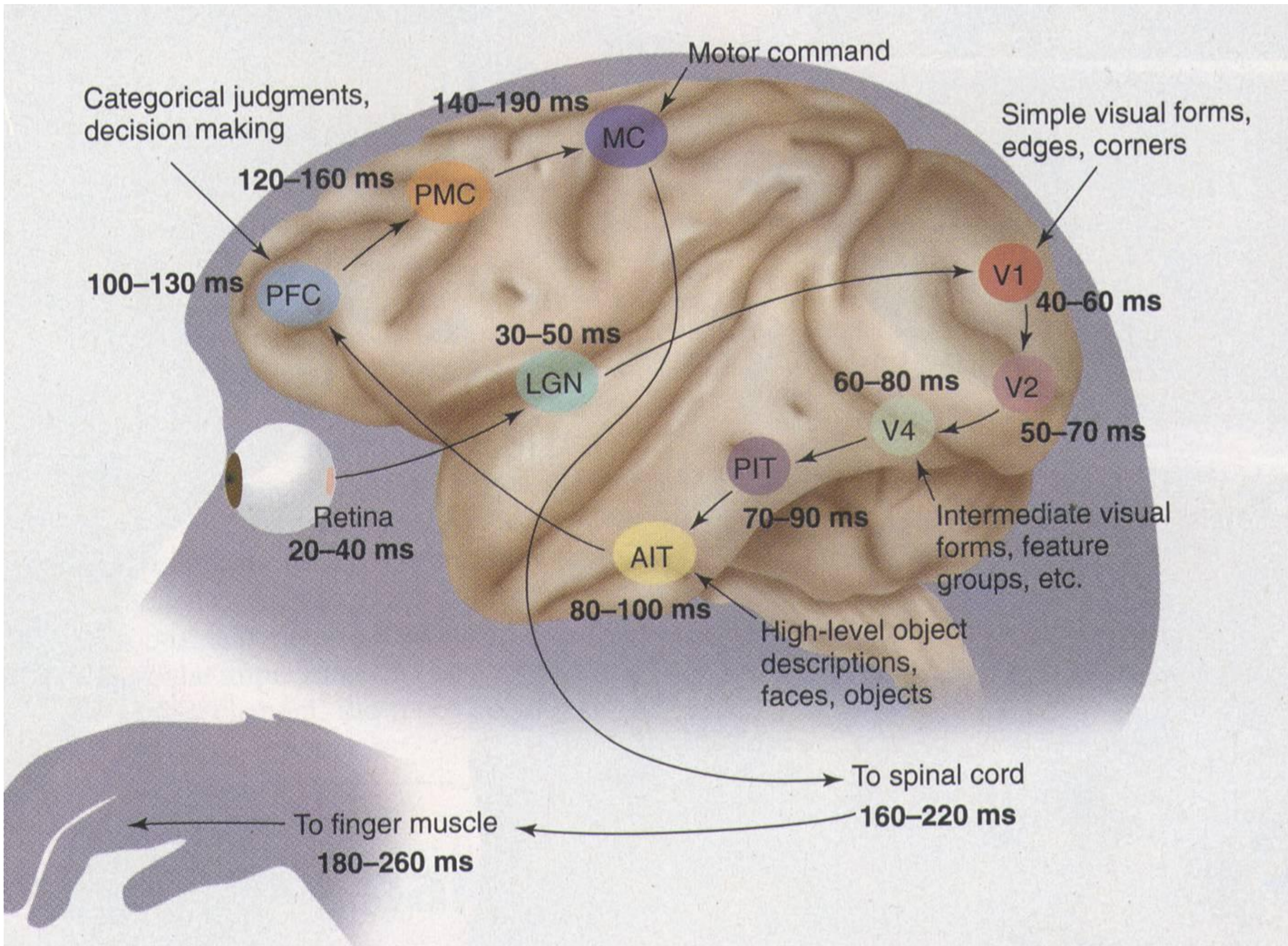
Computer Vision – 3D Shape

Roberto Cipolla
Department of Engineering

http://www.eng.cam.ac.uk/~cipolla/people.html
http://toshiba-europe.com/research/crl/cvg/

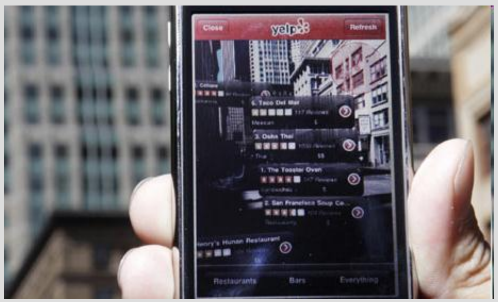
Vision: what is where by looking



Why? Images and Video

Computer vision is now in a wide range of products

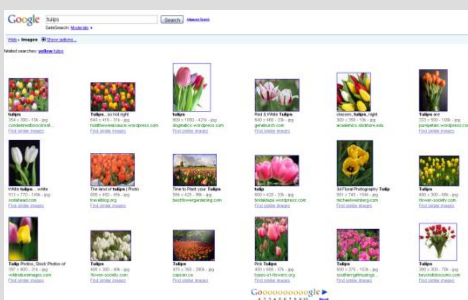
Mobile phones



Cars

Inspection and measurement

Image and video search



Games

Internet and shopping

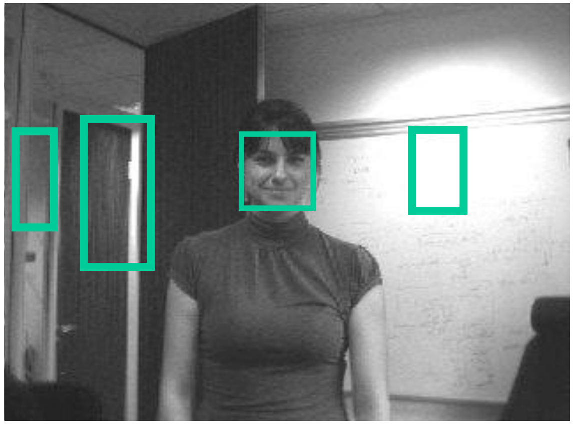
Overview

1. Computer vision at Cambridge
2. Review core technology of 3D shape recovery
3. Work in progress - Novel applications

Computer Vision at Cambridge

Computer Vision: 3R's

Reconstruction | **Recognition** | **Registration**



Reconstruction: Recover 3D shape

Recognition: Identify objects

Registration: Compute their position and pose

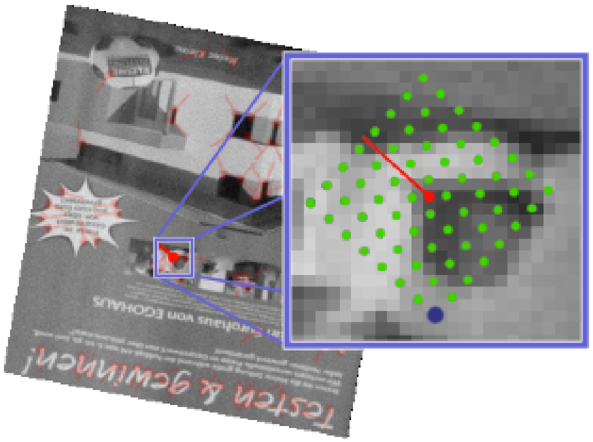
Registration?

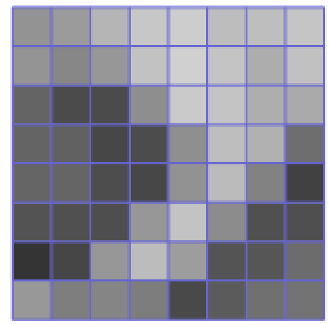
Target detection and pose estimation

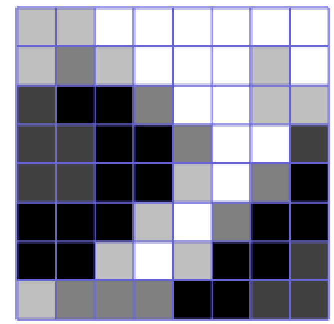
Registration

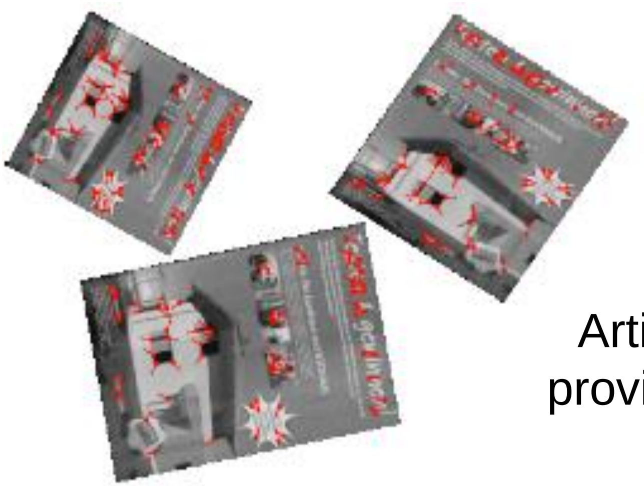
Quantised Patches

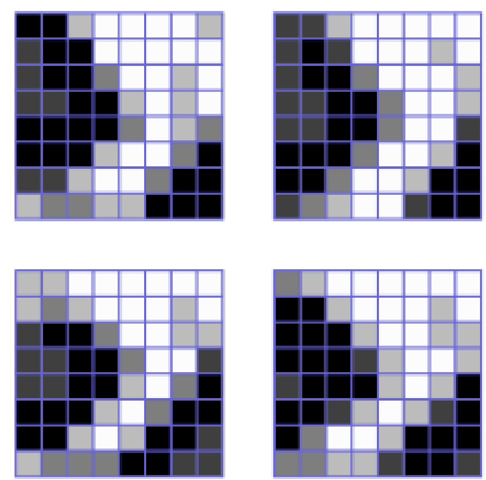
UNIVERSITY OF CAMBRIDGE





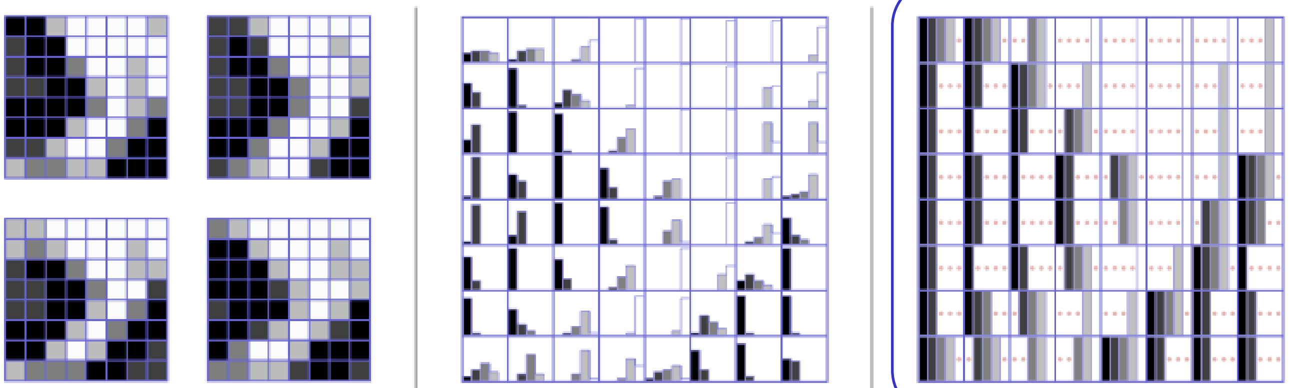




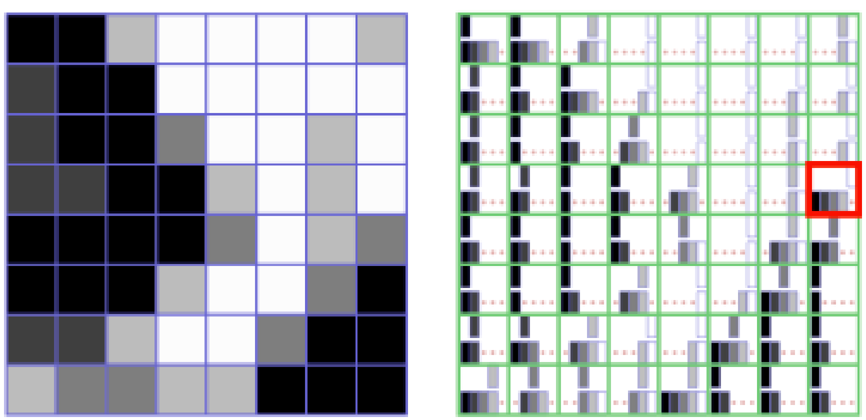


Artificially warped images provide multiple views of the same feature

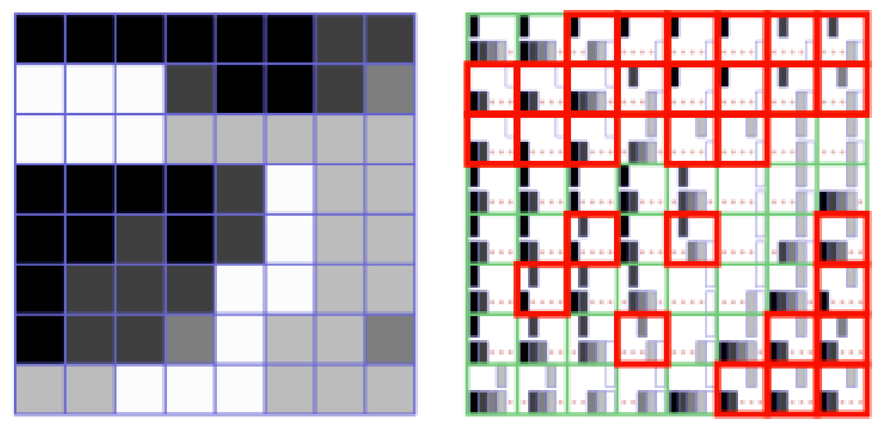
Histogram Intensity Patches



Binary representation of range of possible feature appearances permits fast error score computation

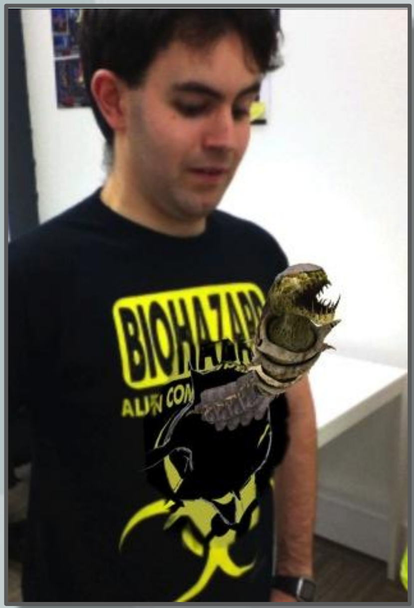


Matching patch: 1 error

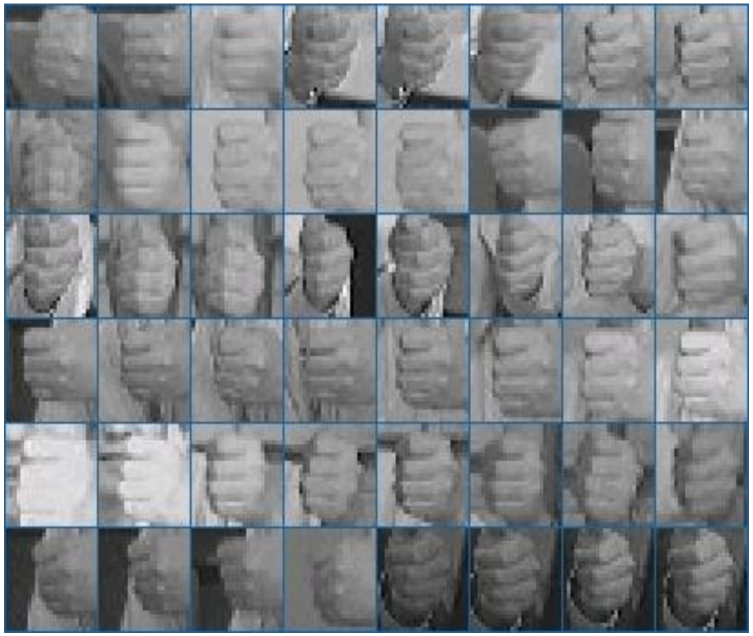


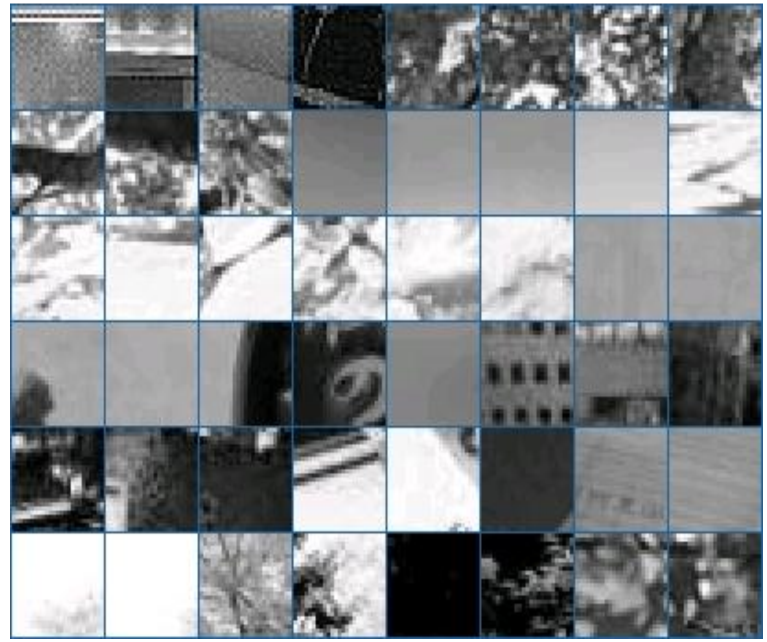
Non-matching patch: 30 errors

zappar
You're seeing things



Hand detection -Examples of training data



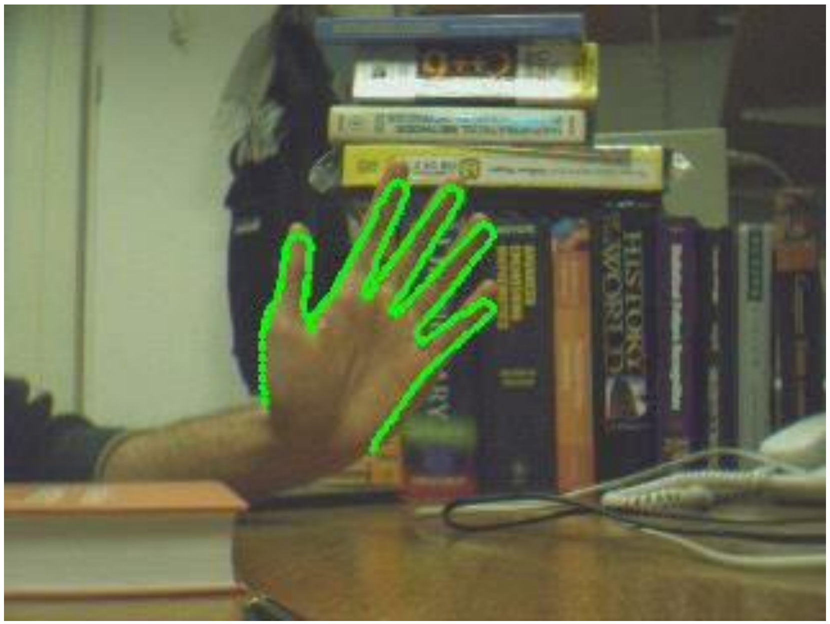


Registration - Hand detection and tracking

Hand Gesture UI

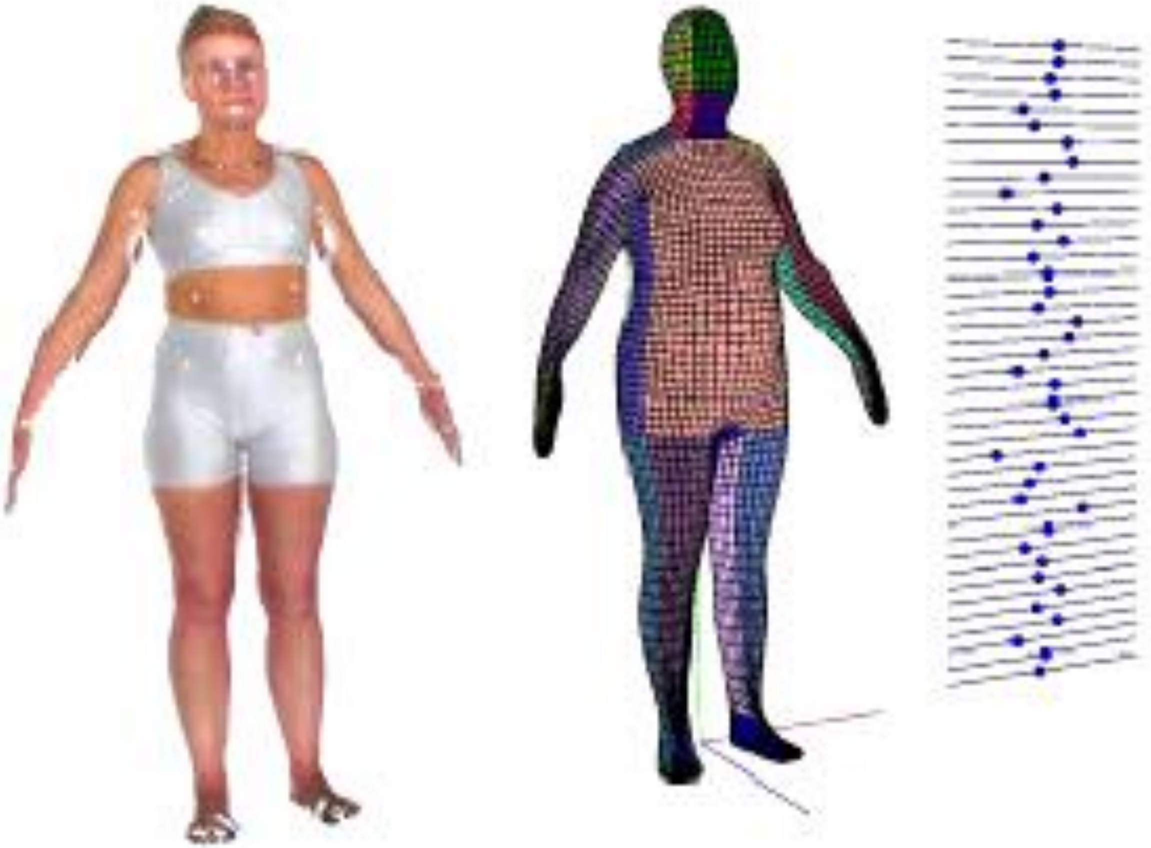
Hand detection system

UNIVERSITY OF CAMBRIDGE



Registration – human pose

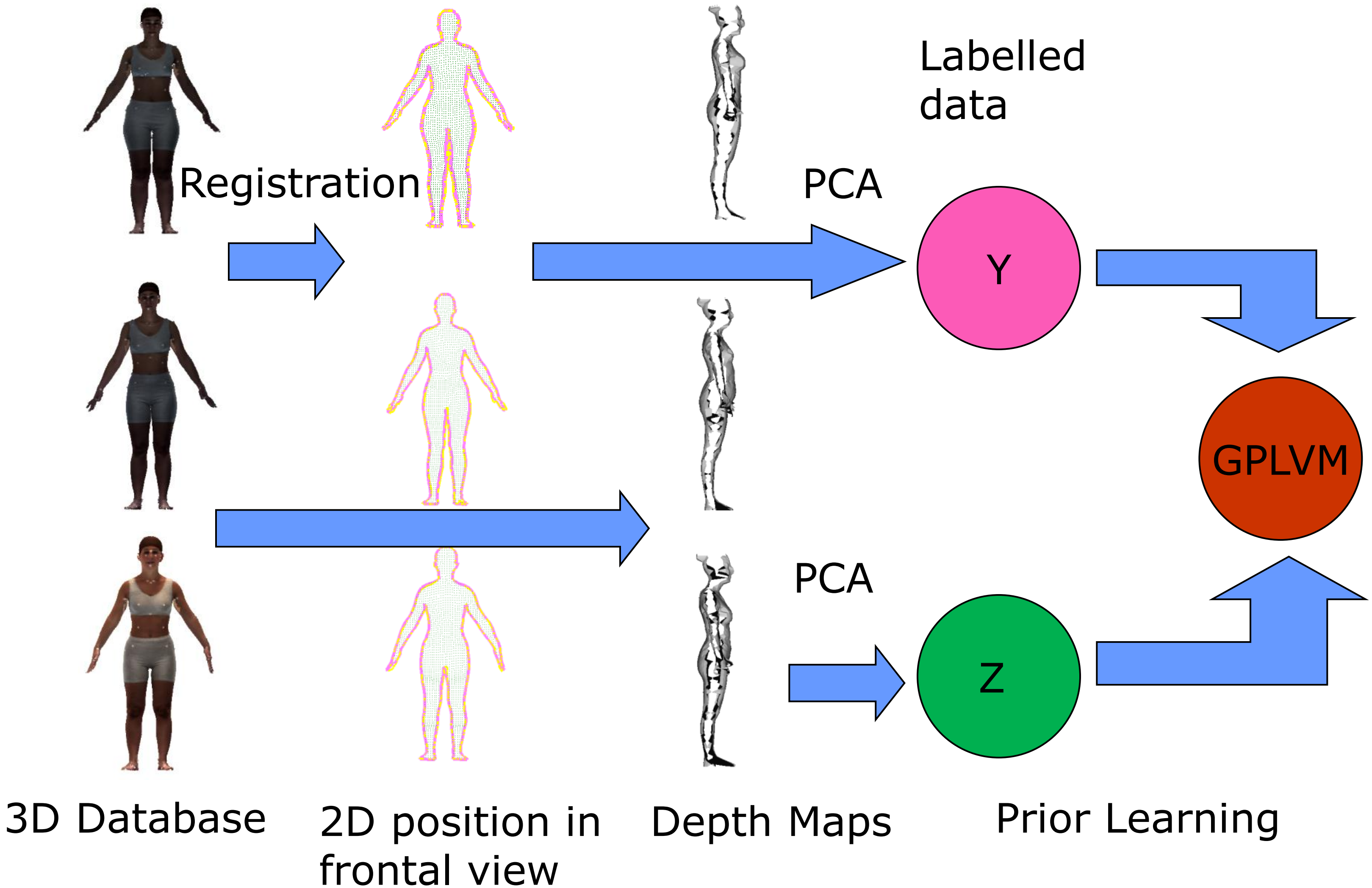
Registration – 3D shape



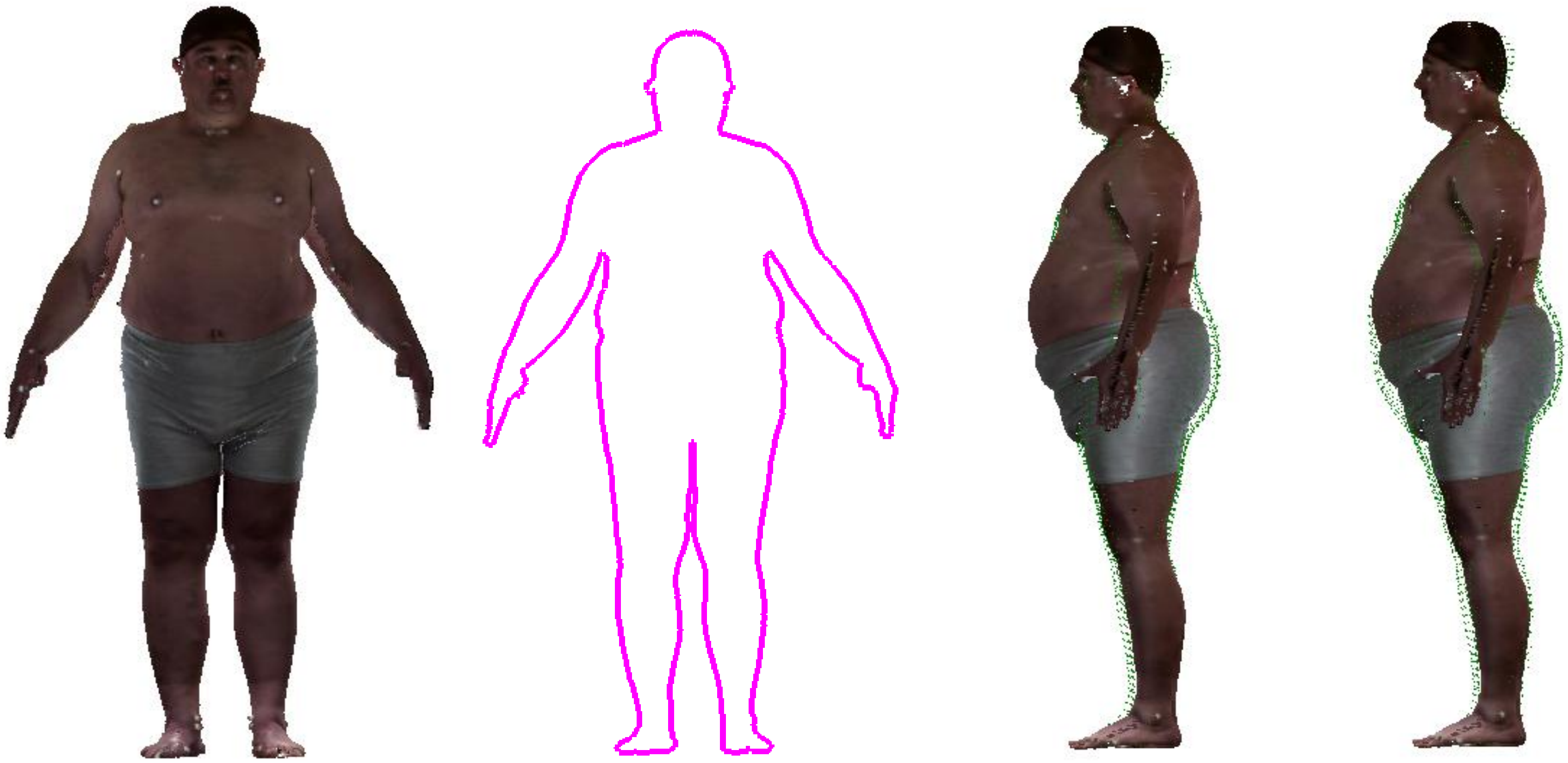
Single view reconstruction Exploiting 3D shape priors

Chen and Cipolla 2009 - 2011

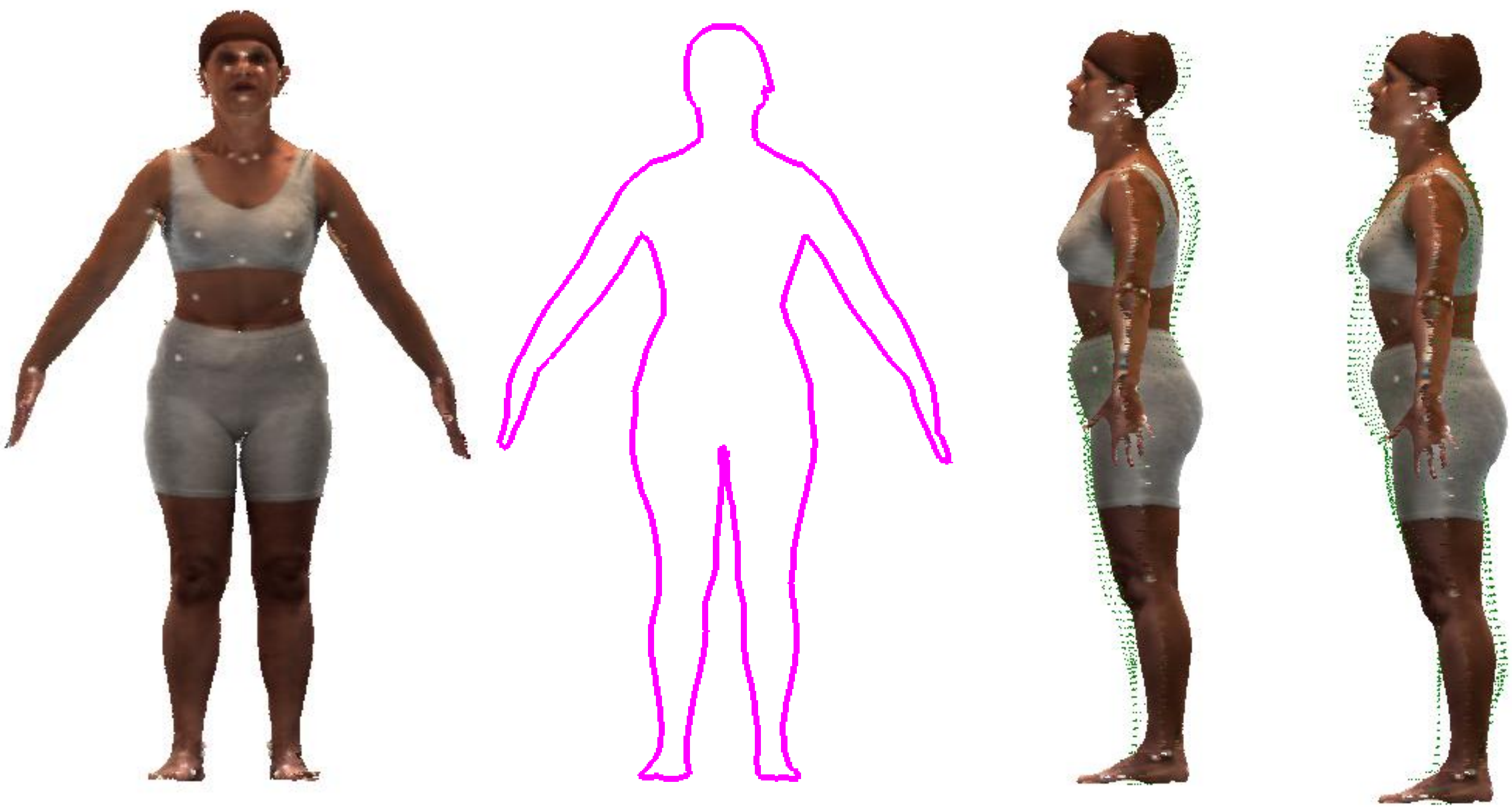
Learning Shape Priors



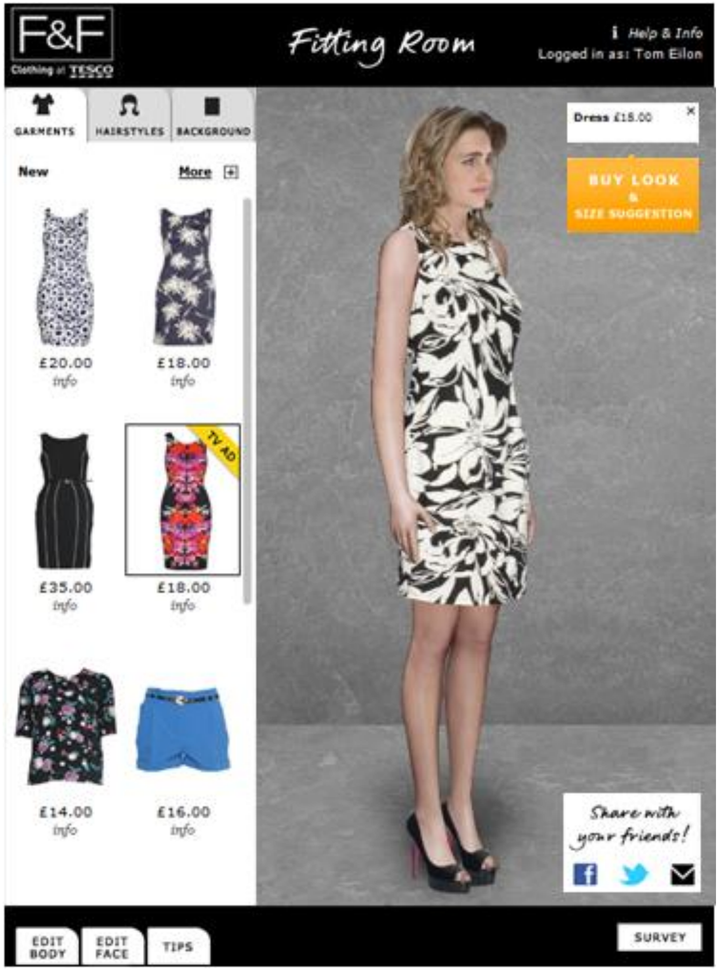
Results - Human Body Data

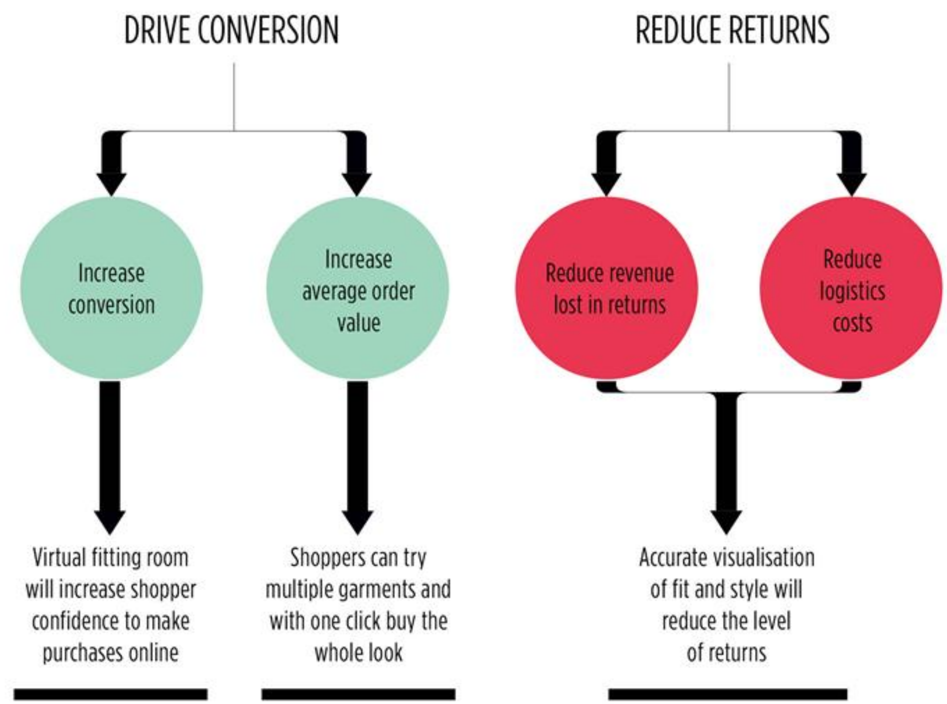


Results - Human Body Data



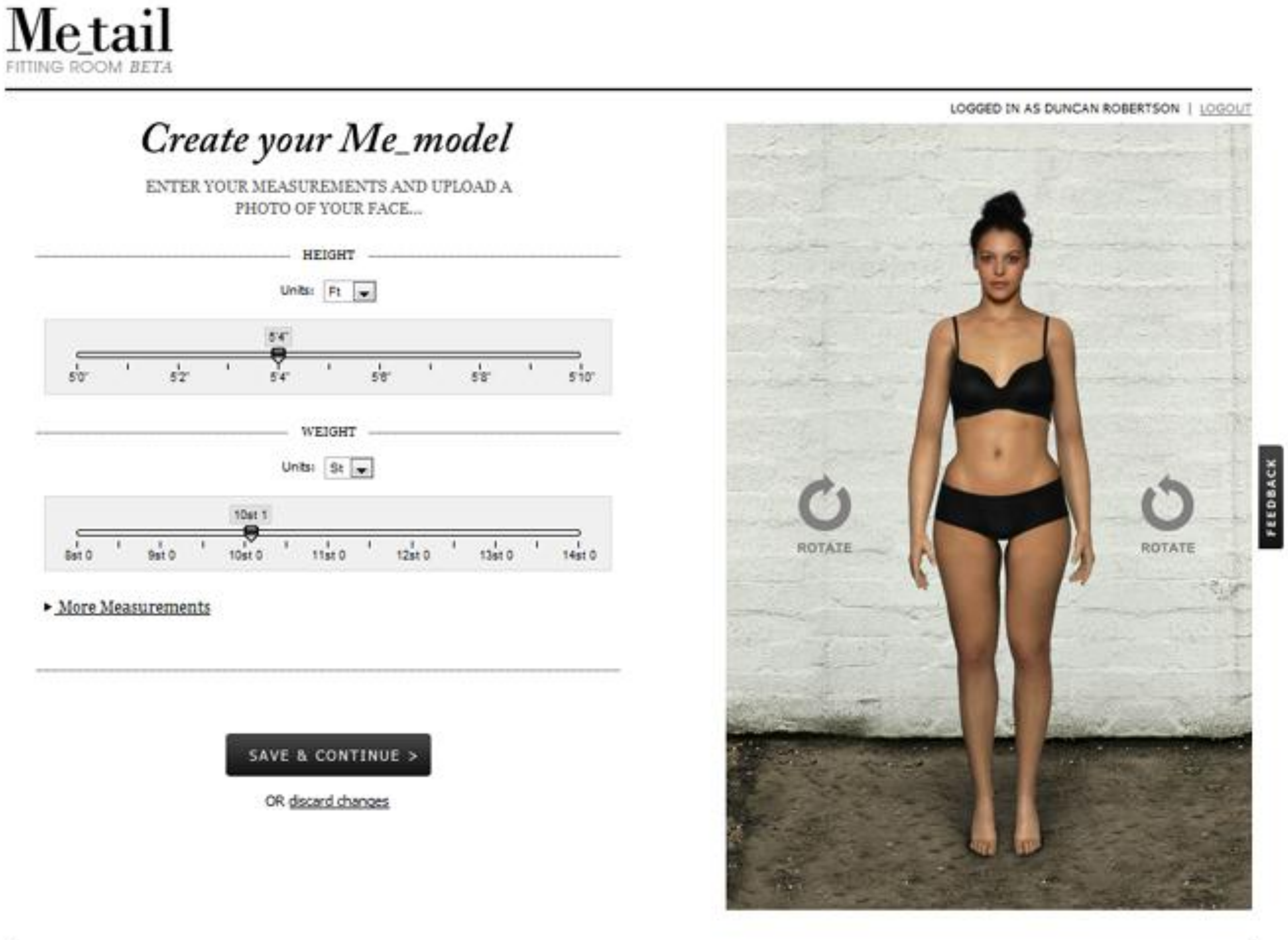
METAIL – A COMPUTER VISION BASED VIRTUAL FITTING ROOM

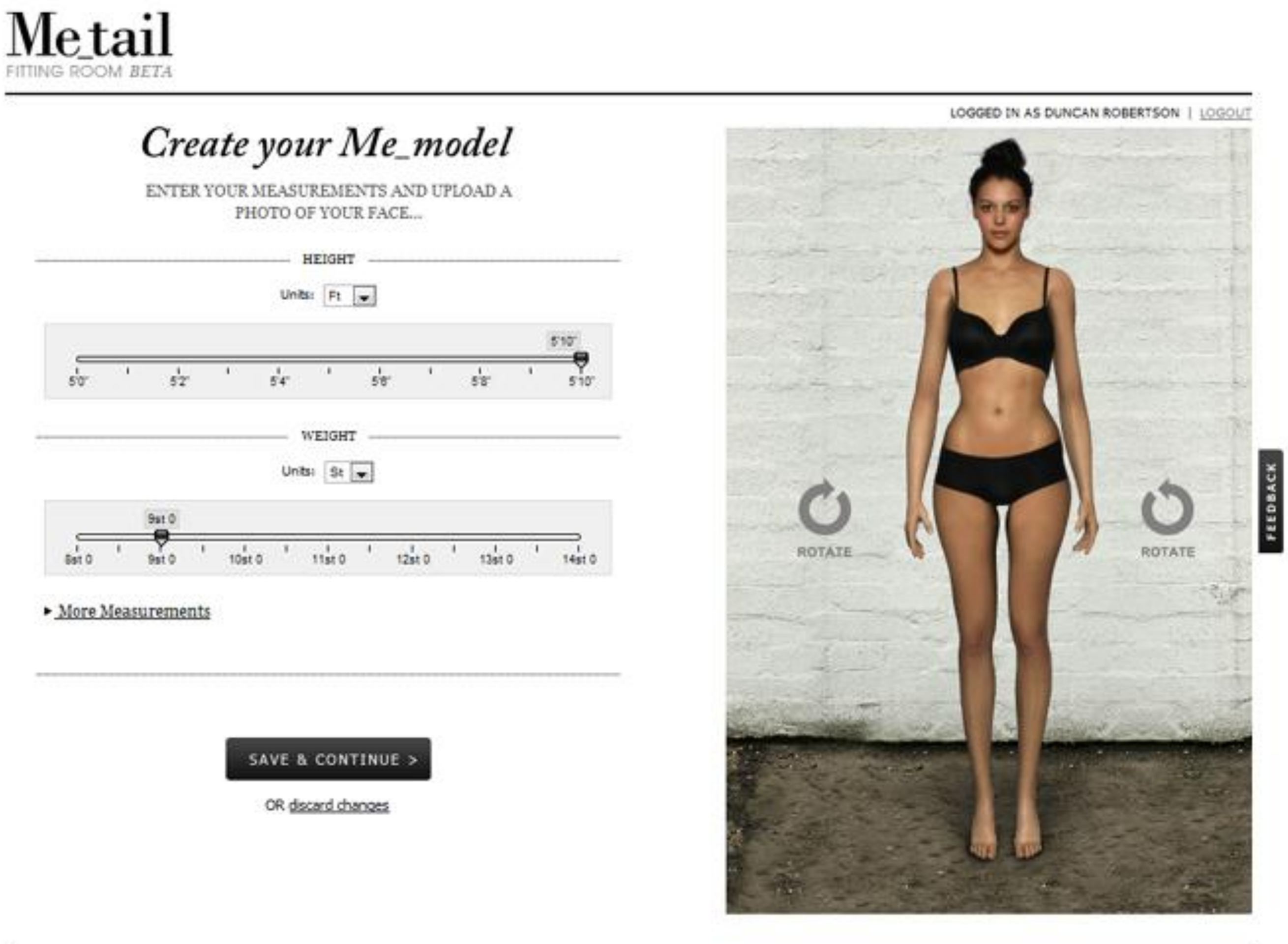




CLOTHING VISUALIZATION

Shoppers can visualize themselves wearing complete outfits.


