UNIVERSITY OF CAMBRIDGE

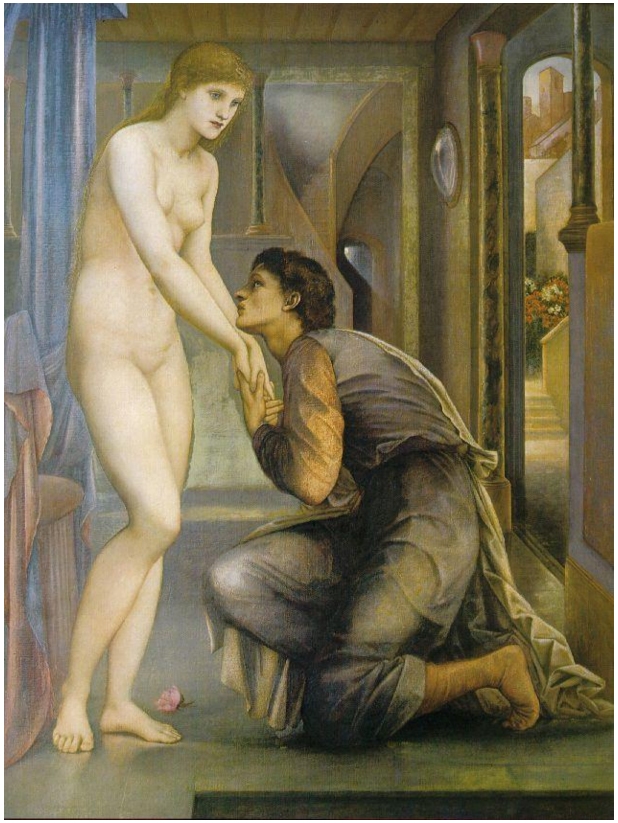
Digital Pygmalion

Accurate 3D reconstruction from uncalibrated images

Roberto Cipolla

Carlos Hernandez and George Vogiatzis
Department of Engineering

Background



Background

Background

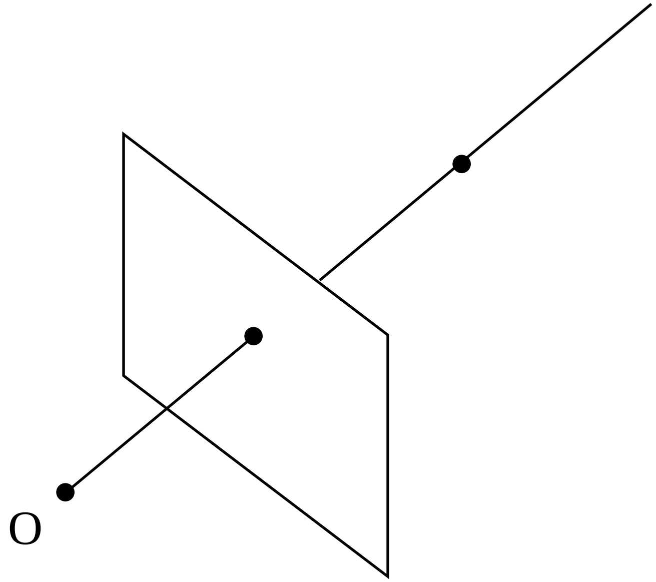
Overview

UNIVERSITY OF CAMBRIDGE

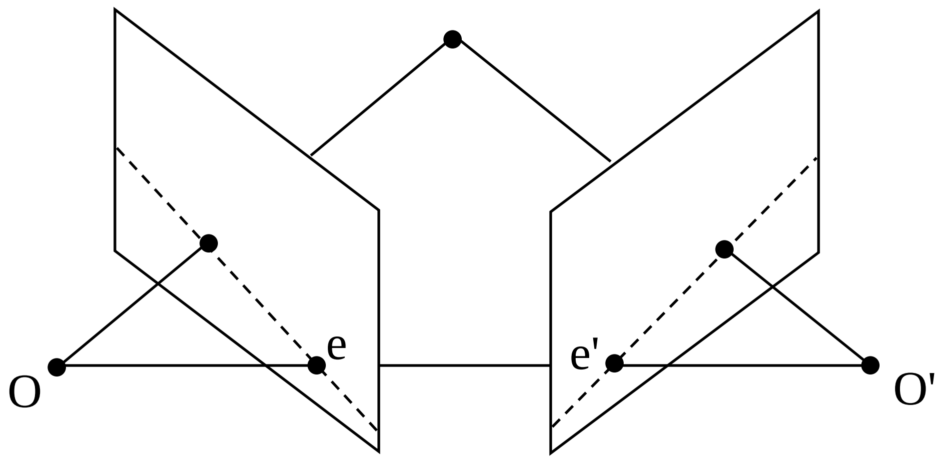
- Accurate 3D reconstruction from uncalibrated cameras (motion and lighting)
- Multi-view stereo - 3D shape from uncalibrated images (review)
- Multi-view photometric stereo with uncalibrated lights (CVPR2006)
- Object detection and tracking (summary)

