

Using a neighbourhood graph based on Voronoï tessellation with DMOS, a generic method for structured document recognition

Aurélie Lemaitre, Bertrand B. Coüasnon, Ivan Leplumey

▶ **To cite this version:**

Aurélie Lemaitre, Bertrand B. Coüasnon, Ivan Leplumey. Using a neighbourhood graph based on Voronoï tessellation with DMOS, a generic method for structured document recognition. Sixth IAPR International Workshop on Graphics Recognition, 2005, Hong Kong SAR China. pp.260-271. hal-00300322

HAL Id: hal-00300322
https://hal.science/hal-00300322

Submitted on 2 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Using a neighbourhood graph based on Voronoï tessellation with DMOS, a generic method for structured document recognition

Aurélie Lemaitre*, Bertrand Coüasnon*, and Ivan Leplumey**

* *IRISA / INRIA,* ** *IRISA / INSA*
Campus universitaire de Beaulieu, F-35042 Rennes Cedex, France
E-mail:couasnon@irisa.fr

Abstract

To develop a method for structured document recognition, it is necessary to know the relative position of the graphical elements in a document. In order to deal with this notion, we build a neighbourhood graph based on Voronoï tessellation. We propose to combine the use of this interesting notion of neighbourhood with an existing generic document recognition method, DMOS, which has been used to describe various kinds of documents. This association allows exploiting different aspects of the neighbourhood graph, separating the graph analysis from the knowledge linked to a kind of document.

Thus, we set up a cooperation between neighbourhood graph, that contains numerical information, and DMOS method, that can convert this data into symbolic information. Moreover, DMOS method can specify a context of analysis when exploiting neighbourhood graph, which makes it possible to establish a bi-directional relation between the analyser and the graph. We apply this method on the analysis of various documents.

Keywords: Voronoï diagram, neighbourhood graph, structured document recognition, generic recognition system.

1 Introduction

In the field of structured document recognition, the knowledge on relative position between the graphical elements of a document is often necessary. Voronoï tessellation of image, and the dual Delaunay graph, provide an interesting description of this concept of neighbourhood. This method is used in several papers for structure recognition of document images, in the context of specific applications: detection of lines, words, segments. We propose to exploit such information in a generic context, using an existing document recognition method, DMOS.

Indeed, in the standard version of DMOS method, the relative position of elements is given with an approximation of components to their bounding box, which is not always precise enough. That is why we have proposed to introduce neighbourhood graph based on Voronoï tessellation, which offers a precise notion of relative position. Furthermore, using the graph with DMOS makes it possible to extract information depending on the context. Indeed, we set up a cooperation between neighbourhood graph, that contains numerical information, and DMOS method, that can convert this data into symbolic information.

In a first part, we will present relative work on Voronoï diagram in the field of structured document recognition, and the associated neighbourhood graph that we have implemented. Then, we will present DMOS method and the integration of neighbourhood that has been realized. We will expose afterwards applications that have been set up in order to validate these tools. We will end by a discussion.

2 Related work on area Voronoï diagram

First of all, we present a few definitions of Voronoï Diagram, and we describe the usual applications that are made on document recognition.

2.1 Definitions

We present two definitions of Voronoï Diagram. More details are given in [7].

2.1.1 Classical Voronoï diagram

The classical Voronoï diagram cuts up the area into influence regions of points. Let P =$\{p_1, ..., p_n\}$ be a set of points from the plan, called *generators*, and $d(p, q)$ be the Euclidean distance between points *p* and *q*. Then, the *Voronoï region* of a point p_i is given by

$$V(p_i) = \{p | d(p, p_i) \leq d(p, p_j), \forall j \neq i\}$$

It is the set of points that is nearest to this generator than any other.

The set of the boundaries of the Voronoï region is called $\delta V(p_i)$.