Me tail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'4"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
10st 1
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE >
OR discard changes
ROTATE
ROTATE
FEEDBACK



Me_tail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'10"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
9st 0
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE >
OR discard changes
ROTATE
ROTATE
FEEDBACK

Me_tail
FITTING ROOM *BETA*

LOGGED IN AS DUNCAN ROBERTSON | LOGOUT

Create your Me_model

ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...

HEIGHT

Units: Ft

5'10"

5'0" 5'2" 5'4" 5'6" 5'8" 5'10"

WEIGHT

Units: St

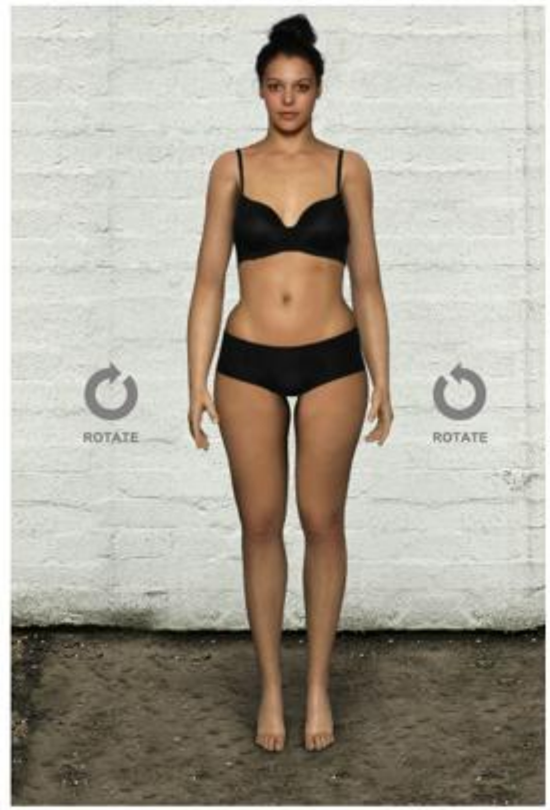
13st 0

8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0

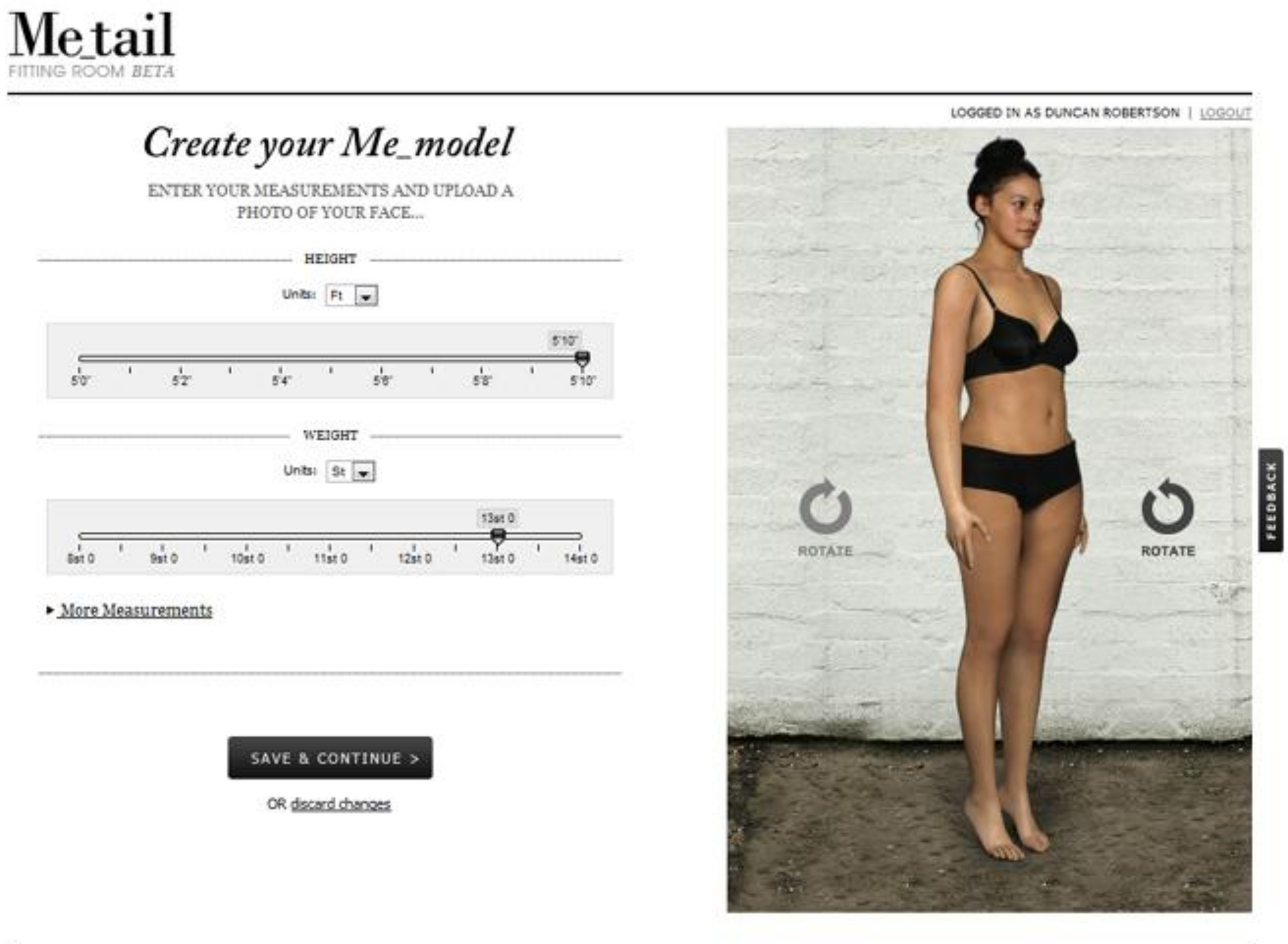
▸ More Measurements

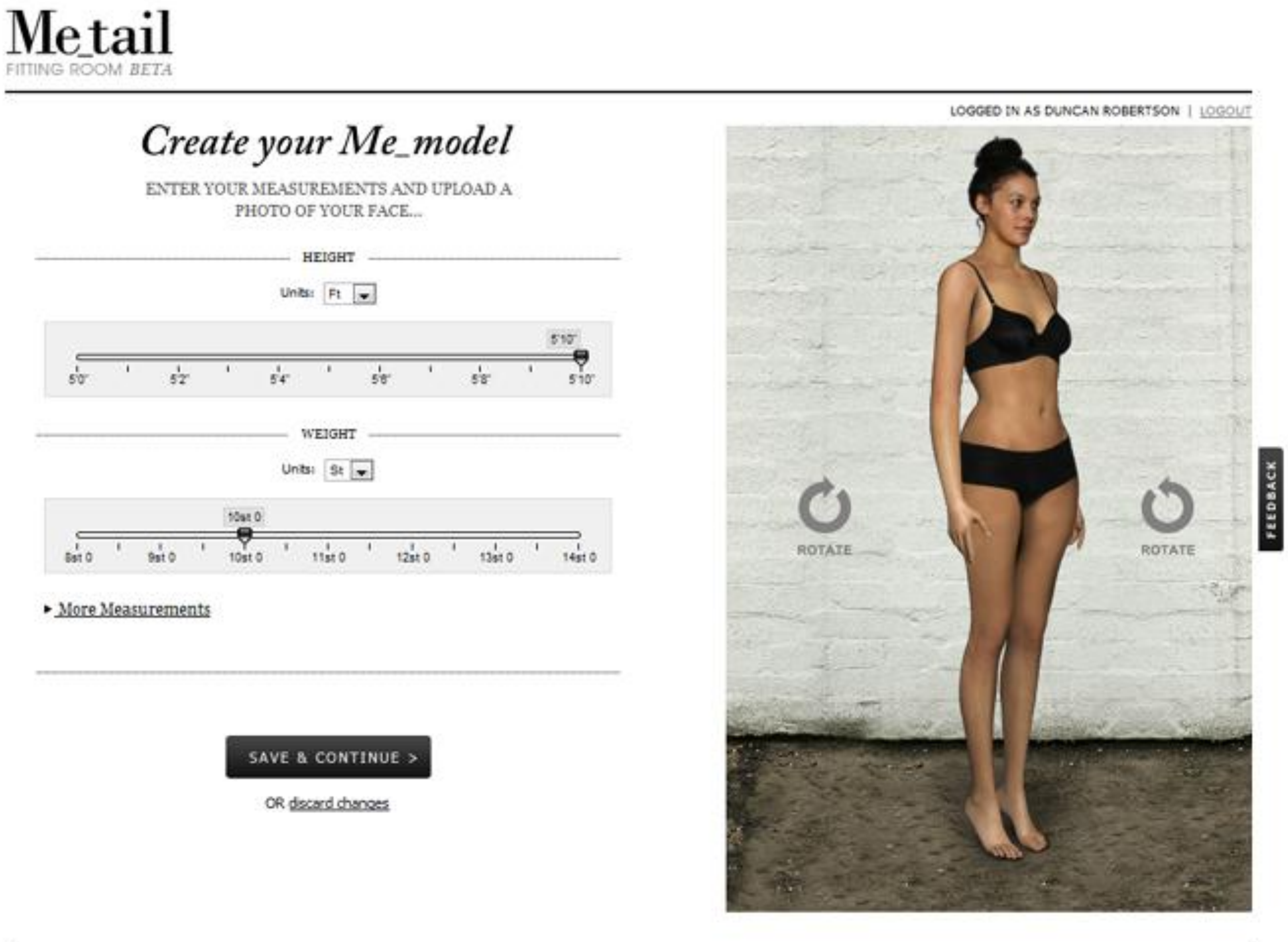
SAVE & CONTINUE >

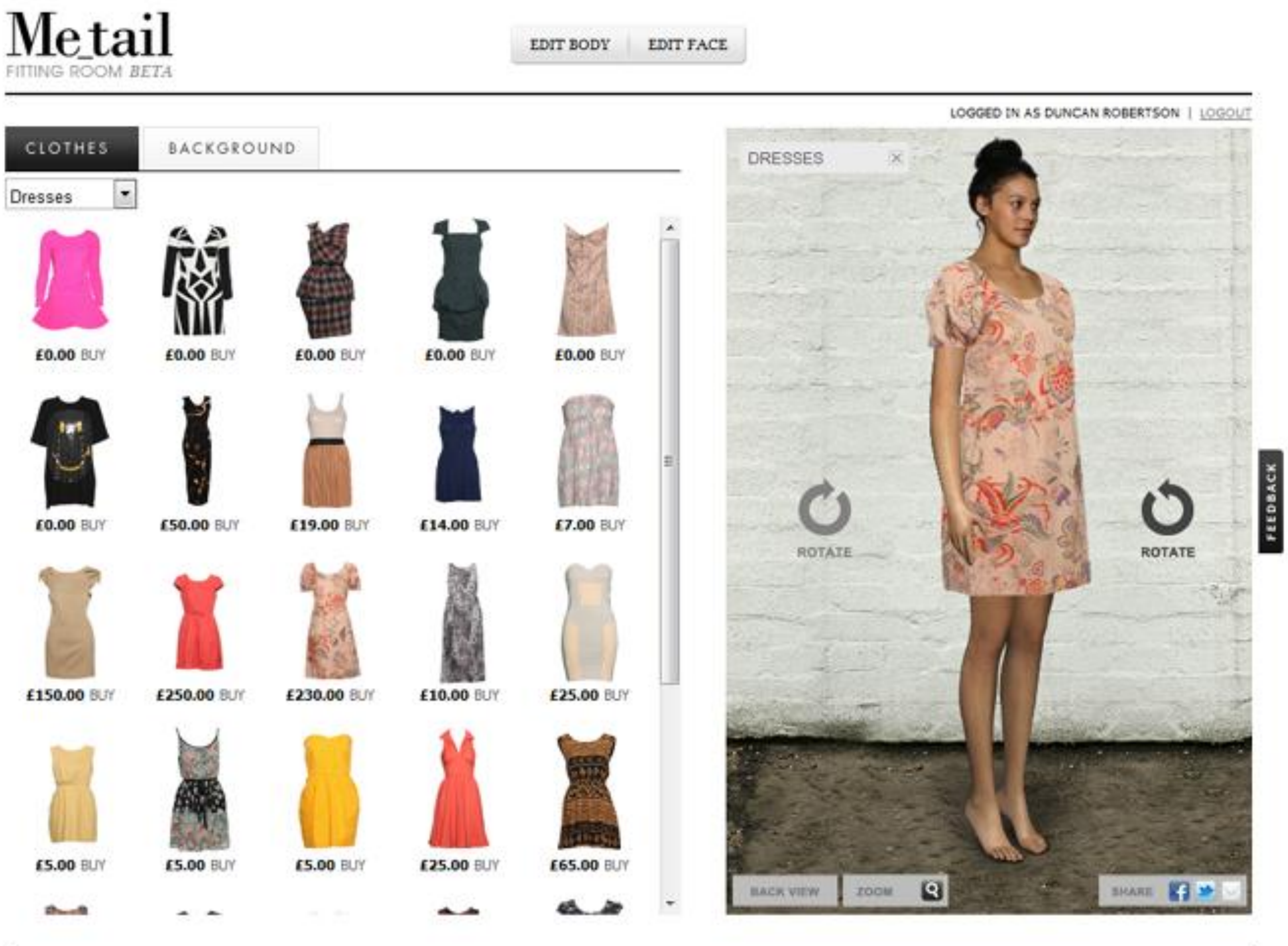
OR discard changes



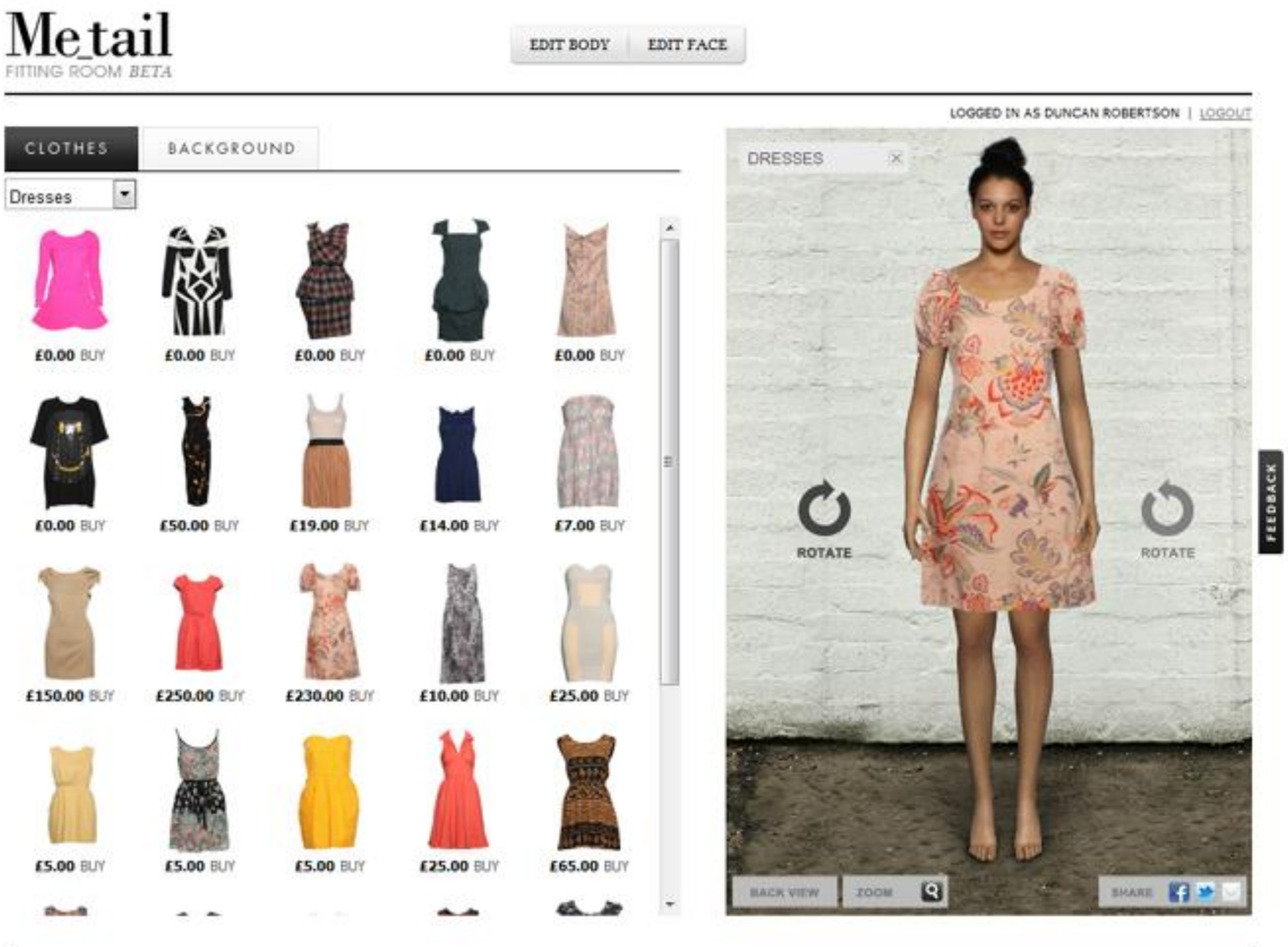
FEEDBACK


Me_tail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'10"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
13st 0
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE >
OR discard changes
ROTATE
ROTATE
FEEDBACK

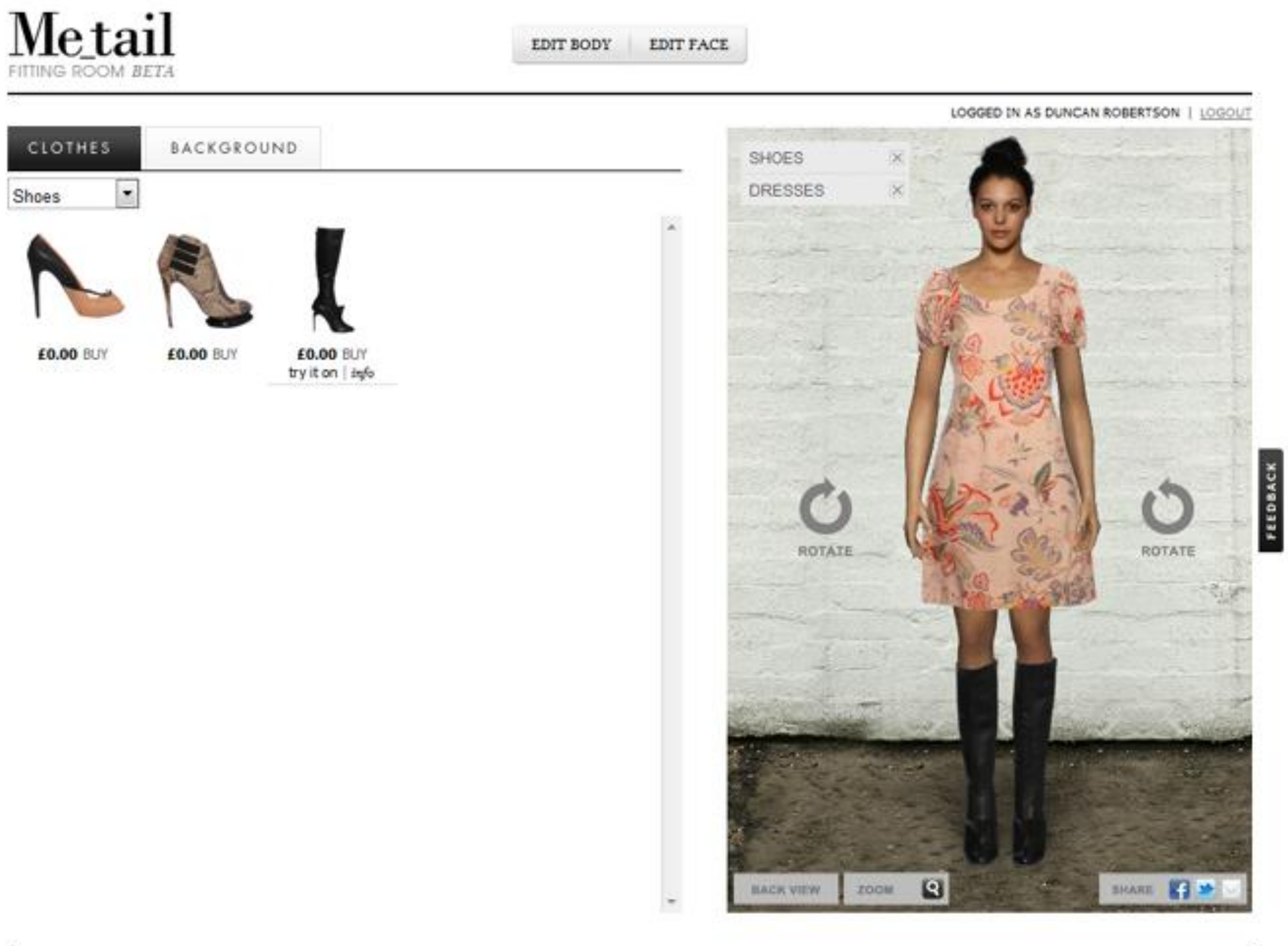

Me_tail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'10"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
10st 0
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE >
OR discard changes
ROTATE
ROTATE
FEEDBACK

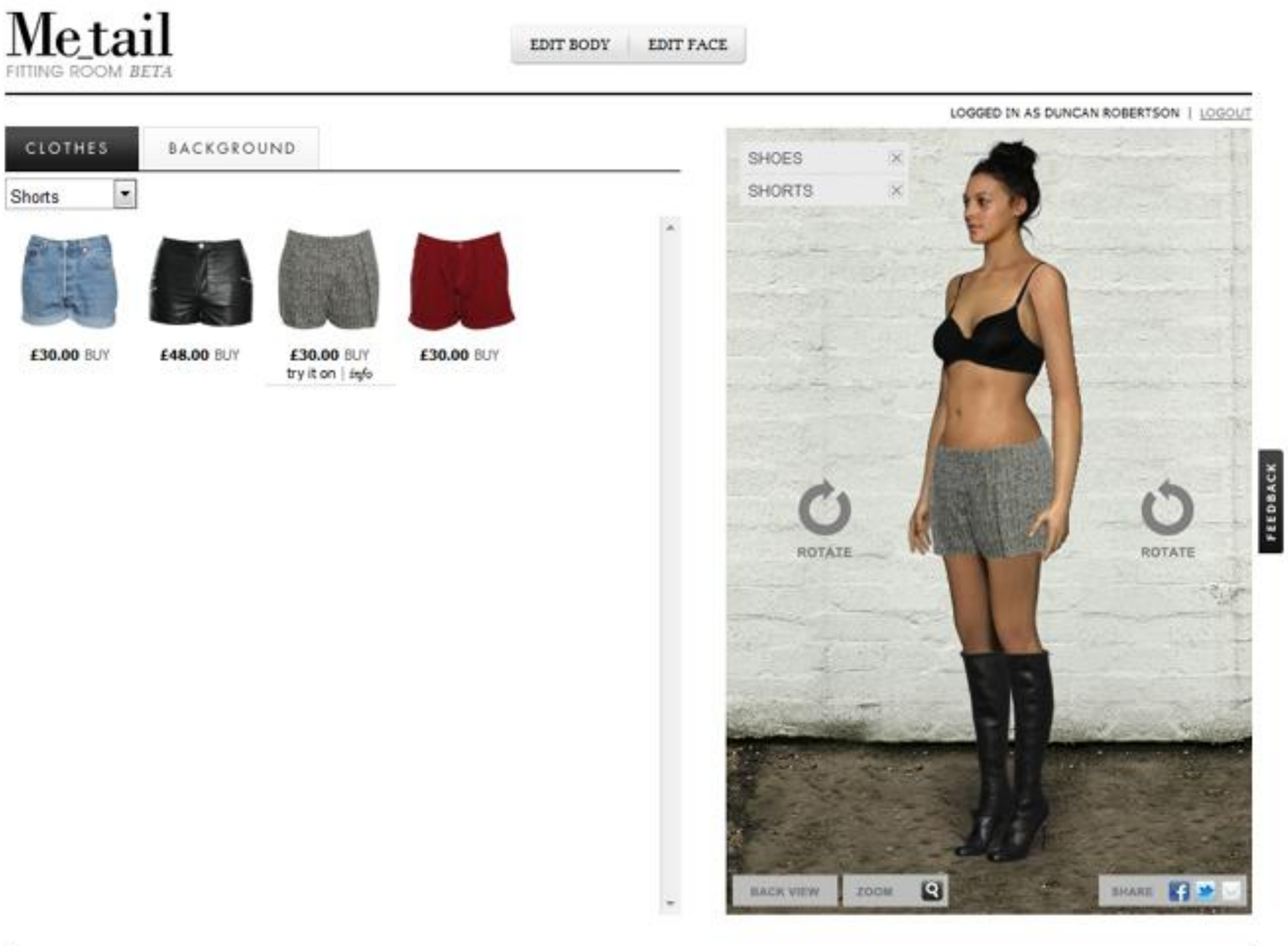


Metail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Dresses
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£50.00 BUY
£19.00 BUY
£14.00 BUY
£7.00 BUY
£150.00 BUY
£250.00 BUY
£230.00 BUY
£10.00 BUY
£25.00 BUY
£5.00 BUY
£5.00 BUY
£5.00 BUY
£25.00 BUY
£65.00 BUY
DRESSES
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE

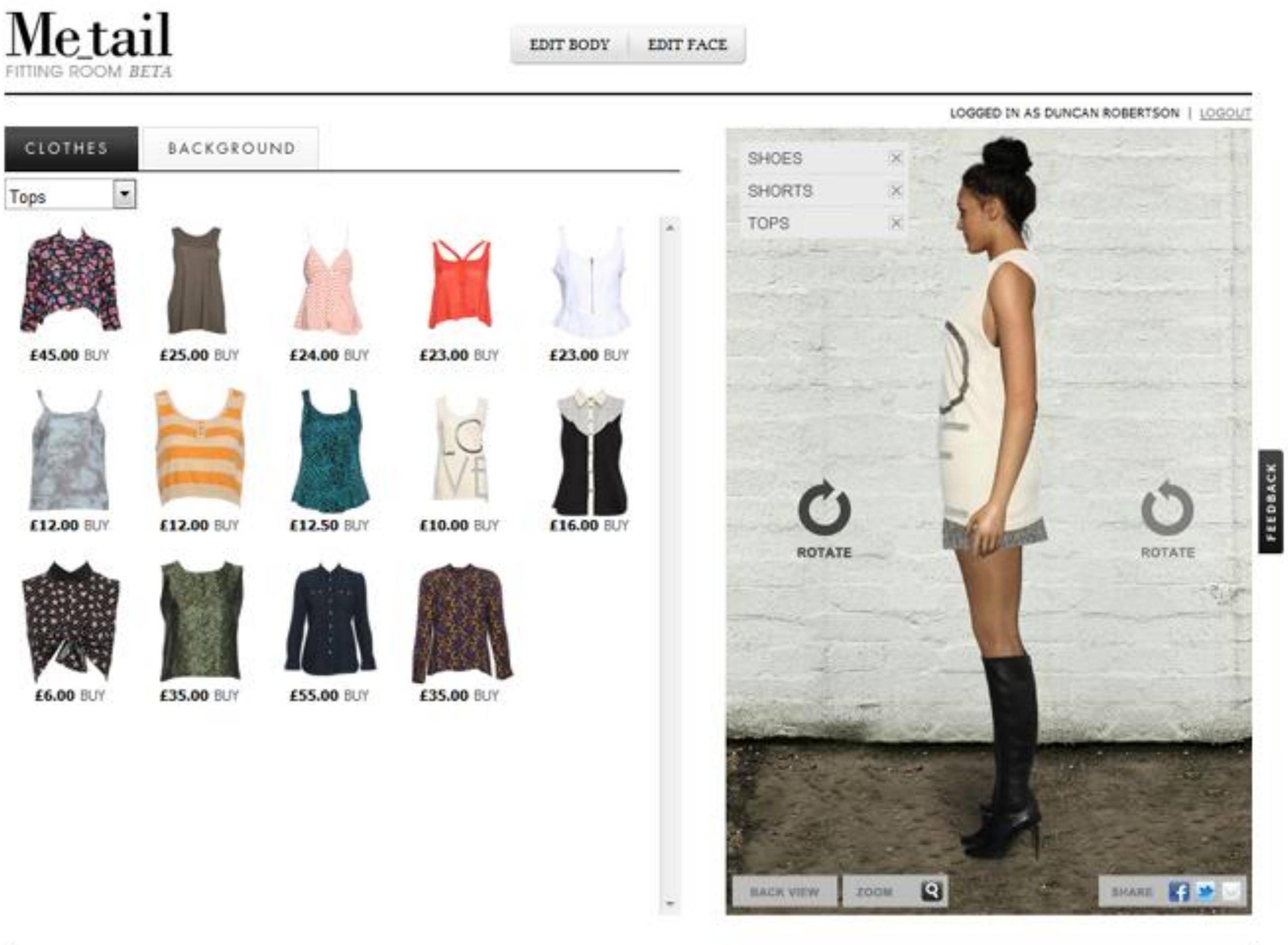


Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Dresses
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£50.00 BUY
£19.00 BUY
£14.00 BUY
£7.00 BUY
£150.00 BUY
£250.00 BUY
£230.00 BUY
£10.00 BUY
£25.00 BUY
£5.00 BUY
£5.00 BUY
£5.00 BUY
£25.00 BUY
£65.00 BUY
DRESSES
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE

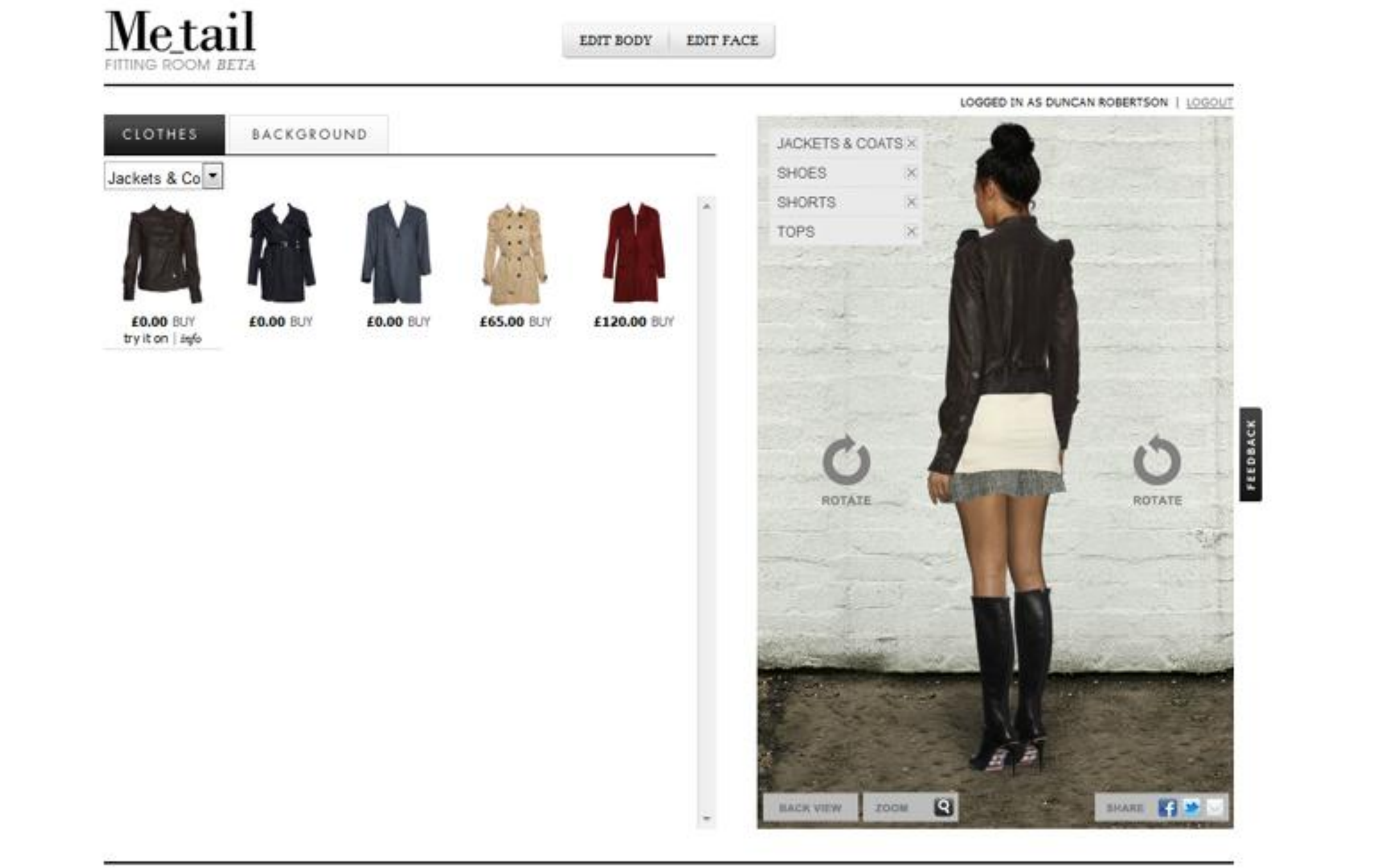

Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Shoes
£0.00 BUY
£0.00 BUY
£0.00 BUY
try it on | info
SHOES
DRESSES
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE



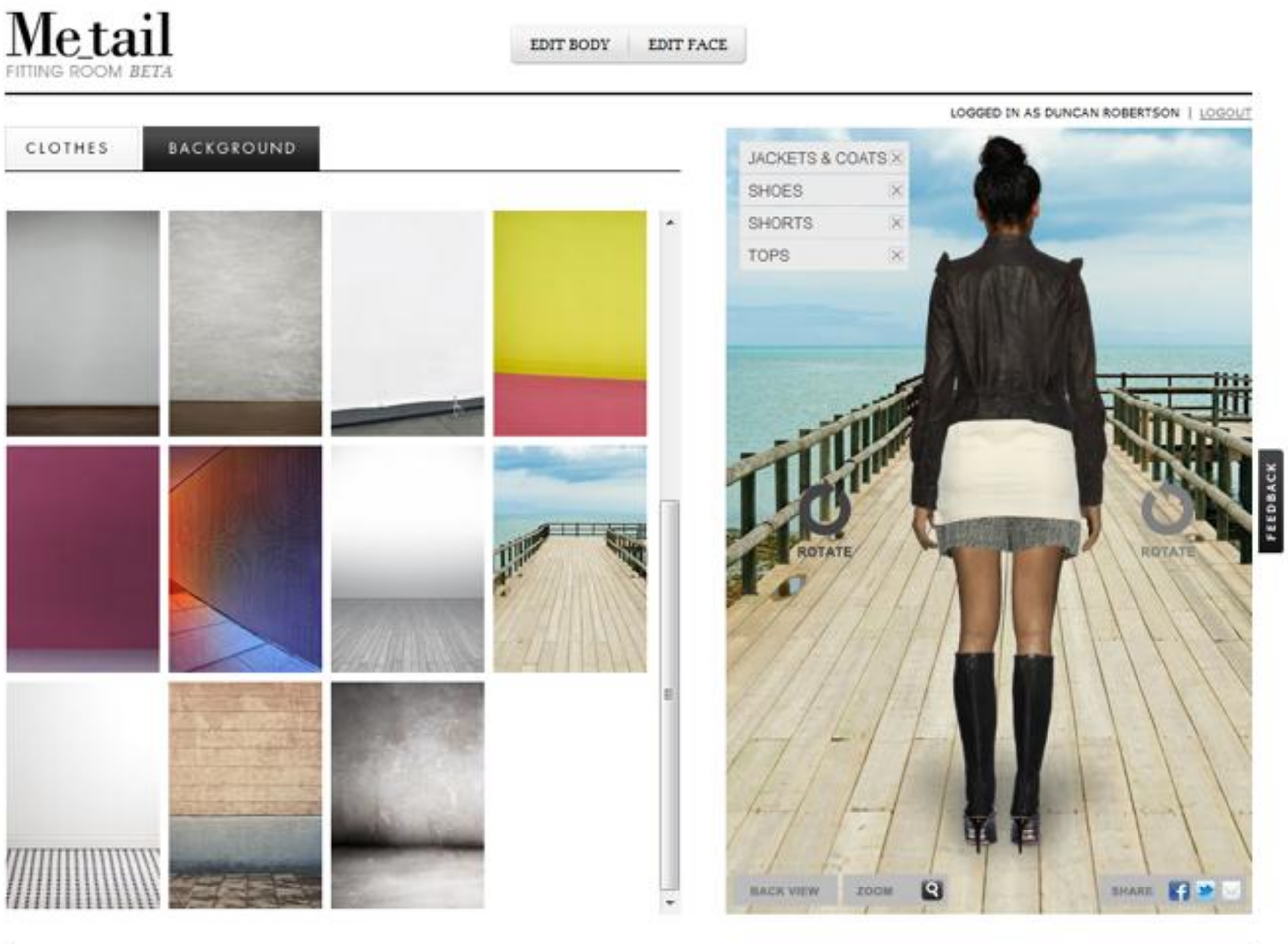
Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Shorts
£30.00 BUY
£48.00 BUY
£30.00 BUY
try it on | info
£30.00 BUY
SHOES
SHORTS
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE

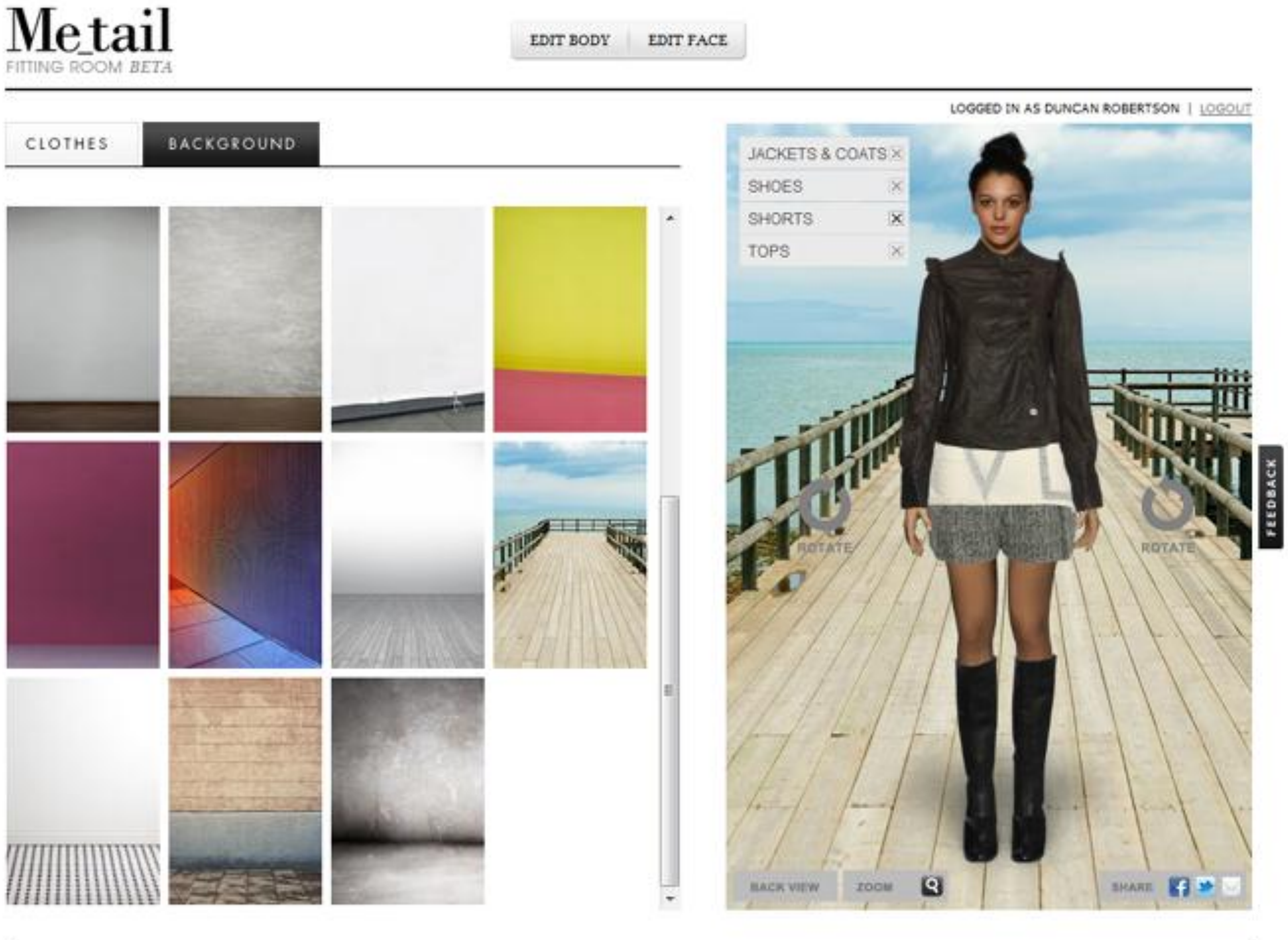


Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Tops
£45.00 BUY
£25.00 BUY
£24.00 BUY
£23.00 BUY
£23.00 BUY
£12.00 BUY
£12.00 BUY
£12.50 BUY
£10.00 BUY
£16.00 BUY
£6.00 BUY
£35.00 BUY
£55.00 BUY
£35.00 BUY
SHOES
SHORTS
TOPS
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE



Metail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Jackets & Co
£0.00 BUY
try it on | info
£0.00 BUY
£0.00 BUY
£65.00 BUY
£120.00 BUY
JACKETS & COATS
SHOES
SHORTS
TOPS
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE


Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
JACKETS & COATS
SHOES
SHORTS
TOPS
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE

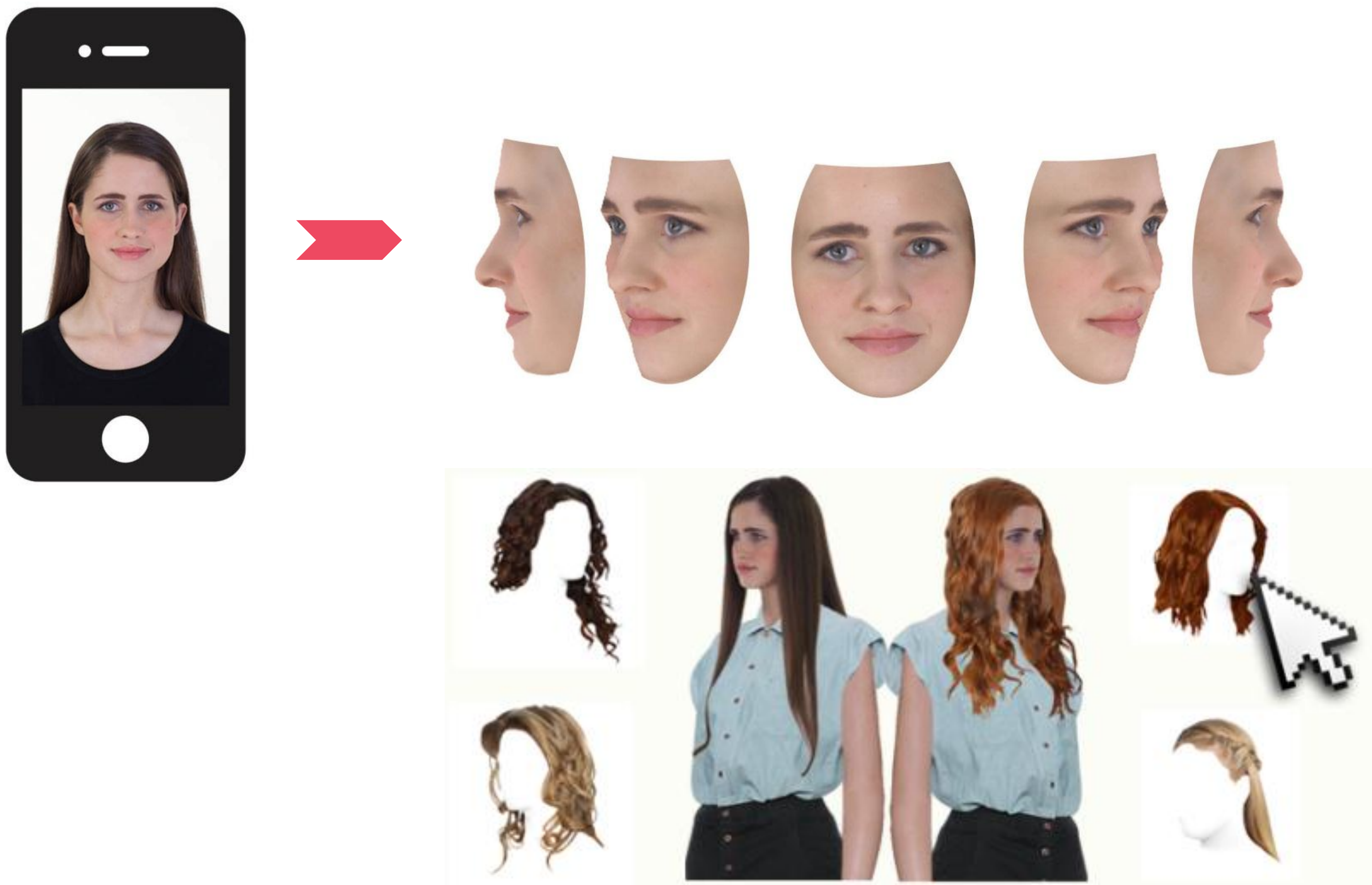

Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
JACKETS & COATS
SHOES
SHORTS
TOPS
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE

Metail – online shopping

(a) (b) (c)

FACE FROM PHOTO

Patented technology is used to create an accurate 3D model of the face, personalising the shopper's model.



BODY SHAPE FROM SILHOUETTE

The shape prior also facilitates 3D shape recovery from the body's silhouette in a 2D photograph.

HUMAN BODY MODELLING
OUTFIT VISUALIZATION FOR ONLINE SHOPPING

Me_

Me_tail

Body tracking with 3D model

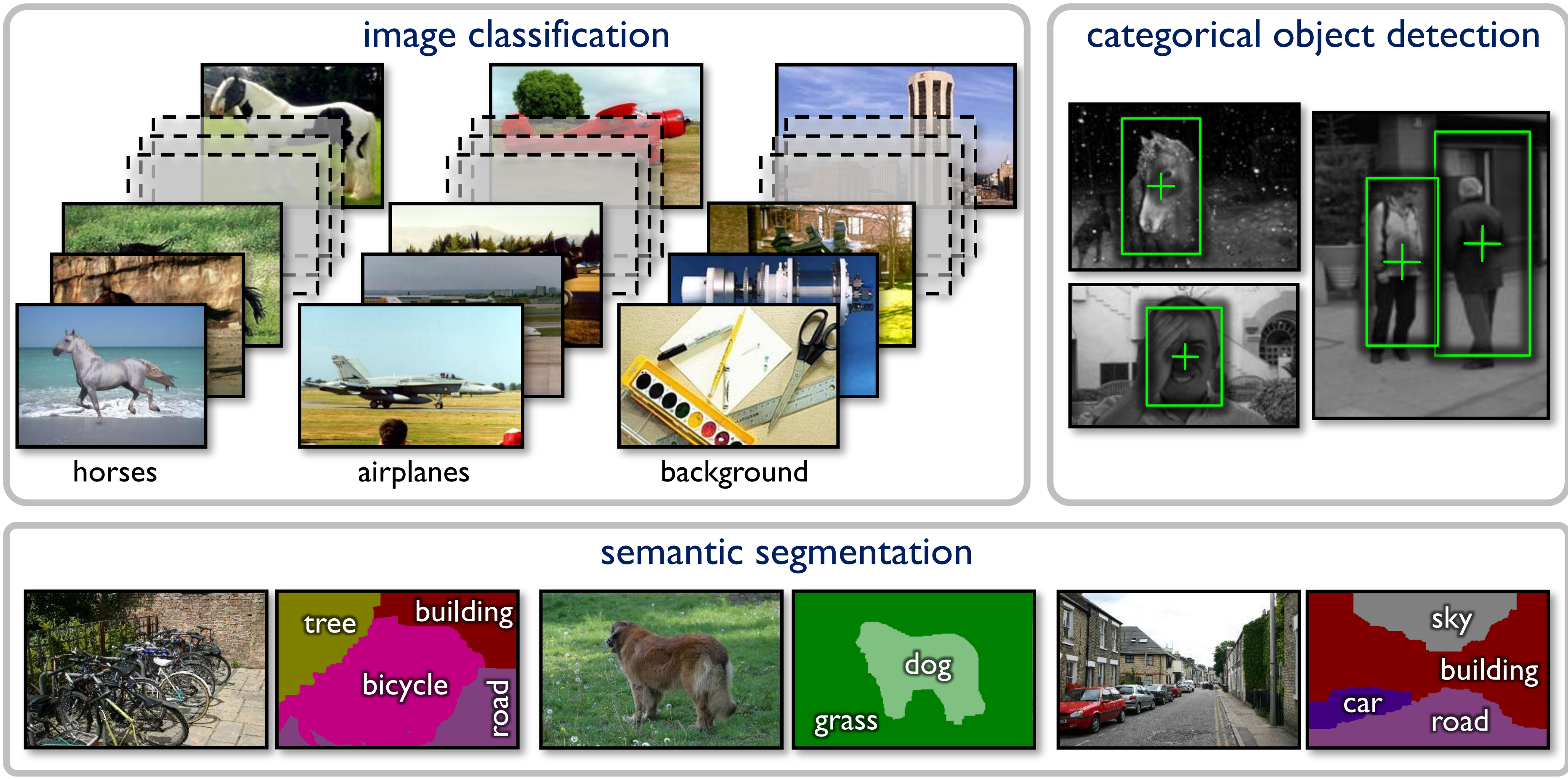
- Cloth simulation and rendering

Virtual Fashion Show

Recognition?

Recognition

UNIVERSITY OF CAMBRIDGE



Segmentation in Video

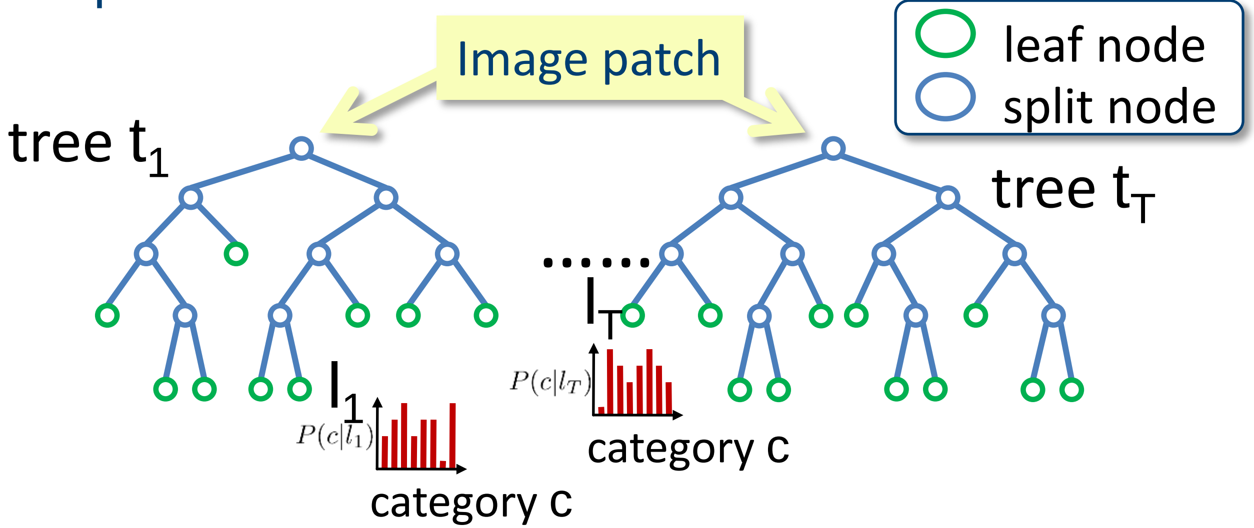
Classification –

Randomized Decision Forests and SVM

Label propagation – Semi-supervised learning with GPS and label transfer

Semantic Texton Forests for classification

- Learn a set of tree structured classifiers which take an *image patch* as input and output a label distribution of its *centre-pixel*.



- Forest is ensemble of T trees
  - classification is $P(c|L) = \sum_{t=1}^{T} P(c|l_t)$

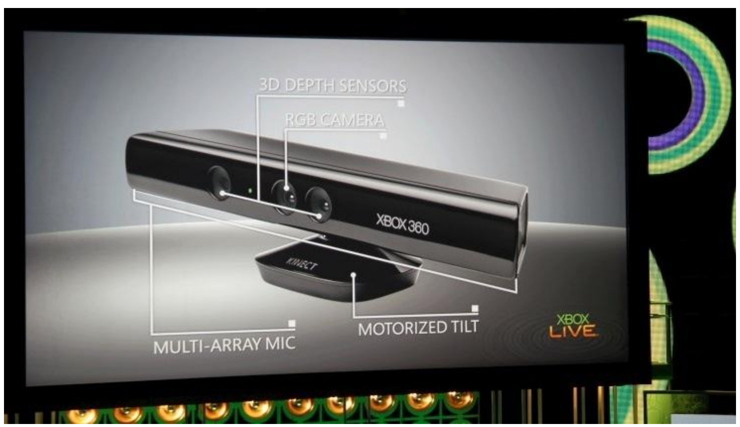
[Amit & Geman 97]
[Lepetit *et al.* 06]

Slide courtesy Jamie Shotton

Semantic Texton Forests for classification

- Huge commercial success for Randomized Decision Forests! Microsoft Xbox 360 gaming.

kinect sensor



Depth image

Body part recognition

Shotton, Fitzgibbon et. al, Real-Time Human Pose Recognition in Parts from a Single Depth Image, CVPR'11.
Shotton, Johnson & Cipolla, Semantic Texton Forests for Image Categorization and Segmentation, CVPR'08.

Interactive Video Segmentation

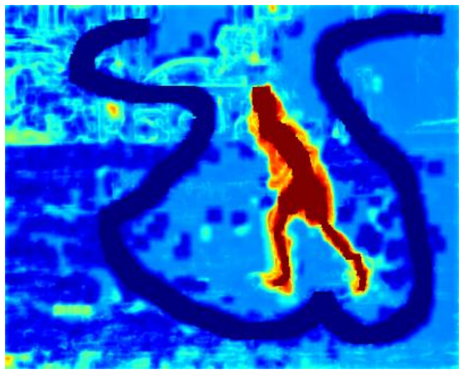
Video

User labels

Our framework

Video

Temporal label propagation

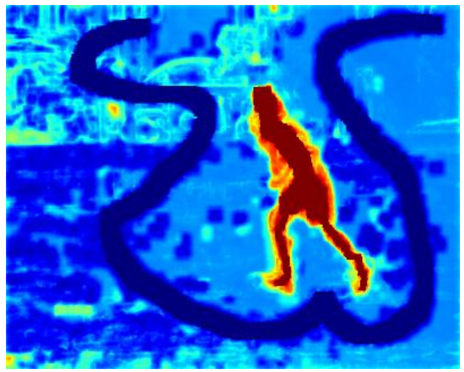


User labels

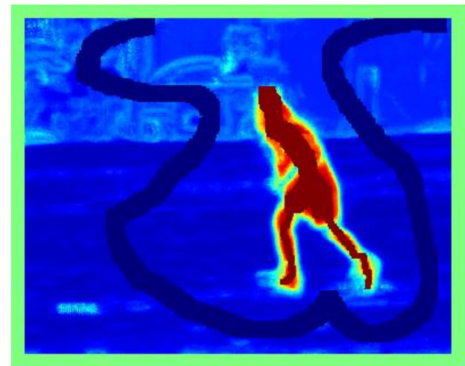
Our framework

Video

Temporal label propagation



Semi-supervised classifier learning

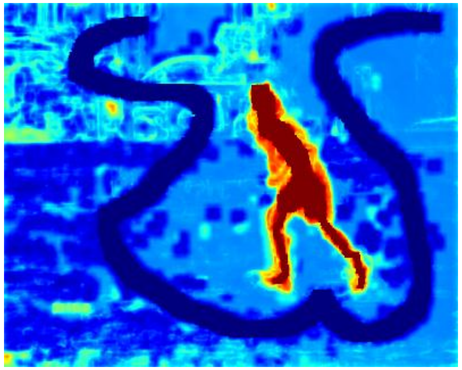


User labels

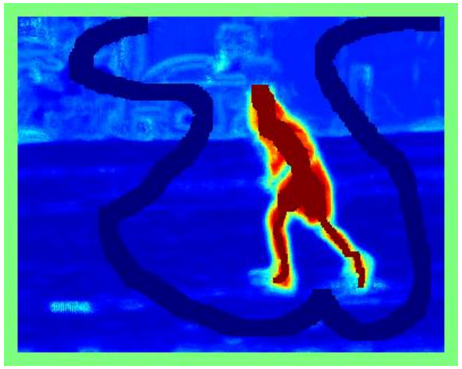
Our framework

Video

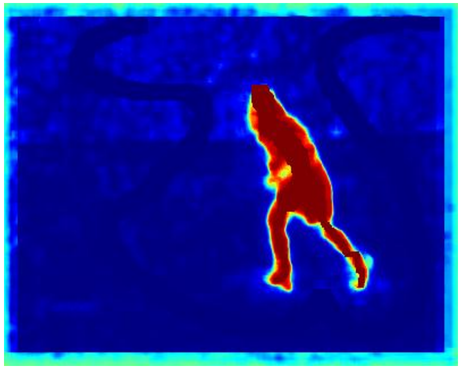
Temporal label propagation



Semi-supervised classifier learning

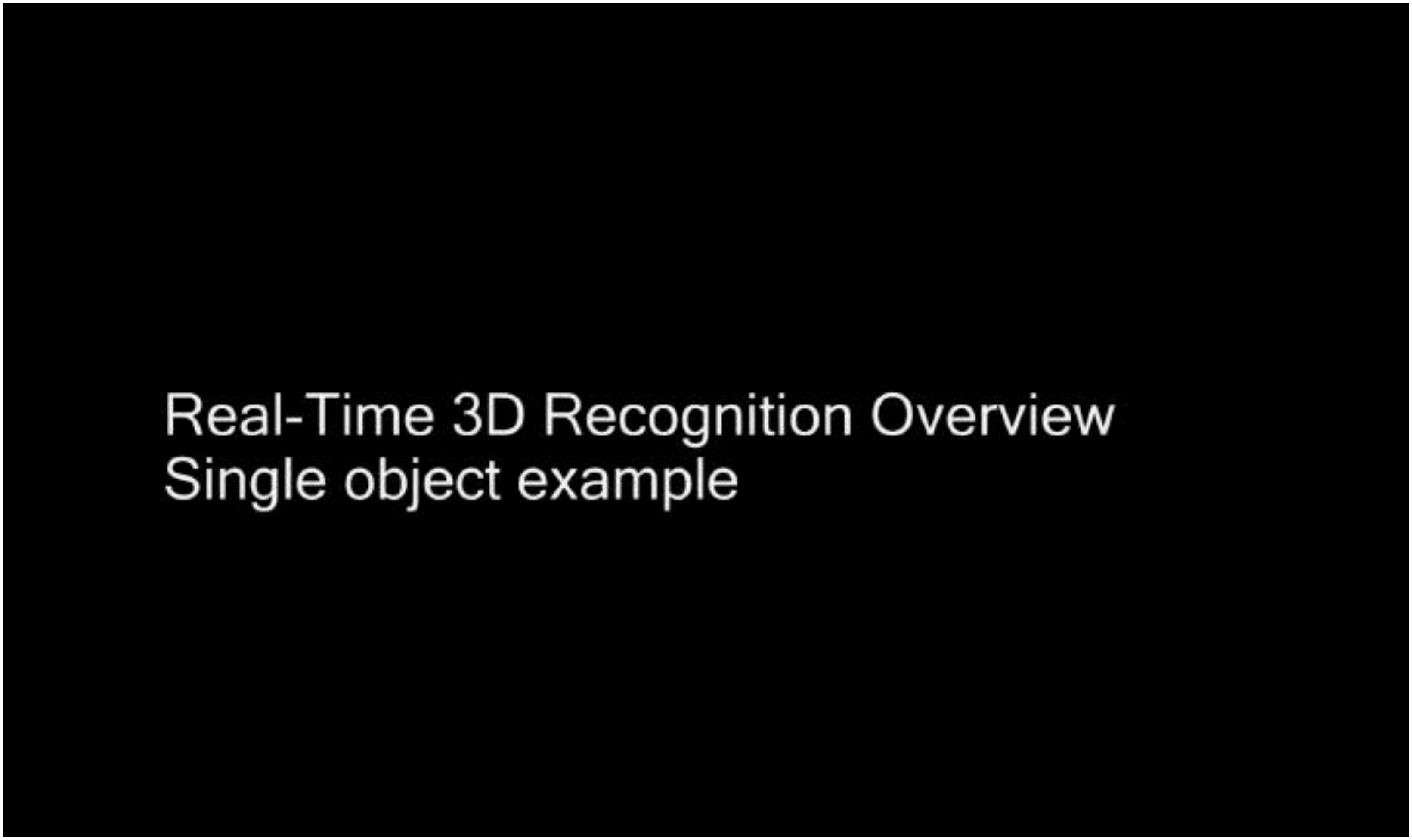


Bootstrapping the classifier



User labels

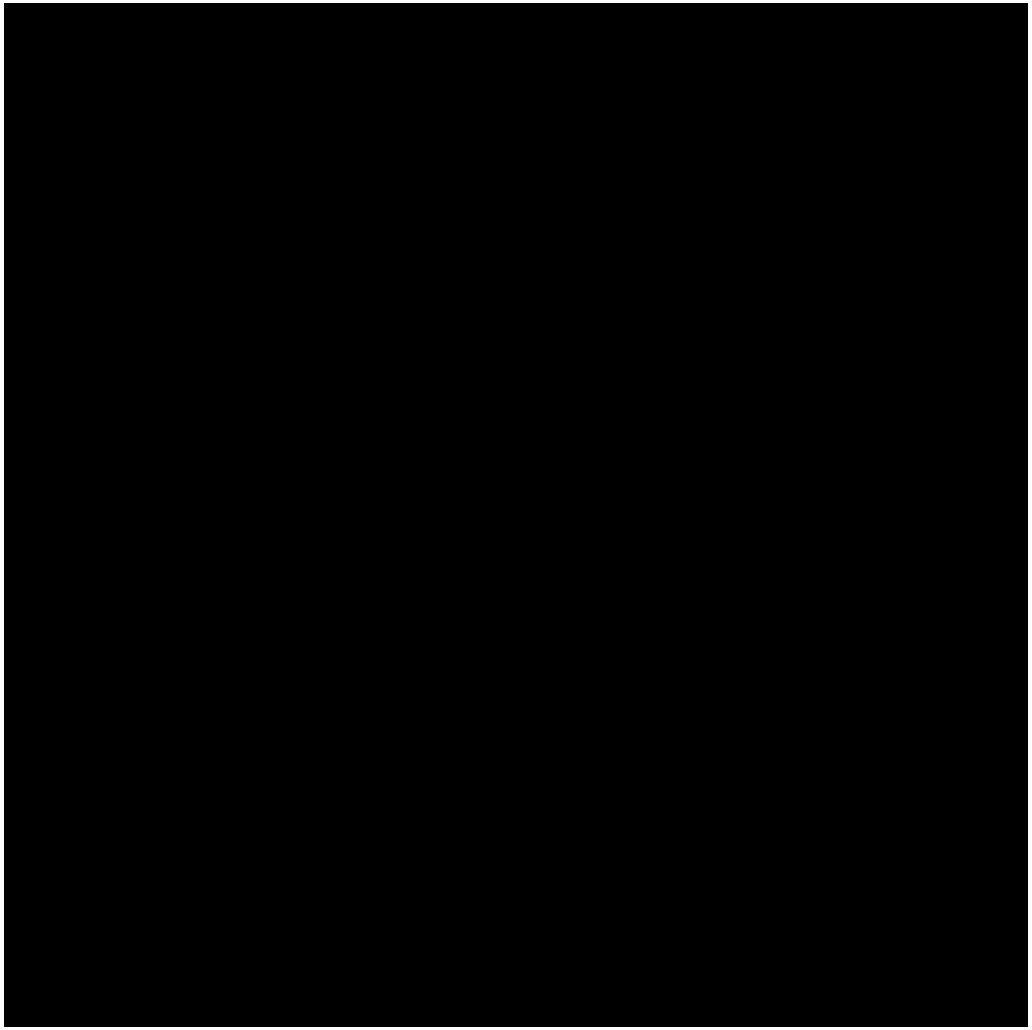
3D object recognition



Reconstruction?

Recovery of 3D shape from images

Reconstruction



Review

Recovery of accurate 3D shape from images

3D shape from photographs

Textured rigid object

Untextured rigid object

Untextured deformable object

3D shape from photographs

Multi-view
stereo

Multi-view
photometric stereo

Single view coloured
photometric stereo

Overview

1. Multi-view stereo
2. Multi-view photometric stereo
3. Single-view colour photometric stereo

UNIVERSITY OF CAMBRIDGE

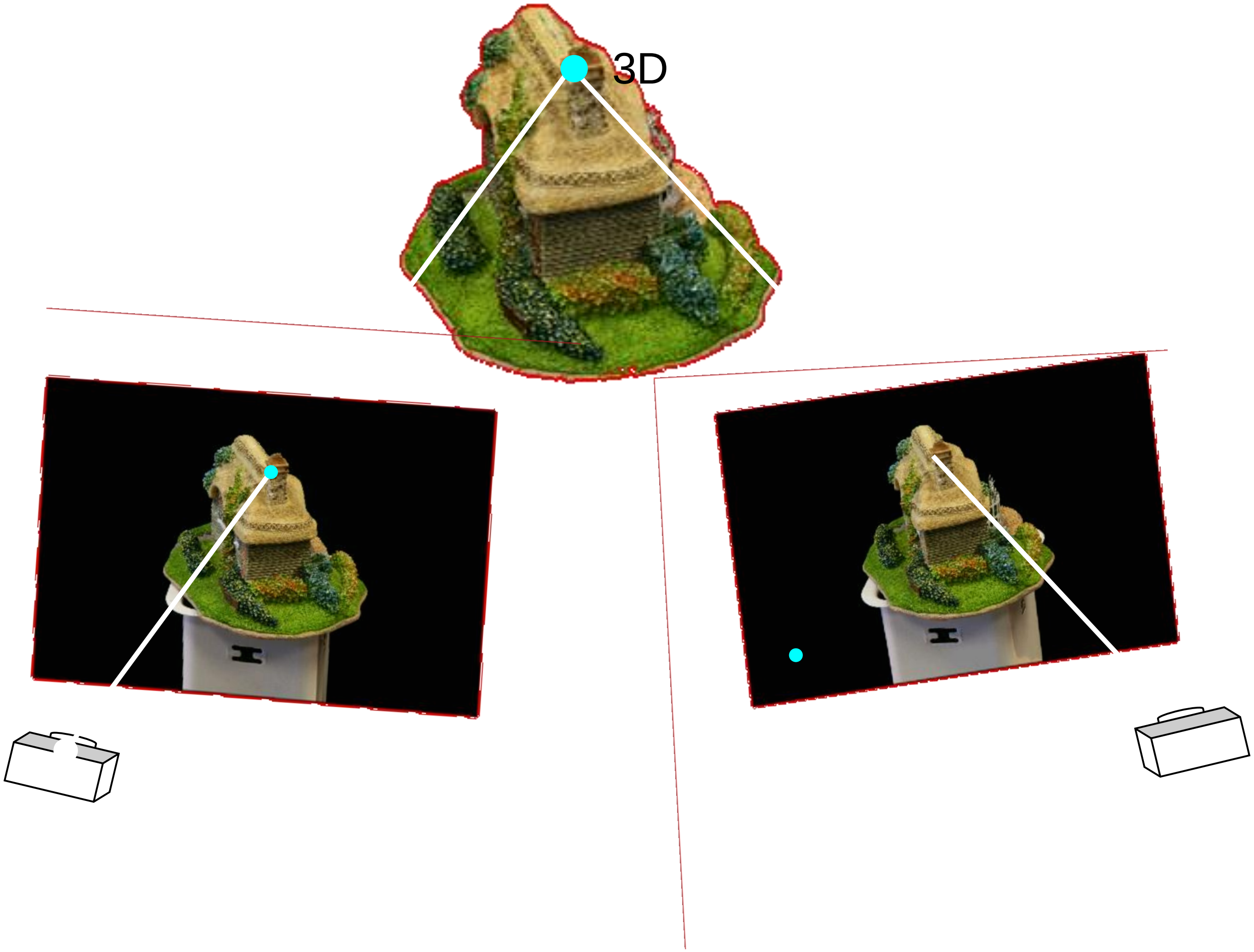
Multi-view stereo

Cipolla and Blake 1992
Cipolla and Giblin 1999
Mendonca, Wong and Cipolla 1999-2005
Vogiatzis, Hernandez and Cipolla 2006-2007
Campbell, Vogiatzis, Hernandez and Cipolla 2008-2011

1. Textured object

- Key assumptions for object surface
  - Smooth and rigid
  - Lambertian reflectance
  - Richly textured

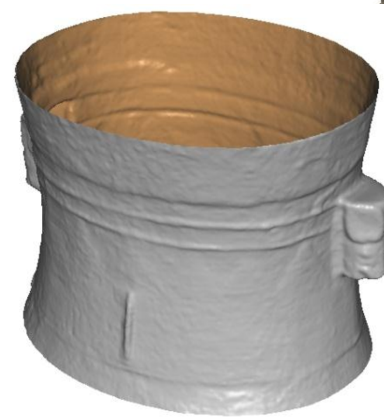
Stereo vision



3D shape of textured objects

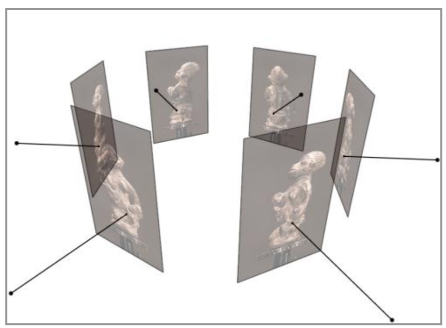
3D models – multiview stereo

Digital Pygmalion Project



Multi-view stereo pipeline

Image acquisition



Camera calibration

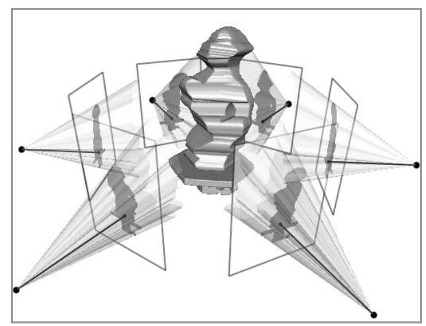
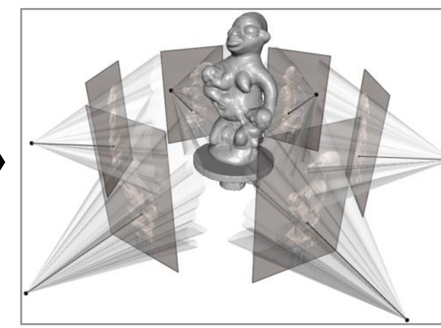


image segmentation



3D reconstruction

Image acquisition

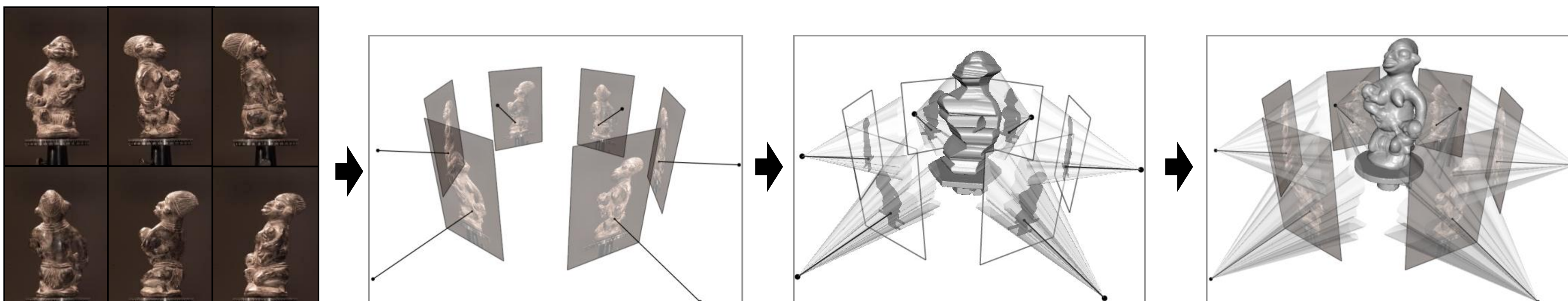


Image acquisition | Camera calibration | image segmentation | 3D reconstruction

Image acquisition

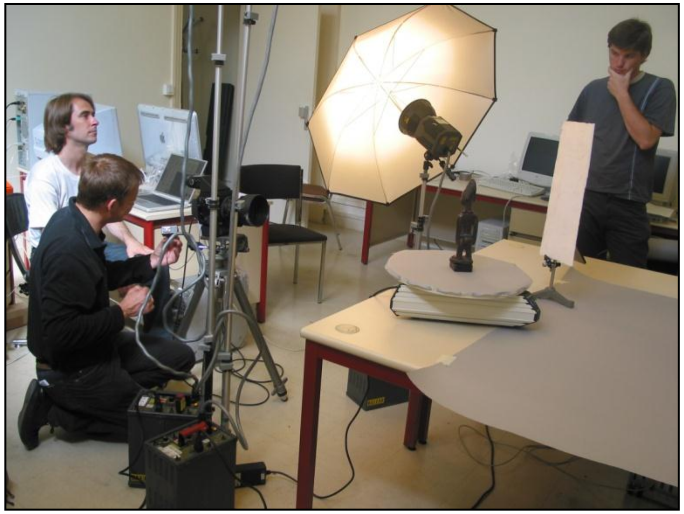
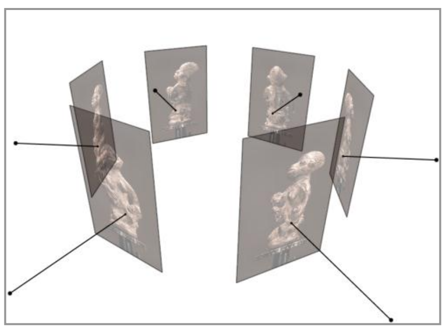