1. 3D shape recovery from uncalibrated images

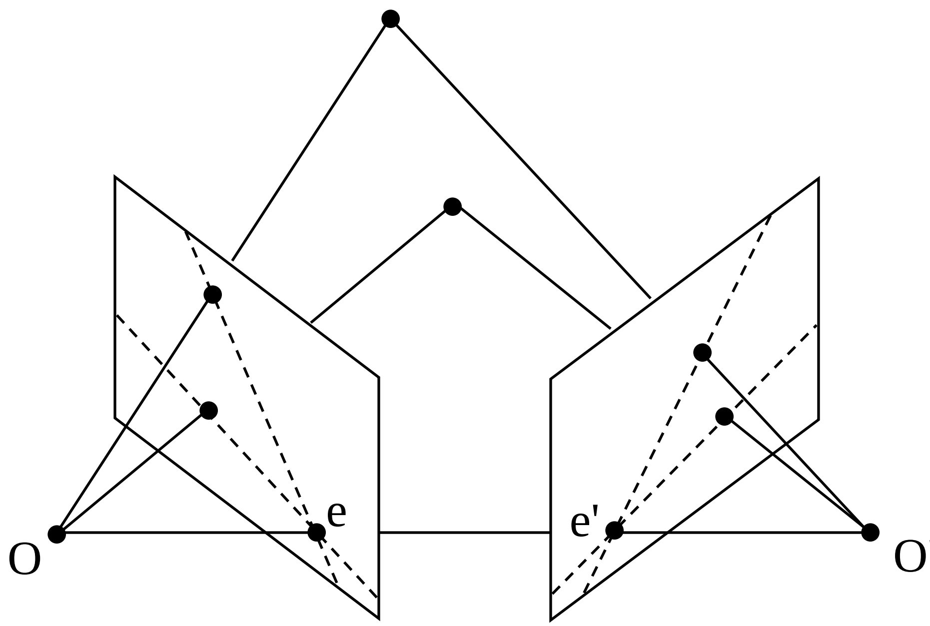
Ambiguity in a single view



Stereo vision



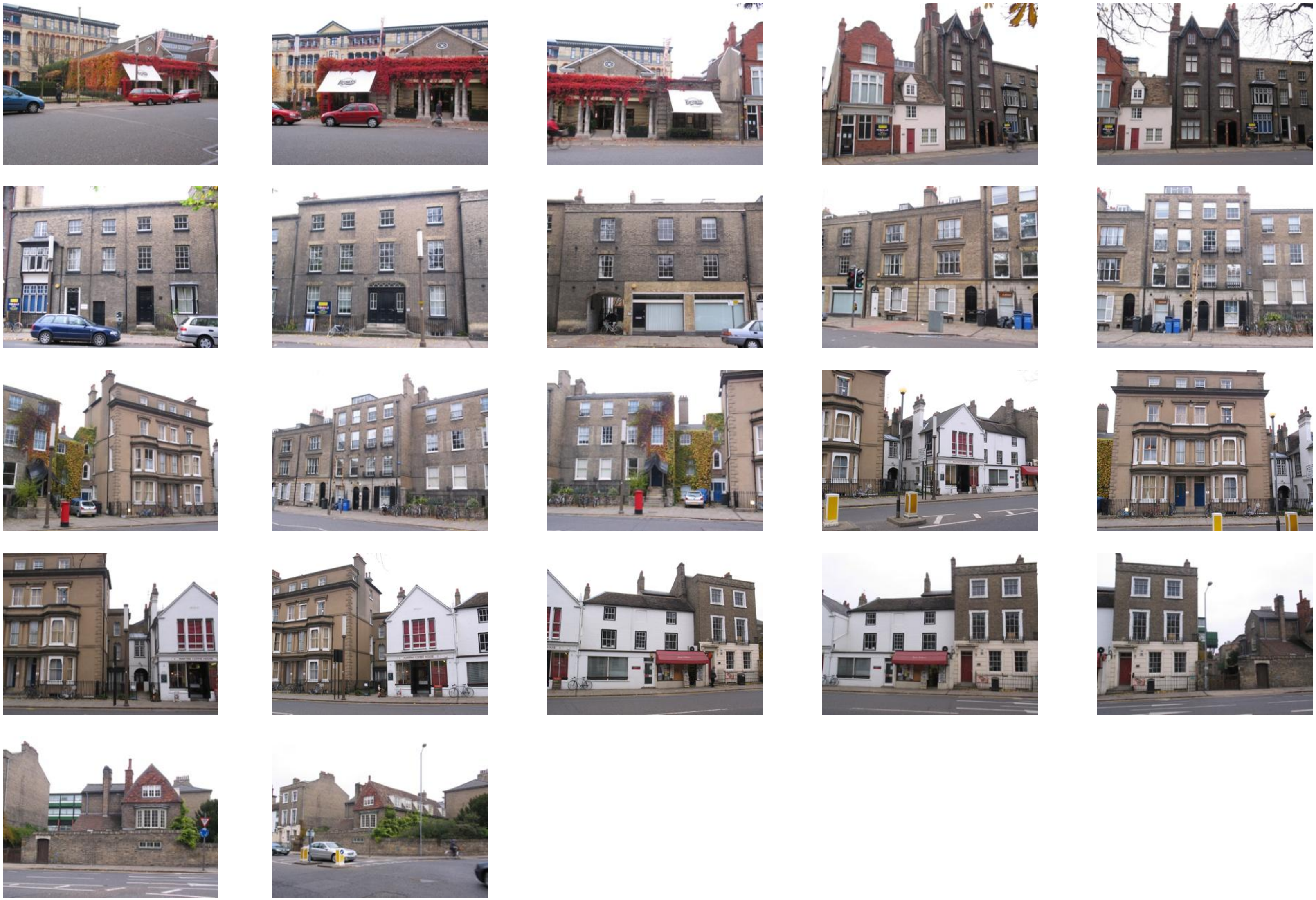
Stereo vision



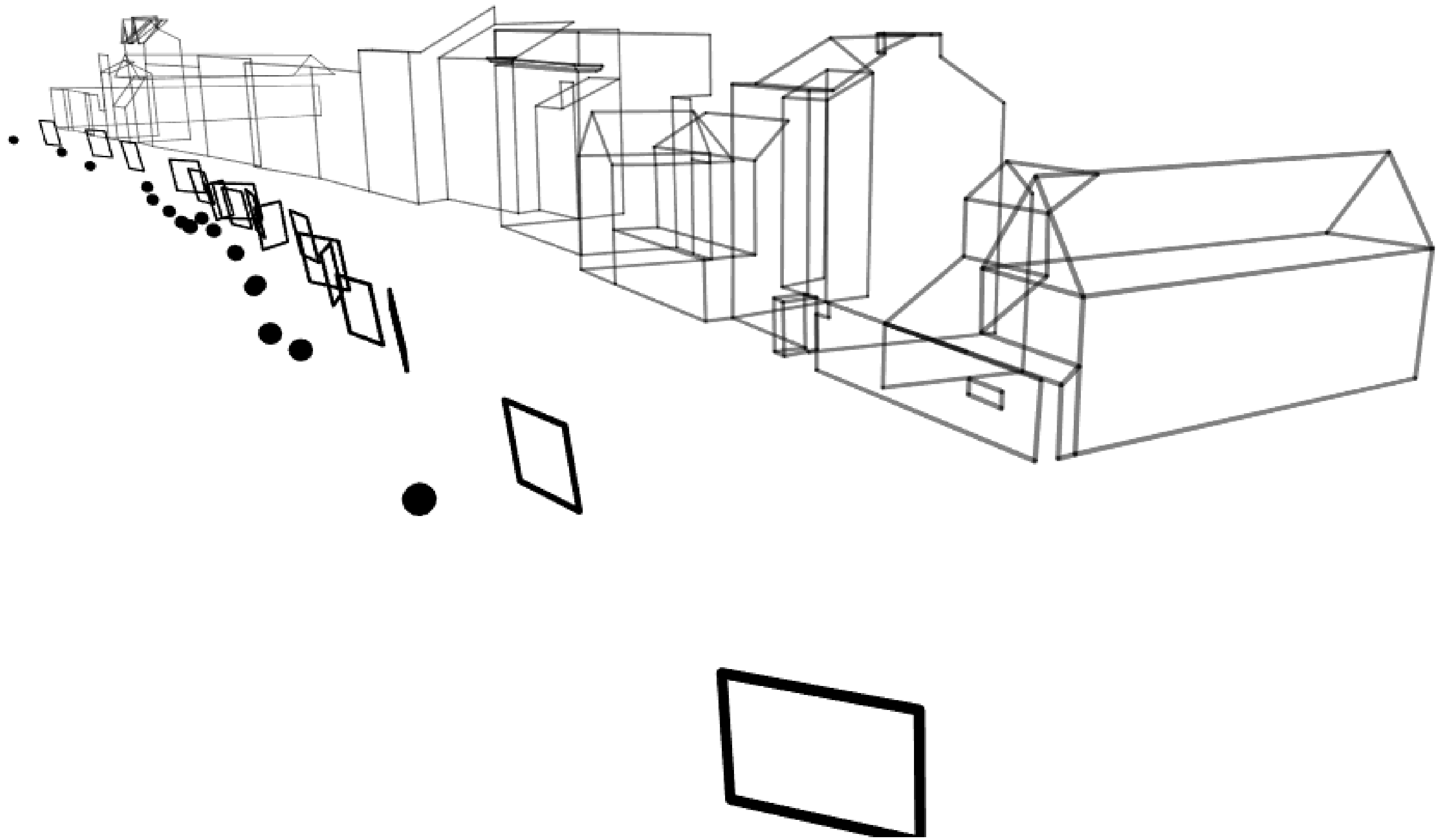
3D reconstruction of streets

UNIVERSITY OF CAMBRIDGE

Trumpington Street Data

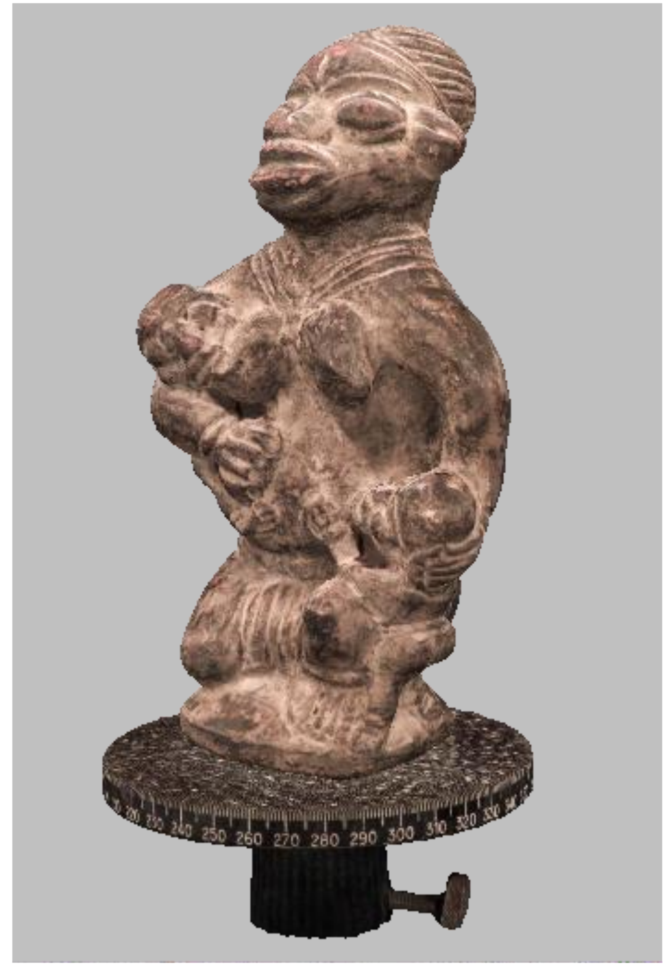


3D reconstruction



Reconstruction texture mapped

Digital Pygmalion project



3D Shape from Images

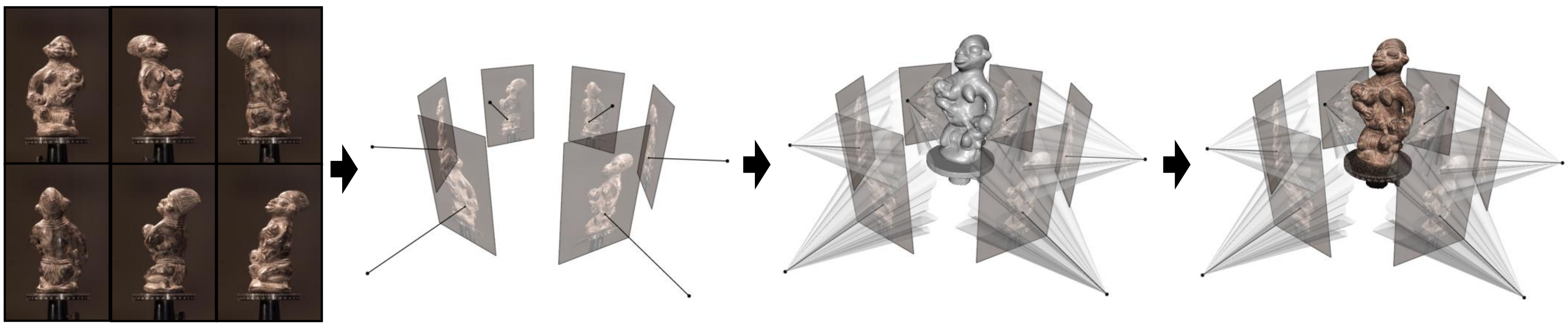
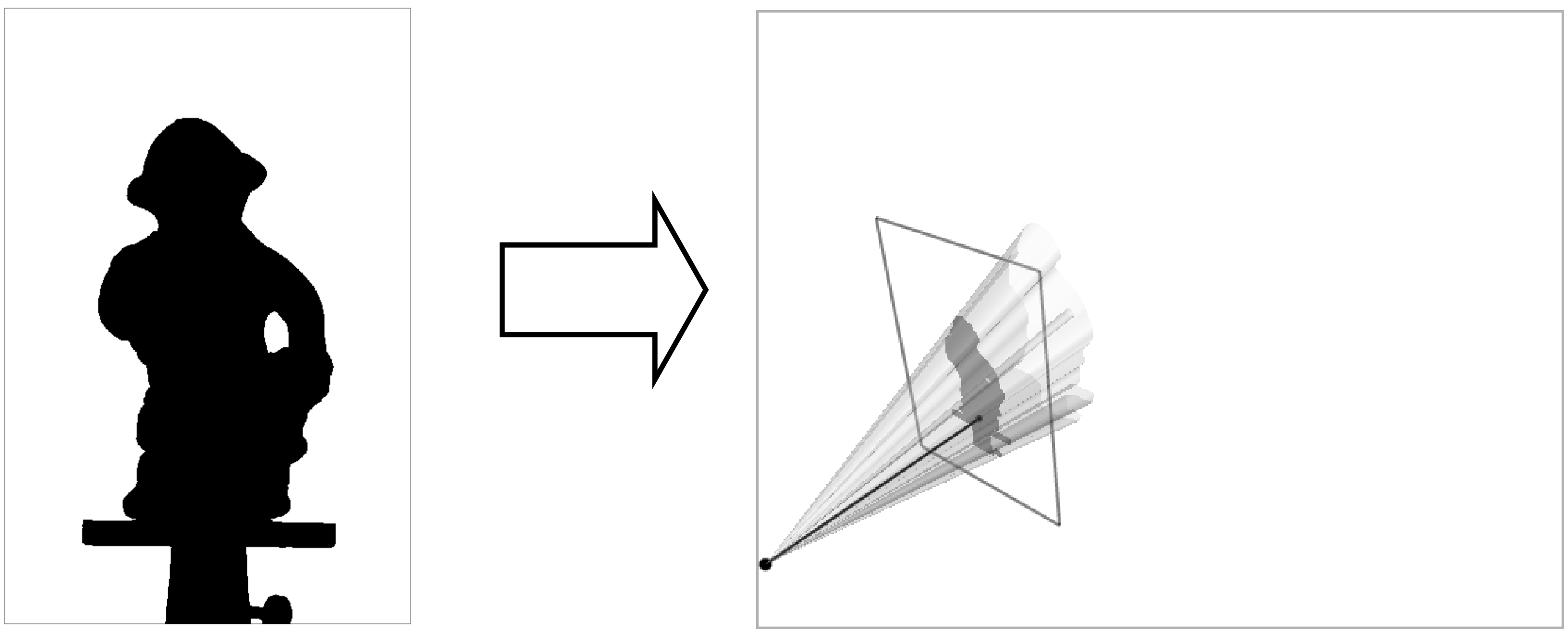


Image acquisition | Camera calibration | Geometry reconstruction | Texture map creation

