Parameter-free/Pareto-driven Procedural 3D Reconstruction of Buildings from Ground-Level Sequences

Loïc Simon   Olivier Teboul   Panos Koutsourakis   Luc Van Gool   Nikos Paragios

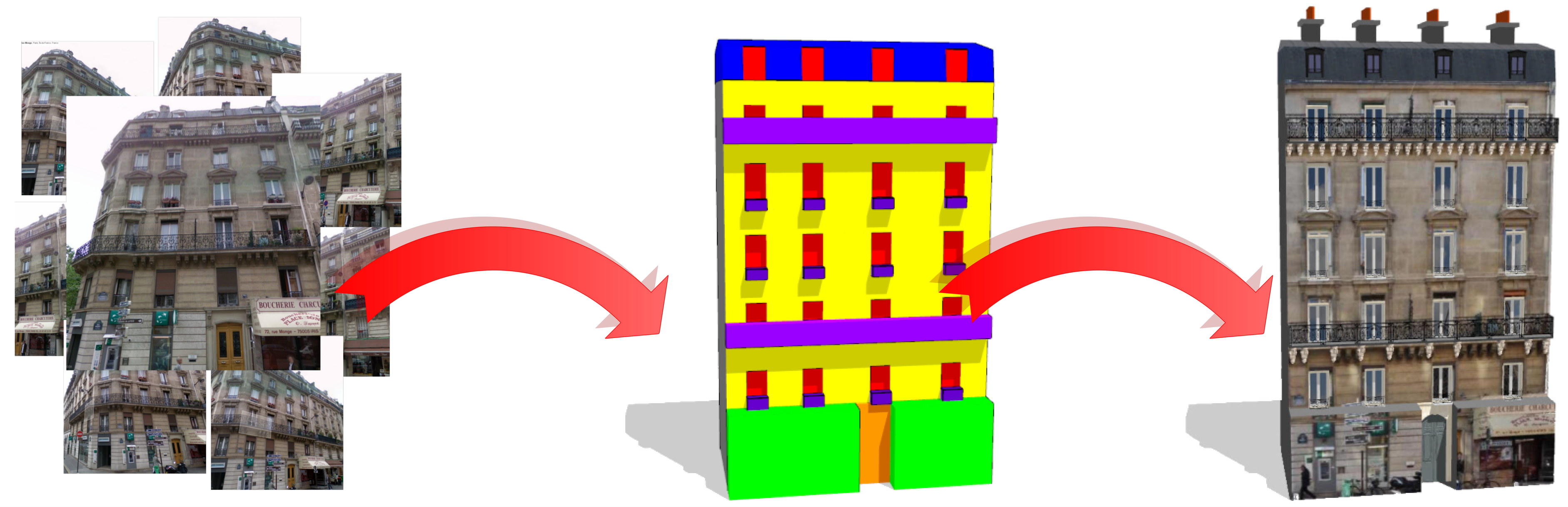
Center for Visual Computing, Ecole Centrale Paris – CMLA, ENS Cachan

