Multi-view stereo via volumetric graph-cuts

George Vogiatzis, Carlos Hernandez and Roberto Cipolla

Cambridge Research Laboratory
Toshiba Europe Research Laboratory

3D models from images

- State-of-the-art improving rapidly in recent years
- Various 3D cues have been exploited:
  - Silhouettes
  - Stereopsis (photo-consistency, multi-view stereo)
  - Shading (single/multiple images)
  - Texture
  - Defocus

Photo-consistency

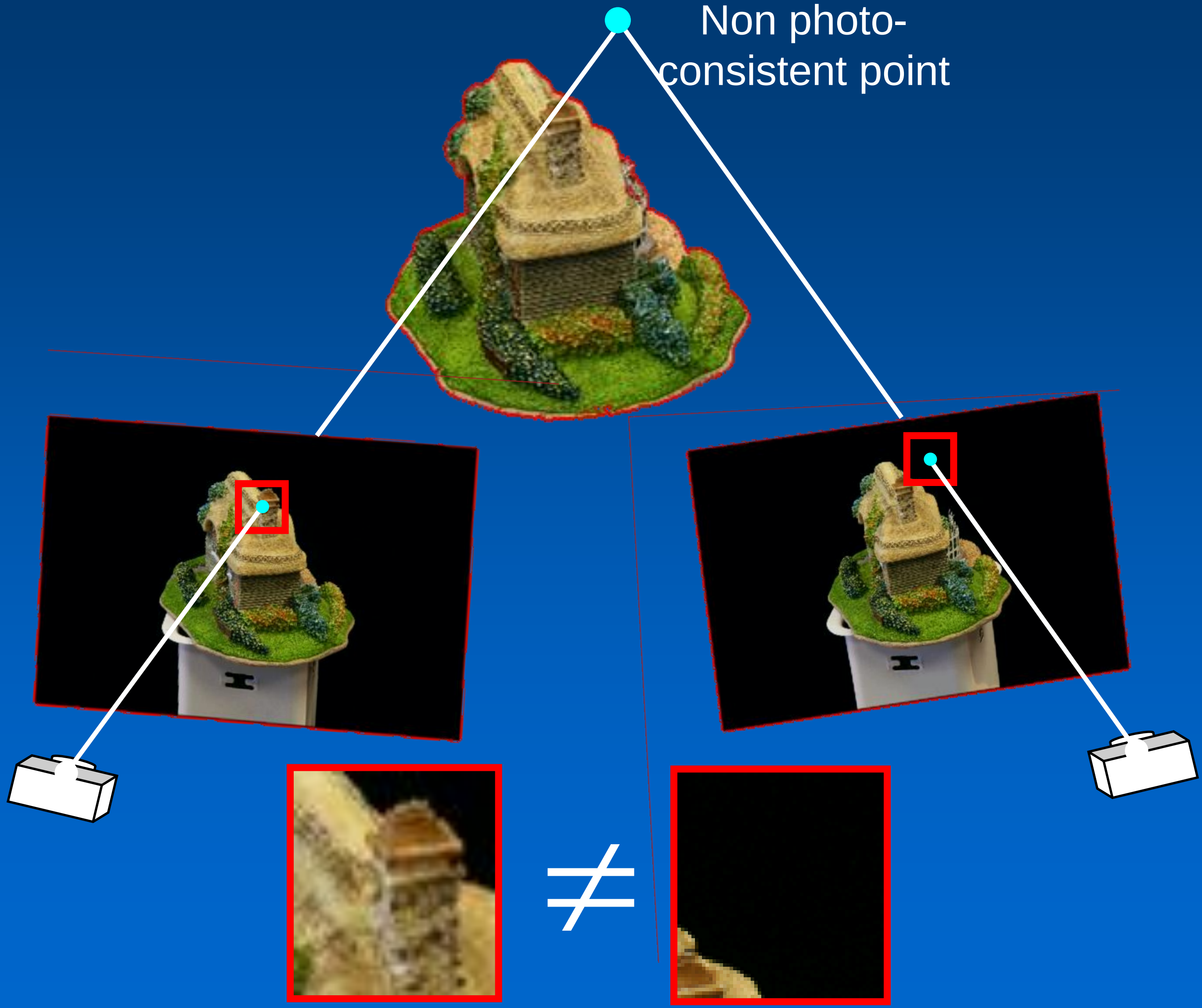
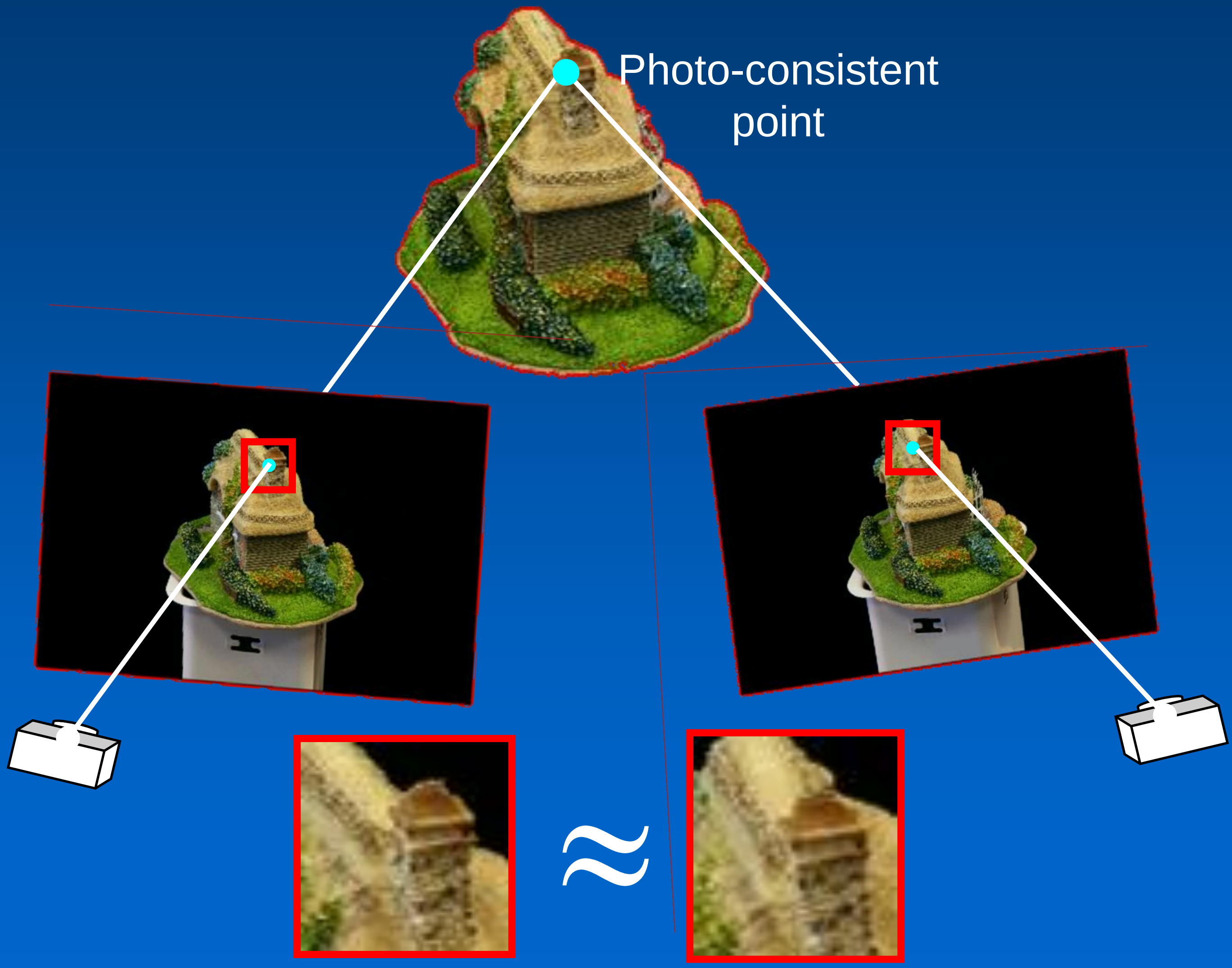
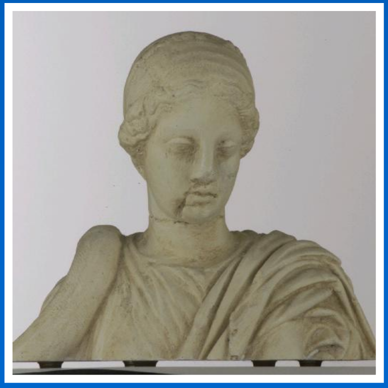