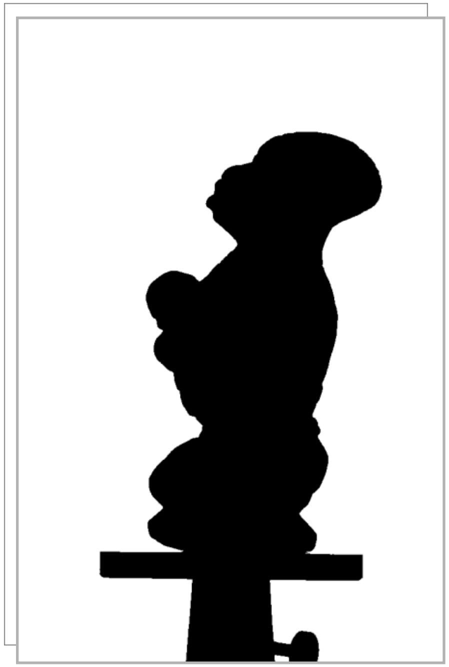
Input Images

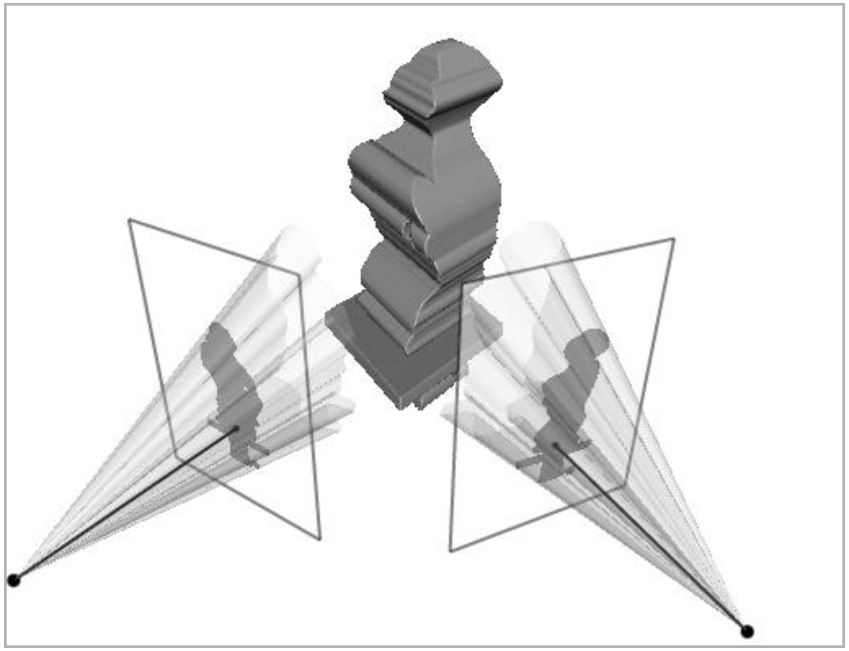
Input Images

Carving the visual hull

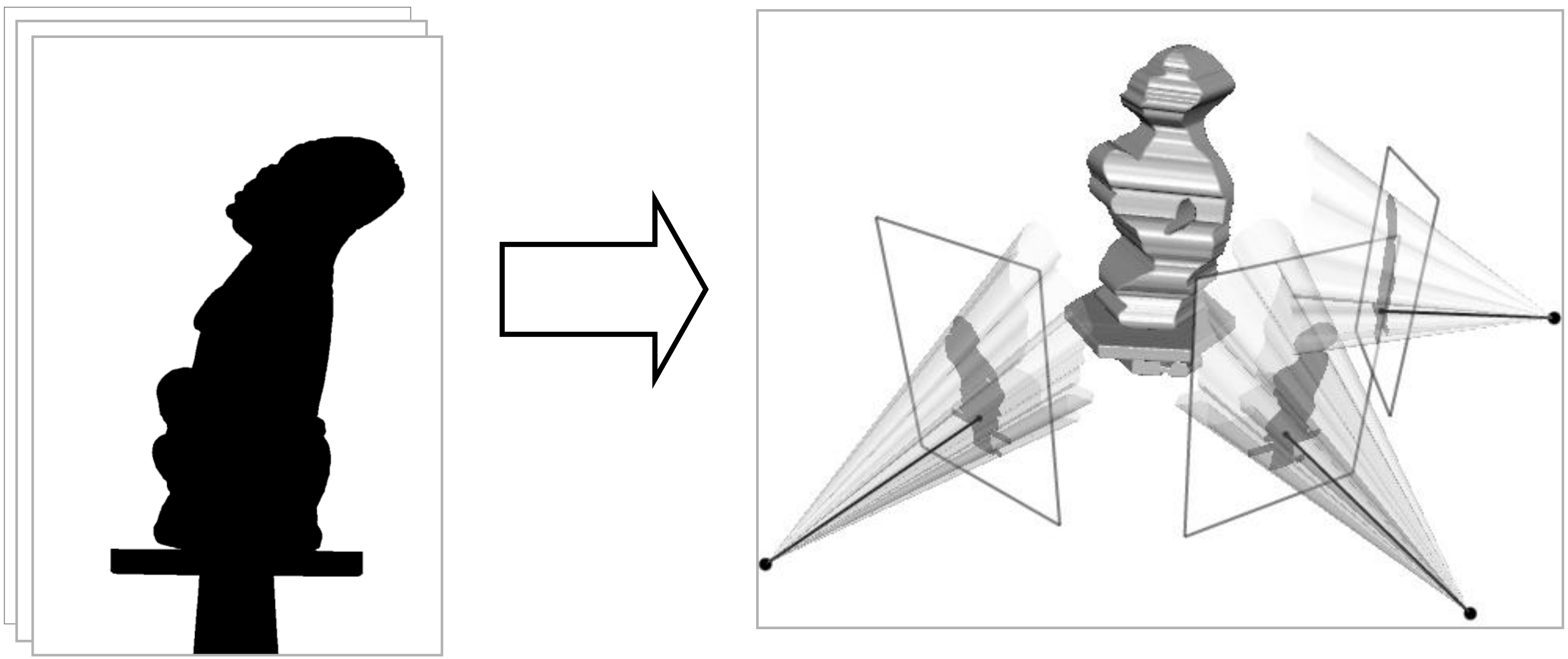


Carving the visual hull

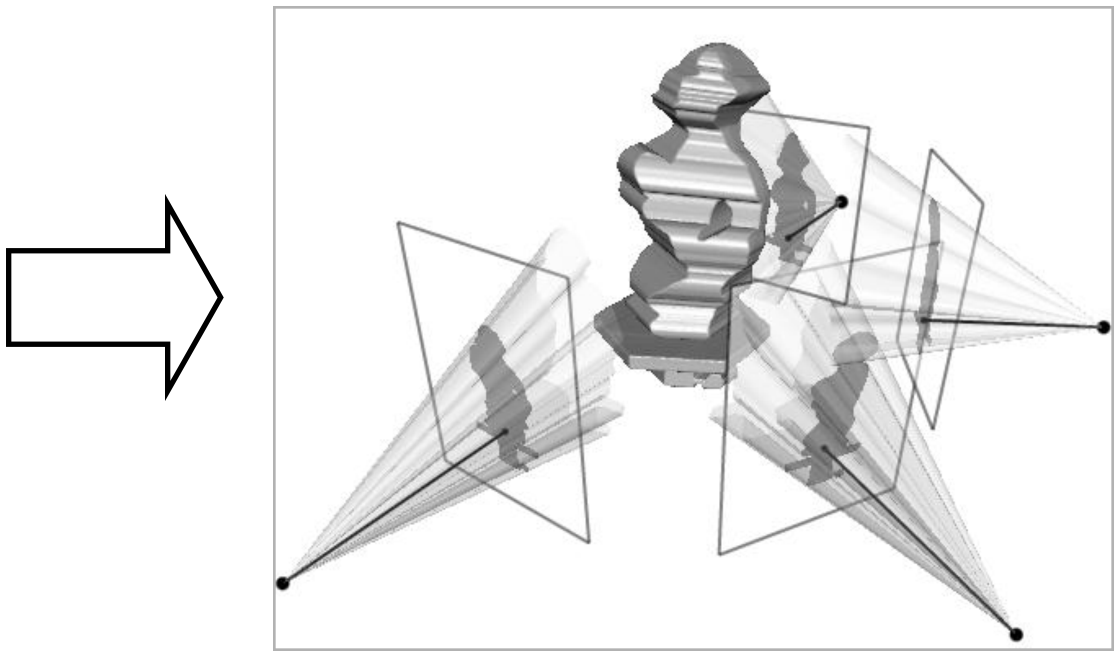




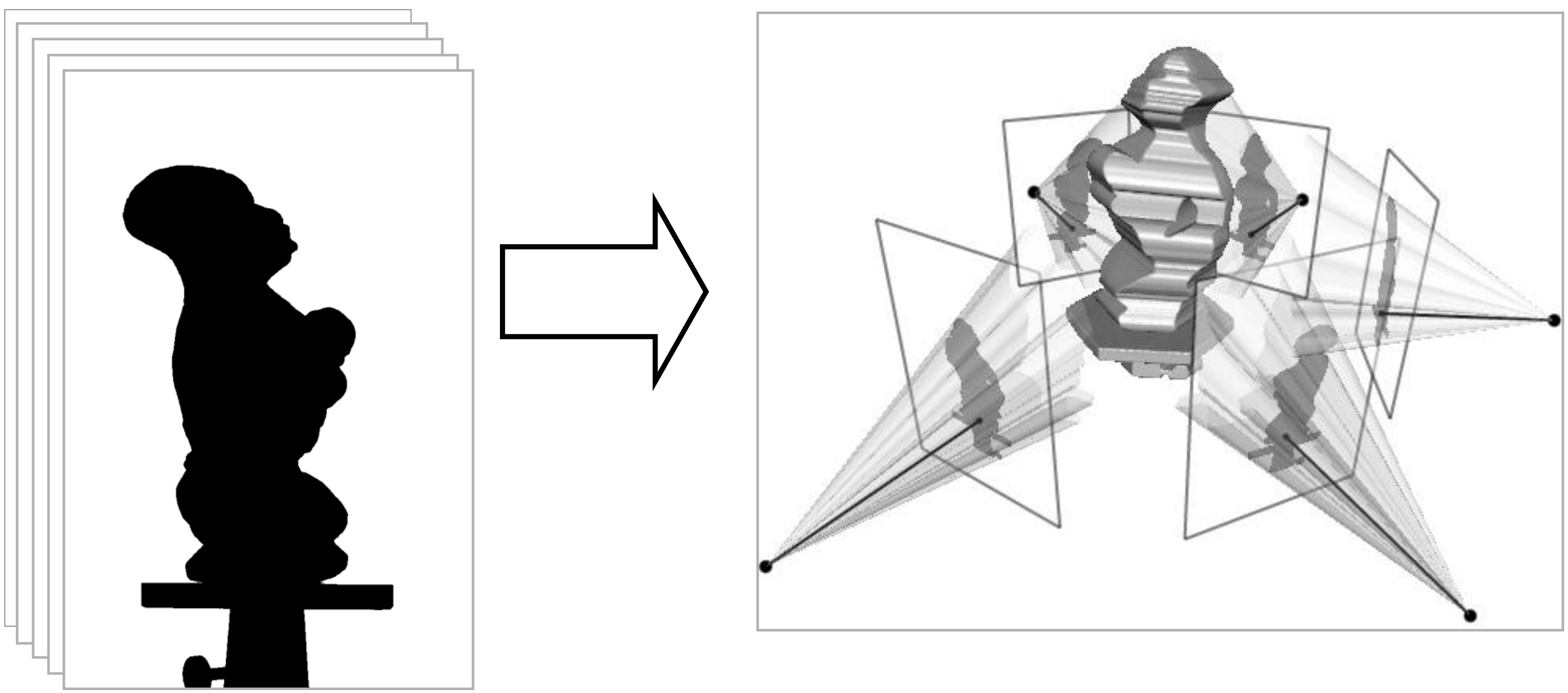
Carving the visual hull



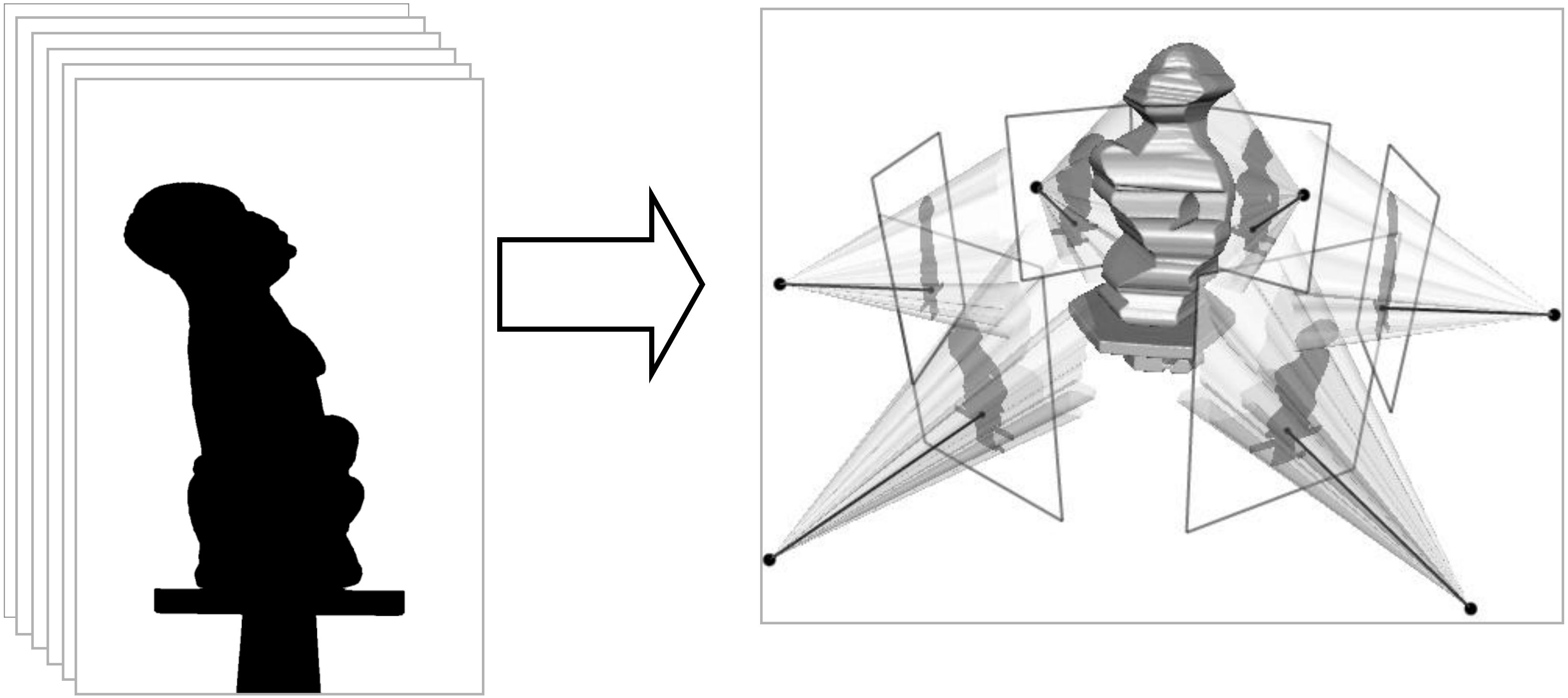
Carving the visual hull



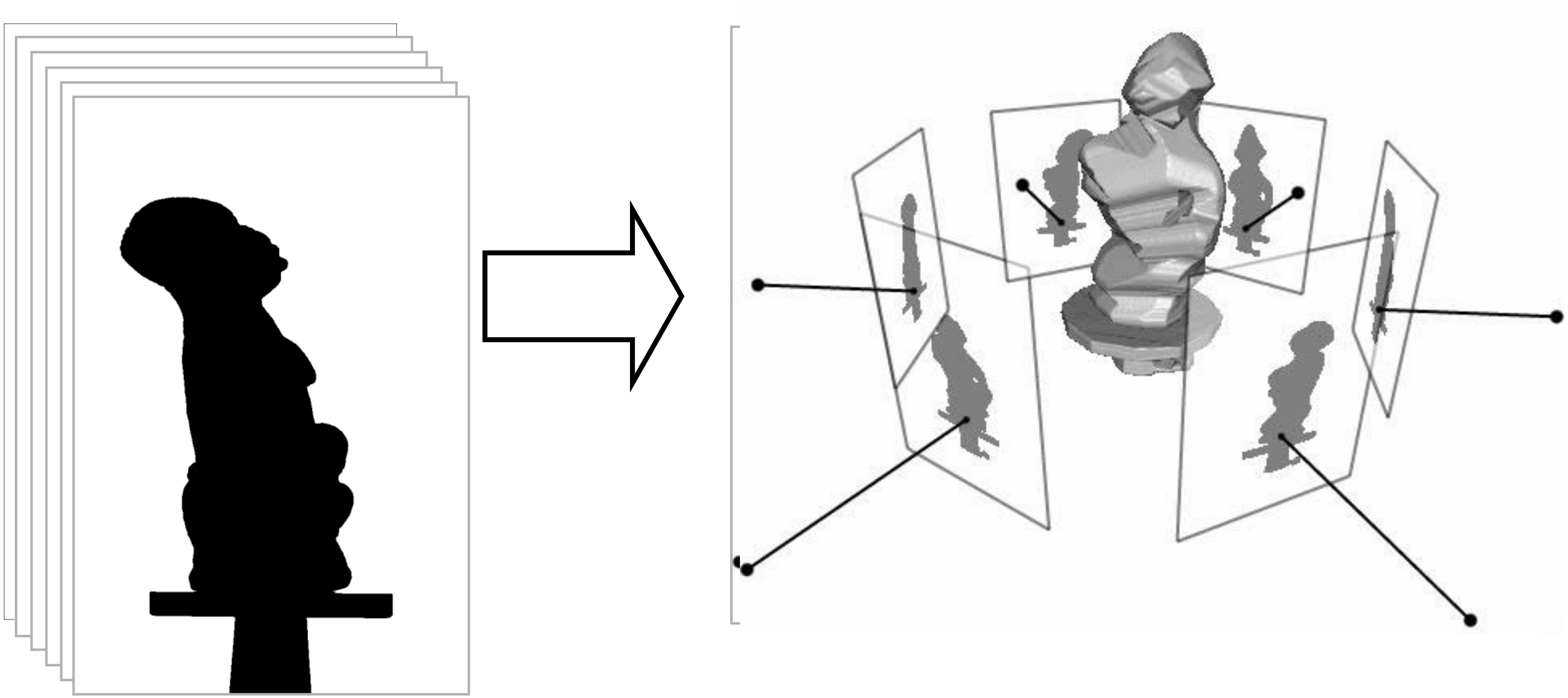
Carving the visual hull



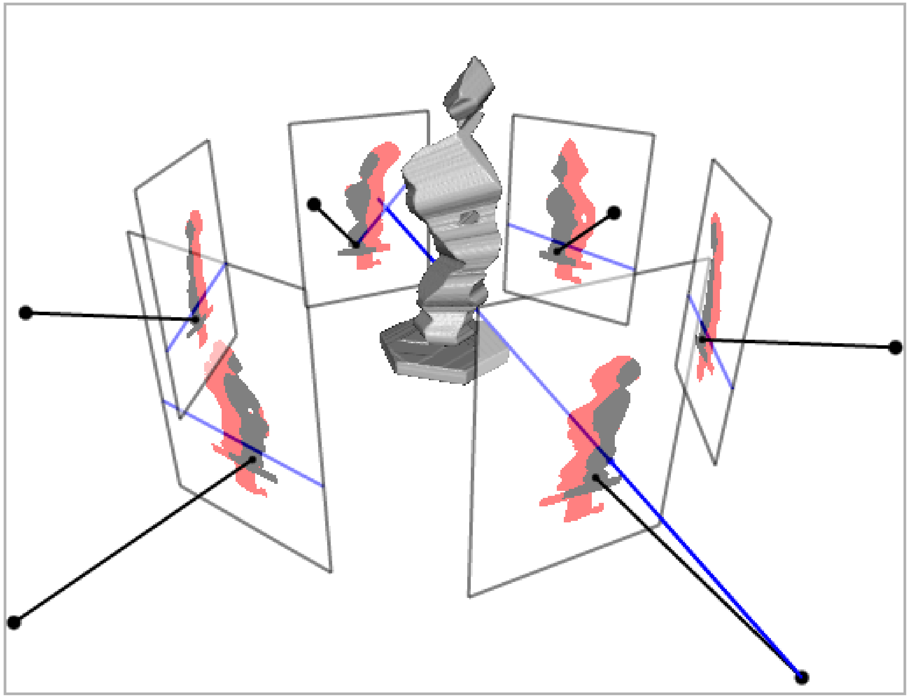
Carving the visual hull



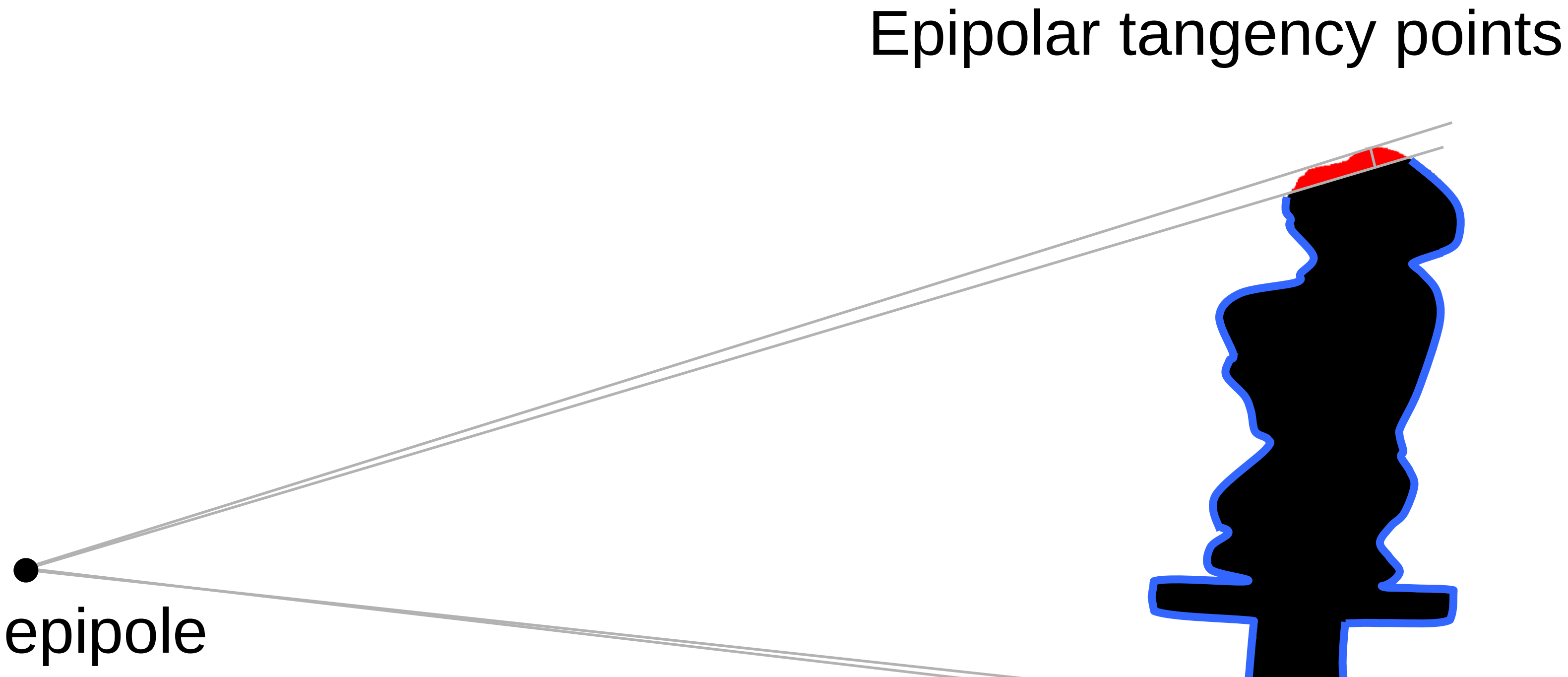
Camera calibration



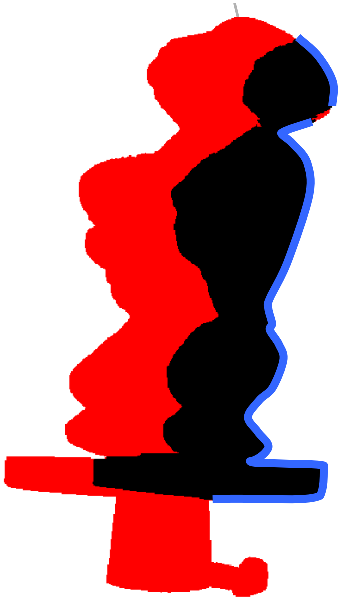
Camera calibration



