# 11

# Segmentation of Handwritten Document Images into Text Lines

Vassilis Katsouros and Vassilis Papavassiliou
*Institute for Language and Speech Processing/R.C. "Athena"*
*Greece*

## 1. Introduction

There are many governmental, cultural, commercial and educational organizations that manage large number of manuscript textual information. Since the management of information recorded on paper or scanned documents is a hard and time-consuming task, Document Image Analysis (DIA) aims to extract the intended information as a human would (Nagy, 2000). The main subtasks of DIA (Mao et al. 2003) are: i) the document layout analysis, which aims to locate the "physical" components of the document such as columns, paragraphs, text lines, words, tables and figures, ii) the document content analysis, for understanding/labelling these components as titles, legends, footnotes, etc. iii) the optical character recognition (OCR) and iv) the reconstruction of the corresponding electronic document.

The proposed algorithms that address the above-mentioned processing stages come mainly from the fields of image processing, computer vision, machine learning and pattern recognition. Actually, some of these algorithms are very effective in processing machine-printed document images and therefore they have been incorporated in the workflows of well-known OCR systems. On the contrary, no such efficient systems have been developed for handling handwritten documents. The main reason is that the format of a handwritten manuscript and the writing style depend solely on the author's choices. For example, one could consider that text lines in a machine-printed document are of the same skew, while handwritten text lines may be curvilinear.

Text line segmentation is a critical stage in layout analysis, upon which further tasks such as word segmentation, grouping of text lines into paragraphs, characterization of text lines as titles, headings, footnotes, etc. may be developed. For instance, a task for text-line segmentation is involved in the pipeline of the Handwritten Address Interpretation System (HWAIS), which takes a postal address image and determines a unique delivery point (Cohen et al., 1994). Another application, in which text line extraction is considered as a pre-processing step, is the indexing of George Washington papers at the Library of Congress as detailed by Manmatha & Rothfeder, 2005. A similar document analysis project, called the Bovary Project, includes a text-line segmentation stage towards the transcription of the manuscripts of Gustave Flaubert (Nicolas et al., 2004a). In addition, many recent projects, which focus on digitisation of archives, include activities for document image understanding in terms of automatic or semi-automatic extraction and indexing of metadata such as titles, subtitles, keywords, etc. (Antonacopoulos & Karatzas, 2004, Tomai et al., 2002). Obviously, these activities include text-line extraction.

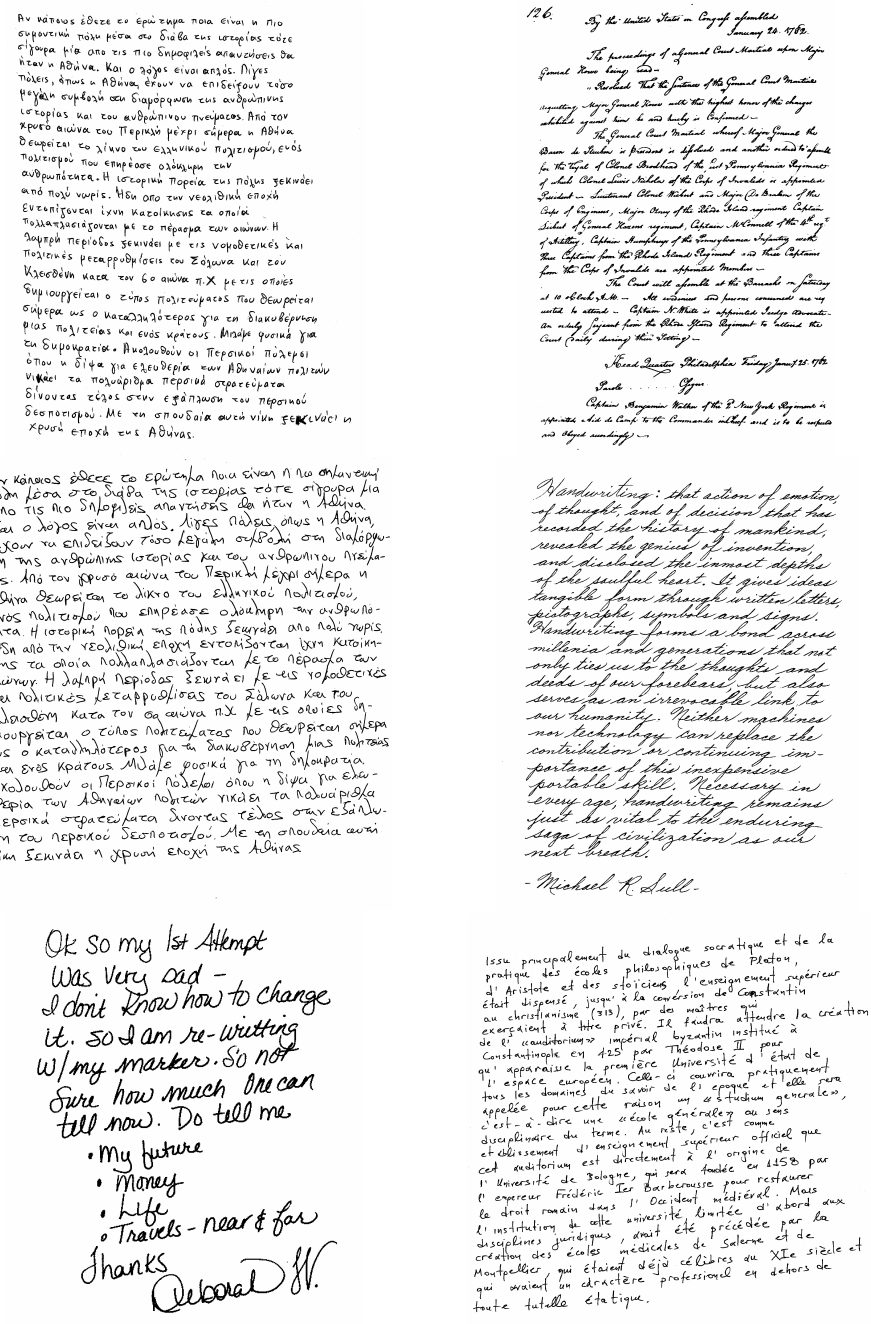

Fig. 1. Handwritten document images included in the training set of ICDAR 2007 Handwriting Segmentation: 024.tif, 016.tif, 010.tif, 044.tif, 068.tif and 026.tif

This chapter is a comprehensive survey of methods exploited to segment handwritten document images into text lines during the last two decades. The main underlying assumption is that the non-textual information has been removed and the document image comprises only plain text. Even though this hypothesis seems to simplify the task, text-line segmentation has to face many challenges, such as the touching or overlapping text lines and the variation of skew angles. Some typical examples of handwritten document images are illustrated in fig. 1. In the next sections, we will describe, in detail, how the proposed methods try to overcome these difficulties.