The ordinary Voronoï diagram is given by the set of Voronoï region:

$$V(P) = \{V(p_i), ..., V(p_n)\}$$

We usually associate to Voronoï diagram the dual Delaunay graph that is composed of the same set of vertex P, and which contains a edge between points p and q if they are neighbours in the Voronoï diagram, that is to say if they have a common boundary:

$$\delta V(p) \cap \delta V(q) \neq 0$$

2.1.2 Area Voronoï diagram

The basic Voronoï diagram has been generalized in several directions. One of the possible generalizations consists in replacing the set of points, the *generators*, by a set of connected components. A connected component is a set of black pixels that are in contact. We present an example figure 1(c). Such a diagram is called *area Voronoï diagram*.

2.2 Using Voronoï diagram for structured document recognition

The area Voronoï diagram has been used a lot for structure detection of images. Indeed, the Voronoï diagram enables to obtain for each component of an image the list of his nearest neighbours.

Thanks to this information, some authors have proposed various methods to segment documents into words, lines, paragraphs, columns. Generally, like in [4] or [6], the knowledge that is necessary for the analysis is included by learning thresholds like inter-character, inter-word and inter-textline gaps.

However, these thresholds are learnt statistically on the whole document. Thus, all the knowledge that is introduced is relative to the global image, and it is not dissociated from the exploitation of the Voronoï tessellation. Consequently, the analysis is limited to quite homogeneous documents, and the knowledge is reduced and appropriated just for one kind of document.

In order to extract more information from Voronoï diagram, we propose to separate the neighbourhood graph analysis from the necessary knowledge. That is why we use the neighbourhood graph with a generic method for structured document recognition: DMOS. The first part of this work has consisted in implementing a neighbourhood graph based on Voronoï tessellation.

3 Implementation of a neighbourhood graph based on discrete distance

We have chosen to implement a neighbourhood graph which implementation has been detailed in [5]. The aim is to build a neighbourhood graph based on an area Voronoï diagram and labelled with discrete distances.

The principle is to apply chamfer distance by propagation on the initial image. We obtain, as a result, three new pictures that contain respectively, for each pixel:

- the discrete distance to the nearest connected component (figure 1(b));
- the name of the nearest connected component (figure 1(c)). Indeed, this image is the approximated Voronoï area diagram;
- the coordinates of the nearest point of the nearest component.



(a) Image (b) Distance image (c) Area Voronoï diagram (d) Neighbourhood graph

Figure 1: Construction of neighbourhood graph

Thanks to these three images, we can build a neighbourhood graph (figure 1(d)), labelled with distances, more complete than a mere Voronoï tessellation. This graph can be exploited for document analysis with method DMOS.

4 Integration of neighbourhood graph in a generic method for structured document recognition

4.1 DMOS method

We presented in various papers ([2], [1]) DMOS (Description and Modification of Segmentation), a generic method for structured documents recognition. This method is made of:

- the grammatical formalism EPF (Enhanced Position Formalism), which can be seen as a description language for structured documents. EPF makes possible at the same time a graphical, a syntactic and even a semantic description of a class of documents;
- the associated parser which is able to change the parsed structure during the parsing. This allows the system to try other segmentations with the help of context to improve recognition.

We have implemented this DMOS method to build an automatic generator of structured document recognition systems. By only changing the EPF grammar, and when needed by training a classifier, we produced automatically various recognition systems: one on musical scores, one on mathematical formulae, one for recursive table structures or one for archive documents.

4.2 EPF formalism

The EPF formalism can be seen as a grammatical language to describe a structured class of documents. It is based on several operators that make possible a two-dimension document description. Thanks to those operators we can specify the position of each element relatively to the others, and the description of a document is based on this relative position. We first introduce the main operators on a simple example and then explain the way they are used to analyse the document.

4.2.1 Example on a simple grammar

The simplified grammar presented here describes a mathematical formula based on an integral, like on figure 1(a).

Intuitively, we can describe such a document by:

- an integral symbol, on the left part of the picture;
- integration bounds, on the top and bottom part of the integral;
- an expression, on the right of the integral.

