

Competitive hybrid exploration for off-line sketches structure recognition

Achraf Ghorbel
Irisa - INSA Rennes
Campus de Beaulieu
35042 Rennes Cedex
achraf.ghorbel@irisa.fr

Aurélie Lemaitre
Irisa - Université de Rennes 2
Campus de Beaulieu
35042 Rennes Cedex
aurelie.lemaitre@irisa.fr

Eric Anquetil
Irisa - INSA Rennes
Campus de Beaulieu
35042 Rennes Cedex
eric.anquetil@irisa.fr

Abstract

We work on new strategies of exploration for interpretation of off-line sketches. A first approach (call IMISketch) was based on a competitive breadth-first exploration of the analysis tree allowing to evaluate simultaneously several possible hypotheses of recognition in a dynamic local context of document. A great advantage of this strategy is to be able to solicit the user during the decision process to avoid error accumulation in the analysis step. IMISketch strategy is very interesting but it can lead combinatory problems when addressing complex sketches. In this paper, we propose a new hybrid strategy for exploration. The recognition process alternates between a breadth-first and depth-first exploration. The strategy is totally driven by the grammatical description of the document. The paper demonstrates the interest of this new hybrid strategy method on handwritten 2D architectural floor plans containing walls, opening and furnitures.

1. Introduction

IMISketch [5] is an interactive syntactic method for interpretation of structured documents. A priori structural knowledge of the document in IMISketch is expressed through a two-dimensional visual language based on production rules. This analyzer aims to interpret the image primitives. These primitives are line-segments and polygons. IMISketch adopts a breadth-first exploration unlike classical syntactic analysis methods [1] [2] [3] [8] based on depth-first exploration. Each branch of the tree is a possible hypothesis. The uncertainty is formalized by the attribution of scores to each hypothesis. If the ambiguities can not be resolved in an automatic manner, the user will be solicited by the analyzer to resolve the ambiguity. The user integration in the recognition loop avoids error ac-

cumulation and thus a fastidious a posteriori correction. The first results of this method performed on the architectural plans containing walls and opening are promising [5]. The presence of the user in the loop analysis requires a reasonable analysis time. For this, we proposed in [4] some optimizations to accelerate the interpretation.

When we analyze some architectural plans containing more complex furnitures, we note that the current IMISketch generates a lot of combinatory due to the presence of many basic primitives in a small zone.

In this paper, we focus on a new strategy of exploration to reduce the combinatory to obtain an analyzer that is really exploitable in the context of an interactive analysis. This strategy consists in passing from a method based only on breadth-first exploration to a hybrid method that alternates between breadth-first and depth-first exploration. At each step of analysis, the type of exploration will be induced by production rules of the two-dimensional grammar that will drive the analysis. The choice of exploration type is induced by the risk of the current interpretation relatively to its neighbours. An interpretation is called low-risk when the analyzer is sure to combine the right primitives that form an entity. Depth-first exploration is adopted in the low-risk interpretation cases. In other cases, the analyzer adopts a breadth-first exploration allowing to evaluate simultaneously several possible hypotheses of recognition in the current local context of the document. The complete system can be applied to off-line documents (images). We validate this new hybrid strategy on handwritten architectural plans. Some examples are illustrated in Figure 9.

The remaining of the paper is organized as follows. In the section 2, we introduce the principles of our existing interactive analysis method for interpreting sketches, IMISketch. The new hybrid strategy is described in section 3. Experimental results are reported in section 4 and finally, section 5 concludes the paper.

