Biological inspired Tools for Patrimonial Handwriting Denoising and Categorization

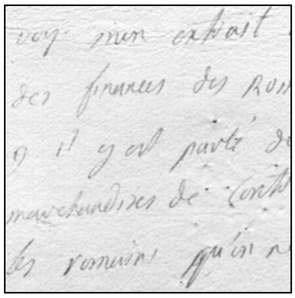
Véronique EGLIN, Stéphane BRES, Carlos RIVERO, Hubert EMPTOZ
LIRIS UMR 5205 INSA Lyon, 20 av. Albert Einstein 69621 Villeurbanne Cedex
e-mail : {veronique.eglin,stephane.bres,hubert.emptoz@liris.cnrs.fr}

Abstract

In this paper, we propose a global segmentation free methodology for patrimonial documents denoising, handwriting characterization and categorization that are based on biological inspired approaches. We widely used here the spectral domain of handwritten images by frequency decompositions: Hermite transforms and Gabor bank filters. Handwritten pages are described by a multiscale signature that is based on orientation features and that is at the basis of a similarity measure. The current results of handwriting categorization and indexing are very promising and show that it is possible to analyze handwritten drawings without any a priori graphemes segmentation.

1. Ancient handwritings and biological inspired tools

Handwritten documents that are considered in our study have been extracted from gigantic patrimonial collections that gathered numerous kinds of writings from the medieval to the contemporary periods of 18th and 19th centuries. For this last period, we can give in example famous French authors like Montesquieu or Flaubert that are at the origin of rare collections for the most preserved in libraries or specialized institutes and recently associated to innovating digitalization projects (ACI MADONNE, CNRS *Patterns and Colors* Project[1]). Among those manuscript collections through different historical periods, we have been interested by those that have been marked by intensive uses and manipulations that can be expressed by a poor documents visual quality, with handwritten texts often erased with multi-writer annotations, corrections, with back ground spots, coins, and different delocalized folds and asperities, [4], [7], see fig.1. Ancient documents of our corpus are characterized by handwritten irregularities that lead to a real difficulty to separate text and non text areas (insufficient pen pressure, drawing luminance/colour, noisy background…).



1.1. 1.2.

Figure 1. Typical background noises. 1.1 Ink of reverse side. 1.2. Granularity of the support.

The second difficulty is linked to the pages layout variability and the presence of irregular background noises that increase the complexity of lines and words segmentation. So as to avoid those difficulties, we have chosen a segmentation free approach that leads to a selective pavement of the page image guided by the textual informative areas.

2. Perceptual approach of background denoising

2. 1 Hermite decompositions

Many digital images of documents and more generally ancient manuscripts are degraded by the presence of strong artefacts in the background. This can either affect the readability of the text and in our case it compromises a relevant handwriting characterization. Background artefacts can derive from many different kinds of degradations such as scan optical blur and noise, spots, underwriting or overwriting, time wearing, intensive use, bad preservation conditions, see figure 1, [7]. We focus here on different damages that can be due to seeping of ink from the reverse side, or due to the ink degradation (attenuation of the ink marks that compromises a correct text reading) or palimpsests (an earlier text has been suppressed and the vellum or parchment reused for another). All those degradations cause difficulties to localize and segment interest handwritten areas. For all those situations, we consider the documents as a multilayer signal that is mostly composed with a textured background and a high frequency handwriting signal. The principle of the background denoising is based on a polynomial transform, the Hermite transform [5], which is a good model of the set of chan-

[1] ACI MADONNE - ACI Masse de donnees of the pluridisciplinary thematic network, RTP-Doc, 2003-2006.
Patterns and Colors, Project, interdisciplinary program «Information Society»,09/2003-02/2005.

nels describing the Human Visual System, [6]. In general, a polynomial transform decomposes locally a signal into a set of orthogonal polynomials with respect to the window W used for localizing the signal. Hermite transform uses a Gaussian window W and is exploited here to decompose the initial signal into polynomial separable maps. The preservation of high frequencies domains (corresponding to textual areas) and the suppression (cleaning) of low frequencies regions (in background regions) are two complementary processes that restore visual pages appearance.