Image acquisition




Camera calibration

Image acquisition



Camera calibration

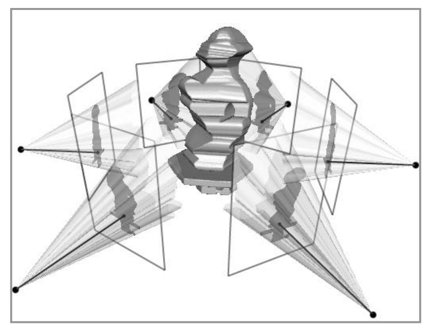
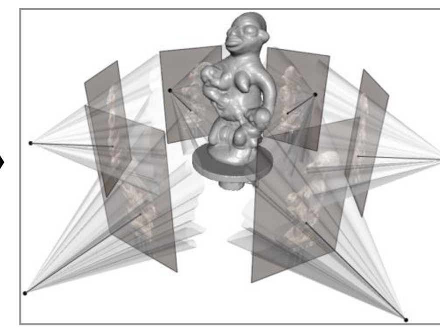
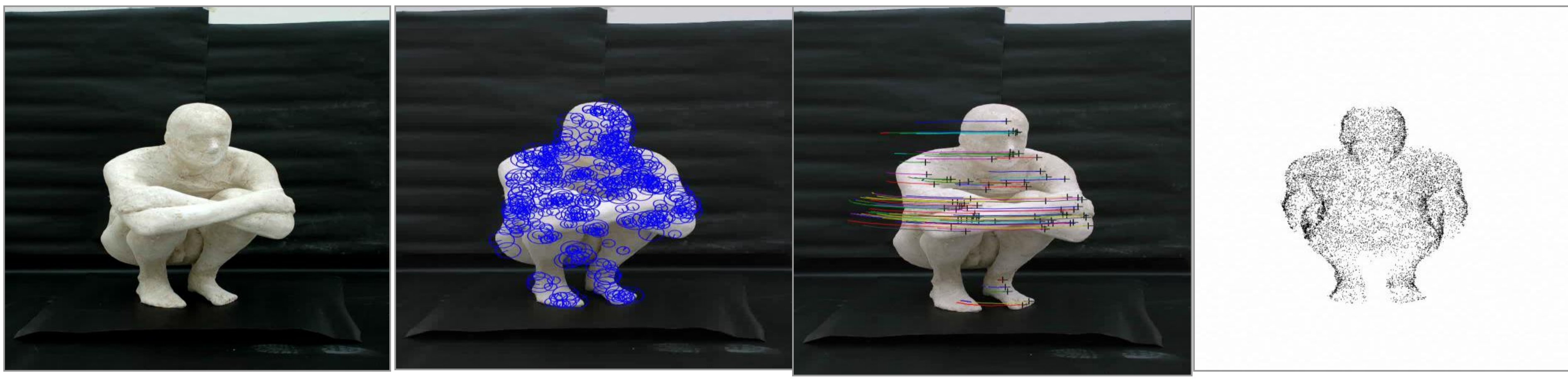


image segmentation



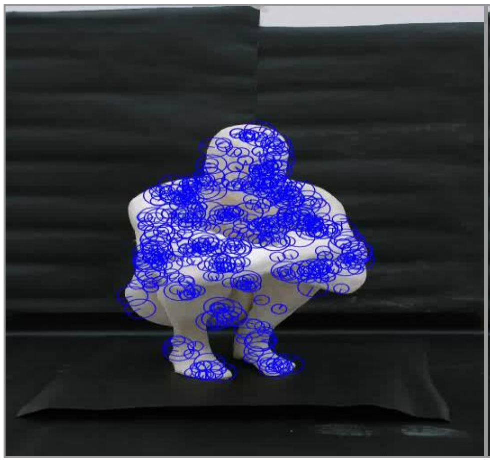
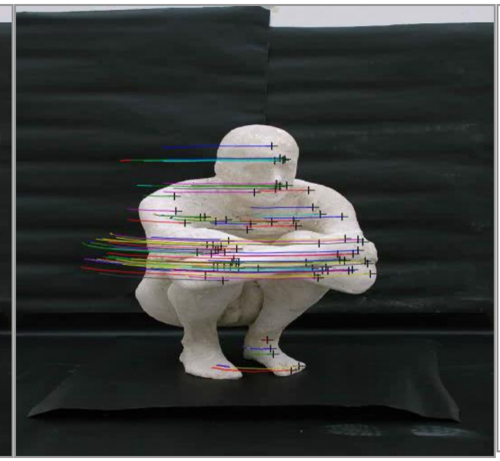
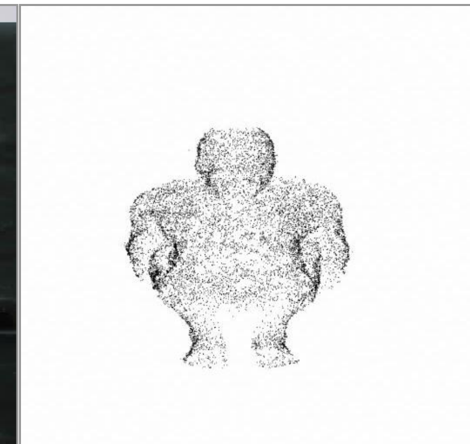
3D reconstruction

Structure from motion



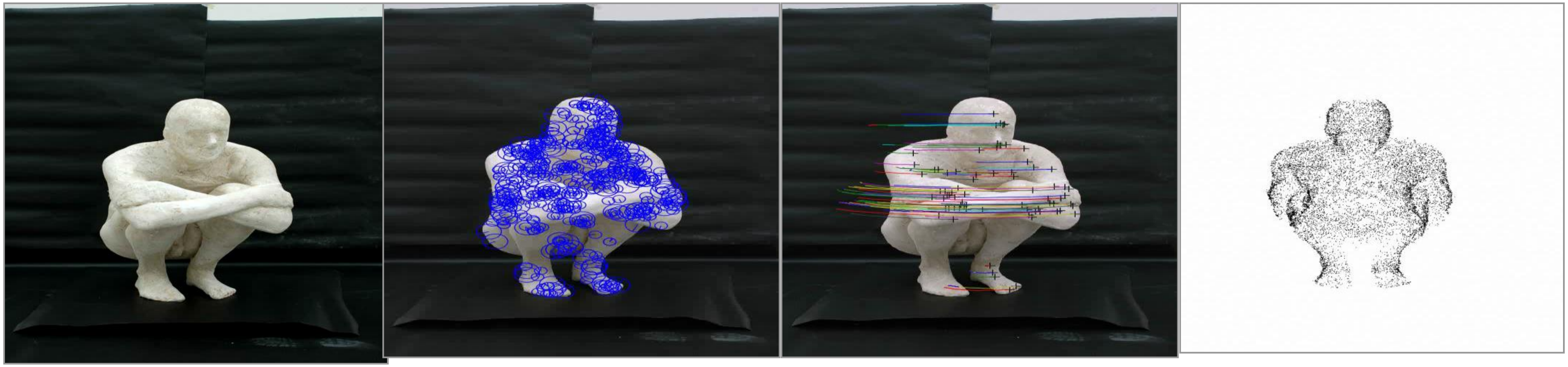
Input sequence 2D features 2D track 3D points

Structure from motion

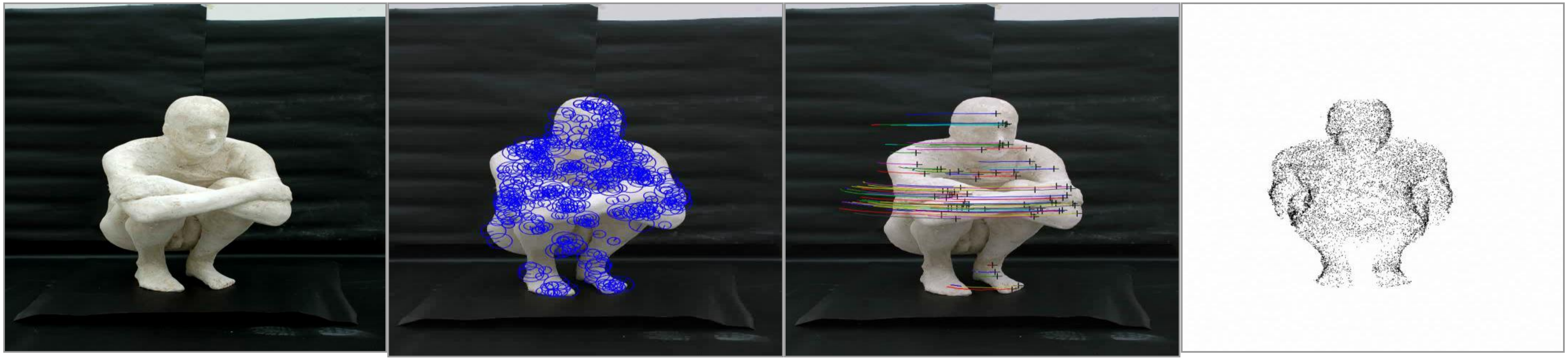
Input sequence 2D features 2D track 3D points

Structure from motion



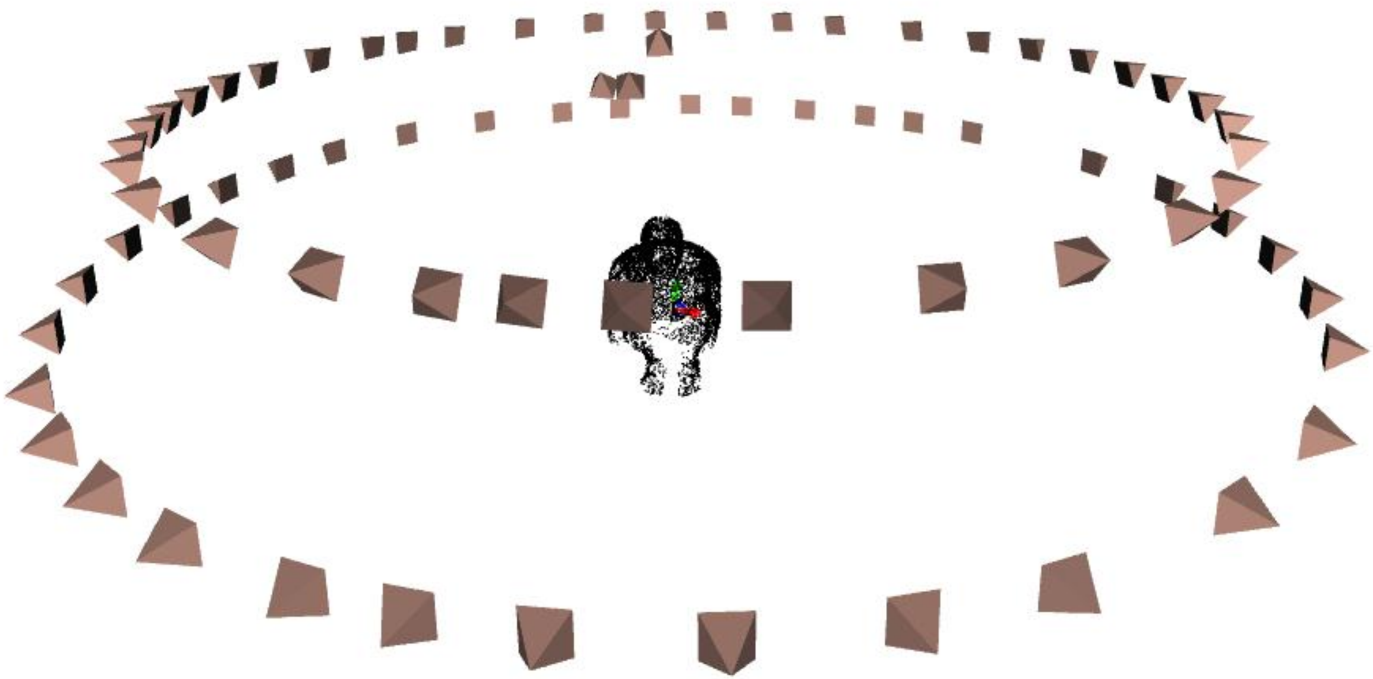
Input sequence 2D features 2D track 3D points

Structure from motion

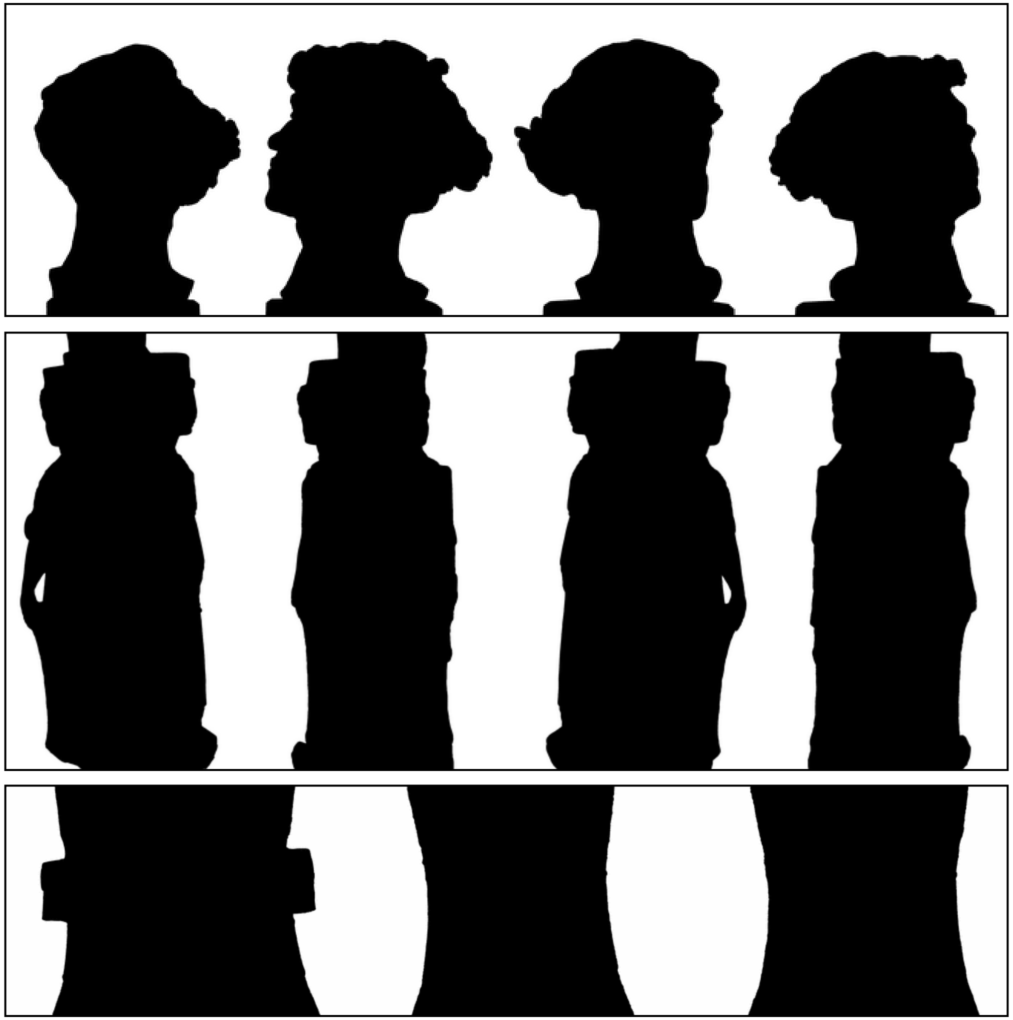


Input sequence 2D features 2D track **3D points**

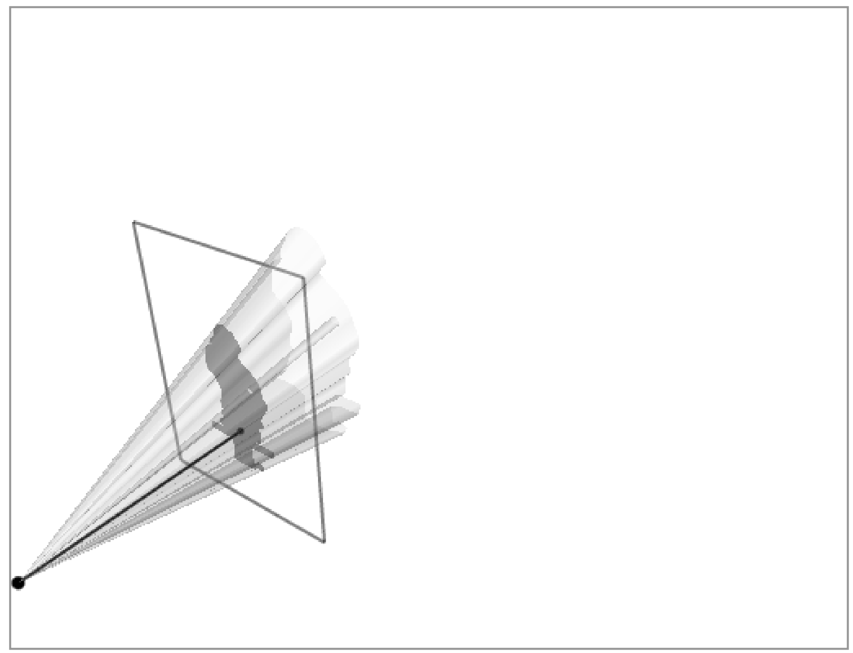
Motion estimation result



Silhouette-based calibration

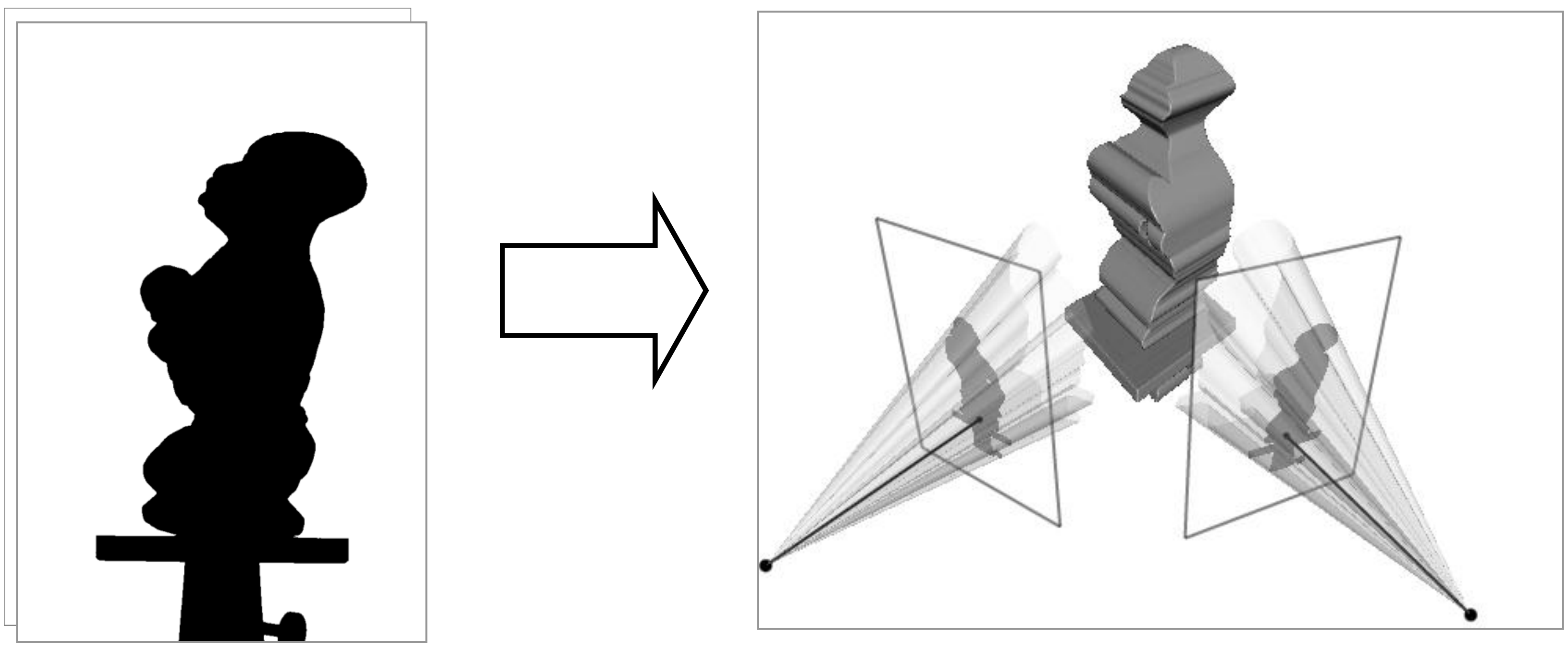


Visual hull concept



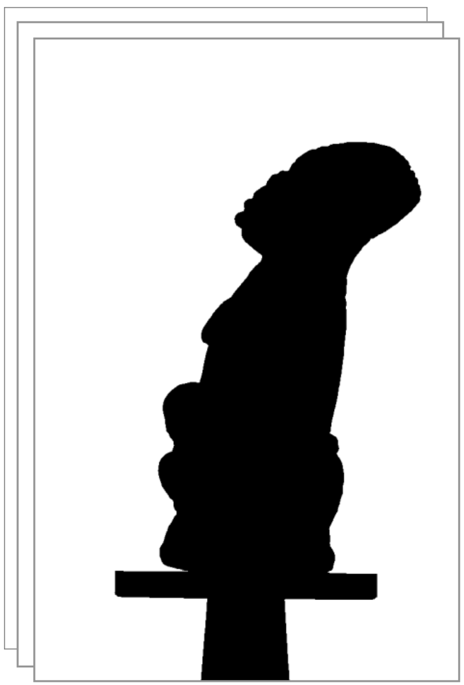
3D silhouette intersection

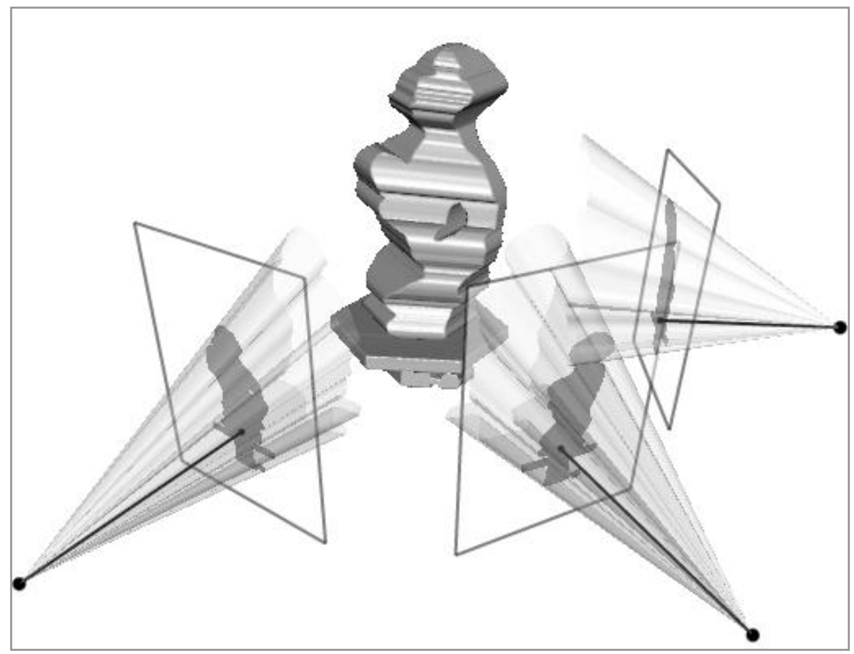
Visual hull concept



3D silhouette intersection

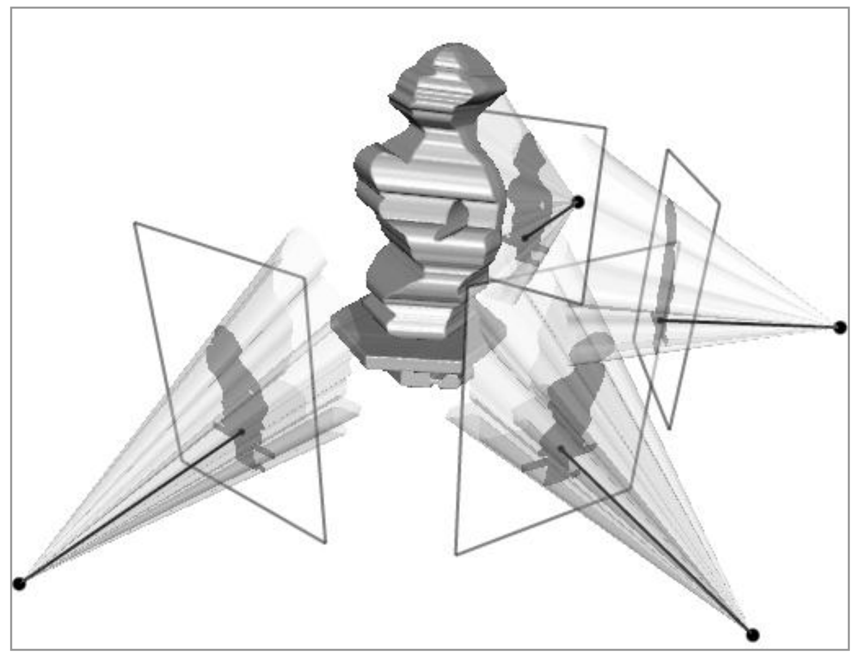
Visual hull concept





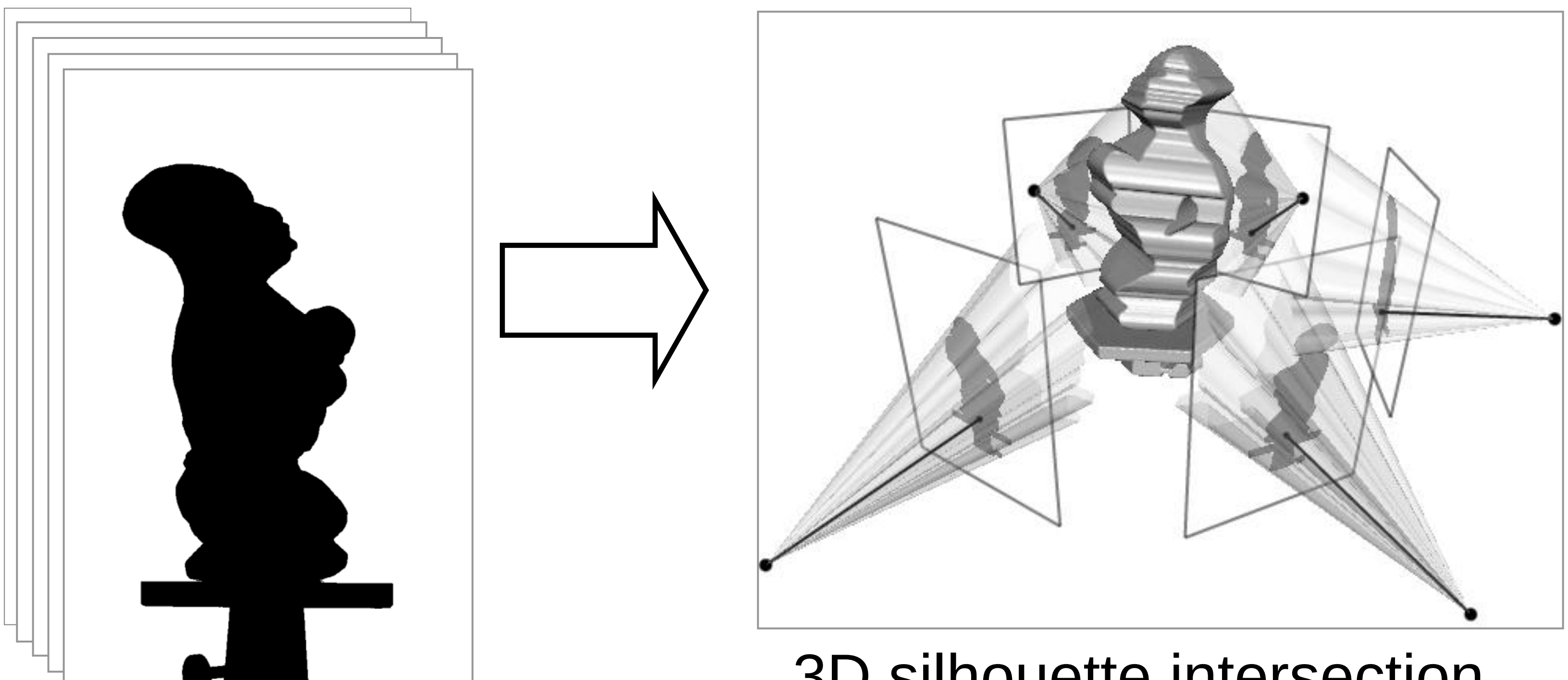
3D silhouette intersection

Visual hull concept



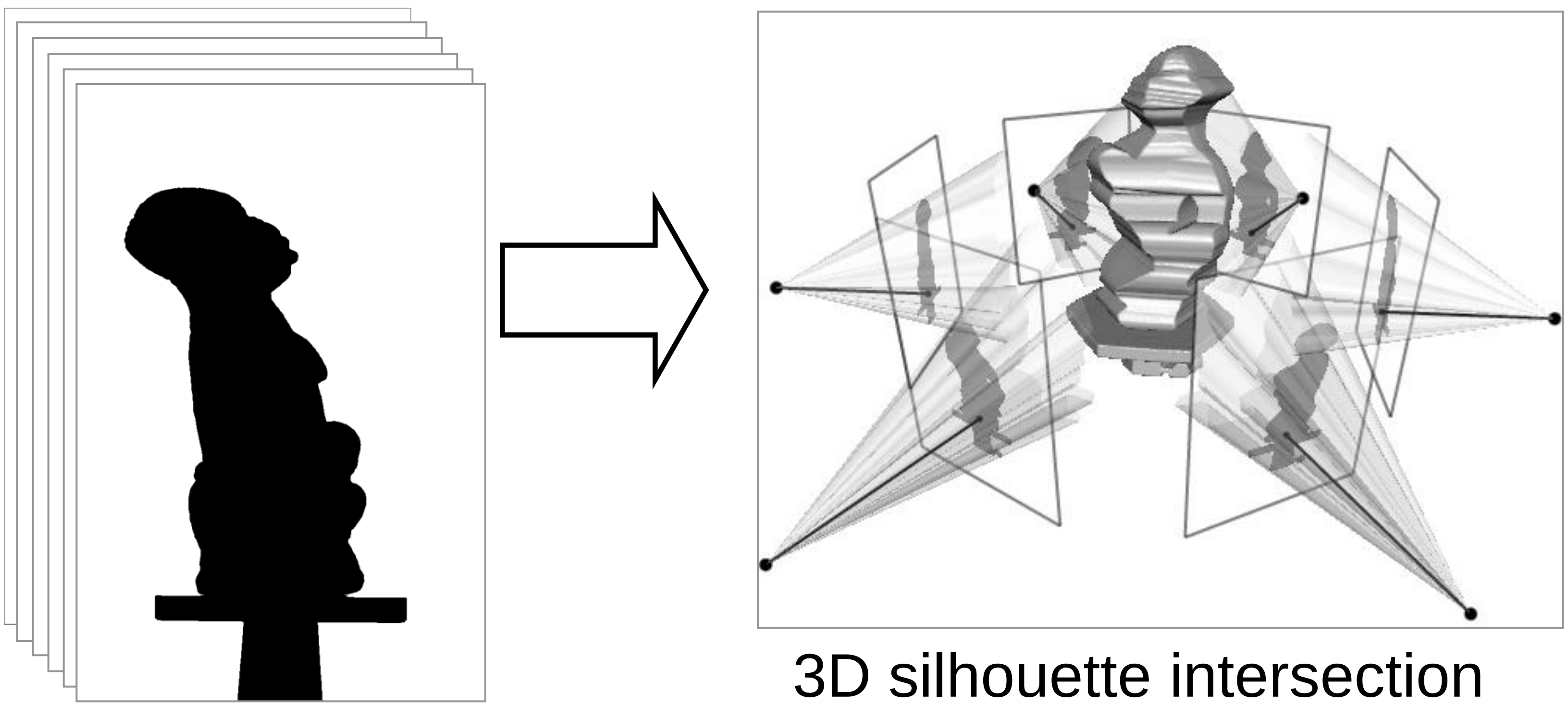
3D silhouette intersection

Visual hull concept



3D silhouette intersection

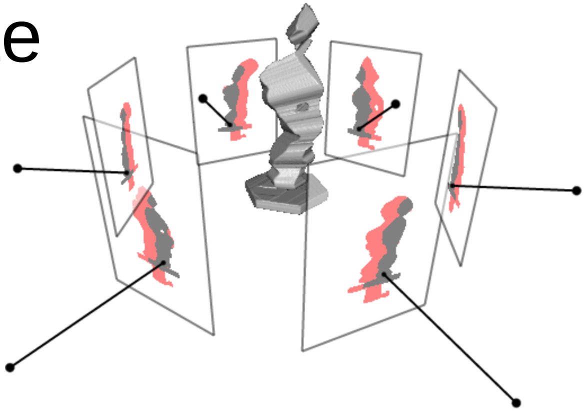
Visual hull concept



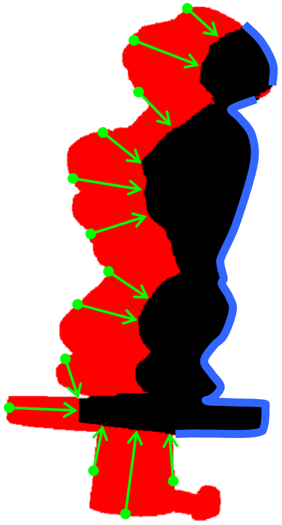
3D silhouette intersection

Silhouette consistency

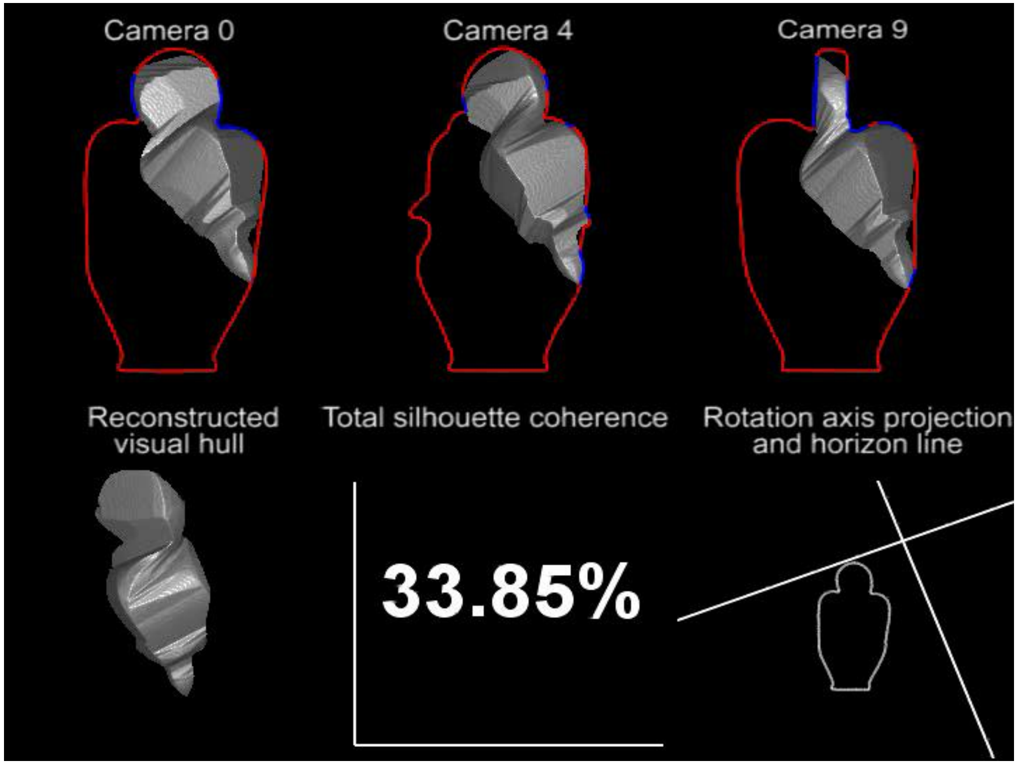
- Compares each original silhouette with the visual hull outline



- How to measure silhouette consistency?



