UNIVERSITY OF CAMBRIDGE

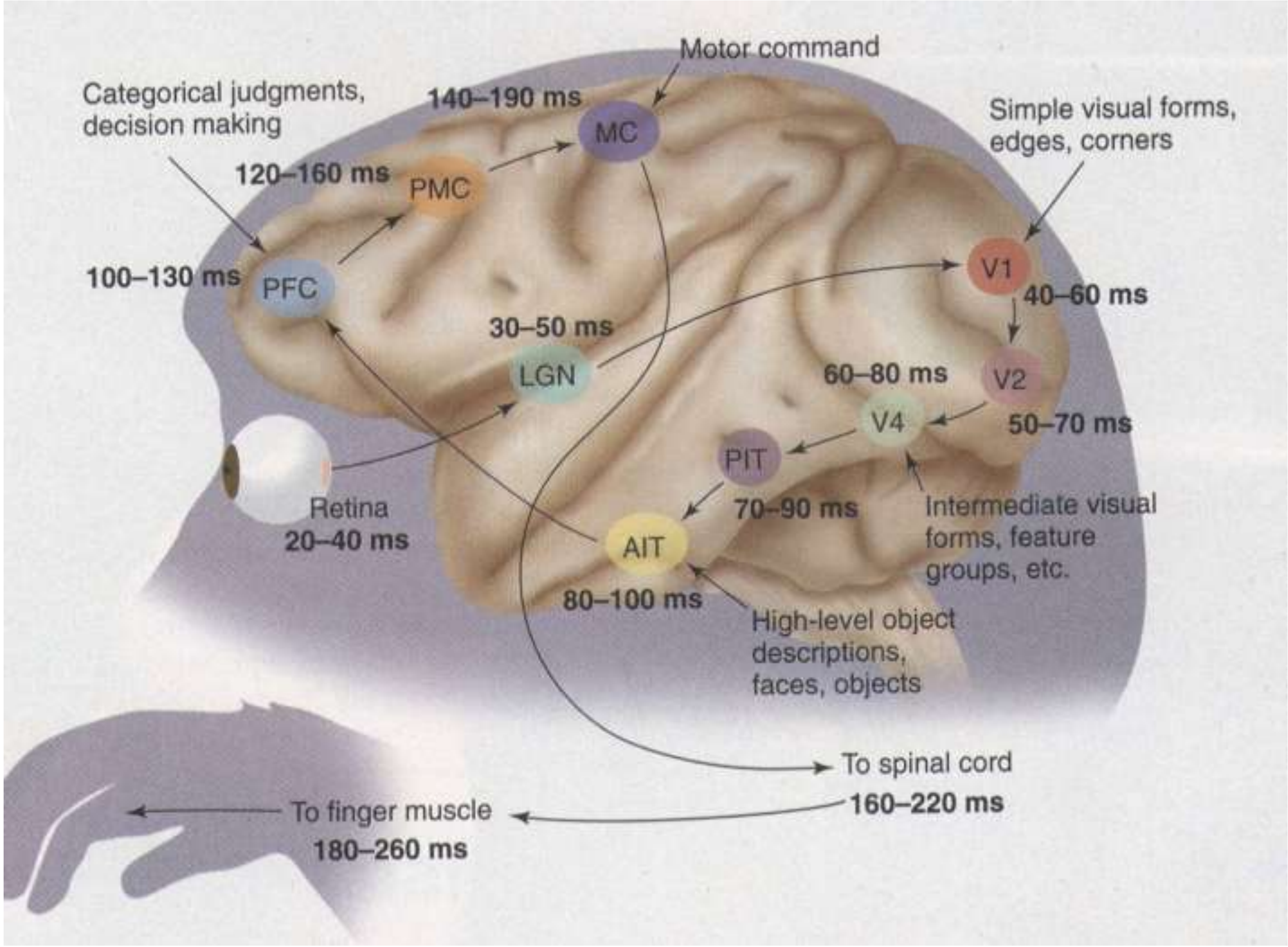
# Computer Vision at Cambridge


Roberto Cipolla
Department of Engineering


http://www.eng.cam.ac.uk/~cipolla/people.html

http://www.toshiba.eu/eu/Cambridge-Research-Laboratory/

# Vision: what is where by looking



# Computer Vision – What?

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 230 | 230 | 227 | 227 | 226 | 222 | 231 | 233 | 233 | 230 | 233 | 231 | 235 | 234 | 233 | 234 | 235 | 234 | 234 | 233 |
| 209 | 208 | 205 | 208 | 212 | 213 | 215 | 216 | 216 | 217 | 215 | 217 | 219 | 219 | 219 | 219 | 220 | 221 | 219 | 218 |
| 210 | 210 | 209 | 211 | 213 | 213 | 214 | 214 | 214 | 214 | 214 | 215 | 215 | 216 | 215 | 215 | 216 | 216 | 216 | 216 |
| 214 | 213 | 212 | 214 | 215 | 212 | 212 | 212 | 212 | 214 | 214 | 214 | 214 | 214 | 214 | 213 | 214 | 213 | 213 | 213 |
| 213 | 212 | 212 | 214 | 214 | 212 | 214 | 213 | 212 | 213 | 216 | 215 | 216 | 216 | 214 | 213 | 211 | 209 | 210 | 212 |
| 210 | 210 | 210 | 210 | 210 | 209 | 210 | 210 | 211 | 210 | 210 | 212 | 213 | 213 | 212 | 210 | 209 | 210 | 212 | 213 |
| 208 | 208 | 207 | 208 | 208 | 207 | 208 | 208 | 208 | 208 | 209 | 209 | 209 | 209 | 210 | 209 | 210 | 211 | 211 | 212 |
| 205 | 206 | 205 | 205 | 205 | 205 | 205 | 206 | 205 | 206 | 206 | 206 | 207 | 207 | 207 | 207 | 207 | 207 | 208 | 208 |
| 202 | 202 | 204 | 202 | 202 | 201 | 202 | 203 | 202 | 203 | 203 | 203 | 204 | 204 | 204 | 204 | 204 | 205 | 205 | 205 |
| 199 | 198 | 198 | 198 | 200 | 201 | 200 | 201 | 192 | 196 | 195 | 197 | 201 | 202 | 201 | 202 | 202 | 202 | 202 | 203 |
| 161 | 160 | 155 | 175 | 194 | 194 | 192 | 176 | 209 | 138 | 090 | 113 | 166 | 198 | 198 | 199 | 199 | 199 | 199 | 197 |
| 208 | 212 | 149 | 170 | 166 | 160 | 158 | 194 | 184 | 086 | 077 | 077 | 068 | 118 | 177 | 166 | 197 | 196 | 196 | 188 |
| 105 | 136 | 103 | 140 | 139 | 144 | 179 | 212 | 182 | 127 | 065 | 063 | 076 | 054 | 167 | 163 | 173 | 190 | 190 | 181 |
| 068 | 090 | 164 | 186 | 185 | 186 | 184 | 178 | 129 | 105 | 062 | 059 | 084 | 165 | 085 | 080 | 069 | 092 | 117 | 126 |
| 069 | 062 | 145 | 066 | 073 | 090 | 093 | 074 | 081 | 054 | 053 | 073 | 114 | 142 | 070 | 064 | 058 | 058 | 060 | 072 |
| 145 | 083 | 098 | 069 | 061 | 057 | 068 | 041 | 034 | 040 | 048 | 064 | 073 | 067 | 067 | 059 | 052 | 049 | 057 | 067 |
| 238 | 067 | 074 | 065 | 069 | 072 | 065 | 043 | 047 | 032 | 063 | 057 | 068 | 067 | 066 | 058 | 066 | 058 | 078 | 079 |
| 236 | 153 | 183 | 068 | 070 | 051 | 061 | 047 | 125 | 027 | 106 | 066 | 074 | 068 | 082 | 092 | 086 | 094 | 090 | 073 |
| 238 | 191 | 185 | 146 | 052 | 066 | 063 | 061 | 126 | 041 | 113 | 086 | 058 | 074 | 076 | 099 | 078 | 089 | 091 | 079 |
| 124 | 176 | 207 | 206 | 188 | 161 | 117 | 185 | 100 | 082 | 126 | 182 | 211 | 210 | 218 | 205 | 210 | 205 | 199 | 193 |
| 091 | 093 | 091 | 092 | 091 | 093 | 095 | 094 | 095 | 133 | 127 | 120 | 102 | 101 | 099 | 096 | 096 | 096 | 095 | 099 |
| 097 | 096 | 093 | 094 | 093 | 091 | 099 | 086 | 077 | 089 | 087 | 121 | 086 | 084 | 078 | 081 | 082 | 081 | 080 | 080 |
| 091 | 089 | 086 | 088 | 090 | 087 | 096 | 088 | 059 | 086 | 070 | 117 | 086 | 087 | 084 | 084 | 082 | 083 | 082 | 089 |
| 101 | 091 | 088 | 087 | 091 | 099 | 084 | 083 | 073 | 095 | 076 | 111 | 090 | 088 | 084 | 084 | 083 | 082 | 079 | 078 |
| 102 | 114 | 103 | 087 | 081 | 080 | 080 | 086 | 094 | 098 | 079 | 112 | 085 | 082 | 081 | 078 | 078 | 076 | 072 | 074 |
| 089 | 093 | 092 | 083 | 078 | 062 | 053 | 053 | 108 | 074 | 076 | 084 | 082 | 080 | 081 | 076 | 077 | 075 | 073 | 075 |
| 099 | 094 | 089 | 087 | 091 | 086 | 082 | 079 | 149 | 075 | 064 | 067 | 101 | 081 | 080 | 076 | 076 | 073 | 074 | 074 |
| 096 | 092 | 088 | 087 | 088 | 096 | 085 | 076 | 078 | 083 | 083 | 081 | 076 | 079 | 078 | 081 | 077 | 077 | 075 | 074 |
| 105 | 104 | 090 | 096 | 084 | 082 | 080 | 078 | 081 | 079 | 078 | 080 | 079 | 084 | 079 | 080 | 079 | 084 | 081 | 082 |
| 086 | 090 | 079 | 093 | 090 | 091 | 106 | 083 | 076 | 083 | 073 | 084 | 081 | 081 | 081 | 084 | 081 | 074 | 071 | 072 |

# Overview

UNIVERSITY OF CAMBRIDGE

1. Background: why and how?

2. 3R's of Computer Vision:
   - Registration
   - Reconstruction
   - Recognition

3. Algorithms and Applications

# Why? Images and Video

UNIVERSITY OF CAMBRIDGE

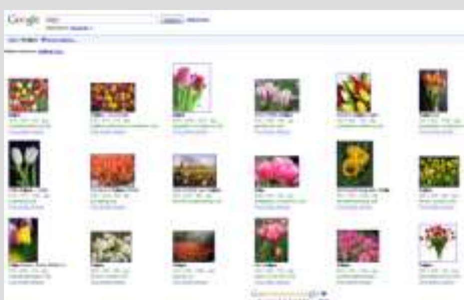
Computer vision is now in a wide range of products

**Mobile phones**

**Cars**

**Inspection and measurement**

**Image and video search**



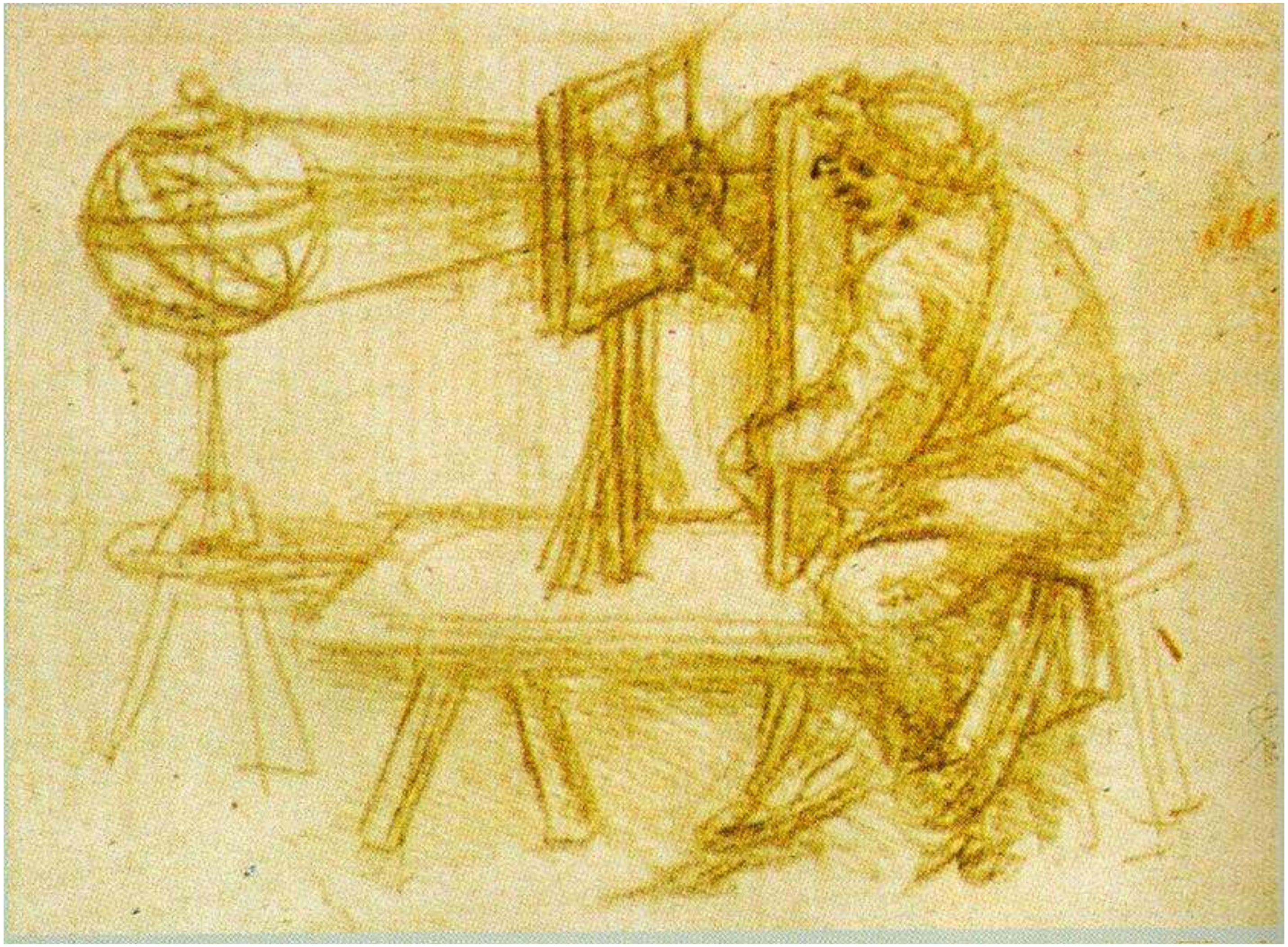
**Games**

**Internet and shopping**

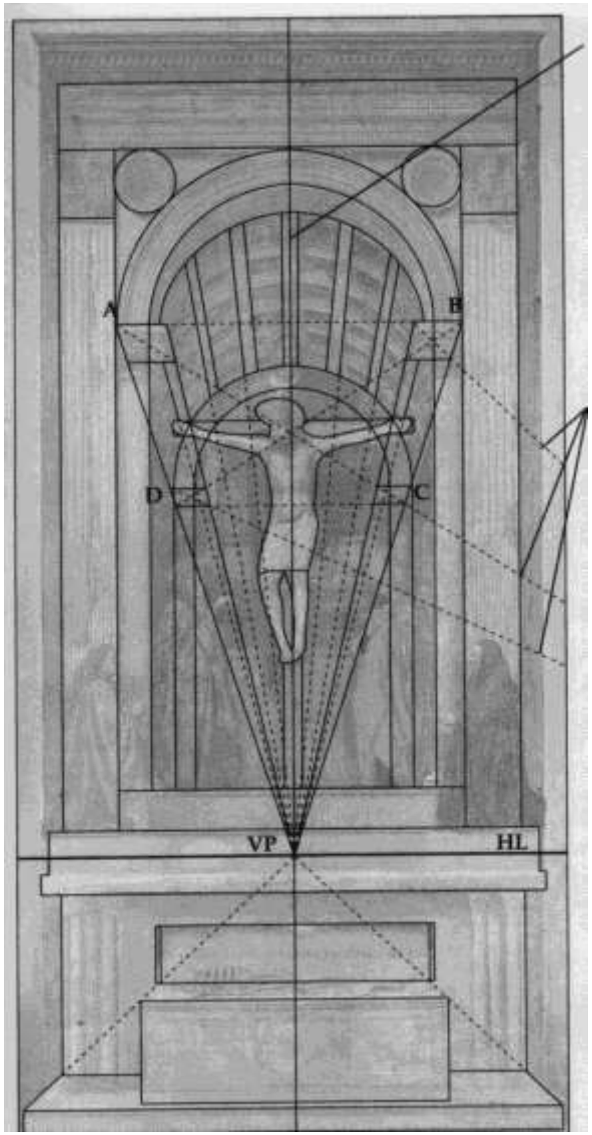
# 1. How to make machines that see?

# How?

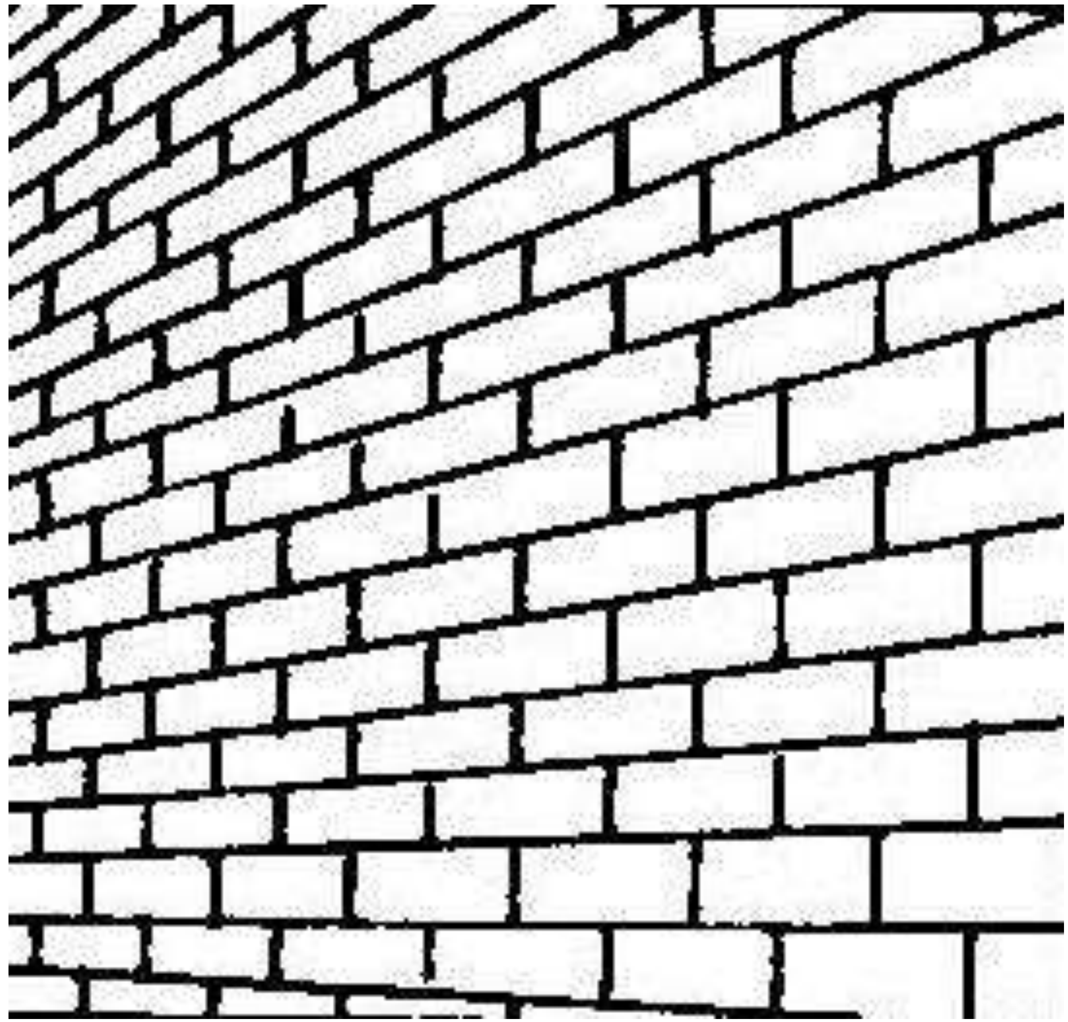
# 1 Geometry - Perspective



# Geometry - Perspective

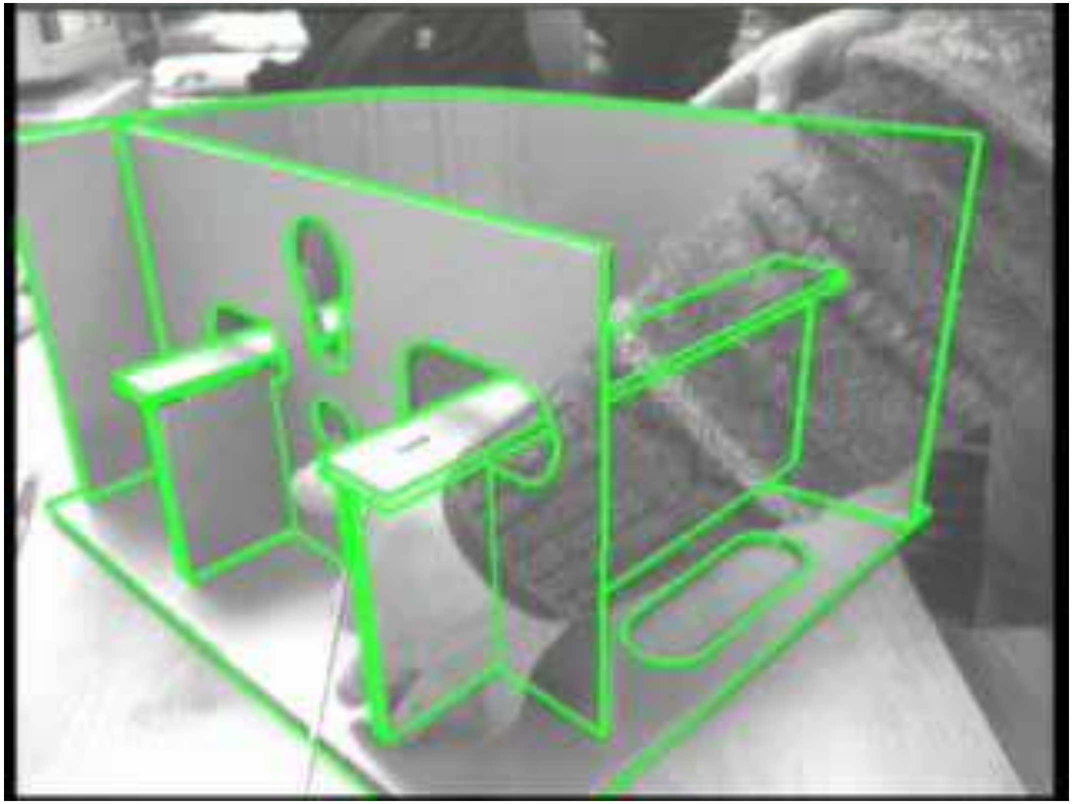


# Geometry - Transformations



# Time to contact

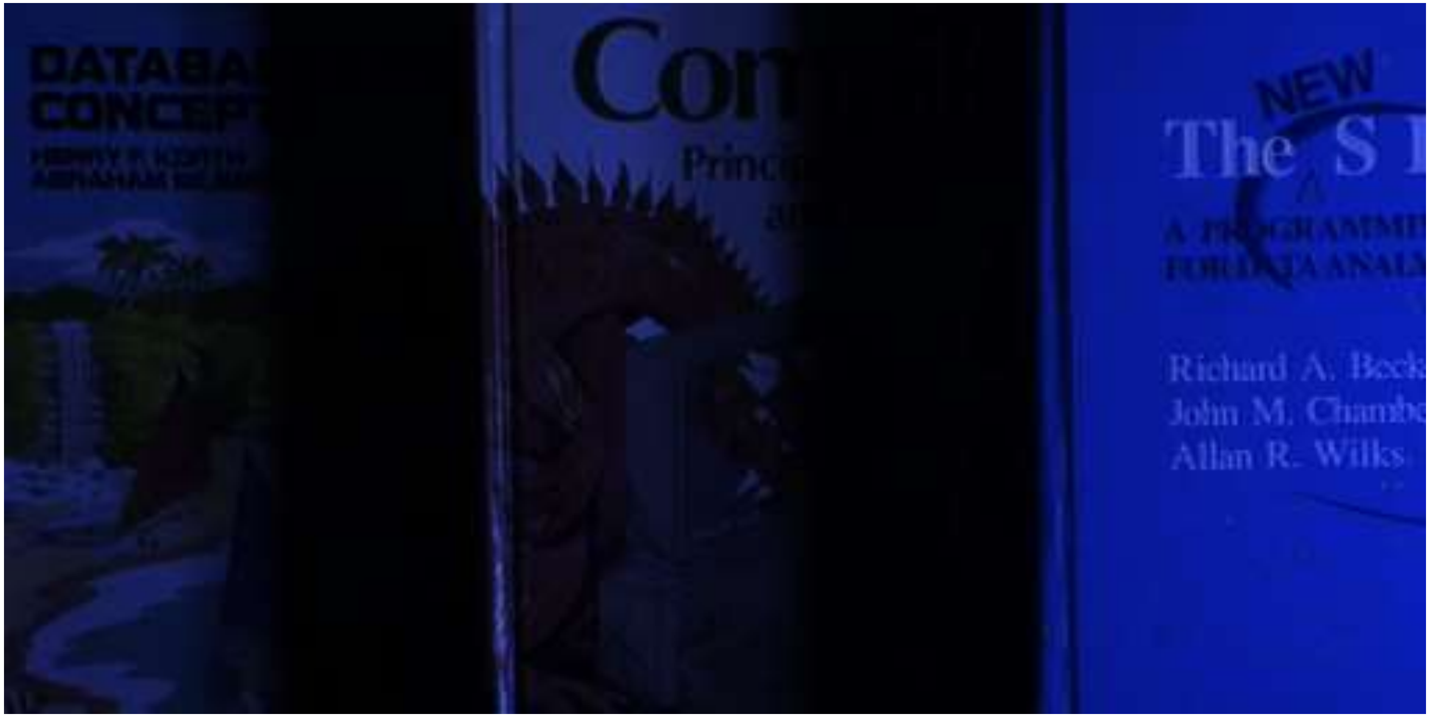
# Geometry – 3D shape



# Geometry – Shape

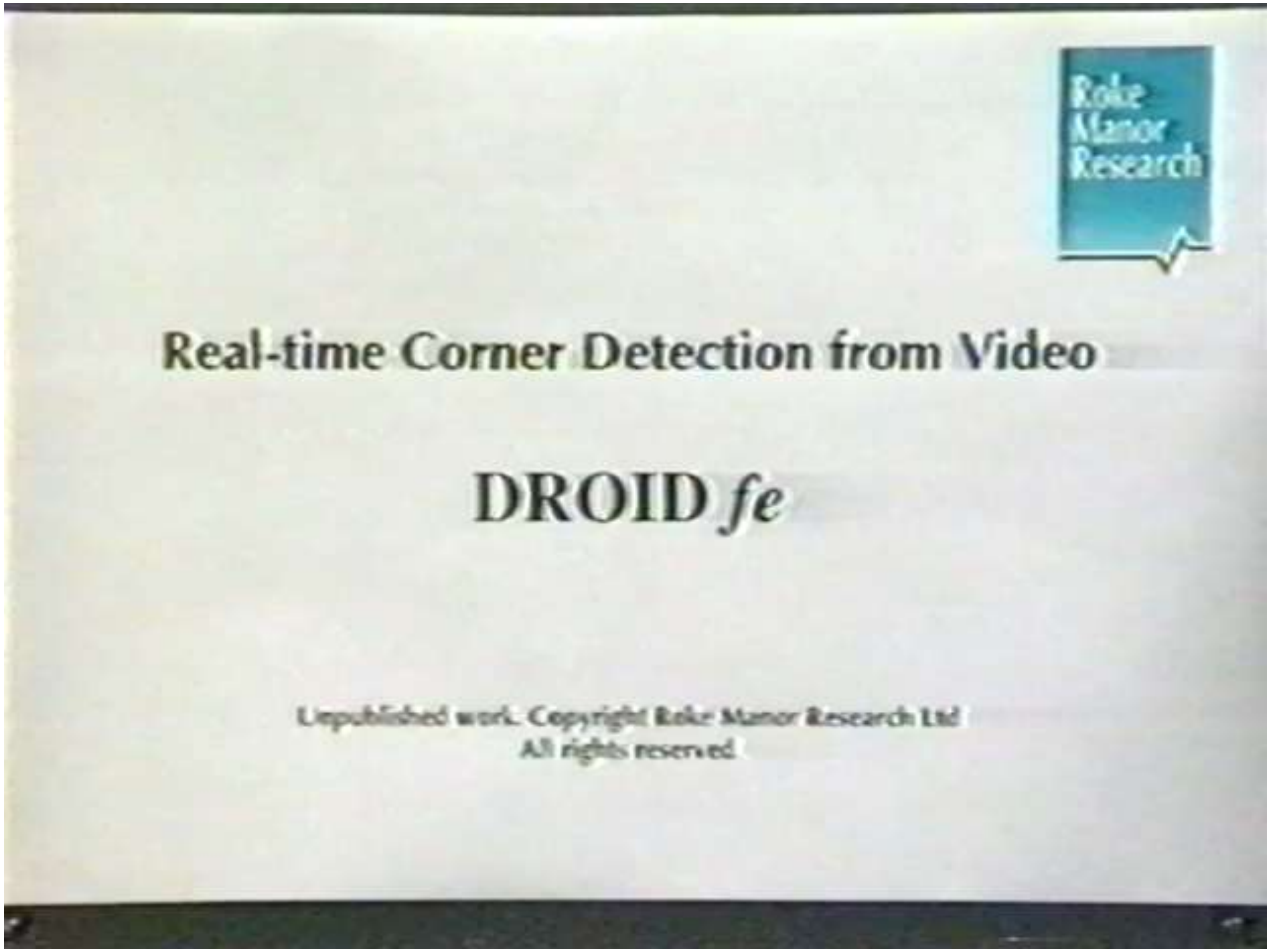
UNIVERSITY OF CAMBRIDGE

# 2 Photometric Features

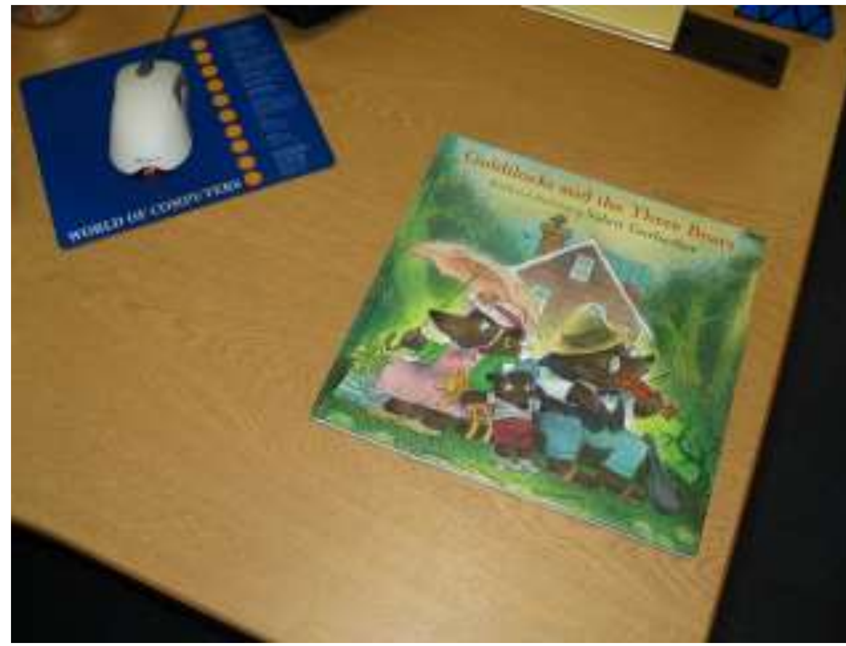


# Raw pixels - Invariance?

# Image features



# Matching – "visual words"

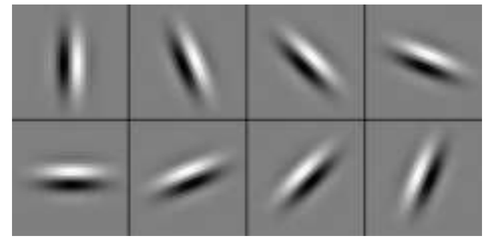


[Feature extraction and matching demo]

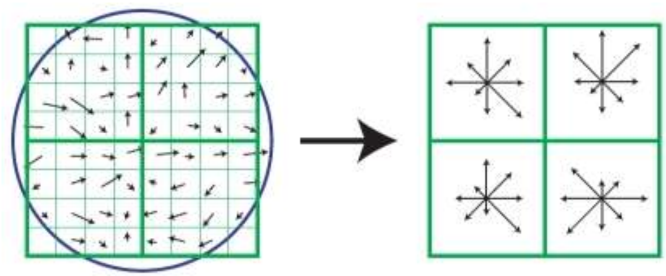
# SIFT Descriptor

Image Pixels

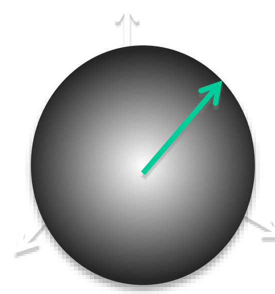
**Filtering**



**Spatial pool (Sum)**

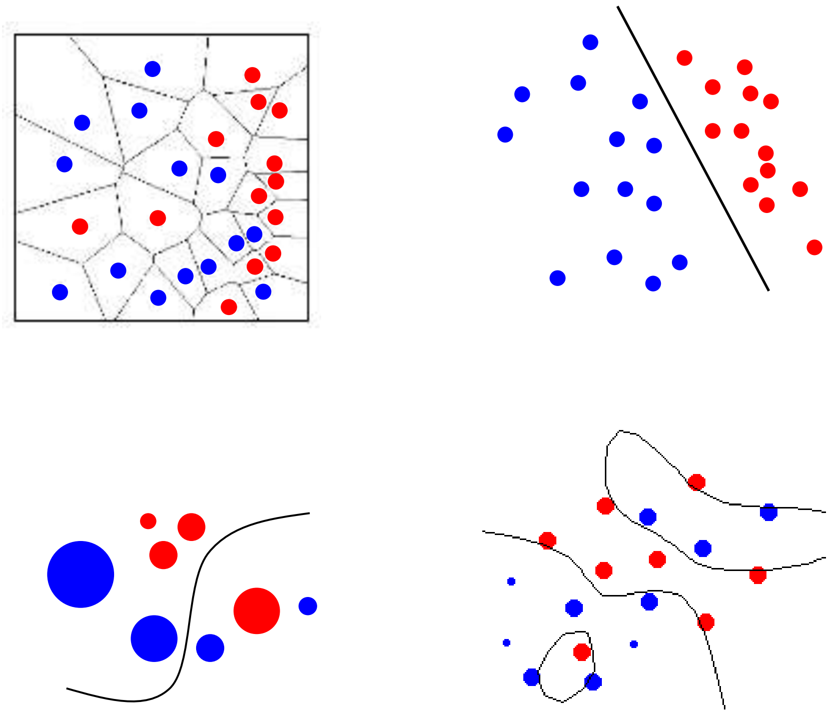


**Local Contrast Normalisation (LCN)**

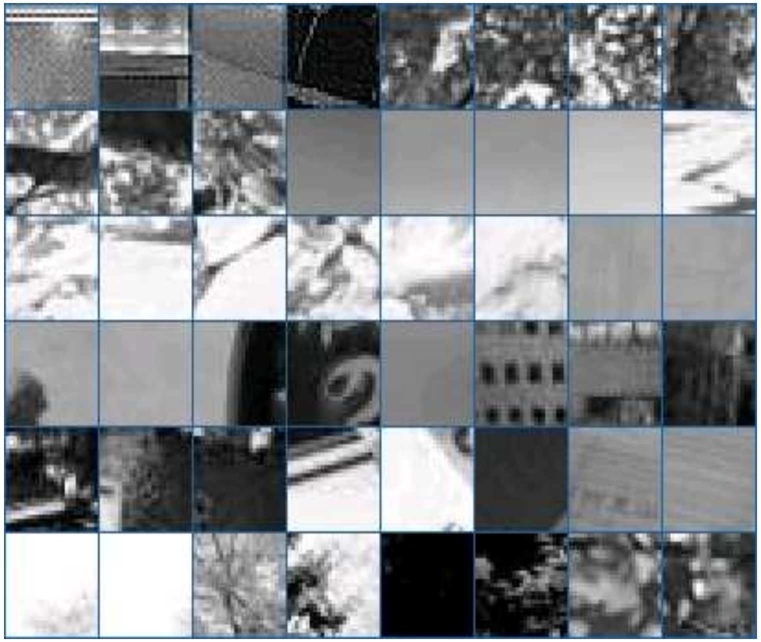


Feature Vector

# 3 Machine Learning



# Training data – supervised learning



# Real-time classifiers - Hand detection and tracking

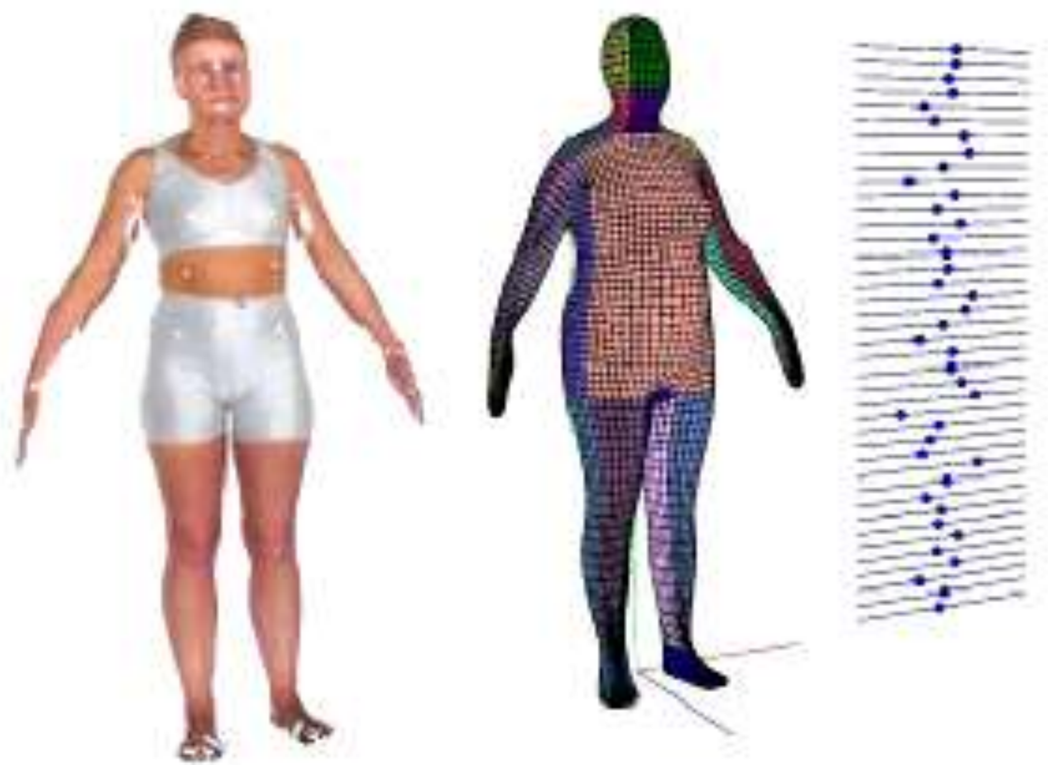
# Real-time classifiers - Hand detection and tracking

# 4 Dealing with ambiguity

Ames (1946) Room

# Dealing with incomplete sensory data

TOSHIBA
Leading Innovation >>>



Helmholtz (1866)-
"Perception is our best guess as to what is in the world, given our current sensory input and our prior experience."

