UNIVERSITY OF CAMBRIDGE

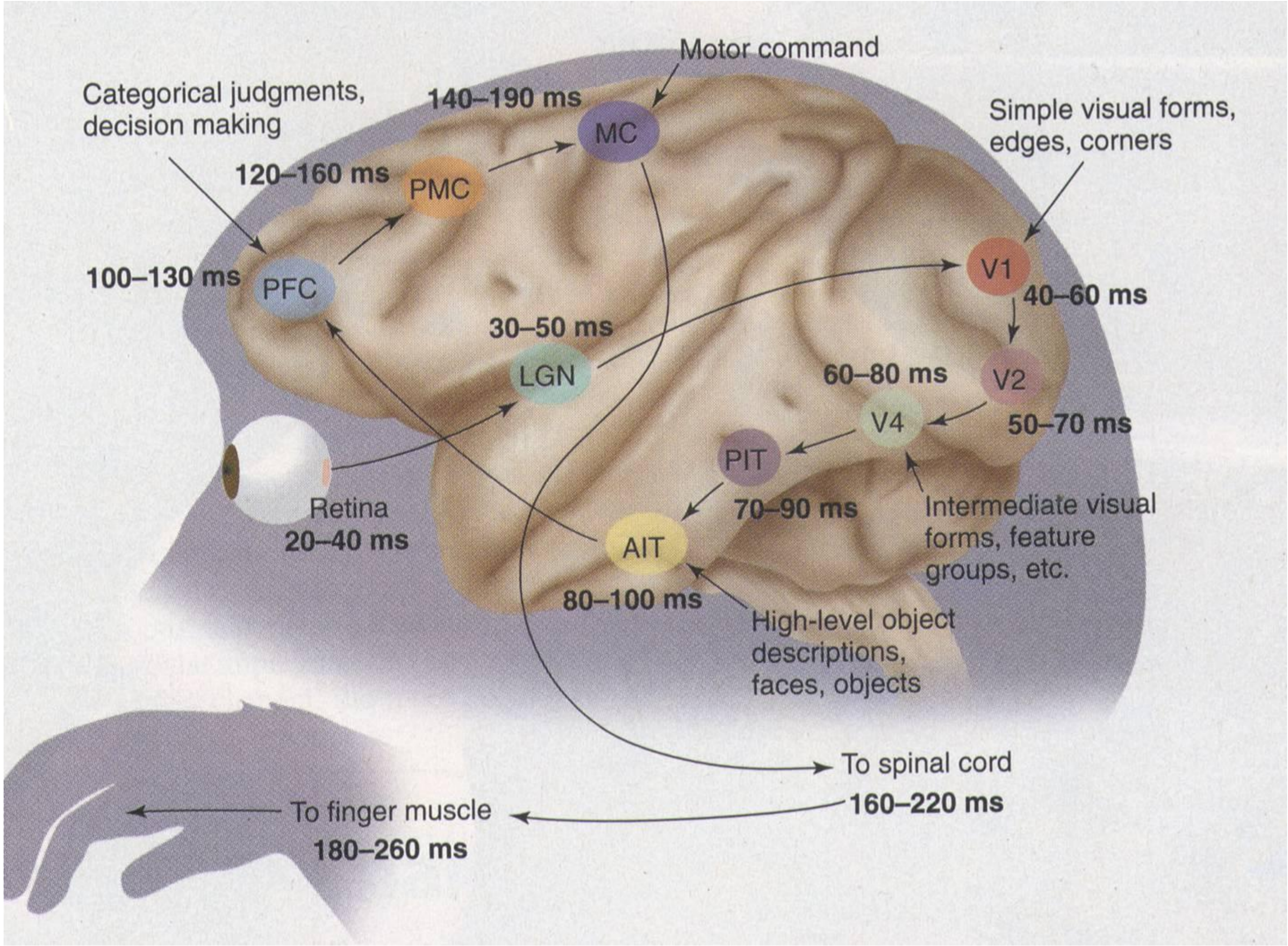
Making Machines See

Roberto Cipolla
Department of Engineering

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

Cognitive Systems Engineering

UNIVERSITY OF CAMBRIDGE



Making machines see

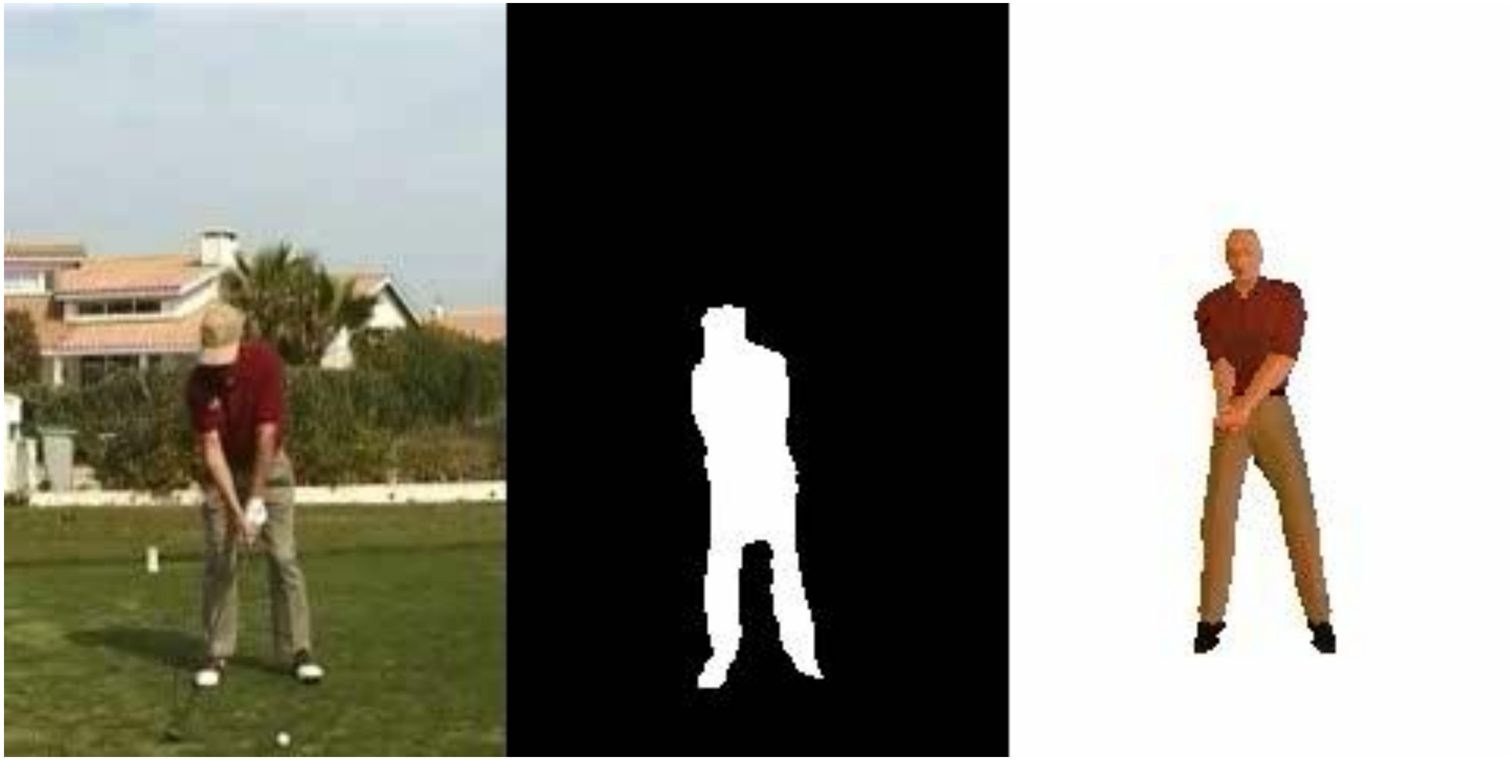
- Vision: What, Why and How?
- 3Rs of computer vision:
  - Reconstruction
  - Registration
  - Recognition

Registration?

Target detection and pose estimation

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

Registration

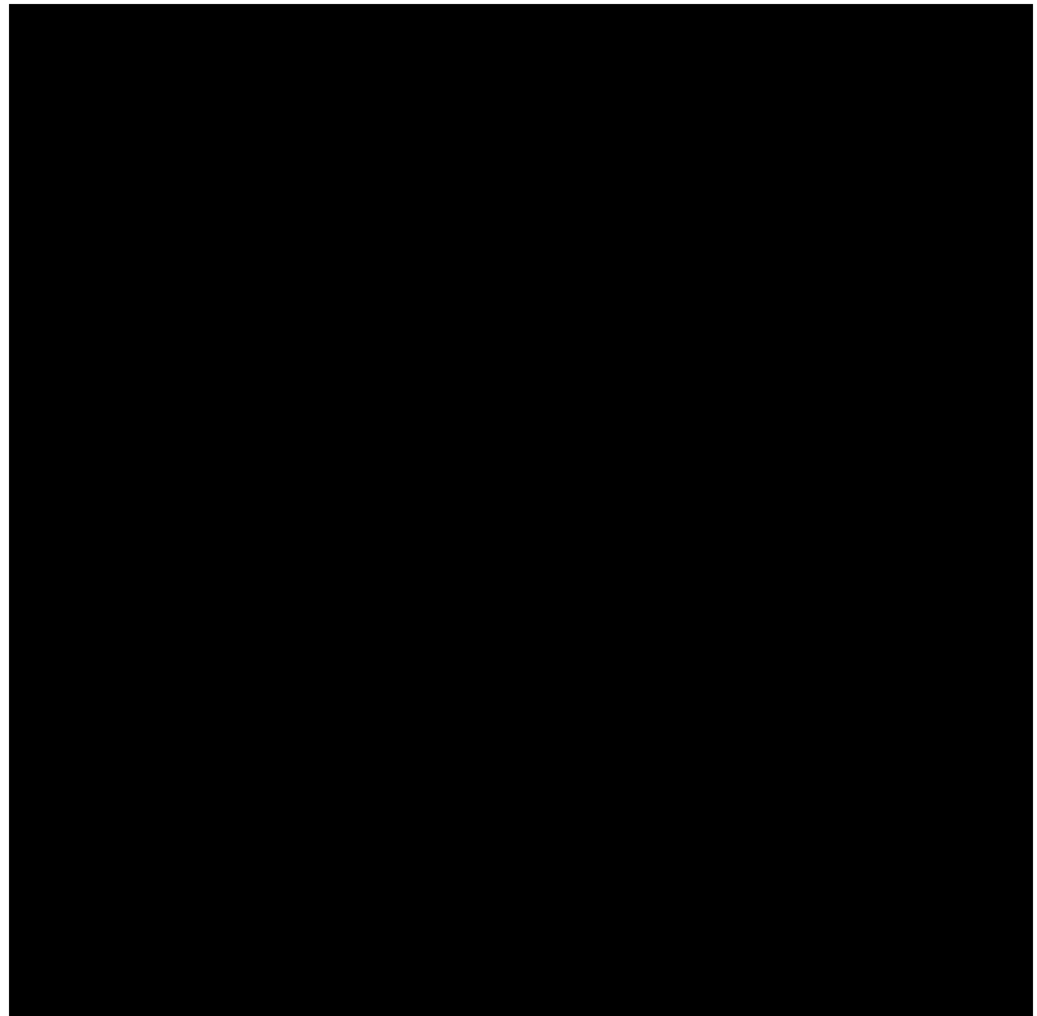


Registration

Reconstruction?

Recovery of 3D shape from images

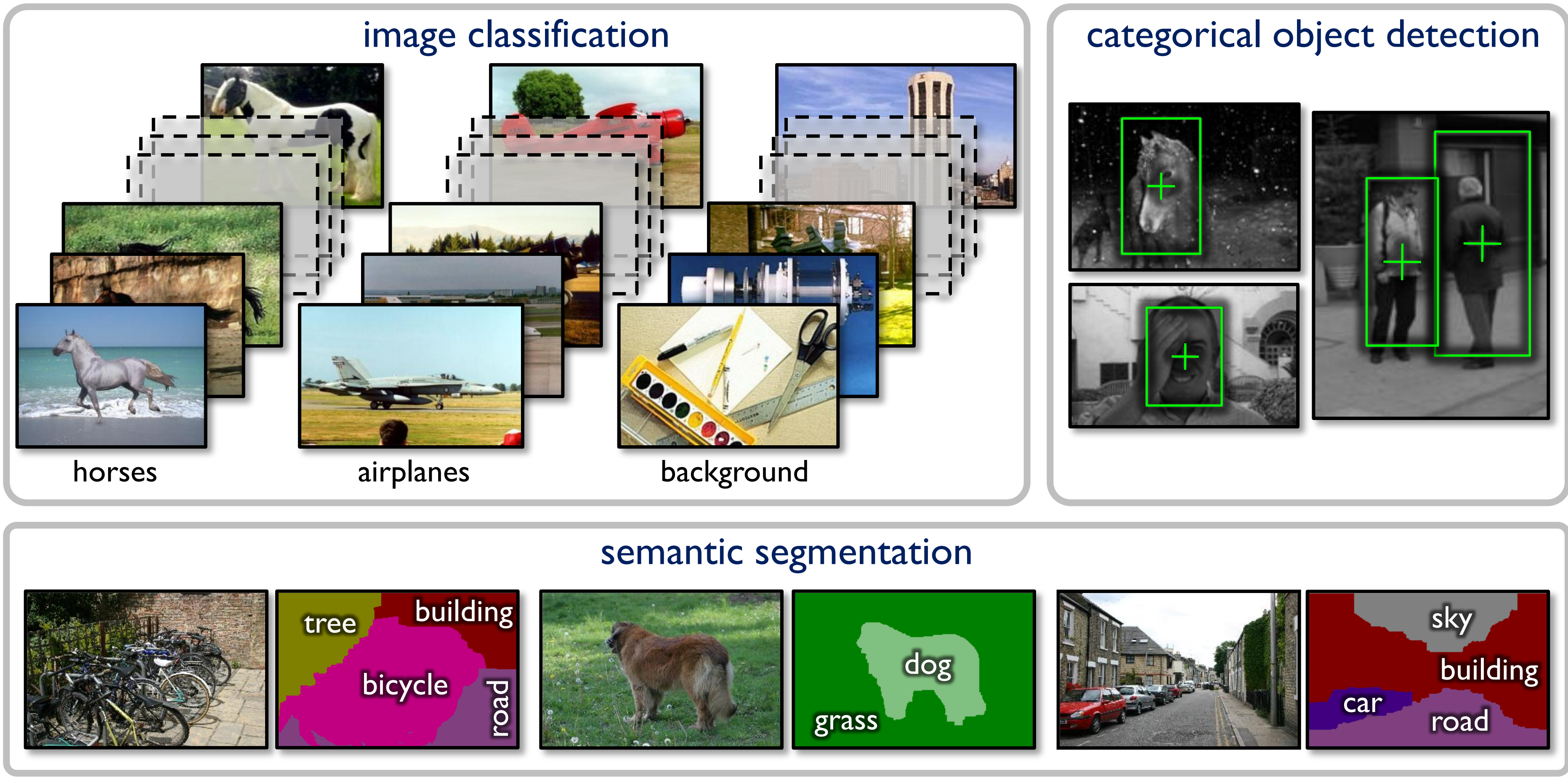
Reconstruction



3D models

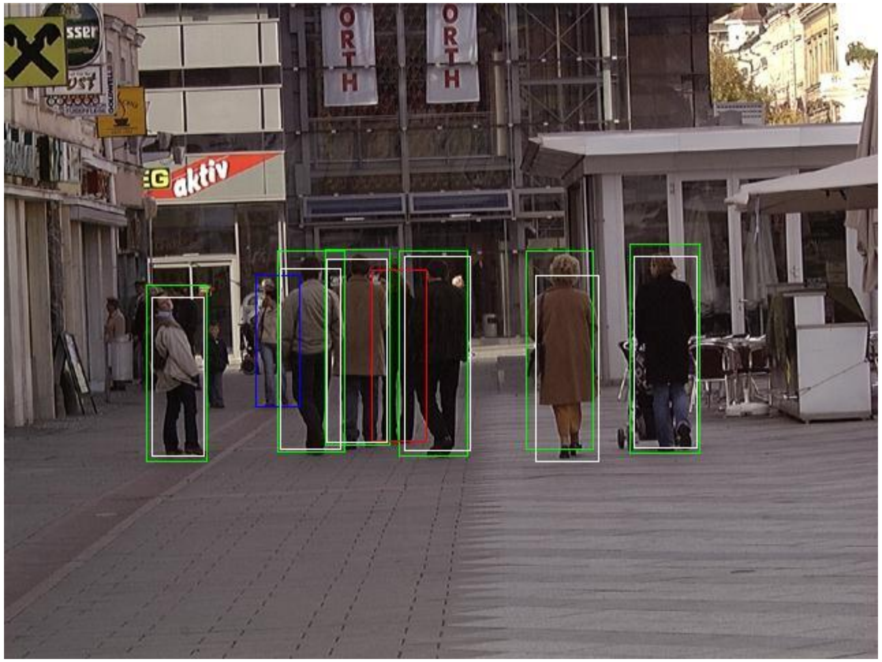
Recognition?

