Syntactic Model for Semantic Document Analysis

Lluís-Pere de las Heras
Advisor: **Gemma Sánchez**
Computer Vision Center – Dept. Ciències de la Computació, UAB, Barcelona, Spain
{lpheras,gemma}@cvc.uab.es

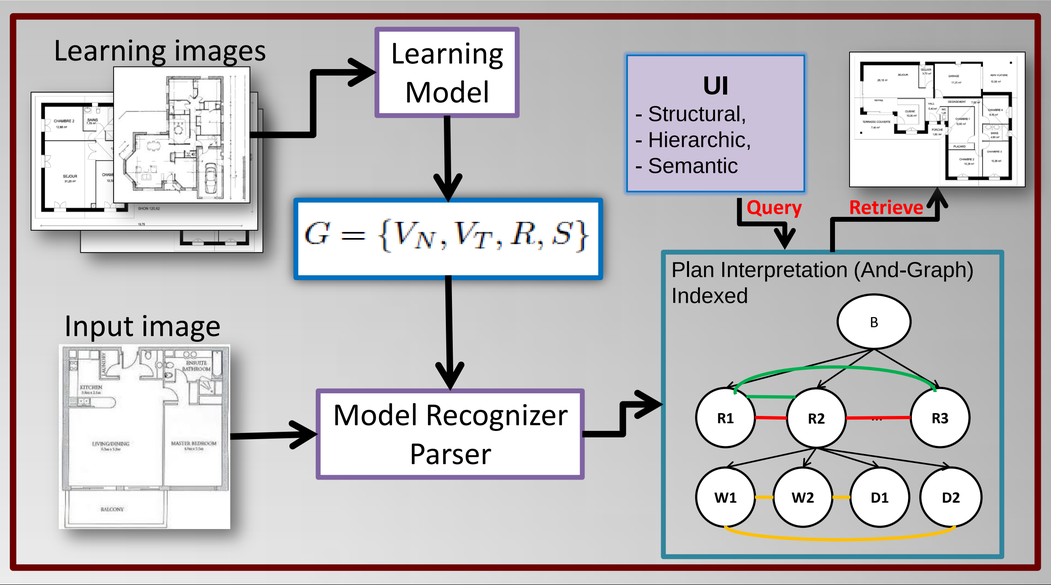
Centre de Visió per Computador

Universitat Autònoma de Barcelona

Problem statement

Problem: Still nowadays, floorplan interpretation is a non-solved problem, essentially because there is no standard notation defined. Elements such as walls, windows, furniture, indications, etc. are drawn distinctively depending on the architect and the country. Therefore, existing approaches are usually focused in one specific notation convention and, in order to deal with new notations, these methods need to be readapted specifically each time.

Project: The aim of this project is to create a general syntactic approach that, by taking into account the hierarchical and structural information between elements, will be capable to interpret and recognize different kinds of floorplan documents. Moreover, floorplan interpretations should be stored and indexed in order to be able to retrieve those plans accordingly to the hierarchic, structural and semantic user queries.



And-Or Graph Grammar for Architectural Floorplan Representation, Learning and Recognition [1]

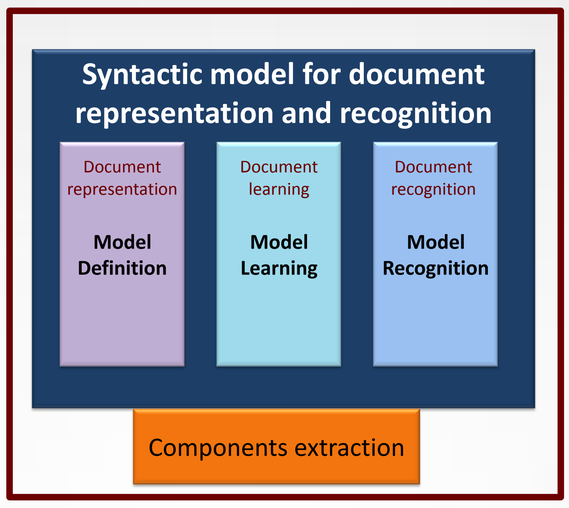
Learning Model:
- SCFG visual grammar over And-Or Graph.