# Probabilistic framework

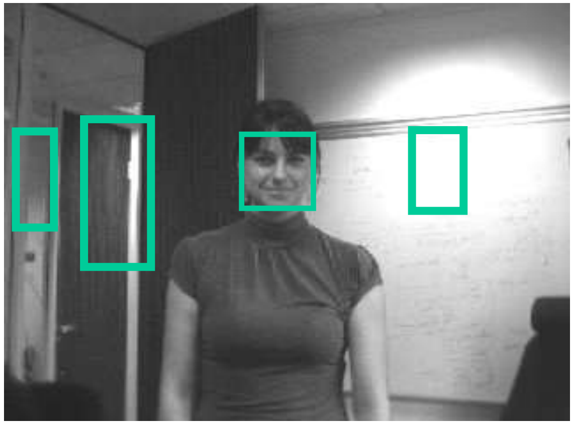
Probabilistic framework to understanding vision and for building systems:

1. Deal with the ambiguity of the visual world
2. Are able to fuse information
3. Have the ability to learn

# 2 Computer Vision at Cambridge

# Computer Vision: 3R's

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



**Reconstruction: Recover 3D shape**

**Recognition: Identify objects** (example)

**Registration: Compute their position and pose**

# Registration?

# Target detection and pose estimation

# 2D Registration

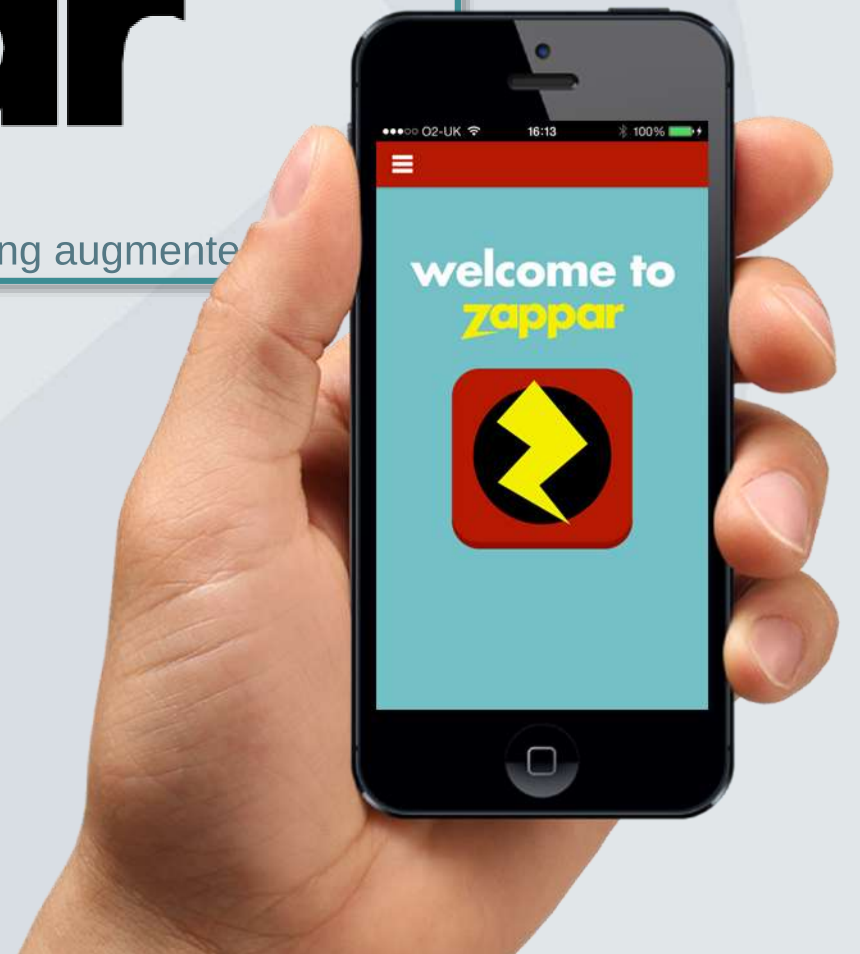
zappar
You're seeing things

zappar

1. An app

2. Using augmente

3. To 'zap' your surroundings



# With Zappar everything communicates

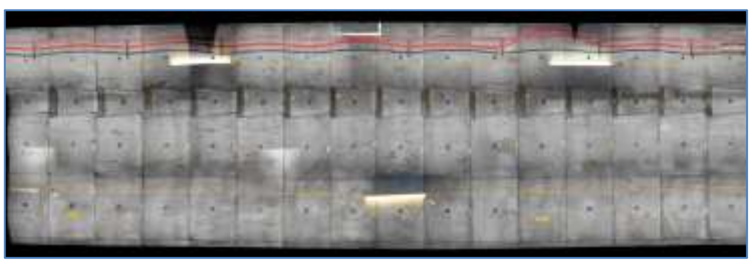

zappar
You're seeing things

# The Zappar Rush’™

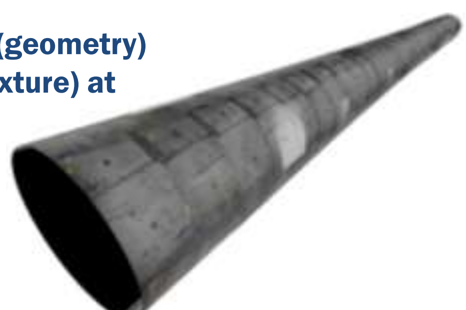

“Zapparと呼ばれるこの偉大な新しいアプリがあります！”

# Ageing Infrastructure – visual inspection

OFFLINE

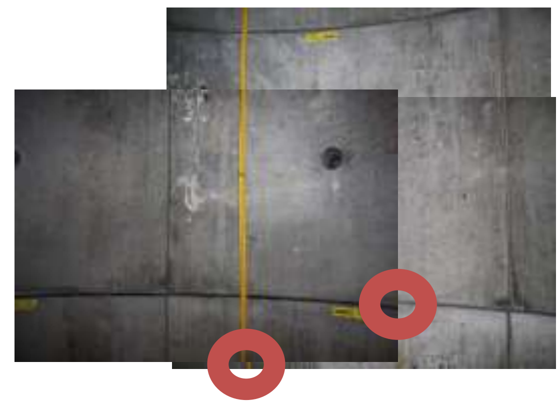


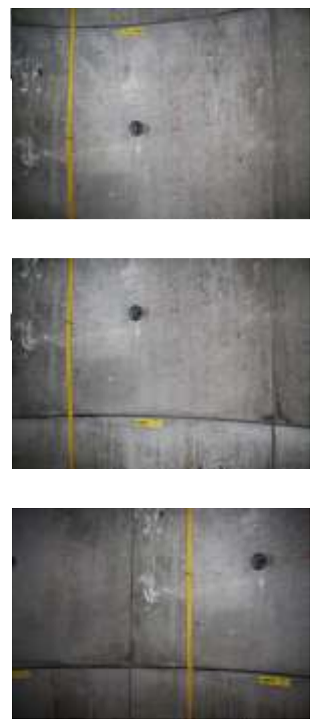
**Database of shape (geometry) and appearance (texture) at time $t_0$**

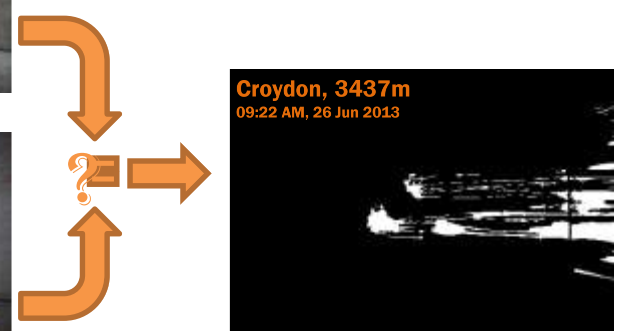


National Rail
nationalgrid
Transport for London

ONLINE







# Visual inspection – demo

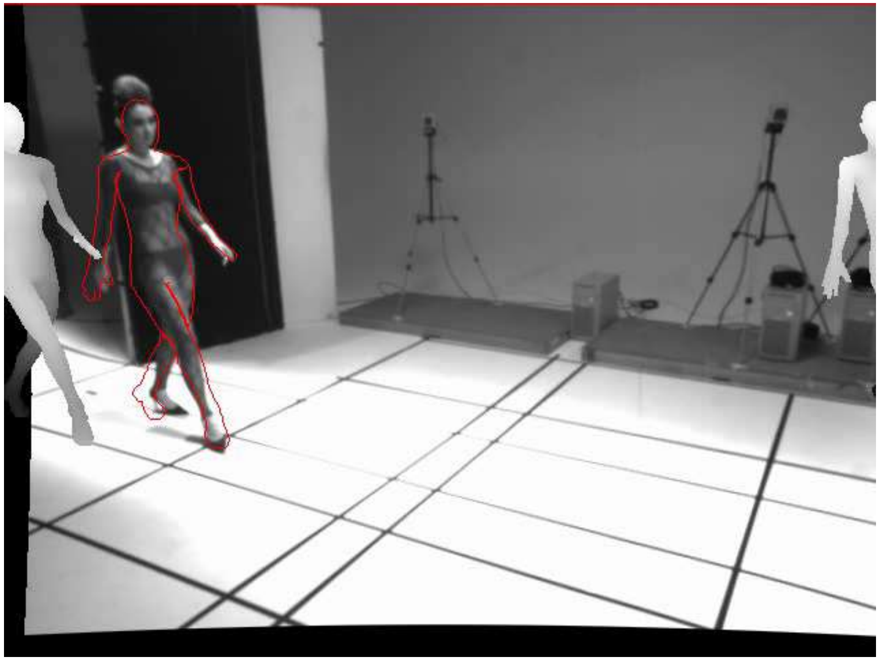
# Visual inspection

# 3D Registration - Magic Mirrors

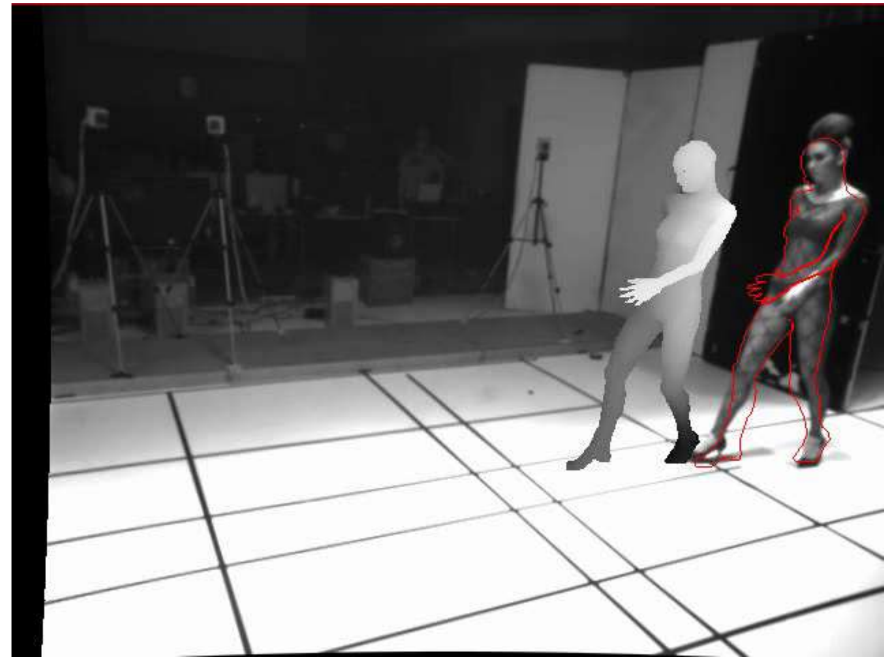
# Body tracking with 3D model

- Cloth simulation and rendering

# Tracking body in 3D



View 1

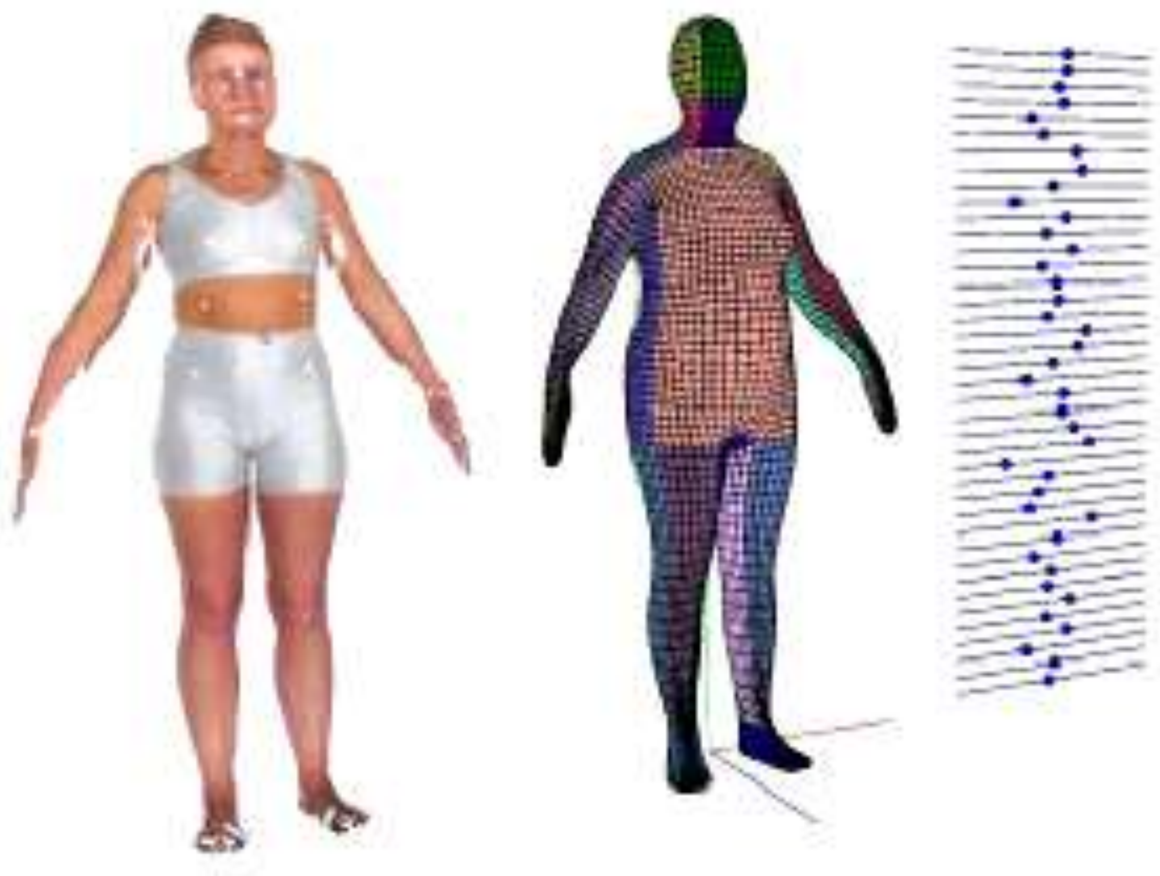


View 2

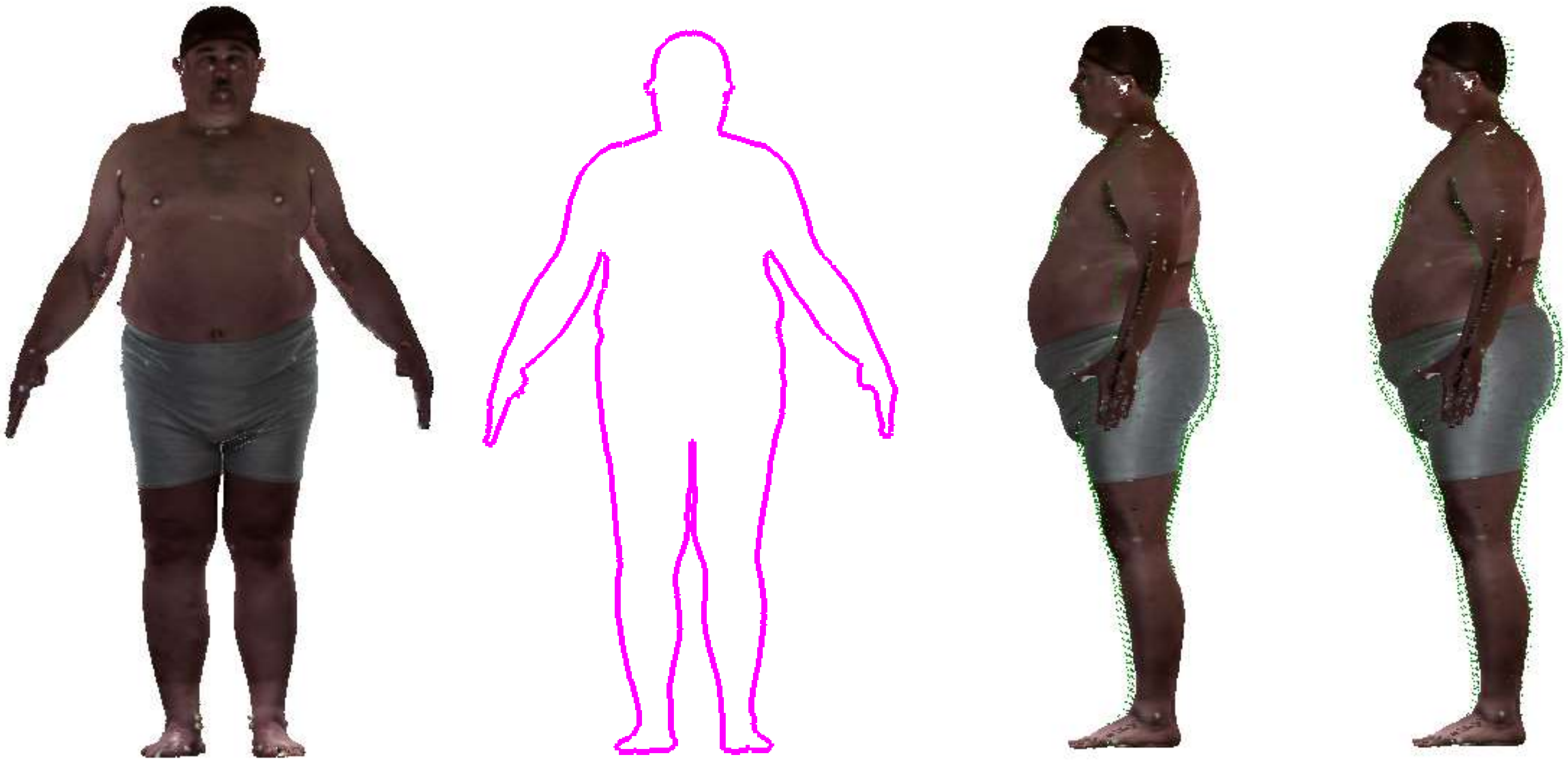
# Virtual Fashion Show

# Registration – Body shape

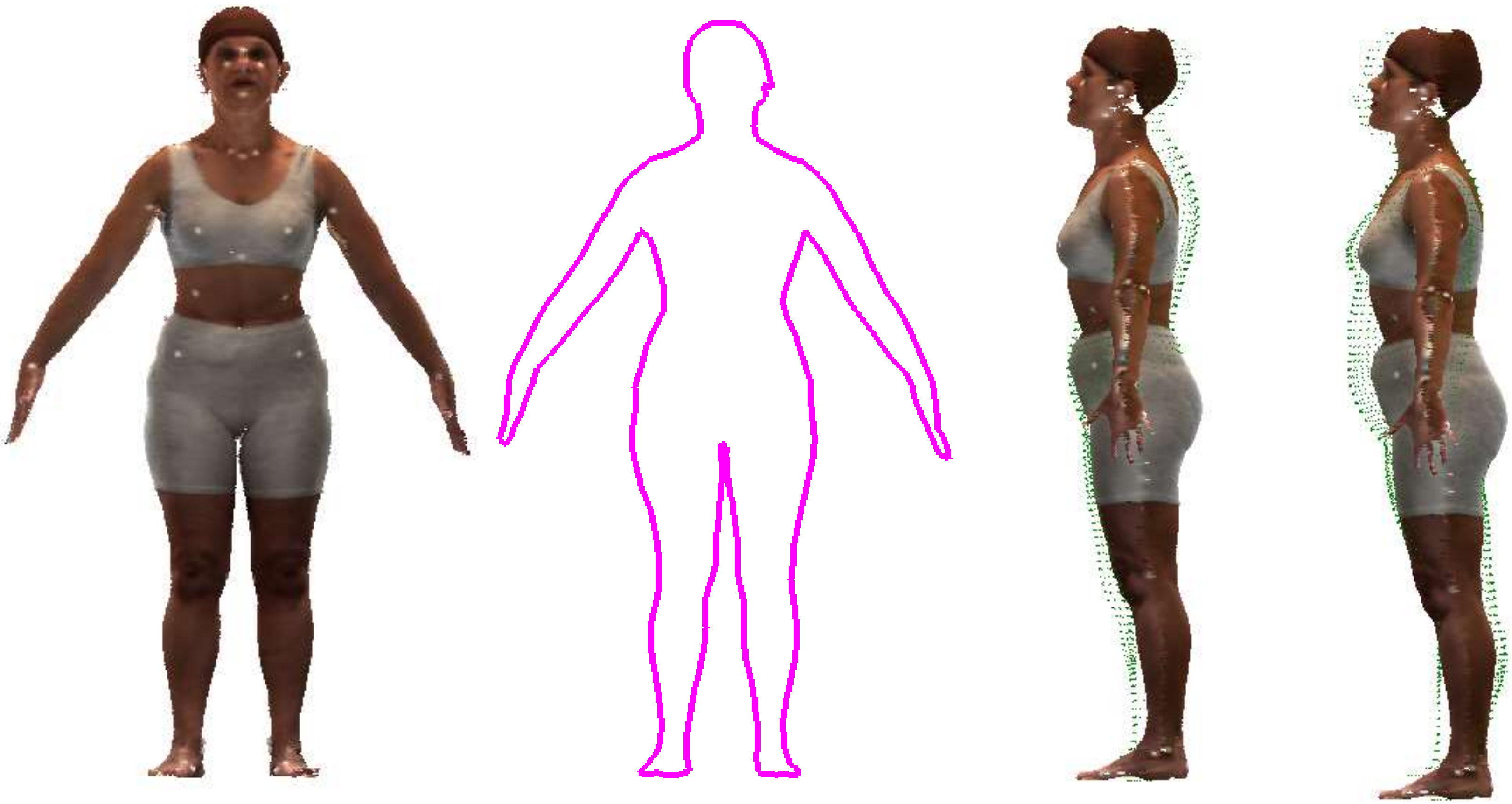
TOSHIBA
Leading Innovation >>>



# Single view - Human Body Data

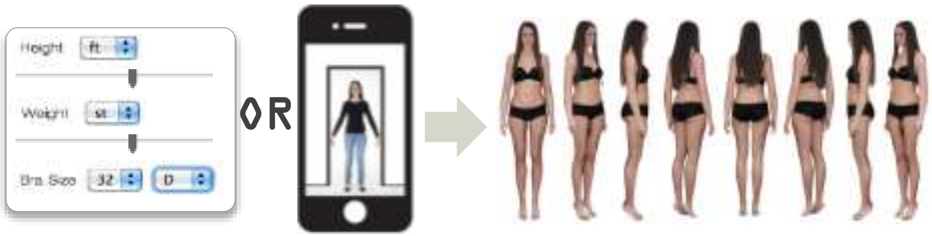


# Single view - Human Body Data



# METAIL ALLOWS USERS TO SEE HOW GARMENTS WOULD LOOK AND FIT ON THEIR BODY SHAPE

QUICKLY MODEL YOUR BODY




A database of thousands of 3D scans used to predict user's 3D body shape from either a few measurements entered or deduced from a single photo

MODEL YOUR FACE

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

METAIL DIGITISES GARMENTS IN 3D

3D garment photography has been developed into a fast and cheap process - taking 10 minutes per garment and integrated into the retailers' existing processes

THE RESULT
AN ACCURATE VISUALISATION OF SIZE AND FIT


Our proprietary software models the fit of the garment on to the customer's specific model, giving them confidence in size selection

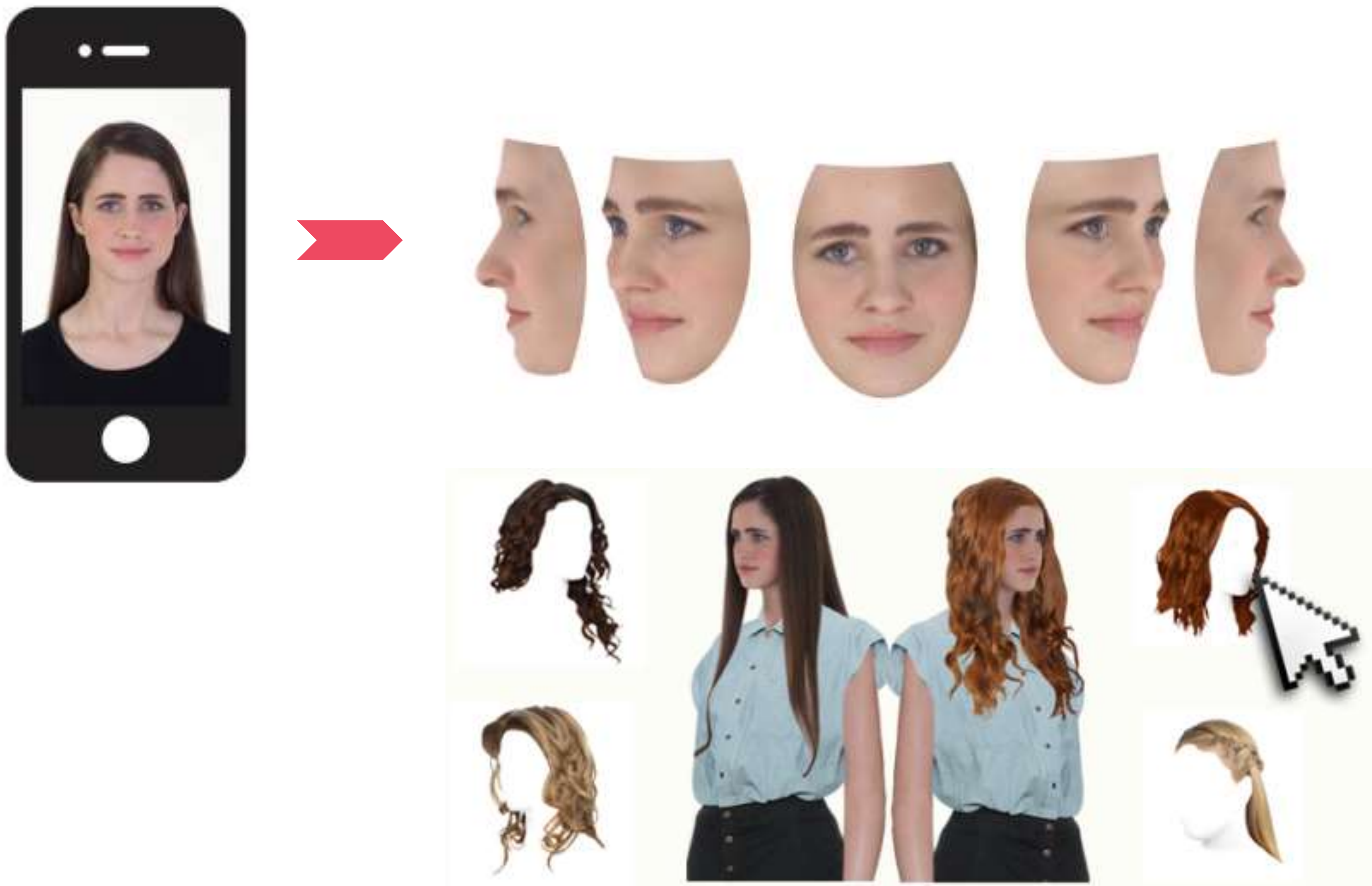
# METAIL CAN ACCURATELY MODEL YOUR BODY FROM A SINGLE PHOTO OR A FEW MEASUREMENTS

A database of scans of human bodies allows users' 3D body shapes to be predicted using a few measurements or a single photo


***or***

# FACE FROM PHOTO

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



# USP: ACCURATE VISUALISATION OF STYLE AND FIT

Our proprietary software models the fit of the garment on to the customer's specific model, giving them confidence in size selection

We recommend

12

Bust: *good fit*

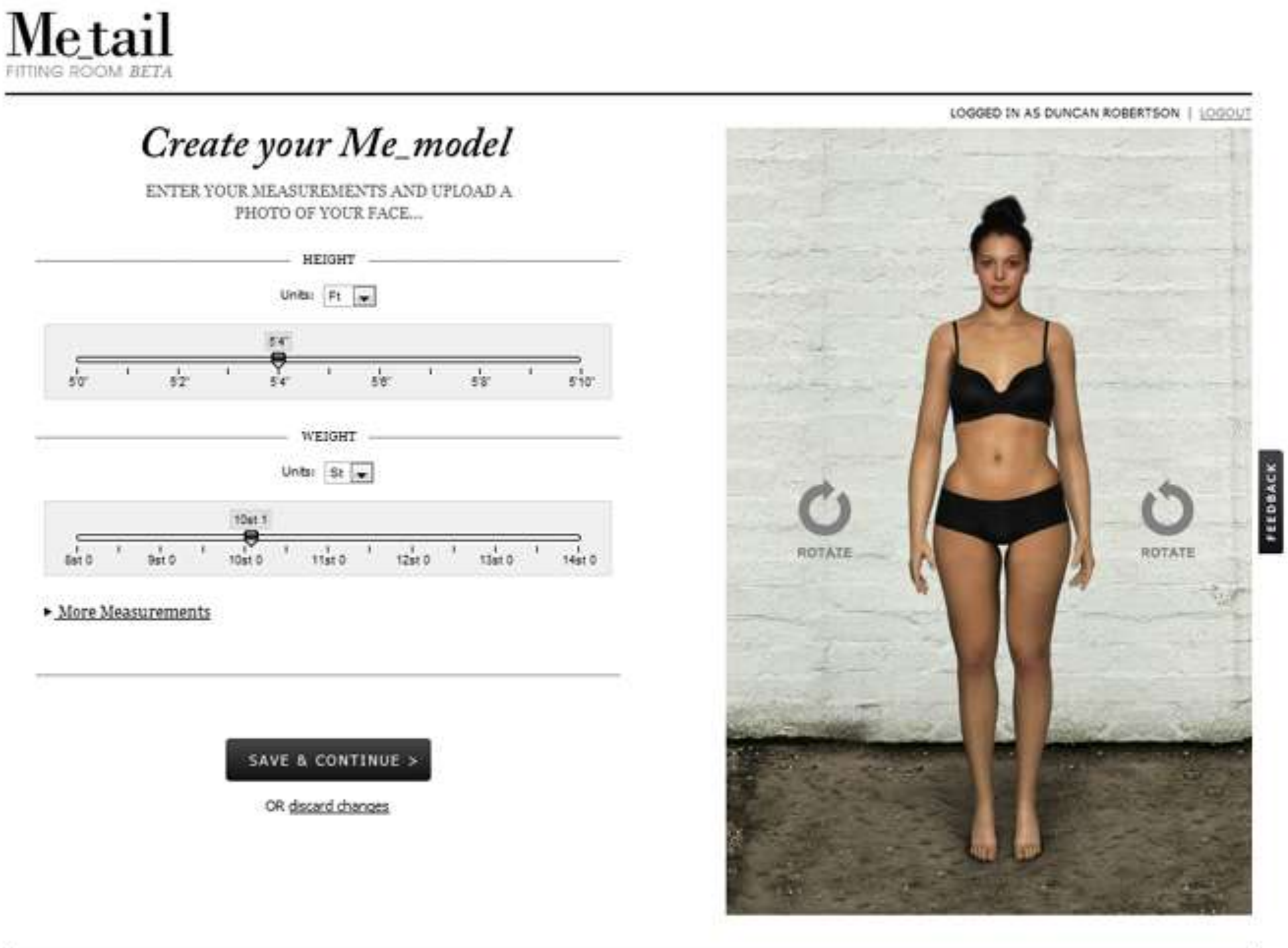
Waist: *tight fit*

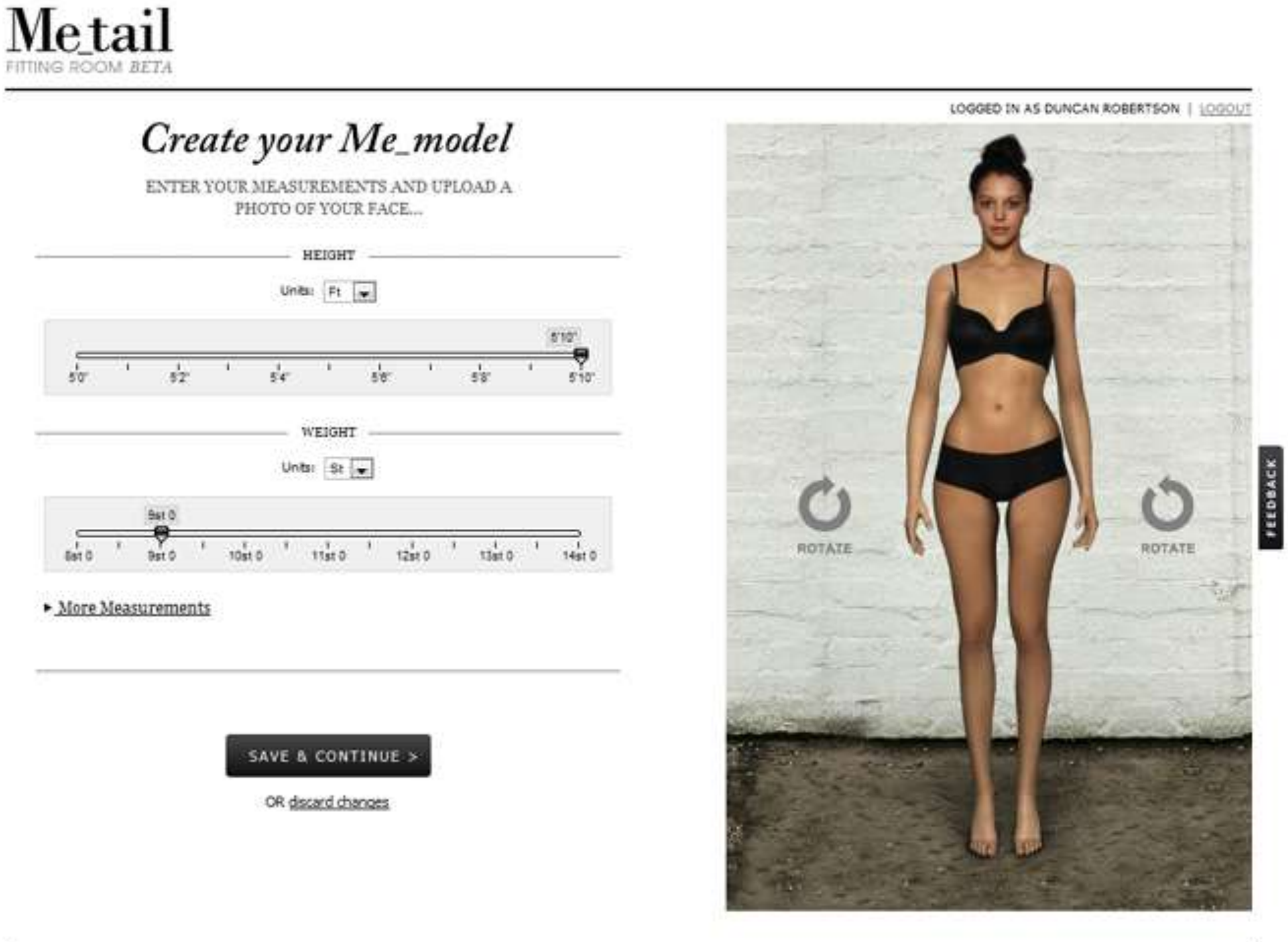
Hips: *good fit*

Daily Mail

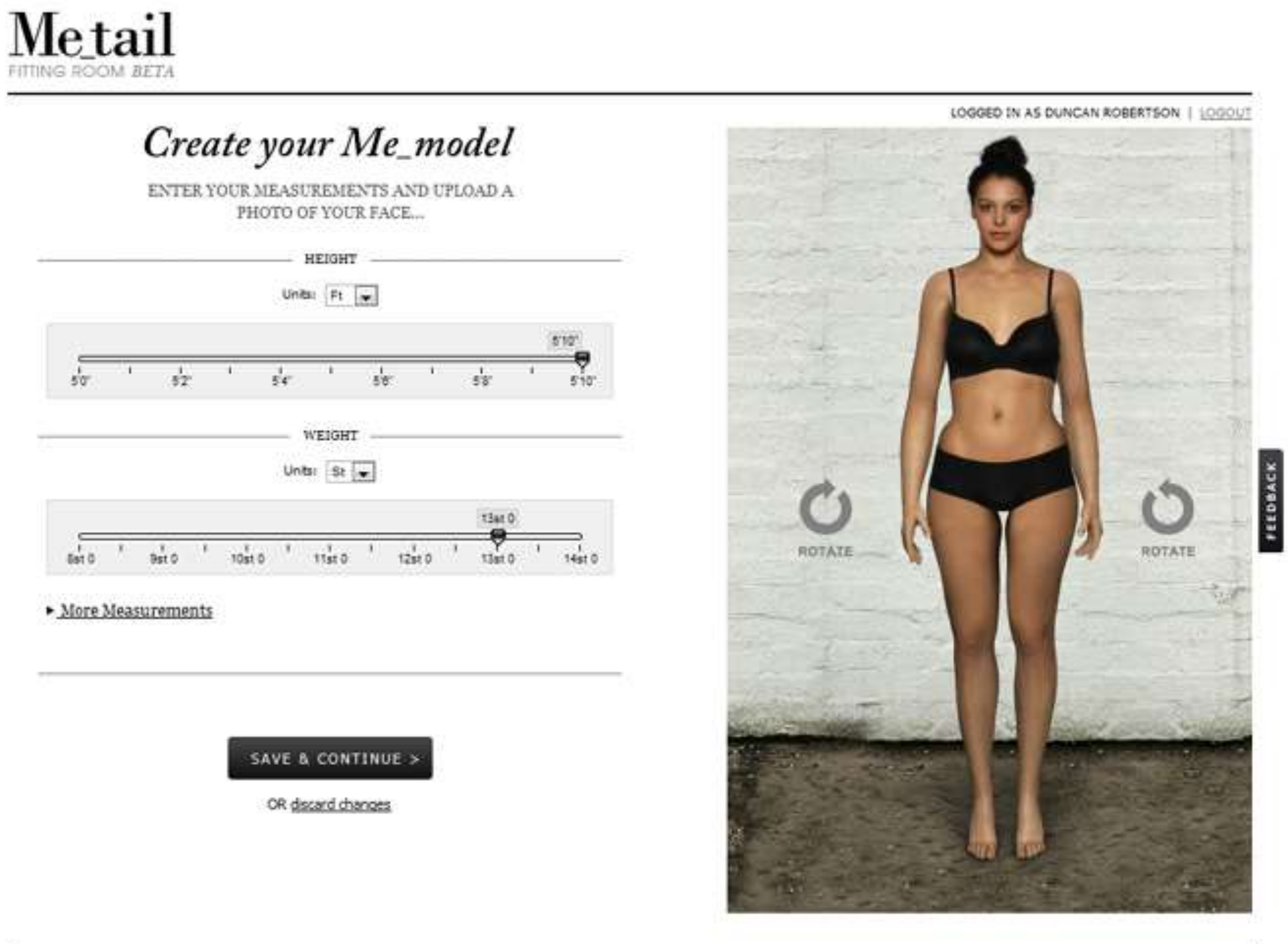
*"I ordered it and a few days later was trying on the real deal. It couldn't have fitted better. I take my hat off ... If this is the way online shopping is going I might just be on board."*

