Hybrid Multi-view reconstruction by Jump-Diffusion

Florent Lafarge¹,²  Renaud Keriven¹  Mathieu Brédif³  Vu Hoang Hiep¹

¹ Imagine group/LIGM, Université Paris Est
² INRIA Sophia Antipolis
³ French Mapping Agency
Urban scene modeling

Mesh-based representations

- **details**
- **generality**

[Seitz et al., cvpr06] [Furukawa and Ponce, iccv07] [Goesele et al. iccv07] [Strecha et al., cvpr08] [Pollefeys et al., ijcv08] [Vu et al., cvpr09]...
Urban scene modeling

Mesh-based representations

♦ details
♦ generality

[Seitz et al., cvpr06] [Furukawa and Ponce, iccv07]
[Goesele et al. iccv07] [Strecha et al., cvpr08]
[Pollefeys et al., ijcv08] [Vu et al., cvpr09]...

Primitive-based representations

♦ compaction
♦ semantic

[Dick et al., ijcv04] [Muller et al., siggraph06]
[Zebedin et al., eccv08] [Lafarge et al., cvpr08]
[Sinha et al., iccv09] [Zhou and Neumann, cvpr09] ...
Urban scene modeling

Mesh-based representations

- details
- generality

[Seitz et al., cvpr06] [Furukawa and Ponce, iccv07] [Goesele et al. iccv07] [Strecha et al., cvpr08] [Pollefeys et al., ijcv08] [Vu et al., cvpr09]...

Primitive-based representations

- compaction
- semantic

[Dick et al., ijcv04] [Muller et al., siggraph06] [Zebedin et al., eccv08] [Lafarge et al., cvpr08] [Sinha et al., iccv09] [Zhou and Neumann, cvpr09]...

Why not combining these two types of representations?
Hybrid representation
## Hybrid representation

*Primitives to describe regular structures (walls, columns...)*
Hybrid representation

Primitives to describe regular structures (walls, columns...)

Mesh patches to describe irregular elements (statues, ornaments...)

[Introduction] [Mesh segmentation] [Hybrid sampling] [Iterative refinement] [Experiments]
Hybrid representation

*Primitives to describe regular structures (walls, columns...)*

*Mesh patches to describe irregular elements (statues, ornaments...)*

*compaction while preserving details*
Hybrid representation

*Primitives to describe regular structures (walls, columns...)*

*Mesh patches to describe irregular elements (statues, ornaments...)*

compaction while preserving details

An idea still lightly explored

- Ransac-based primitive extraction from point clouds + meshing [Labatut et al., 3DIM09]
  - fails to describe details (outliers)
  - no priors on primitive layouts
Hybrid representation

Primitives to describe regular structures (walls, columns...)

Mesh patches to describe irregular elements (statues, ornaments...)

compaction while preserving details

An idea still lightly explored

♦ Ransac-based primitive extraction from point clouds + meshing [Labatut et al., 3DIM09]
  • fails to describe details (outliers)
  • no priors on primitive layouts

♦ Primitive insertion in meshes [Lafarge et al., BMVC09]
  • limited accuracy (no photo-consistency)
  • high quality mesh required as input
## Contributions

- **mesh and 3D-primitive joint sampler**
  
  mesh patches and 3D-primitives evolve and interact in a common framework
Contributions

♦ **mesh and 3D-primitive joint sampler**
  mesh patches and 3D-primitives evolve and interact in a common framework

♦ **shape layout prior in urban scenes**
  introduction of urban knowledge by favoring certain shape layouts according to parallelism/perpendicularity and repetitiveness properties
Contributions

♦ mesh and 3D-primitive joint sampler
  mesh patches and 3D-primitives evolve and interact in a common framework

♦ shape layout prior in urban scenes
  introduction of urban knowledge by favoring certain shape layouts according to parallelism/perpendicularity and repetitiveness properties

♦ efficient global optimization
  Jump-Diffusion allows the escape from local minima thanks to stochastic relaxation while gradient descent based dynamics guarantee fast local explorations
System overview

**INPUT**

- multi-view stereo images
- a rough initial surface
System overview

INPUT

♦ multi-view stereo images
♦ a rough initial surface

OUTPUT

a hybrid surface combining mesh patches and 3D-primitives
**System overview**

**INPUT**
- multi-view stereo images
- a rough initial surface

**OUTPUT**
a hybrid surface combining mesh patches and 3D-primitives

**Introduction**

**[Mesh segmentation]**

**[Hybrid sampling]**

**[Iterative refinement]**

**[Experiments]**
Mesh segmentation

Initial mesh based surface (1st iteration)