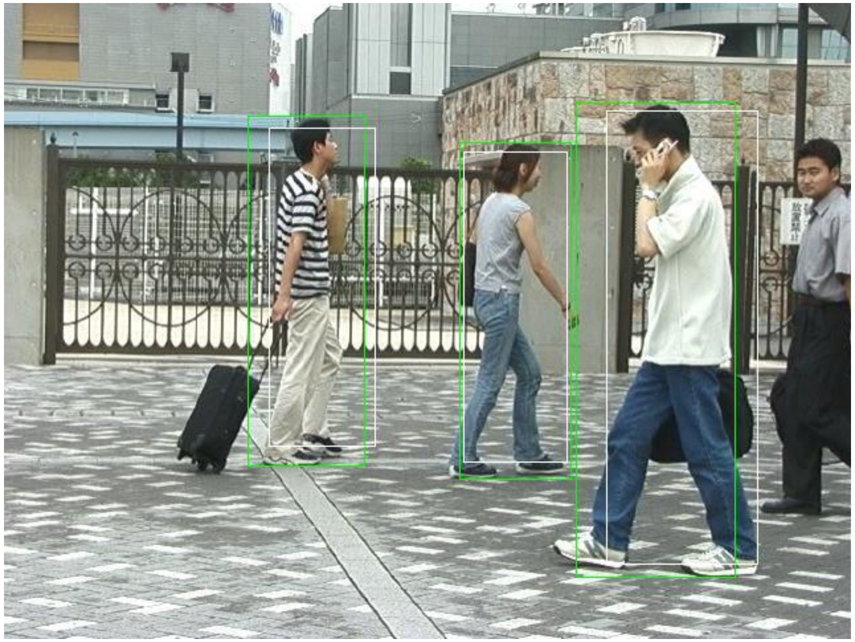
Recognition



Pedestrian detection

UNIVERSITY OF CAMBRIDGE





Existing applications

UNIVERSITY OF CAMBRIDGE

Traditional research in Computer vision developed for:

- Visual inspection
- Medical imaging
- Remote sensing
- Surveillance and biometrics
- Target detection and tracking

New applications

Computer vision has now found a place in consumer products

- Mobile phones and PDAs
- Games
- Cars
- Image and video search
- Internet and shopping

Smart erase on a mobile phone

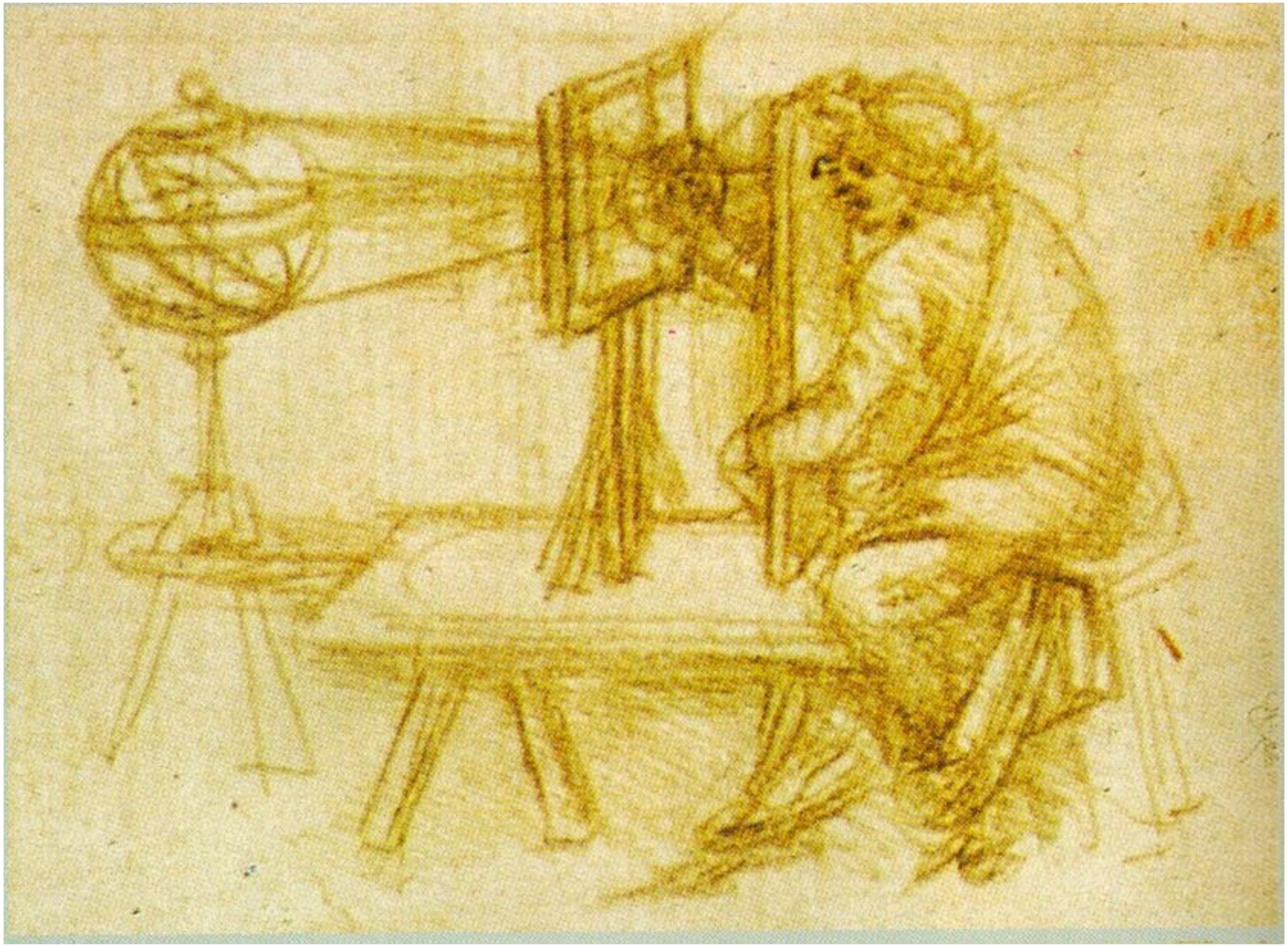
UNIVERSITY OF CAMBRIDGE

How to make machines that see?

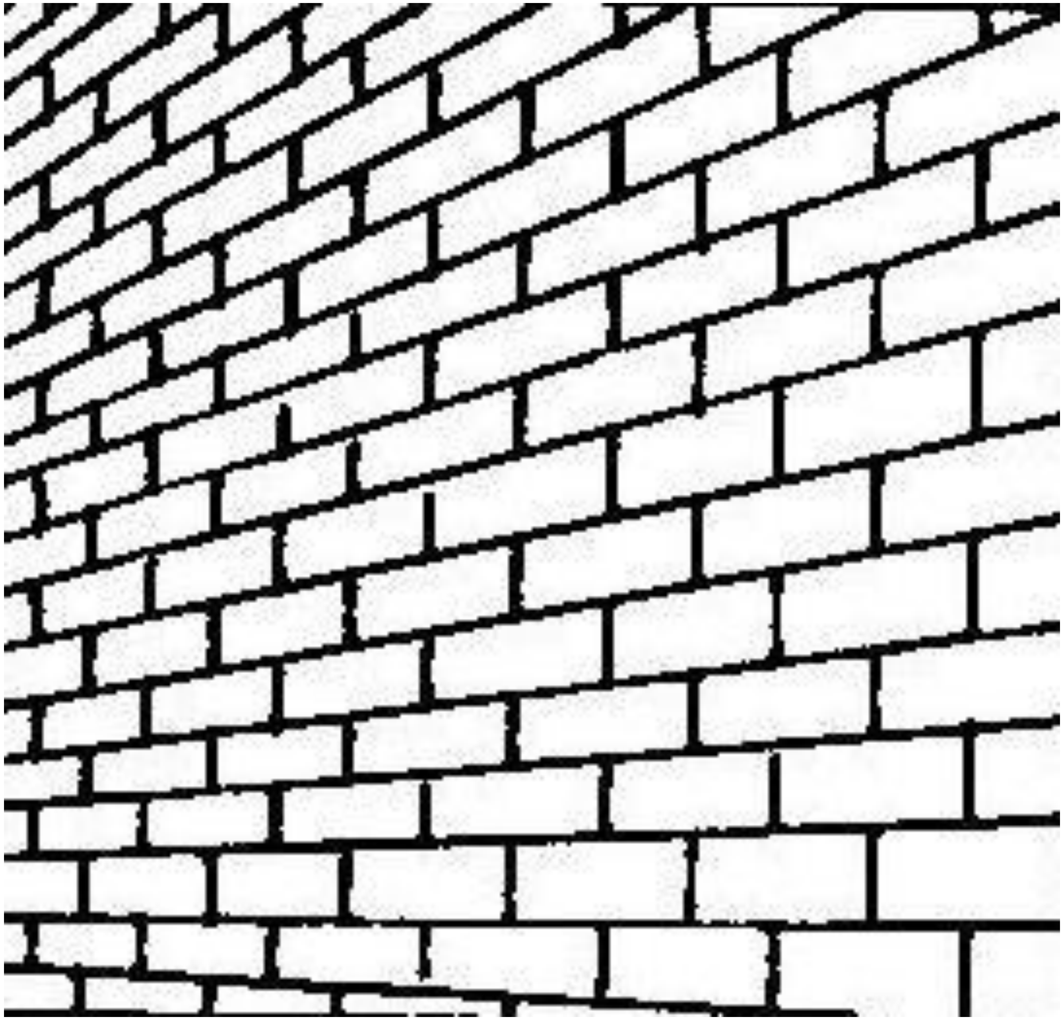
Why not study biology?

UNIVERSITY OF CAMBRIDGE

Perspective



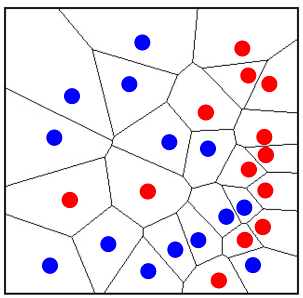
Transformations

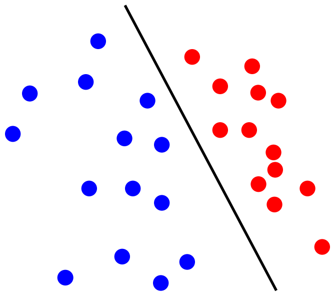


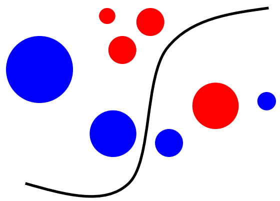
Shape

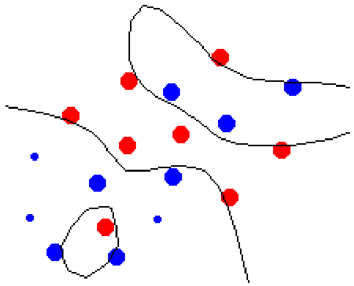
Machine Learning

Machine Learning









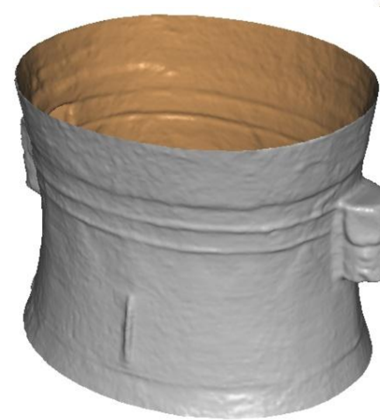
UNIVERSITY OF CAMBRIDGE

I. Reconstruction:

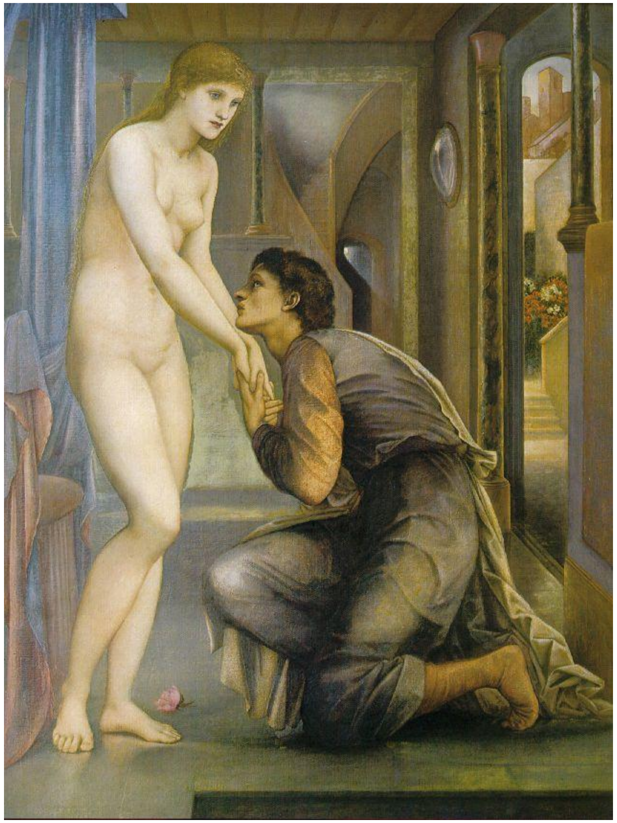
Recovery of accurate 3D shape from uncalibrated images

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

Digital Pygmalion Project

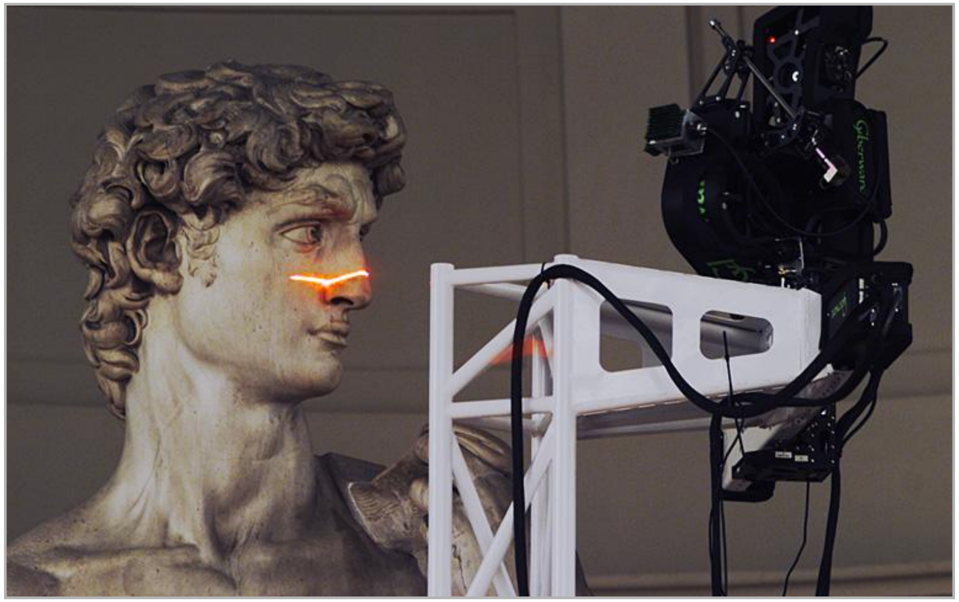


Digital Pygmalion – the myth



Scanning technologies

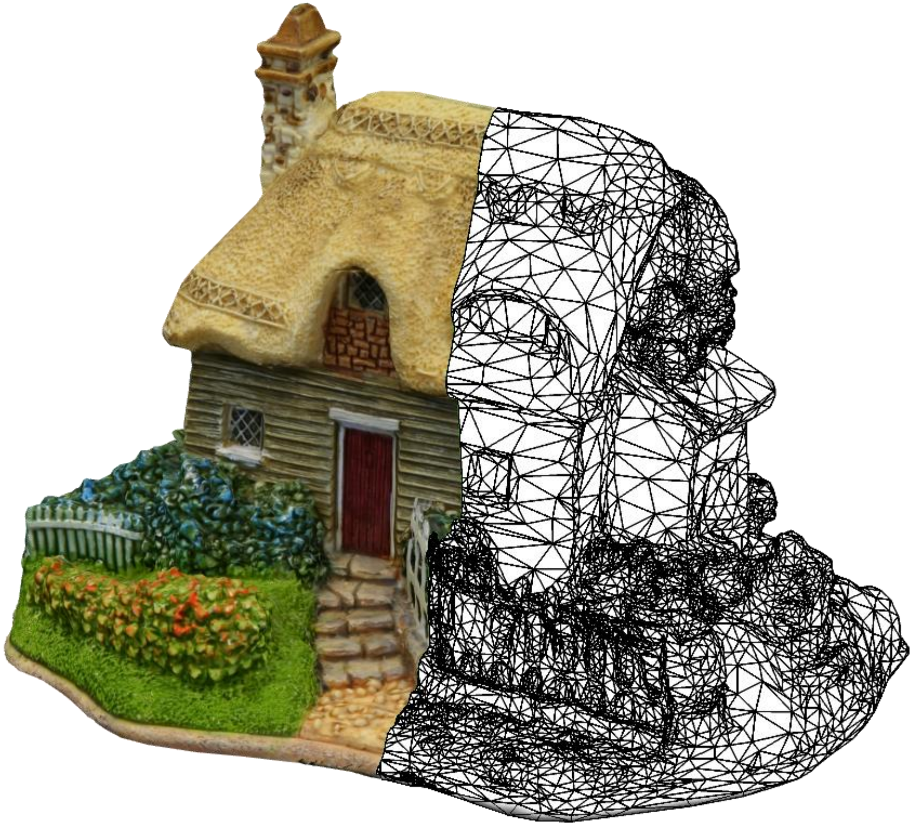
- Laser range finders
  - Very accurate
  - Very expensive
  - Complicated to use



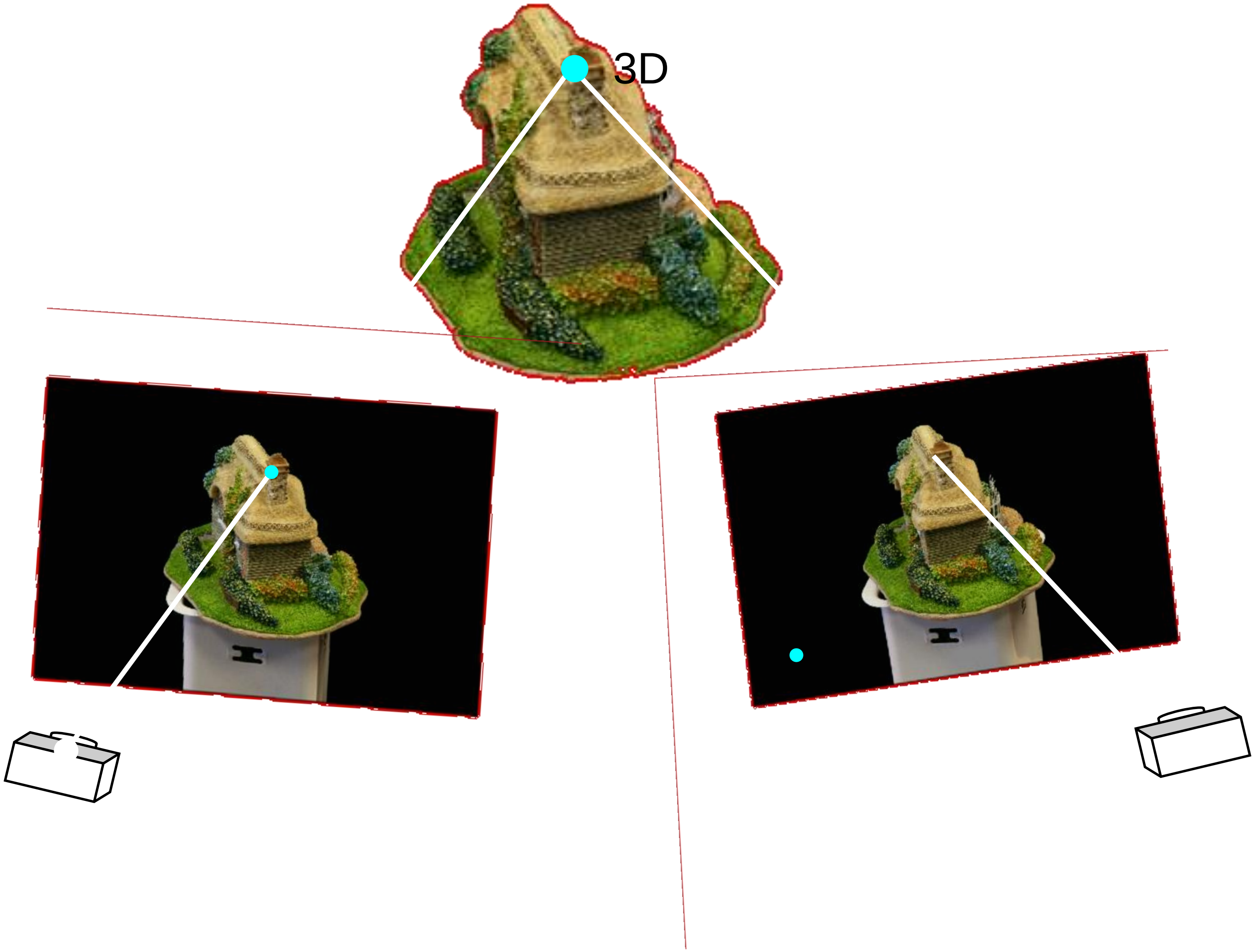
3D models

UNIVERSITY OF CAMBRIDGE

- We need a way to get them that is
  - practical
  - fast
  - non-intrusive
  - low cost