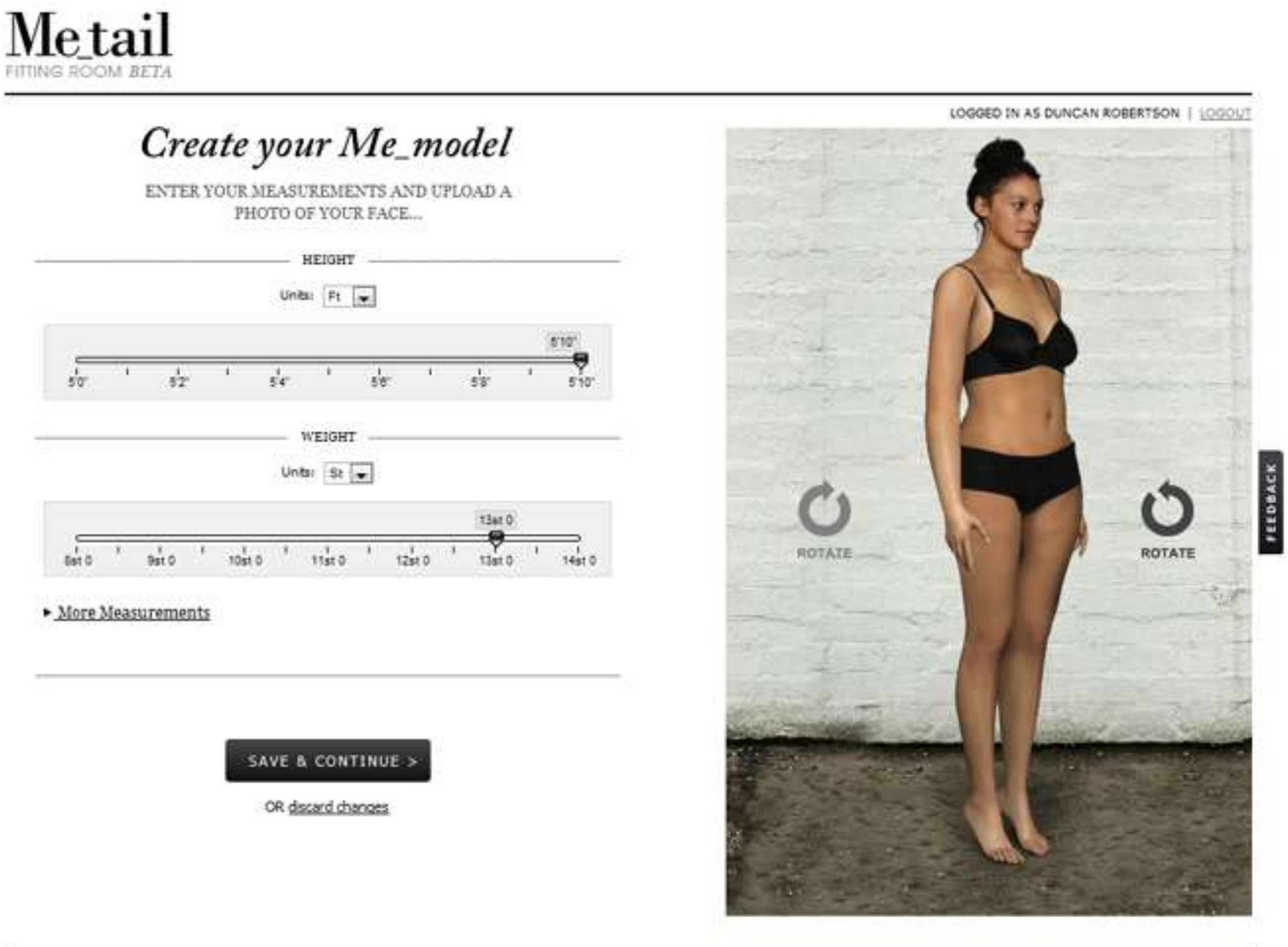
# BODY MODEL ACCURACY - DEMO

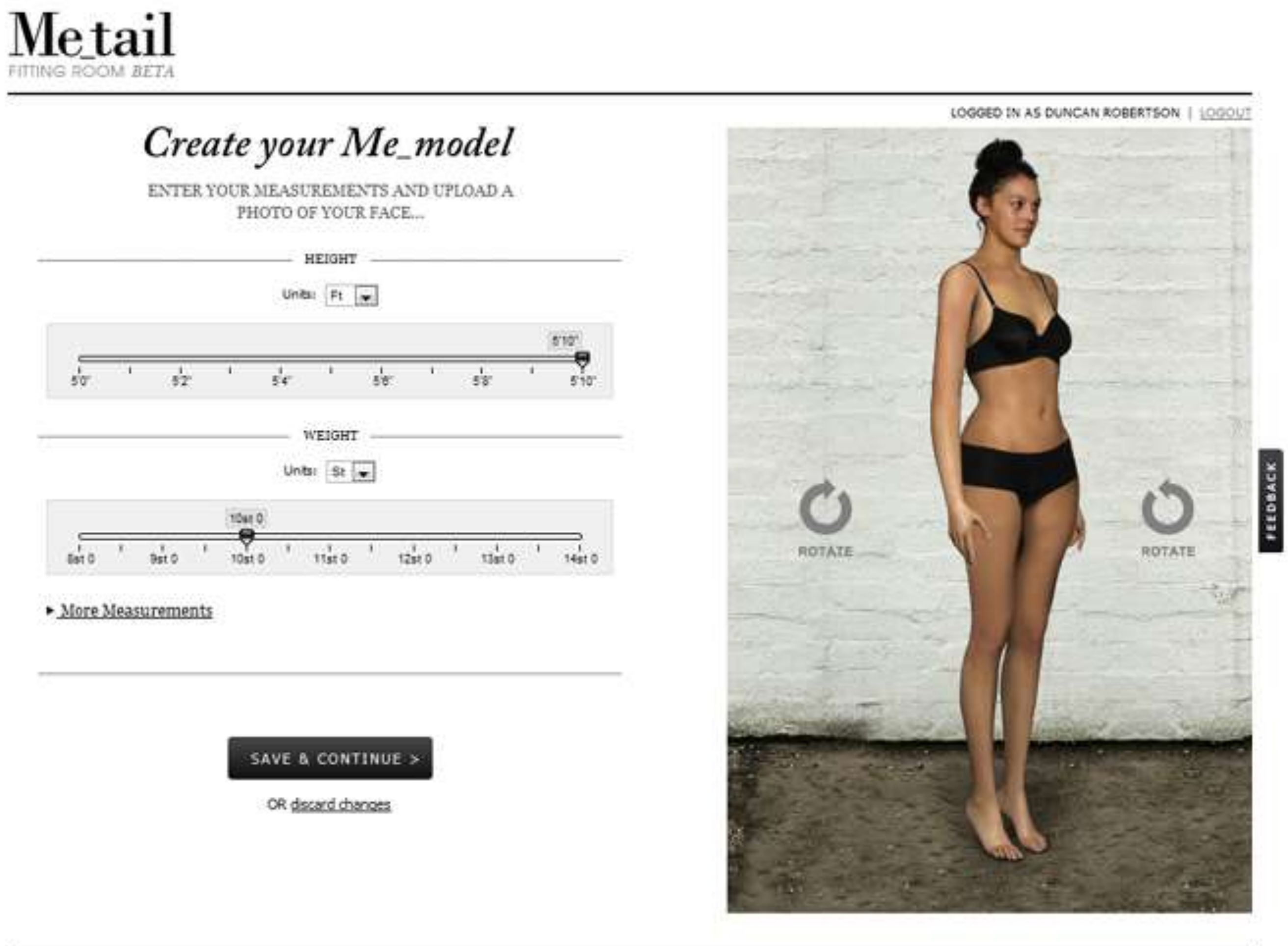

Me_tail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'4"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
10st 1
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE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

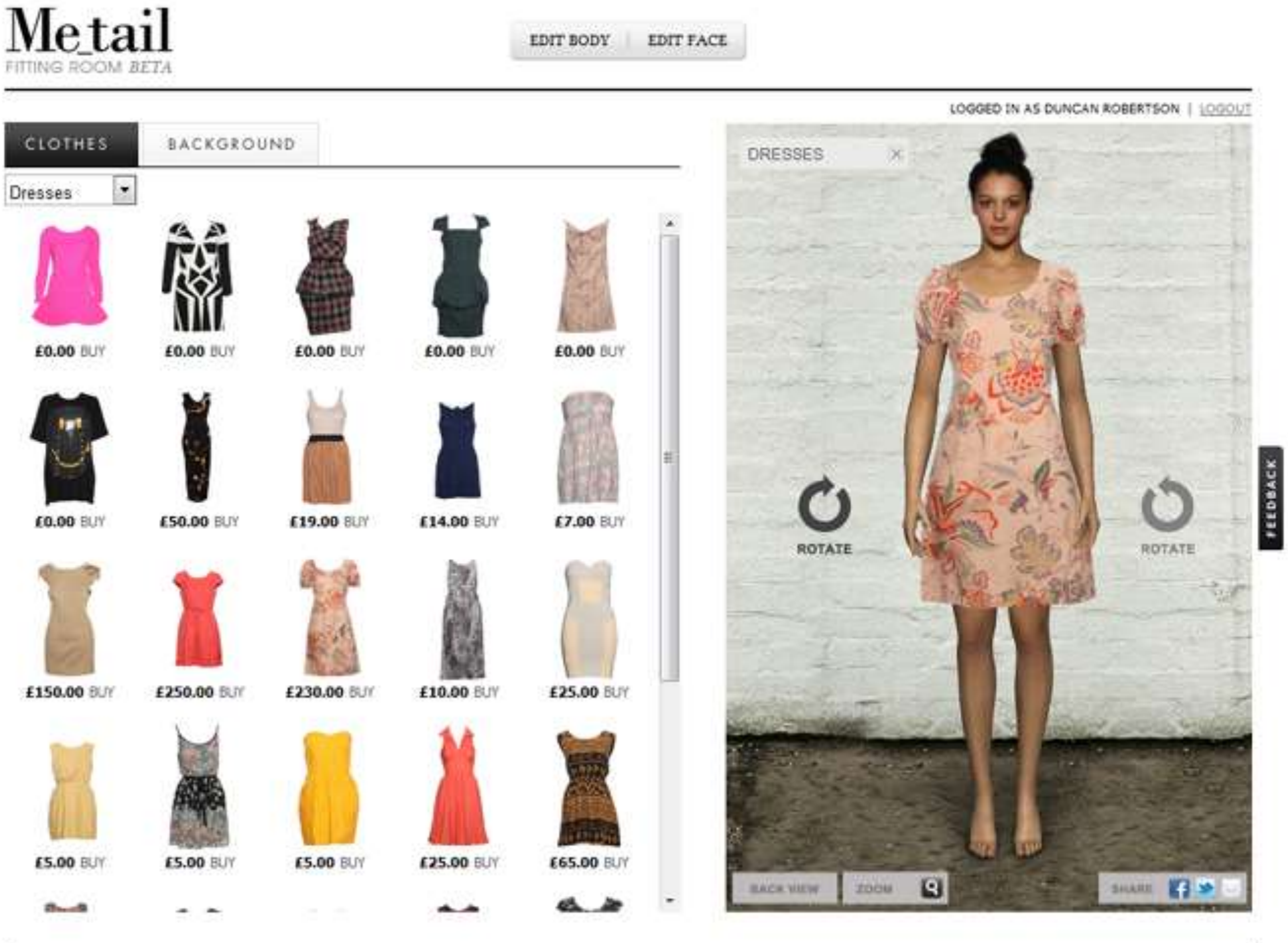


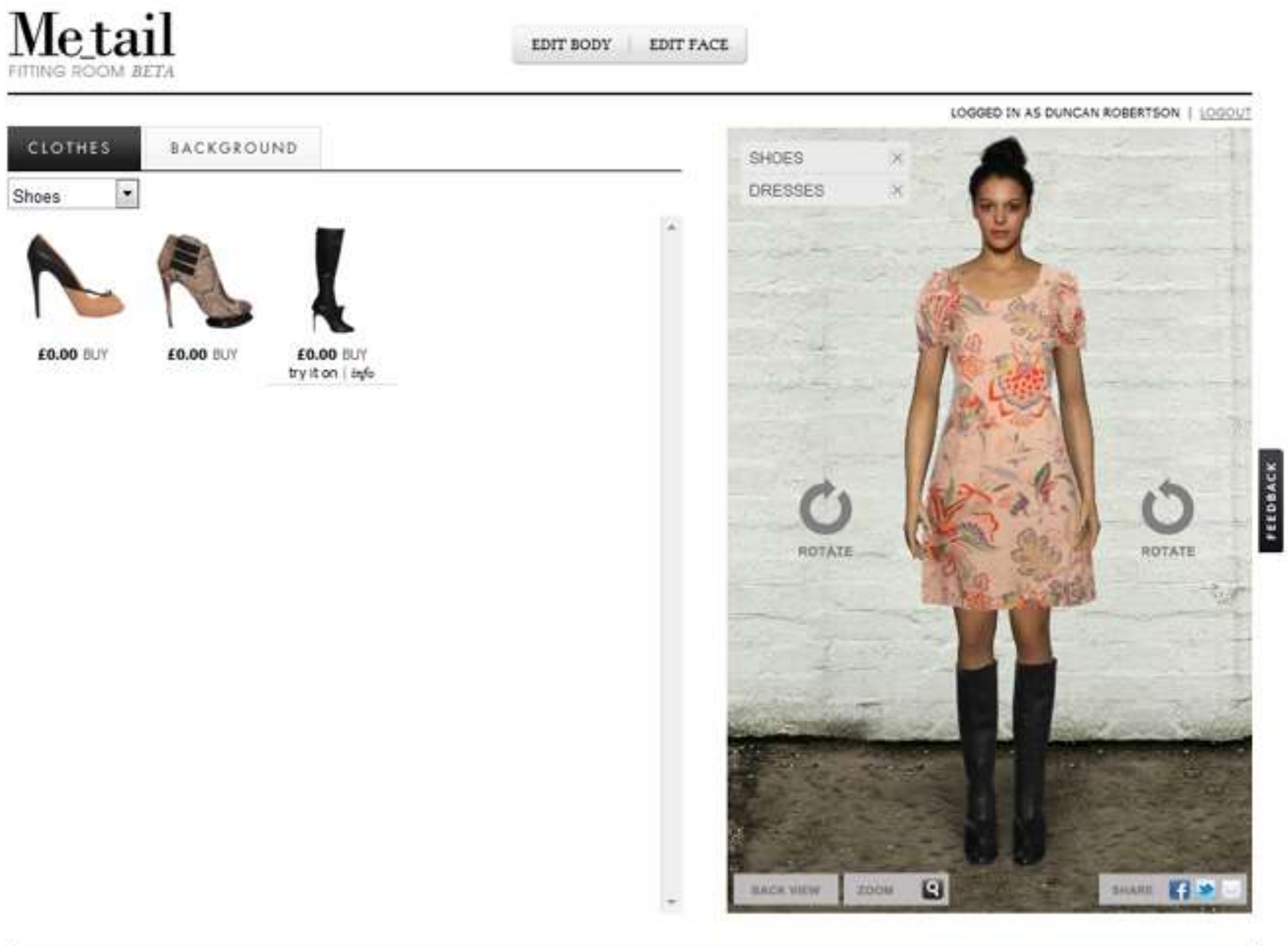
Me_tail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'10"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
9st 0
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE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

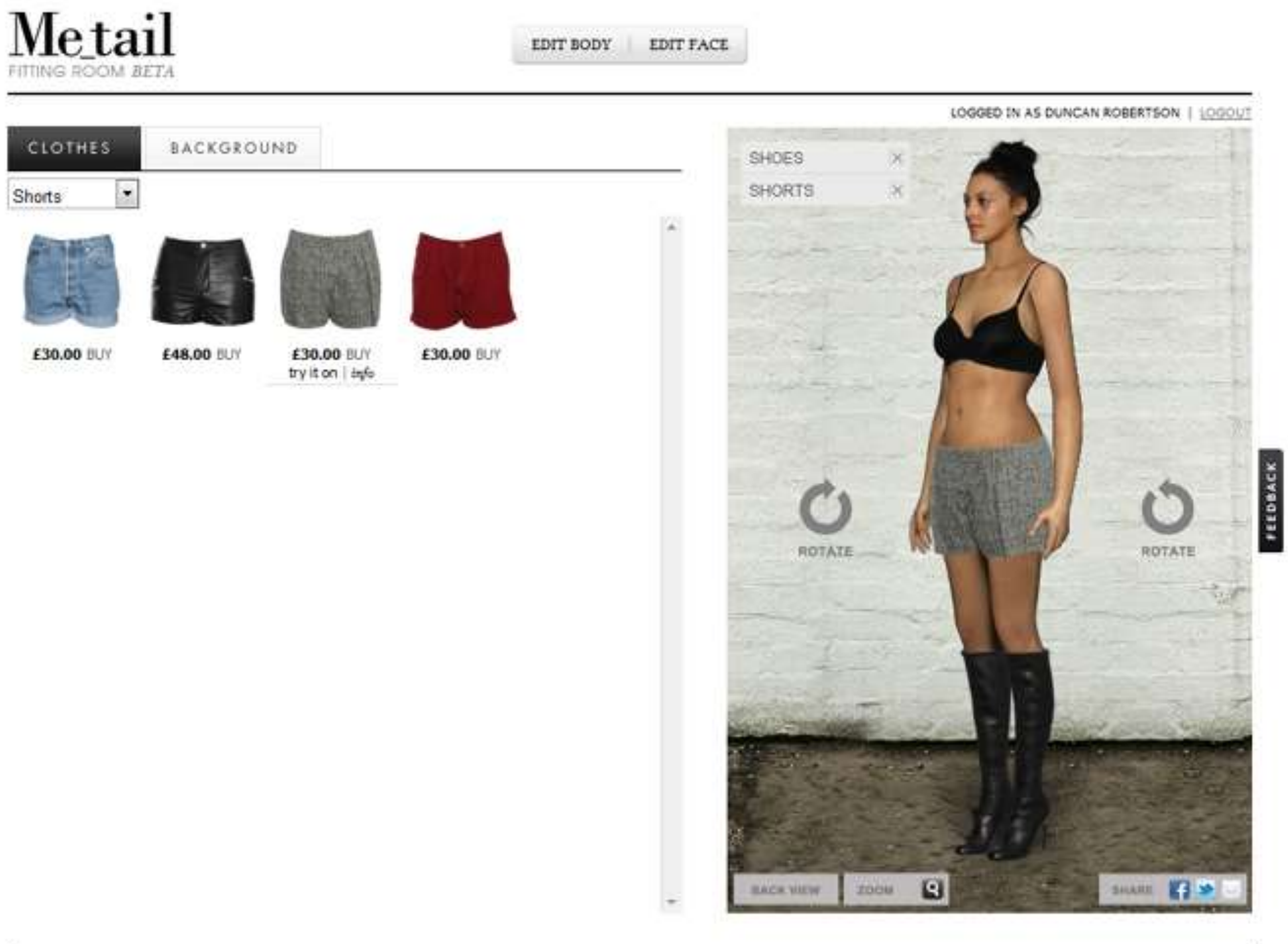

Metail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'10"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
13st 0
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE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


Metail
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

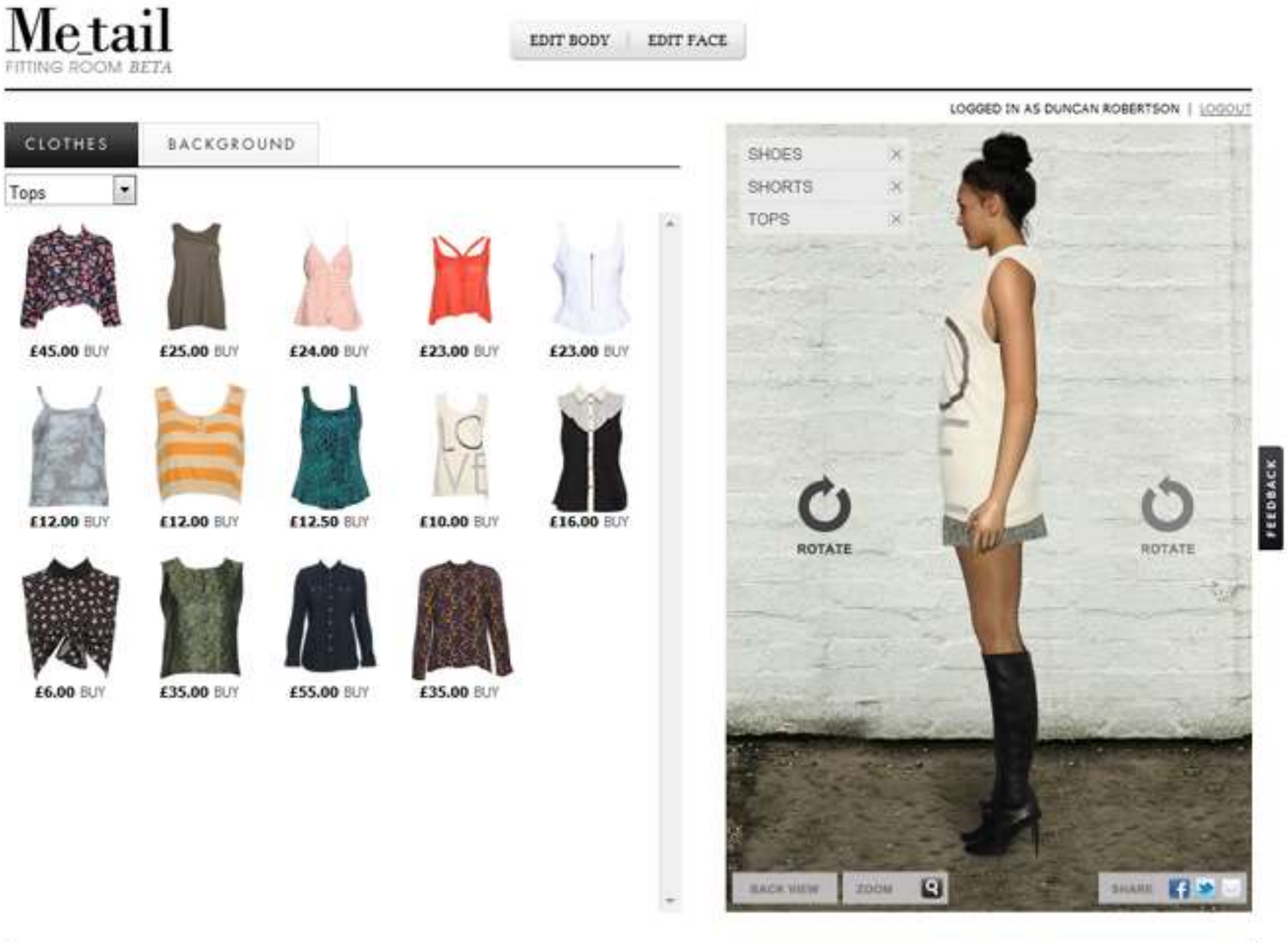

Me_tail
FITTING ROOM BETA
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
Create your Me_model
ENTER YOUR MEASUREMENTS AND UPLOAD A PHOTO OF YOUR FACE...
HEIGHT
Units: Ft
5'10"
5'0" 5'2" 5'4" 5'6" 5'8" 5'10"
WEIGHT
Units: St
10st 0
8st 0 9st 0 10st 0 11st 0 12st 0 13st 0 14st 0
More Measurements
SAVE & CONTINUE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

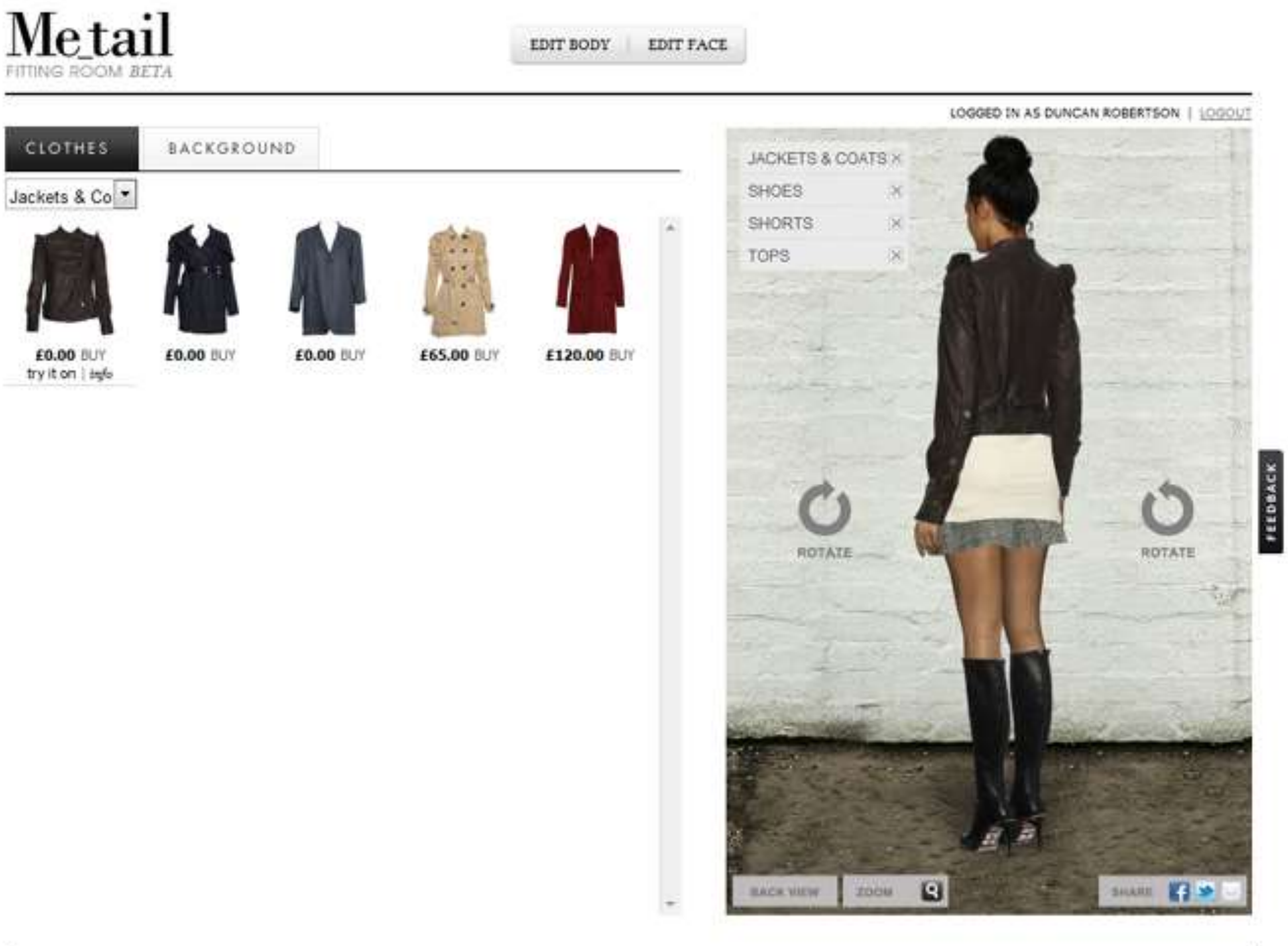

Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Dresses
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£0.00 BUY
£50.00 BUY
£19.00 BUY
£14.00 BUY
£7.00 BUY
£150.00 BUY
£250.00 BUY
£230.00 BUY
£10.00 BUY
£25.00 BUY
£5.00 BUY
£5.00 BUY
£5.00 BUY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

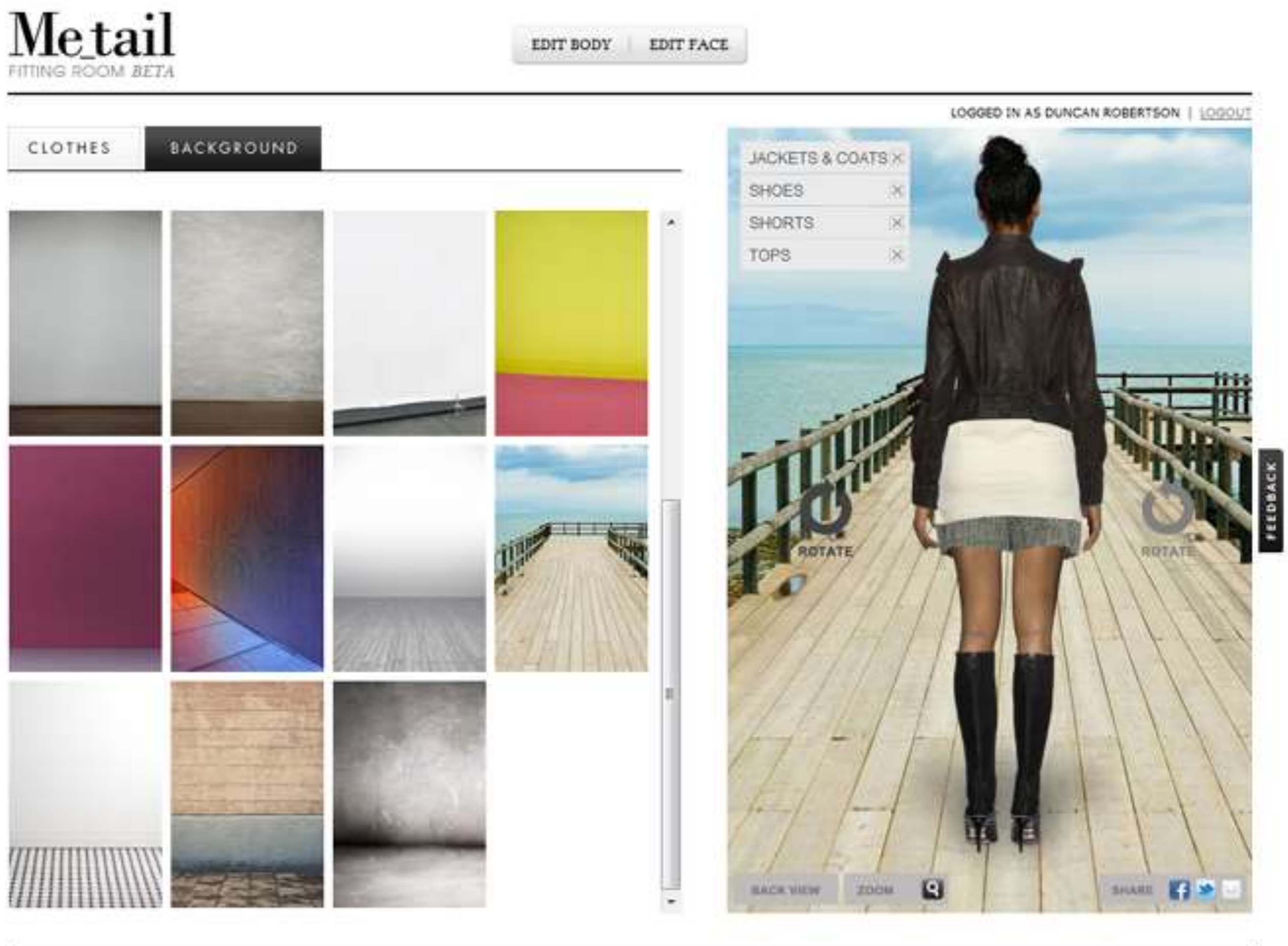

Me_tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Shoes
£0.00 BUY
£0.00 BUY
£0.00 BUY
try it on | info
SHOES
DRESSES
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE

