# Ray Potentials for Dense Semantic 3D Reconstruction (published at CVPR'16)

Nikolay Savinov

ETH zürich

Computer Vision and Geometry Lab

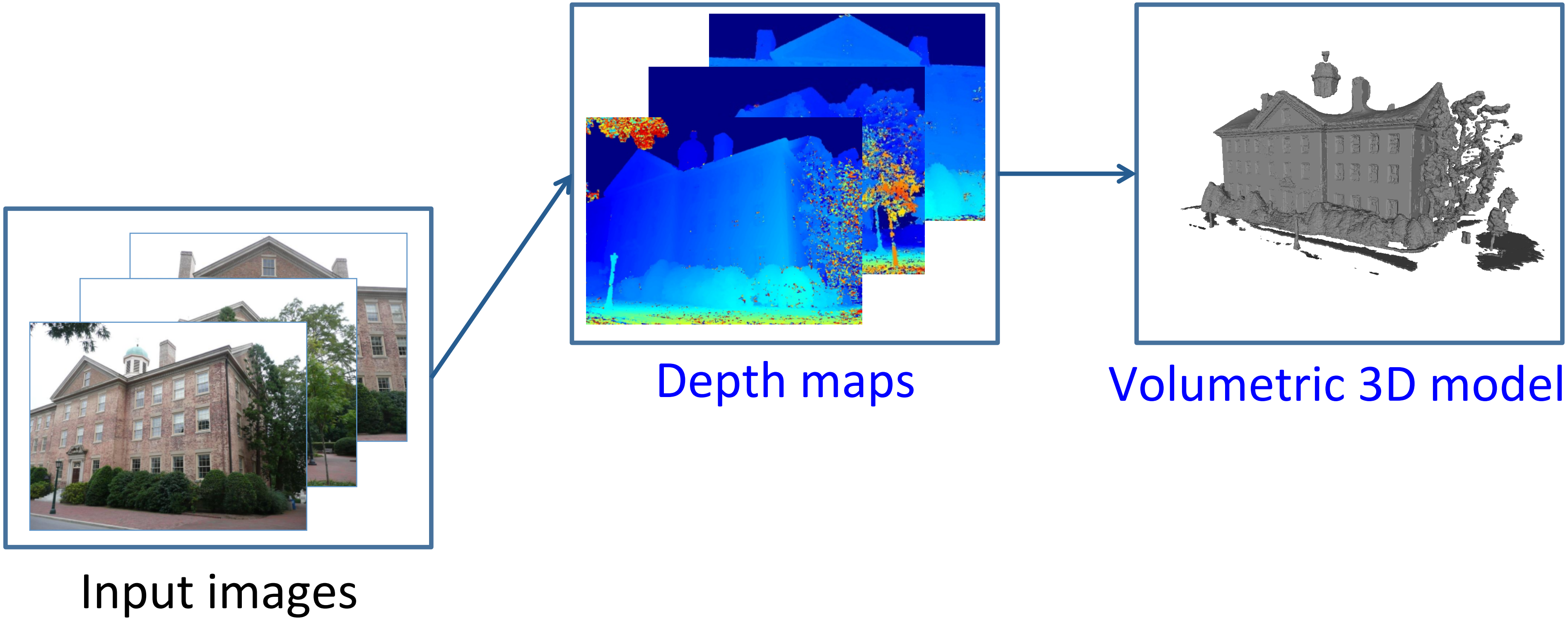
# Joint work with:

Christian Häne,
Postdoctoral Researcher at UC Berkeley

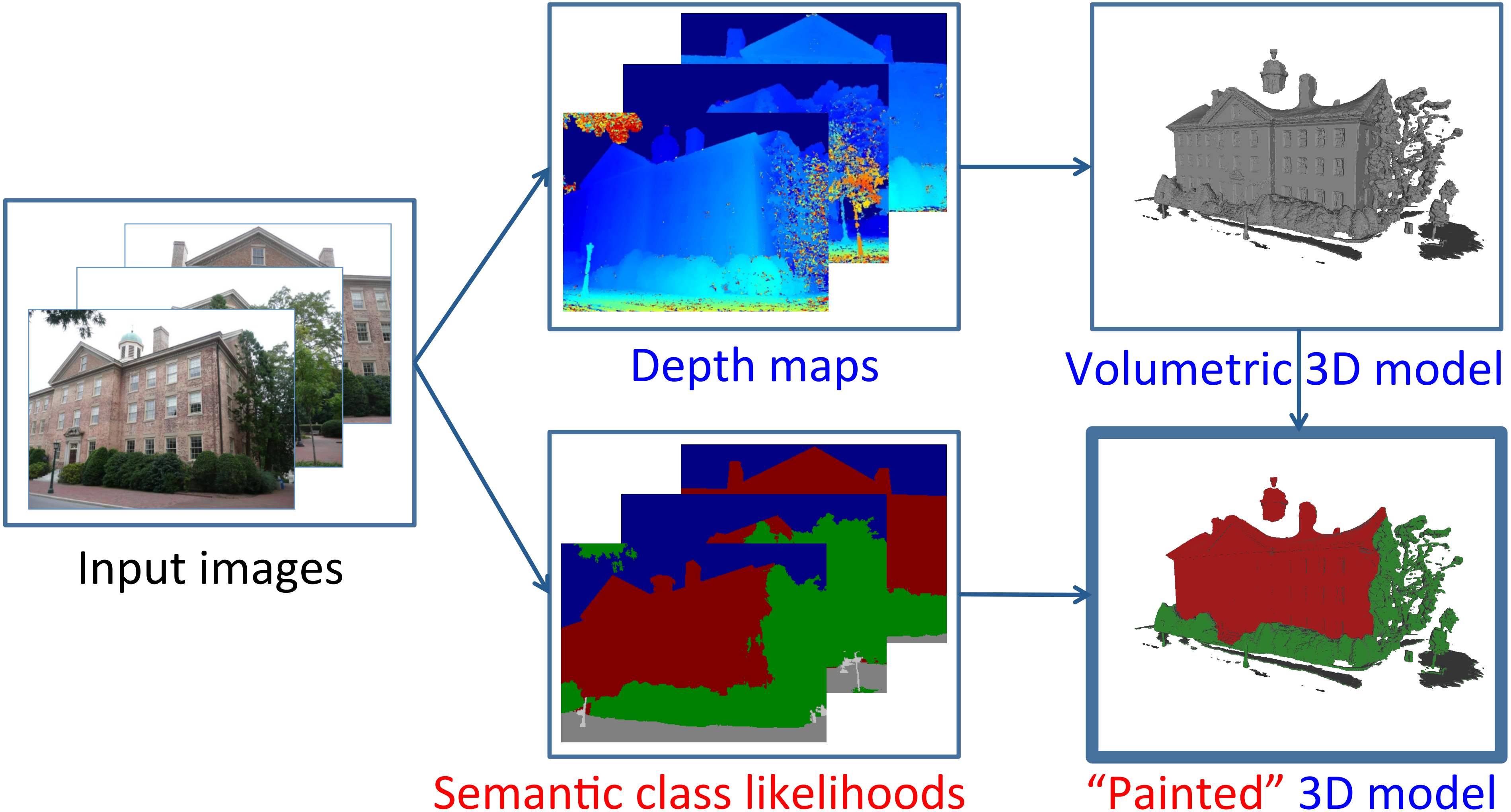
Lubor Ladicky,
Postdoctoral Researcher at ETH Zurich

