Bayesian 3D Modeling from Images using Multiple Depth Maps

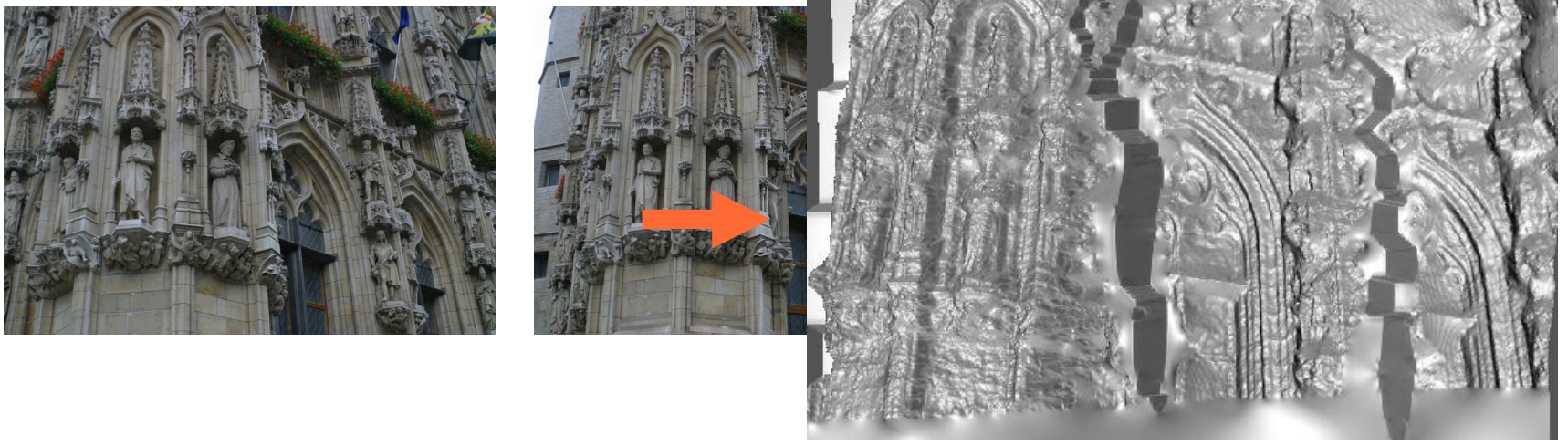
Pau Gargallo and Peter Sturm

INRIA Rhône-Alpes

France

The Problem

- Compute a 3D representation of the geometry and reflectance of a Lambertian scene from a collection of calibrated images.
- Ill-posed problem.
- Solutions need:
  - Scene representation
  - Prior knowledge
  - Search method



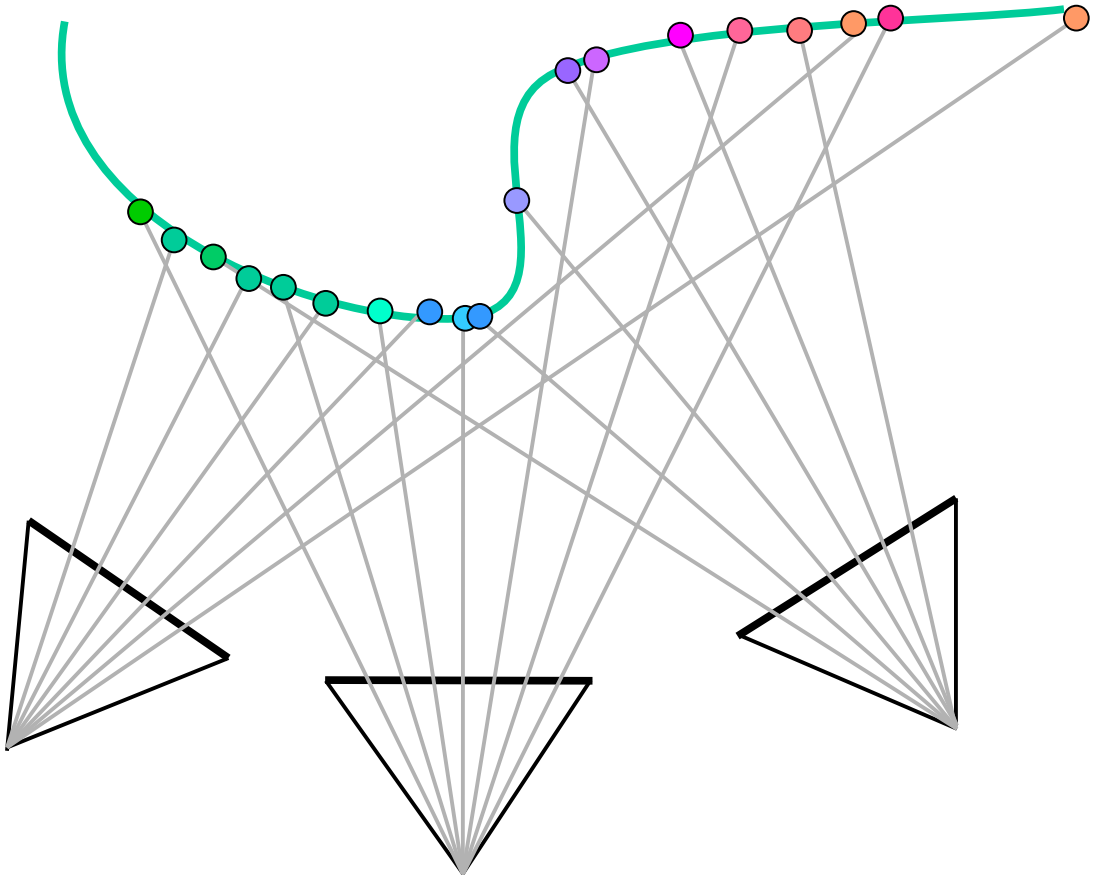
Outline

- Previous work.
- Model and relevant variables.
- Joint probability definition:
  - Likelihood.
  - Visibility prior.
  - Depth Maps prior.
- MAP optimization.
- Experimental results.

Previous Work

- Largely studied problem!
- From classical stereo methods computing disparities based on SSD or other [Scharstein02].
- To more recent approaches mainly based on energy minimization: $E = E_{data} + E_{model}$
  - Voxels: Voxel Carving [Kutulakos00], Graph Cuts [Paris04].
  - Level Sets [Keriven98].
  - Meshes [Fua93,Morris01,Vogiatzis03].
  - Depth Maps: Graph Cuts [Kolmogorov03], Belief Propagation [Sun03], other [Fitzgibbon03, Strecha04].

