphi^{i}_{j,\boldsymbol{\eta},n}$ oscillations denoted $\theta(\boldsymbol{\eta})$ is given by:

$$\theta(\boldsymbol{\eta}) = \arctan(\frac{\eta_2}{\eta_1}) \tag{12}$$

Now, by producing a simple analogy with the theoretical analysis, one can model the coefficients subset G_i and its orientation $\vartheta(G_i)$ by the wave atom cell $C(j, \boldsymbol{\eta})$ and its oscillation angle $\theta(\boldsymbol{\eta})$, respectively. With aim to exploit the directional feature of the wave atom cells $C(j, \boldsymbol{\eta})$, let

us introduce the distance $\psi_{\alpha}^{\theta(\eta)}$ expressed as follows:

$$\psi_{\alpha}^{\theta(\boldsymbol{\eta})} = e^{\frac{-|\ \theta(\boldsymbol{\eta})-\alpha|^2}{\sigma}}, \tag{13}$$

This is as a measure of the distance between the considered orientation α and the oscillation angle $\theta(\boldsymbol{\eta})$ of the wave atoms cell $C(j,\boldsymbol{\eta})$. The coefficient $\sigma \in \mathbf{R}$ controls the exponential tight, and intervene to characterize the selectivity degree of this distance. In other words, the more σ is high the more the difference between α and $\theta(\boldsymbol{\eta})$ should be low to reach a well defined value of the distance $\psi_{\alpha}^{\theta(\boldsymbol{\eta})}$.

The function $\phi(\vartheta(G_i),\Omega_F)$ which allows to quantify the relevance of the G_i orientation angle according the foreground angles vector Ω_F, has been introduced previously. Through the analogy used above, this function could be written now as $\phi(\theta(\boldsymbol{\eta}),\Omega_F)$, and one can compute it by the sum of all distances between the oscillation angles $C(j,\boldsymbol{\eta})$ and each Ω_F component. This is done by estimating the distance of $\theta(\boldsymbol{\eta})$ according to each Ω_i^F thanks to equ. (13). Therefore, we propose to write $\phi(\theta(\boldsymbol{\eta}),\Omega_F)$ as follows:

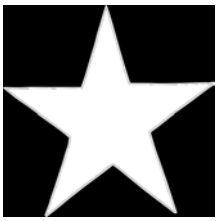
$$\phi(\theta(\boldsymbol{\eta}),\Omega_F) = \sum_{\Omega_i^F \in \Omega_F} \psi_{\Omega_i^F}^{\theta(\boldsymbol{\eta})} \tag{14}$$

or it could be written as:

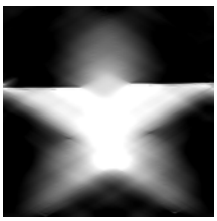
$$\phi(\theta(\boldsymbol{\eta}),\Omega_F) = e^{\frac{\arctan(\frac{\eta_2}{\eta_1})^2}{\sigma}} \sum_{\Omega_i^F \in \Omega_F} e^{\frac{\Omega_i^F}{\sigma}\left(\Omega_i^F - 2\arctan(\frac{\eta_2}{\eta_1})\right)} \tag{15}$$

The function $\psi_{\Omega_i^F}^{\theta(\boldsymbol{\eta})}$ could be seen as the relevance metric of $\theta(\boldsymbol{\eta})$ compared to the vector Ω_F. Then, if $C(j,\boldsymbol{\eta})$ contains coefficients which represent mostly strokes oriented on minimum one of the foreground angles within Ω_F, $\phi(\theta(\boldsymbol{\eta}),\Omega_F)$ takes high value, and vice versa. Having regard this characteristic, $\phi(\theta(\boldsymbol{\eta}),\Omega_F)$ is used to select WA coefficients representing strokes oriented on directions in Ω_F. Then, these coefficients are all stored in N_Ω thanks to equ. (3), with $T_d = \xi \times \mathrm{Card}(\Omega_F)$, $\xi \in \mathbf{R}$.

Concerning the WA transform implementation, six distinct algorithms were proposed in [30]. In our study, the parabolic-directional implementation is used although it is the most redundant of all algorithms proposed in [30]. On the other hand, this implementation allows calibrating the $\boldsymbol{\eta}$ indexes in the largest band, which permits to provide the largest distinctive orientation vector.



(a): Toy image



(b): $\Omega_F = 0$

(c): $\Omega_F = \frac{\pi}{3}$

(d): $\Omega_F = -\frac{\pi}{3}$

(e): $\Omega_F = \{0, \frac{\pi}{3}\}$

(f): $\Omega_F = \{0, -\frac{\pi}{3}\}$

(g): $\Omega_F = \{0, -\frac{\pi}{3}, \frac{\pi}{3}\}$

Figure 6: SFIs obtained through different values of Ω_F for a toy image.

Finally, the reconstructed image $\hat{I}_r$ called Shadow Foreground Image (SFI) is computed by using equation (4). For illustration, we present in Figure 6, SFIs obtained through different values of Ω_F for a toy image sample.

This study allows us to notify two main cases. Firstly, I_d strokes oriented in one of directions defined in Ω_F. This case is is the optimal one, since the corresponding strokes in the SFI preserve their original contrast. Secondly, I_d strokes which are not oriented in directions within Ω_F. Corresponding strokes in the SFI exhibit low contrast, and may be completely dropped.

We might assume that information to detect (foreground pixels) are more contrasted than artifacts. Then, a contrast estimation in each area (block or neighborhood area) in the degraded image should be sufficient to distinguish areas containing foreground and background strokes. However, in several DHW documents, some foreground strokes are not more contrasted than

artifacts. In this case, directional information given by $\hat{I}_r$ should be useful. However, using the directional information, only, to highlight foreground strokes areas is likewise insufficient. Indeed, as stressed in the theoretical analysis of the SF image, although if some foreground strokes are presenting a quite relevant contrast in the degraded document, a decay on their contrast is observed in $\hat{I}_r$ if these strokes are not oriented in directions defined in Ω_F. Therefore, taking into account both the local contrast information and the directional features is a prerequisite to perform efficiently this discrimination process. For that, it is necessary to develop a scheme that uses both the original degraded image I_d and its corresponding $\hat{I}_r$, to mark foreground strokes from the original degraded image.

