Relational Models for Visual Understanding of Graphical Documents.
Application to Architectural Drawings.

Lluís-Pere de las Heras

Computer Vision Center, Bellaterra, Catalonia / Spain
Advisor/s: Gemma Sánchez
Date and location of PhD thesis defense: 15 December 2014, Universitat Autònoma de Barcelona

Abstract

Graphical documents express complex concepts using a visual language. This language consists of a vocabulary (symbols) and a syntax (structural relations between symbols) that articulate a semantic meaning in a certain context. Therefore, the automatic interpretation by computers of these sort of documents entails three main steps: the detection of the symbols, the extraction of the structural relations between these symbols, and the modeling of the knowledge that permits the extraction of the semantics. Different domains in graphical documents include: architectural and engineering drawings, maps, flowcharts, etc.

Graphics Recognition in particular and Document Image Analysis in general are born from the industrial need of interpreting a massive amount of digitalized documents after the emergence of the scanner. Although many years have passed, the graphical document understanding problem still seems to be far from being solved. The main reason is that the vast majority of the systems in the literature focus on very specific problems, where the domain of the document dictates the implementation of the interpretation. As a result, it is difficult to reuse these strategies on different data and on different contexts, hindering thus the natural progress in the field.

In this thesis, we face the graphical document understanding problem by proposing several relational models at different levels that are designed from a generic perspective, see Figure 1. Firstly, we introduce three different strategies for the detection of symbols. The first method tackles the problem structurally, wherein general knowledge of the domain guides the detection [3]. The second is a statistical method that learns the graphical appearance of the symbols and easily adapts to the big variability of the problem [4]. The third method is a combination of the previous two methods that inherits their respective strengths, i.e. copes the big variability and does not need annotated data [2]. Secondly, we present two relational strategies that tackle the problem of the visual context extraction. The first one is a full bottom up method that heuristically searches in a graph representation the contextual relations between symbols [1]. Contrarily, the second is syntactic method that models probabilistically the structure of the documents. It automatically learns the model, which guides the inference algorithm to encounter the best structural representation for a given input. Finally, we construct
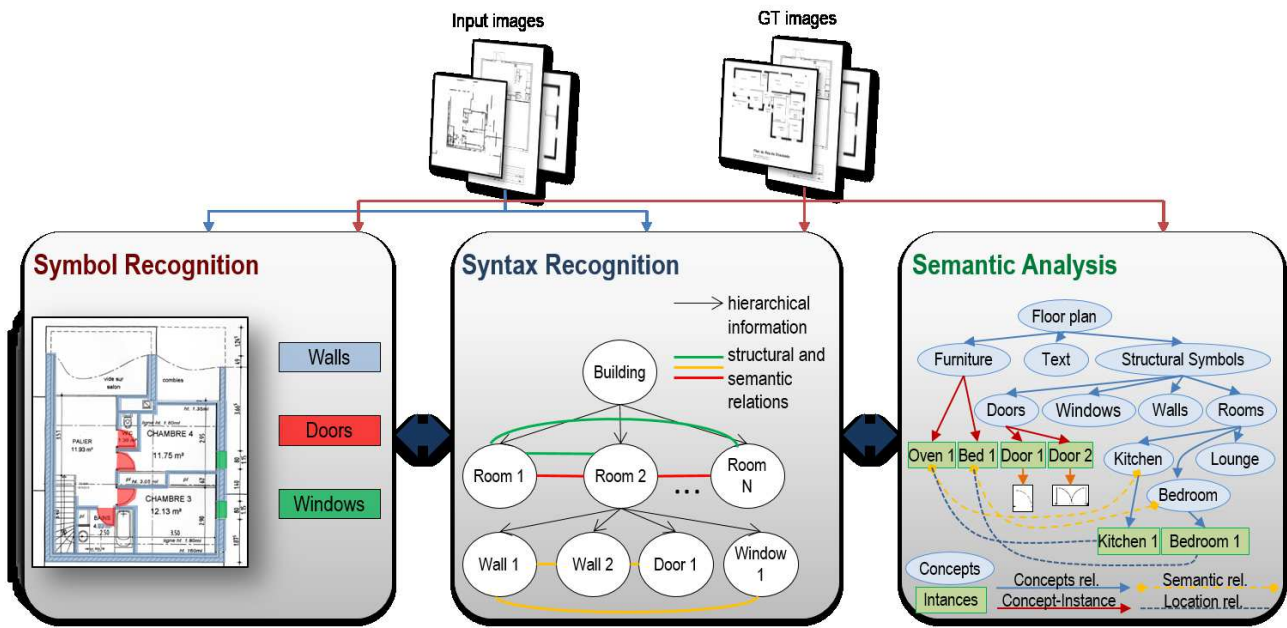
Figure 1: System overview.

a knowledge-based model consisting of an ontological definition of the domain and real data. This model permits to perform contextual reasoning and to detect semantic inconsistencies within the data. We evaluate the suitability of the proposed contributions in the framework of floor plan interpretation. Since there is no standard in the modeling of these documents there exists an enormous notation variability from plan to plan in terms of vocabulary and syntax. Therefore, floor plan interpretation is a relevant task in the graphical document understanding problem. It is also worth to mention that we make freely available all the resources used in this thesis –the data, the tool used to generate the data, and the evaluation scripts– with the aim of fostering research in the graphical document understanding task [5].

References

[1] L.-P. de las Heras, S. Ahmed, M. Liwicki, E. Valveny, and G. Sánchez. Statistical segmentation and structural recognition for floor plan interpretation. *International Journal on Document Analysis and Recognition*, vol. 17 pages 221-237, 2014.

[2] L-P. de las Heras, E. Valveny, G. Sánchez. Unsupervised and Notation-Independent Wall Segmentation in Floor Plans Using a Combination of Statistical and Structural Strategies. *Graphics Recognition. Current Trends and Challenges. (LNCS)*, B. Lamiroy, and J.-M. Ogier, (Eds.), pages 109-121. 2014.

[3] L-P. de las Heras, D. Fernández, E. Valveny, J. Lladós, and G. Sánchez. Unsupervised wall detector in architectural floor plans. *In International Conference on Document Analysis and Recognition*, pages 1245-1249. 2013.

[4] L-P. de las Heras, J. Mas, G. Sánchez, and E. Valveny. Wall Patch-Based Segmentation in Architectural Floorplans. *International Conference on Document Analysis and Recognition*, pages 1270-1274. 2011.

[5] L-P. de las Heras, O. Ramos Terrades, S. Robles, G. Sánchez. CVC-FP and SGT: a new database for structural floor plan analysis and its groundtruthing tool *International Journal on Document Analysis and Recognition*, vol. 18 pages 38-70. 2015.