

# Mean Shift Tracking Algorithm

Aljaž Konec

## I. INTRODUCTION

In this report we present the implementation of the Mean Shift tracking algorithm. The algorithm uses a kernel density estimation to find the mode of a probability density function PDF. In the context of tracking, the PDF represents the color distribution of the object to be tracked. The algorithm iteratively shifts the kernel towards the mode of the PDF until convergence. For the kernel density we use the Epachenik kernel for its simplicity and computational efficiency. The report is split into two parts. In the first part we describe and test mean shift mode seeking and in the second part we apply mode seeking for tracking and showcase the results on the VOT 2014 dataset.

## II. MEAN SHIFT MODE SEEKING

Figure 1 shows the results of mean shift mode seeking on a synthetic image. We test the algorithm with four different kernel sizes and multiple starting points. For kernel size 15 it can be seen that some starting positions do not converge to the global maximum. Increasing the kernel size to more than 31 completely negates this local convergence as the kernel is large enough to include the higher values of the global maximum. To improve convergence speed and accuracy we

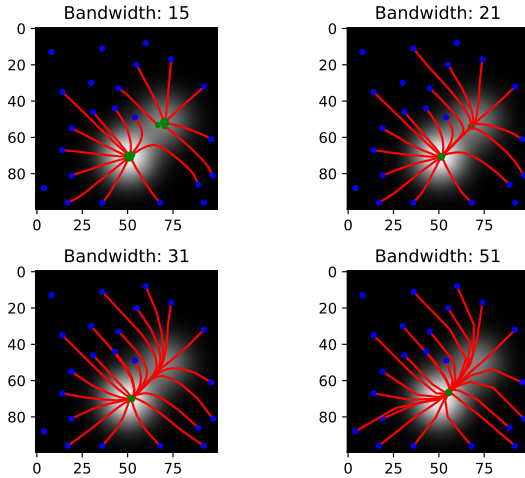


Figure 1: Mean shift mode seeking.

start decreasing the kernel size by 2 as soon as the length of the shift vector is smaller than 0.4 pixels. We completely terminate the algorithm when the shift is smaller than 0.1 pixels. Using these convergence criteria we get the following convergence speeds:

- Kernel size 15: 17.0 iterations
- Kernel size 21: 17.5 iterations
- Kernel size 31: 13.7 iterations
- Kernel size 51: 12.3 iterations

### A. Mean Shift Mode seeking on custom sythetic image

Figure 2 shows the results of mean shift mode seeking on a custom synthetic image to show the robustness of the algorithm. This case showcases how to small of an initial kernel size

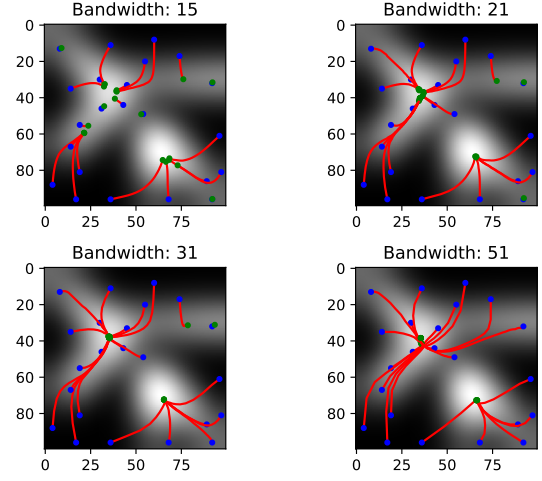


Figure 2: Mean shift mode seeking on custom synthetic image.

can lead to convergence in a local maximum. For most of the starting positions, a kernel size of 15 is too small to converge to the global maximum. Kernels with size 31 and 51 cover the appropriately sized area and converge to acceptable maximums. The paths from starting positions to the maximums are also longer because of the longer ridges which have to be traversed. For this reason the convergence speed is the following:

- Kernel size 15: 23.3 iterations
- Kernel size 21: 28.4 iterations
- Kernel size 31: 23 iterations
- Kernel size 51: 25 iterations