The main idea in text-line segmentation is to consider the foreground pixel density and employ one of the following three broad classes of techniques (Razak et al., 2008). The first includes traditional methods that have been applied to printed documents and is based on the analysis of projection profiles. The second class incorporates grouping techniques, also known as bottom-up strategies that attempt to build text lines by considering the alignments of foreground pixels or connected components. The third category includes smearing approaches that aim to enhance the text lines structure by applying linear or morphological filters and exploiting well-known image segmentation methods, such as level sets, scale-space analysis (Lindeberg, & Eklundh, 1992), etc. Moreover, there are some methods that exploit a combination of these techniques with the purpose to improve further the segmentation results.

In section 2, the problem definition and the main challenges of the task are described. Several techniques and the contributions of the most effective algorithms within each class are presented in section 3. The available recourses for validating relative methods and the comparative results of recent contests are reported in section 4. Finally, the chapter is concluded with a discussion of the main outcomes.

## 2. Background

When creating a manuscript, the author selects the writing instrument and the paper in such a way as to produce a readable document, namely a document with high contrast between the traces of the pen (foreground) and the paper (background). As a consequence, the digitisation of these documents in most of the cases generates binary images. In the case of grey-scale document images, most of the proposed methods for text-line extraction incorporate an initial processing stage of binarization by employing global (Otsu, 1979) or local thresholding (Niblack, 1986, Sauvola & Pietikäinen, 2000). However, some recent techniques combine the results of processing both the grey-scale and the binary versions of the document image.

In binary document images, the traces of the writing instrument are represented by pixels that have value one and constitute the text. The other pixels have value zero, corresponding to the background. The convention for using the values 1 and 0 for foreground and background pixels respectively is very common in studies related to the binary images.

### 2.1 Definitions

Considering that a two-dimensional binary image is defined on the discrete plane $\mathbb{Z}^2$ and by selecting a square grid and a certain type of connectivity (e.g. 8-n denotes all the neighbours of a pixel, while 4-n indicates only the cross neighbours), we could represent objects or shapes of the image as groups of neighbouring pixels with the same value. In the view of set theory, a binary image is modelled by the corresponding set $S$ as follows:

$$S = \left\{ \mathbf{x} \in \mathbb{Z}^2 : s(\mathbf{x}) = 1 \right\} \quad (1)$$

where $\mathbf{x}$ denotes the coordinates of a pixel and $s$ is a binary function $s : \mathbb{Z}^2 \to \{0,1\}$ (Soille, 2004). Then, the shapes in the binary image are defined as the maximal connected subsets of the image foreground pixels, called connected components (CCs). Therefore, the CCs in a document image could be noise specks, dots, single symbols, groups of touching characters, parts of a character that is broken, etc. The extraction of CCs is accomplished by applying a connected component operator that assigns the same value to every pixel of each distinct CC. A common algorithm for identifying CCs is outlined in (Haralick & Shapiro, 1992).
A text line could be considered as a group of CCs that are adjacent, relative close to each other and correspond to occurrences of text elements. By adopting this simple definition, text-line segmentation produces an image, in which each text pixel has a value that identifies the proper text-line (fig. 2, left). Alternatively, a text line could be represented by a large CC that covers the corresponding part of the image, or by a closed curve that represents the boundary of each text line (fig. 2, right).


Fig. 2. Representation of text lines as: groups of text pixels with the same value (left) and parts of the document image (right)

### 2.2 Challenges

The main challenges of text-line segmentation of handwritten documents arise from the variation of the skew angle between text lines or along the same text line, the existence of overlapping and/or touching lines, the variable character size, the variation of intra-line and inter-line gaps and the non-manhattan layout. To overcome these difficulties, an efficient text-line segmentation algorithm should represent the boundaries of each text line by a closed curve instead of enclosing a text line with a rectangular. In addition, such an algorithm should incorporate procedures for cutting CCs which running along two or more text lines. These are the major differences in handling manuscripts rather than machine-printed documents.

Actually, text line segmentation in printed documents could be seen as a solved problem (Plamondon & Srihari, 2000), which is equivalent with the estimation of the document's skew angle. A comprehensive survey and the annotated bibliography on skew detection of printed document images are presented in (Hull, 1998). To this point, it is supposed that text lines in a printed document have a unique skew angle. Thus, the proper rotation of the image will result to horizontal text lines that could be easily located. A well-known and efficient method for layout analysis of printed documents, called Docstrum, is outlined in (O'Gorman, 1993). The main assumption is that the distances between characters of the same text line are smaller than the distances between characters of successive text lines. The fact that this assumption does not hold for manuscripts, explains why Docstrum cannot handle handwritten documents successfully, as shown in fig. 3.

As mentioned above, this task focuses on text elements only. Therefore, noise removal is the first pre-processing step. In binary document images which incorporate merely textual information, noise removal is equivalent with the elimination of CCs that do not represent text elements but mainly occur due to misses at the digitisation phase. In document images which are considered as normally "clear", simple methods adopting median filters or heuristics based on geometrical and topological properties of the CCs are employed to remove the noisy data. For example, a large CC, which is lying on the edges of the image arises due to the inaccurate placement of the manuscript to the scanner and need to be removed. However, the extraction of actual text elements from digitised historical archives might be a significant issue. Actually, historical documents suffer from smudges, smears, faded print and bleed-through of writing from the opposite side of a page (Likforman-Sulem et al., 2007).

Document images are captured in high resolution (about 300dpi) in order to be suitable for OCR engines. However, text lines have an underlying texture that is manifest in printed documents at low resolutions about 40dpi (Bloomberg, 1996). Hence, subsampling methods that prevent aliasing are also applied in text-line segmentation.

## 3. Proposed methods

Handwritten documents are characterised by high variability of writing styles. Thus, most of the existing methods adapt to the properties of a document image and eliminate the use of prior knowledge. According to the adopted strategy, the existing methods are classified in three categories, which are discussed in this section. In general, text-line segmentation techniques are script independent. However, some special scripts such as Indian and Arabic incorporate many characters with diacritical points that require great care in CCs assignment.