CVPR'12
Providence, Rhode Island
June 16-21, 2012

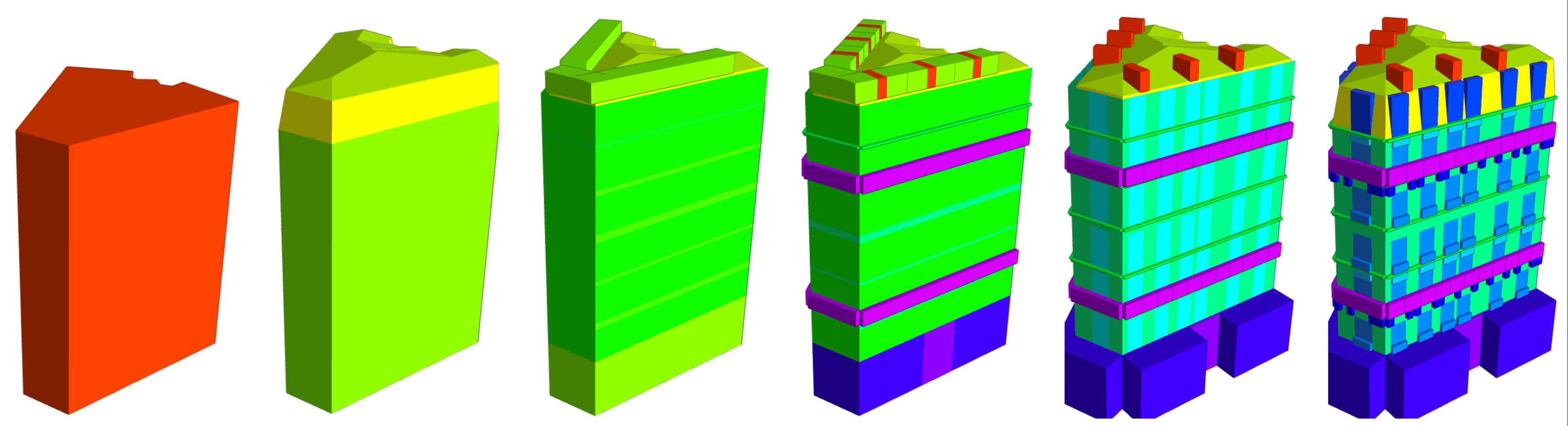
Overview

We address **multi-view reconstruction** of urban environments using 3D **shape grammars**. Our formulation expresses the solution to the problem as a shape grammar **parse tree** where both the tree and the corresponding derivation parameters are unknown. Besides the grammar constraint, the solution is guided by an image support that is twofold. First, we seek for a derivation that induces optimal **semantic partitions** in the different views. Second, using structure-from-motion, noisy **depth maps** can be determined towards minimizing their distance from to the ones predicted by any potential solution. We show how the underlying data structure can be efficiently optimized using **evolutionary algorithms** with automatic parameter selection. To the best of our knowledge, it is the first time that the multi-view 3D procedural modeling problem is tackled.

keywords: *Multi-view reconstruction, Shape Grammar, Evolutionary Algorithms*



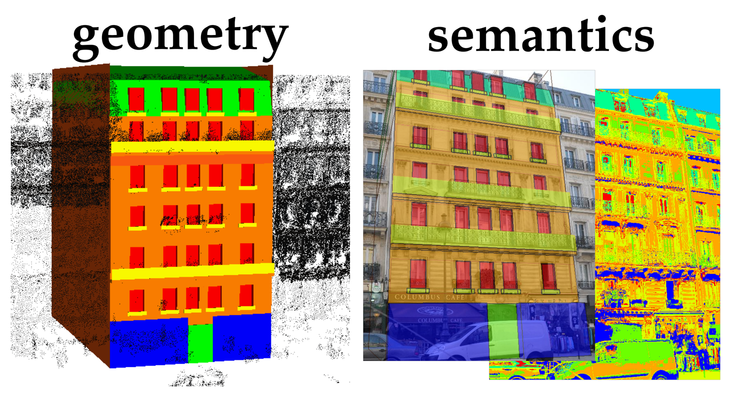
Shape Grammar Representations [1]