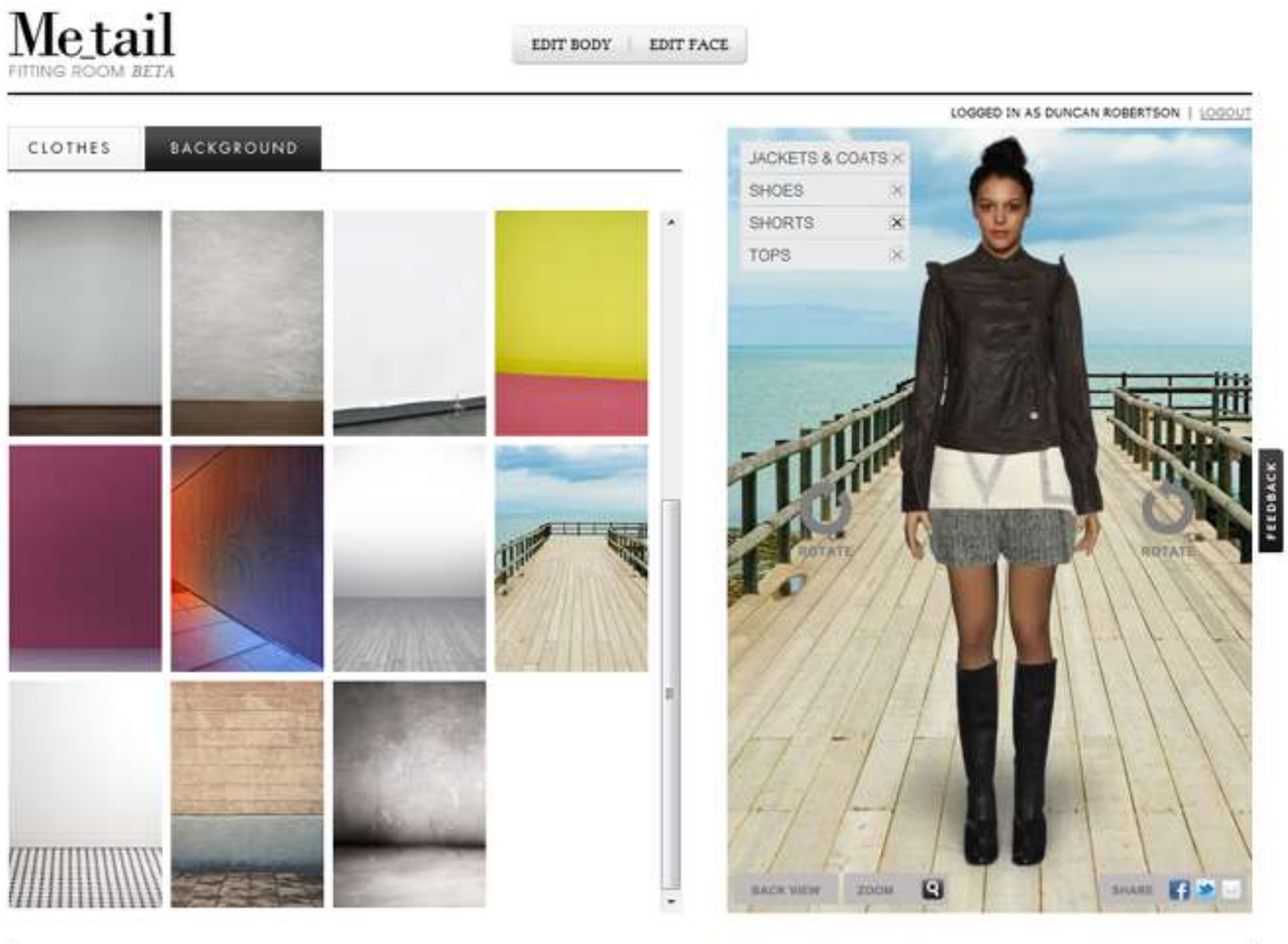


Me tail
FITTING ROOM BETA
EDIT BODY
EDIT FACE
LOGGED IN AS DUNCAN ROBERTSON | LOGOUT
CLOTHES
BACKGROUND
Shorts
£30.00 BUY
£48.00 BUY
£30.00 BUY
try it on | info
£30.00 BUY
SHOES
SHORTS
ROTATE
ROTATE
FEEDBACK
BACK VIEW
ZOOM
SHARE


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


Me tail
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


Metail
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


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

# Registration:

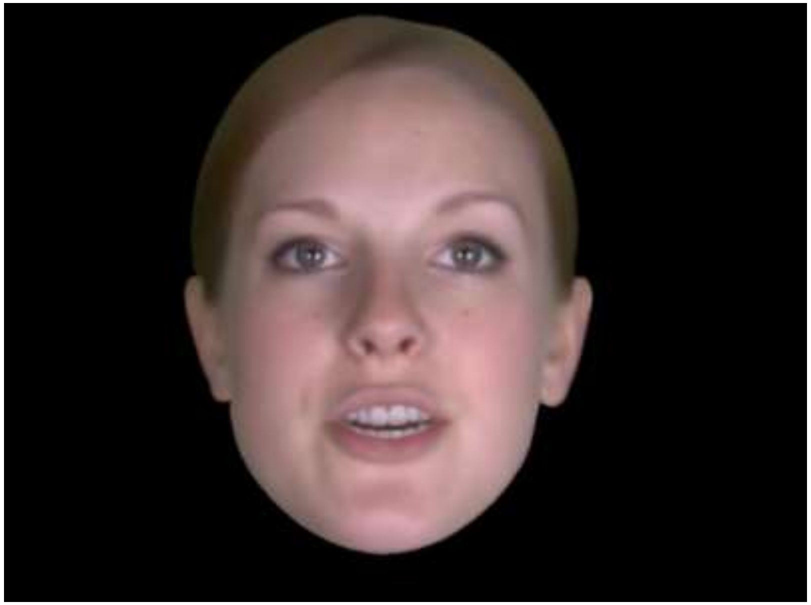
# Expressive Visual Text-to-Speech

# Registration – alignment of training data

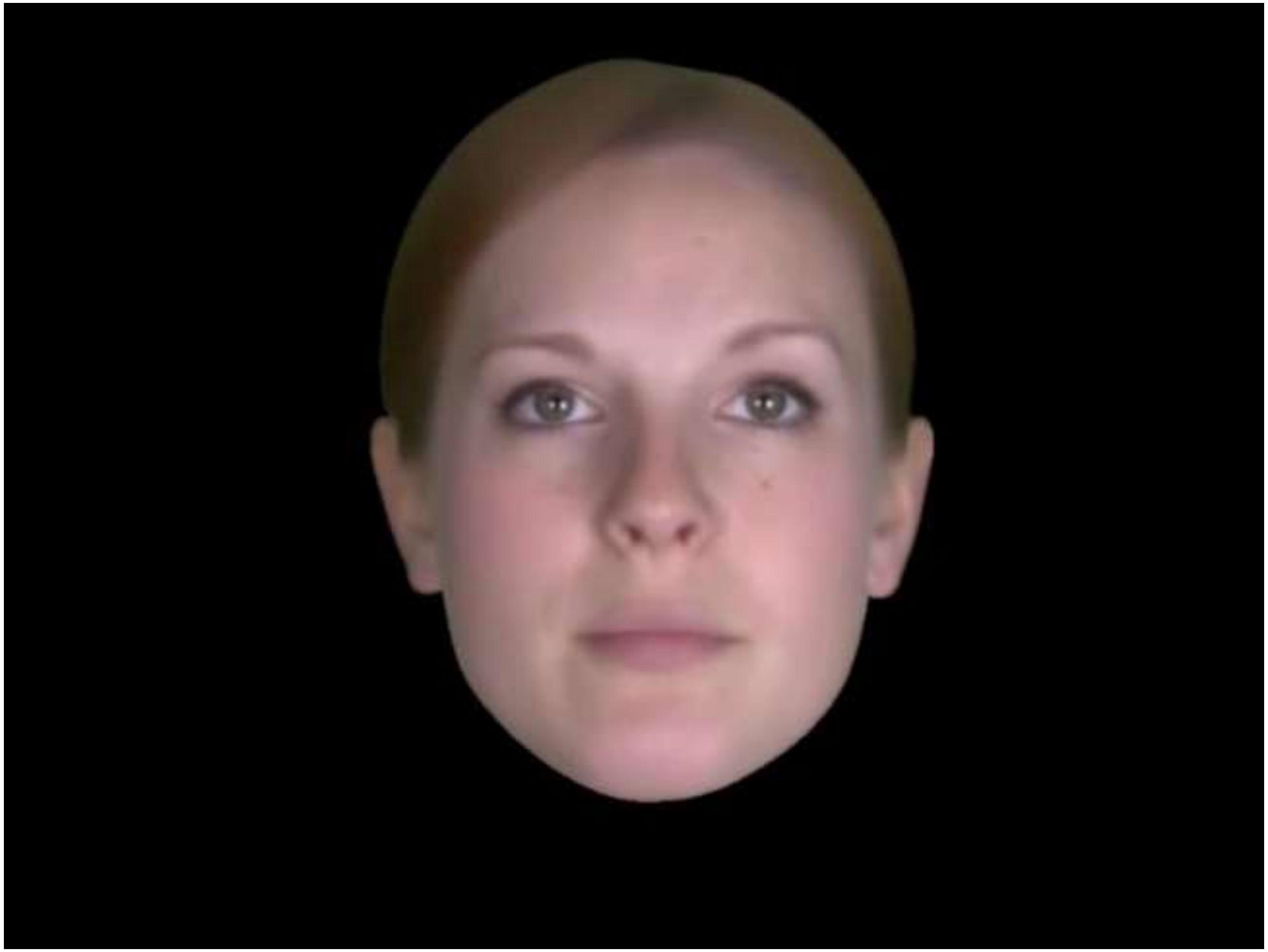
# What is an expressive talking head?

> User inputs a sentence which they wish to be uttered
> User specifies an emotion

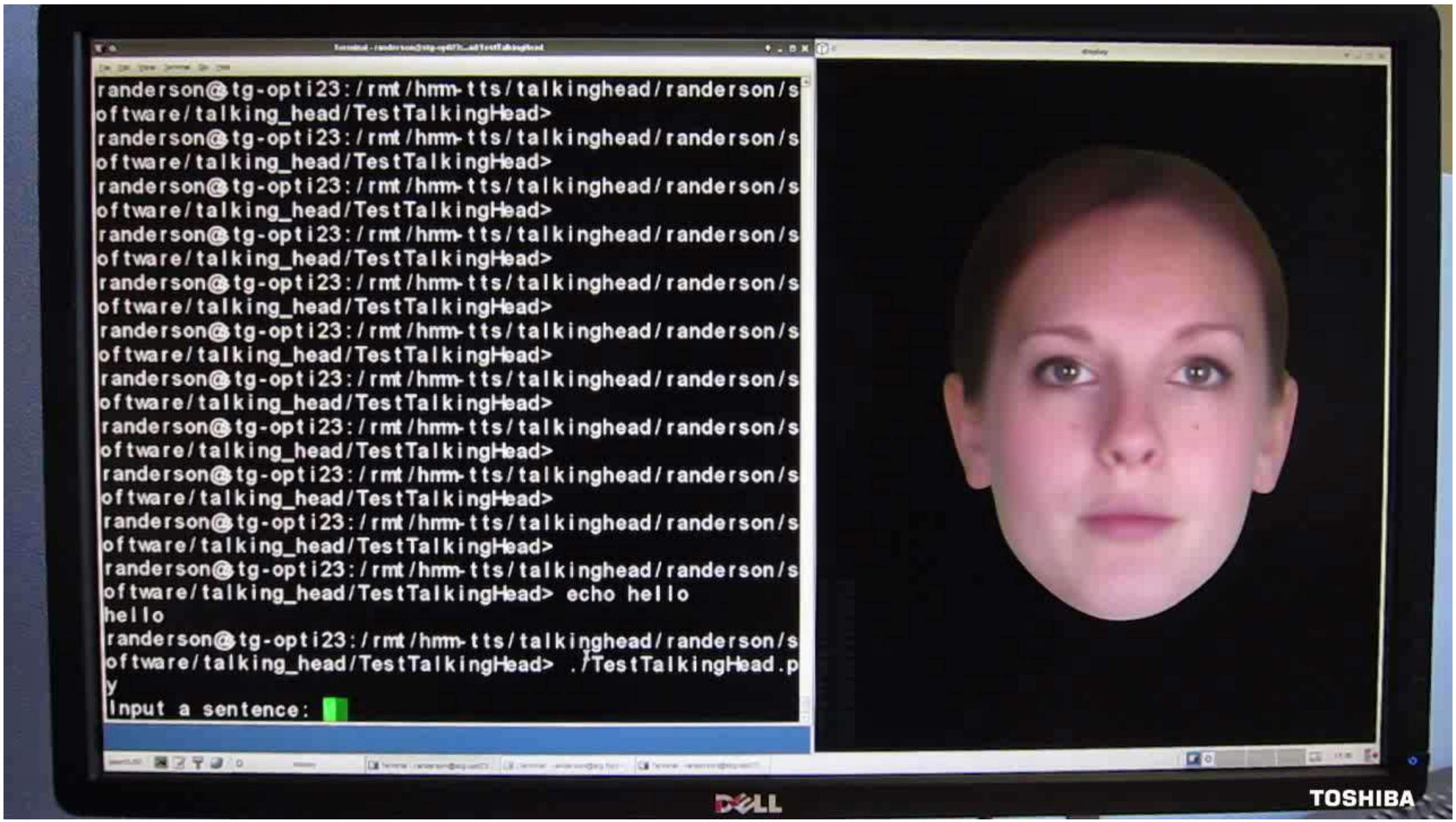
Video output is generated



# Our current talking head



# Our current talking head

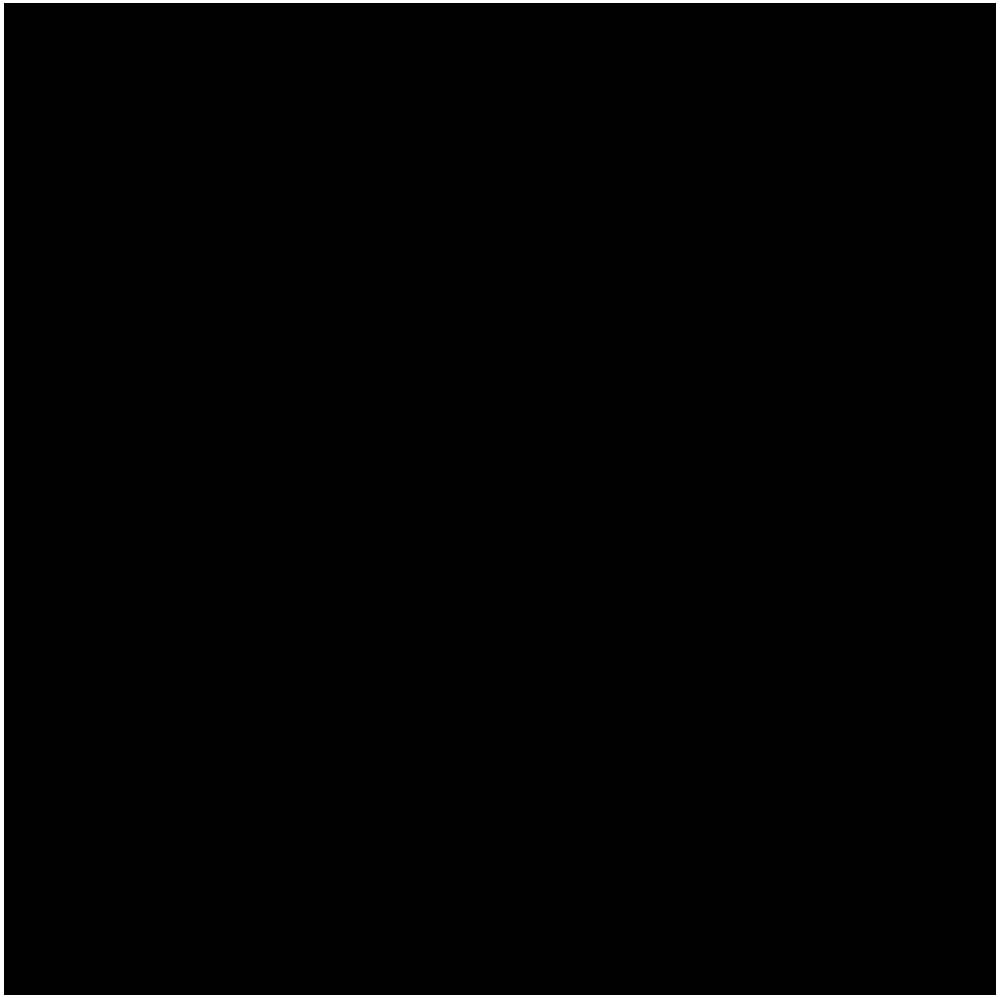


# EVTTS - Xpressive Talk

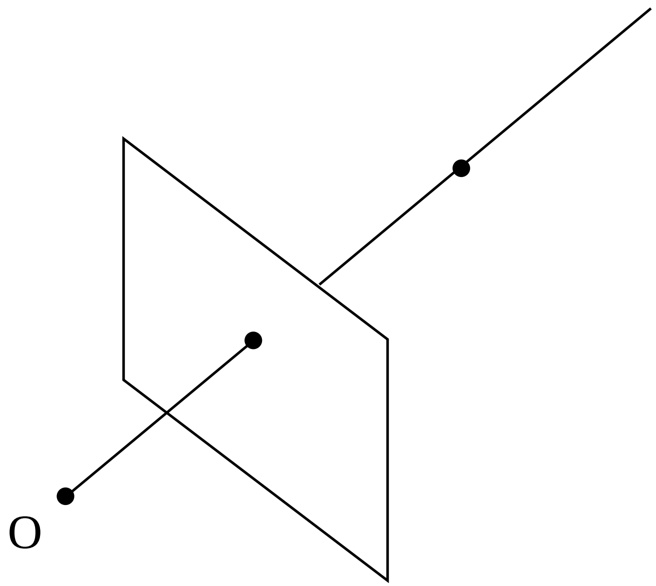
# Reconstruction?

Recovery of 3D shape from images

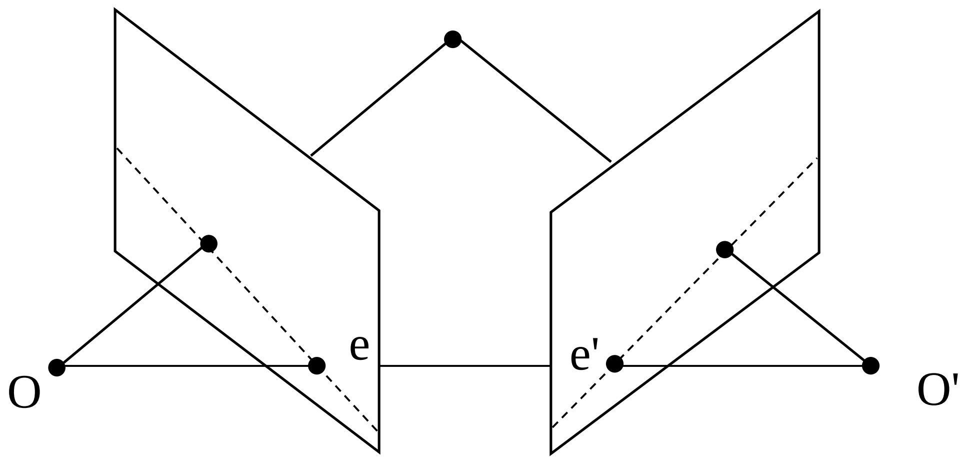
# Reconstruction



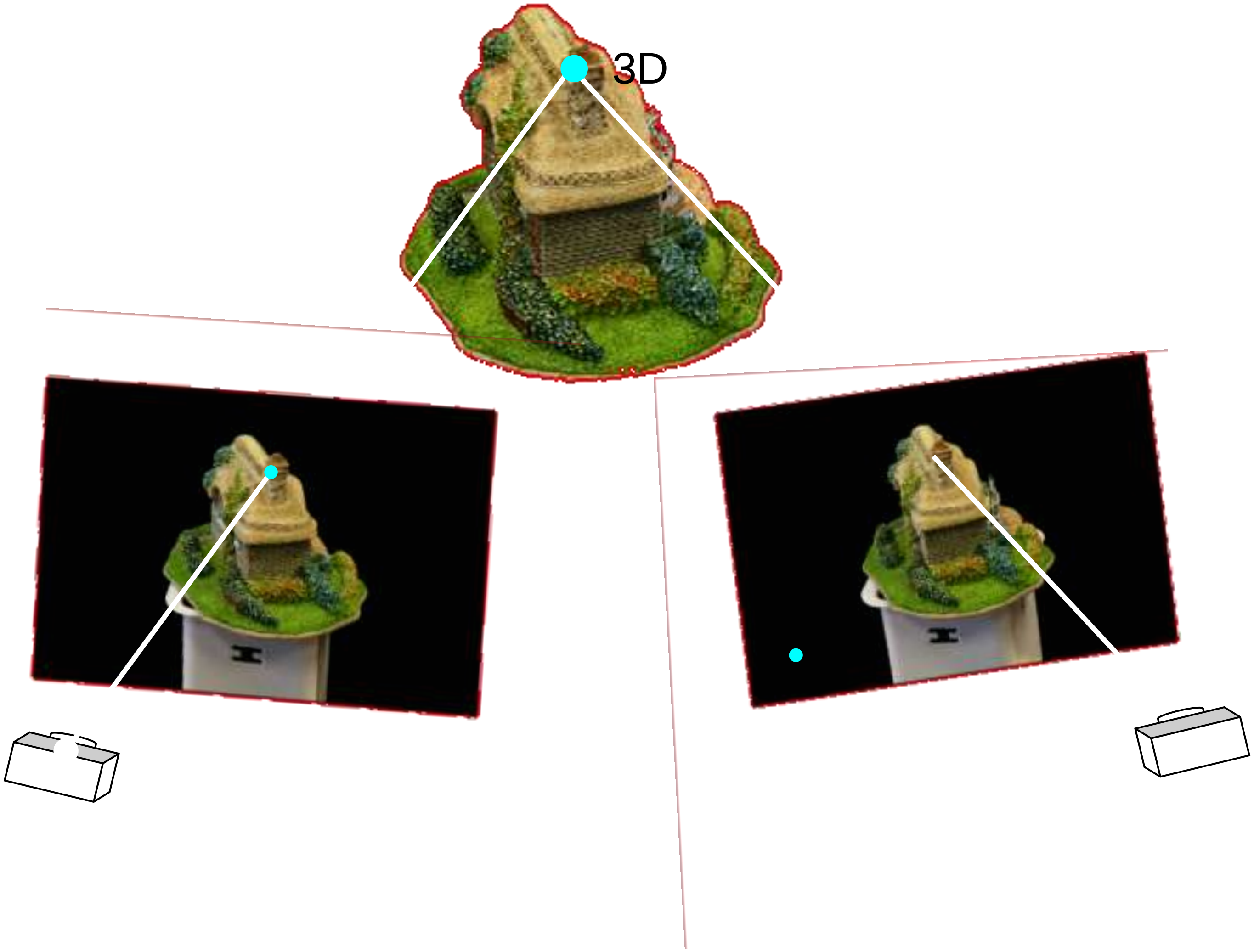
# Ambiquity in a single view



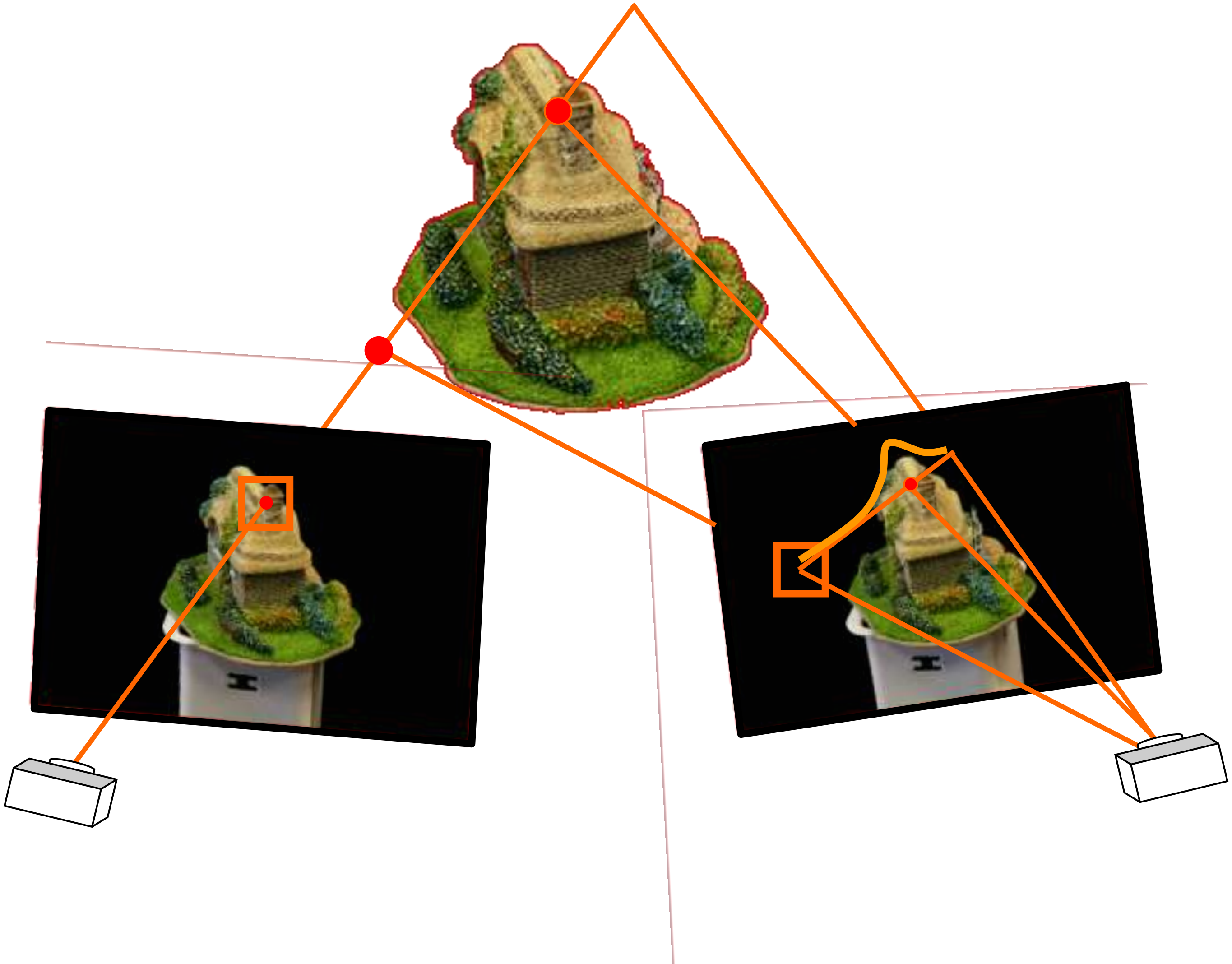
# Stereo vision



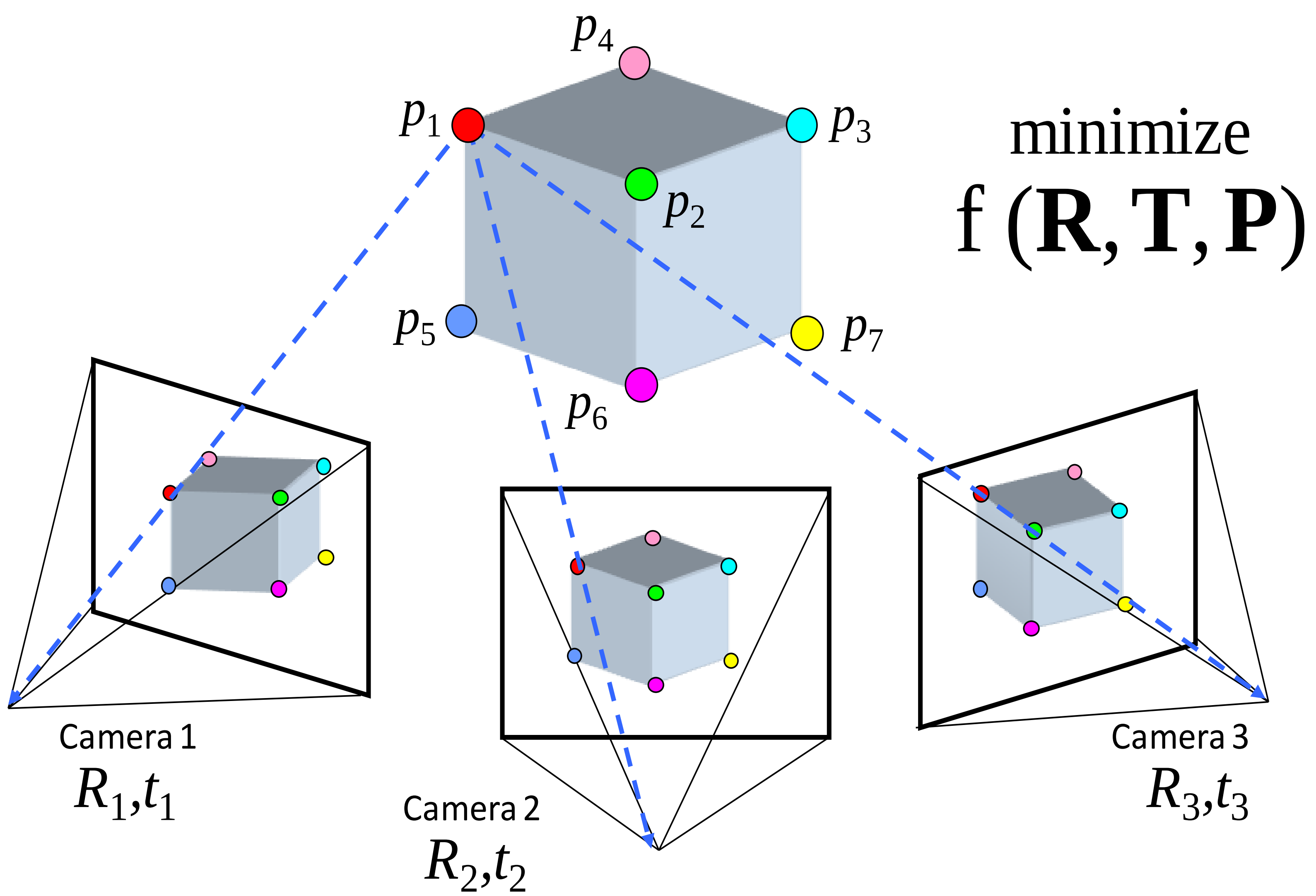
# Stereo vision



# Find optimal depth



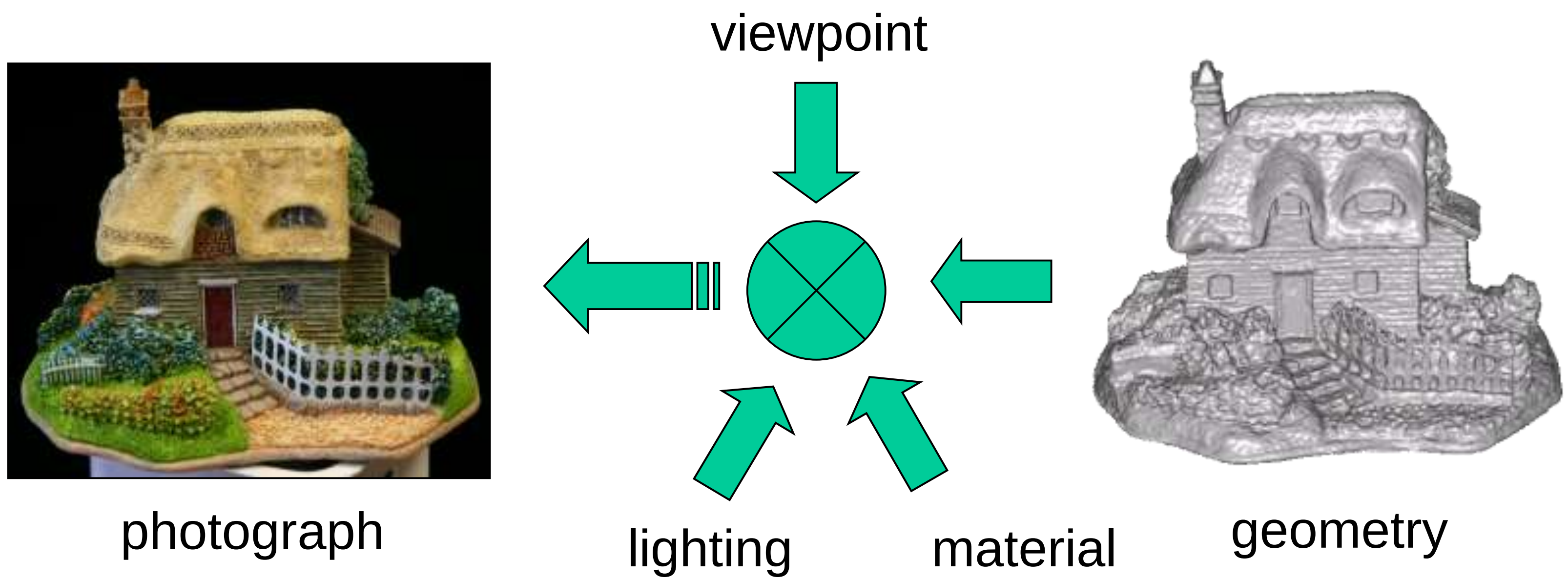
# Multi-view stereo



# Review

# Recovery of accurate 3D shape from images

# 3D shape from photographs



# 3D shape from photographs

Textured
rigid object

Untextured
rigid object

Untextured
deformable object

# 3D shape from photographs

UNIVERSITY OF CAMBRIDGE

Multi-view
stereo

Multi-view
photometric stereo

Single view coloured
photometric stereo

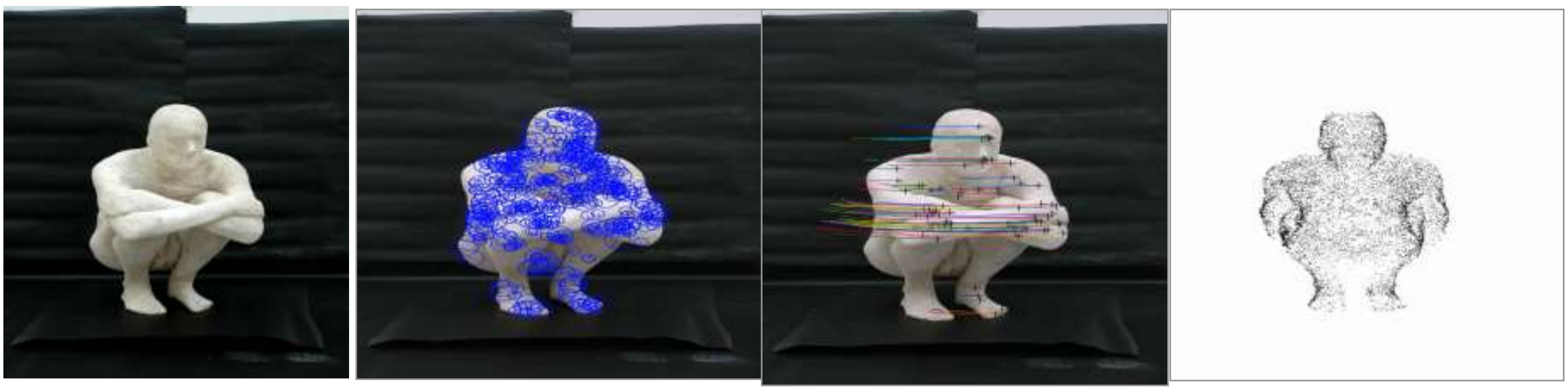
# Overview of reconstruction

UNIVERSITY OF CAMBRIDGE

1. Multi-view stereo
2. Multi-view photometric stereo
3. Single-view colour photometric stereo
4. Large scale outdoor reconstructions

# Digital Pygmalion Project

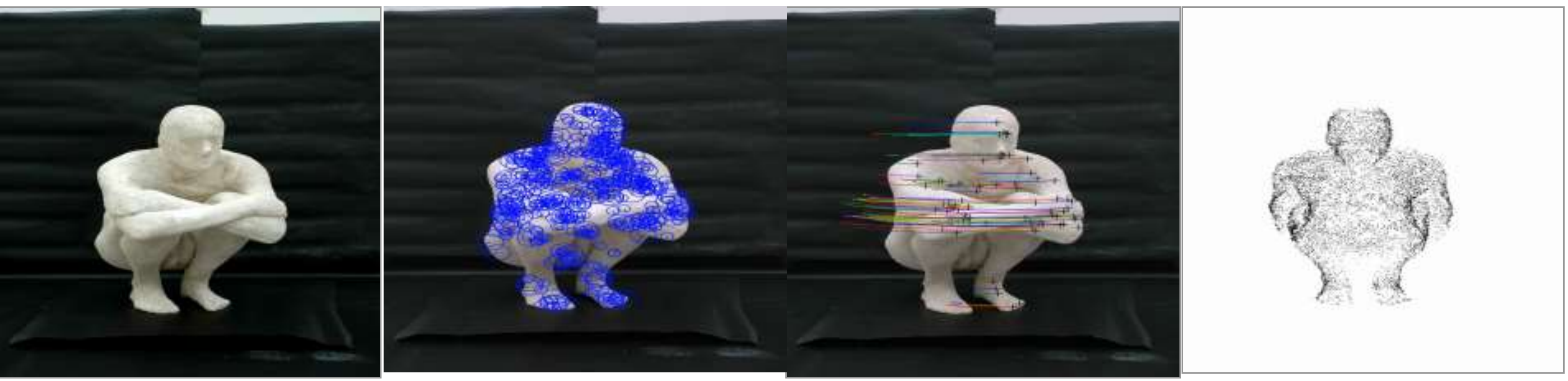
# Structure from motion



Input sequence 2D features 2D track 3D points

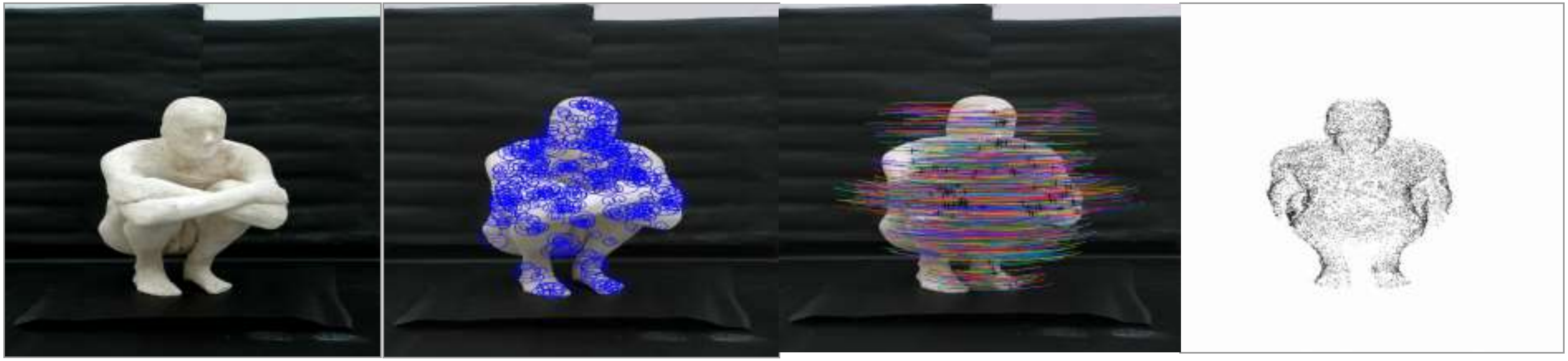
# Structure from motion

UNIVERSITY OF CAMBRIDGE



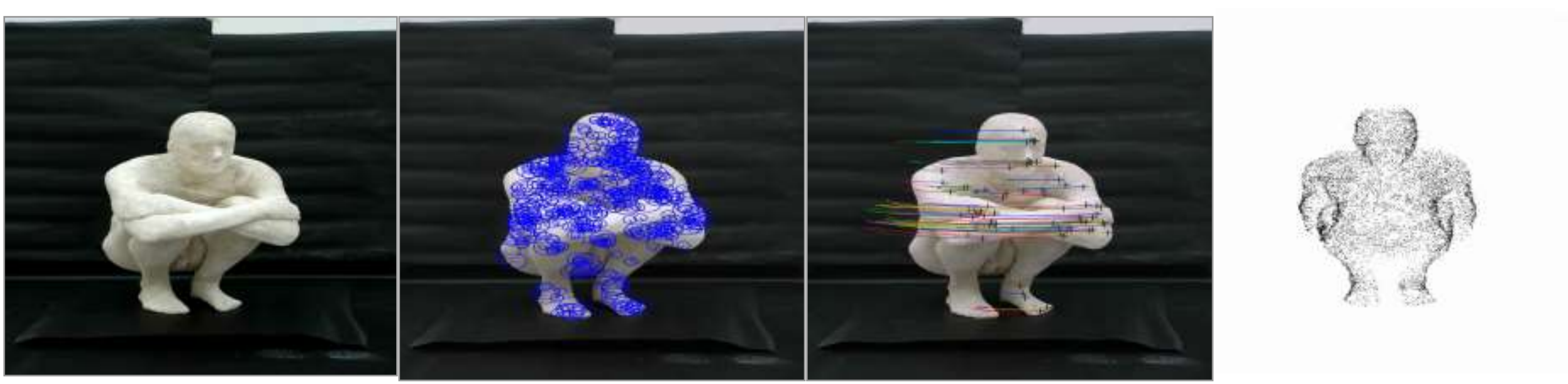
Input sequence    2D features    2D track    3D points