The grammar rules will follow this intuitive description. In order to translate the way the elements are positioned in the document, we use two position operators:

- AT gives the position of an element relatively to a previous one;
- AT_ABS gives an absolute position of the element.

Then, the simplified main rule is:

```
integralFormula::=
      AT_ABS(leftPicture) && integralSymbol && (
            AT(topRight integralSymbol) && bound ##
            AT(bottomRight integralSymbol) && bound ##
            AT(right integralSymbol) && expression).
```

The concatenation operator in the grammar is &&. The operator ## is the factorization operator, it means here that each of the three last lines is relative to the first one. The rule `expression` is a non-terminal that is not detailed here. The rules `integralSymbol` and `bound` consist in extracting terminals. This is expressed thanks to the operator `TERM_CMP` that makes it possible to extract a connected component, if necessary fulfilling a certain condition. We present here the simplified rule for the detection of a bound:

```
bound::= TERM_CMP noCond character.
```

No condition is required here (`noCond`) for the detection of a bound; we could specify here a condition about the kind or the size of the component.

4.2.2 Mechanism of neighbourhood

We have seen on the previous part that the common use of a position operator AT and a component detection operator TERM_CMP, was necessary to detect an element. We present here the mechanism of neighbourhood that is used by the analyser to extract a component.

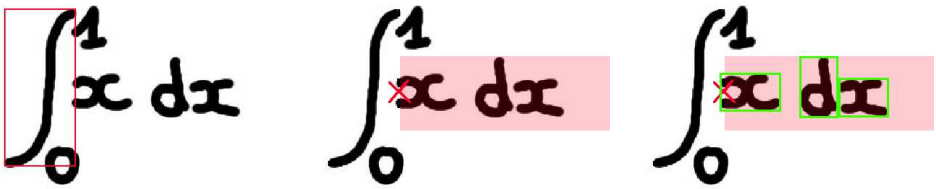
First, the operator AT makes it possible to choose a reference position (a point) and a research zone, depending on the last component found. Then, the operator TERM_CMP extracts a component:

- in the research zone;
- the one which bounding box is the nearest from the reference position;
- fulfilling the condition.

The example on figure shows the application of the rules:

```
AT(right integralSymbol) && TERM_CMP noCond character.
```

The `integralSymbol` has just been recognized, represented by his bounding box on figure 2(a). The operator AT sets the reference position and the research zone corresponding to `right` of the `integralSymbol` bounding box (figure 2(b)). Then, the instruction TERM_CMP makes the analyser look for the nearest component in the research zone, using the distance between reference point and bounding boxes (figure 2(c)).



(a) Bounding box of integral

(b) Definition of the reference position and the research zone

(c) Detection of components depending on their bounding box

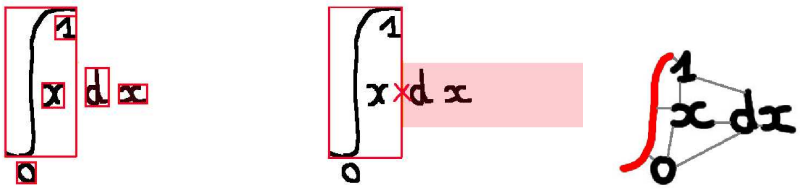
Figure 2: Detection of the component *x* from the *integralSymbol*

4.2.3 Limits of this version

The use of the mechanism of neighbourhood presented in the previous paragraph is not always appropriate. Indeed, to build this neighbourhood, the elements are compared to their bounding box, which is quite vague in certain cases, and most particularly in handwritten documents.

Let us take the example of the previous grammar with another document, presented figure 3(a). With the same detection mechanism, the research zone and the reference pointer are set like on figure 3(b), and the nearest component that will be detected is the *d* instead of the x. This is due to overlapping bounding boxes of components.

