Computer Vision

António J. R. Neves / Paulo Dias

Departamento de Electrónica, Telecomunicações e Informática

Universidade de Aveiro

an@ua.pt / paulo.dias@ua.pt
http://elearning.ua.pt/

Contents

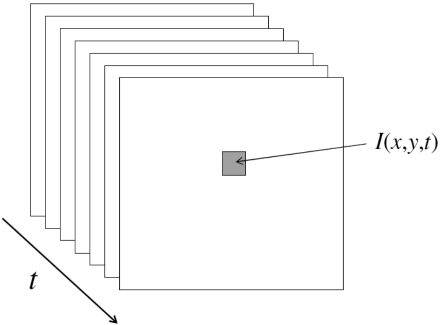
1. Motion
2. Background subtraction
3. Tracking

Contents

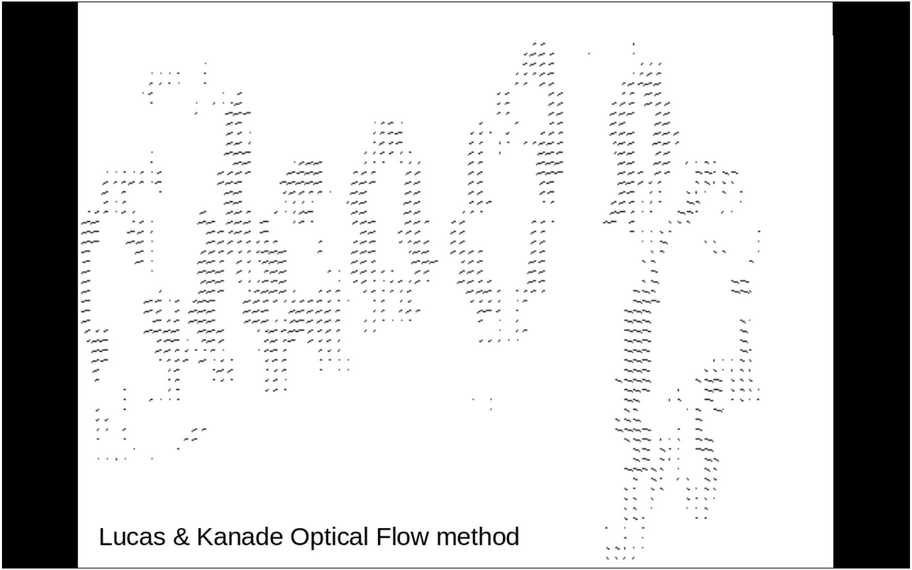
1. Motion
2. Background subtraction
3. Tracking

Video

- A video is a sequence of frames captured over time
- The image data is a function of space (x, y) and time (t)



Motion (1)


Lucas & Kanade Optical Flow method

Motion analysis

- Several information can be extracted from time varying sequences of images:
  - Camouflaged objects are only easily seen when they move
  - The relative sizes and position of objects are more easily determined when the objects move
  - Even simple image differencing provides an edge detector for the silhouettes of texture-free objects moving over any static background.

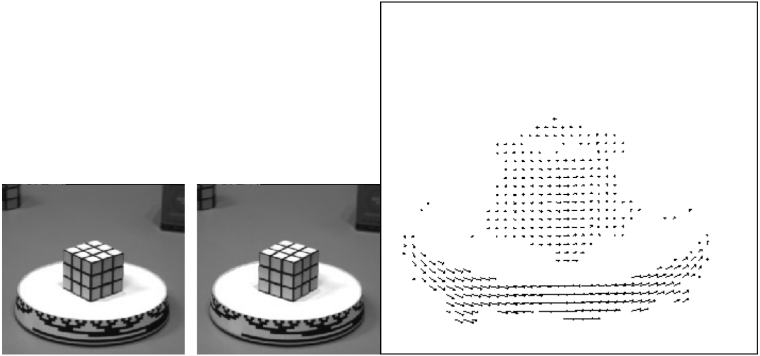
Motion analysis

- The analysis of visual motion can be divided into two stages:
  - the measurement of the motion
  - the use of motion data to segment the scene into distinct objects and to extract three dimensional information about the shape and motion of the objects.
- There are two types of motion to consider:
  - movement in the scene with a static camera,
  - and movement of the camera, or ego motion.
- Since motion is relative, these types of motion should be the same. However, this is not always the case, since if the scene moves relative to the illumination, shadow and specularities effects need to be dealt with.