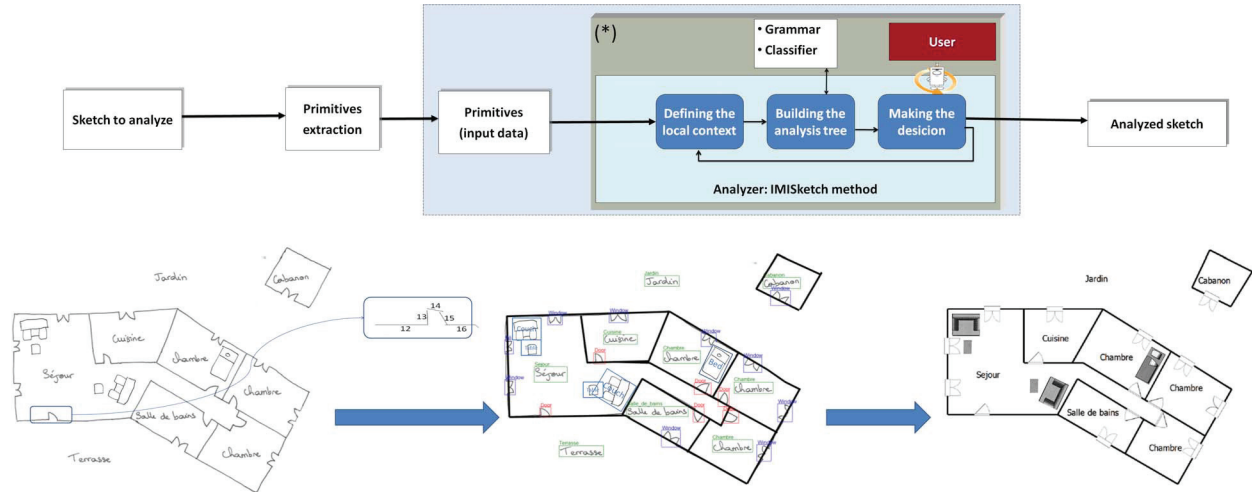


Figure 1. Synthetic scheme of IMISketch method for structured documents interpretation

2. Overview of IMISketch

In this section, we present the different steps of the existing method IMISketch (cf. Figure 1). The first step is the extraction of basic primitives required for analysis. The second main step is to analyze these primitives and to use their structural arrangement in the document to identify the symbols. This phase is interactive: the analyzer can solicit the user to resolve ambiguity cases. The recognition of the document structure is managed by a grammatical analyzer that drives the calls of symbol classifiers to evaluate a scoring for each hypotheses. We now detail each step of analysis.

2.1 Preprocessing phase: Primitive extraction

The first step consists of extracting the necessary information from the structured document. This phase is generic and off-line and does not depend on the type of document to interpret. We have chosen to work only with line-segments, which represents the primitives of our analysis. We adopt the Kalman filter to extract these primitives [6] presented in figure 2.

2.2 Interactive breadth-first exploration: IMISketch

IMISketch analyzer explores the matching between the document structure defined by grammatical rules and the set of primitives contained in the document. The grammatical rules are expressed using the context-driven constraint multiset grammars (CD-CMG) [7]. In CD-CMG, when a production reduction occurs, it

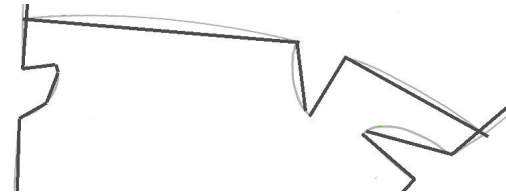


Figure 2. Extraction of primitives. The original drawing is in light gray, and the extracted primitives are in black.

means that a multiset of elements is replaced by another one α . This reduction has impacts on the analysis of the document: now that new α elements exists in the document, it may be possible to add new others, positioned relatively to these α elements. The interpretation of α creates a new zone, called *relative position* to search the next elements to analyse in the document.

The primitive interpretation depends on its neighbourhood in structured documents: the structured document analysis requires a two-dimensional context. The analyzer begins by defining a spatial contextual focus that aims to limit the combinatory exploration due to the breadth-first exploration of the analysis tree.

Each primitive can be interpreted in several ways which led to a construction of an analysis tree. In the building of the analysis tree, the analyzer explores all the possible hypotheses of interpretation using breadth-first exploration in the spatial context with the algorithm described in [4]. Each root is the production rule that would consume this primitive. Each node or leaf is the

application of a production rule deduced from the previous node. The number of analysis trees corresponds to the number of possible interpretations for the current primitive.

Once the tree is constructed, we start the decision phase. The role of the decision process is to validate the right hypothesis among a set of competing hypotheses generated with a breadth first analysis. The analyzer solicits the user in case of ambiguity, i.e. when more than one hypothesis can be validate with almost the same score. Otherwise, the analyzer validates the best hypothesis.

When the correct root is validated, other roots are put on hold and the new roots are either the sons of this root if exists, or the waiting roots otherwise. The analyzer goes back to the first step (defining the local context). The analysis is complete when no more production rule is applicable.

2.3 IMISketch application on furniture

In this section, we present an example of interpretation produced by IMISketch. We want to analyze the primitives $\{1, \dots, 6\}$ illustrated in Figure 3 with the grammar described in Figure 4. In the grammar, each new interpreted element induces the creation of a new *relative position* for the search of the following elements to recognize. We are mainly interested in the combinatory at the analysis tree level. According to Figure 4, there are two ways to interpret the primitive '1':

- As wall: the Figure 5 shows the analysis tree corresponding to all the hypotheses generated by this interpretation.
- As a part of furniture: the Figure 6 shows the part of analysis tree corresponding to some hypotheses generated by this interpretation. We present just a part of the tree analysis for more clarity. This analysis tree generates a lot of combinatory, and we obtain several equivalent hypotheses.