That is why we proposed to include new operators for this grammar, based on the use of the neighbourhood graph that has been presented on section 3. In this graph, the relative position of two graphical components is then given by the existence of an edge in the graph and the associated distance, which is more precise than a relative position of bounding boxes.



(a) Bounding boxes

(b) Definition of the reference position and the research zone

(c) Memorization of reference component

Figure 3: Detection of the component *x* from the *integralSymbol*

4.3 Integration of neighbourhood graph in DMOS

The work has consisted in inserting, in the existing formalism, the data that could be extracted from the graph and that could be interesting, generally speaking, for structured document recognition.

We have shown previously that the component detection mechanism, based on the common use of the operators `AT` and `TERM_CMP`, was sometimes mistaken. That is why our first idea was to propose another component detection mechanism based on neighbourhood graph. In order to complete the expression of component detections, we have introduced new basic conditions based on the graph.

Then, we wanted to use the fact that method DMOS makes it possible to analyse documents locally, to propose new information, based on the graph, depending on the context. Consequently, we have set up a bi-directional communication between the analyser and the neighbourhood graph. Indeed, the graph transmits numerical data to DMOS that translates it into symbolic data. Besides, DMOS sends information to the graph by specifying the context of analysis. This is one of the main interests of the method.

We will detail now the different parts of our work.

4.3.1 New component detection mechanism

The first part consist in replacing the mechanism of component detection, presented on paragraph 4.2.2, by a new one, based on a neighbourhood graph. Thanks to the graph, we have set up a new component extraction operator and conditions for recognition.

New operator We introduce a new operator, `TERM_CMP_GRAPH`, that can be used in the same conditions as `TERM_CMP`, but which mechanism is based on the neighbourhood graph.

The grammar rule presented section 4.2.1 becomes:

```
AT(right integralSymbol) && TERM_CMP_GRAPH noCond character.
```

When executing the instructions, the analyser leans on neighbourhood graph. Indeed, the operator `AT` set the research zone, like previously. However, the reference position is a component instead of a point in the previous version. On figure 3(c), the `integralSymbol` element is memorized as reference component.

With the `TERM_CMP_GRAPH` operator, we can detect an element:

- in the research zone;
- the nearest in the neighbourhood graph to the reference component;
- respecting the conditions.

The introduction of this operator `TERM_CMP_GRAPH` has been made similarly as the existent `TERM_CMP`, which makes it possible to use indifferently one of the two operators in the description of a kind of document.

New conditions for the detection Thanks to the neighbourhood graph, we have introduced new conditions for searching elements. We present here a few ones.

Edge between components When detecting a component among several in a research zone, we can reduce the choice to the elements that are linked with the reference component in the neighbourhood graph. The syntax is:

```
condExistDirectLink ReferenceComponent ComponentToAnalyse
```

Edge direction These conditions make it possible to choose the direction of the link between components. Indeed, it can be useful to specify the relative position of two components. That is why we propose, for example, the condition:

```
condOnTheRightGraph ReferenceComponent ComponentToAnalyse
```

In order to check this condition, we extract the path between the components `ReferenceComponent` and `ComponentToAnalyse`. The condition succeeds if the extremity of the path is situated on the right of the origin. The same condition exists with other directions.

Edge distance This condition makes it possible to limit the distance between two components. Indeed the condition

```
condMaxDistance ReferenceComponent MaxDistance ComponentToAnalyse
```

succeeds if distance between `ReferenceComponent` and `ComponentToAnalyse` is inferior to `MaxDistance`. This condition makes it possible to detect elements only if they are close enough.

These are just examples of the most common used conditions. However, the user can develop specific ones when necessary.

4.3.2 Extraction of local statistical information

Using Voronoï neighbourhood with DMOS makes it possible to extract local information from the graph. Indeed, in an area determined by the context, we can examine only the corresponding part of the neighbourhood graph. Thus, we propose an extractor of statistics about distances in the neighbourhood graph. The syntax of this extractor is:

```
calculateStatDistGraph Area FavoriteDirection RequiredStatistic
```

In the selected zone `Area`, we extract distances of a chosen set of edges, depending on `FavoriteDirection`: every edge included in the zone, only the vertical or horizontal ones. Then, we calculate the chosen statistic `RequiredStatistic` that can be average, median or threshold for separating data into classes.