Uses of motion

- Estimating 3D structure
- Segmenting objects based on motion cues
- Learning dynamical models
- Recognizing events and activities
- Improving video quality (motion stabilization)

- The motion field is the projection of the 3D scene motion into the image.



Motion field + camera motion

- Length of flow vectors inversely proportional to depth Z of 3D point - points closer to the camera move more quickly across the image plane.

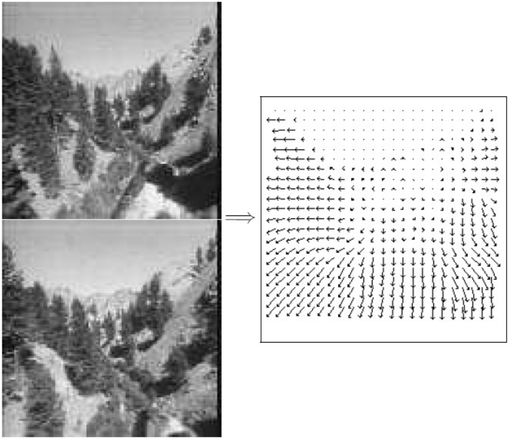


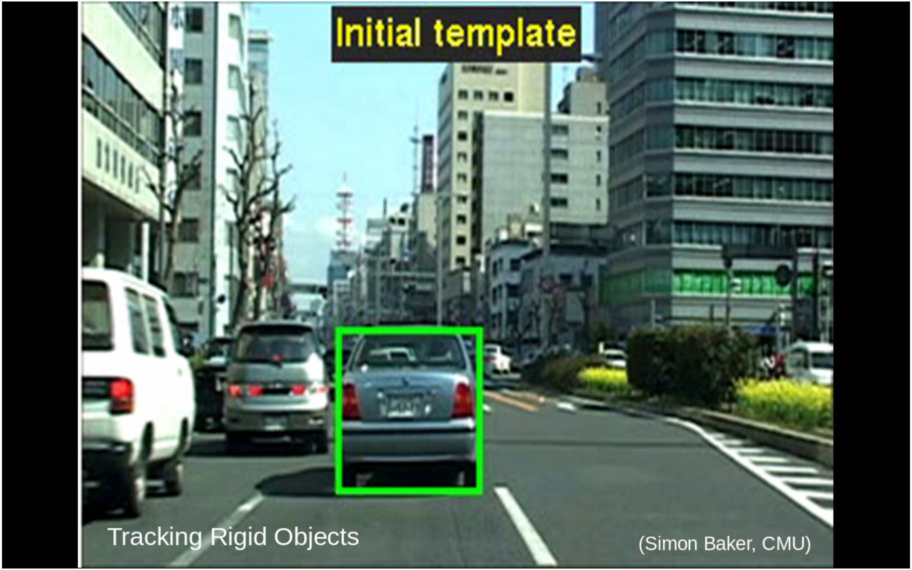
Figure 1.2: Two images taken from a helicopter flying through a canyon and the computed optical flow field.

Optical flow

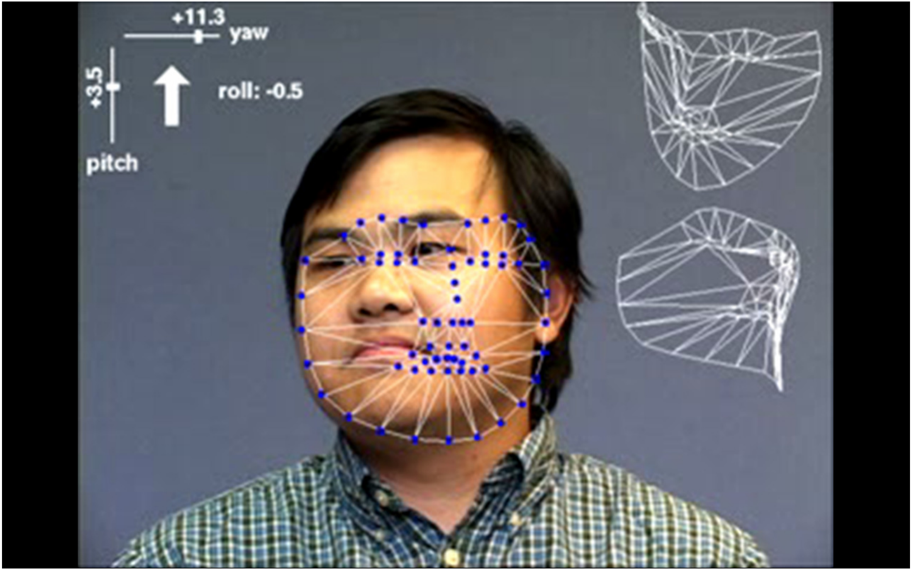
- Definition: optical flow is the apparent motion of brightness patterns (or colors) in the image.
- Ideally, optical flow would be the same as the motion field.
- Have to be careful: apparent motion can be caused by lighting changes without any actual motion.
- To estimate pixel motion from image we have to solve the pixel correspondence problem.
- Given a pixel in frame t, look for nearby pixels with same characteristics (color, brightness, ...) in frame $t-1$.