### 3.1 Projection-based methods

Given that an image $\mathbf{A}$ with height $M$ and width $N$ could be considered as a matrix of the same dimensions, the projection profile of the image is defined as follows:

$$P(i) = \sum_{j=1}^{N} \mathbf{A}(i, j), \quad i = 1, ..., M \tag{2}$$

Therefore, the projection is a one-dimensional signal that denotes the amount of text pixels per row. Consequently, the lobes (valleys) of the projection correspond to foreground

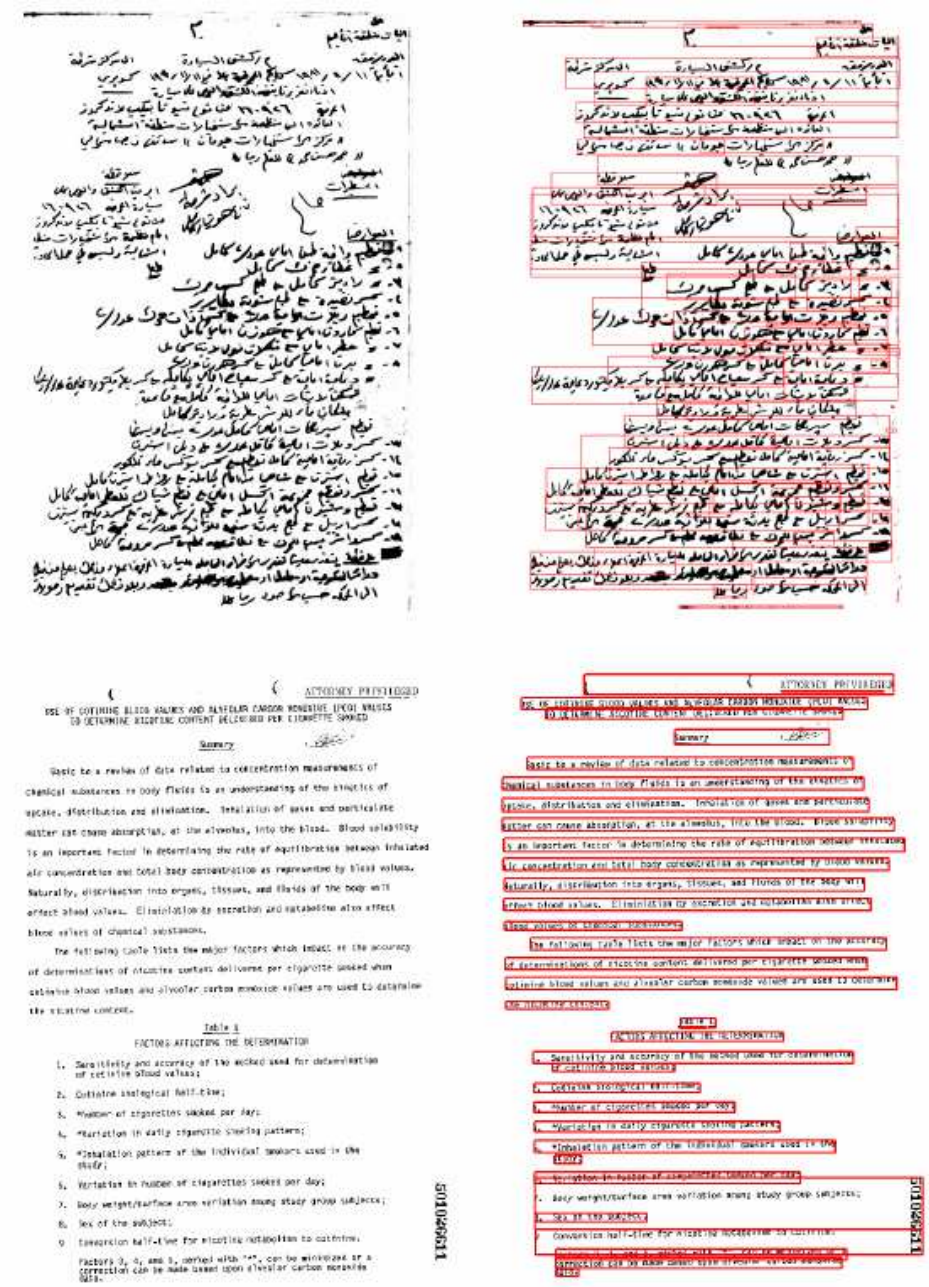

Fig. 3. Text-line segmentation of a machine-printed and a handwritten document image by exploiting the Docstrum method (O'Gorman, 1993). The figure is reprinted (Li et al., 2008) with permission from the author.

(background) areas of the image. Supposing that the text lines have the same skew angle, the amplitude and the frequency of the projection are maximized when the skew of the text is zero. Based on this characteristic, many proposed approaches rotate the image through a range of angles, calculate the projection for each angle and estimate the global skew angle according to a properly selected criterion. Such a criterion could be based on the variance of the projections (Bloomberg et al., 1993) and the sum of the coefficients of the power spectrum (Postl, 1988). After estimating the unique skew angle and rotating the image appropriately, the local minima of the projection allocate the positions of text-line separators.

In machine-printed documents, the separators will be horizontal lines lying within the space between adjacent text lines. Apparently, this might be occurred in some manuscripts written with great care and consistency. In fact, this technique is adopted for the process of 1000 sampled documents from the George Washington corpus at the Library of Congress (Manmatha & Rothfeder, 2005). The additional processing step is the smoothing of the projection by applying a Gaussian low pass filter in order to eliminate false alarms (i.e. insignificant minima) and reduce the noise. A similar method (Santos et al., 2009) includes a post-processing stage for labeling candidate text lines as false or actual, according to their geometrical features (i.e. lines, which correspond to very narrow lobes of the projection, should be removed). Although this approach has been tested in 150 images from the IAM off-line handwritten database (Marti & Bunke, 2002) and showed almost excellent results, it is worth to mention that the text-line segmentation in documents of this database seems to be a straightforward task.