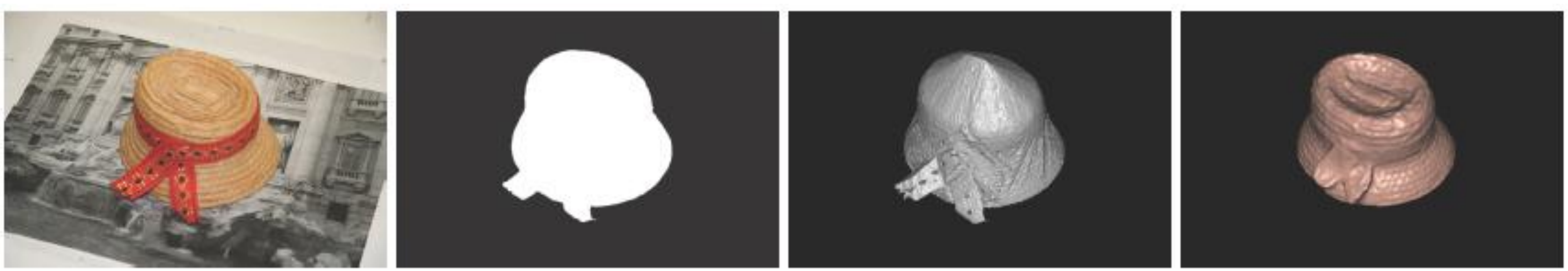
Stereo vision



Automatic 3d modeller

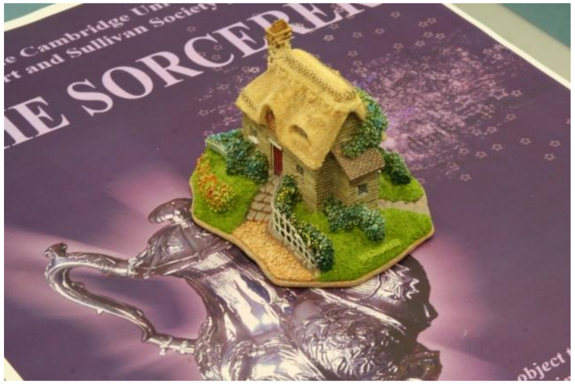
- Camera motion
- Segmentation of object outline
- Multi-view volumetric stereo
- 3D segmentation

input images → silhouettes → visual hull → 3D model



UNIVERSITY OF CAMBRIDGE

From images to model

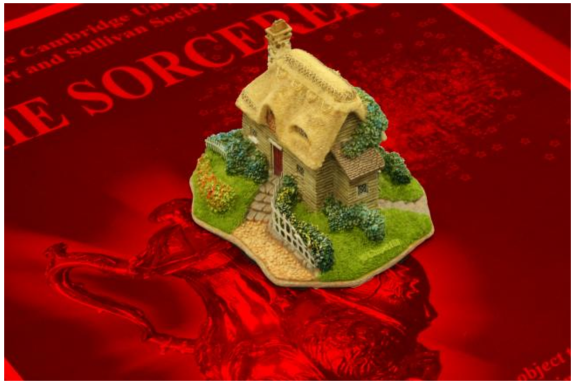


- capture images of object on top of a coloured sheet

From images to model

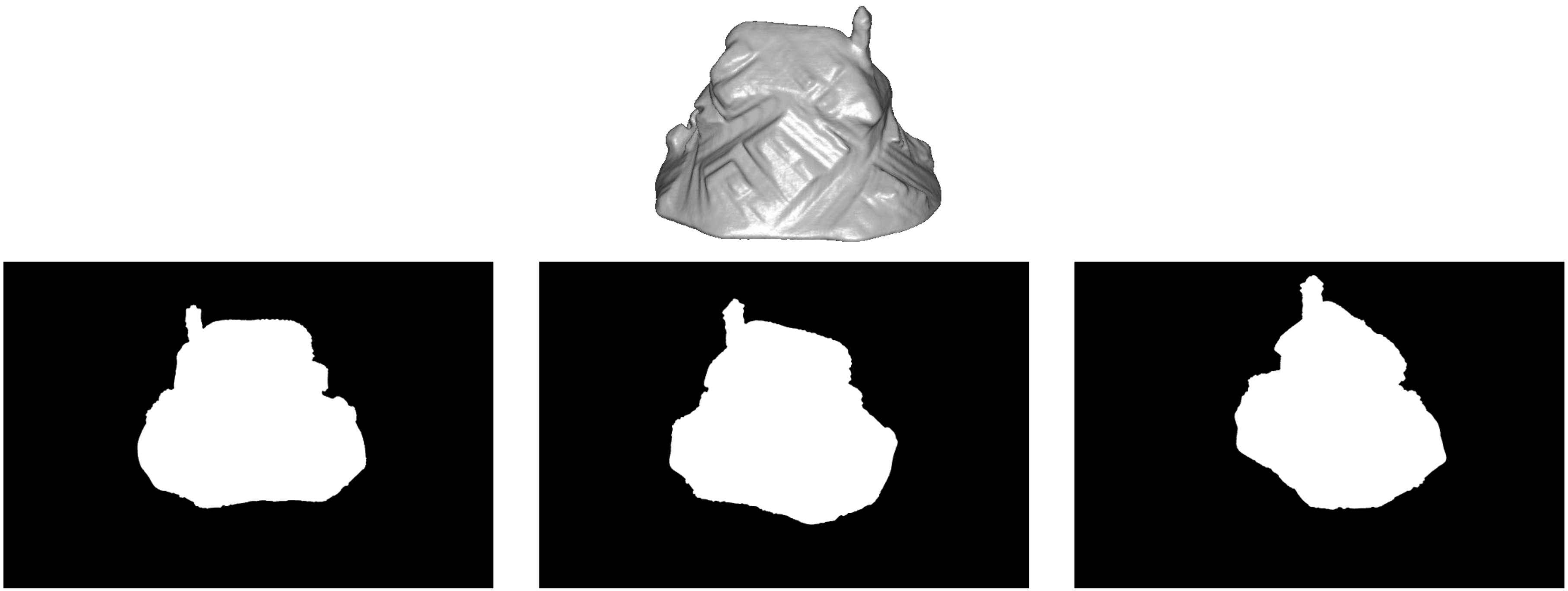
- calibrate cameras (i.e. estimate position, pose and focal length of camera in each photo) using pattern on sheet

From images to model



- identify object of interest by using *fixation constraint* and simultaneous 3D

From images to model



- construct visuall hull (largest object that can fit inside silhouettes)

Photo-consistency

Non-photo-consistent point

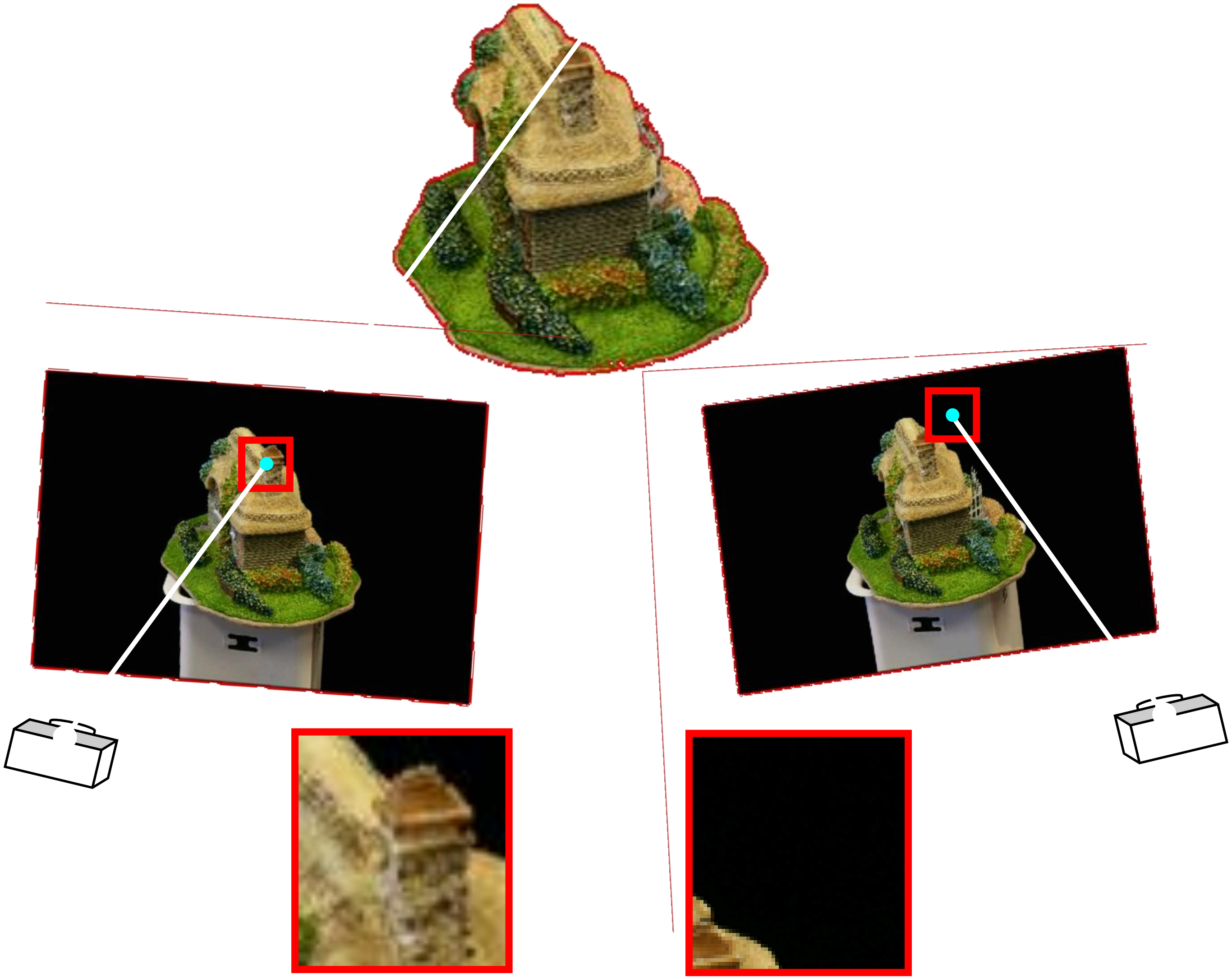
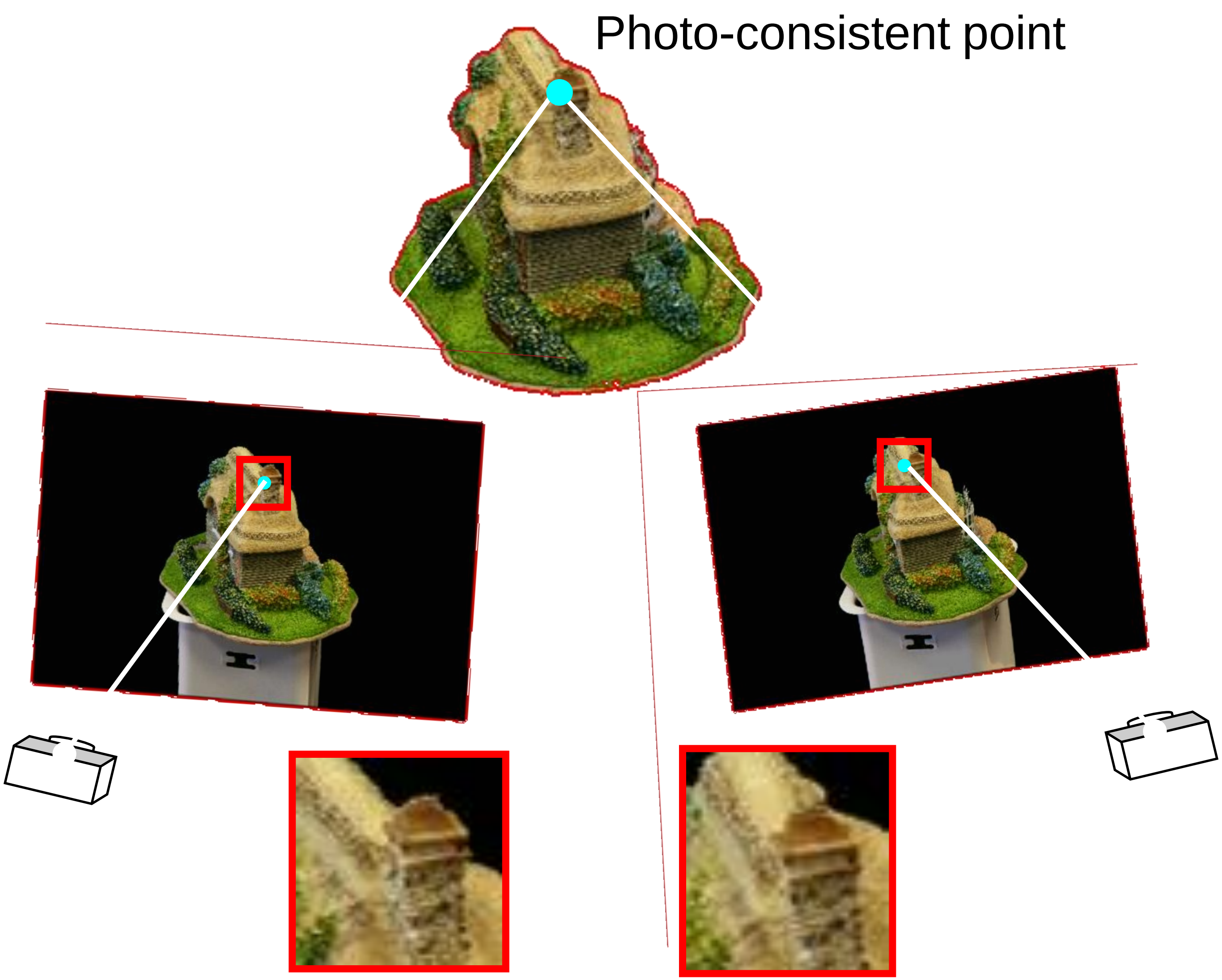
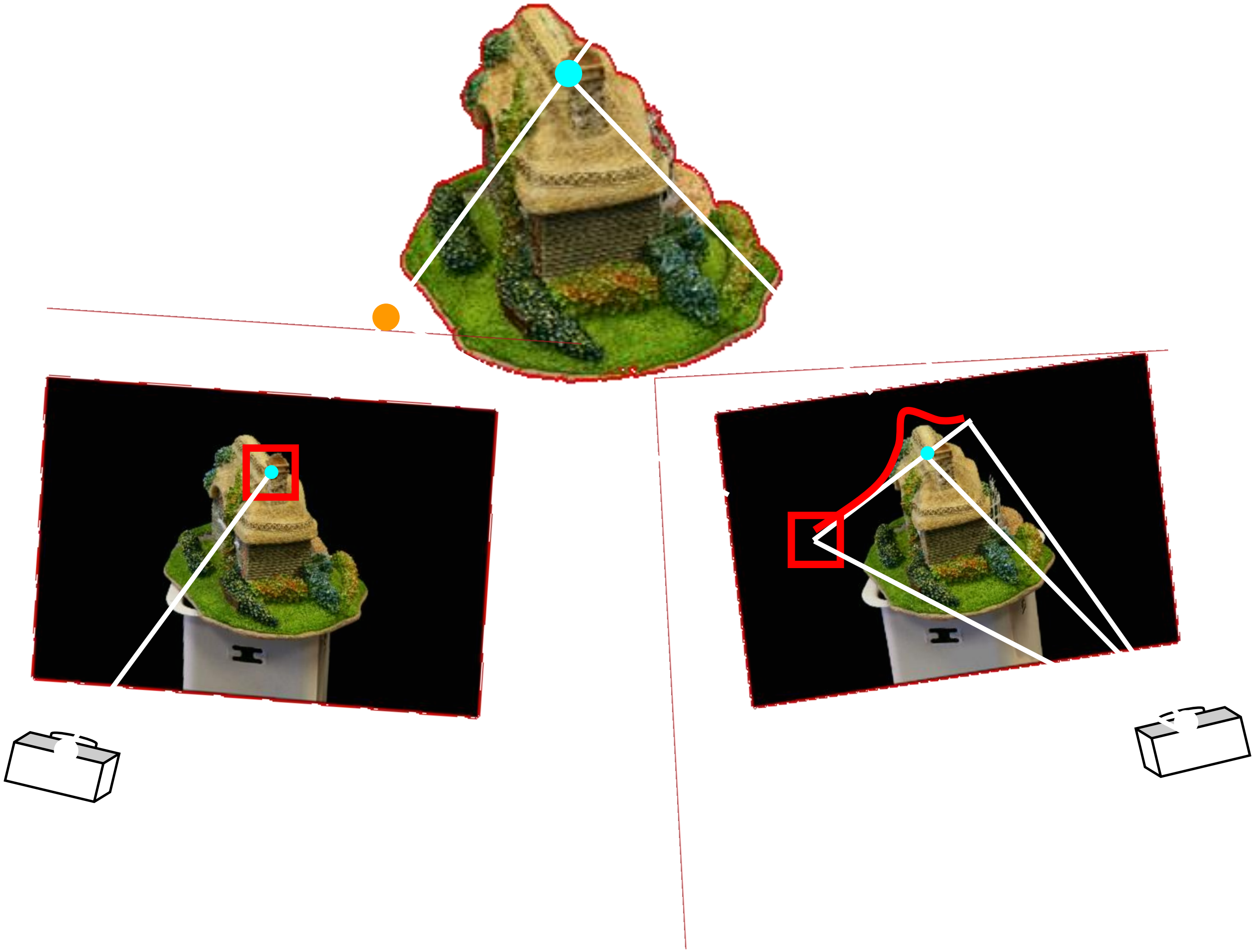