# Structure from motion



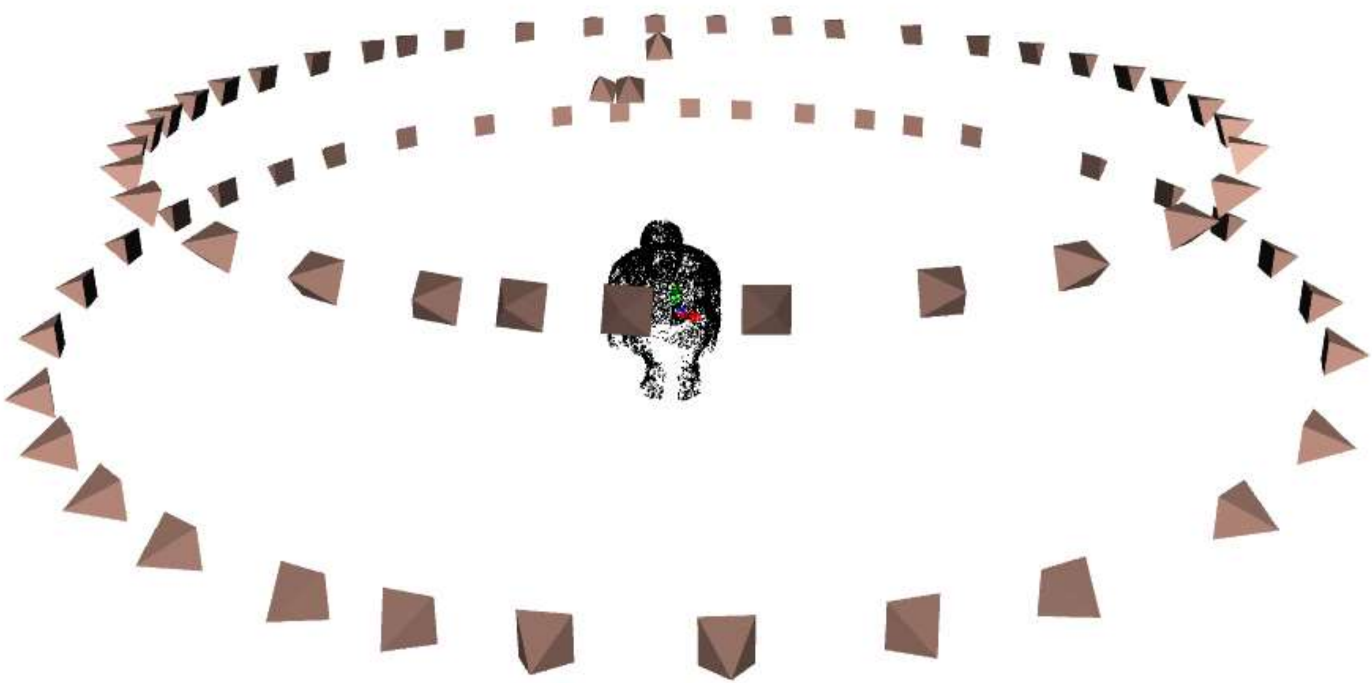
Input sequence    2D features    2D track    3D points

# Structure from motion

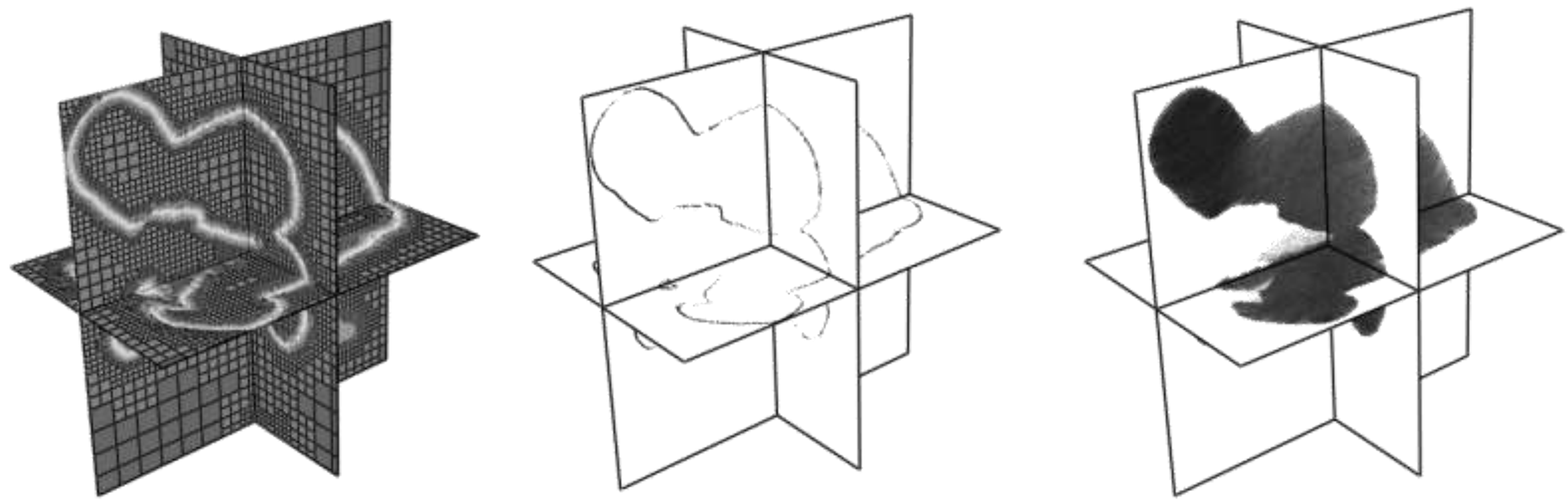


Input sequence   2D features   2D track   3D points

# Motion estimation result



# 3D MRF for 3D modelling

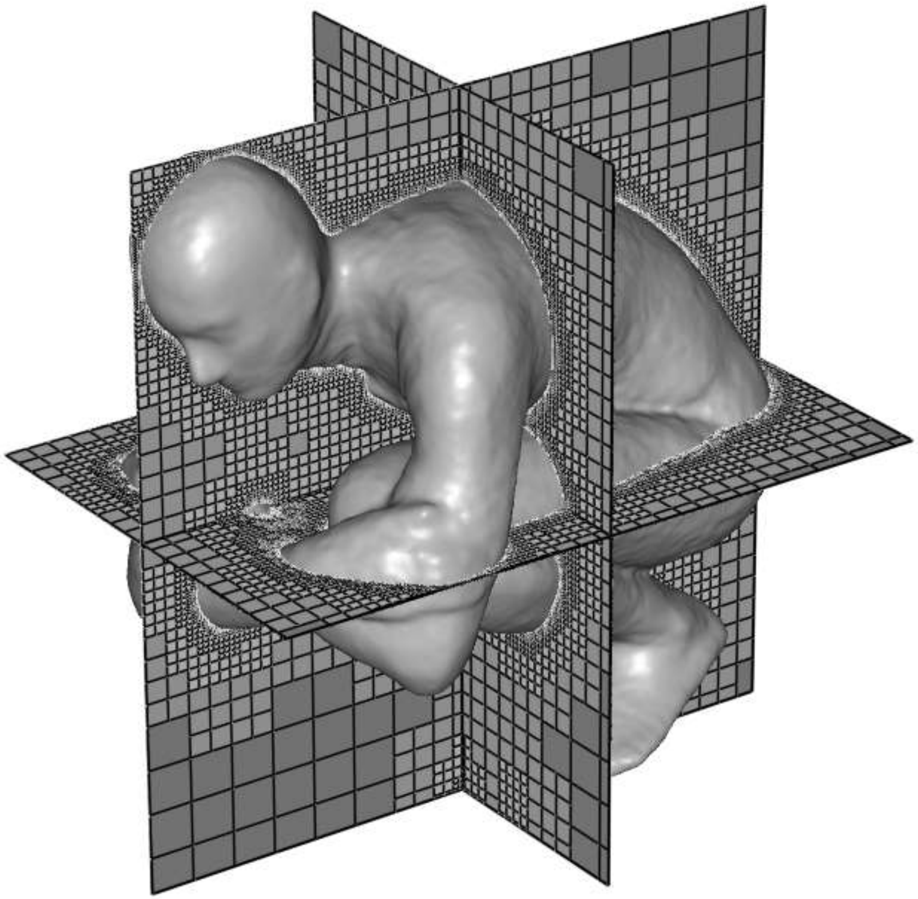


Multi-resolution grid

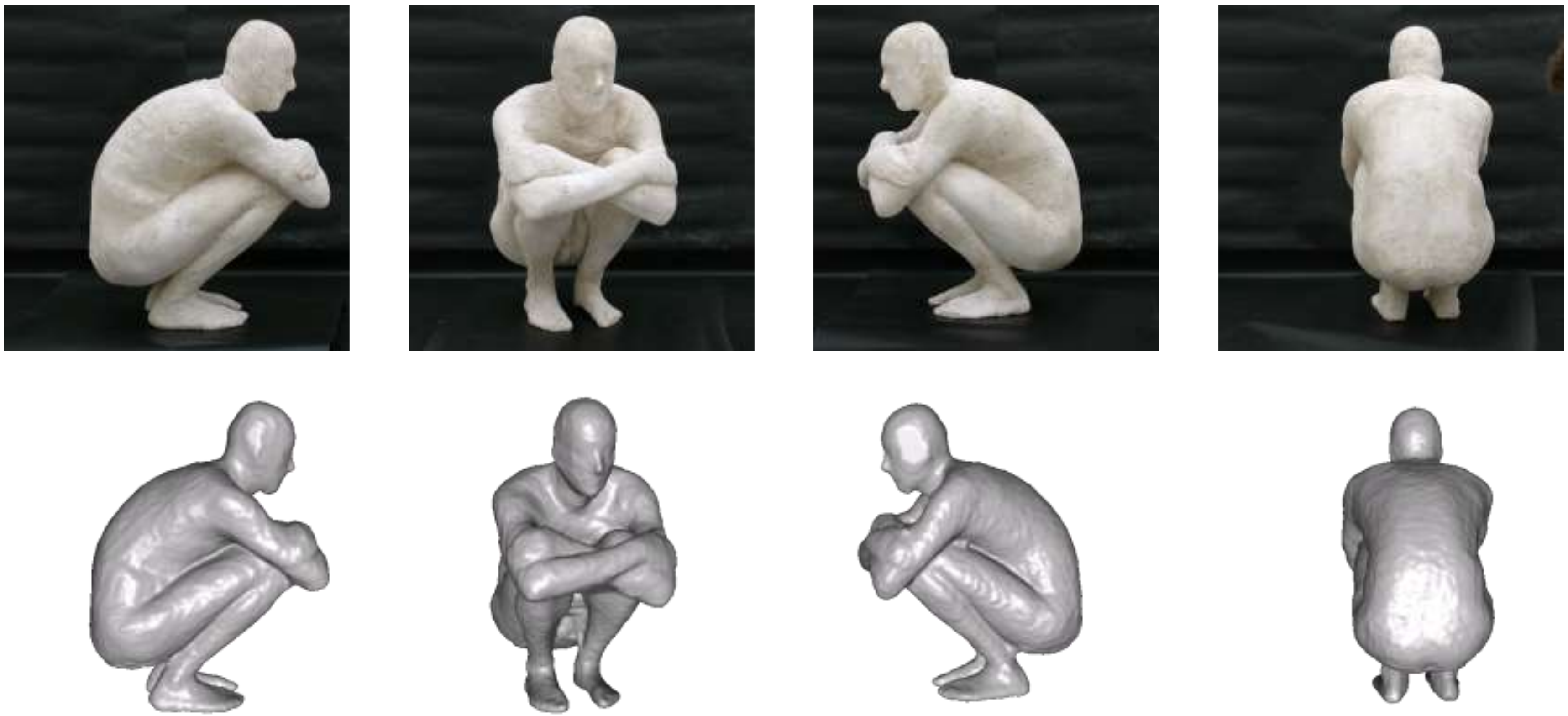
Edge cost

Foreground/ background cost

# 3D MRF for 3D modelling

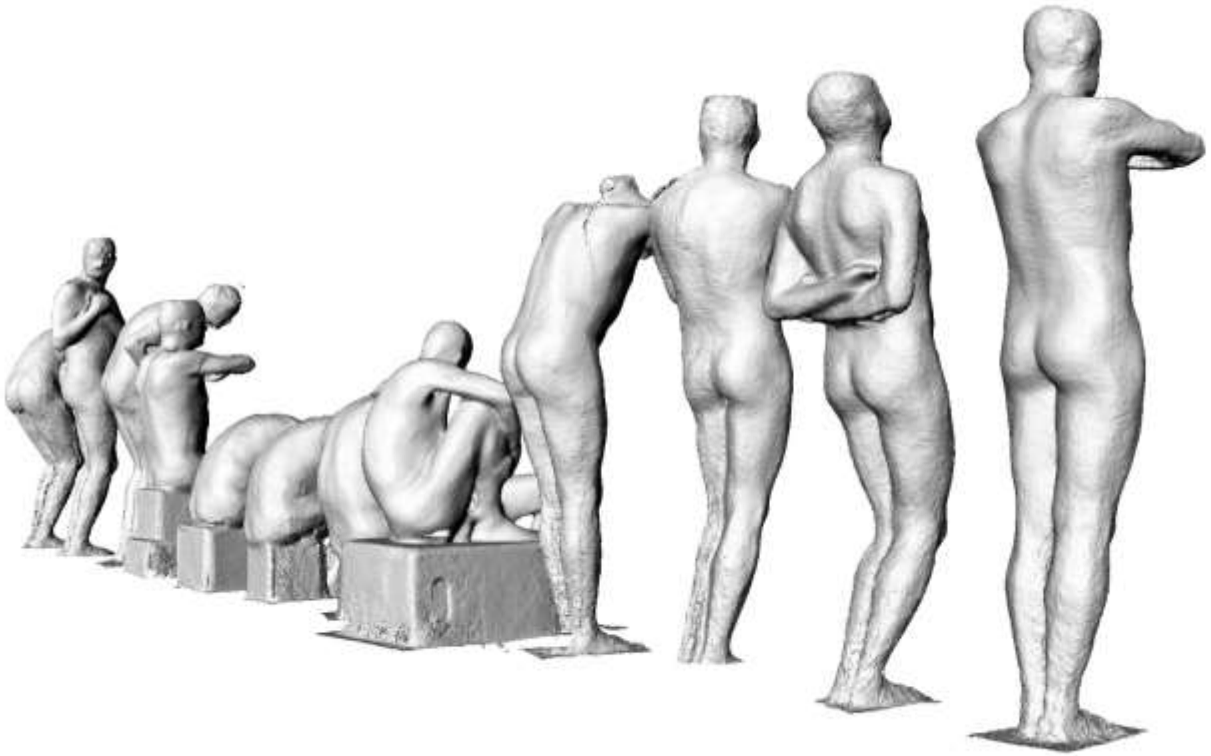


# 3D Models

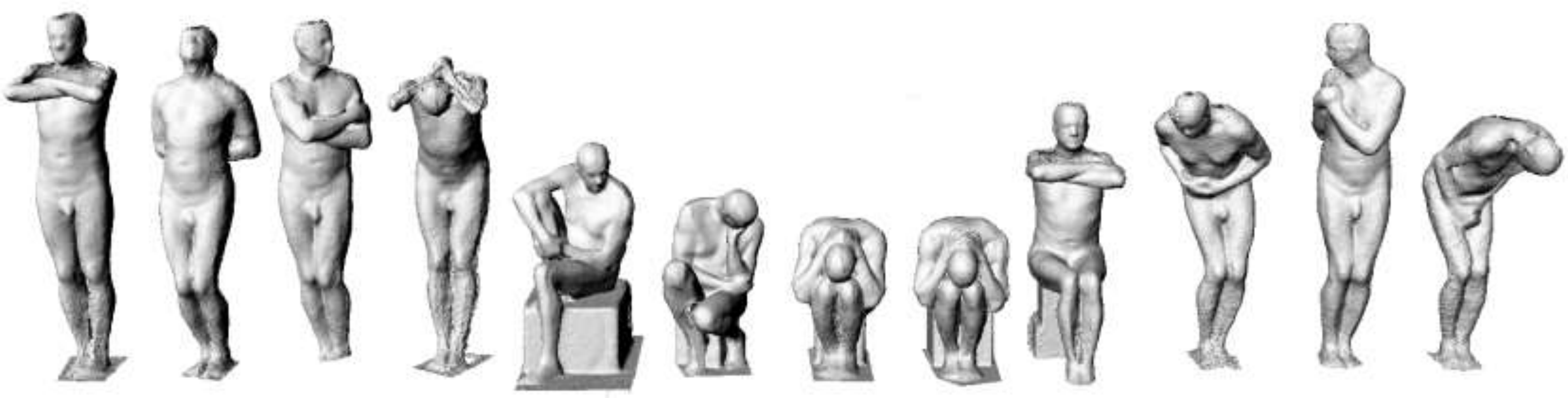


# Final installation

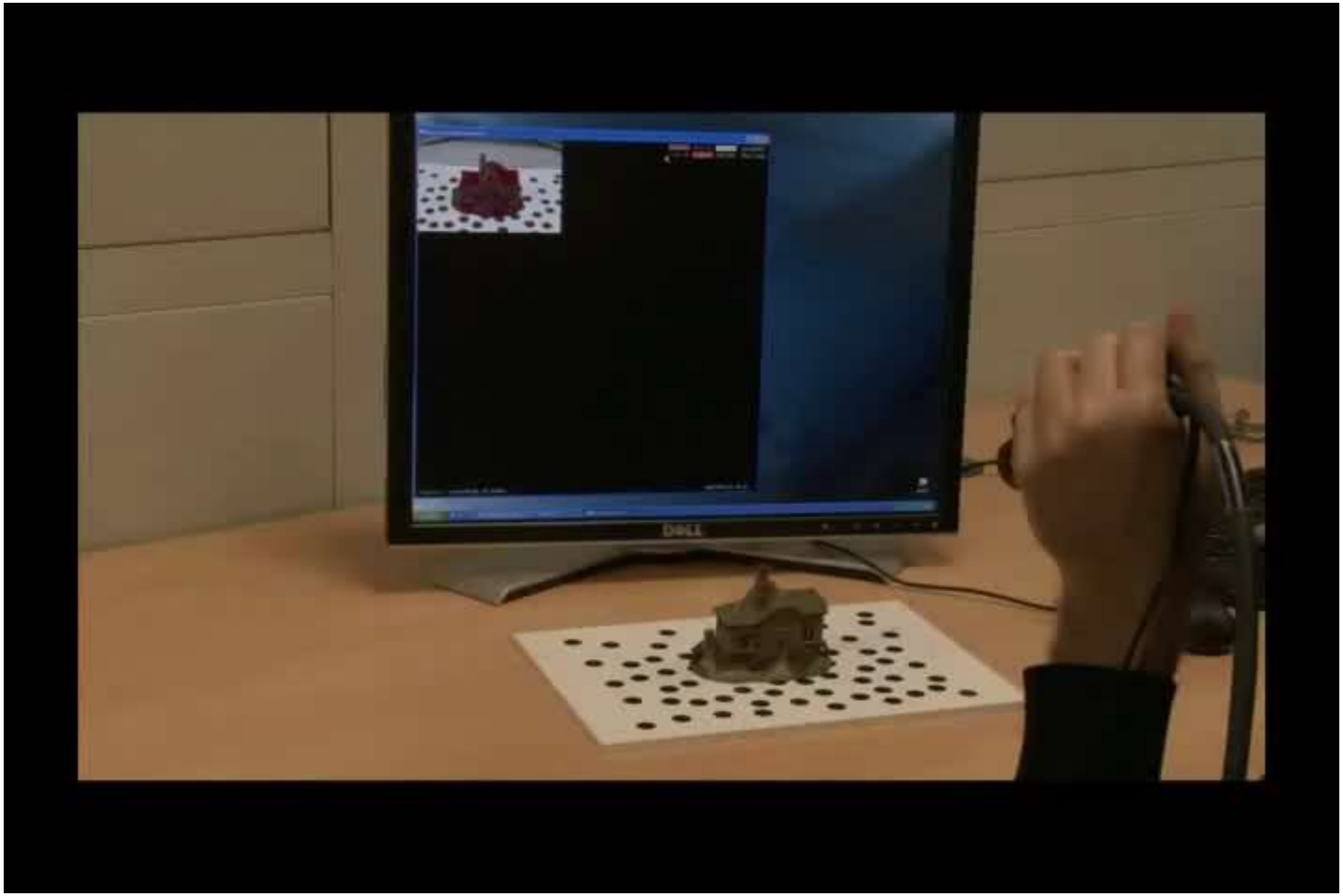
# 3D scans - Antony Gormley



# 3D scans - Antony Gormley



# Real-time depth

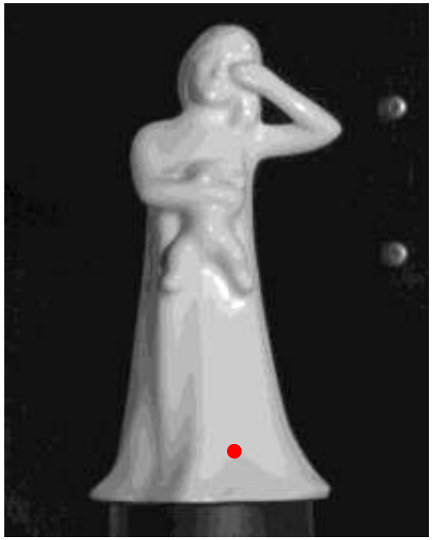


# Multiview photometric stereo

Vogiatzis, Hernandez and Cipolla 2006 and 2008

# 2. Untextured objects

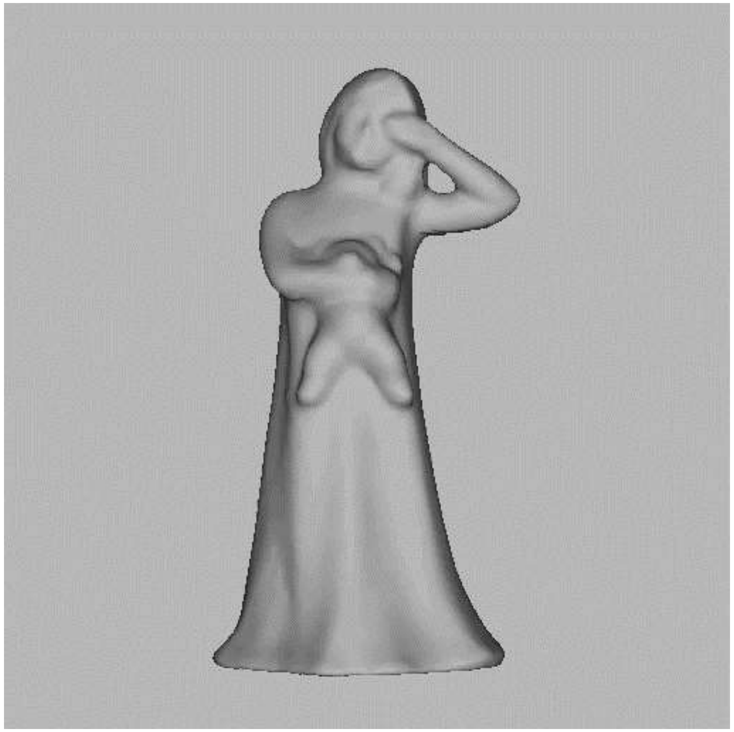
- Almost impossible to establish correspondence



