IJDAR (2007) 9:101–122
DOI 10.1007/s10032-007-0039-z

ORIGINAL PAPER

Hermite and Gabor transforms for noise reduction and handwriting classification in ancient manuscripts

Véronique Eglin · Stéphane Bres · Carlos Rivero

Received: 15 February 2005 / Revised: 26 January 2007 / Accepted: 30 January 2007 / Published online: 27 March 2007

Abstract In this paper, we propose a biologically inspired, global and segmentation free methodology for manuscript noise reduction and classification. Our method consists of developing well-adapted tools for writing enhancement, background noise, text and drawing separation and handwritten patterns characterization with *orientation* features. We have used here analysis of handwritten images in the spectral domain by frequency decompositions (Hermite transforms) and Gabor filtering for selective text information extraction. We have tested our approach of writing classification on ancient manuscripts corpus, mainly composed of 18th century authors' documents. The current results are very promising: they show that our biologically inspired methodology can be efficiently used for handwriting analysis without any a priori grapheme segmentation.

Keywords Handwriting characterization · Patrimonial manuscripts · Background noise reduction · Hermite polynomial decomposition · Gabor filtering · Orientations signature · Similarity measure · Classification · Writer identification

1 Introduction

History is full of people who spent their life writing, for pleasure or obligation, all possible sorts of text. These texts are of great interest because they testify to the lifestyle, the ideas and thoughts of people of those times. These are the reasons why these documents are so precious now, and the subject of so careful attention.

Most of these documents, and the ones that interest us in this work, are made of paper. Unfortunately, even if paper is a proven method to preserve writing through time and space, these documents are very fragile and easily damaged. Because they were handled and manipulated so many times, most of them are in a really bad condition. To avoid more damage, old documents are most often not directly accessible and kept in public or private collections. Converting them to digital formats is a good solution to give access to these documents without more damage, but it takes time and effort, and done little by little and step by step. Various research projects have been undertaken to get things moving. For example, various projects are supported by French national financing to help local libraries to digitalize documents of inheritance. The most famous research projects are based on the development of image processing tools like: BAMBI, DEBORA, Philectre, METAe, DMOS, Agora, etc. Other projects are mainly based on digitization with manual textual annotation and metadata, like the Gallica project from the BNF (you can find more references on the web sites www.bnf.fr or http://www.culture.gouv.fr/mrt/numerisation). Another project, on Word Spotting, has been sponsored by the National Science Foundation Digital Libraries II program. This project develops innovative techniques for indexing handwritten historical manuscripts written by a single author. This work has essentially been developed by Toni M. Rath, R. Manmatha and Victor Lavrenko. We are involved in the project "Culture, Inheritance and

V. Eglin (✉) · S. Bres · C. Rivero
LIRIS–UMR 5205, INSA Lyon, 20 Av. A. Einstein,
69621 Villeurbanne Cedex, France
e-mail: veronique.eglin@liris.cnrs.fr

S. Bres
e-mail: stephane.bres@liris.cnrs.fr

Creation" [1] but the documents that we could collect are not as numerous as we would like, especially for specific studies. In such cases, we cannot treat documents coming from various sources, but need to focus on few documents of a precise origin (e.g., one author or one period).

Documents for our study came from a very recently digitized corpus of handwritten pages from the contemporary period of 18th and 19th centuries. From this period, we have some famous French authors like Montesquieu and Flaubert, whose works are found in rare collections in libraries or specialized institutes. Among those collections, we focused on manuscripts that have been extensively handled, that can also contain multi-writer annotations or corrections, with sometimes background spots or delocalized folds, see Fig. 1.

These documents are the subject of many studies and much research. In our case, we want to identify the authors of some of these manuscripts, or at least identify and group together the manuscripts written by the same author. This identification of author/writer is of great interest because it gives important clues for the comprehension of the genesis of these manuscripts. This is especially the case for collective documents, even if some writers were only copyists or secretaries. A typical example is that of Montesquieu's work. His manuscripts are characterized by great diversity of writers. They were written by more than 20 secretaries having different visual characteristics of handwriting [33]. The pages of this collection can be considered as *drafts* with lots of corrections, cross-outs and scribbles.

Our goal here is to prove that it is possible to characterize handwritings and to classify them into families that share some common visual properties and that give similar visual impression to the reader at first glance or after a short inspection. For that purpose, we chose to extract features on these documents using tools whose principles are linked to the human visual system (HVS) of perception.

It is well known that the HVS codes efficiently visual stimuli. Both neurophysiology and psychophysics support the hypothesis that early visual processing can be described by a set of channels operating in parallel that transform the input signal to obtain a coded version of the stimulus characteristics. This code can subsequently be used as the basis for all kinds of perceptual attributes. It is thus desirable to have a feature extractor that approximates, in such a way, the channels used to describe the HVS.

One of the most famous models of the perceptive fields of the visual system is the family of Gabor filters. Another very interesting model is the family of Hermite filters. We will see in this paper that both families (Hermite and Gabor) are interesting in the context of our application. As good models of the same system, they share common properties, but they also some have differences that we will exploit to perform different tasks.

From the large diversity of handwriting features, Gabor or Hermite filters focus on *the orientations*. We do believe in the relevance of this feature for handwriting characterization, because it reveals both global and local handwriting characteristics. According to our point of view, this is discriminative enough for a well-identified corpus. In practice, we compute a *signature* for each analysed handwritten image and define a similarity measure to compare different samples.

The directional analysis we propose in this paper is based on the exploitation of Gabor filter banks. It could be Hermite based as well, but Gabor filter banks offer more parameterization possibilities that are especially interesting here. Gabor filters are then parameterized to detect relevant handwriting orientations. This is performed through the analysis of the autocorrelation function. Using this *signature*, we create handwriting families (grouping handwritings with similar visual features) for the authentication and the identification of copyists. This approach is a relevant technique to evaluate the traceability of handwritings all over a book or a work and to authenticate different writers who have taken part in the realization of a single book [8,21,26]. In that sense, we are able to make a precise discrimination between writers (different authors who are involved in the writing of a single book, through different periods, from different geographic places and sometimes different historical and social contexts, etc.). We have tested the method on almost 500 handwritten pages coming from 48 different writers.

But before the computation of these characteristics, we try to reduce the noise of the images as much as possible. As explained earlier, the documents we treat are most of the time damaged for different reasons, and these degradations are mainly visible on paper (the background of the image). As this background represents the most important part of the image, its influence with the computed signatures should be as weak as possible in order to characterize the authors by their handwritings and not by the type of degradation we see on their manuscripts. We shall demonstrate the interest of Hermite decompositions for that purpose. Hermite

[1] Project "Digitization and Documents Recognition" in the field of a regional project "Cluster Région Rhône Alpes: culture, Inheritance and Creation" in collaboration with literary partners, Lyon (2006).

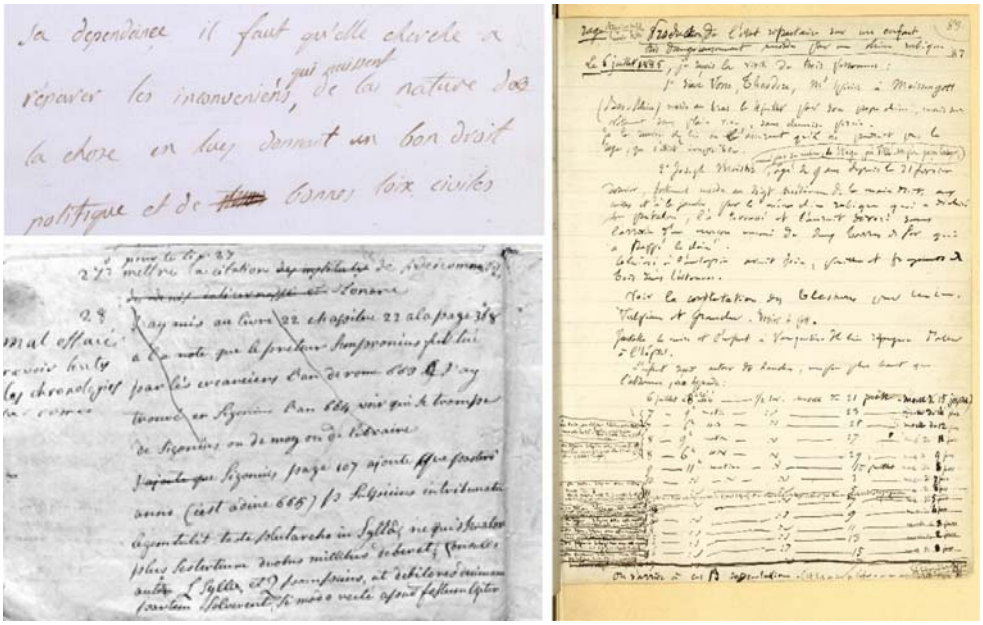

Fig. 1 Examples of old manuscripts. Montesquieu's autograph (De l'Esprit des Lois 1789). Montesquieu's secretary (1780), Bordeaux Library. Pasteur correspondences [11] (GALLICA, BNF[2])

decompositions are reversible and allow a reconstruction of the image after filtering (modifications applied directly on the decomposition). Gabor filters do not perform a reversible decomposition and they are more interesting for analysis purposes than reconstruction after filtering. These decompositions are based on sets of Hermite filters that analyse the image in the frequency domain where signal (writings) and noise (damages) can be more easily identified. After noise reduction, the image is reconstructed (see Sect. 3).

2 Existing approaches for handwriting analysis and our contribution

In the field of writer identification, it is *neither* required to transcribe the texts *nor* to recognize word content, because the graphical aspect of the written shapes is the focus of all attention.

2.1 Existing approaches

In that context, we mainly find two types of approaches: the first type is based on the definition of structural local features that mainly describe structural properties of the shapes, like the height, the width or the legibility of characters [19], or gradient, curvature of the contours and concavity-based features [5]. For this kind of approach, it is mainly required to segment the text in characters or graphemes or to localize it precisely. The graphemes are actually elementary patterns of handwriting, extracted by a segmentation algorithm. Nosary et al. [21] proposes a characterization of different levels of graphemes based on the analysis of the minima on their upper contour.

Some authors [4–6] propose to detect some typical handwriting invariants (presence of characteristic loops, typical words cuttings, recurrent orthography errors, etc.). In [6] and in [24], Srihari et al. build their analysis on a PCA-based approach with a set of macro and micro-features that characterize the handwritings. The elements that capture the local characteristics of the writer's individual writing style are regarded as micro-features and correspond to gradient, concavity and structure. In [4], Bulacu and Schomaker propose a system that evaluates edge-based directional probability distributions extracted from handwriting contours for the uppercase and lowercase handwriting discrimination.

But, usual handwriting segmentation approaches become inefficient in ancient degraded corpus (connected components analysis, directional filtering like RLSA, lines and columns projection profiles [15]). Moreover, a thresholding step degrades the handwritten regions by merging words and text lines together. In that context, the existing segmentation methods show their limits because handwritten text areas are sometimes multi-oriented or not written on straight lines. We can

[2] GALLICA, digital library of the BNF (Bilbiothèque Nationale de France).

find also marginal annotations and irregular body paragraphs (see Fig. 1). These characteristics lead to unpredictable page layouts that cannot be modelled by any formal representation technique. The second difficulty deals with the irregularities of handwritten shapes, which can have small interline spaces and frequent word contacts. The separation of text and non-text areas becomes also difficult in case of insufficient pen pressure. Some techniques have very good accuracy, like Zois's work [37] where the author reports a correct writer identification performance of 92% among 50 writers by using 45 samples of the same word from modern and not degraded manuscripts. Precision could not be found in the context of degraded source images.

A second type of approach for writer identification is based on global features that are based on statistical measurements, and extracted from whole blocks of text. This kind of approach is usually based on the extraction of *features from texture*: in this case, the manuscript is considered as a whole image and not as handwriting, and is not segmented into characters. We can give as examples, the work of Kuckuck [14], who considers handwriting as a texture with visually strong properties, broadly used by human experts. Some texture-based approaches have also been later developed by Said [25] with multi-spectral text images decomposition and co-occurrence matrix. In this work, the authors report a correct accuracy of 95% on 40 writers with only some handwriting text lines and on clean handwriting images.

Finally, we can notice that it is also possible to combine the two different types of approaches, as proposed by some authors like Catalin et al. [5]. In their previous works [28], they present global statistical macro-features at the document level and micro-features at the character-level.

The evaluation of the performance of those works is quite difficult to establish because the conditions and parameters for the tests (number of writers, volume of the corpus, existence and size of the training set, size of the tested image samples, etc.) are broadly different. Nevertheless, some studies have been proposed these last years [2] to compare global and local approaches for writer identification. They have proposed to categorize writer identification works according to the number of writers and the nature of the training samples (whole text pages, paragraphs of text, single lines and few words). As a conclusion, we retain that the best performances belong to the systems using a great number of writers, with consequent training sets and a significant number of samples. Srihari et al. [28] currently holds the best performance results with more than 1,000 writers and with the same text samples three times written by each hand.