Photo-consistency



3D shape from *photo-consistency*

- Integrate this cue on surface
- Algorithms try to find the most *photo-consistent* 3D surface
- Key assumptions for object surface
  - Lambertian
  - Richly textured
  - Smooth

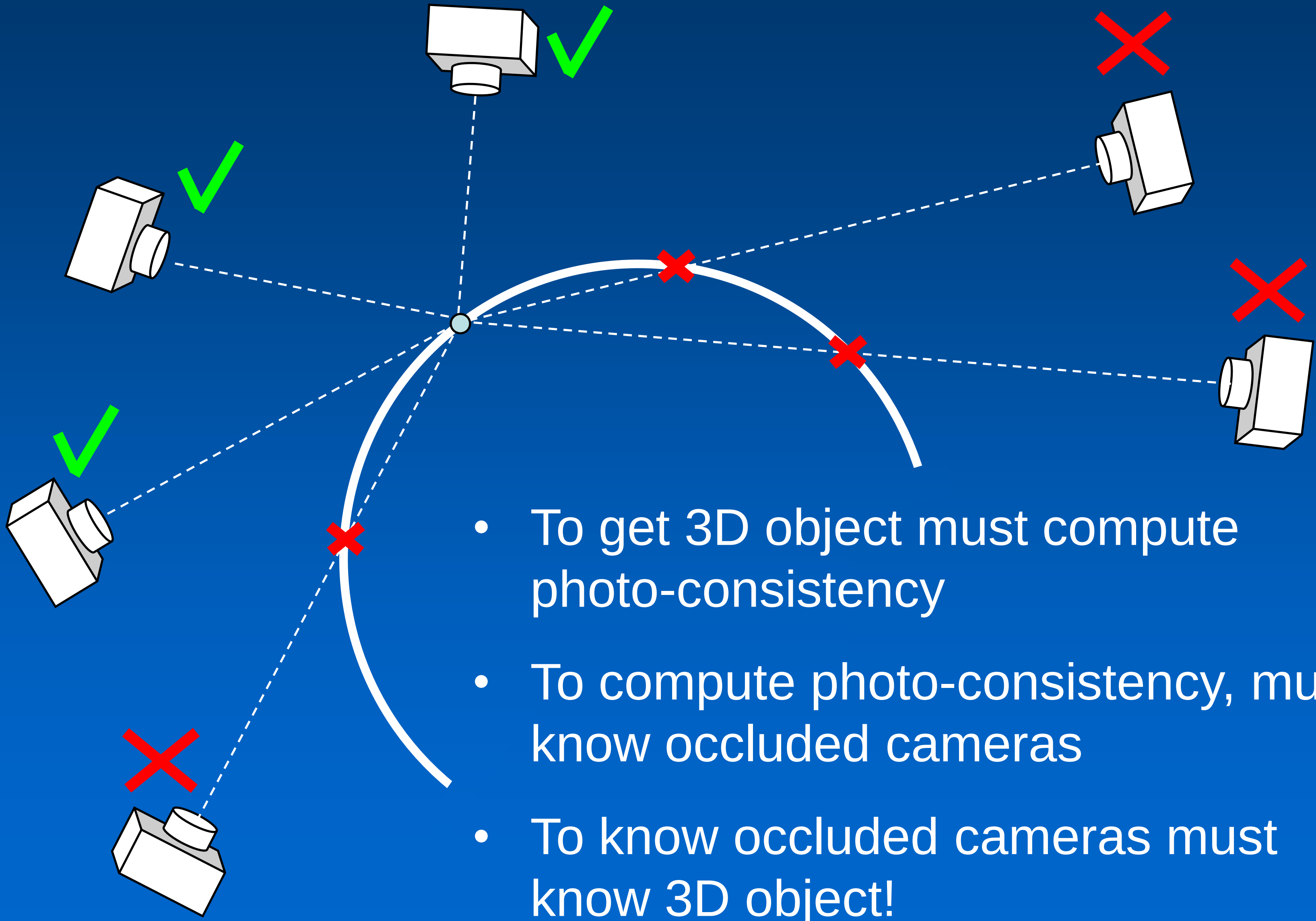

3D shape from *photo-consistency*

- Scene representation
  - Meshes, level-sets, voxel occupancy
- Shape prior
  - Local smoothness
- Photo-consistency metric
