City modeling

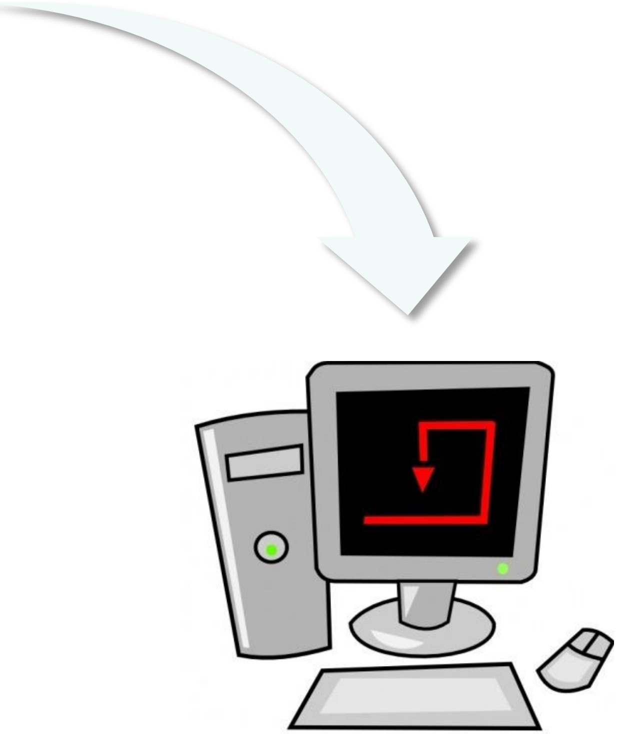
Florent Lafarge
Inria Sophia Antipolis - Mediterranee

What is city modeling ?

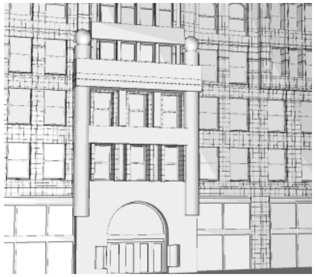
What is city modeling ?

What is city modeling ?

geometry



shape



What is city modeling ?

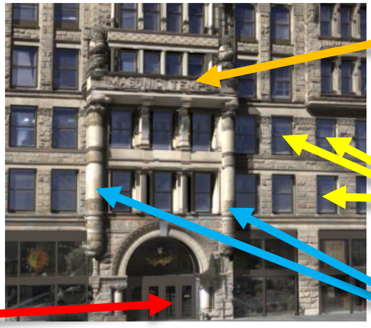
geometry
radiometry

color

What is city modeling ?

geometry
radiometry
semantics

meaning



Urban objects

- Permanent elements: Buildings, roads, bridges, trees…
- Temporary elements: cars, fences, cranes…

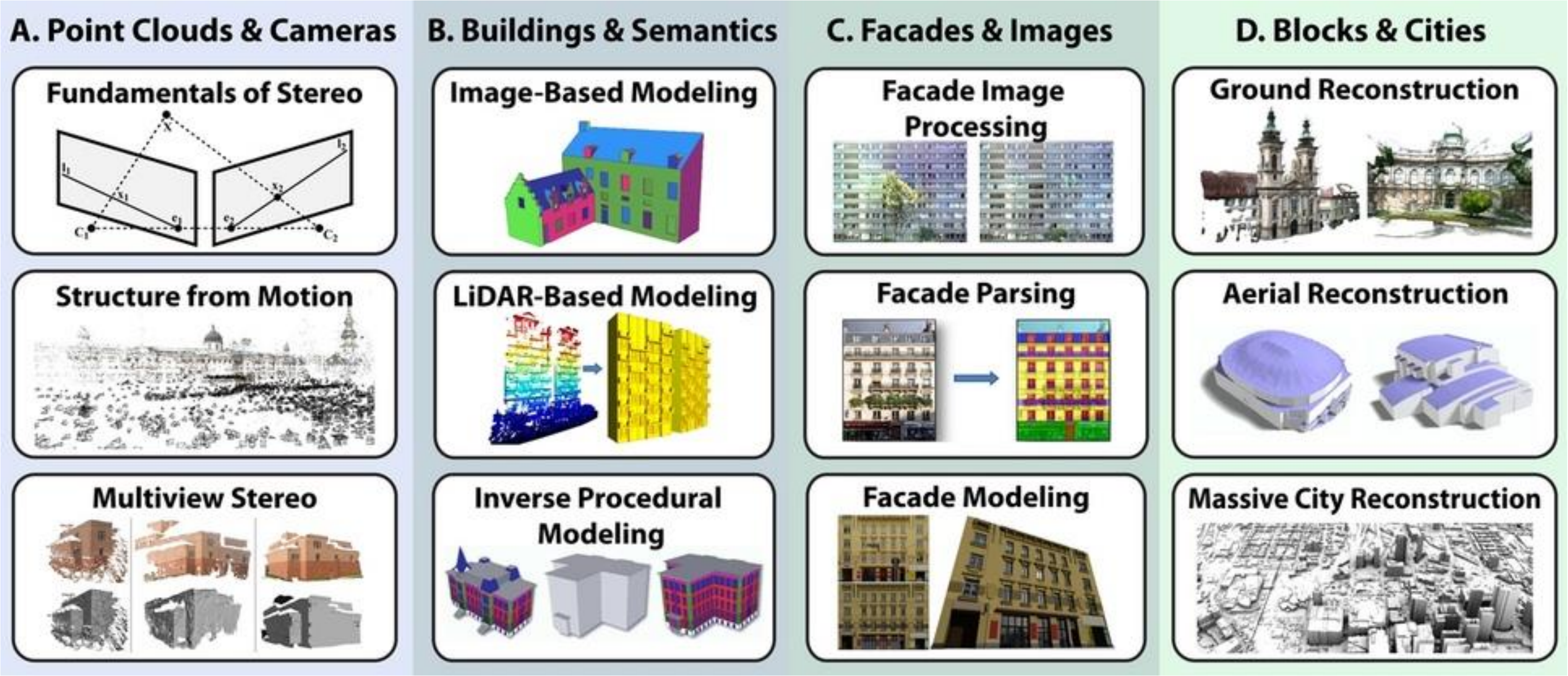
Urban objects

Objects differ in terms of:

- density
- diversity
- dependence with each other

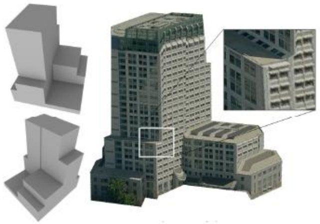
- Permanent elements: Buildings, roads, bridges, trees…
- Temporary elements: cars, fences, cranes…

Categories of problems



Building reconstruction

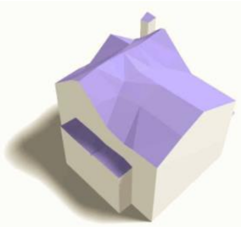
Manhattan-world
[Furukawa et al., 2009] [Vanegas et al., 2010]
[Poullis et al., 2009] [Matei et al., 2008]

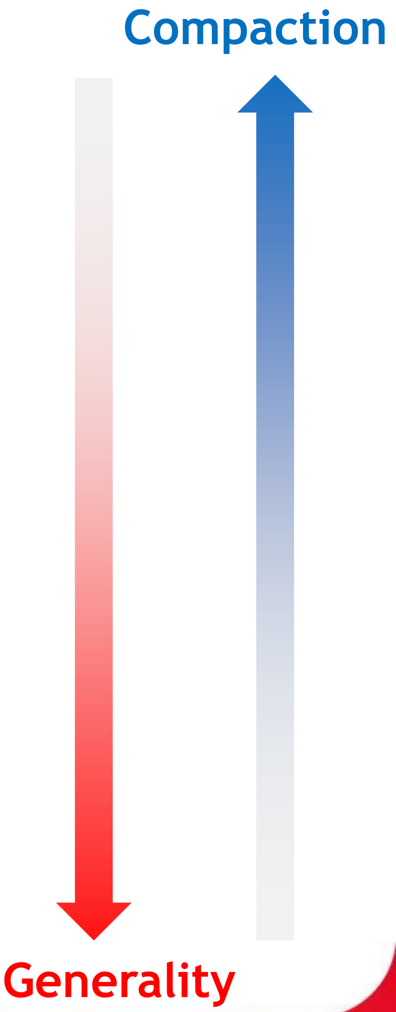


Piecewise planar structures
[Zebedin et al., 2008] [Brédif et al., 2007]

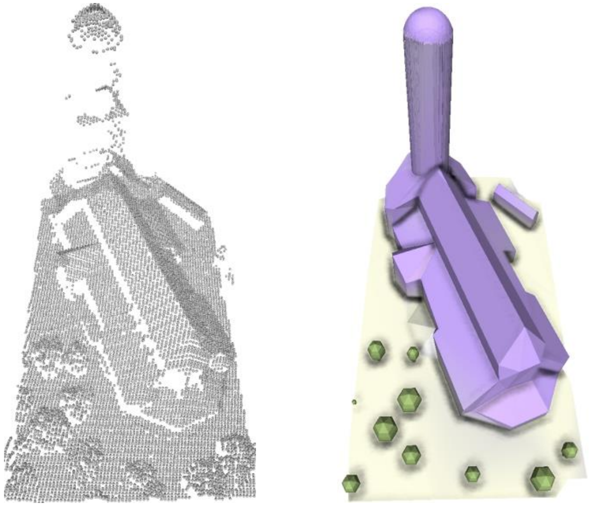
Block assembling / grammars
[Verma et al., 2006] [Lafarge et al., 2008]

Mesh simplification
[Zhou et al., 2010] [Verdié et al., 2011]

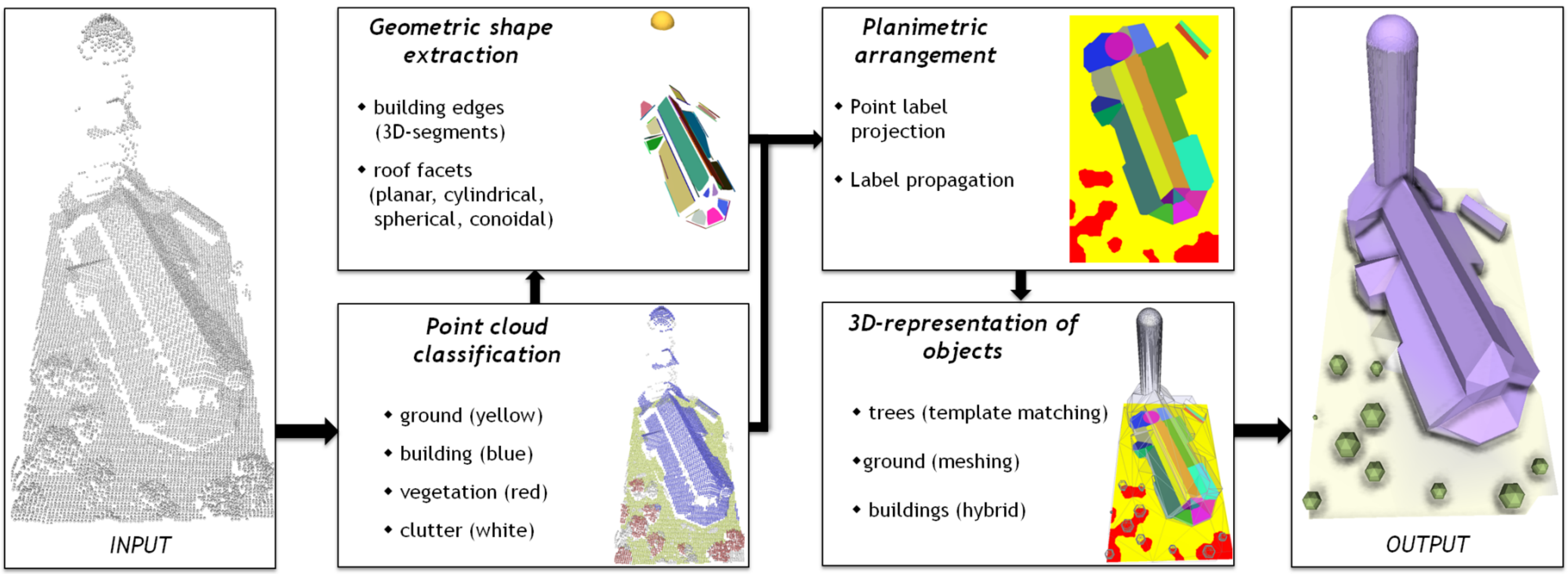




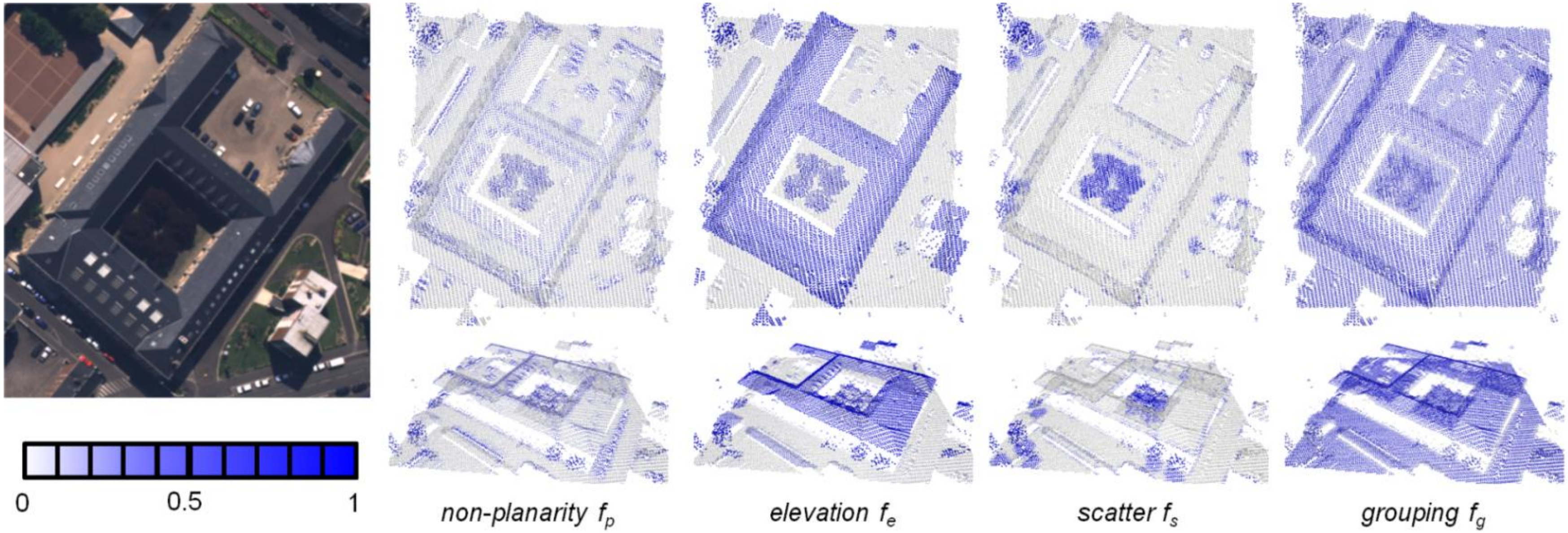
Example 1:
Reconstruction of cities from airborne Lidar



System overview



Discriminative attributes



classification

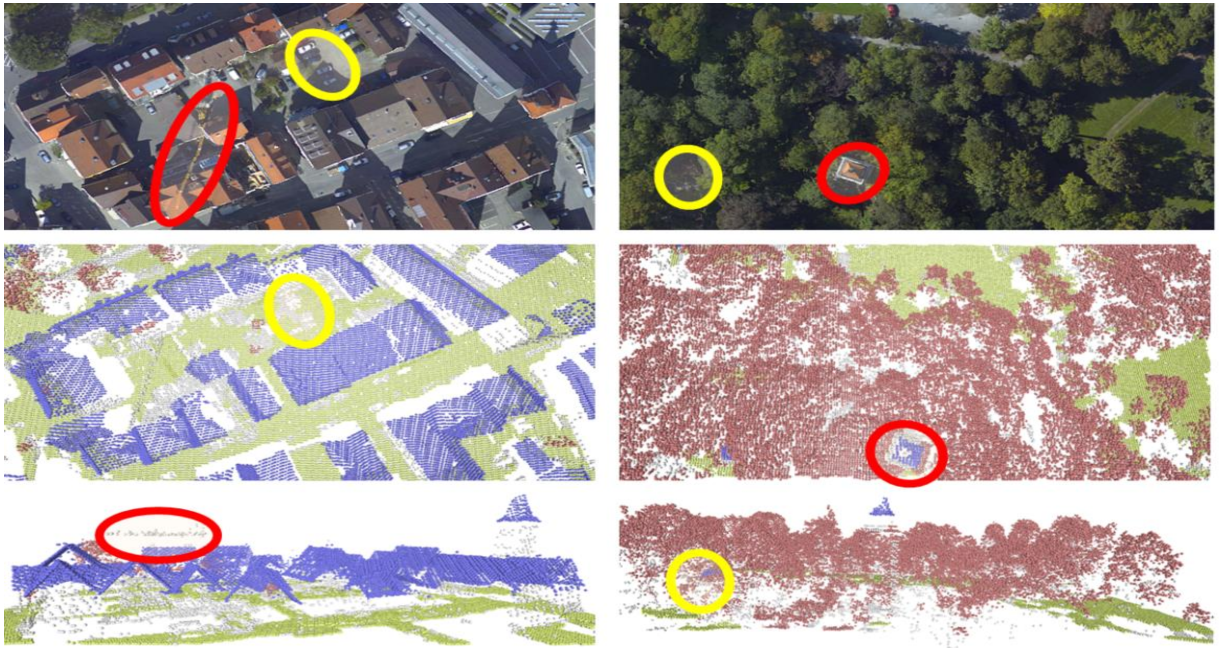
4 classes: building [blue], ground [yellow], vegetation [red], clutter [white]

Energy minimization: combination of the point descriptors + Potts model + Graph-Cuts

$$E(x) = \sum_{i=1..N_c} E_{di}(x_i) + \gamma \sum_{i \sim j} \mathbb{1}_{\{x_i \neq x_j\}}$$

$$E_{di}(x_i) = \begin{cases} (1-f_e).f_p.f_s & \text{if } x_i = \textit{building} \\ (1-f_e).(1-f_p).(1-f_s) & \text{if } x_i = \textit{vegetation} \\ f_e.f_p.f_s & \text{if } x_i = \textit{ground} \\ (1-f_p).f_s.f_g & \text{if } x_i = \textit{clutter} \end{cases}$$

classification

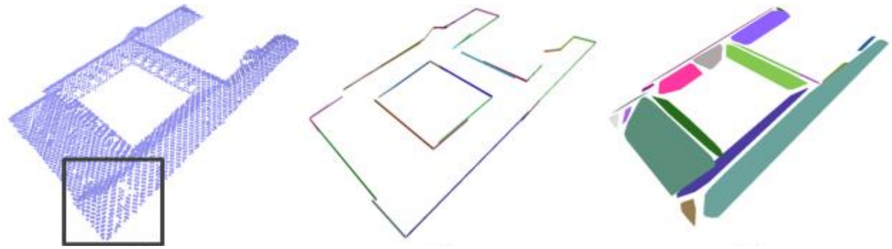


Color code: building [blue], ground [yellow], vegetation [red], clutter [white]

Clutter class includes

- outliers
- points of non significant urban objects (cars, wires, cranes, fences…)
- points of vertical structures (facades)

Primitive extraction

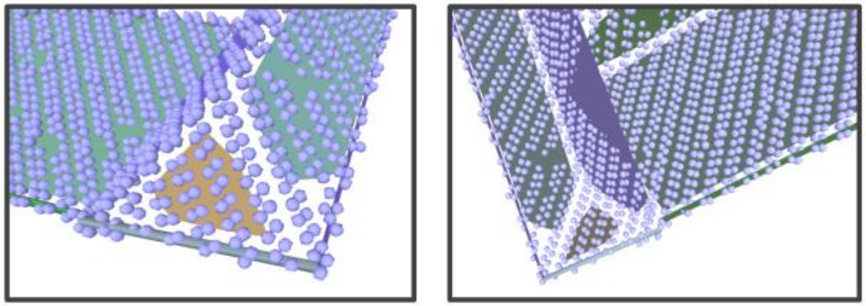


'*building*' points   3D-segments   surface primitives

Building contours by 3D-segments

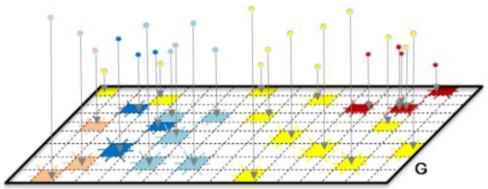
Roof sections

- by planes (region growing)
- by cylinders, spheres and cones on remaining points



Crop: top view   Crop: bottom view

Planimetric labeling

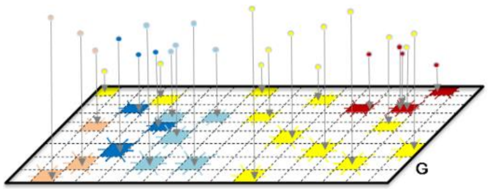


Configuration space L: point labels projected in a 2D-grid G

$$L = \{\text{ground}, \text{vegetation}, \text{plane}^{(l)}, \text{cylinder}^{(m)}, \text{sphere}^{(n)}, \text{cone}^{(o)}, \text{roof}\}^{card(G)}$$

Energy of standard form: $U(l) = \sum_{i \in G} D_i(l_i) + \beta \sum_{\{i,j\} \in E} V_{ij}(l_i, l_j)$

Planimetric labeling



Data term

$$D_i(l_i) = \begin{cases} c & \text{if } l_i = roof \\ \min(1, |z_{l_i} - z_{p_i}|) & \text{else if } i \in G^{(proj)} \\ 0 & \text{otherwise} \end{cases}$$

- altimetric error between the surface associated with l_i and the highest point of the cell i
- c controls the occurrence of irregular roof sections w.r.t. regular ones

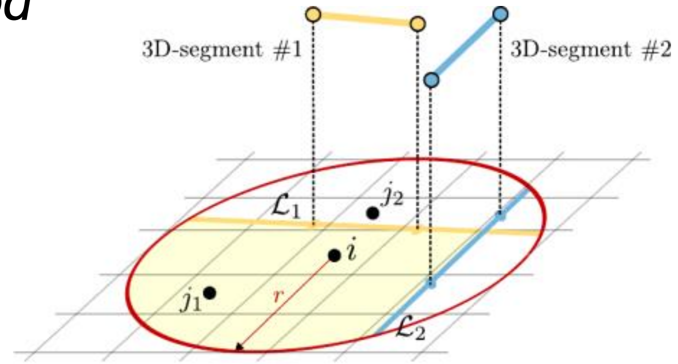
Planimetric labeling

Propagation constraints

$$V_{ij}(l_i, l_j) = \begin{cases} \epsilon_1 \text{ if } l_i \bowtie l_j \\ \epsilon_2 \text{ if } l_i = l_j \\ 1 \quad \text{otherwise} \end{cases}$$

- *breakline-dependent neighborhood*

$$\{i, j\} \in E \Leftrightarrow \begin{cases} \|i - j\|_2 \leq r \\ \mathcal{O}(i, \mathcal{L}_k) = \mathcal{O}(j, \mathcal{L}_k) \end{cases}$$



