Character Recognition in Historical Documents

Louloudis Georgios*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
louloud@mm.di.uoa.gr

Abstract. "Character recognition" refers to the procedure of 'reading' text using a computer, taking as input a document image as well as to the conversion of the document image to electronic text. This dissertation focuses on the segmentation of handwritten document images to the basic semantic units that comprise them, namely text lines and words.

Concerning the problem of text line segmentation, a new methodology was developed whose novelties are: (i) an efficient block-based Hough transform in which voting occurs on the basis of equally spaced blocks after splitting of the connected components' bounding box; (ii) a partitioning of the connected component domain into three spatial sub-domains, for which a different processing strategy of the corresponding connected components can be employed; and (iii) the efficient separation of vertically connected parts of text lines. The proposed text line segmentation methodology has been evaluated against other state-of-the-art text line segmentation methodologies and has proven to achieve better results.

Concerning the word segmentation stage, two different methodologies were developed. Concerning the first methodology, the decision on whether a gap is between two words or inside a single word, a threshold was proposed which is calculated making use of several characteristics of the document image. On the second approach, we make use of a well-known methodology in the field of unsupervised clustering, the Gaussian mixture modeling in order to classify the gaps into each class. Experimental results prove the efficiency of the proposed methodologies.

Keywords: Handwritten document image segmentation, Text line segmentation, Word segmentation, Hough transform.

1. Introduction

One of the early tasks in a handwriting recognition system is the segmentation of a handwritten document image into text lines and words. The overall performance of a handwritten character recognition system strongly relies on the results of the text line and the word segmentation process. If the quality of the results produced by the stages of text line and word segmentation is poor, this will affect the accuracy of the text recognition procedure. Thus, the algorithms employed for these two stages are critical for the overall recognition procedure.

Segmentation of a document image into its basic entities, namely, text lines and words, is considered as a non trivial problem in the field of handwritten document recognition. The difficulties that arise in handwritten documents make the segmentation procedure a challenging task. For the text line

* Dissertation Advisor: Constantin Halatsis, Professor

segmentation procedure, major difficulties include the difference in the skew angle between lines on the page or even along the same text line, overlapping words and adjacent text lines touching each other. Furthermore, the appearance of accents in many languages (e.g. French, Greek) further complicates the text line segmentation procedure. Regarding word segmentation, the challenges include the appearance of skew and slant in the text line, the existence of punctuation marks along the text line and the non-uniform spacing of words, which is a common residual in handwritten documents.

A wide variety of segmentation methods for handwritten documents has been reported in the literature. We categorize these methods depending on whether they refer to text line segmentation, word segmentation or both text line and word segmentation. To this end, a brief description of the related work on the text line segmentation and word segmentation problem is presented in Section 2.

2. Related Work

2.1 Line segmentation

There are mainly three basic categories that line segmentation methods lie in: methods making use of the projection profiles, methods that are based on the Hough transform and, finally, smearing methods. Also, several methods exist that cannot be clearly classified in a specific category, since they employ particular techniques.

Methods that make use of the projection profiles include [1, 2]. In [1], the initial image is partitioned into vertical strips. At each vertical strip, the histogram of horizontal runs is calculated. This technique assumes that text appearing in a single strip is almost parallel to each other. Srihari et al. [2] partitions the initial image into vertical strips called chunks. The projection profile of every chunk is calculated. The first candidate lines are extracted among the first chunks. These lines traverse around any obstructing handwritten connected component by associating it to the text line above or below. This decision is made by either (i) modeling the text lines as bivariate Gaussian densities and evaluating the probability of the component for each Gaussian or (ii) the probability obtained from a distance metric.

Methods that make use of the Hough transform include [3-5]. The Hough transform is a powerful tool used in many areas of document analysis that is able to locate skewed lines of text. By starting from some points of the initial image, the method extracts the lines that fit best to these points. The points considered in the voting procedure of the Hough transform are usually either the gravity centers [3], [4] or minima points [5] of the connected components.