2.2 Our contribution

Basically, the accuracy of writer identification techniques depends on the above parameters and also on the image quality, and we are working on degraded documents with mostly bad appearance and noisy background. In that context, we propose an alternative that combines a reduction noise step at the beginning and, afterwards, an analysis step that takes advantage of both global feature extraction and local shape analysis. Indeed, the originality of our approach comes from the consideration of both texture properties of handwriting and local oriented variations along pattern contours.

Because degraded handwritten documents cannot be easily segmented into lines or words, we have chosen a "segmentation-free" method which does not need any separation of characters or graphemes.

The global scheme of our proposition is illustrated in Fig. 2. All steps of the scheme are described in the next sections: Sect. 3 is dedicated to the Hermite-based noise reduction and Sect. 4 presents the Gabor filtering for handwriting characterization.

3 Noise reduction step

3.1 Existing approaches

Many digital images of documents and more generally ancient manuscripts are degraded by the presence of strong artifacts in the background. This can either affect

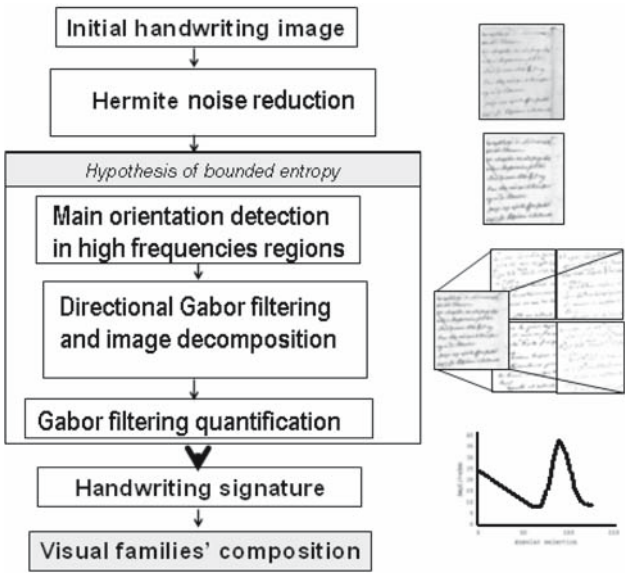

Fig. 2 Global scheme of image noise reduction and handwriting characterization

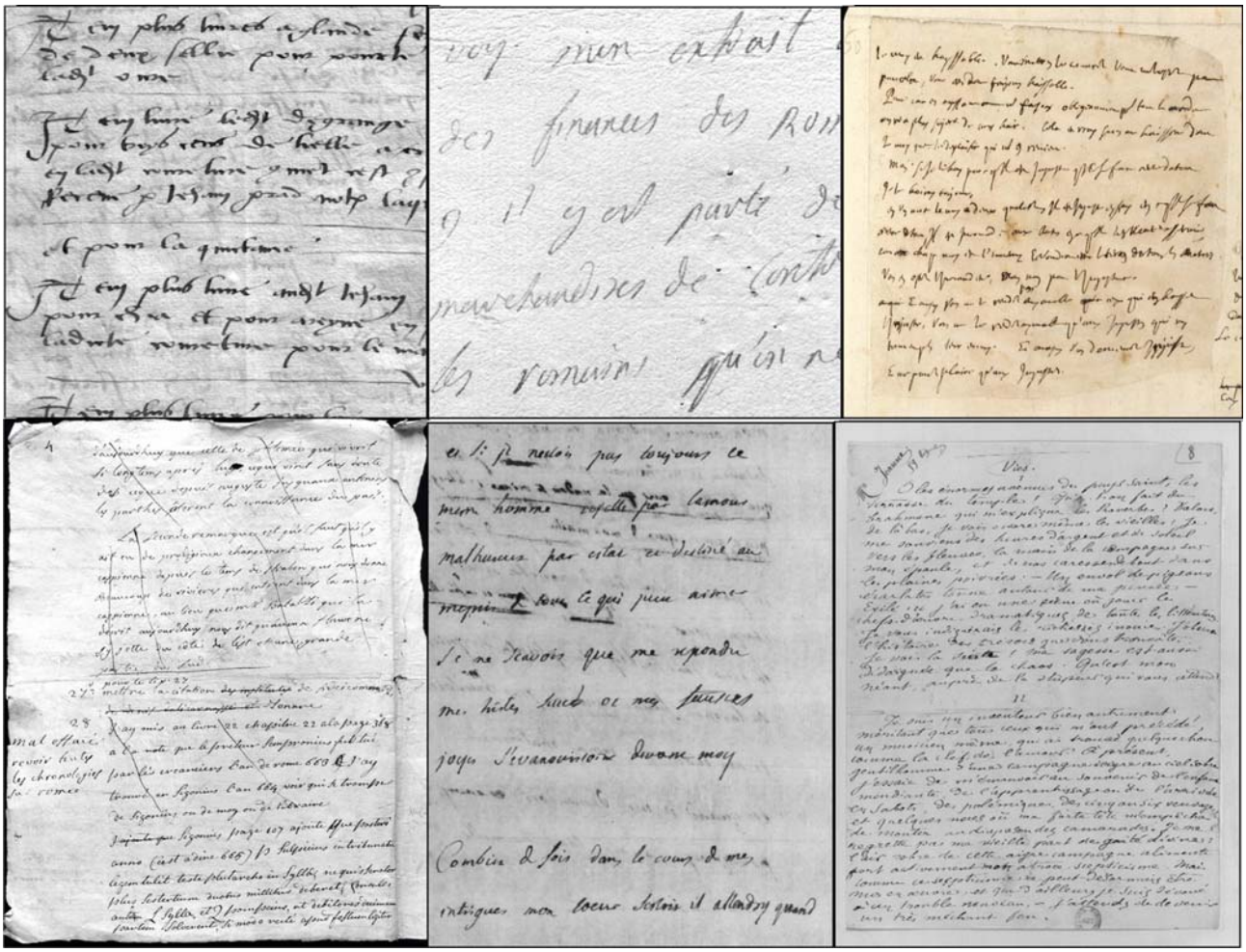

Fig. 3 Examples of handwritings with typical noisy background (multi-spots, ink on the reverse side and granularity of the support) [32]

the readability of the text or, in some cases, the relevance of our handwriting characterization. Background artifacts can arise from many different kinds of degradations such as scan optical blur and noise, spots, underwriting or overwriting, time wearing, intensive use or bad preservation conditions (see Fig. 3). All these degradations create dark areas most of the time, with more or less uniform colours and different sizes.

Among all the possible degradations of the background, the visibility on the reverse side of the page (or bleeding of ink), ink degradation (attenuation of the ink marks which affects correct text reading) or palimpsests (an earlier text was erased and another writer had reused the vellum or parchment) were especially studied in literature. The purpose here is to reduce the influence of these damages, included in the background, and to highlight the handwriting, which is the interesting and informative part of the document. More generally, we can consider documents as a combination of a textured background and a handwriting signal or foreground. A lot of different approaches exist to extract this foreground part of the signal. The most naïve methods use thresholding techniques. A comparative study of global thresholding appears in Leedham et al. [17] for text and background separation. The authors conclude that "no single global thresholding algorithm/scheme can work with degraded document images". The main reason is the peculiar characteristics of these images (varying background/foreground intensities, varying contrast, bleeding of ink from the other side of the page, etc.). Other techniques based on adaptative filtering have been tested on forensic documents to separate homogeneous textured background from handwriting marks [9]. But for [17], these local adaptative thresholding algorithms are not effective with these types of images.

Many authors consider the specific problem of bleed-through handwriting pages, which is very often encountered in ancient documents: writings on the backside are

visible through the paper on the front side [20,26]. Most of the time, the backside image is not available. Some approaches consider a physical model of degradation for text enhancement and background cleaning [26]. In Tonazzini et al. [32], the authors propose to decompose the signal into two blind sources where the overlapping texts and the supports (paper) texture are the unknown sources to be recovered with different spectral bands of the documents. Fairly recently, we can find in Nishida and Suzuki [20], a method of frontside/backside separation using colour information by local adaptative thresholding. Here again, the backside image is not needed. Finally, recent research uses wavelet analysis to perform foreground extraction [30,31]. In Tonazzini et al. [31], the decomposition of wavelets is used to filter the high frequencies containing most of the background noise energy. This filter process is based on the estimation of the noise level in the background.

This approach is very close to the one we use. Our proposition uses a decomposition of the original image by Hermite transforms, which cannot be considered as a wavelet transform (even if it realizes a localized frequencies decomposition), because it realizes an overcomplete decomposition. Overcomplete decomposition means that redundant information is present and that is not the case for wavelet decompositions. Before the explanation of our noise reduction process, we will present the Hermite transform.

3.2 The Hermite theory

Basically, a polynomial transform locally decomposes a signal into a set of orthogonal polynomials. A local version of the signal is computed through a multiplication of the whole signal by a window W (signal null outside a given interval). Each point of the signal can be reached through translations of the window W. In the following subsection, we present in detail the definition of the Hermite transform and its discrete representation. This discrete representation, called Krawtchouk transform, is the one we use in practice in our treatments.

3.2.1 Cartesian Hermite filters

We present the definitions of Hermite filters, which agree with the Gaussian derivative model of the HVS [18,24]. We will focus on the Cartesian representation, which is more oriented to extract spatial primitives such as edges, lines, bars and corners, into the vertical, horizontal and oblique directions rather than oriented textures. However, this is similar to Gabor filters [18], which are more often used, for texture, in image processing and feature extraction. Hermite and Gabor filters are equivalent models of receptive field profiles (RFPs) of the HVS [18,24]. Hermite filters $d_{n-m,m}(x,y)$ decompose a localized signal $l_v(x-p,y-q) = v^2(x-p,y-q)l(x,y)$ by a Gaussian window $v(x,y)$ with spread σ and unit energy, defined as

$$v(x,y) = \frac{1}{(\sigma\sqrt{\pi})\mathrm{e}^{-(x^2+y^2)/(2\sigma^2)}} \tag{2.1}$$

into a set of Hermite orthogonal polynomials $H_{n-m,m}(x/\sigma, y/\sigma)$. Coefficients $l_{n-m,m}(p,q)$ at lattice positions $(p,q) \in P$ are then derived from the signal $l(x,y)$ by convolving with the Hermite filters. These filters are equal to Gaussian derivatives where $n-m$ and m are, respectively, the derivative orders in x- and y-directions, for $n = 0,\ldots,D$ and $m = 0,\ldots,n$. Thus, the two parameters of Hermite filters are the maximum derivatives of order D (or polynomial degree) and scale σ.

Hermite filters are separable both in spatial and polar coordinates, so they can be implemented very efficiently. Thus, $d_{n-m,m}(x,y) = d_{n-m}(x)d_m(y)$, where each 1D filter is

$$d_n(x) = \left((-1)^n \Big/ (\sqrt{2^n \cdot n!}\sqrt{\pi}\sigma)\right) H_n(x/\sigma)\mathrm{e}^{-x^2/\sigma^2}, \tag{2.2}$$

where Hermite polynomials are represented by $H_n(x)$, which are orthogonal with respect to the weighting function $\exp(-\mathrm{x}^2)$, and are defined by Rodrigues' formula:

$$H_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{d^n}{dx^n}\mathrm{e}^{-x^2}. \tag{2.3}$$

In the frequency domain, these filters are Gaussian-like band-pass filters with extreme value for $(\omega\sigma)^2 = 2n$ [18,24], and hence filters of increasing order analyse successively higher frequencies in the signal.

3.2.2 Discrete implemention: Krawtchouk filters

Krawtchouk filters are the discrete equivalents of Hermite filters. They are equal to Krawtchouk polynomials multiplied by a binomial window:

$$v^2(x) = \frac{C_N^x}{2^N}, \tag{2.4}$$

which is the discrete counterpart of a Gaussian window. These polynomials are orthonormal with respect to this window and are defined as [18]:

$$K_n(x) = \frac{1}{\sqrt{C_N^n}} \sum_{\tau=0}^{n} (-1)^{n-\tau} C_{N-x}^{n-\tau} C_x^{\tau} \tag{2.5}$$

for $x = 0,\ldots,N$ and $n = 0,\ldots,D$ with $D \leq N$.