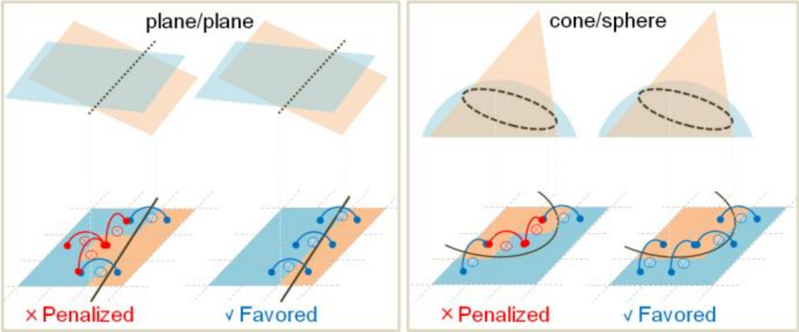
Planimetric labeling

Propagation constraints

$$V_{ij}(l_i, l_j) = \begin{cases} \epsilon_1 \text{ if } l_i \bowtie l_j \\ \epsilon_2 \text{ if } l_i = l_j \\ 1 \quad \text{otherwise} \end{cases}$$

- *breakline-dependent neighborhood*
- *structure arrangement law*

$$l_i \bowtie l_j \Leftrightarrow \mathcal{O}(i, \mathcal{I}_{l_i,l_j}) \neq \mathcal{O}(j, \mathcal{I}_{l_i,l_j})$$



Planimetric labeling

Propagation constraints

$$V_{ij}(l_i, l_j) = \begin{cases} \epsilon_1 \text{ if } l_i \bowtie l_j \\ \epsilon_2 \text{ if } l_i = l_j \\ 1 \quad \text{otherwise} \end{cases}$$

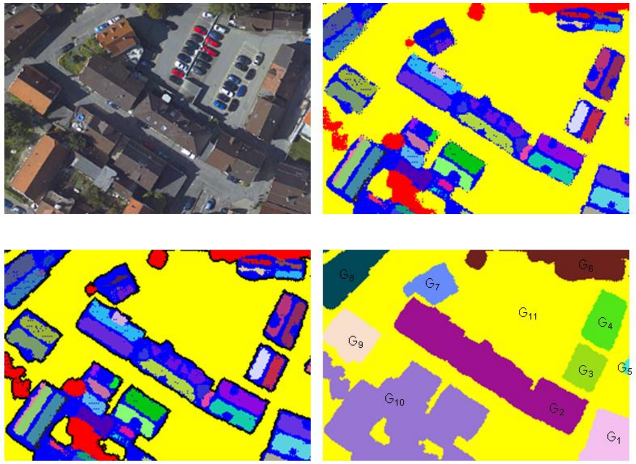
- *breakline-dependent neighborhood*
- *structure arrangement law*
- *label smoothness*

Optimization with parallelization scheme

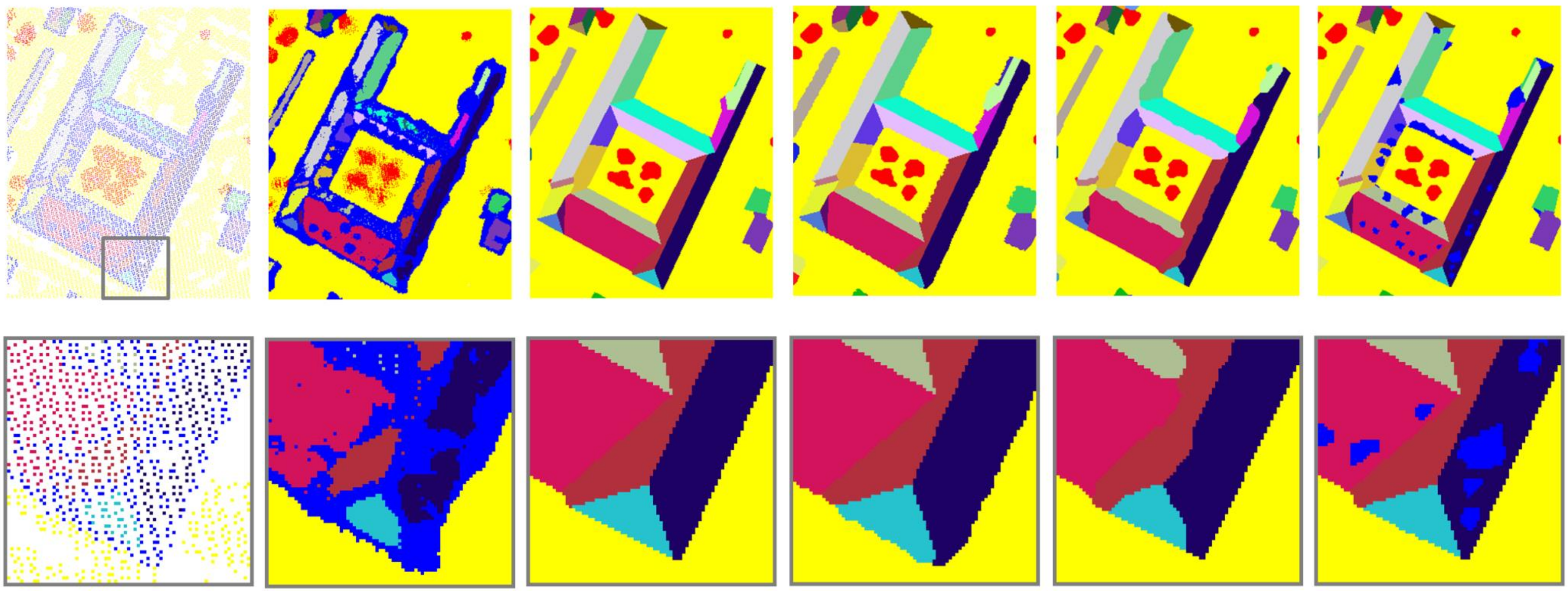
Minimizing *U* on a large scene: extremely time consuming !!
(>6 hrs for a 1km² urban area)

Decomposition in independent sub-problems :

$$\min_{l \in L} U(l) \Leftrightarrow \begin{cases} \min_{l_{/G_k} \in L_k} U(l_{/G_k}) \ , \ \forall k \in [1, N-1] \\ l_{/G_N} = \{\text{ground}\}^{card(G_N)} \end{cases}$$



Impact of the various energy composants



labels originally projected on the 2D-grid

initial label map

label map after minimizing U

label map after minimizing U without breakline-dependent neighborhood

label map after minimizing a variant of U without structure arrangement

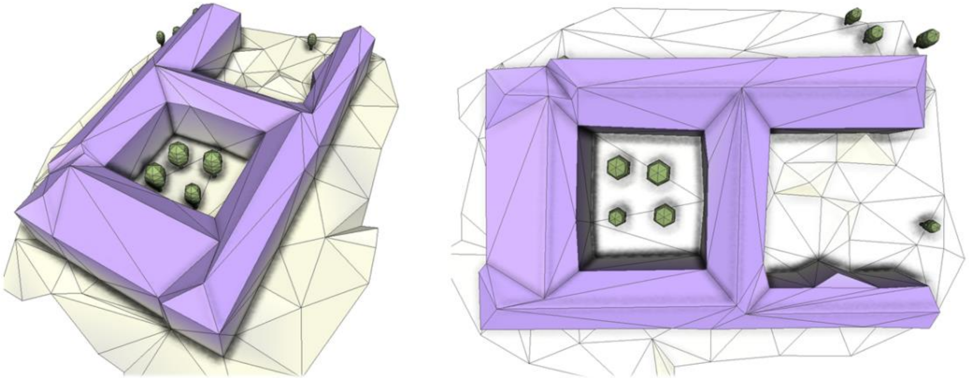
label map after minimizing U whose parameter c has been significantly decreased

Color code: roof [blue], ground [yellow], vegetation [red], empty cell [white] , surface primitives [random color]

Object modeling

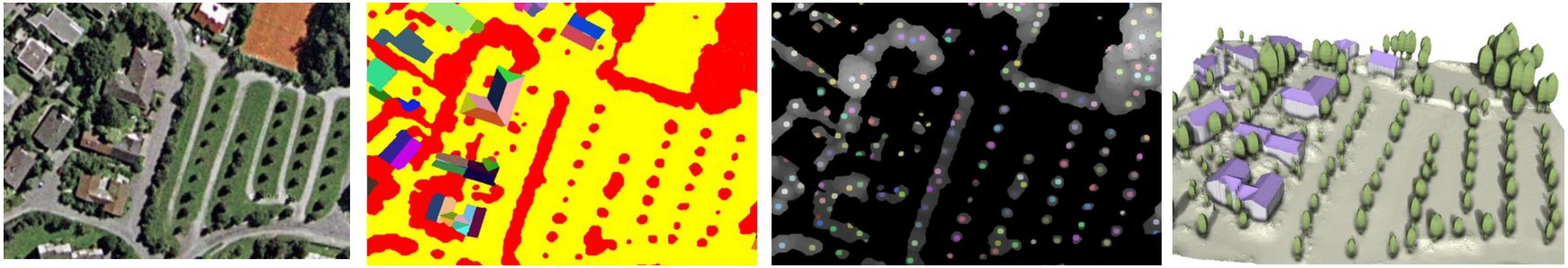
Buildings

- Hybrid repesentation (mesh+3D-primitives)



Trees

- Template matching (ellipsoid)

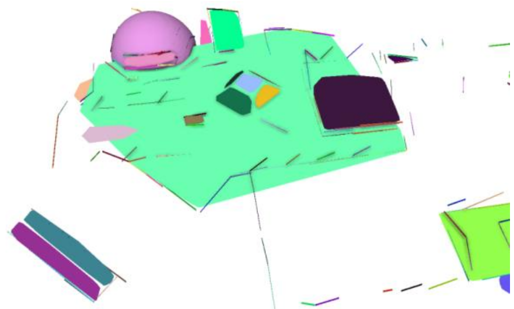


Ground

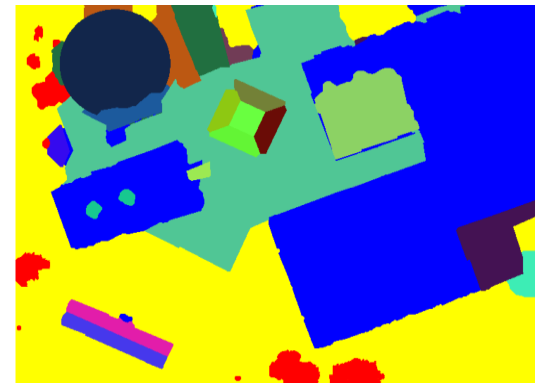
- mesh

Results

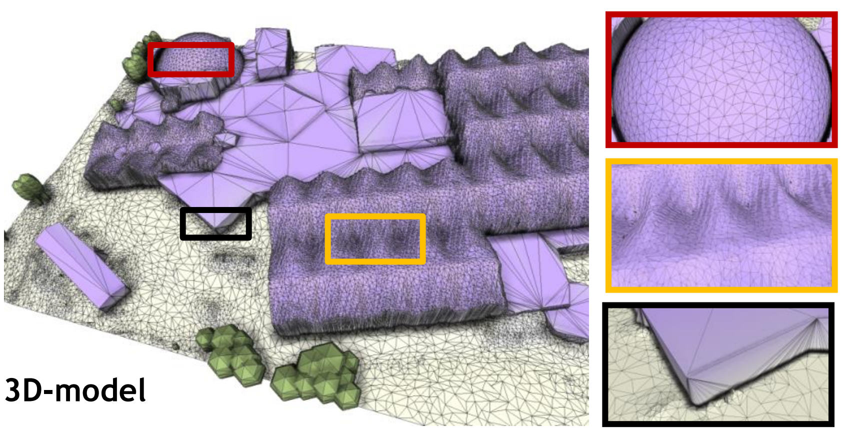
Aerial image



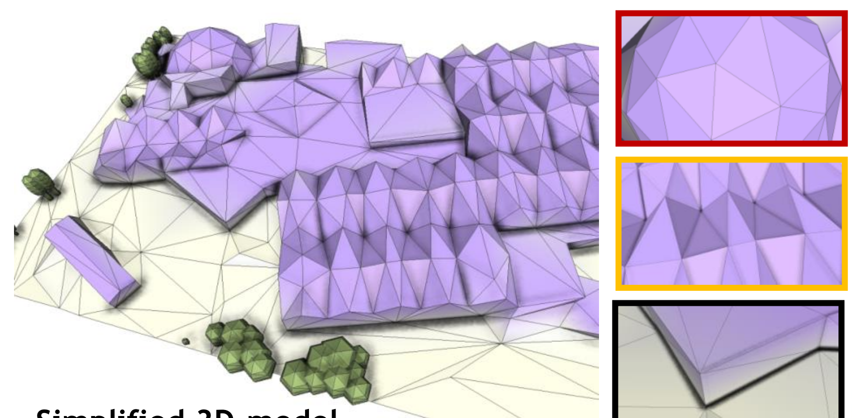
Extracted primitives



Label map



3D-model



Simplified 3D-model

Biberach, Germany (1km², 2.3M points)

Aerial image (from Google Maps)

Classified point cloud [color code: blue=building, red=vegetation, yellow=ground, white=clutter]

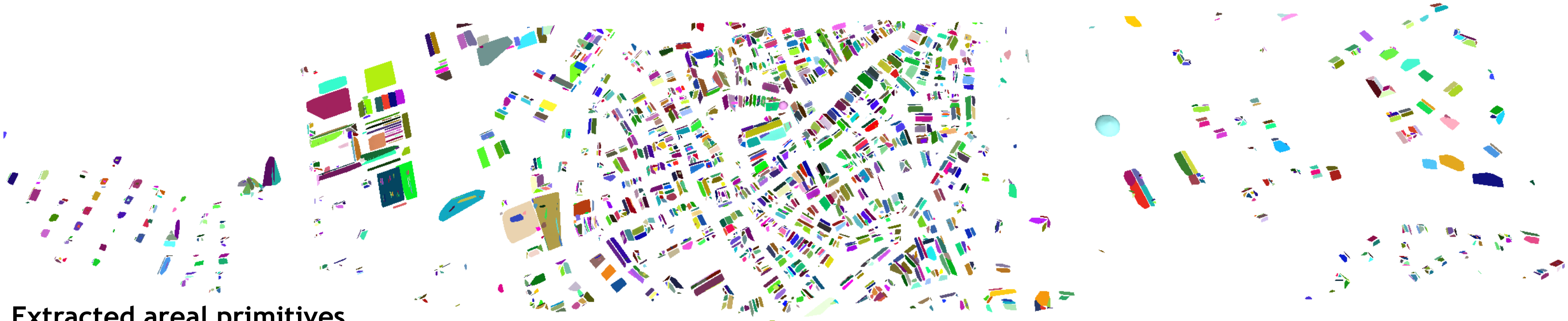
Biberach, Germany (1km², 2.3M points)

Aerial image (from Google Maps)

Extracted 3D-segments

Biberach, Germany (1km², 2.3M points)

Aerial image (from Google Maps)



Extracted areal primitives

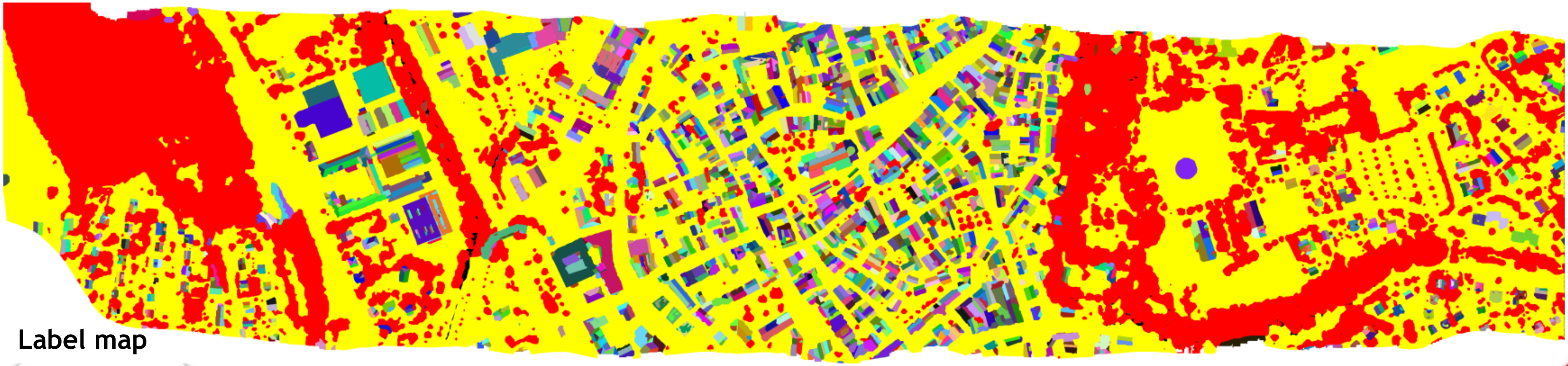
Biberach, Germany (1km², 2.3M points)

Aerial image (from Google Maps)

Block decomposition

Biberach, Germany (1km², 2.3M points)

Aerial image (from Google Maps)



Label map

Inria

Biberach, Germany (1km², 2.3M points)

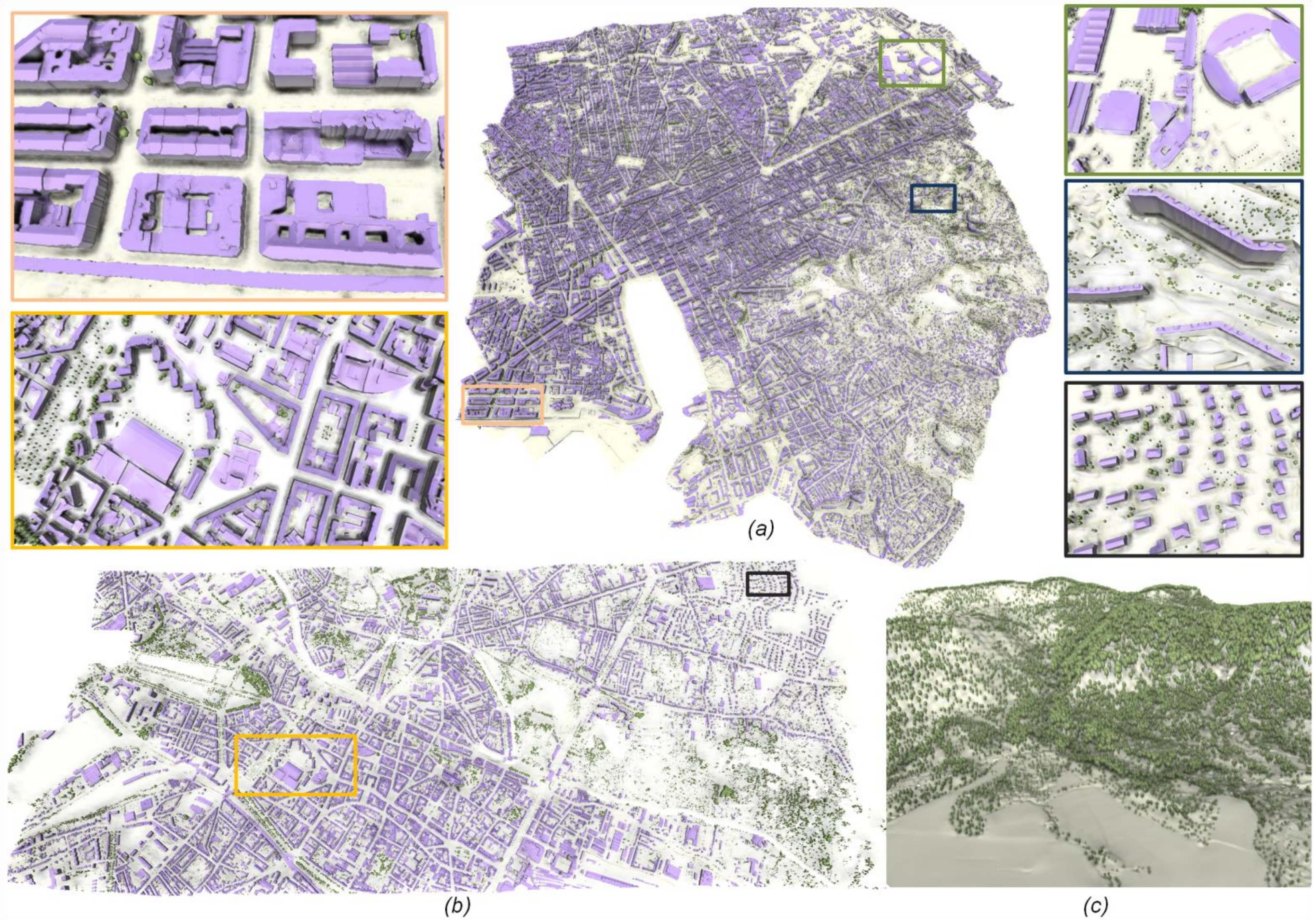
Aerial image (from Google Maps)

3D-model

Biberach, Germany (1km², 2.3M points)

Aerial image (from Google Maps)

3D-model with mesh visualization



	#input points ($\times 10^6$)	area (km^2)	altimetric variation (m)	#primitives ($\times 10^3$)	#trees ($\times 10^3$)	computing time (hour)	compaction (Mo)
Marseille, France *(a)*	38.67	19.8	192	108.6	35.7	2.52	131
Amiens, France *(b)*	24.52	11.57	76	56.7	22.8	1.34	93
Mountain area *(c)*	22.67	3.41	525	0.01	21.1	0.31	34

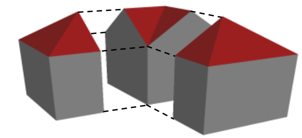
Example 2:

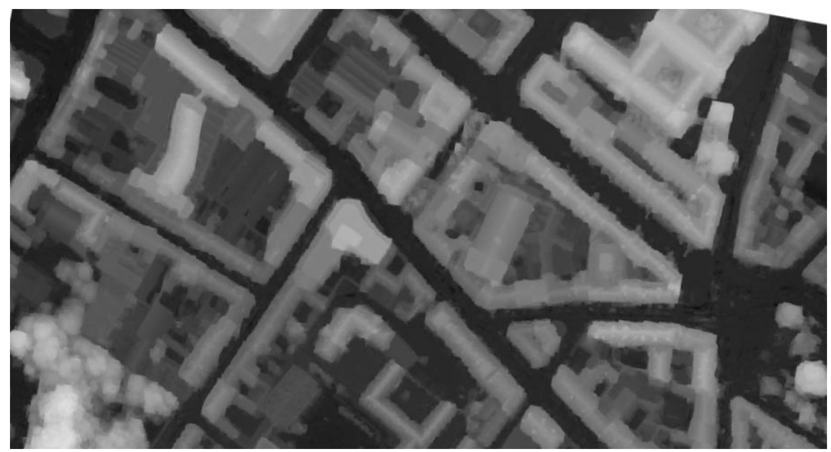
Reconstruction of buildings from MVS images