3.1.2. Foreground/Background discrimination process

Herein, the foreground/background separation algorithm is constructed by dividing the original image I_d and its corresponding SFI into blocks of size $[z \times z]$ to discriminate between blocks contain original strokes from those containing only artifacts.

First, we may assume that if the contrast is enhanced in certain blocks in $\hat{I}_r$ regarding to their corresponding block in I_d, these areas are considered as foreground ones. However, if this condition is not verified, a local contrast coefficient called λ_i depending on both I_d and $\hat{I}_r$ is computed and used to separate blocks containing original strokes from the others. The local contrast is then estimated as follows..

$$\lambda_i = \frac{\max(bl_i + \hat{bl}_i) - \min(bl_i + \hat{bl}_i)}{\max(bl_i + \hat{bl}_i) + \min(bl_i + \hat{bl}_i)} \tag{16}$$

where : bl_i and $\hat{bl}_i$ are, the SFI values of the ith block in the original image I_d and its corresponding one in $\hat{I}_r$, respectively

The window size $z \times z$ is chosen so as to encompass the different stroke sizes. A given block bl_i is finally considered as foreground one, if and only if it is identified as foreground through at least 2/3 of the considered sizes of z taken in $\{8, 16, 24\}$.

Now, let us define $K = \{1, ..., MN/z^2\}$ and consider $D = \{d_{k,i} : k \in K, \quad d_{k,i} >= d_{k+1,j} \quad and \quad d_{k,i} = \lambda_i\}$ vector containing all λ_i coefficient in the decreasing order. Afterward, for each $d_{k,i}$ except $k = 1$, we consider the local relative variation δd_k expressed as follows:

$$\delta d_k = \frac{d_{k,i} - d_{k-1,j}}{d_{k-1,j}}, \tag{17}$$

Relying to the features considered to compute λ_i, foreground side blocks should present a strong and very close values. Hence, to highlight them, the problem boils down to find a threshold value λ_T which separates foreground and background local contrast coefficient values. In other words, if one sort λ_i in a decreasing order and compute the relative variation of reordered coefficients, the threshold λ_T corresponds to the first maximum relative variation. Moreover, if the processed document suffers from bleed-through, a second maximum variation may exist and should correspond to the transition "reverse-side image"/"background image". However, as noticed earlier, the limit between these images is ill-defined. Indeed, in practice, the second maximum relative variation does not correspond to the real frontier separating these images.

Let us define t_1 and t_2 ($t_1 < t_2$ and $(t_1, t_2) \in K^2$) as the indexes maximizing the ratio δd_k. These subscripts represent, respectively, the "foreground"/"reverse side" and "reverse side"/"background artifact" transition thresholds. Hence, the threshold value λ_T which separates original strokes from other is nothing else than λ_{t_1}. This process is performed for $z = 8, 16, 32$ to consider all strokes widths. The finale segmented image I_r is then computed as the union of the resulted images from these three configurations. The deduced algorithm algo. 1 summarizes this step. An example illustrating performances of the proposed segmentation approach on two images from the most arduous-to-restore within all DIBCO datasets is shown in Figure 7.

3.2. Joint enhancement-compression process

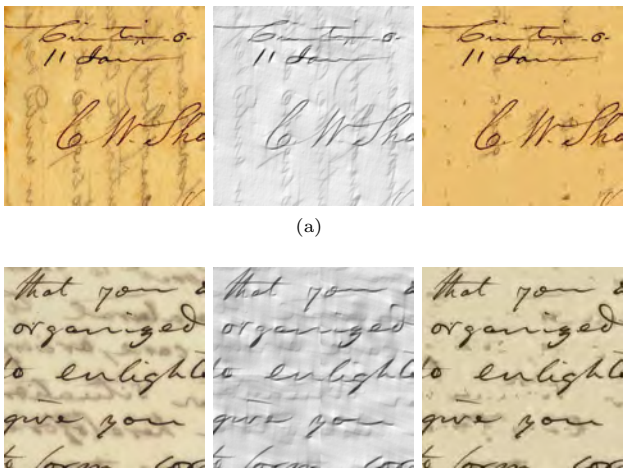
DjVu is the most accurate and robust document images compression algorithm, as proven in [2]. This explains its massive use in real commercial applications and projects such as the Wikipedia project: Wikisource[2], which is a an online digital library of free textual forms in many languages and translations. Moreover, this encoder is an open source, and both binary and release versions are available[3].

The main motivations behind the use of DjVu encoder in the proposed method are its per-

[2] http://en.wikipedia.org/wiki/Wikisource
[3] http://djvu.sourceforge.net/

input : I_d, I_r from equ. (4)
output: I_s,I_m: segmented image and its mask respectively.
// $0 \leftarrow whitepixel$, $1 \leftarrow blackpixel$;
Initialize I_s and I_m to 0;
for $i = 1$ *to* MN/z^2 **do**
  if $contrast(bl_i) < contrast(\hat{bl}_i)$ **then**
    $\lambda_i = +\infty$;
  else
    Compute λ_i through equ. (16);
  end
end
$D \leftarrow$ Decreasing sort of λ_i coefficients;
for $k = 2$ *to* MN/z^2 **do**
  Compute δd_k;
end
$t_1 \leftarrow \arg\max_{k \in K}(\delta d_k)$;
$\delta d_{t_1} \leftarrow 0$;
$t_2 \leftarrow \arg\max_{k \in K}(\delta d_k)$;
$t_1 \leftarrow \min\{t_1, t_2\}$;
$\lambda_T \leftarrow \lambda_{t_1}$;
for $i = 1$ *to* MN/z^2 **do**
  if $\lambda_i \geq \lambda_T$ **then**
    $bl_i \leftarrow Foreground$;
    $I_s^i \leftarrow bl_i$;
    $I_m^i \leftarrow 1$;
  else
    $bl_i \leftarrow Background$;
  end
end

Algorithm 1: Foreground/Background segmentation



(a)

(b)

Figure 7: In (a) and (b), and from left to right: DHW document, SFI version and the segmented version.