Mesh segmentation

Hybrid sampling

Mesh subdivision

Primitives

Mesh patches

Primitive accumulation

Hybrid model (after n iterations)
Partitioning non-synthetic meshes

Surface characteristics

♦ frequent meshing degeneracies
Partitioning non-synthetic meshes

**Surface characteristics**

- frequent meshing degeneracies
- significant noise corruption
Partitioning non-synthetic meshes

Surface characteristics

♦ frequent meshing degeneracies

♦ significant noise corruption

segmentation algorithm for non-synthetic meshes [Lafarge et al., BMVC09]
A multi-label MRF model

\[ U_{\text{seg}}(l) = \sum_{i \in V} D_i(l_i) + \beta \sum_{\{i,j\} \in E} V_{ij}(l_i, l_j) \]
A multi-label MRF model

\[ U_{\text{seg}}(l) = \sum_{i \in V} D_i(l_i) + \beta \sum_{\{i,j\} \in E} V_{ij}(l_i, l_j) \]

♦ data term
principal curvature distributions
A multi-label MRF model

\[ U_{\text{seg}}(l) = \sum_{i \in V} D_i(l_i) + \beta \sum_{\{i,j\} \in E} V_{ij}(l_i, l_j) \]

- data term
  - principal curvature distributions

- propagation constraints
  - label consistency
  - edge preservation
A multi-label MRF model

\[ U_{\text{seg}}(l) = \sum_{i \in V} D_i(l_i) + \beta \sum_{\{i,j\} \in E} V_{ij}(l_i, l_j) \]

- data term
  principal curvature distributions

- propagation constraints
  - label consistency
  - edge preservation

- optimization
  \( \alpha \)-expansion [Boykov et al., pami01]
Interesting points

- robust when faced with noise corruption and facet degeneracies
Interesting points

- robust when faced with noise corruption and facet degeneracies
- adapted to different mesh densities
Interesting points

- robust when faced with noise corruption and facet degeneracies
- adapted to different mesh densities
- eventual errors do not have critical consequences on the final result
Hybrid sampling
3D-objects and configuration space

- 6 types of 3D-objects
  - *plane, cylinder, cone, sphere, torus* and *mesh*
3D-objects and configuration space

♦ 6 types of 3D-objects
   plane, cylinder, cone, sphere, torus and mesh

♦ a hybrid model x
   a set of 3D-objects, each of them associated with a cluster of the segmented initial surface
3D-objects and configuration space

- **6 types of 3D-objects**
  - \( plane, cylinder, cone, sphere, torus \) and \( mesh \)

- **a hybrid model** \( x \)
  - a set of 3D-objects, each of them associated with a cluster of the segmented initial surface

- **the configuration space**
  - a union of \( 6^N \) continuous subspaces, each subspace containing a predefined object type per cluster
3D-objects and configuration space

- 6 types of 3D-objects
  plane, cylinder, cone, sphere, torus and mesh

- a hybrid model \( x \)
  a set of 3D-objects, each of them associated with a cluster of the segmented initial surface

- the configuration space
  a union of \( 6^N \) continuous subspaces, each subspace containing a predefined object type per cluster

- an energy \( U \)
  a measure of the quality of a hybrid model \( x \)
Energy formulation

\[ U(x) = U_{\text{photo}}(x) + U_{\text{smooth}}(x) + U_{\text{shape}}(x) \]
Energy formulation

\[ U(x) = U_{\text{photo}}(x) + U_{\text{smooth}}(x) + U_{\text{shape}}(x) \]

- Photo-consistency
  based on computation of the image back-projection error with respect to the object surface [Pons et al., ijcv07]
Energy formulation

\[ U(x) = U_{\text{photo}}(x) + U_{\text{smooth}}(x) + U_{\text{shape}}(x) \]

- **Photo-consistency**
  based on computation of the image back-projection error with respect to the object surface [Pons et al., ijcv07]

- **mesh smoothness**
  smoothness constraints penalizing strong bending
Energy formulation

\[ U(x) = U_{\text{photo}}(x) + U_{\text{smooth}}(x) + U_{\text{shape}}(x) \]

- **Photo-consistency**
  based on computation of the image back-projection error with respect to the object surface [Pons et al., ijcv07]

- **mesh smoothness**
  smoothness constraints penalizing strong bending

- **shape layout**
  urban knowledge favoring special layouts of primitives:
  - perpendicular and parallel structures
  - primitive repetition
Energy formulation

\[ U(x) = U_{photo}(x) + U_{smooth}(x) + U_{shape}(x) \]

- Photo-consistency
  based on computation of the image back-projection error with respect to the object surface [Pons et al., ijcv07]

- mesh smoothness
  smoothness constraints penalizing strong bending

- shape layout
  urban knowledge favoring special layouts of primitives:
  - perpendicular and parallel structures
  - primitive repetition

→ Energy not convex!
Jump-Diffusion

Algorithm [Grenander and Miller, 94]