2.2 Presentation of Hermite Filters

We present the definitions of Hermite filters which agree with the Gaussian derivative model of the HVS [15] [16]. We will focus on their cartesian representations 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, they have similarities to Gabor filters [11], which are more used, essentially for texture, in image processing and feature extraction. Hermite and Gabor filters are equivalent models of receptive field profiles (RFPs) of the HVS [13] [14].

2.2.1 Cartesian Hermite Filters

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, which is defined as:

$$v(x,y) = 1/(\sigma\sqrt{\pi})e^{-(x^2+y^2)/(2\sigma^2)} \quad . \qquad (1)$$

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 derivative order D (or polynomial degree) and the 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 1-D filter is:

$$d_n(x) = \left((-1)^n \Big/ (\sqrt{2^n \cdot n!}\sqrt{\pi}\sigma)\right) H_n(x/\sigma)e^{-x^2/\sigma^2} \quad . \qquad (2)$$

where Hermite polynomials $H_n(x)$, which are orthogonal with respect to the weighting function $exp(-x^2)$, are defined by Rodrigues' formula [7] as:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \quad . \qquad (3)$$

In the frequency domain, these filters are Gaussian-like band-pass filters with extreme value for $(\omega\sigma)^2 = 2n$ [13] [14], and hence filters of increasing order analyze successively higher frequencies in the signal.

2.2.2 Krawtchouk Filters

Krawtchouk filters are the discrete equivalent of Hermite filters. They are equal to Krawtchouk polynomials multiplied by a binomial window $v^2(x) = C_N^x / 2^N$, which is the discrete counterpart of a Gaussian window. These polynomials are orthonormal with respect to this window and they are defined as [7] :

$$K_n(x) = \frac{1}{\sqrt{C_N^n}} \sum_{\tau=0}^{n} (-1)^{n-\tau} C_{N-x}^{n-\tau} C_x^{\tau} \quad . \qquad (4)$$

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

It can be shown that the Krawtchouk filters of length N approximates 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:

$$K_{n+1}(x) = \frac{1}{\sqrt{(N-n)(n+1)}}\left[(2x-N)K_n(x) - \sqrt{n(N-n+1)}K_{n-1}(x)\right] \qquad (5)$$

for $n \ge 1$ and with initial conditions $K_0(x) = 1$, $K_1(x) = \frac{2}{\sqrt{N}}\left(x - \frac{N}{2}\right)$.

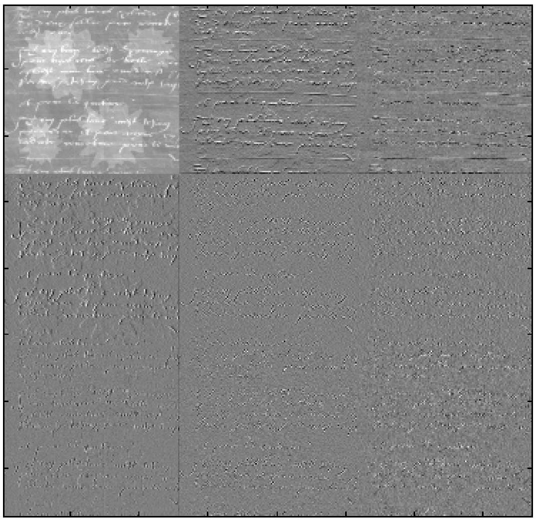


Figure 2.1 Hermite decomposition of a document at a given scale and up to degree.

Figure 2.1 presents the Hermite decomposition of an original noisy document at a given scale, depending on the size of window W and up to polynoms of degree 2 as well as the reconstructed image I_R, using this decomposition and some adaptative thresholdings of the different quadrants of the Hermite transform. Consequently, handwritten regions can be processed globally without background artifacts that compromise the system performance. The result of image denoising after Hermite transform process is presented in figure 2.2.

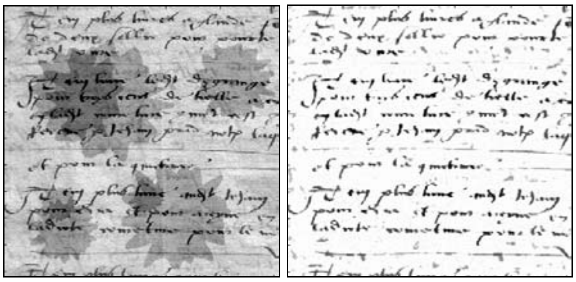


Figure 2.2 Original document. Denoised image using the Hermite decomposition.