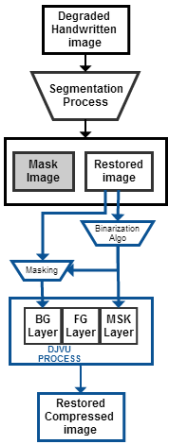


Figure 8: Proposed joint enhancement-compression scheme- a summary

formance and its extensive use in many national archive libraries. Thus, developing a technique which could be easily integrated in the existing digitization frameworks without important modifications on the system architecture offers many advantages including flexibility, adaptability and cost savings.

Roughly speaking, we propose to substitute the segmentation algorithm used by this standard, based on HMM, by the proposed F/B segmentation algorithm which is suitable to deal with historical documents. The idea is then to use lossless compression for foreground strokes, so as to preserve the most relevant information and performing a highly lossy compression in the less relevant remaining parts of the image document.

Herein, the proposed enhancement-compression process is described. First, the segmented image is binarized using Su et al. algorithm [35]. The output binary image is then inserted in the mask layer and its associated color image, taken from the segmented one, is included in the foreground layer. It is worth noticing that any binarization algorithms may suffer from some limitations, especially pixel miss-classification errors [3]. Thus, the segmented image with the proposed approach (see algo. 1) is poured through the mask layer onto the background layer in order to store the rest of pixels which are not selected in the binarized image. The foreground characters affected by the binarization process could be found in the background layer. Likewise, storing these missed pixels in the background layer allows preserving the look and the feel of the original document in the output one. This improve both the visual quality and the readability of

the processed image. Finally, the mask layer is compressed through JBIG2 lossless compression algorithm, and both foreground/backtround layers are compressed with a lossy IW44 compression algorithm. Therefore, performing a lossless compression algorithm in the mask layer allows to keep the readability of the processed document. Figure 8 summarizes the most relevant parts of the proposed method.

Parameters q_{fg} and q_{bg} are used to tune the foreground and background layers bitrate, respectively. In practice, since most of document images exhibit low variations in the background, q_{bg} is set to a low value without any significant impact on the output document. Consequently, the output document image size depends mainly on q_{fg}. It is worth noticing that the variation of q_{fg} affects only the color nuances, and the output document patterns are literally not affected when tuning q_{fg}. Hence, q_{fg} could be set to low values without striking the readability.

4. Experiment and discussion

In this section, the obtained results are shown and discussed to illustrate the performance of the proposed joint enhancement-compression approach. All DHW document images of DIBCO 09,10,11,13 databases[4] are used. The dataset consists of 31 DHW document images with their corresponding binary ground truth images. From this set, 9 suffer from bleed-through and the rest suffers from other background artifacts, such as: ink/coffee smudges or illumination variation. Herein, the whole parameters introduced in the proposed approach are first tuned. The performance of the proposed F/B segmentation algorithm against the most recent document image enhancement algorithm [7] is evaluated subjectively and objectively. Finally, we carry a strong subjective and objective evaluation of the proposed enhancement-compression approach in terms of quality/bitrate against the most efficient scanned-document compression approach [1] and the well-known JPEG [44] and JPEG 2000 [44] standards.

4.1. Parameters calibration

Tuning the parameters is based on some observations and preliminary studies. In the following, we provide some guidelines on the tuning of the whole parameters introduced in this work.

[4] http://users.iit.demokritos.gr/∼kntir/HDIBCO2014

It has been noticed in several studies, such in [13], that reverse side strokes are oriented differently than original strokes when the degraded document is observed through the front side. Often, most people without noticing bow slightly their pens when writing vertical strokes, while horizontal oriented strokes are often written in their real direction. From this observation, we propose to set $\Omega_F = \{0, -25, -50, -75\}$. Parameters σ and ξ which intervene in: equ. (13) and (3) could take values in [100; 600], [0.01; 0.5], respectively. Nevertheless, from experiments, we have noticed that these parameters have not a very significant impact on the segmentation algorithm performance. They are then fixed to 300 and 0.05, respectively.

In the compression step, the parameter q_{bg} used in the background layer is set to 0.01 bit/pixel. This is due to the fact that, mainly, background in document images shows less variations. Thus, the global bitrate on which the processed document is compressed depends only on the foreground layer bitrate q_{fg} and the mask matrix size which is lossless compressed.

Concerning the implementation of the WA Transform, it is worth to notice that, due to some specificities within the WA basis, the whole provided versions by Demanet *et al.* [5] could be only performed on images of sizes $2^N * 2^N$. For that, a significant effort has been done in this direction to generalize their implementations to any image size. This implementation will be available on the web, as well as the proposed enhancement-compression framework.

4.2. Foreground/Background segmentation evaluation

Many approaches have been proposed for image segmentation evolution [36, 37]. Here, it is worth to point out that we do not consider the proposed F/B segmentation algorithm as a document image enhancement algorithm, since the aspect and the look of the processed document image are altered by the block division step performed in algo. 1, then a visual subjective comparison could not be performed. However, the restored version with the same look and feel of the original document is given after performing the full proposed enhancement-compression process (see Fig. 11).

Mainly, document images enhancement approaches try to improve the visual quality of the document without corrupting its original aspect, so as to facilitate its readability. On the other

[5] http://www.waveatom.org/

hand, the aim of the proposed F/B segmentation algorithm is to detect foreground side pixels in the processed document and setting all other pixels to zero value, which improves the document readability also. Hence, the proposed F/B segmentation algorithm as well as the document enhancement approaches converge globally towards the same purpose, which consists in improving the processed document readability. According to this ascertainment, an objective evaluation of the readability between the proposed F/B algorithm and existing document image enhancement techniques in the state-of-the-art could be performed.

To evaluate the proposed segmentation algorithm performance, an objective quality assessment through some commonly used document readability measures is performed. The evaluation set includes one of the most recent restoration approach proposed by Drira et al. in [7] and the most recent automatic approach proposed as well by Drira et al. in [17].