Smearing methods mainly include the fuzzy RLSA [6] and the adaptive RLSA [7]. The fuzzy RLSA measure is calculated for every pixel on the initial image and describes “how far one can see when standing at a pixel along horizontal direction”. By applying this measure, a new grayscale image is created which is binarized and the lines of text are extracted from the new image. The adaptive RLSA [7] is an extension of the classical RLSA, in the sense that additional smoothing constraints are set in regard to the geometrical properties of neighboring connected components. The replacement of background pixels with foreground pixels is performed when these constraints are satisfied.

There are also metholodologies that cannot be included in a certain category because they do not share a common guideline. For a more detailed description of text line segmentation methodologies the interested reader should read [8].

Although the above mentioned techniques have been proved efficient for certain problems, there are more challenges found in a text line detection process. For example, none of the above techniques deals with the problem of accents. Although accents don't appear in English documents, it is a common constituent in documents of many languages, e.g. French and Greek. Furthermore, most of these techniques do not strive towards solving the problem of vertically connected characters that results to text line merging. In some of the above techniques, an assumption is made that all text lines have no skew angle or they have the same skew angle. Finally, in the Hough transform-based approach of [4], only one point from every connected component votes in the Hough domain. This may cause a serious problem in cursive multi-accented documents, where one connected component can be a whole word. In this case, a whole word and a small accent have the same contribution in the Hough domain and this may lead to erroneous results.

2.3.2 Word segmentation

Algorithms dealing with word segmentation in the literature are based primarily on analysis of geometric relationship of adjacent components. Components are either connected components or overlapped components. An overlapped component is defined as a set of connected components whose projection profiles overlap in the vertical direction. Related work for the problem of word segmentation differs in two aspects. The first aspect is the way the distance of adjacent components is calculated, while the second aspect concerns the approach used to classify the previously calculated distances as either between-word gaps or within-word gaps. Most of the methodologies described in the literature have a preprocessing stage which includes noise removal, skew and slant correction.

Many distance metrics are defined in the literature. Seni et al. [9] presented eight different distance metrics. These include the bounding box distance, the minimum and average run-length distance, the Euclidean distance and different combinations of them which depend on several heuristics. A thorough evaluation of the proposed metrics was described.

A different distance metric was defined by Mahadevan [10] which was called convex hull-based metric. The author after comparing this metric with some of the metrics of [9] concludes that the convex hull-based metric performs better than the other ones. Kim et al. [11], investigated the problem of word segmentation in handwritten Korean text lines. To this end, they used three well-known metrics in their experiments: the bounding box distance, the run-length/Euclidean distance and the convex hull-based distance. For the classification of the distances, the authors considered three clustering techniques: the average linkage method, the modified Max method and the sequential clustering. Their experimental results showed that the best performance was obtained by the sequential clustering technique using all three gap metrics. Varga and Bunke [12], tried to extend classical word extraction techniques by incorporating a tree structure. Since classical word segmentation techniques depend solely on a single threshold value, they tried to improve the existent theory by letting the decision about a gap to be taken not only in terms of a threshold, but also in terms of its context i.e. considering the relatives sizes of the surrounding gaps. Experiments conducted with different gap metrics as well as threshold types showed that their methodology yielded improvements over conventional word extraction methods.

In all the aforementioned methodologies, the gap classification threshold used derives: (i) from the processing of the calculated distances, (ii) from the processing of the whole text line image or (iii) after the application of a clustering technique over the estimated distances. There also exist methodologies in the literature that make use of classifiers for the final decision of whether a gap is a between-word gap or a within-word gap [13-15]. An early published work making use of classifiers for the word segmentation problem is the work of Kim and Govindaraju [13]. A similar work was presented in [14] by Huang and Srihari. This approach claimed two differences from previous methods: (i) the gap metric was computed by combining three different distance measures, which avoided the weakness of each of the individual one and thus provided a more reliable distance measure and (ii) besides the local features, such as the current gap, a new set of global features were also extracted to help the classifier make a better decision. Finally, a different approach was presented from Luthy et al. [15]. The problem of segmenting a text line into words was considered as a text line recognition task, adapted to the characteristics of segmentation. That is, at a certain position of a text line, it had to be decided whether the considered position belonged to a letter of a word, or to a space between two words. For this purpose, three different recognizers based on Hidden Markov Models were designed, and results from writer-dependent as well as writer-independent experiments were reported.