Epipolar tangency points

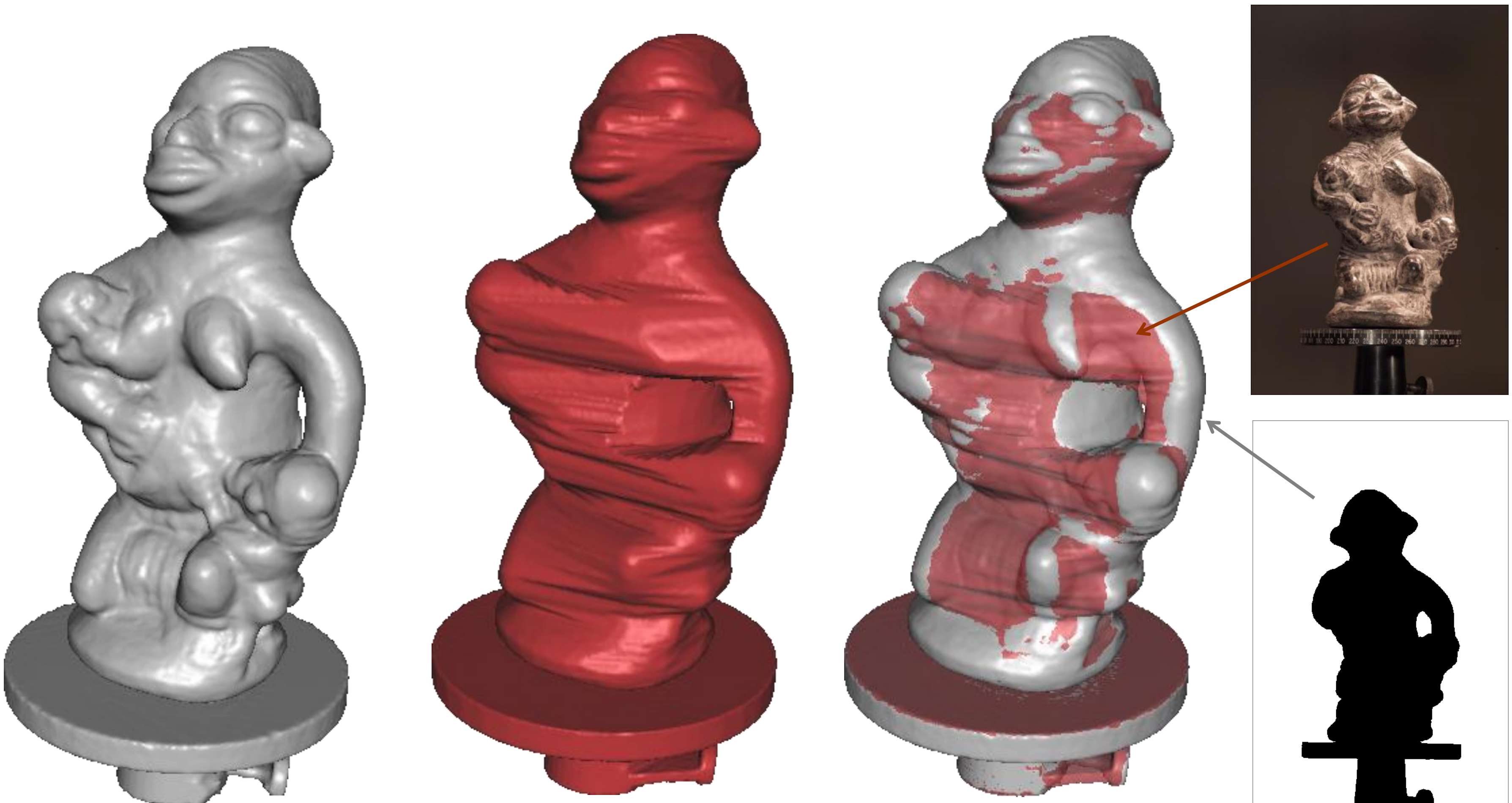


Silhouette coherency



UNIVERSITY OF CAMBRIDGE

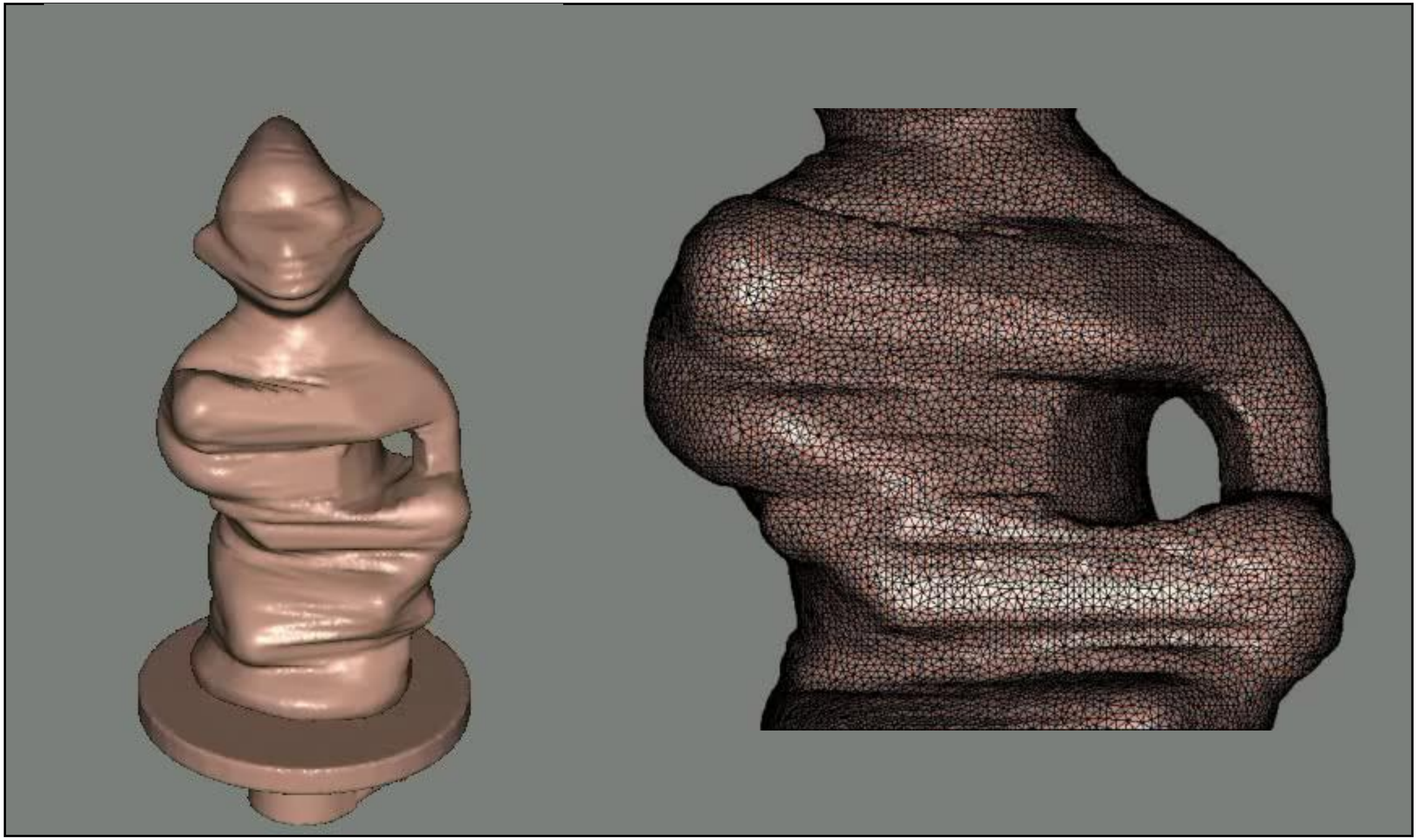
Recovery of concavity



Real surface

Visual hull surface

Refining the mesh



Texture mapping

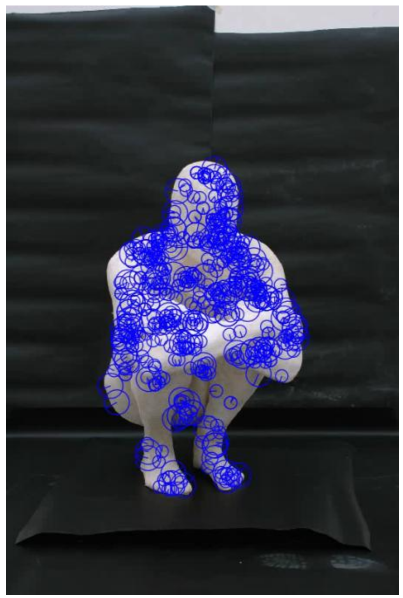
83241 vertices, 166482 triangles

Input Images

Input images

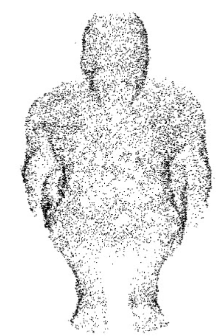
Recovery of camera motion

Input images



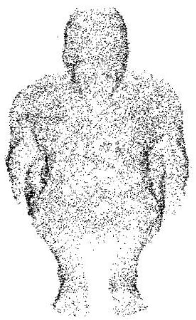
Feature extraction

Feature matching

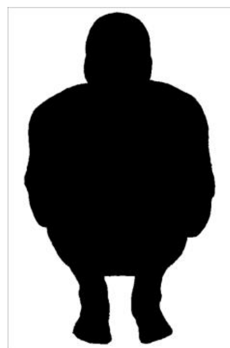


Bundle adjustment

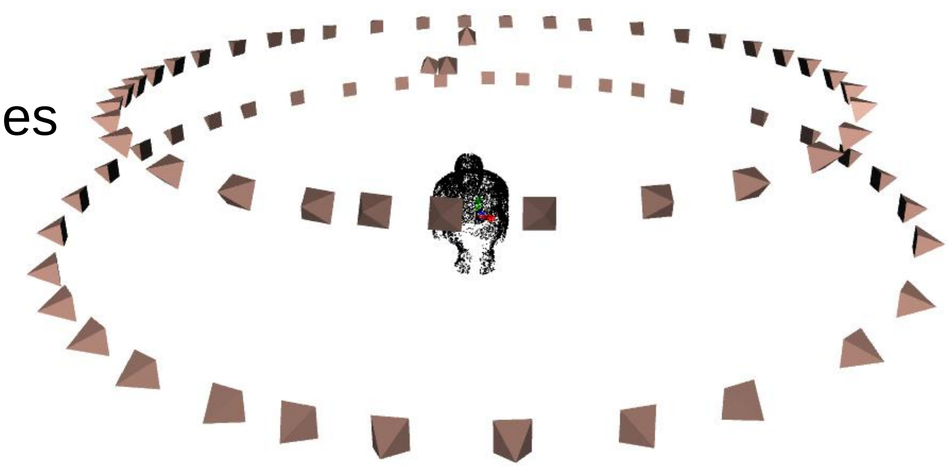
Refine with profiles