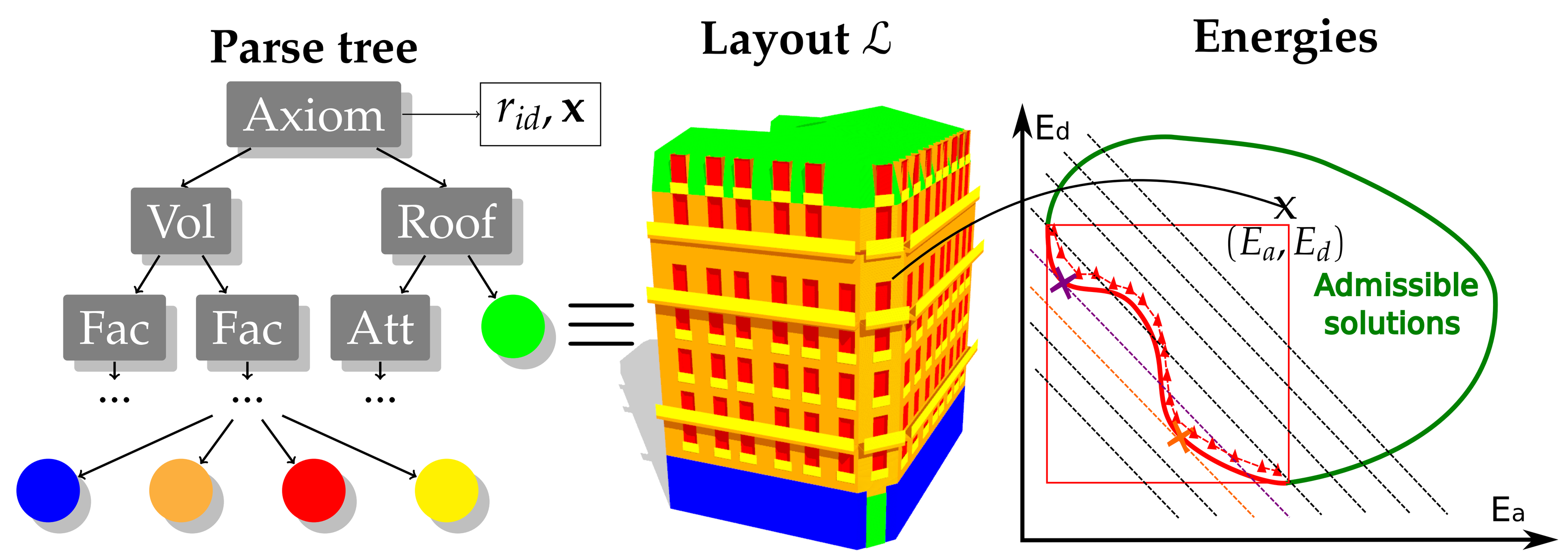
Data terms

Two concurrent energies:

- **Depth cues ↔ geometry**
  $E_d(\mathcal{L}) = \sum_m \sum_x \|D^m_{\mathcal{L}}(x) - D^m_{ref}(x) - \bar{d}^m\|^2$
- **Appearance cues ↔ semantics**
  $E_a(\mathcal{L}) = \sum_k \sum_x -\log P\left(I_k(x)|\mathcal{L}\right)$



Inference



Difficulties:

- variable length description
- discrete/continuous
- huge parse-tree
- multi-objective

Goal:

- estimate **Pareto front**
- optimal solution

Evolutionary Algorithms [2]

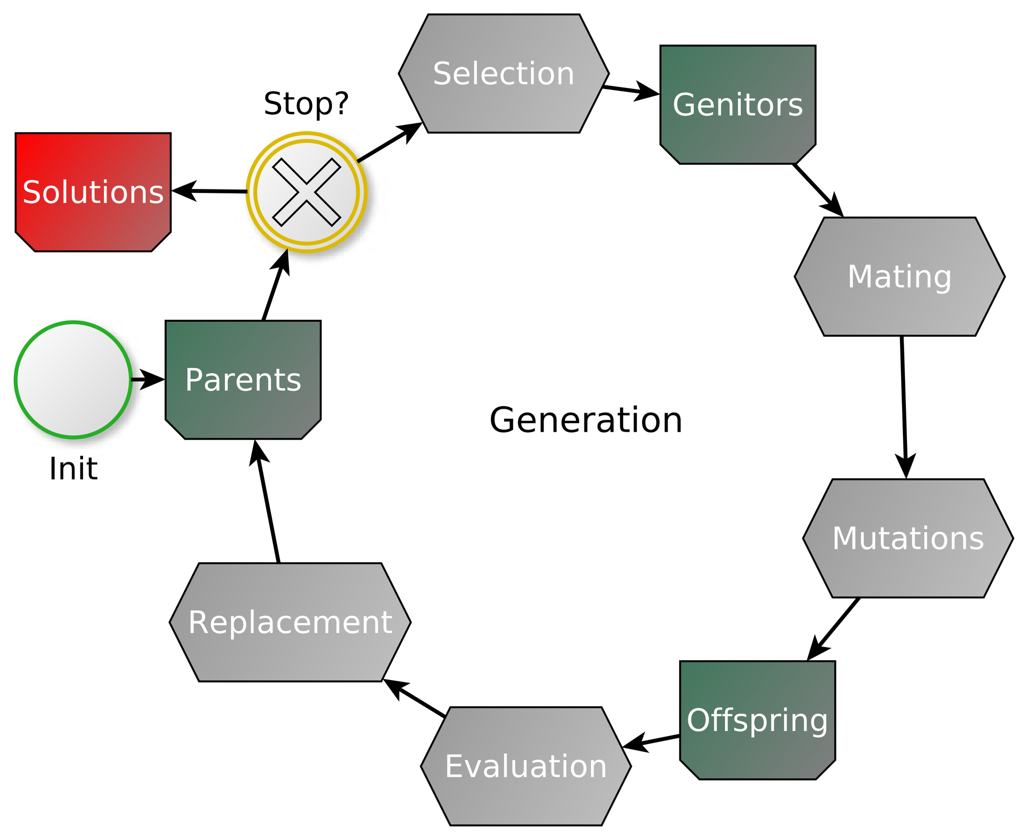


Figure: Typical pipeline of EAs.

