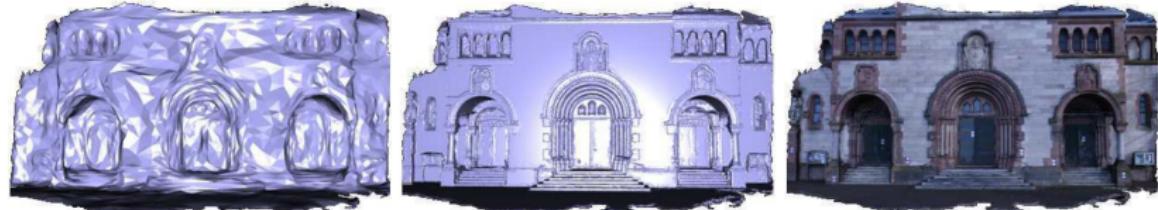


A hybrid multi-view stereo algorithm for modeling urban scenes

Sasha Pagani, Julia Giger, Prashanth Chandran

Swiss Federal Institute of Technology

April 24, 2017



Overview

1 Introduction: meshes and primitives

2 Hybrid model: 2-step strategy with iterative refinement

- Step 1: Mesh-based surface segmentation
- Step 2: Stochastic hybrid reconstruction

3 Experiments

4 Conclusions

Overview

1 Introduction: meshes and primitives

2 Hybrid model: 2-step strategy with iterative refinement

- Step 1: Mesh-based surface segmentation
- Step 2: Stochastic hybrid reconstruction

3 Experiments

4 Conclusions

Introduction

- Urban scenes are difficult to analyze
 - Structures are very different within the same scene...
 - ... but created with man-made rules (piecewise planar, shape repetition)
- Urban scene reconstruction: 2 approaches
 - Meshes
 - 3D primitive arrangements

Introduction: Meshes

- **Meshes:** set of vertices, edges and faces that defines the shape of polyhedral objects
- Several methods for generating a mesh from multi-view stereo images
 - Region growing based methods
 - Fusing a set of depth maps
 - Variational methods
 - 3D Delaunay triangulations
- **Pro:** highly detailed scene reconstruction
- **Cons:**
 - Regular structures not optimally modeled
 - Semantic not taken into account

Introduction: 3D-primitive arrangements

- **3D primitives:** planes, spheres, cylinders, cones and tori
- Exploit the **Manhattan-world assumption**
 - Predominance of three mutually orthogonal directions in the scenes
- **Pro:**
 - particularly well adapted to describe urban environments
 - Semantic taken into account
 - Easy to store and render
- **Cons:** Fail to model fine details and irregular shapes

Solution: hybrid (meshes + primitives) model approach!

Overview

1 Introduction: meshes and primitives

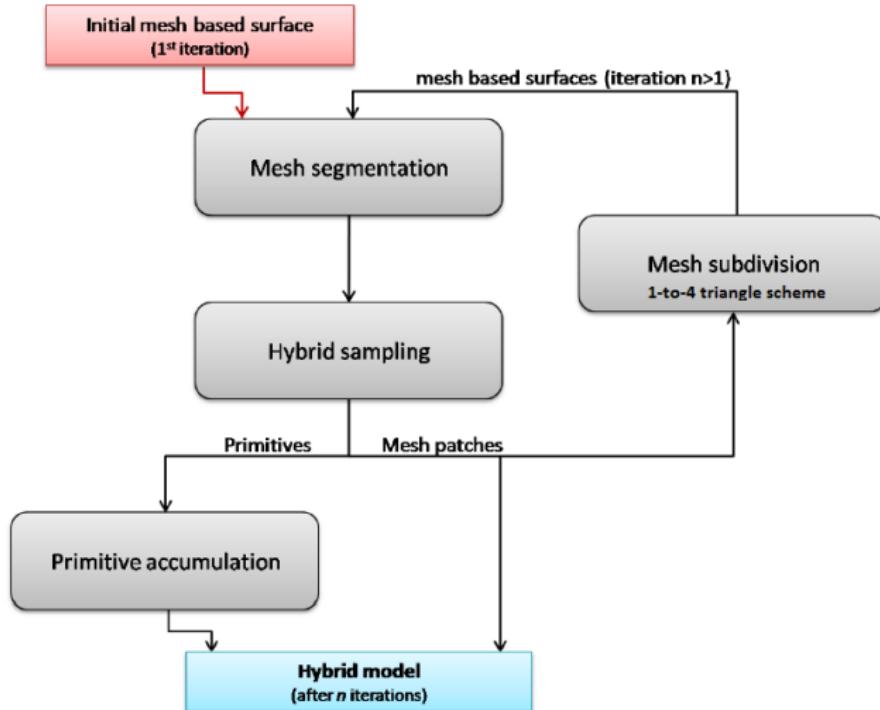
2 Hybrid model: 2-step strategy with iterative refinement