Marc Pollefeys, Full Professor at ETH Zurich,
Director of Science at Microsoft HoloLens

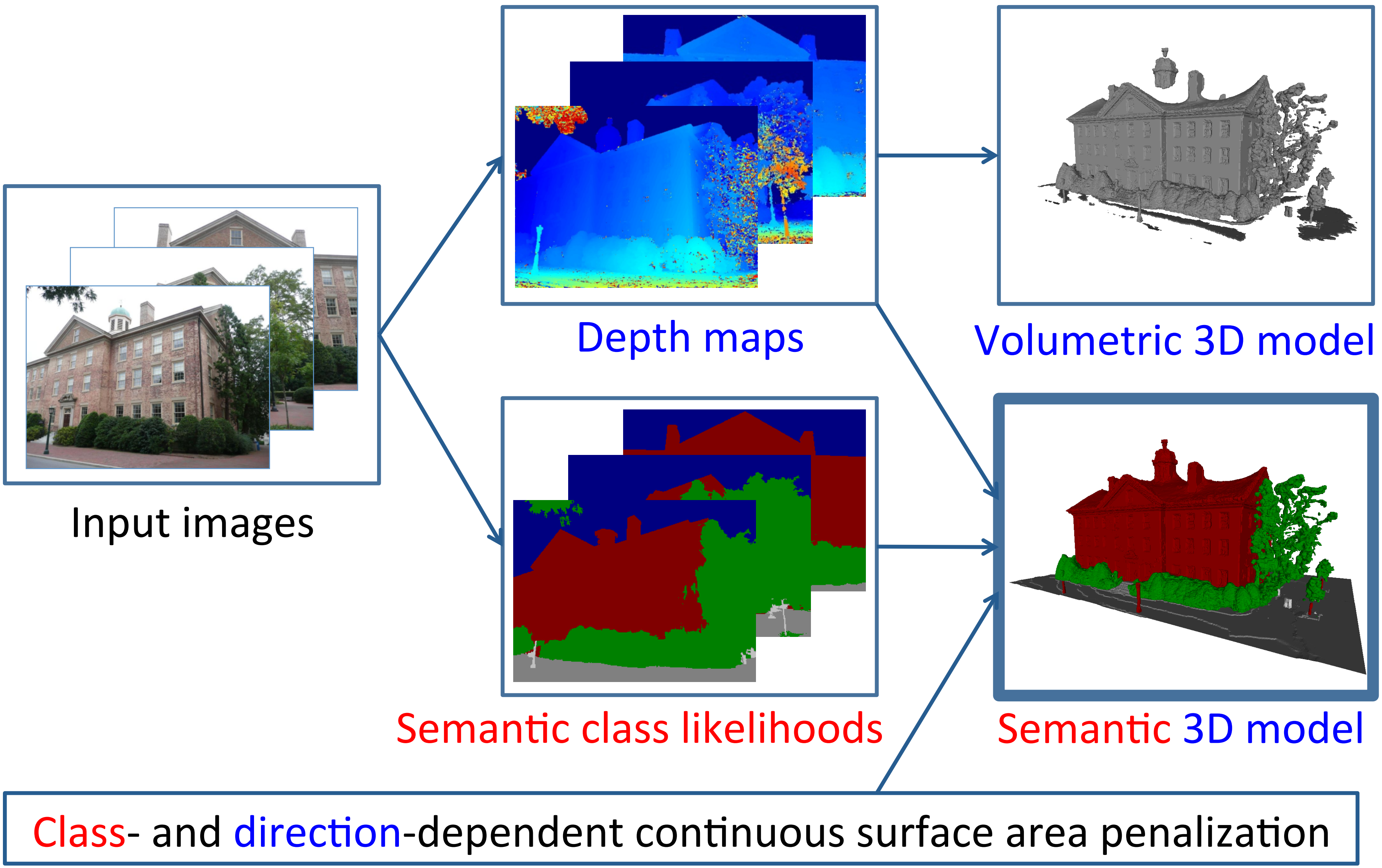
# Dense 3D Reconstruction



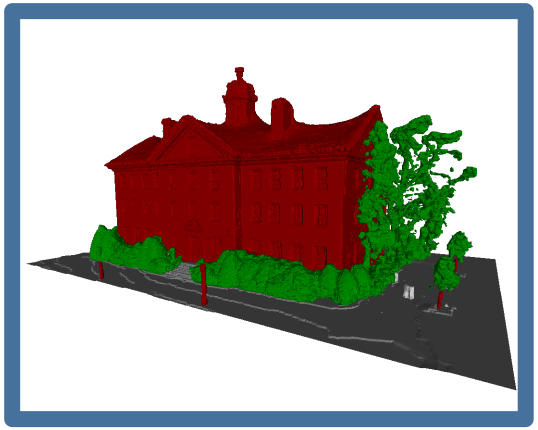
# “Painted” Dense 3D Reconstruction



# Semantic Dense 3D Reconstruction

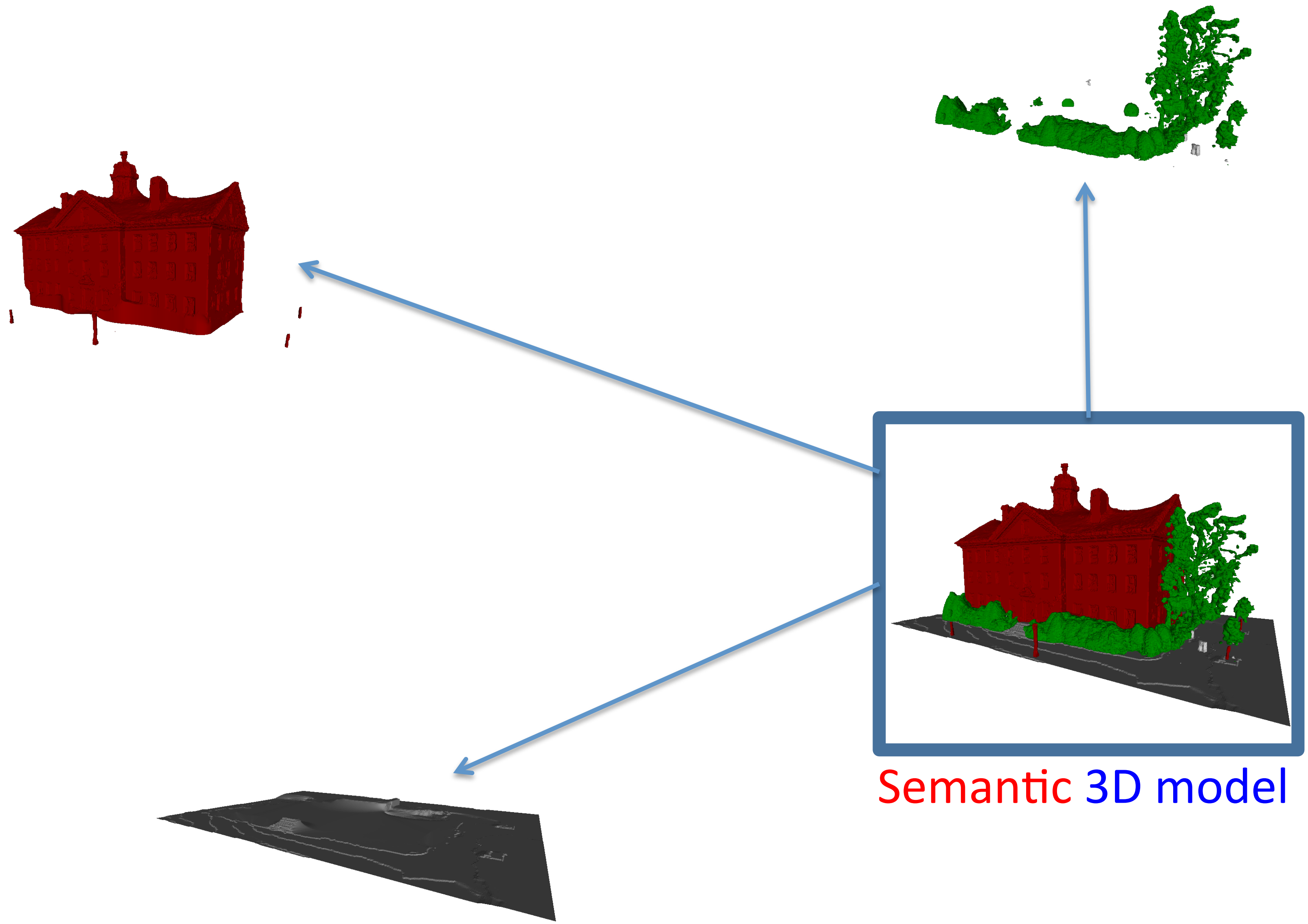