Initialize with features



Refine with profiles

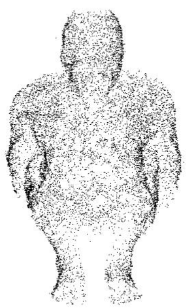


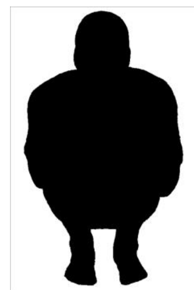
Final camera motion

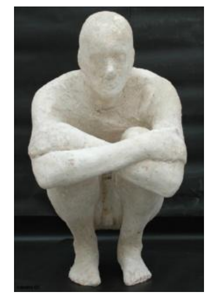
Recovery of surface geometry

UNIVERSITY OF CAMBRIDGE

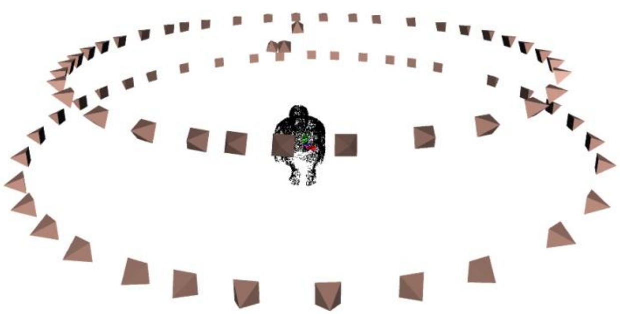
Input data



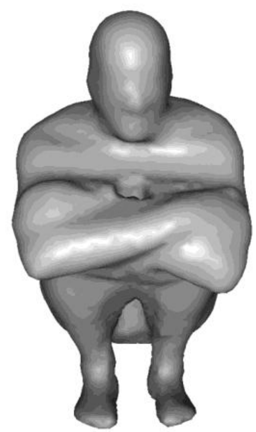




Process



camera motion



visual hull

rough geometry

detailed geometry

texture map

2 Reconstruction in the round with photometric normals

Background

Uniform albedo object

- Challenging objects
- Lack of features makes correspondences hard
- Silhouette and shading are only available cues