3. Text line segmentation methodology

The proposed methodology for text line detection in handwritten document images [16-20] deals with the following challenges: (i) each text line that appears in the document may have an arbitrary skew angle and converse skew angle along the text line; (ii) text lines may have different skew directions; (iii) accents may be cited either above or below the text line and (iv) parts of neighboring text lines may be connected.

To meet the aforementioned challenges, we propose a methodology which consists of three main steps. The first step includes binarization and image enhancement, connected component extraction, average character height estimation and partitioning of the connected component domain into three distinct spatial sub-domains. In the second step, a block-based Hough transform is used for the detection of potential text lines, while a third step is used to correct possible splitting, to detect possible text lines which the previous step did not reveal and, finally, to separate vertically connected parts and assign them to text lines. A detailed description of these stages is given in the following subsections 3.1 – 3. 3.

3.1 Preprocessing

The preprocessing step consists of four stages. First, an adaptive binarization and image enhancement technique is applied. Then, the connected components of the binary image are extracted and the bounding box coordinates for each connected component are calculated. The average character height AH for the whole document image is calculated. We assume that the average character height equals to the average character width AW.

The connected components domain includes components of a different profile with respect to width and height since it is frequent to have components describing one character, multiple characters, a whole word, accents and characters from adjacent touching text lines. The aforementioned connected components variability has motivated us to divide the connected component domain into different sub-domains, in order to deal with these categories separately. More specifically, in the proposed approach,

we divide the connected components domain into three distinct spatial sub-domains denoted as “Subset 1”, “Subset 2” and “Subset3” (Figure 1).

“Subset 1” contains all components which correspond to the majority of the characters with size which satisfies the following constraints:

$$\left(0.5*AH \le H < 3*AH\right) \text{ AND } \left(0.5*AW \le W\right) \tag{1}$$

where H, W denote the component’s height and width, respectively, and AH, AW denote the average character height and the average character width, respectively. The motivation for ‘Subset 1’ definition is to exclude accents and components that are large in height and belong to more than one text line.

“Subset 2” contains all large connected components. Large components are either capital letters or characters from adjacent text lines touching. The size of these components is described by the following equation:

$$H \ge 3*AH \tag{2}$$

The motivation for ‘Subset 2’ definition is to grasp all connected components that exist due to touching text lines. We assume that the corresponding height will exceed three times the average character height.

Finally, “Subset 3” should contain characters as accents, punctuation marks and small characters. The equation describing this set is:

$$\left(\left(H < 3*AH\right)\text{AND}\left(0.5*AW > W\right)\right)\text{OR}\left(\left(H < 0.5*AH\right)\text{AND}\left(0.5*AW < W\right)\right) \tag{3}$$

The motivation for ‘Subset 3’ definition is that accents usually have width less than half the average character width or height less than half the average character height.

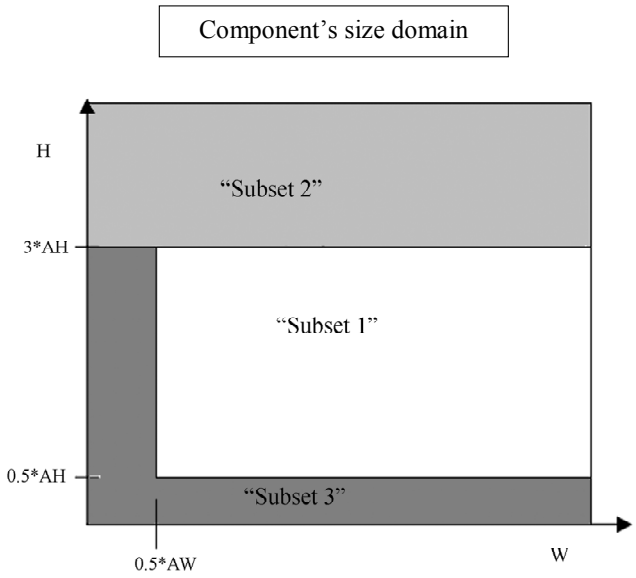

Figure 1: The connected component domain partitioned to 3 sub-domains denoted as “Subset 1”, “Subset 2” and “Subset 3”, respectively.