# Semantic Dense 3D Reconstruction

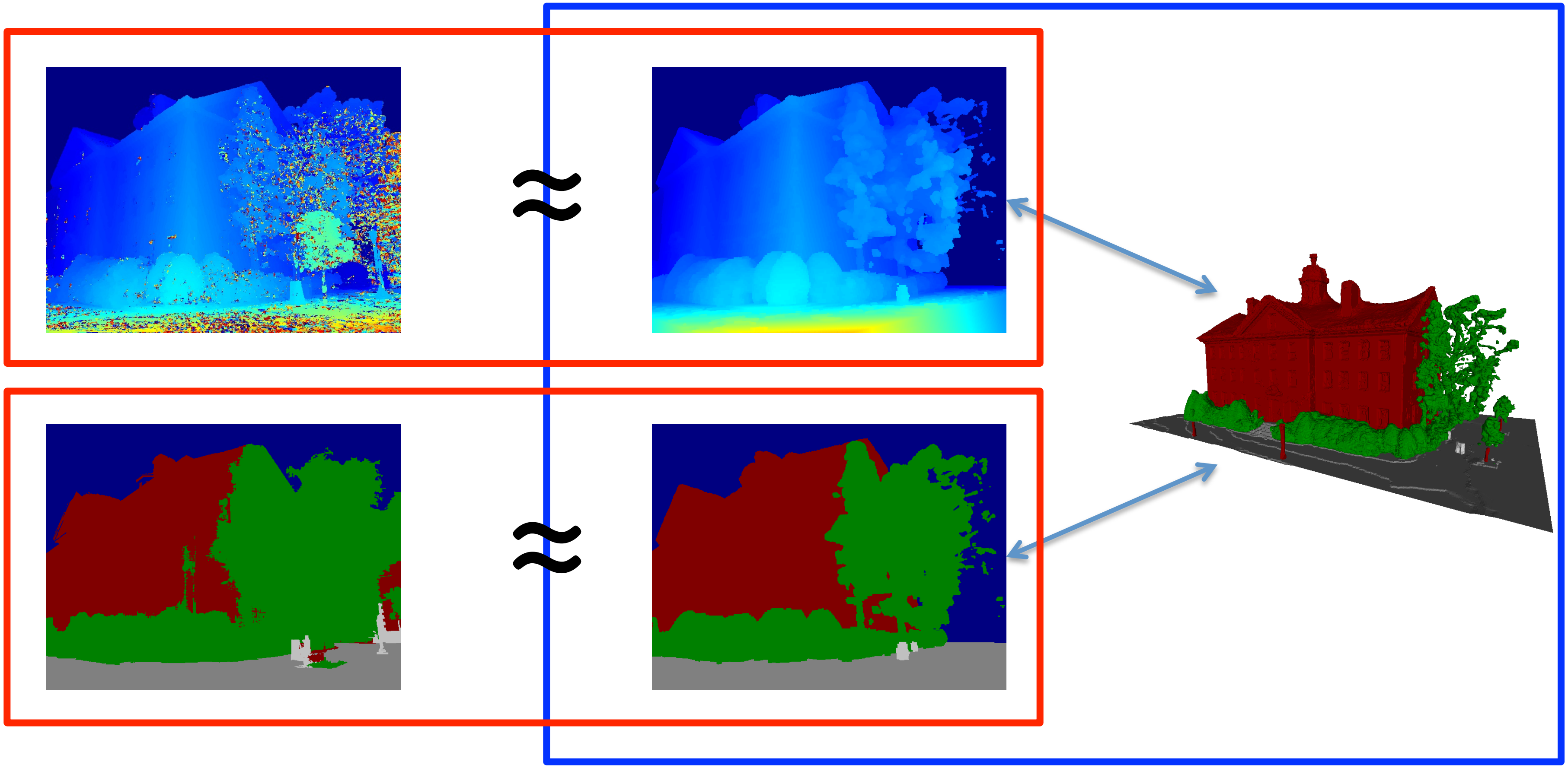


Semantic 3D model

# Semantic Dense 3D Reconstruction



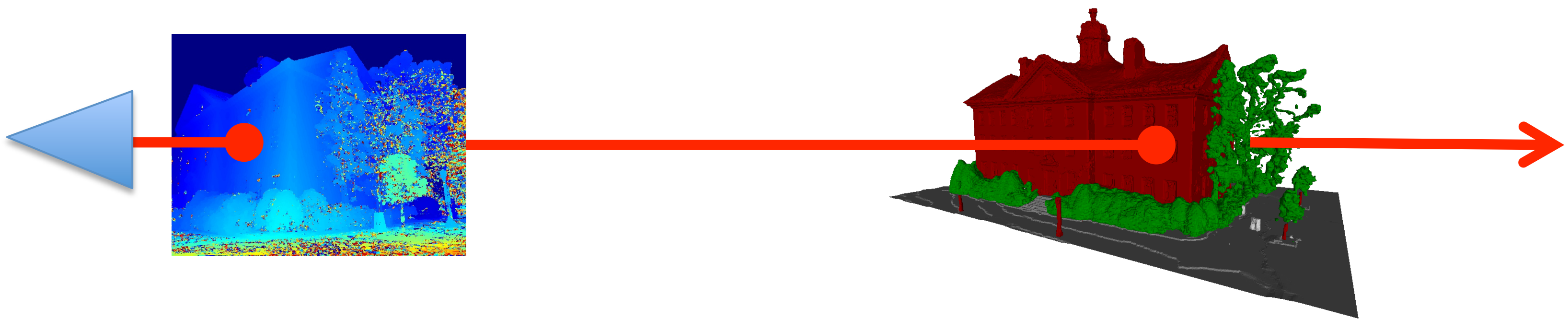
# “Ideal” formulation: data term in 2D, smoothness in 3D