Indeed, it is worth noticing that if a restoration/segmentation processes are successfully done, performance of the binarization algorithm obtained after performing restoration/segmentation processes should be similar or better than performance obtained when performing a binarization algorithm directly on the degraded documents. Hence, to evaluate the readability improvement given by restoration/segmentation algorithms, it seems to be sufficient to compare the binarization quality obtained before/after performing these restoration/segmentation algorithms, as done by Hedjam et al. in [12]. Then, for tests, a set of binarization algorithms have been used to carry out the objective evaluation. We have then considered two kinds of binarization algorithms:

a)- binarization algorithms adapted to degraded document images: This group, contains one of the most used binarization algorithm in the literature introduced by Otsu in [38], and to our best knowledge, the most robust binarization algorithm proposed by Su et al. in [35] and detailed in [39].

b)- binarization algorithms not adapted to degraded document images: In this group, we have considered the widely used Tsai [40] binarization algorithm specialized on natural color document images binarization.

As noticed earlier, the Drira et al. approach proposed in [7] is not blind (see sec. 2.1) and requires empirical manual interventions to set the $K_{\pm}$ thresholds. Thus, with the aim to optimize its performance, these thresholds have been tuned to obtain the best $\mathrm{F}_{\mathrm{Measure}}$ result in each DHW document image belonging to the constructed dataset.

4.2.1. Objective evaluation criteria

The performance of the binarization outputs is evaluated by means of three measures widely used in the DIBCO competitions.

1. The classical $\mathrm{F}_{\mathrm{Measure}}$ criterion, which is widely used for document image binarization.

2. The $\mathrm{PF}_{\mathrm{Measure}}$ criterion: This criterion is computed as well as the $\mathrm{F}_{\mathrm{Measure}}$ while substituting the RC score by the Pseudo_RC score described in equ. 19.

3. The Distance Reciprocal Distortion (DRD) score [41]: This measure correlates with the human visual perception. Since 2012, this strong measure is performed to evaluate the binarization algorithms in DIBCO competitions.

$$\mathrm{F}_{\mathrm{Measure}} = 100 * \frac{2 * \mathrm{PR} * \mathrm{RC}}{\mathrm{PR} + \mathrm{RC}} \tag{18}$$

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

$$\mathrm{Pseudo\_RC} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} I_{bin}(i,j) I_{sg}(i,j)}{\sum_{i=1}^{M} \sum_{j=1}^{N} I_{sg}(i,j)} \tag{19}$$

where I_{bin} is the binarized document image and I_{sg} represents its skeletonized ground truth image.

4.2.2. Performance evaluation and discussion

First, proposed F/B segmentation and both Drira et al. automatic and manual approaches are performed to segment and restore, respectively, all DHW document images in the constructed dataset. Afterward, we perform the three binarization algorithms presented above in four distinct datasets: constructed dataset, enhanced document images with the manual Drira et al. approach dataset, enhanced document images with the automatic Drira et al. approach dataset and segmented document images with the proposed F/B segmentation algorithm dataset.

The obtained results are summarized in Tab. 1 in terms of Recall, Precision, PSNR and the three robust readability measures.

Table 1: Recall, Precision, $F_{Measure}$, $PF_{Measure}$, DRD and PSNR obtained through several binarization algorithms applied on: Original DHW document images, enhanced documents with the **manual** Drira et al. filter [7], enhanced documents with the **automatic** Drira et al. filter [17] and enhanced documents with the proposed segmentation approach.

Binarization tech.	Dataset	Recall	Precision	F_MESURE	PF_MESURE	DRD	PSNR
Tsai	Direct Binarization (DB)	0,694	0,886	74,143	76,619	16,747	14,971
	Drira et .al **manual** approach (MDA) [7]	0,766	0,859	77,216	80,233	13,338	15,947
	Drira et .al **automatic** approach (ADA) [17]	0,810	0,734	72,614	72,777	14,490	15,058
	Proposed **automatic** approach	0,790	0,876	80,762	83,440	9,445	16,787
	Gain against DB			**6,620**	**6,821**	**7,302**	**1,816**
	Gain against MDA			**3,547**	**3,207**	**3,893**	**0,841**
	Gain against ADA			**8,148**	**10,663**	**5,045**	**1,729**
Otsu	Direct Binarization (DB)	0,727	0,870	74,480	77,163	16,690	15,442
	Drira et .al **manual** approach (MDA) [7]	0,819	0,844	79,413	82,991	12,346	16,961
	Drira et .al **automatic** approach (ADA) [17]	0,808	0,763	73,919	74,420	14,328	15,746
	Proposed **automatic** approach	0,850	0,826	81,013	85,102	8,836	17,317
	Gain against DB			**6,533**	**7,939**	**7,854**	**1,875**
	Gain against MDA			**1,600**	**2,111**	**3,510**	**0,356**
	Gain against ADA			**7,094**	**11,183**	**5,510**	**1,571**
Su et al.	Direct Binarization (DB)	0,954	0,887	91,645	93,873	2,561	20,799
	Drira et .al **manual** approach (MDA) [7]	0,956	0,893	92,202	94,337	2,235	20,849
	Drira et .al **automatic** approach (ADA) [17]	0,775	0,938	83,534	83,536	3,882	18,743
	Proposed **automatic** approach	0,949	0,897	92,033	94,466	2,566	20,821
	Gain against DB			**0,388**	**0,593**	**-0,004**	**0,022**
	Gain against MDA			**-0,169**	**0,129**	**-0,330**	**-0,028**
	Gain against ADA			**8,499**	**10,930**	**1,647**	**2,078**

From Tab. 1 one could firstly note that the proposed segmentation approach improves readability performances given by both adapted and non-adapted binarization techniques to document images. In addition, performances of the non-adapted to historical document technique are significantly enhanced when associated with the proposed scheme.

Relaying to the most accurate document image quality measures, Tab. 1 shows objectively that the proposed automatic technique greatly outperforms the commonly used automatic technique proposed by Drira et al. [17]. The major contribution of the proposal against existing approaches consists in a considerable improvement of the readability as confirmed by the $PF_{Measure}$. Indeed, the $PF_{Measure}$, estimated through the ground-truth skeleton image, evaluates the readability by estimating the edges sharpness degree within the document image. This characteristic makes it in the top of the most representative readability measures [42] as well as the DRD score. In addition to the readability amelioration, the historical aspect of the processed document is well preserved when using the proposed approach. This could be demonstrated by using the DRD score. In fact, so far, the Distance Reciprocal Distortion (DRD) [41] remains the most representative measure of the human visual perception in the case of historical documents,

as explained in [42].