- Step 1: Mesh-based surface segmentation
- Step 2: Stochastic hybrid reconstruction

3 Experiments

4 Conclusions

Hybrid model: pipeline

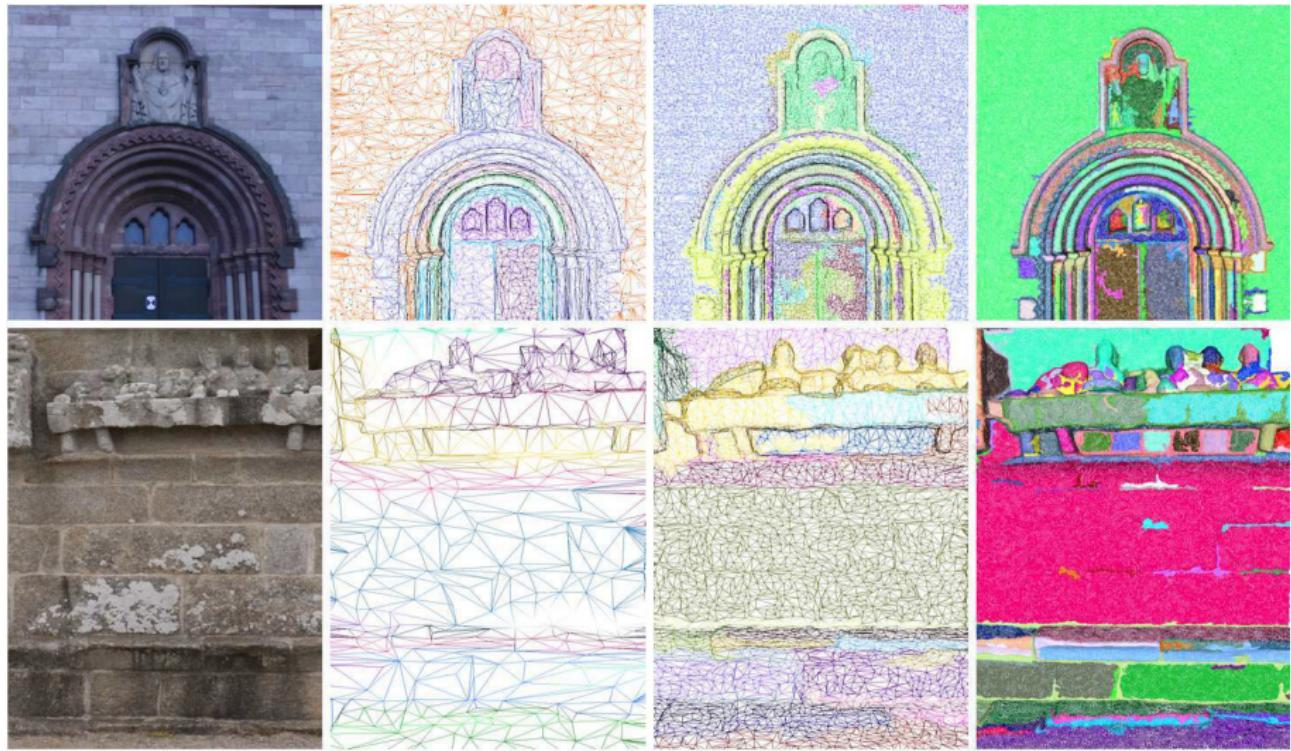


Preliminary segmentation: reduces the complexity of the problem
The procedure stops when the generated facets become too small

Step 1: Mesh-based surface segmentation

- Segmentation problem formulated with a Markov Random Field
 - Vertex set and edge adjacency as in the mesh
- Label for each vertex
 - Planar, developable convex, developable concave, non developable
- Quality of a label configuration: energy
 - Energy given by curvature distributions, label homogeneity and edge preservation
- Subdivision into clusters (using label as discriminant)

