Frequencies Decomposition and Partial Similarities Retrieval for Ancient Handwriting Documents Compression

Abir EL ABED, Véronique EGLIN, Frank LEBOURGEOIS, Hubert EMPTOZ
LIRIS UMR 5205 INSA Lyon, 20 av. Albert Einstein 69621 Villeurbanne Cedex
E-mail : abir_elabed@yahoo.fr, {veronique.eglin,frank.lebourgeois@liris.cnrs.fr}

Abstract

This paper presents a new segmentation free approach of partial similarities retrieval in ancient handwritten documents. The method has been developed to improve usual handwritings compression approaches that are not adapted to patrimonial images specificities. We present here the similarities characterization that lies on oriented handwriting shapes decomposition. The frequencies page decomposition realizes a pavement of handwritten regions stored in directional maps *where partial similarities are estimated. This decomposition is obtained by a frequencies analysis implying Gabor bank filters with an adequate parameter setting based on the most significant directions of the text. For each map, we compute a similarity graph that reveals redundant shapes and determines a resulting redundancy rate. The resulting graph is the first part of the compression system currently under development.*

1. Introduction

1.1Ancient handwritten documents specificities

The digitalization of the patrimonial scientific and cultural documents facilitates the access to a larger audience and offers new services like the online consultation of rare documents, the fast and economic duplication of works, a quick information retrieval and the possibility of knowledge sharing with other readers, [1]. We can also notice that digitalized documents cause difficulties in terms of storage and transmission on a flow-limited network. Those needs and remarks are at the basis of an innovating digitalization project called ACI MADONNE[1] and our contribution is presented here. From there, we prove the needs to develop compression and information retrieval methods that are well adapted to images content. They generally require new data format definitions for remote contents inquiries. In that context, we can easily show that the images compression became a necessity. Only lossy compression with an acceptable perceptible loss of information may reduce the weights of images. Only users can really evaluate the fidelity of broadcasted images compared to the originals, the acceptable loss of information and the speed of access through the network. The existing compression formats like JPEG, DJVU and DEBORA are unfortunately not effective on handwritten documents images, [2]. The limitations of these methods lie on the great complexity of the features present in the handwritten texts and on the difficulty to localize precisely handwritten parts of the page. Indeed, we can notice that a robust handwritten documents segmentation algorithm and shape recognition need at least shapes redundancy, noticeable relevant similarity and some handwritten regularities that can not be easily found in our corpus, see figure 1. With all these difficulties, it is necessary to develop specific methods to locate precisely handwritten text areas and redundant shapes without any segmentation, [3]. Existing approaches of regular handwritten text and drawing location (connected components analysis, directional and morphological filtering, simple directional projections, [4]) are inefficient in noisy and variable documents.

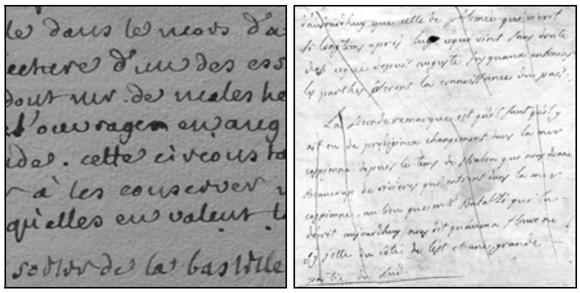

Figure 1: Pages with irregular background and a significant non linearity of drawing, [3].

1.2 Some traditional compression formats

The existing traditional compression formats are not effective on handwritten document images. Indeed these formats do not take into account the specificity of the handwritten layouts, the frequency of shapes location and the shapes characteristics. For example,

[1] ACI MADONNE - ACI Masse de données of the pluridisciplinary thematic network, RTP-Doc, 2003-2006.

we notice a delocalization of the pixels gray levels for an image compressed with JPEG. With this compression method, we can also notice a reduction of the gray levels due to low-pass frequential filtering and unexpected appearance of gray levels areas due to the specific square blocks cutting, see figure 2. In that context, we can also mention other methods like DjVu that applies a compression with and without loss of information depending on the use. Handwritten documents present a real difficulty in separating the different information layers (graphic, textual…).