The Model: Multiple Colored Depth Maps



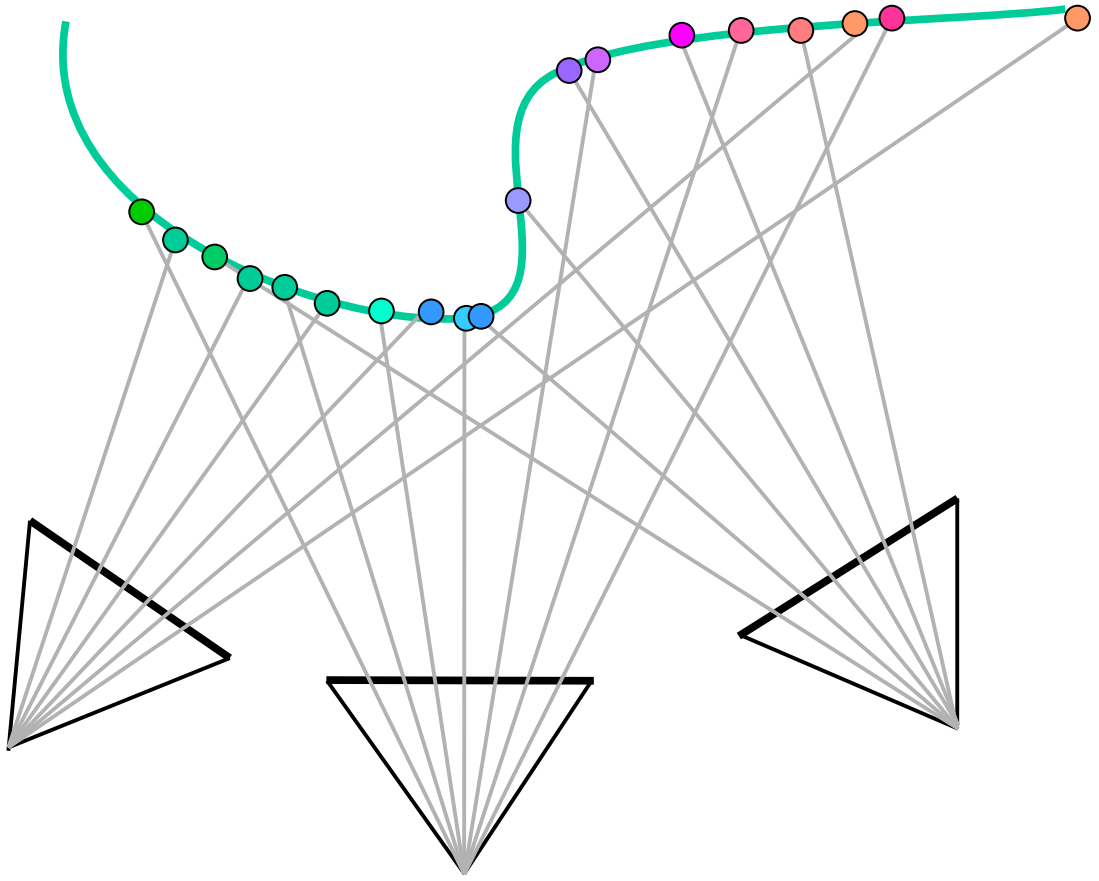
The Model: Multiple Colored Depth Maps

- One depth map per input image parameterizing a 3D point cloud.
- A color associated to every modeled 3D point.
  - Advantages:
    - All input is used.
    - Higher resolution than volumetric models.
    - Easier to deal with than meshes.
    - Occlusions can be efficiently detected.
  - Drawbacks:
    - High number of unknowns.
    - Coherence between different depth maps must be imposed.
    - The output is just a set of points, not a surface.

The Approach

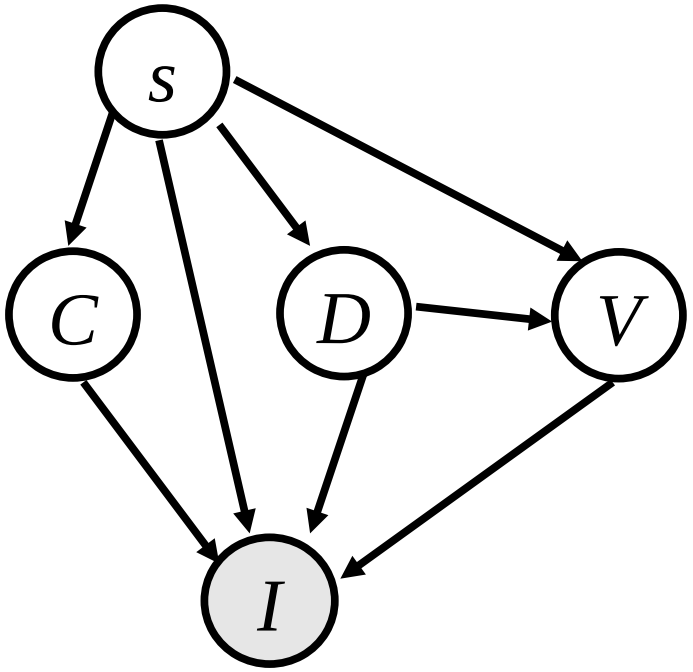
- Probabilistic framework:
  - Definition of the relevant random variables.
  - Decomposition of their joint probability density function.
  - Definition of each term in the decomposition:
    - Likelihood
    - Visibility prior
    - Depth Maps prior
  - Research of the *Maximum A Posteriori* using the EM-algorithm.
- Why probabilities?
  - To deal with noise, the missing data and *outliers*.
  - To be able to use specific optimization methods.

The Variables



- Depth Maps D
- Colors C
- Visibility V
- Input images I

The Decomposition



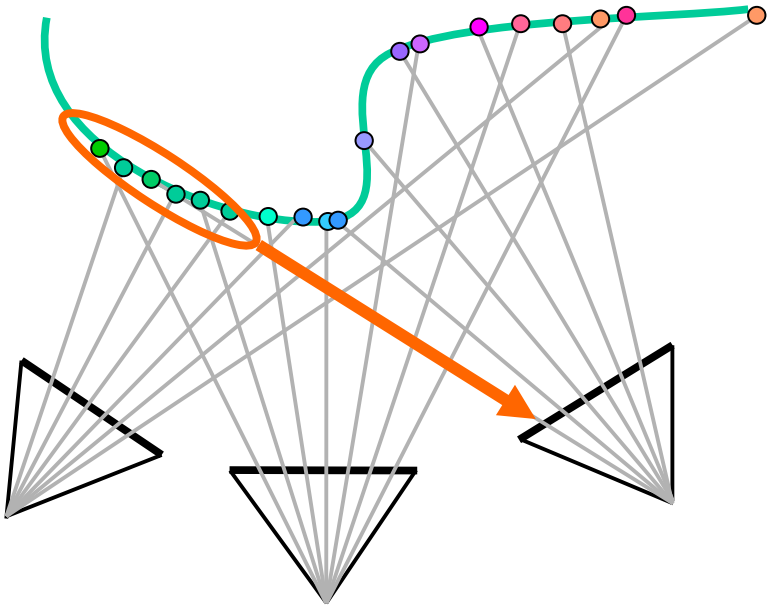
- Posteriori:

$$p(D,C,s \mid I) = \frac{1}{Z} \int p(I,C,D,s,V)dV$$

- $p(s)$ prior on the parameters.
- $p(C|s)$ prior on estimated colors.
- $p(D|s)$ depth maps prior.
- $p(V|D,s)$ visibility prior.
- $p(I|C,D,V,s)$ likelihood.

