# Unsupervised Wall Detector in Architectural Floor Plans

Lluís-Pere de las Heras, David Fernández, Ernest Valveny, Josep Lladós and Gemma Sánchez

Computer Vision Center - Dept. Ciències de la Computació

Universitat Autònoma de Barcelona

Barcelona, Catalonia, Spain

Email: {lpheras,dfernandez,ernest,josep,gemma}@cvc.uab.cat

*Abstract*—**Wall detection in floor plans is a crucial step in a complete floor plan recognition system. Walls define the main structure of buildings and convey essential information for the detection of other structural elements. Nevertheless, wall segmentation is a difficult task, mainly because of the lack of a standard graphical notation. The existing approaches are restricted to small group of similar notations or require the existence of pre-annotated corpus of input images to learn each new notation. In this paper we present an automatic wall segmentation system, with the ability to handle completely different notations without the need of any annotated dataset. It only takes advantage of the general knowledge that walls are a repetitive element, naturally distributed within the plan and commonly modeled by straight parallel lines. The method has been tested on four datasets of real floor plans with different notations, and compared with the state-of-the-art. The results show its suitability for different graphical notations, achieving higher recall rates than the rest of the methods while keeping a high average precision.**

## I. Introduction

The analysis of floor plans is one of the central applications of Graphics Recognition. A number of automatic tools have been incorporated into CAD platforms. Examples are the classical raster to CAD conversion systems [1], 3D visualization from printed floor plans [2], [3], or hand drawn sketch interpretation [4] in the early conceptualization stages of the architects work. Very recently, several works have been proposed beyond the analysis of floor plan for creativity activities but at functional level. Here, the interpretation of the structure and the space is very relevant, so the detection of walls is the key in this step [5], [6], [7], [8], since they bear inherent information of other structural elements, such as rooms, windows, doors, etc.

A number of wall detection techniques have been presented in the literature, most of them based on the structural grouping of some basic primitives. In [3], the parallel pairs of lines are firstly detected. Then, text information and an established set of graphical rules are used to guide the semantic wall detection. Contrarily, in [2], walls are modeled by dashed lines and detected by applying a morphological filtering. In [5], walls are recognized after finding the parallel lines encountered using a combination of Hough Transform and image vectorization. Only the couple of parallel lines detected with black pixels in between are considered as final walls. More recently, in [7] walls are segmented by iteratively performing erosions followed by dilatations, permitting to differentiate between thick, medium and thin walls.

These approaches are able to detect walls successfully in their own tested datasets. Nevertheless, the lack of a standard notation in floor plans leads to different graphical modeling depending on the architectural offices, or countries. Thus, walls can be modeled very differently from plan to plan. This fact provokes that state-of-the-art strategies, which are strongly notation oriented, need to be reformulated for every floor plan that contains a new graphical notation. With the aim of giving solution to this issue, we presented in [9] a supervised patch-based wall segmentation method able to deal with different notations. The main drawback of this approach is given by the need of ground-truthed data for each new notation to learn the graphical appearance of walls. Since the manual task of labeling is tedious and subject to errors, this approach is appropriate for a controlled set of notations but fails into being a reasonable approach to generally solve the problem.

In this paper we present, to the best of our knowledge, the first unsupervised segmentation system which is able to segment walls independently to their notation. Conversely to [9], it automatically adapts to every wall notation; without the need of labeled data to learn their graphical appearance. The approach, based on some general properties of walls inherent in all floor plans, generate multiple segmentation candidates and select the one which better characterizes this element. We compare its performance on four labeled datasets of different notations with two recent wall segmentation strategies [7], [9], showing that it surpass them in recall terms. Moreover, to enhance the strength of our method, which is the ability of segmenting walls of any plan when no ground-truth is available, we qualitatively show its robustness on some challenging real images of different notations and resolutions extracted directly from the Internet.

The rest of the paper is organized as follows. Firstly, in section II the methodology is explained in deep. Secondly, in section III the datasets where our approach have been tested, the evaluation protocol used and the quantitative and qualitative results are presented. We finally conclude the paper in section IV.

## II. Methodology

The underlying idea of the method for segmentation of walls is based on a flexible combination of 6 general premises for characterizing walls, called *wall-assumptions*:

1) Walls are modeled by parallel lines.
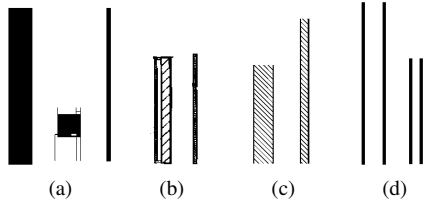2) They are rectangular; longer than thicker.