♦ combination of 2 types of dynamics
  • MCMC: performs jumps between the subspaces
  • Langevin equations: diffusions within each continuous subspace

♦ a relaxation parameter controls the global process
Jump-Diffusion

Algorithm [Grenander and Miller, 94]

- combination of 2 types of dynamics
  - MCMC: performs jumps between the subspaces
  - Langevin equations: diffusions within each continuous subspace
- a relaxation parameter controls the global process

In our case,

- switching kernel: jumps from an object type to another
- mesh adaptation: evolution of the mesh-based object by variational considerations
- primitive competition: selection of relevant parameters for the primitive-based objects
Iterative refinement
Primitive accumulation and mesh subdivision

- **Hybrid models at different scales**
  
  *First iterations*: main regular structures and rough mesh patches  
  *Last iterations*: structures at various scales and fine meshing
Primitive accumulation and mesh subdivision

- **Hybrid models at different scales**
  
  *First iterations*: main regular structures and rough mesh patches
  
  *Last iterations*: structures at various scales and fine meshing

- **Self-correction**
  
  The irrelevant clusters are corrected at the next iterations as a result of a more accurate re-meshing
Experiments
Facades and roofs

COLOR CODE

white: mesh  purple: plane  pink: cylinder  blue: cone  yellow: sphere  green: torus
Facades and roofs

COLOR CODE
white: mesh  purple: plane  pink: cylinder  blue: cone  yellow: sphere  green: torus
Facades and roofs

COLOR CODE

white: mesh    purple: plane    pink: cylinder    blue: cone    yellow: sphere    green: torus

images

low resolution (first iterations)

high resolution (last iterations)
Rock sculture

- statues: mesh patches
- rock facets and pipes: primitives
Rock sculpture

statues: partially described by small primitives
Comparison with mesh-based multi-view algorithms

Accuracy

[strecha et al., cvpr08]

♦ 1\textsuperscript{st} and 2\textsuperscript{nd} best accuracies on Herz-Jesu-P25 and Entry-P10 data sets
Comparison with mesh-based multi-view algorithms

Accuracy

Trompe l’œil textures

Wall correctly modeled by a plane
Comparison with mesh-based multi-view algorithms

Accuracy

Trompe l’œil textures

- wall correctly modeled by a plane
- wall corrupted by an important noise
Comparison with mesh-based multi-view algorithms

Accuracy

Partially occluded component

- Ground truth
- Hybrid model
- Vu et al.
- Salman et al.

← column correctly modeled by a cylinder
Comparison with mesh-based multi-view algorithms

Accuracy

Partially occluded component

- column correctly modeled by a cylinder
- column with missing parts
Comparison with mesh-based multi-view algorithms

Compaction

<table>
<thead>
<tr>
<th></th>
<th>initial surface</th>
<th>LR hybrid model</th>
<th>HR hybrid model</th>
<th>storage saving rates w.r.t. [vu et al.,09]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry-P10 (10 images)</td>
<td>9K vert. 16K fac.</td>
<td>51 prim. 20K vert. 37K fac.</td>
<td>342 prim. 0.33M vert. 0.62M fac.</td>
<td>LR: 26.4 HR: 5.7</td>
</tr>
<tr>
<td>Calvary (27 images)</td>
<td>23K vert. 47K fac.</td>
<td>37 prim. 56K vert. 0.11M fac.</td>
<td>426 prim. 0.55M vert. 1.04M fac.</td>
<td>LR: 17.7 HR: 4.3</td>
</tr>
<tr>
<td>Herz-Jesu-P25 (25 images)</td>
<td>14K vert. 17K fac.</td>
<td>41 prim. 42K vert. 77K fac.</td>
<td>263 prim. 0.38M vert. 0.74M fac.</td>
<td>LR: 21 HR: 5.2</td>
</tr>
<tr>
<td>Church (37 images)</td>
<td>21K vert. 34K fac.</td>
<td>143prim. 82K vert. 0.15M fac.</td>
<td>406 prim. 0.13M vert. 0.22M fac.</td>
<td>LR: 22.9 HR: 4.9</td>
</tr>
</tbody>
</table>

♦ storage saving rates: around 5 at high resolution
Other interesting points

occurrence of primitives w.r.t. mesh patches

high

low
Other interesting points

occurrence of primitives w.r.t. mesh patches
Conclusion

- both compact and detailed
- accuracy similar to the best mesh based multi-view algorithms
- partially semantized (occluded structures, *trompe l’œil* textures...)

Perspectives

- improving the shape layout prior (more constraints on structure repetitions)
- embedding the segmentation step into the sampling procedure
Thank you!

(Also in poster version this evening)

Acknowledgements

- Funding: EADS foundation
- Datasets: C. Strecha and B. Curless
initial rough surface

hybrid model

textured hybrid model