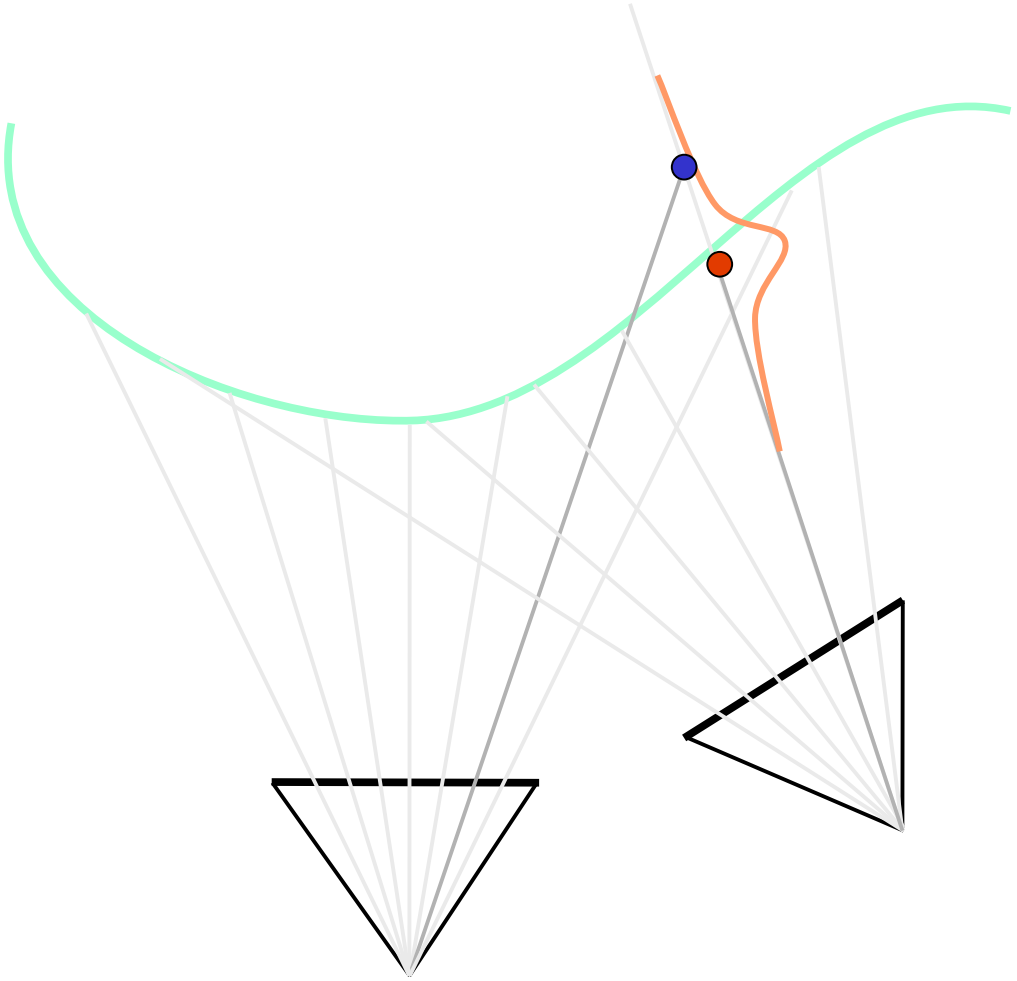
The Likelihood

- Measures the similarity between the model and input data.



- Desired:
  - Render model on all images
  - Compute likelihood per pixel
  - Hard to compute derivatives.
- Typical Approximation:
  - Compute likelihood per 3D point
  - Overuse of pixel information.
- Our approximation:
  - Distribute pixel information equally among 3D points.

Visibility Variables Prior



Geometric Prior

- Modeling all the depth maps at the same time enables geometric reasoning about visibility.
- A 3D point is likely to be visible in an image if its depth is similar to the one of the pixel it projects to [Szeliski99].

$$p(\mathcal{V}_{i,\mathbf{X}} = 1|\mathcal{D}) = v \exp\left(-\frac{(d_i(\mathbf{X}) - \mathcal{D}_i(\mathbf{X}))^2}{2\sigma^2}\right)$$