3. Multiscale approach for orientation characterization

3.1 Emergent handwriting directions selection

We have chosen to characterize handwritten text by considering a visual salient property: the orientation in the way of human visual expertise like it has been introduced by Kuckuck in [4]. The originality of our approach is to consider a multiscale approach from a macroscopic (linked to the page dimension and the global text lines orientations) to a microscopic (linked to the grapheme or the letter dimension) analysis. Precisely here, handwritings can be described by different pop-out orientations at different scales, see figure 3. By zooming progressively on the handwriting, we can underline the evolution of the different orientations that are basic handwriting characteristics. The principle of orientation selection lies on a frequencies decomposition based on the detection of most representative directions obtained in the *directional rose* by the application of the autocorrelation function on the entire document spectrum. The directional rose computation lies on the use of the *autocorrelation function*, which correlates the image with itself, highlights periodicities and orientations of texture. This function has been widely used in a context of texture characterization,[12]. Its definition for a bi-dimensional signal is the following:

$$C_{xx}(k,l) = \sum_{k'=-\infty}^{+\infty} \sum_{l'=-\infty}^{+\infty} x(k',l').x(k'+k,l'+l)$$

The autocorrelation function $C_{II}(i,j)$, applied to an image I, combines this, 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 privileged directions of the image. The data that are relative to a same direction will be located in a same line. With this principle, it is possible to detect orientations of the texture blocks.

3.2 Multi channel Gabor filtering for interest regions selection

From this step, we apply on handwriting images selected Gabor bank filters that enhance directional interest regions. Gabor approaches are commonly used in perception based methodologies, [2], [8].

We used here the multi-channel filtering technique to localize precisely directional information of handwritten data. This approach is inspired by the multichannel filtering theory for processing visual information in the early stages of the human visual system, [17, 18]. Gabor filters are well known for the capacity too model the receptive filed profiles of simple cells in the primate visual cortex well. They are generally used for texture segmentation by tuning the filters to the image's dominant spectral information. This filter function is given by the following formula:

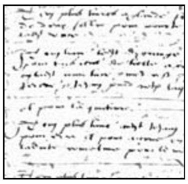
$$G(u,v) = A\left(\exp\left\{-\frac{1}{2}\left[\frac{(u-u_0)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\}+\exp\left\{-\frac{1}{2}\left[\frac{(u+u_0)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\}\right)$$

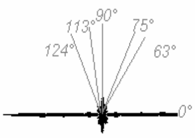
where $\sigma_u=1/2\pi\sigma_x$, $\sigma_v=1/2\pi\sigma_y$ and $A = 2\pi\sigma_x\sigma_y$ where σ_x and σ_y are standard deviations in the x and y-axis. U_o is the sinusoidal bandwidth in the x-axis (corresponding to the 0° orientation). In our work, directional Gabor responses are binarized so as to enhance preserved high frequencies of handwritings. The resulting directional pattern sketches are finally quantified in surface so as to weight the orientation relevance (see figure 3).

Figure 3: Scale variant decomposition in two scales of a handwritten extract and their Gabor filters responses in significant directions.

3.3 Multiscale handwriting signature

The image signature lies on the hierarchy of Gabor responses quantification. This approach is a relevant technique to evaluate the variability of writer handwritings (like the traceability of a same handwriting all along a book or a work). And in the same way, it allows making a precise discrimination between writers. This signature is considered here as an indexation and classification tool. The proposed methodology for handwriting classification simulates the human visual behaviour in its ability to observe and discriminate handwritten patterns with different perceptive points of views. A multiscale signature is computed for each text block images and is derived in a similarity measure to compare different extracts, see figure 4.





Gabor density	θ°
36.43	90°
24.33	0°
14.35	75°
11.34	113°
8.96	124°
7.89	63°

4.1 **4.2** **4.3**

Size : 512* 512		
Θ (in degrees)	Gabor density	Rose amplitude
51	7.68	1.00
25	7.04	0.24
3	5.23	0.16
94	4.57	0.16
132	3.37	0.16
103	2.83	0.15
Size : 256*256		
Θ (in degrees)	Gabor density	Rose amplitude
51	8.30	1.00
2	4.71	0.18
77	5.35	0.29
64	7.43	0.30
127	4.88	0.14
108	3.26	0.15
Size : 128*128		
Θ (in degrees)	Gabor density	Rose amplitude
2	6.81	0.25
88	6.51	0.24
57	10.13	0.76
45	10.47	0.65
34	10.77	0.92
171	4.58	0.36

4.4

Figure 4: Signature of a handwritten fragment at full scale (4.3) obtained from the directions distribution (4.2). Results in three successive scales (4.4).