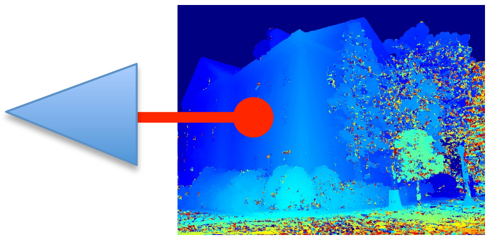

2D input data

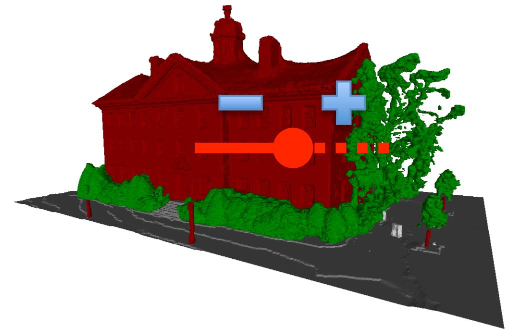
MAP solution

# “Ideal” data term: higher-order ray potentials



# Traditional data term: unary potentials



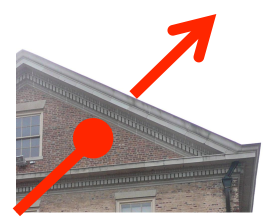


input

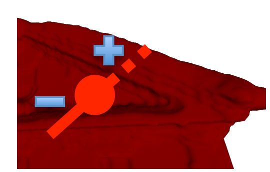
unary

Wrongly closed holes

input



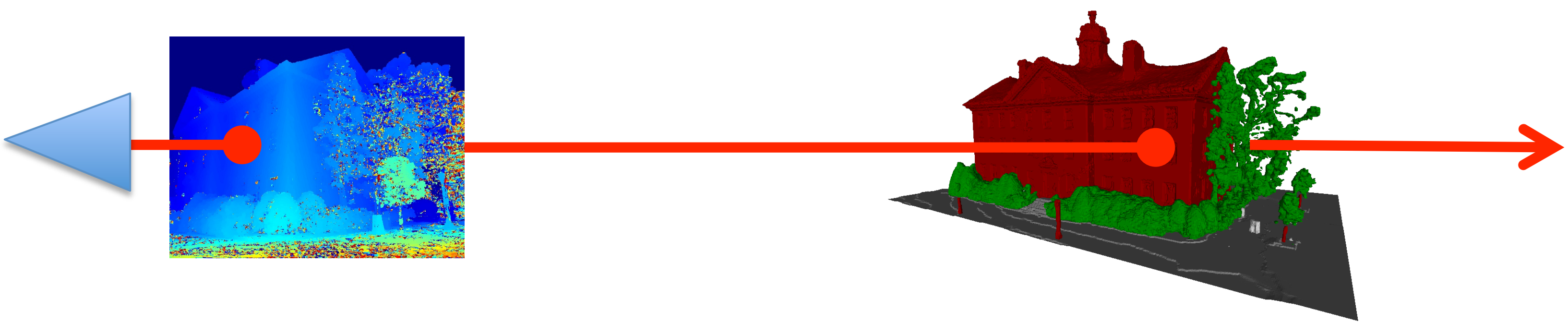
unary



Ridges/corners inflation

# Traditional data term: unary potentials

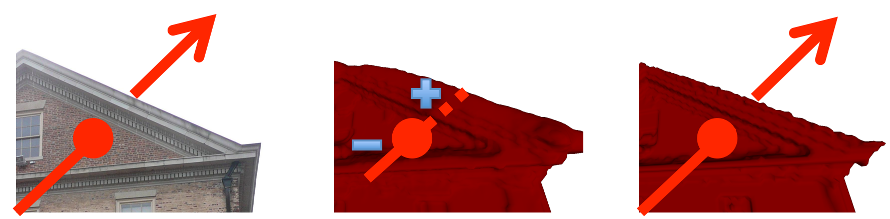
# We do: higher-order ray potentials



input unary ray

Fixing wrongly closed holes

input unary ray



Avoiding ridges/corners inflation

# Naïve approach to ray potentials

**Simple formulation**

Convex relaxation

# Naïve approach to ray potentials

**Simple formulation**

Convex relaxation

“Easy” to solve

# Naïve approach to ray potentials

**Simple formulation**

Convex relaxation

"Easy" to solve

Weak relaxation,
can't be used

# Convex relaxation (naïve)

$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} \left( \psi_S(\mathbf{x}) + \sum_{r \in \mathcal{R}} \sum_{\ell \in \mathcal{L}} \sum_{i=0}^{N_r} c_{ri}^{\ell} y_{ri}^{\ell} \right)$$

$$\text{s.t.} \quad y_{ri}^{\ell} \leq y_{r(i-1)}^{f}, \; y_{ri}^{\ell} \leq x_{s_{ri}}^{\ell}, \; y_{ri}^{\ell} \geq 0 \;\; \forall r, \forall \ell, \forall i$$

$$\sum_{\ell \in \mathcal{L}} x_s^{\ell} = 1, \qquad x_s^{\ell} \in [0,1].$$

Convex regularizer

Rays

Labels

Positions on ray

Cost taken if first visible label along ray r is $\ell$ at position i

Indicator if first visible label along ray r is $\ell$ at position i

Indicator if voxel s occupied with label $\ell$

# Convex relaxation (naïve)

$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} \left( \psi_S(\mathbf{x}) + \sum_{r\in\mathcal{R}} \sum_{\ell\in\mathcal{L}} \sum_{i=0}^{N_r} c_{ri}^{\ell} y_{ri}^{\ell} \right)$$

$$\text{s.t.} \quad y_{ri}^{\ell} \leq y_{r(i-1)}^{f}, \; y_{ri}^{\ell} \leq x_{s_{ri}}^{\ell}, \; y_{ri}^{\ell} \geq 0 \;\; \forall r, \forall \ell, \forall i$$

$$\sum_{\ell\in\mathcal{L}} x_s^{\ell} = 1, \qquad x_s^{\ell} \in [0,1].$$

Convex regularizer

Rays

Labels

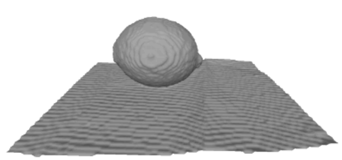
