Charlie's notes

December 3, 2015

1 CS31310 - Agile

2 CS36110 - Machine Learning

3 CS34110 - Computer Vision

3.1 November 20: Motion Models

Modelling Change & Tracking

Motion:

- Background Subtraction
- Optical Flow

Mixture of Gaussians (MoG):

- Robust to noise
- Handles shadows ok
- Common first step

3.1.1 Tracking: Modelling change

Video:

- detections in each frame
- detections are noisy & computationally expensive
- tracking mitigates both issues

Noise can occur if the camera on a robot/car is moving up/down

3.1.2 A general framework for tracking

Recursively:

- An idea about how something will change (*Model*)
- Make a prediction (*Predict*)
- See what happens (*Measure*)
- Update model (*Update*)

Advantages:

- Smooths the data
  - estimate location upon predictions & the measurement
- Constrains search
  - start looking for target in the location it was last seen

3.1.3 Kalman Filter

- Like predict, measure, update from earlier
- Useful for tracking
- Copes well with missing information (occlusions)

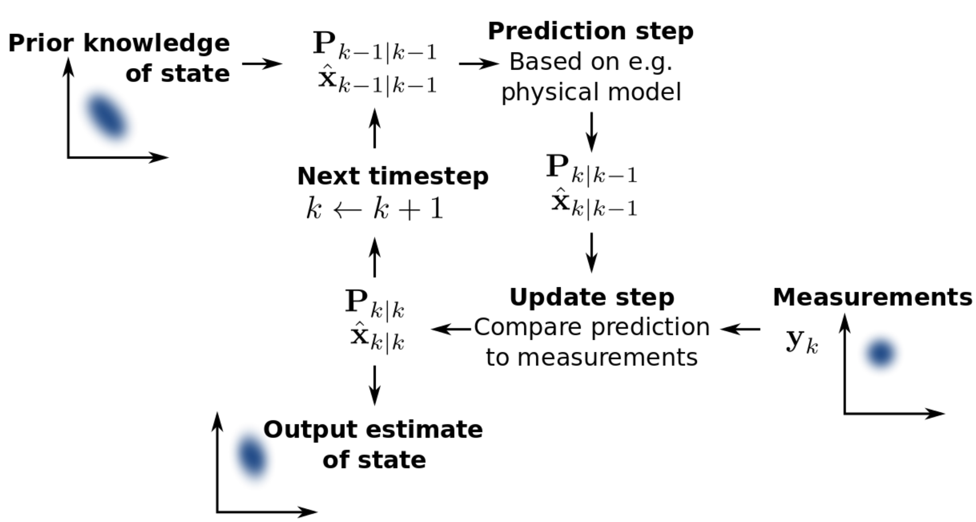


Figure 3.1: Sourced from Wikipedia

Background subtraction	→	Pixels grouped into objects	→	Tracker
Sparse Optical Flow	→	Features grouped into objects	→	
Face Detection	→	→	→	

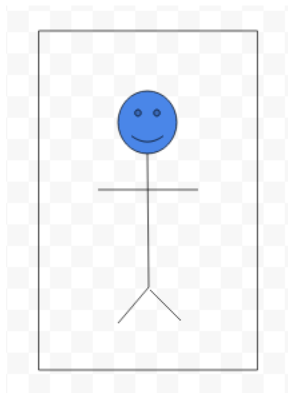
Use Kalman to smooth any measurement

- X,Y location
- size
- colour

See also: Particle filtering: works with combining and splitting objects (e.g. people holding hands, then letting go)
Hannah's video: `https://www.youtube.com/watch?v=NYdwpX1a7-Y`

3.1.4 Mean Shift

Computer the mean of the data within the window
Shift the window to the mean every time



Notes:

- Changes size - can use CAM-Shift to mitigate?
- Lighting change - not really, gradually changes mean over time
- If it picks up something you're not looking for, will slowly drift off

3.1.5 Problems with Tracking

- Initialisation (what are you tracking?)
- Having more than 1 item to track
- Losing target due to motion / occlusion
- Losing target due to appearance change

Usually initialise from a detector of some sort
Useful speed up for detectors & accuracy
Look into: TLD: Tracking Learning Description

3.1.6 Closing note

Vision Systems tend to have multiple layers
Tracking is extremely common in anything which deals with change.

HOG for example, has several layers:

- 2D filters
- Tangent
- Histogram
- Superimpose grid
- SVM

May be more, however couldn't write quick enough... You get the idea...

3.2 November 24: Shape from Shading

3.2.1 Shape

3D structure of:

- Object
- Scenes

Why?

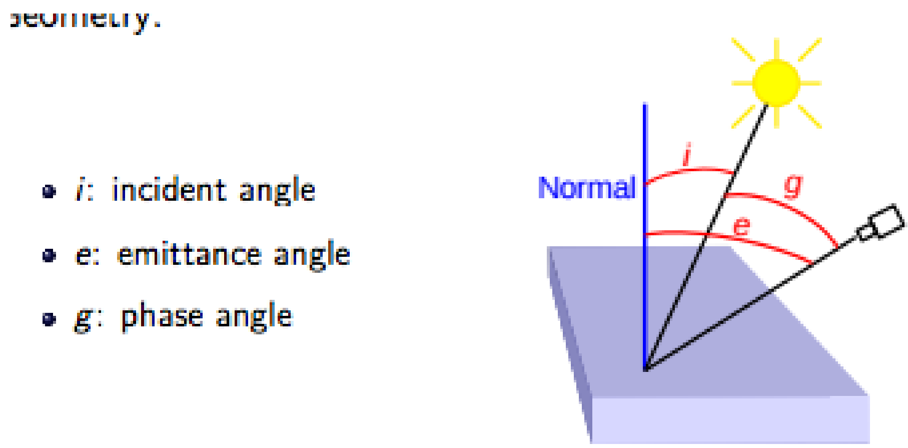
- Science (surface of Mars)
- Graphical (3D model of heritage)