3.3.1 Recursive handwriting blocks analysis

A hierarchical analysis is applied on the original handwritten text block that is recursively cut into four smaller blocks until a non significant text block size. A block is recursively cut starting from the entire image to a minimal-sized image. The first scale (in the entire image) is called *scale0*. Recursively, each scale is renamed by, *scale1*, *scale2*…*scaleS* until the blocks sizes contain quantitatively significant handwritten patterns that are estimated by a *minimal* entropy value. The minimal size is then estimated by the threshold $TH_E(S)$ corresponding to the *minimal* entropy value of the sub-scaled block (in the corresponding *ScaleS*). That means that the image must contain a significant text lines number to be exploited by the method. The initial image size (*scale0*) can not be less than the minimal entropy threshold TH_E (0) (in our corpus it generally corresponds to a 512*512 image with a entropy value equal to 0.22) for a multiscale relevant analysis. Precisely here, the entropy is directly correlated to the visual impression of "complexity" we have during the observation. A text made of small letters seams more "complex" than a text with big letters. Our study quantifies this complexity with a measure of entropy. For that purpose, we compute the number of transitions from the background to the text that can be found on random oriented lines (so as to privileged no particular direction). That 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 because it is representative of how much complex the analyzed text block can be (or the grapheme in a reduced analysis scale).

For each sub-scale, we have an increasing number of signatures that is proportional to the number of cuts (one for scale 0, 2^2 for scale 1, 2^4 for scale 2…). Those signatures must be merged into a unique numerical vector of significant directions. The fusion principle consists in merging all redundant orientations into a unique value (with the corresponding merged Gabor density): all similar values whose differences are less than 8° are considered as being equal. All dual values are suppressed. This vector is completed with all singular values if they are representing in more than 50% of small blocks (so as to neglect background residual noise, artefacts, and punctual overwriting). Finally, the resulting signature is composed in *S* scales (*S* cutting levels) with ordered directions values weighted by their corresponding averaged Gabor densities, see figure 4.4

3.3.2 Dynamic comparison for handwriting categorization and indexing

The comparison between two signatures lies on a warping function that allows possible fusion and fraction operations between two signatures. The warping function consists in a non linear matching: 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 (a vector with only 6 values can be compared with a vector with 8 values). It is possible to compare two signatures in fixed scales (simple vector containing no more than 8 values) or in different scales (large multiscale vectors). In both cases, the number of scales in each signature must be the same. Within this approach it is also possible to compare two signatures that represent two different scales. According to that principle, it is possible to compare two handwritings of different text sizes (big or little writings) that generally need a normalization step to assure a well text size and text block correspondence. The goal of the warping function is to make a correspondence between two signatures $S_{\theta I}$ and $S_{\theta J}$ between the *I* values of $S_{\theta I}$ and the *J* values of $S_{\theta J}$. The difference $D(S_{\theta I}, S_{\theta J})$ between two signatures is expressed by the sum of minimal deformation existing between the two vectors.

The warping distance is interpreted in this study as an expression of *similarity* that leads to a possible handwriting images ordering. The figure 5 proposes a direct application of the warping measure by the presentation of a list of handwriting extracts ordered by increasing distances for the only *full* scale signature. The query is formulated with a reference handwriting image. For legibility reasons, we have resized resulting images having all a similar entropy value, see [1] for the entropy definition and use. We have tested the whole system on our personal database composed of documents coming from different authors but mainly patrimonial handwritings documents. Most of the time, we have full pages of the same author and for evaluation purpose.