Step 1: Mesh-based surface segmentation



Step 2: Stochastic hybrid reconstruction

- H : configuration space of the hybrid models given the segmented initial mesh-based surface $\mathbf{x}^{(0)}$
- $U(\mathbf{x})$: energy measuring the quality of a hybrid model $\mathbf{x} \in H$
- Optimal hybrid model: $\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in H} U(\mathbf{x})$
- Multi-object energy model:

$$U(\mathbf{x}) = \sum_{i=1}^N U_{pc}(x_i) + \beta_1 \sum_{i=1}^{N_m} U_s(m_i) + \beta_2 \sum_{i \bowtie i'} U_a(p_i, p_{i'})$$

with m_i, p_i meshes respectively primitives in \mathbf{x} and $N_m + N_p = N$

Step 2: Photo-consistency U_{pc}

$$U(\mathbf{x}) = \sum_{i=1}^N U_{pc}(x_i) + \beta_1 \sum_{i=1}^{N_m} U_s(m_i) + \beta_2 \sum_{i \bowtie i'} U_a(p_i, p_{i'})$$

U_{pc} is the image back-projection error with respect to the object surface

$$U_{pc}(x_i) = A(x_i) \sum_{\tau, \tau'} \int_{\Omega_{\tau, \tau'}^S} f(I_\tau, I_{\tau, \tau'}^S)(s) ds$$

with $A(x_i)$ a function tuning the occurrence of primitives/meshes

Step 2: Mesh smoothness U_s

$$U(\mathbf{x}) = \sum_{i=1}^N U_{pc}(x_i) + \beta_1 \sum_{i=1}^{N_m} U_s(m_i) + \beta_2 \sum_{i \bowtie i'} U_a(p_i, p_{i'})$$

U_s allows the regularization of mesh patches and penalizes strong bending using the thin plate energy E_{TP}

$$U_s(m_i) = \sum_{\nu \in V_{m_i}} E_{TP}(\nu, \{\bar{\nu}\})$$

with $\{\bar{\nu}\}$ the set of adjacent vertices to the vertex ν

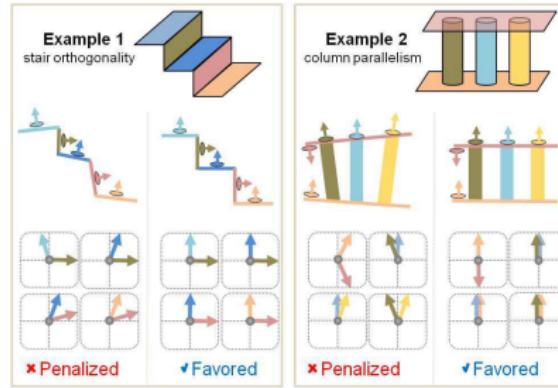
Step 2: Priors on shape layout U_a

$$U(\mathbf{x}) = \sum_{i=1}^N U_{pc}(x_i) + \beta_1 \sum_{i=1}^{N_m} U_s(m_i) + \beta_2 \sum_{i \bowtie i'} U_a(p_i, p_{i'})$$

U_a favors both perpendicular and parallel primitive layouts and object repetition in a scene

$$U_a(p_i, p_{i'}) = \omega_{ii'}(1 - \cos(2\gamma_{ii'}))^{2\alpha}$$

with $\gamma_{ii'}$ the angle between the direction of revolution of the two primitives and $\omega_{ii'}$ weight which favors repetitiveness