Thanks to this operator, we can compensate for the problem presented on section 2.2. Indeed, the fact that DMOS method knows the analysis context makes is possible to extract from the graph only the appropriate information. Consequently, Voronoï tessellation can be exploited depending on the context and not necessarily on homogeneous document.

We give in part 5.2 an example of application of such statistics.

5 Validation of neighbourhood graph integration

The different tools based on neighbourhood graph, integrated with method DMOS, have been applied for the description of various kinds of documents. The aim was to prove their genericity and to determine the cases the neighbourhood graph could be useful.

We have first shown the interest of the basic operators. Then, we have experimented the statistic tools on the detection of words in papers. We have finally studied a definition of a grammar using both bounding boxes and neighbourhood graph on handwritten register of the 19th century.

5.1 Interest of basic operators

Applications on several documents, that will be presented later, have made it possible to show the interest of the new component detection mechanism. The main improvement is the possible detection of overlapping bounding box components with a simple expression of the grammar rule.

We have seen, like on figure 3(b), that the standard mechanism of component detection was not dealing with overlapping bounding boxes. Thus, when the user wanted to take this into account, he had to specify his grammar so that overlapping could be detected, which was not intuitive and required the use of many position operators and conditions for detecting components.

On the opposite, writing a grammar rule based on operator `TERM_CMP_GRAPH` makes it possible to describe intuitively a document; the user does not have to worry about overlapping bounding boxes.

For example, let us present the grammar rule for the description of an expression. It is composed of a succession of letters, from left to right.

```
expression::= letter && AT(right letter) && expression.
expression.  %stop case
```

The standard definition of a letter is given by:

```
letter::= TERM_CMP noCond element.
```

In that case, this will not be able to detect elements that are overlapping horizontally. Each letter will not be extracted. However, with the definition:

```
letter::= TERM_CMP_GRAPH noCond element.
```

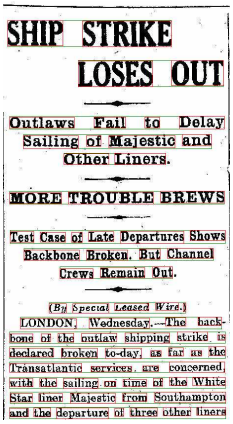
each element will be successively detected.

Consequently, thanks to the neighbourhood graph, on the one hand the result is more precise for equivalent grammars and on the other hand, for the same precision in the result, the grammar will be easier to write because it needn't deal with overlapping boxes.

5.2 Validation of local statistic extraction

Voronoï diagram has been used a lot for structured document recognition. Generally, like in [4] or [6], the knowledge that is necessary for the analysis is included by learning thresholds like inter-character, inter-word and inter-textline gaps. However, this statistic analysis is made on the whole document, which requires the document to be homogeneous enough.

Thanks to DMOS, we propose to extract thresholds depending on the context, using the statistic tool presented previously. We have chosen to apply this on the detection of words in columns extracted from the *International Herald Tribune*. We can see on the example presented figure 4 the interest of a local analysis, because of the large variation in police size.



SHIP STRIKE LOSES OUT

Outlaws Fail to Delay Sailing of Majestic and Other Liners.

MORE TROUBLE BREWS

Test Case of Late Departures Shows Backbone Broken, But Channel Crews Remain Out.

(By Special Leased Wire.)