Fig. 1: Real graphical examples of vertical walls for *Dataset Black* in (a), *Dataset Textured* in (b), *Dataset Textured2* in (c), and *Dataset Parallel* in (d).

3) Walls appear in orthogonal directions.
4) Different thickness is used for external and internal walls.
5) The walls of a document are filled by the same pattern (hatched, tiled, solid, empty, etc.).
6) They appear repetitively and naturally distributed among the plan.

As it is shown in Figure 1, this set of *wall-assumptions* are far from being a collection of unbreakable statements that perfectly define walls in their graphic composition. Nevertheless, the appropriate combination of them is used to guide the final segmentation.

Figure 2 shows the pipeline of our approach. Firstly, the image is preprocessed to filter unnecessary information and to facilitate the posterior segmentation. After that, the edges of the input image are computed just in case black thick walls are present. Otherwise, the raster image is considered as input. Then, using run length analysis, the parallel lines in the plan are detected and the distances between them are quantized in a histogram. Outstanding values of the histogram correspond to frequent runs likely to define wall segments. Finally, the final wall segmentation is given by the combination of wall image candidates according to the assumptions postulated above.

### A. Pre-process

The input images are preprocessed as in [6] and [9]. Firstly, all images are binarized to reduce the dimensionality of the input space. Then, since textual information is not relevant for wall segmentation, it is filtered out using [10]. In addition to that, we detect and correct possible deviations in floor plan orientation by adapting the approach for handwritten text deskewing [11]. Finally, floor plans with resolutions higher than 4000×4000 are down scaled for efficiency issues.

### B. Black-walls detection

Despite walls are usually drawn by parallel lines with a repetitive graphical pattern –or texture– inbetween them or a lack of it, there are floor plans that include walls graphically composed by black thick lines –called black-walls for clarity–, as the ones shown in Figure 1a. Since we base the detection in *wall-assumption 1*, that asserts that they are composed by at least two parallel lines, we need to detect the floor plans that contain this sort of walls and transform them into a more suitable input; the edge image.

In order to detect the existence of this kind of walls, run lengths over the foreground pixels in the horizontal and vertical directions are quantized in a histogram. Floor plans with black walls present more sparse frequencies with significant out-layers in high positions. Contrarily, the runs in images with a lack of black-walls are distributed normally-like in the lower bins. A Gaussian Mixture is fitted into the 1D data using the Expectation Maximization algorithm. Then, a relaxed boundary on the $sigma$ parameter $\sigma^{thw}$ of the normal distribution estimation, which tend to be a big deal higher in images with black walls, is used to detect plans containing these kind of walls. Finally, documents with black-walls are transformed to their corresponding edge image using the Canny edge detector.

### C. Wall-segment candidates generation

Wall segment candidates of different widths are generated in this step according to the *wall-assumption 1* defined above. Firstly, the parallel lines at different image orientations $\alpha$ are detected by foreground runs of a certain minimum length $rl^b_{min}$. Then, the distances between each parallel candidates are calculated by the background runs in their orthogonal direction. The runs are quantized into a histogram $\text{hist}_{RL}$, where high frequencies state for repetitive runs among black lines, and thus, possible widths modeling walls. On the other hand, lower frequencies, which are the vast majority, are produced by more infrequent objects also modeled by parallel lines. The histogram is smoothed and the bins with the maximum frequencies according a predefined threshold are grouped into a set of adjacent runs. This is done to reduce the noise dependency when walls in the same floor plan have a slightly breath difference. In the end, a segmentation image is generated for each one of the groups of widths by retrieving the foreground lines involved. They are considered as segments which possibly belong to walls, or part of them; from now on, called wall-candidates. The different steps implicated in this process are illustrated in Figure 3.

### D. Wall Segmentation Ranking

Wall-candidates are combined generating multiple wall segmentation hypothesis. The resulting hypothesis are ranked according to the properties involved in the *wall-assumptions*. The final segmentation adopted is the one with the highest score.