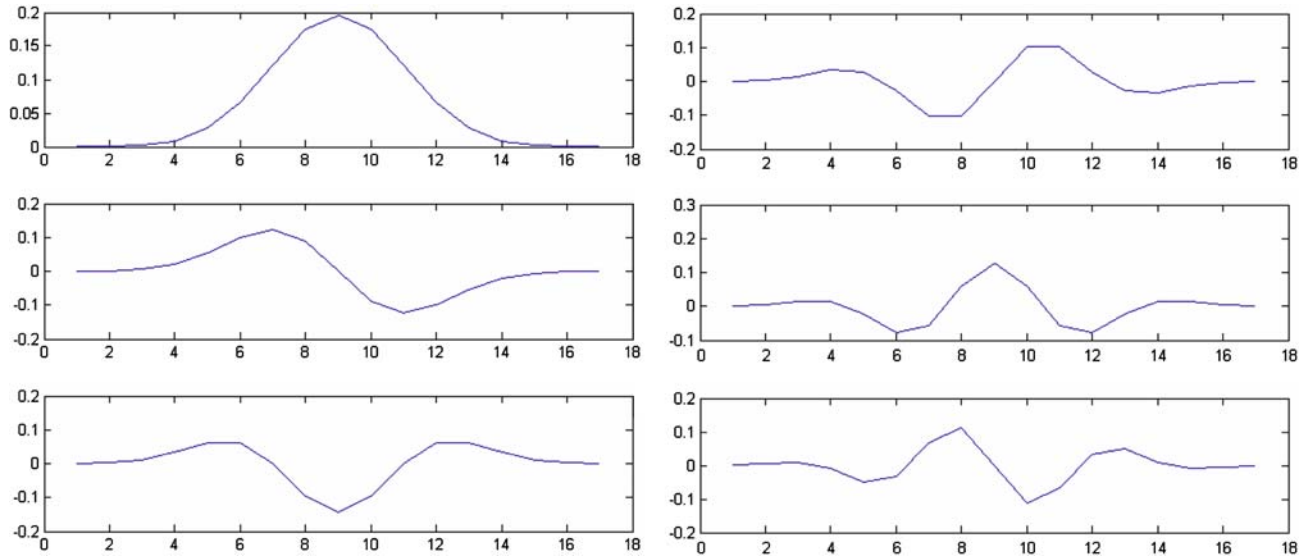

Fig. 4 1D Krawtchouk filters for $N = 16$ and up to degree $n = 5$

It can be shown that the Krawtchouk filters of length N approximate the Hermite filters of spread $\sigma = \sqrt{N/2}$. In order to achieve fast computations, we present a normalized recurrence relation to compute these filters

$$K_{n+1}(x) = \frac{1}{\sqrt{(N-n)(n+1)}} \left[(2x-N)K_n(x) - \sqrt{n(N-n+1)}K_{n-1}(x)\right] \tag{2.6}$$

for $n > 0$ and with initial conditions

$$K_0(x) = 1, \quad K_1(x) = \frac{2}{\sqrt{N}}\left[x - \frac{N}{2}\right].$$

In practice, we use these Krawtchouk polynomials and filters to compute the Hermite transform. These formulas lead to a set of orthonormal filters of length N that decomposes a local area of the original signal in the frequencies domain (see Fig. 4). We apply a translation T to the set of filters to treat the next (and neighbour) area and so on, on the whole signal. This way we have a frequencies decomposition of every treated area.

It is possible to have an exact reconstruction of the original signal if the translation T leads to overlapped areas. All the filter results (from $n = 0$ to N) are then necessary. With a minimal overlap, there is no redundancy in the decomposition. With a higher overlap (lower value for T), some information is redundant, and the decomposition becomes overcomplete. Then, it presents smoother proprieties and gives a more continuous analysis of the signal.

3.2.3 2D version of Krawtchouk filters

These formulas can be generalized in 2D to obtain the filters we use in practice on images. The 2D discrete Hermite transform is built on the Krawtchouk 1D filters, using the separability property. Consequently, the parameters N and T can be chosen independently of the rows and the columns, and are not necessarily equal. In our document application, we use a longer weighting window to treat the rows because of the word shape, which is most of the time larger in the horizontal direction. Figure 6 presents an example of discrete Hermite transform (or Krawtchouk transform) of the image of Fig. 5. As we described earlier, this is an overcomplete transformation because of the redundancy introduced by the undersampling parameter T. This property allows a much smoother analysis and reconstruction of the original signal or image after filtering, without block effects or discontinuities that we observe in wavelet-based reconstructions.

3.3 Application of Hermite to handwriting image noise reduction

Most of the time, the noise or degradations we can see on ancient documents have lower frequency characteristics, while the writing by itself is composed of higher frequencies. Degradations coming from darker areas, spots or even writings visible from the backside have a more blurred aspect than the writings from the font side. Thus, they contain lower frequencies. Moreover, degradations have a smaller contrast than the writing on

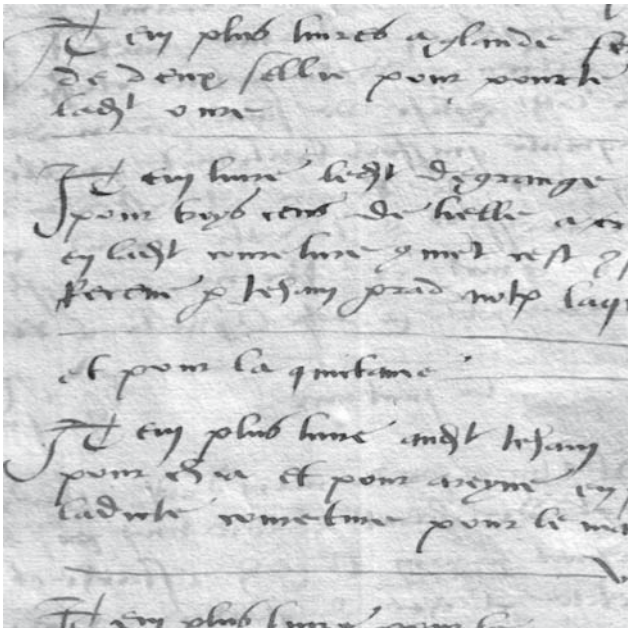

Fig. 5 An example of handwritten document with different kinds of noises or damages

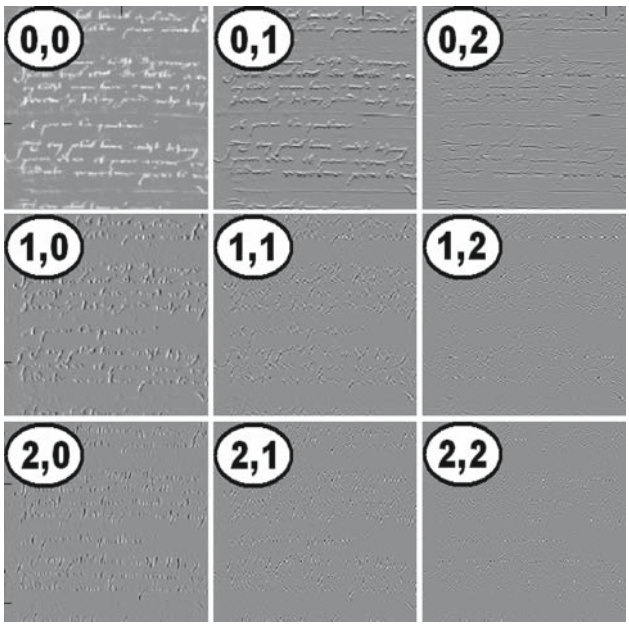

Fig. 6 2D Krawtchouk filters for $N = 6$ and up to degree $n = 2$ for the rows and the columns

the front side. This is also an important characteristic. If we take into account even higher frequencies with a sufficiently high level, we can extract the contours of the writings. The threshold value depends on the original contrast. These are the hypotheses we assume to distinguish handwritings on the front side from the noise on the background. If this noise does not verify these characteristics (like cross out on the writings, for example), our method would give poor results. Fortunately, the documents we treat mainly contain noise of this type.

Consequently, if we assume that high frequencies with a sufficiently high level represent the contour of the writings we want to keep, and lower frequency regions (in background regions) mostly contain degradations, it is interesting to separate them. Keeping the first one and suppressing the second one will then restore the degraded page. This is exactly what we can achieve with the Hermite transform. The size of the localization window selects the range of frequencies used for the analysis. The smaller the size of the window, the higher the analysis frequencies are.

Figure 6 presents the Hermite decomposition of a document at a given scale $N = 6$ and up to degree 2, and with an undersampling parameter $T = 3$. The quadrant (0,0) is equivalent to a Gaussian filtered image using a $(N+1) \times (N+1) = 7 \times 7$ filter and undersampled by three (one sample for each window position). The other quadrants correspond to frequencies analysed in that 7×7 window. Consequently, the complete decomposition contains 7×7 quadrants. The analysed frequencies are thus relatively high frequencies of the original image. Middle grey values correspond to zeros, darker values are negative and bright values are positive.

Our noise reduction process uses the Hermite decomposition. In a first step, we localize the writing areas using the energy contained in quadrants $(1,0)$ and $(0,1)$ (see Fig. 6). This information is very close to gradient energy. The second step uses the normalized energy map M (values between 0 and 1) as a mask to filter all the quadrant of the decomposition. The normalized energy map gives, for each position, a probability to contain writing. An example of normalized energy mask is given in Fig. 7a. The coefficients of the Hermite decomposition are then thresholded using the following method:

$$C_{i,j}(x,y) = \begin{cases} \operatorname{sign}(C_{i,j}(x,y)) \cdot (\left|C_{i,j}(x,y)\right| - \bar{\sigma}_{i,j}(x,y)) \times K_{i,j}, \\ 0 \quad \text{if } \left|C_{i,j}(x,y)\right| < \bar{\sigma}_{i,j}(x,y) \end{cases} \tag{2.7}$$

and

$$\bar{\sigma}_{i,j}(x,y) = \sigma_{i,j} \cdot (1 - M_{i,j}(x,y)),$$

where:

- $C_{i,j}(x,y)$ is the coefficient at position (x,y) in the quadrant (i,j),
- $\sigma_{i,j}$ the weighted level of the noise in the quadrant (i,j), estimated in an area that does not contain writings,

- $M_{i,j}(x, y)$ the normalized coefficient at position (x, y) in the mask quadrant of gradient energy,
- $K_{i,j}$ is a normalization coefficient to keep the original maximum level to its original value in the quadrant (i, j).

In the third step, we reconstruct the image using the thresholded Hermite quadrants. We obtain an image with a cleaner background. Figure 7b presents an example of the image on Fig. 5 after noise reduction by Hermite reconstruction. A more detailed view is presented in Fig. 7d. Another example, coming from an artificial degradation of image 5 is presented in Fig. 7e and f. In this case, simulations of ink spots were added. The main part of this method is based on adaptative thresholding steps. Nevertheless, the results we obtain are different from the results of adaptative thresholding methods, because these adaptative thresholding steps are applied independently on each quadrant, i.e., on each frequency domain. Consequently, smooth edges with low contrast, like those coming from backside writing are treated differently than sharp edges with higher contrast. In that point of view, our method has some similarities with wavelet-based methods.

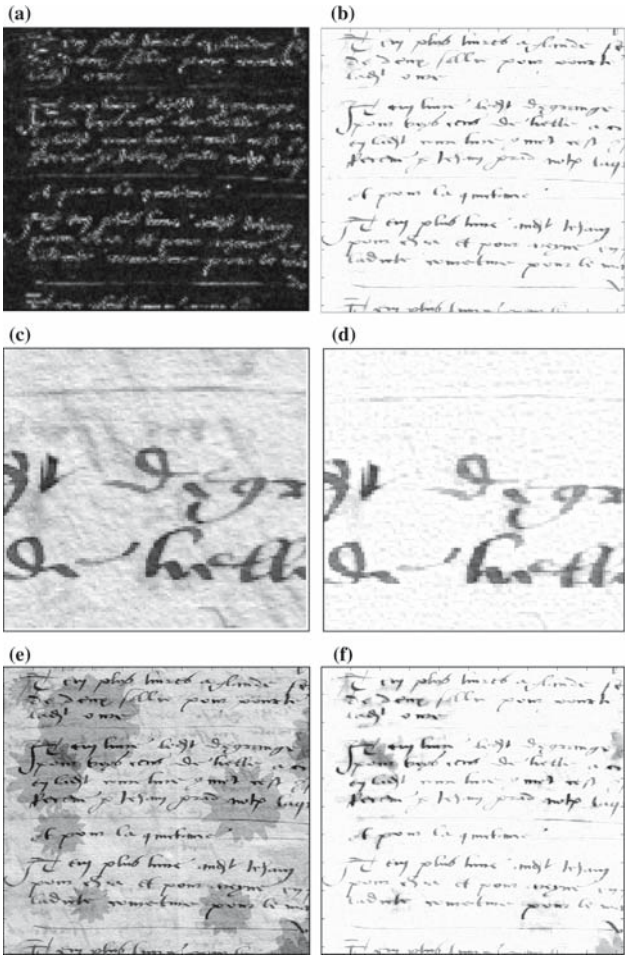

Fig. 7 Example of document noise reduction. Energy mask used to localize writings (**a**), denoised document (**b**), details of an original document (**c**), details of a denoised document (**d**), another example with artificial noise effect (**e**) and its denoised version (**f**)

The main interest of this noise reduction step in the context of this paper is to suppress as much as possible the information coming from the background, which will modify the following analysis. An image with reduced noise allows focusing on the handwriting by itself.