Positions on ray

Cost taken if first visible label along ray r is $\ell$ at position i

Indicator if first visible label along ray r is $\ell$ at position i

Indicator if voxel s occupied with label $\ell$

What we want 🙂



Lemon dataset

Desired (strong relaxation)

# Convex relaxation (naïve)

Convex regularizer

Rays

Labels

Positions on ray

Cost taken if first visible label along ray r is $\ell$ at position i

$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} \left( \psi_S(\mathbf{x}) + \sum_{r\in\mathcal{R}} \sum_{\ell\in\mathcal{L}} \sum_{i=0}^{N_r} c_{ri}^{\ell} y_{ri}^{\ell} \right)$$

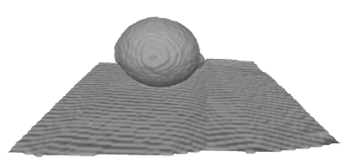
$$\text{s.t.} \quad y_{ri}^{\ell} \le y_{r(i-1)}^{f}, \; y_{ri}^{\ell} \le x_{s_{ri}}^{\ell}, \; y_{ri}^{\ell} \ge 0 \;\; \forall r, \forall \ell, \forall i$$

$$\sum_{\ell\in\mathcal{L}} x_s^{\ell} = 1, \qquad x_s^{\ell} \in [0,1].$$

Indicator if first visible label along ray r is $\ell$ at position i

Indicator if voxel s occupied with label $\ell$

What we get



Lemon dataset

**Obtained (weak relaxation)**

# Why so weak?

# Simple example (no regularizer)

$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} \left(-2y_0^o - 3y_1^o - 2y_2^o\right)$$

One ray with 3 positions

$$\text{s.t.} \quad y_i^o \leq y_{i-1}^f,\ y_i^f \leq y_{i-1}^f,$$
$$y_i^o \leq x_i^o,\ y_i^f \leq 1 - x_i^o,$$
$$x_i^o \in [0, 1], \quad \forall i.$$

Two labels: occupied o, free f

# Simple example (no regularizer)

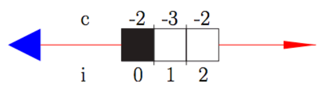
$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} \left(-2y_0^o - 3y_1^o - 2y_2^o\right)$$

$$\text{s.t.} \quad y_i^o \le y_{i-1}^f, \; y_i^f \le y_{i-1}^f,$$
$$y_i^o \le x_i^o, \; y_i^f \le 1 - x_i^o,$$
$$x_i^o \in [0,1], \quad \forall i.$$

One ray with 3 positions

Two labels: occupied o, free f

Desired solution — taking best-cost position $i = 1$ (cost $-3$):




$$x_0^o = 0, \; x_1^o = 1, \; x_2^o = 0, \qquad y_0^o = 0, \; y_1^o = 1, \; y_2^o = 0, \qquad y_0^f = 1, \; y_1^f = 0, \; y_2^f = 0.$$

# Simple example (no regularizer)

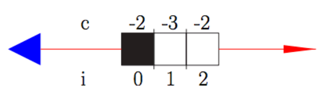
$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} (-2y_0^o - 3y_1^o - 2y_2^o)$$