#### Wall-candidates combination

Multiple segmentation hypothesis are generated from the set of wall-candidates because generally, in floor plans, inner and outer walls have different widths. There are also some inner walls which usually are slightly wider than the rest, mainly because of structural purposes in the building architecture. Moreover, some walls are graphically modeled by more than two single parallel lines. Hence, is likely that more than one wall-candidate lead to the correct segmentation.

The $k$-combinations for the $n$ wall-candidates for all possible $k$ subsets, except for the empty set, are generated spreading into $2^n - 1$ final combination subsets. The final segmentation
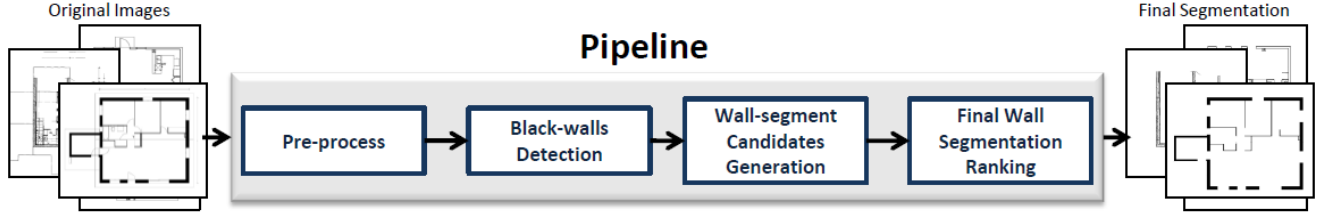
Fig. 2: The pipeline of our approach

hypothesis set $S$ is given by the logical disjunction function over the wall-candidates $w_i$ in every subset:

$$S = |\{w_1\}; ...\{w_n\}; ...\{w_1 \vee w_n\}; ...\{w_1 \vee ... \vee w_n\}|, \quad (1)$$

renamed as,

$$S = \{hyp_1, ..., hyp_h, ..., hyp_{2^n-1}\}, \quad (2)$$

being $hyp_h$ a final segmentation image hypothesis.

*Wall general attributes*

For each one of the final segmentation hypothesis in $S$, four different generalist attribute scores based on the *wall-assumptions* are extracted to determine their likelihood on being the correct solution:

- **SH** is the summation of frequencies in the histogram of runs for the widths involved in the segmentation hypothesis:

$$\mathrm{SH}_{hyp_h} = \sum_i \mathrm{hist}_{\mathrm{RL}}(w_i), \forall i | w_i \in hyp_h. \quad (3)$$

SH benefits those hypothesis formed by several wall-candidates –segmentations with multiple thickness–, which agrees with *wall-assumption 4*.

- **CC**: is the summation of the number of individual connected components in each wall-candidate involved in the segmentation hypothesis:

$$\mathrm{CC}_{hyp_h} = \sum_i \#\mathrm{CC}(w_i), \forall i | w_i \in hyp_h, \quad (4)$$

where $\#\mathrm{CC}(w_i)$ is the number of isolated connected components in the wall-candidate $w_i$. The more number of CC's, the higher is the score. This attribute score avails segmentations with multiple components, which agrees with *wall-assumption 6* when mentioning that walls should appear *repetitively*.

- **AR** states for the mean longness aspect ratio (longitude / width) of the CC in each of the wall-candidates:

$$\mathrm{AR}_{hyp_h} = \overline{\mathrm{long}(\mathrm{CC}_j(w_i))/\mathrm{width}(\mathrm{CC}_j(w_i))}, \quad (5)$$
$$\forall j | CC_j \in w_i, \forall i | w_i \in hyp_h$$

According *wall-assumption 2*, walls are longer than wider, and then, longer aspect ratios are enhanced in the final segmentation.

- **DiffD** accounts on the difference of black pixel distribution between the original image with respect to each

segmentation hypothesis in the different equalized rectangular regions $r$ they are split. The proportional difference is calculated as:

$$\mathrm{DiffD}_{hyp_h} = \sum_{n=1}^{r} \sum_{m=1}^{r} p_{nm} - p_{nm}^h, \quad (6)$$

where $p_{mn}$ and $p_{mn}^h$ are the percentage of the black pixels in the $mn^{th}$ region of the original image and $hyp_h$ respectively. DiffD enforces segmentations distributed similarly to the input image throughout the plan, agreeing with *wall-assumption 6* in terms of walls location, and allows to filter dispersedly located elements.