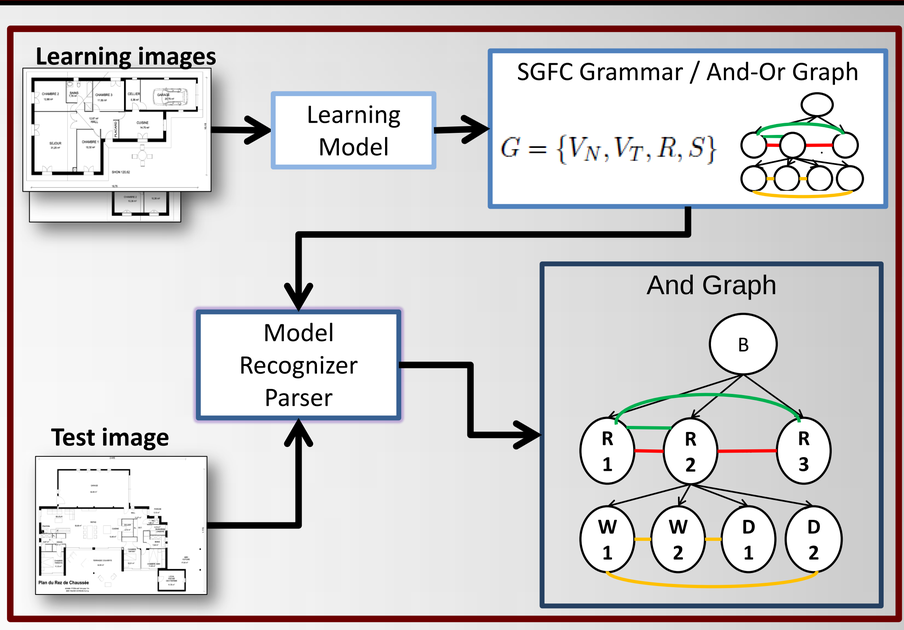
Model Recognition:
- Bottom-Up/Top-Down Parser.
- Probabilistic and Semantic Pruning.

Floorplan Interpretation:
- Hierarchical, semantic and structural interpretation.
- Validation of consistent floorplans.
- Interpretation at different levels of abstraction.

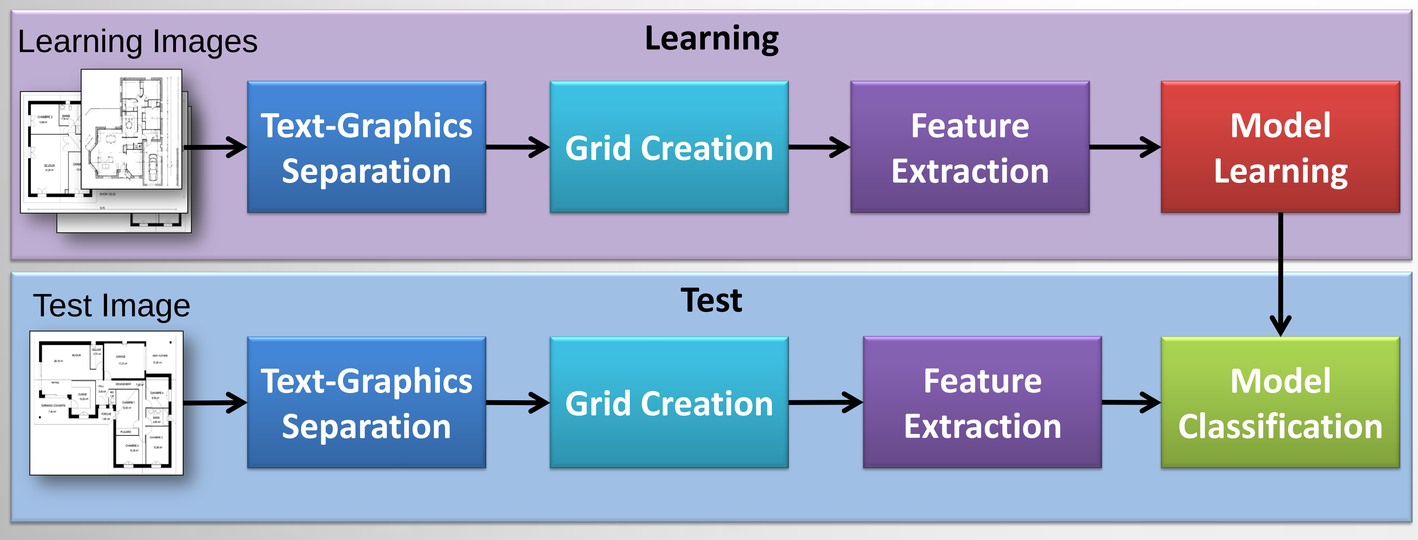
Problems:
- Primitive extraction dependant.
- Completely supervised and notation-oriented due to the extraction methods used.