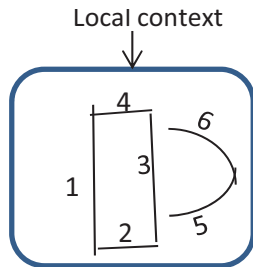


Figure 3. Primitives to interpret

Name	Created elements	Consumed elements
P _s	Starting wall	Longest line in the document
P ₁	Wall	Line at the end of another wall
P ₂	Sequence_Opening	Line at the end of an interpreted wall or sequence_Opening
P ₃	Sequence_Furniture	A sequence and two colinear walls or a sequence and wall and a line wich are colinear
P ₄	Door	A sequence and two colinear walls or a sequence and wall and a line wich are colinear
P ₅	Window	A sequence and two colinear walls or a sequence and wall and a line wich are colinear
P ₆	Furniture	A Sequence_Furniture

Figure 4. Example of production rules used for the architectural plans

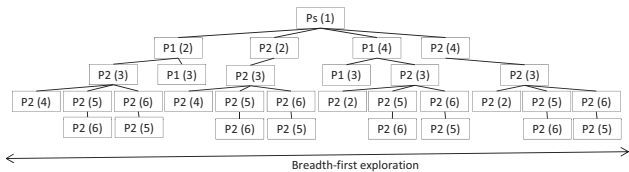


Figure 5. Analysis tree resulting from the interpretation of the primitive '1' as wall

Despite our previous optimizations on IMISketch [4], the combinatorial problem remains a major problem especially when the documents are more complex than walls and openings. For example, the analysis tree for interpreting the primitive '1' in a component of furniture requires more than 50 nodes (Figure 6). We present in the next section, a new hybrid strategy to improve the exploration of the analysis tree.

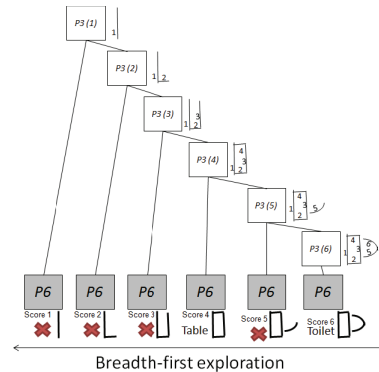


Figure 6. Part of the analysis tree resulting from the interpretation of the primitive '1' as furniture (should contain about 50 nodes)

3. Hybrid exploration

Symbols with a complex structure, such as furniture in the case of architectural floor plan, generates a lot of combinatory because they are composed of a big number of primitives interconnected in a reduced zone. Despite breadth-first exploration offers the possibility to choose between several possible hypotheses for interpreting primitives, this exploration generates a lot of combinatory. In opposite, the depth-first exploration avoids combinatorial problems but does not generate all hypotheses. This exploration may suggest a wrong interpretation of primitives.

In this section, we propose a dynamic strategy to switch between a breadth-first exploration and a depth-first exploration to reduce the combinatory. We improve the use of the existing CD-CMG grammar to drive this new strategy analysis. In our hybrid strategy, CD-CMG is not restricted to the modeling of the document but also used for the choice of the exploration strategy: either breadth-first exploration or depth-first exploration.

3.1 Breadth-first / Depth-first exploration

The CD-CMG grammar drives the exploration of the analysis tree. It enables to choose between breadth-first and depth-first exploration at each step of analysis.

- *depth-first exploration*: this exploration strategy is chosen if the production rule applied at the root of the analysis tree generates **only one way** to interpret the other primitives of the same group (interconnected primitives).
- *breath-first exploration*: if the interpretation of the root of the analysis tree generates **several ways** to interpret the primitives of the same group.

As we have said previously the depth-first exploration may not generate all the hypotheses. Consequently, we propose to reduce the risks by limiting the possible zone of application of the depth-first analysis. We implement this idea using the concept of *relative position*.

3.2 Reduced relative position Vs extended relative position

We recall that the *relative position* is the search zone that is created after the interpretation of each element in order to continue the analysis. In this analyzer, we combine two kinds of relative positions (Figure 7):

- *Extended relative position*: this position is the same to that used in previously IMISKetch. When

this position is activated, the analyzer adopts the classical breadth-first exploration.

- *Reduced relative position*: this zone search is smaller and enables to adopt the depth-first exploration. This position is intended to collect interconnected primitives that are very close.

In the grammatical description, each interpreted elements can create the both kinds of relative positions. Thanks to these positions, the hybrid exploration is totally led by the grammatical description and can be adapted for the description of each element present in the document. If a created element is associated to two positions, the reduced position is used in priority. The transition from reduced position to extended position is established only after no rules of productions is applicable.