Furthermore, even if we propose a full-automatic technique, close performances are observed between the manual Drira et al. algorithm and the proposed method. However, it is worth to point out that the manual Drira et al. algorithm needs a significant effort for tuning the $K_{\pm}$ coefficients to reach such results and this operation becomes tedious when dealing with a considerable number of degraded documents. In addition, it is well-known that time complexity is the most important drawback of diffusion techniques (see Tab. 2). Then, approaching or beating the Drira et al. performances with a simple-to-implement, fast and automatic technique could be seen as an achievement. Tab. 2 shows the minimum, the maximum and the average executing time when using the proposed and Drira et al. automatic approaches to restore document images of the considered dataset. We note that both of them have been executed in the same environment (Intel(R) Core i7 at 2,6Ghz).

Table 2: A runtime comparison of the evaluated automatic approaches

Execution time (s)	**Proposed approach**	**Drira et al. auto. approach [17]**
Min	3,658	13,692
Max	84,116	222,539
Average	11,233	37,890

Nevertheless, it is clear from Tab. 1 that the proposed method fails to improve Precision performances given by some binarization algorithms (eg. [40, 38]). This is due to the fact that the proposed segmentation algorithm focuses on highlighting foreground blocks instead of foreground pixels in the document. Consequently, even if this choice is highly interesting in terms of computational load saving, precision is doubtless lost when dealing with blocks instead of pixels.

4.3. Joint enhancement-compression method evaluation

To our best knowledge, the proposed joint compression-enhancement approach for document images is the first of this kind. In fact, so far, document image compression issue is treated with classic image compression standards. Therefore, we propose to compare performances of the proposed compression approach with three robust and well-known compression encoders:

- DjVu [21]: As discussed earlier, DjVu is the most accurate document image compression

scheme.

- JPEG 2000 [43]: JPEG2000 is one of the most efficient lossy image compression standard.
- JPEG [44] : So far, due to its implementation simplicity, JPEG remains the most used image compression standard.

We start by compressing the "constructed dataset" with DjVu, JPEG and JPEG2000 by varying the compression bitrate. Thereafter, by varying only the foreground layer aggressiveness compression parameter q_{fg}, the proposed joint restoration-compression approach is performed with various compression bitrates on the same dataset.

Afterward, as well as we have done to assess the proposed F/B segmentation scheme, the compressed documents by the whole compared methods are binarized using the Su et al. binarization technique to estimate the readability accuracy for each bitrate. Figure 9 shows the readability evolution in terms of $\mathrm{F}_{\mathrm{Measure}}$, $\mathrm{PF}_{\mathrm{Measure}}$ and DRD across the compression bitrate on the whole tested approaches. Figure 10 depicts some document images obtained by performing the evaluated approaches.

Discussion

We can easily observe from Figures 9 and 10 that varying bitrates from high to low when using the proposed joint-enhancement compression approach keeps an acceptable visual aspect, look, and feel of the original document and does not affect the readability. In addition, the proposed joint enhancement-compression method greatly outperforms the classic DjVu approach as well as the well-known JPEG and JPEG 2000 standards when facing historical documents.

As noticed previously, facing the historical document compression issue with common image compression standards such as JPEG or JPEG2000 or even the specialized on scanned-document compression encoder DjVu seems to be insufficient and inappropriate. Indeed, it is clear from the presented results that readability is corrupted when decreasing the bitrate using classic compression approaches (see Fig. 9 and 10). The limitations of JPEG and JPEG2000 when dealing with historical documents is undoubtedly due to the compression uniformity across image pixels. Within these standards, the encoding process does not differentiate relevant/useless information in the image, so as pixels representing foreground edges are treated as well as the others pix-

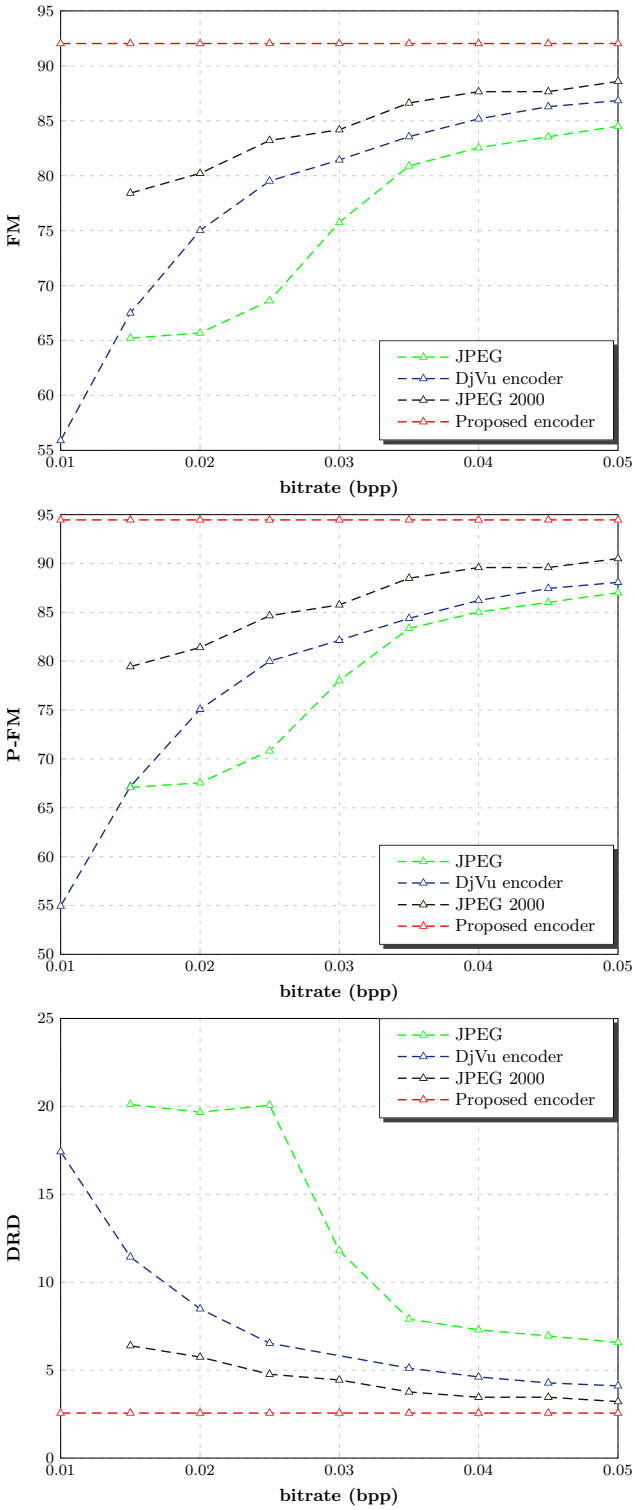
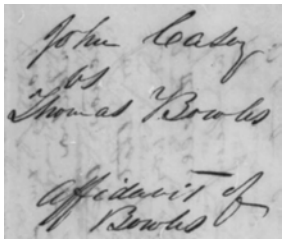
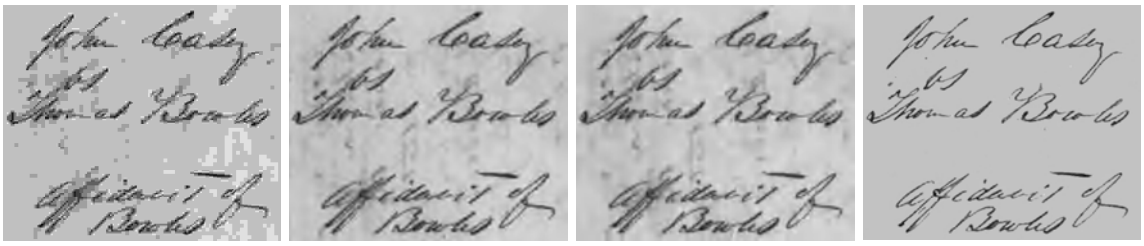