  - Normalised cross-correlation (NCC), sum of squared differences (SSD)
- Occlusion reasoning
  - How to determine visible images
- Reconstruction algorithm
  - How to obtain globally optimal solution

Our solution

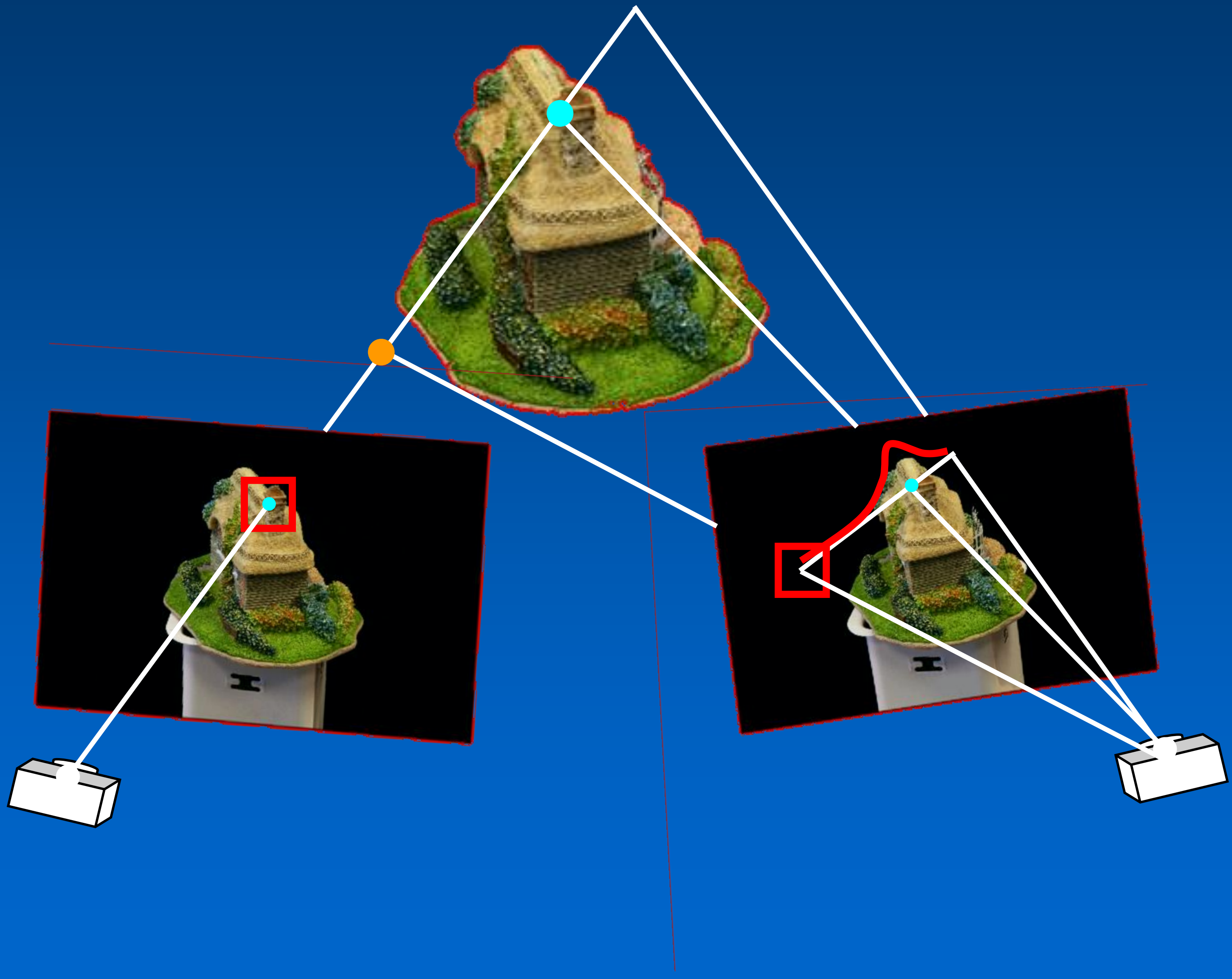
- Volumetric Graph-cuts
  - Uses an occlusion-robust photo-consistency
  - Casts the problem as discrete Markov Random Field (MRF) optimisation, obtaining global solution