For these reasons, it is difficult to apply a compression method based on pattern redundancy. The lack of similarities between patterns disturbs the application of compression methods based on pattern redundancy.

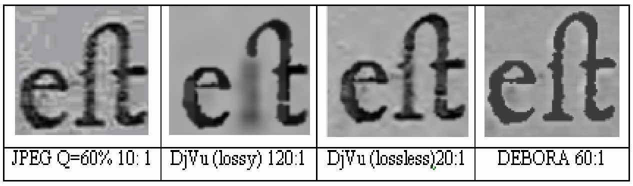

Figure 2: Visual differences and compression rates between most significant methods, [2].

These limitations allows us to conclude that the methods based on layers separation (like DjVu and DEBORA) are inefficient in ancient handwritten documents, see figure 2. It is necessary to consider a more robust approach that can support great patterns variability in even noisy environment.

1.3 Segmentation free proposition

An effective handwriting compression approach needs at first a robust and performing method for text location with all its complexity and variability and excluding background noises. Our proposition concerns the first stage of the global compression process, and we propose here to retrieve partial shapes similarities of degraded handwritings without any preliminary global segmentation process, see figure 3.

We have chosen to locate all handwritten interest zones with the analysis of patterns directions. Interest areas are considered through their significant orientations with redundant intern sketches. These redundant shapes are not necessarily words or letters: they can simply be composed in inner loops of letters and can be localized anywhere in the page with some elementary geometrical transforms (scale changing, rotation or translation). Our approach is based on a frequencies analysis implying Gabor bank filters. The global scheme of the approach is illustrated in figure 3. The principle consists in breaking up the original image into directional maps, using Gabor filters with an adequate parameter setting (see *Step2*) based on the most significant directions of the image. These directions are initially evaluated from the image spectrum by the exploitation of the autocorrelation function (Perseval Theorem, see *Step1)*.

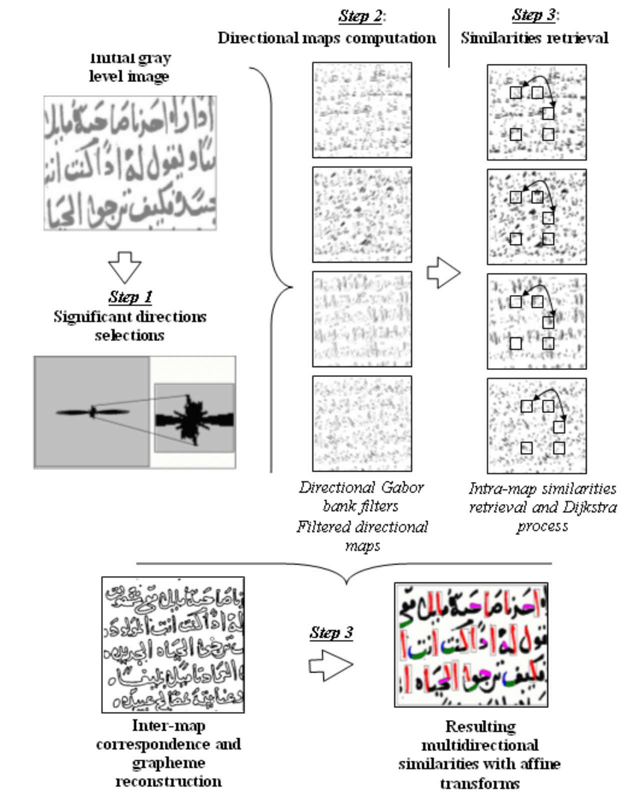

Figure 3: Global scheme for partial similarity retrieval.

A classification of oriented shapes on each directional map is then proposed to define similarities between them respecting translation, rotation, and scale changing invariance (see *Step3*). Those similarities quantify handwritten pattern redundancies. The exploitation of the method in a context of handwriting compression is presented at the end of the paper. It is the direct application of this work.

2. Segmentation-free text retrieval

2.1 Emergent directions selection

In a first step, we have chosen to retrieve the most significant directions of the handwriting without taking into account secondary texture orientations that are naturally present in the image background and that are considered as noise. The principle of orientation selection lies on a frequencies decomposition based on the detection of main directions obtained in the *directional rose* by the application of the autocorrelation function on the entire document spectrum, see figure 4. The *autocorrelation function* correlates the image with itself, highlights periodicities and orientations of texture. This function has been widely used in a context of texture characterization, [5]. The autocorrelation result is used here for the construction of a corresponding directional rose.