A common feature of manuscripts is the overlapping of successive text lines due to the ascenders and/or descenders of some characters. Hence, the formulation of a horizontal line as a separator is often not feasible. With the purpose to overcome this difficulty, some researchers exploit the projections in order to locate the areas (i.e. the areas between two successive maxima) in which the separators should be allocated. Considering ascenders/descenders as obstacles, the algorithms try to find a path from the left to the right edge in each area, by attempting to move around the obstacles (fig. 4). If the deviation is too high, the algorithm intersects the character and continues forward. Such segmenters could be based on predefined constrains (Yanikoglu & Sandon, 1998) or on the minimization of a proper cost function (Weliwitage et al., 2005).

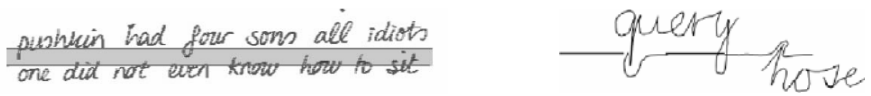


Fig. 4. The line separator should be lying in the gray area between the text lines (left, reprinted from Yanikoglu & Sandon, 1998). The final line separator (right, reprinted from Weliwitage et al., 2005).

Since variation of skew angles between text lines or along the same text line is common in manuscripts, the global projections based approaches cannot provide a general solution. Piece-wise projections can be seen as a modification of global projections to the properties of handwritten documents, by which the separators between text lines are drawn in staircase function fashion along the width of the document page. The main idea of piece-wise projections is to divide the document image in vertical non-overlapping equi-width zones

and find the critical local minima of each projection. The selection of the width of the zones is a trade-off between the local skew and the text density. In other words, if the width was large enough, the skew should not be considered as constant. Furthermore, a narrow width would produce zones, which do not include adequate amount of text. Relative experiments showed that a zone width equal to 5% of the document image width seems to be an appropriate value.

However, the non-manhattan layout of manuscripts will result in vertical zones without enough foreground pixels for every text line. In such cases, some local minima may be lost and the results of two adjacent zones may be ambiguous. To deal with these problems, we calculate a smooth version of the projections influenced by the neighboring zones and introduce a separator-drawing algorithm that combines separators of consecutive zones according to their proximity and the local text density as shown in fig. 5 (Papavassiliou et al. 2010).

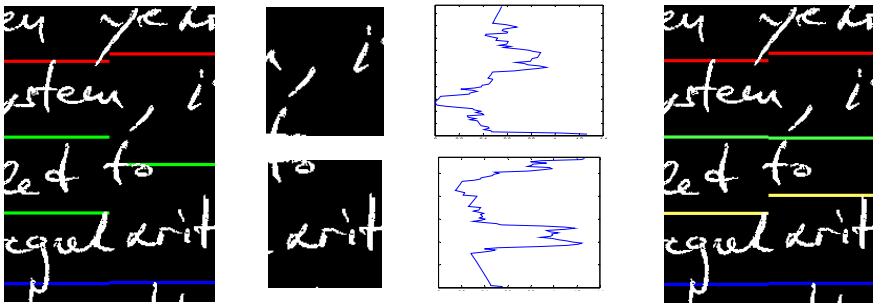

Fig. 5. The separator-drawing algorithm (reprinted from Papavassiliou et al., 2010). The separators in green are ambiguous (first column). New separators in these areas (second column) should be located at the global minima of the metric function (third column) influenced by the local foreground density and the proximity of the separators. The separators with the same colors are associated (fourth column).

The last challenge that projection-based methods have to face is the assignment of CCs to the proper text-lines. In most of the cases, this is a straightforward task since the majority of CCs lie between two line separators. However, some CCs either overlap with two text lines (i.e. characters with ascenders/descenders) or run along two text lines (touching lines). In order to preserve the ascending/descending symbols from been corrupted by arbitrary cuts, several heuristics based on the geometrical and topological properties such as the height, the length, the distance of neighboring CCs, etc. have been proposed. An interesting approach models the text lines by bivariate Gaussian densities considering the coordinates of the pixels of the CCs that have been already assigned. Then, the probabilities that the CC under consideration belongs to the upper or lower text lines are estimated and the decision is made by comparing the probabilities (Arivazhagan et al., 2007). This method has been tested on 720 documents including English, Arabic and children's handwriting and performed a detection rate of 97.31%. In the case that a character should be split, the segmentation occurs at a proper cross point of the skeleton (Lam et al., 1992) by taking account the distance from the separator (fig. 6) as well as the slope and curvature of the stroke (Kuzhinjedathu et al., 2008).

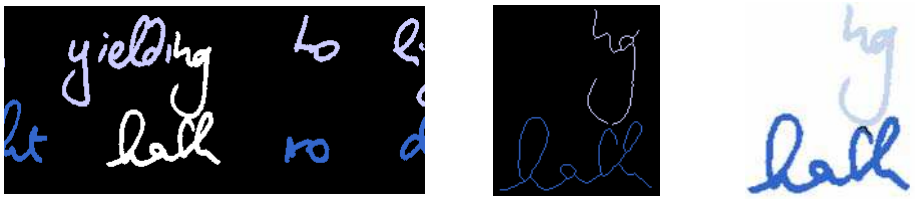

Fig. 6. Segmentation of a CC running along two text lines (reprinted from Papavassiliou et al., 2010).

### 3.2 Grouping methods

Grouping approaches, also known as bottom-up techniques, were very popular in text-line segmentation of handwritten documents due to their success in prior tasks concerning the process of machine-printed documents (Simon et al., 1997). These methods try to group CCs considering geometrical and topological characteristics of the CCs such as their distances, locations and orientation. The common strategy is to represent each CC with an appropriate vector (e.g. the coordinates of its gravity centre), calculate the distances between that point and the corresponding points of its neighbouring CCs and compare the distances with a proper predefined value. If the constrain is satisfied the CCs are grouped (Khandelwal et al., 2009). Since such methods strongly depend on the values of the thresholds, they cannot handle variation in writing styles. In fact, (Feldbach & Tonnies, 2001) report that a similar method tested on historical church registers achieved a 97% recall rate when the thresholds values are adjusted to specific authors but decreased to 90% when these parameters remained constant for various authors. As a result, many recent methods produce an adjacency graph constructed by linking the pairs of neighbouring CCs with edges. Then, they try recursively to find the minimum spanning tree, which likely crosses CCs of the same text line (Nicolas et al., 2004b).

Additionally, the orientations of the edges that connect these points are also examined. Supposing that CCs in the same text line could be represented by almost collinear points, Hough transform (Duda, & Hart, 1972) has been applied on handwritten documents. Although Hough-based approaches locate text lines with different skew angles correctly, they are not flexible to follow variation of skew along the same text line (fig. 7 left).

