Jack Hanke
MSAI 495 Computer Vision
MP #7

Template-Matching Based Target Tracking Report

In this assignment we implement the template-matching based target tracking algorithm, using 500 frames of a woman moving around a room. The method is initialized with the first frame, and a given bounding box around the object to be tracked. The algorithm then proceeds to the next frame, searching around the bounding box of the last frame (some square region around the center of the bounding box) for a “similar” region of the same size. We explored multiple metrics for similarity, namely the *sum of squared difference, cross-correlation, and normalized cross-correlation* metrics. This newly found bounding box is then rendered on the frame, and the method repeats with the newly labelled frame as the old frame.

In my implementation, I chose to fix a square 41 by 41 pixel bounding box, and search for matches in a 11x11 pixel grid around the last target. I convert the image to grayscale for similarity comparisons to improve speed and accuracy. I chose to implement the naive approach, where only the last frame’s target is considered, as opposed to considering a longer history of frames. This resulted in the bounding box appearing “jumpy”, especially when using the SSD and CC metrics for similarity. My best results using this approach, as expected, was the NCC method. All three of the methods, with this naive approach, were fooled by the occlusion of the other person, and occasionally was “distracted” by the patterns in the background. I hope to further improve this algorithm in my own time by computing a match score over some past window of frames, as opposed to the single frame.

I was surprised by when the algorithm worked and didn’t work. The “jumpiness” seemed to come out of the blue, and I’m not really sure as to why these deviations arise out of the blue. I am certain that considering more frames will improve performance.