4 Handwriting signature using Gabor filters

4.1 Principle of the Gabor-based approach

In this part, we develop an original method for the classification of handwritings in visual separable families. Writer identification is the task that consists in determining the author of a given document. In this case, it is essential to repeat individual verification between the tested sample and all individual identified handwriting families among known writers [5]. The first step in the process consists in defining a similarity measure to compare two handwritings. The second step consists in taking a decision, which must answer the following question: "Is there any intra-class stability (the within-writer stability) and does any visible difference between the tested handwriting and the training set exist?" To be sound and consistent, the similarity measures must minimize wrong acceptations and wrong rejections.

Writer identification needs a characterization as relevant as possible of the graphical proprieties of its handwriting. It is difficult to pretend to be exhaustive in such descriptions. We chose a biologically inspired approach using adapted Gabor-based filtering. Thus, the feature we focus on is *the orientation*, which expresses both global and local handwritings proprieties. In this analysis, we evaluate the ratios, which exist between the main orientations found in the handwriting shapes. In that way, two distinct handwritings, even identically skewed, will not be considered as similar because all other detected orientations will show significant differences. Of course, for a word-based writer identification, as proposed by Srihari et al. [28], the orientation is not a sufficient dimension and must be completed by other features.

Our method lies on the evaluation of a compact *signature* for each handwriting. The signature is obtained by the estimation of Gabor filter coefficients, which reveal the presence of salient orientations. The Hermite-based noise reduction process is essential here to separate the background information to the writing

itself. Nevertheless, the orientation is no more relevant if the background is textured too much and if it contains too many oriented noisy strokes, or when the samples contain very badly written texts with too many irregularities (non-constancy of a same writer). But those limitations are also valuable for other techniques.

A small written text with great stems and downstrokes does not present the same orientation distribution as a curved handwriting. The approach is based on image frequencies decomposition with the application of directional Gabor bank filters. The frequencies decomposition is based on the detection of most regular directions obtained by the application of the autocorrelation function on a sample of the entire initial image, which contains a minimum of five handwritten lines. This image sample must be as homogeneous as possible: it is chosen so as to contain normalized text density and bounded entropy (this measure is presented in Sect. 4.2). The selection of the sample has been automated for all digitalized pages of the corpus. According to this process, the Gabor analysis needs homogeneous handwriting image samples with no empty areas or noisy strokes line regions. For a given image, Gabor filters are computed in all significant directions of the handwriting. In this work, the scale of Gabor filters bank is constant in order to produce readable results (a good compromise between a too blurred response and a not significant filtering; see Fig. 12).

The directional Gabor filters produce directional maps that reveal oriented patterns (graphemes). For each θ-direction, these graphemes are then quantified by a density measure that reveals the contribution of the θ-direction in the handwriting. The validity of the approach lies on a within- and a between-writer stability analysis.

4.2 Initial hypotheses of handwriting density

The analysed handwriting blocks must contain quantitatively significant handwritten patterns that are estimated by two extreme entropy values: a *minimal* entropy value E_{MIN} and a *maximal* one E_{MAX}. That means that the image must contain a significant number of text lines to be exploited by the method and inversely it must not contain numerous black strokes (often visible in draft pages, which the noise reduction step cannot suppress; see Fig. 9).

The entropy is directly correlated to the visual impression of "complexity" that we have during the observation. A text made of small letters seems more "complex" than a text with big letters. Our study quantifies this complexity with a measure of entropy. Practically, we compute the number of transitions from the background to the text that can be found on random oriented lines, which leads to the estimation of transition probability occurrence on a pixel for each horizontal line. We only keep the maximum probability p in a considered text block T_{B} because it is representative of how complex the analysed text block can be (or the grapheme in a reduced analysis scale). The Entropy $E(T_{\mathrm{B}})$ is then defined for each block by the following formula:

$$E(T_B) = p\,Log\frac{1}{p} + (1-p)Log\frac{1}{(1-p)}. \tag{3.1}$$

Figure 8 presents the hierarchy of entropy values that are estimated in a set of representative handwriting images of the Montesquieu's corpus.

In practice, an initial handwriting image must contain no less than five text lines to be interesting, with a minimal entropy value corresponding to the threshold $E_{\mathrm{MIN}} = 0.18$ and a maximal value corresponding to $E_{\mathrm{MAX}} = 0.6$. These two values have been selected so as to keep significant homogeneous samples and to reject unreadable draft blocks. In each original handwriting page, we have also considered a $\mathrm{W_I}$ *window* with an entropy value that verifies the condition $E_{\mathrm{MIN}} < E_{\mathrm{W_1}} < E_{\mathrm{MAX}}$. This condition is automatically applied to all tested samples because it has been shown that it was difficult in practice to normalize ancient handwritten texts (in size and in density) and that a normalization process could be responsible for irreversible visual damages in handwritten patterns [26]. The $\mathrm{W_I}$ cutting is based on a first global density estimation (naively computed as the ratio between dark and light points): for an original entire page, we consider $2n \times 2m$ sized samples that recover the initial page area—the samples can be square ($n = m$) or rectangular ($n \neq m$)—and we

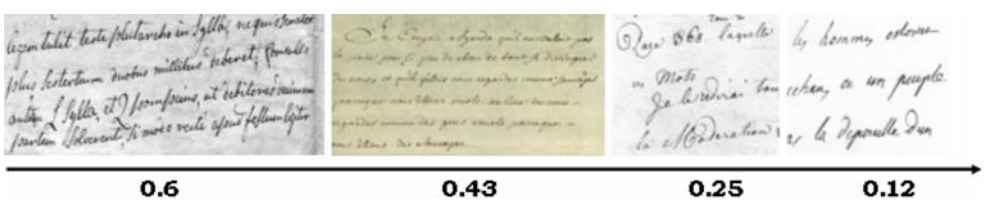

Fig. 8 Linear scale of complexity for significant handwriting samples

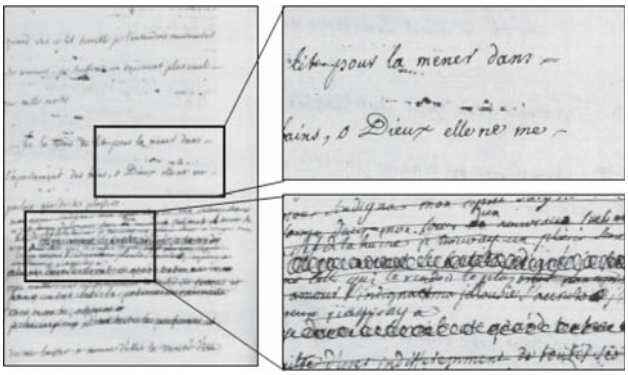

Fig. 9 Entire handwritten page and selected 256×512 samples with weak ($E = 0.12$) and high ($E = 0.75$) entropies

only keep *the* ones that have a density equal or superior to the entire image density. If this condition cannot be satisfied, we consider the whole image. Within this principle, we keep away quite empty background areas in the frontier of the main text. Finally, we keep *the* window that has the maximal entropy value and that also verifies the condition $E_{\mathrm{MIN}} < E_{\mathrm{W}_1} < E_{\mathrm{MAX}}$. There are mainly several samples that verify the conditions, but we chose only one for the tests. In conclusion, we allow comparing two handwriting samples only if they have near Grey level densities and entropy values included in the $[E_{\mathrm{MIN}}; E_{\mathrm{MAX}}]$ interval. In that way, comparison is possible even if the sample sizes are not equal. Figure 9 shows examples of page samples that were rejected because they contained too weak or too high entropy.

4.3 Selection of salient handwriting directions

The salient directions are highlighted by a directional rose computed through the *autocorrelation function*. This function correlates the image with itself, highlights periodicities and orientations of texture. It has been widely used in a context of texture characterization [7,29]. The definition for a bi-dimensional signal is:

$$C_{II}(i,j) = \sum_{i'=-\infty}^{+\infty} \sum_{j'=-\infty}^{+\infty} I(i',j') \cdot I(i'+i, j'+j). \tag{3.2}$$

The autocorrelation function $C_{II}(i,j)$, applied to an image I, combines image I with itself after a translation of vector (i,j). The different translations that are considered by the function give information on the different directions of the image. The data relative to the same direction will be located in the same line. With this principle, it is possible to detect orientations of the texture blocks. For example, the translation of a line in its direction leads to a complete correspondence and is expressed by a great value of autocorrelation in the line direction. In the orthogonal direction of this line, the resulting value will be low. The autocorrelation underlines objects overlapping obtained by translation. This principle can be generalized to a set of objects having a common direction: in our work, we use it to show that text lines can be characterized by the horizontal direction and can be also considered with a possible skew variation. The autocorrelation result can be analysed by the construction of a corresponding directions rose. This rose gives, with a great precision, the main orientations of the block. In Bres [3], we propose an approach for directions rose computation. It is based on the mean value that is computed from the autocorrelation result. Let us consider I, the image block, and $\{(x,y)\}$ the set of coordinates in this image. We also consider θ as a salient direction of the block. The mean value E_θ is then defined by the following formula:

$$E_\theta = \sum_{X=1}^{N} \sum_{Y'=1}^{N} I(x,y) \cdot I(x+a, y+b) \quad \text{with } \mathrm{Arctan}(b/a) = \theta. \tag{3.3}$$

E_θ is the mean value of consecutive products $I(x,y) \times I(x+a, y+b)$ for all sets of couples (x,y) and values a and b that verifiy Arctan $(b/a) = \theta$. So, a point $C(a,b)$ in the autocorrelation function contains the sum of grey level products of overlapping points after a translation (a,b). The autocorrelation function gives values that are proportional to this mean value E_θ. The directional rose represents the sum $R(\theta_i)$ of different values $C_{II}(i,j)$ (defined in 4.2.1) in a given θ_i-direction. So, the directional rose corresponds to the polar diagram where each direction θ_i is represented by the sum $R(\theta_i)$. For all points (a,b) in the θ_i-direction along the $D_{\theta i}$ line, we have the following relation:

$$R(\theta_i) = \sum_{D_{\theta i}} C_{II}(a,b). \tag{3.4}$$

From this set of values, we only keep relative variations of all contributions for each direction. So, the relative sum $R'(\theta_i)$ is the following:

$$R'(\theta_i) = \frac{R(\theta_i) - R_{\min}}{R_{\max} - R_{\min}}. \tag{3.5}$$

From the polar angular rose representation, we decide to keep only significant directions and to neglect microscopic orientations, which are naturally present in the background image (even after the Hermite process). The principle of significant directions extraction lies on the location of rose petals centres (local extreme amplitude values that are greater than the extreme average), neglecting all secondary non-interesting orientations.

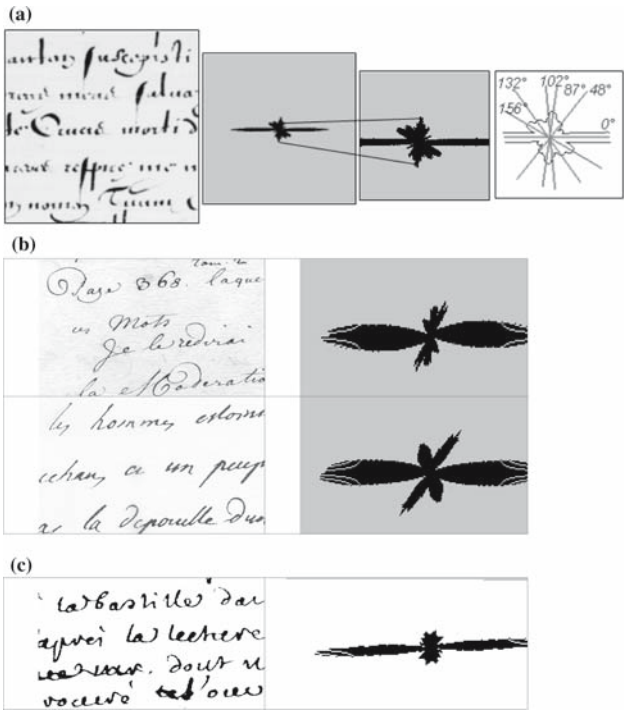

Fig. 10 Directional rose and zooms in significant directions of the rose petals (**a**). Visual differences between two copyists (**b**). Regular distribution of orientations for a curved handwriting (**c**)

This technique reveals the visual differences existing between writings in only the significant directions. Especially here, the presence of loops and curves is highlighted by a regular orientation distribution (except in the 0°-direction that represents the horizontal text lines) as different from most compact and irregular handwritings (see Fig. 10). From this representation, we only keep the eight most significant values in the interval $[0°, 180°]$. The relevance of those significant directions is linked to their local amplitudes (which are directly determined in the rose) and the area quantization of the corresponding Gabor filter responses in the significant directions (see next section).

Gabor quantization is mainly necessary to express the *real* contribution of the horizontal direction (near 0°). This direction is often overvalued in the tested samples, because it does not only express the directions of all cumulated local horizontal strokes but also the orientation of the global dominant text lines.