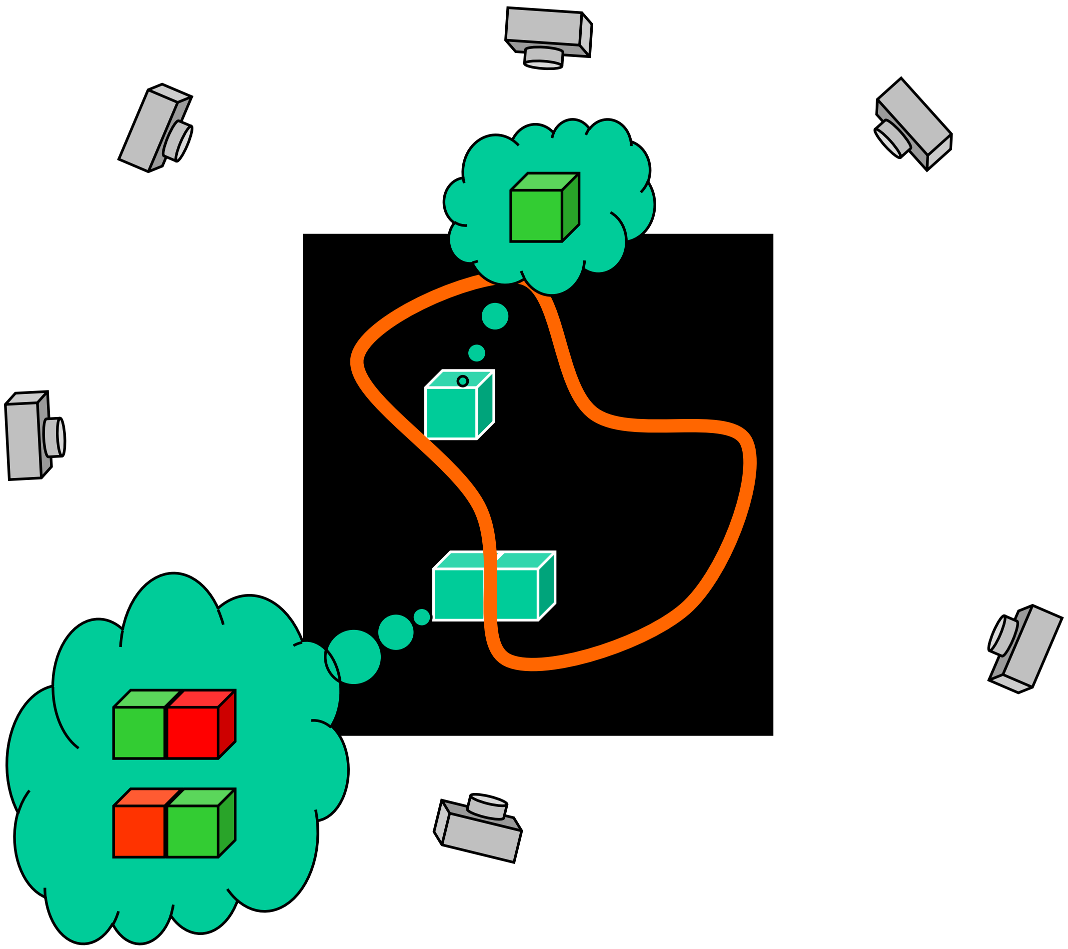
Photo-consistency



Finding the surface



3D segmentation

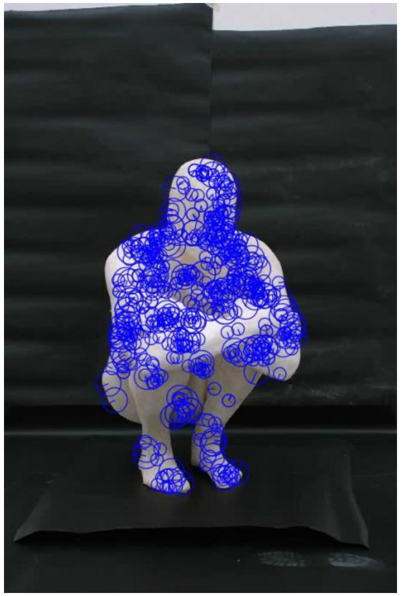


3D Models

Gormley - input Images

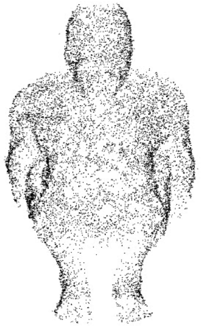
Recovery of camera motion

Input images



Feature extraction

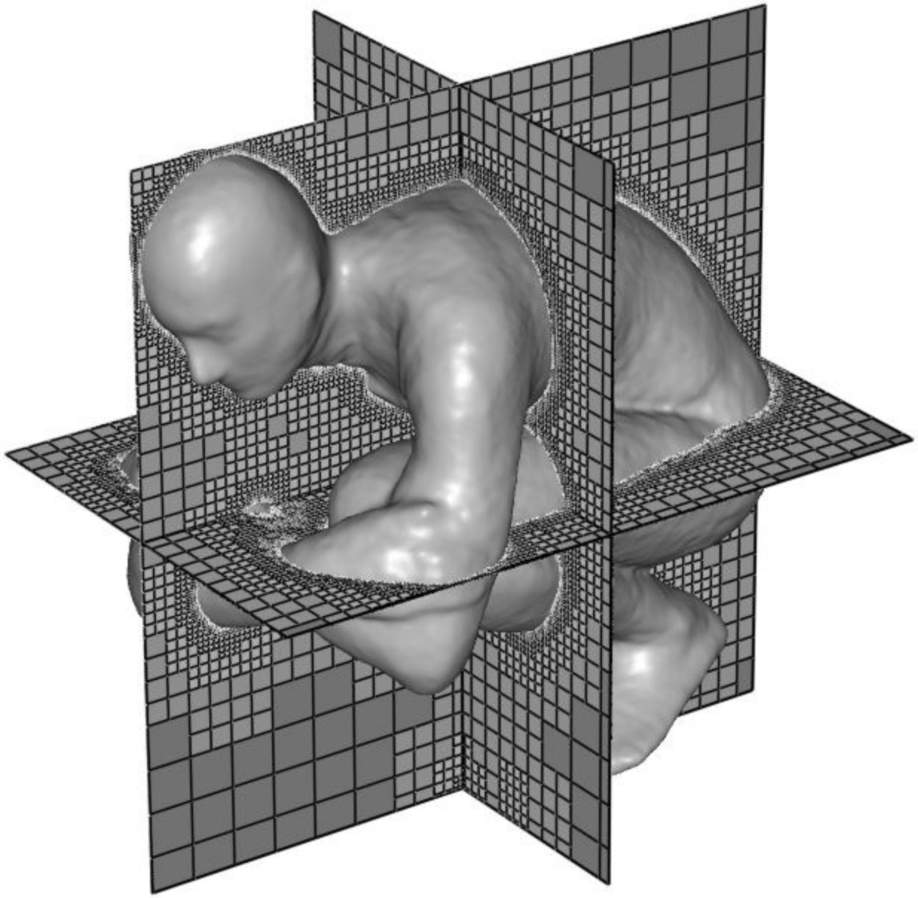
Feature matching



Bundle adjustment

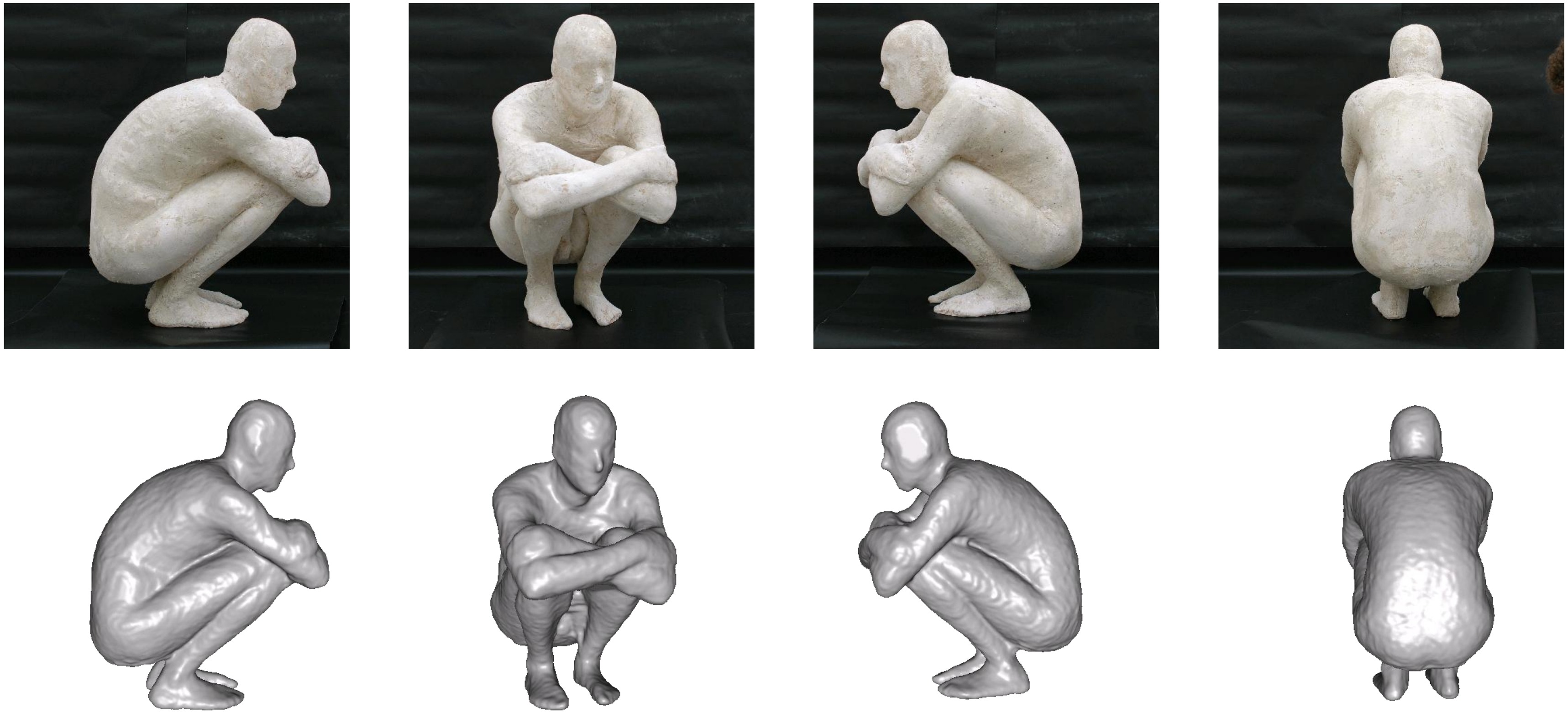
Probabilistic 3D segmentation

UNIVERSITY OF CAMBRIDGE



3D Models

UNIVERSITY OF CAMBRIDGE



Final installation

3D models

UNIVERSITY OF CAMBRIDGE

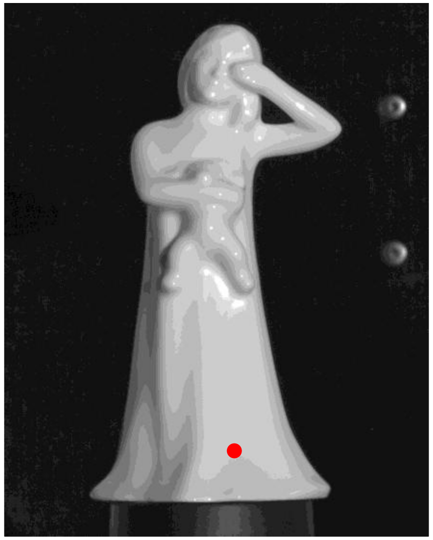
Multiview photometric stereo

Vogiatzis, Hernandez and Cipolla 2006 and 2008

Untextured objects

UNIVERSITY OF CAMBRIDGE

- Almost impossible to establish correspondence



Use shading cue

Untextured objects

Changing lighting uncovers fine geometric detail

- Assumptions:
  - Single, distant light-source
  - Silhouettes can be extracted
  - No texture, single colour

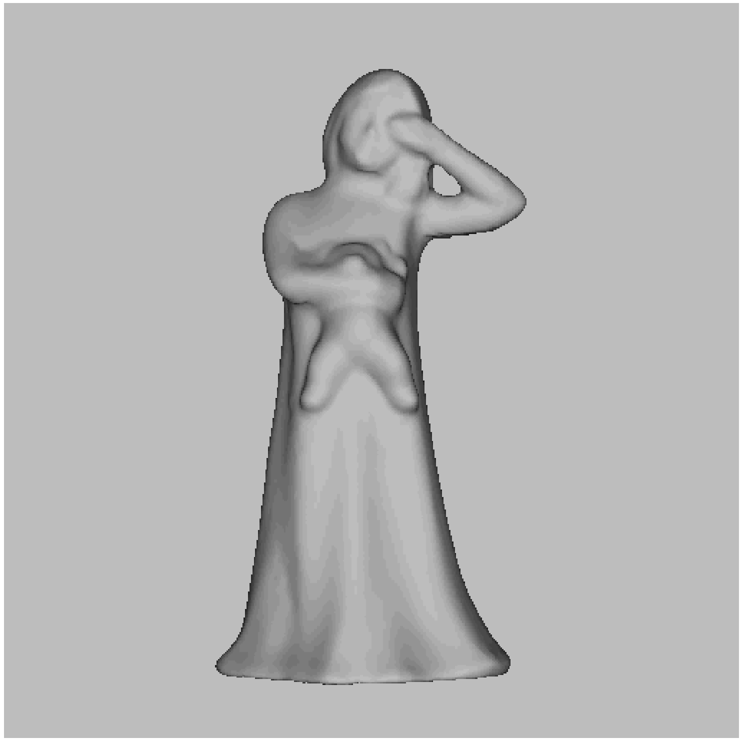
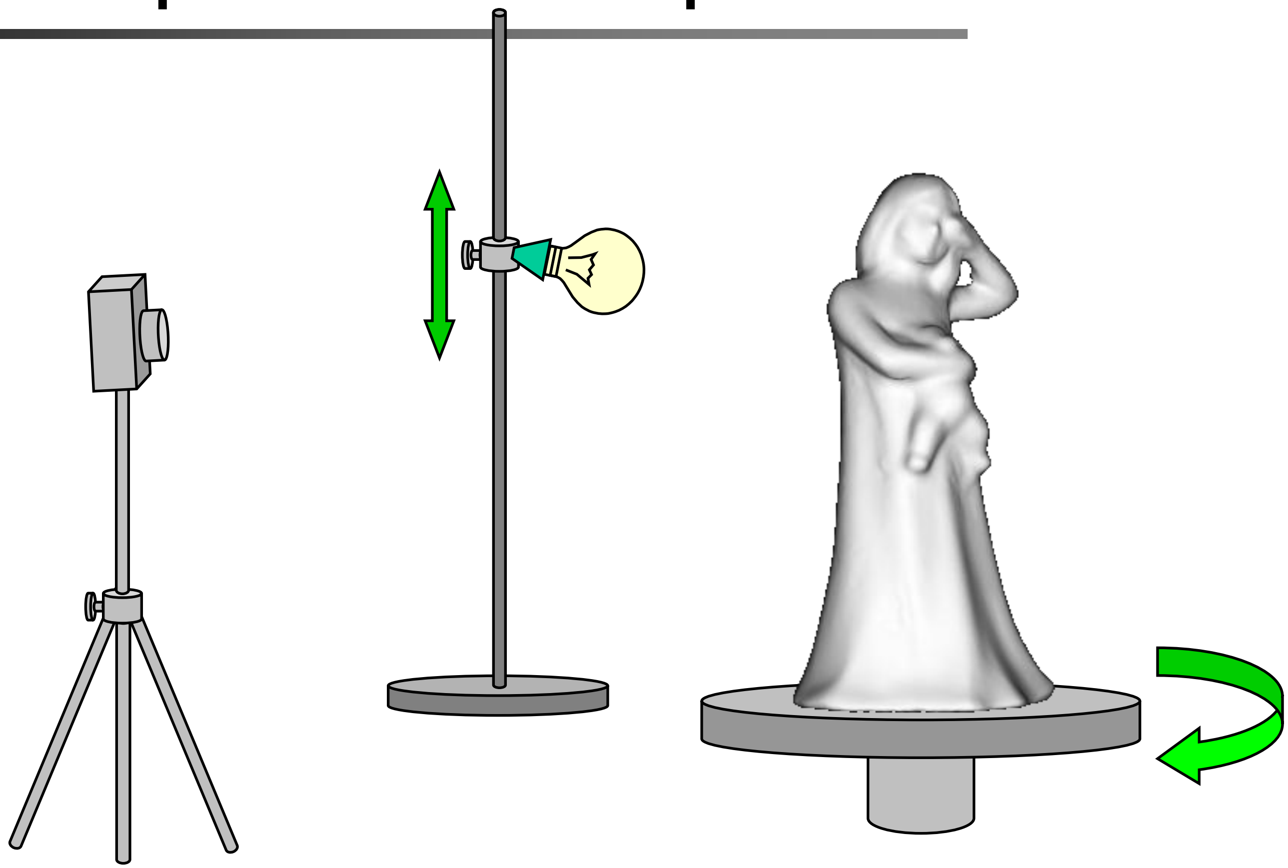
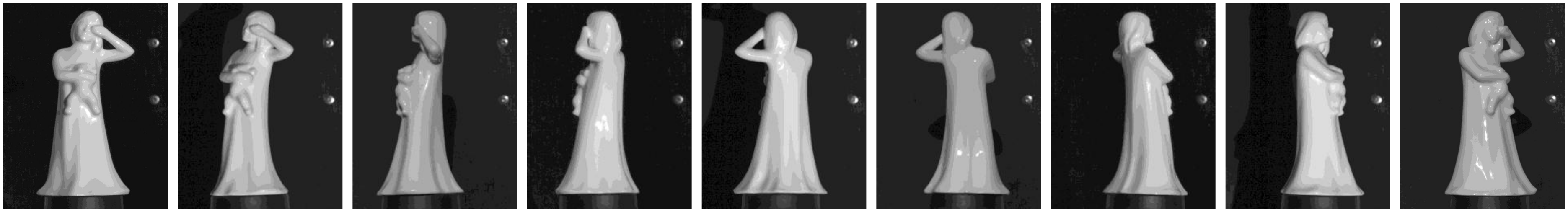


