Making Machines See

Roberto Cipolla
Department of Engineering

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

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

Target detection and pose estimation
Registration
Registration
Reconstruction?

Recovery of 3D shape from images
Reconstruction
3D models
Recognition?
Recognition

image classification

horses
airplanes
background

categorical object detection

car
building
sky

semantic segmentation

tree
building
bicycle
road

dog
grass
car
buildings
road
Pedestrian detection
Existing applications

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
How to make machines that see?
Why not study biology?
Perspective
Machine Learning
Machine Learning
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

- 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
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 visual hull (largest object that can fit inside silhouettes)
Photo-consistency

Non-photo-consistent point
Photo-consistency

Photo-consistent point
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
3D Models
Final installation
3D models
Multiview photometric stereo

Vogiatzis, Hernandez and Cipolla 2006 and 2008
Untextured objects

• 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
Surface Evolution of 3D Mesh
3D Models
Making physical copies

Real

Replica
3D Models
3D Models
3D Models
3D Models
3D Models
Deformable objects:

Real-time photometric stereo using colour lighting

Hernandez et al 2007
Photometric stereo with colour
Deformable objects?

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

• 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

Tracking Untextured Deformable Surfaces Under Multispectral Lighting

Paper ID 2192
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

Oriented Chamfer Matching
Hand detection system
Tracking - 3D mouse
Real-time visual controller for Dasher
People and pose detection
People and pose detection
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
- car
- sky
- building
- road
Using interest points and visual words

Johnson and Cipolla
Image matching
Using contour and shape

Shotton, Blake and Cipolla 2005-2007
Learning and Adaptability

Training Data

Class
Segmented: $D_S$

Unsegmented: $D_U$

Background: $D_B$
Object Model

\[ F = (T, \sigma, \lambda, \theta, a, b) \]
Shape

Exemplar  Centroid votes  Exemplar  Sub-cluster members  Centroid votes
Recognition in video

Raw Footage

Ground Truth

Boosted on motion-filters

Boosted on shape-filters
Making machines see

- 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

• Image registration and matching

• 3D shape from uncalibrated images.

• Object detection and tracking