Silhouette consistency



Silhouette consistency

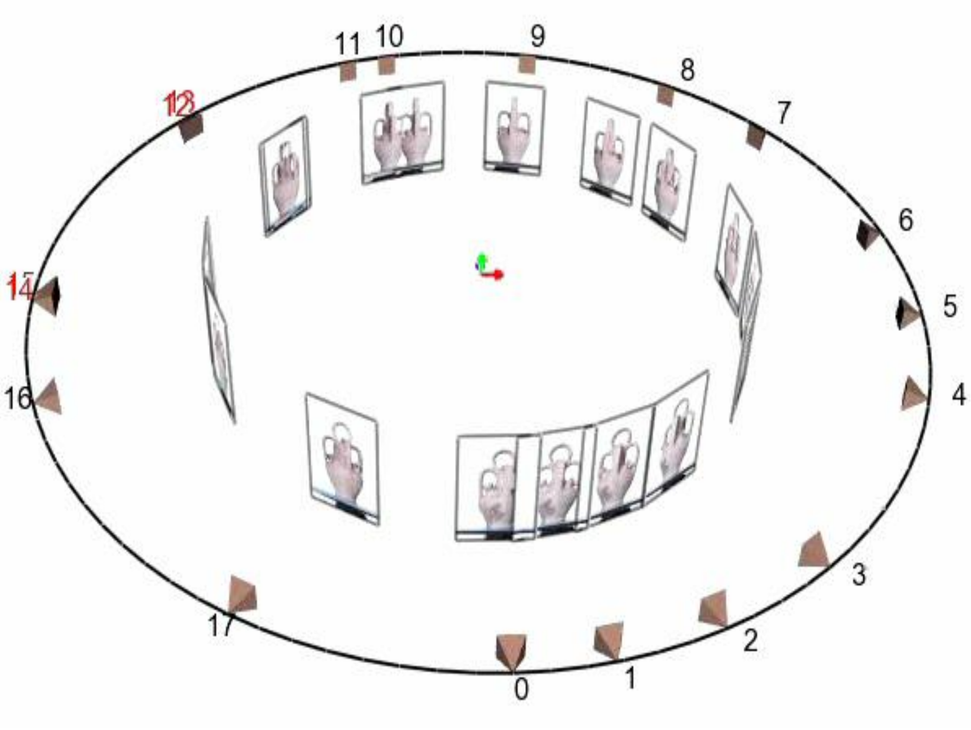


Image segmentation

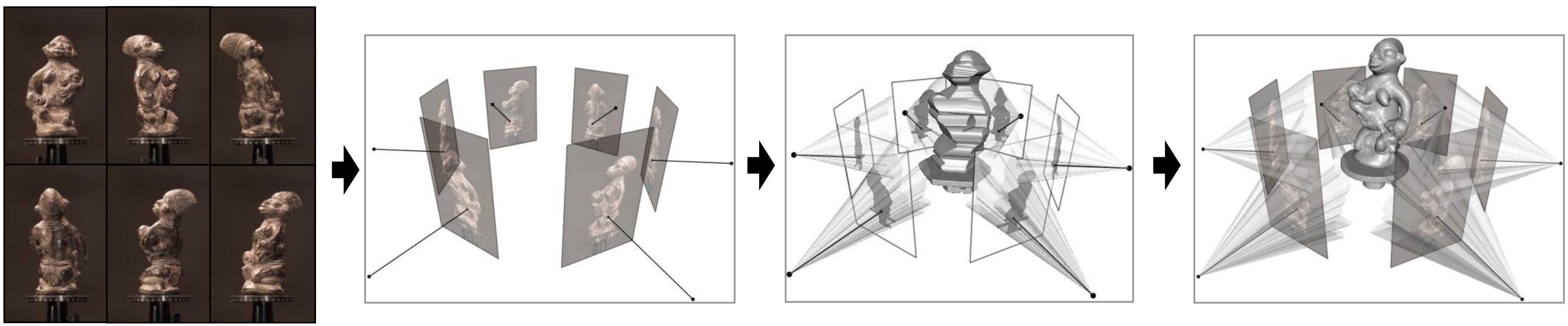
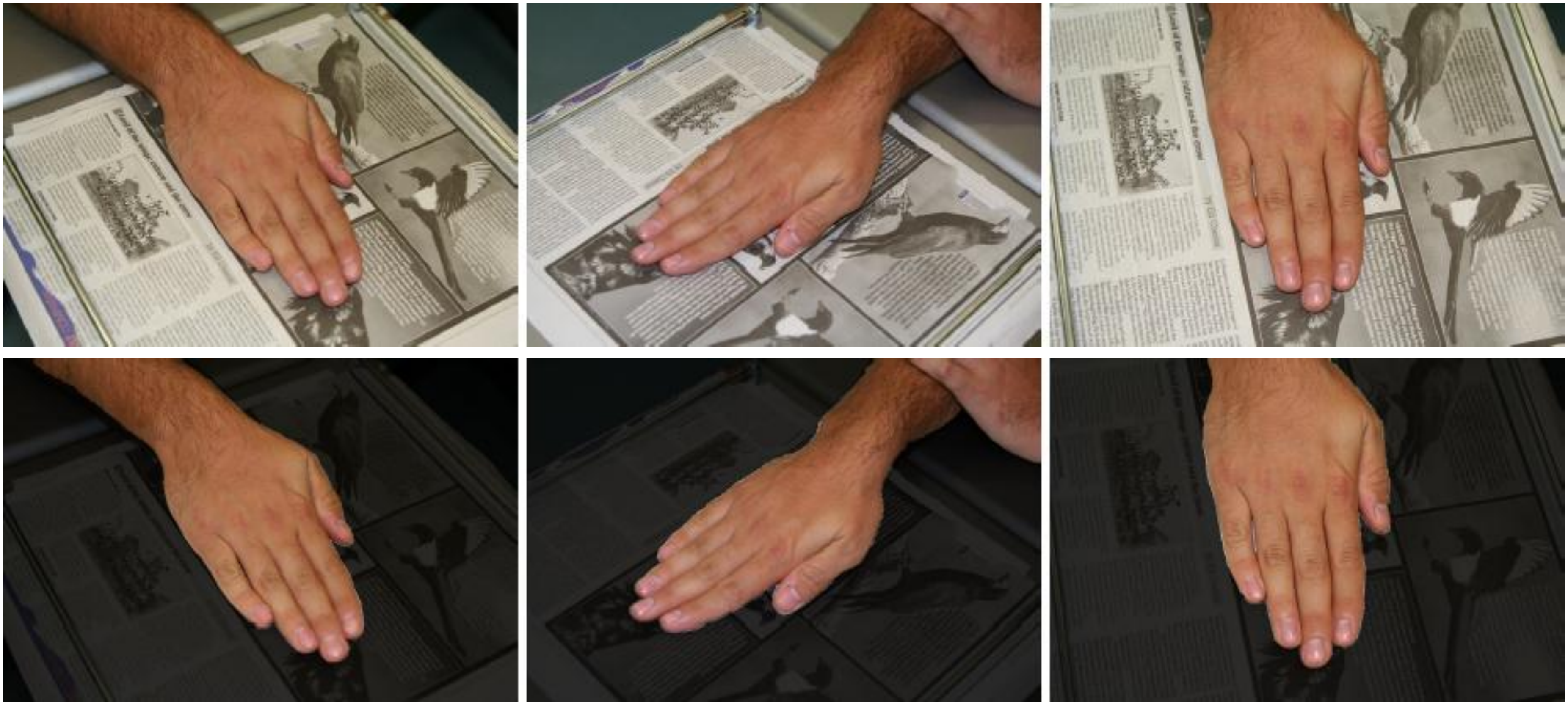


Image acquisition | Camera calibration | Image segmentation | 3D reconstruction

Automatic segmentation



Hand results

Hand results

3D Object reconstruction

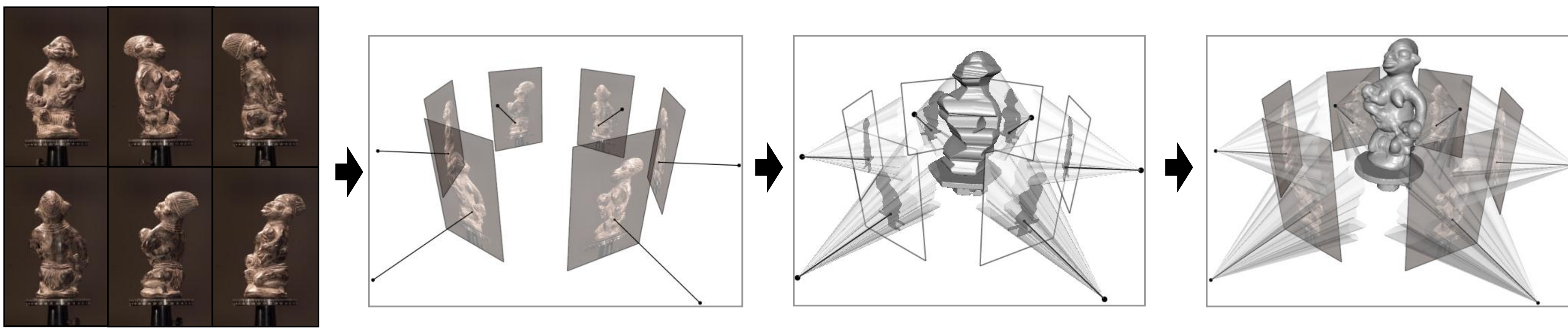


Image acquisition → Camera calibration → image segmentation → 3D reconstruction

Finding the surface

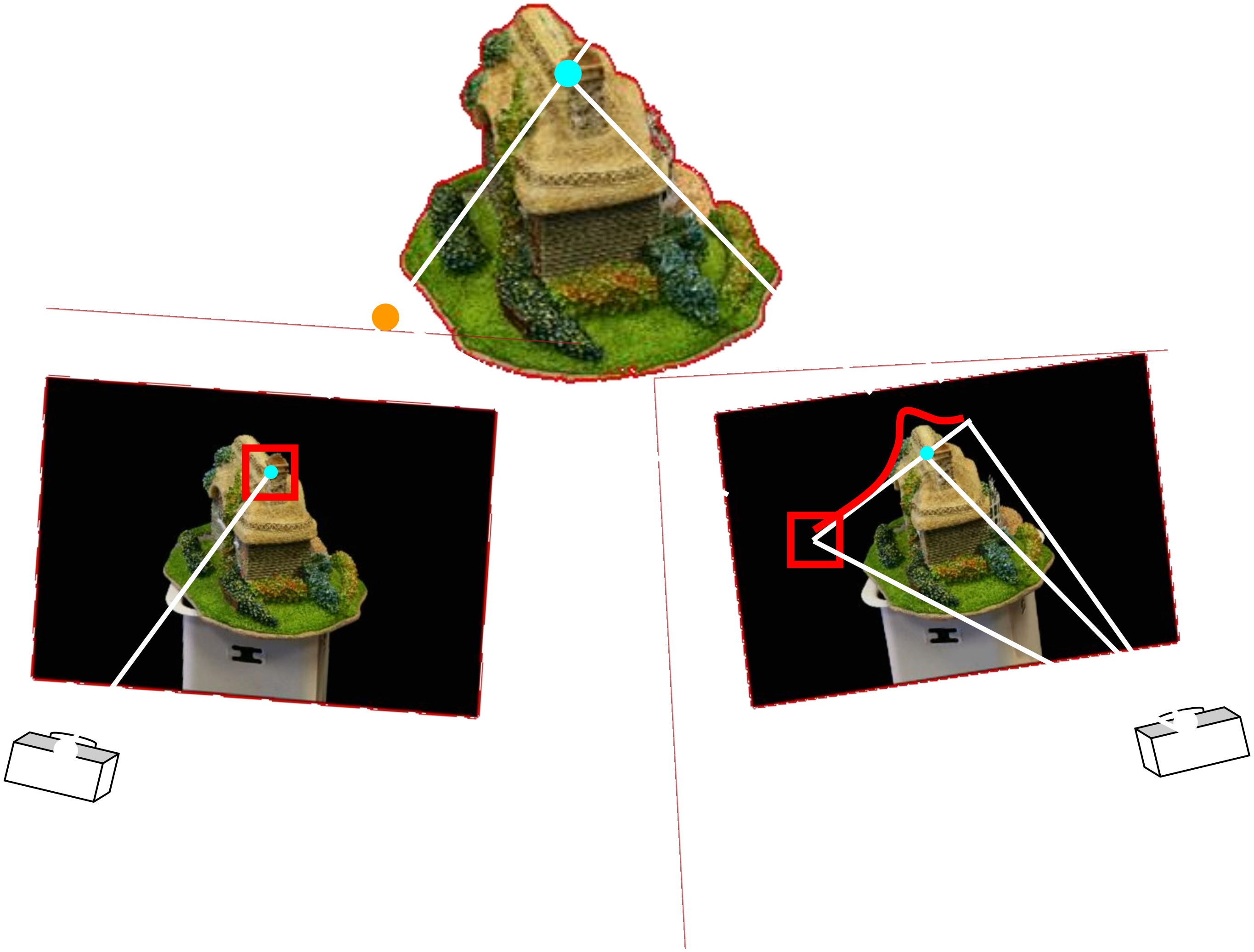


Photo-consistency

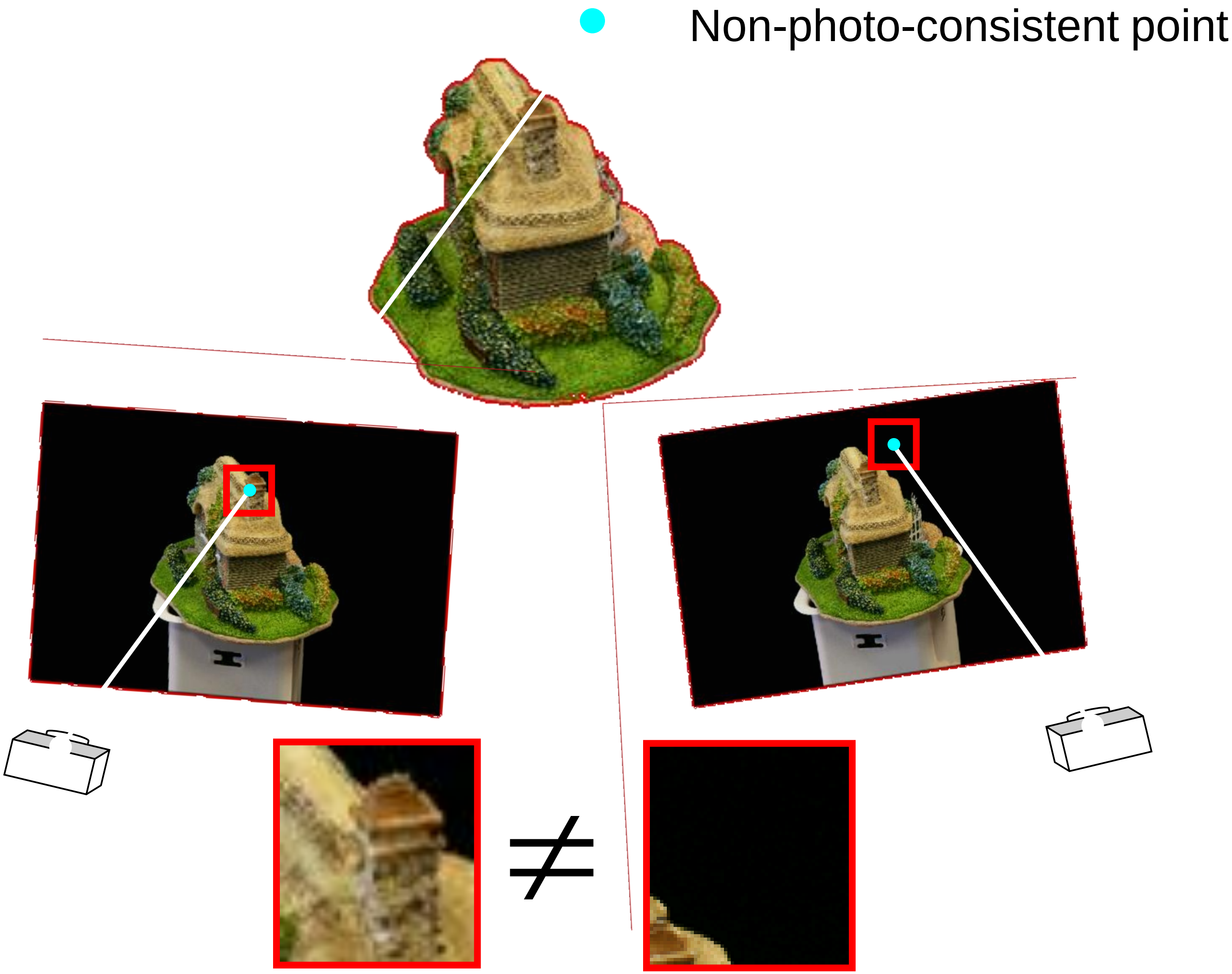
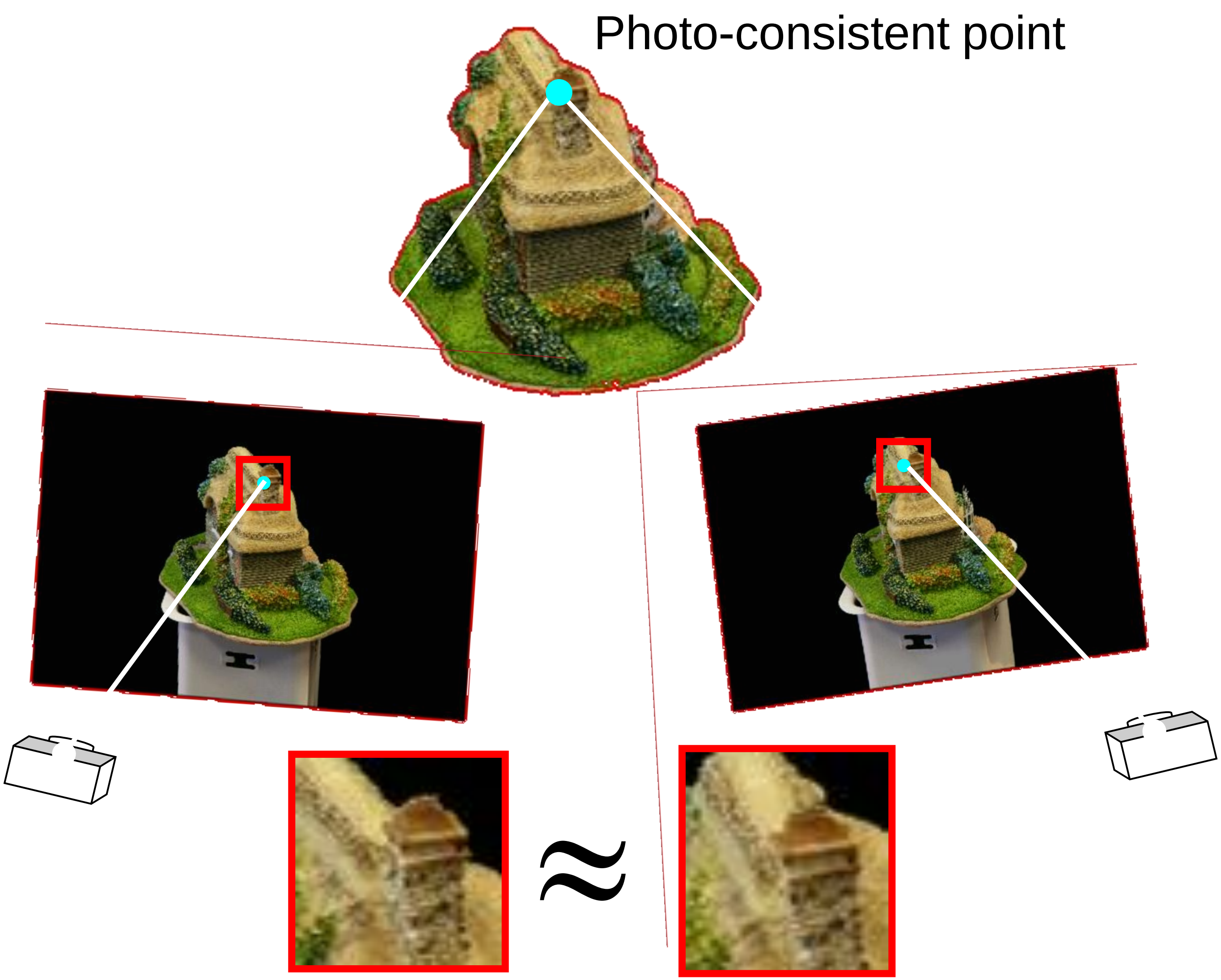
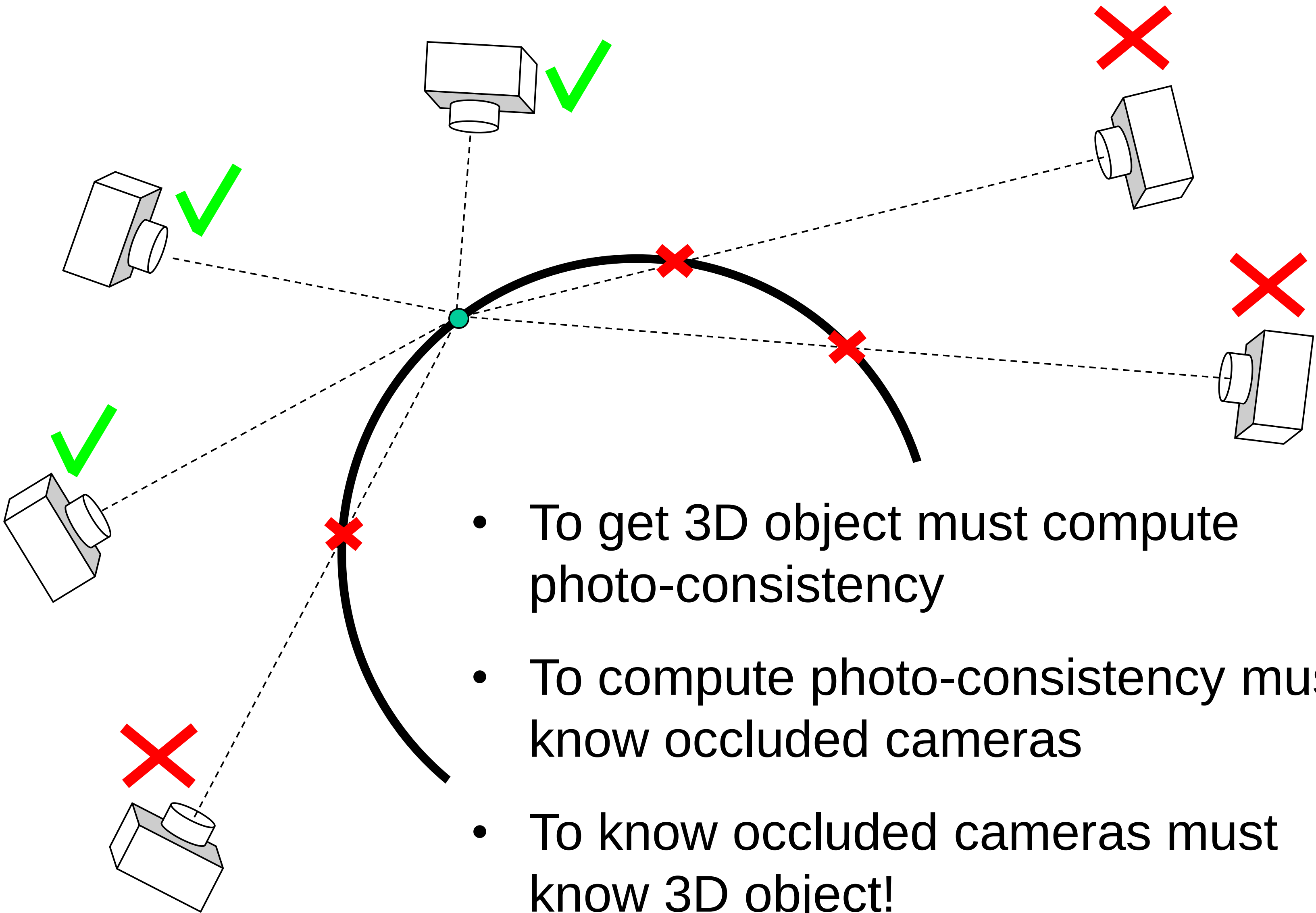


Photo-consistency



The occlusion problem

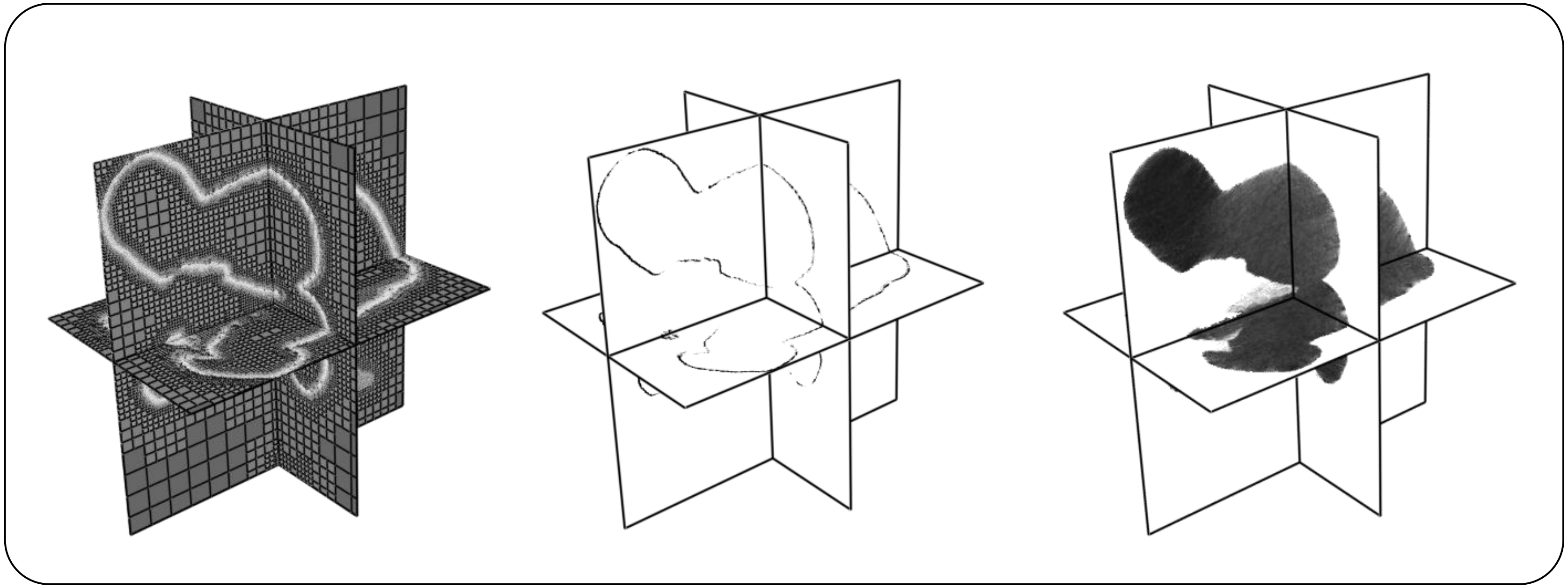


- To get 3D object must compute photo-consistency
- To compute photo-consistency must know occluded cameras
- To know occluded cameras must know 3D object!

Multi-view stereo algorithm

- Make photo-consistency robust to occlusion
- Casts the problem as discrete Markov Random Field (MRF) optimisation, obtaining global solution
- Use a triangle mesh as final representation

3D MRF for 3D modelling



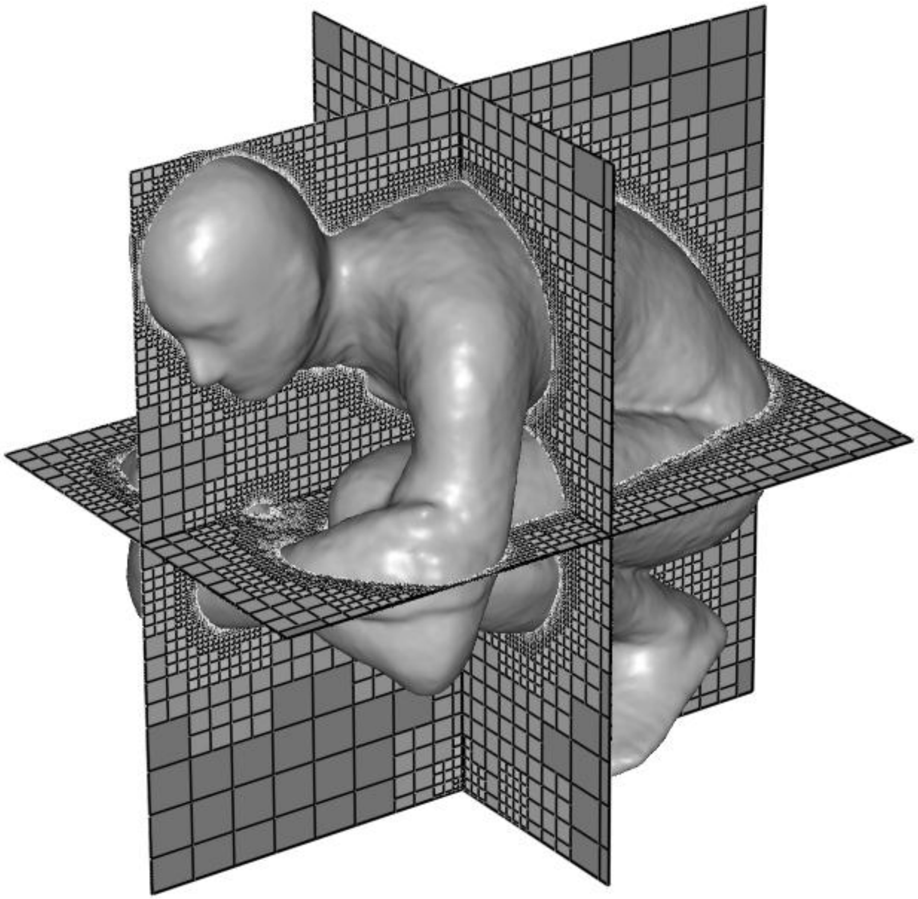
Multi-resolution grid

Edge cost

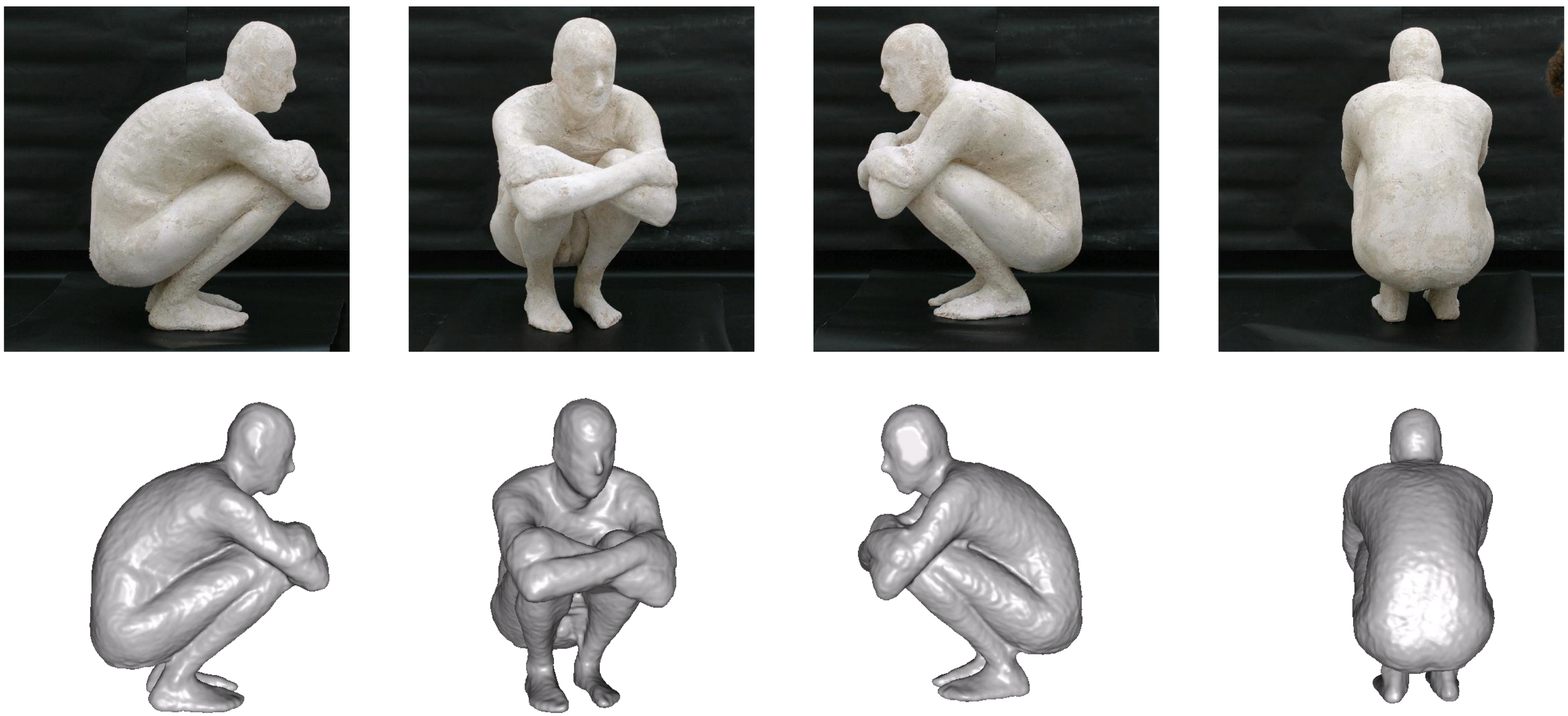
Foreground/ background cost

3D MRF for 3D modelling

UNIVERSITY OF CAMBRIDGE



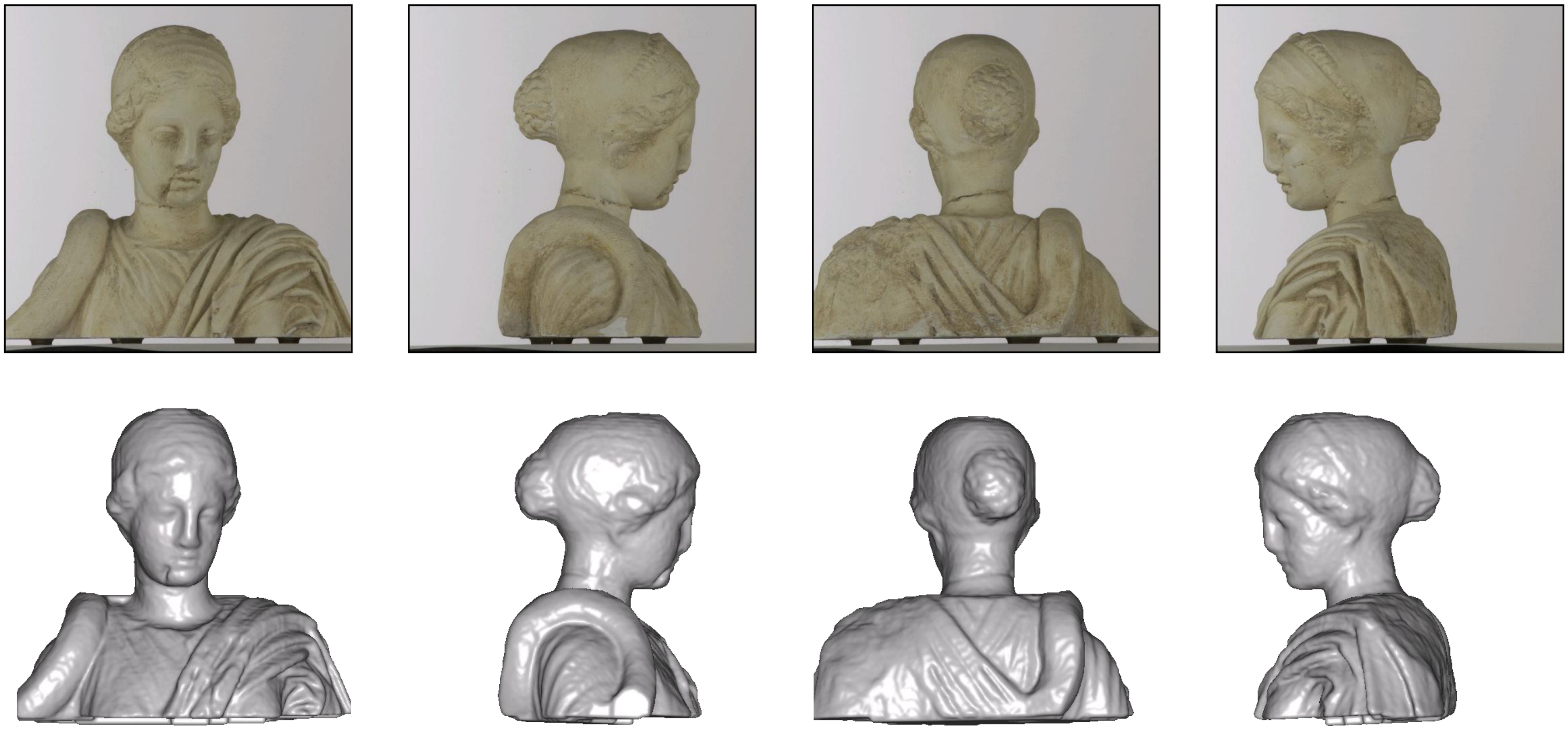
3D Models



Final installation

Results

Results



Real-time depth

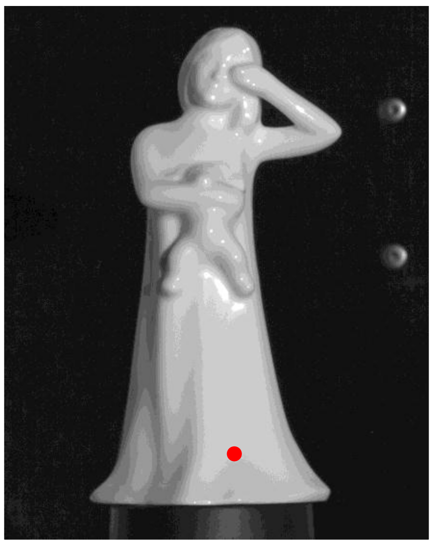
UNIVERSITY OF CAMBRIDGE

Multiview photometric stereo

Vogiatzis, Hernandez and Cipolla 2006 and 2008

2. Untextured objects

- Almost impossible to establish correspondence



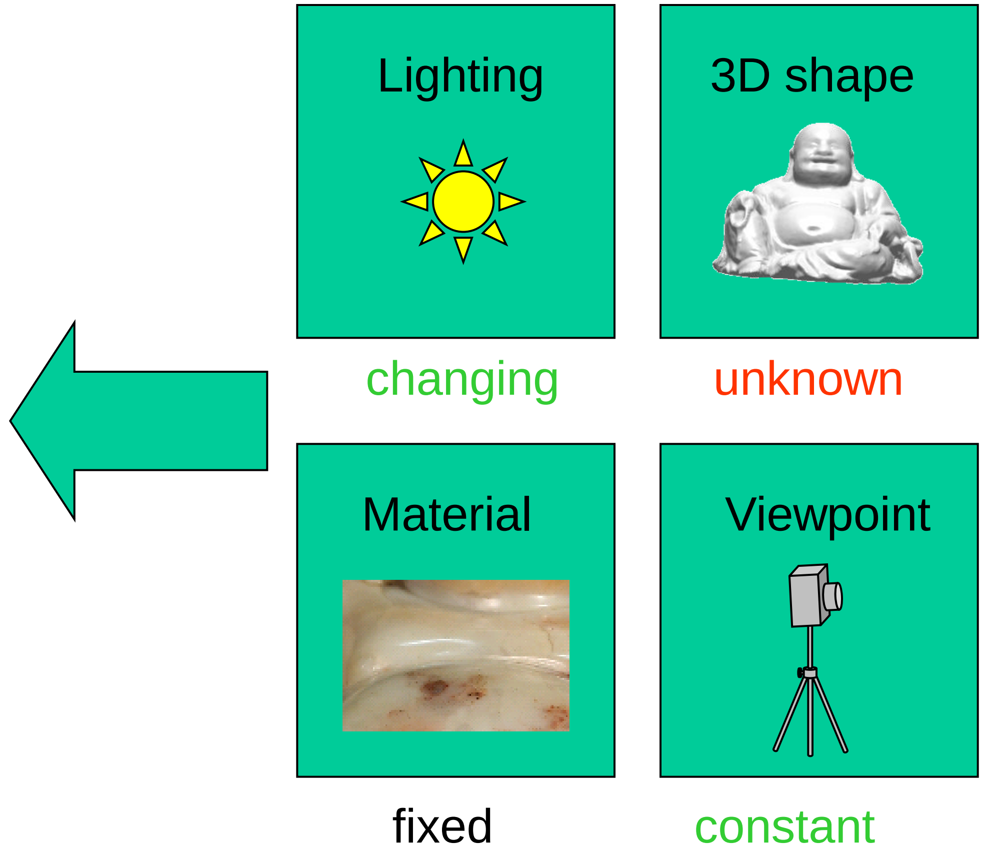
Use shading cue

Photometric stereo

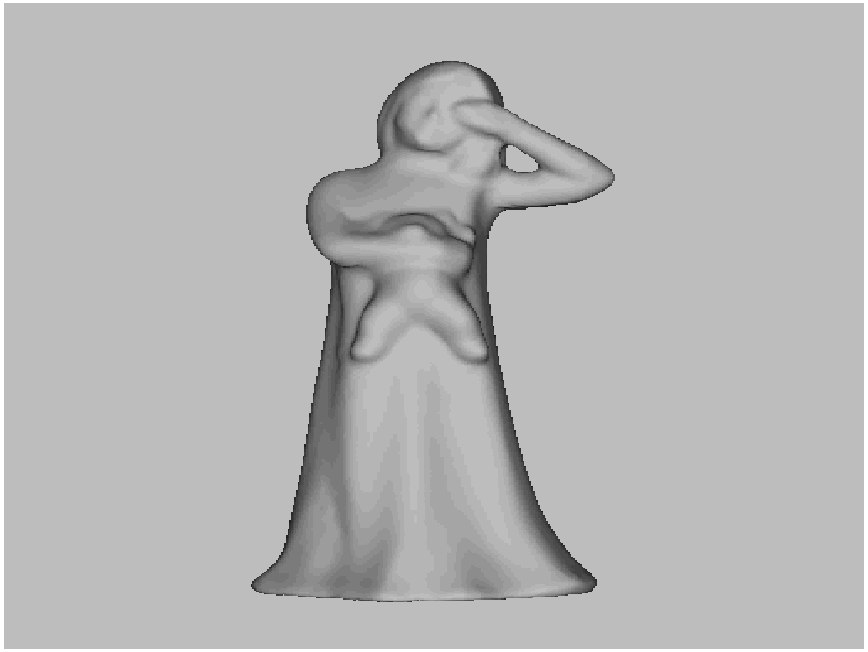
- Assumptions:
  - Single, distant light-source
  - No texture, single colour
  - Lambertian with few highlights

Changing lighting uncovers geometric detail

Photometric stereo

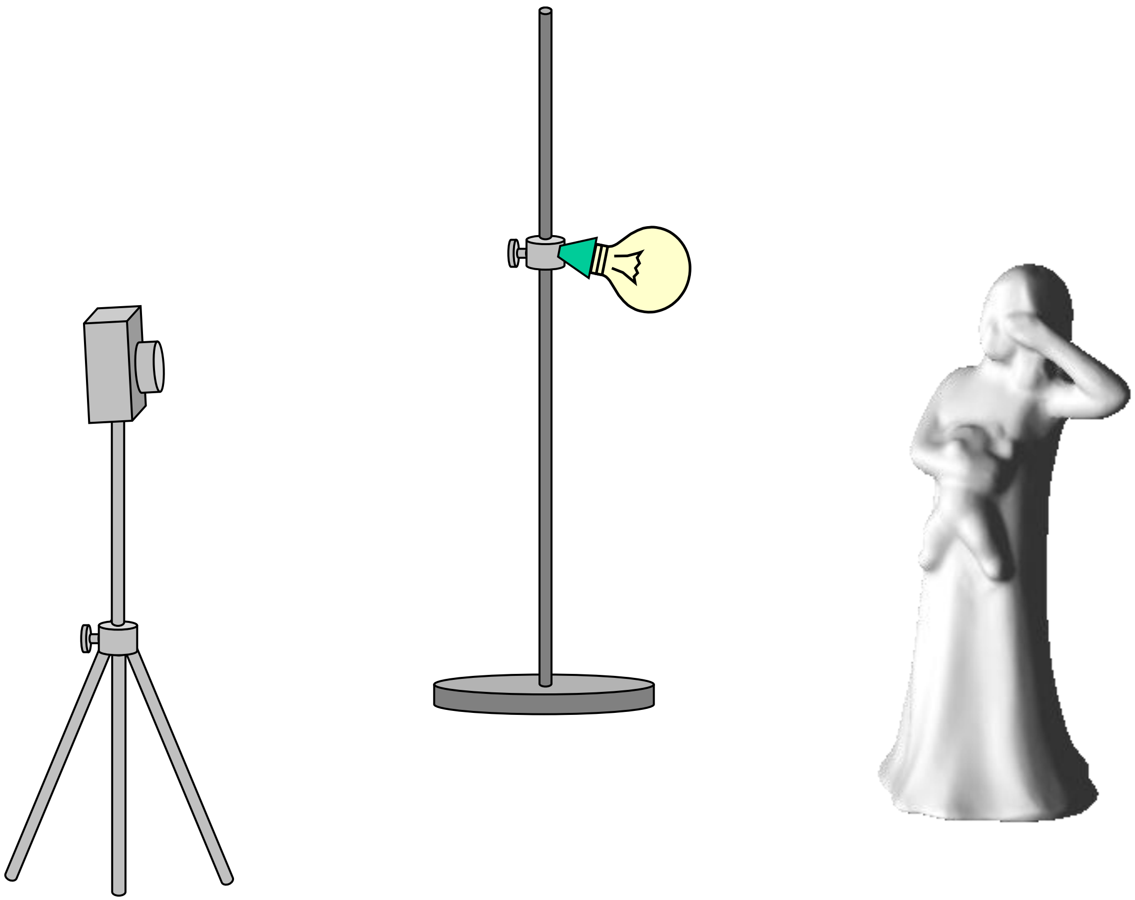


Shading cue

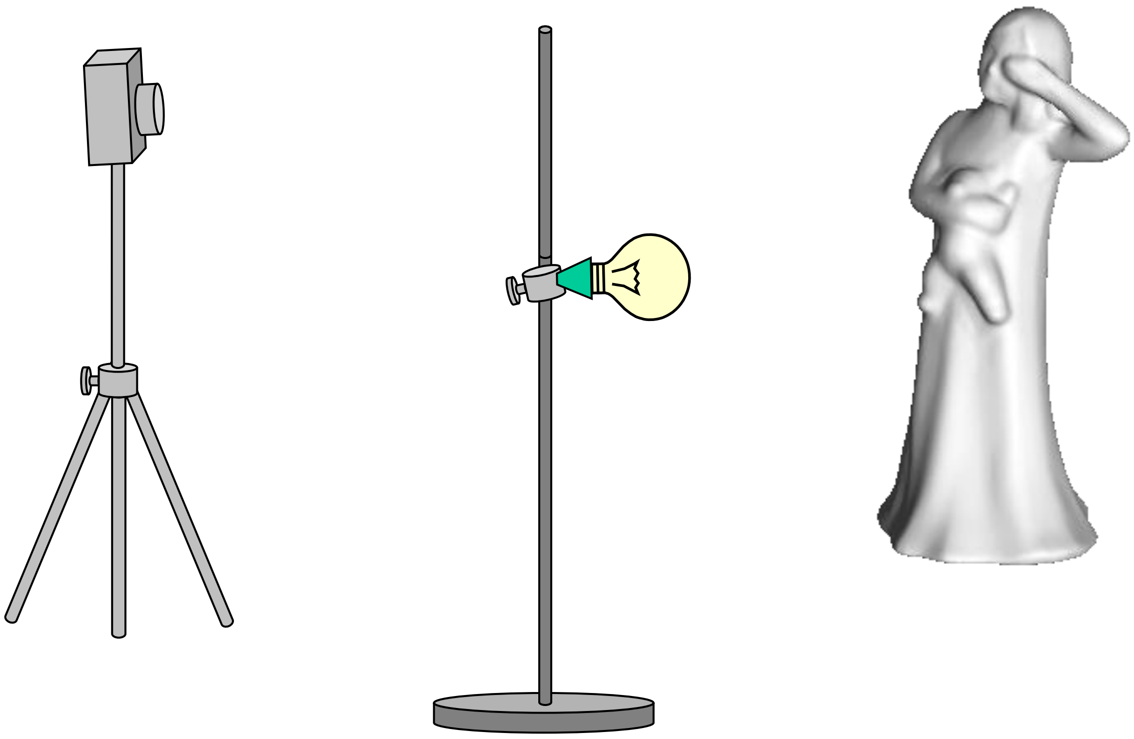


Surface reflectance is following Lambert's cosine law
Assume a distant point source

