ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Computer Vision
AS 2022

Prof. Marc Pollefeys *Prof. Siyu Tang* *Prof. Fisher Yu*

Exam Preparation Exam: Tracking

December, 2022

Question 1: Questions from previous years part 1

a) What are the two main problems with tracking a point with the energy term $E(h) = [I_0(x+h) - I_1(x)]^2$ (2 pts.)? Describe the scenario when each problem occurs (1 pt each). What is the characteristic of a good point to track (1 pt)? **5 pts.**

ANSWER ANSWER

1. *Zero Gradient: The image region is flat. There is no difference between a point feature and its neighbor in the next frame. E.g Aperture problem.*
2. *Local Minima: There are multiple solutions for the motion. E.g Checker board pattern.*

The point should be a corner point (or more generally a distinctive point).

ANSWER ANSWER

b) To track an object that might change its appearance over time, we could learn the appearance model online. However, this method could gradually change the target object that we are tracking. How can we prevent the model from tracking the wrong object? **2 pts.**

ANSWER ANSWER

We can anchor the tracking model with the initial appearance of the object. In other word, we incorporate the initial appearance into the model every time we update the objective patch.

ANSWER ANSWER

c) In a tracking algorithm, we sometimes use the constant-velocity heuristic to predict the potential locations of the object. By doing so, what assumption do we make about the camera? In particular, what information do we need (about the camera) for the velocity heuristic to work properly? **2 pts.**

ANSWER ANSWER

The camera needs to be stationary, or appears to be stationary relative to the object, the global locations of the camera in each time step are known, or the initial position and the velocities of the camera in each time step are known.

ANSWER ANSWER

d) Judge the correctness of the following statements and **select** the corresponding box (☑). For each statement, 1 pt, 0 pt and **-1 pt** are given for a correct answer, both empty/selected boxes, and an incorrect answer, respectively. The minimum number of points is 0. **3 pts.**

		True	**False**
1)	Histogram of Oriented Gradients can be made a rotation invariant feature descriptor	☐	☐
2)	Hungarian Algorithm can be used to match candidate regions between two frames based on similarity scores between each pair.	☐	☐
3)	To track by detection, we do not need to know the type or shape of the object to be tracked beforehand.	☐	☐

ANSWER ANSWER

	True	**False**	*Reason*
1)	☑	☐	
2)	☑	☐	*The matching can be seen as a bipartite matching problem.*
3)	☐	☑	

ANSWER ANSWER

Question 2: Questions from previous years part 2

a) When we track a point with the energy term $E(h) = [I_0(x+h) - I_1(x)]^2$, we could have a problem with local minima. Explain when the local minima problem occurs (1 pt) and how can we prevent it by changing the frame-rate (2 pts)? **3 pts.**

ANSWER ANSWER

1. *Local Minima happens when there are high frequency details that look the same in multiple direction e.g checker board.*
2. *The problem can be prevented by increasing the frame-rate to be faster than the motion Nyquist rate or more than half of the motion wavelength.*

ANSWER ANSWER

b) Explain technical pros and cons of the Lucas-Kanade template tracker. Give at least 2 examples for each category. **4 pts.**

ANSWER ANSWER

1. *Good: It can handle different parameter space (e.g . different motion or transformation models).*
2. *Good: It can converge fast in a high-frame rate video.*
3. *Bad: It is not robust to noise and large displacement (low-frame rate video).*
4. *Bad: Some transformations are not possible to parameterize, thus can not be used in the framework.*

ANSWER ANSWER

c) Explain how we can use a template to track an object that changes its appearance over time (1 pt). What are the problems that might occur if we only keep the information of the template from the previous step T_{t-1} (1 pt)? And how can we adjust the algorithm to prevent such problem? Write a simple equation in term of T_i and initial target T_0 (1pt). Explain beyond using a single keyword as an answer. **3 pts.**

ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER

1. *We can update the target template over time using the current tracked bounding box.*
2. *However, the target template could gradually change into something else that is not intended to be tracked (drifting).*
3. *This could be prevent by also using the initial template to update the current target.* $T_i = \alpha T_{i-1} + (1 - \alpha)T_0$

ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER

d) In a tracking algorithm, we sometimes use the constant-velocity heuristic to track object motion. Explain what the constant-velocity heuristic is (1 pt) and provide 2 reasons why it can help improving the tracking results (3 pts). **4 pts.**

ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER

The constant-velocity heuristic assume that all objects of interest move at a constant velocity relative to the camera. It can help with motion tracking to

1. *Predict the probable location of the object,*
2. *Reduce noise in the motion,*
3. *Disambiguate multiple objects using their trajectories.*

ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER

e) Explain the four steps in the general framework for offline multiple object tracking. Give a brief (one-sentenced) explanations for each step. **4 pts.**

ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER

1. *Object Detection: Find objects in each frame with your best object detection algorithms.*
2. *Motion Estimation: Propagate the objects from Frame T to Frame T+1. It may not depend on Frame T+1.*
3. *Initial Association: Associate objects with estimated motion and appearance features.*
4. *Association Optimization: Optimize the association with matching constraints using Hungarian matching or GNN.*

ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER ANSWER

f) Judge the correctness of the following statements and **select** the corresponding box (☑). For each statement, 1 pt, 0 pt and **-1 pt** are given for a correct answer, both empty/selected boxes, and an incorrect answer, respectively. The minimum number of points is 0. **2 pts.**

		True	False
1)	For multi-object tracking, the method's result does not depend on the feature representation.	☐	☐
2)	To do tracking by detection with a neural network, we must know the category of the object to track.	☐	☐

ANSWER ANSWER

	True	False	*Reason*
1)	☐	☑	*Bounding box size can greatly influence the results.*
2)	☑	☐	

ANSWER ANSWER