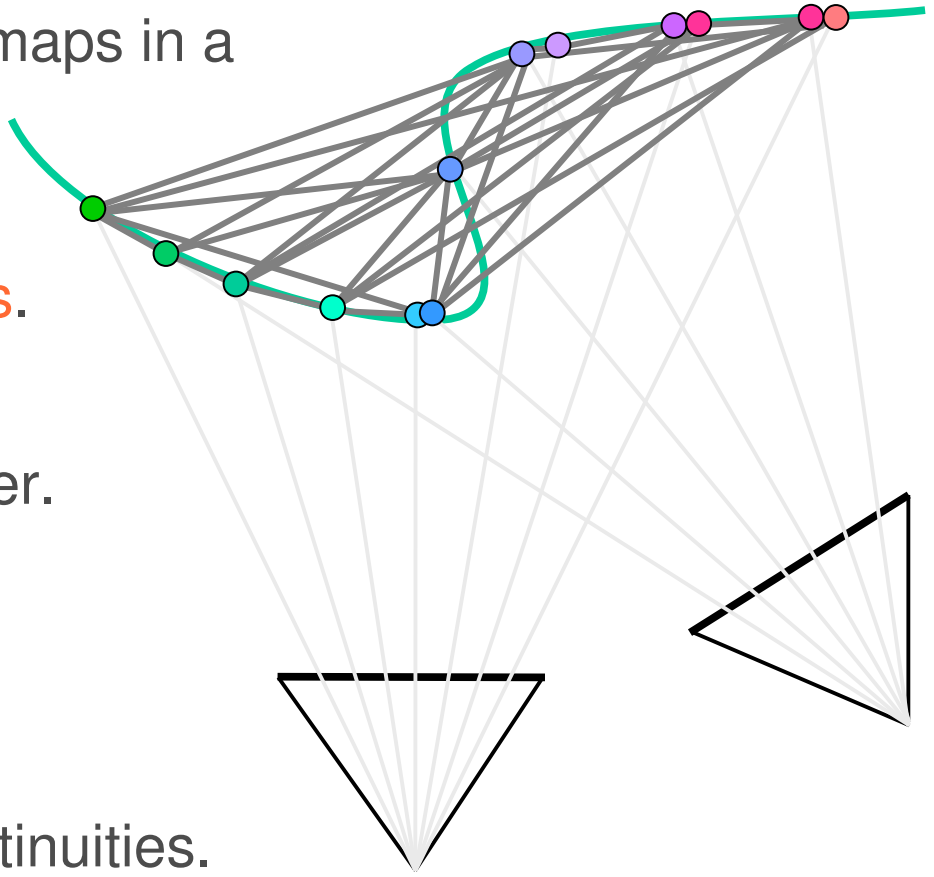
Posterior of Visibility

- Mixes prior with photometric evidence

$$f_{i,\mathbf{X}} = \frac{p(\mathcal{V}_{i,\mathbf{X}} = 1|\mathcal{D})\mathcal{N}}{p(\mathcal{V}_{i,\mathbf{X}} = 1|\mathcal{D})\mathcal{N} + p(\mathcal{V}_{i,\mathbf{X}} = 0|\mathcal{D})\mathcal{U}}$$

Multiple Depth Maps Prior

- Three requirements:
  - Integrate the different depth maps in a single surface.
  - Smooth this surface.
  - But allow large discontinuities.
- Idea:
  - Close points attract each other.
  - Far points do not interact.
- Formalism:
  - Complete Markov network.
  - Line process to signal discontinuities.



Optimization

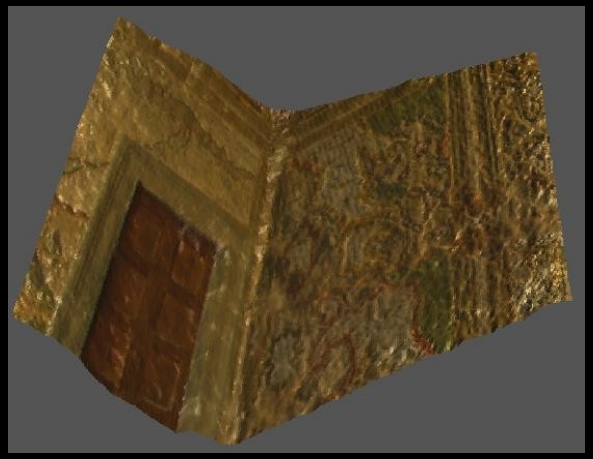
- Expectation-Maximization:
  - E-step: Compute expectation of visibility variables.

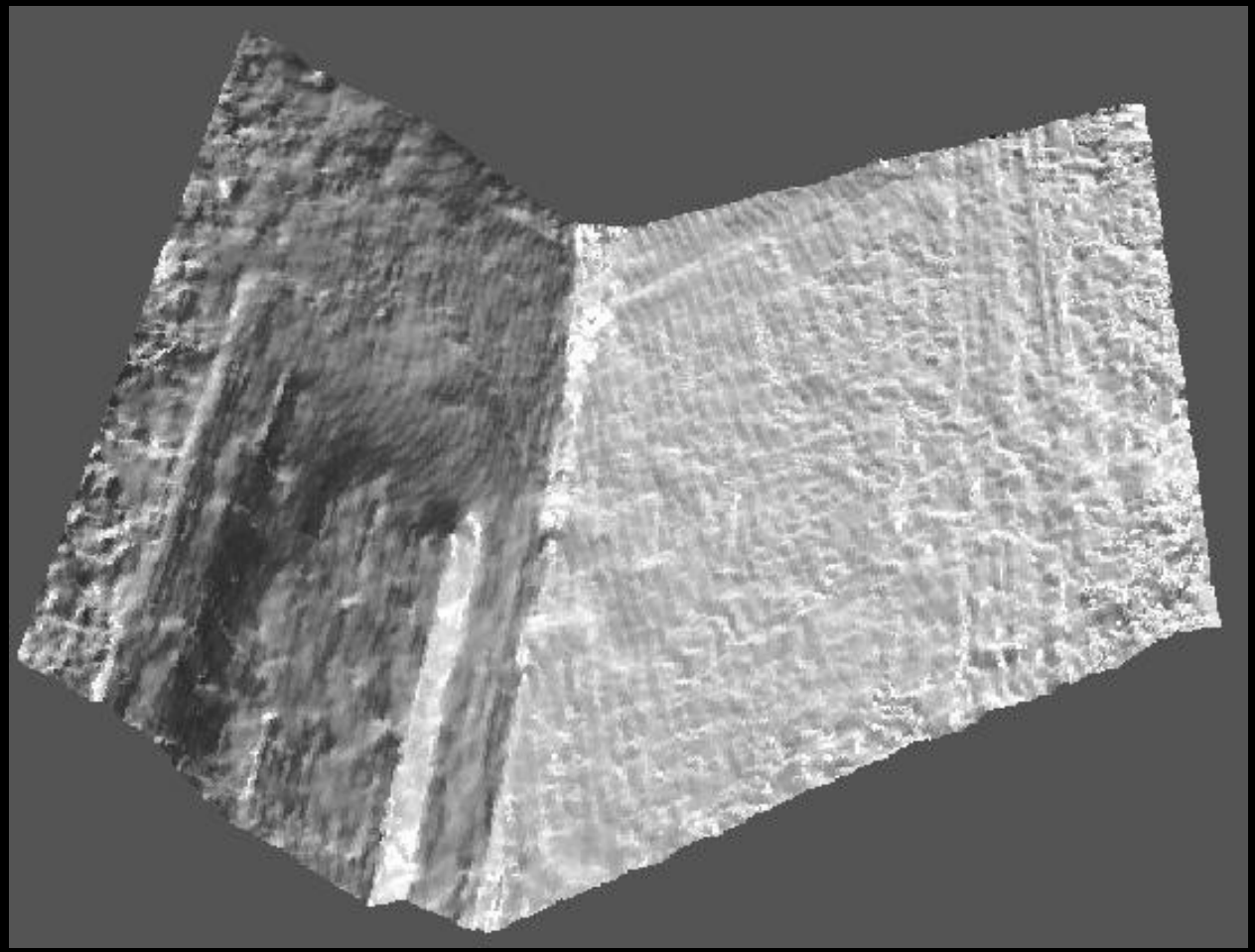
$$f_{i,\mathbf{X}} = \frac{p(\mathcal{V}_{i,\mathbf{X}} = 1|\mathcal{D})\mathcal{N}}{p(\mathcal{V}_{i,\mathbf{X}} = 1|\mathcal{D})\mathcal{N} + p(\mathcal{V}_{i,\mathbf{X}} = 0|\mathcal{D})\mathcal{U}}$$

  - M-step: Maximization of the expected log-*posteriori*.
    - Gradient descent.