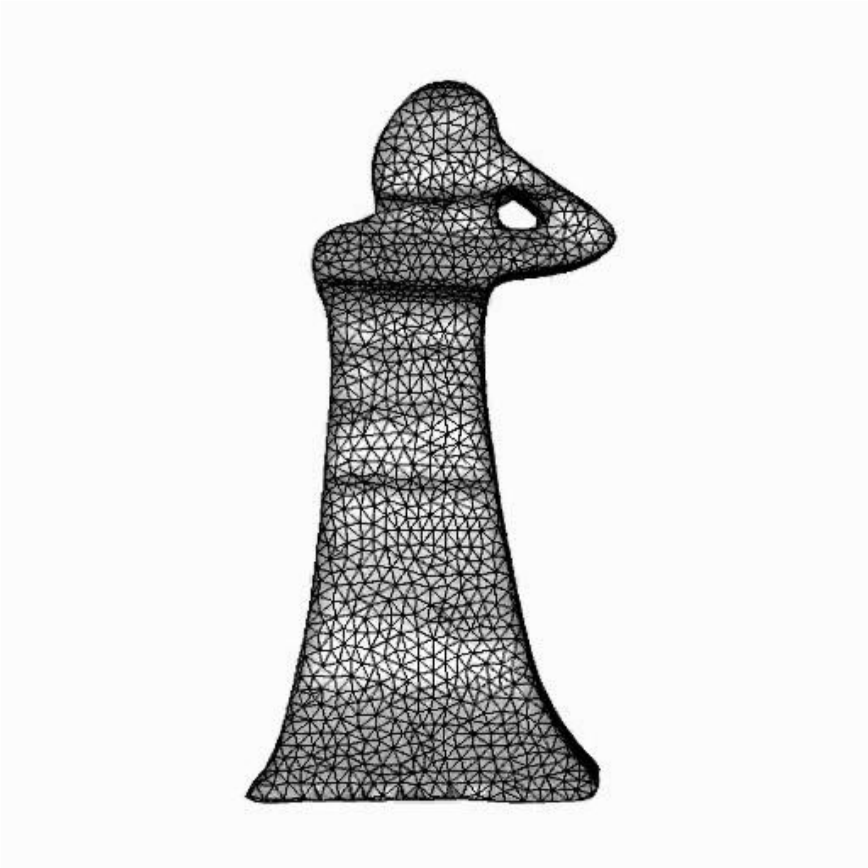
Image acquisition setup

UNIVERSITY OF CAMBRIDGE



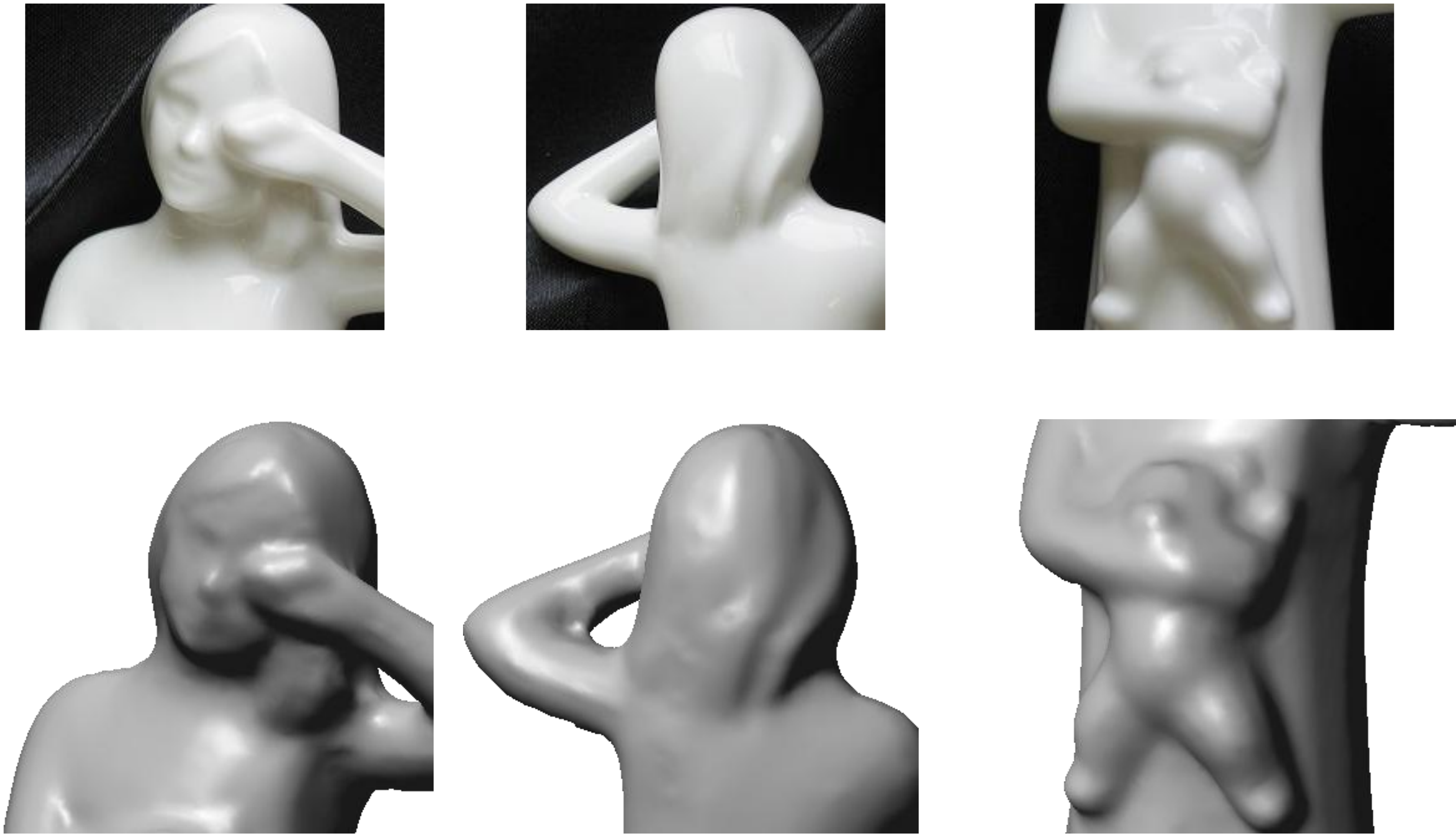


Surface Evolution of 3D Mesh



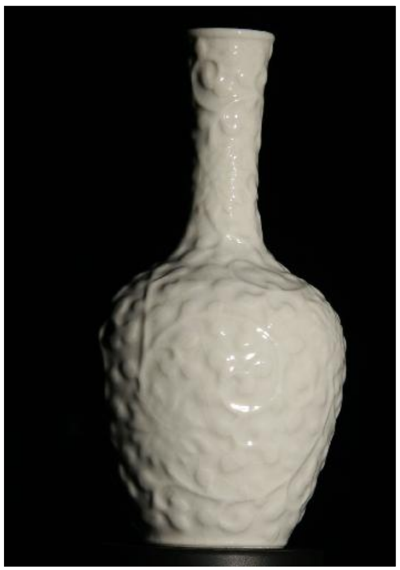
3D Models

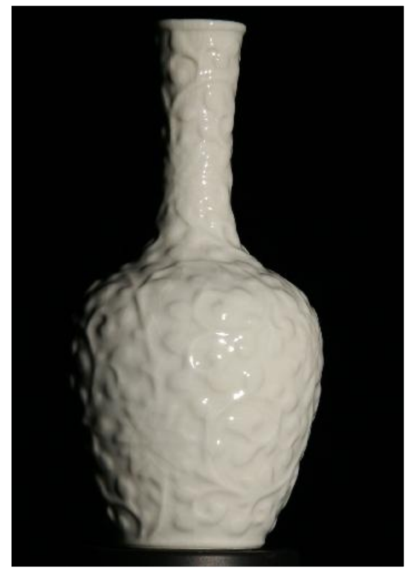
UNIVERSITY OF CAMBRIDGE



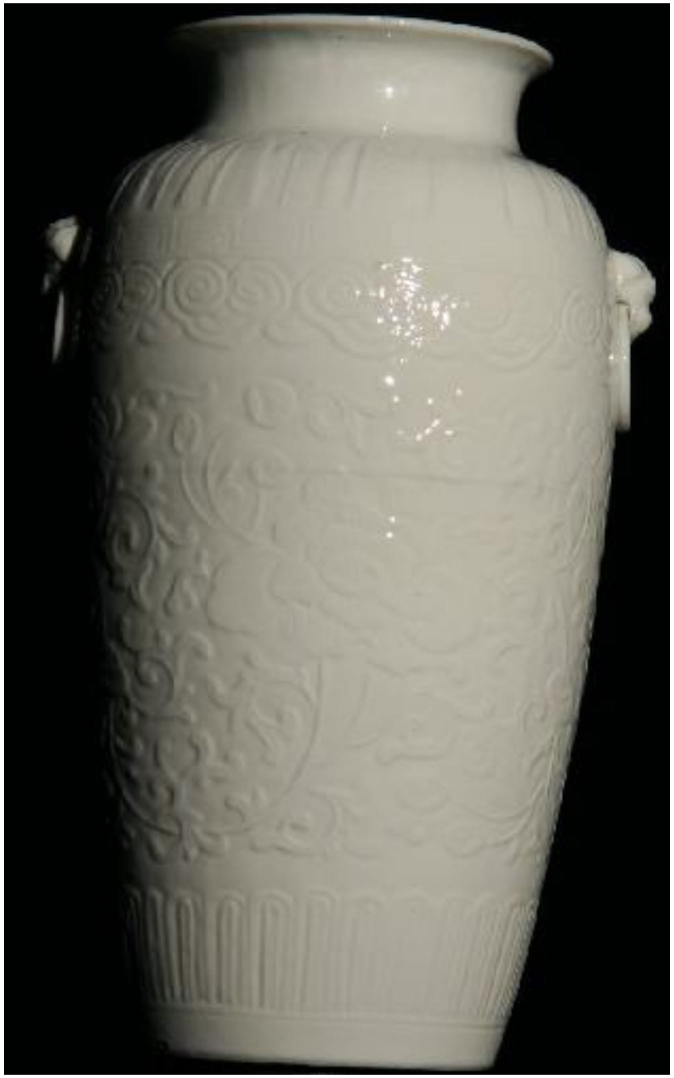
Making physical copies

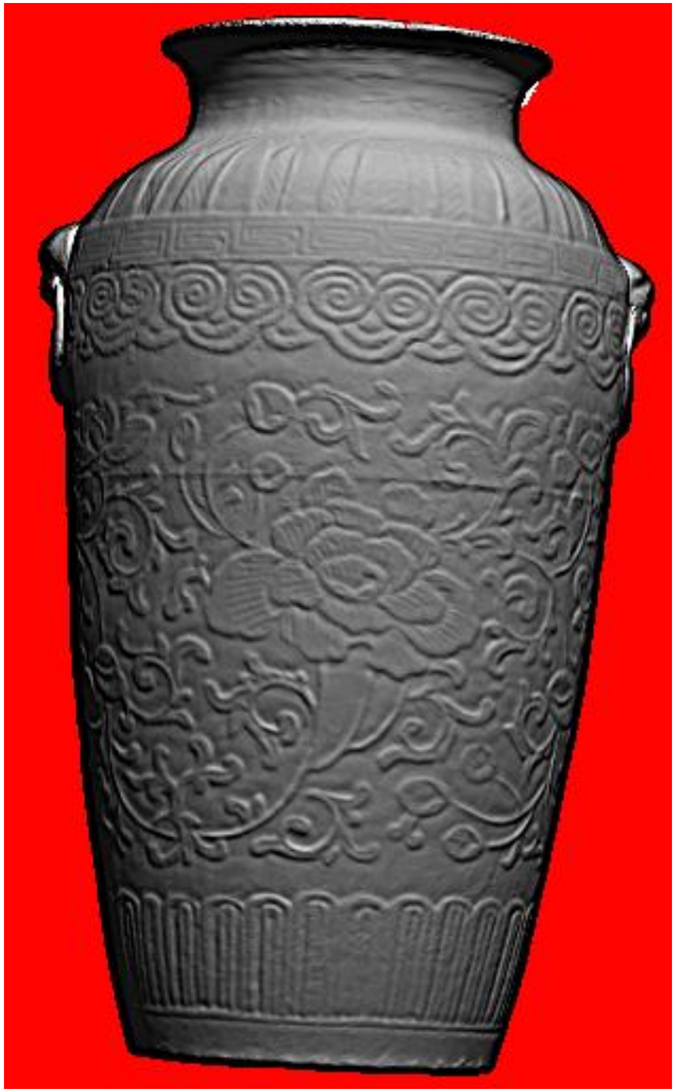
3D Models





3D Models

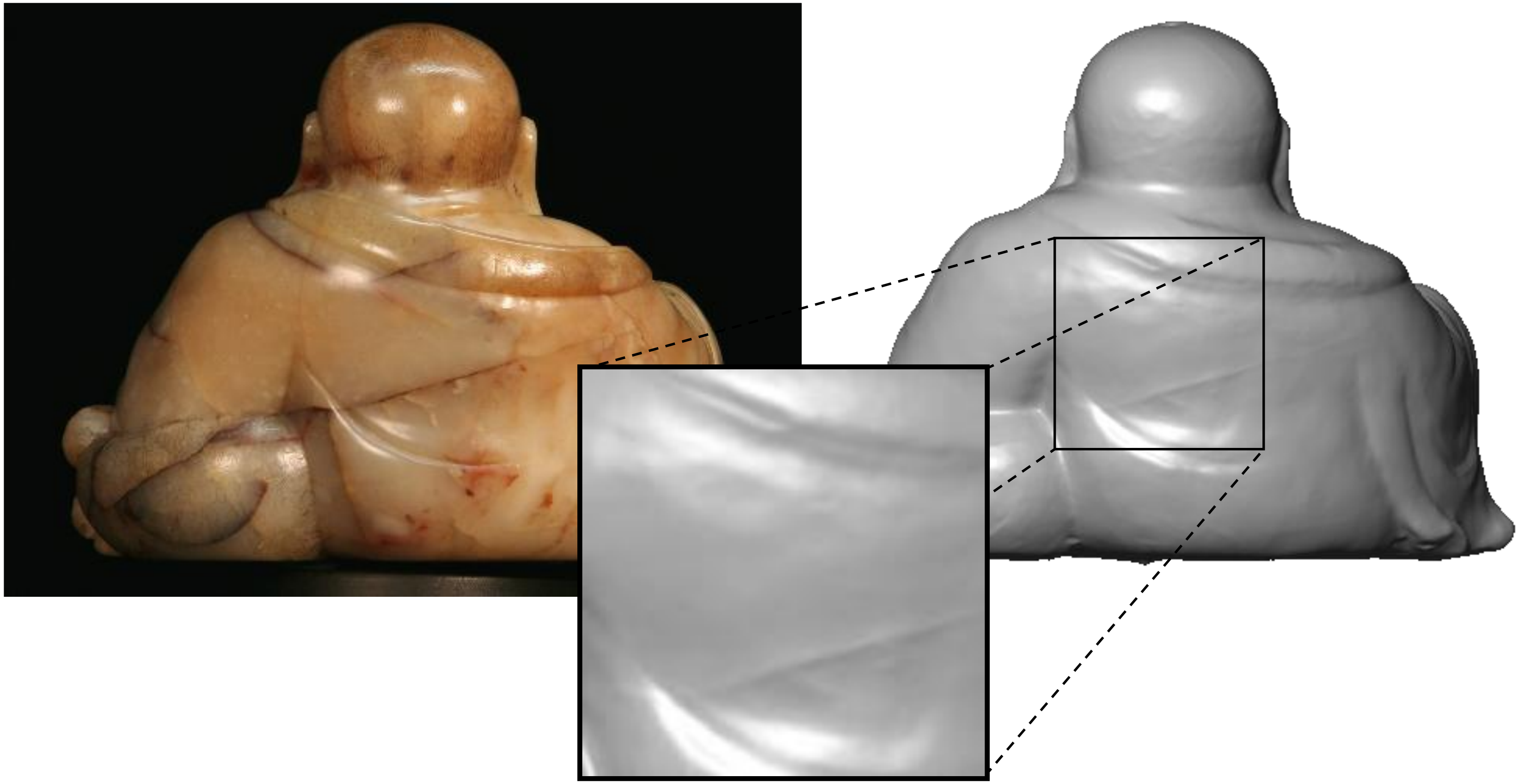




3D Models

3D Models

3D Models



UNIVERSITY OF CAMBRIDGE

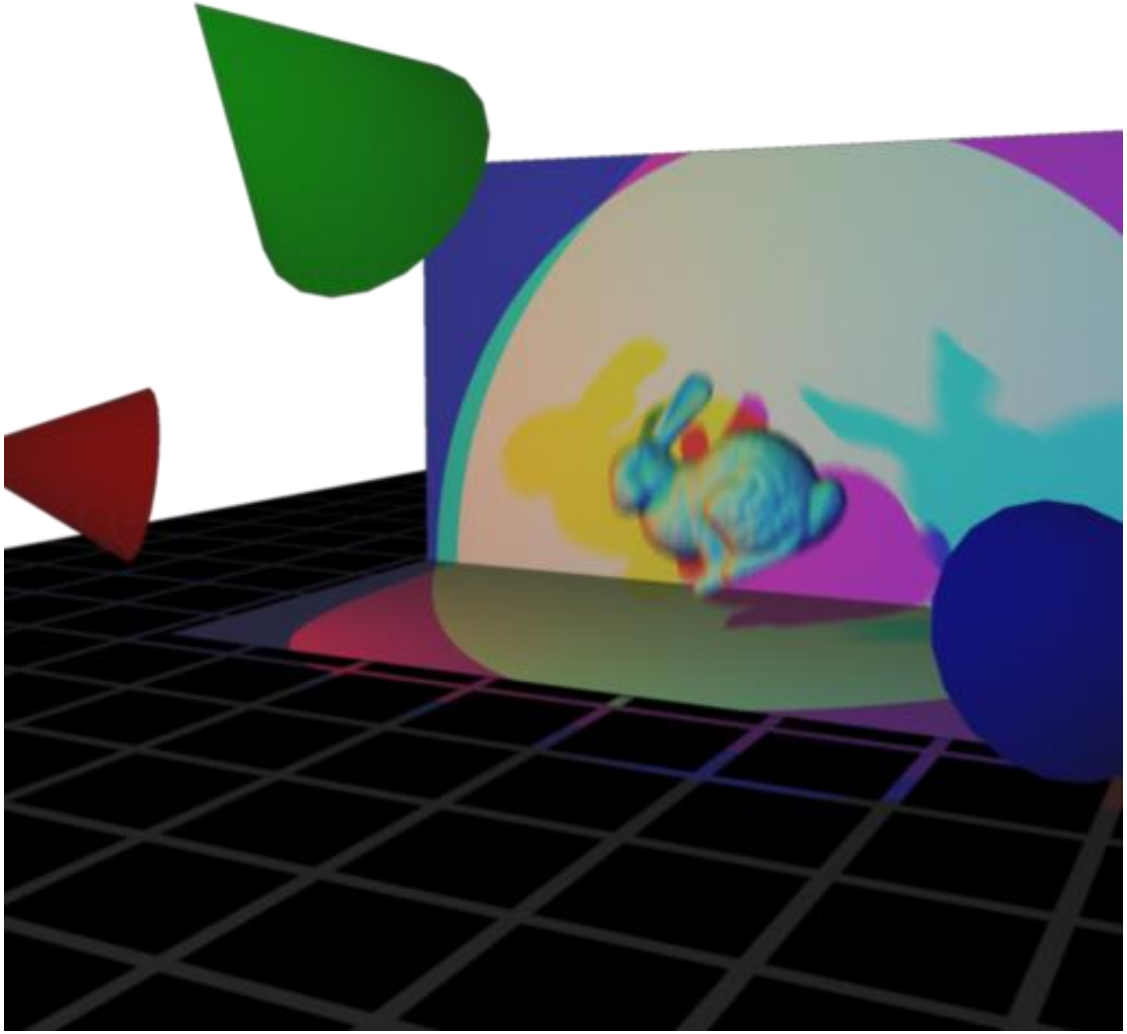
Deformable objects:

Real-time photometric stereo using colour lighting

Hernandez et al 2007

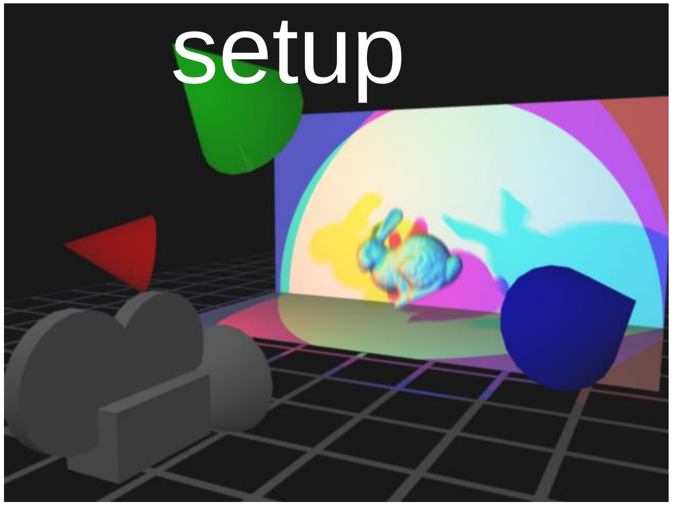
Photometric stereo with colour

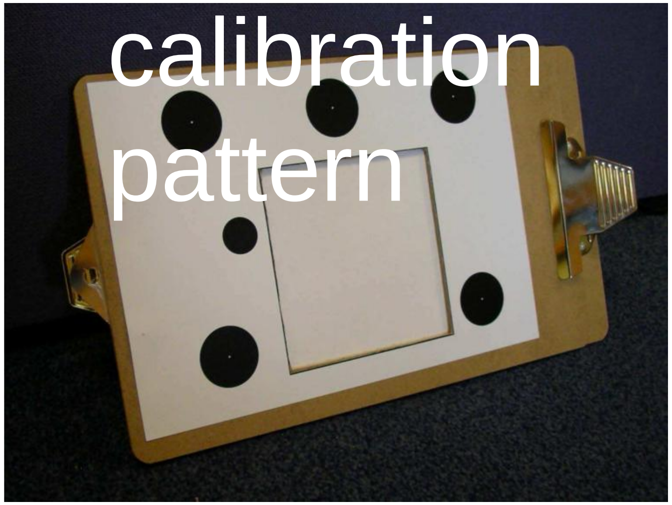
UNIVERSITY OF CAMBRIDGE



Deformable objects?