Photometric stereo

- Single Viewpoint
- Move light-source for each image
- Same pixel always corresponds to same surface point
- With known light directions can estimate **n**

$$i = \mathbf{l}^{\mathbf{T}}\mathbf{n}$$

- Integrate normals to get depth map

Photometric stereo

- To get more than depth-maps, we need multiple viewpoints…
- … and in that case pixels are no longer automatically in correspondence
- However, if some correspondence is given, photometric stereo can proceed as usual

Photometric stereo

UNIVERSITY OF CAMBRIDGE

- Our strategy:

1. Estimate light direction and intensity
2. Evolve a surface using photometric stereo with approximate correspondences from the current surface (starting from visual hull)

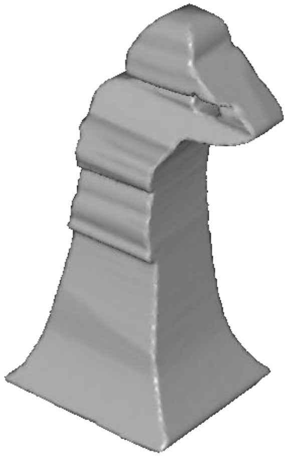
Light estimation

- Three surface points with known surface normals and their image intensities are enough to estimate a directional light source

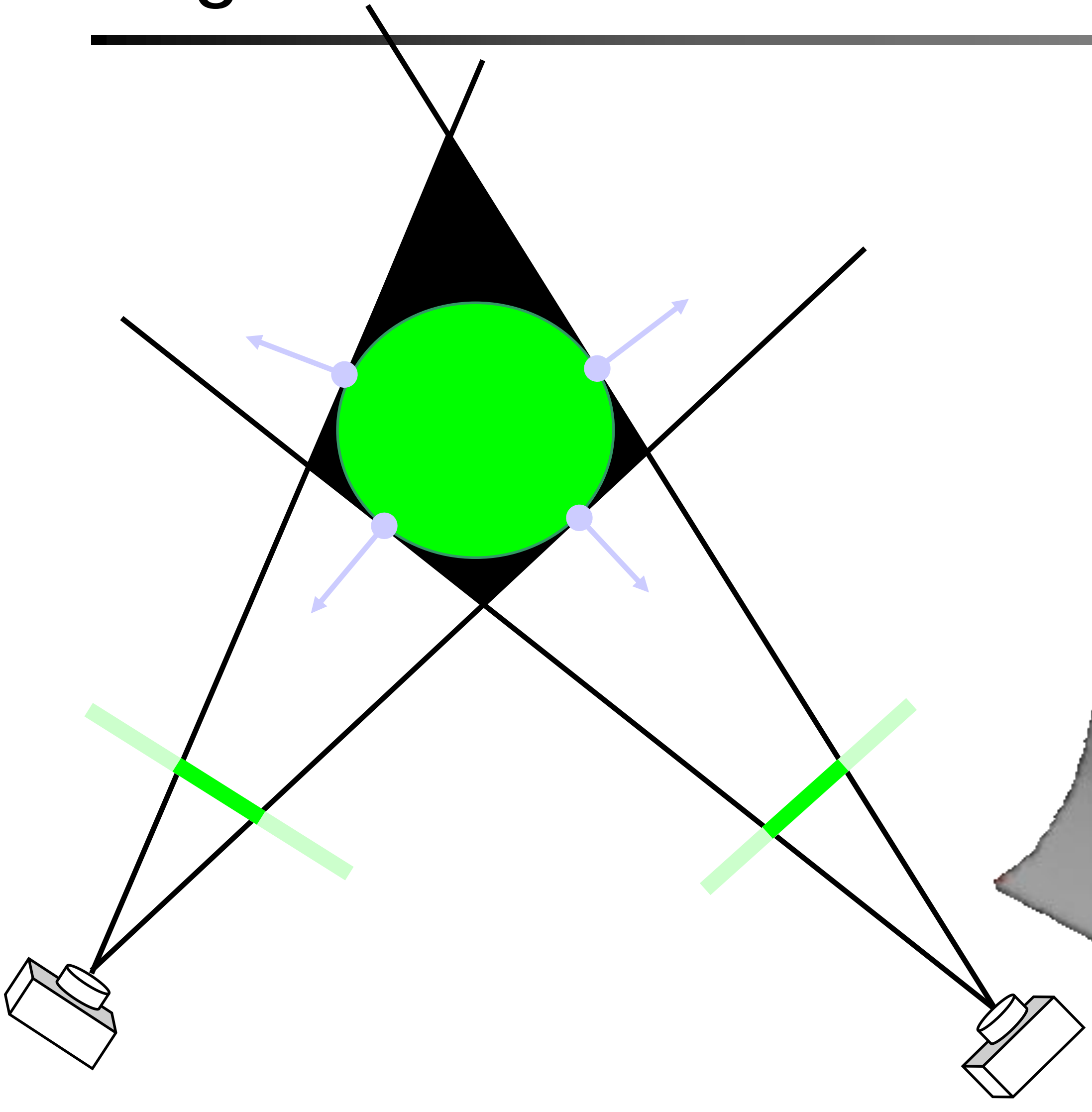
$$\mathbf{l} = [\mathbf{n_a}\,\mathbf{n_b}\,\mathbf{n_c}]^{-1} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}$$

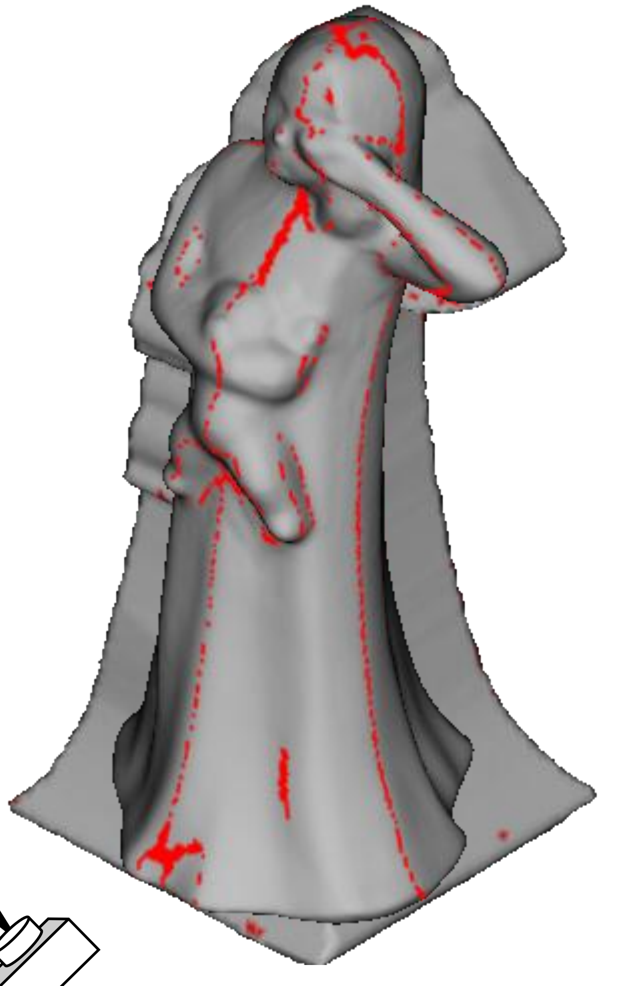
- But where do you get these three points ?