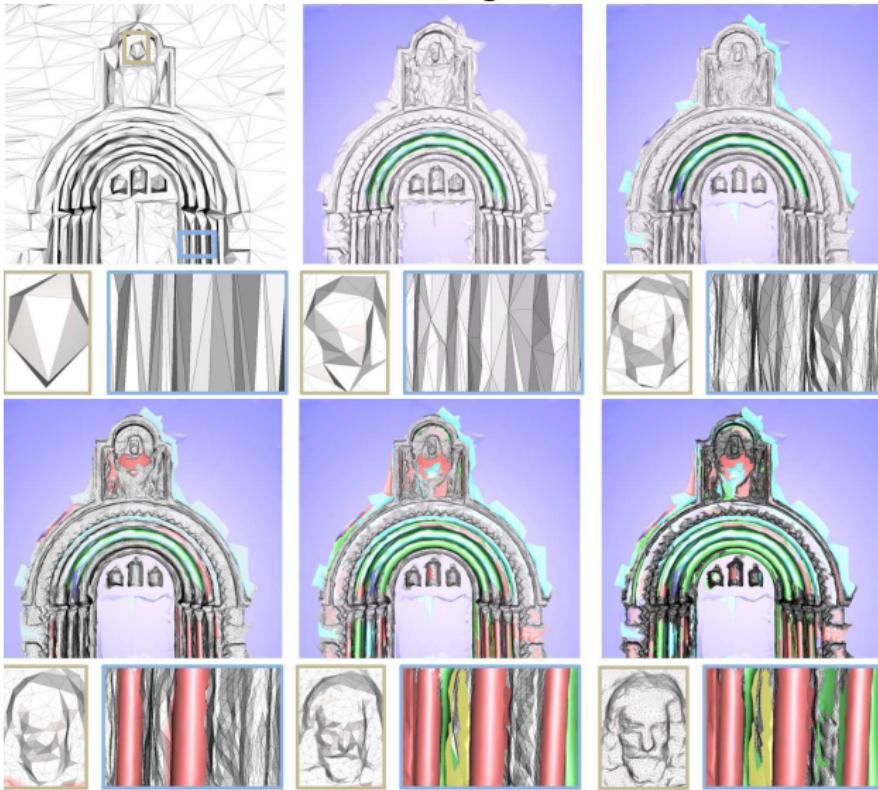
Step 2: Jump-Diffusion based sampling

$$U(\mathbf{x}) = \sum_{i=1}^N U_{pc}(x_i) + \beta_1 \sum_{i=1}^{N_m} U_s(m_i) + \beta_2 \sum_{i \bowtie i'} U_a(p_i, p_{i'})$$

- Can sample from two different types of structures
- **Jump dynamics:** $\min(1, \frac{Q(\mathbf{y} \rightarrow \mathbf{x})}{Q(\mathbf{x} \rightarrow \mathbf{y})} e^{-\frac{U(\mathbf{y}) - U(\mathbf{x})}{T}})$
 - Performs a switch on the type of an object
 - Convergence necessary condition: jump must be reversible
- **Diffusion dynamics:** $d\mathbf{x}(t) = -\frac{dU(\mathbf{x})}{d\mathbf{x}} dt + \sqrt{2T(t)} d\omega_t$
 - Optimizes the parameters for the chosen object
 - $T \gg 0$: local minima avoided
 - $T \simeq 0$: diffusion dynamics act as a gradient descent
 - Mesh adaptation: $\nabla_{m_i} U = \nabla_{m_i} U_{pc} + \beta_1 \nabla_{m_i} U_s$
 - Primitive adaptation: $\nabla_{\theta_i} U = \nabla_{\theta_i} U_{pc} + \beta_2 \sum_{i'} \nabla_{\theta_i} U_a$

Hybrid model: iterative refinement

It allows the extraction of the main regular structures at low resolutions



Overview

1 Introduction: meshes and primitives

2 Hybrid model: 2-step strategy with iterative refinement

- Step 1: Mesh-based surface segmentation
- Step 2: Stochastic hybrid reconstruction

3 Experiments

4 Conclusions

Experiments

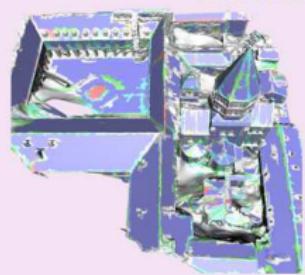
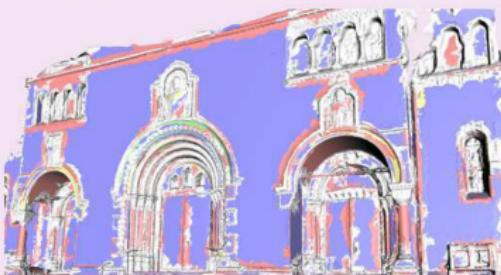
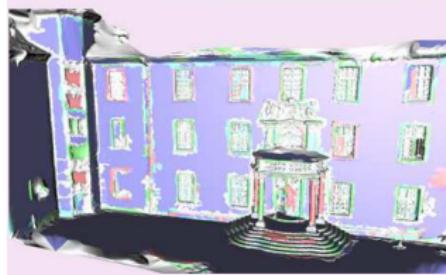
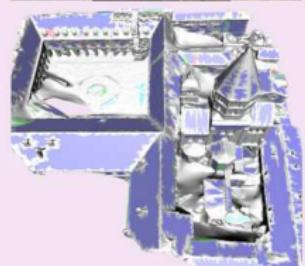
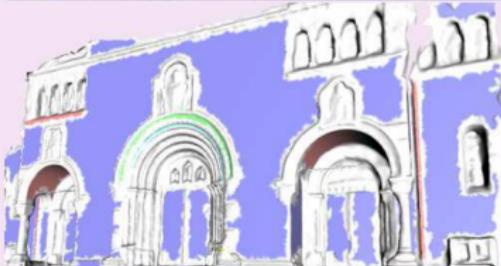
Entry-P10