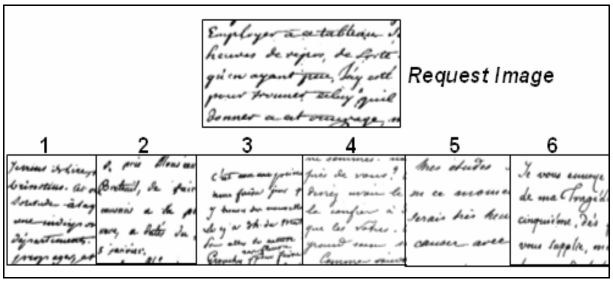


Figure 5. Visual query results at full scale.

The comparison between handwriting images can be realized through different scales. We can compute the warping distance between two handwriting images in a single scale level with similar entropies. The choice of scales that is initially made for the warping distance computation depends on the analysis purpose. If we want to build a global handwriting images classification in distinct visual families lying on a global point of view, we select the part of the signature that is representative to the global first scales. Practically, the selection of scales depends on the classification purpose. It can be a complete handwriting image classification for which we use the different scales and we compute the warping distance on the entire signatures. This approach is used for a precise intra writer analysis: it shows the possible variability that exists in handwritings of a same writer. It can be a precise discrimination at grapheme level where we can select high scales parts of the signatures (i.e local measures on smallest blocks), especially for the comparison between two identical signatures. Table 1 summarizes results of multiscale classification and gives percentage of correct retrieval with a comparison at different scaled parts of the signature.

Number of scales	1	2	3	4	5
Good retrieval	54%	65%	76%	82%	64%

Table 1. Percentages of image retrieval through different scales

The initial images sizes have been all chosen with sizes generally greater than 1024*1024 so as to justify the recursive cutting in 5 successive scales. In practice we consider the first 4 scales that is the best compromise that we can obtain between time computing and retrieval quality. Our database contains 1438 images coming from 189 different authors, in different languages and alphabets. We can notice in this table that taking into account all different scales, we can give high precision on handwriting local graphemes and retry a maximal number of text blocks for the same writer.

4. Conclusion

This work presents the first part of a global indexing, retrieval and characterization project that is applied on degraded ancient handwriting documents in noisy environment. We propose here a biological inspired approach for images denoising (especially background cleaning) and handwriting characterization (especially here with the orientation). The orientation feature is currently completed by other spatial primitives of text drawings (curvature, complexity, linearity, and patterns invariance) that are currently developed. Two perceptive based models have been used for that purpose: the Hermite frequencial decomposition (for the denoising and the handwriting enhancement) and the Gabor bank filtering (for the multiscale orientation characterization). Our motivation is directly linked to the difficulty to perform efficient image segmentation on degraded handwriting historical documents. In that way, we have chosen a segmentation free perceptive approach that also leads to a selective pavement of the page in textual areas. The results of classification with the multiscale orientation are very promising but also show that the single dimension of orientation must be completed with other discriminative pattern features to define a selected handwriting with an accurate precision.

5. References

[1] Eglin V., Volpilhac-Auger, C., *Handwriting multiscale characterization for writers classification* (in french), in Proc. CIFED, 2004, pp. 106-114.

[2] Jain, A.K., Bhattacharjee, S., *Text segmentation using Gabor filters for automatic document processing*, Machine Vision & Appli., 1992, vol.5, no.3, pp.169-184.

[3] Kuckuck, W., *Writer recognition by spectra analysis,* Int. Conf. In Security th. Sci. Eng., 1980, pp.1-3.

[4] Lebourgeois, F., Trinh, E., Allier, B., Eglin, V., Emptoz, H. *Doc.ument Image Analysis solutions for Digital libraries, In DIAL*, pp. 20-32, 2004.

[5] Martens J.-B., *The Hermite transform – Theory,* IEEE Trans. Acoust., Speech, Signal Processing, vol. 38, no. 9, pp. 1595-1606, 1990.

[6] Rivero C. Bres, S. *Conditions of similarity between Hermite and Gabor filters as models of the human visual system*, In Petkov, N.& Westenberg, M.A. (eds.): Comp. Anal. of Images & Patterns, pp.762-769, 2003.

[7] Tonazzini A., Bedini L., Salerno E., *Independent component analysis for document restoration*, Inter. Jour. on Doc. Anal. and Recog, vol.7, N°1, 2004, pp. 17-27.

[8] Weldon, T.P., Higgins, W.E. *Algorithm for designing multiple Gabor filters for segmenting multi-textured images*, Inter. Conf. on Image Proc., Chicago, Illinois, 4-7, 1998.