Computer Vision 2019 - 2020

Xin Wang
March 18, 2022

1 Question 2

Question 2.1:

1. In the context of image sequence processing, what is a feature and describe why local features are desired for tracking?
   - A feature is a piece of information relevant for solving a computational task related to a certain application e.g. points, edges or objects.
   - Local features (Keypoint features) are features that are in specific locations of the images e.g. mountain peaks or building corners. These local features are unique and disinctive enough from any angle, lighting and scale.
2. What are feature descriptors and why are they used over the original features?

   Feature descriptors are interesting information encoded into a series of numbers that acts as a numerical fingerprint to be used to differentiate one feature from another.

   This is used over original features since feature descriptors are much lower dimensions than the original image and, this reduction in dimensionality, reduces the overheads in processing the images.

Question 2.2:

1. What is meant by the following when measuring tracking errors?
   - TP: True Positive - An outcome where the model correctly predicts the positive class
   - TN: True Negative - An outcome where the model correctly predicts the negative class
   - FN: False Negative - An outcome where the model incorrectly predicts the negative class
   - FP: False Positive - An outcome where the model incorrectly predicts the positive class
2. How are precision and recall defined?
   - Precision: Fraction of relevant instances among the retrieved instances
   - Recall: Fraction of relevant instances that were retrieved

Question 2.3: Describe the steps of the Harris corner detector, providing the necessary equations in matrix form.

The process of the harris corner detector:

1. Compute x and y derivatives of an image
$$I_x = G_x * I; \quad I_y = G_y * I$$
   where G is a filter e.g. Sobel filter
2. At each pixel, compute the matrix M
$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$
3. Calculate detector response
$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2$$
4. Detect interest points which are local maxima and whose response R are above a threshold

Question 2.4: Propose a suitable framework to track the balloons in this sequence, taking into consideration:

- What features you will use and how you will extract them.
- How you would distinguish between balloons and clouds
- A possible way to deal with a balloon disappearing from a few frames due to it being blocked by another. This problem is known as occlusion.

Possible features are patterns on the balloon, the balloon basket and the outline of the balloon. These can be extracted using the SIFT feature descriptor that takes in a 16×16 window to return a 4×4 keypoint descriptor.

The balloons and clouds can be differentiated using shape and colour.