One ray with 3 positions

$$\text{s.t.} \quad y_i^o \le y_{i-1}^f,\ y_i^f \le y_{i-1}^f,$$
$$y_i^o \le x_i^o,\ y_i^f \le 1 - x_i^o,$$
$$x_i^o \in [0,1], \quad \forall i.$$

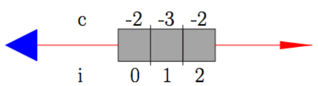
Two labels: occupied o, free f

Desired solution — taking best-cost position $i = 1$ (cost $-3$):




$$x_0^o = 0,\ x_1^o = 1,\ x_2^o = 0, \qquad y_0^o = 0,\ y_1^o = 1,\ y_2^o = 0, \qquad y_0^f = 1,\ y_1^f = 0,\ y_2^f = 0.$$

Weak relaxation solution (cost $-\mathbf{3.5}$):




$$x_0^o = x_1^o = x_2^o = y_0^o = y_1^o = y_2^o = y_0^f = y_1^f = y_2^f = \mathbf{0.5}.$$

## Non-convex constraint necessary for ray potentials

**Our formulation**

Convex relaxation

Non-convex constraint

Can still be efficiently solved in practice

Strong relaxation

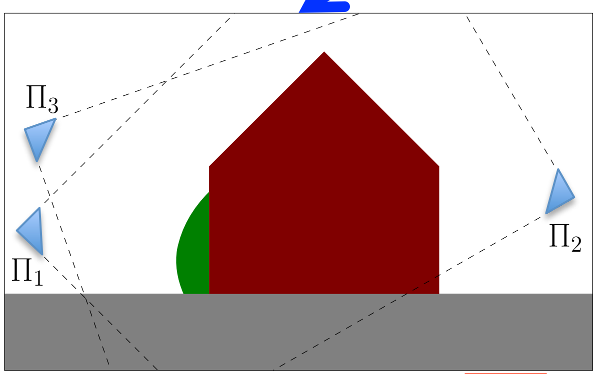
# Non-convex constraint necessary for ray potentials

**Our formulation**

Convex relaxation

Non-convex constraint

Visibility consistency constraint:
local views consistent with global view




Global view X


Local views from cameras $\Pi_1$ and $\Pi_2$ Y

# Simple example fixed by visibility consistency constraint!

$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} (-2y_0^o - 3y_1^o - 2y_2^o)$$

$$\text{s.t.} \quad y_i^o \leq y_{i-1}^f, \; y_i^f \leq y_{i-1}^f,$$
$$y_i^o \leq x_i^o, \; y_i^f \leq 1 - x_i^o,$$
$$x_i^o \in [0, 1], \quad \forall i.$$

One ray with 3 positions

Two labels: occupied o, free f

$$y_i^o \leq max(0, y_{i-1}^f - x_i^f), \forall i$$

# Simple example fixed by visibility consistency constraint!

$$(\mathbf{x}^*, \mathbf{y}^*) = \arg\min_{(\mathbf{x},\mathbf{y})} (-2y_0^o - 3y_1^o - 2y_2^o)$$

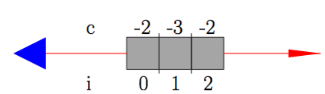
One ray with 3 positions

$$\text{s.t.} \quad y_i^o \leq y_{i-1}^f, \; y_i^f \leq y_{i-1}^f,$$
$$y_i^o \leq x_i^o, \; y_i^f \leq 1 - x_i^o,$$
$$x_i^o \in [0,1], \quad \forall i.$$

Two labels: occupied o, free f

$$y_i^o \leq max(0, y_{i-1}^f - x_i^f), \forall i$$

Weak relaxation solution (cost $-\mathbf{3.5}$):




$$x_0^o = x_1^o = x_2^o = y_0^o = y_1^o = y_2^o = y_0^f = y_1^f = y_2^f = \mathbf{0.5}.$$

# Properties of $y_i^o \leq max(0, y_{i-1}^f - x_i^f), \forall i$

- Generalizes to **multi-label**: $y_i^o \rightarrow \sum_{\ell \in \mathcal{L} \setminus \{f\}} y_i^\ell$

# Properties of $y_i^o \leq max(0, y_{i-1}^f - x_i^f), \forall i$

- Generalizes to **multi-label**: $y_i^o \rightarrow \sum_{\ell \in \mathcal{L} \setminus \{f\}} y_i^\ell$
- **Optimizable** with majorize-minimize: take the active branch of max operator (0 or linear), solve, take the active branch etc.

# Properties of $y_i^o \leq max(0, y_{i-1}^f - x_i^f), \forall i$

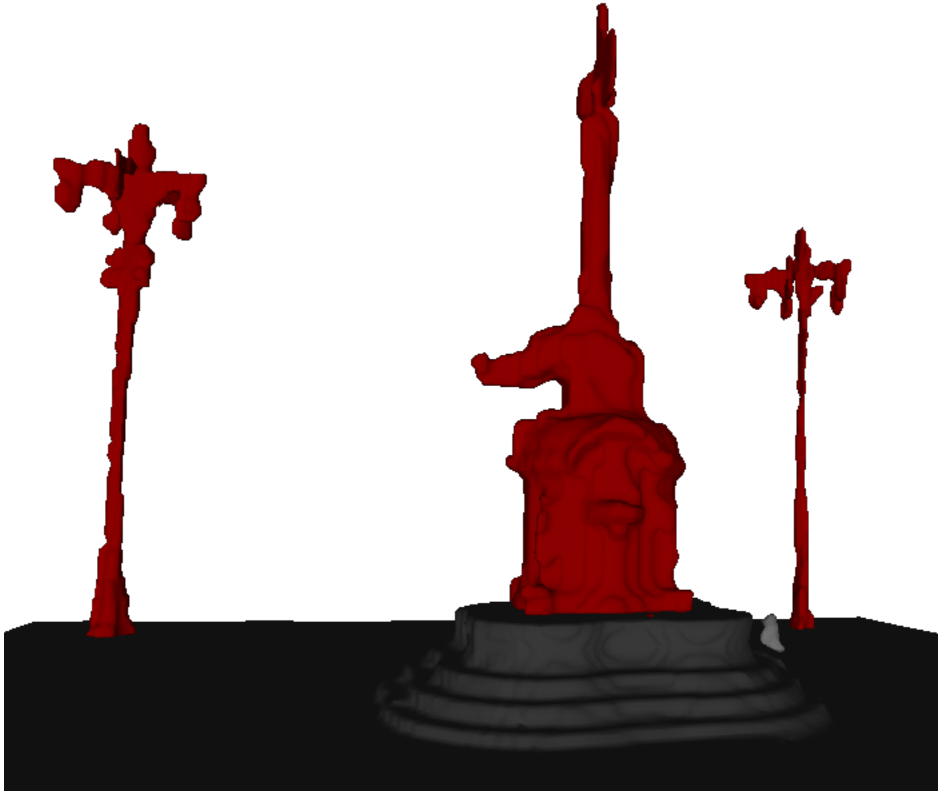
- Generalizes to **multi-label**: $y_i^o \to \sum_{\ell \in \mathcal{L} \setminus \{f\}} y_i^\ell$
- **Optimizable** with majorize-minimize: take the active branch of max operator (0 or linear), solve, take the active branch etc.
- **Provably tight** for 2-label without regularizer: constraint polyhedron is integral (totally-unimodular matrix)