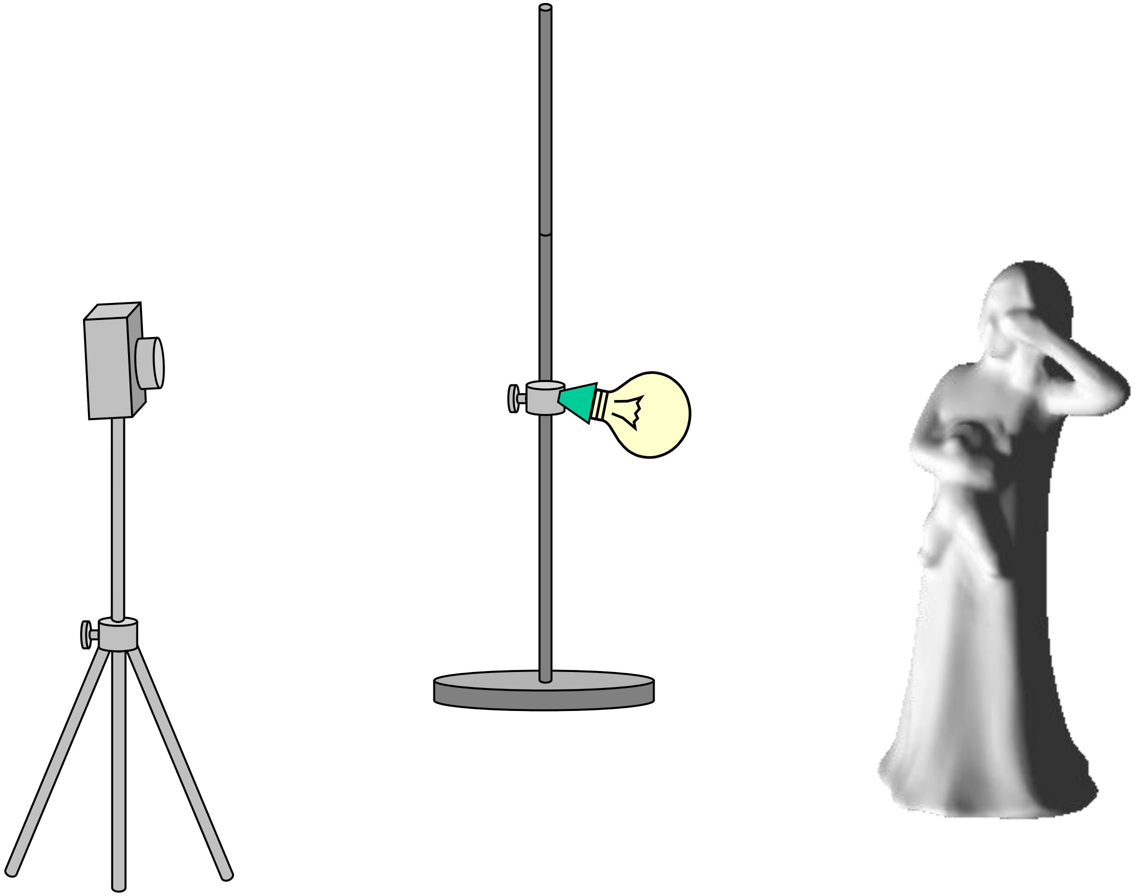
Classic photometric stereo



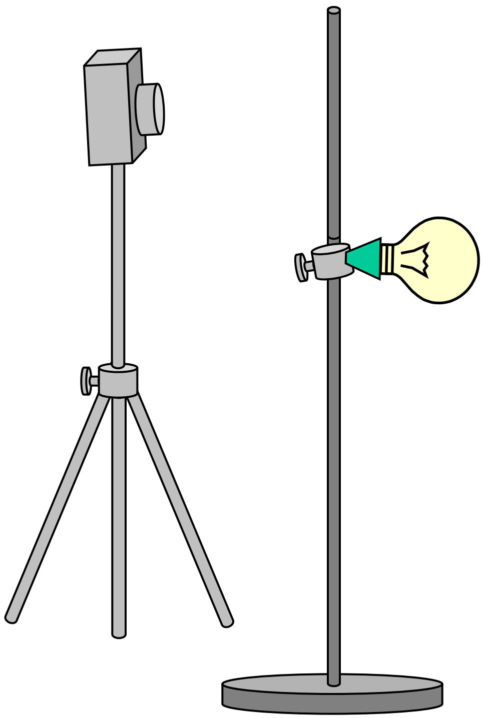
Photometric stereo



Photometric stereo

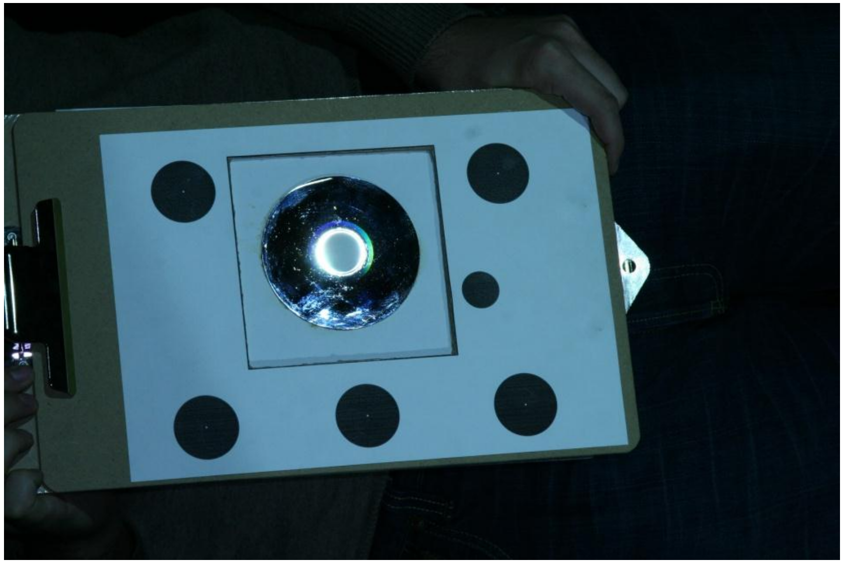


Photometric stereo

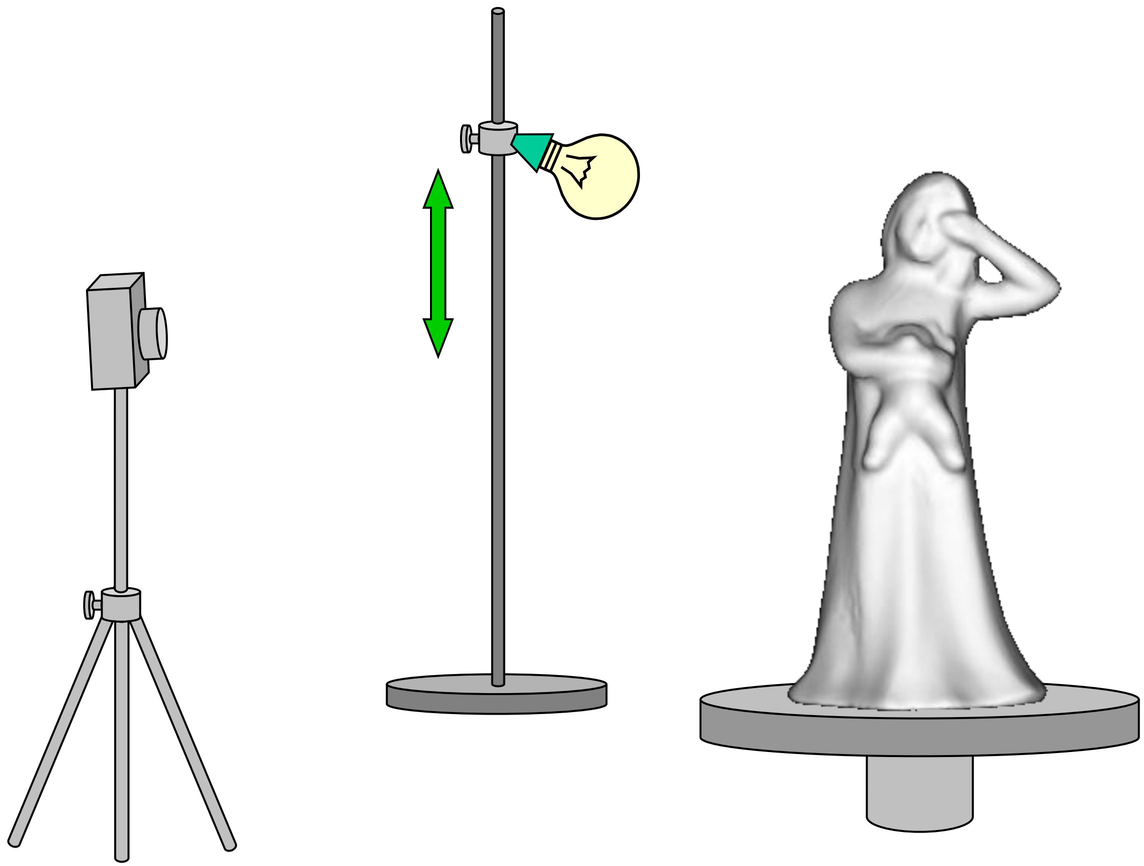


Example of calibration

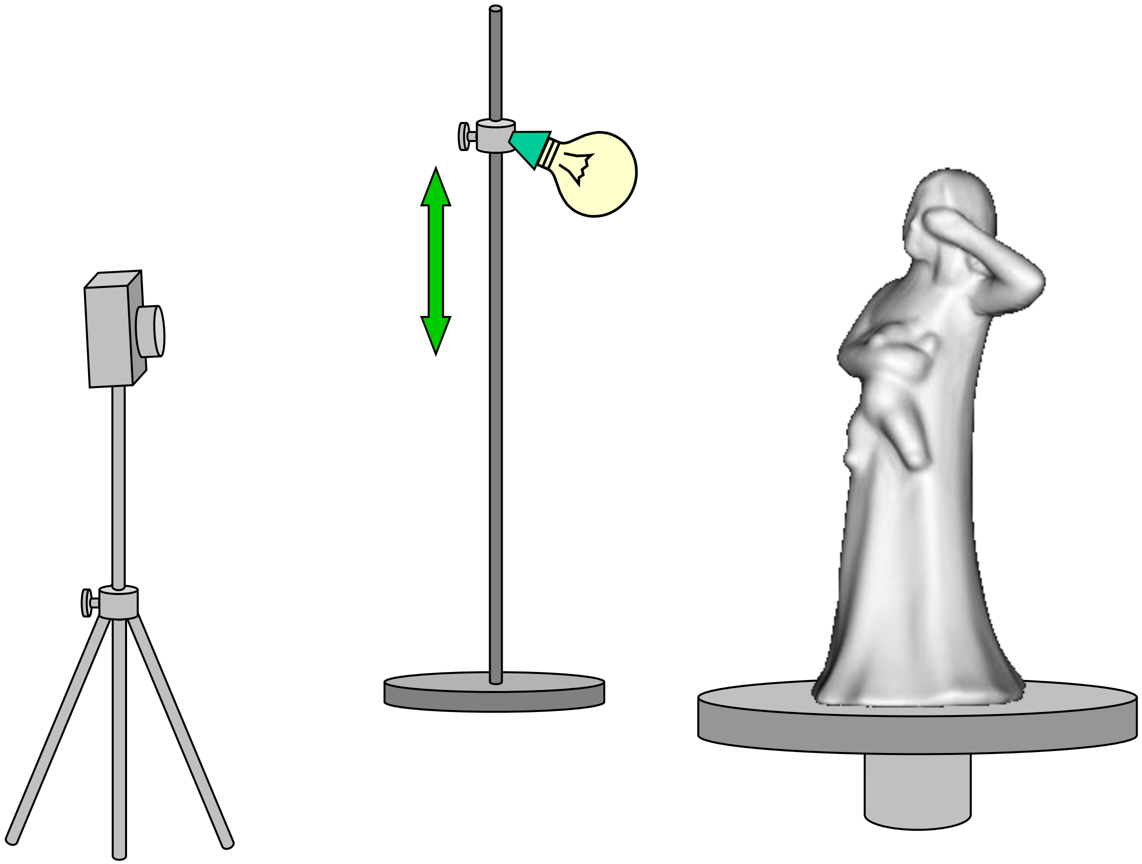
- Can use mirror spheres
- Can also use a mirror attached to known planar pattern