- As hidden variables can be analytically marginalized, we can also optimize the posterior directly without using EM.
- To address the problem of local minima, *global* minimization methods should be used:
  - Loopy Belief Propagation or Simulated Annealing.

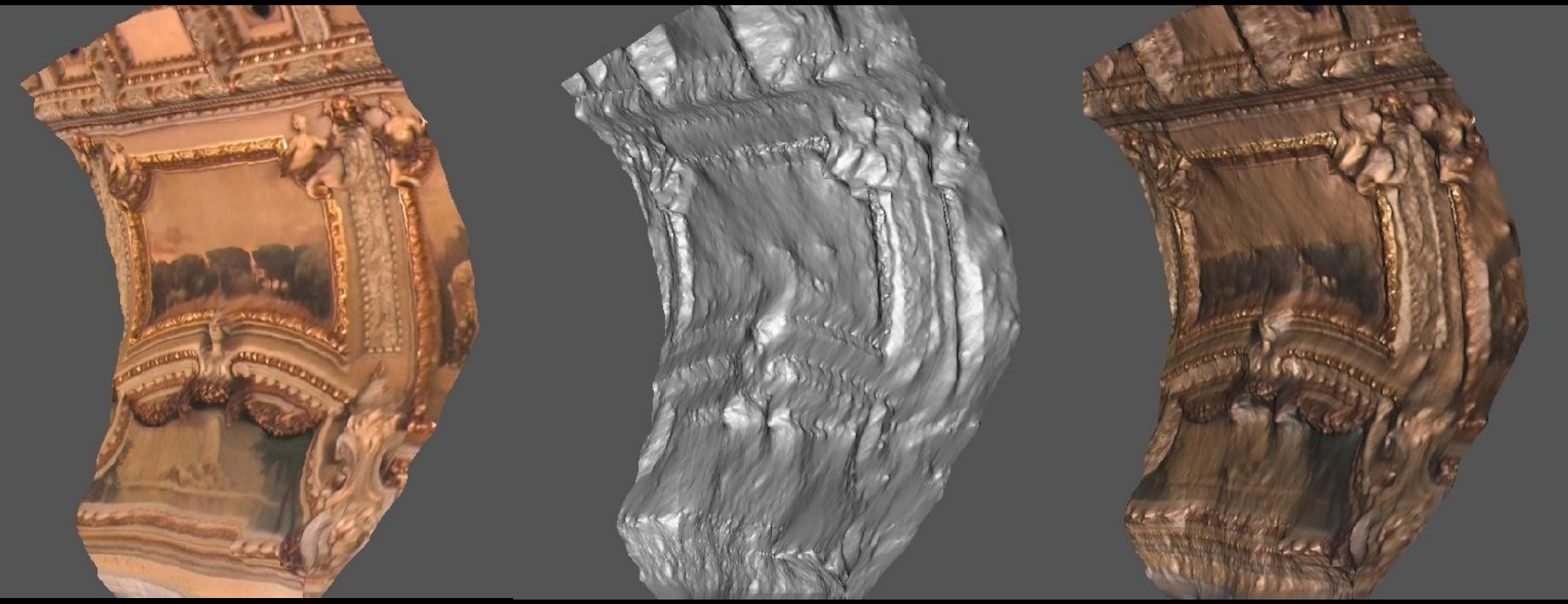












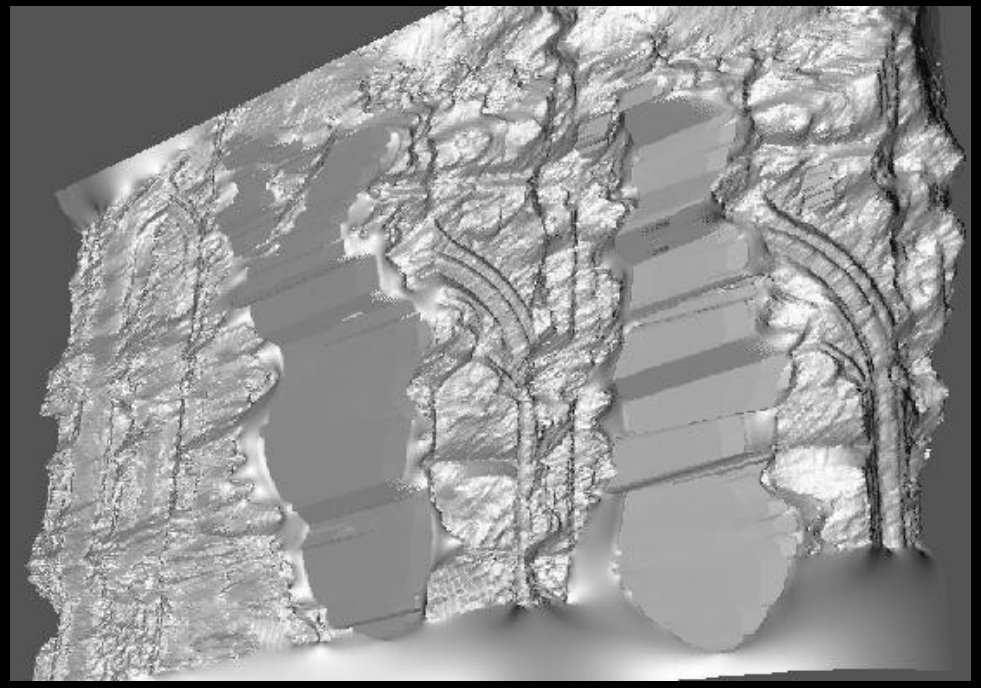