The occlusion problem

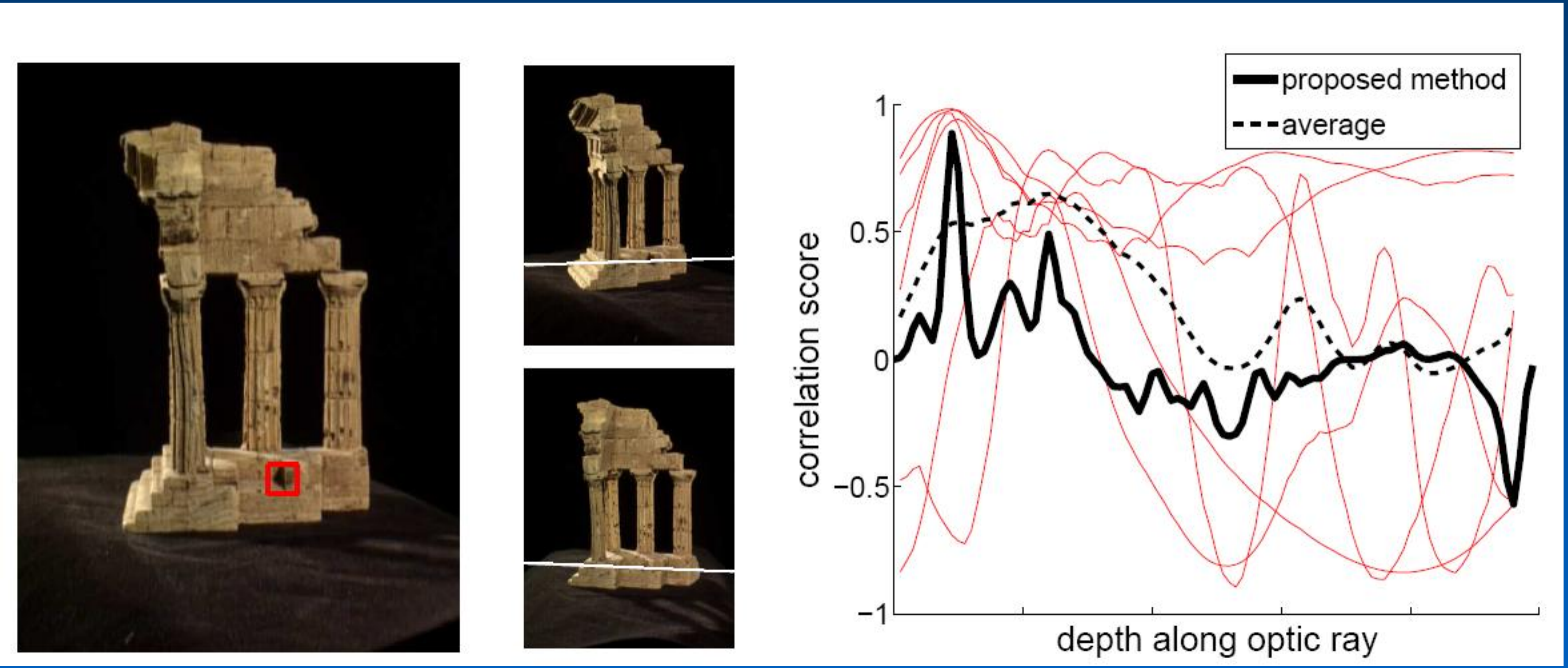


- To get 3D object must compute photo-consistency
- To compute photo-consistency, must know occluded cameras
- To know occluded cameras must know 3D object!