Aim: Create a detection method capable to deal with multiple notations and resolutions.





Component Extraction Methods [2][3]



Quantitative results for methods [2] and [3] tested on two datasets with completely different notations and resolutions

Method	Dataset	GT	PS	D	LS	DS	ϕ_o	*JI*
[3]	*Dataset-1*	Non-overlapped	15×15	BSM_3	50000	-	-	0.9571
		Overlapped	15×15	BSM_3	50000	-	3	0.9667
[2]	*Dataset-1*	Non-overlapped	10×10	PCA	-	100	-	0.9543
		Overlapped	8×8	PCA	-	100	4	0.9673
[3]	*Dataset-2*	Non-overlapped	15×15	BSM_8	7500	-	-	0.7441
		Overlapped	15×15	BSM_8	7500	-	3	0.8233
[2]	*Dataset-2*	Non-overlapped	20×20	PCA	-	2000	-	0.7085
		Overlapped	20×20	PCA	-	2000	5	0.8241

Jaccard Index (JI) is calculated at pixel-level

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

Text Graphics Separation *[2][3]:*
- Otsu method for binarization.
- Text Graphics separation using [4].

Grid Creation
- Squared-rigid grid. *[2][3]*
- Squared-overlapped grid. *[2][3]*
- Squared-deformable grid (based on [6]). *[2]*

Feature Extraction
- Pixel Image Descriptor (PID). *[3]*
- PCA patch Descriptor. *[2][3]*
- Blurred Shape Model Descriptor (BSM). *[3]*

Model Learning
- **Vocabulary Creation + Word Probability Assignation *[2]:***
  - Fast version of K-means using [5].
  - Probability of belonging to each class for every visual word.
- **SVM Model Learning *[3]:***
  - Select randomly equal number of descriptors for each class.
  - Cross validation strategy.
  - RBF Kernel.

Model Classification
- **Bag-of-patches object recognition *[2]:***
  - Nearest Neighbor (NN) to the closest visual word in the dictionary.
  - The input patch inherits the probabilities of the closest visual word.
- **SVM Model Classification *[3]:***
  - Classification over all patch descriptors of the input image.

Conclusions

Complete Syntactic Model for floorplan interpretation [1]:
- Model works excellently with perfect component extraction techniques.
- Notation dependant on the component extraction techniques.

Component extraction methods [2], [3]:
- Patch-based methods for wall segmentation.
- For a new notation, only the parameters need to be retrained, without changing the methods themselves.
- High segmentation rates with floorplans with different notations and resolutions.

Future work: semi-unsupervised floorplan interpreter that deals with different notations.

References

[1] LP. de las Heras and Gemma Sánchez: "And-Or Graph Grammar for Architectural Floorplan Representation, Learning and Recognition. A Semantic, Structural and Hiararchical Model". *In Iberian Conference in Pattern Recognition and Image Analysis (IbPRIA). 2011.*

[2] LP. de las Heras, J. Mas, G. Sánchez and E. Valveny: "Wall Patch-Based Segmentation in Architectural Floorplans". In *International Conference in Document Analysis and Recognition (ICDAR2011). In press, 2011.*

[3] LP. de las Heras, J. Mas, G. Sánchez and E. Valveny: "Descritor-based Wall Detection on floorplans". In *Graphic Recognition (GREC2011). In press, 2011.*

[4] K. Tombre, S. Tabbone, L. Plissier, 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.

[5] C. Elkan, "Using the triangle inequality to accelerate kmeans," in the Twentieth International Conference on Machine Learning, 2003, pp. 147–153.

[6] D. Keysers, T. Deselaers, C. Gollan, and H. Ney, "Deformation models for image recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, pp. 1422–1435, 2007.