*Final Wall Segmentation*

The global scoring function is given by the aggregated summation of the different normalized attribute scores:

$$W(hyp_h) = \mathrm{SH}_{hyp_h} + \mathrm{CC}_{hyp_h} + \mathrm{AR}_{hyp_h} + \mathrm{DiffD}_{hyp_h}. \quad (7)$$

The final wall segmentation adopted is that hypothesis with the highest score.

## III. EXPERIMENTS

In this section we firstly present the datasets used to evaluate our approach. Secondly, we briefly describe the evaluation protocol used to rate the segmentation. And finally, we analyze both, the quantitative and qualitative results.

### A. Datasets

Two reference datasets used to evaluate recent floor plan analysis systems [6], [7], [8], [9], are adopted to test and compare our method. These datasets, named as *Dataset Black* and *Dataset Textured* contain respectively 90 and 10 real architectural images with completely different graphical notations and resolutions. Moreover, we have collected and ground-truthed two new datasets with different notations for walls. These datasets are called Textured2 and Parallel:

- **Textured2** contains 18 real floor plan images of 5671×7383 pixels. They contain text, legends, stair-side-views and symbols of different domains, such as electrical, furniture, etc. The texture for walls (see Fig. 1c) is composed by hatched lines with a big deal higher frequency and opposite direction than images in Dataset2.

- The 4 real floor plans of **Parallel** are 2550×3300 pixels. Walls are modeled by parallel lines (see Fig.
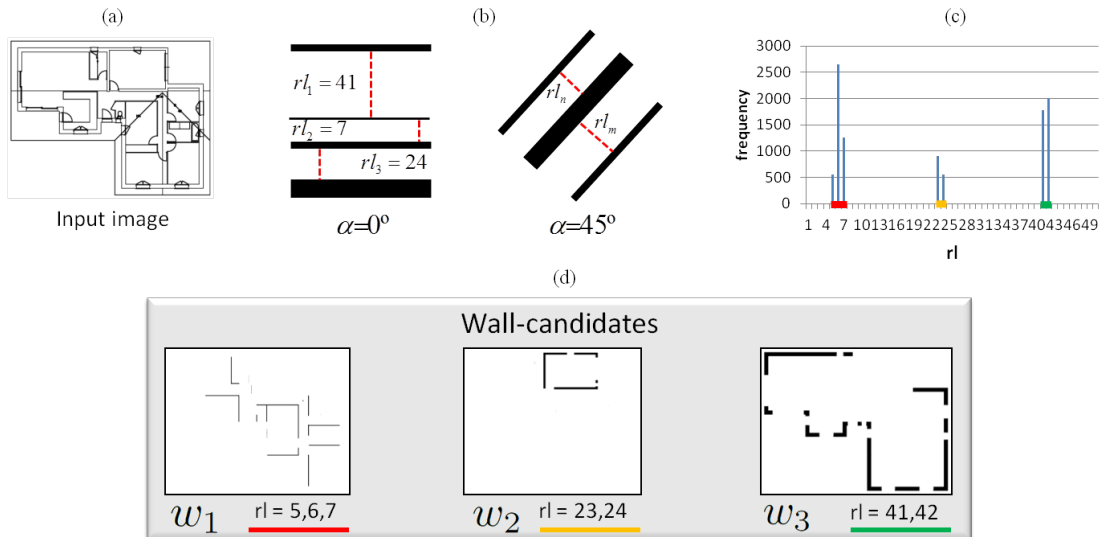
Fig. 3: Wall-candidates generation. In (a), the input image is shown. The extraction of the runlengths for different orientations $\alpha$ is zoomed in (b). These runs are quantized in the histogram $hist_{RL}$ shown in (c) and grouped into three colored clusters. Each cluster represents a common parallel line thickness in the input image, generating a wall-image-candidate as it is shown in (d).

1d), either for interior and exterior. The images contain text, text-tables and furniture.

### B. Evaluation

The evaluation protocol adopted to evaluate our method is the same as in [9]. The evaluation is at pixel level only on the foreground pixels of the original image. The results are expressed using Jaccard Index (JI):

$$JI = \frac{TruePos}{TruePos + FalsePos + FalseNeg} \quad (8)$$

In addition to that, since this method is thought as an initial step for a complete floor plan interpretation system, Recall is also taken into account; it is more straightforward and effortless to post-process an over-segmented result, than finding some lost walls in later processes of a global floor plan analysis system.