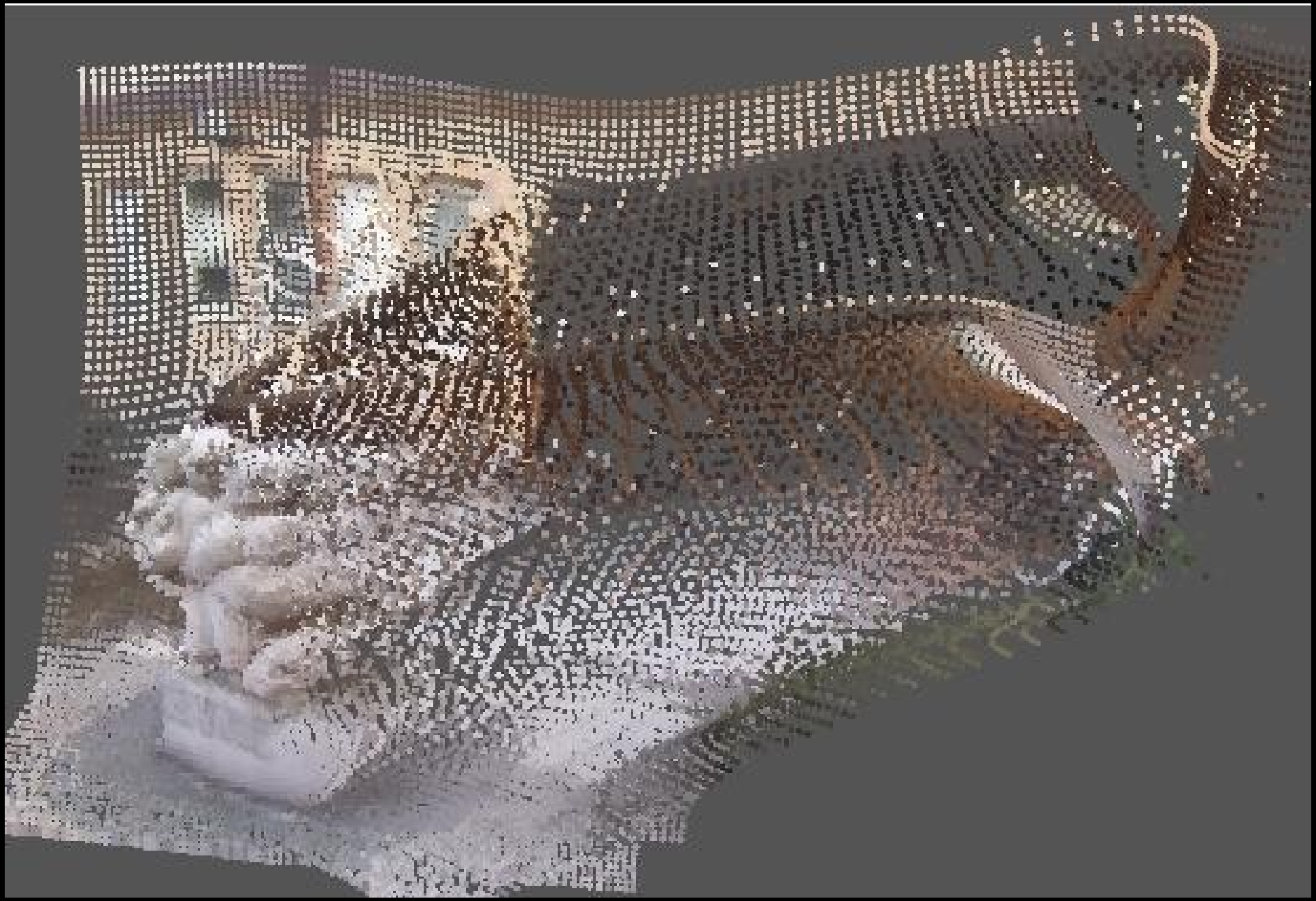

DEN OFFENTLIGA MAKTEN SKALL UTÖVAS
MED RESPEKT FÖR ALLA MÄNNISKORS
LIKA VÄRDE OCH FÖR DEN ENSKILDA
MÄNNISKANS FRIHET OCH VÄRDIGHET

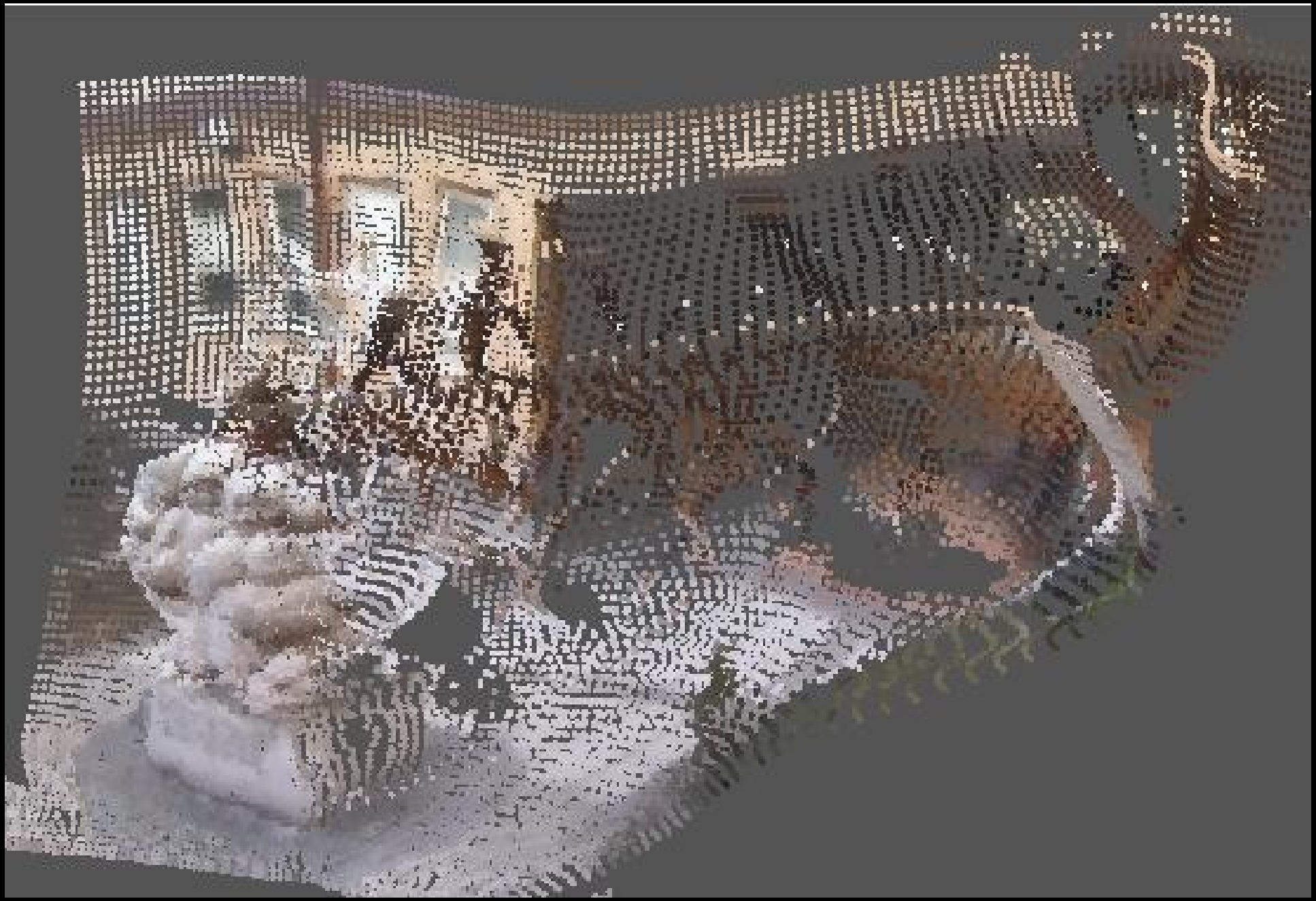

DEN OFFENTLIGA MAKTEN SKALL UTÖVAS
MED RESPEKT FÖR ALLA MÄNNISKORS
LIKA VÄRDE OCH FÖR DEN ENSKILDA
MÄNNISKANS FRIHET OCH VÄRDIGHET