# Properties of $y_i^o \leq max(0, y_{i-1}^f - x_i^f), \forall i$

- Generalizes to **multi-label**: $y_i^o \rightarrow \sum_{\ell \in \mathcal{L} \setminus \{f\}} y_i^\ell$
- **Optimizable** with majorize-minimize: take the active branch of max operator (0 or linear), solve, take the active branch etc.
- **Provably tight** for 2-label without regularizer: constraint polyhedron is integral (totally-unimodular matrix)
- **Provably 2-approximation** for 2-label without regularizer and 1-best position along the ray

# Our method

Input image example



Reconstruction

# Unary potentials [Häne et al., CVPR'13]

Input image example

Reconstruction

# We set state of the art on Middlebury MVS

- Best accuracy on 2 out of 6 datasets
- Completeness $\geq 99.5\%$ on 5 out of 6 datasets

**Multi-View Stereo** | **Evaluation** • Datasets • Submit • Code

Acc. Threshold: 90%
Comp. Threshold: 1.25 mm

Data in new window | Open Data Window
Data: View 1 and Ground Truth | Image Size: Small

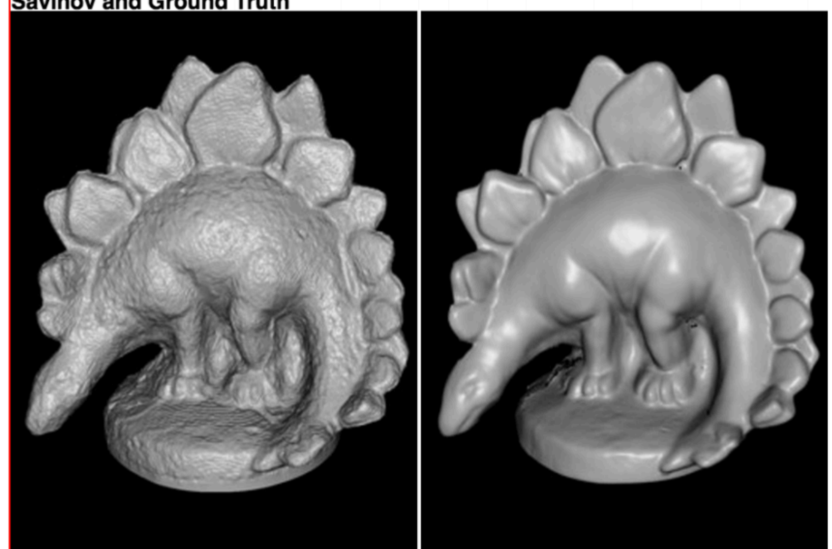
Tip: Mousing over any portion of a method's row will show its reference

| Sort By | Temple Full 312 views Acc [mm] | Comp [%] | Temple Ring 47 views Acc [mm] | Comp [%] | Temple Sparse 16 views Acc [mm] | Comp [%] | Dino Full 363 views Acc [mm] | Comp [%] | Dino Ring 48 views Acc [mm] | Comp [%] | Dino Sparse 16 views Acc [mm] | Comp [%] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Savinov | 0.41 | 99.7 | 0.5 | 99.5 | 0.69 | 97.8 | 0.26 | 99.8 | 0.25 | 99.9 | 0.34 | 99.7 |
| Furukawa 3 | 0.49 | 99.6 | 0.47 | 99.6 | 0.63 | 99.3 | 0.33 | 99.8 | 0.28 | 99.8 | 0.37 | 99.2 |
| DCV | | | 0.73 | 98.2 | 0.66 | 97.3 | | | 0.28 | 100 | 0.3 | 100 |
| Galliani | 0.39 | 99.2 | 0.48 | 99.1 | 0.53 | 97.0 | 0.31 | 99.9 | 0.3 | 99.4 | 0.38 | 98.6 |
| ECCV2016_624 | 0.37 | 98.9 | 0.49 | 97.6 | 1.27 | 39.2 | 0.26 | 97.8 | 0.31 | 99.5 | 0.28 | 98.1 |
| 3DV2014_25 | | | 0.51 | 96.4 | 1.23 | 90.2 | | | 0.32 | 97.3 | 0.42 | 96.7 |
| Furukawa 2 | 0.54 | 99.3 | 0.55 | 99.1 | 0.62 | 99.2 | 0.32 | 99.9 | 0.33 | 99.6 | 0.42 | 99.2 |
| Schroers | 0.57 | 99.1 | 0.64 | 96.4 | 2.12 | 62.9 | 0.33 | 99.7 | 0.33 | 99.7 | 0.54 | 98.6 |
| Kostrikov | | | 0.57 | 99.1 | 0.79 | 95.8 | | | 0.35 | 99.6 | 0.37 | 99.3 |
| Yichao Li | 0.46 | 96.4 | 0.56 | 89.6 | | | 0.4 | 94.9 | 0.37 | 80.6 | | |
| Song | | | 0.61 | 98.3 | | | | | 0.38 | 99.4 | 0.54 | 95.5 |
| Khuboni | | | 0.67 | 98.3 | | | | | 0.38 | 99.5 | | |
| Zhu | | | 0.4 | 99.2 | 0.45 | 95.7 | | | 0.38 | 98.3 | 0.48 | 95.4 |

**Reference:** N. Savinov, C. Haene, L. Ladicky, and M. Pollefeys. Semantic 3D reconstruction with continuous regularization and ray potentials using a visibility consistency constraint. CVPR 2016.