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





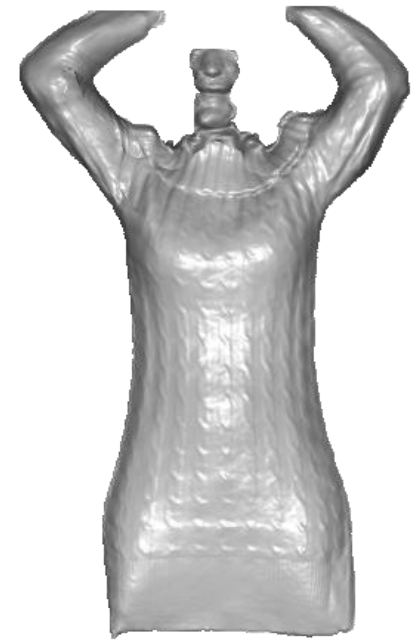
Shape from colour

UNIVERSITY OF CAMBRIDGE

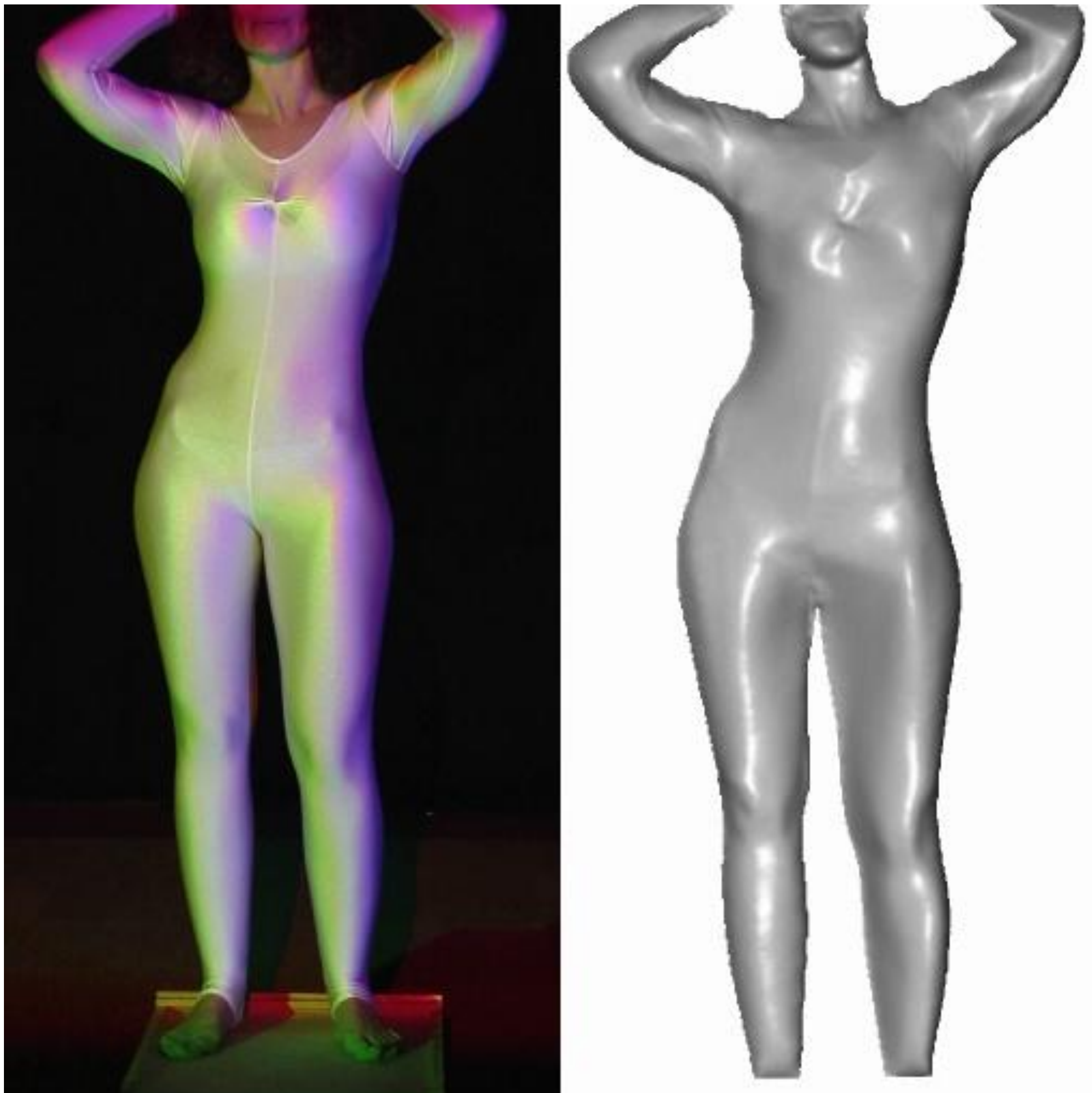
- observation: **1-1 mapping between colour and surface orientation**

- get map of surface orientations from colour image
- integrate orientations to get depth map

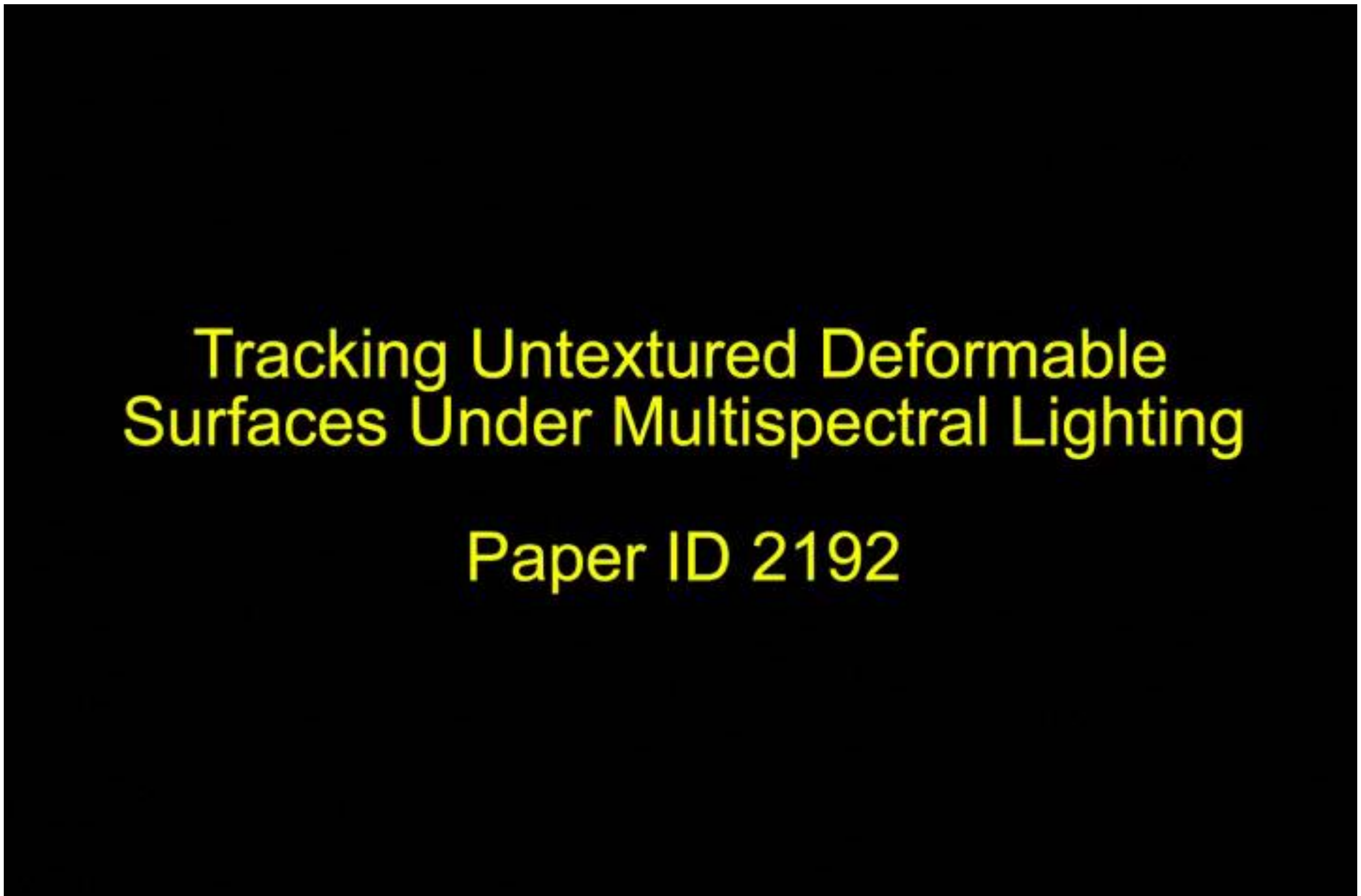
- do this for colour video to get 2.5d reconstruction of deforming object!



Photometric stereo with colour



shape from colour



II. Registration:

Target detection and pose estimation

Image matching

Registration:

Where am I?
What am I looking at?

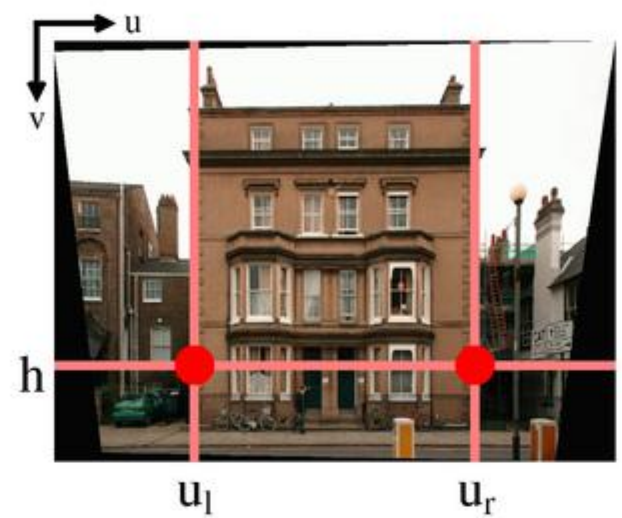
Johansson and Cipolla 2002
Robertson and Cipolla 2004
Cipolla et al 2004

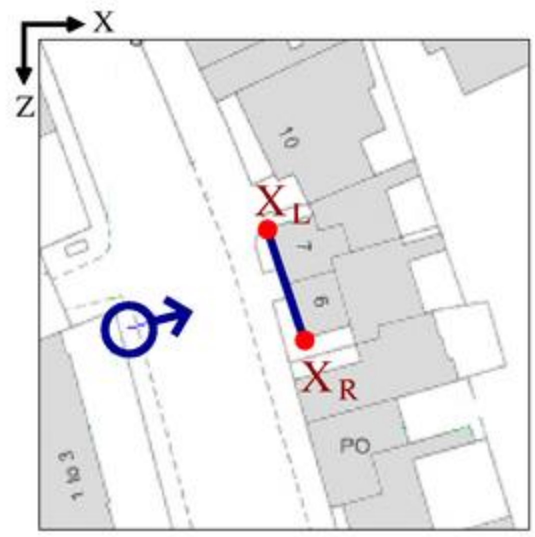
Where I am?

Determine pose from single image by matching

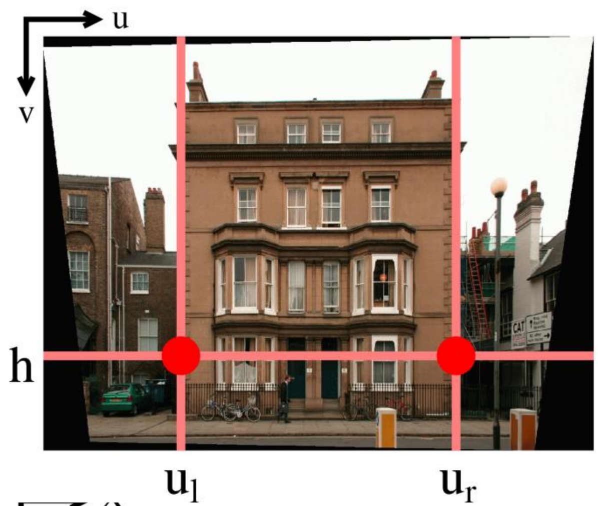
Register database view

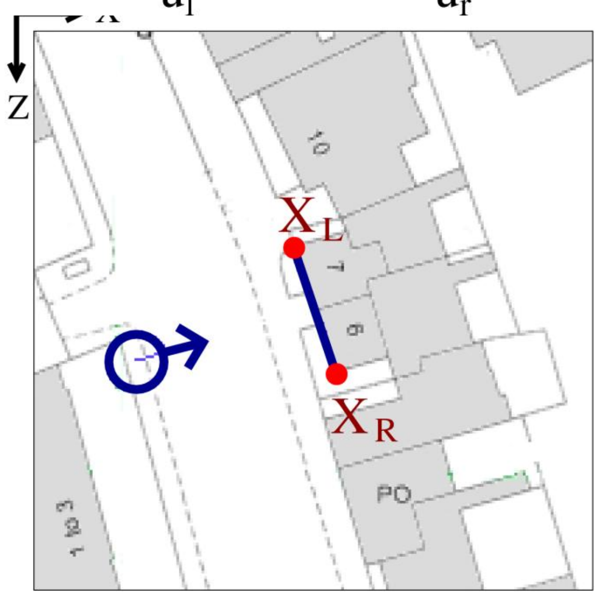
First align database view to map





Localisation of query view





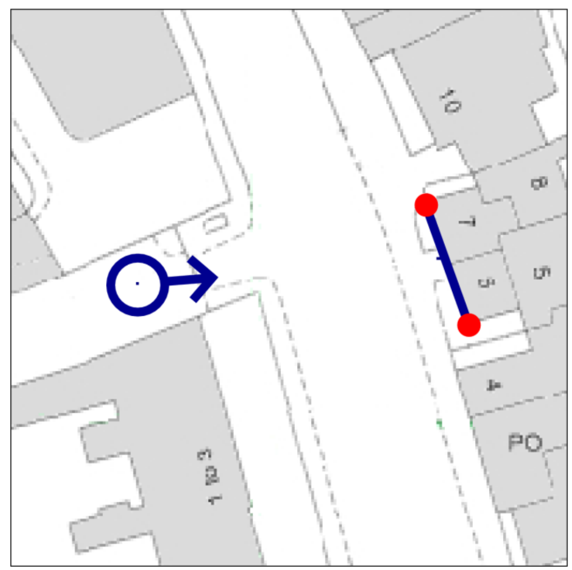


Image-based localisation

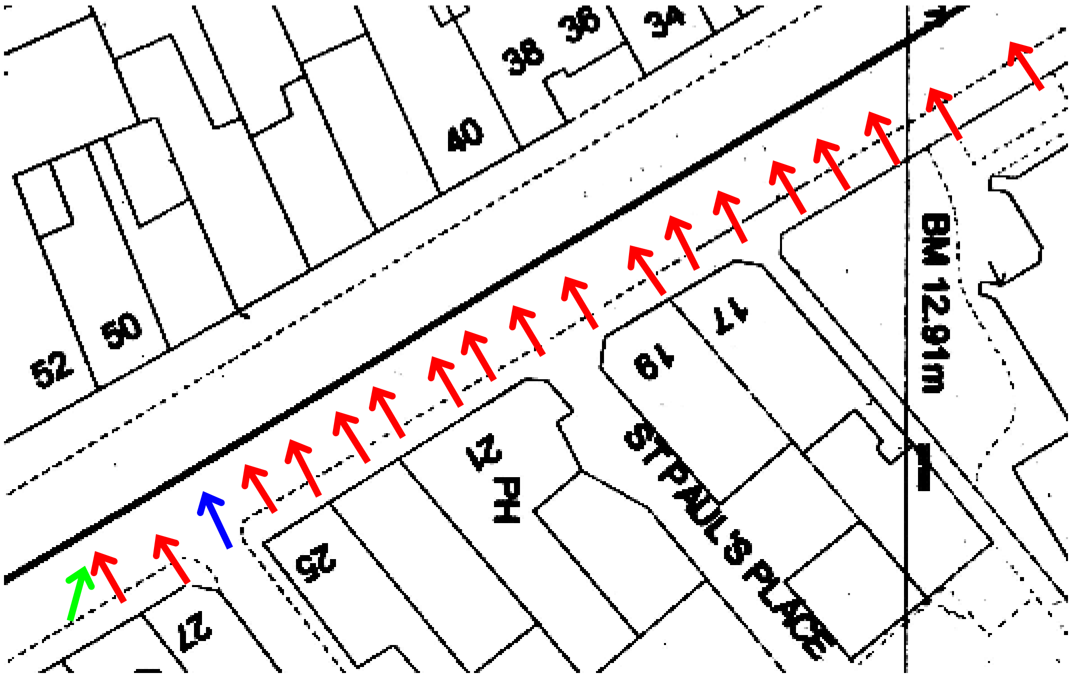


Image-based localisation

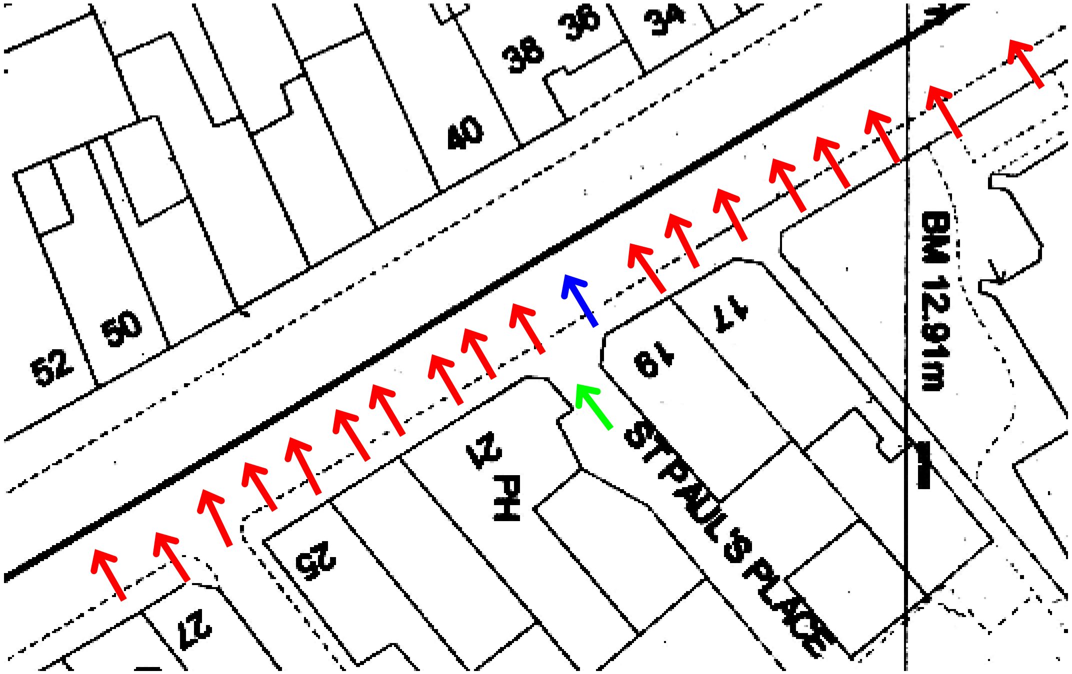
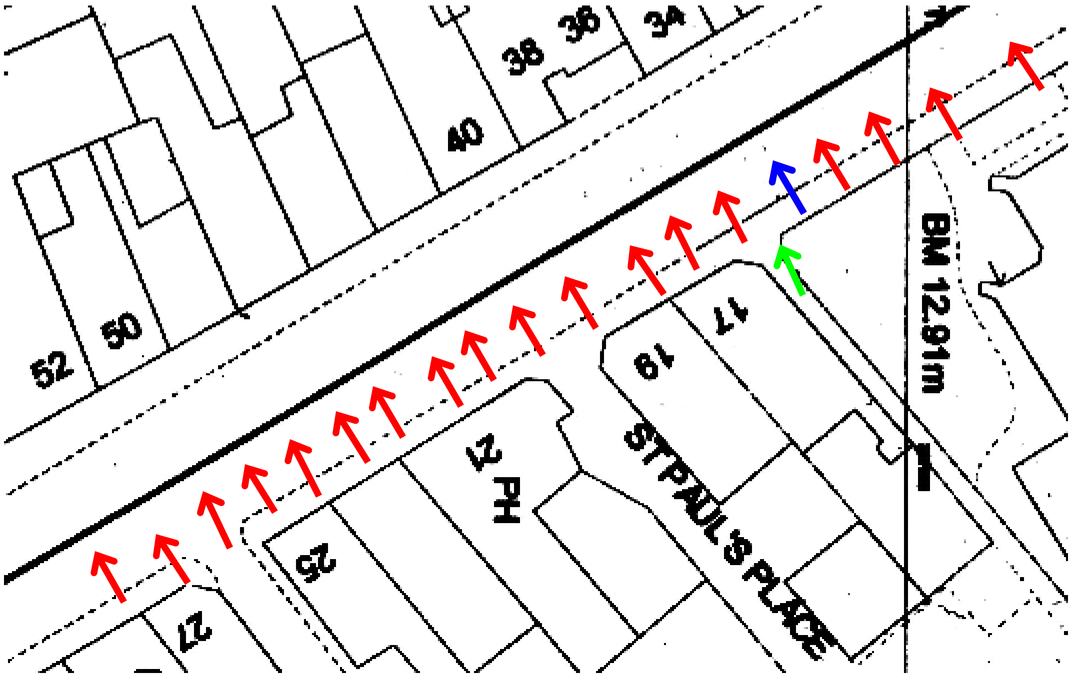


Image-based localisation

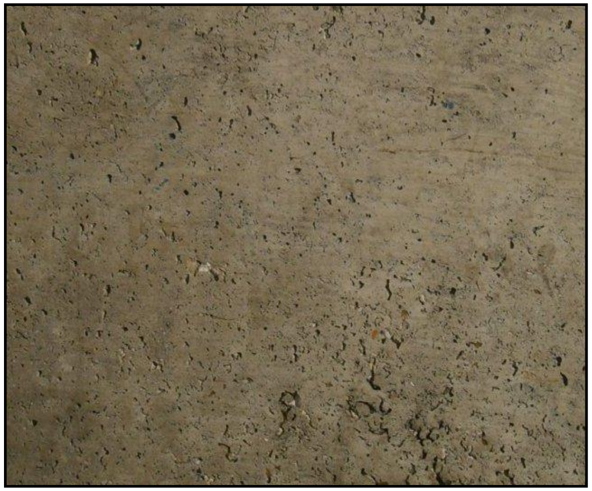


Registration:

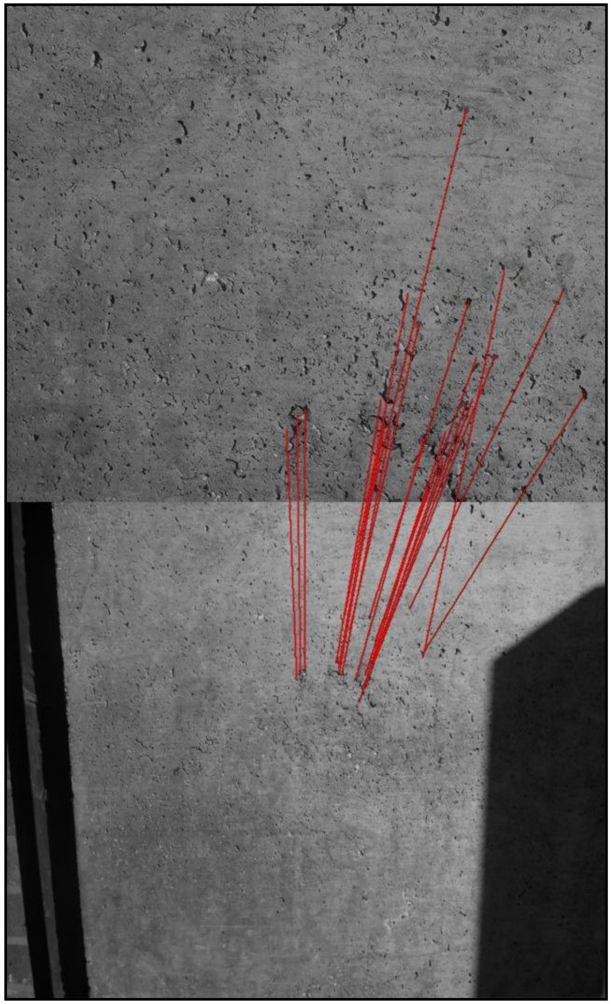
Ageing infrastructure inspection

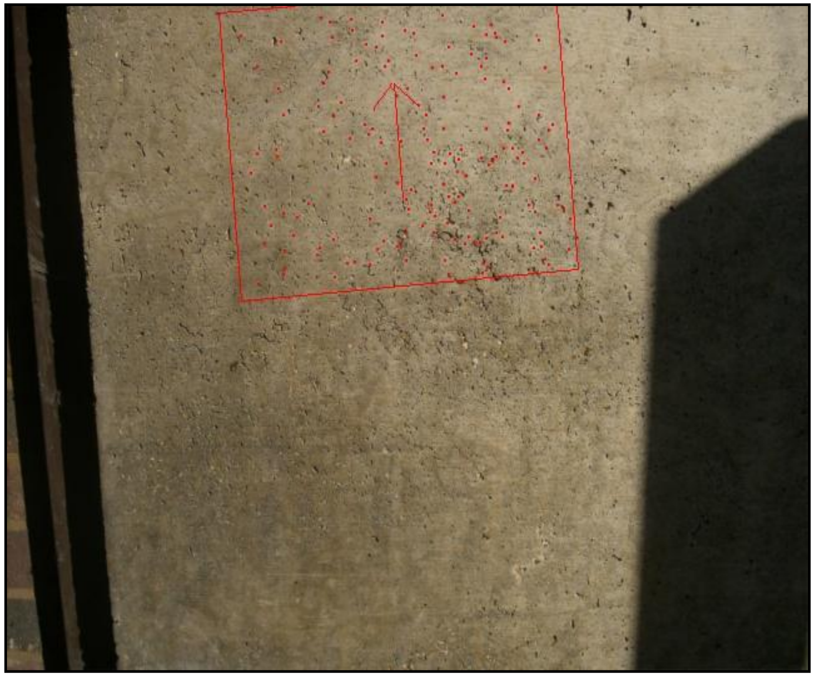
Registration with concrete

- Can appear very repetitive to the eye
- However, plenty of distinguished features can be extracted
- Very accurate matching is possible



Registration with concrete





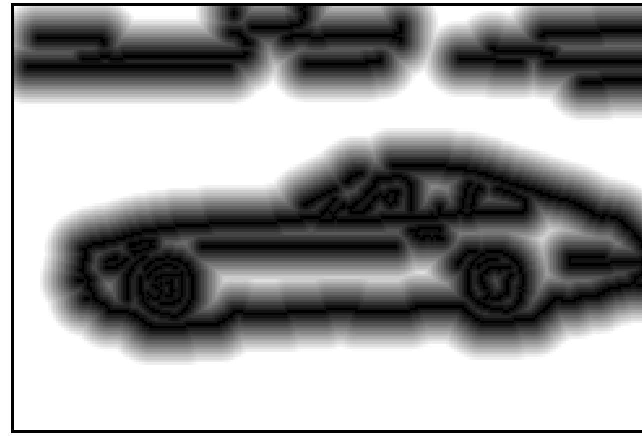
Finding 2D shapes and applications to HCI

Stenger, Thayananthan, Torr and Cipolla 2003
Williams, Blake and Cipolla 2003 and 2006
Ramanan, Fitzgibbon and Cipolla 2006-2007

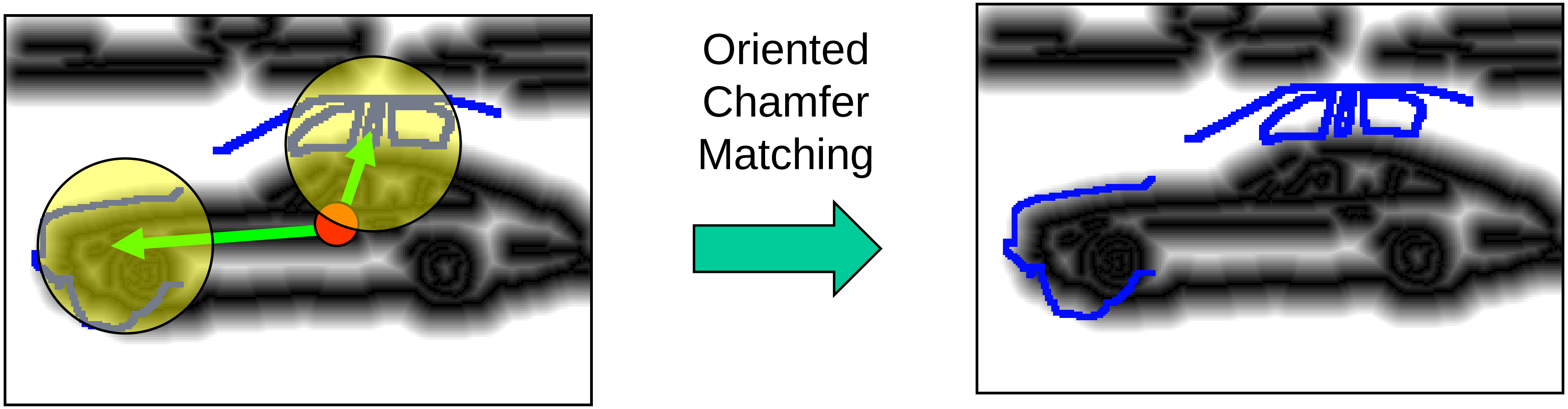
Matching shape templates

Oriented
Canny
Edge
Detector

Oriented
Distance
Transform



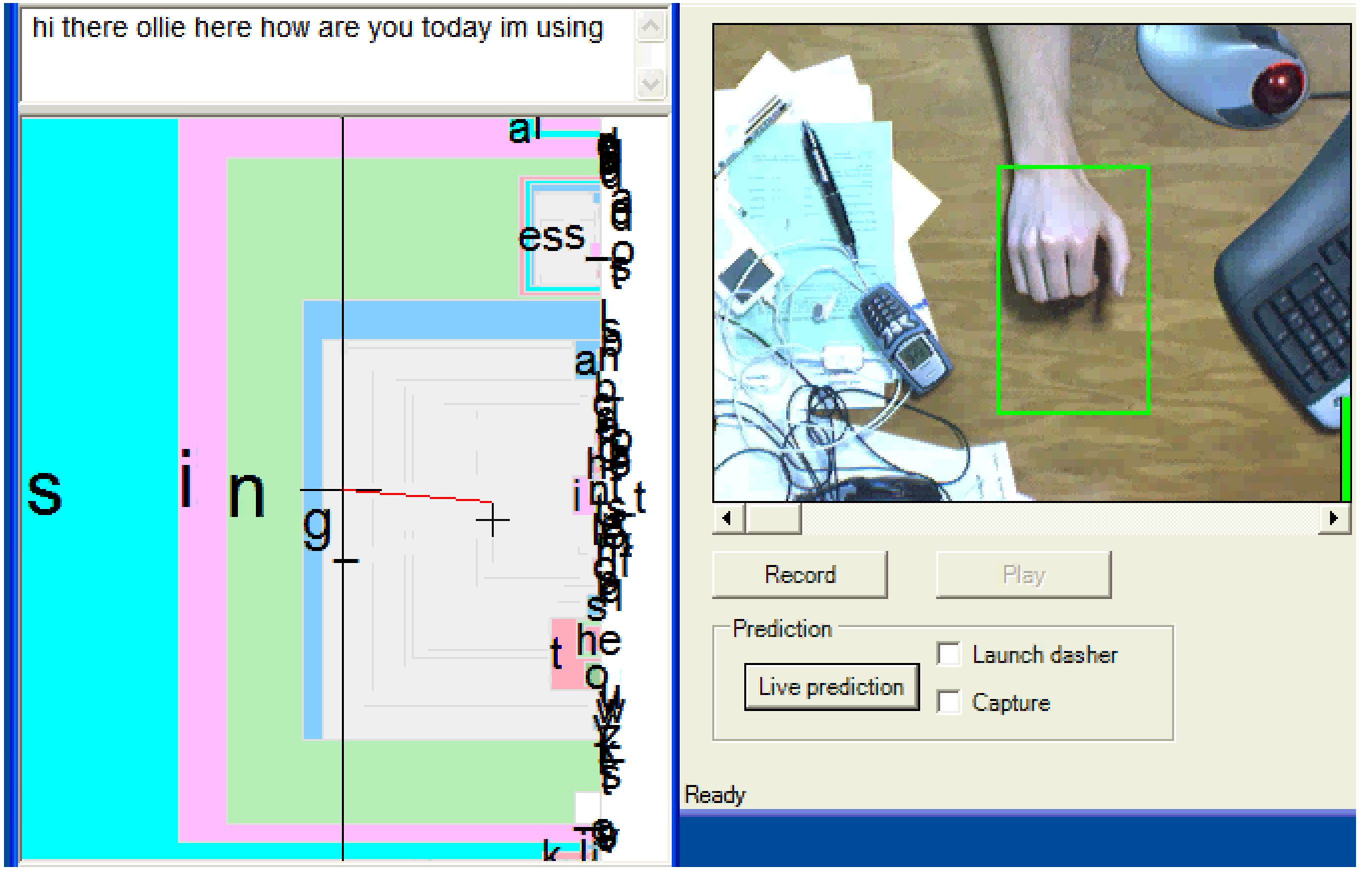
Matching shape templates



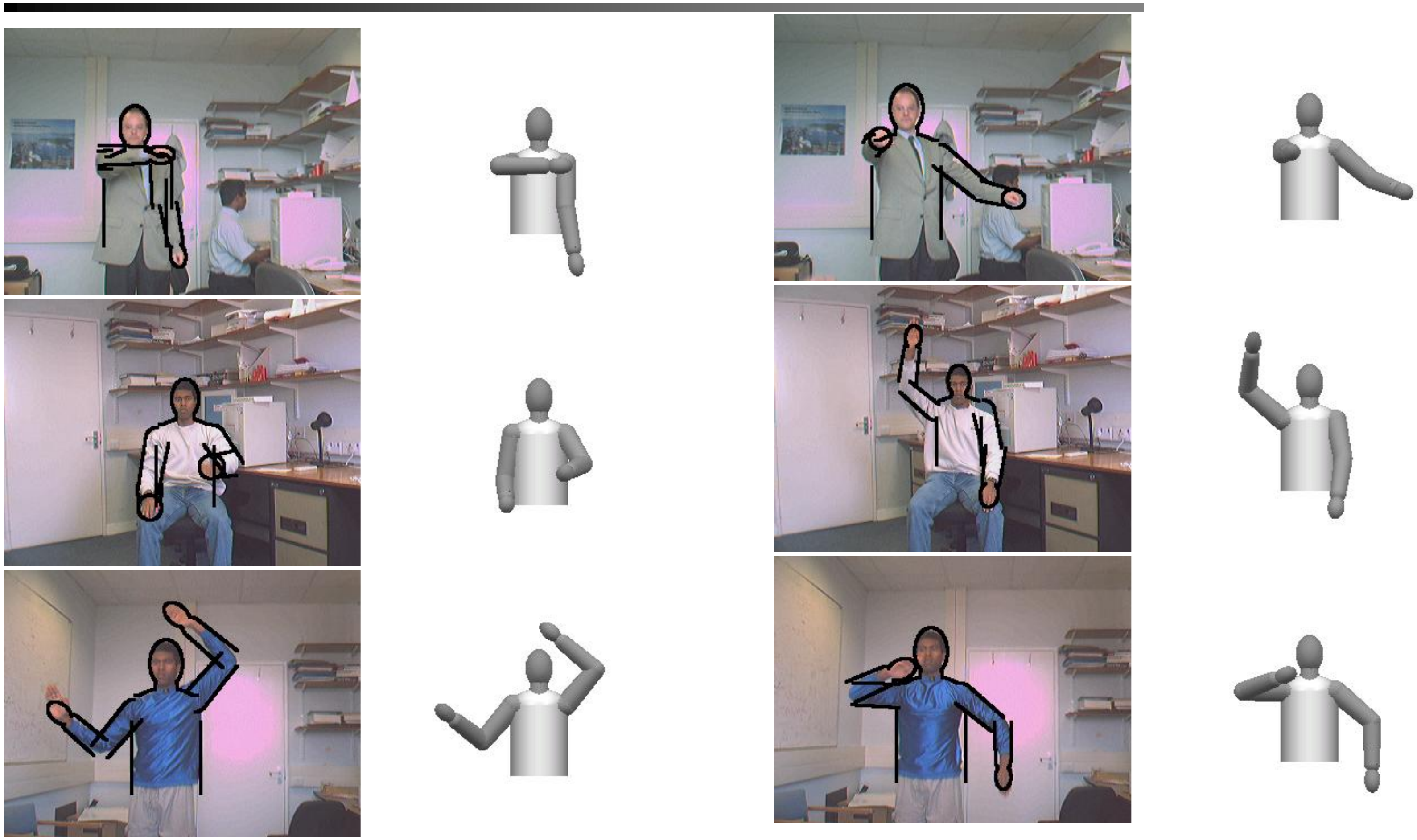
Hand detection system

Tracking - 3D mouse

Real-time visual controller for Dasher



People and pose detection



People and pose detection

UNIVERSITY OF CAMBRIDGE

III. Object recognition and machine learning

Shotton, Blake and Cipolla 2005-2007

Kim, Kittler and Cipolla 2006

Wong and Cipolla 2007

Overview

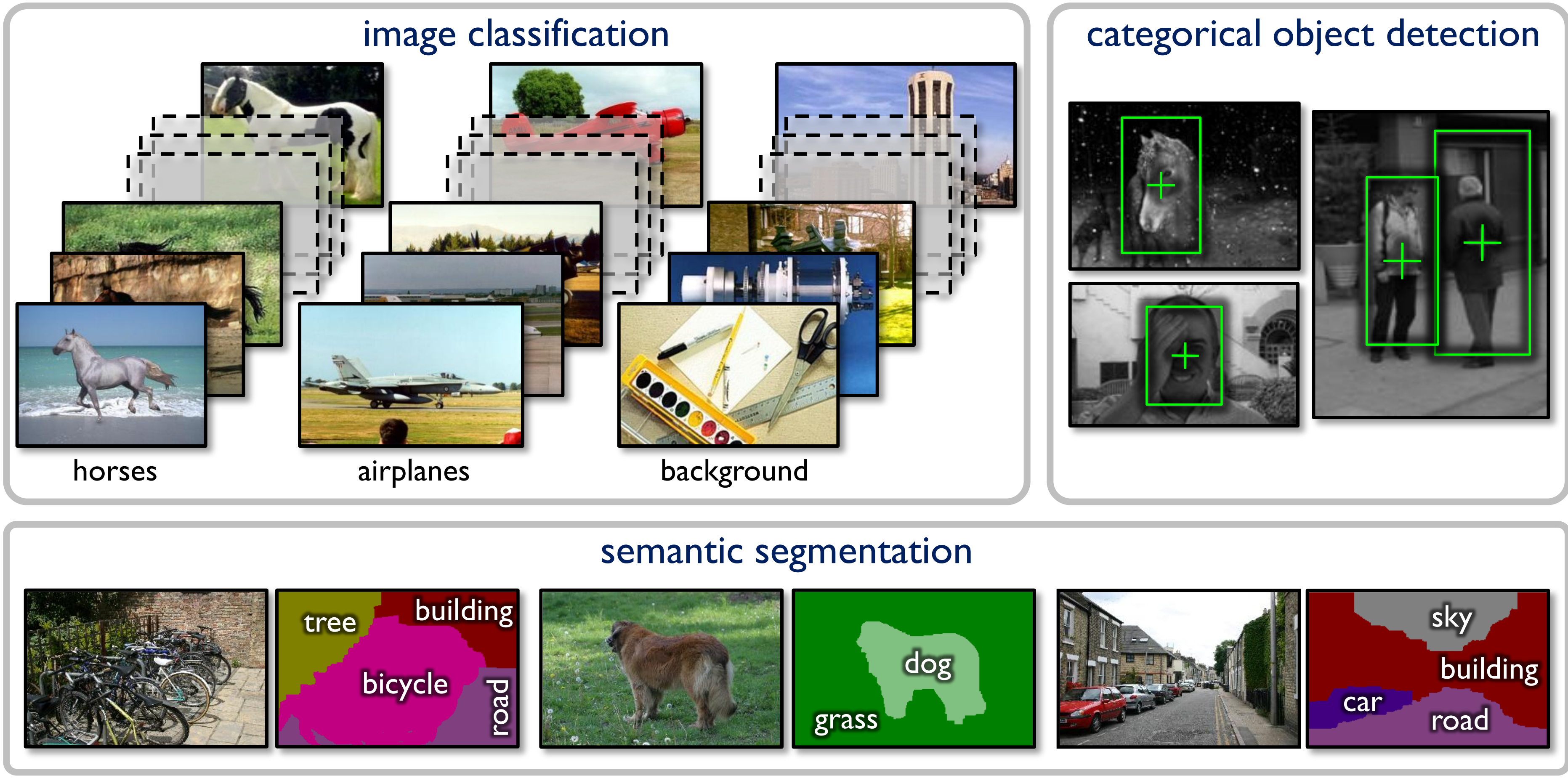


image classification
horses
airplanes
background
categorical object detection
semantic segmentation
tree
building
bicycle
road
dog
grass
sky
building
car
road