LONDON, Wednesday.—The backbone of the outlaw shipping strike is declared broken to-day, as far as the Transatlantic services are concerned, with the sailing on time of the White Star liner Majestic from Southampton and the departure of three other liners

Figure 4: Example of paper

5.2.1 Recognition principle

We have chosen to calculate the threshold line after line. The work is split up into three parts:

- First, the approximate area `Area` that contains the line is detected.
- The threshold that distinguish inter-words and inter-letters gap is extracted, thanks to the operator presented on section 4.3.2. We study each edge contained in the area and we ask for a threshold separating 2 classes thanks to the k-average method:

  ```
  calculateStatDistGraph Area EveryLink KAverage
  ```

- The words are extracted, thanks to the application of the rules below and taking the calculated threshold into account.

We consider that a word is a succession of letters, from left to right, linked in the neighbourhood graph, and spaced by a distance inferior to the threshold. The associated rules are:

```
word Threshold::=
     firstLetter MyLetter &&
     AT(rigthLine MyLetter) &&
     endOfWord Threshold MyLetter.
endOfWord Threshold LastLetter::=
     letterOfWord Threshold LastLetter ThisLetter &&
     AT(rightLine ThisLetter) &&
     endOfWord Threshold ThisLetter.
endOfWord _ _. %Stop case
```

The recurrence ends when the next letter is too far from the previous one or when there is no more letters on the line. The detection of terminals is given by the rules below:

```
firstLetter MyLetter::= TERM_CMP_GRAPH noCond MyLetter.
letterOfWord Threshold LastLetter ThisLetter::=
     TERM_CMP_GRAPH
          [(condExistDirectLink LastLetter),
           (condMaxDistance LastLetter Threshold)],
          ThisLetter.
```

Thanks to those grammar rules, we can apply our work on neighbourhood graph for the local detection of words.

5.2.2 Results

We applied this grammar on columns of papers extracted from *International Herald Tribune* of years 1900, 1925 and 1950; those documents had been proposed for a contest at ICDAR 2001. An example is given on picture 4. Our base was composed of 2588 words to recognize; we managed to detect 98.53% of them.

Using local threshold is very useful because our documents are not homogeneous. However, it has got limits. Indeed, the local analysis zone must not be too small, so that the threshold keeps correct. In our application, we have chosen to extract a threshold for each line. When the line is composed of just one word, each letter is considered as a word. This is the only reason for the remaining 2.47% undetected words.

In order to solve this problem, we could extend the local application to the whole paragraph with the same character size. This could be done easily thanks to DMOS method.

5.3 Using both Voronoï and bounding box neighbourhoods

The implementation of operators linked with neighbourhood graph makes it possible to use in DMOS either the bounding box based distances or the Voronoï tessellation neighbourhood. Right now, when defining a new grammar, the user keeps the possibility to use either the bounding box neighbourhood or the Voronoï based graph.

In order to give an example of choice that could be done, we have worked on the description of handwritten registers of naturalization decree from the end of the 19th century. A previous grammar, defined with EPF, made it possible to extract the columns, the registration numbers and the names from such a document (example on picture 5). The way to analyse this kind of documents has been presented in [3]. When necessary, we have modified this grammar in order to introduce neighbourhood graph. We will detail each part of the recognition.

5.3.1 Detection of alignments

The first part of the work consists in detecting the position of the margin in the document. Indeed, the numbers are in the margin area whereas the names are fronting the numbers, in the body of the text. Consequently, the first task of the analysis is to detect vertical alignments of characters, in order to find the column inside of which the number will be researched.

The detection of alignments is based on a global research. Yet, the neighbourhood graph is more useful for local analysis, when precise positioning between two components is required. That's why, for this global phase of the detection, we have chosen to keep using the bounding box distances.

