

Template-matching based Target Tracking

Algorithm:

This algorithm implements basic ideas of template matching based visual target tracking. 500 frame video is used to test various image matching metrics. It is implemented with the following steps:

Step 1) Selected a target object area that needs to be tracked from the video frame and create an initial bounding box of the target in the first frame. This will be used as a template for further processing.

Step 2) Defined a search window around the template. This is the area in which local exhaustive search is performed to find the best match.

Step 3) Within this search window, for every bounding box, calculated any one of the metrics - Sum of Squared Difference(SSD), NCC(Normalized Cross Correlation), or Cross Correlation(CC) with the template. Selected the best matching bounding box.

Step 4) Draw the best matching bounding box from step 3) on the current image frame, append it to an image_array (and also save it in a target folder).

Step 5) Repeat steps 2-4 for the next frame image until we reach the end of frames, setting the template to be the best matching image found in the previous step each time.

Step) Once all the frames are processed, save the image_array to an output video.

Results & Analysis:

The target tracking algorithm was able to successfully track the selected target and identify it in all the frames with a good accuracy. [The video results are attached with this submission.]

Overall NCC performed the best, as it goes through a normalization process and hence has reduced sensitivity to noise. SSD performed second to NCC, accurate in majority of the scenarios, while CC was slightly worse out of the three.

The bounding boxes selected based on the initial frame are point1: (56 27) point2: (89 63) [with +/- 2 for all (x,y)]

Analysis based on face tracking through some specific movements are given below.

Moving left-right-left

In this case, both SSD and NCC get very good results, tracking the face accurately. CC does slightly worse compared to the other two, but it's still able to track the face partially(at least 2/3rd of the face].

| NCC | SSD | CC |
|-----|-----|----|
| | | |

Moving Back

Next the girl can be seen moving back from the camera. In this case also, both SSD and NCC get very good results, tracking the face accurately. CC again does slightly worse, but it is tracking the face partially(at least 1/3rd).

| NCC | SSD | CC |
|-----|-----|----|
| | | |

Rotation

When the girl is rotating on the chair, NCC gives the best results, and does stable tracking of the head. SSD performs slightly worse, and loses track of the head in a few frames. CC does much worse compared to the other two and the tracking is not very stable during rotation for CC.

| NCC | SSD | CC |
|-----|-----|----|
| | | |

OPTION

When the frame is occluded, NCC still is able to do fairly stable tracking of the face, while SSD and CC are not able to track accurately.