4.4 Adaptative directional Gabor filtering

4.4.1 Gabor functions and orientations quantification

Multi-channel Gabor filtering is inspired by the psychophysical findings of the cortex that has a set of parallel and quasi-independent mechanisms usually modelled by bandpass filters [34,35]. Here, we use this multi-channel filtering technique to precisely localize directional information of handwritten data. Those filters are mainly used for texture segmentation by tuning the filters to the image dominant spectral information [12]. This filter function is given by the following formula:

$$G(u,v) = A\left\{\exp\left[-\frac{1}{2}\left(\frac{(u-u_0)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right] + \exp\left[-\frac{1}{2}\left(\frac{(u+u_0)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right]\right\}, \quad (3.6)$$

where $\sigma_u = 1/2\pi\sigma_x, \sigma_v = 1/2\pi\sigma_y$ and $A = 2\pi\sigma_x\sigma_y$, where σ_x and σ_y are standard deviations in the x- and y-axis. U_o is the sinusoidal bandwidth in the x-axis (corresponding to the 0° orientation). We have implemented adaptative bank filters with a very precise selection of parameters for frequency, orientation and bandwidth. This selection is highly dependent on the image and an automatic parameterization is a non-trivial process in image analysis because it needs to parameterize the filters in each selected direction θ [7,34,35]. The implementation of a complete Gabor expansion entails an impractical number of filters. In our work, we have proposed an automatic process of bank filters selection.

4.4.2 Multi-channel Gabor filtering for the selection of interesting patterns

We have limited the number of filters by selecting relevant directions in the extremes of the directional rose. Handwriting images have the specificities to contain a typical frequencies distribution: the handwritten patterns are globally contained in the high frequencies, whereas the background is in majority contained in the low frequencies. As for noise frequencies, they can be both on high and low frequencies that will depend on its type. Considering here again that the noise in our documents is mainly concentrated in the low frequencies, we favoured high frequencies of the outlines of patterns and massively filtered low frequencies of the residual background (after the Hermite noise reduction step). The filtering produces a set of directional maps that are then quantified to sort the responses according to their increasing relevance. We are interested in four different parameters in Gabor functions that represent the selection in frequencies and in orientations. The scale factor selection is determined by the amplitude of the standard deviations in Gaussian functions of Gabor expressions. In all tested samples, we apply a generic formula to determine Gabor deviations for all

significant directions in the interval [0°, 180°]:

$$\begin{aligned} \mathrm{Dev_H}(\theta) &= \sqrt{\mathrm{Dev_H}\left(\frac{\pi}{4}\right)^2 + \mathrm{Dev_V}\left(\frac{\pi}{4}\right)^2} \times \cos(\theta), \\ \mathrm{Dev_V}(\theta) &= \sqrt{\mathrm{Dev_H}\left(\frac{\pi}{4}\right)^2 + \mathrm{Dev_V}\left(\frac{\pi}{4}\right)^2} \times \sin(\theta). \end{aligned} \tag{3.7}$$

$\mathrm{Dev_H}(\theta)$ and $\mathrm{Dev_V}(\theta)$, respectively, correspond to horizontal and vertical deviations for a direction θ. $\mathrm{Dev_H}(\pi/4)$ et $\mathrm{Dev_V}(\pi/4)$ correspond to reference horizontal and vertical deviations with $\theta = \pi/4$ and is represented in Fig. 11.

Figure 11a shows the graphical representation of a filters bank in four fixed directions in the frequency domain. The choice of the standard deviation $\sigma_{u,v}$ (also called scale factor) is fundamental and modifies the diameter of the non-filtered regions. It is proportional to the tested image size. In Fig. 12, we have illustrated

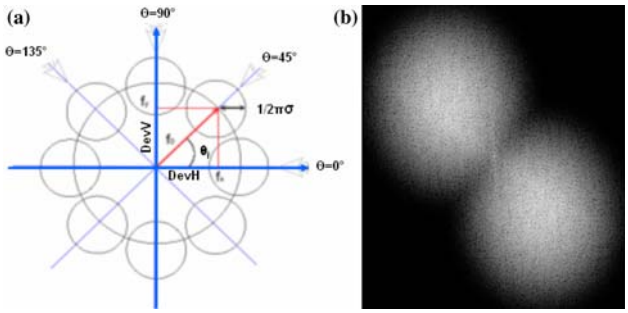

Fig. 11 Filters bank in four standard directions ($\theta = 0°, 45°, 90°$ and 135°) (**a**). Gabor spectrum filter for $\theta = 45°$ with $\sigma_{u,v} = 1/256$ (in a 256×256 image) (**b**)

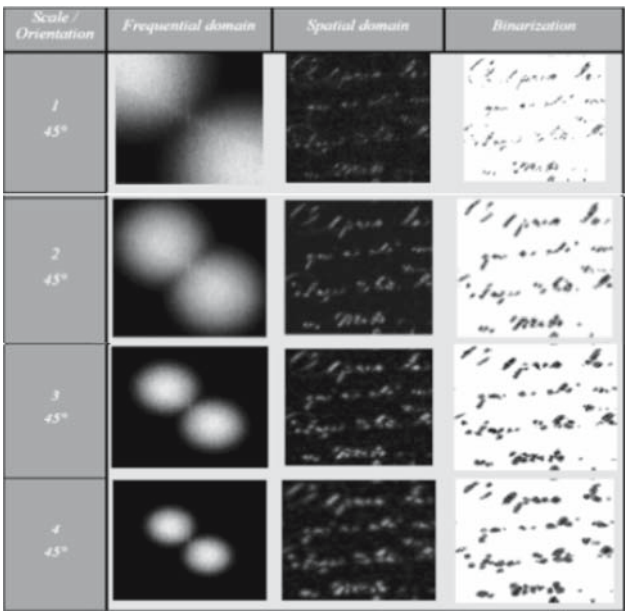

Fig. 12 Examples of Gabor filter responses for $\theta = 45°$ and their bi-level versions for a sample of Montesquieu's draft

Gabor filter responses with four increasing scale factors and different orientations. A satisfying scale has been obtained by an evaluation of the ratio between Gabor response and the initial handwriting thickness.

The more the filtered areas are close to the FFT centre (in the low frequencies), the more the background of the image is filtered. Inversely, when high frequencies are weakly attenuated, outlines of handwriting regions are significantly underlined. In this work, we have been working with a constant standard deviation.

4.4.3 Quantification and directional sketches analysis

In our work, the relevance of a direction is estimated by the measure of the quantity of bi-level Gabor filter response to the direction in consideration. We can notice from Fig. 12 that Gabor response (white regions of the images) are concentrated in the high frequencies regions and are present in the handwriting outlines. The background is mainly black, filled without significant response for each orientation. Now, we are able to evaluate Gabor response by quantifying the thresholded regions. Binarization thresholds are fixed for all handwritten pages of the same book. Gabor responses allow decomposing the initial image into a set of separable directional maps containing oriented patterns (see Fig. 13). The significant orientations are given by the directional rose analysis and determine the basic parameters of Gabor bank filters. In the example of Fig. 13, the represented directions are 2°, 55°, 90° and 145°. The successive AND-logical operations that are applied between the four maps guarantee that the initial selection of significant directions is relevant. The handwriting outline reconstruction and the complete recovering of handwriting patterns allow us to neglect a lot of secondary insignificant directions (we only keep

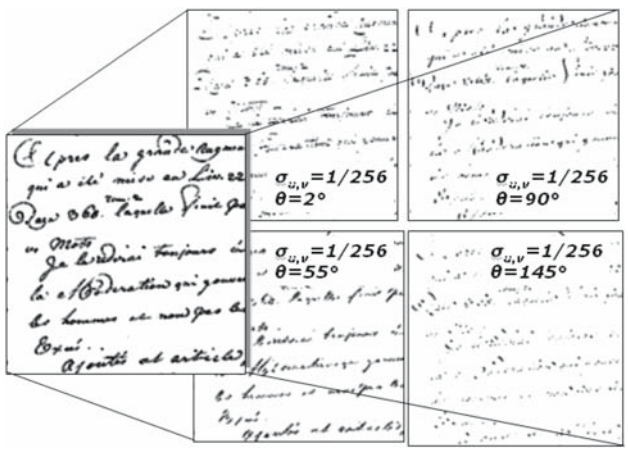

Fig. 13 Handwriting reconstruction from only four directional maps

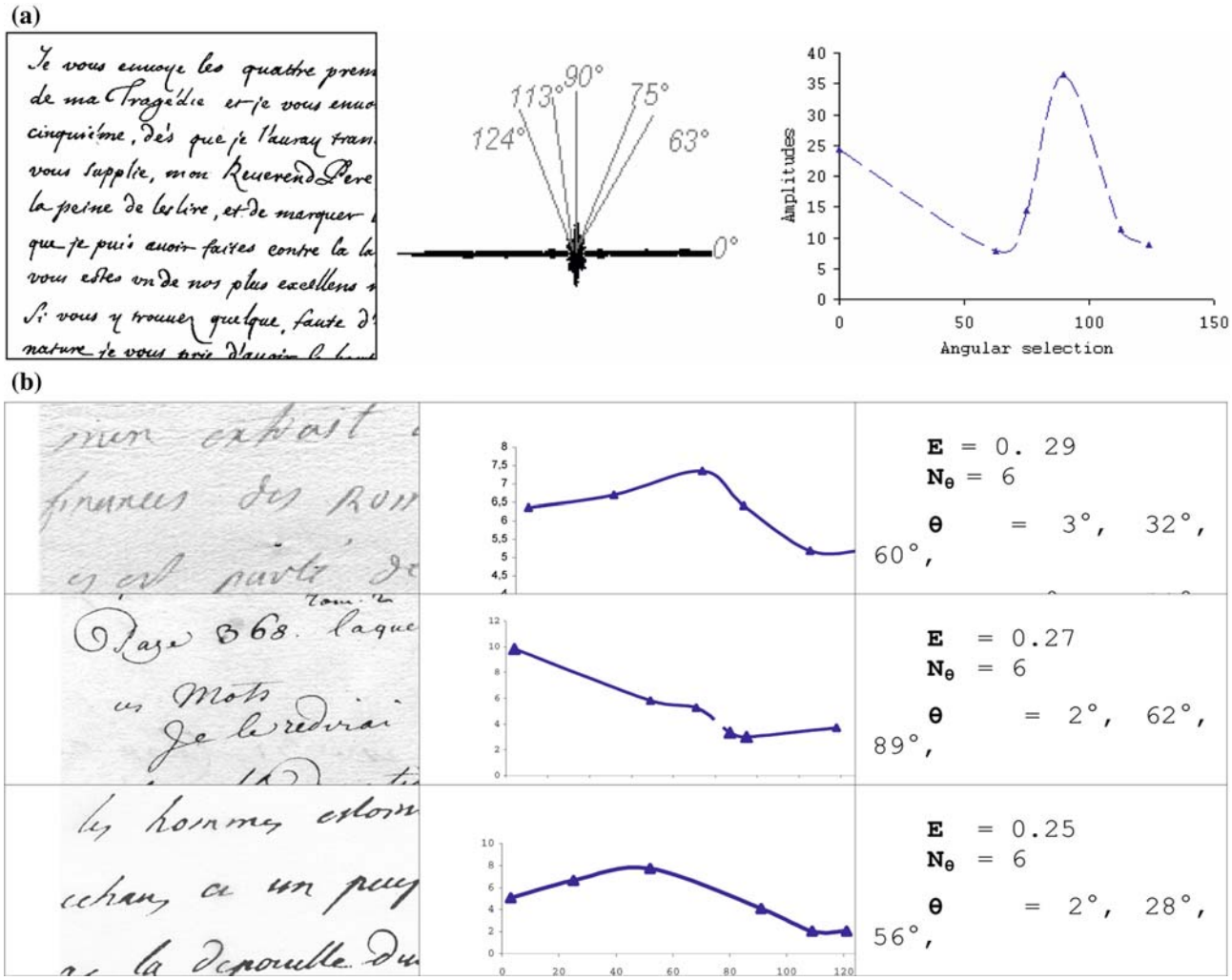

Fig. 14 **a** Sample signature with $E = 0.59$ and six emergent orientations. **b** Examples of numerical full scale signatures of handwritten text block extracted from our corpus (here Montesquieu's secretaries' writing). E entropy, N_θ: number of angular values, θ list of angular values

four directions here). After the binarization step, we evaluate the density of object pixels in the directional maps. Each direction is then weighted and ordered in a list, which is called *the handwriting signature*.

4.5 Individual handwriting signature

The signature is expressed by a list of significant orientations and their corresponding Gabor densities measured over the entire handwriting sample. Figure 14 shows numerical signatures of a handwriting block that have been obtained by quantifying Gabor densities in all significant directions. In the x-axis, we have the angular θ-values and in the y-axis the corresponding Gabor quantification.