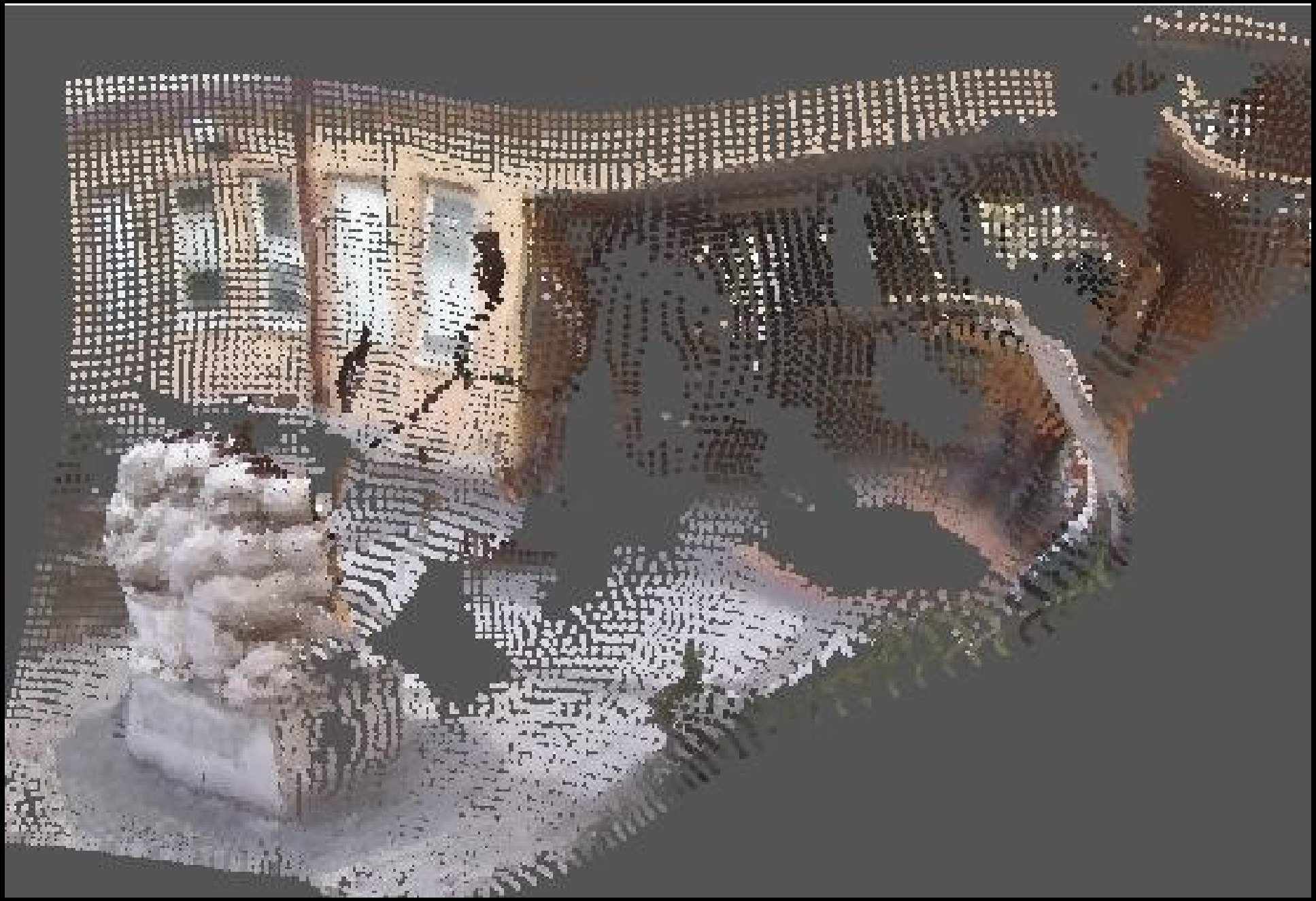

DEN OFFENTLIGA MAKTEN SKALL UTÖVAS
MED RESPEKT FÖR ALLA MÄNNISKORS
LIKA VÄRDE OCH FÖR DEN ENSKILDA
MÄNNISKANS FRIHET OCH VÄRDIGHET

