Correlation Filter Tracking

Stefanela Stevanović, 63220492

I. Introduction

The goal of this homework is to implement simplified version of MOSSE correlation filter tracker. Correlation filter tracking is a computer vision method that uses a correlation filter to locate an object in a video sequence. The filter is trained on the target object's feature representation in the first frame, and then used to locate the object in subsequent frames by comparing it with the current frame's feature representation. The performance of the tracker is evaluated on the sequences of the dataset VOT2014 based on average overlap, total failures and average speed in frames-per-second.

II. Experiments

A. Implementation of the correlation filter tracker

In the first experiment we implemented simplified version of MOSSE correlation filter tracker and tested it on VOT2014 sequences with parameters: α = 0.1, σ = 2, enlarge_factor = 1 and λ = 1. The results are shown in table I.

TABLE I: Tracking performance with parameters: α = 0.1, σ = 2, enlarge_factor = 1 and λ = 1.

Average Overlap	Total Failures	Average Speed (FPS)
0.44	72	254

The first reported metric in the table I is the average overlap, which measures how much the predicted bounding box overlaps with the ground truth bounding box. An average overlap of 1 indicates perfect tracking, while a value of 0 indicates complete failure to track the object. Average overlap of 0.44 suggests that the tracker is able to maintain reasonably good overlap with the object throughout the different sequences. Total number of failures of 72 indicated that our tracker has satisfying accuracy and average speed of 254 frames-per-seconds shows computational efficiency.

B. Tracking performance with different parameters

Update speed α is the parameter that controls the rate at which the correlation filter is updated in response to changes in the object appearance. A smaller value of alpha corresponds to a slower update rate and may make the tracker less responsive to sudden changes in object appearance. On the other hand, a high update speed means that the filter adapts more quickly to changes in the appearance of the object being tracked, but it can an make the tracker more susceptible to external factors such as occlusions or changes in lighting conditions, which can cause the filter to update inappropriately and cause the tracker to lose track of the object. The performance results with different are shown in table II.

From the table II we can see that we achieve optimal results in terms of total number of failures with α = 0.3. Average speed over all sequences, represented in frames-per-second (FPS), also varies with the change of α. In terms of computational efficiency, a lower value of α can be beneficial because it requires fewer filter updates per frame, reducing the overall computational load on the tracker. However, this can also lead to slower tracking speeds, as the filter may take longer to adapt to changes in the target's appearance. A higher value of α can result in faster tracking speeds but may require more frequent filter updates, leading to higher computational load and potentially slower overall performance.

TABLE II: Tracking performance with different parameter α and constant parameters σ = 1, enlarge_factor = 1, λ = 1.

Parameter α	Average Overlap	Total Failures	Average Speed (FPS)
0.1	0.46	84	236
0.2	0.46	85	173
0.3	0.46	74	226
0.4	0.46	89	170
0.5	0.43	78	178

In correlation filter tracking, σ refers to a parameter that controls the bandwidth of the Gaussian function used to construct the filter. A low value of sigma will result in a sharper filter response, which can improve the accuracy of the tracking by focusing more closely on the target object's features, but it can also make tracker less robust to changes in appearance or environmental conditions. A high value of sigma will result in a smoother filter response, which can help improve robustness, but may also result in a loss of detail, making the tracker less precise. Tracking performance with different σ values is shown in table III.

TABLE III: Tracking performance with different parameter σ and constant parameters α = 0.1, enlarge_factor = 1, λ = 1.

Parameter σ	Average Overlap	Total Failures	Average Speed (FPS)
1	0.46	84	203
2	0.44	72	157
3	0.44	76	185
4	0.47	86	160
5	0.45	88	156

From the table III we can see that we achieve optimal performance with σ = 2, with the lowest number of total failures of 72. By further increasing the σ values, we get increase in the total number of failures, indicating to poorer performance of the tracker due to loss of detail.

In correlation filter tracking, the enlarge factor is a parameter that controls the size of the search window used to track the target object. Specifically, the search window is expanded by a factor of enlarge_factor around the location of the target object in the previous frame to ensure that the target remains within the search region even if it moves. Table IV shows tracking performance with different enlarge_factors.

TABLE IV: Tracking performance with different parameter enlarge_factor and constant parameters $\alpha = 0.1$, $\sigma = 1$, $\lambda = 1$.

enlarge_factor	Average Overlap	Total Failures	Average Speed (FPS)
1	0.46	84	181
1.5	0.47	91	89
2	0.46	105	57
3	0.44	160	37

From the results in the table IV we can see that the average overlap decreases and total number of failures increases with larger enlarge_factors. With larger search window, more background is captured, so this may cause tracker to become susceptible to tracking false positives or other objects in the scene that have similar visual features to the target. Also, from the table IV we can see that increase in enlarge_factors increases the computational load on the tracker and significantly decrease tracking speed.

C. Tracking speed over different sequences

In this experiment, we measured the average tracking speed in frames-per-second on different sequences of the VOT2014 dataset. The results are presented in the table V.

TABLE V: Average tracking speed in frames-per-second for different sequences.

Sequence Name	Tracking Speed (FPS)
ball	511
basketball	144
bicycle	576
bolt	132
car	708
david	69
diving	80
drunk	90
fernando	33
fish1	351
fish2	178
gymnastics	184
hand1	338
hand2	499
jogging	110
motocross	65
polarbear	132
skating	104
sphere	420
sunshade	356
surfing	394
torus	371
trellis	431
tunnel	144
woman	231

From the table V we can see that the tracking speed varies greatly between the different sequences, ranging from 33 FPS to 576 FPS. These differences in tracking speed come from the different complexity of the tracking task in each sequence. Some of the factors that can impact the tracking speed are: the size of the object being tracked, the amount of movement and changes in appearance of the object, the resolution of the frames, etc.