Herz-Jesu-P25

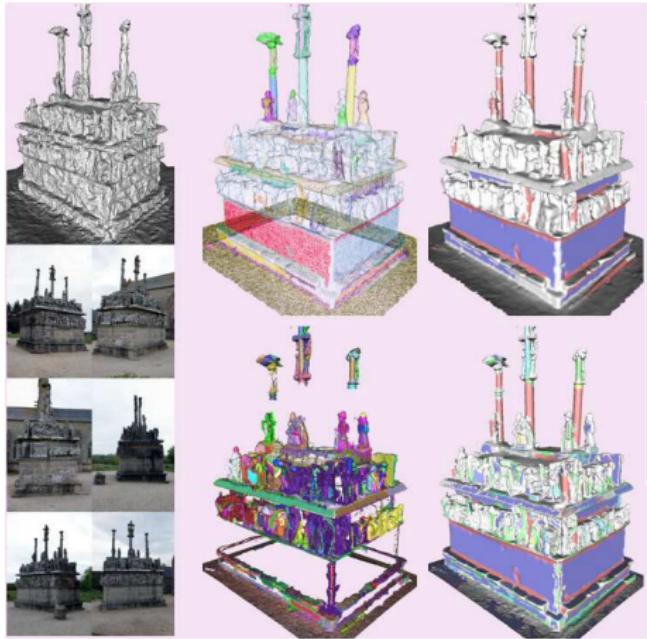


Church



Experiments

Calvary



	initial surface	LR hybrid model	HR hybrid model
Entry-P10 (10 images)	9K vert. 16K fac.	51 prim. 20K vert. 37K fac.	342 prim. 0.33M vert. 0.62M fac.
Calvary (27 images)	23Kvert. 47K fac.	37 prim. 56K vert. 0.11M fac.	426 prim. 0.55M vert. 1.04M fac.
Herz-Jesu-P25 (25 images)	14K vert. 17K fac.	41 prim. 42K vert. 77K fac.	263 prim. 0.38M vert. 0.74M fac.
Church (37 images)	21K vert. 34K fac.	143prim. 82K vert. 0.15M fac.	406 prim. 0.13M vert. 0.22M fac.

Experiments: Observations

- Segmentation stage: defines the partitions of the initial surface to sample → not crucial
- Hybrid sampling re-meshes patches which aren't relevant enough
- Tromp l'oeil structures correctly reconstructed
- Partially occluded regular structures more accurately reconstructed
- Reduction of storage capacity for similar accuracy
- Better preservation of the details

Experiments: Limitations and possible improvements

- Some details can be lost (photo-consistency tuning parameter)
 - **could initialize next iteration with full hybrid model**
 - **would considerably increase computing time**
- Shape layout prior may fail
 - **consider additional primitive attributes/more primitives**
- Embedding segmentation step into sampling procedure
 - **a better segmentation could be chosen**
 - **algorithm would become less robust and efficient**
- Rough initial surface has to be topologically close to the real one

Overview

1 Introduction: meshes and primitives

2 Hybrid model: 2-step strategy with iterative refinement

- Step 1: Mesh-based surface segmentation
- Step 2: Stochastic hybrid reconstruction

3 Experiments

4 Conclusions

Conclusions

- Original multi-view reconstruction algorithm (primitives for regular structures, mesh patches for irregular components)
- High storage savings but accuracy similar to the best mesh-based algorithms
- Takes into account semantic knowledge (some occlusions can be solved)
- Jump-Diffusion sampler explores complex configuration spaces fast

Thanks for your attention!

Questions?