3.2 Hough Trasform Mapping

In this step, the Hough transform takes into consideration only connected components that belong to sub-domain "Subset 1". Selection of this sub-domain (Figure 1) is done for the following reasons: (i) It is guaranteed that components which span across more than one text line will not vote in the Hough domain; (ii) it rejects components, such as accents, which have a small size. This avoids false text line detection by connecting all the accents above the core text line.

In our approach, instead of having only one representative point for every connected component, a partitioning is applied for each connected component lying in "Subset 1", in order to have more representative points voting in the Hough domain. In particular, every connected component lying in this subset is partitioned to equally-sized blocks. The number of the blocks is defined by the following equation:

$$N_b = \left\lceil \frac{W_c}{AW} \right\rceil \tag{4}$$

where W_c denotes the width of the connected component and AW the average character width of all connected components in the image. For a detailed description of this step see [16, 18].

3.3 Postprocessing

The post-processing step consists of two stages. At the first stage (i) a merging technique over the result of the Hough transform is applied to correct possible false alarms and (ii) connected components of "Subset 1" that were not clustered to any text line are checked to see whether they create a new text line that the Hough transform did not reveal.

The second post-processing stage deals with components lying in "Subset 2". This subset includes components whose height exceeds three times the average height AH. All components of this subset mainly belong to n detected text lines (n>1). Our novel methodology for splitting these components [17] consists of the following:

(A) Calculate yi, which are the average y values of the intersection of detected line i and the connected component's bounding box (i =1..n).

(B) Exclude from the procedure the last line n if the condition described by eq. (5) is not satisfied. In equation 5, (xs, ys), (xe, ye) are the coordinates of the bounding box of the component and I is the image of the component (its value is 1 for foreground and 0 for background pixels). Eq. (5) verifies that the component area near line n is due to a vertical character merging and not due to a long character descender from text line n-1. The denominator of Eq. (5) corresponds to the sum of all black pixels of the connected component whereas the numerator corresponds to the sum of black pixels of the connected component on the lower area below line defined by $y = y_n - \frac{(y_n - y_{n-1})}{10}$.

$$\frac{\sum_{x=x_s}^{x_e}\sum_{y=y_n-\frac{(y_n-y_{n-1})}{10}}^{y_e} I(x,y)}{\sum_{x=x_s}^{x_e}\sum_{y=y_{n-1}}^{y_e} I(x,y)} > 0.08 \quad (5)$$

(C) For every line i, i =1..n-1, we define zones Zi taking into account the following constraint:

$$y_i + \frac{y_{i+1} - y_i}{2} < y < y_{i+1} \quad (6)$$

Then, we compute the skeleton of the connected component, detect all junction points and remove them from the skeleton, if they lie inside zone Zi. If no junction point exists in the segmentation zone Zi we remove all skeleton points in the center of the zone.

(D) For every zone zi, the skeleton parts that intersect with line i are flagged with id '1', while all other parts are flagged with id '2'. Finally, in each zone zi, separation of the initial connected component into different segments is accomplished by assigning to a pixel the id of the closest skeleton pixel.

3.4 Performance Evaluation Methodology

The text line performance evaluation is based on counting the number of matches between the areas detected by the algorithm and the areas in the manually created ground truth. The metrics used in order to compare different text line segmentation methodologies are the detection rate (*DR*), recognition accuracy (*RA*) and the F-Measure (*FM*) which is a combination of the previous two metrics. The set used to check the effectiveness of the proposed text line segmentation methodology was the test set of the ICDAR2007 Handwriting Segmentation Contest. The detailed comparative results for text line segmentation in terms of detection rate (*DR*), recognition accuracy (*RA*), F-Measure (*FM*) and the number of matches are shown in Table 1. In this table we also include all methodologies that participated in the handwriting segmentation contest of ICDAR2007. It is worth noting that our methodology outperforms all the other approaches (marginally in the case of ILSP_LWSeg, more clearly in all other cases) achieving detection rate 97.4% and recognition accuracy 97.4%.