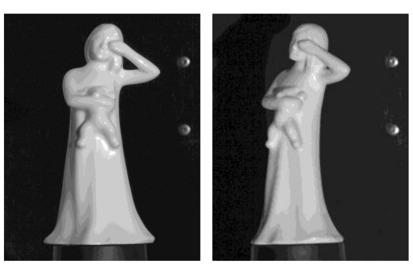


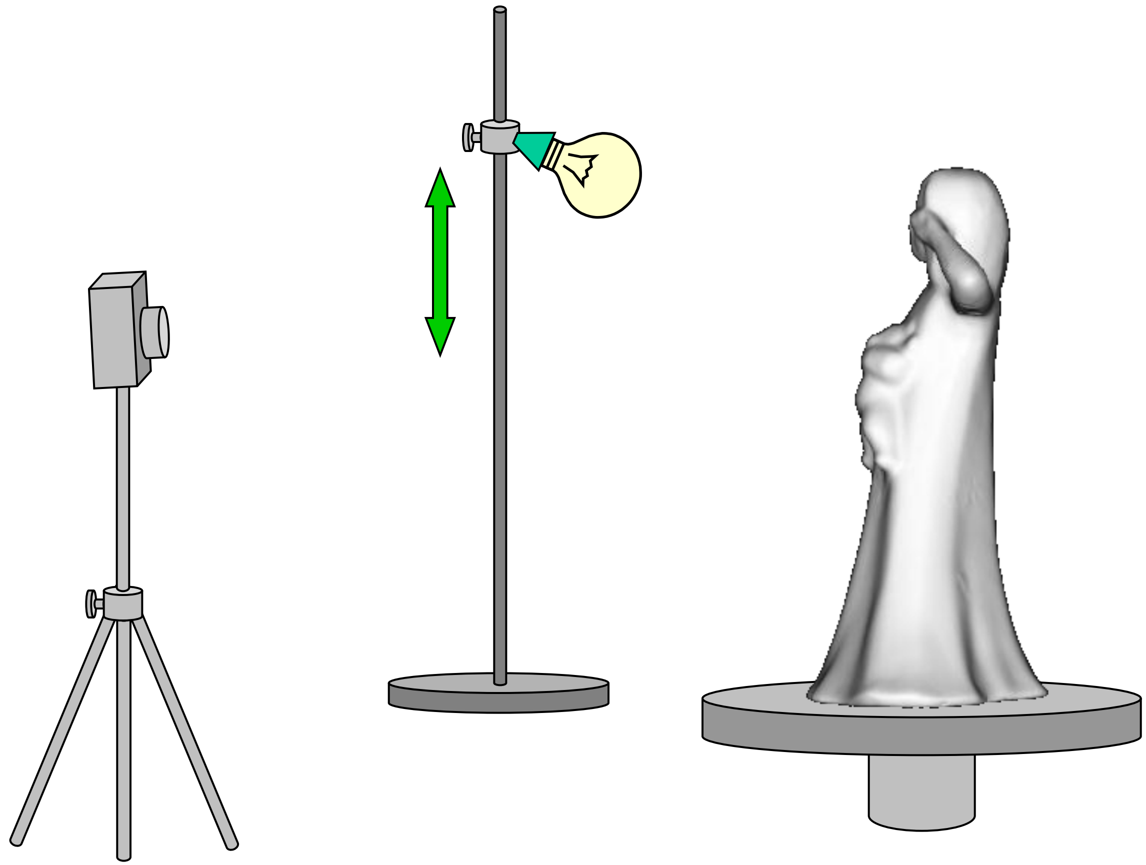
UNIVERSITY OF CAMBRIDGE

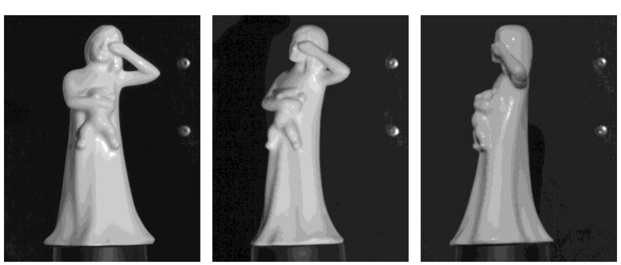


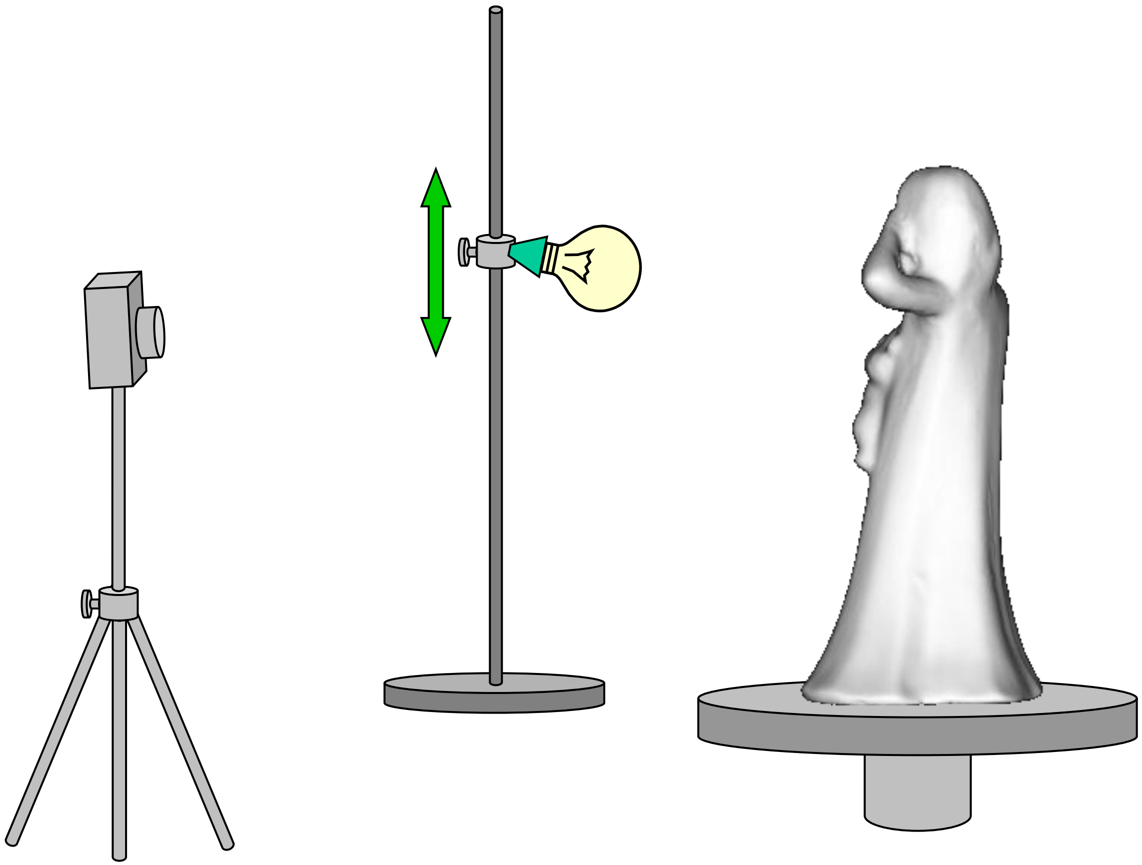
UNIVERSITY OF CAMBRIDGE

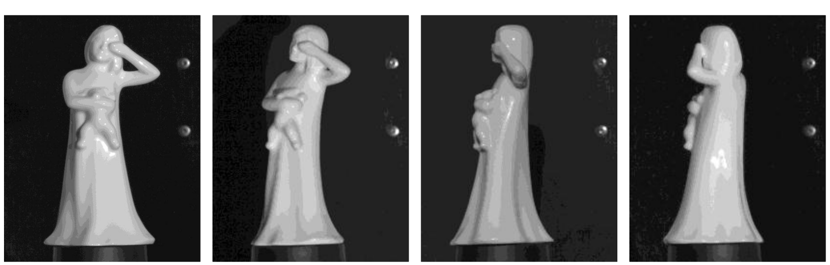


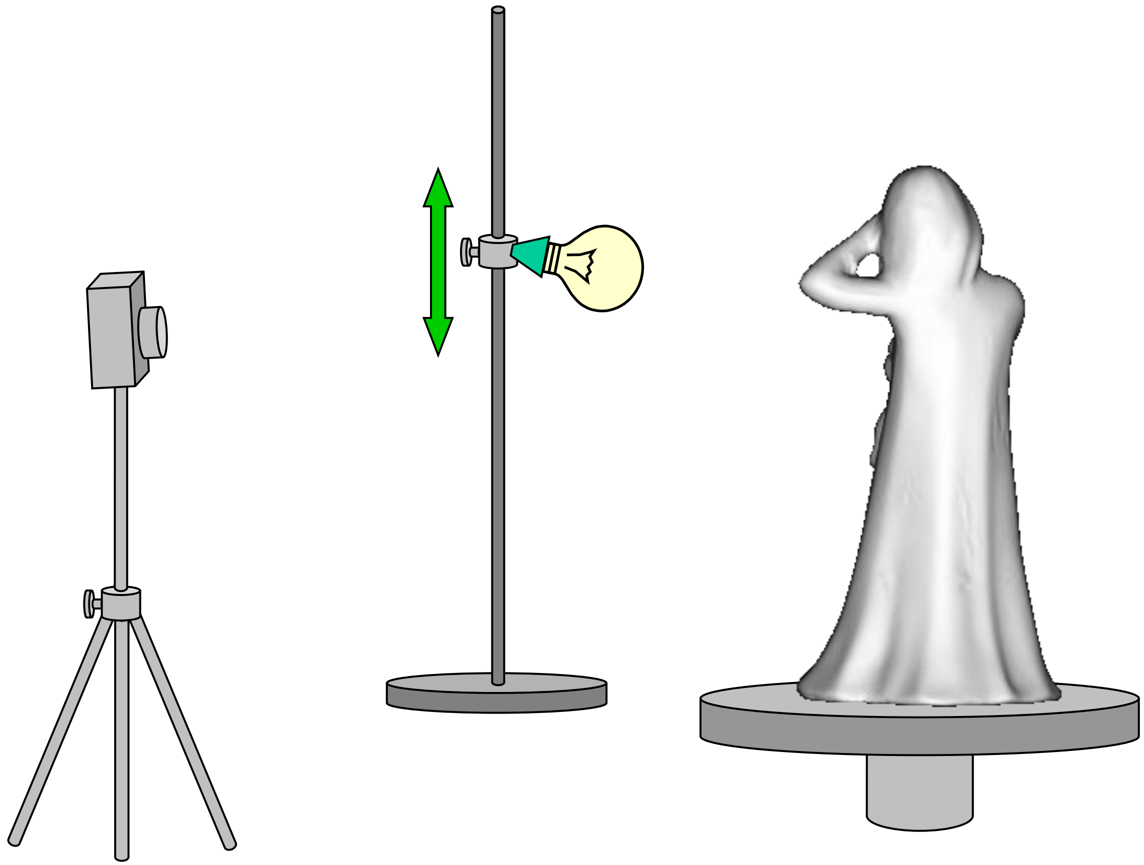


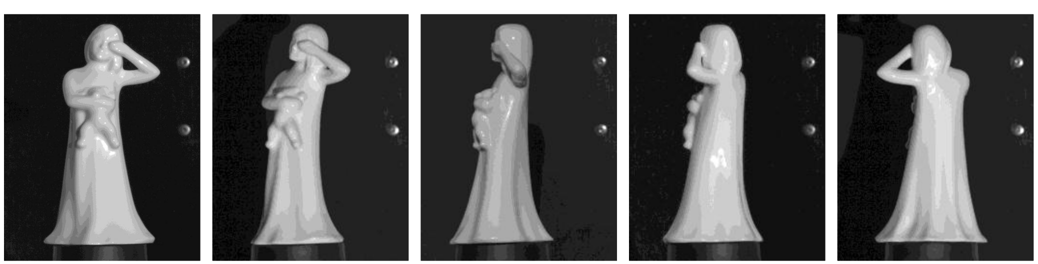




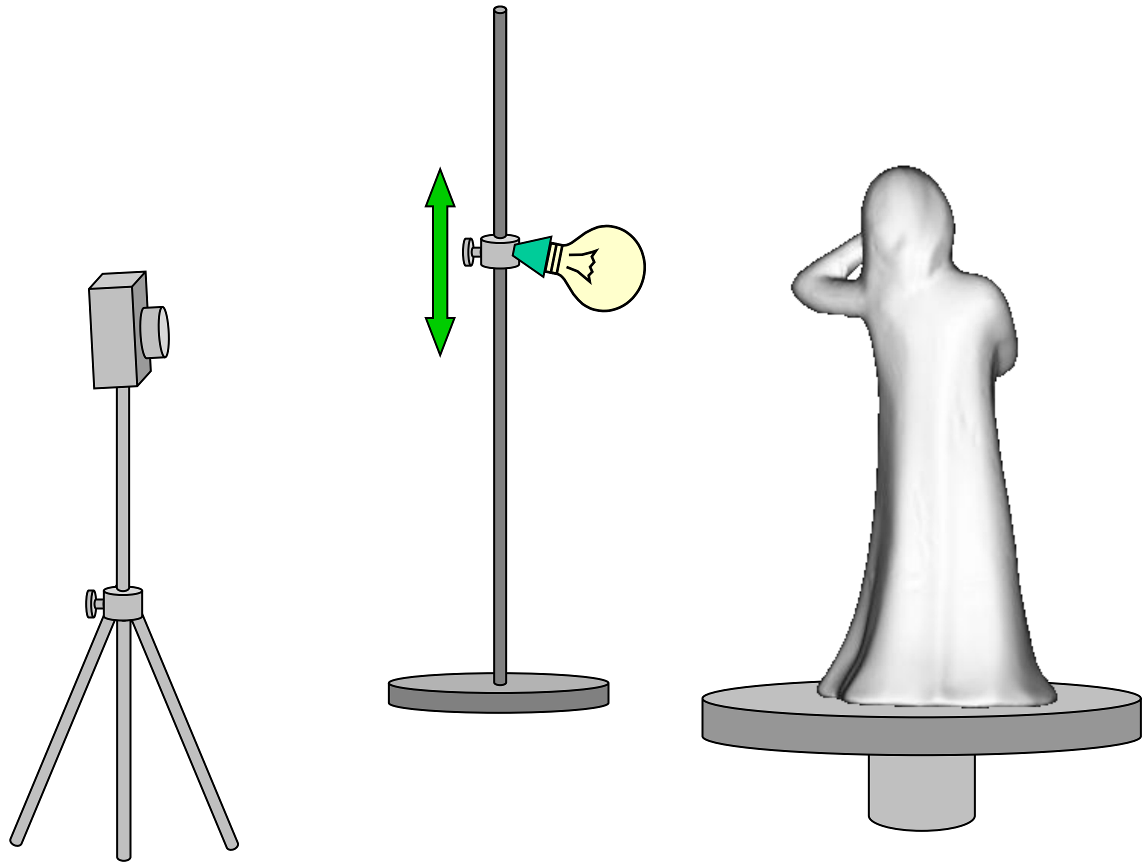


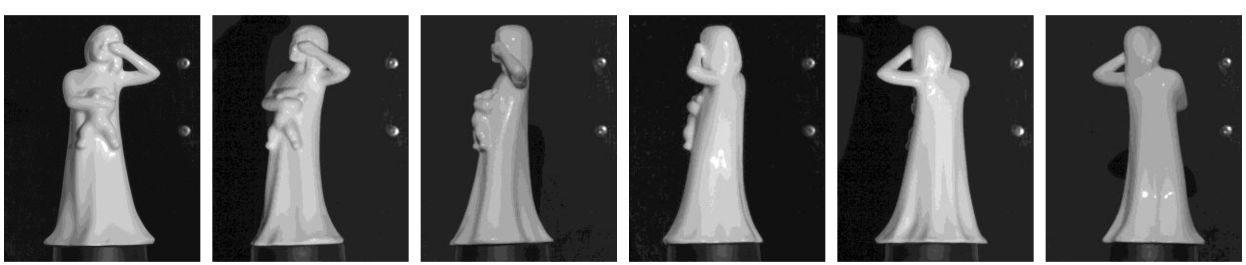


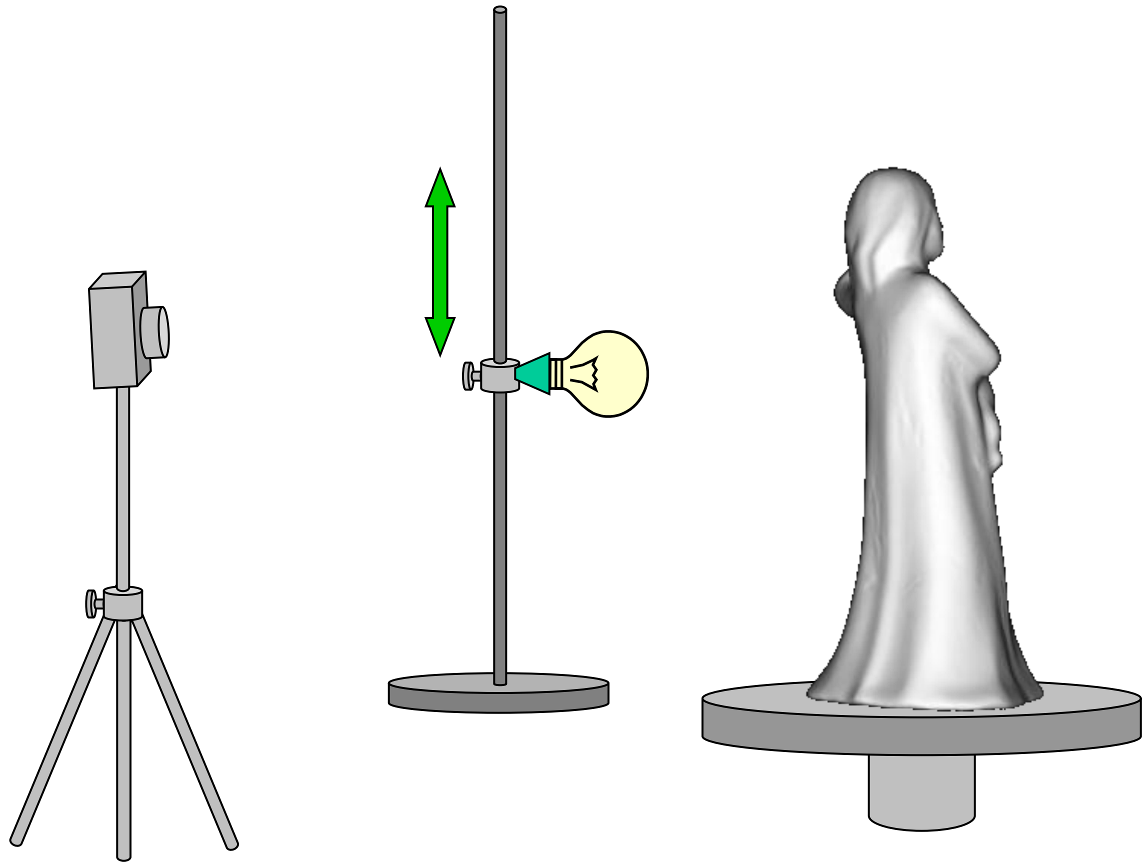


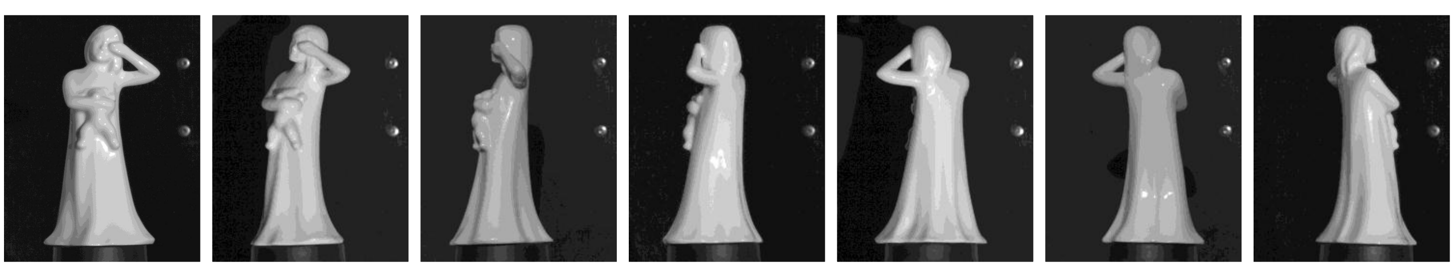


UNIVERSITY OF CAMBRIDGE

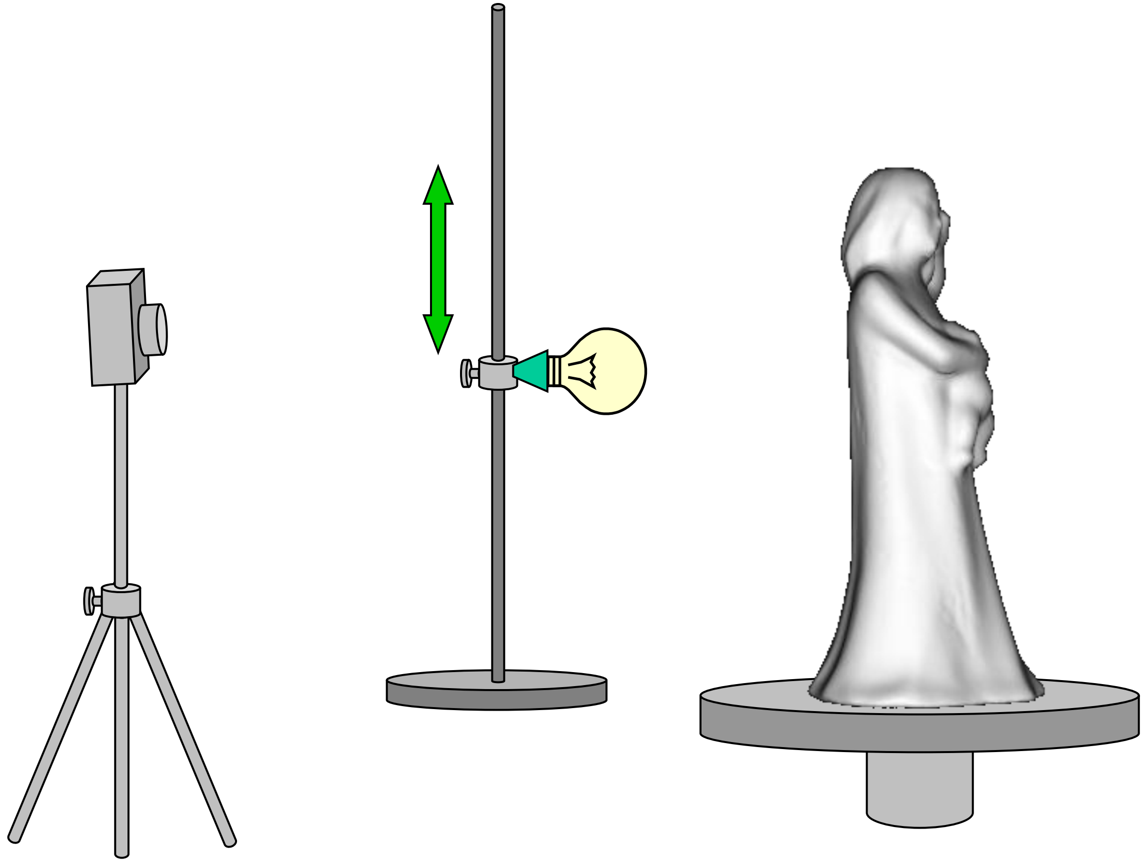


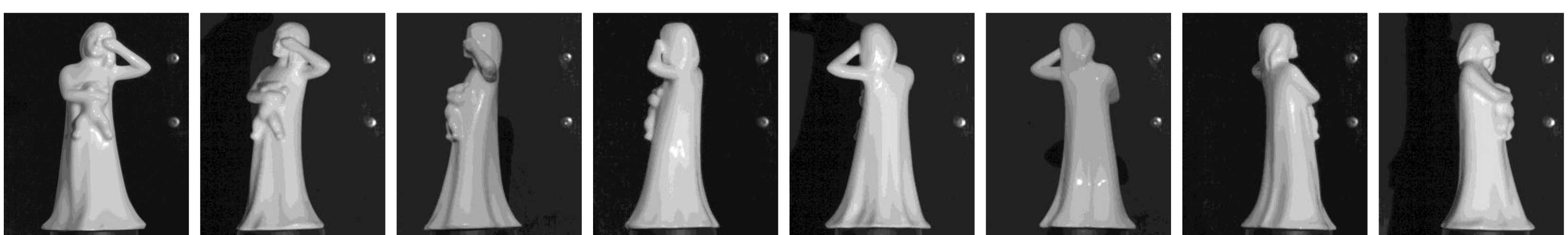




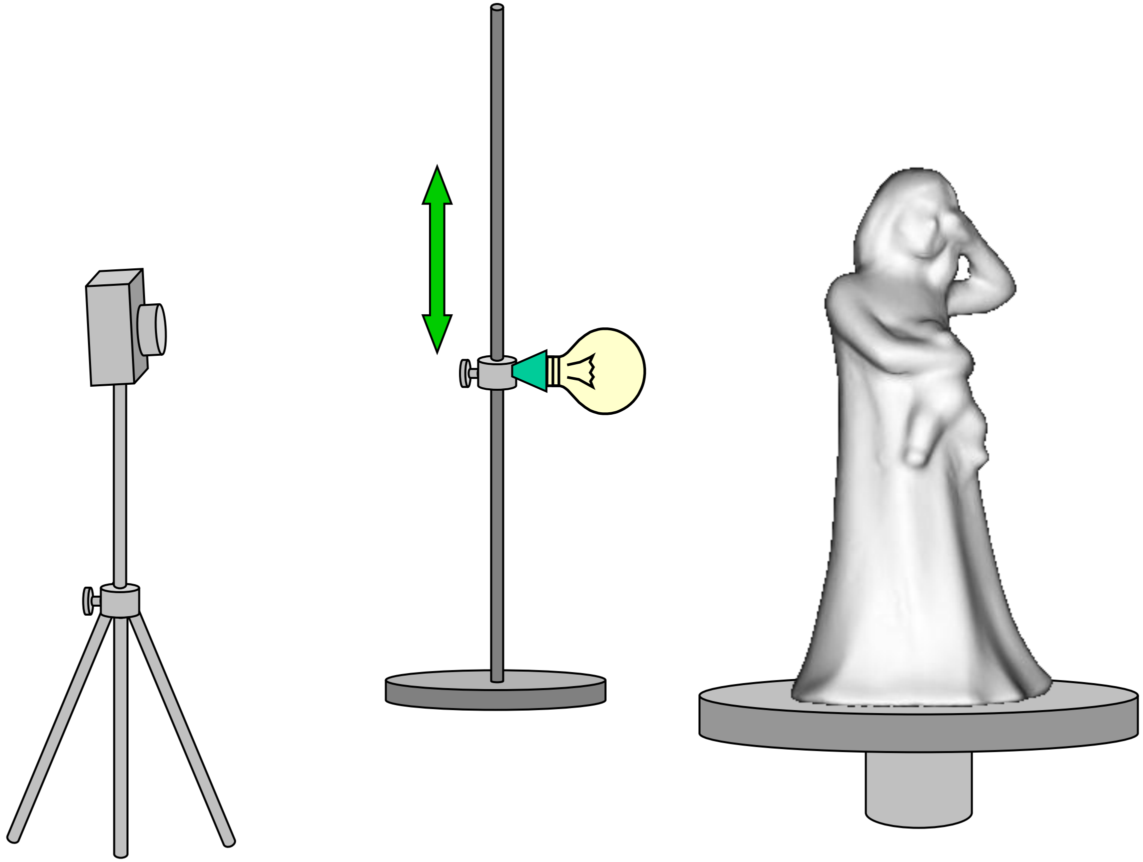


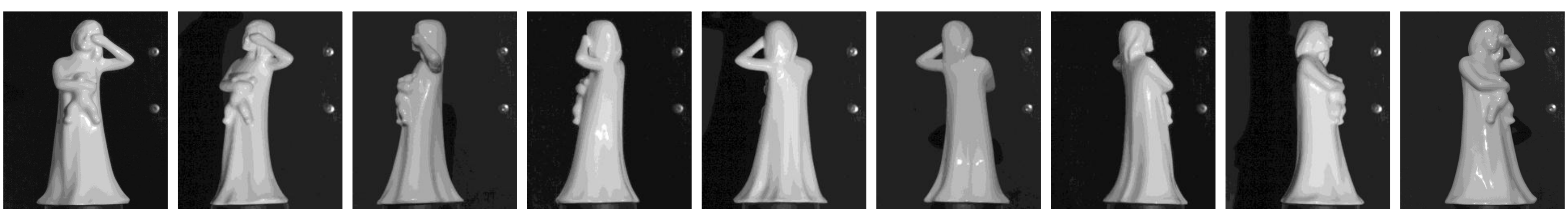
UNIVERSITY OF CAMBRIDGE





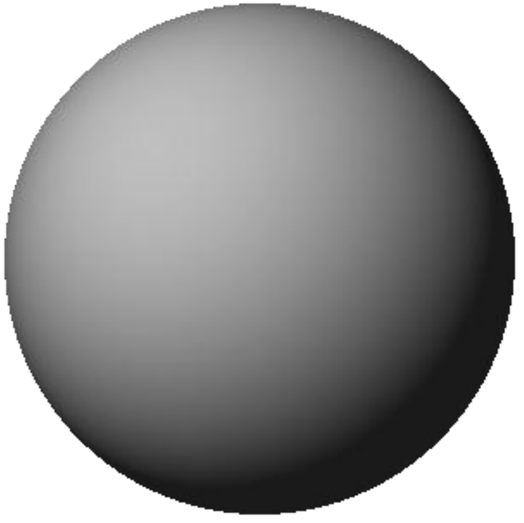
UNIVERSITY OF CAMBRIDGE



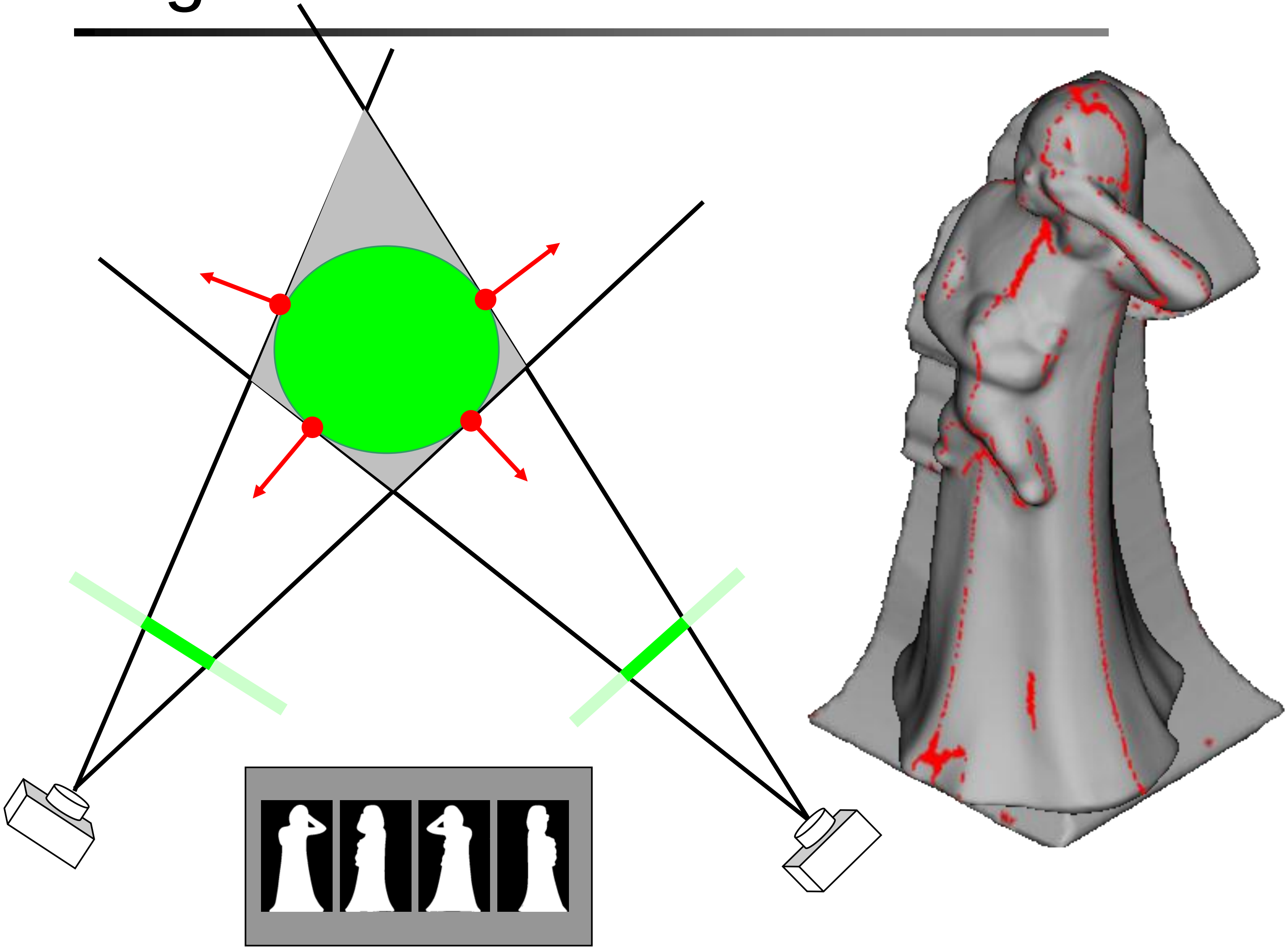


Light estimation

- Calibration object
  - Fully known geometry
  - Light can be estimated from intensity of all points



Light estimation



Light estimation

- Virtual calibration object
  - Partially known geometry
  - Light can be estimated from intensity of correct points
  - How do we eliminate incorrect points?

Voting approach

- Given a light direction, a point is **consistent** if its predicted appearance matches observation

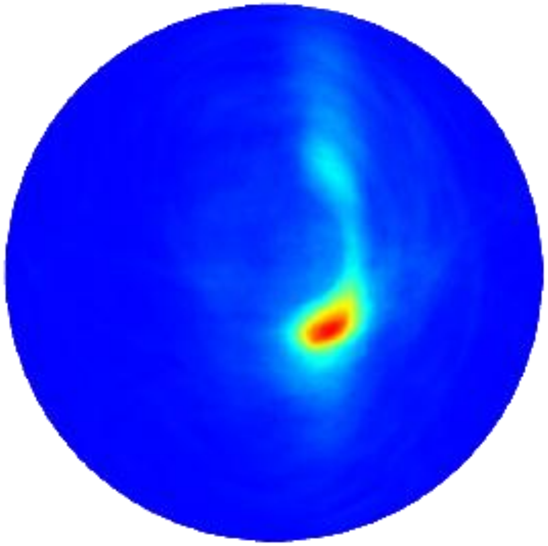
- For each possible light direction count **consistent** points on the visual hull

- Optimal light is obtained when number of consistent points is maximised

Number of consistent points

Visual hull

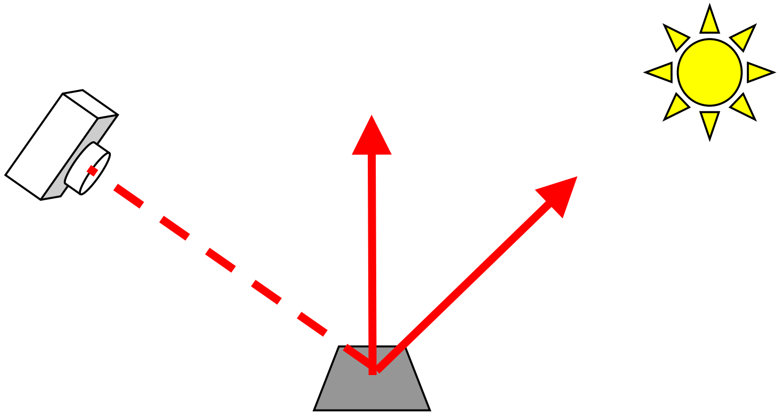
Sample input image



Consensus according to light direction

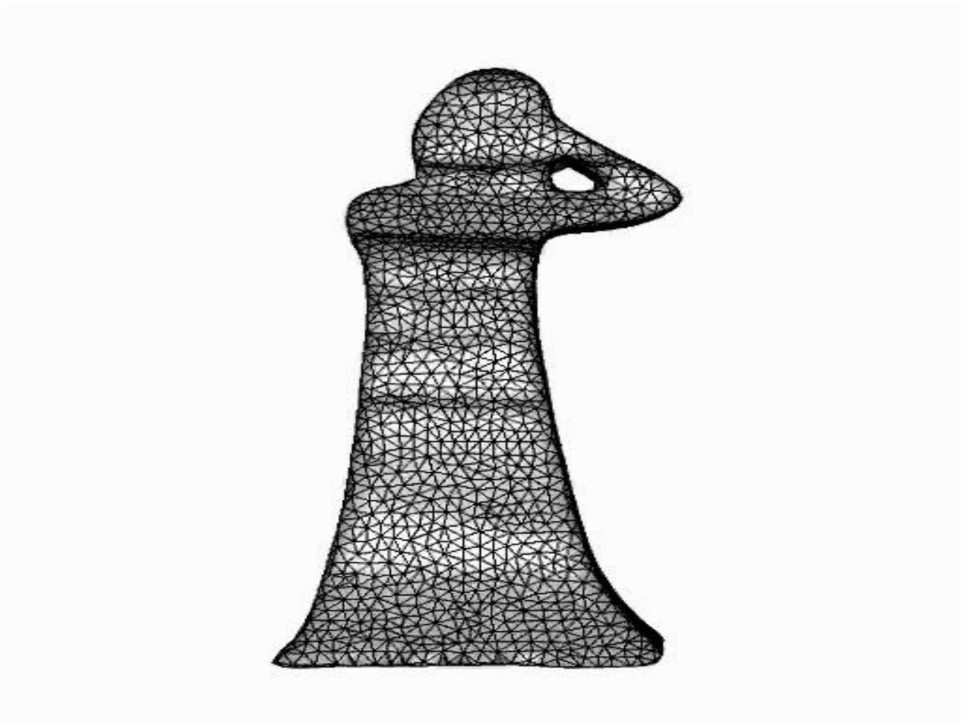
Algorithm

- Estimate light direction $\boldsymbol{l_k}$ in image k
- Evolve surface until predicted appearance under illumination $\boldsymbol{l_k}$ matches image k



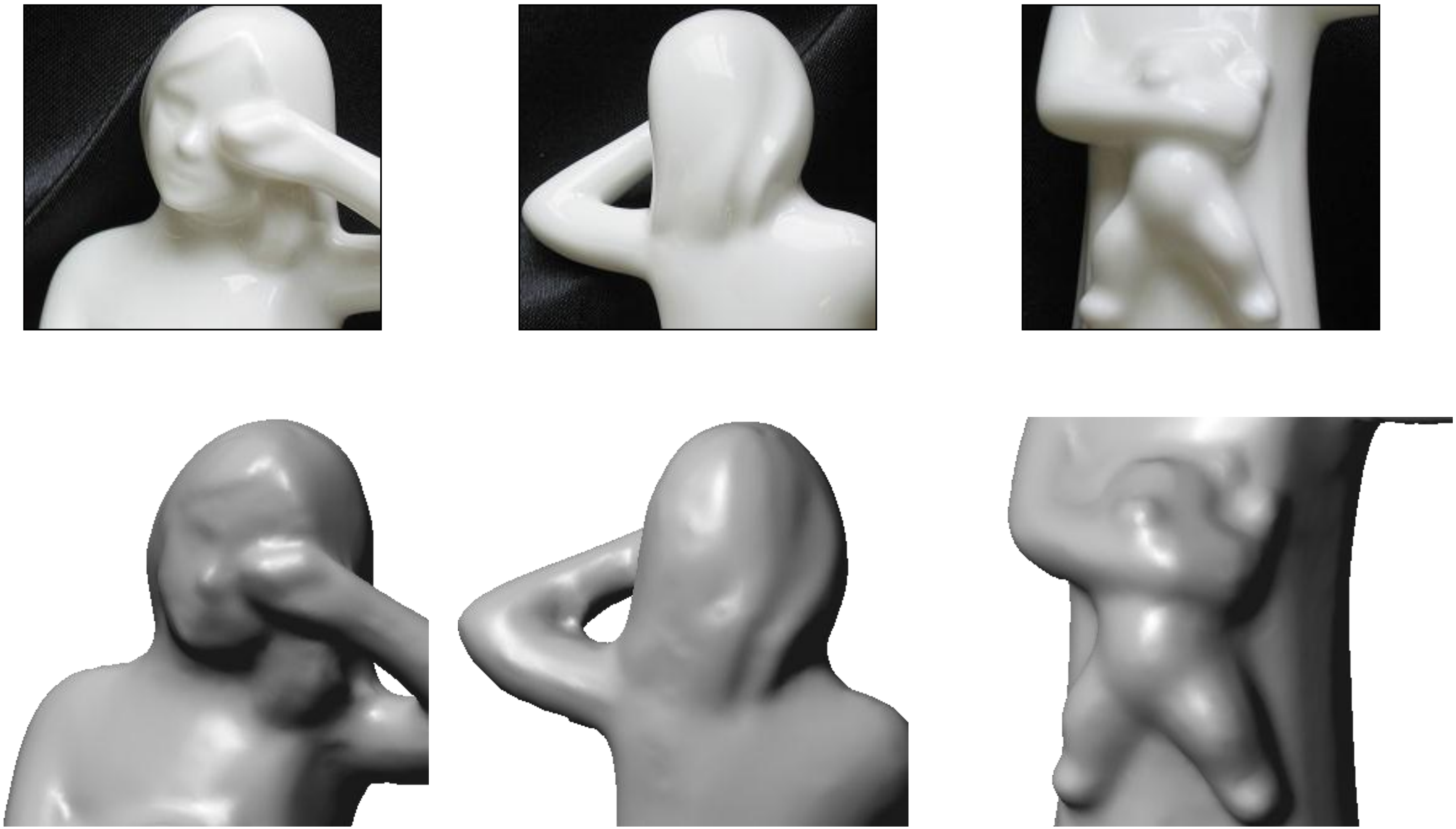
$$\boxed{i = \mathbf{l}^{\mathbf{T}}\mathbf{n}}$$

Surface Evolution: 3D Mesh



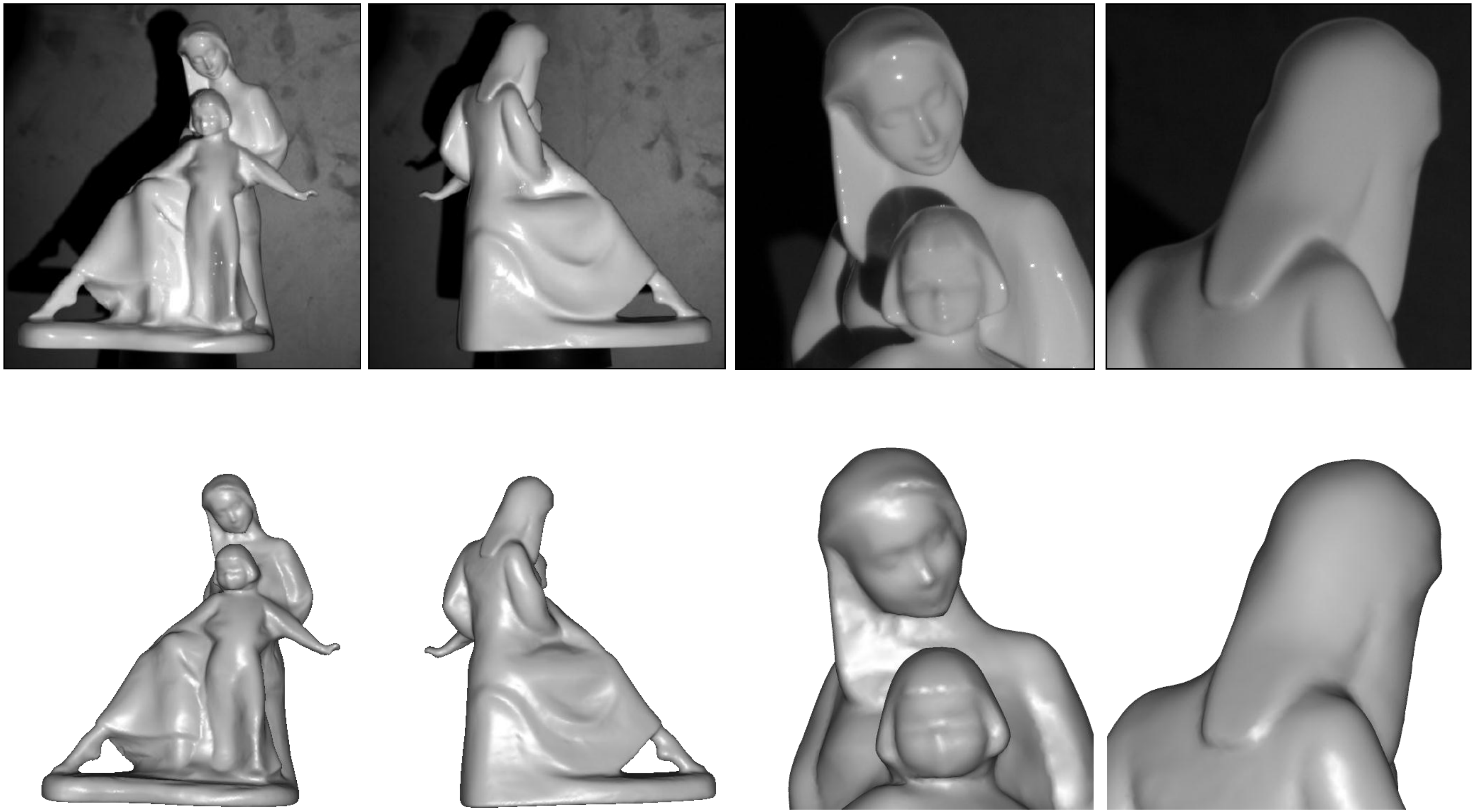
Evolve mesh until it is predicted appearance under recovered illumination matches images

3D Models



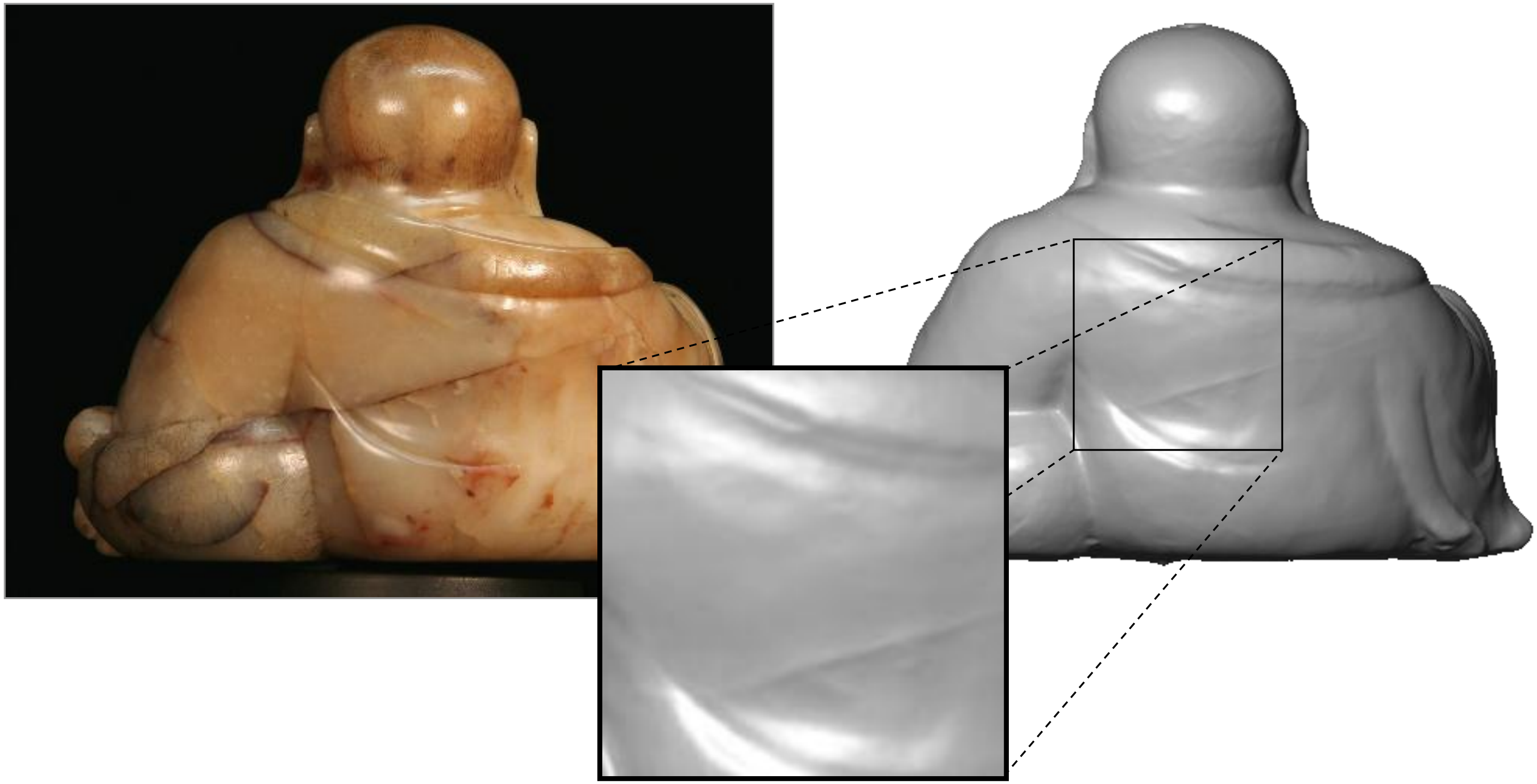
Making physical copies

3D Models



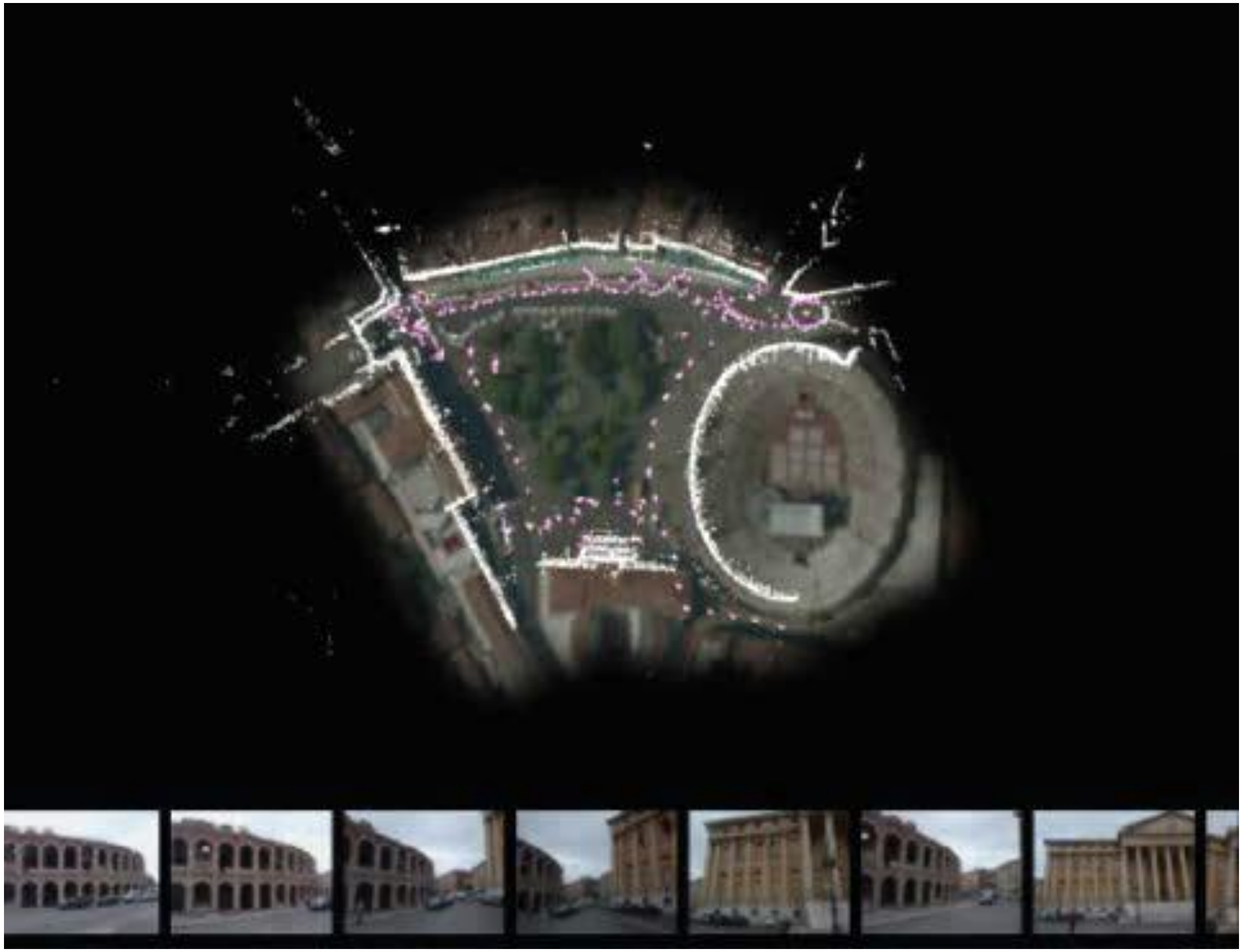
3D Models

3D Models



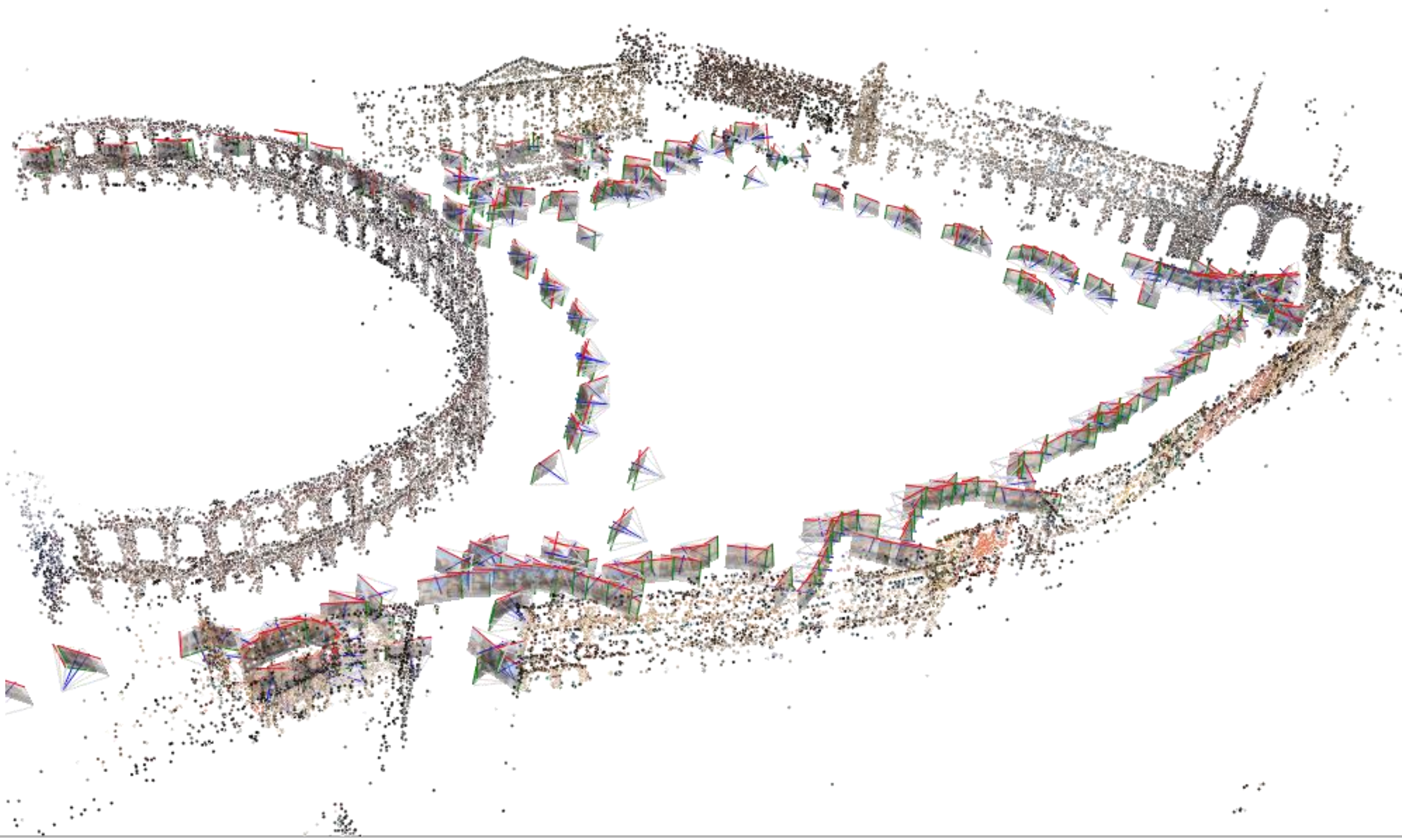
Large-scale reconstructions

Large Scale Reconstruction



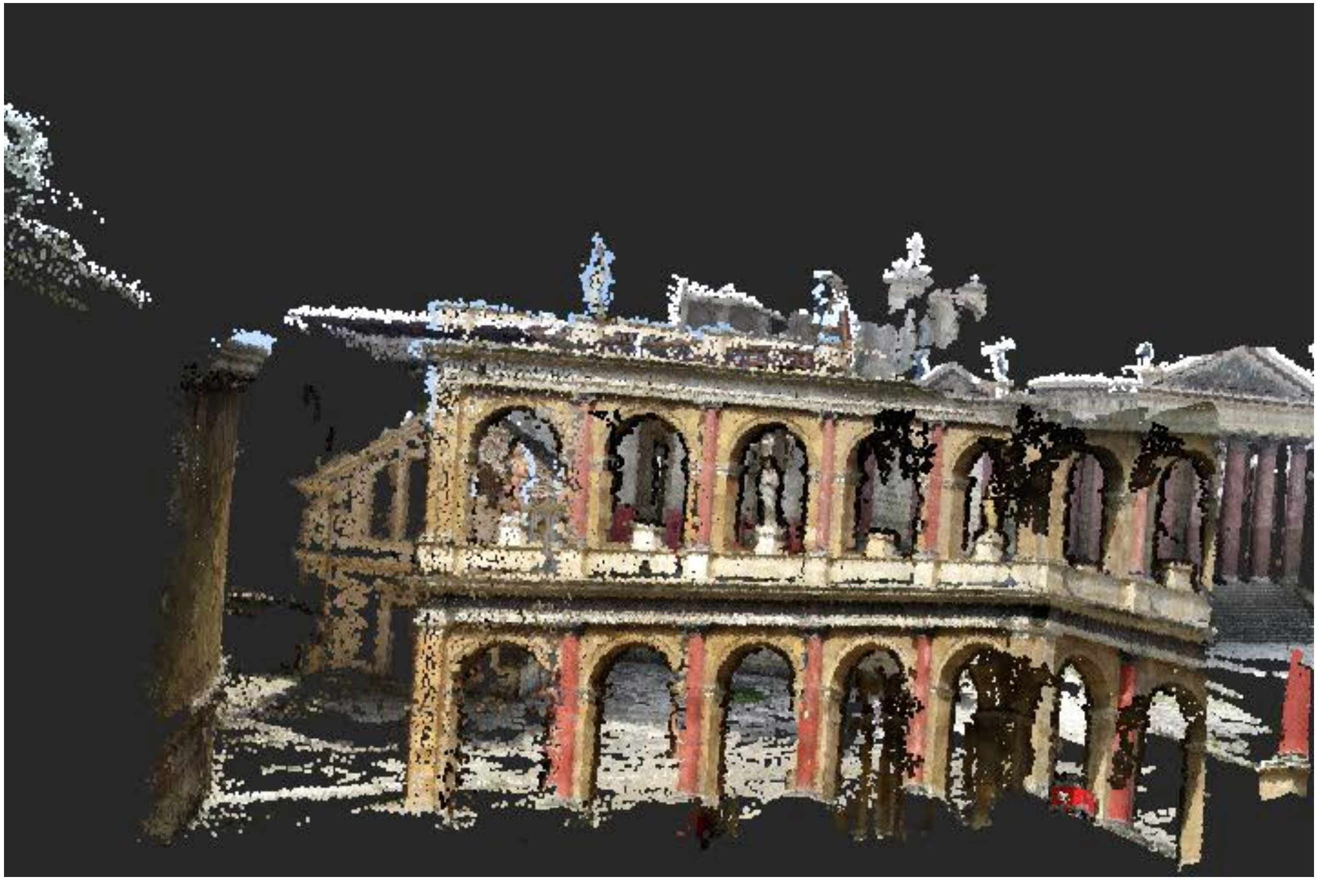
Large Scale Reconstruction

Dataset “Piazza Bra”, 300 images, 100K points, 3 hours, error 0.1-0.2%

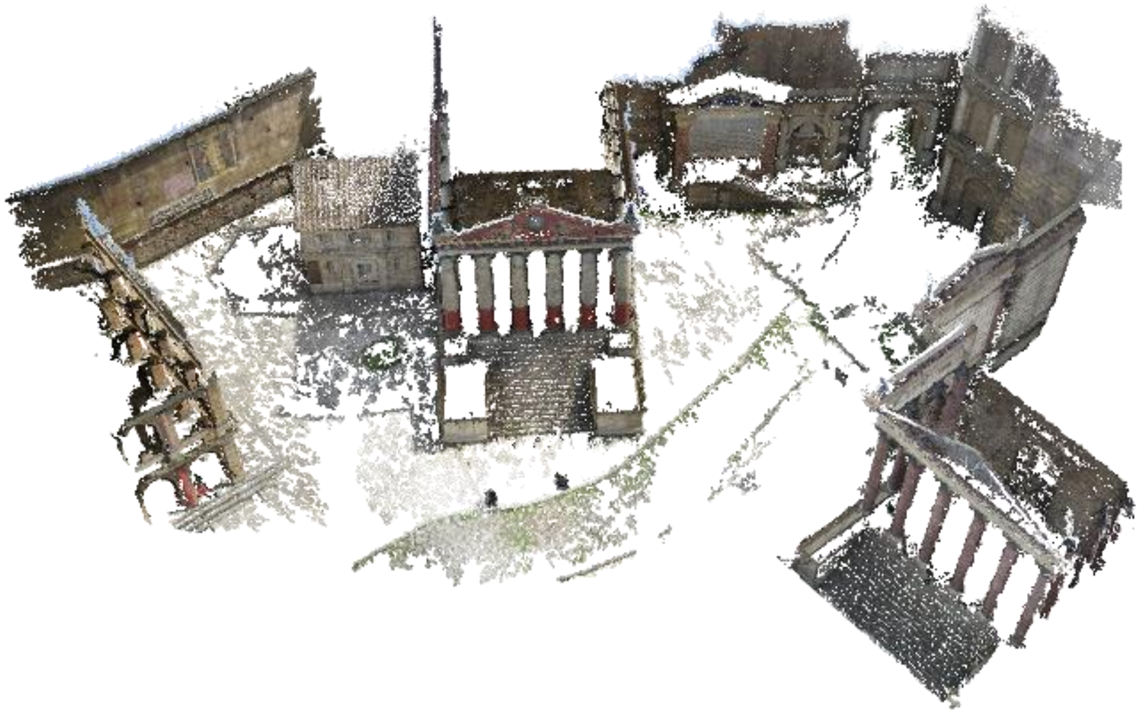


Large Scale: Reconstruction of Forum Romanum

Large-scale reconstruction

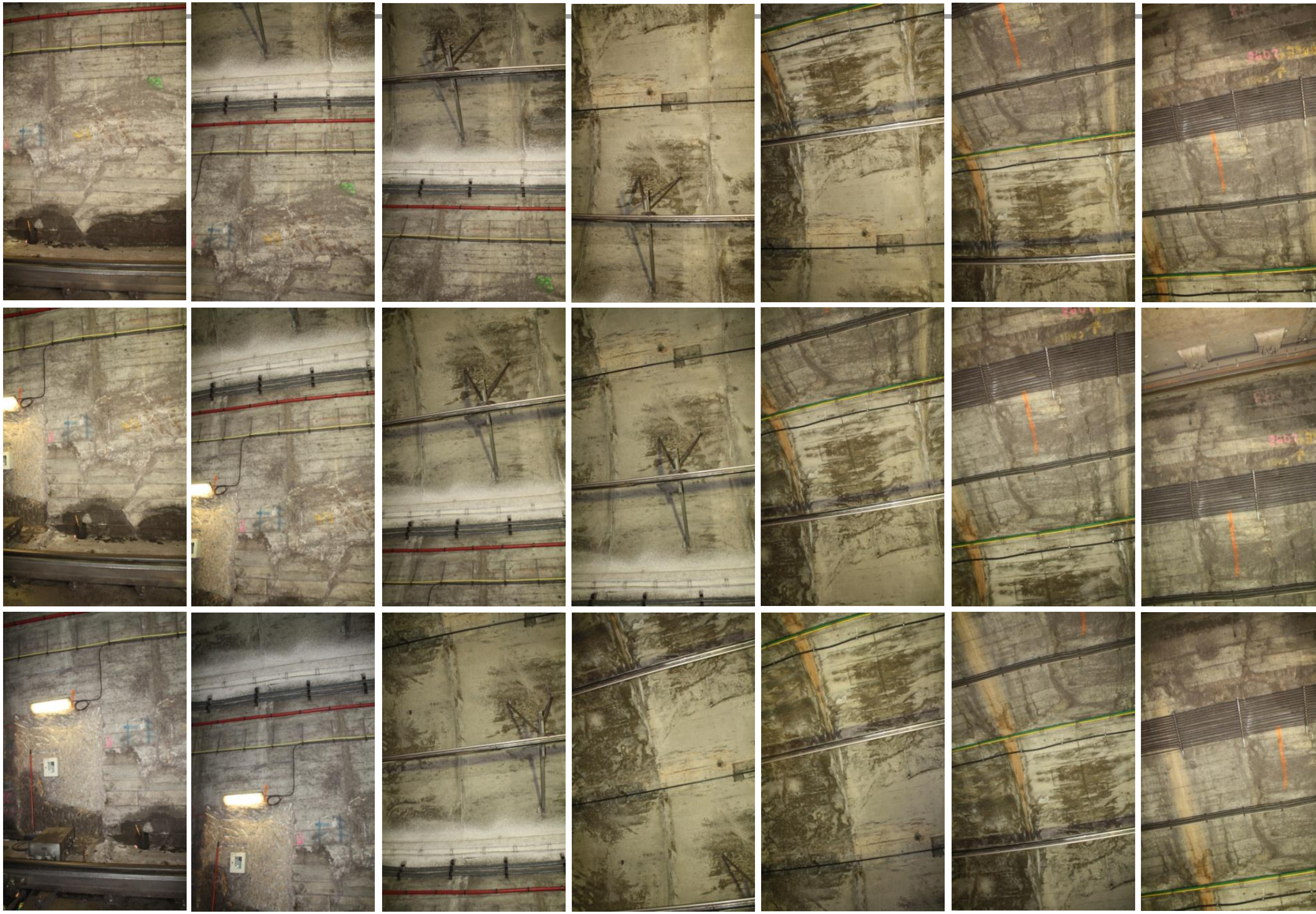


Large-scale reconstruction



Sagrada Familia Station

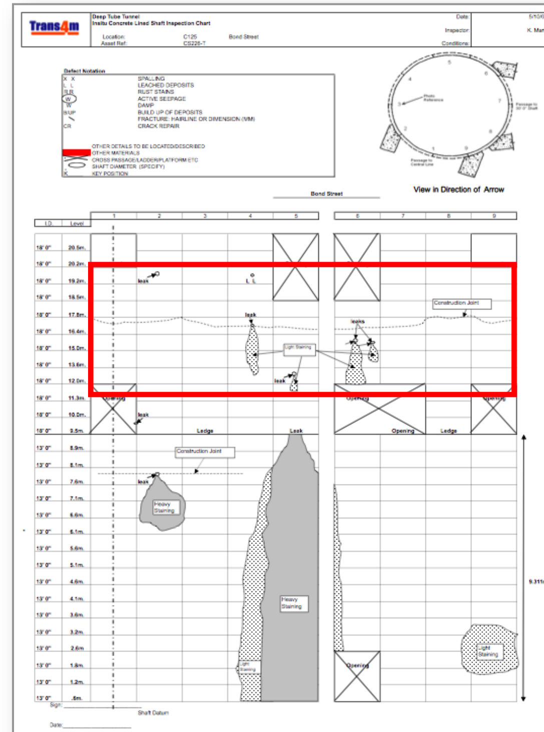
Barcelona



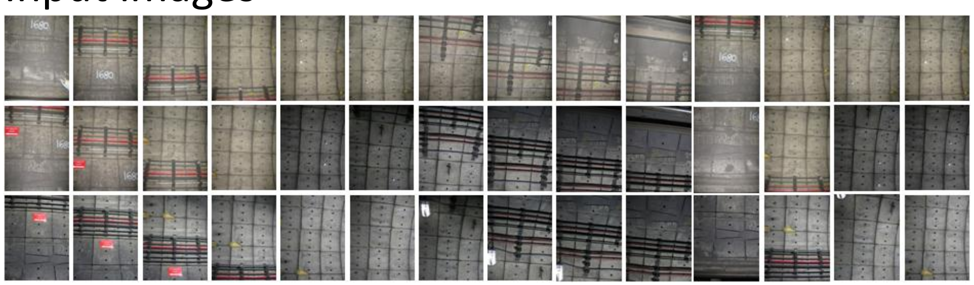
Infrastructure Reconstruction (CUED)