Overview

A/ Generation of a DEM from multi-view images

Structural concept
A building = an assemblage of elementary urban models

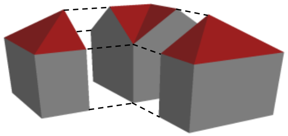


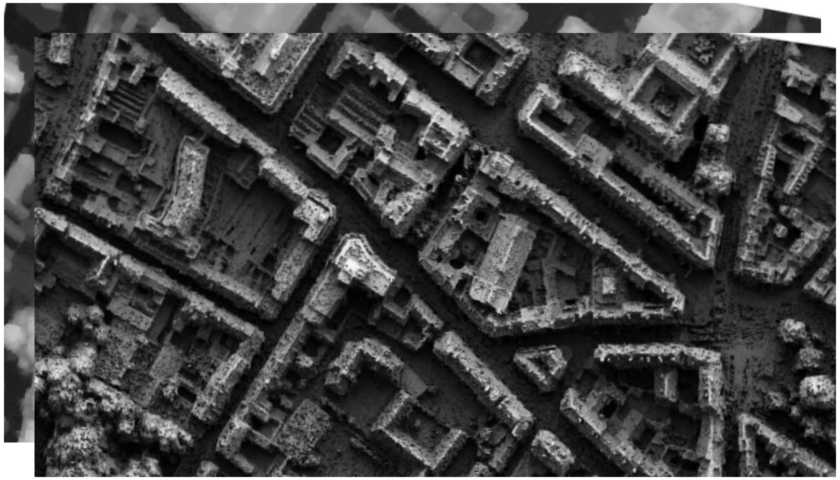


Overview

A/ Generation of a DEM from multi-view images

Structural concept
A building = an assemblage of elementary urban models

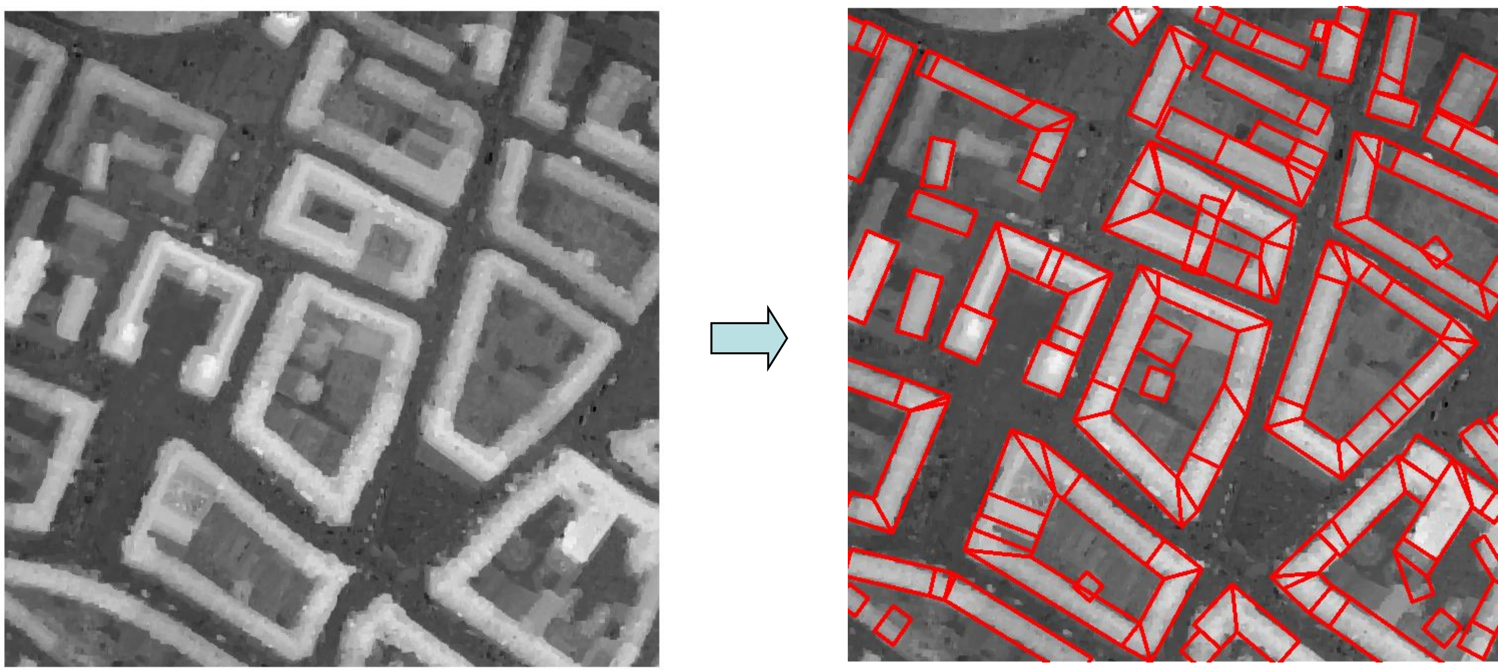




Overview

A/ Generation of a DEM from multi-view images

B/ 2D-extraction of building supports

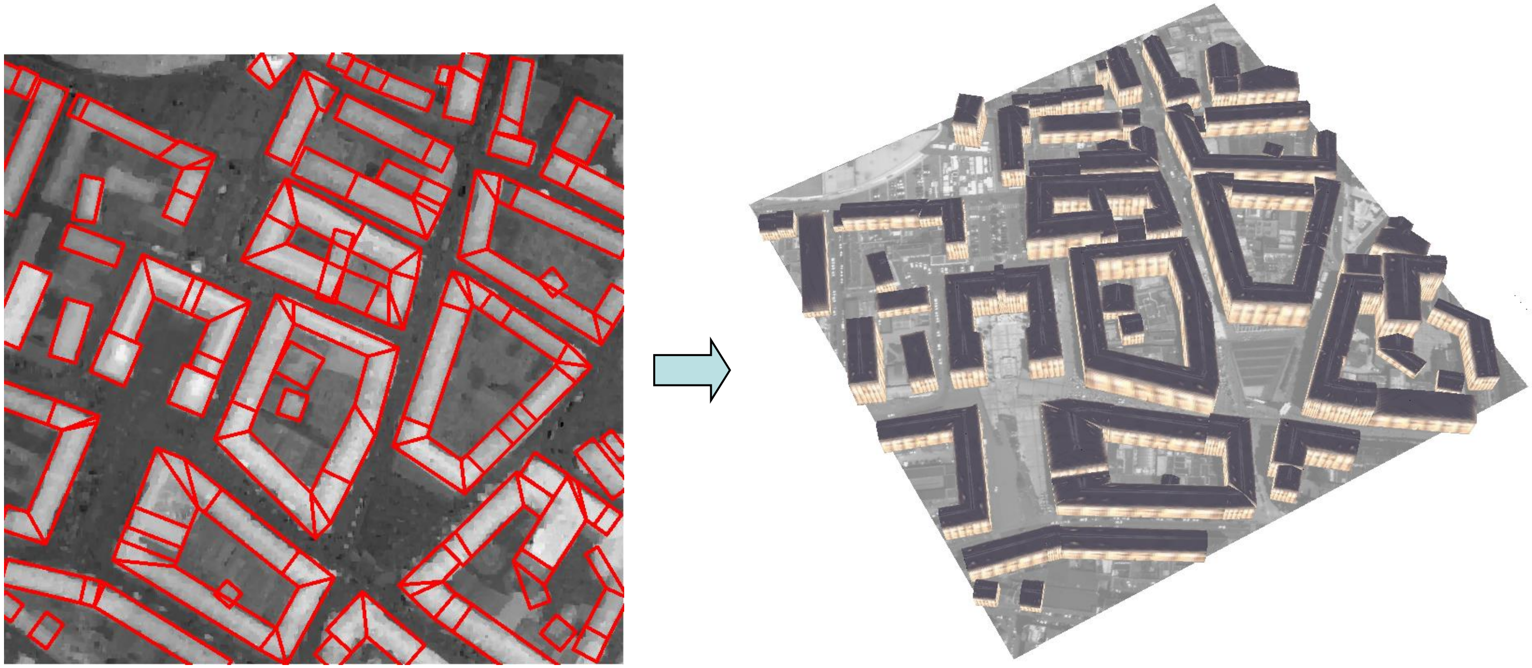


Overview

A/ Generation of a DEM from multi-view images

B/ 2D-extraction of building supports

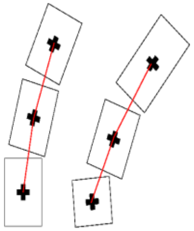
C/ 3D-reconstruction of buildings



Building footprint extraction

- configuration space: set of configurations of rectangles
- formulation of a energy: $U = \rho U_{ext} + U_{int}$
  - U_{ext} : data term
    coherence between a rectangle and the DEM discontinuities

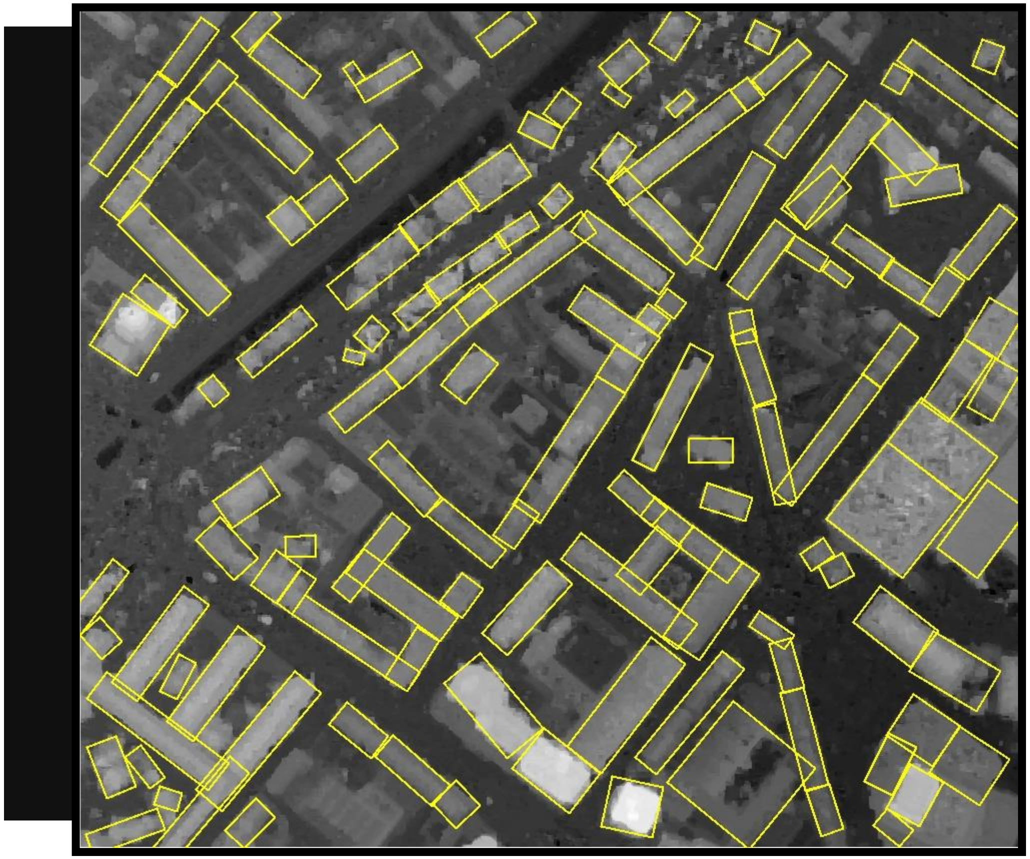
  - U_{int} : regularization term
    prior knowledge on the rectangle layout (alignment, paving, completion)



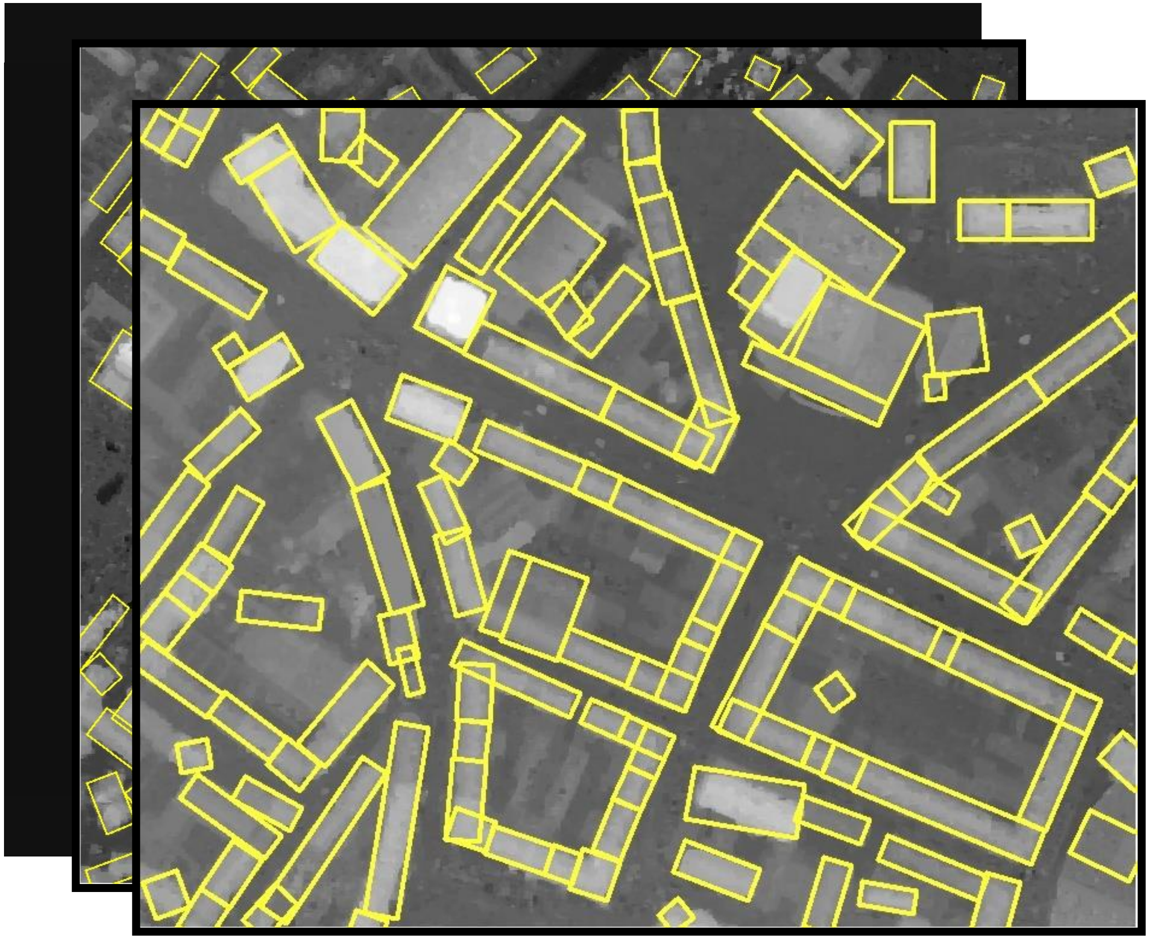
- optimization: MCMC with birth and death kernels

Building footprint extraction

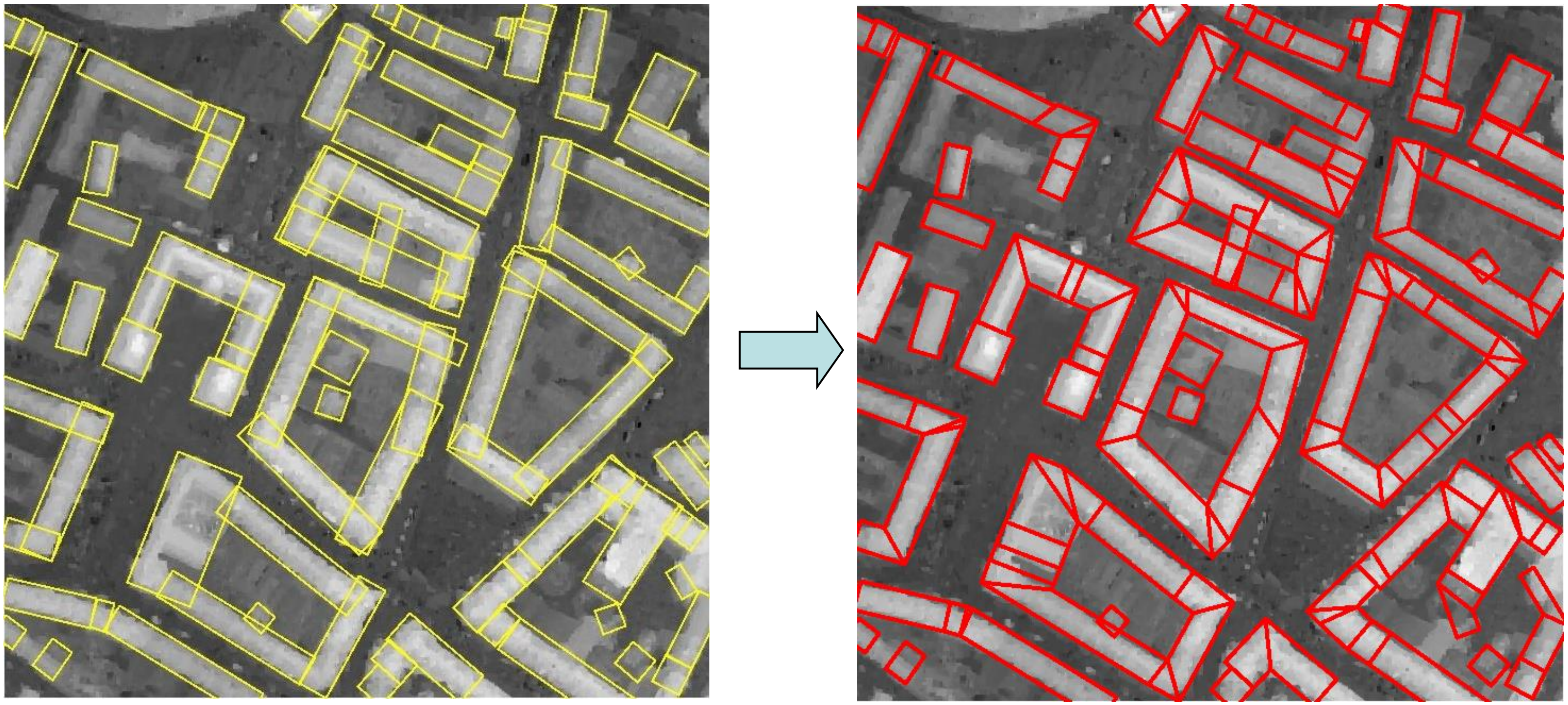
Building footprint extraction



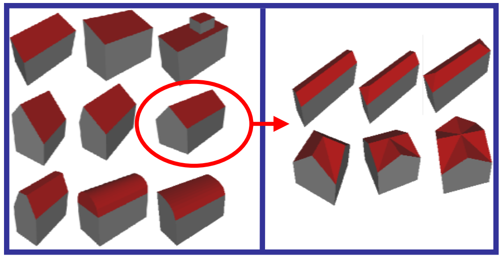
Building footprint extraction



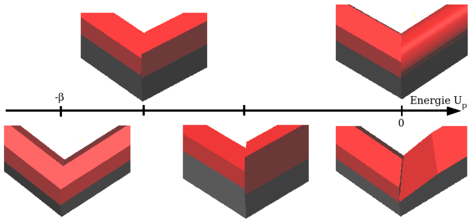
Building footprint extraction



Block assembling

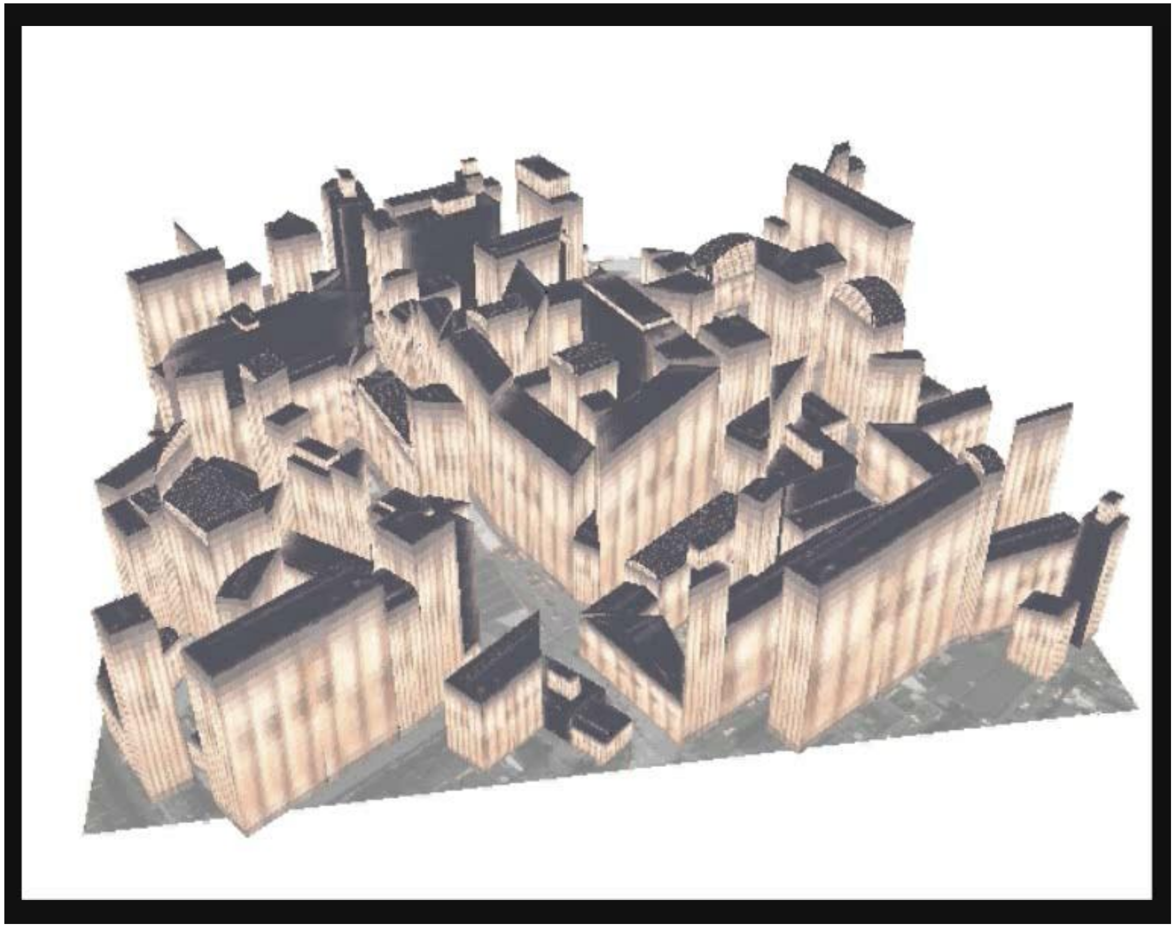


- configuration space : set of elementary urban structures
- formulation of a Bayesian energy
  - Likelihood
    coherence between the 3D-object and the DEM
  - Priors
    assemblage rules between neighboring urban objects (form of the roofs, connexion of rooftops…)