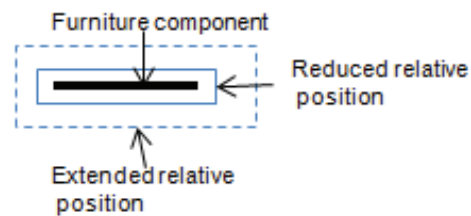


Figure 7. Example of two relative position for a furniture component

3.3 Example of hybrid exploration

In this section we return to the same example of section 2.3. Figure 3 illustrates the primitives to interpret and Figure 4 describes the production rules.

In the case of architectural floor plans, we consider that the close interconnected set of basic primitives belong either to the same furniture (toilet, table, bed) or to one or more walls and openings (doors, windows); the not interconnected primitives elements belong to the same or different elements. For this, we have described in the grammar that the walls create only *extended* relative positions. The furniture rule now creates two relative positions: *reduced* and *extended*.

Primitives {1, 2, 3, 4} are interconnected and form the group G1. Primitives {5, 6} are interconnected and form the group G2. By applying the production rules described in Figure 4, there are two ways to interpret the primitive '1': a wall or a component of furniture. Since the wall creates only an extended relative position, the analysis tree thus constructed is similar to IMISKetch

(Figure 5). Figure 8 shows the analysis tree obtained with the new exploration strategy about how to obtain furniture. The analyzer uses the reduced relative position to group primitives that can form a furniture by adopting a depth-first exploration. The group {1, 2, 3, 4} forms a table. The extended position is used to find another hypothesis of furniture. {1, 2, 3, 4, 5, 6} may form another furniture: toilet.

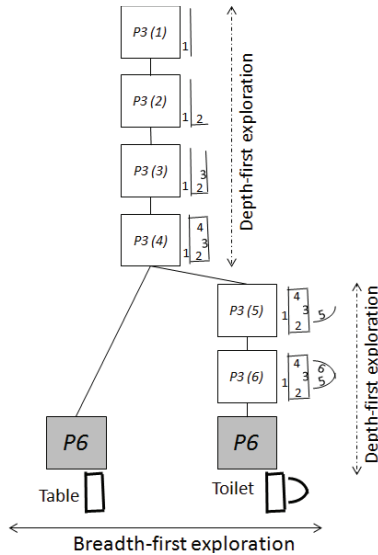


Figure 8. Analysis tree with Hybrid exploration: transformation of primitive '1' to table or toilet.

With this new strategy, to interpret the primitive '1' as a component of a furniture, we pass from around 50 nodes with IMISketch to 7 nodes with the hybrid exploration. We have shown in this simple example, by applying our new strategy on six basic primitives, we won 86% of exploration of nodes. This new hybrid strategy also lets alleviate the competitive hypotheses. In our example, this new strategy allows to pass from 6 competing hypotheses (Figure 6) to 2 hypotheses (Figure 8).

4 Experimental results

In this section we report results obtained from the hybrid exploration strategy described in section 3 on the interactive method of document analysis. The aim of this strategy is to reduce the analysis time of the document to meet the criteria of acceptability and usability of the system.

The experiments were carried on 24 architectural plans of different complexity drawn by six different people. Each architectural plan is composed of a set of walls, opening (doors, windows and sliding windows) and furniture (single sink, double sinks, toilet, bathtub,

square table, table with chairs, single bed, double bed, shower, couch). Figure 9 shows some examples of analyzed architectural plans and Figure 10 reports the total number of symbols of the database.

To evaluate our new exploration strategy, we adopt two experiments. The first one consists of comparing our new hybrid approach to IMISketch method (based on breadth-first exploration). In this experiment, we analyzed 7 architectural plans using IMISketch. This interpretation requires more than 5300 nodes per image and therefore a large computation time. The interpretation of the same plans with our new exploration strategy requires an average of 1713 nodes and thus acceptable computation time for an interactive method. This represents a gain of 70% (Figure 11).

Today, the average computation time is 4 min per image. In future work, we will try to reduce the computation time to attain around 1 min 30: experimentally, it seems the ideal timing that allows the user to follow in real time the recognition process.

The second experiment evaluate the rate of structural recognition, i.e. the interpretation of primitives into walls, openings and furnitures, using our hybrid strategy and the presence of the user in the loop of analysis.