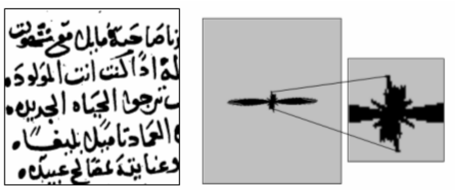

Figure 4: Directional rose of an Arabic handwritten extract

This rose gives with a great precision the main orientations of the block. In [5], the author proposes an approach of directional rose computation. It is based on the mean value that is computed from the autocorrelation result.

2.2 Multi channel Gabor filtering for handwriting decomposition

The selected directions are then quantified by selected Gabor filters responses. It is important here to take into account the only relevant directions of the image so as to neglect noisy part of the image and revealed in the directional rose. We used here the multi-channel filtering technique to localize precisely directional information of handwritten data. This approach is inspired by the multi-channel filtering theory for processing visual information in the early stages of the human visual system, [6]. Those filters are able to suppress low frequencies patterns by underlying handwriting text lines with a precise selection of parameters for frequency, orientation and bandwidth. The selection is highly dependent on the image. But this automatic parameterization is generally a non trivial process in image analysis because it needs to parameterize the filters in each selected direction θ, [7]. The accurate implementation of a complete Gabor expansion entails a generally impractical number of filters. In our work, we propose an automatic process of bank filters selection.

In our study, we have limited the number of filters by selecting relevant directions contained in the extremes of the directional rose. Handwriting images have the specificities to contain a typical frequencies distribution: the handwritten drawings are globally contained in the highest spectrum frequencies whereas the background is in majority contained in the low frequencies. This frequencial paving map is a guide for the automatic filtered frequencies selection. In that context, we keep high frequencies of the patterns outlines and we massively filter low frequencies of the background. 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. The bank filter produces a set of directional multiscale maps that can be quantified to order the responses according to an increasing rank of relevance. In this work the scale has been fixed to produce readable results (a well compromise between a too blurred response and a not significant filtering), see figure 5.

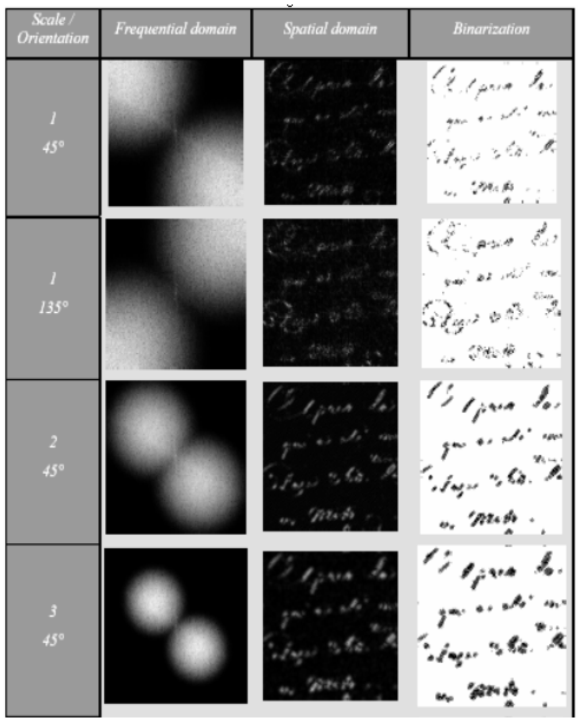

Figure 5:Bi-level Gabor bank filters responses.

This figure illustrates Gabor filtering responses with four increasing scale factors and different orientations. The handwriting text has been selected in Montesquieu corpus. More the filtered areas are close to the FFT center (in the low frequencies domain), more the background of the image is filtered. Inversely, when high frequencies are weakly attenuated, outlines of handwriting regions are highly underlined. As it is illustrated in figure 5, the scale factor modified the diameter of the non filtered region.

Gabor responses are then bi-leveled with a fixed threshold that is estimated for a same work (for an entire book or limited to a single page). A similar approach using four fixed Gabor orientations has been developed in [8] for handwritten signature recognition.