**Normalized Time (H:M:S):** 43:12:00

**Savinov and Ground Truth**

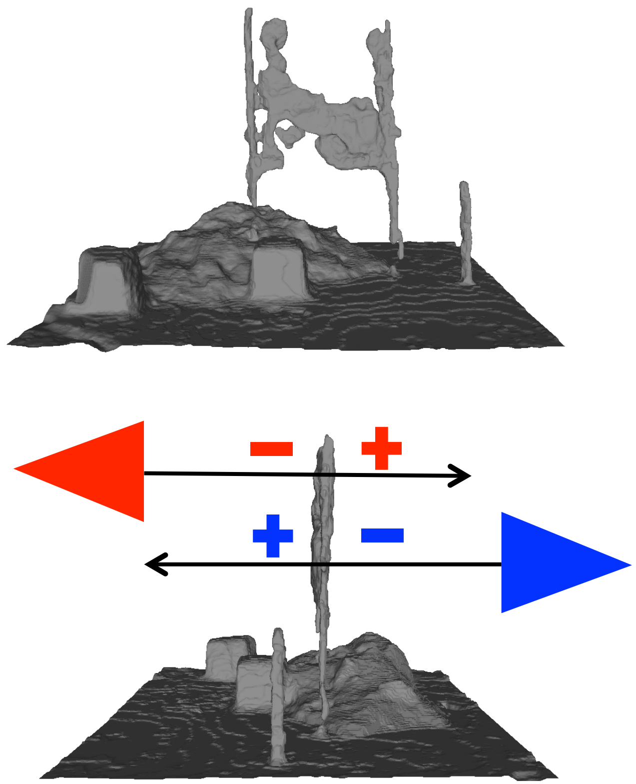


http://vision.middlebury.edu/mview/eval/

# Thin objects reconstructed properly (why?)

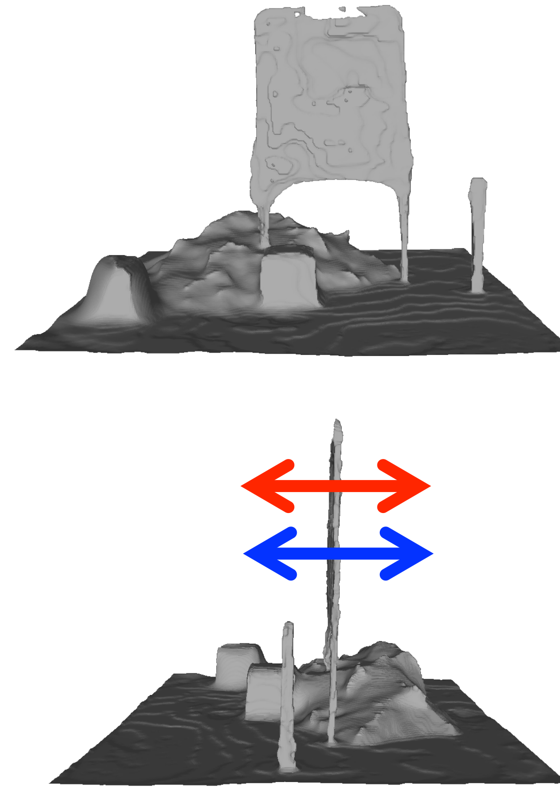
[dataset from Ummenhofer & Brox, ICCV2013]


Example images



Unary potentials

Data cancel out (bad)!



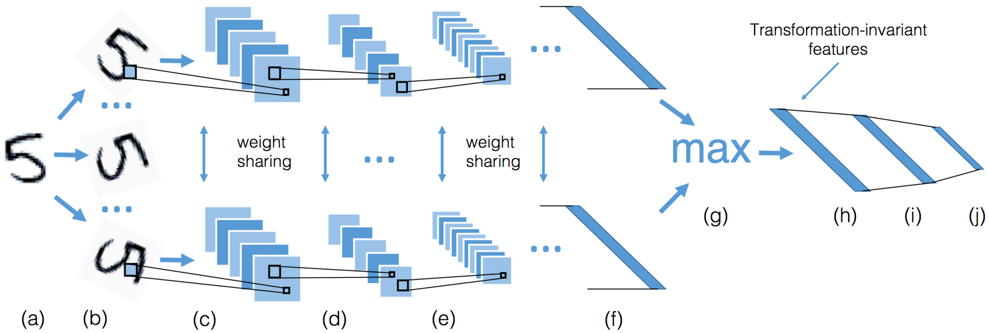
Our method

Data remain (good)!

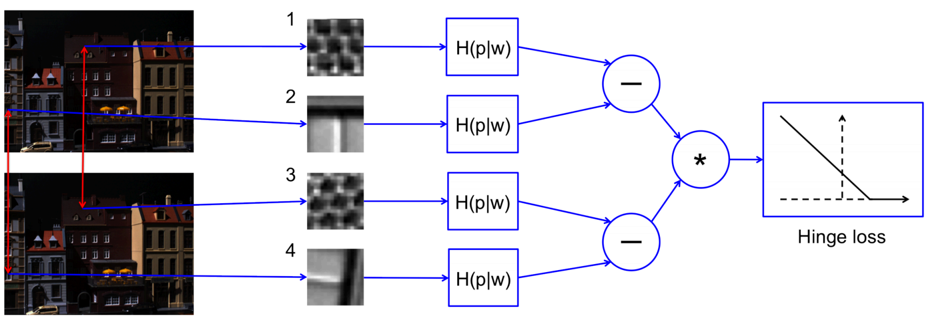
# Practical recommendations

- Don't use Ray Potentials for whole scenes
- Pick sub-box to reconstruct accurately
- Very good for thin surfaces and objects

# Advertisement (other papers)




TI-POOLING: transformation-invariant pooling for feature learning in Convolutional Neural Networks (CVPR'16)




Quad-networks: unsupervised learning to rank for interest point detection (under submission to CVPR'17, on arxiv)

Questions?