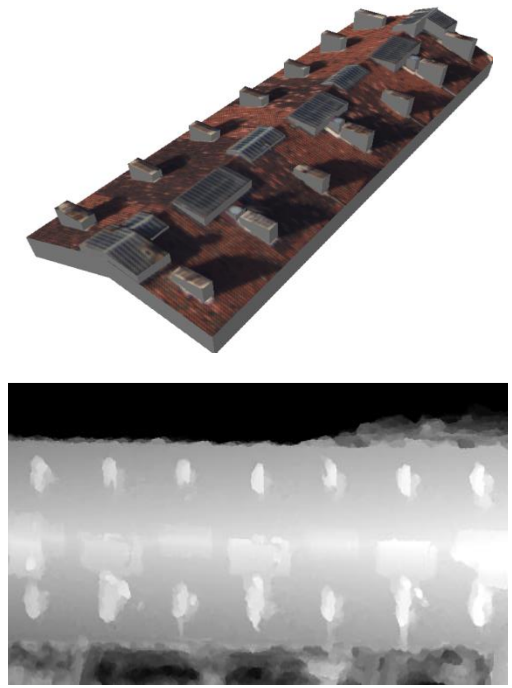
- optimization: MCMC

Block assembling

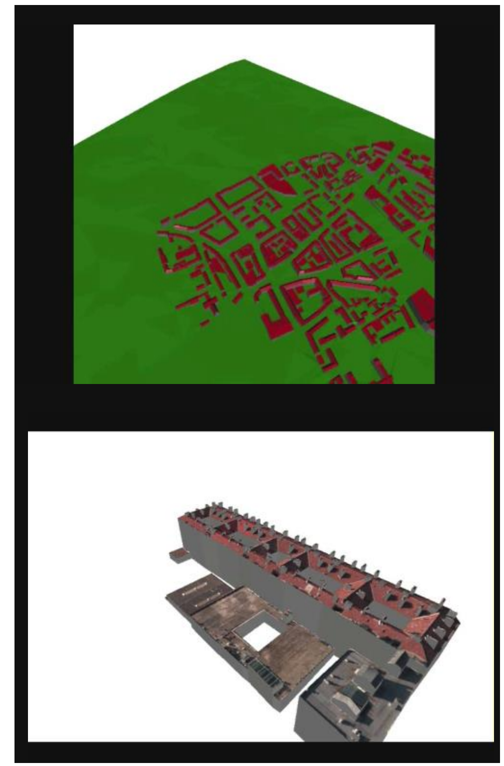


Results

0.70 m



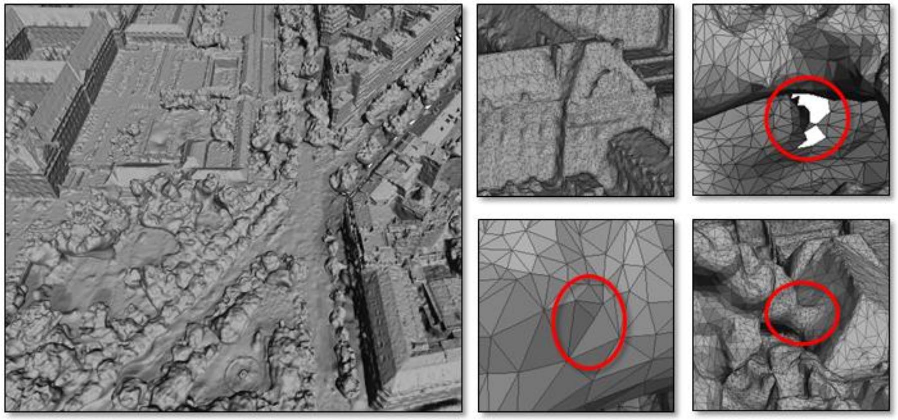
0.10 m



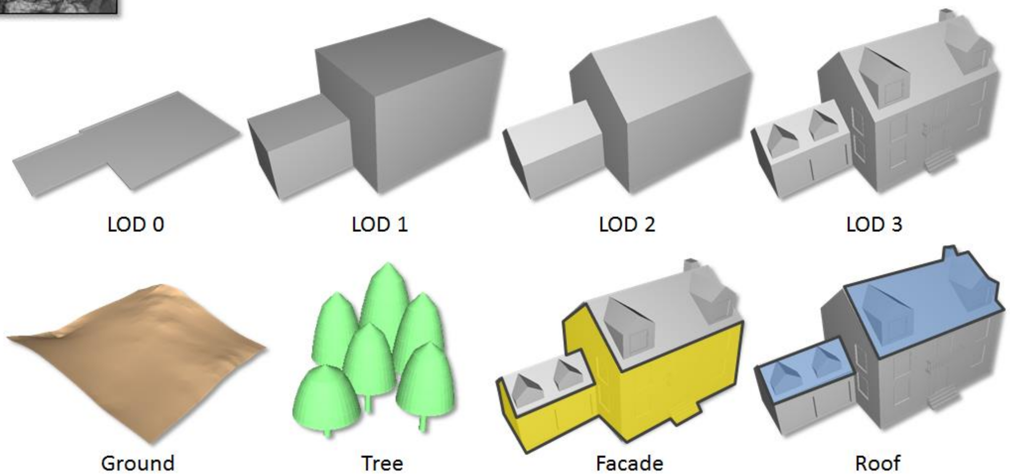
Example 3:

Generation of urban Levels Of Detail from raw meshes

Overview

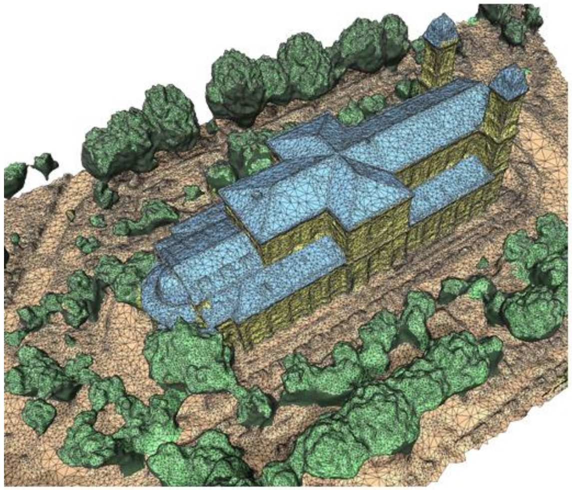


Input: raw meshes fom MVS

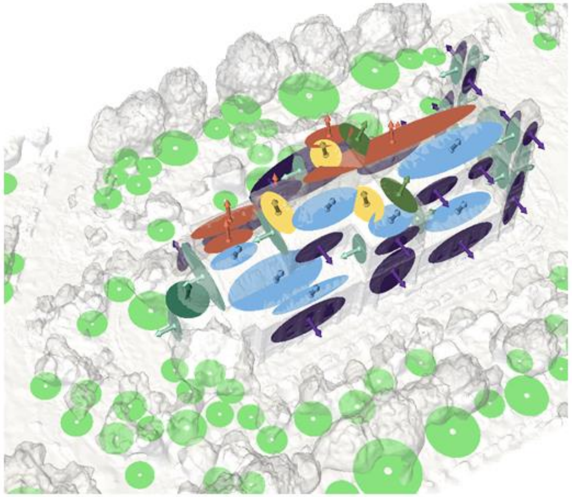


output: LOD models

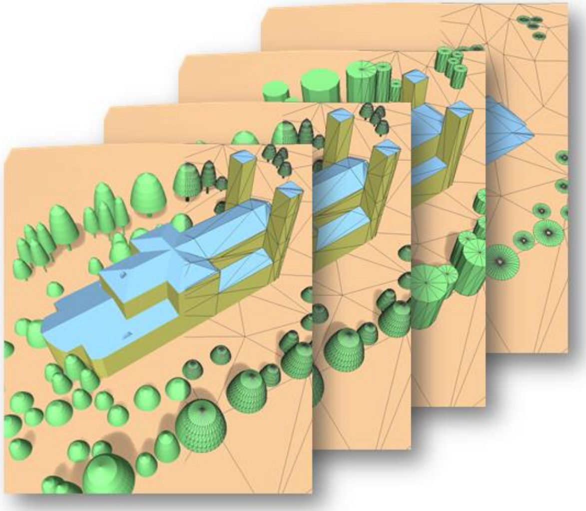
Overview



Classification



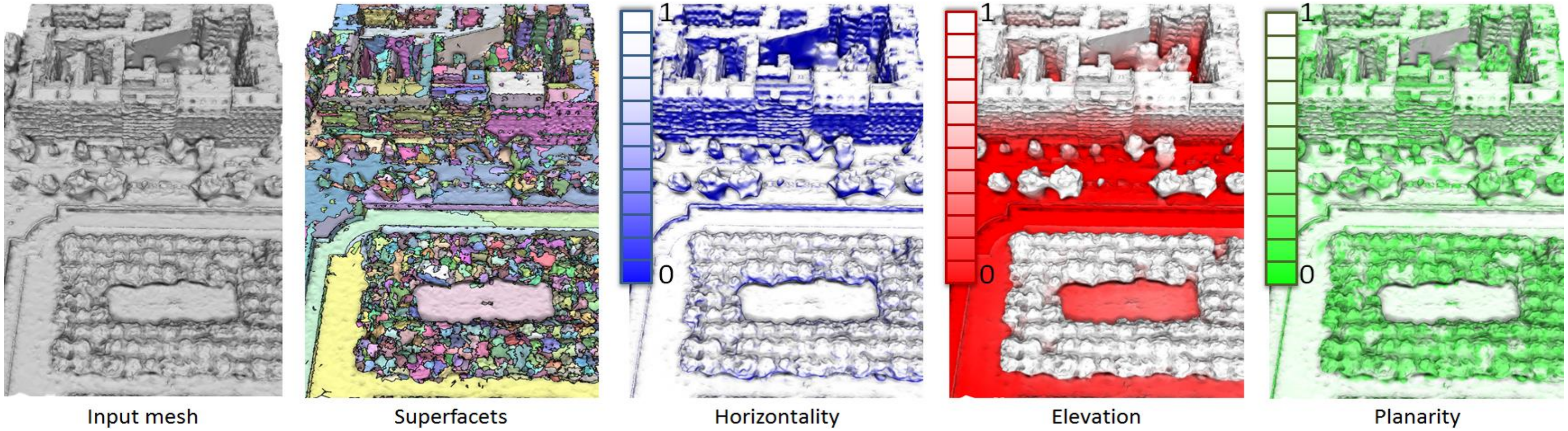
Abstraction



Reconstruction

Classification

Extracting relevant geometric attributes

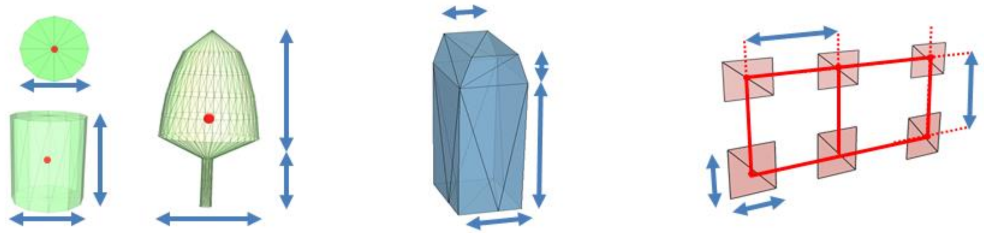


Labeling facets by MRF as
roof, facade, ground or trees

Geometric representations

Facades and roofs
→ planar proxies

Roof superstructures, facade components and trees
→ icons [Verdié2014]

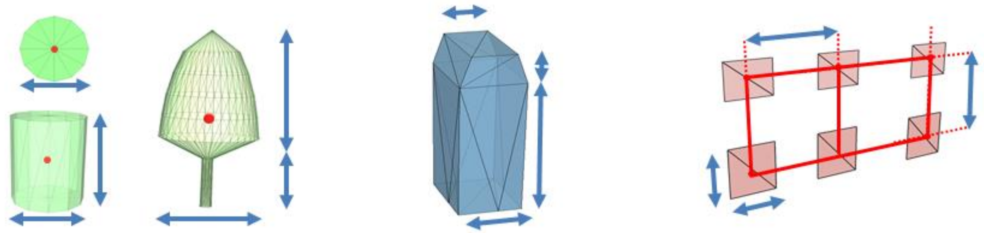


Ground
→ Delaunay triangulation

Geometric representations

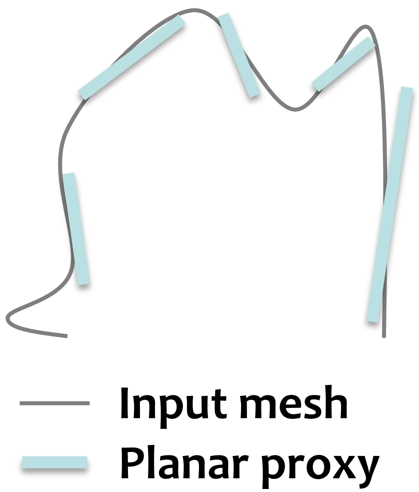
Facades and roofs
→ planar proxies

Roof superstructures, facade components and trees
→ icons [Verdié2014]

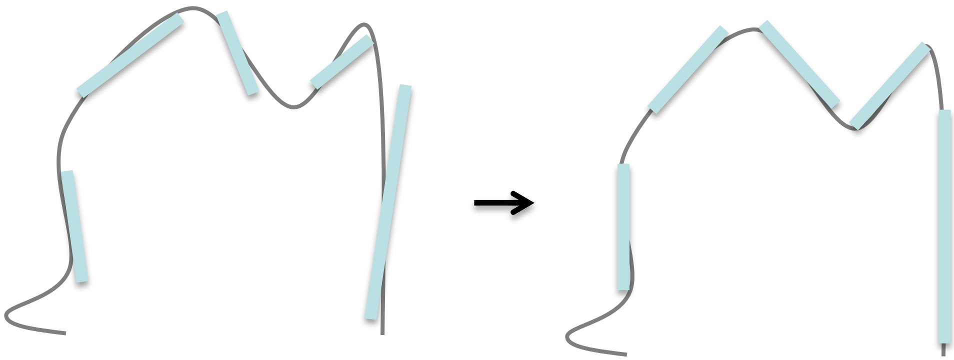


Ground
→ Delaunay triangulation

Detection and regularization of planar proxies



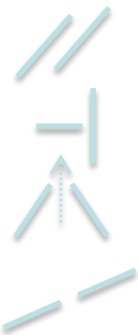
Detection and regularization of planar proxies



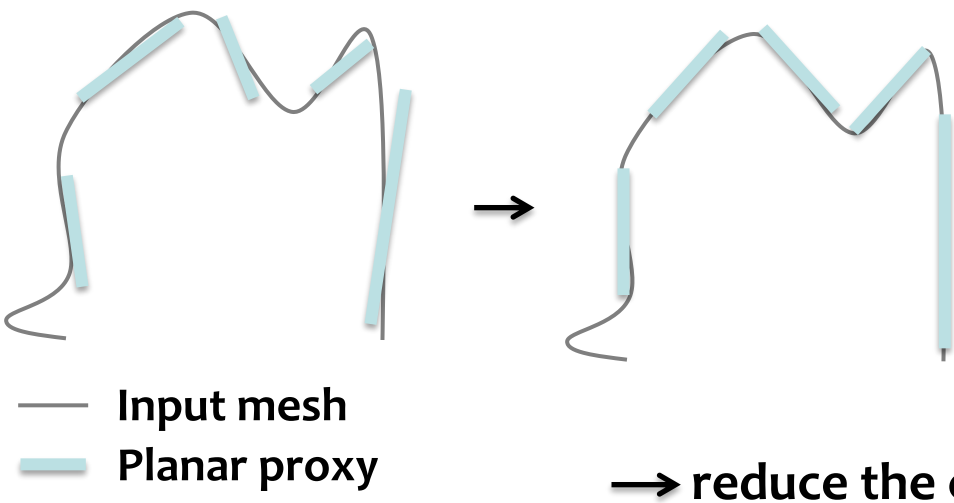
Input mesh
Planar proxy

Regularities

parallelism
orthogonality
Z-symmetry
coplanarity

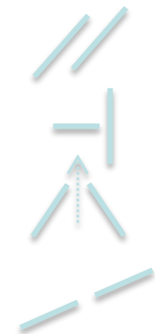


Detection and regularization of planar proxies



Regularities

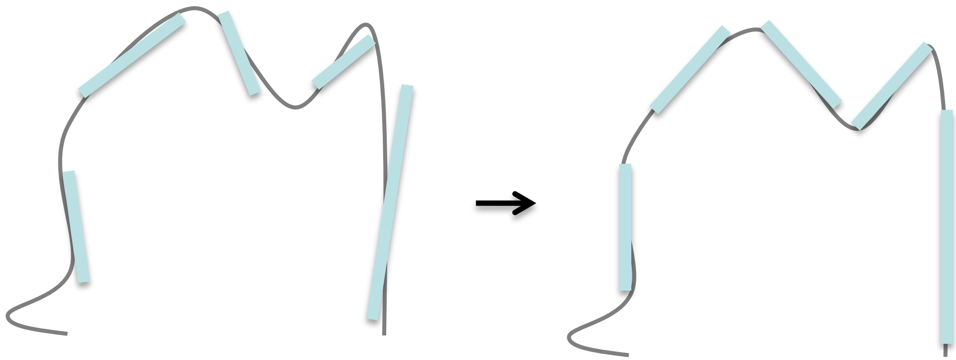
parallelism
orthogonality
Z-symmetry
coplanarity



→ reduce the complexity of subsequent reconstruction

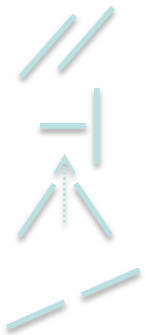
→ increase the visual quality of output surfaces

Detection and regularization of planar proxies



Input mesh
Planar proxy

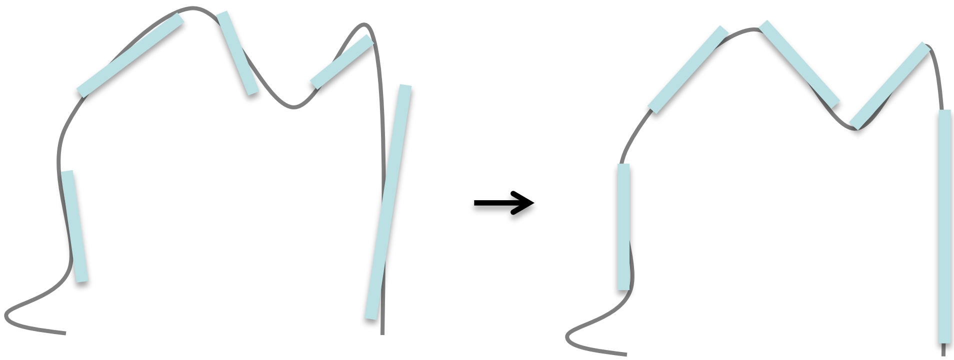
Regularities
parallelism
orthogonality
Z-symmetry
coplanarity



→ reduce the complexity of subsequent reconstruction
→ increase the visual quality of output surfaces

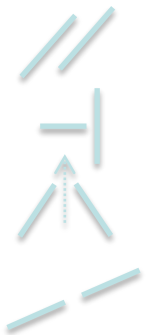
must be fast, scalable and urban-specific

Detection and regularization of planar proxies



Input mesh
Planar proxy

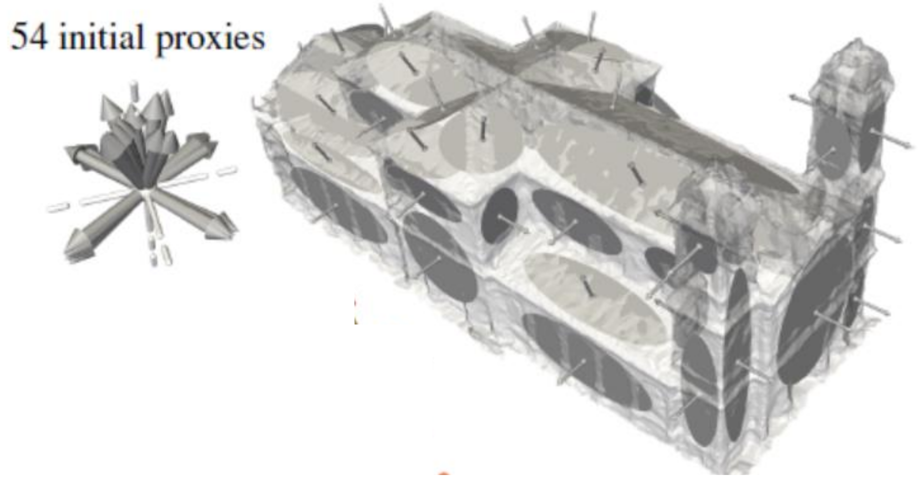
Regularities
parallelism
orthogonality
Z-symmetry
coplanarity



Idea: create a hierarchy between regularities within a *detection-then-regularization* approach

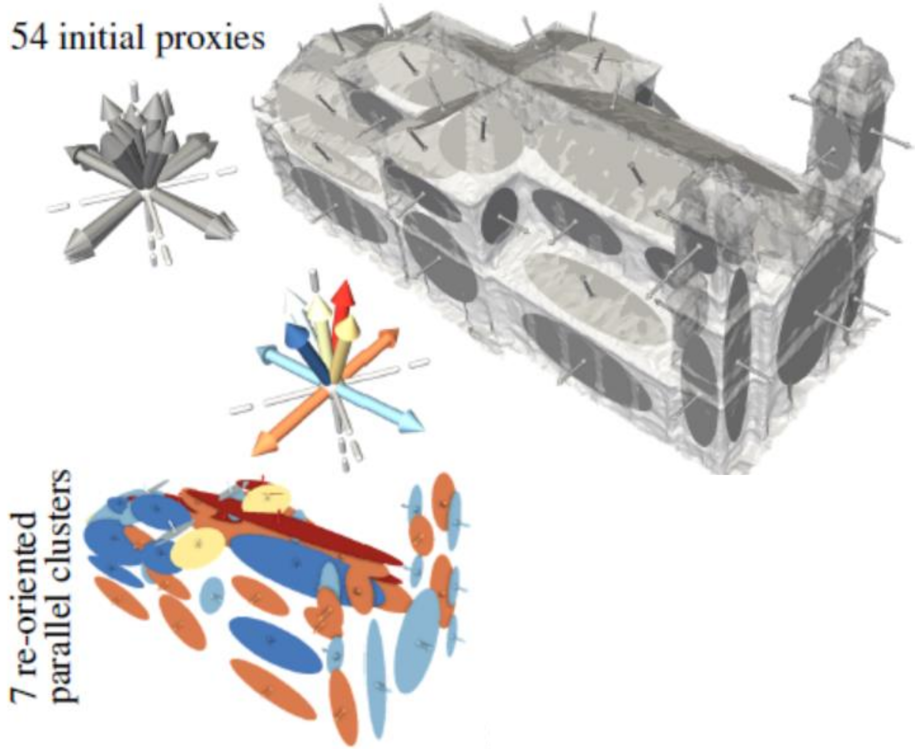
Detection and regularization of planar proxies

- **initial planar proxy from large superfacets**



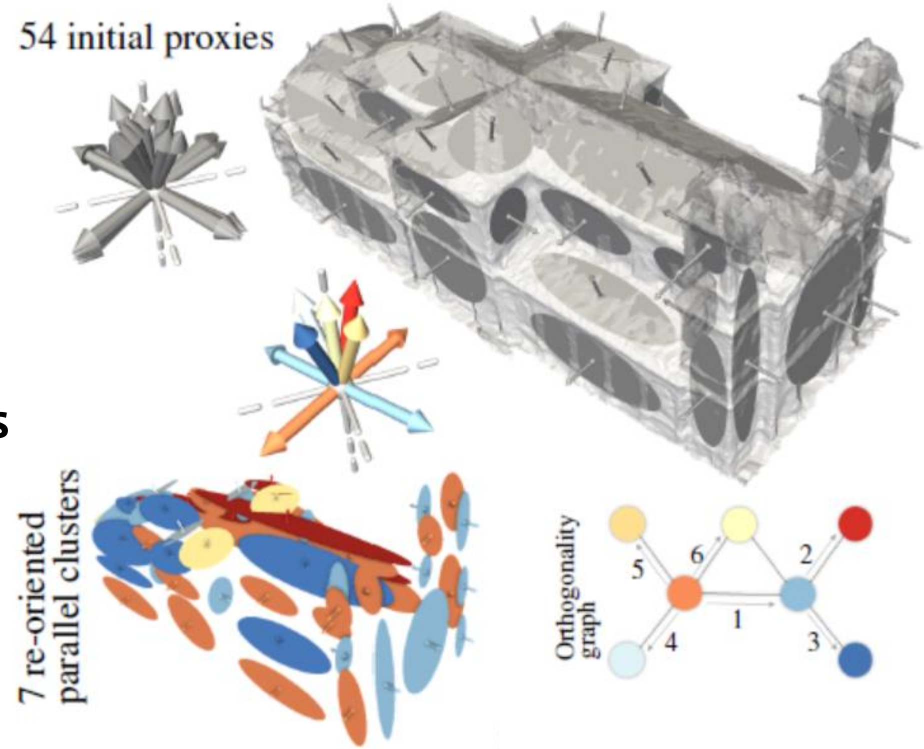
Detection and regularization of planar proxies

- **initial planar proxy from large superfacets**
- **grouping of proxies wrt parallelism**