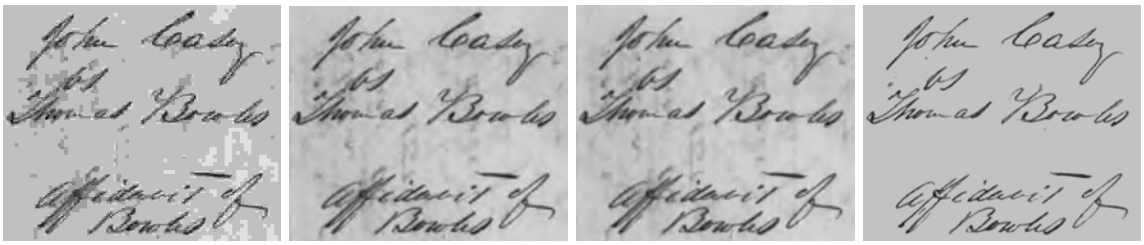
Figure 9: Evolution of FM, PFM and DRD in all DHW document images included within the DIBCO datasets compressed in various bitrates with JPEG, JPEG2000, DjVu and the proposed joint compression-enhancement scheme.



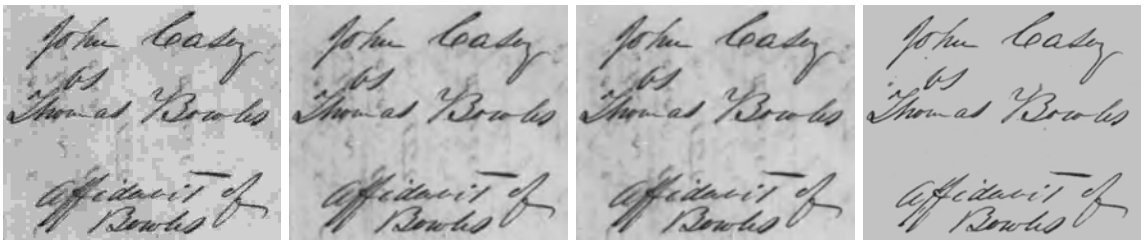
(a): DHW document



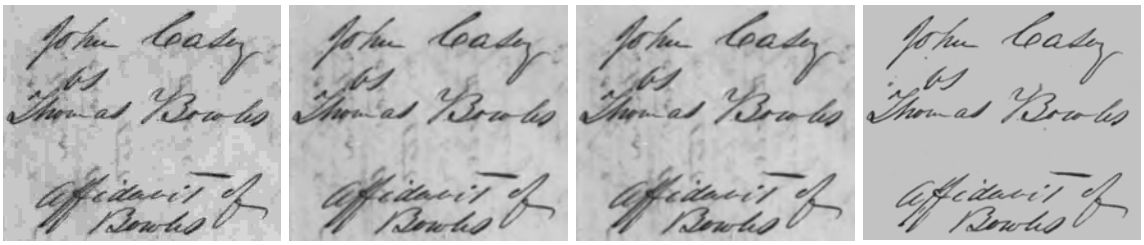
(b): bitrate=0.02bpp



(c): bitrate=0.03bpp

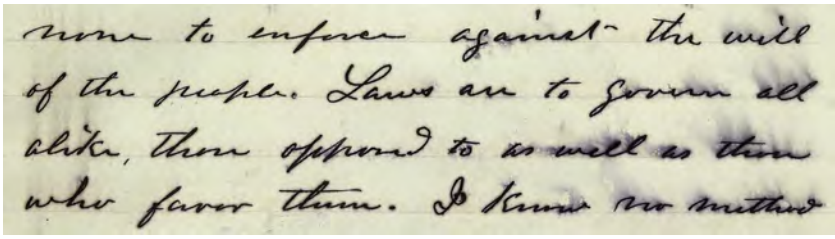


(d): bitrate=0.04bpp

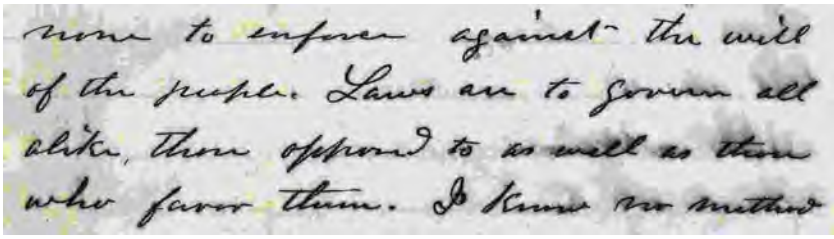


(e): bitrate=0.05bpp

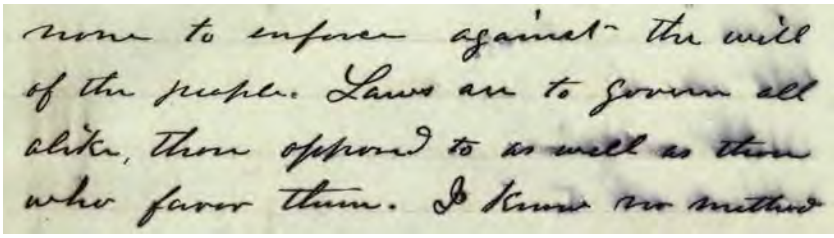
Figure 10: From left to right : Results given by JPEG, DjVu , JPEG2000 and the proposed joint compression-enhancement approach.



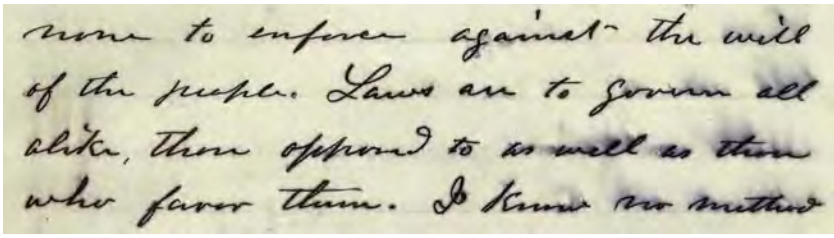
(a): DHW document image



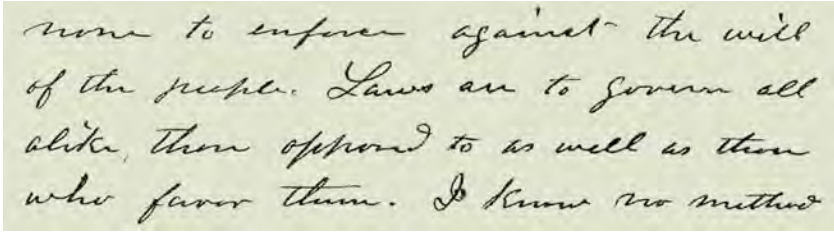
(b): JPEG



(c): DjVu



(d): JPEG 2000



(e): Proposed compression-enhancement approach

Figure 11: DHW document image sample compressed in a medium bitrate.

els; this provides an important degradation to the readability when reducing the compression bitrate. Concerning DjVu, the unsuitable HMM-based F/B segmentation algorithm embedded within this encoder is behind its inability when compressing historical degraded documents. Through the objective evaluation performed in Figure 9, we demonstrate that performing a joint enhancement-compression method is the most appropriate way to solve the degraded documents archiving issue. The segmentation algorithm incorporated within the proposal drives properly the compression process across the document image. Consequently, the visual quality and the historical aspect of the document are fully preserved since we compress slightly across foreground edges while we perform an aggressive compression around artifacts.

Besides, neither degradation relative to the from of characters is perceptible when decreasing bitrates using the proposed approach, since the global compression bitrate depends mostly on the foreground layer bitrate q_{fr} and the binarized version of the document is lossless-compressed and stored in the background layer. Consequently, the readable version of the document is kept whatever the compression rate. In other words, decreasing the bitrate when using the proposed scheme affects only the color tones of the foreground strokes and the readability performance is totally preserved regardless of the compression bitrate. This observation could be clearly noticed by the constant lines obtained through each readability measure in Figure 9. Finally, for medium and higher bitrates (≥ 0.04 in document image compression), the proposed approach beats once again the classic image compression standards since the degradations are still visible and corrupting the readability when performing those approaches; Figures 11 depicts results given by the evaluated approaches on a document image sample extracted from DIBCO datasets.