### 3.3 Smearing methods

In general, smearing approaches include two main processing steps. The first stage aims to enhance text areas by blurring the input image. The second step concerns the modification and use of well-known image segmentation methods in order to formulate the text lines. Li et al. (2008) apply an anisotropic Gaussian filter to smooth the image and provide a grey-scale "probabilistic" image that denotes the text line distribution. It is worth to mention that the horizontal dimension of the filter is greater than the vertical in order to advance the mainly horizontal text-line orientation. Then, they locate the initial boundaries of text lines or parts of them using Niblack's algorithm for binarization and finding the contours of the CCs in the produced binary image. Next, the level set method (Osher & Fedkiw, 2003) is adopted and the boundaries evolve concerning the local curvature and density with the purpose of moving fast along the horizontal direction and towards areas with high

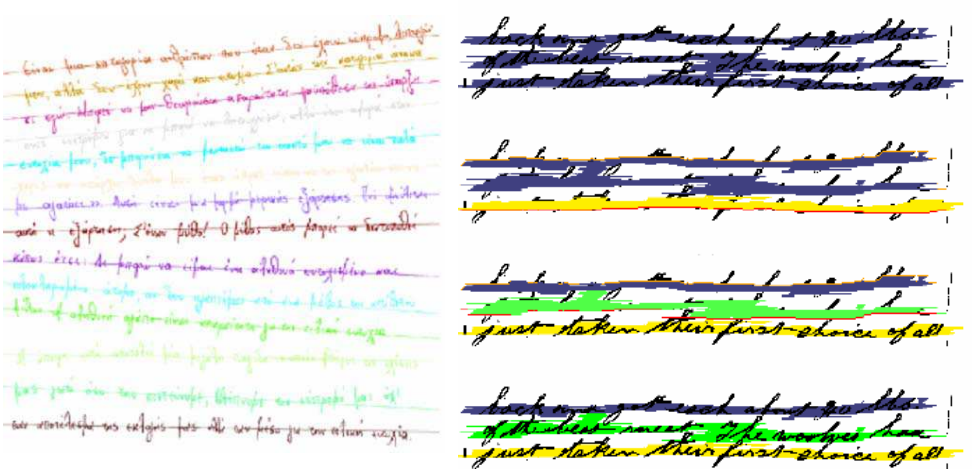

Fig. 7. (left) Results of Hough-based transform (reprinted from Louloudis et al., 2008); (right) Segmentation of CCs based on min-cut/max-flow algorithm (reprinted from Kennard & Barret, 2006).

probability being text. This method has been tested on several manuscripts in different scripts and performed pixel-level hit rates varying from 92% to 98%. As expected, the proposed method fails when the gap between two neighbouring text lines is smaller than the vertical dimension of the Gaussian filter. A similar approach adopts the Mumford-Shah model to locate text lines and then applies morphological operations to either segment merged text lines or join parts of the same text line (Du et al., 2008). Instead of using erosions and dilations Yin & Liu (2009) apply a modification of the variational Bayes framework to the downsampled input image. The binary image (after down-sampling, smoothing and binarization) is considered as a mixture model in which each CC is a Gaussian component. Since, a CC may correspond to more than one text lines, a CC is split according to the second eigenvalue of the covariance (i.e. thick CCs are candidate to be segmented).

Other smearing strategies enhance the text areas by estimating an adaptive local connectivity map (ALCM). This map is actually a grey-scale image in which each pixel has a value that denotes the amount of text pixels in the proximity of the pixel under consideration. By converting this image to a binary one, the resulting CCs represent the text areas of the document image. In most cases, these CCs include many text lines and should be split. Considering such a CC as a graph with candidate source (sink) nodes the pixels in the upper (lower) part, the min-cut/max flow algorithm has been proposed for segmenting the CC to its main components (Kennard & Barret, 2006), as illustrated in fig. 7 right. Alternatively, the ALCM could be replaced by another image produced by applying the run length smoothing algorithm (RLSA). In this image, the value of a pixel is the distance of that pixel from the nearest text pixel. As previously, text areas are represented by pixels with small (dark) values while background corresponds to bright areas. Then, dark areas are grouped according to their orientation and proximity in order to formulate text lines (Shi & Govindaraju, 2004).

## 4. Evaluation

Research groups test their method either on their own collection of handwritten documents or on a public available database. As a result, many such test sets have been constructed. In this section, we refer to three well-known collections that could be used for evaluating various processing steps such as text-line segmentation, word segmentation and character recognition. The IAM handwriting database[1] consists of 1539 pages of scanned text containing 13353 and 115320 isolated and labelled text lines and words respectively. Although, this database is an excellent resource for validating word segmentation and character recognition algorithms, the text-line extraction seems not to be a complex task due to significant gaps between successive text lines in many images. Another famous database is the NIST Handprinted Forms and Characters Database[2], which includes handwritten sample forms from 3600 writers. This collection is mainly used for evaluating character recognition techniques but it could be also employed to assess text-line segmentation algorithms.

The other two benchmarking databases are the training and test sets constructed for the Handwriting Segmentation Contests in the context of ICDAR 2007[3] and 2009[4]. The first collection consists of 100 images (20 and 80 for training and test, respectively). The second database includes these images (as the training set) and 200 images that construct the test set. The documents are either modern manuscripts written by several writers in several languages (English, French, German and Greek) or historical handwritten archives, or document samples selected from the web. It is worth to mention that none of the documents includes any non-text elements (lines, drawings, etc.)

The comparative results of the algorithms, which participated in ICDAR 2007 Handwriting Segmentation Contest (Gatos et al., 2007) or have been tested on this dataset, are presented in Table 1. Detection Rate (DR) denotes the ratio between the number of text lines detected correctly and the number of ground-truth lines (1771). Similarly, Recognition Accuracy (RA) is calculated by dividing the number of correctly detected lines with the total number of detected text lines. FM denotes the harmonic mean of DR and RA.

| | DR(%) | RA(%) | FM% |
|---|---|---|---|
| **BESUS (Das et al., 1997)** | 86.6 | 79.7 | 83.0 |
| **DUTH-ARLSA** | 73.9 | 70.2 | 72.0 |
| **ILSP-LWSeg (Papavassiliou et al., 2010)** | 97.3 | 97.0 | 97.1 |
| **PARC** | 92.2 | 93.0 | 92.6 |
| **UoA-HT (Louloudis et al., 2008)** | 95.5 | 95.4 | 95.4 |
| **PROJECTIONS** | 68.8 | 63.2 | 65.9 |
| **Ridges-Snakes (Bukhari et al., 2009)** | 97.3 | 95.4 | 96.3 |
| **Shredding (Nikolaou & Gatos, 2009)** | 98.9 | 98.3 | 98.6 |

Table 1. Evaluation results of algorithms tested on the database of ICDAR2007 Handwriting Segmentation Contest.