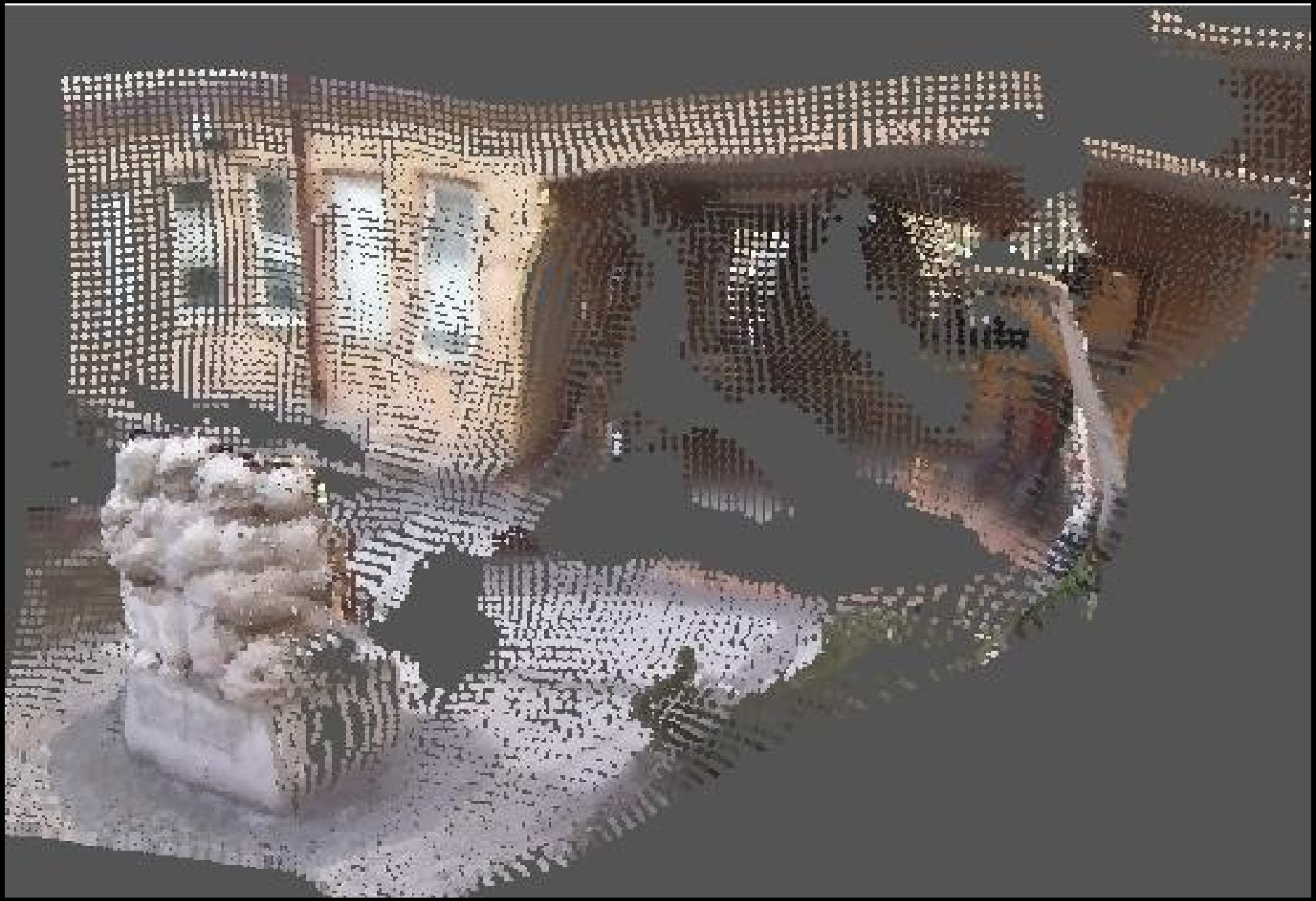

DEN OFFENTLIGA MAKTEN SKALL UTÖVAS
MED RESPEKT FÖR ALLA MÄNNISKORS
LIKA VÄRDE OCH FÖR DEN ENSKILDA
MÄNNISKANS FRIHET OCH VÄRDIGHET


DEN OFFENTLIGA MAKTEN SKALL UTÖVAS
MED RESPEKT FÖR ALLA MÄNNISKORS
LIKA VÄRDE OCH FÖR DEN ENSKILDA
MÄNNISKANS FRIHET OCH VÄRDIGHET





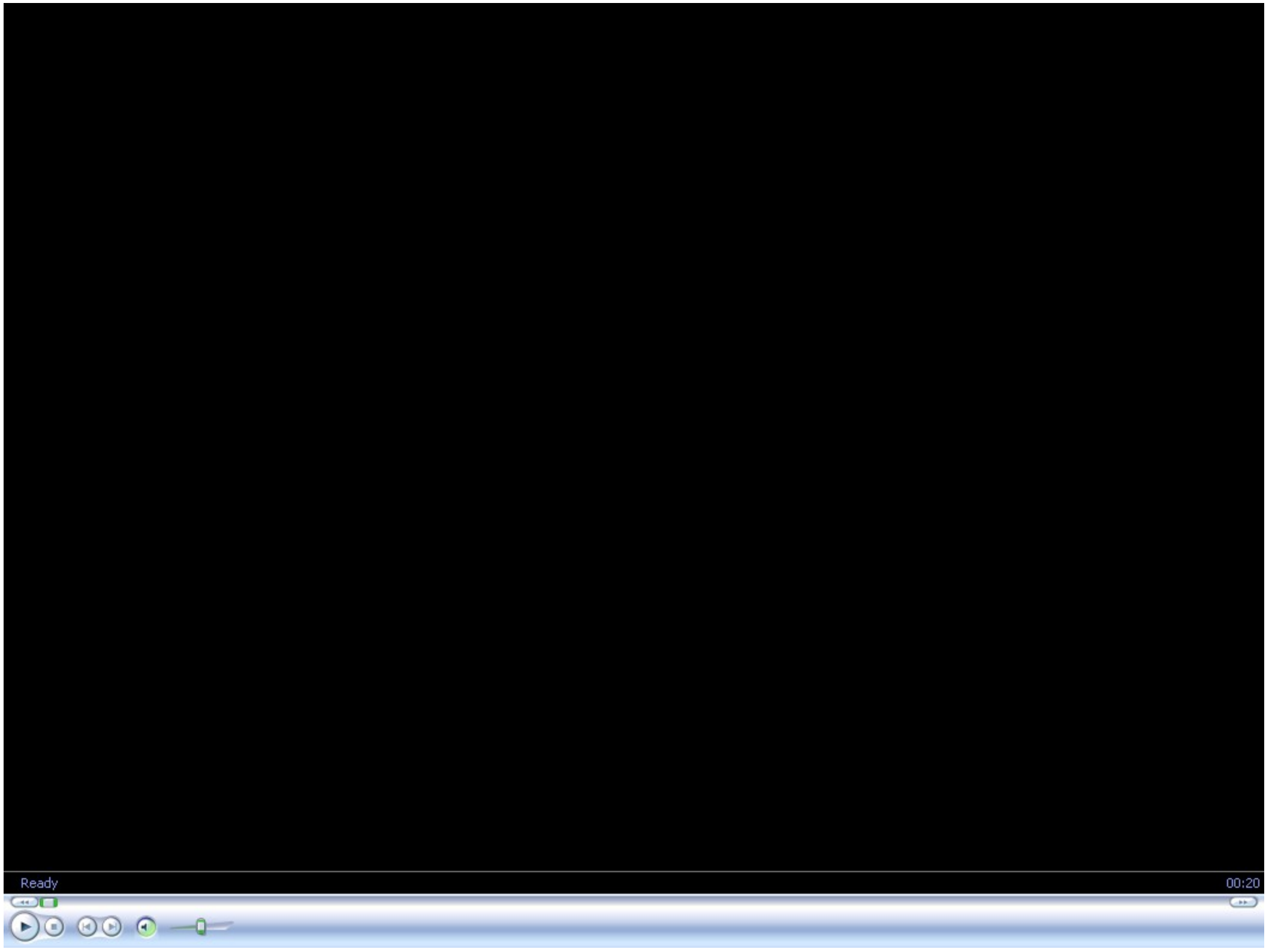




Conclusion

- Multiple colored depth maps as scene representation.
- Geometric occlusion reasoning.
- Prior on the set of depth maps.
- Needs global optimization.

Thank you very much!


Ready
00:20