Table 1: Comparative experimental results for line segmentation over the test set of ICDAR2007 handwriting segmentation contest.

Line Segmentation		*N*	*M*	*DR*	*RA*	*FM*	*o2o*	*go2m*	*gm2o*	*do2m*	*dm2o*
Projection [1]		1773	1610	58.3	66.5	62.1	925	222	216	102	482
Fuzzy RLSA [6]		1773	1813	77.6	75.3	76.4	1288	76	277	126	186
Hough [4]		1773	1984	88.5	78.4	83.1	1532	14	136	66	29
ICDAR2007 Handwriting Competition	ILSP_LWSeg	1773	1773	97.3	97	97.1	1713	5	34	17	10
	PARC	1773	1756	92.2	93	92.6	1604	40	76	34	85
	UOA-HT	1773	1770	95.5	95.4	95.4	1674	14	54	27	28
	BESUS	1773	1904	86.6	79.7	83	1494	9	151	72	21
	DUTH-ARLSA	1773	1894	73.9	70.2	72	1214	149	227	107	354
	RLSA	1773	1877	44.3	45.4	44.8	632	264	346	122	757
UOA-HT_EXT		**1773**	**1770**	**97.4**	**97.4**	**97.4**	**1717**	**6**	**34**	**17**	**13**

4. Word Segmentation

4.1 Introduction

Word segmentation refers to the process of detecting the word boundaries starting from a text line image. Although the reader may think that the solution to this problem is trivial, it is still considered an open problem among researchers of handwriting recognition and document analysis areas due to several challenges that need to be addressed. These challenges include: i) the appearance of skew along a single text line, ii) the existence of slant, iii) the existence of punctuation marks and iv) the non-uniform spacing of words.

In most of the methodologies presented in the literature, the word segmentation procedure is divided into two steps. The first step deals with the computation of the distances of adjacent components in the text line image and the second step concerns the classification of the previously computed distances as either inter-word gaps or inter-character gaps. For the first step, we propose the average of two different metrics: the Euclidean distance metric and the convex hull-based metric [17]. For the classification of the computed distances two different strategies are proposed: (i) a thresholding methodology which makes use of several geometrical characteristics of the document image [20] and

(ii) a well-known methodology from the area of unsupervised clustering techniques, namely the Gaussian Mixtures [17].

4.2 Distance Computation

In order to calculate the distance of adjacent components in the text line image, a pre-processing procedure is applied. The pre-processing procedure concerns the correction of the skew angle as well as the dominant slant angle of the text line image. The computation of the gap metric is considered not on the connected components (CCs) but on the overlapped components (OCs), where an OC is defined as a set of CCs whose projection profiles overlap in the vertical direction.

We define as distance of two adjacent overlapped components (OCs) the average value of the Euclidean distance and the convex hull - based distance. The Euclidean distance between two adjacent overlapped components is defined as the minimum among the Euclidean distances of all pairs of points of the two adjacent overlapped components.

We calculate the convex hull-based metric as follows: Given a pair of adjacent overlapped components C_i and C_{i+1}, let H_i and H_{i+1} be their convex hulls. Let L be the line joining the centers of gravity (or centroid) of H_i and H_{i+1}. Let P_i and P_{i+1} be the points of intersection of L with the hulls H_i and H_{i+1}, respectively. The gap between the two hulls is defined as the Euclidean distance between the points P_i and P_{i+1}.

4.3 Gap Classification

4.3.1 Methodology based on threshold

For the gap classification, we define a global threshold in the image. To compute this threshold we calculate the black-to-white transitions in every scanline of the text line image. We focus on the scanline with the maximum number of black-to-white transitions. In this particular scanline we calculate and store all the lengths of the white runs and sort them in a descending order. Finally, we use the median of the sorted list in the line threshold equation which is defined as:

$$LT = 1.8 * M_v \tag{7}$$

where Mv is the median value of the sorted list. The weighting factor 1.8 has been determined after experimental work.

In order to define the global threshold we calculate the average of the temporary threshold along all text lines of the document image.