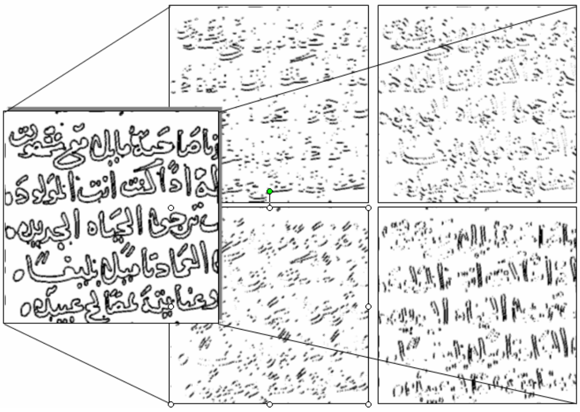

Figure 6: Directional maps separation and handwriting reconstruction.

The combination of different directional maps that have been computed with only significant directions allows reconstructing with a great precision, the entire high frequency handwriting. Figure 6 give an example of handwriting reconstruction with only four directional maps in the selected directions (0°, 42°, 89°and 129°). By considering only the most significant handwriting orientations, we have concluded that it is not necessary to retrieve secondary orientations (that can also be included in the handwritten texts).

3. Partial similarities retrieval

The handwriting is decomposed into partial sketches that are then processed with some morphological operations to eliminate noisy connected components (background noise or local handwritten shapes irregularities) that are not relevant for the similarities retrieval. From this step, each directional sketch can be localized and matched to the others to determine partial similarities and produce redundancy rates.

3.1 Directional sketches pre-processing

A very low thickness of directional sketches is an obstacle for matching operations. To solve this problem we apply to directional sketches a particular morphological treatment adapted to the types of writings. We apply successively an adaptive dilation to initial Gabor filters responses followed by adaptative erosion. The dilation consists in highlighting most significant sketches by filling the existing "holes" in the related shapes and enlarging some local thin shapes. The erosion consists in eliminating the smallest shapes having an area A lower than a selected threshold like A < (*Minimal Area* + 10% of the variation amplitude). The threshold estimation depends on different criteria and depends on handwriting initial size, see figure 7.

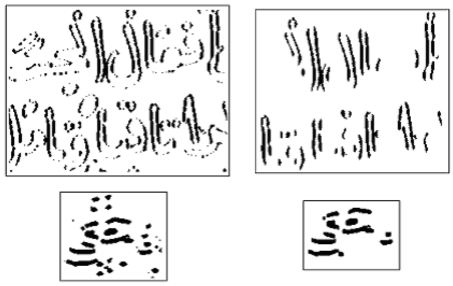

Figure 7: Morphological erosion

The localization and numbering of processed sketches is finally based on the Growing Area Algorithm of recursive filling. A simple indexing makes it possible to compare the shapes one to each other.

3.2 Intra and inter - maps similarities and resulting redundancy rate estimation

The similarity rate is the best adapted indicator to compare and quantify differences between sketches in a same map and between different directional maps. The similarity rate is obtained with a flexible comparison between all size normalized sketches icons pixels P_i and JPL. The similarity is equal to a matching value that is normalized by the maximal sketches area.

$$T_{i,j} = \frac{\sum_{k,l} gl(k,l)}{Max(Area_i, Area_j)}$$

$$with \quad \begin{cases} gl(k,l) = 1 & if\ P_i(k,l) = P_j(k,l) = 0 \\ gl(k,l) = 0 & elsewhere \end{cases}$$

So as to increase the similarity rate, we allow some affine transforms (translation and rotation) that can be applied to sketches before the global similarity rate estimation. The rotation can be applied in a same map (with an angular variation ξ that can varies from –10° to + 10°) but also between two different maps (with an enlarged interval included in [0°,180°]). This flexible estimation for the similarity rate computation guaranties better results. Figure 9 presents an example of similarity retrieval for sketches revealed in the direction 89° with intra map rotations (left part) and an illustration of similarities for three other directions (0°, 42° and 129°) in the right part of the figure. So as to speed up the similarity retrieval process, we apply the similarity operations only on the *reference* sketches that represent the most significant and redundant shapes.