The local maxima of the curves for θ-directions signify that the corresponding θ-oriented shapes of the handwriting samples are significantly represented. The local minima of the curves (the valleys) show that other directions are less important in the handwriting. The more the curve is horizontal, the more the handwriting is curved, with a balanced distribution of angular direction in the handwriting. On the other hand, when angular Gabor densities are strongly contrasted with high maxima, the handwriting presents eye-catching shape properties with generally skewed and thin handwritten lines.

5 Application to writers' identification

5.1 Dynamic comparisons of handwritings

The comparison between two signatures uses a warping function that allows possible fusion and fraction operations between them. The warping function consists of non-linear matching.

It is found that in different domains, it is necessary to match sequences with tolerance of small local misalignments. In that context, Dynamic Time Warping has been shown to be an efficient tool for this [13,22,23]. It solves the problem of correspondences between two sequences by searching the optimal warping path, along which the accumulated distance or distortion is minimized. This distance has been widely used in handwriting and document recognition [1], because it allows series to be locally stretched or shrunk before applying the base distance measure.

Considering two signatures, $S_{\theta I}$ and $S_{\theta J}$, the goal of the warping function is to make a correspondence (with a *matching curve C* between two signatures $S_{\theta I}$ and $S_{\theta J}$) between the I-values of $S_{\theta I}$ and the J-values of $S_{\theta J}$.

Within it, it is possible to compare two signatures $S_{\theta I}$ (with I different values) and $S_{\theta J}$ (with J different values) that have non-identical sizes. It is possible to compare two signatures that characterize two image samples having different sizes with, for example, big or small writings. With the entropy-based selection, it is not necessary to normalize text blocks. In practice, the condition of a relevant comparison consists in comparing entropy values of the set of handwriting samples before computing their warping distances, DTW. This measure is considered in the following section as the within-writer and the between-writer distance. In this work, we retain the definition proposed by Fu et al. [10].

Definition Given two sequences, $S_{\theta I} = S_{\theta I}(1), S_{\theta I}(2), \ldots, S_{\theta I}(I)$ and $S_{\theta J} = S_{\theta J}(1), S_{\theta J}(2), \ldots, S_{\theta J}(J)$, the warping distance DTW is defined recursively as follows:

$$\mathrm{DTW}(\emptyset, \emptyset) = 0$$

$$\mathrm{DTW}(S_{\theta I}, S_{\theta J}) = \left\{ \frac{1}{I+J}\left(d(\mathrm{First}(S_{\theta 1}), \mathrm{First}(S_{\theta J})) + \min \left\{ \begin{array}{l} \mathrm{DTW}(S_{\theta I}, \mathrm{rest}(S_{\theta J})) \\ \mathrm{DTW}(\mathrm{rest}(S_{\theta I}), S_{\theta J}) \\ \mathrm{DTW}(\mathrm{rest}(S_{\theta I}), \mathrm{rest}(S_{\theta J})) \end{array} \right\} \right\},$$

where $\emptyset$ is the empty sequence, First $(S_{\theta I}) = S_{\theta I}$ (1), rest $(S_{\theta I}) = S_{\theta I}(2), \ldots, S_{\theta I}$ (I) and d denotes the distance between two entries. Several metrics can be used for d, such as Manhattan Distance [36] and squared Euclidean Distance [13]. Here, we used the Euclidean distance to ensure the symmetry of the result, that is:

$$d(S_{\theta I}, S_{\theta J}) = \sqrt{(S_{\theta I}(i) - S_{\theta J}(j))^2}.$$

The difference $\mathrm{DTW}(S_{\theta I}, S_{\theta J})$ between two signatures $S_{\theta I}$ and $S_{\theta J}$ can also be expressed by the sum of the differences that exist between the two values θ_i and θ_j,. So, the warping path can be compared to the *minimal* deformation that exists between the two vectors. If the warping path between the two signatures $S_{\theta I}$ and $S_{\theta J}$ is completely linear (and it is not the case in Fig. 15), that means that both signatures present similar angular densities, and in that context they can be considered as similar.

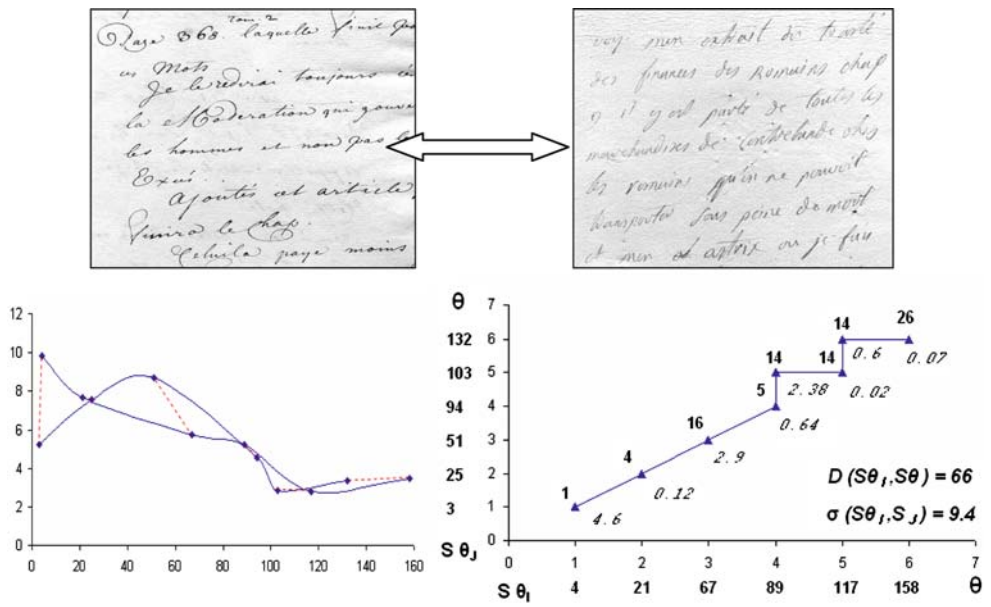

Fig. 15 Comparison of signatures and warping path between $S_{\theta I}$ and $S_{\theta J}$

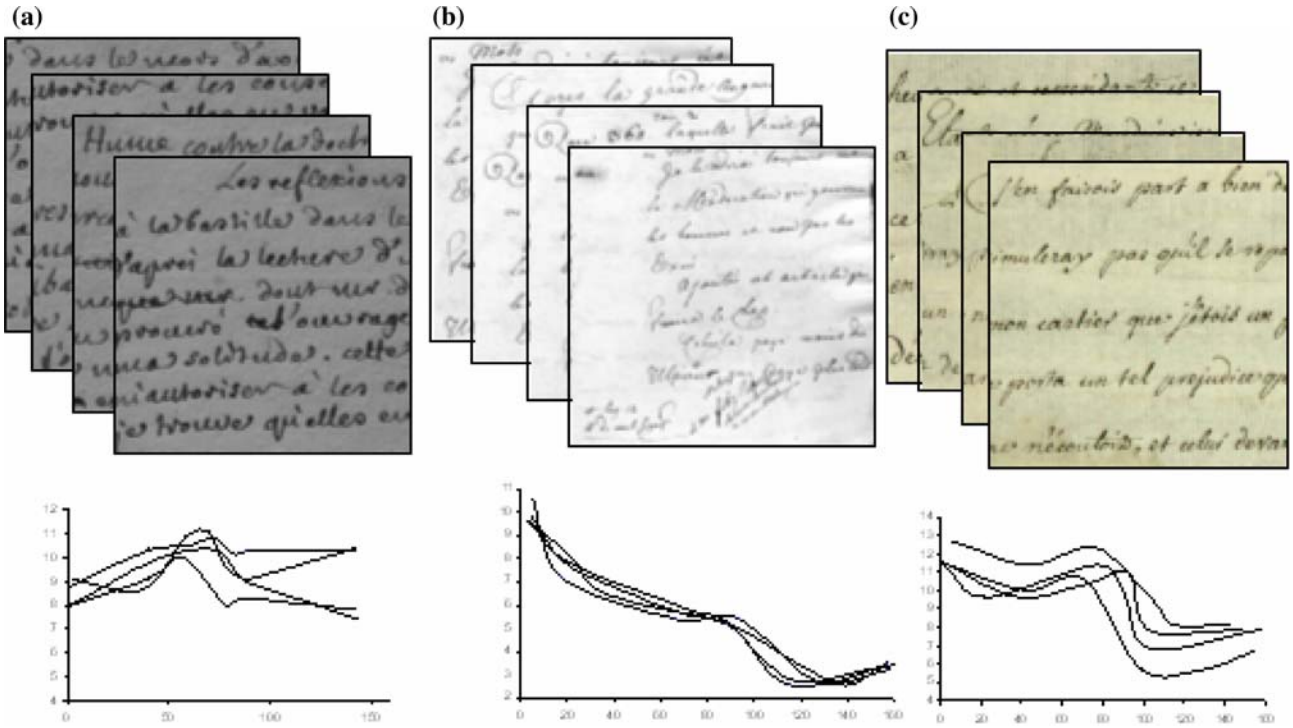

Fig. 16 Within-writer variability. **a** For an Economist of 1789 in the Bastille (Library of the Municipality of Lyon, France). **b** For a Montesquieu's copyist. **c** For Montesquieu himself, *Histoire Véritable* 1750

The complete comparison between two signatures lies on three criteria: the DTW($S_{\theta I}, S_{\theta J}$), the differences of Gabor-based quantification DGQ only used for the quantization of the writing thickness (useful for the within-writer stability analysis) and the variance of the differences $\sigma(S_{\theta I}, S_{\theta J})$that is estimated over the set of angular differences and that quantifies the dispersion of the differences between two signatures.

The tolerance threshold between two handwriting samples that belong to the same writer has been chosen for a maximum value DTW$(S_{\theta I}, S_{\theta J}) = 18$ and a maximum standard deviation $\sigma(S_{\theta I}, S_{\theta J}) = 2$. The distance, DTW, the deviation, σ, and the differences of Gabor quantification, DGQ, are all three necessary to express the resulting similarity between two handwritings.

5.2 Experiments

5.2.1 Within-writer stability

The system is required to deal with as many writers as possible. This study deals with 48 different writers with sometimes very similar styles that have been identified by literary experts. The training samples for each writer represent several lines of text with a certain entropy value. The training set has been created so as to recover the entire set of writers. For the training set, we use ten different samples per writer.

With D and σ, we can notify the similarities that exist between writers. This similarity is used here to distintinguish different writing styles and to characterize the stability of a particular writing (the *within-writer* stability). We show here that the warping distance D, the deviation σ and the *Gabor quantification* are efficient indicators for the within-writer stability evaluation. These measures will also be used to generalize the *between-writer* discrimination.

The warping distance is computed between two handwritten samples having similar entropy values. Figure 16 shows examples of within-writer stability evaluation, all signatures are superposed and warping distances, standard deviations and Gabor quantifications are automatically computed. Ten separate pages were involved in each of the within-writer tests. In Fig. 16b, the samples are characterized by weak variations that reveal a rather curved handwriting. In all the samples, we can notice that the signatures present similar tendencies with possible top or bottom curve translations that can be quantified by Gabor differences and that reveal handwriting thickness. The main orientations are close to a maximal average standard deviation of 0.6 for all tested samples.

The global results are presented in Table 1 according to the 3D vector [$D(S_{\theta I}, S_{\theta J})$; Gabor Diff; $\sigma(S_{\theta I}, S_{\theta J})$] with the following notations: *Mtq* for Montesquieu, S_i for Montesquieu's secretaries, E_i for 18th century econ-

Table 1 Results of ten separate handwritings of the corpus

	Mtq	S1	S2	S3	S4	E1	E2	A1	A2	A3
Mtq	**13-1.5-0.9**									
S1	67-9.6-8.4	**10-2.4-0.3**								
S2	43-12.7-2.3	56-8.9-2.7	**14-1.1-0.6**							
S3	54-10.6-10.2	45-12.6-5.4	39-4.9-2.1	**8-1.7-0.9**						
S4	21-5.8-1.9	78-26.2-5.7	67-5.2-1.9	74-5.8-8.9	**11-3.4-0.2**					
E1	89-22.6-14.5	67-14.2-4.6	49-9.2-1.8	52-12.7-2.3	67-14.2-4.6	**16-1.8-0.4**				
E2	59-14.9-9.8	59-12.5-1.8	77-22.6-4.7	69-5.9-8.4	45-8.4-1.9	49-10.6-5.4	**11-1.8-0.6**			
A1	43-5.8-8.9	45-8.4-1.9	27-12.2-2.8	29-12.8-2.8	27-12.2-2.8	21-4.8-1.9	39-4.9-3.1	**13-1.2-0.8**		
A2	77-32.2-14.7	37-8.9-2.1	18-1.2-1.6	34-7.8-1.9	52-12.7-2.3	29-8.6-4.6	50-8.9-2.7	21-5.8-2.9	**12-2.8-0.9**	
A3	22-8.6-4.6	67-10.2-3.1	46-16.7-3.1	43-4.6-3.2	77-32.2-14.7	53-8.4-3.9	29-12.8-6.8	41-12.7-1.3	89-22.6-14.5	**9-1.2-0.4**

omists, A_i for 18th century independent authors. In this reduced corpus, the within-writer analysis is synthesized in the first bold diagonal: the average warping distance is around 12 (i.e., less than 2° differences between each main orientation), the average Gabor quantification differences are equal to 1.9 and the average standard deviation is equal to 0.6 (i.e., the main angular values present a very small difference near 0.6 on average).