Reconstruction context:

- *genotype:* parse tree
- *phenotype:* layout
- *fitness:* $(E_d(\mathcal{L}), E_a(\mathcal{L}))$
- *selection:* SPEA-II

Mutation of 1 tree:

- select a random node
- change r_{id} or **x**

Mating of 2 trees:

- select equivalent nodes
- swap content

⇒ Partial rederivation

SPEA-II selection & Optimal solution

Inputs:

P: current population
A: archive of non-dominated solutions
a: maximum size of archive

Outputs:

$P*$: genitors
A: updated archive

Execute:

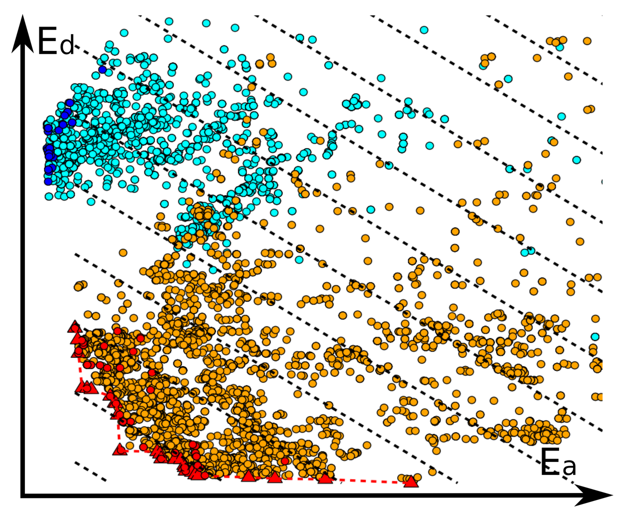
$A \leftarrow$ non-dominated solutions of $P \cup A$
if $\#A < a$:
  extend A with $(a - \#A)$ fittest in $P \setminus A$
elsif $\#A > a$:
  remove $(\#A - a)$ most dense from A
$P^* \leftarrow$ tournament selection on A

Optimal solution:

$$\mathcal{L}^* = \operatorname*{argmin}_A E^\alpha(\mathcal{L})$$

where,
$E^\alpha(\mathcal{L}) := \alpha E_a(\mathcal{L}) + (1-\alpha)E_d(\mathcal{L})$

SPEA-II vs tournament



Results

Settings:

- 10 buildings with GT layouts $\mathcal{L}_{gt}$
- 2 views per facade
- OSM footprints
- Bundler+PMVS

Performance:

primitive	geometry	semantics					
window	11 cm	**70**	24	5	0	1	0
wall	4 cm	3	**83**	13	0	0	0
balcony	13 cm	10	7	**82**	0	1	0
door	1 cm	0	2	0	**84**	0	14
roof	31 cm	8	6	7	0	**79**	0
shop	27 cm	0	4	0	2	0	**94**

Reconstruction results

Conclusion

Contributions:

- multi-view procedural reconstruction
- EA for grammar parsing
- multi-objective / automatic selection

Perspectives:

- data driven EA
- range data
- aerial

Bibliography

[1] **P. Müller, P. Wonka, S. Haegler et al.** *Procedural modeling of buildings.* ACMTOG, 2006

[2] **S. Luke.** *Essentials of Metaheuristics.* Lulu, 2009

[3] **F. Alegre and F. Dellaert.** *A probabilistic approach to the semantic interpretation of building facades.* In *International Workshop on VTARCC*, 2004

[4] **L. Simon, O. Teboul, P. Koutsourakis et al.** *Random exploration of the procedural space for single-view 3d modeling of buildings.* IJCV, 2011

[5] **O. Teboul, I. Kokkinos, L. Simon et al.** *Shape grammar parsing via reinforcement learning.* In *CVPR*, 2011