[1] http://www.iam.unibe.ch/fki/databases/iam-handwriting-database

[2] http://www.nist.gov/srd/nistsd19.cfm

[3] http://users.iit.demokritos.gr/~bgat/HandSegmCont2007/resources.htm

[4] http://users.iit.demokritos.gr/~bgat/HandSegmCont2009/resources.html

The evaluation results of algorithms participated in ICDAR2009 Handwriting Segmentation Contest are presented in Table 2. We mention that the test set consists of 200 binary images and their dimensions vary from 650x825 to 2500x3500 pixels. The total number of text lines included in this dataset is 4034.

| | DR(%) | RA(%) | FM(%) |
|---|---|---|---|
| **CASIA-MSTSeg (Yin & Liu, 2008)** | 95.86 | 95.51 | 95.68 |
| **CMM** | 98.54 | 98.29 | 98.42 |
| **CUBS (Shi et al., 2009)** | 99.55 | 99.50 | 99.53 |
| **ETS** | 86.66 | 86.68 | 86.67 |
| **ILSP-LWSeg-09** | 99.16 | 98.94 | 99.05 |
| **Jadavpur Univ** | 87.78 | 86.90 | 87.34 |
| **LRDE** | 96.70 | 88.20 | 92.25 |
| **PAIS** | 98.49 | 98.56 | 98.52 |
| **AegeanUniv (Kavallieratou et al., 2003)** | 77.59 | 77.21 | 77.40 |
| **PortoUniv (Cardoso et al., 2008)** | 94.47 | 94.61 | 94.54 |
| **PPSL** | 94.00 | 92.85 | 93.42 |
| **REGIM** | 40.38 | 35.70 | 37.90 |

Table 2. Comparative results of ICDAR2009 Handwriting Segmentation Contest.

PAIS and ILSP are based on piece-wise projections and achieved high results. On the other hand, similar methods presented poor results because they either adopt global projections (PROJECTIONS) or divide the image into only three vertical zones (AegeanUniv).

Ten participating methods are classified as grouping approaches. In particular, five methods (Jadavpur Univ, CASIA-MSTSeg, CMM, PPSL and REGIM) introduce constrains on the topological and geometrical properties of the CCs in order to create groups of CCs that correspond to text lines. Since, these approaches require many predefined thresholds, the selection of appropriate (improper) values results in good (poor) results. Three approaches (BESUS, ETS and PARC) apply morphological operations to produce new CCs by merging the initial neighbouring CCs and then adopt similar constrains. Another grouping approach is UoA-HT, which exploits the Hough transform. As expected, the algorithm is not very effective when the skew of a text line varies along its width. Although, two methods (DUTH-ARLSA and CUBS) exploit the RLSA algorithm, their results differ significantly. The reason is that CUBS applies the RLSA algorithm in five directions ($-20^o$, $-10^o$, $0^o$, $10^o$, and $20^o$) and combines the results in order to calculate the local skew of each text line.

PortoUniv proposes a tracing algorithm that tries to find proper paths that connect the edges of the image without cutting the textual elements. A similar approach (Shredding) includes a pre-processing step for blurring and then exploits the tracing algorithm. LRDE is a fast algorithm that enhances the test areas by anisotropic Gaussian filtering, smoothes the image by applying morphological operations and segments it by using the watershed transform (Vincent & Soille, 1991). A recent method (Ridges-Snakes) uses a multi-oriented anisotropic Gaussian filter bank for smoothing, approximates the ridges as the central lines of the text parts and then the ridges evolve until they overlap the CCs of the manuscript.

## 5. Conclusions

After reviewing the existing methods for text-line segmentation we conclude that there are pros and cons for each approach. For example, piece-wise projection based methods can handle text lines with varying skew angles, but fail when the document includes high degree of curl text lines. In addition, the benefit of some grouping strategies is that they succeed to extract text lines from a complex layout but may fail to segment touching text lines. Regarding smearing approaches, some of them seem to be promising since they exploit image segmentation algorithms that have been already applied on other kinds of images. However, they may merge two successive text lines if the gap between them is not large enough. As a conclusion, we report that the existing methods do not generalize very well to all possible variations encountered in handwritten documents.

Thus, text-line segmentation of handwritten documents remains an open issue. This fact explains why the number of relative papers and contests is increasing. Since different methods can face different challenges of this task, we foresee that a combination of complementary techniques could result in a generalized solution.

## 6. References

Antonacopoulos A. & Karatzas D (2004), Document Image analysis for World War II personal records, *Proceedings of International Workshop on Document Image Analysis for Libraries DIAL'04*, pp. 336-341, ISBN 0-7695-2088-X, Palo Alto, USA, January 23-24, 2004

Arivazhagan, M.; Srinivasan, H. & Srihari, S. (2007). A statistical approach to line segmentation in handwritten documents, *Proceedings of Document Recognition and Retrieval XIV SPIE*, Vol.6500, No.1, ISSN 0277-786X, San Jose, CA, USA, January 30, 2007

Bloomberg, D.S. & Kopec, G. (1993). Method and apparatus for identification and correction of document skew. Xerox Corporation, U.S. Patent 5,563,403, October 8, 1996.