Light estimation



Light estimation



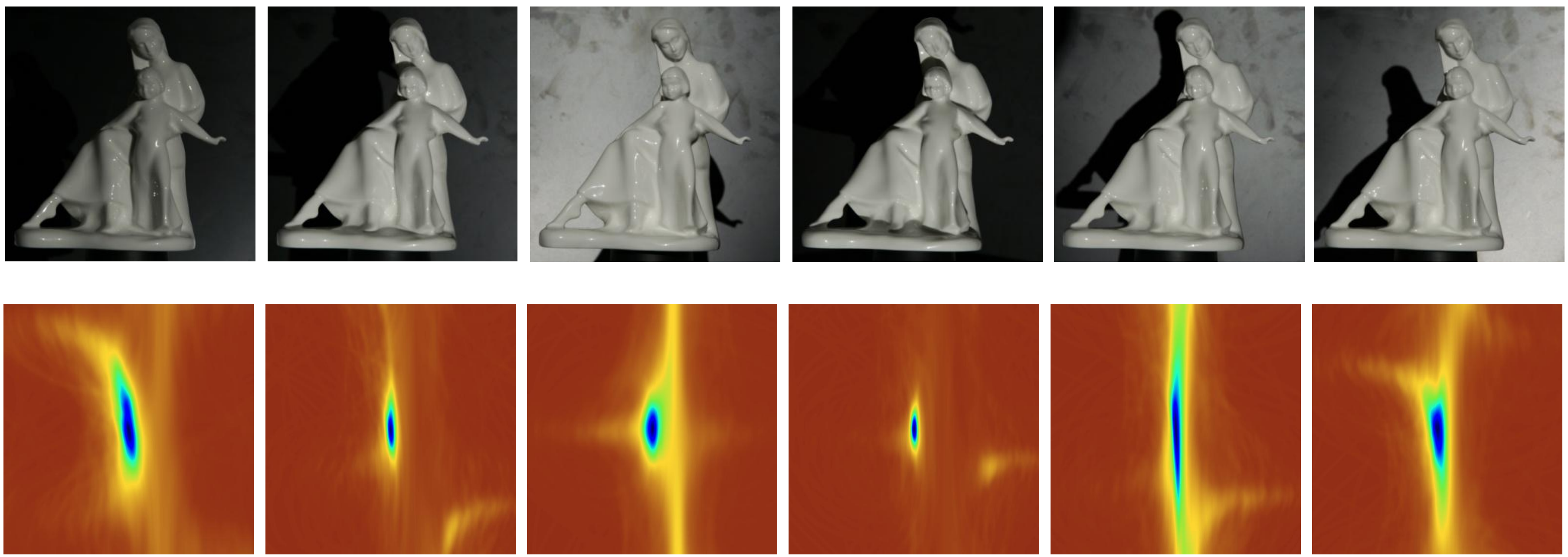


Light estimation

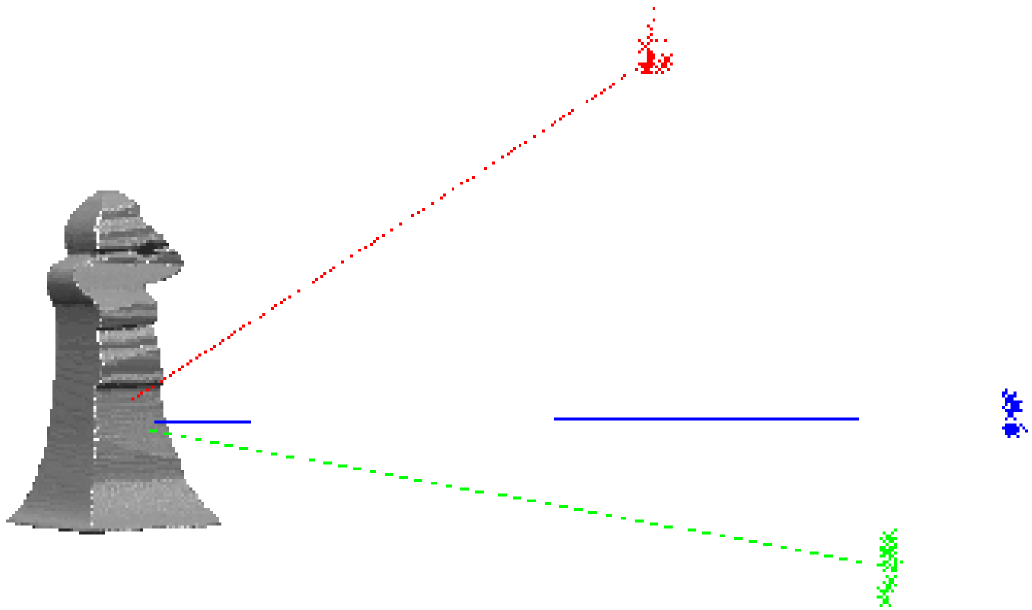
UNIVERSITY OF CAMBRIDGE

- Recover contour generators by random sampling

Accuracy of light estimation



Light estimation



Multi-view photometric stereo

UNIVERSITY OF CAMBRIDGE

- Mesh with vertices $x_1,\ldots,x_M$
- And faces $f=1,\ldots,F$
- Define photometric normals $v_1,\ldots,v_F$
- Minimise sum of two energies
  - E_m with respect to $x_1,\ldots,x_M$

$$E_m\left(\mathbf{x_1},\ldots,\mathbf{x_M};\mathbf{v_1},\ldots,\mathbf{v_F}\right)=\sum_{f=1}^{F}\left\|\mathbf{n_f}-\mathbf{v_f}\right\|^2 A_f$$

$$E_v\left(\mathbf{v_1},\ldots,\mathbf{v_F};\mathbf{x_1},\ldots,\mathbf{x_M}\right)=\sum_{f=1}^{F}\sum_{k\in\mathcal{V}_f}\left(\mathbf{l_k}^T\mathbf{v_f}-i_{f,k}\right)^2$$

Multi-view photometric stereo

UNIVERSITY OF CAMBRIDGE

Reconstruction in the Round
Using Photometric Normals

Paper ID #548

Mesh Evolution

Full algorithm for uniform albedo

UNIVERSITY OF CAMBRIDGE

Capture images of object.
Extract silhouettes.
Recover camera motion and compute visual hull.
Estimate light directions and intensities in every image

Initialise a mesh with vertices $\mathbf{x}_1 \dots \mathbf{x}_\mathrm{M}$ and faces $f = 1 \dots F$ to the object's visual hull.

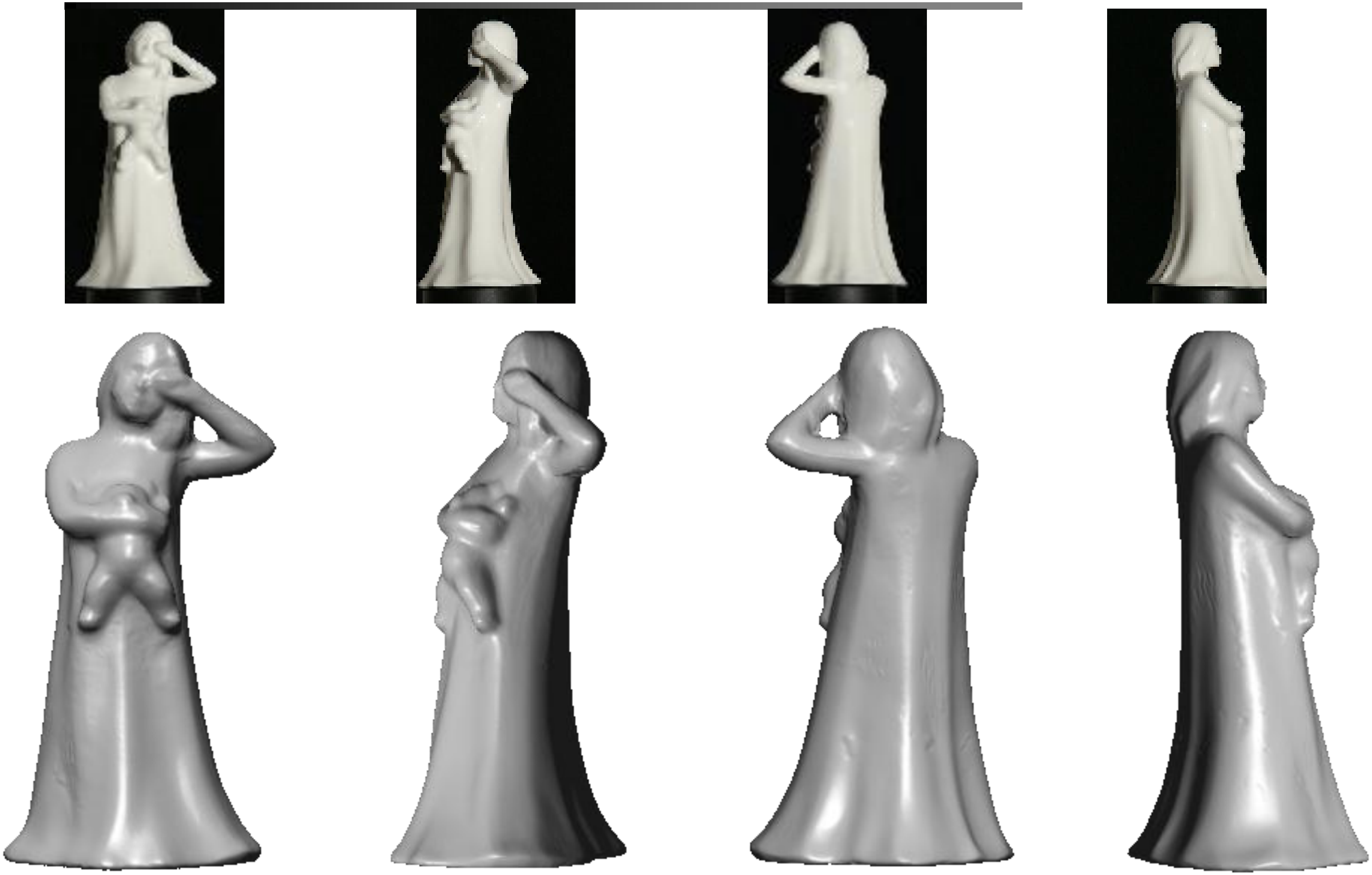
while mesh-not-converged **do**

  Optimise E_v with respect to $\mathbf{v}_1 \dots \mathbf{v}_\mathrm{F}$.

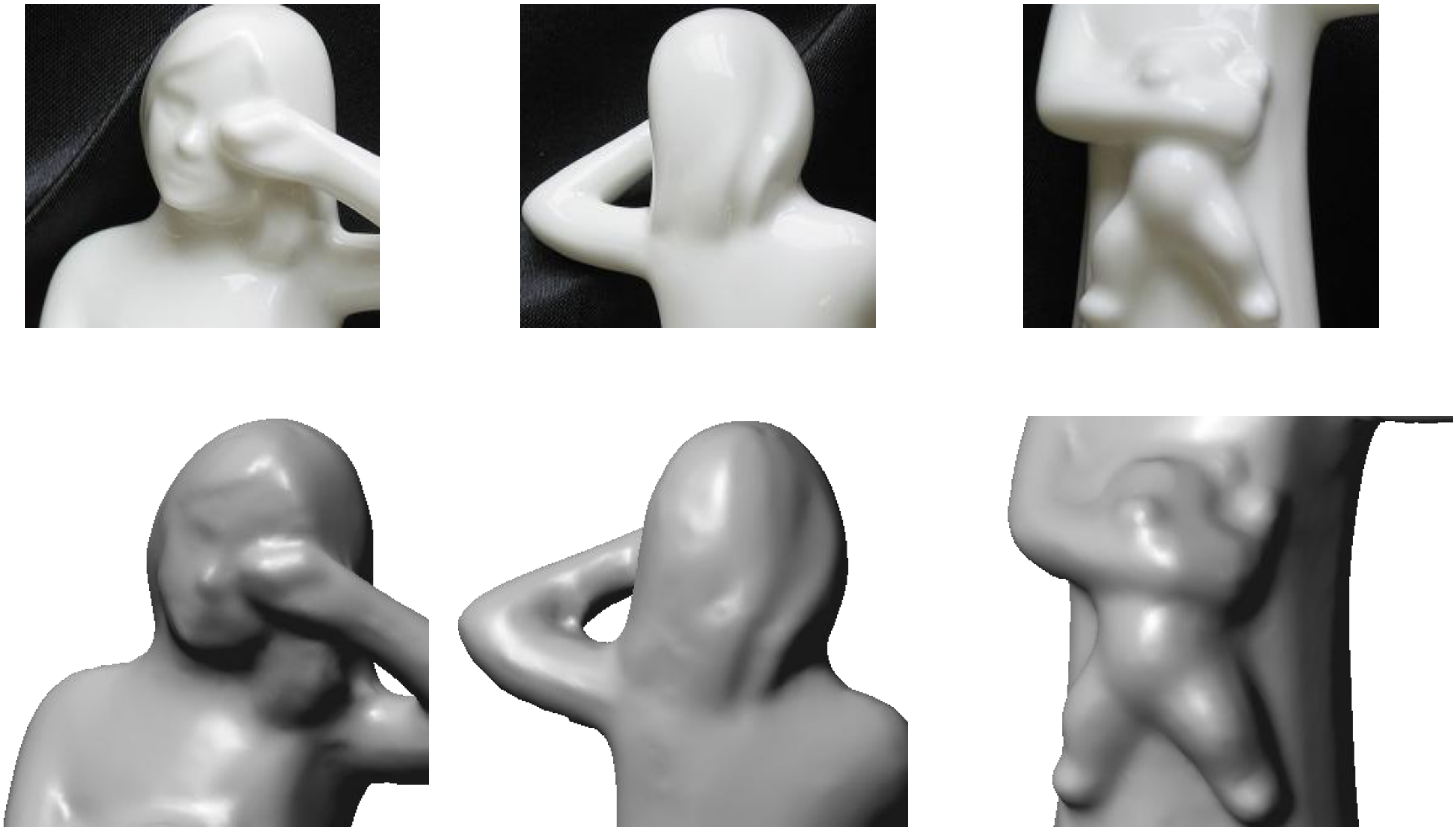
  Optimise E_m with respect to $\mathbf{x}_1 \dots \mathbf{x}_\mathrm{M}$.