Find optimal depth with NCC

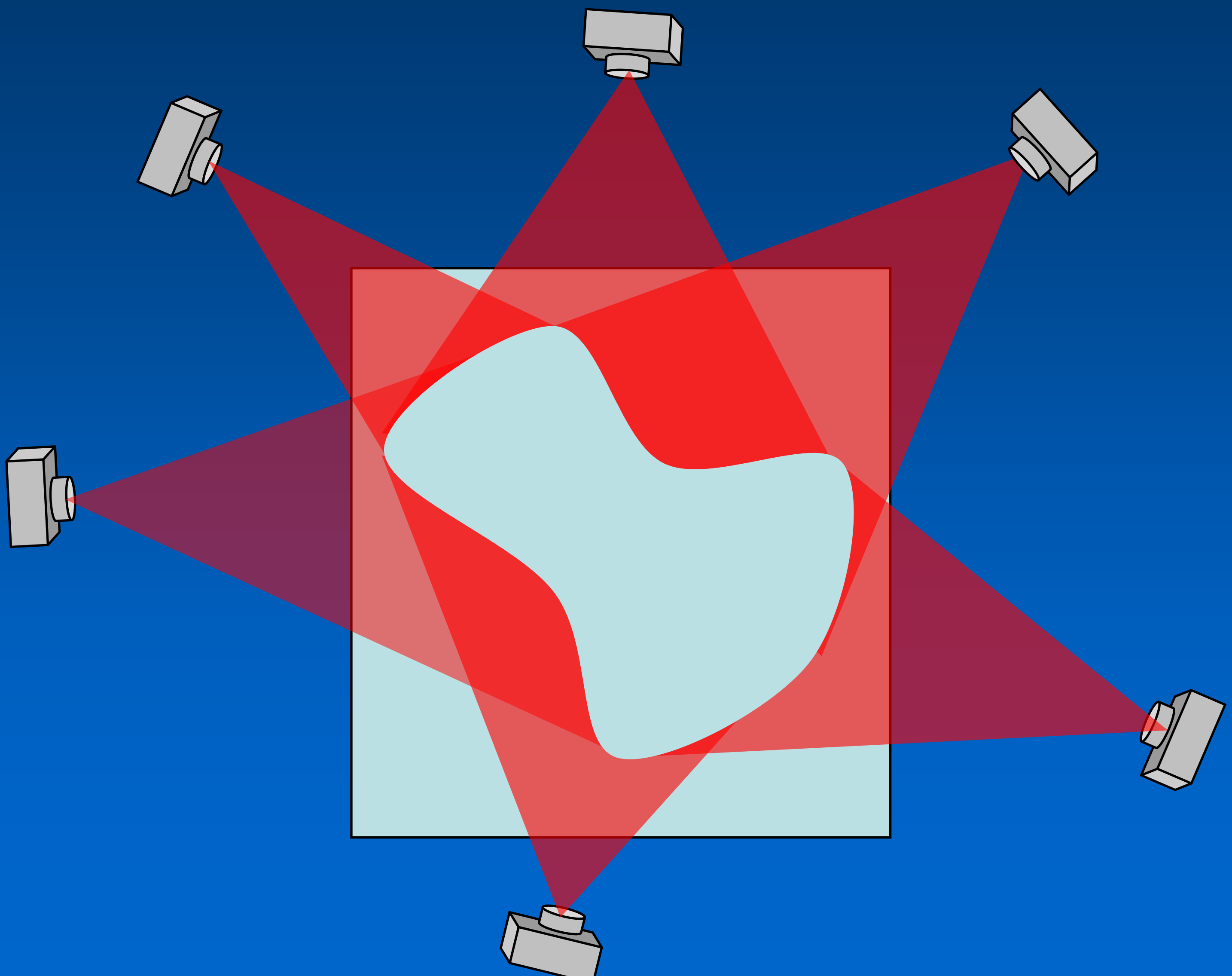


Find optimal depth with NCC

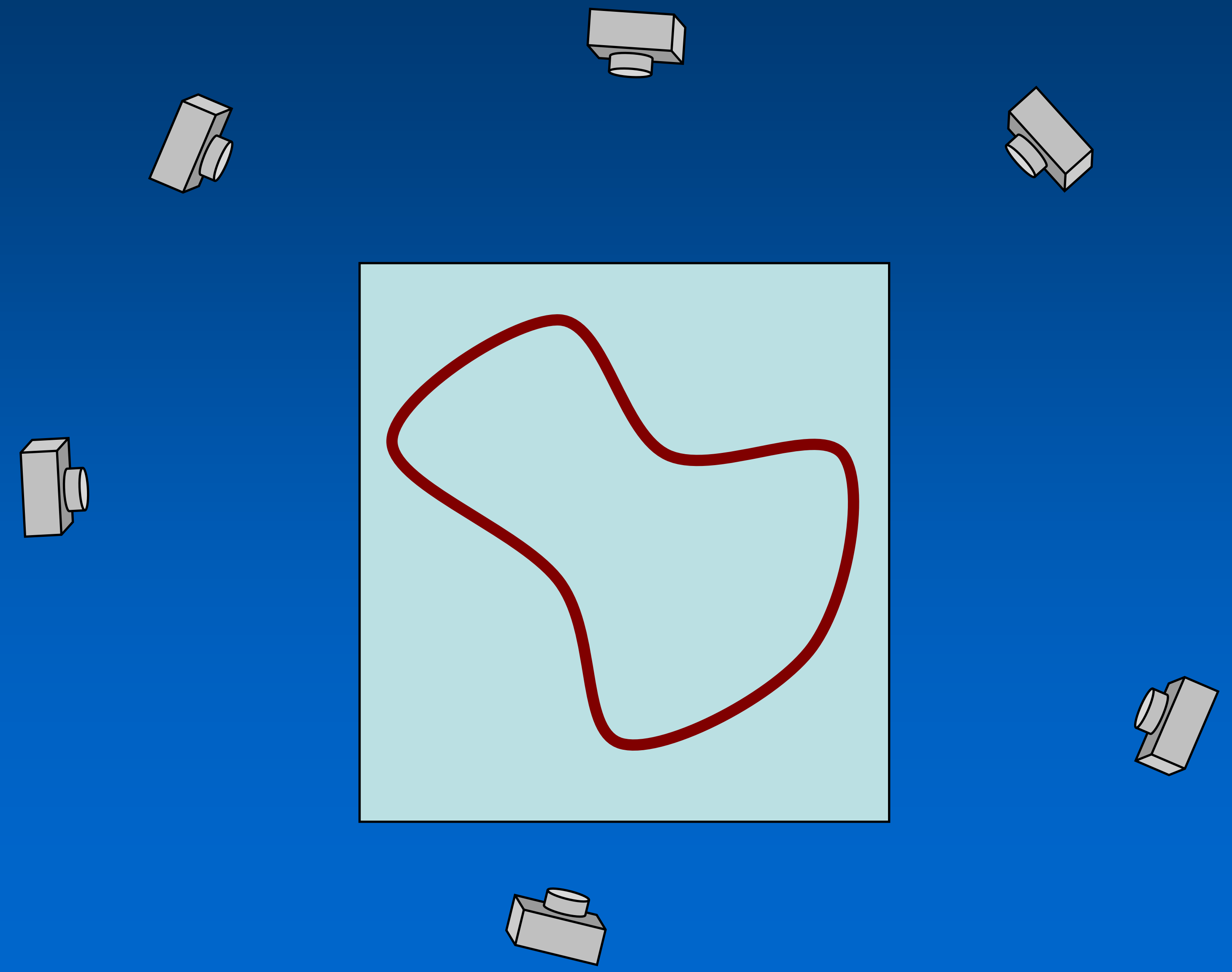


- Count number of local maxima of red curves for each depth (in small interval)