Bloomberg, D.S. (1996). Textured Reductions for Document Image Analysis, *Proceedings of IS&T/SPIE EI '96, Conference 2660: Document Recognition III*, pp. 160-174, ISBN: 9780819420343, San Jose, USA, March 7, 1996

Bukhari, S.S.; Shafait, F. & Bruel, T.M. (2009). Script-Independent Handwritten Textlines Segmentation using Active Contours, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 446-450, ISSN 1520-5363, Barcelona, Spain, July 26-29, 2009

Cardoso, J.S.; Capela, A.; Rebelo, A. & Guedes, C. (2008). A Connected Path Approach for Staff Detection on a Music Score, *Proceedings of International Conference on Image Processing*, pp.1005-1008, ISSN 1522-4880, San Diego, CA, USA, October 2-15, 2008

Cohen, E.; Hull, J.J. & Srihari, S.N. (1994). Control Structure for Interpreting Handwritten Addresses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.16, No.10, (October 1994), pp. 1049- 1055, ISSN 0162-8828

Das, A.K.; Gupta, A. & Chanda, B. (1997). A Fast Algorithm for Text Line & Word Extraction from Handwritten Documents, *Image Processing & Communications*, Vol. 3, No. 1-2, pp. 85-94

Du, X.; Pan, W. & Bui, T.D. (2009). Text Line Segmentation in Handwritten Documents Using Mumford-Shah Model, *Pattern Recognition,* Vol. 42, No. 12, (December 2009), pp. 3136-3145, ISSN 0031-3203

Duda, R.O. & Hart, P. E. (1972). Use of the Hough Transformation to Detect Lines and Curves in Pictures, *Communications of the ACM*, Vol. 15, No. 1, (January 1972), pp. 11-15,

Feldbach M. & Tonnies, K.D. (2001) Line Detection and Segmentation in Historical Church Registers, Proceedings of International Conference on Document Analysis and Recognition, pp. 743-747, ISBN 0-7695-1263-1, Seattle, WA , USA, September 10-13, 2001

Gatos, B.; Antonacopoulos, A. & Stamatopoulos, N. (2007). ICDAR2007 Handwriting Segmentation Contest, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 1284-1288, Curitiba, Brazil, September 23-26, 2007

Gatos, B.; Stamatopoulos, N. & Louloudis, G. (2009). ICDAR2009 Handwriting Segmentation Contest, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 1393-1397, ISSN 1520-5363, Barcelona, Spain, July 26-29, 2009

Haralick, R.M. & Shapiro, L.G. (1992). Computer and Robot Vision, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, ISBN:0201569434

Hull, J.J. (1998). Document Image Skew Detection: Survey and Annotated Bibliography, In: Document Analysis Systems II, Hull, J.J. & Taylor, S.L. (Eds.), pp. 40-64, World Scientific Publishing Co. Pte. Ltd, ISBN 978-981-02-3103-3

Kavallieratou, E.; Dromazou, N.; Fakotakis, N. & Kokkinakis, G. (2003). An Integrated System for Handwritten Document Image Processing, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 17, No. 4, pp. 101-120, DOI 10.1.1.107.7988

Khandelwal, A.; Choudhury, P.; Sarkar, R.; Basu, S.; Nasipuri, M. & Das, N. (2009). Text Line Segmentation for Unconstrained handwritten Document Images Using Neighborhood Connected Components Analysis, *Lecture Notes in Computer Science*, Vol. 5909/2009, pp. 369-374, DOI: 10.1007/978-3-642-11164-8_60

Kuzhinjedathu, K.; Srinivansan, H. & Srihari, S. (2008). Robust Line Segmentation for Handwritten Documents, *Proceedings of Document Recognition and Retrieval XV ST/SPIE Annual Symposium*, Vol. 6815, San Jose, CA, January 2008

Lam, L.; Lee, S.W. & Suen, C.Y. (1992). Thinning Methodologies-A Comprehensive Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 14, No. 9, (September 1992), pp. 869-885, ISSN 0162-8828

Li, Y.; Zheng, Y.; Doermann, D. & Jaeger, S. (2008). Script-Independent Text Line Segmentation in Freestyle Handwritten Documents, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.30, No.8, (August 2008), pp. 1313-1329, ISSN 0162-8828

Likforman-Sulem, L.; Zahour, A. & Taconet, B. (2007). Text-line Segmentation of Historical Documents: a Survey, *International Journal on Document Analysis and Recognition,* Vol. 9, No. 2, (April 2007), pp. 123–138, ISSN:1433-2833

Lindeberg, T. & Eklundh, J.O. (1992). Scale-Space Primal Sketch: construction and experiments, *Image and Vision Computing*, Vol. 10, No.1, (January-February 1992), pp. 3–18, doi 10.1016/0262-8856(92)90079-I

Louloudis, G.; Gatos, B. & Halatsis, C. (2008). Text line detection in handwritten documents, *Pattern Recognition*, Vol.41, No.12, (December 2008), pp. 3758-3772, ISSN 0031-3203

Manmatha, R. & Rothfeder, J.L. (2005). A Scale Space Approach for Automatically Segmenting Words from Historical Handwritten Documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No.8, (August 2005), pp. 1212-1225, ISSN 0162-8828

Mao, S.; Rosenfeld, A. & Kanungo, T. (2003). Document Structure Analysis Algorithms: A Literature Survey, *Proceedings of Document Recognition and Retrieval X SPIE*, Vol. 5010, pp. 197-207, ISBN 0-8194-4810-9, Santa Clara, California, USA, January 22-23, 2003

Marti, U.V. & Bunke, H. (2002). The IAM-Database: an English sentence database for off-line handwriting recognition, *International Journal on Document Analysis and Recognition*, Vol.5, No. 1, (), pp. 39-46, DOI: 10.1007/s100320200071

Nagy, G. (2000). Twenty Years of Document Image Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, (January 2000), pp. 38–62, ISSN 0162-8828

Niblack, W. (1986). An Introduction to Digital Image Processing, pp. 115-116, Prentice Hall, ISBN-13: 978-0134806747

Nicolas, S.; Paquet, T. & Heutte, L. (2004a). Enriching Historical Manuscripts: The Bovary Project, In: Document Analysis Systems VI, LNCS 3163, pp. 135-146, Marinai, S. & Dengel, A. (Eds.) Springer-Vergal, ISBN 3-540-23060-2, Berlin/Heidelberg

Nicolas, S.; Paquet, T. & Heutte, L. (2004b). Text Line Segmentation in Handwritten Document Using a Production System, *Proceedings of International Workshop on Frontiers in Handwriting Recognition,* pp. 245–250, ISSN 1550-5235, Tokyo, Japan, October 26-29, 2004

Nikolaou, A. & Gatos, B. (2009). Handwritten Text Line Segmentation by Shredding Text into its Lines, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 1393-1397, ISSN 1520-5363, Barcelona, Spain, July 26-29, 2009

O'Gorman, L. (1993). The document spectrum for page layout analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No.11, (November 1993), pp. 1162–1173, ISSN 0162-8828