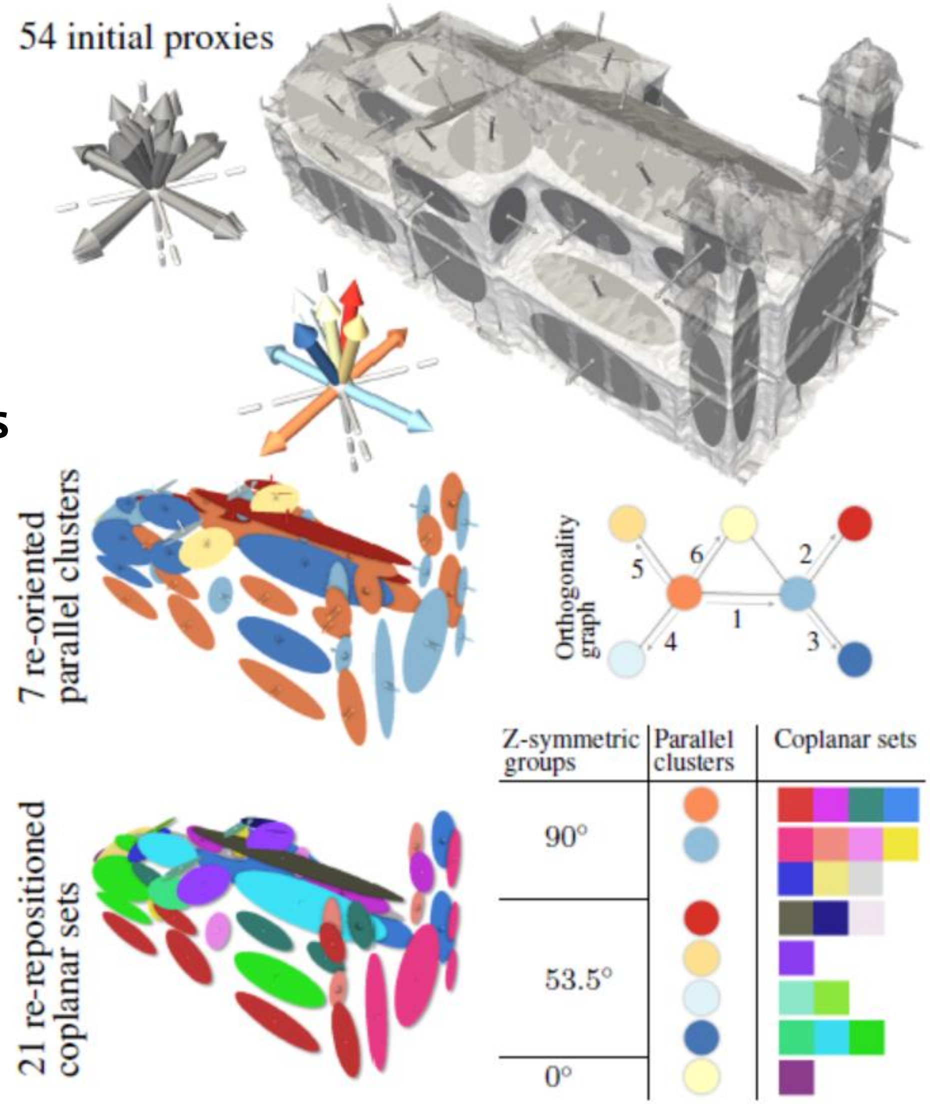
Detection and regularization of planar proxies

- initial planar proxy from large superfacets
- grouping of proxies wrt parallelism
- re-orientation parallel clusters wrt orthogonality and Z-symmetry

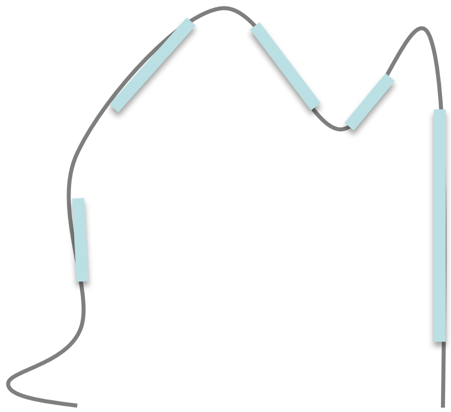


Detection and regularization of planar proxies

- initial planar proxy from large superfacets
- grouping of proxies wrt parallelism
- re-orientation parallel clusters wrt orthogonality and Z-symmetry
- Re-positioning of proxies wrt coplanarity



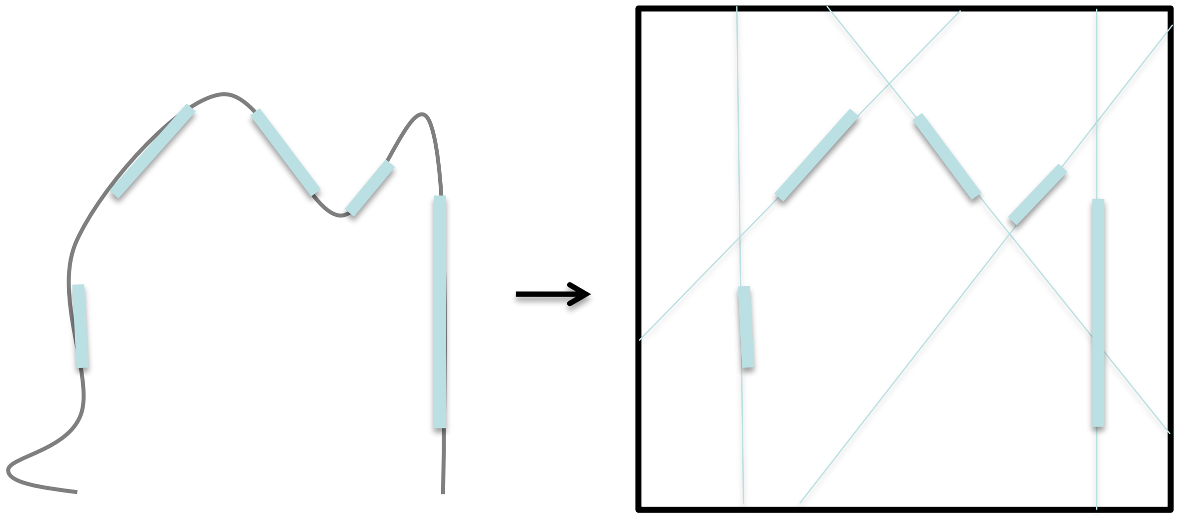
Discrete 3D arrangements



Input mesh
Planar proxy

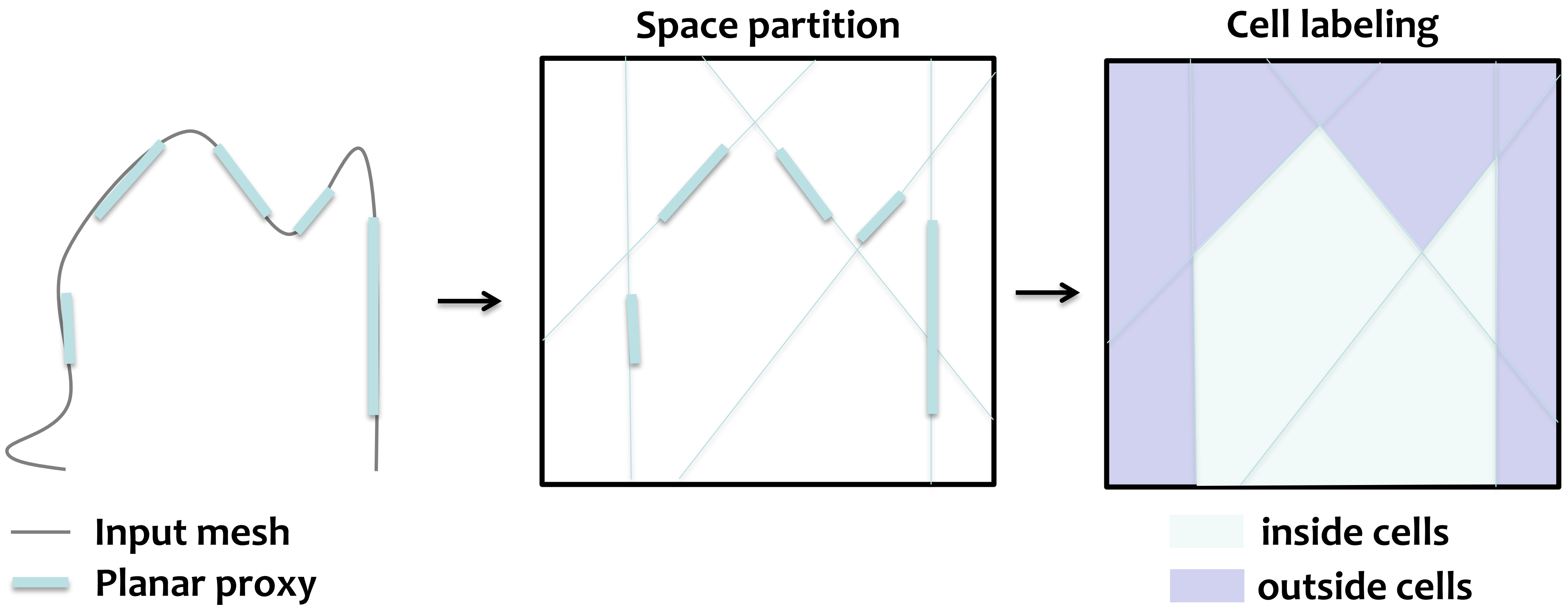
Discrete 3D arrangements

Space partition

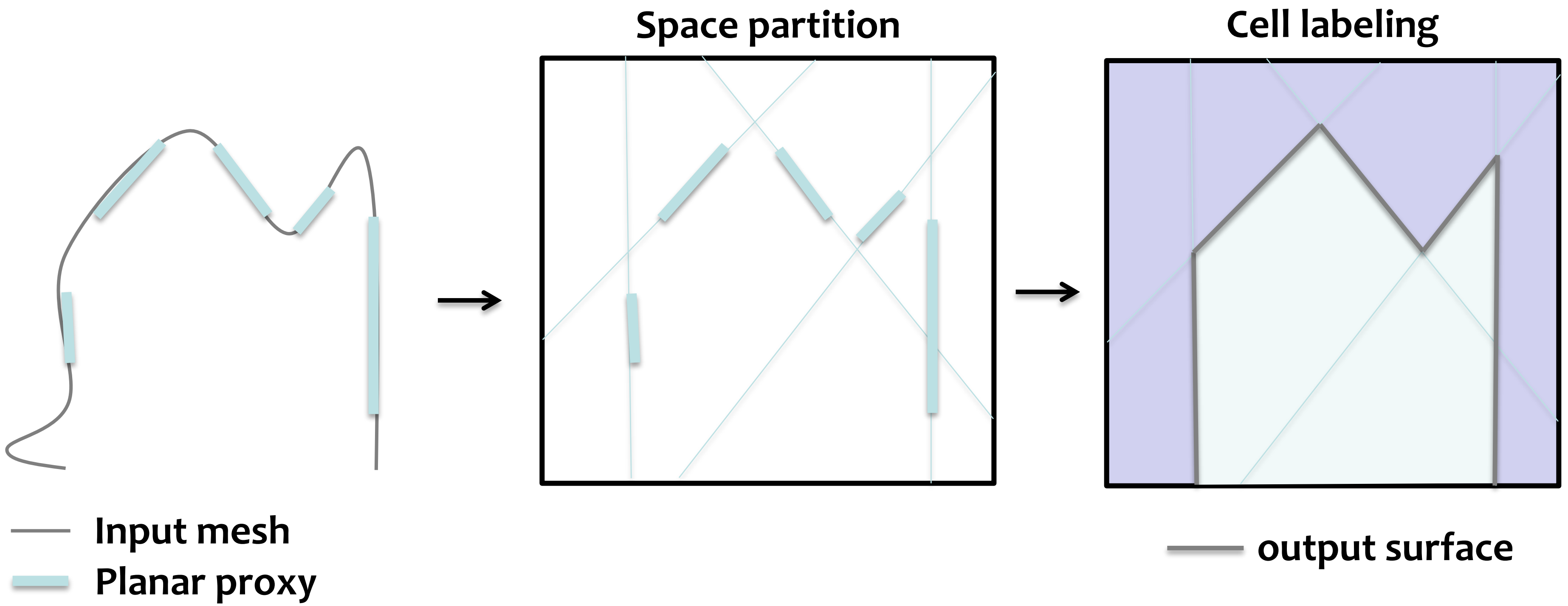


Input mesh
Planar proxy

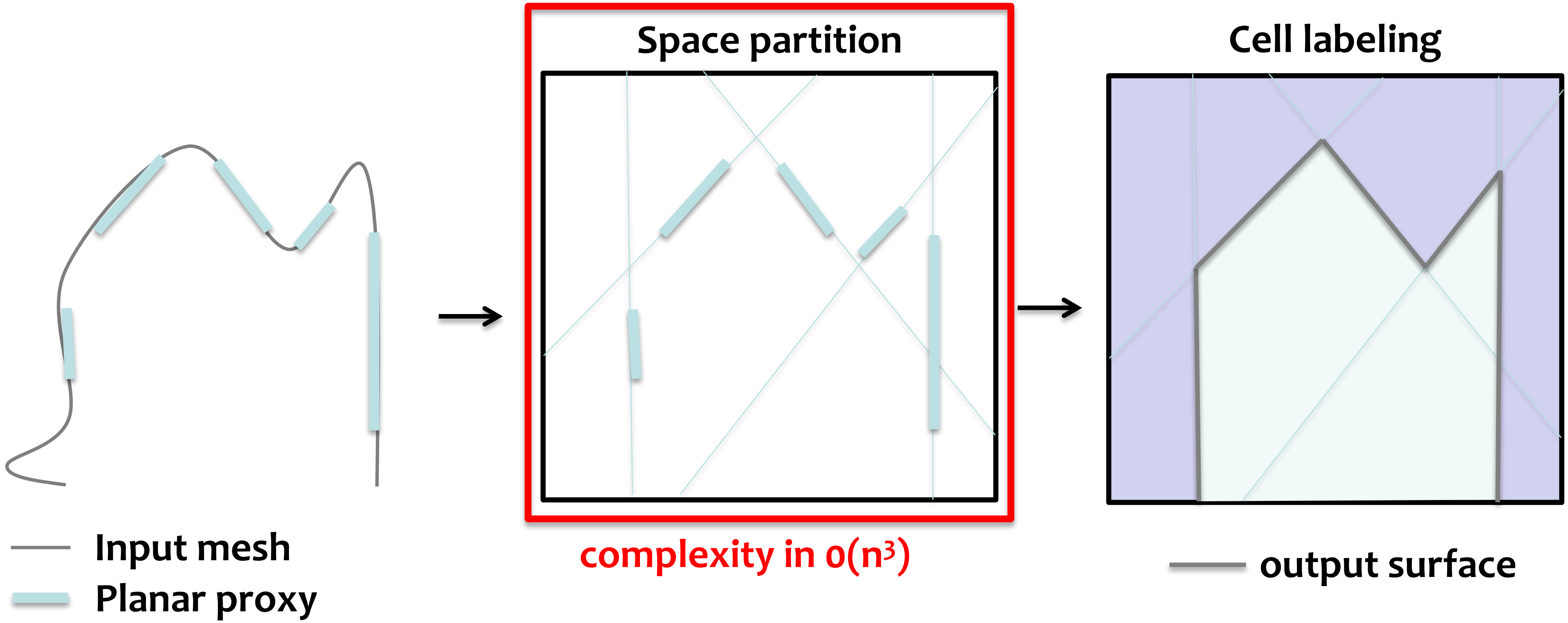
Discrete 3D arrangements



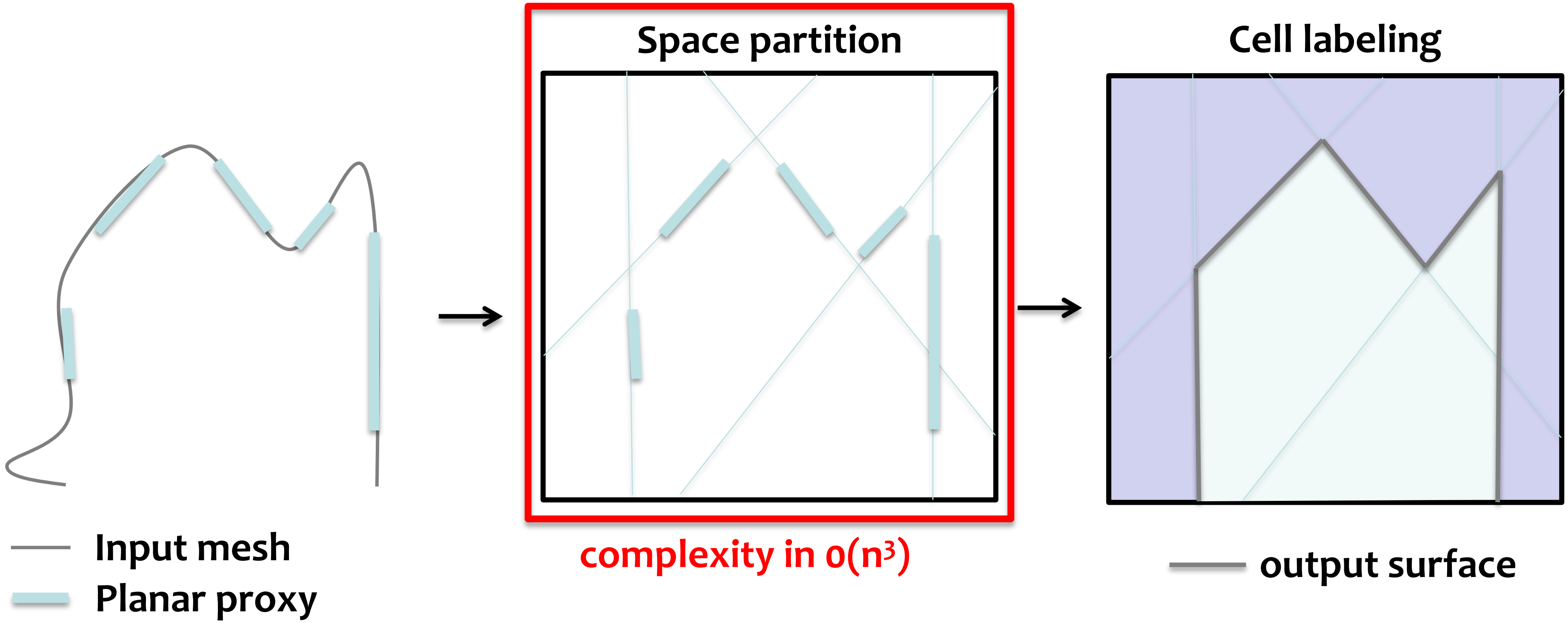
Discrete 3D arrangements



Discrete 3D arrangements



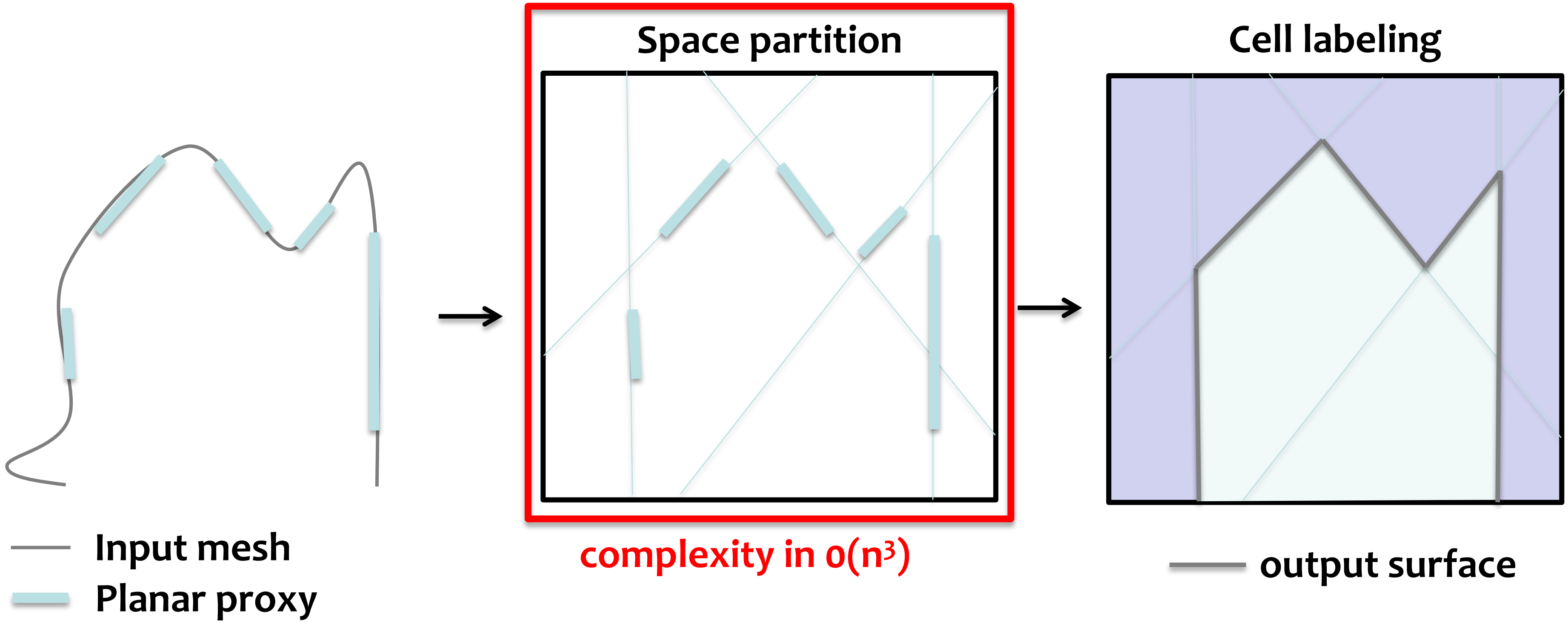
Discrete 3D arrangements



Use of strong geometric assumptions

- restriction to axis-aligned planar proxies [Furukawa2009]
- multi-layer of 2D arrangements [Oesau2014]
- convex polyhedral cell decomposition [Chauve2010]