Use shading cue

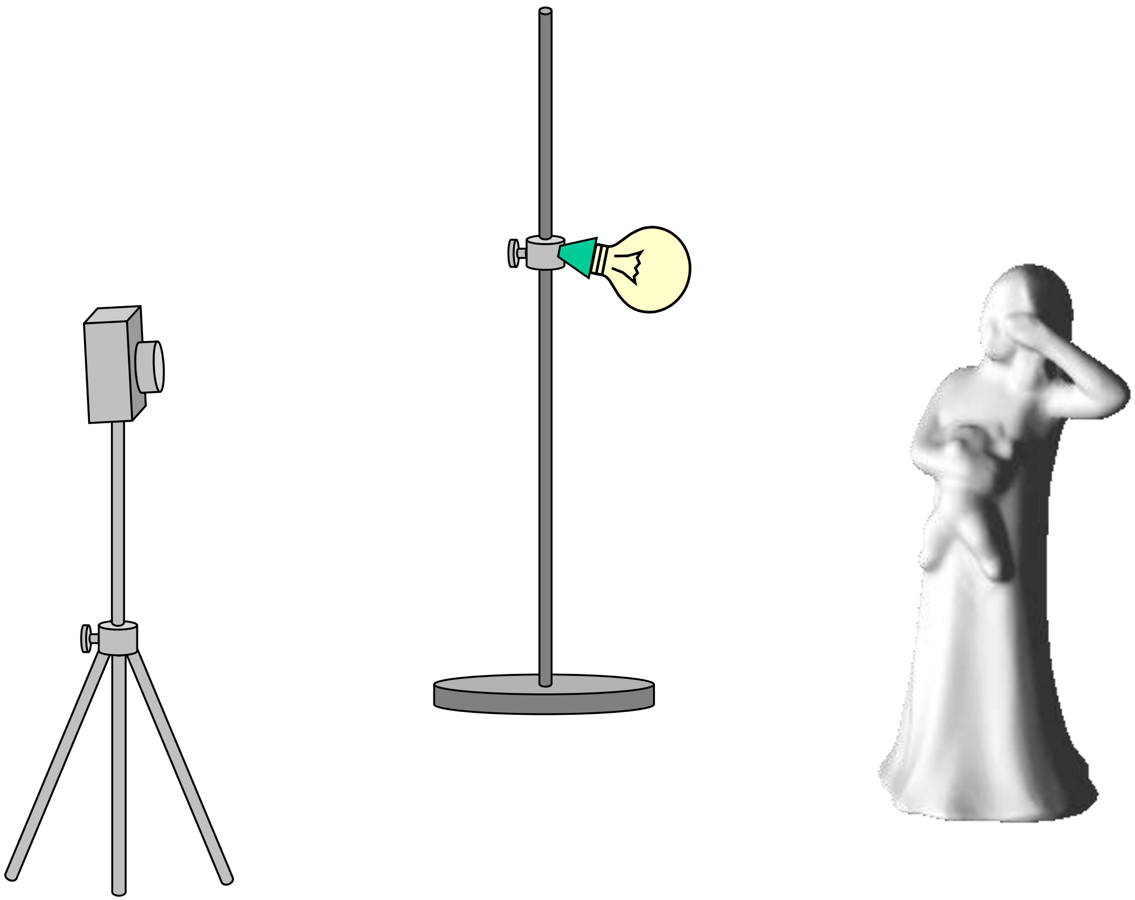
# Photometric stereo

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

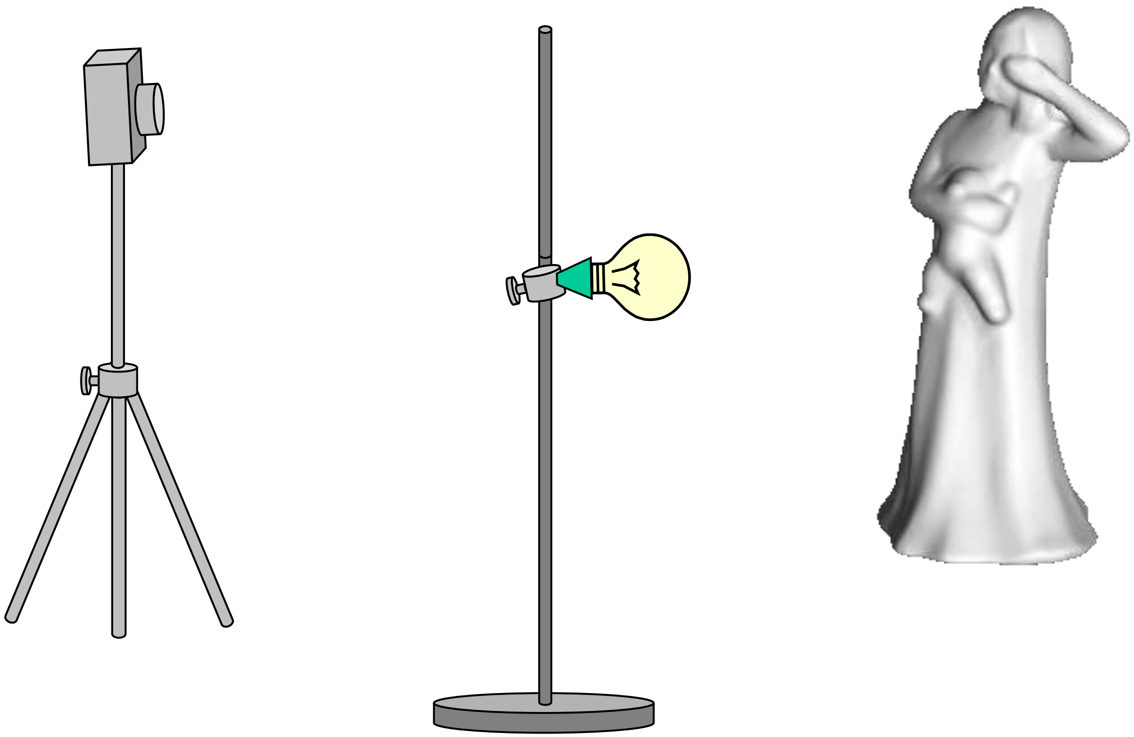


Changing lighting uncovers geometric detail

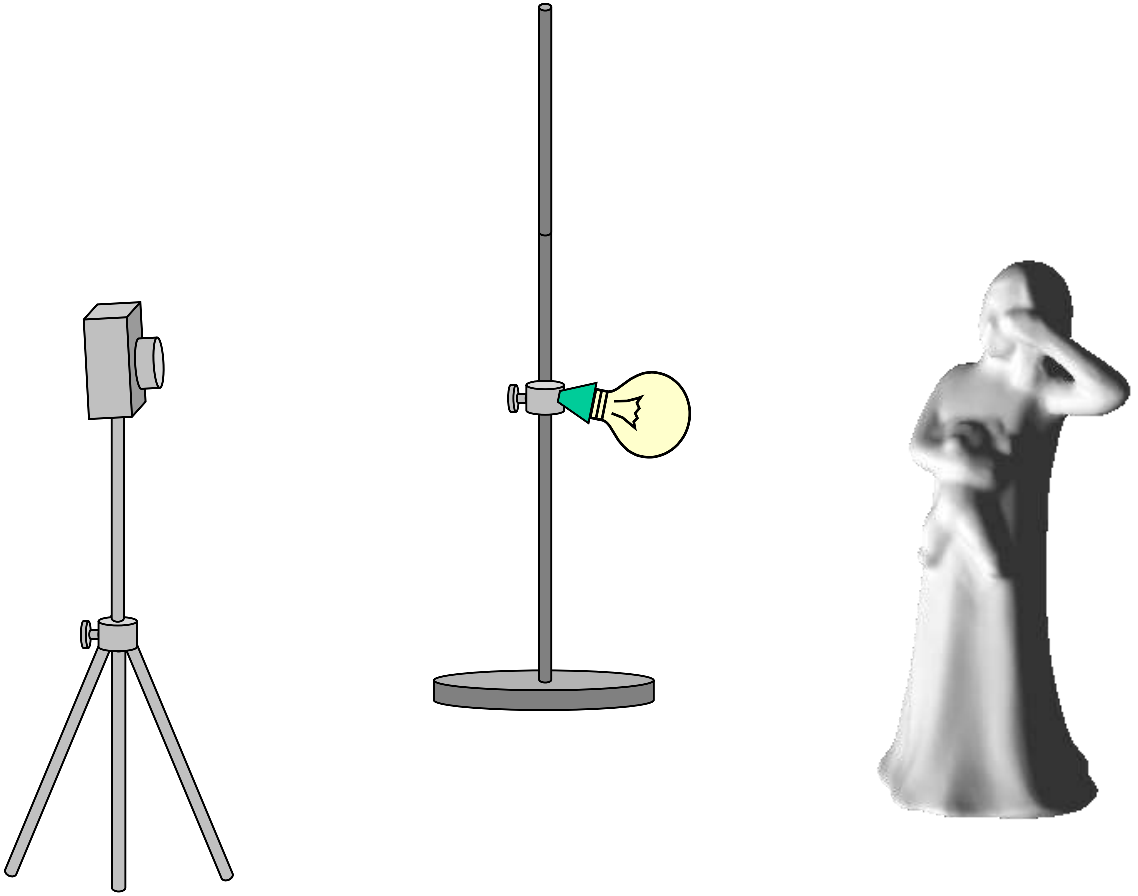
# Classic photometric stereo



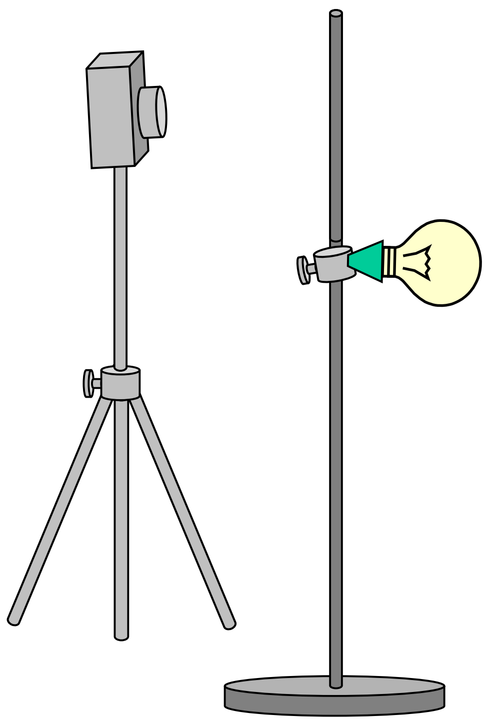
# Photometric stereo

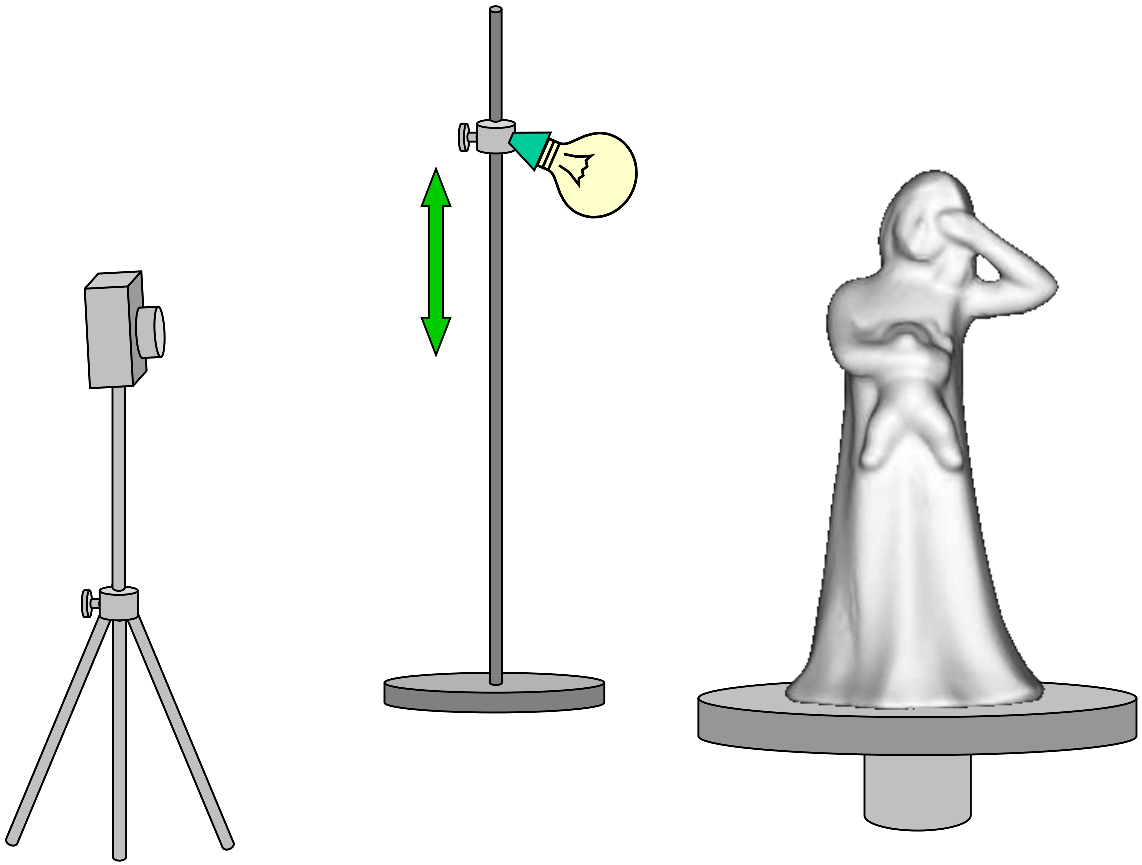


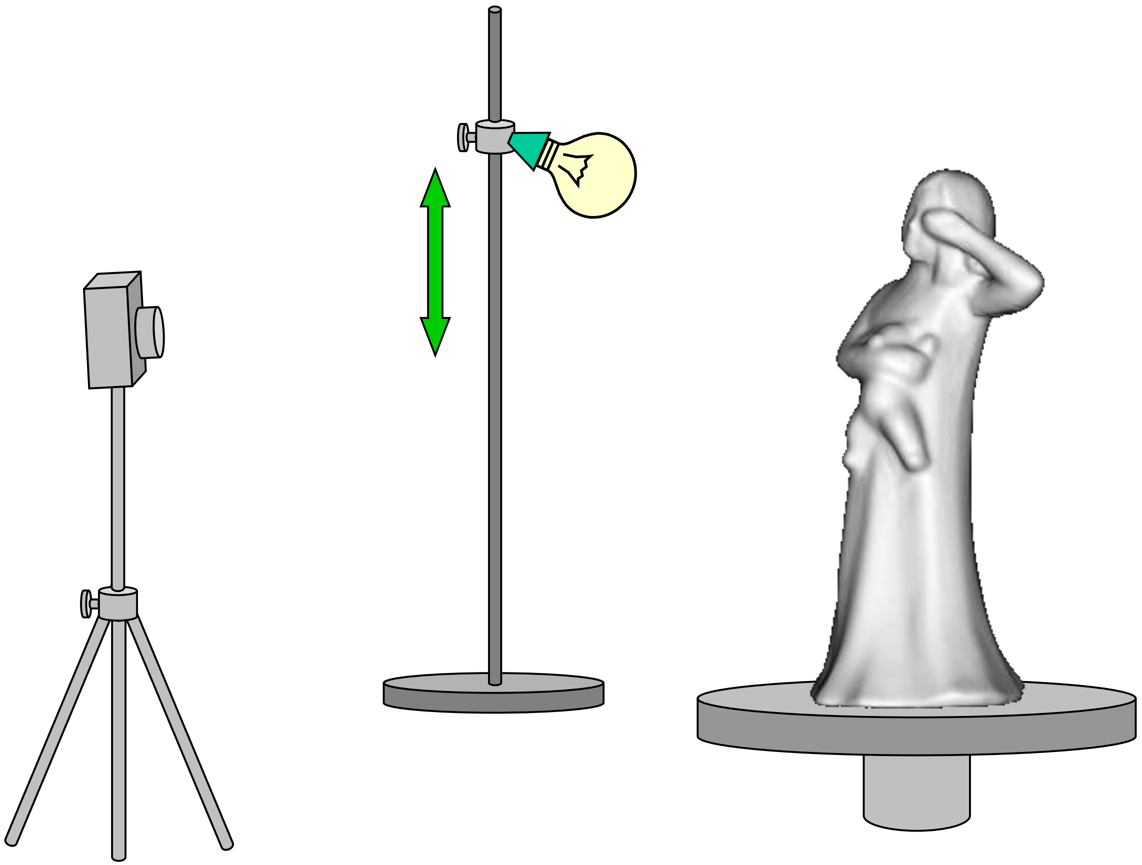
# Photometric stereo

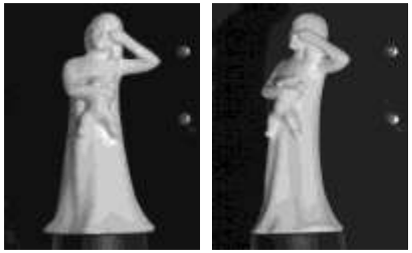


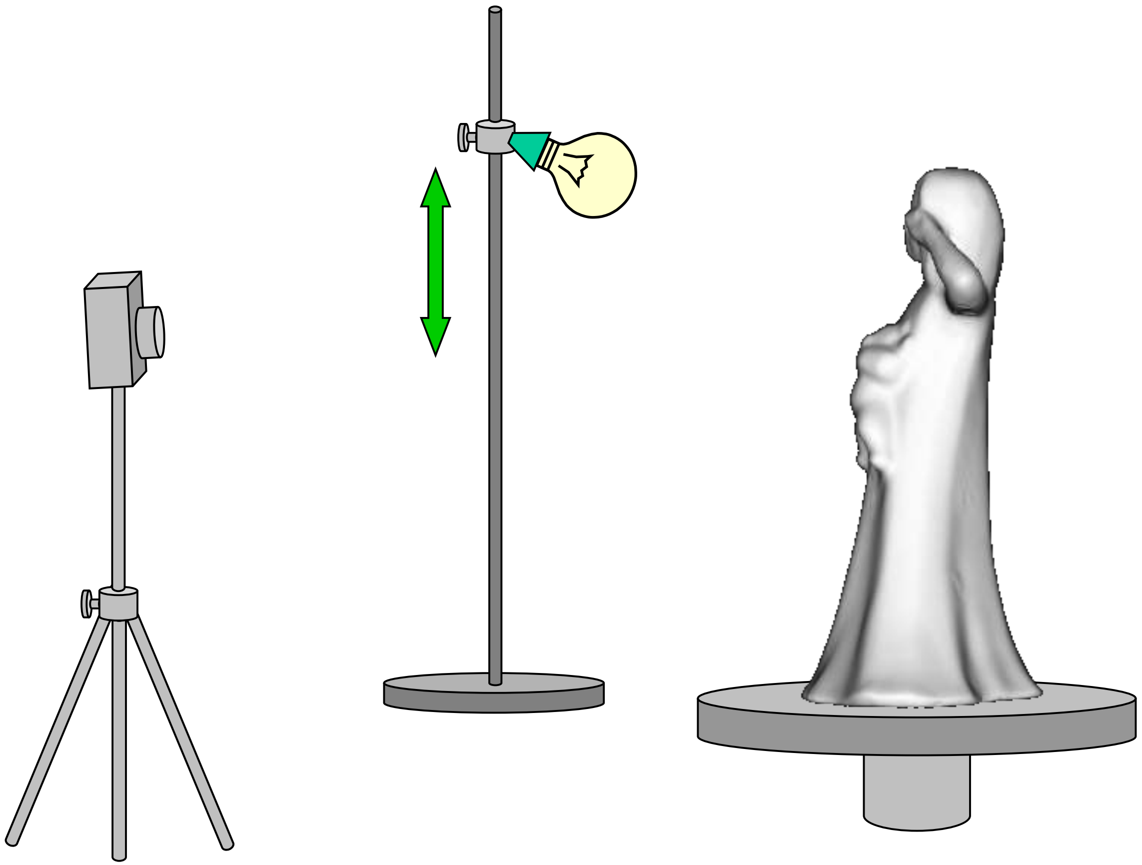
# Photometric stereo



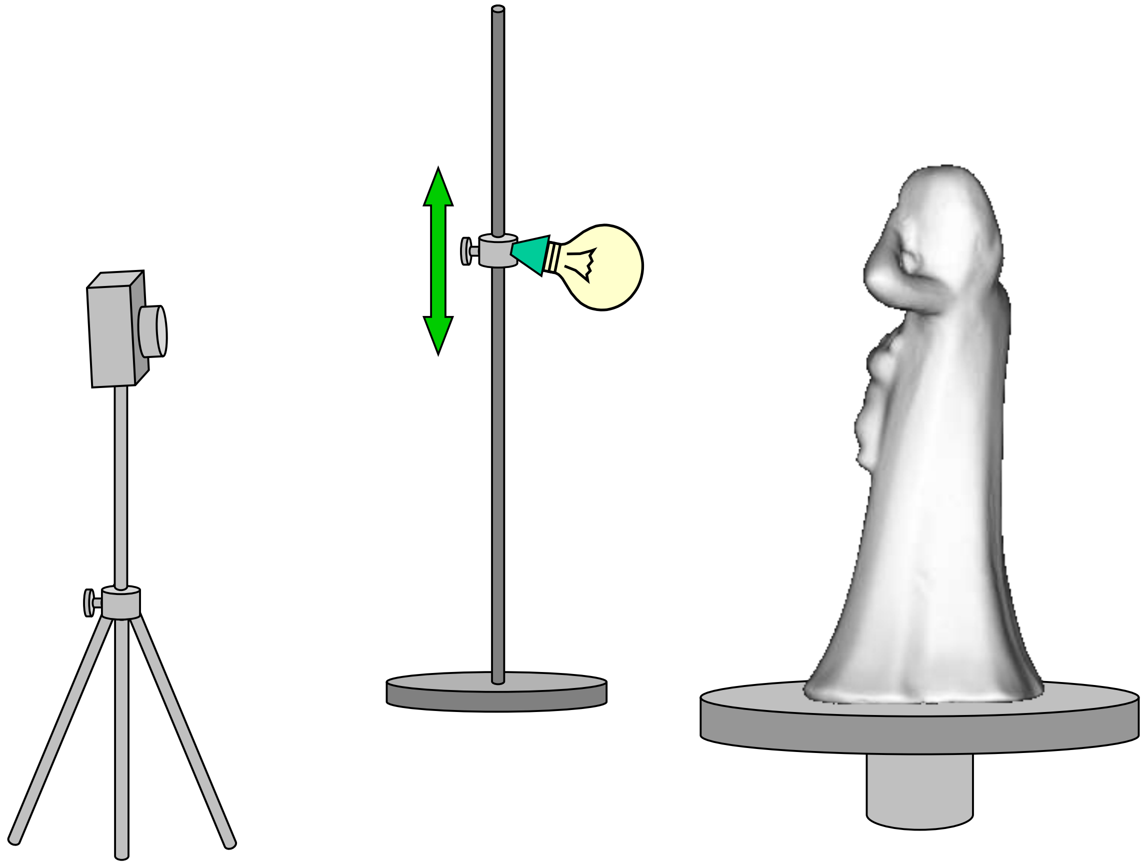


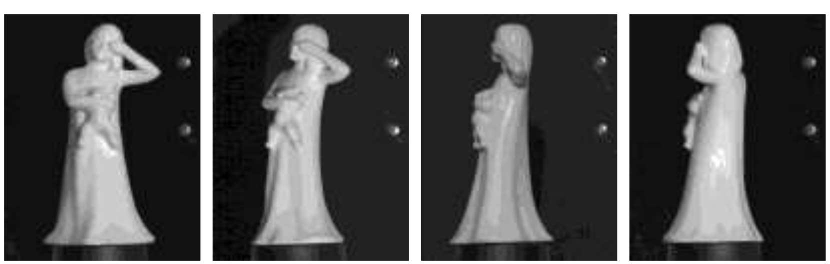




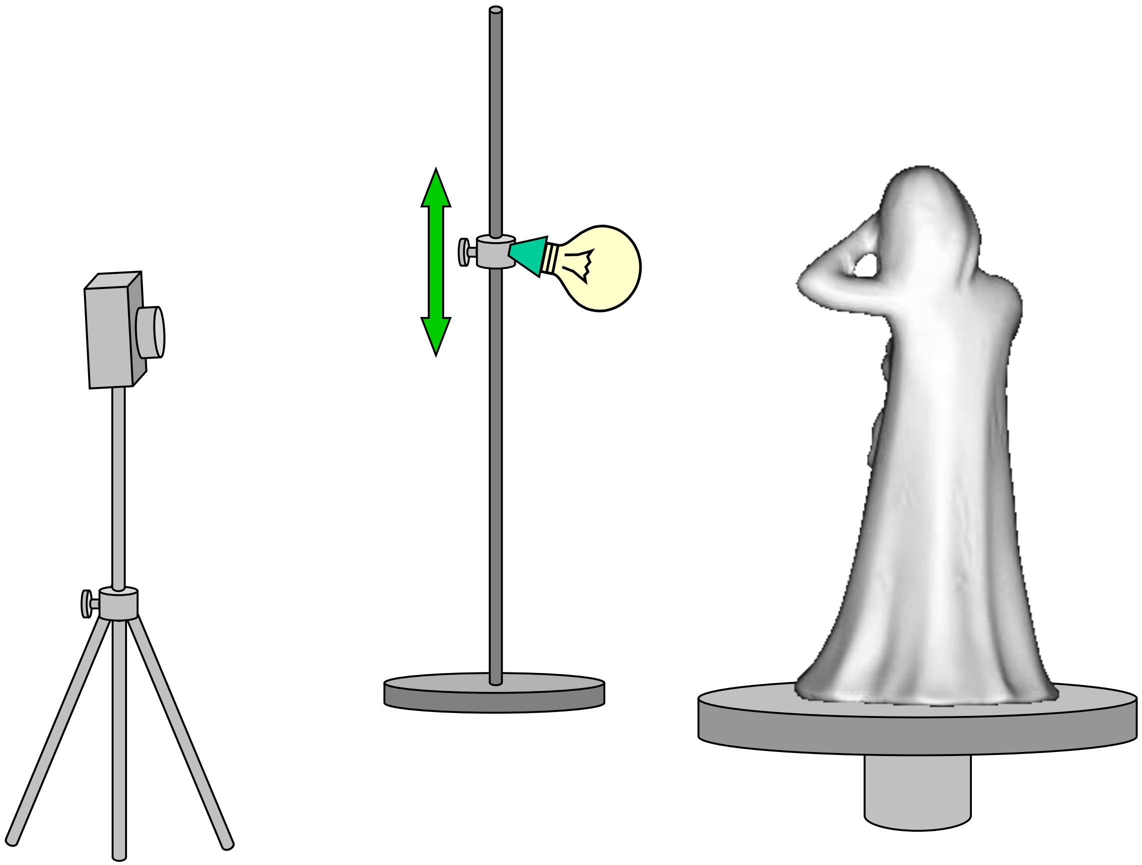


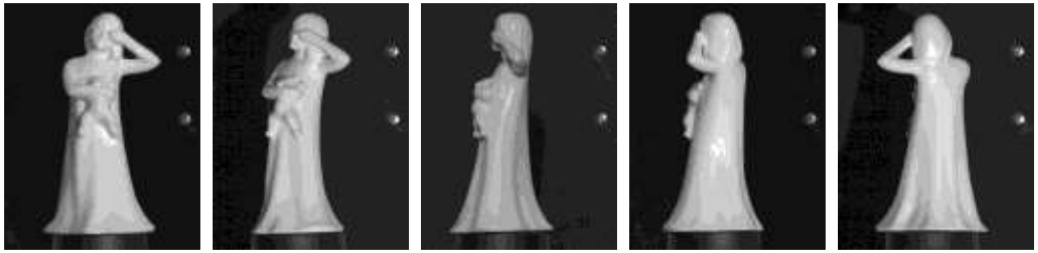
UNIVERSITY OF CAMBRIDGE



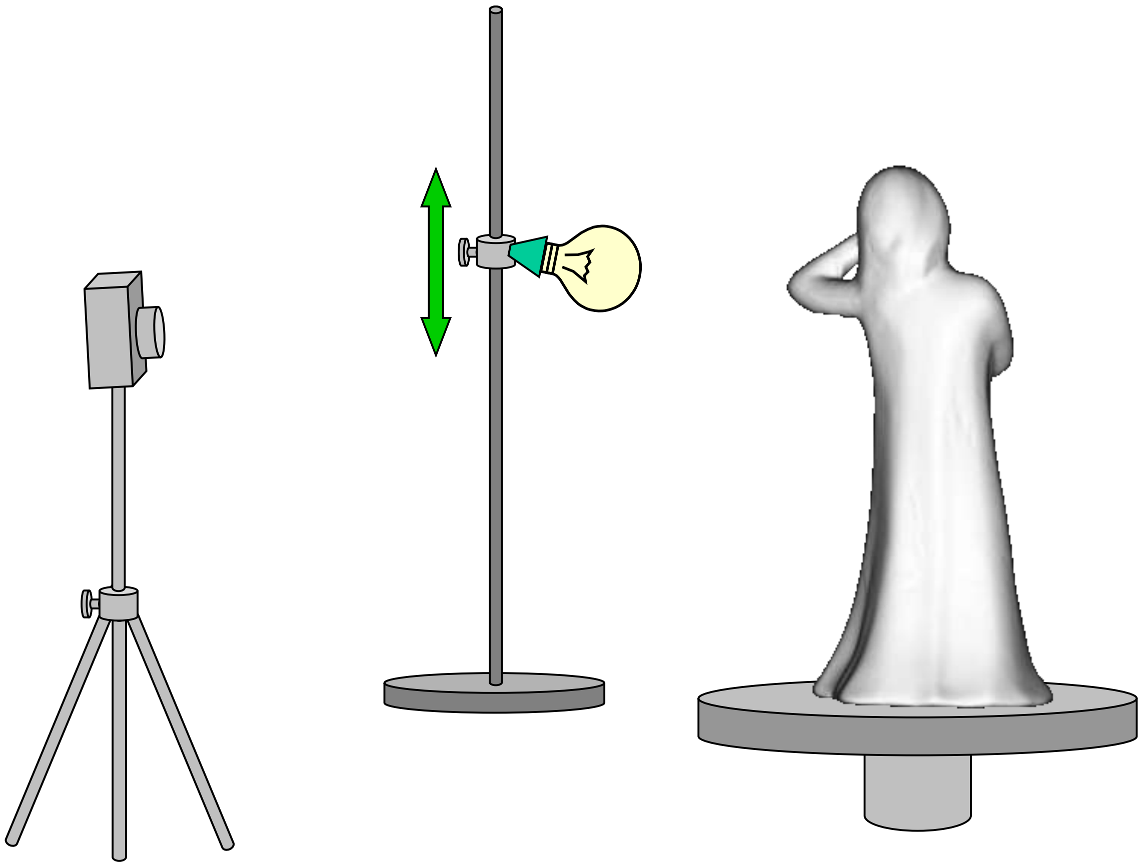


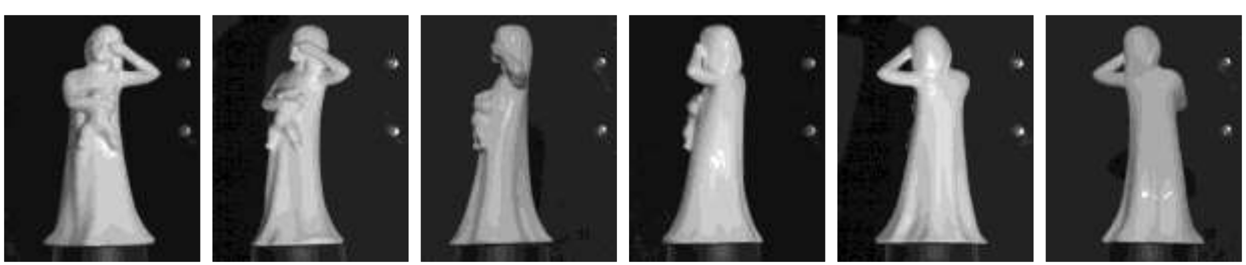
UNIVERSITY OF CAMBRIDGE



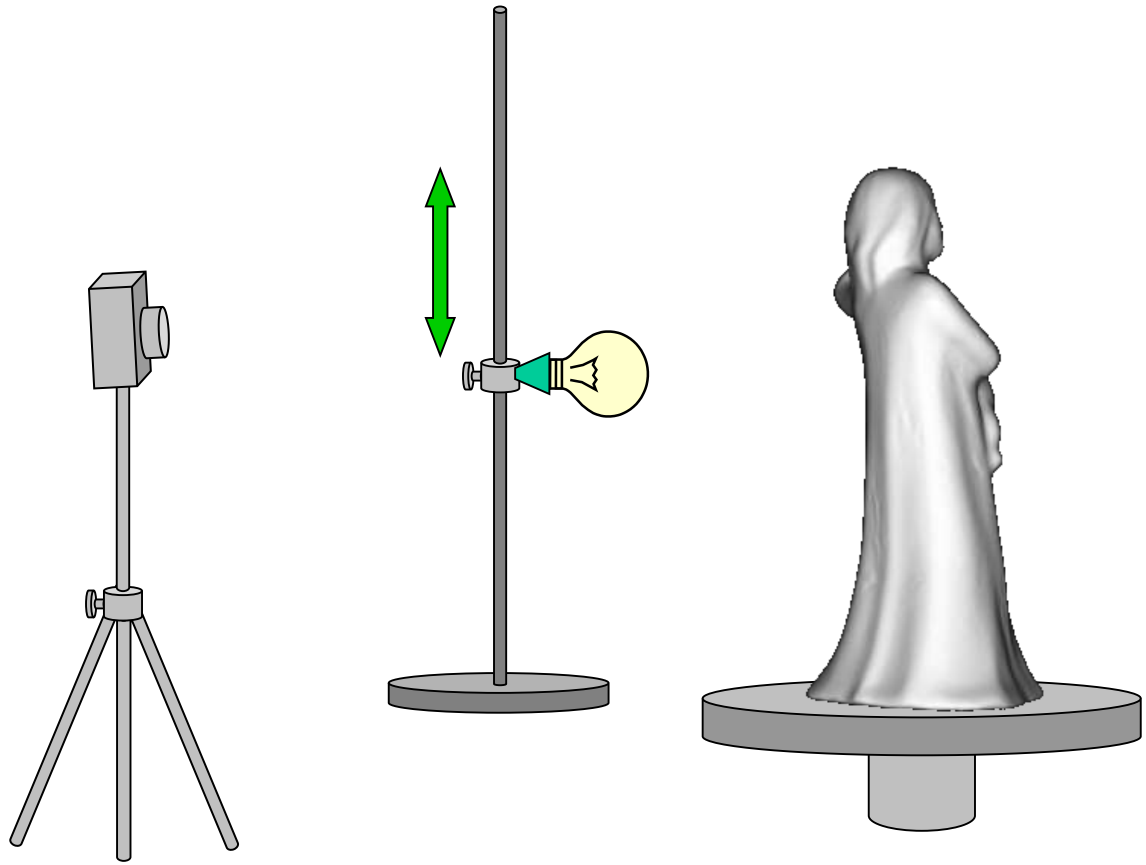


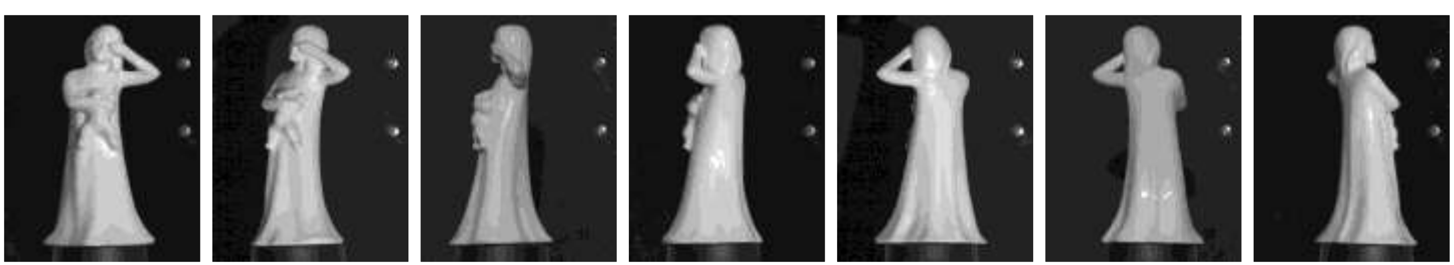
UNIVERSITY OF CAMBRIDGE

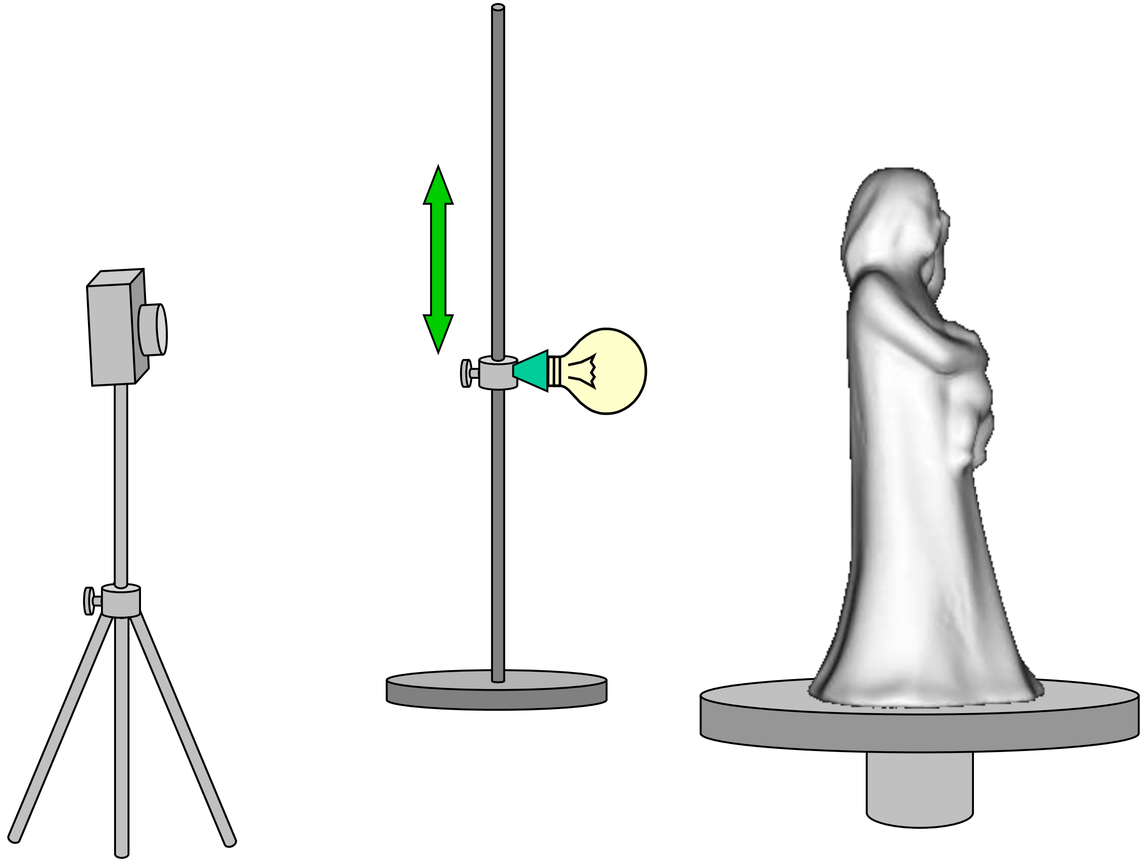


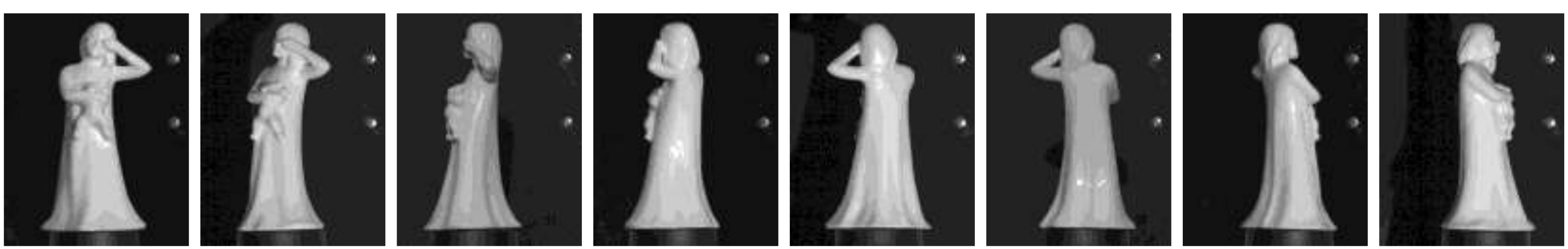


UNIVERSITY OF CAMBRIDGE

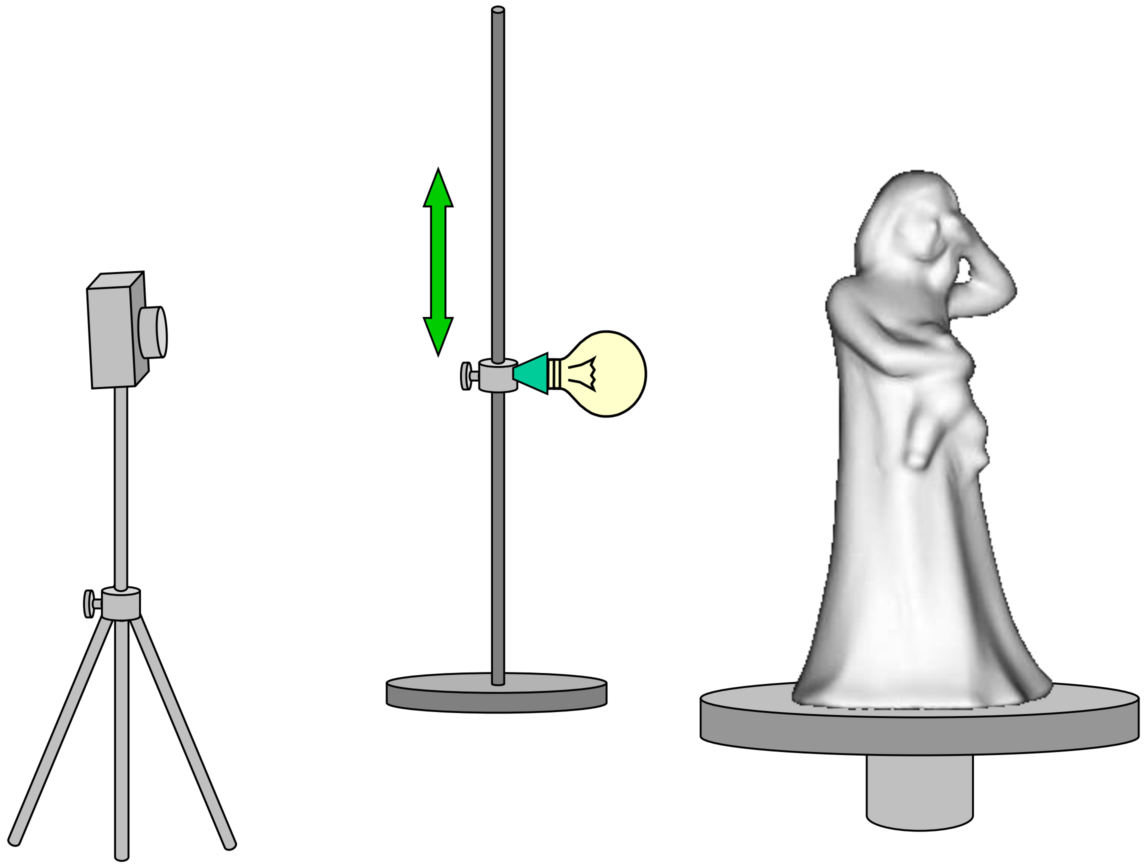


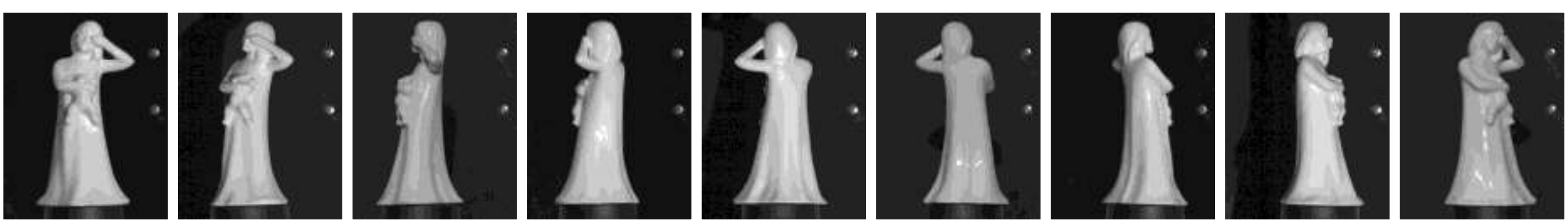




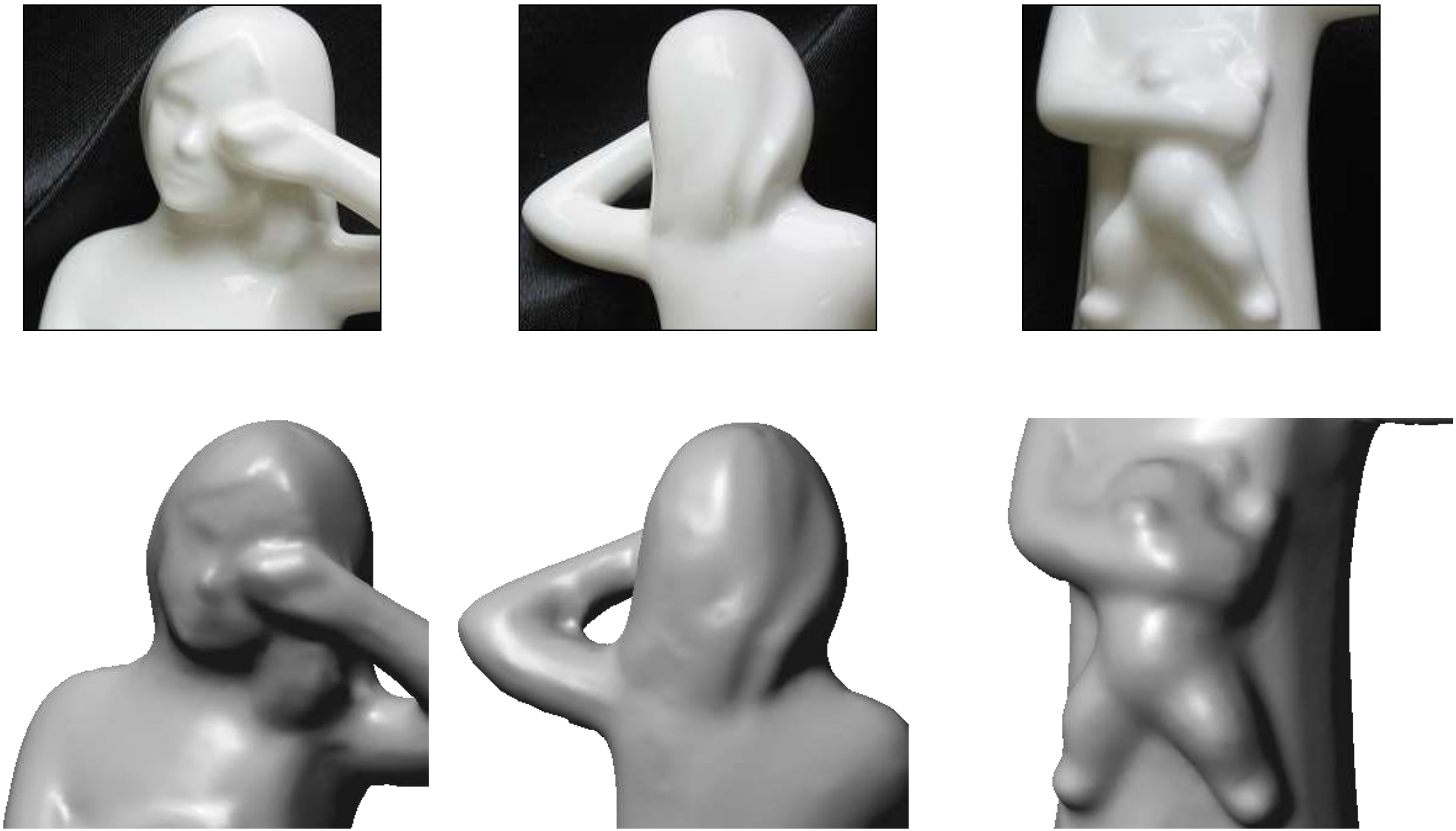


UNIVERSITY OF CAMBRIDGE





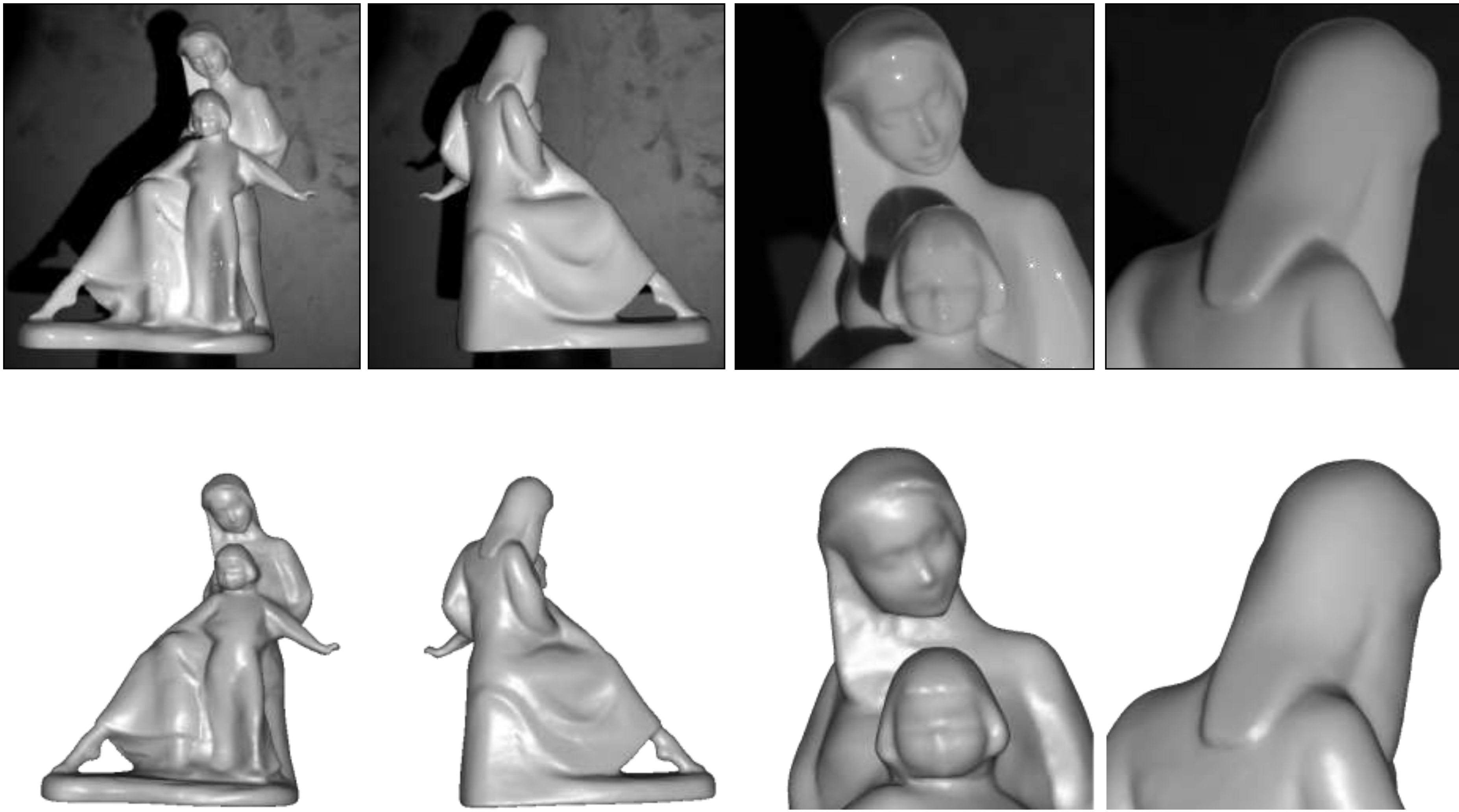
# 3D Models



# Making physical copies

# 3D Models

UNIVERSITY OF CAMBRIDGE



# 3D Models

UNIVERSITY OF CAMBRIDGE