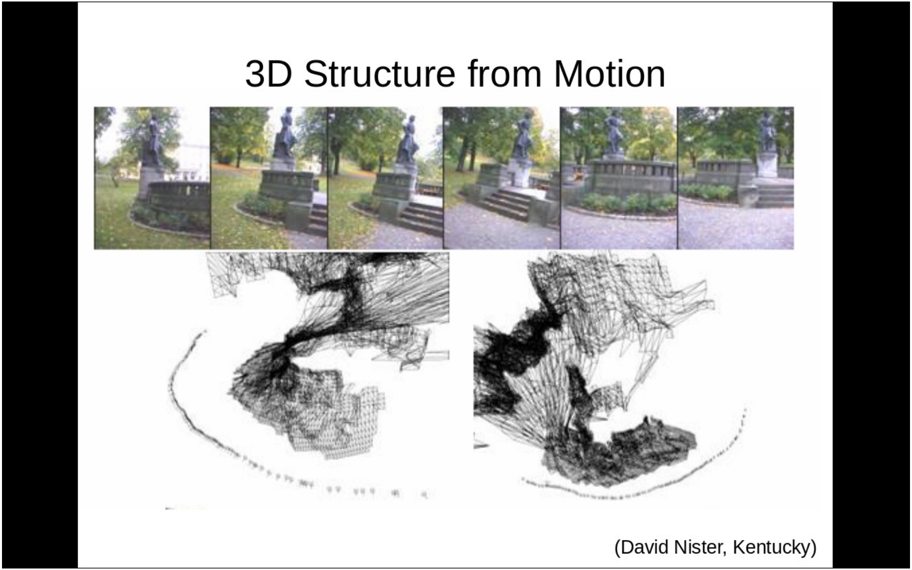
What can i use it for?

Tracking rigid objects

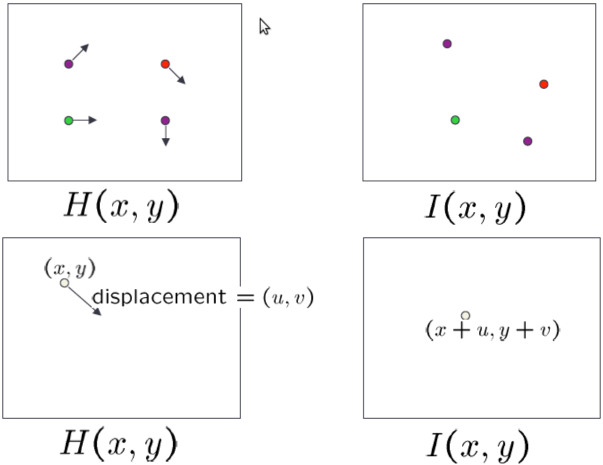


Tracking deformable objects (1)


+11.3
yaw
+3.5
roll: -0.5
pitch


3D Structure from Motion
(David Nister, Kentucky)

Optical flow constraints

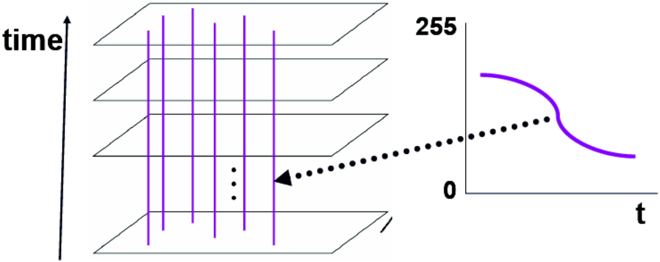


- $I(x,y) = H(x+u, y+v)$
- brightness constancy and small motion

Contents

1. Motion
2. Background subtraction
3. Tracking

- It is possible to look at video data as a spatio-temporal volume.
- If camera is stationary, each line through time corresponds to a single ray in space.