Discrete 3D arrangements

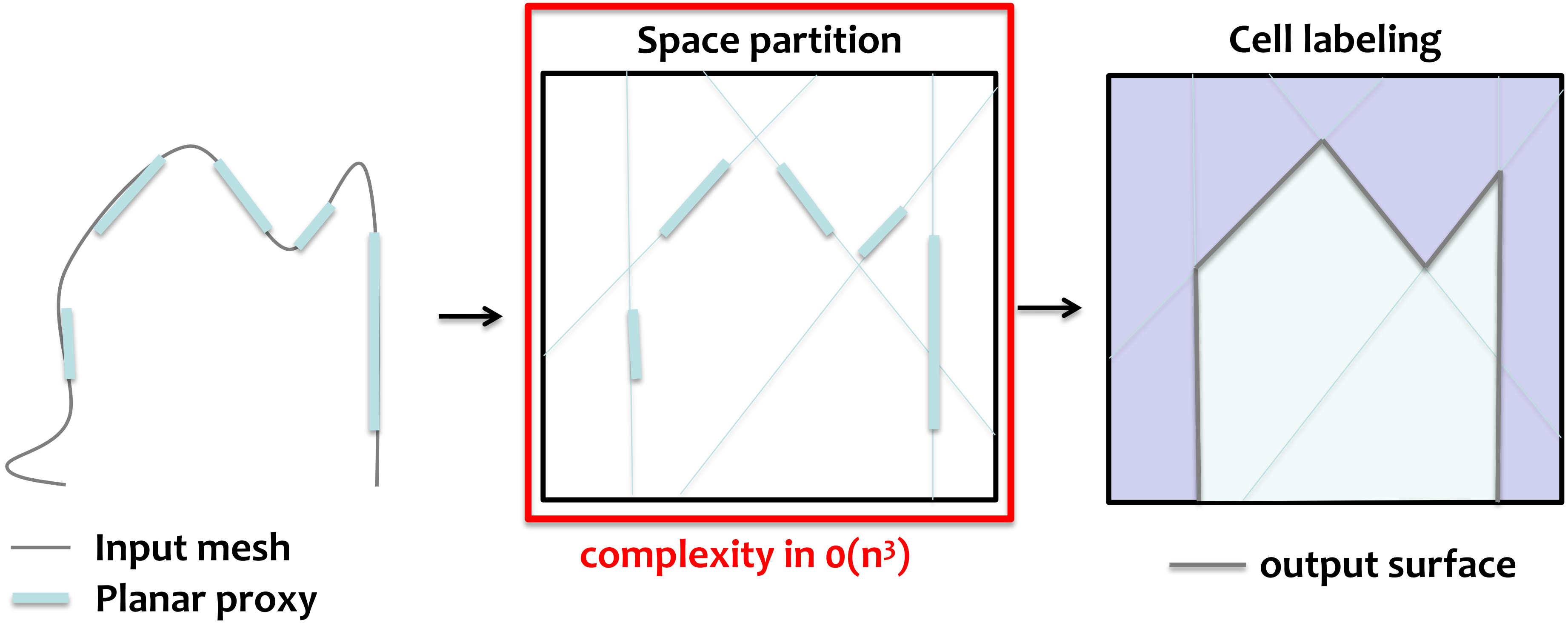


Use of strong geometric assumptions

- restriction to axis-aligned planar proxies [Furukawa2009]
- multi-layer of 2D arrangements [Oesau2014]
- convex polyhedral cell decomposition [Chauve2010]

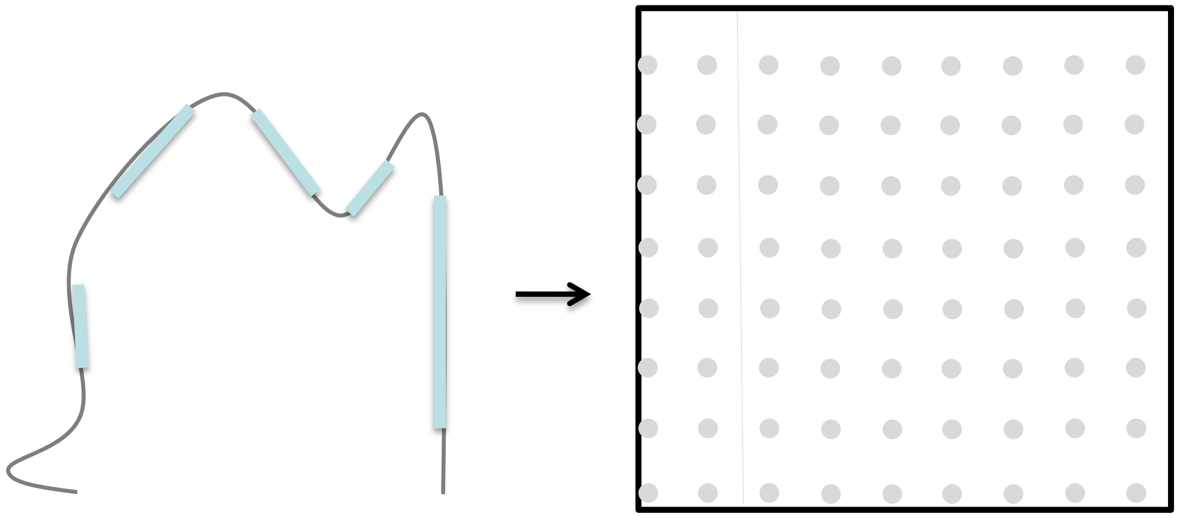
valid only in specifc cases

Discrete 3D arrangements



Idea: use a discrete partition to avoid computing the exact geometry

Discrete 3D arrangements

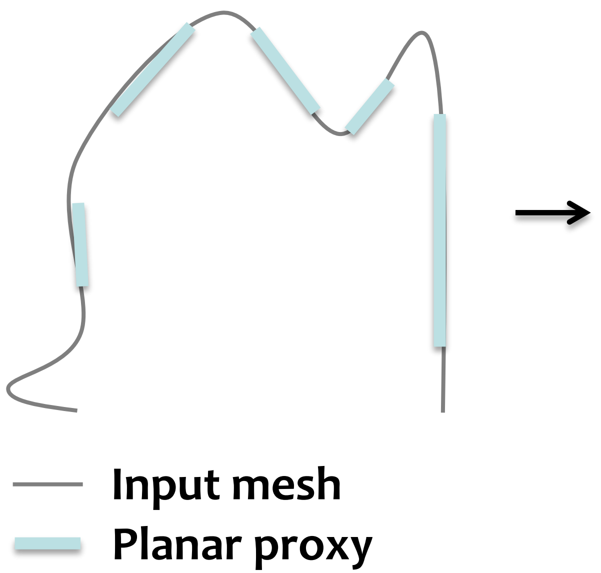


Input mesh
Planar proxy

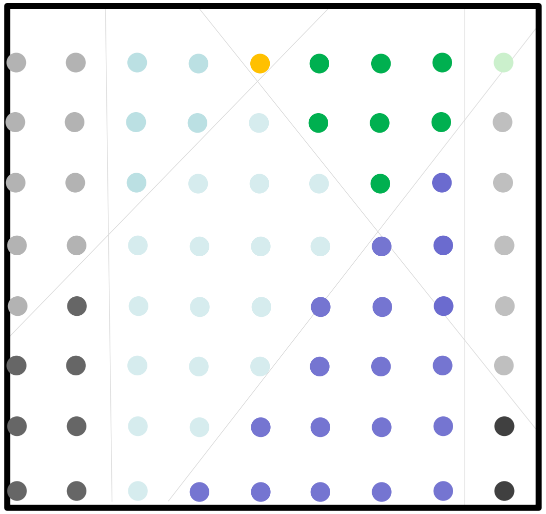
anchor

Idea: use a discrete partition to avoid computing the exact geometry

Discrete 3D arrangements



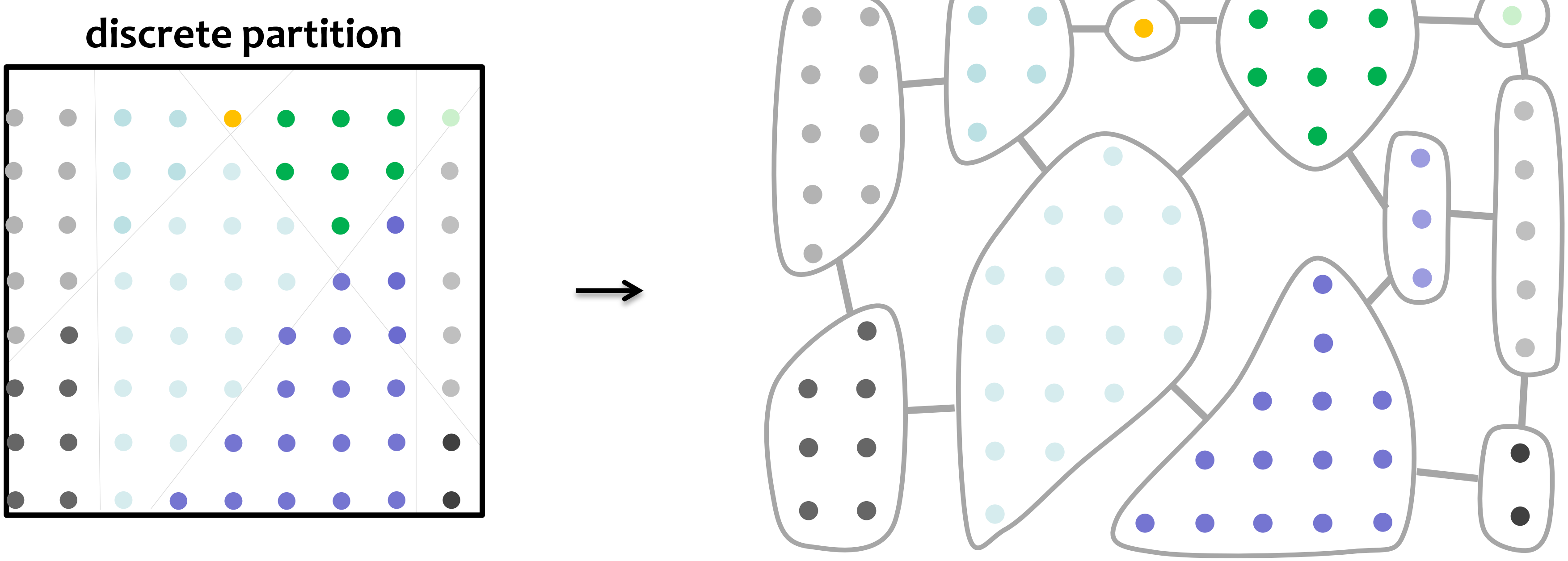
discrete partition



complexity in o(n log(k))

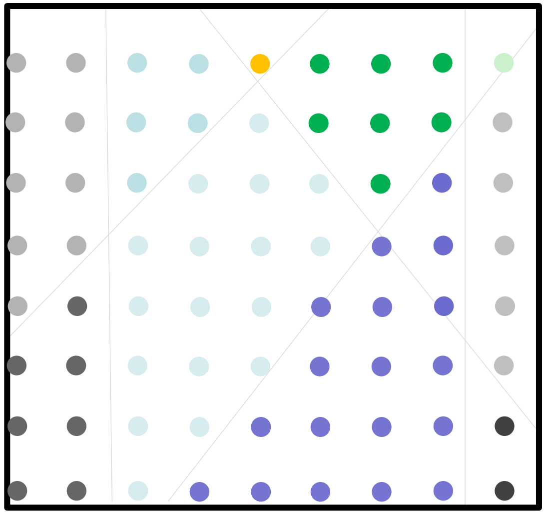
Idea: use a discrete partition to avoid computing the exact geometry

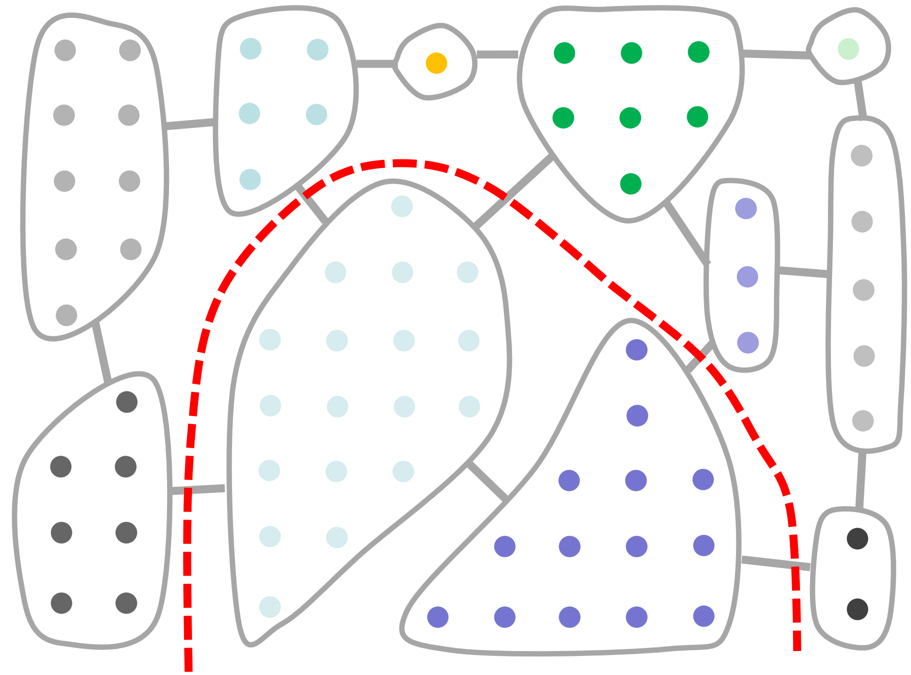
Min-cut formulation



Min-cut formulation

discrete partition



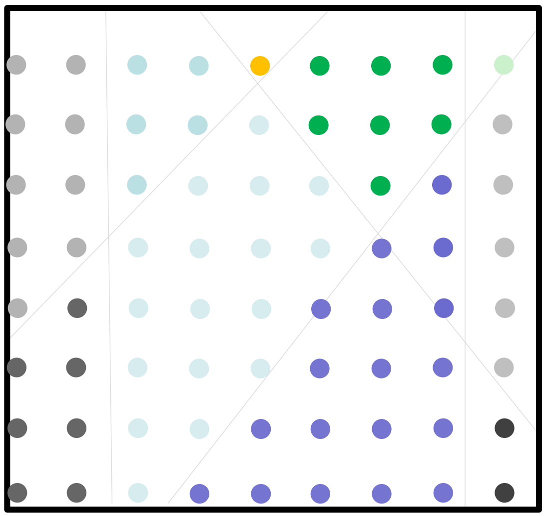


Quality of a cut S={C_{in},C_{out}}

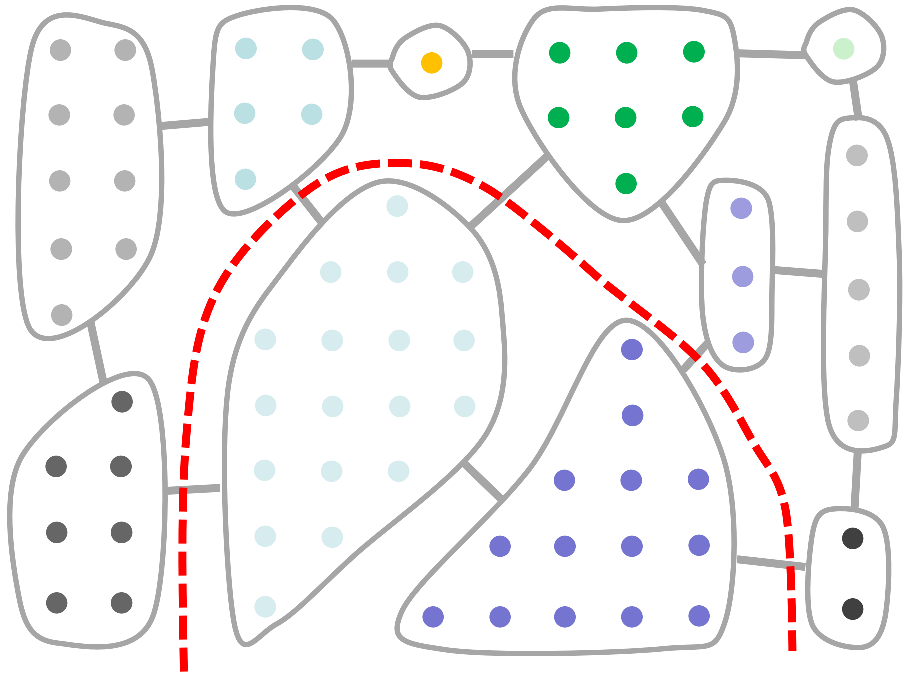
$$C(\mathcal{S}) = \sum_{c_k \in \mathcal{C}_{out}} V_{c_k}\, g(c_k) + \sum_{c_k \in \mathcal{C}_{in}} V_{c_k}\, (1 - g(c_k)) + \beta \sum_{f_i \in \mathcal{S}} A_{f_i}$$

Min-cut formulation

discrete partition



→

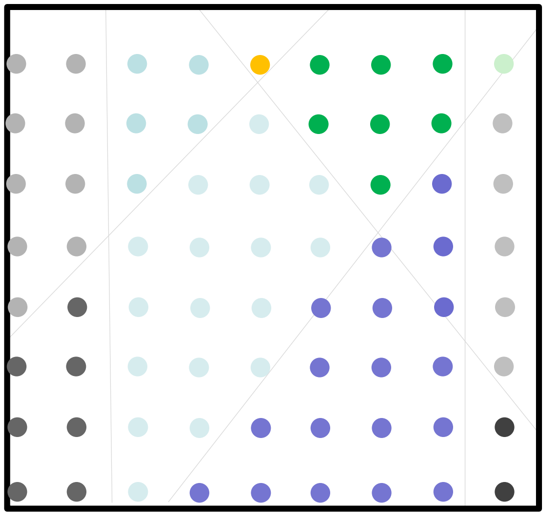


Quality of a cut S={C_{in},C_{out}}

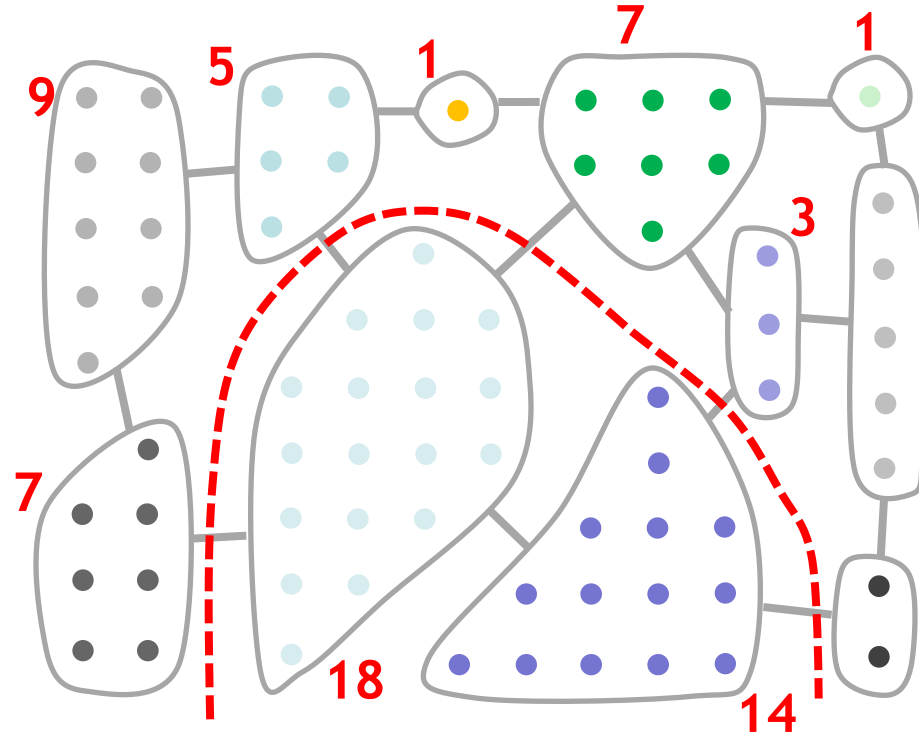
$$C(\mathcal{S}) = \underbrace{\sum_{c_k \in \mathcal{C}_{out}} V_{c_k}\, g(c_k) + \sum_{c_k \in \mathcal{C}_{in}} V_{c_k}\, (1 - g(c_k))}_{\text{Data term}} + \beta \sum_{f_i \in \mathcal{S}} A_{f_i}$$

Min-cut formulation

discrete partition



→



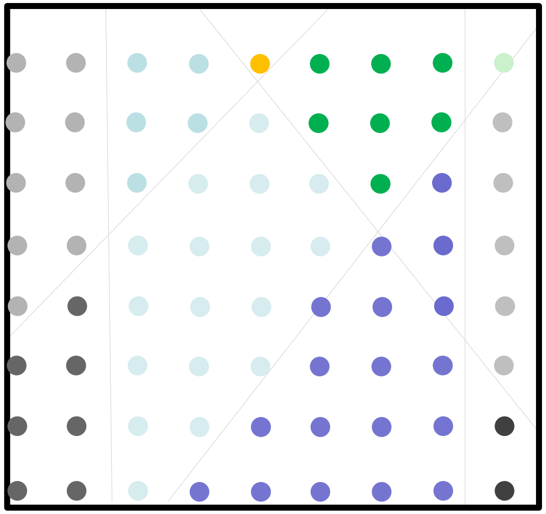
Quality of a cut $S=\{C_{in},C_{out}\}$

$$C(\mathcal{S}) = \underbrace{\sum_{c_k \in \mathcal{C}_{out}} \boxed{V_{c_k}} g(c_k) + \sum_{c_k \in \mathcal{C}_{in}} \boxed{V_{c_k}} (1-g(c_k))}_{\text{Data term}} + \beta \sum_{f_i \in \mathcal{S}} A_{f_i}$$

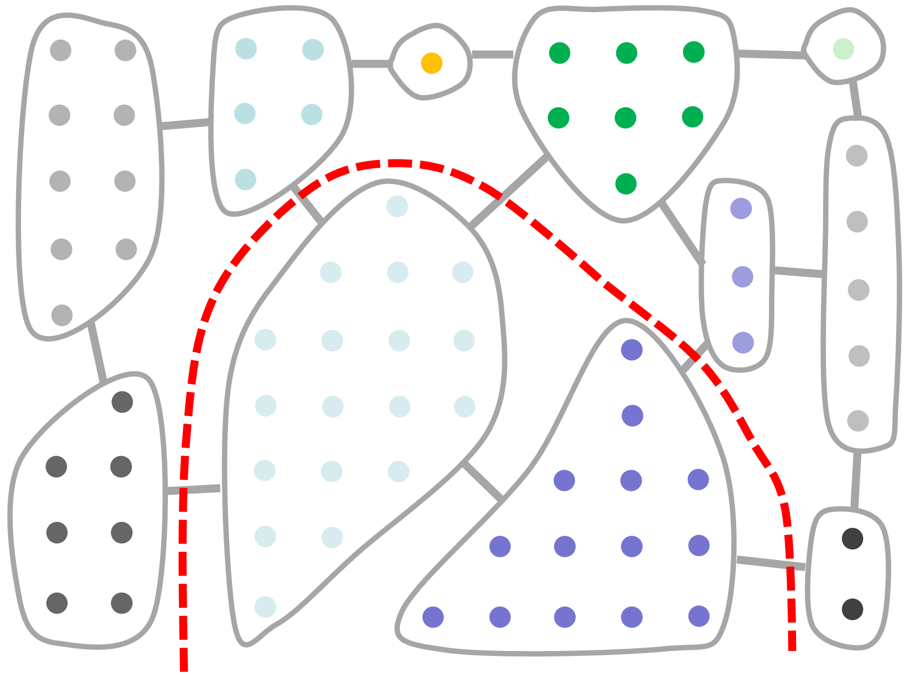
Data term

Min-cut formulation

discrete partition



→

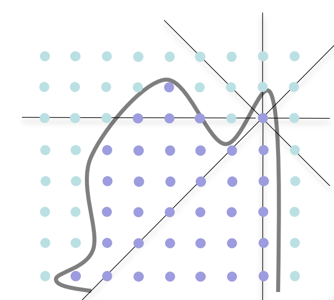


Quality of a cut $S=\{C_{in},C_{out}\}$

$$C(\mathcal{S}) = \sum_{c_k \in \mathcal{C}_{out}} V_{c_k} \boxed{g(c_k)} + \sum_{c_k \in \mathcal{C}_{in}} V_{c_k} (1 - \boxed{g(c_k)}) + \beta \sum_{f_i \in \mathcal{S}} A_{f_i}$$