# Deformable objects:

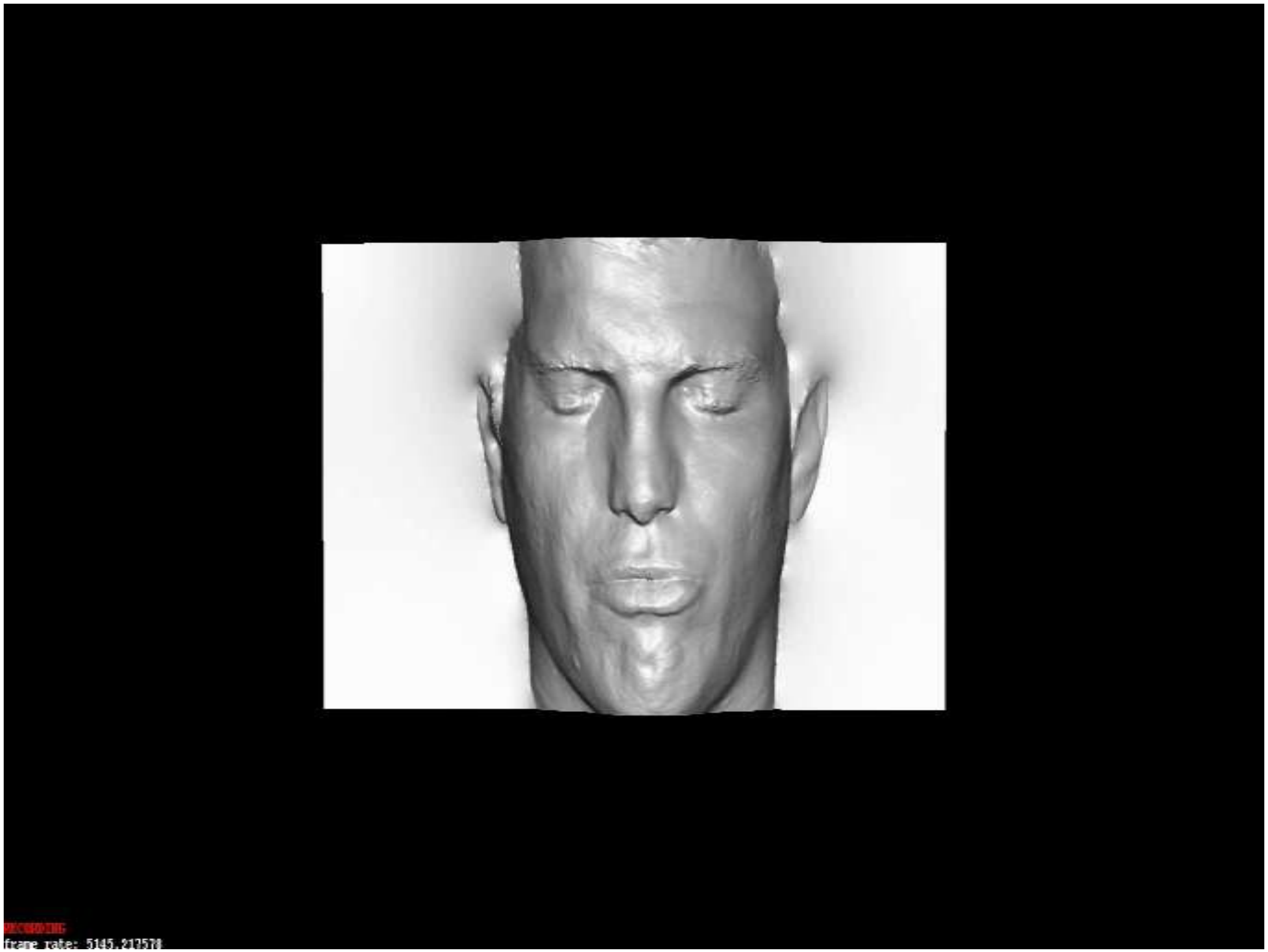
# Real-time photometric stereo using colour lighting

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

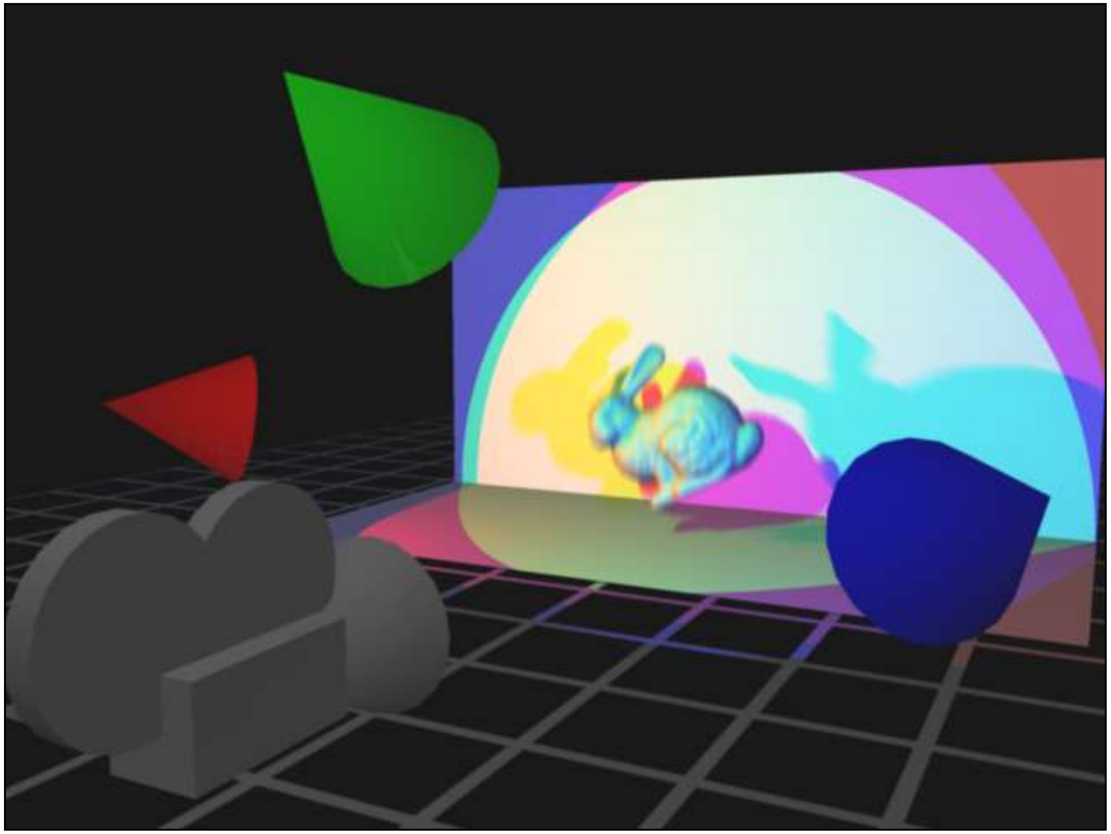
# 3 Untextured and deforming

# Colour Photometric Stereo

# Real-time deformable surfaces



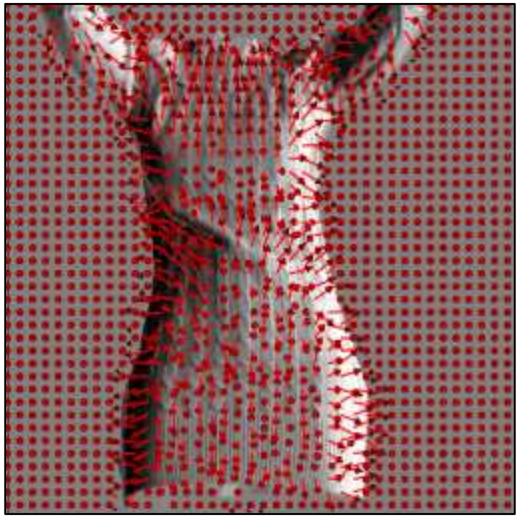
# Textureless deforming objects



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

# Coloured photometric stereo

Single frame
from video



RGB Color is converted to
a normal at each pixel

Normals integrated using
FFT Poisson solver

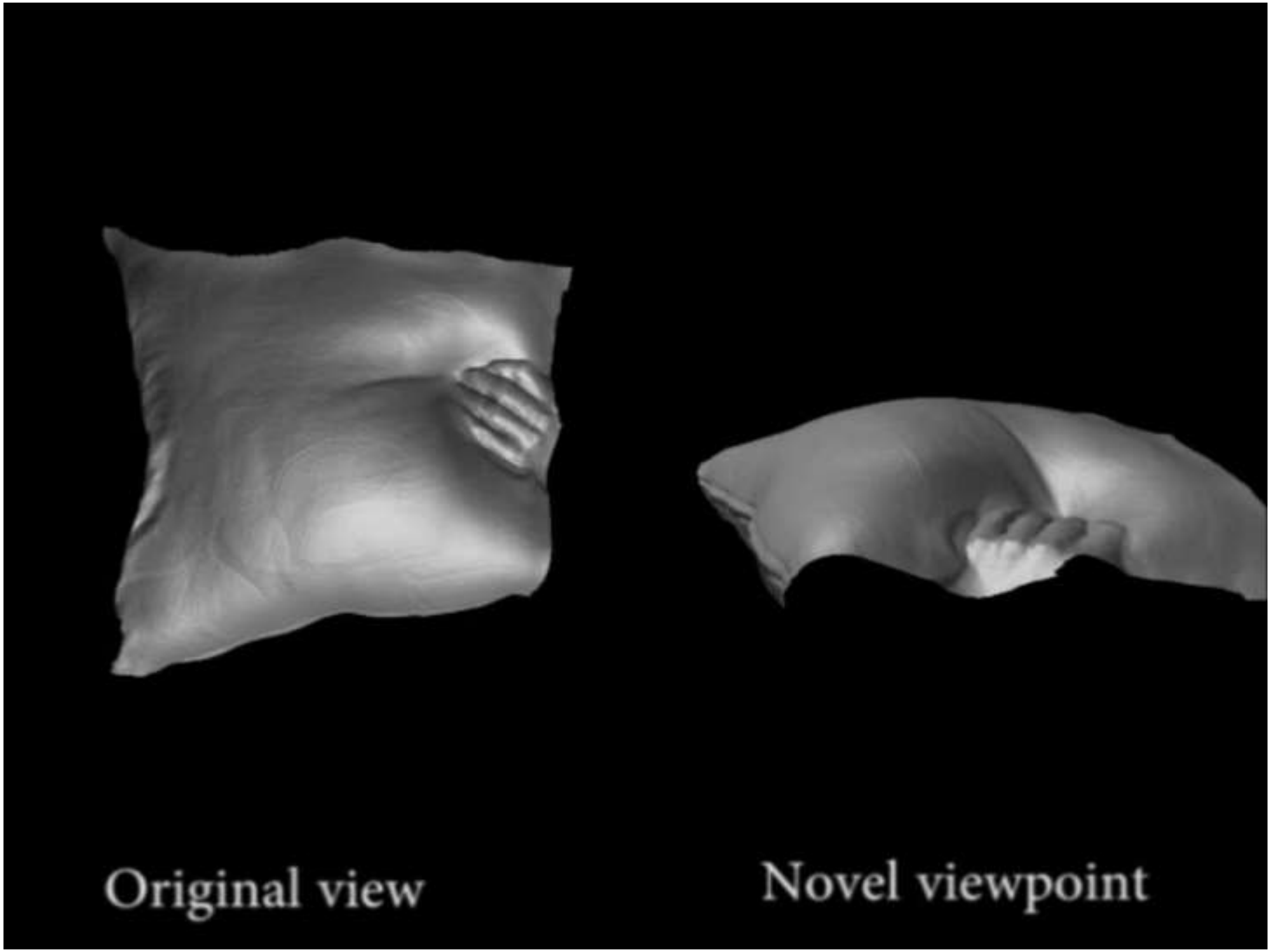
# Results

classic photometric stereo

coloured photometric stereo

# Multicoloured surfaces

# Multicoloured surfaces

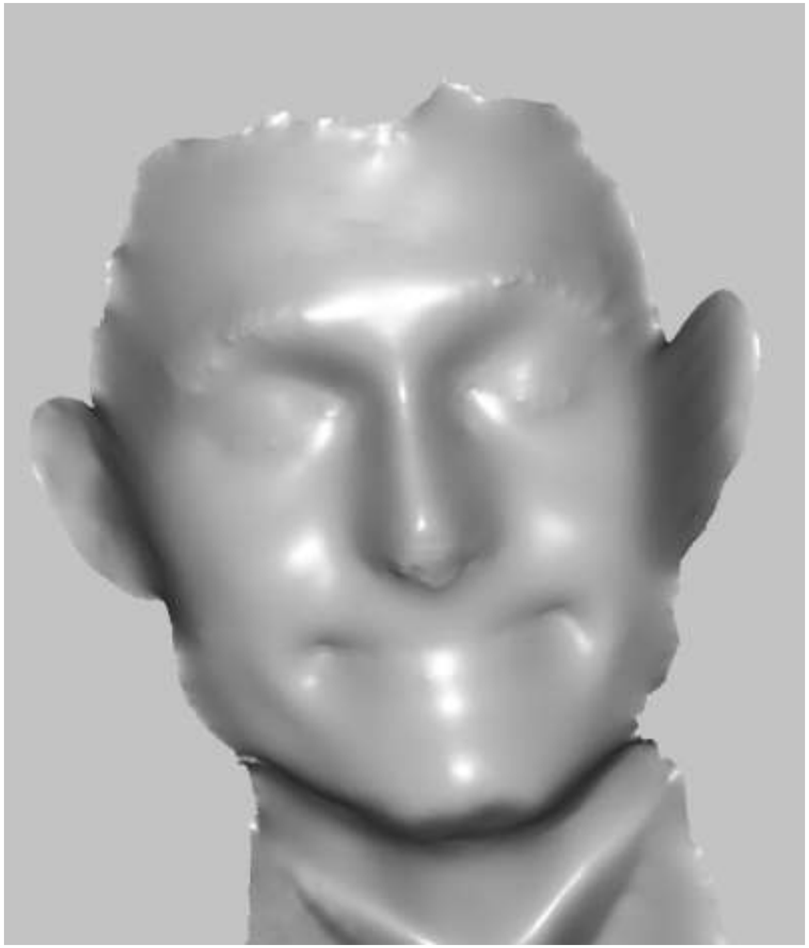



Anderson, Stenger, Cipolla ICCV 2011

# Dynamic Face Capture

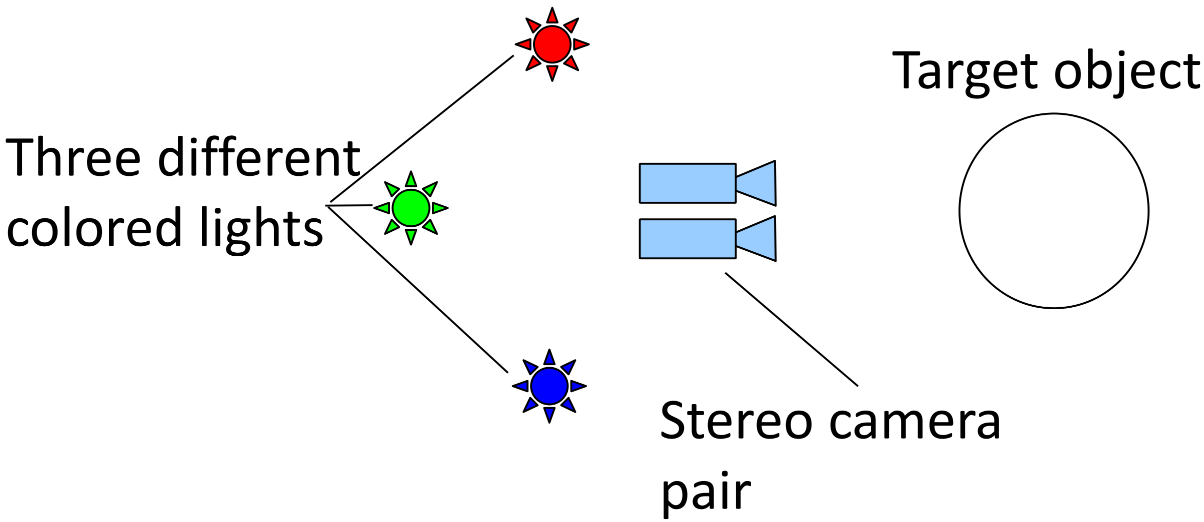
UNIVERSITY OF CAMBRIDGE

- Multiview stereo using two cameras can provide coarse geometry.



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

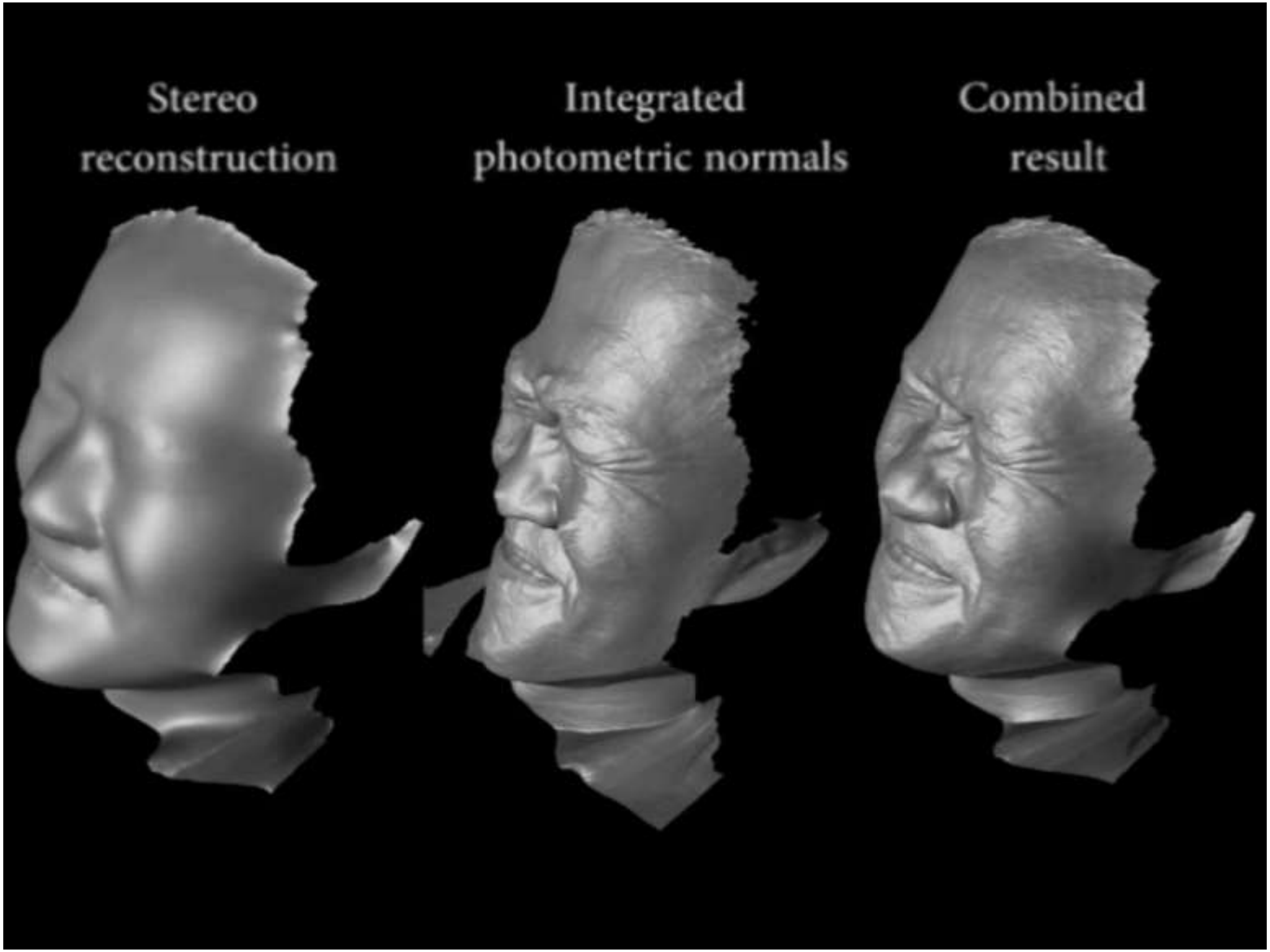
# Equipment



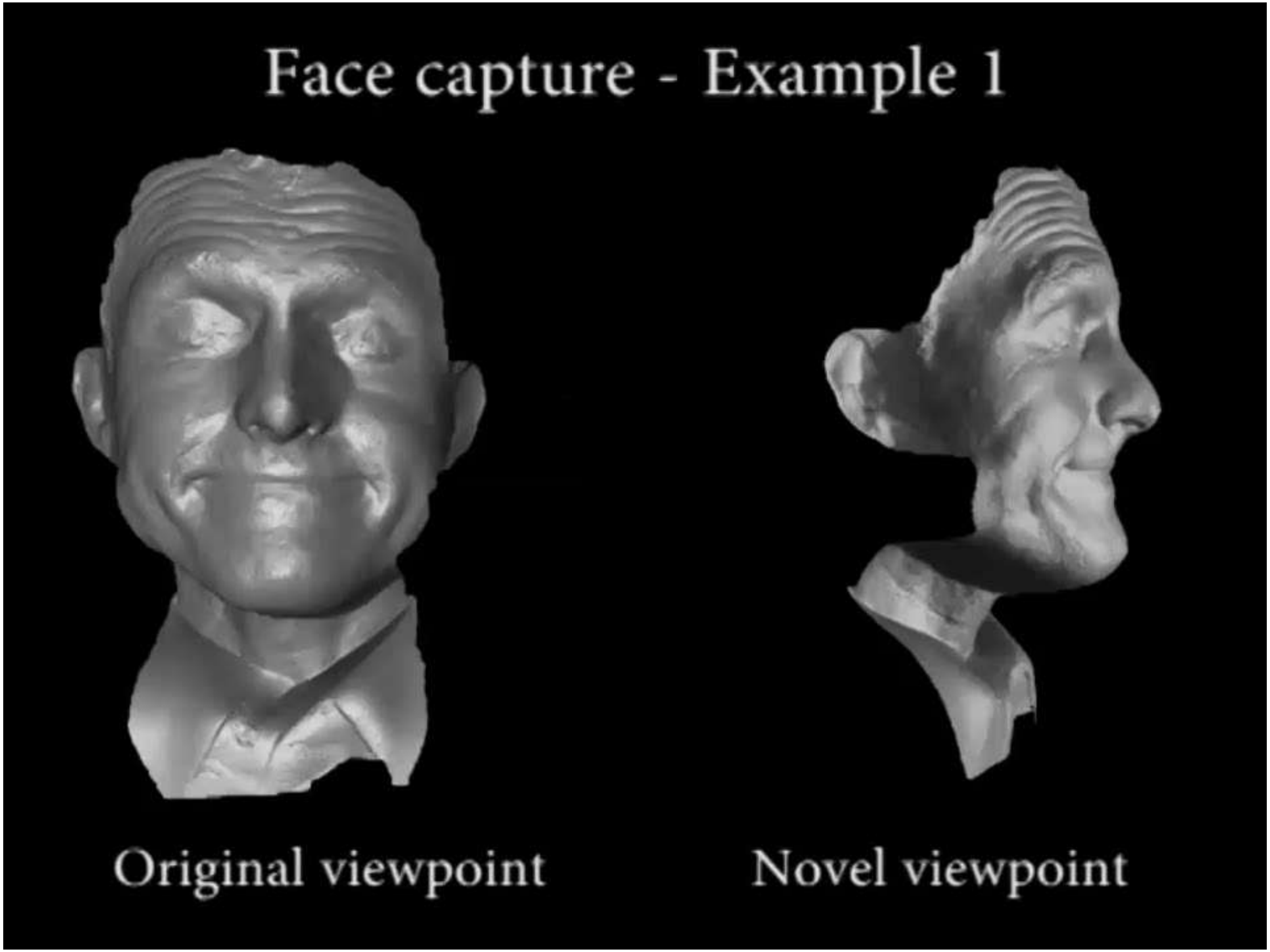

Sample input pair

- 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

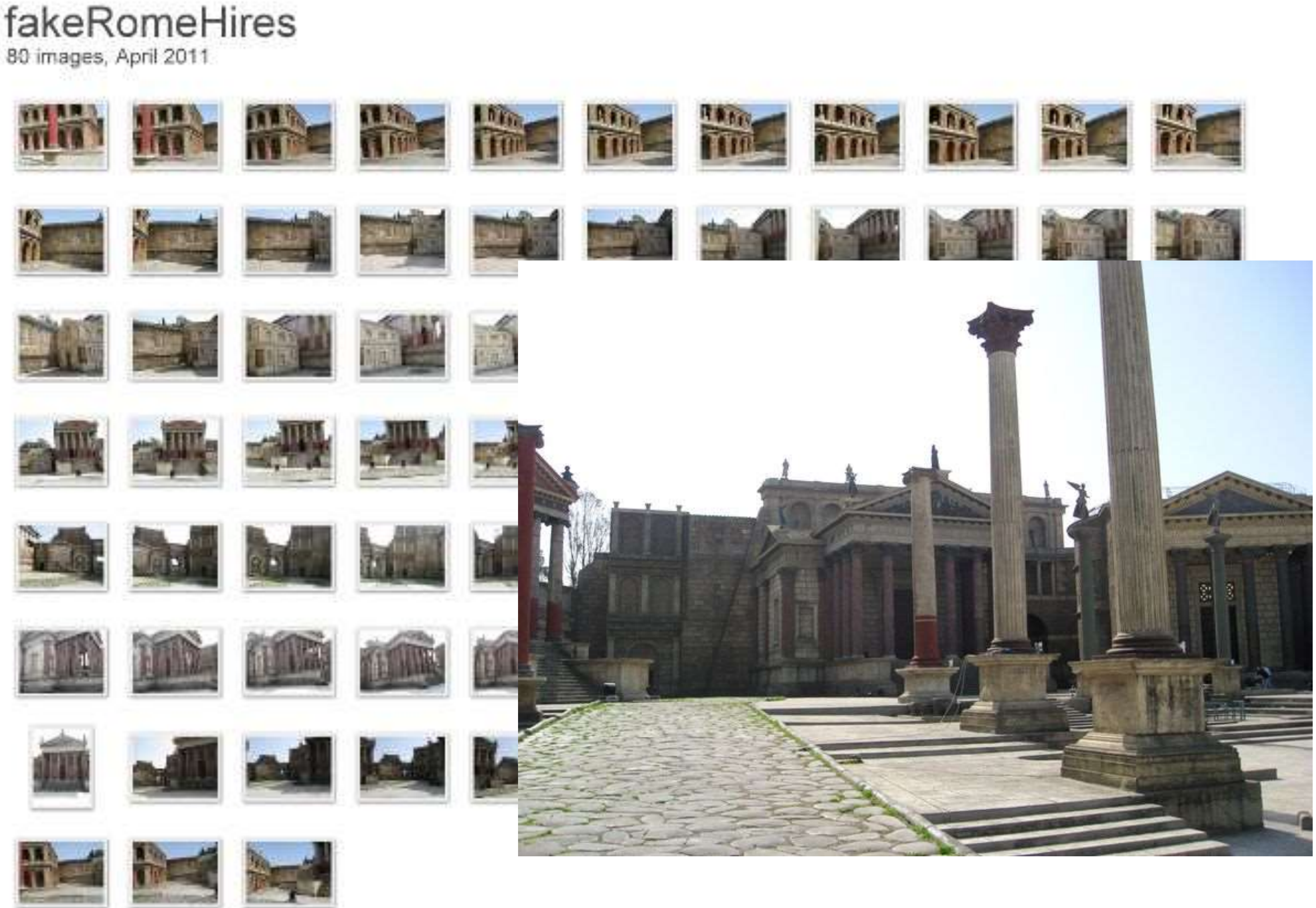


# Sample Reconstructions

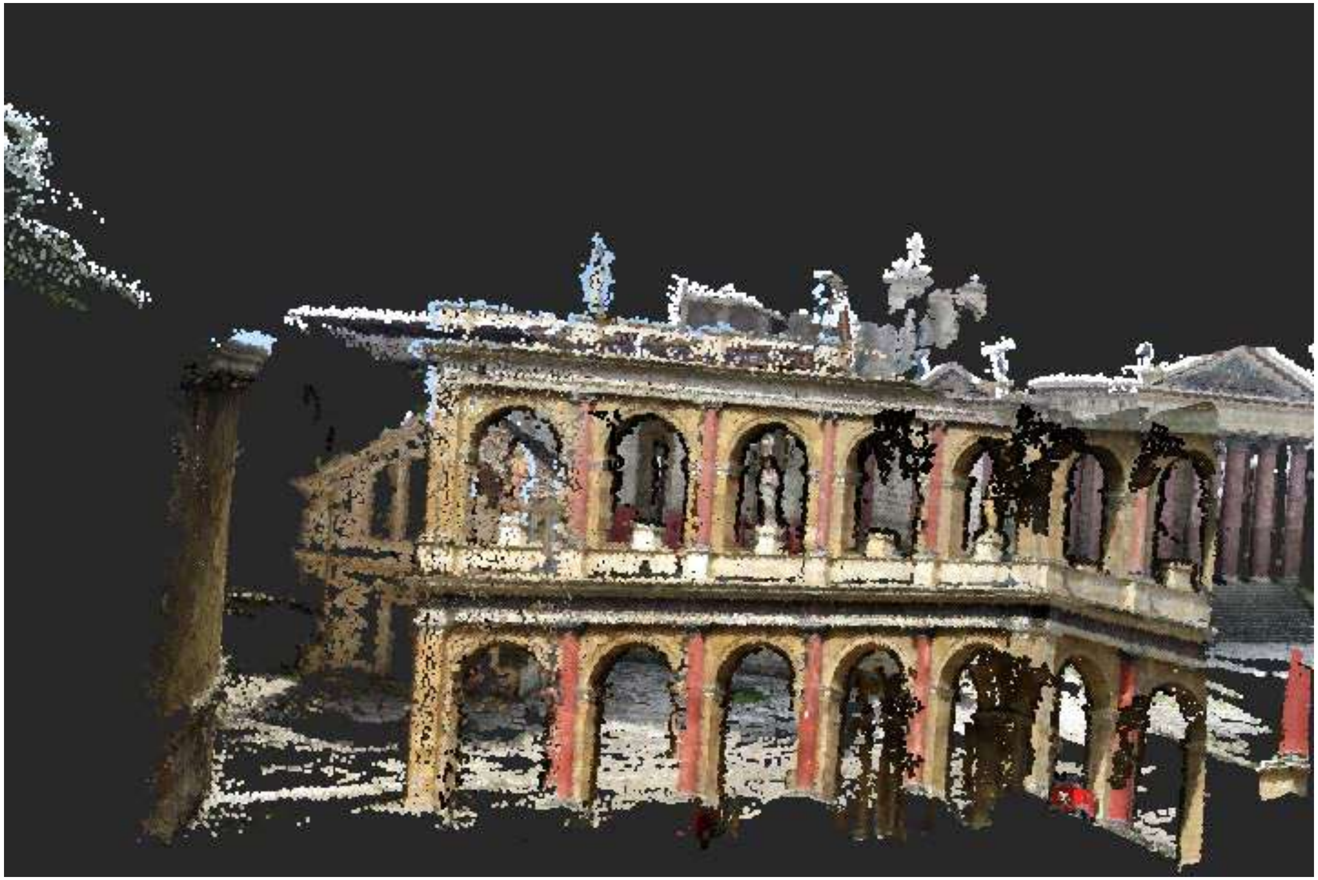


# Large-scale reconstructions

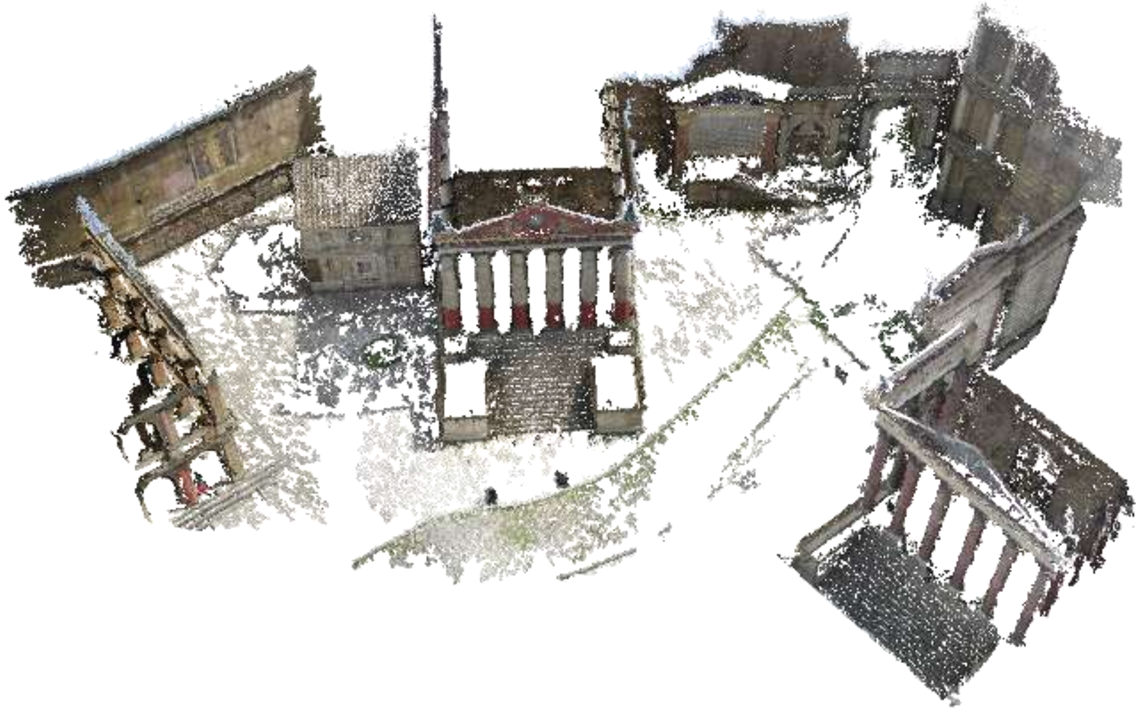
# Large Scale: Reconstruction of Forum Romanum