### C. Results

Our method is influenced by four parameters ($rl_{min}^b$, $\alpha$, $\sigma^{tw}$ and $r$) set experimentally in a very relaxed way for the multiple plans tested. The parameter $rl_{min}^b$ states for the minimum run length in the black horizontal line generation for being considered as a possible line. $rl_{min}^b$ is set to 10 pixels, which is sufficiently small to cope with low resolution documents, and adequately high for efficiency issues. The angle interval $\alpha$ specifies in which rotation of the input image lines can be detected. It has a strong impact when diagonal walls occur in the image. Yet, the lower $\alpha$, the more image-lines are generated and thus, the slower is the global performance. Experimentally, we set $\alpha$ increment in $15°$, which is a good trade of between performance and speed. The sensitivity boundary over the estimated $\sigma^{thw}$ is used to detect plans with thick-walls. The results obtained for the 4 different datasets have demonstrated that $\sigma^{thw}$ values for plans with thick-walls are, at least, 75

times higher than in plans without this kind of walls. Therefore, in a very relaxed way, we decided that floor plans with $\sigma^{thw}$ estimation values over 25, are classified as documents containing black-walls. The last parameter to be set is the number of equal-size regions $r$ used to calculate the black pixels distribution difference DiffD. Experimental tests have shown that the performance for $r = \{9, 16, 25\}$ varies at most 0.02 in terms of JI. For other close values to them, the rates drop significantly. $r = 9$ is adopted since is the configuration with the best global performance.

Table I shows the quantitative results compared with the notation oriented strategy used [7], the notation invariant patch-based detector [9] and our unsupervised notation invariant approach. It shows the average JI score and the recall obtained particularly for each dataset, and globally for all of them. In the case of [7], exclusively results for *Dataset Black* are presented since this approach is specifically oriented to extract black-walls and then, useless in floor plans with different notations. Finally, it is also worth to point out that we can only compare with [9] on the labeled datasets we have created, which contain multiple images using similar notations on purpose, as this method requires the existence of a ground-truth for training purposes.

Quantitatively, our method performs more effectively than [7] on *Dataset Black* in both JI and recall, bearing in mind that this approach was specifically thought for this dataset. Comparing to [9], our average JI score is modestly worse; yet, only 3 points below taking into account that [9] is a supervised method. On the other hand, our approach performs slightly better in global recall terms, remarking that, conversely to [9] the local recall scores in every dataset are always higher than 0.9.

In addition to that, we show in Figure 4 some qualitative results on floor plans extracted from the Internet[1]. We have

---

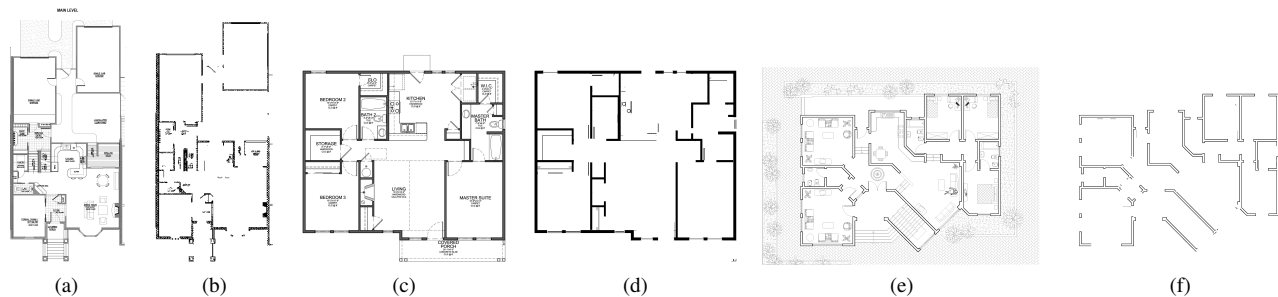[1]https://www.google.es/imghp?q=floor%20plan

Fig. 4: Qualitative results for images extracted from the Internet. In (b), (d) and (f), the segmented walls are shown from their corresponding original images (a), (c) and (e).

