

# Computer Vision Question Pack

Xin Wang  
March 18, 2022

## 1 Year 2018

### Question 1(b):

1. Explain how the Hough transform could be used to find these curve segments.

This is possible using Generalised Hough Transform. As the curves are defined by the equation  $y = ax^2 + bx + c$ , the equation can be used to create a discrete voting space and perform voting in the parameter space.

2. Discuss how to limit the voting space to 2D to solve the above problem. What are the main advantages of using low dimensional voting spaces?

Low dimensional voting spaces reduces the computational complexity required. This reduction in processing load means a faster feature extraction.

**Question 4(a):** Explain the term optical flow and the usual assumptions in using it for motion analysis in an image sequence.

- Optical flow is the motion (flow) of brightness patterns (optic) in videos.
- Assumptions in optic flow:
  - Brightness consistency: Pixel has constant brightness across time
  - Small motion: Between frames, motion is small
  - Spatial coherence: Pixels move like their neighbours

## 2 Year 2017

**Question 1(c):** The features detected in the images will end up in a database used for object recognition.

1. What are the two main challenges in object recognition? Provide two examples of each challenge.
  - Viewpoint variation: Objects viewed from different angles look very different. E.g. cakes look different from the top and side.
  - Occlusion: Objects can be obscured by other things that makes it difficult to process and identify these objects. E.g. a hand covering a cup.
2. List the 4 main steps for k-means clustering used in bag-of-features.
  - Randomly initialize  $K$  cluster centers
  - Iterate until convergence:
  - Assign each data point to the nearest center
  - Recompute each cluster center as the mean of all points assigned to it

### Question 2

1.
  - In the context of image sequence processing, what is feature and describe why local features are desired for tracking?

A feature is a piece of information that is relevant for solving a computational task related to a certain application e.g. points, edges and objects for computer vision.

Local features are features in specific locations of the image e.g. mountain peakss that are distinct and unique from any angles and brightness.

- What are feature descriptors and why are they used over the original features?

Feature descriptors are interest points encoded into a series of numbers that acts as a numerical "fingerprint" to differentiate one feature from another.

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

2.
  - What is meant by the following when measuring tracking errors: TP, RN, FP, TN?
    - TP: True positive - Model correctly predicts a positive class
    - FP: False positive - Model incorrectly predicts a positive class
    - FN: False negative - Model incorrectly predicts a negative class
    - TN: True negative - Model correctly predicts a negative class
  - How are precision and recall defined?
    - Precision: Fraction of TP to sum of TP and FP
    - Recall: Fraction of TP to sum of TP and FN