This quantification is stable for 93% of the corpus even on samples that do not have the same sizes, the same grapheme size or the same contrast. The seven residual percentages concern very badly written samples with too many irregularities, non-constancy of the same writer, large variations between uppercase and lowercase characters for the same writer, poor draft quality and heterogeneity of the page layout (see Fig. 17).

The approach presents some limitations that are foreseeable because it concerns ancient degraded pages whose appearance cannot be sufficiently improved by a noise reduction process. Moreover, as we compute a global analysis on a document, without any previous segmentation step, the signature we obtain is representative of the whole document, and each handwritten text area gives a contribution to this signature. Consequently, we cannot analyse documents containing multiple writers. In such a case, the result will be a signature that will not represent correctly any of the writings. We have to suppose that an expert makes a selection

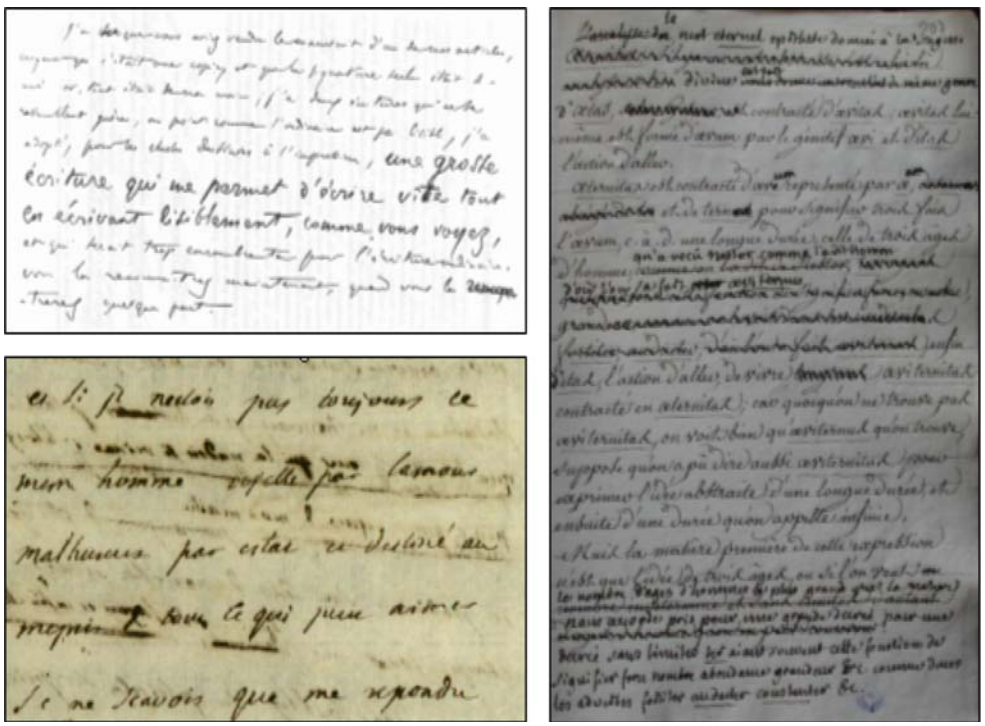

Fig. 17 Typical irregularities in handwriting samples: style inconstancy of a writer (Saint-Saens 1860) and low draft quality

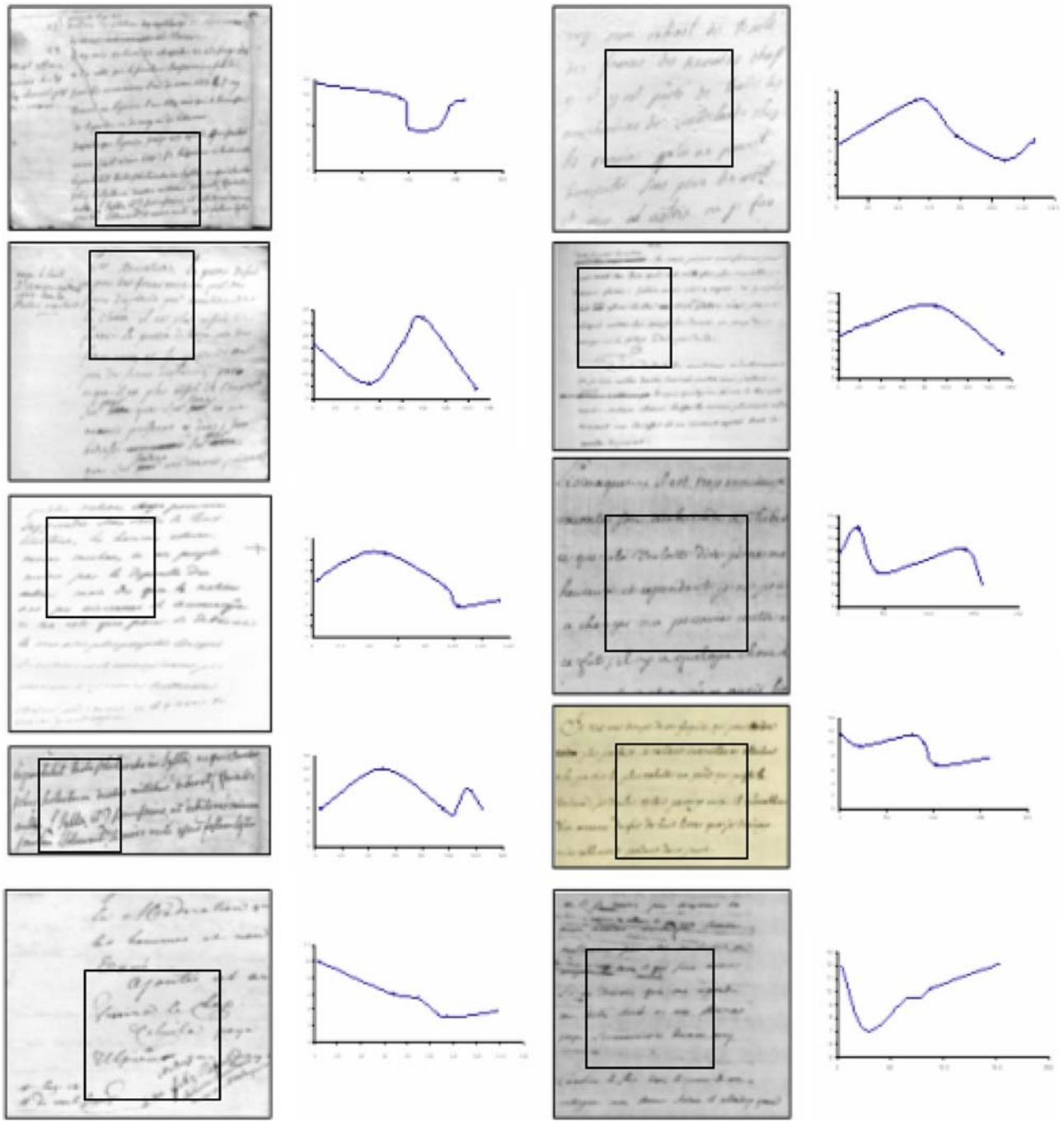

Fig. 18 Individual handwriting signatures for ten selected writers of the corpus. Each signature is computed on an inner selected sample (*black bordered window*)

of areas containing only one kind of writing before the analysis.

5.2.2 Between-writer analysis and classification results

Figure 18 shows a simplified set of relevant visual handwriting classes (or families) of the test corpus. Each class is represented here by a single handwriting reference that is characterized by an individual typical signature. The samples of Fig. 18 have been chosen here among the 48 different writers of the database, because they are specific to the Montesquieu's collection only. As for the within-writer stability analysis, the writer classification step lies on the computation of the similarity measure, based on the 3D vector $[D(S_{\theta I}, S_{\theta J})$, Gabor Diff, $\sigma(S_{\theta I}, S_{\theta J})]$. This measure is estimated between the

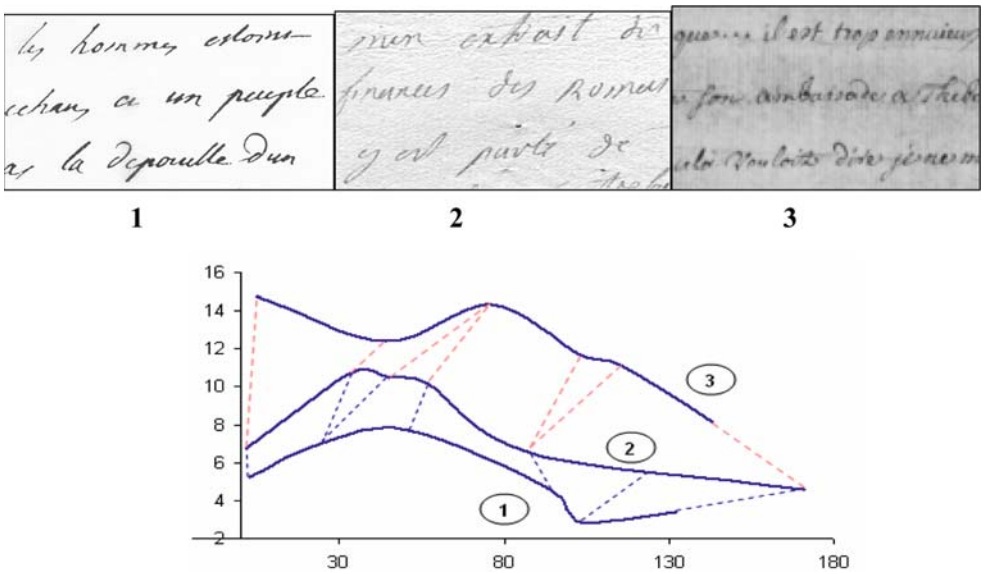

Fig. 19 Illustration of individual comparisons between two sets of samples

handwriting query and the handwriting references. The classification decision lies on a set of individual verifications between two samples, the query sample and each individual reference samples of the database. Figure 19 illustrates the distances between two sets from sample (1–2) and (2–3).

For each query handwritten page, we compute the signature on a reduced region that verifies the initial conditions of entropy and density. The relevance of the analysis is systematically evaluated by a priori knowledge of the writers' styles that historians have taught us. In practice, we observe that two handwritten images can be compared if their size ratio are not inferior to 1/2. Over this limit, the distance measure and the similarity value are no more relevant.

The output of this process is a list of images that are ordered according to their similarity with the query. Figure 20 presents an example of classification: the pages have been extracted from an entire book entitled " *Histoire Véritable*" written by Montesquieu in 1750. By comparing the test sample with all other reference images, we order the images according to the resulting warping distances and the standard deviations. The within-writer average values are used as thresholds to choose the final class. In this example where the writing is very stable, we have produced 98% of correct classification.

With this methodology, we statistically obtained 91% of correct classification with the correct class as first response. The remaining 9% corresponds to queries, which are not homogeneous or which contain too many irregularities (essentially on draft pages containing many crossings out).

By enlarging the problem to a larger corpus, we can decide to generate automatically a new class when the warping distance between the tested handwriting and the reference models exceeds the predetermined threshold.

5.3 Discussion

Orientation is a psycho-visual feature with strong perceptive properties. The results underline the ability of the system to categorize handwritings with a single directional analysis and without any graphemes segmentation. The proposed approach is based on a small assumption on the content of the handwritten block, minimal entropy and a significant occupation rate. One of the advantages of the approach is that two extracts of different sizes (and also of different writing sizes) can be compared. The second original point of this study deals with the background noise reduction. Considering all background frequencies, the signature contains more than 50% of background orientations (with the hypothesis that the background represents more than 50% of the information). In that context, the Hermite noise reduction step is considered here as a necessary preprocess for the undesirable low frequencies reduction.