There is a table in the slides of the topics that will be covered in the 'Shape' series

3.2.2 Shape from Shading

Brightness depends on:

- where light source is
- where viewer is
- local orientation of surface
- properties of surface (matt vs glossy)



BRDF:

- Bidirectional Reflectance Distribution Function
- fraction of light reflected in the direction of the viewer
- Lambertian: $\Phi(i, e, g) = \cos i$
- Specular: $\Phi(i, e, g) = 1$ when $i = eandg = i + e$
- Largely property of the surface
- Used in graphics a lot
- Difficult to establish wavelength of light & properties of surface
- Determined experimentally

3D Surface information **Look up - 'Normal Mapping' wikipedia page**
Orientation of surface described by a normal vector (x,y,z):

- Change in z when x changes (p)
- Change in z when y changes (q)

3.2.3 Gradient Space

$p = \dfrac{\delta z}{\delta x}$
$q = \dfrac{\delta z}{\delta y}$

Reflectance Map:

- For surface of 'x' orientation, we expect this reflection back

Light Change: 1 = lightest, 0 = darkest

- Global solution found by integration in gradient and image space + smoothness assumption:
  - Maths in confusing!
  - Don't need to know the maths for the exam

3.2.4 Photometric Stereo

We can reduce assumptions if we have multiple lighting conditions.
Example: 2 overlapping reflectance maps on sphere

- Use same point in image plane under different lighting
- You know where the lights are (calibrated)
- Nothing moves - just the lights on/off
- Calibration with sphere of similar reflectance
- There is always noise
- Reasonably robust

This is at the heart of light-stage imaging

- Illuminate with 1 source at a time
- Calibrate & results in look-up table

Have to reduce inter-reflections!

x,y,z mapped to RGB figures:

- gives diagram of 3D structure

Fewer images → More assumptions
⋆ **Look up ISO-contours**

3.3 November 27: Shape from X

3.3.1 Shape from Texture

1 + images, with a lot of assumptions.

Marr 1982:

- 2 textured gradients
- Limited information
- Circles turn into ellipses (illusion of depth)

Types of textures:

Isotropic: rotation perpendicular to texture plane does not 'change' the texture

Homogeneous: texture looks 'the same' everywhere

Texel: basic texture elements, the repetition of which creates the texture

Local Binary Patterns

- 3x3 filters
- Capture changes in greyscale

1	1	0
0	T	1
0	1	0

0	1	1
0	T	1
0	0	1

Figure 3.2: Top: Standard output. Bottom: Could be used to detect edges

0 = darker than T 1 = lighter than T

Global planes

- Find the plane, uphill, downhill etc
- Gradients in texture
- Isotropic texture:
  - circles $\rightarrow$ ellipses

- Major axes of ellipses to work out slant & tilt

- Homogeneous texture:
  - textures made of dots
  - look for a density change

Generally you fit methods locally rather than globally

3.3.2 Shape from Motion

Look up Marr (again)
Can also do structure from Motion: this is next week!
The examples on BB are really good: biomotionlab!
Things further away move more slowly - can infer things from this

Assumes:

- rigidity
- parallel projection

Works both long and short range...

- Initial guess on 3D model of world (flat screen)
- + new image
- update on those which disagree with the initial model

Very popular!

3.3.3 Shape from Occlusions

Infer shapes from which bits are in front of other parts

Occlusion contours:

- discontinuity in depth signal
- corresponds to silhouette (shadow)

Assumptions (1 image):

- each point on contour corresponds to a single point on the object (except when a 'T' junction)
- nearby points on the contour correspond to nearby points on the object
- points on contour correspond to planar points on the object

When you have 2+ images:

- Space Carving
- Volumetric reconstruction
- Gives where the object is isn't
- Object is intersection of several generalised cones

Watch the Youtube video on BB - Greenvision

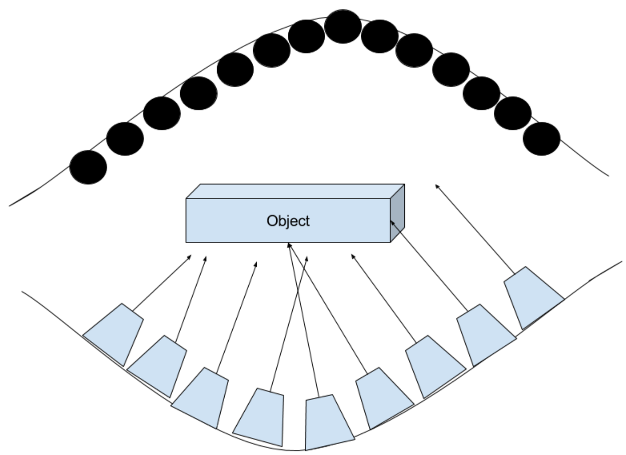


Figure 3.3: Can the cameras see the points on the other side of the object? Make a map of what it can / can't see to build a shape of the object using multiple cameras (bottom row).

From a video sequence:

- inferred from motion & occlusion
- what occludes people in the scene?

Things at the front = black, as it's never occluded from you
Lots of motion = white

3.3.4 Shape from focal length

Depth from defocus
Things are blurry past a depth of focus Take 10 pictures, at different lengths, with different focuses

- Texture really helps focus
- Assumptions about smoothness

Markov-Random Field

How to tell if in focus?

- Look for sharp edges = ✓
- Only low-frequency components = ✗

---

4 CS32310 - Advanced Computer Graphics

5 SE31520 - Internet-based Applications

Fred Long's lecture slides contain all the content.

6 Other