Figure 13 reports recognition rates based on 24 architectural handwritten plans with varying threshold of user solicitations. This figure shows the evolution of recognition rates related to the number of user solicitations. We can notice that the best compromise (recognition rates / user solicitations) is obtained with 12 user solicitations per image: it means that 4% of primitive interpretations require the user solicitation. 49% of user solicitations are useful to take the right decision which is not the best hypothesis proposed by the analyser. The structural recognition rates increase from 91% without user solicitations to 96% with user solicitations. The remaining errors (about 4%) are mostly due to badly drawn symbols.

With this strategy, we can deal with complex problems where the furniture is attached to wall. Figure 12(a) and Figure 12(b) present examples of furniture that share a component with a wall. Figure 12(c) illustrates an example of furniture that has a component attached to wall. Figure 12(d) and Figure 12(e) show examples of furniture that has a component attached to an opening. The interpretation of these cases is very difficult with IMISketch because large combinatory is generated ; but in these complex cases, the furniture are localized with the hybrid strategy.



Figure 9. Examples of architectural floor plans

Number of architectural plans	24
Number of walls	961
Number of opening	414
Number of furniture	523

Figure 10. Database properties

IMISketch	> 5300
Our new hybrid strategies	1713

Figure 11. Average of explored nodes number

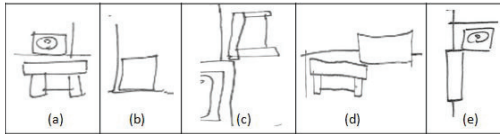


Figure 12. Example of complex interpretations

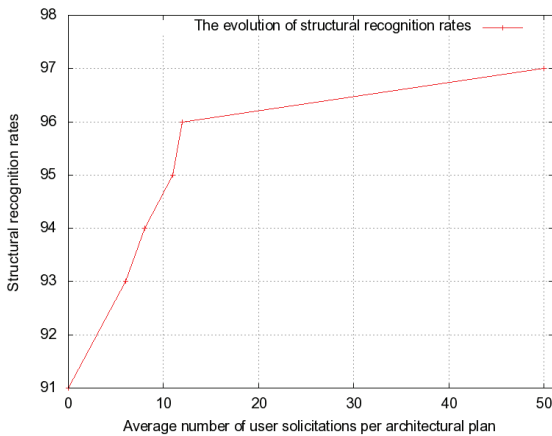


Figure 13. The evaluation of structural recognition rate

5 Conclusion

In this paper, we have presented a hybrid strategy for solving combinatory generated by an interactive analysis for the sketch recognition. This strategy alternates between a breadth-first and depth-first exploration. Note that this hybrid exploration is driven by the grammar. This strategy has been validated and tested on an interactive analyzer for interpreting handwritten architectural floor plans.

Future work will focus on extending the experimental results on large image databases containing printed and vectored document architectural plans... We will also validate the criteria of acceptability and usability of the system by doing usage tests that will be conducted in collaboration with experts from the laboratory uses "Loustic" (<http://www.loustic.net/rennes>).

Acknowledgment

The authors would like to thank all the people who took part in the experiments. This work partially benefits from the financial support of the ANR Project MobicSketch.

References

- [1] K. Chan and D. Yeung. An efficient syntactic approach to structural analysis of on-line handwritten mathematical expressions. *Pattern Recognition*, 33(3):375–384, 2000.
- [2] B. Coïtasnon. Dmos, a generic document recognition method: Application to table structure analysis in a general and in a specific way. *IJDAR 2006*, 8(2):111–122, 2006.
- [3] J. Fitzgerald, F. Geiselbrechtinger, and T. Kechadi. Mathpad: A fuzzy logic-based recognition system for handwritten mathematics. In *ICDAR 2007*, volume 2, pages 694–698, 2007.
- [4] A. Ghorbel, E. Anquetil, and A. Lemaitre. Optimization analysis based on a breadth-first exploration for a structural approach of sketches interpretation. In *10th IAPR International Workshop on Document Analysis Systems (DAS'12)*, 2012.
- [5] A. Ghorbel, S. Macé, A. Lemaitre, and E. Anquetil. Interactive competitive breadth-first exploration for sketch interpretation. *ICDAR 2011*, pages 1195–1199, 2011.
- [6] A. Lemaitre and J. Camillerapp. Text line extraction in handwritten document with kalman filter applied on low resolution image. *Document Image Analysis for Libraries, International Workshop on*, 0:38–45, 2006.
- [7] S. Macé and E. Anquetil. Eager interpretation of on-line hand-drawn structured documents: The dali methodology. *Pattern Recognition*, 42(12):3202–3214, 2009.
- [8] S. Mao, A. Rosenfeld, and T. Kanungo. Document structure analysis algorithms: a literature survey. In *Proc. SPIE Electronic Imaging*, volume 5010, pages 197–207, 2003.