Video Analytics: Template Matching (Week 6)

Video Analytics

- Video is a display of moving visual media, which can be seen as a sequence of still images.
  - In a general sense: Video is metadata i.e. data that provides information about other data.
- In video analytics, the metadata is processed in real-time and transformed into intelligent data. Examples include:
  - Generation of descriptions of what is happening in the video.
  - Detect and track objects.
- Potential applications are: Motion tracking, object tracking, license plate recognition, facial recognition, crowd detection, indoor people tracking, etc.

Template Matching: Motion/Object tracking

- Template matching (TM) method detects different types of objects with the help of a template.
  - The generated templates from detection module are passed on to the tracking module.
  - TM initiates the tracking process with a given input reference template.
- TM works without tedious training procedures.

Template Matching

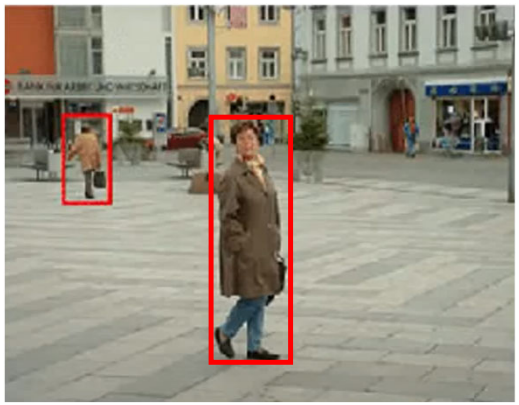
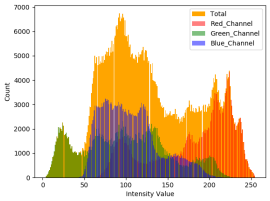


Image template



Histogram template

HoG descriptor template

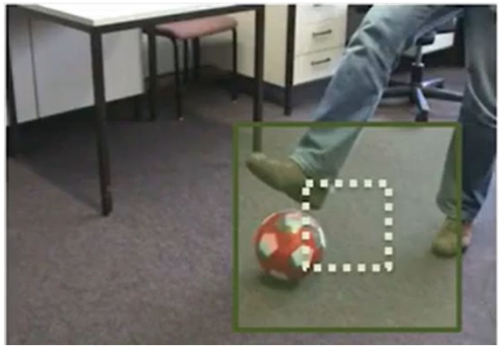
Template Matching: Appearance

- Given a template window *S* in video frame t_{-1}.
  - Extract the template appearance matrix of football.
- Use the appearance template, and find a matching template in frame t_0 using buffer region.

How to reduce complexity for matching the template?

Frame t_{-1}

Template



Frame t_0

How wide the buffer region can be?

Template Matching: Appearance: DIY

- How wide the buffer region can be?
- Assume you want to develop a template matching algorithm for a company that wants to develop a low-cost portable solution, which uses a cheap high-speed camera for monitoring the lawbreakers. Find the quickest solution that can find a lawbreaker.
  - Assume a far-mounted 1MP (megapixel) camera captures 6000 fps.
  - Assume dimension as 1000 x 1000 pixels.
  - Assume a camera is calibrated to monitor 100 m distance horizontally.
  - Assume a lawbreaker drives a car at 180 km/hr.
- Distance covered in each consecutive frame by lawbreaker is 5 cm/frame
- 10cm horizontal distance is covered by each pixel.
- Hence 0.5 pixel of horizontal movement between consecutive frames signifies a lawbreaker.

Template Matching: Appearance

- For template matching, the appearance matrix can be used as it is, or the following matrices can be used:
  - Sum of Absolute Differences (SAD)

    Find set of pixels $(k, l) \in S$ with min/max difference.

$$SAD(k,l) = \sum_{(i,j)\in T} |I_1(i,j) - I_2(i+k, j+l)|$$

  - Sum of Squared Difference (SSD)

    Find set of pixels $(k, l) \in S$ with min/max difference.

$$SSD(k,l) = \sum_{(i,j)\in T} |I_1(i,j) - I_2(i+k, j+l)|^2$$

To find a perfect match, SAD and SSD have to be Min or Max?

Template Matching: Appearance

- Normalised Cross Correlation (NCC)
  Find set of pixels $(k, l) \in S$ with min/max difference.

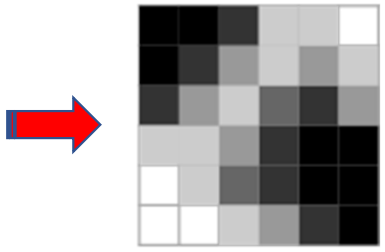
$$NCC(k,l) = \frac{\sum_{(i,j)\in T} I_1(i,j) \cdot I_2(i+k, j+l)}{\sqrt{\sum_{(i,j)\in T} I_1(i,j)^2 \cdot \sum_{(i,j)\in T} I_2(i+k, j+l)^2}}$$

To find a perfect match, NCC has to be Min or Max?

Template Matching: Appearance: DIY

- Assume a 6x6 matrix, and 3x3 template.

0	0	1	4	4	5
0	1	3	4	3	4
1	3	4	2	1	3
4	4	3	1	0	0
5	4	2	1	0	0
5	5	4	3	1	0



1	0	0
1	0	0
3	1	0

- Form a group, and evaluate SAD, SSD, and NCC.

$$SAD(k,l) = \sum_{(i,j)\in T} |I_1(i,j) - I_2(i+k, j+l)|$$

$$SSD(k,l) = \sum_{(i,j)\in T} |I_1(i,j) - I_2(i+k, j+l)|^2$$

$$NCC(k,l) = \frac{\sum_{(i,j)\in T} I_1(i,j) \cdot I_2(i+k, j+l)}{\sqrt{\sum_{(i,j)\in T} I_1(i,j)^2 \cdot \sum_{(i,j)\in T} I_2(i+k, j+l)^2}}$$