end while

Results



Results

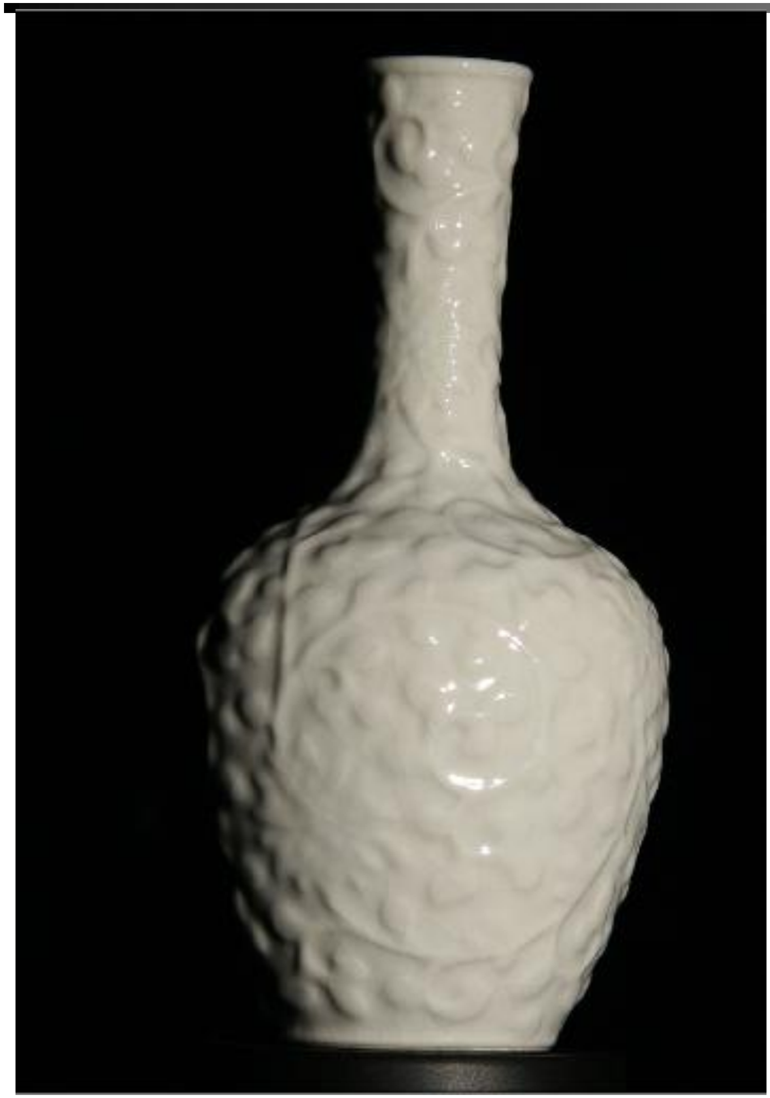


Results

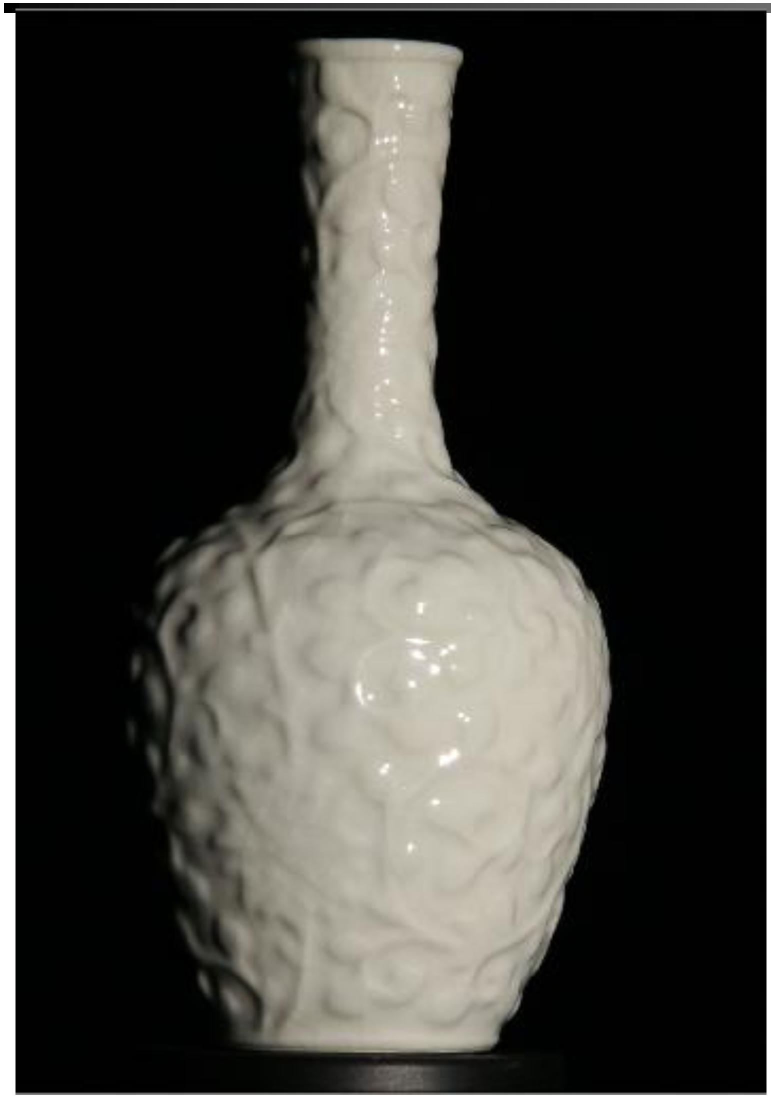
Results

Results – Fitzwilliam museum

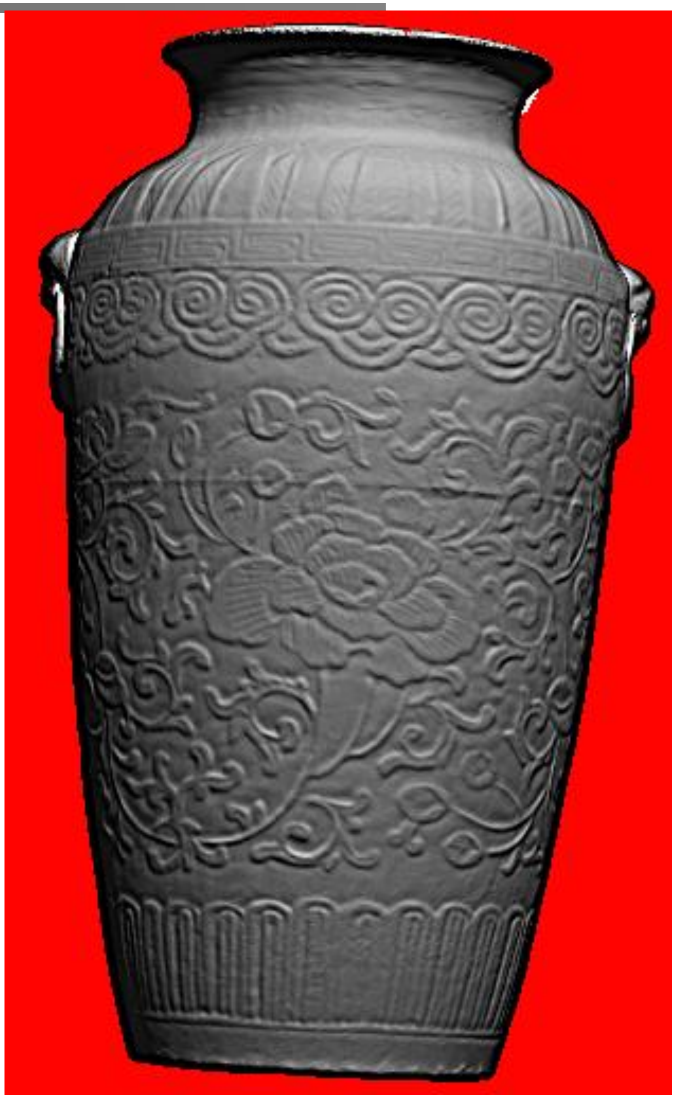
UNIVERSITY OF CAMBRIDGE



Results – Fitzwilliam museum



Results – Fitzwilliam museum



Results – Fitzwilliam museum

Results – Fitzwilliam museum

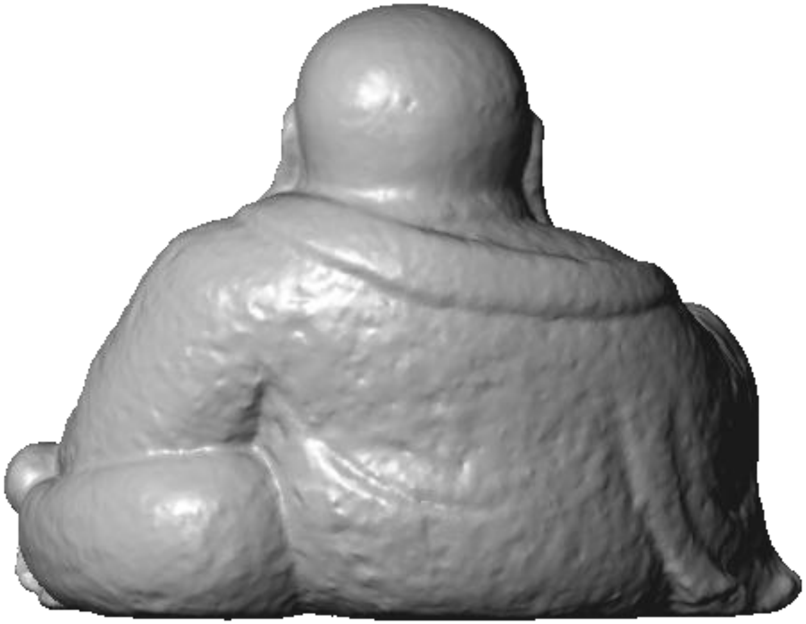
Multi-view Dense Stereo

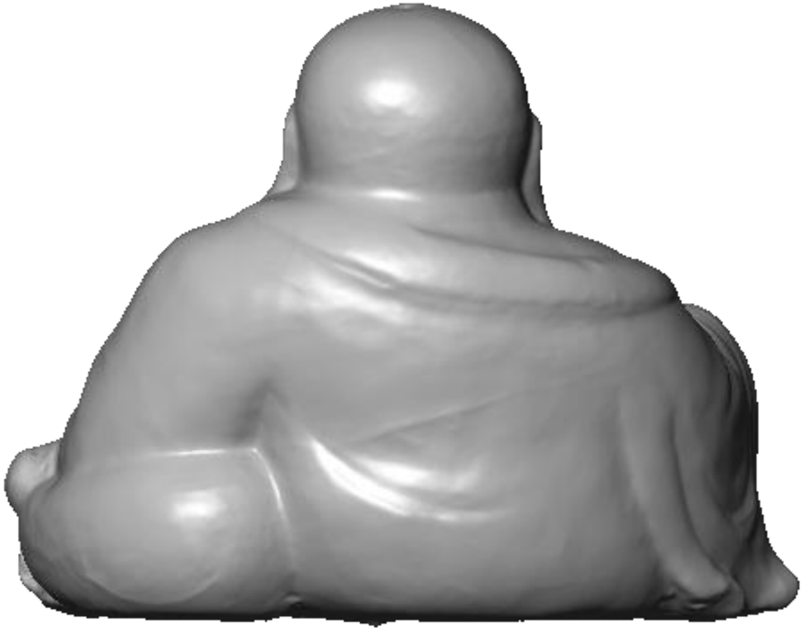
Multi-view Photometric Stereo

Results – Fitzwilliam museum

Multi-view Dense Stereo

Multi-view Photometric Stereo





Summary

UNIVERSITY OF CAMBRIDGE

- Accurate 3D shape from uncalibrated images

- Multi-view photometric stereo with uncalibrated lights