Osher, S. & Fedkiw, R. (2003). Level Set Methods and Dynamic Implicit Surfaces, Springer, ISBN 978-0-387-95482-0

Otsu, N. (1979). A Threshold Selection Method From Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, Vol.9, No.1, (January 1979), pp. 62-66, ISSN 0018-9472

Papavassiliou, V.; Stafylakis, T.; Katsouros, V. & Carayannis, G. (2010). Handwritten Document Image Segmentation into Text Lines and Words, *Pattern Recognition*, Vol.43, No.1, (January 2010), pp. 369-377, ISSN 0031-3203

Plamondon, R. & Srihari, S.N. (2000). On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey, *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol.22, No.1, (January 2000), pp. 63–84, ISSN 0162-8828

Postl, W. (1988). Method for automatic correction of character skew in the acquisition of a text original in the form of digital scan results. Siemens AG, U.S. Patent 4,723,297, February 2, 1988

Razak, Z.; Zulkifee, K.; Idris, M.Y.I.; Tamil, E.M.; Noor, M.N.M.; Salleh, M.; Yaakob, M.; Yusof, Z.M. & Yaacob, M. (2008). Off-line Handwriting Text Line Segmentation: A

Review, *International Journal of Computer Science and Network Security*, Vol. 8, No. 7, (July 2008), pp. 12-20, http://paper.ijcsns.org/07_book/200807/20080703.pdf

Santos, R.P.; Clemente, G.S.; Ren, T.I. & Calvalcanti, G.D.C. (2009). Text Line Segmentation Based on Morphology and Histogram Projection, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 651-655, ISBN 978-1-4244-4500-4, Barcelona, Spain, July 26-29, 2009

Sauvola, J. & Pietikäinen, M. (2000). Adaptive document image binarization, *Pattern Recognition*, Vol. 33, No. 2, (February 2000), pp. 225-236, doi:10.1016/S0031-3203(99)00055-2

Shi, Z. & Govindaraju, V. (2004). Line Separation for Complex Document Images Using Fuzzy Runlength, Proceedings of International Workshop on Document Image Analysis for Libraries, pp. 306-312, ISBN : 0-7695-2088-X, Palo Alto, USA, January 23-24, 2004.

Shi, Z.; Seltur, S. & Govindaraju, V. (2009). A Steerable Directional Local Profile Technique for Extraction of Arabic Text Lines, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 176-180, ISBN 978-1-4244-4500-4, Barcelona, Spain, July 26-29, 2009

Simon, A.; Pret, J.C. & Johnson, A.P. (1997). A Fast Algorithm for Bottom-Up Document Layout Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 3, (March 1997) pp. 273-277, ISSN 0162-8828

Soille, P. (2004). Morphological Image Analysis Principles and Applications, (2nd ed.), Springer-Verlag New York, Inc. Secaucus, NJ, USA, ISBN:3540429883

Tomai C. I.; Zhang B. & Govindaraju V. (2002), Transcript Mapping for Historic Handwritten Document Images, *Proceedings of International Workshop on Frontiers in Handwriting Recognition IWFHR2002*, pp. 413 - 418, ISBN 0-7695-1692-0, Ontario, Canada, August 6-8, 2002.

Vincent, L. & Soille, P. (1991). Watershed in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 6, (June 1991) pp. 583-598, ISSN 0162-8828

Weliwitage, C.; Harvey, A.L. & Jennings, A.B. (2005). Handwritten Document Offline Text Line Segmentation, *Proceedings of Digital Image Computing: Techniques and Applications Conference (DICTA2005)*, pp. 27, ISBN: 0-7695-2467-2, Cairns, Australia, December 6-8, 2005

Yanikoglu, B. & Sandon, P.A. (1998). Segmentation of Off-line Cursive Handwriting Using Linear Programming, *Pattern Recognition*, Vol.31, No.12, (December 1998), pp. 1825-1833, doi:10.1016/S0031-3203(98)00081-8

Yin, F. & Liu, C.L. (2008). Handwritten Text Line Segmentation by Clustering with Distance Metric Learning, *Proceedings of International Conference on Frontiers in Handwriting Recognition (ICFHR 2008)*, pp.229-234, Montreal, Quebec, Canada, August 19-21, 2008

Yin, F. & Liu, C.L. (2009). A Variational Bayes Method for Handwritten Text Line Segmentation, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 436-440, ISSN: 1520-5363, Barcelona, Spain, July 26-29, 2009

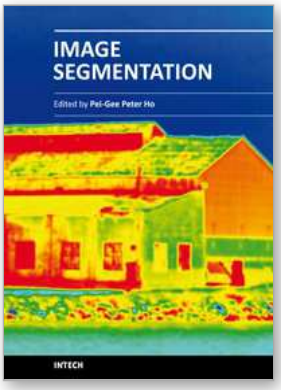


**Image Segmentation**
Edited by Dr. Pei-Gee Ho

ISBN 978-953-307-228-9
Hard cover, 538 pages
**Publisher** InTech
**Published online** 19, April, 2011
**Published in print edition** April, 2011

It was estimated that 80% of the information received by human is visual. Image processing is evolving fast and continually. During the past 10 years, there has been a significant research increase in image segmentation. To study a specific object in an image, its boundary can be highlighted by an image segmentation procedure. The objective of the image segmentation is to simplify the representation of pictures into meaningful information by partitioning into image regions. Image segmentation is a technique to locate certain objects or boundaries within an image. There are many algorithms and techniques have been developed to solve image segmentation problems, the research topics in this book such as level set, active contour, AR time series image modeling, Support Vector Machines, Pixon based image segmentations, region similarity metric based technique, statistical ANN and JSEG algorithm were written in details. This book brings together many different aspects of the current research on several fields associated to digital image segmentation. Four parts allowed gathering the 27 chapters around the following topics: Survey of Image Segmentation Algorithms, Image Segmentation methods, Image Segmentation Applications and Hardware Implementation. The readers will find the contents in this book enjoyable and get many helpful ideas and overviews on their own study.

**How to reference**
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Vassilis Katsouros and Vassilis Papavassiliou (2011). Segmentation of Handwritten Document Images into Text Lines, Image Segmentation, Dr. Pei-Gee Ho (Ed.), ISBN: 978-953-307-228-9, InTech, Available from: http://www.intechopen.com/books/image-segmentation/segmentation-of-handwritten-document-images-into-text-lines

INTECH
open science | open minds

**InTech Europe**
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

**InTech China**
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821