4.3.2 Methodology based on Gaussian mixtures

A mixture model based clustering is based on the idea that each cluster is mathematically presented by a parametric distribution. We have a two clusters problem (inter-word and intra-word gaps) so every cluster is modeled with a Gaussian distribution. The algorithm that is used to calculate the parameters for the Gaussians is the Expectation Maximization (EM) algorithm. We use this methodology since the Gaussian Mixtures is a well-known unsupervised clustering technique with many advantages which include: (i) the mixture model covers the data well, (ii) an estimation of the density for each cluster can be obtained and (iii) a “soft” classification is available. For a detailed description on the Gaussian Mixtures, the interested reader is referred to. For the calculation of the number of parameters and the number of Gaussians we used software CLUSTER, which is an unsupervised algorithm for modeling Gaussian mixtures.

4.4 Experimental Results

Both word segmentation methodologies were also tested on the test set of the ICDAR2007 handwriting segmentation contest. Table 2 contains experimental results for word segmentation algorithms in terms of detection rate (*DR*), recognition accuracy (*RA*), F-Measure (*FM*) and the number of matches. The average value of both distance metrics when using the Gaussian mixtures as the gap classification methodology yields the best results (DR: 93.9% - RA: 90.8%). The word segmentation module takes as input the result of the proposed text line segmentation technique (Section 3). Although, there is a small improvement in performance between the Gaussian mixture classification methodology and the methodology based on threshold, the advantage of the Gaussian mixture classification methodology is that it is parameter-free, contrary to the methodology based on threshold where an experimentally defined factor is required for the calculation of the final threshold.

Table 2: Experimental results for word segmentation using combinations of distance metrics and gap classification methodologies over the test set of ICDAR2007 handwriting segmentation contest.

Distance Metric		Classification methods										
Euclidean	Convex	LT	GM	*N*	*M*	*DR*	*RA*	*FM*	*o2o*	*Go2m*	*gm2o*	*do2m*	*dm2o*
✓		✓		13311	13322	91.8	91.7	91.7	11933	326	869	410	732
	✓	✓		13311	13249	91.8	92.3	92.0	11953	334	779	367	764
✓	✓	✓		13311	13334	92.4	92.1	92.2	12018	302	828	390	673
✓			✓	13311	13666	93.4	90.3	91.8	12106	202	1137	535	427
	✓		✓	13311	13622	93.2	90.5	91.8	12093	206	1082	514	454
✓	✓		✓	13311	13655	93.9	90.8	92.3	12190	175	1062	503	371

5 Conclusions

This thesis proposed novel methodologies for the segmentation of a handwritten document image into text lines and words which outperform state-of-the-art methodologies. Concerning the text line segmentation stage, the proposed text line segmentation methodology (UOA-HT-EXT) outperformed all existing text line segmentation methodologies on the modern handwritten set, achieving F-Measure of 97.4%, while the second-best text line segmentation methodology scored 97.1%. The modern handwritten set is the test set of ICDAR2007 Handwriting Segmentation Contest.

Regarding the word segmentation stages, both proposed word segmentation methodologies outperformed state-of-the-art word segmentation methodologies. The experiment conducted took as input the text line segmentation result of the UOA-HT-EXT methodology. The Gaussian mixtures methodology achieved F-Measure 92.3% and the methodology based on threshold presented F-Measure 92.2%. The best state-of-the-art methodology achieved F-Measure 91.3%.

Acknowledgments. This thesis has been concluded under the PENED 2001 framework.

References

1. E. Bruzzone, and M. C. Coffetti, "An Algorithm for Extracting Cursive Text Lines", *Proc. 5th Int'l Conf. on Document Analysis and Recognition* (ICDAR'99), 1999, pp. 749-752.

2. M. Arivazhagan, H. Srinivasan, and S. N. Srihari, "A Statistical Approach to Handwritten Line Segmentation", *Document Recognition and Retrieval XIV, Proceedings of SPIE*, 2007, pp. 6500T-1-11.

3. L. A. Fletcher, and R. Kasturi, "A Robust Algorithm for Text String Separation from Mixed Text/Graphics Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.10, no.6, Nov. 1988, pp. 910-918.