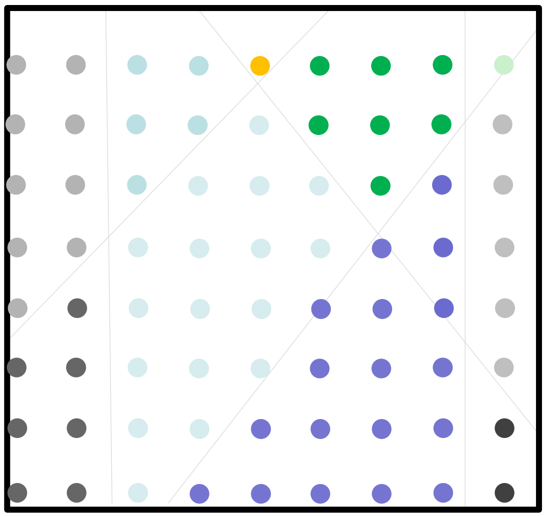
Data term

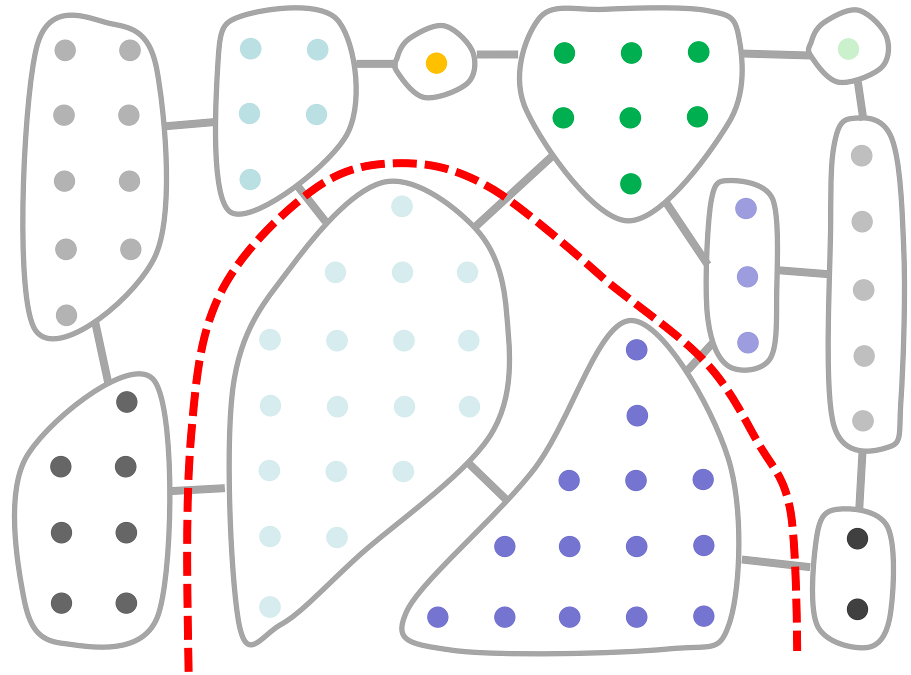
inside/outside prediction function estimated by ray casting



Min-cut formulation

discrete partition

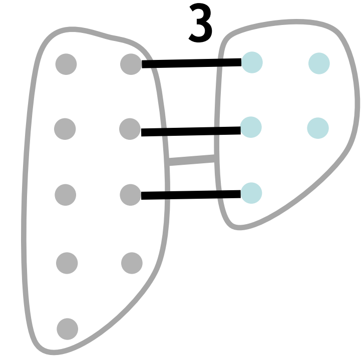




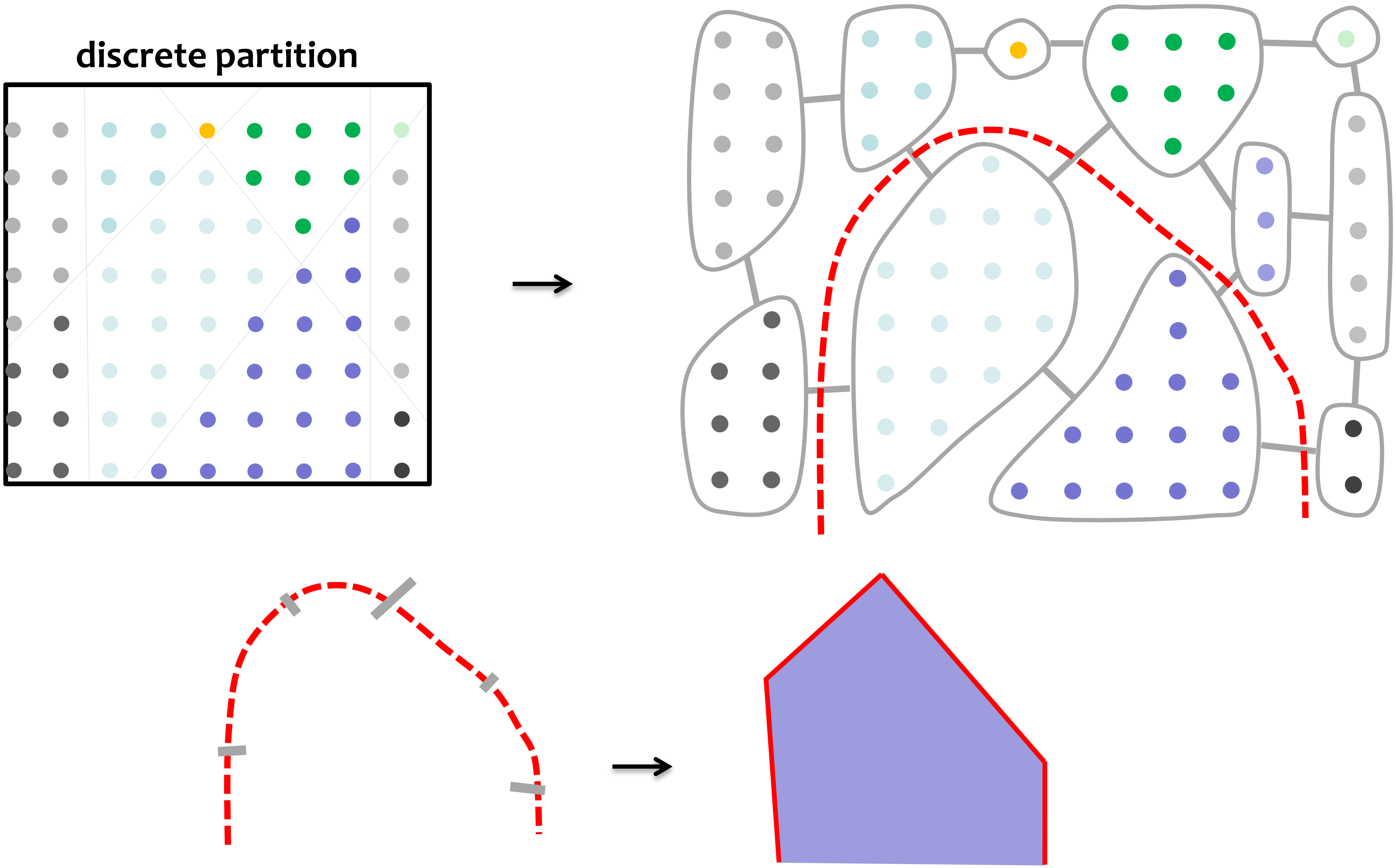
Quality of a cut S={C_{in},C_{out}}

$$C(\mathcal{S}) = \sum_{c_k \in \mathcal{C}_{out}} V_{c_k}\, g(c_k) + \sum_{c_k \in \mathcal{C}_{in}} V_{c_k}\, (1 - g(c_k)) + \beta \sum_{f_i \in \mathcal{S}} A_{f_i}$$

complexity

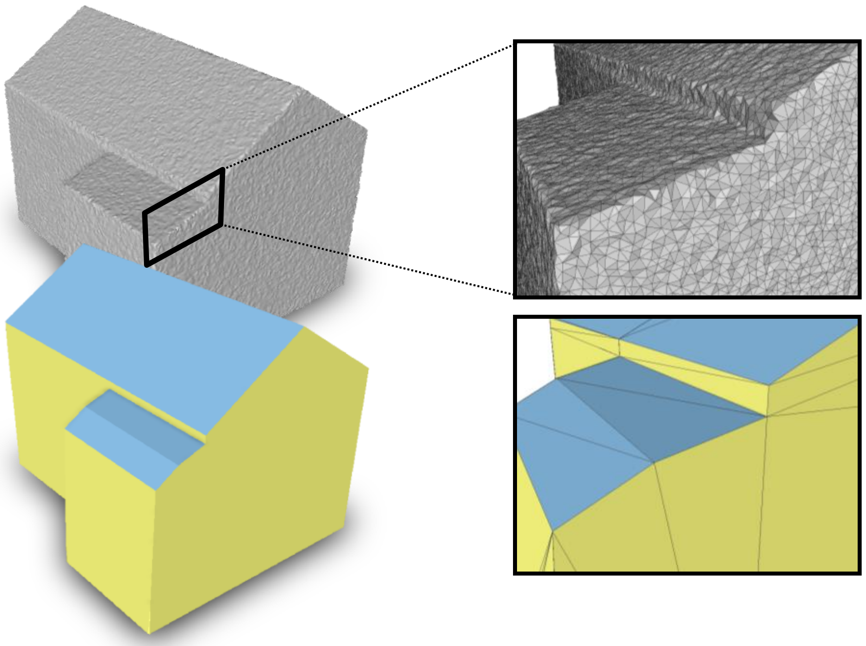


Min-cut formulation

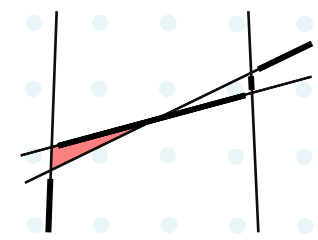


Anchor spacing setting

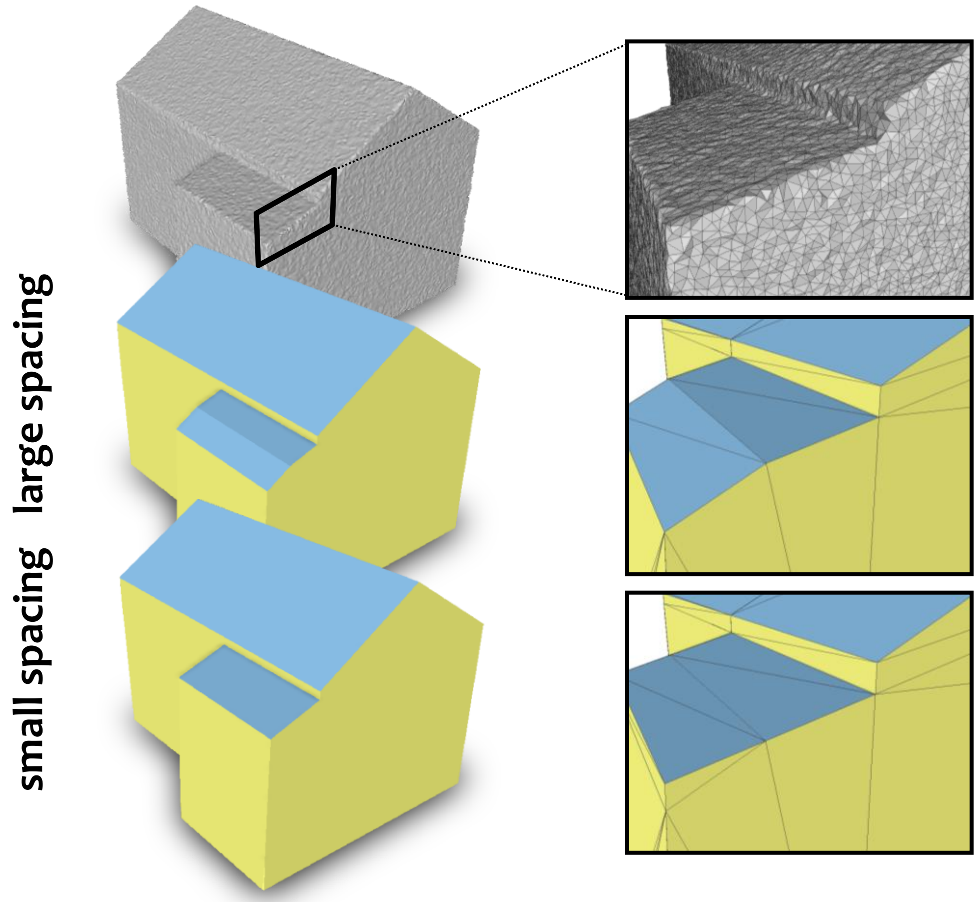
large spacing



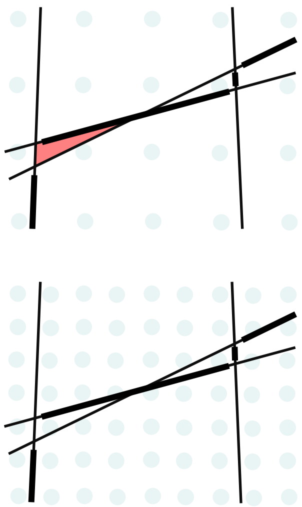
Trade-off between accuracy and time



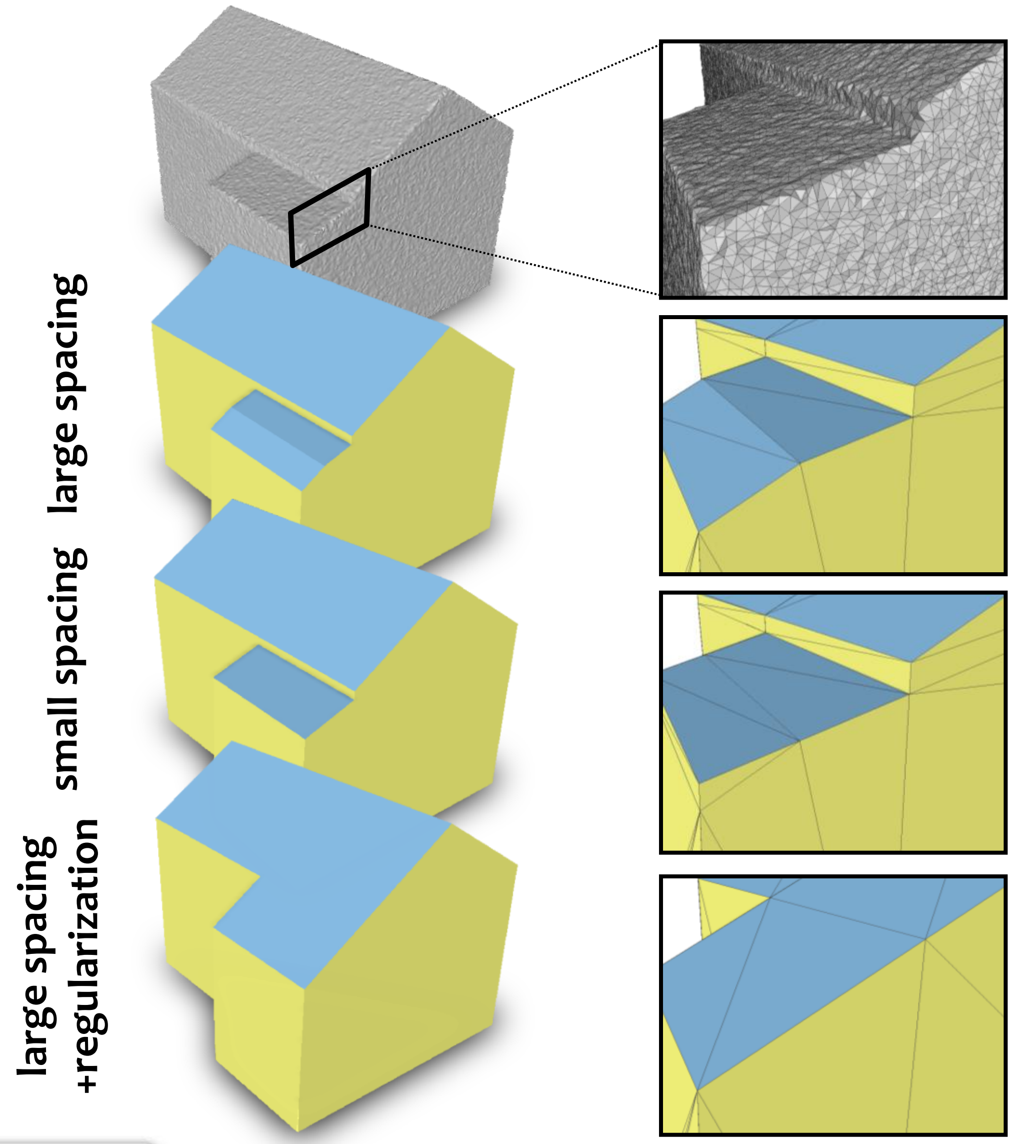
Anchor spacing setting



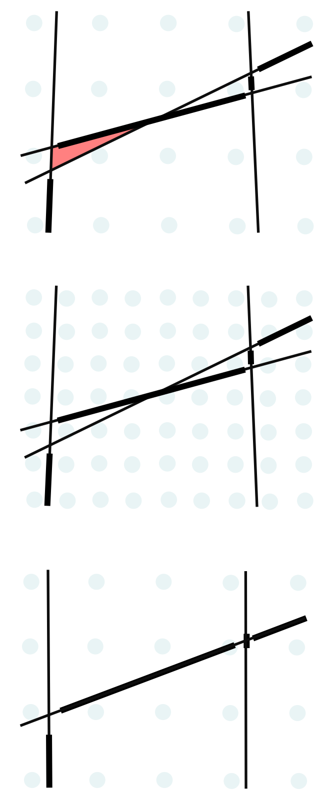
Trade-off between accuracy and time



Anchor spacing setting



Trade-off between accuracy and time



Reconstruction at various LOD

Planar proxy filtering

LOD0
→ only *facade* planar proxies

LOD1
→ LOD0 + constant roof height estimation

LOD2
→ all planar proxies

LOD3
→ LOD2 + roof and facade icons

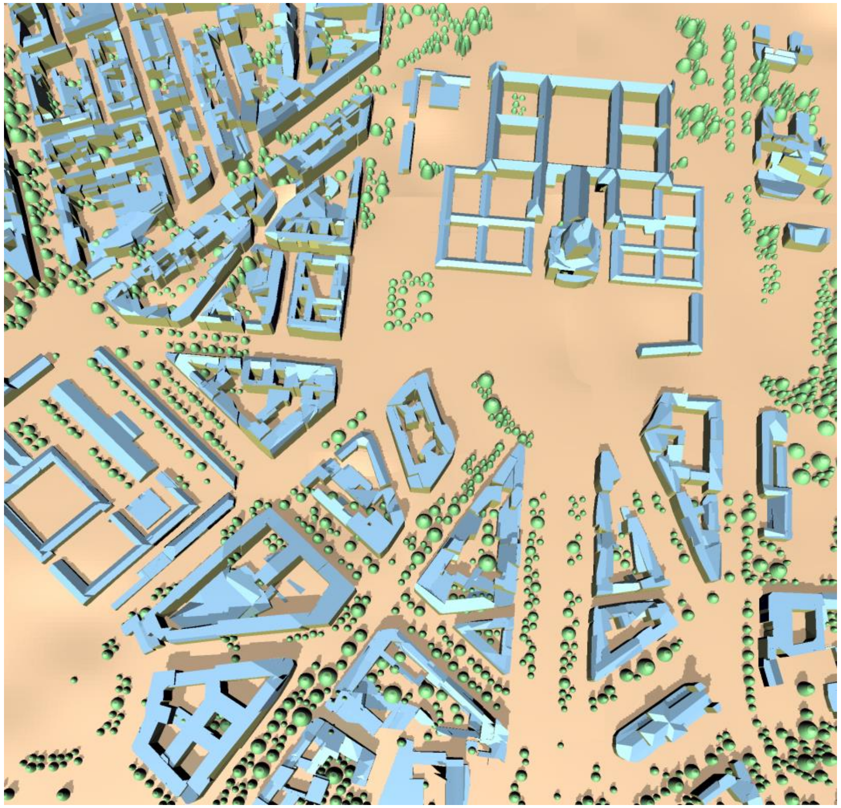
Large-scale reconstruction

input mesh
(11M facets)

LOD1
(10K facets for buildings)

Large-scale reconstruction

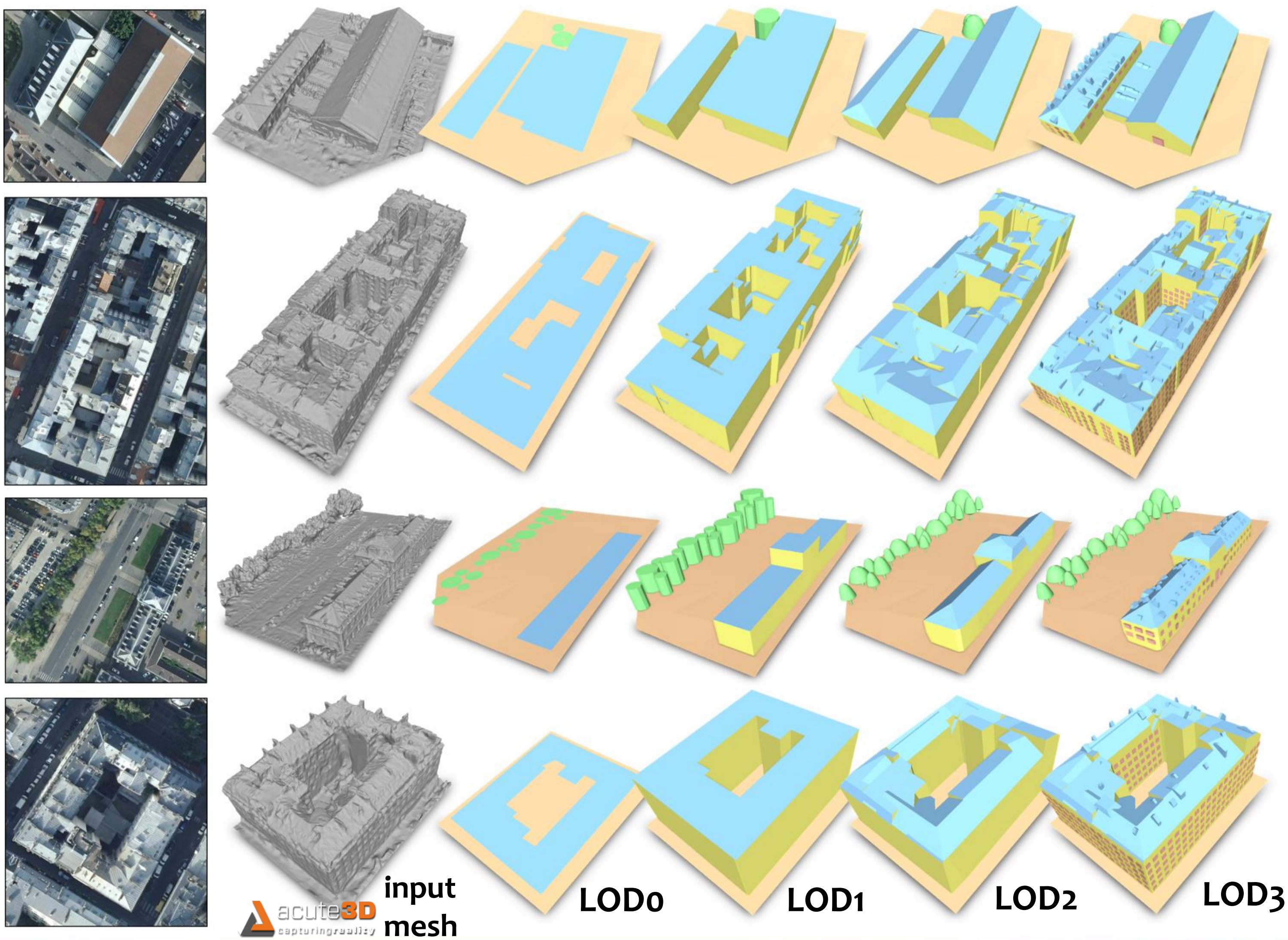
input mesh
(11M facets)