Background subtraction

- Background subtraction is a commonly used class of techniques for segmenting out objects of interest in a scene for applications such as:
  - Surveillance
  - Robot vision
  - Object tracking
  - Traffic applications
  - Human motion capture
  - Augmented reality

- It involves comparing an observed image with an estimate of the image if it contained no objects of interest.
- The areas of the image plane where there is a significant difference between the observed and estimated images indicate the location of the objects of interest.
- The name background subtraction comes from the simple technique of subtracting the observed image from the estimated image and thresholding the result to generate the objects of interest.

Important issues

- foreground detection – how the object areas are distinguished from the background;
- background maintenance – how the background is maintained over time;
- post-processing – how the segmented object areas are postprocessed to reject false positives.

$\Rightarrow$

Generic algorithm

- Create an image of the stationary background by averaging a long sequence.
- Difference a frame from the known background frame
- Motion detection algorithms such as these only work if the camera is stationary and objects are moving against a fixed background

Image at time t: $I(x,y,t)$ $\Downarrow$

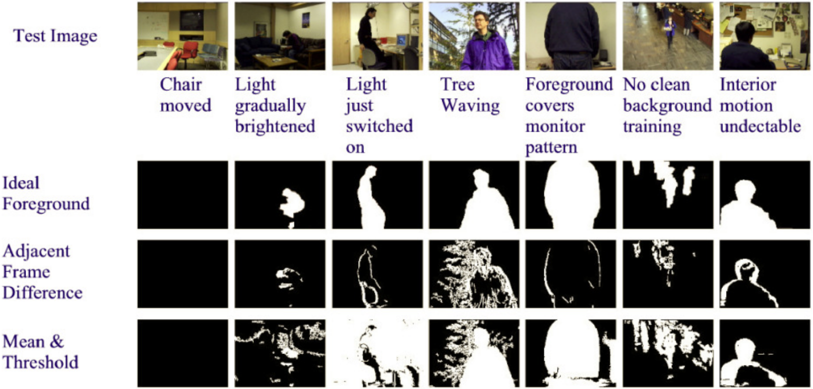
$-$

Background at time t: $B(x,y,t)$ $\Downarrow$

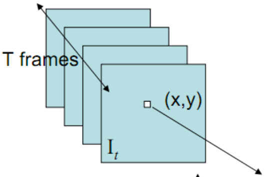
$| > Th$

- With frame differencing, background is estimated to be the previous frame. Background subtraction equation becomes:
  $B(x, y, t) = I(x, y, t-1)$ and $|I(x, y, t) - I(x, y, t-1)| > Th$
- Depending on the object structure, speed, frame rate and global threshold may or may not be useful (usually not).
- Another approach is to model the background using a running average. A pixel is marked as foreground if $|I_t - B_t| > \epsilon$, where ϵ is a predefined threshold. The thresholding is followed by morphological closing with a 3x3 kernel and the discarding of small regions
- The background update is
  $B_{t+1} = \alpha I_t + (1-\alpha)B_t$,
  where α is kept small to prevent artificial tails forming behind moving objects.

Some examples



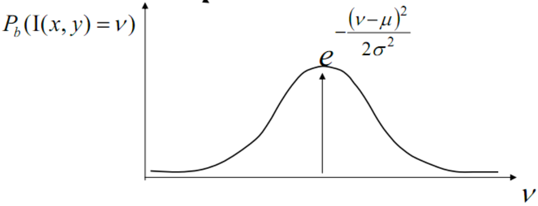
- Adaptive Mixture of Gaussians

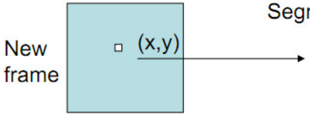


Background model

$$\mu(x,y) = \frac{1}{T}\sum_{t} I_t(x,y)$$

$$\sigma^2(x,y) = \frac{1}{T}\sum_{t} (I_t(x,y) - \mu(x,y))^2$$





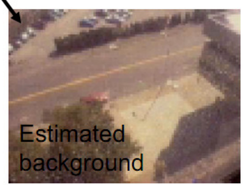
Segmentation - moving objects:

$$\left| \mathrm{I}(x, y) - \mu(x, y) \right| > 2\sigma$$

Estimated background

The most probable background image

dominant Gaussian mean for each pixel's mixture model



Contents

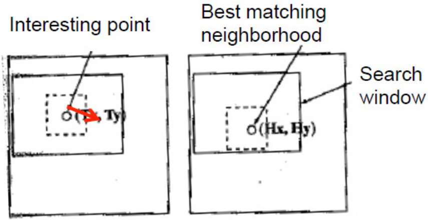
1 Motion

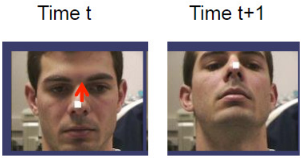
2 Background subtraction

3 Tracking

- Object tracking is a crucial research issue in computer vision, especially for the applications where the environment is in continuous changing:
  - Robot Vision - mobile robot navigation, applications that must deal with unstable grasps, ...
  - Surveillance
  - Traffic applications
  - Human motion capture

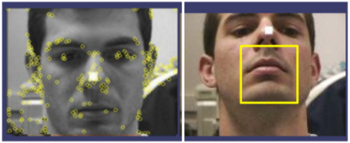
- Search window is centered at the point where we last saw the feature, in image I_t.
- Best match = position where we have the highest normalized cross-correlation value.

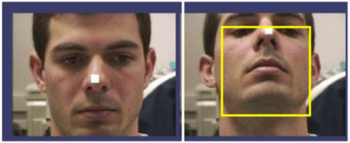




Feature-based matching for motion (2)

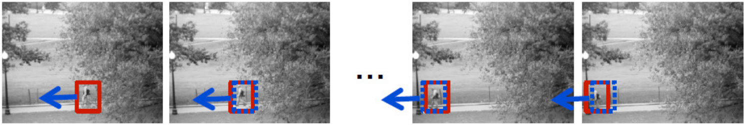
- For a discrete matching search, what are the tradeoffs of the chosen search window size?
- Which points to track? - Select interest points – e.g. corners, edges, etc.
- Where should the search window be placed? - Near match at previous frame; More generally, taking into account the expected dynamics of the object.





Detection vs. tracking

- Detection: We detect the object independently in each frame and can record its position over time, e.g., based on blob's centroid or detection window coordinates
- Tracking with dynamics: We use image measurements to estimate the position of the object, but also incorporate position predicted by dynamics, i.e., our expectation of object's motion pattern.



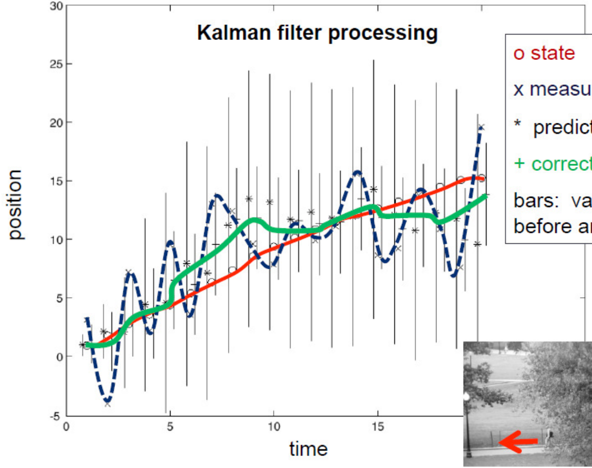
Tracking as inference

- The hidden state consists of the true parameters we care about, denoted X.
- The measurement is our noisy observation that results from the underlying state, denoted Y.
- At each time step, state changes (from X_{t-1} to X_t) and we get a new observation Y_t.
- Hidden state : parameters of interest
- Measurement : what we get to directly observe
- Our goal: recover most likely state X_t given all observations seen so far and the knowledge about dynamics of state transitions.

Kalman filter

- Method for tracking linear dynamical models in Gaussian noise
- The predicted/corrected state distributions are Gaussian
- Only need to maintain the mean and covariance
- The calculations are easy

Example: Constant velocity model



o state

x measurement

* predicted mean estimate

+ corrected mean estimate

bars: variance estimates
before and after measurements

Time t

Time t+1

Kristen Grauman

Tracking: issues (1)

- Initialization - Often done manually (Background subtraction, detection can also be used)
- Data association, multiple tracked objects - occlusions, clutter

Tracking: issues (2)

- Deformable and articulated objects
- Constructing accurate models of dynamics - (example: Fitting parameters for a linear dynamics model)
- Drift - accumulation of errors over time