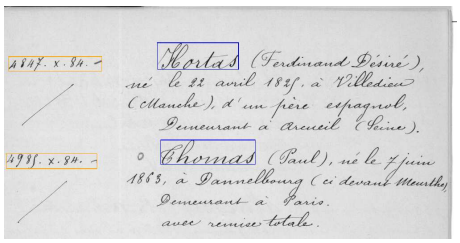


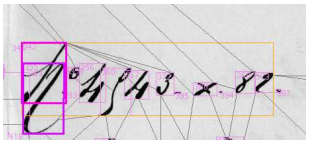
Figure 5: Example of handwritten register of naturalization decree

5.3.2 Detection of numbers

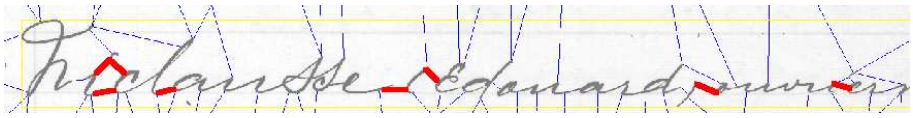
The detection of numbers is restricted in the margin that has been detected previously. It consists in finding one component that belongs to the number, and then detecting at least three other aligned components on the right or on the left. The notion of alignment is based on bounding boxes. Then, the global position of the number is built thanks to the extremities of the bounding box. However, each component is not detected, specially when the bounding boxes are overlapping.

We propose to use a grammar based on neighbourhood graph to describe a number. Indeed, it seems natural to describe a number as a succession of characters, linked with an edge in the graph. This makes it possible to detect each component contained in the number and to overcome the difficulty of bounding boxes.

For example, on picture 6(b), the bounding boxes of the right parts of the *N* are totally overlapping horizontally. With the standard version of the grammar, the second part of the *N* was not detected. With the new one, based on neighbourhood graph, we can detect each component.



(a) Example of overlapping bounding boxes



(b) Example of irregular text line

Figure 6: Limits of handwritten documents

5.3.3 Detection of surnames

In order to extract surnames from naturalization decrees, the previous grammar was detecting a global line of text, made of at least five aligned components. Then, the surname was supposed to be contained in the first half of the text line. Consequently, the area given as surname was the first half of the line. Nevertheless, this is approximate, because when the surname is very short, the analysis returns a lot of noise, whereas a long surname will be cut.

In order to improve this detection, we propose to introduce the grammar for detecting words that has been presented section 5.2.1. Indeed, we can extract statistic information from the approximate zone of the line that has been detected. Then, we consider that the surname is compound of the two first words.

This method should be better for detecting the surname. However, handwritten documents are not very suitable for statistic studies. Indeed, the thresholds for words detection, that were easy to extract on printed documents, are not so significant for irregular handwritten text. For example, on figure 6(a), we have highlighted edges that had a similar size. Some of them are intra-words distances whereas some are inter-words.

The irregularities of hand writing increase the difficulty of word detection.

Nevertheless, the combined use of neighbourhoods in DMOS makes it possible to extract information from both graph and bounding boxes. Indeed, bounding boxes are more interesting for a global detection of the document whereas using neighbourhood graph is necessary to detect components precisely.

5.3.4 Results

This grammar has been applied on 1130 naturalization decree pages from the end of the 19th century. The global recognition rate is similar to the one corresponding to the previous grammar, that is to say around 98%. Indeed, what is important for this application is that the number and surname are contained in the final extracted zone. With our new method, these areas have not been significantly modified. However, the elements that are detected inside the areas of numbers and names are more precise, because we can detect each component of the number and only the first words for the name.

6 Discussion

6.1 A contextual utilization of Voronoï diagram

The main particularity of the exploitation of Voronoï diagram is that the data contained in the graph is separated from the knowledge. This gives mainly one advantage: the grammar-based description of the kind of document makes it possible to exploit data contained in the graph according to the context of analysis. It means that, depending on the circumstances of the analysis, the user can, on one hand, choose which information should be extracted from the graph, and on the other hand, determine how this characteristic should be interpreted into symbolic information.

We have set up a cooperation between Voronoï diagram, which contains numerical and statistic information, and the analyser that knows the context of recognition and which is able to extract symbolic information from neighbourhood graph. This makes it possible to extract more information from Voronoï diagram than in classical applications.