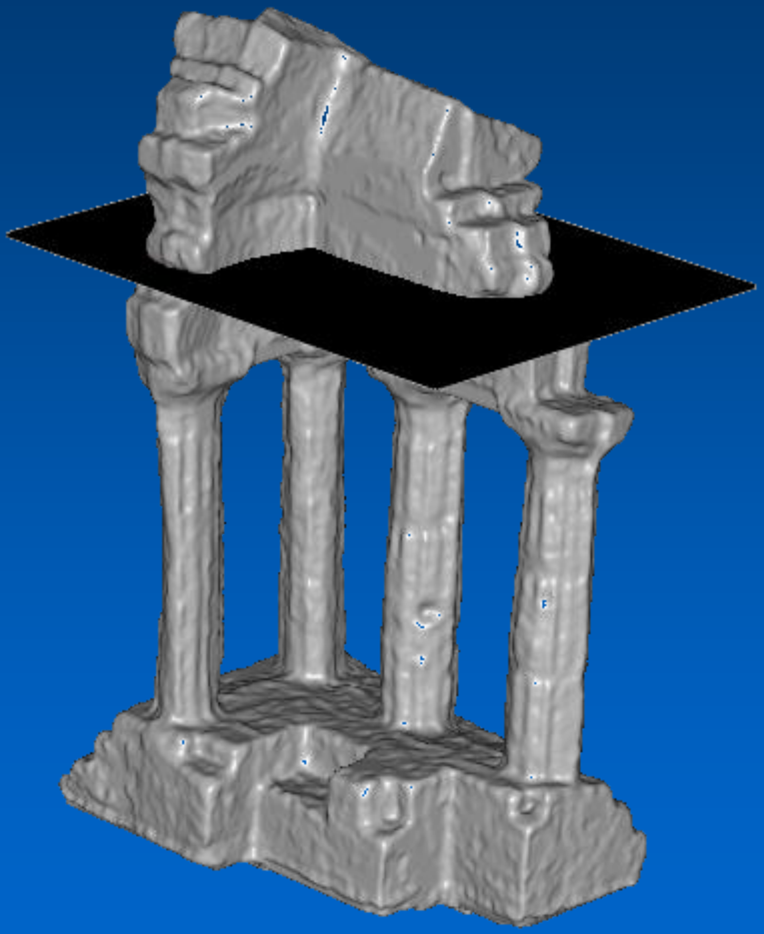
Combining depth-maps

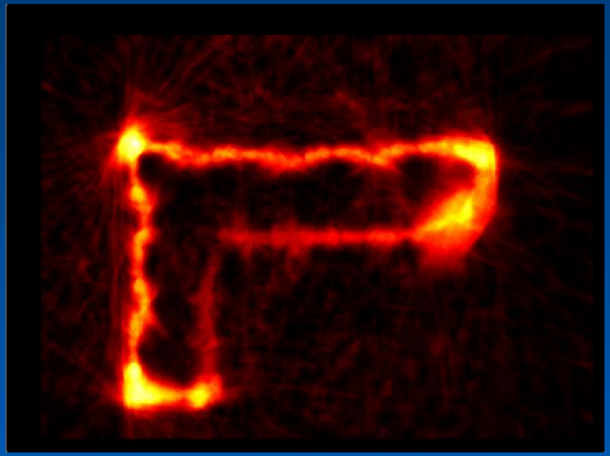


Aggregating information

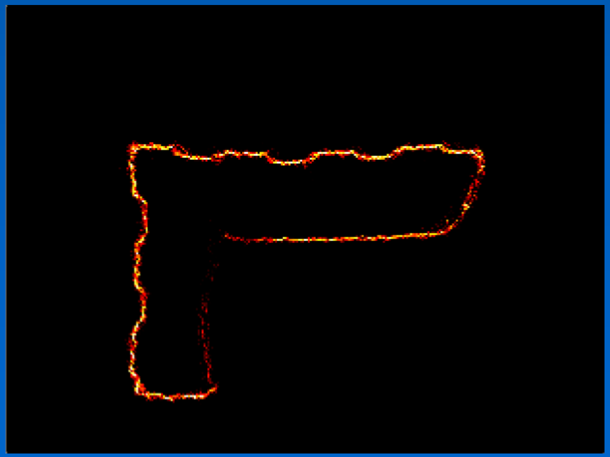


Occlusion-robust NCC



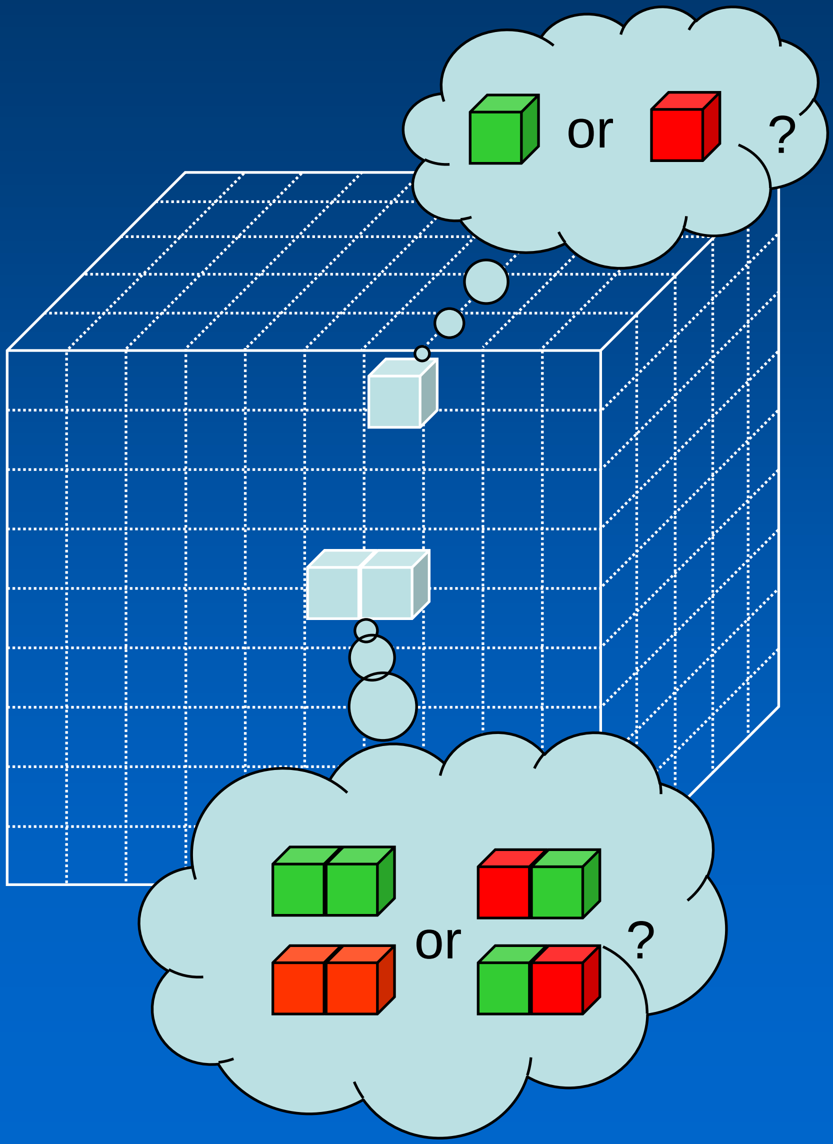


Averaged NCC



Robust NCC

3D MRF models



Labelling cost:

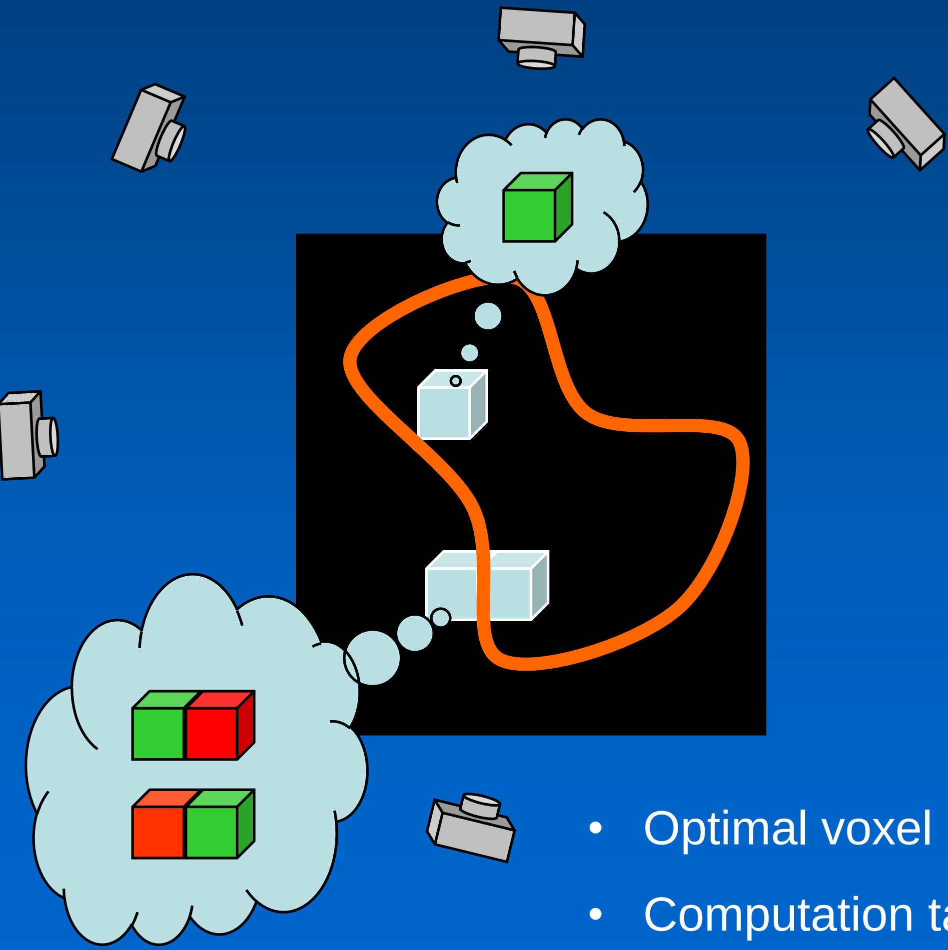
- Every voxel has a certain preference for being **foreground** or **background**

Compatibility cost:

- Every pair of neighbour voxels has a certain preference for being given the *same* or *oposite* labels

Optimal labelling can under some simple condition be obtained in polynomial time using “graph-cuts” (*Boykov & Kolmogorov, ICCV 2003*)

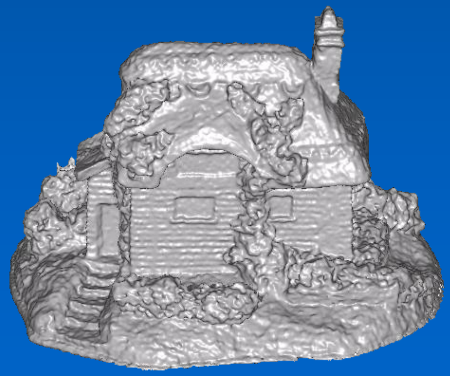
3D MRF for 3D modelling



- Labelling cost:
  - Constant bias towards being **foreground**
- Compatibility cost:
  - Pair of neighbour voxels prefers having oposite labels if photo-consistent region is between them

- Optimal voxel labelling can be computed using graph-cuts
- Computation takes approx. *7mins* for *512x512x512* grid on Pentium IV 2.6Ghz

3D Models



3D Models

3D models

Middlebury evaluation (temple)

	Accuracy / Completeness		
	Full (312 images)	Ring (47 images)	SparseRing (16 images)
Hernandez [10]	0.36mm / 99.7%	0.52mm / 99.5%	0.75mm / 95.3%
Goesele [9]	0.42mm / 98.0%	0.61mm / 86.2%	0.87mm / 56.6%
Hornung [12]	0.58mm / 98.7%	–	–
Pons [20]	–	0.60mm / 99.5%	0.90mm / 95.4%
Furukawa [8]	0.65mm / 98.7%	0.58mm / 98.5%	0.82mm / 94.3%
Vogiatzis [29]	1.07mm / 90.7%	0.76mm / 96.2%	2.77mm / 79.4%
Present method	**0.50mm / 98.4%**	**0.64mm / 99.2%**	**0.69mm / 96.9%**

Advantages

- Accurate
  – sub-millimetre accuracy on sequence with ground truth
- Simple
  – Can work with about 15-30 images
- Fast
  – Approximately 45’ of computation for these models
  – We believe we can bring this down to few minutes

Publications

- **Silhouette and stereo fusion for 3D object modeling.** Hernández *et al, CVIU 96(3):367–392, 2004*.
  - Described occlusion-robust photo-consistency

- **Multi-view Stereo via Volumetric Graph-cuts**, Vogiatzis *et al, CVPR, 391–398, 2005*.
  - Formulated the multi-view stereo problem as MRF inference

- New and improved version appears in PAMI (December 2007)