5. Conclusion

A first method for joint enhancement-compression of historical document images has been proposed. Through this work, it has been shown that by taking into account some specific characteristics of these images, good performance could be achieved. A segmentation guided the enhancement scheme embedded in the DjVu encoder has been proposed and evaluated on real databases. In addition, this solution gives an efficient, and an easy-to-integrate scheme to many national archives and libraries using already the DjVu framework. The performance evaluation

of the proposed scheme has been conducted on the whole DIBCO databases. We have shown that applying the proposed segmentation algorithm, before performing binarization algorithms, allows to circumvent some of their inherent drawbacks. Further, concerning the compression issue, promising results are obtained once again. In fact, both subjective/objective evaluations have demonstrated the efficiency of the proposed joint approach against the most efficient scanned-document image encoder and the most popular image compression standards as shown through Figures 9, 10 and 11. However, both enhancement/compression performance could be improved. This could be achieved by using more distinguishing characteristics of foreground strokes. Reducing the computational time is another issue to be considered in a near future.

Bibliography

[1] P. Haffner, P. Howard, P. Simard, Y. Bengio, Y. Lecun, et al. High quality document image compression with djvu. *Journal of Electronic Imaging*, 7(3):410–425, 1998.

[2] E. Haneda and C.A. Bouman. Text segmentation for mrc document compression. *Image Processing, IEEE Transactions on*, 20(6):1611–1626, June 2011.

[3] M. Sezgin et al. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 13(1):146–168, 2004.

[4] W. Lee and K. Fan. Document image preprocessing based on optimal boolean filters. *Signal processing*, 80(1):45–55, 2000.

[5] G. Ramponi and P. Fontanot. Enhancing document images with a quadratic filter. *Signal Processing*, 33(1):23–34, 1993.

[6] A. Le Négrate, A. Beghdadi, and H. Dupoisot. An image enhancement technique and its evaluation through bimodality analysis. *CVGIP: Graphical Models and Image Processing*, 54(1):13 – 22, 1992.