Tunnel inspection

Inspection report



Input images



Panoramic mosaic

- 35 images of Antrim Coast Road cliff, taken with off-the-shelf DSLR camera (£400)

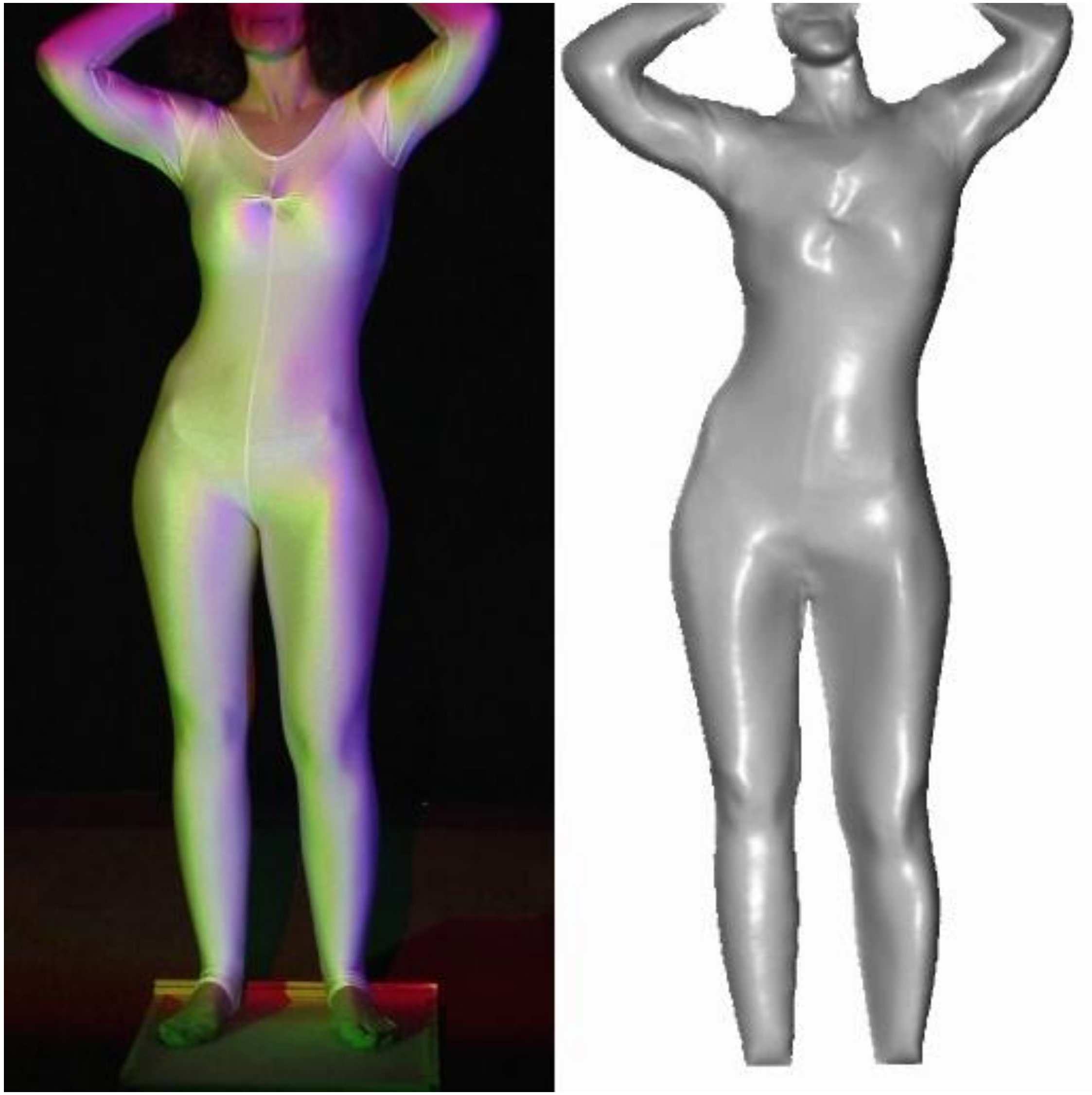
UNIVERSITY OF CAMBRIDGE

Deformable objects:

Real-time photometric stereo using colour lighting

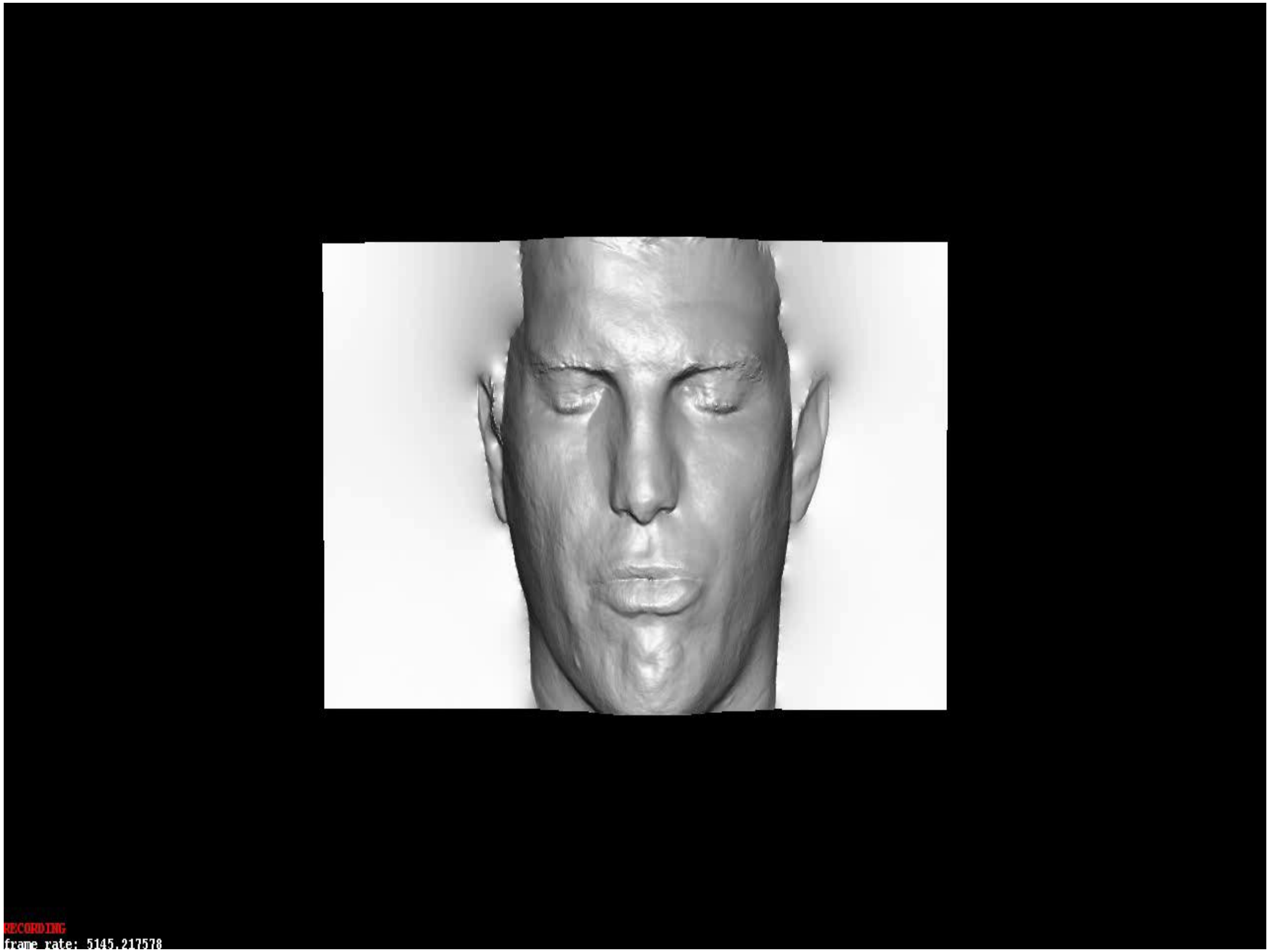
Hernandez et al 2007
Anderson, Stenger and Cipolla 2010-2011

3 Untextured and deforming

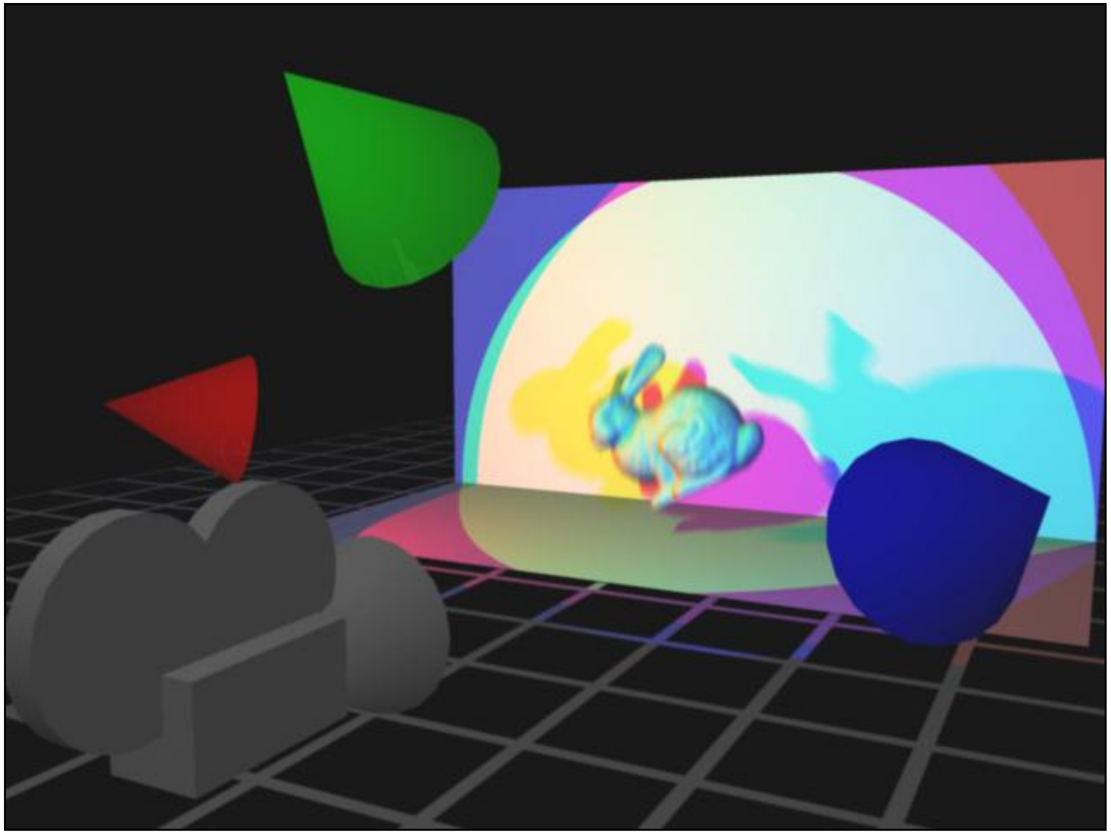


Colour Photometric Stereo

Real-time deformable surfaces

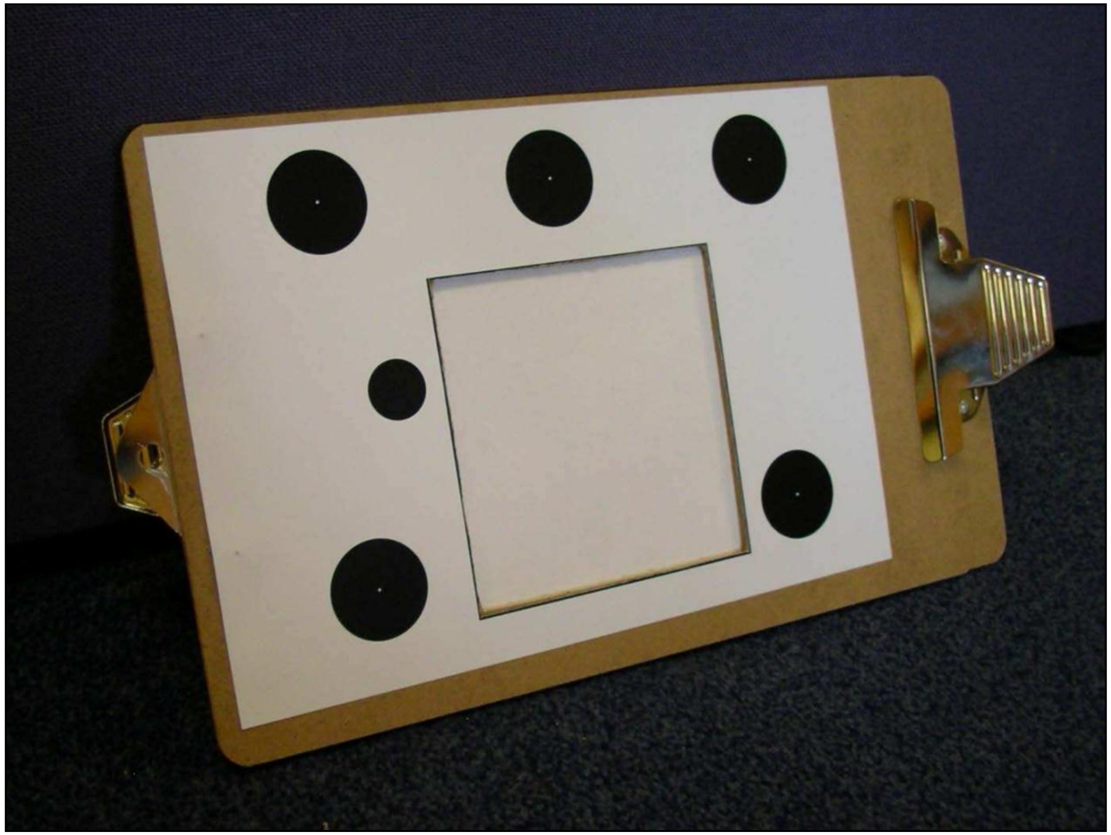


Textureless deforming objects



- a method for reconstructing a textureless *deforming* object in 2.5d

Textureless deforming objects

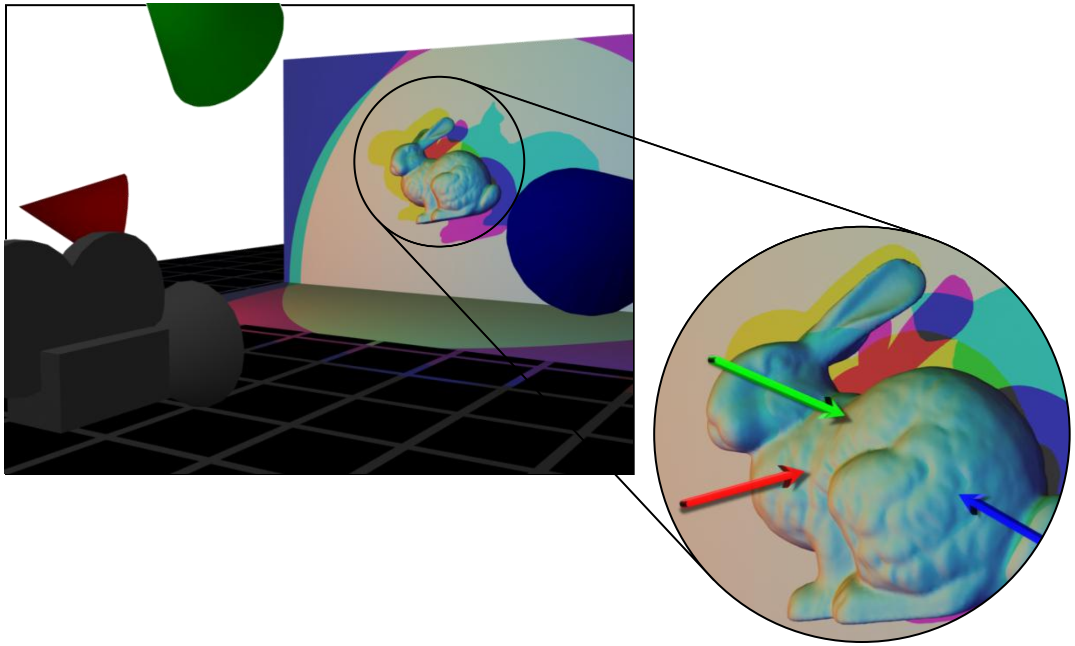


- a method for reconstructing a textureless *deforming* object in 2.5d

Untextured deforming objects

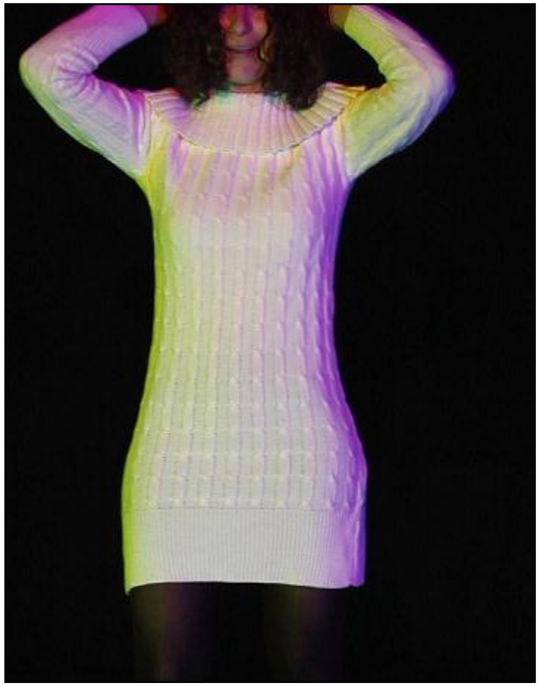
If a white object is illuminated by a red, a green and a blue light source, the color reflected by a point on the surface is in 1-1 correspondence with the local orientation.

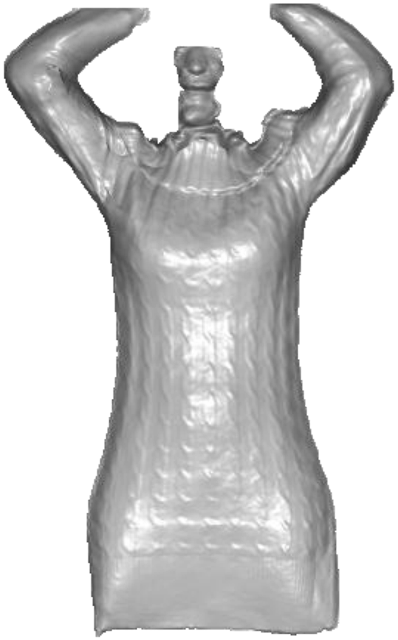
A. Petrov. Light, color and shape. Cognitive Processes and their Simulation (in Russian), pages 350–358, 1987



Coloured photometric stereo

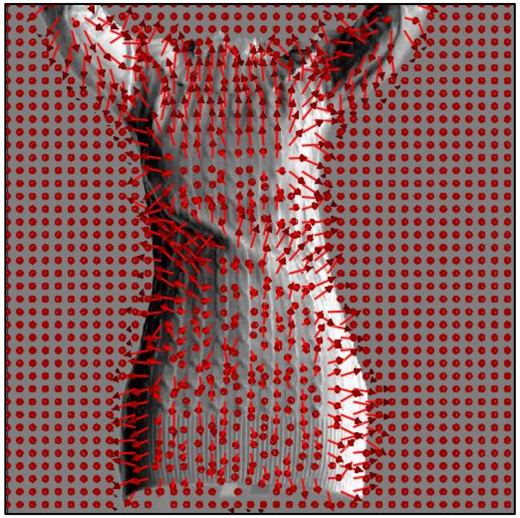
- observation: 1-1 mapping between colour and surface orientation
- get map of surface orientations from colour image
- integrate orientations to get depth map
- do this for colour video to get 2.5d reconstruction of deforming object!





Coloured photometric stereo

Single frame
from video



RGB Color is converted to
a normal at each pixel

Normals integrated using
FFT Poisson solver

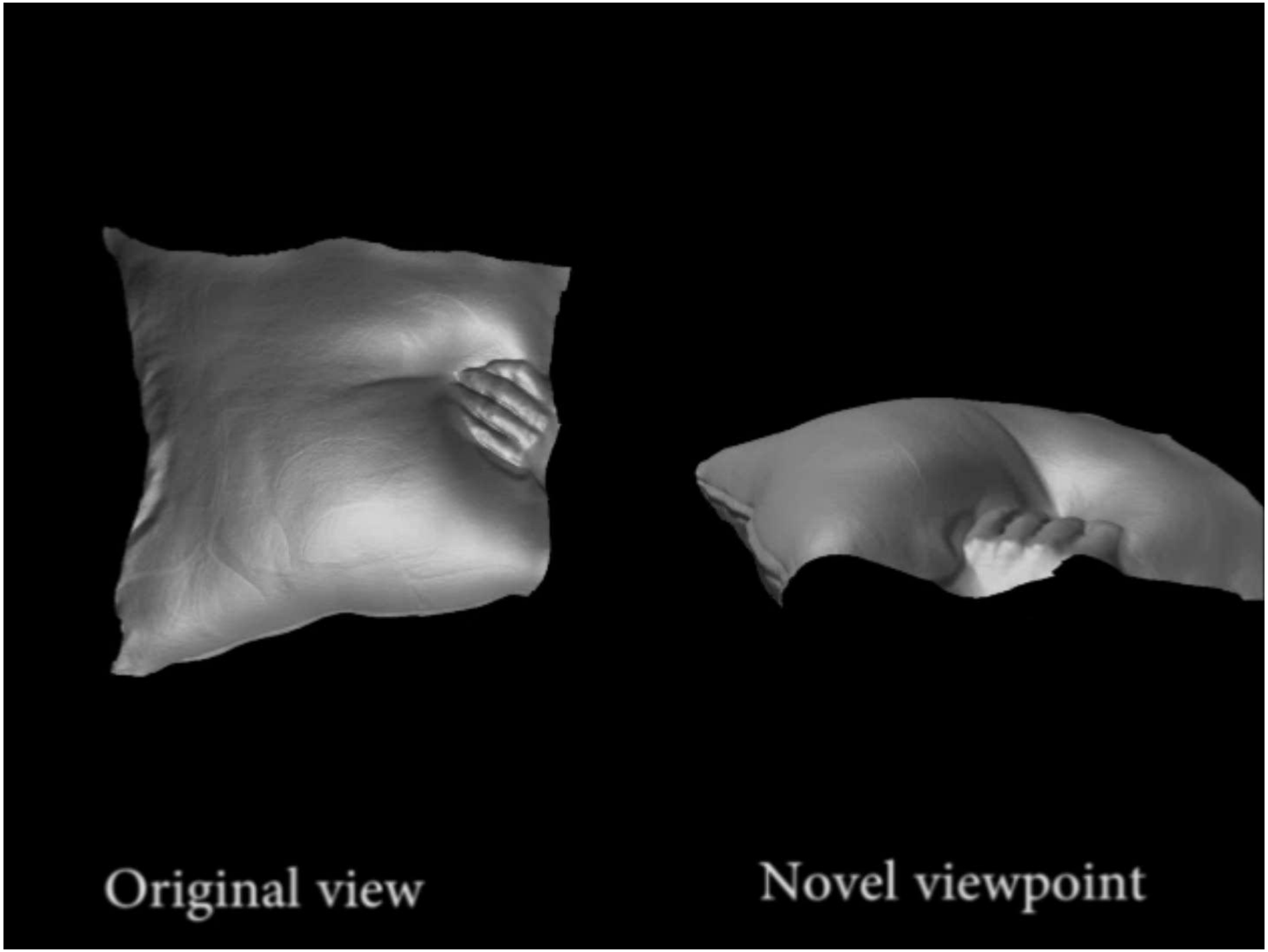
Results

classic photometric stereo

coloured photometric stereo

Multicoloured surfaces

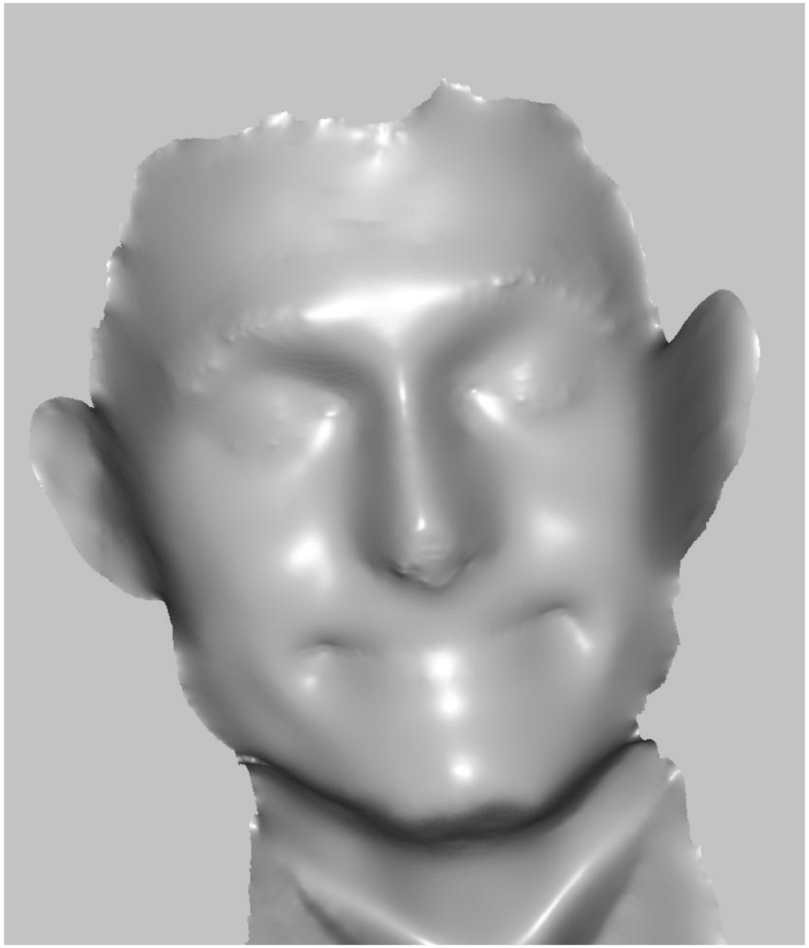
Multicoloured surfaces



Anderson, Stenger, Cipolla ICCV 2011

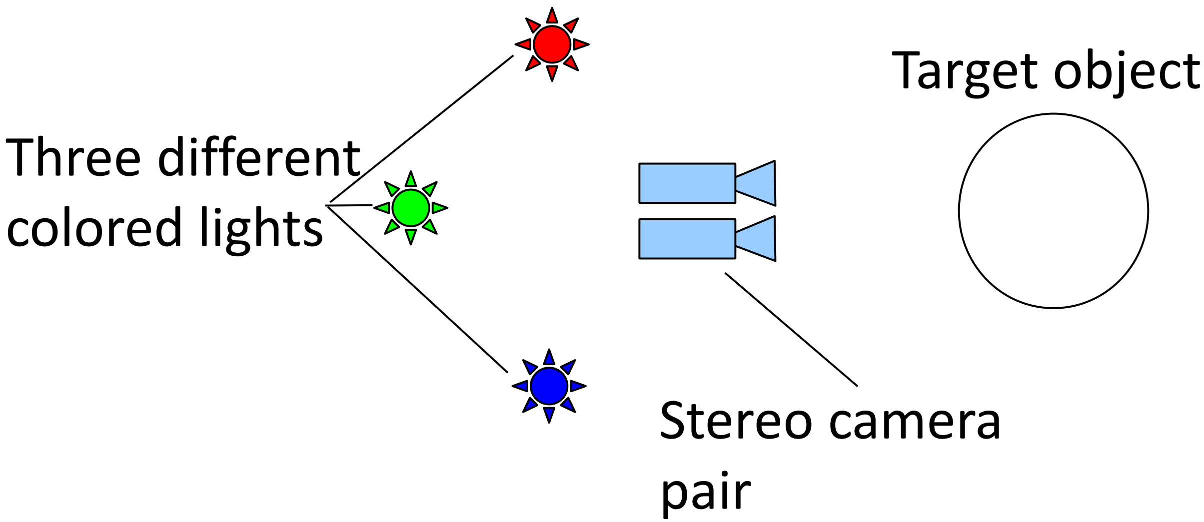
Dynamic Face Capture

- Multiview stereo using two cameras can provide coarse geometry.



- Photometric stereo can add much more detail to the reconstruction.

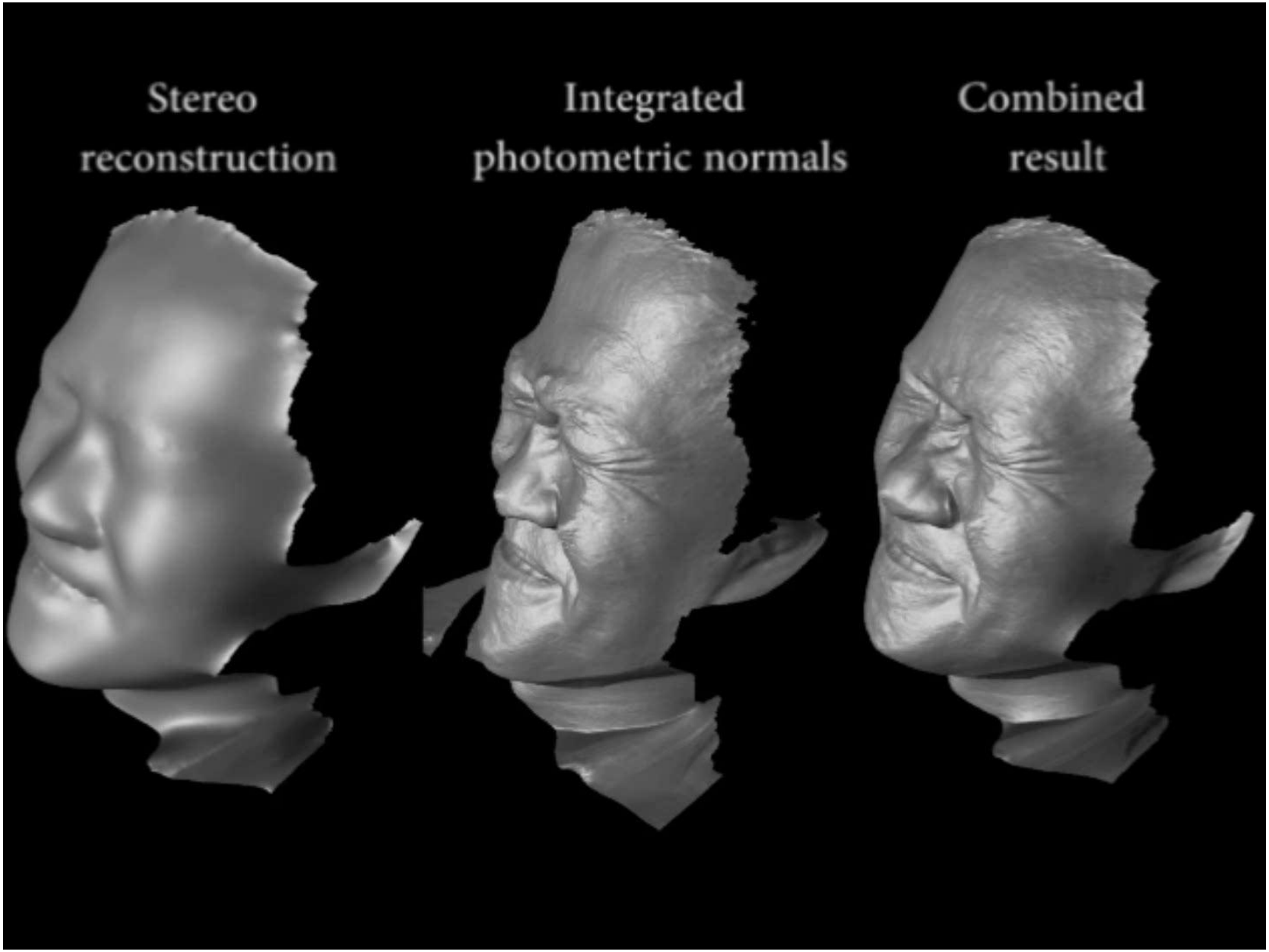
Equipment



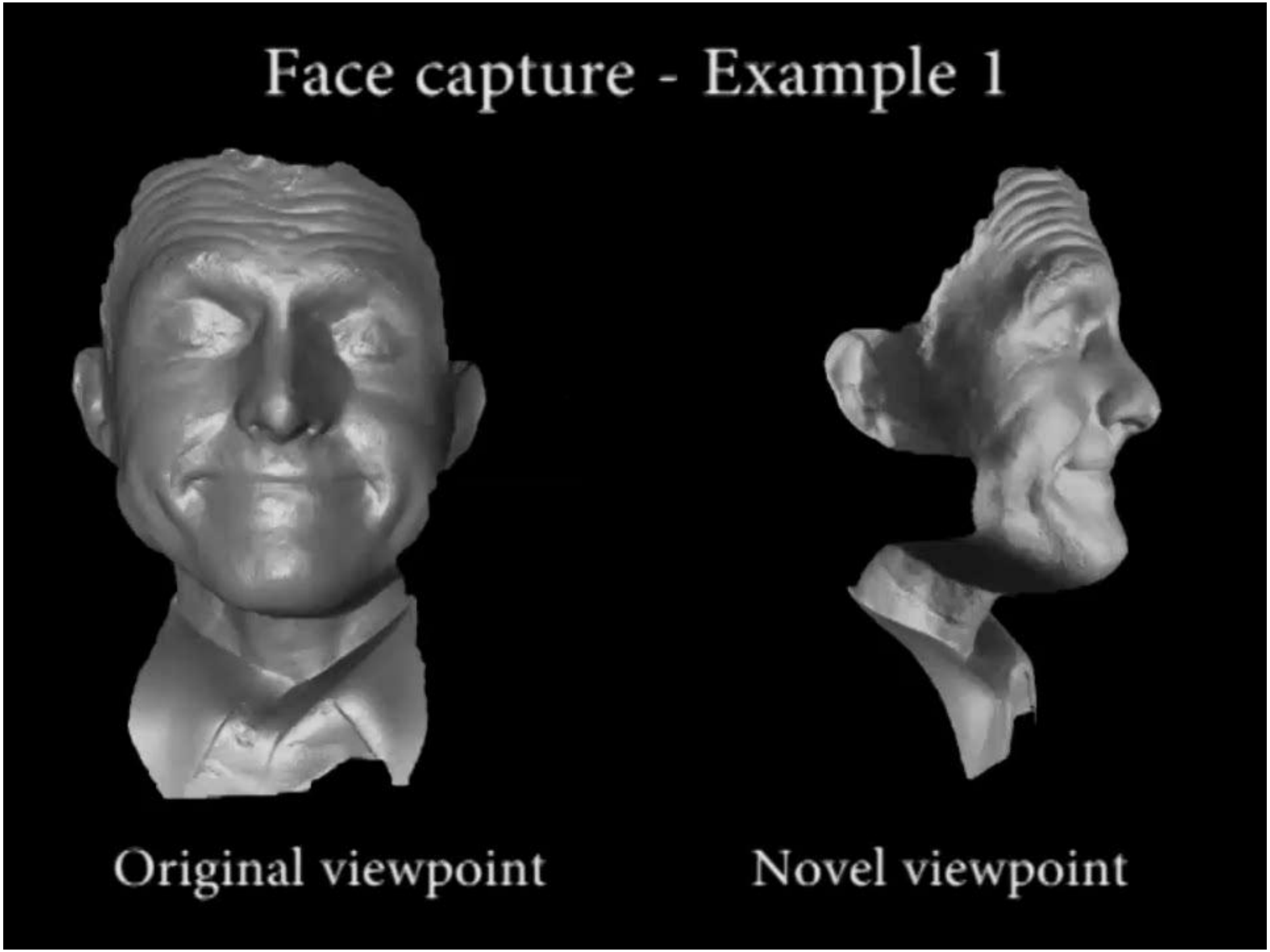

Sample input pair

- Capture takes place at 30 fps.
- Three different colored lights allow photometric stereo to be performed on each frame individually.
- The stereo camera pair allows low frequency geometry to be computed using standard stereo techniques.

Combining Data Modalities



Sample Reconstructions



Overview

- Multi-view stereo
- Multi-view photometric stereo
- Single-view colour photometric stereo

Future work

UNIVERSITY OF CAMBRIDGE

- Need to reconstruct with fewer images
- Exploit 3D examples
- Learn priors for single view reconstruction

Summary

1. 3R's of computer vision at Cambridge

2. Accurate 3D shape recovery

3. Challenges:
   - Large-scale and outdoors;
   - real-time
   - fewer images.

UNIVERSITY OF CAMBRIDGE

http://www.eng.cam.ac.uk/~cipolla/people.html
http://toshiba-europe.com/research/crl/cvg/

Bjorn Stenger, Carlos Hernandez, George Vogiatzis, Rob Anderson, Riccardo Gherardi, Neill Campbell.

Jamie Shotton, Duncan Robertson and Simon Taylor

ECCV 2012