The remaining difficulties we have encountered deal with some writers' instability and carelessness, especially in draft documents. In those cases, it is difficult to classify two handwritings of the same author. In some situations, we have also noticed that some handwriting lines are not necessarily horizontally aligned. The orientation rose expresses a possible shift between the expected

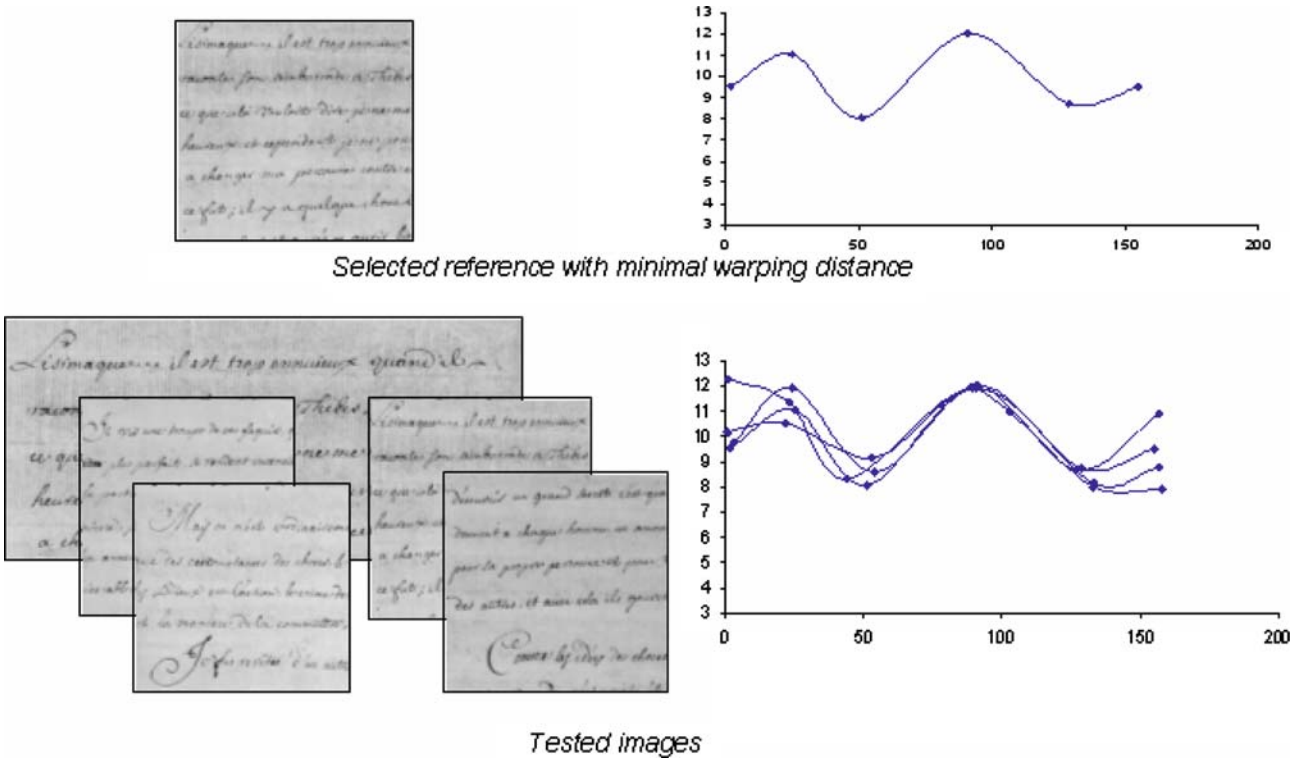

Fig. 20 Examples of handwriting samples of a complete book (*Histoire Véritable* 1750) that must be classified. Best matching with one of the proposed reference models

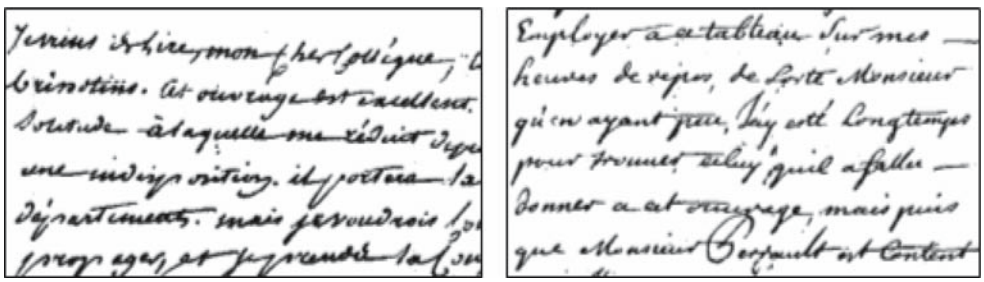

Fig. 21 Example of similar writings that cannot be differentiated with our orientation-based approach

main horizontal direction (the main directional pick) and the real lines' orientations. Consequently, significant angular values must be rotated. In most of the cases, the single orientation is not efficient for a complete robust handwriting categorization (robust to simple transforms like rotations and scale changes; see Fig. 21). In this figure, both samples of different writers cannot be distinguished with our global orientation-based approach.

For these reasons, we are currently trying to implement complementary local grapheme-based features to complete the description (e.g., continuity, compactness and density of the writing) in accordance with the Srihari and Schomaker approaches [4–6]. In their works, particular care has been taken in the detection of some typical handwriting invariants (presence of characteristic loops, of typical words cuttings, of recurrent orthography errors, etc.). But it is important to recall here, that in those systems, the handwriting identification and verification processes are based on no less than 75,000 images written by more than 1,000 writers and do not need a priori knowledge on reference writers' signatures like in our extended Montesquieu's database.

6 Conclusion

This work deals with handwriting categorization in noisy documents and is applied here to writer identification of ancient 18th and 19th century authors' manuscripts. This paper is the first part of a global indexation and classification system for degraded historical handwritten documents. We propose here a biologically inspired approach

for image noise reduction (by background *cleaning*) and handwriting categorization (based on *orientation*). The orientation feature is currently complemented by other spatial primitives of handwritten text (curvature, complexity, linearity and pattern invariance). Two perception-based models have been used for that purpose: the Hermite frequency transforms (for the noise reduction) and the Gabor filter banks (for the multi-scale orientation characterization). Our motivation is directly linked to the difficulty in performing efficient image processing on degraded handwritten historical documents. In this way, we have chosen a segmentation-free approach that also leads to a selective page mapping in textual areas. The results of handwriting classification with only one feature are very promising and show that a unique measure can be a discriminative factor for a relevant visual classification in a reduced and labelled corpus. We are currently working on automatic learning of rejected test images that have not been classified into reference classes.

References

1. Bahlmann, C., Burkhardt, H.: The writer independent online handwriting recognition system flog on hand and cluster generative statistical dynamic time warping. IEEE Trans. PAMI **26**(3), 299–310 (2004)
2. Bensefia, A., Paquet, T., Heutte, L.: Handwritten document analysis for automatic writer recognition. Electron. Lett. Comput. Vision Image Anal. **5**(2), 72–86 (2005)
3. Bres, S.: Contributions à la quantification des critères de transparence et d'anisotropie par une approche globale. Ph.D. Thesis, Lyon (1994)
4. Bulacu, M., Schomaker, L.: Writer style from oriented edge fragments. In: Proceedings of the CAIP Computer Analysis of Images and Patterns, Groningen, The Netherlands, pp. 460–469 (2003)
5. Catalin, I.T., Zhang, B., Srihari, S.N.: Discriminatory power of handwritten words for writer recognition. In: Proceedings of the International Conference on pattern Recognition, IEEE Computer Society, Cambridge, UK, pp. 638–643 (2004)
6. Cha, S.H., Srihari, S.: Multiple feature integration for writer verification. In: Proceedings of the 7th International Workshop on Frontiers in Handwriting Recognition, IWFHR VII, Amsterdam, pp. 333–342 (2000)
7. Chetverikov, D., Liang, J., Komuves, J., Haralick, R.M.: Zone classification using texture features. In: Proceedings of the 13th International Conference on Pattern Recognition, vol. 3, pp. 676–680 (1996)
8. Eglin, V., Volpilhac-Auger, C.: Caractérisation multi-échelle des tracés manuscrits en vue de la catégorisation de scripteurs. In: Proceedings of the CIFED, La Rochelle, France, pp. 106–114 (2004)
9. Franke, K., Koppen, M.: A computer-based system to support forensic studies on handwritten documents. Int. J. Doc. Anal. Recognit. **3**, 218–231 (2001)
10. Fu, A.W.-C., Keogh, E., Yung, L., Lau, H., Ratanamahatana, C.A.: Scaling and time warping in time series querying. In: Proceedings of the 31st VLDB Conferecne (2005)
11. Gallica: digital library of the BNF: http://gallica.bnf.fr
12. Jain, A.K., Bhattacharjee, S.: Text segmentation using Gabor filters for automatic document processing. Mach. Vision Appl. 5(3):169–184 (1992)
13. Keogh, J.: Exact indexing of Dynamic Time Warping. In: VLDB, pp. 406–417 (2002)
14. Kuckuck, W.: Writer recognition by spectra analysis. In: Proceedings of the International Conference in Security through Science Engineering, pp. 1–3 (1980)
15. Lebourgeois, F., Trinh, E., Allier, B., Eglin, V., Emptoz, H.: Document image analysis solutions for digital libraries. In: Proceedings of the DIAL, Palo Alto, pp. 20–32 (2004)
16. Lebourgeois, F., Trinh, E., Emptoz, H.: Compression and accessibility with the images of digitized documents—application to the Debora project, numerical document. Flight **7**(3–4), 103–127 (2003)
17. Leedham, G., Varma, S., Patnkar, A., Govindaraju, V.: Separating text and background in degraded document images- a comparison of global thresholding techniques for multi-stage thresholding. In: Proceedings of the 8th International Workshop on Frontiers in Handwriting Recognition, Niagara-on-the-Lake, Canada, pp. 244–249 (2002)
18. Martens, J.-B.: The Hermite transform—theory. IEEE Trans. Acoust. Speech Signal Processing **38**(9), 1595–1606 (1990)
19. Marti, U.V., Messerli, R., Bunke, H.: Writer identification using text line based features. In: Proceedings of the ICDAR'01, Seattle (WA, USA), pp. 101–105 (2002)
20. Nishida, H., Suzuki, T.: A multiscale approach to restoring scanned color document images with show-through effects. In: Proceedings of the 7th International Conference on Document Analysis and Recognition, pp. 584–88 (2003)
21. Nosary, A., Paquet, T., Heutte, L.: Reconnaissance de textes manuscrits par adaptation au scripteur, CIFED', pp. 365–374 (2002)
22. Ratanamahatana, C.A., Keogh, E.: Everything you know about dynamic time warping is wrong. In: Proceedings of the 3rd Workshop on Mining Temporal and Sequential Data, in Conjunction with the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Seattle (WA, USA) (2004)
23. Rath, T., Manmatha, R.: Word image matching using dynamic time warping. In: Proceedings of the CVPR, Vol. **2**, 521–527 (2003)
24. Rivero-Moreno, C.J., Bres, S.: Conditions of similarity between Hermite and Gabor filters as models of the human visual system. In: Petkov, N., Westenberg, M.A. (eds.) Computer Analysis of Images and Patterns. Lectures Notes Computer Science, (CAIP 2003) Groningen, The Netherlands, Vol. **2756**, 762–769 (2003)
25. Said H.E.S., Peake G.S., Tan T.N., Baker K.D.: Writer identification from non-uniformly skewed handwriting Images. In: Proceedings of the British Machine Vision Conference, pp. 478–489 (1998)
26. Sharma, S.: Show-through cancellation in scans of duplex printed documents. IEEE Trans. Image Process **10**(5), 736–754 (1998)
27. Srihari, S.N., Beal, M.J., Bandi, K., Shah, V.: A statistical model for writer identification. In: Proceedings of the 8th International Conference on Document Analysis and Recognition, pp. 626–630 (2005)
28. Srihari, S.N., Cha, S.-H., Arora, H., Lee, S.: Individuality of handwriting. J. Forensic Sci. **47**(4), 1–17 (2002)
29. Strouthopoulos, C., Papamarkos, N.: Text Identification for document image analysis using a neural network. Image Vision Comput. **16**, 879–896 (1998)

30. Tan, C.L., Cao, R., Shen, P.: Restoration of archival documents using a wavelet technique. Proc. Pattern Anal. Mech. Intell. IEEE Trans. **4**(10), 1399–1404 (2002)
31. Tonazzini, A., Vezzosi, S., Bedini, L.: Analysis and recognition of highly degraded printed characters. Int. J. Doc. Anal. Recognit. **6**, 236–247 (2004)
32. Tonazzini, A., Bedini, L., Salerno, E.: Independant component analysis for document restoration. Int. J. Doc. Anal. Recognit. **7**, 17–21 (2004)
33. Volpilhac-Auger, C., Eglin, V.: La problématique des ouvrage manuscrit ancien: vers une authentification des écritures des secrétaires de Montesquieu, Journée sur la valorisation des documents et numérisation des collections, Ecole Normale Supérieure de Lyon, Lyon, le 7 mars (2002)
34. Weldon, T.P., Higgins, W.E.: Algorithm for designing multiple Gabor filters for segmenting multi-textured images. In: Proceedings of the International Conference on Image Processing, Chicago, IL, pp. 4–7 (1998)
35. Yang, F., Lishman, R.: Land cover change detection using Gabor filter texture. In: Proceedings of the 3rd International Workshop on Texture Analysis and Synthesis, pp. 1024–1029 (2003)
36. Yi, B.-K., Jagadish, H.V., Faloutsos, C.: Effcient retrieval of similar time sequences under time warping. In: Proceedings of the 14th International Conference on Data Engineering (1998)
37. Zois, E.N., Anastassopoulos, V.: Morphological waveform coding for writer identification. Pattern Recognit. **33**(3), 385–398 (2000)