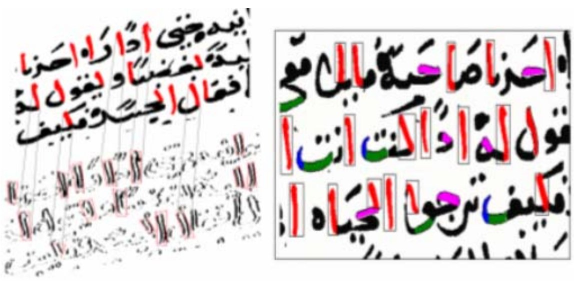

Figure 8 : Some similarities in Arabic handwriting marking for four significant directions (0°, 42°, 89° and 129°)

The similarity rate $T_{i,j}$ is estimated from a matrix representation that is simulated by a genetic tree whose nodes are reference sketches and sons are reconstructed from their reference. A map can be represented by several reference sketches. The determination of references is based on the Dijkstra algorithm that is frequently used in graph theory. In that context, the $T_{i,j}$ value are considered as cost necessary to transform the node i to the node j. The complete original development is presented in [9].

Finally, we can give an estimation of the frequency redundancy in the analyzed handwriting by defining a redundancy rate R_r with the ratio measured between the total reference sketches numbers called N_D and the total number of sketches that are stored in the tree called N_T : $R_r = 100 * (1 - N_D / N_T)$. On the examples of our corpus we have obtained intra and inter map redundancy rates that vary from 67% to 82% for Arabic handwritings and 58% to 72% for Latin handwritings. The amplitude of values is linked to the similarity thresholds that are initially chosen.

The similarity marking leads to a relevant handwriting pavement that is centered in only interest areas. The global compression scheme can not be presented here because it is still under development. It is based on a set of logical references between the distribution of binary sketches similarities and original image shapes. More precisely, it is based on the construction of a similarity graph that reproduces the initial binary sketches graph with an effective link between original Gray level image and bi-level maps. A compensation map must then be computed to correct the reconstructed (decompressed) image by adding the differences (residuals) with the original image. Image of residuals contains pixels that are randomly located into the pattern and around the contours of substituted sketches. The encoding of the residuals is still under investigation.

4. Discussion and conclusion

This work is a first contribution for segmentation free approach of handwriting characterization and compression. We have presented here the first step of partial similarities retrieval that decomposes the original complete signal into directional maps. Those maps contain all primary sketches of initial patterns that are associated with a references graph of the entire analyzed page. As for now, the method is very effective on handwriting documents where graphemes are efficiently separable, visible and contrasted. A first improvement that is currently under development concerns the possibilities to compare different sketches with scale changing (sketches of different sizes with same shapes) and to merge different sketches so as to increase their size (from grapheme to portion of word). This work is currently integrated to a complete compression system that is dedicated to gray level ancient handwriting documents without information loss and that will be integrated in the MADONNE project in the coming days.

5. References

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

[2]European Project *DEBORA*, free access book: http://rfv6.insa-lyon.fr/debora/client.htm, june 2000.

[3]*Handwriting multiscale characterization for writers classification* (in french), in Proc. CIFED'04, 21-25 juin 2004, pp. 106-114.

[4]Lebourgeois, F., Emptoz, H., *Documents Analysis in Gray Level and Typography Extraction Using Character Pattern Redundancies*, 5th ICDAR, 1999, pp. 177 – 180.

[5]Bres, S. *Contributions à la quantification des critères de transparence et d'anisotropie par une approche globale.* PhD Thesis, France, 1994, 231 p.

[6]Weldon, T.P., Eldon, W., Higgins, E. *Algorithm for designing multiple Gabor filters for segmenting multi-textured images*, IEEE ICIP, 1998, pp. 4-7.

[7]Daugman, J.G. *Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters*, Journal of the Optical Society of America A, 1985, vol. 2, pp. 1160-1169.

[8]Fasquel, J.B., Bruynooghe, M., *A hybrid opto-electronic method for fats off-line handwritten signature verification,* In Proc of IJDAR, vol.7, n°1, 2004, pp.56-68.

[9] El Abed A. *Partial similarities retrieval for patrimonial handwritten documents compression*, ACI MADONNE Project, Internal Rapport, September 2004, 46 p.