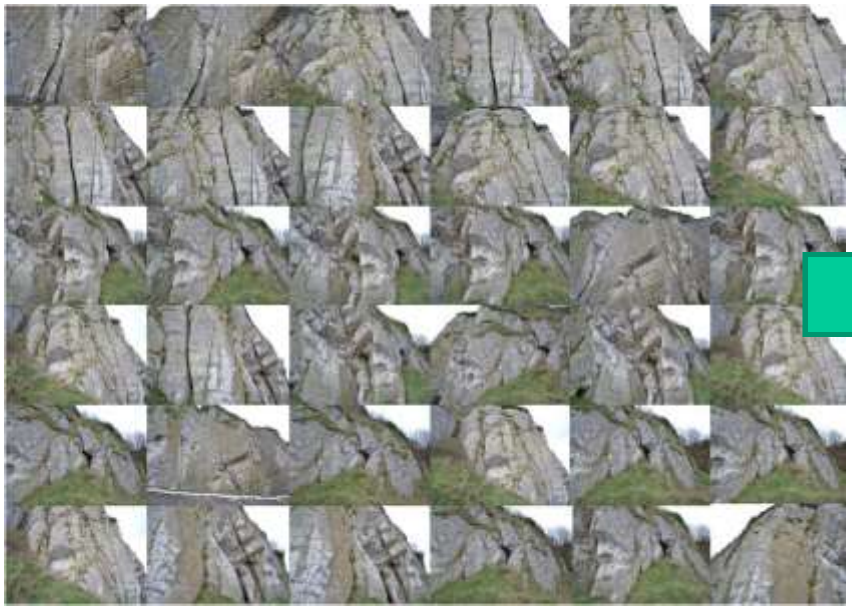
# Large-scale reconstruction



# Large-scale reconstruction



# Outdoor reconstructions



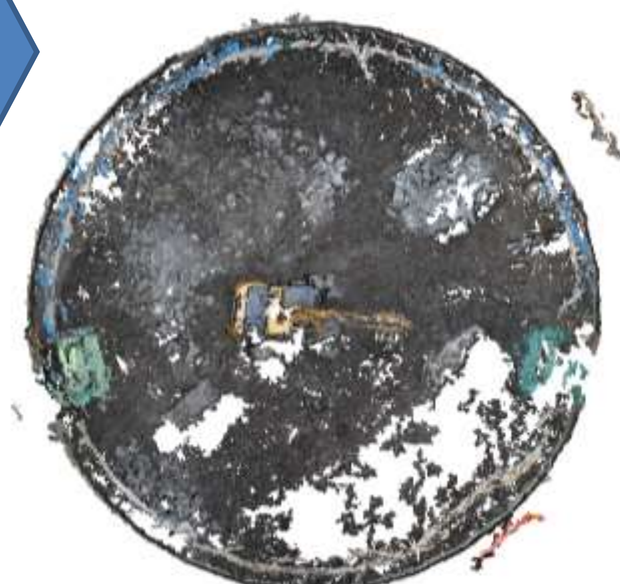
**80 images from digital camera**

**Point cloud reconstruction of cliff-face**

# CSIC - Construction Progress Monitoring

- Cambridge Heath Excavation site

- 100-200 photos collected of site per day over the course of 15+ days of activity



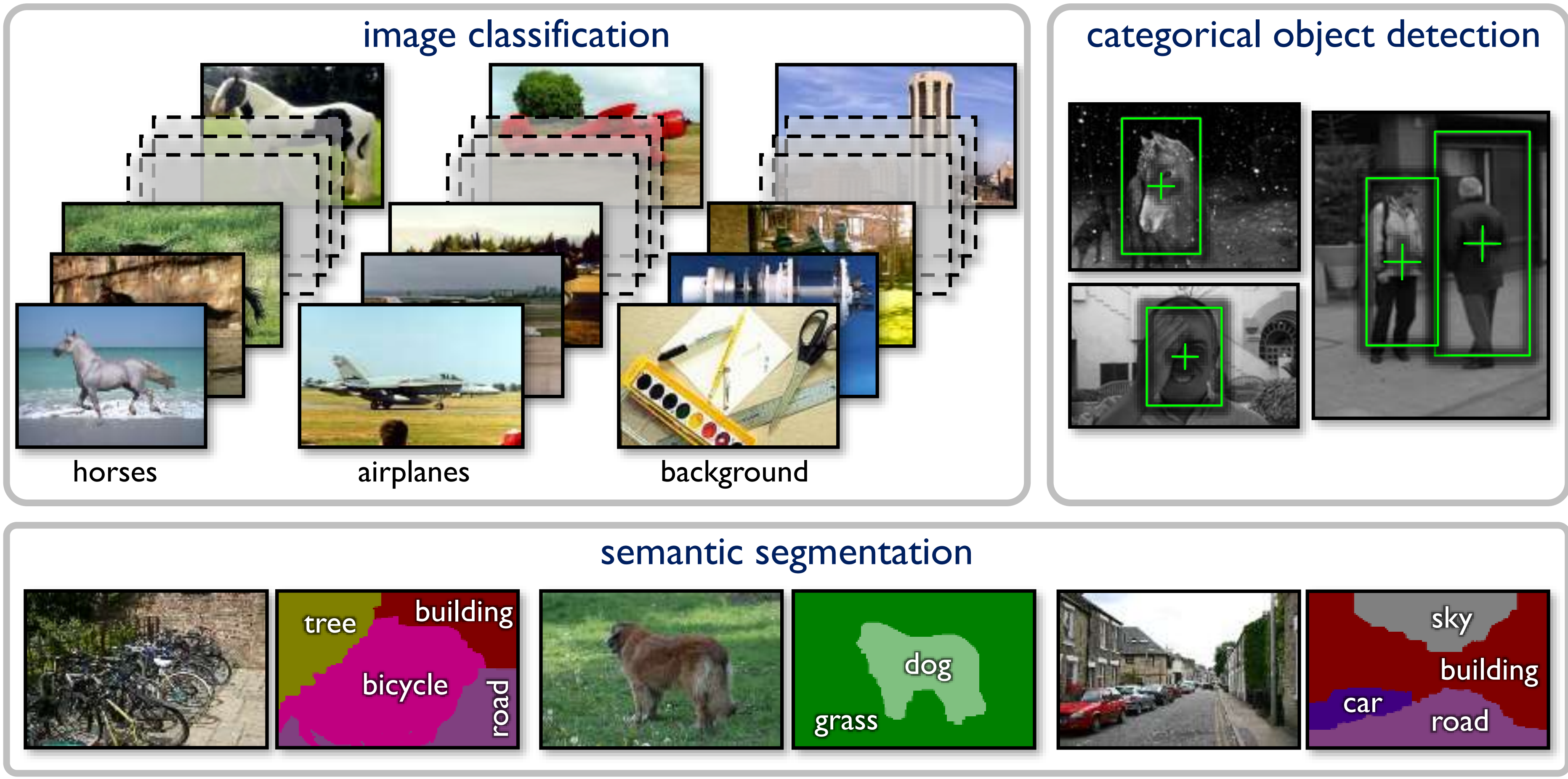
With this unique dataset, we intend to carry out:

- As-excavated volumetric analysis using both image data and LIDAR data
- “As-built” vs. “as-planned” difference detection
- Comparison study between the use of LIDAR vs. images for progress monitoring

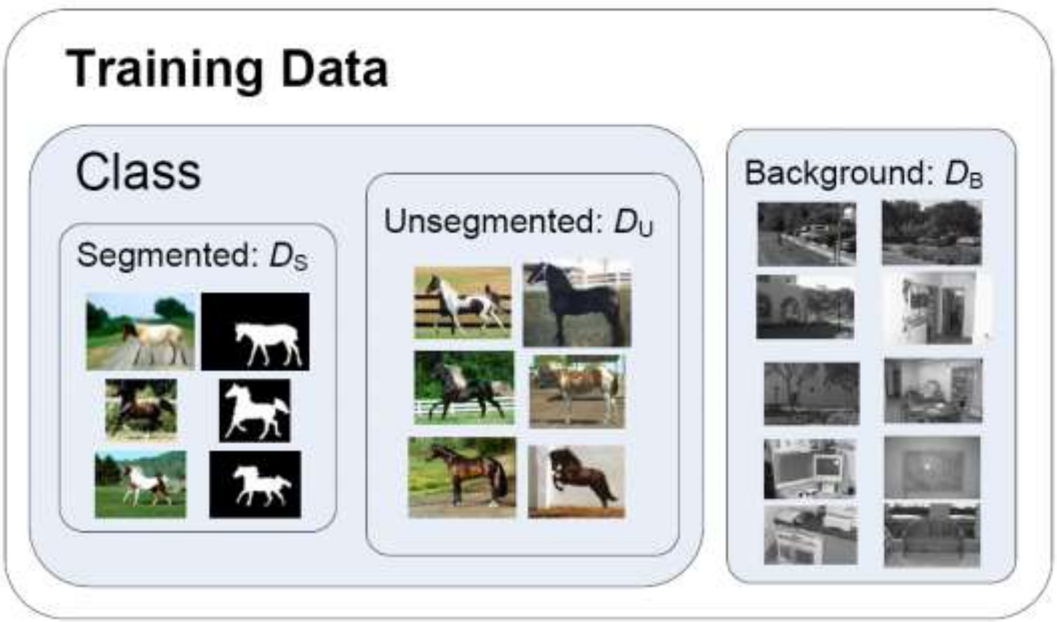
01/15 11-02-13

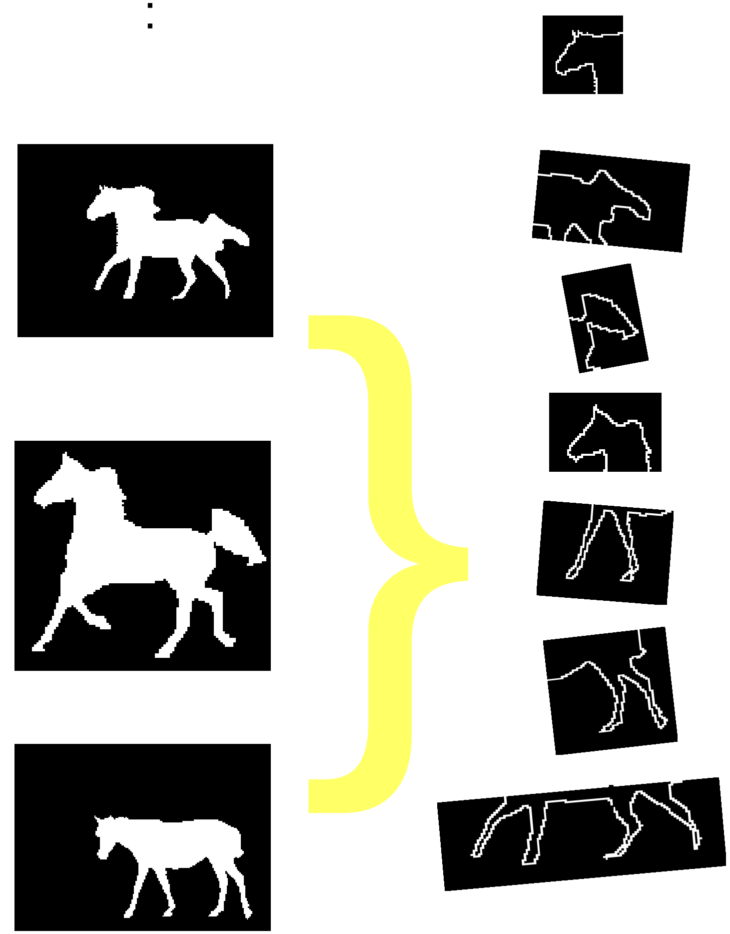

# Recognition?

# Recognition

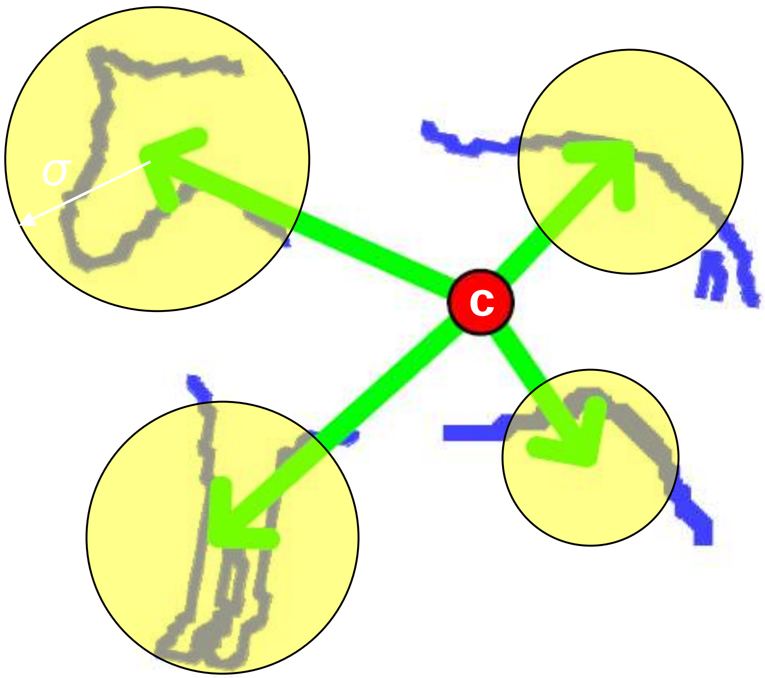


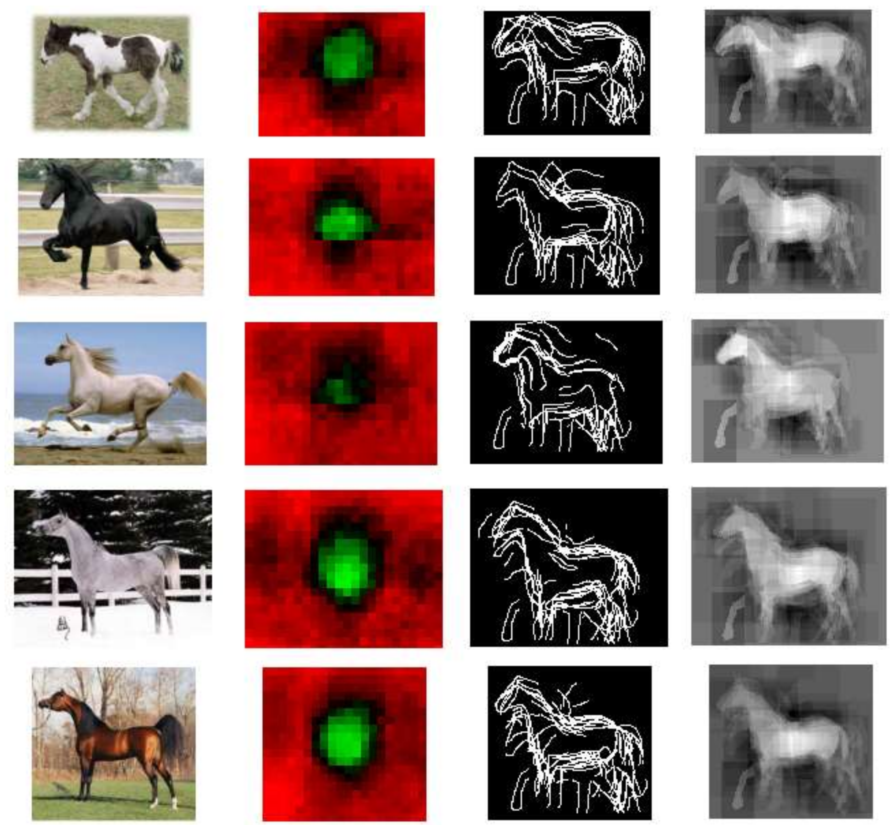
# Supervised Learning

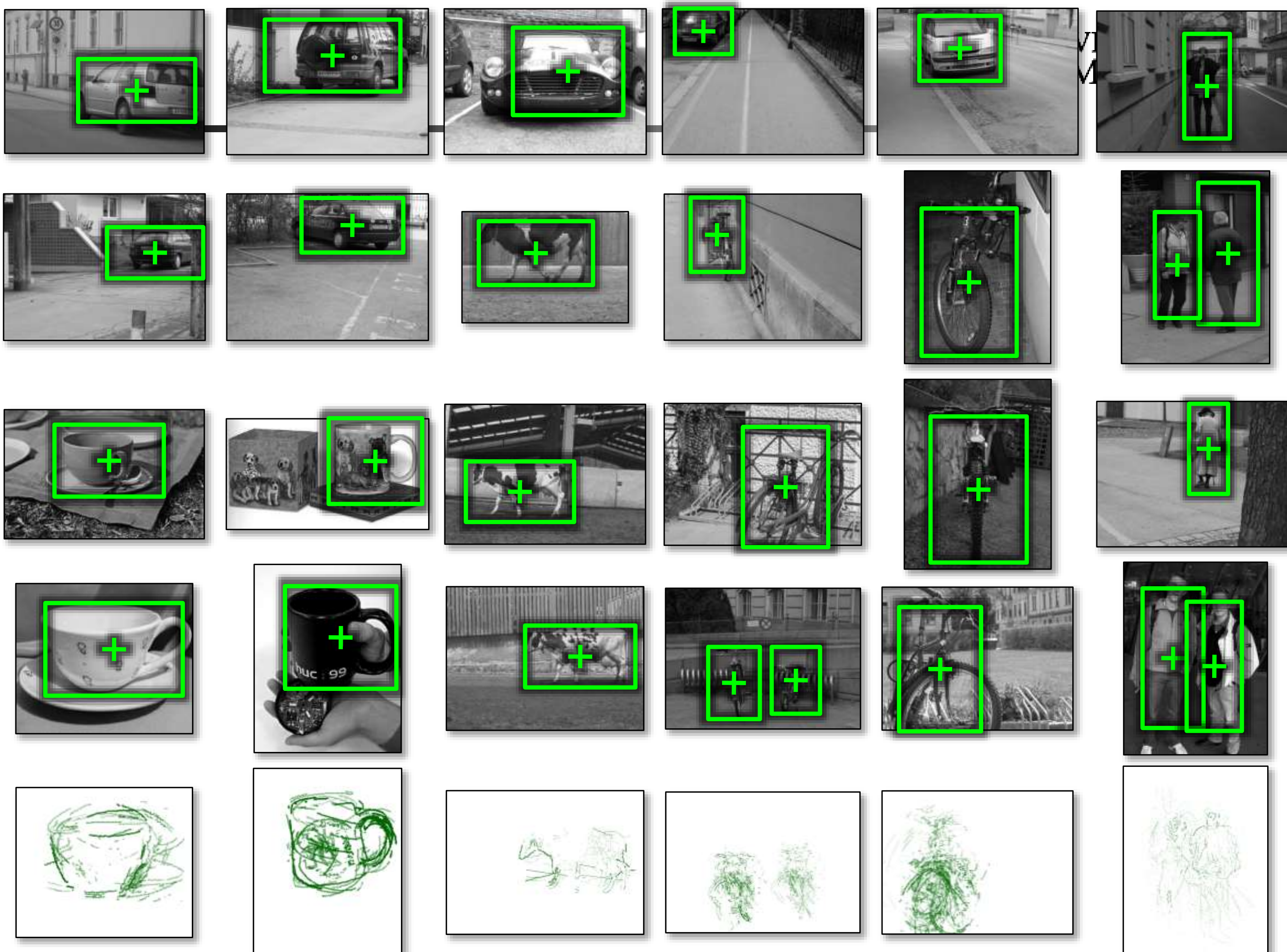


# Object Model

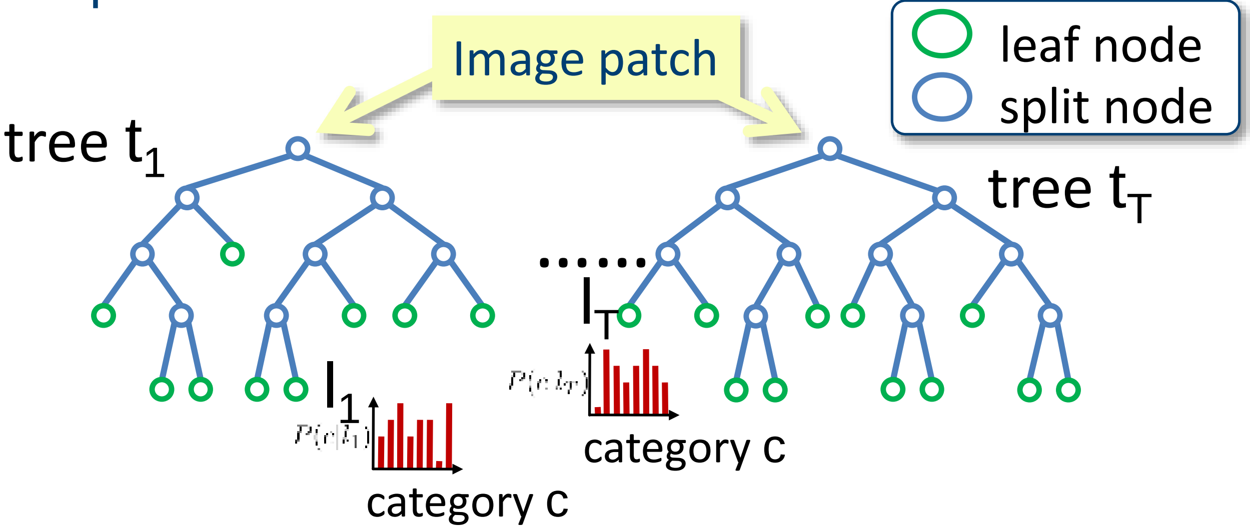




# Semantic Texton Forests for classification

- 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]

# Semantic Texton Forests for classification

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

kinect sensor

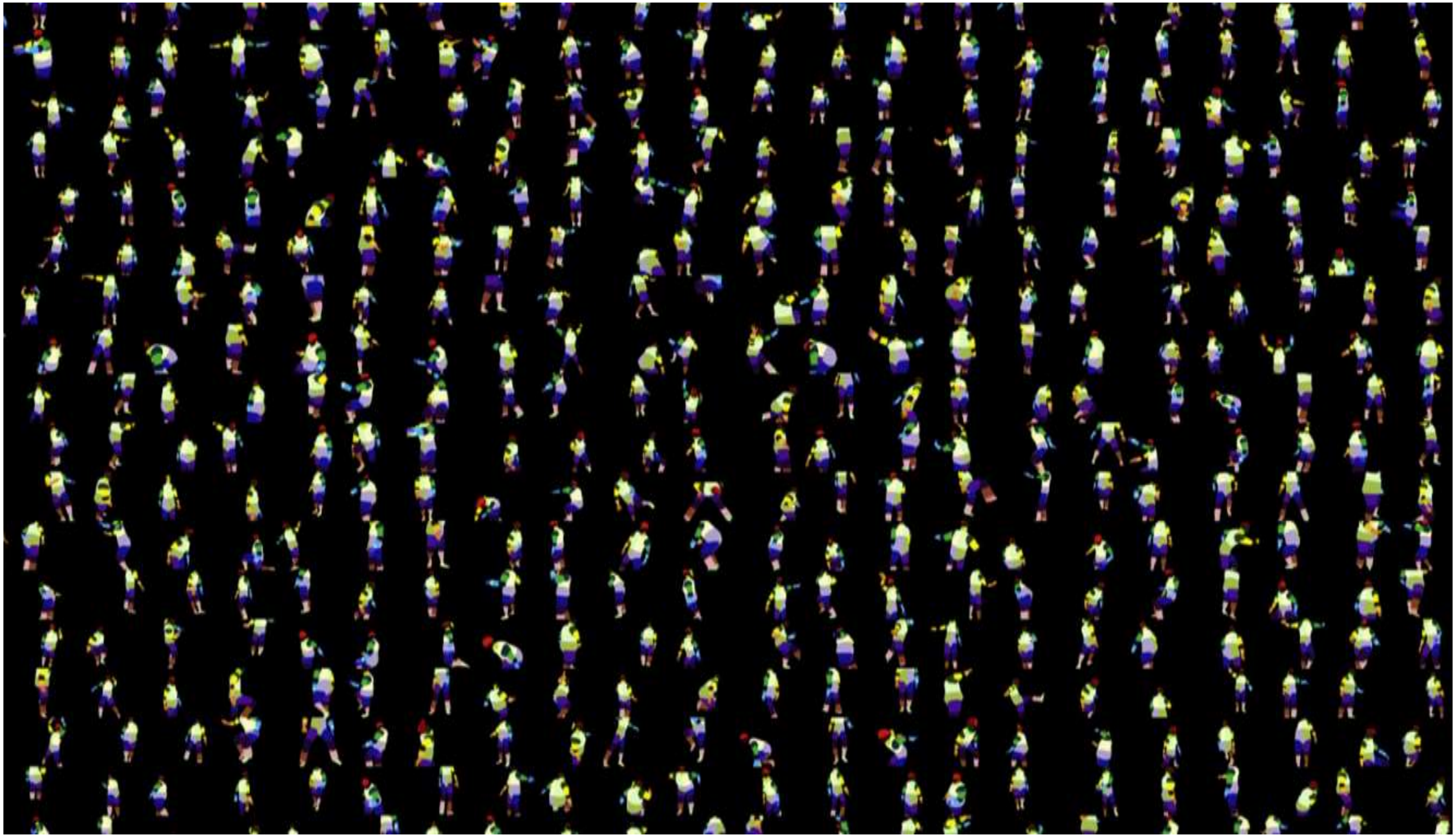

Depth image

Body part recognition

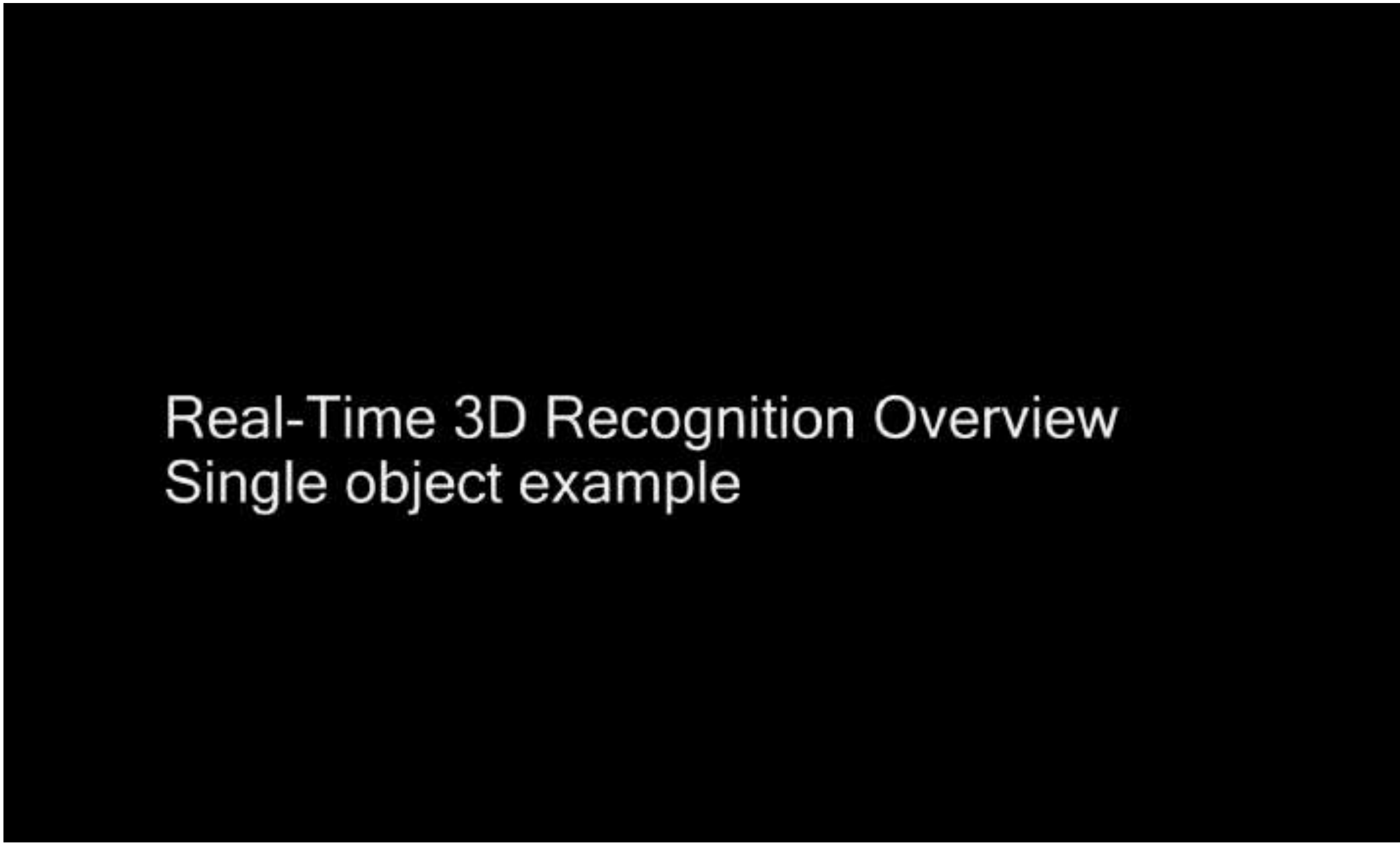
Shotton, Fitzgibbon et. al, Real-Time Human Pose Recognition in Parts from a Single Depth Image, CVPR’11.

Shotton, Johnson & Cipolla, Semantic Texton Forests for Image Categorization and Segmentation, CVPR’08.

# Synthetic training data



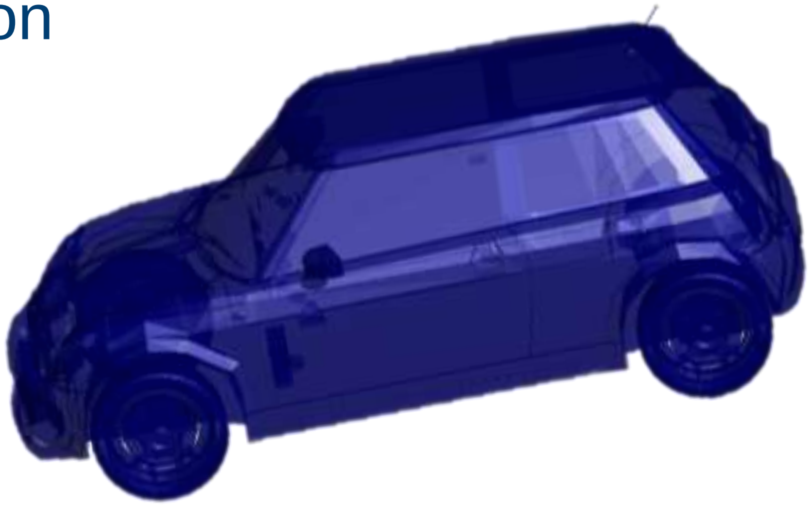
# 3D object recognition



# 3R’s – live demonstration

UNIVERSITY OF CAMBRIDGE

Reconstruction

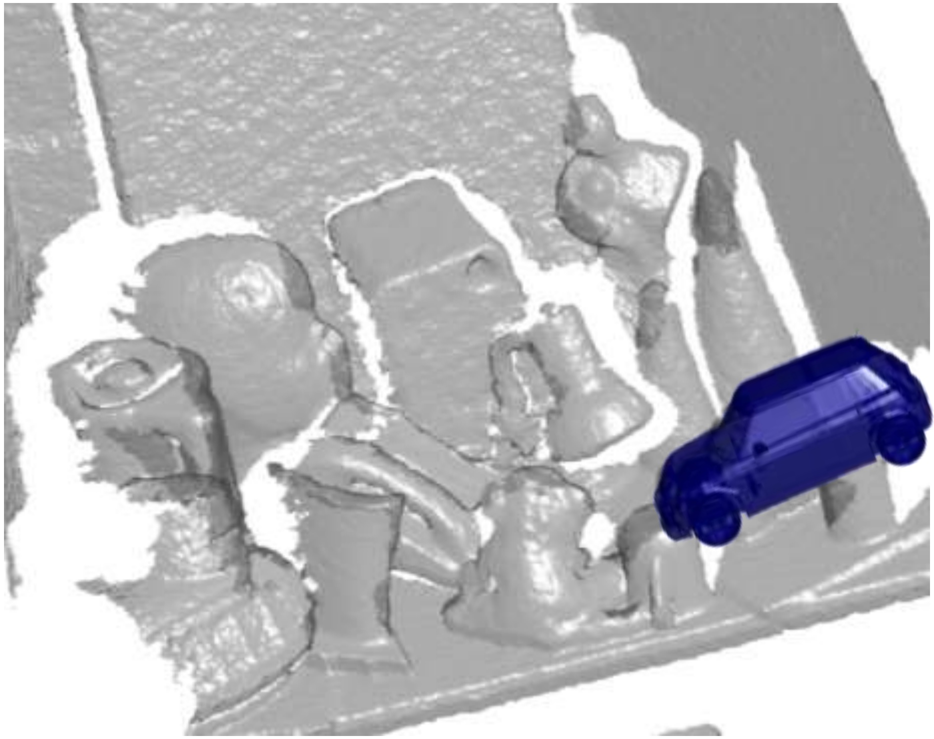


# 3R's – live demonstration

Recognition

# 3R's – live demonstration

Register



# Semantic Segmentation – Label Propagation

# Interactive Video Segmentation

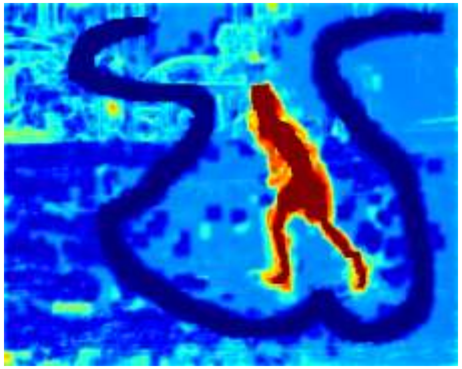
## Video

## User labels


# Our framework

**Video**

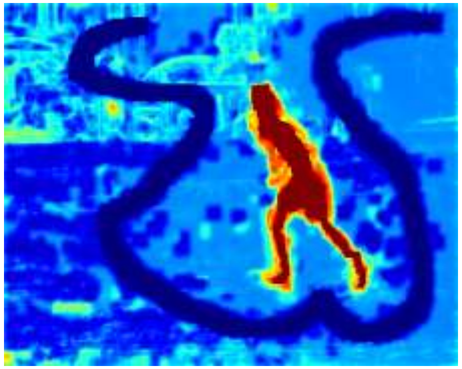
**Temporal label propagation**

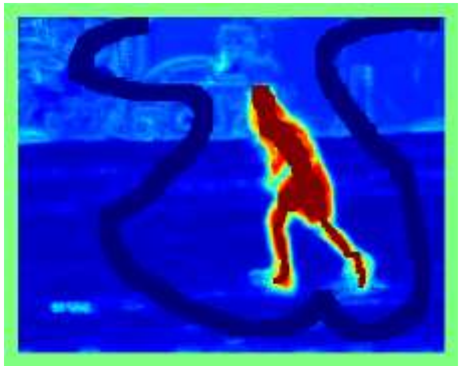


**User labels**


# Our framework

**Video** | **Temporal label propagation** | **Semi-supervised classifier learning**



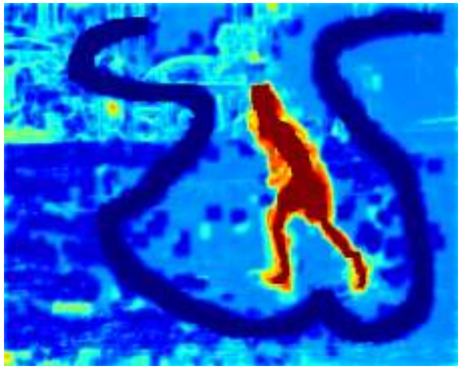


**User labels**

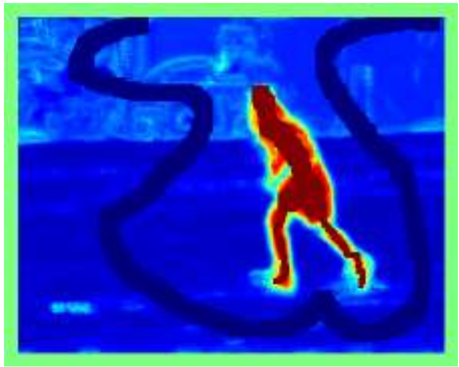

# Our framework

**Video**

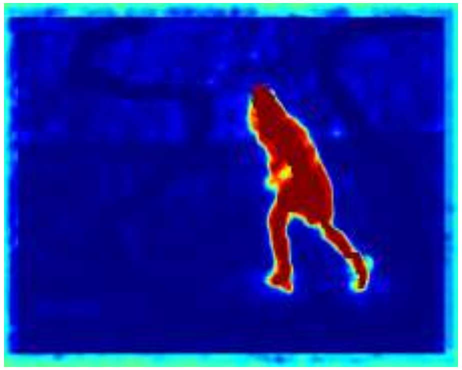
**Temporal label propagation**



**Semi-supervised classifier learning**



**Bootstrapping the classifier**



**User labels**


# Summary – Computer Vision

UNIVERSITY OF CAMBRIDGE

3R's of Computer Vision:

- Registration
- Reconstruction
- Recognition

http://www.eng.cam.ac.uk/~cipolla/people.html

Bjorn Stenger, Carlos Hernandez, George Vogiatzis,
Riccardo Gherardi, Frank Pebert

Ujwal Bonde, Ignas Budyvitis, Yu Chen,
Simon Stent and Simon Taylor

Duncan Robertson and Jamie Shotton