4. L. Likforman-Sulem, A. Hanimyan, and C. Faure, "A Hough Based Algorithm for Extracting Text Lines in Handwritten Documents", *Proc. 3rd Int'l Conf. on Document Analysis and Recognition* (ICDAR'95), 1995, pp. 774-777.

5. Y. Pu, and Z. Shi, " A Natural Learning Algorithm Based on Hough Transform for Text Lines Extraction in Handwritten Documents", *Proc. 6th Int'l Workshop on Frontiers in Handwriting Recognition* (IWFHR'98), 1998, pp. 637-646.

6. Z. Shi, and V. Govindaraju, "Line Separation for Complex Document Images Using Fuzzy Runlength", *Proc. 1st Int'l Workshop on Document Image Analysis for Libraries* (DIAL'04), 2004, pp. 306-312.

7. M. Makridis, N. Nikolaou and B. Gatos, "An Efficient Word Segmentation Technique for Historical and Degraded Machine-Printed Documents", *Proc. 9th Int'l Conf. on Document Analysis and Recognition* (ICDAR'07), 2007, pp. 178-182.

8. L. Likforman - Sulem, A. Zahour, and B Taconet, "Text Line Segmentation of Historical Documents: A Survey", Internation Journal on Document Analysis and Recognition (IJDAR), vol. 9, no.2-4, April 2007, pp. 123-138.

9. G. Seni, and E. Cohen, "External Word Segmentation of Off-line Handwritten Text Lines", Pattern Recognition, vol. 27, no. 1, Jan. 1994, pp. 41-52.

10. U. Mahadevan, and R. C. Nagabushnam, "Gap metrics for word separation in handwritten lines", Proc. 3rd Int'l Conf. on Document Analysis and Recognition (ICDAR'95), 1995, pp. 124-127.

11. S.H. Kim, S. Jeong, G.- S. Lee, C.Y. Suen, "Word segmentation in handwritten Korean text lines based on gap clustering techniques", Proc. 6th Int'l Conf. on Document Analysis and Recognition (ICDAR'01), 2001, pp. 189-193.

12. T. Varga, and H. Bunke, "Tree structure for word extraction from handwritten text lines", Proc. 8th Int'l Conf. on Document Analysis and Recognition (ICDAR'05), 2005, pp. 352-356.

13. G. Kim, and V. Govindaraju, "Handwritten Phrase Recognition as Applied to Street Name Images", Pattern Recognition, vol. 31, no. 1, Jan. 1998, pp. 41-51.

14. C. Huang, and S. Srihari, "Word segmentation of off-line handwritten documents", Proc. Annual Symposium on Document Recognition and Retrieval (DRR) XV, IST/SPIE, 2008.

15. F. Luthy, T. Varga, and H. Bunke, "Using Hidden Markov Models as a Tool for Handwritten Text Line Segmentation", Proc. 9th Int'l Conf. on Document Analysis and Recognition (ICDAR'07), 2007, pp. 8-12.

16. G. Louloudis, B. Gatos, I. Pratikakis, and C. Halatsis, "Text line detection in handwritten documents", Pattern Recognition Volume 41, Issue 12, December 2008, Pages 3758-3772.

17. G. Louloudis, B. Gatos, I. Pratikakis, and C. Halatsis, "Text line and word segmentation of handwritten documents", Pattern Recognition Article in Press, doi:10.1016/j.patcog.2008.12.016.

18. G. Louloudis, K. Halatsis, B. Gatos, and I. Pratikakis, "A Block-Based Hough Transform Mapping for Text Line Detection in Handwritten Documents", 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR 2006), pp. 515-520, La Baule, France, October 2006.

19. G. Louloudis, B. Gatos, and C. Halatsis, "Text Line Detection in Unconstrained Handwritten Documents Using a Block-Based Hough Transform Approach", 9th International Conference on Document Analysis and Recognition (ICDAR'07), pp. 599-603, Curitiba, Brazil, September 2007.

20. G. Louloudis, B. Gatos, I. Pratikakis, and C. Halatsis, "Line and Word Segmentation of Handwritten Documents", 11th International Conference on Frontiers in Handwriting Recognition (ICFHR'08), Montreal, Canada, August 2008.