6.2 Interests for the formalism

Compared with the standard version of DMOS, the introduction of a neighbourhood graph brings new faculties of expression of the knowledge. The operator `TERM_CMP_GRAPH` compensates for the imprecision of bounding box based distance, giving precise information about the existence of a neighbourhood between two components, and their distance.

The tools that have been implemented respect the genericity of the method. This has been shown by the definition of grammars for three kinds of documents.

When defining a new grammar, the user keeps the possibility to use either the bounding box neighbourhood or the Voronoï based graph. Information that are contained in neighbourhood graph are relative to a local vision of the document: a relation between two components. That is why their exploitation is really convenient to detect close components, when a precise relative position is required. This is particularly adapted for handwritten document analysis, liable to overlapping bounding box problems.

In return, in a global study of the document, like extracting columns for example, the neighbourhood graph doesn't seem to bring, for the present, pertinent information. The bounding box neighbourhood still seems more relevant in that case.

6.3 Possible evolutions

In this version of our work, neighbourhood graph makes it possible to position only two components. It would be sometimes interesting to know the relative position of two groups of components. For example, once a line of text has been detected as a set of components, we could gather those elements in order to be able to position

a line relatively to another. This would require a hierarchical structure in order to make the graph evolve during the analysis. This evolution could bring new interesting information for a global document analysis.

7 Conclusion

This paper shows how we have extended the exploitation of a neighbourhood graph based on Voronoï tessellation, by separating the graph analysis from the expression of the necessary knowledge, thanks to the generic method DMOS.

The standard relative position mechanism used in DMOS, based on bounding boxes, was not precise enough in certain cases. That is why we have introduced a new mechanism based on Voronoï tessellation. We set up a cooperation between neighbourhood graph, that contains numerical information, and DMOS method that can convert data into symbolic information. This makes it possible to increase precision when detecting elements with method DMOS.

Furthermore, Voronoï tessellation is usually used globally on a document, which reduces its capacity of exploitation. Thanks to method DMOS, this graph can be exploited depending on the context. Consequently, the data extracted from the neighbourhood graph can be more complete and adapted to required information.

While implementing this integration, we have respected the genericity of the method. Consequently, we have been able to apply this work on two kinds of documents: printed newspapers and handwritten naturalization decrees. The neighbourhood graph could be used for the description of any kind of document. Thanks to the formalization with EPF, the knowledge is more complex, than in a dedicated system linked to a specific application.

References

[1] Bertrand Coüasnon. DMOS: A generic document recognition method to application to an automatic generator of musical scores, mathematical formulae and table structures recognition systems. International Conference on Document Analysis. (ICDAR'01), pages 215-220, 2001.

[2] Bertrand Coüasnon. Dealing with noise in DMOS, a generic method for structured document recognition: an example on a complete grammar. Graphics Recognition: Recent Advances and Perspectives, pages 38-49, 2004.

[3] Bertrand Coüasnon, Jean Camillerapp, and Ivan Leplumey. Making Handwritten Archives Documents accessible to Public with a Generic System of Document Image Analysis. International Workshop on Document Image Analysis for Libraries (DIAL'04), Pages 270-277, 2004.

[4] Koichi Kise, Motoi Iwata, and Keinosuke Matsumoto. On the application of Voronoï diagrams to page segmentation. Document Layout Interpretation and its Application (DLIA'99), 1999.

[5] Ivan Leplumey and Charles Queguiner. Un graphe de voisinage basé sur l'utilisation des distances discrètes. Conférence Internationale Francophone sur l'Ecrit et le Document (CIFED'2002), pages 41-50, 2002.

[6] Yue Lu, Zhe Wang, and Chew Lim Tan. Word grouping in document images based on Voronoï tessellation. Workshop on Document Analysis Systems (DAS'04), 2004.

[7] Atsuyuki Okabe, Barry Boots, Kokichi Sugihara, and Sung Nok Chiu. *Spatial Tessellations: Concepts and Applications of Voronoï Diagrams*. Wiley, 2000.