UNIVERSITY OF CAMBRIDGE

Using interest points and visual words

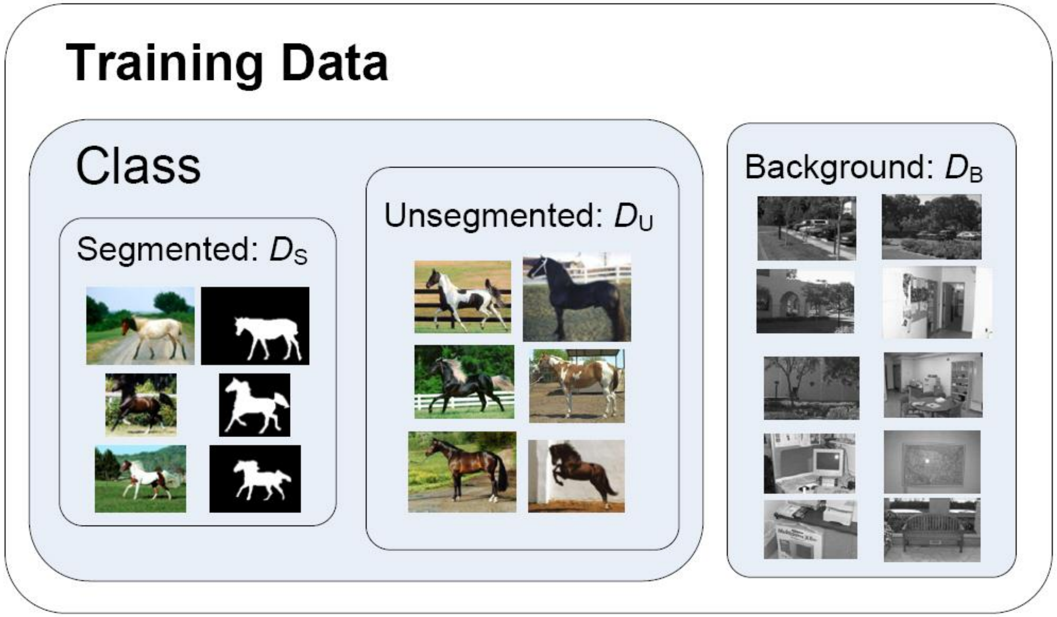
Johnson and Cipolla

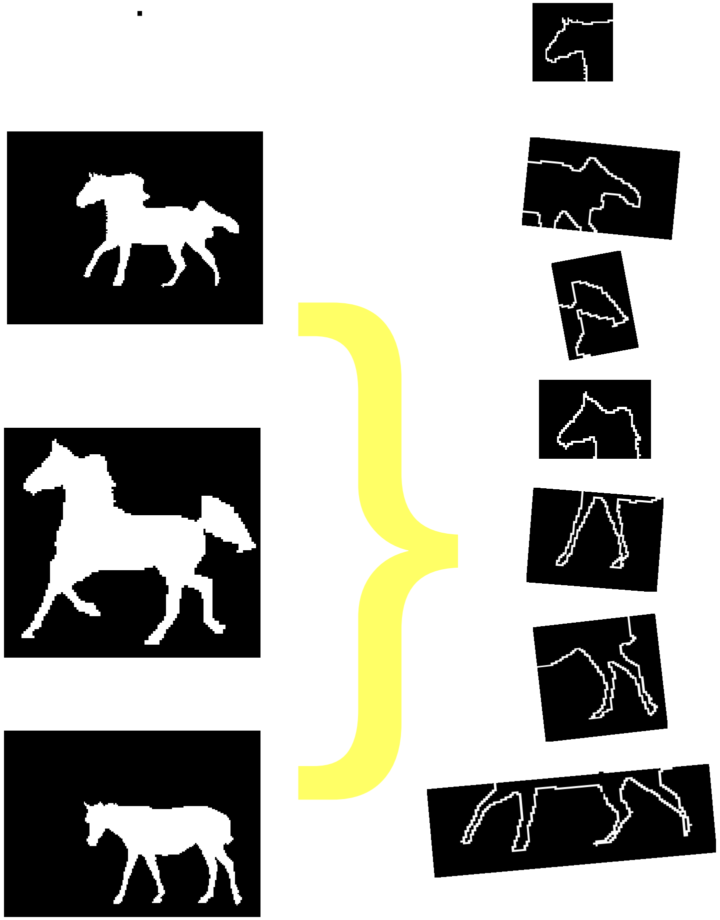
Image matching

Using contour and shape

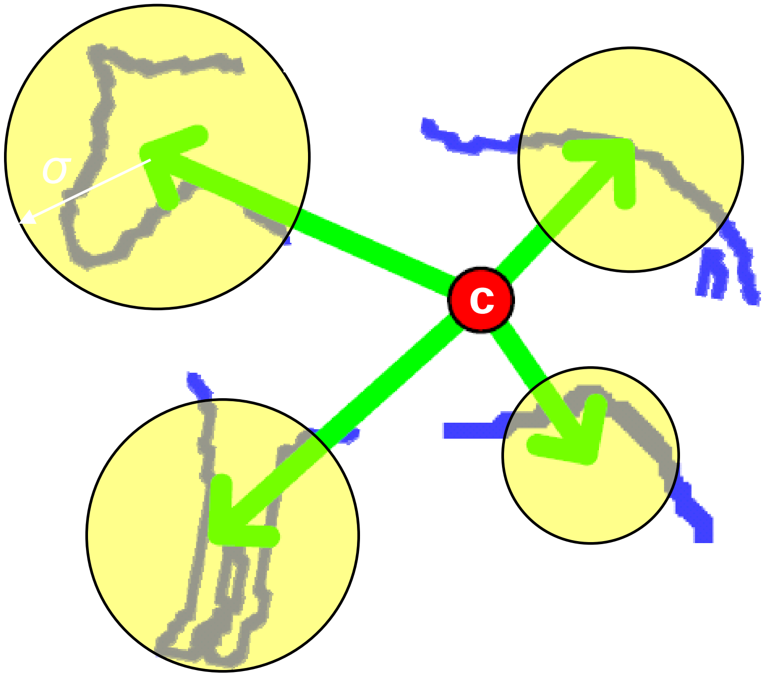
Shotton, Blake and Cipolla 2005-2007

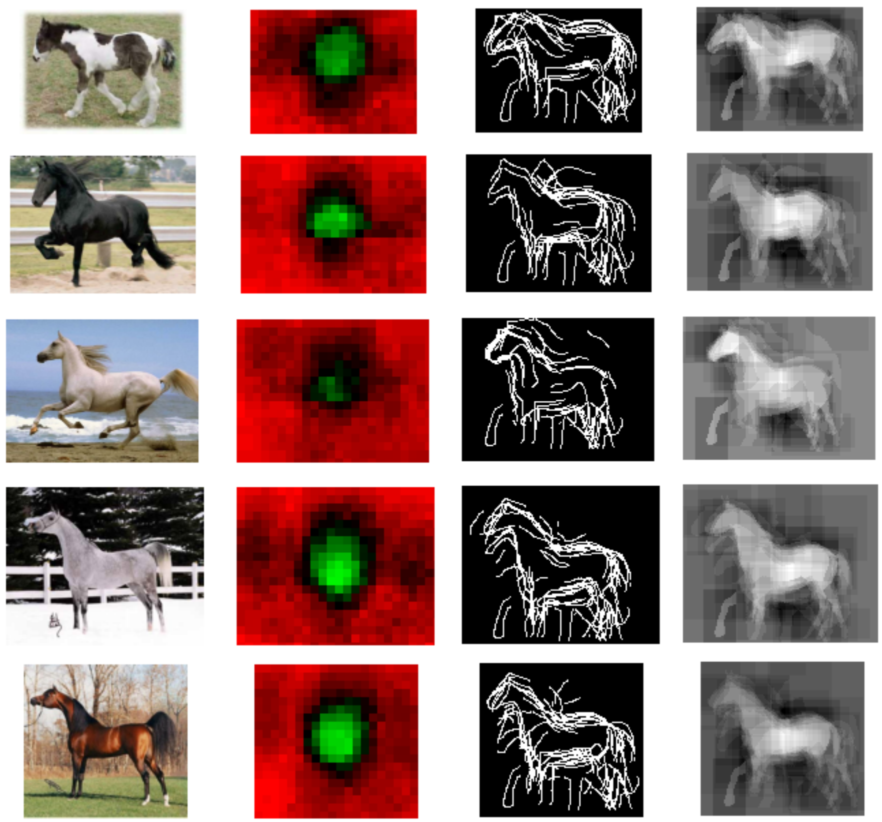
Learning and Adaptability



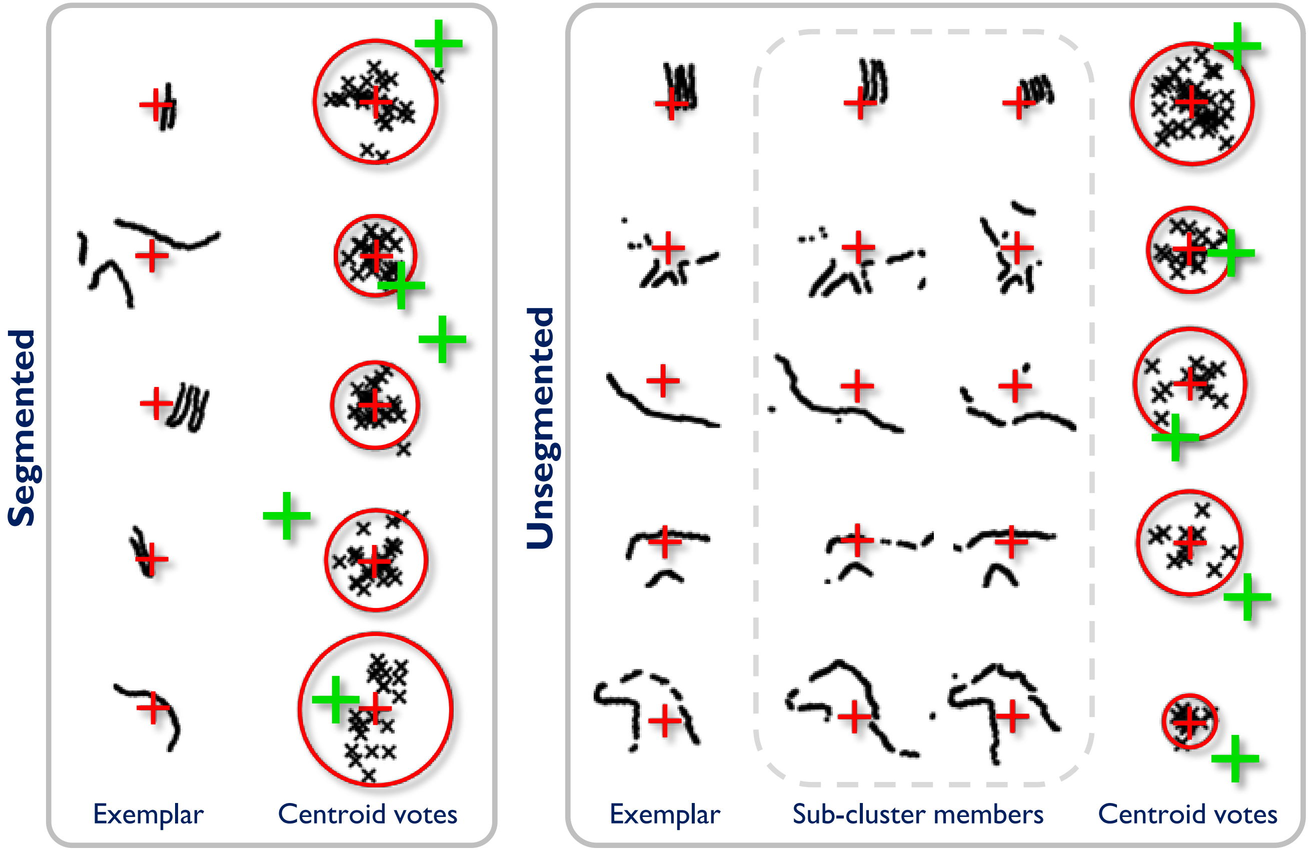


Object Model

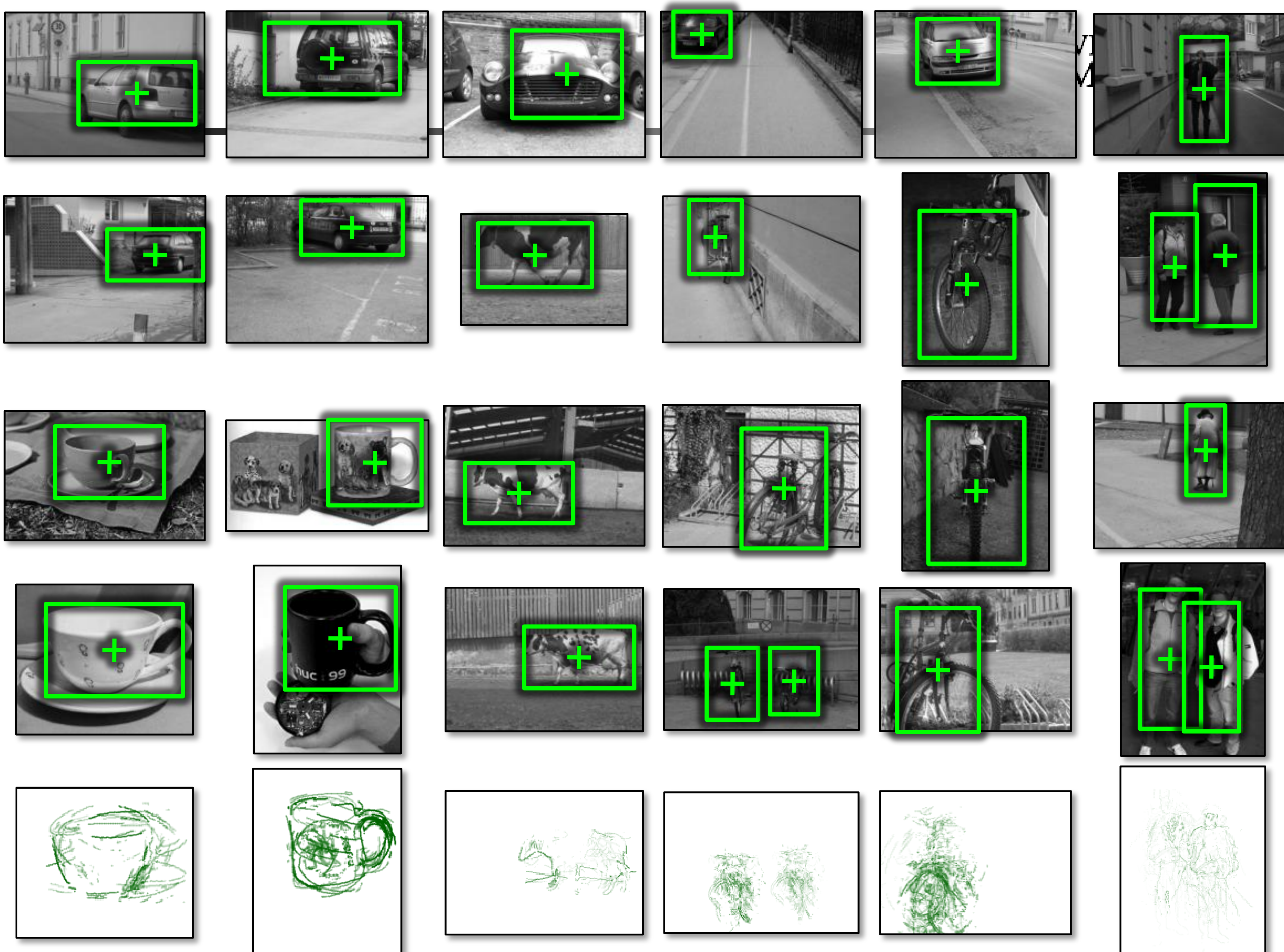




Shape

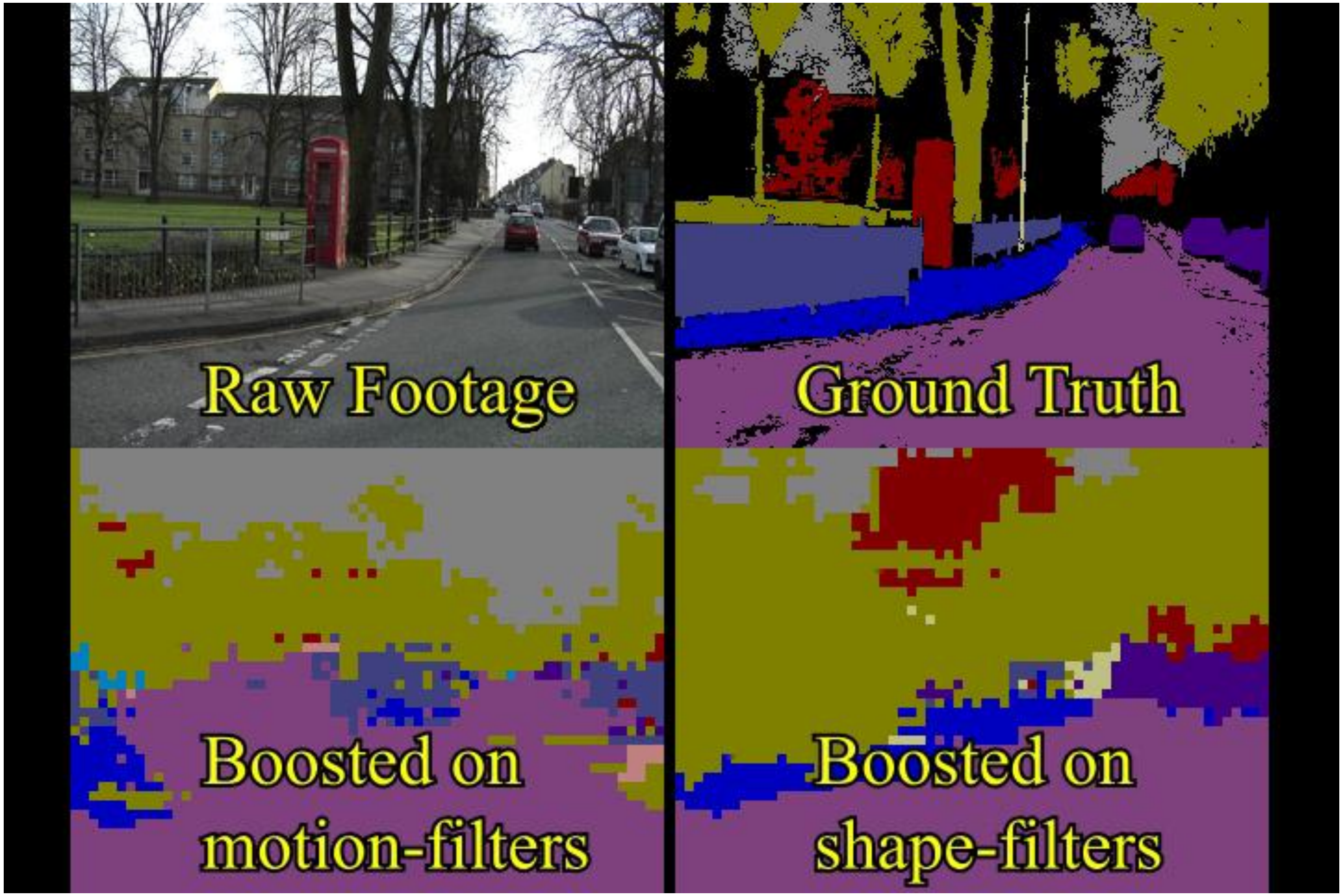


Segmented
Exemplar
Centroid votes
Unsegmented
Exemplar
Sub-cluster members
Centroid votes


huc 99

Recognition in video

UNIVERSITY OF CAMBRIDGE



Making machines see

UNIVERSITY OF CAMBRIDGE

- What is vision and how to duplicate it?
- 3D shape: making digital copies of sculpture from photographs from multiple viewpoints
- Recognition of a painting/picture from a single photo using a mobile (camera) phone
- Detection of objects: hands, faces and people and use in novel man-machine interfaces

Summary

UNIVERSITY OF CAMBRIDGE

- Image registration and matching
- 3D shape from uncalibrated images.
- Object detection and tracking