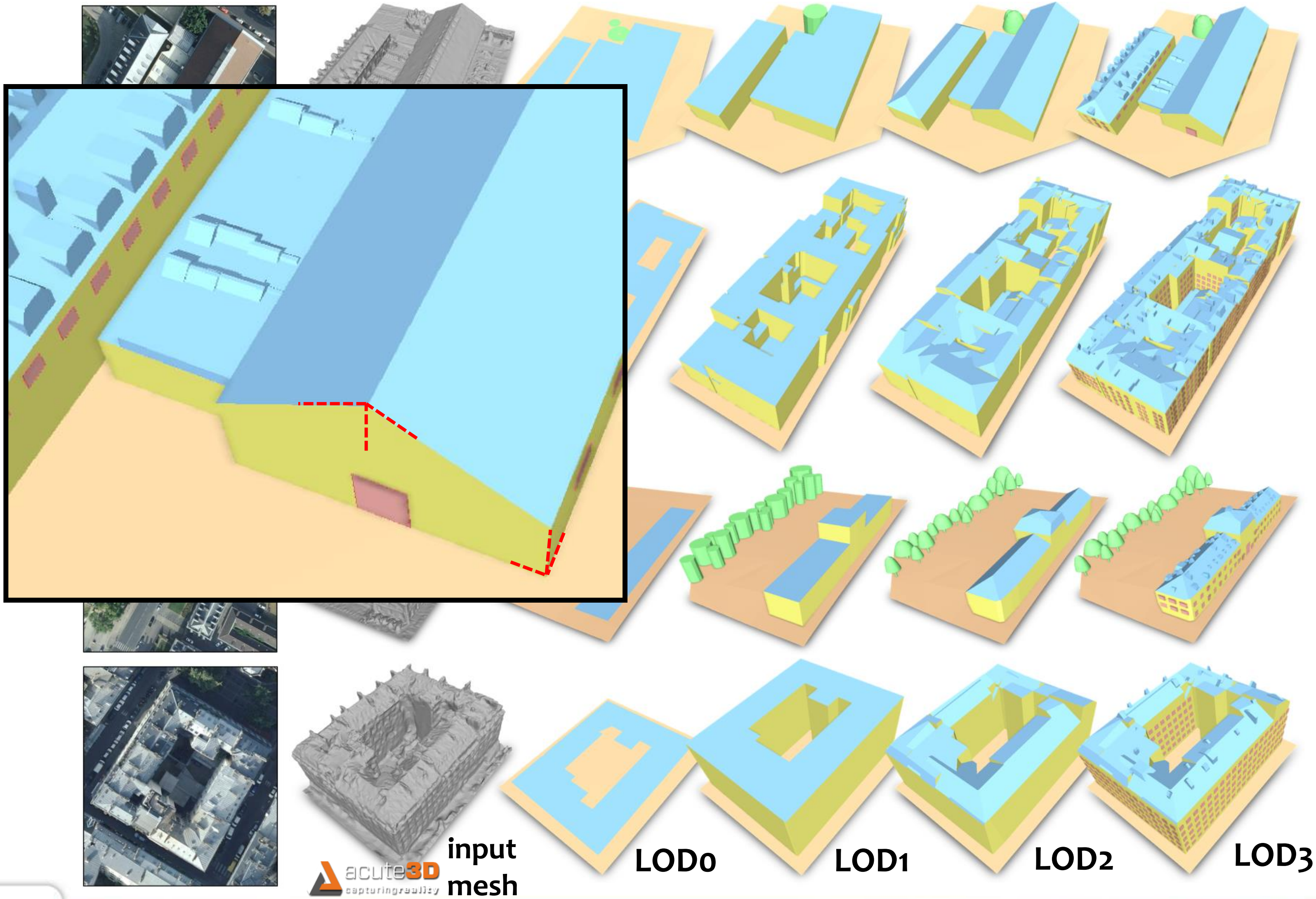
LOD2
(170K facets for buildings)

Inria

Building reconstruction



Building reconstruction

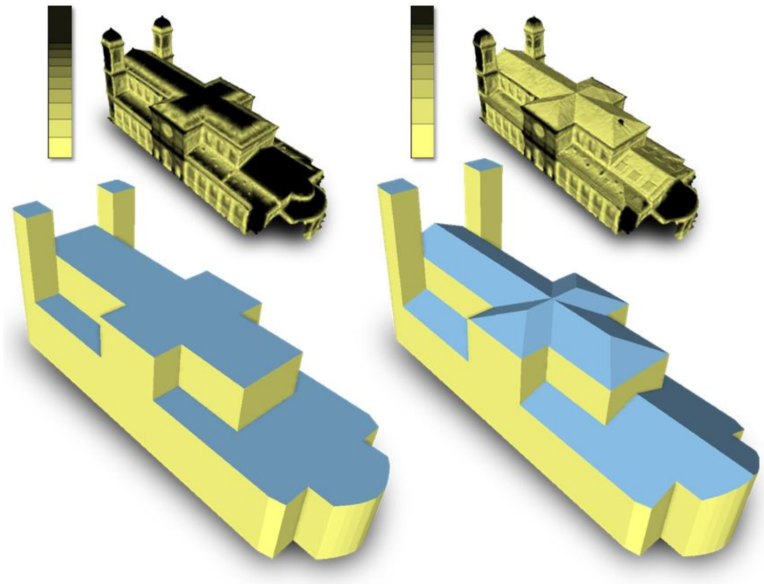


Accuracy and structure-awareness

error (meter)

0 >1

RMS=0.47 RMS=0.39



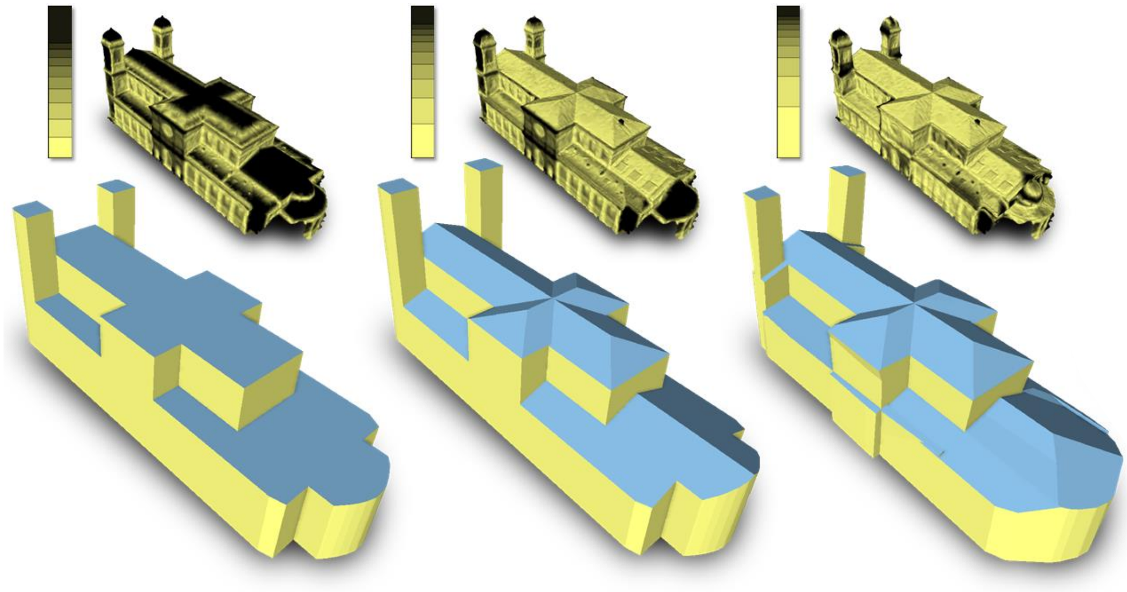
LOD1 LOD2

Accuracy and structure-awareness

error (meter)

0 >1

RMS=0.47 RMS=0.39 RMS=0.33



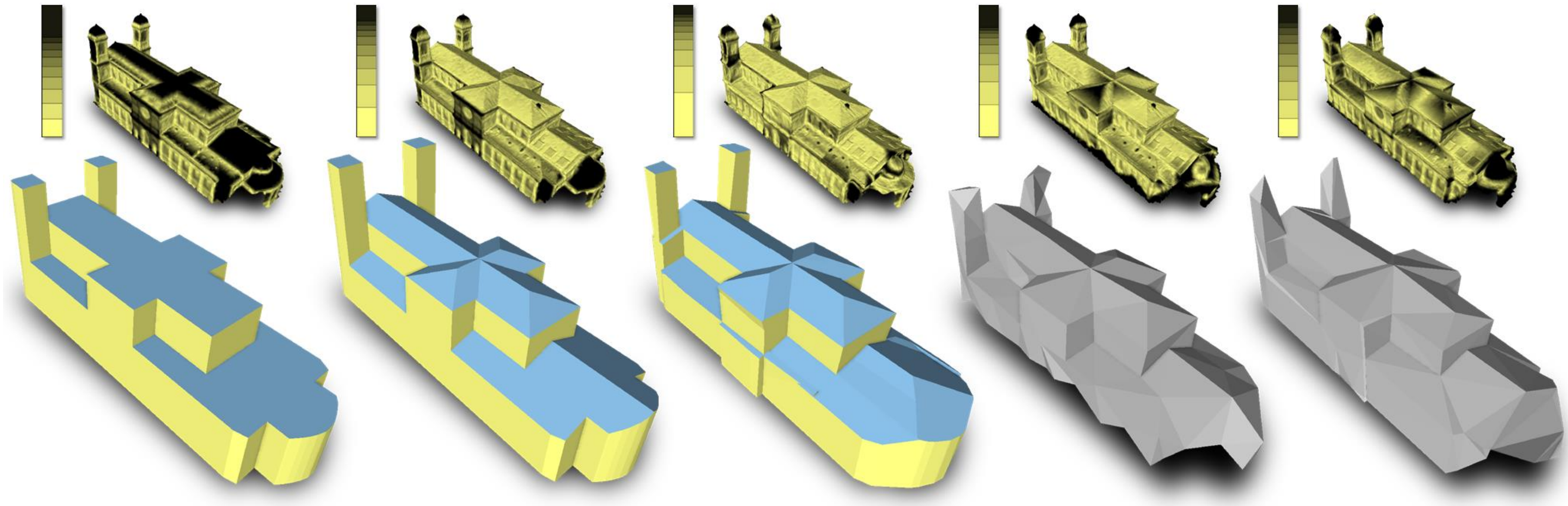
LOD1 LOD2 LOD2 w/o reg.

Accuracy and structure-awareness

error (meter)

0 >1

RMS=0.47 RMS=0.39 RMS=0.33 RMS=0.4 RMS=0.43



LOD1

LOD2

LOD2
w/o reg.

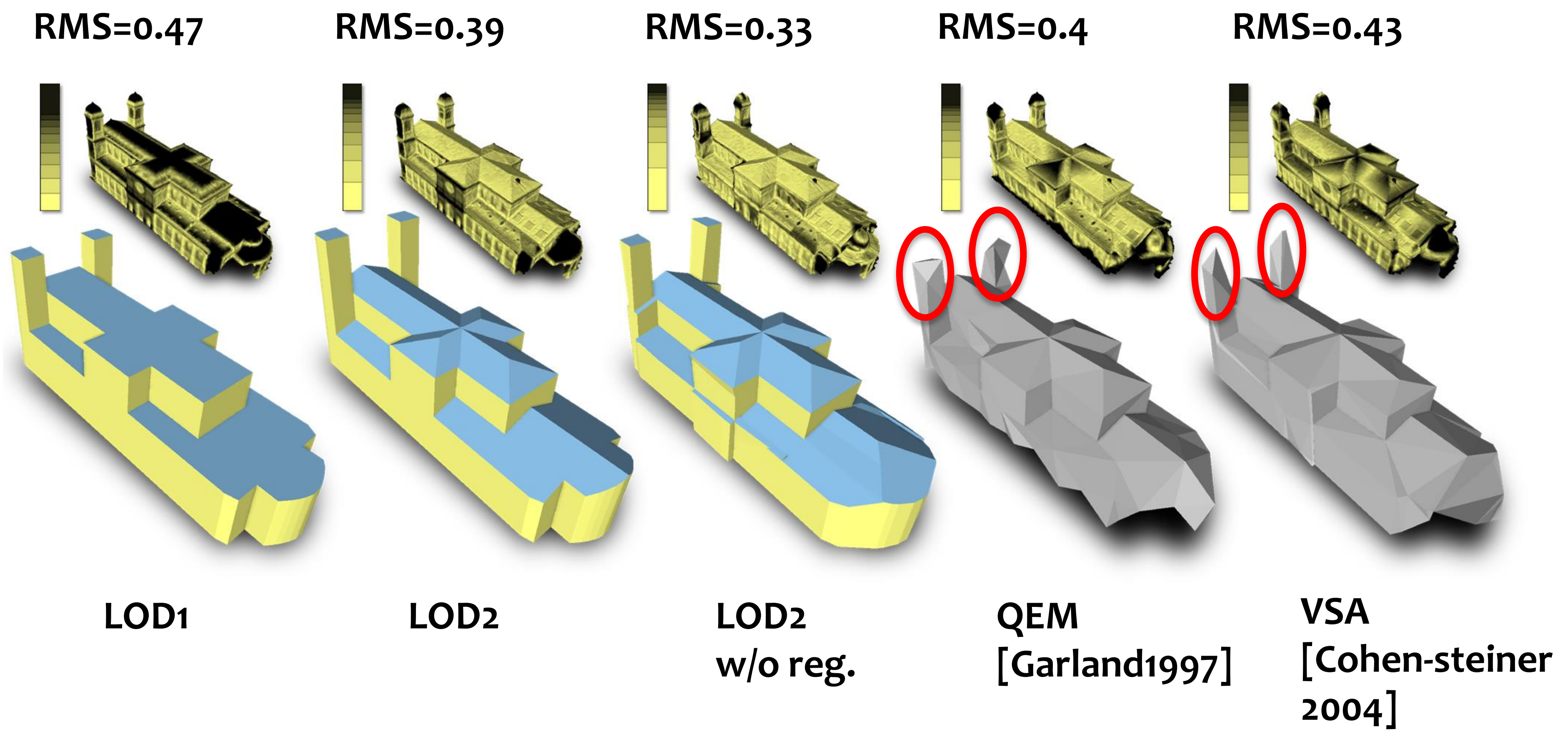
QEM
[Garland1997]

VSA
[Cohen-steiner
2004]

Accuracy and structure-awareness

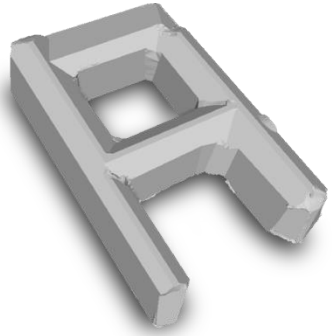
error (meter)

0 >1

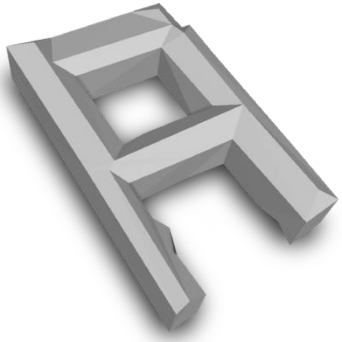


Comparison with Lidar methods

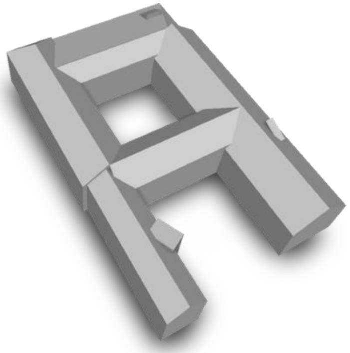
point set structuring
[Lafarge2013]

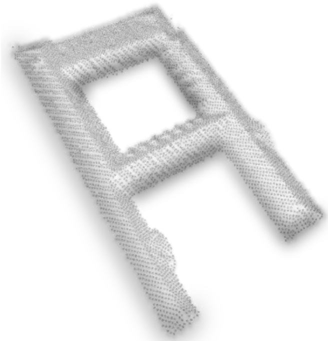


planimetric arrangement
[Lafarge2012]



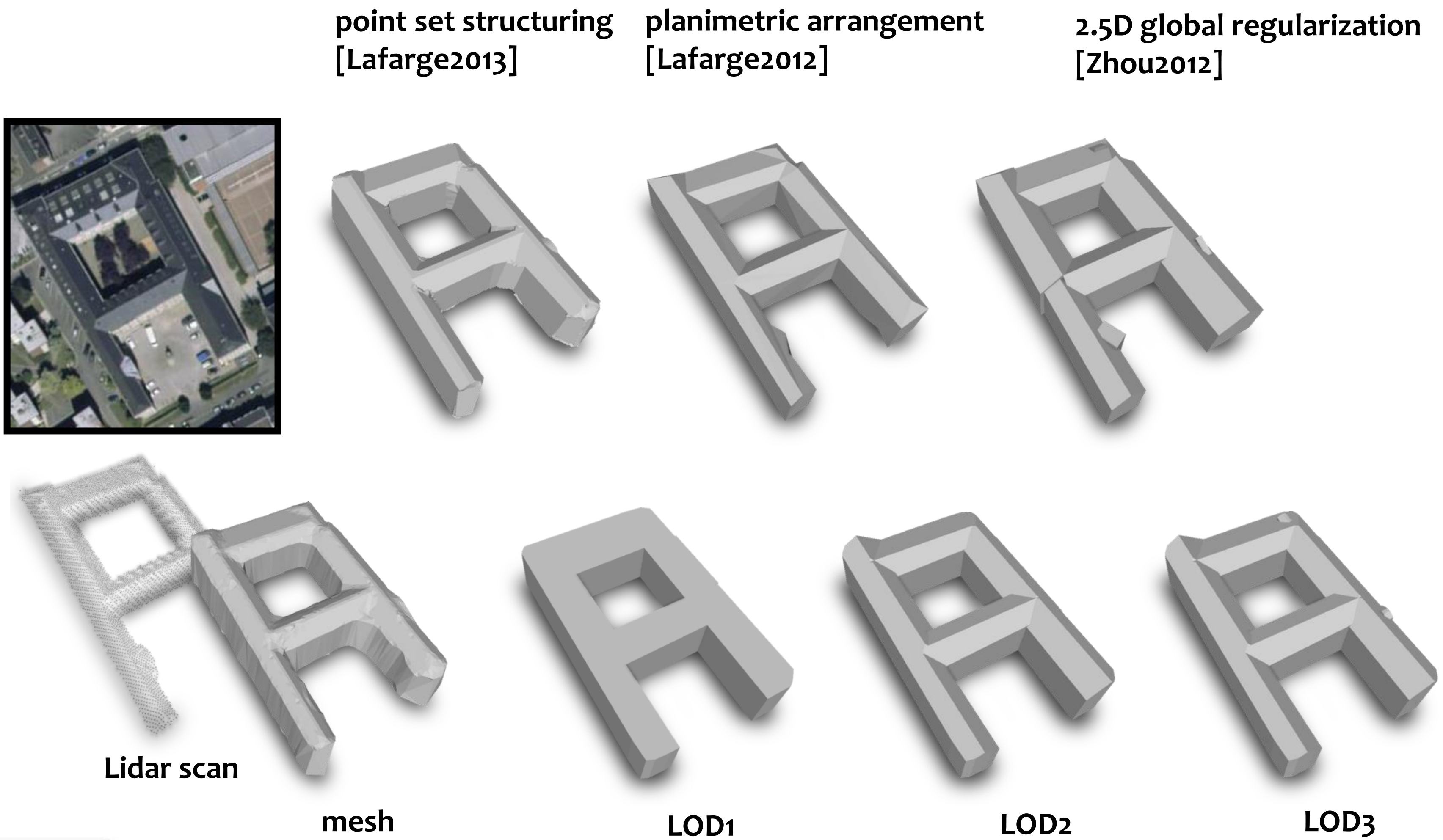
2.5D global regularization
[Zhou2012]



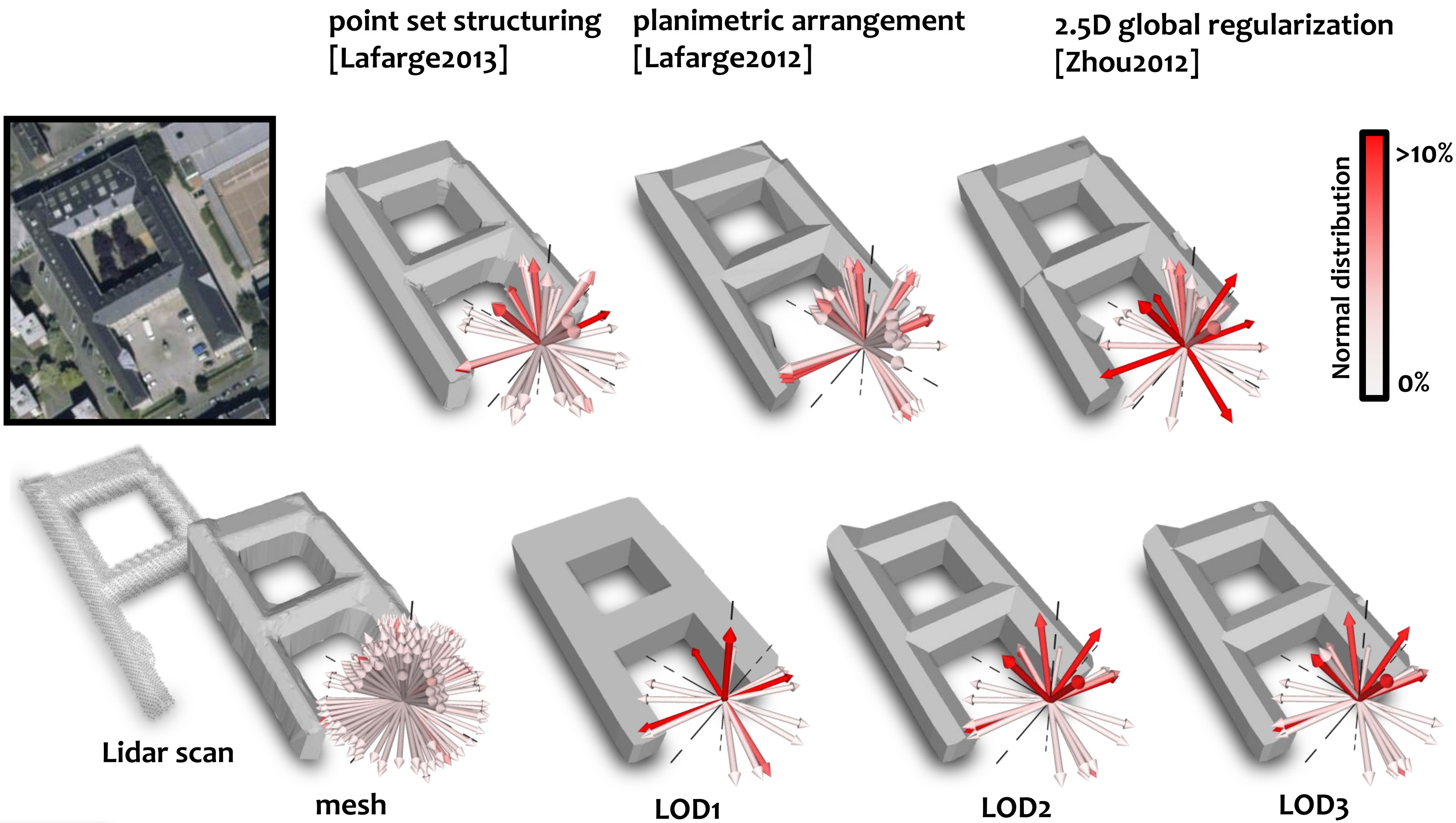


Lidar scan

Comparison with Lidar methods



Comparison with Lidar methods



Comparison with Lidar methods