ECE 332 Introduction to Computer Vision, MP#7

The due date is 11/12/2020 (Th).

1 Template-matching based Target Tracking

The purpose of this MP is to let you have a hands-on experience on visual target tracking, to understand the fundamental components of this problem, and to be aware of the challenges that are still yet to be overcome.

You are given a 500-frame long video, which the target is the head of the girl, as shown in Figure 2(a). The objective is to keep tracking the location and size of the head of the girl, represented by a bounding box of the target), as shown in Figure 2(b).

(a)

(b)

Figure 1: (a) one frame of the input video, where the target is the head of the girl. (b) The objective is to track the head of the girl over all frames, by estimating the bounding box of the target.

- **Initialization**: you don't have to use a face detector to initialize the tracker. You can do it manually, by appointing a bounding box of the target in the first frame;
- **Image matching methods**: The key here is the image matching criteria, between two regions (of the same sizes). You need to try and compare the performance of the following three options:
  1. SSD: sum of squared difference

$$D = \sum_{u,v} [I(u,v) - T(u,v)]^2$$

where $T(u, v)$ is the template (i.e., the target region in the previous frame), and $I(u, v)$ is the matching candidate.

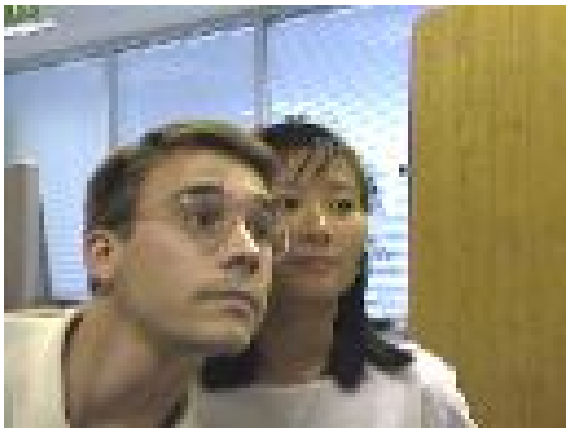
2. CC: cross-correlation
$$C = \sum_{u,v} I(u, v)T(u, v)$$

3. NCC: normalized cross-correlation
$$\hat{I}(u, v) = I(u, v) - \bar{I},\ \hat{T}(u, v) = T(u, v) - \bar{T},$$
where $\bar{I}$ and $\bar{T}$ are the average intensity of I and T.
$$N = \frac{\sum_{u,v} \hat{I}(u, v)\hat{T}(u, v)}{\sqrt{\left[\sum_{u,v} \hat{I}(u, v)^2\right]\left[\sum_{u,v} \hat{T}(u, v)^2\right]}}$$

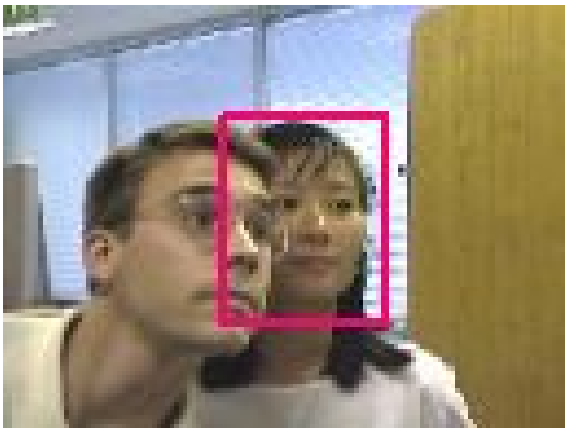
- **Search methods**: Another important issue is how to search for the best match of the target region in the next frame. In this MP, you can simply use local exhaustive search, i.e., search every location within a search window centered on the previous location.
- **Output**: you need to draw the bounding box for each image frame, and then make a video.

2 Option (with significant extra credits)

At the later part of this video, the girl is occluded by a boy. This occlusion creates missing observation of the template, which is likely to fail the template-based tracker.



(a)



(b)

Figure 2: (a) Partial occlusion is presented in this video. (b) Can you overcome this challenge?

The option of this MP is very different from other MPs. This is an open option, meaning that you can come up with any solutions, as long as it can overcome the challenge and can successfully track through the entire video!

You will get up to 30% extra credits!

3 What to turn in

Each individual student should turn in his/her own solution. What you need to turn in includes:

- your code;
- a short report ($\leq$ 1 page is fine). If you choose to do the option, you can write as long as you want;
- your results (video, NOT individual frames).