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

[8] G. Sharma. Cancellation of show through in duplex scanning. In *Image Processing, 2000. Proceedings. 2000 International Conference on*, volume 2, pages 609–612 vol.2, Sept 2000.

[9] C.L. Tan, R. Cao, and P. Shen. Restoration of archival documents using a wavelet technique. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(10):1399–1404, Oct 2002.

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

[11] K. Joo, F. Deng, and M. Brown. Visual enhancement of old documents with hyperspectral imaging. *Pattern Recognition*, 44(7):1461–1469, 2011.

[12] R. Hedjam and M. Cheriet. Historical document image restoration using multispectral imaging system. *Pattern Recognition*, 46(8):2297 – 2312, 2013.

[13] W. Qian, X. Tao, C.L. Tan, and L. Lida. Directional wavelet approach to remove document image interference. In *Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on*, pages 736–740, Aug 2003.

[14] M.R. Yagoubi, A. Serir, and A. Beghdadi. Blind document image enhancement based on diffusion process. In *Visual Information Processing (EUVIP), 2014 5th European Workshop on*, pages 1–6, Dec 2014.

[15] M.R. Yagoubi, A. Serir, and A. Beghdadi. A new automatic framework for document image enhancement process based on anisotropic diffusion. In *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*, pages 1126–1130, Aug 2015.

[16] M. Farrahi and M. Cheriet. Rsldi restoration of single-sided low-quality document images. *Pattern Recognition*, 42(12):3355–3364, 2009.

[17] 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)*, 15(3):183–212, 2012.

[18] 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*, 21(9):940–946, 1999.

[19] J. Weickert. Coherence enhancing diffusion of colour images. *Image and Vision Computing*, 17(3):201–212, 1999.

[20] R. De Queiroz, R. Buckley, and M. Xu. Mixed raster content (mrc) model for compound image compression. In *Electronic Imaging'99*, pages 1106–1117. International Society for Optics and Photonics, 1998.

[21] P. Haffner, L. Bottou, Y. Lecun, and L. Vincent. A general segmentation scheme for djvu document compression. In *ISMM*, volume 2, 2002.

[22] Y. Zheng, H. Li, and D. Doermann. Machine printed text and handwriting identification in noisy document images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(3):337–353, 2004.

[23] John Lafferty, Andrew McCallum, and Fernando CN Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *K. Morgan*, pages 282–389, 2001.

[24] F. Shafait, D. Keysers, and T. Breuel. Performance evaluation and benchmarking of six-page segmentation algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(6):941–954, 2008.

[25] P. Howard, F. Kossentini, B. Martins, S. Forchhammer, and W. Rucklidge. The emerging jbig2 standard. *Circuits and Systems for Video Technology, IEEE Transactions on*, 8(7):838–848, 1998.

[26] K. Panchapakesan, A. Bilgin, M. Marcellin, and B. Hung. Joint compression and restoration of images using wavelets and non-linear interpolative vector quantization. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 5, pages 2649–2652. IEEE, 1998.

[27] E. Dubois and P. Dano. Joint compression and restoration of documents with bleed-through. In *Archiving Conference*, volume 2005, pages 170–174. Society for Imaging Science and Technology, 2005.

[28] J. Emmanuel E. Candes and D. Donoho. Ridgelets: A key to higher-dimensional intermittency? *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1760):2495–2509, 1999.

[29] E. Candes, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, 5(3):861–899, 2006.

[30] L. Demanet and L. Ying. Wave atoms and sparsity of oscillatory patterns. *Applied and Computational Harmonic Analysis*, 23(3):368–387, 2007.

[31] Z. Haddad, A. Beghdadi, A. Serir, and A. Mokraoui. Wave atoms based compression method for fingerprint images. *Pattern Recognition*, 46(9):2450 – 2464, 2013.

[32] F. Liu, L.M. Cheng, H.Y. Leung, and Q.K. Fu. Wave atom transform generated strong image hashing scheme. *Optics Communications*, 285(24):5008 – 5018, 2012.

[33] L. Villemoes. Wavelet packets with uniform time-frequency localization. *Comptes Rendus Mathématique*, 335(10):793–796, 2002.

[34] L. Demanet and L. Ying. Wave atoms and sparsity of oscillatory patterns. *Applied and Computational Harmonic Analysis*, 23(3):368–387, 2007.

[35] B. Su, S. Lu, and C. Tan. Binarization of historical document images using the local maximum and minimum. In *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*, pages 159–166. ACM, 2010.

[36] H. Zhang, J.E. Fritts, and S.A. Goldman. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2):260 – 280, 2008.

[37] A. Beghdadi and W. Souidene. An hvs-inspired approach for image segmentation evaluation. In *Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium on*, pages 1–5, Feb 2007.

[38] N. Otsu. A threshold selection method from gray-level histograms. *Systems, Man and Cybernetics, IEEE Transactions on*, 9(1):62–66, Jan 1979.

[39] B. Su, S. Lu, and C.L. Tan. Robust document image binarization technique for degraded document images. *Image Processing, IEEE Transactions on*, 22(4):1408–1417, April 2013.

[40] C. Tsai and H. Lee. Binarization of color document images via luminance and saturation color features. *Trans. Img. Proc.*, 11(4):434–451, April 2002.

[41] H.Lu, A.Kot, C.Alex, and Y.Shi. Distance-reciprocal distortion measure for binary document images. *IEEE Signal Processing Letters*, 11(2):228–231, 2004.

[42] I. Pratikakis, B. Gatos, and K. Ntirogiannis. Icfhr 2012 competition on handwritten document image binarization (h-dibco 2012). In *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*, pages 817–822. IEEE, 2012.

[43] M.W. Marcellin, M.J. Gormish, A. Bilgin, and M.P. Boliek. An overview of jpeg-2000. In *Data Compression Conference, 2000. Proceedings. DCC 2000*, pages 523–541. IEEE, 2000.

[44] J. Mitchell. Digital compression and coding of continuous-tone still images: Requirements and guidelines. *ITU-T Recommendation T*, 81, 1992.