TABLE I: Wall segmentation results

| | #images | [7] | | [9] | | our new approach | |
|---|---|---|---|---|---|---|---|
| | | JI | Rec. | JI | Rec. | JI | Rec. |
| Dt. Black | 90 | 0.90 | 0.92 | 0.97 | 0.99 | 0.93 | 0.97 |
| Dt. Textured | 10 | – | – | 0.83 | 0.98 | 0.82 | 0.97 |
| Dt. Textured2 | 18 | – | – | 0.81 | 1 | 0.77 | 0.91 |
| Dt. Parallel | 4 | – | – | 0.70 | 0.84 | 0.66 | 0.98 |
| **Average per Dataset** | | – | – | **0.83** | **0.95** | **0.80** | **0.96** |

selected challenging images with different notations, resolutions, containing diagonal walls, repetitive textures, such as terraces or parquet floors, text, stairs, etc. The qualitative results confirm its good performance and adaptability in wall segmentation on floor plan images of any graphical notation and resolution. We visually confirm that very few walls are missed, as it is ratified by its recall score in Table I. This fact makes our approach very attractive when wall detection is a little step in a global floor plan analysis system; when the lose of a wall can imply the misinterpretation of part of the floor plan.

## IV. CONCLUSIONS

We have presented a floor plan wall detector able to deal with any kind of floor plan notation. Contrarily to the state of the art, it automatically adapts to any new graphical notation without the need of annotated data to learn the graphical appearance of walls. The results obtained on 4 real floor plan datasets demonstrate that outperforms recent strategies in recall terms; an important point to consider when walls are used to posteriorly detect other structural elements in floor plan analysis systems. In addition to that, we have presented some qualitative results on floor plans extracted from the Internet to highlight its robustness for any new input document.

We are working in a complete floor plan analysis system which incorporates this wall segmentation strategy. Our intention is to create a global system that is able to interpret floor plans independently from their graphical notation and without the need of labeled data to learn each graphical notation.

## ACKNOWLEDGEMENT

## REFERENCES

[1] Y. Aoki, A. Shio, H. Arai, and K. Odaka, "A prototype system for interpreting hand-sketched floor plans," in *Proceedings of the 13th International Conference on Pattern Recognition, 1996.*, vol. 3, 1996, pp. 747–751.

[2] P. Dosch, K. Tombre, C. Ah-Soon, and G. Masini, "A complete system for the analysis of architectural drawings," *International Journal on Document Analysis and Recognition*, vol. 3, pp. 102–116, 2000.

[3] T. Lu, H. Yang, R. Yang, and S. Cai, "Automatic analysis and integration of architectural drawings," *International Journal on Document Analysis and Recognition*, vol. 9, pp. 31–47, 2007.

[4] J. Juchmes, P. Leclercq, and S. Azar, "A multi-agent system for the interpretation of architectural sketches," in *Eurographics Workshop on Sketch-Based Interfaces and Modeling*, 2004, pp. 53–61.

[5] S. Macé, H. Locteau, E. Valveny, and S. Tabbone, "A system to detect rooms in architectural floor plan images," in *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*, 2010, pp. 167–174.

[6] L.-P. de las Heras and G. Sanchez, "And-or graph grammar for architectural floorplan representation, learning and recognition. a semantic, structural and hierarchical model," in *Proceedings of the 5th Iberian Conference on Pattern Recognition and Image Analysis*, vol. 6669, 2011, pp. 17–24.

[7] S. Ahmed, M. Liwicki, M. Weber, and A. Dengel, "Improved automatic analysis of architectural floor plans," in *Proceedings of the 11th International Conference on Document Analysis and Recognition*, 2011.

[8] S. Ahmed, M. Liwicki, M. Weber, and A. Dengel., "Automatic room detection and room labeling from architectural floor plans," in *Proceedings of the 10th IAPR International Workshop on Document Analysis Systems (DAS-2012)*, 2012, pp. 339–343.

[9] L.-P. de las Heras, J. Mas, G. Sánchez, and E. Valveny, "Wall patch-based segmentation in architectural floorplans," in *Proceedings of the 11th International Conference on Document Analysis and Recognition*, 2011, pp. 1270–1274.

[10] K. Tombre, S. Tabbone, L. Pélissier, B. Lamiroy, and P. Dosch, "Text/graphics separation revisited," in *Document Analysis Systems V*, ser. Lecture Notes in Computer Science, 2002, vol. 2423, pp. 615–620.

[11] N. Ouwayed and A. Belaid, "A general approach for multi-oriented text line extraction of handwritten document," *International Journal on Document Analysis and Recognition*, vol. 14, no. 4, Sep. 2011.