## III. MEAN SHIFT TRACKING

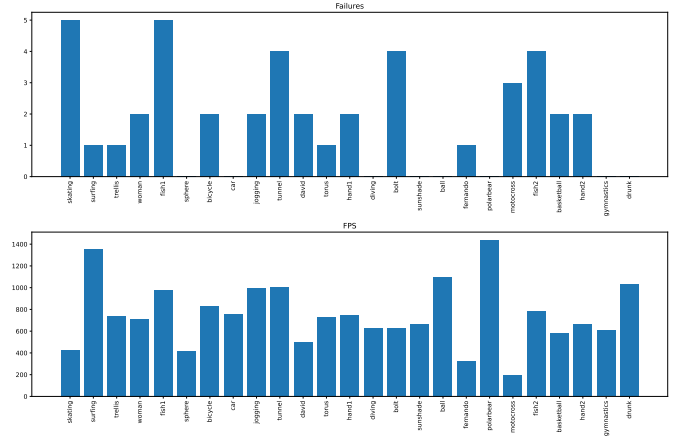


Figure 3: Mean shift tracking on VOT 2014 dataset.

We use the mean shift mode seeking algorithm to track an object in a video sequence. The object is represented as histogram of the color distribution in the object region. To find the object in the next frame we use the mean shift algorithm on a weighted quotient of target and candidate histograms.

When computing the quotient we add a small constant number  $\epsilon$  to the denominator to avoid division by zero. We repeat this process until the shift vector is smaller than 0.2 pixels or the number of iterations exceeds 50. Figure 3 shows the results of the mean shift tracking algorithm on the VOT 2014 dataset. Using the best performing parameters that we discuss in the following subsections we see that the tracker fails at most 5 times in a single video. Combined, the tracker fails 42 times and in average achieves a rate of 750 FPS, if we do not show the frames.

#### Failure cases

In this section we discuss failure cases of the mean shift tracking algorithm. We focus on the following video sequences: bolt, skating, fish1 and tunnel. Figure 4 shows the failure cases of the mean shift tracking algorithm.

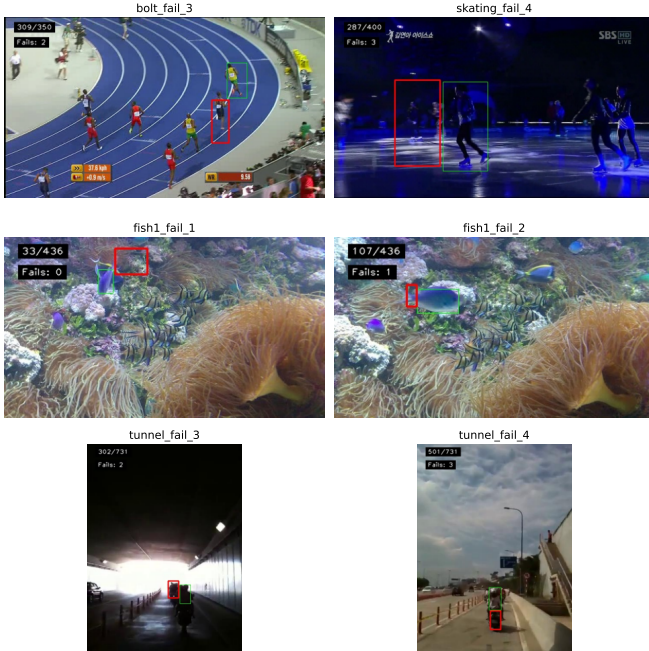


Figure 4: Failure cases of the mean shift tracking algorithm.

For the bolt sequence, the tracker fails because the object (Bolt) changes direction, moves fast and changes shape. Increasing the kernel size could help with the fast moving object, but the shape change is a limitation of the mean shift algorithm. In the skating sequence the lighting conditions change and the skater becomes darker then before, which leads to the tracker failing. For the fish1 sequence the fish drastically changes shape and also moves in a fast jerking way. To improve tracking in this case a larger update rate  $\alpha$  would help with the fast jerking movements. In the tunnel sequence the light changes significantly. The target object in the left image is completely black and also changes lane to the right causing the tracker to fail. After some time the tracker adapts to the darker object color, when the moped comes out of the tunnel and into the light the tracker still tracks dark regions and converges to the (darker) bottom of the moped. To improve tracking in this case using a different colorspace (like HSV) might help.

#### Parameter tuning

As mentioned before we used optimal parameters for the mean shift tracking algorithm. To find these parameters we

performed a grid search over the following parameters:

- Kernel size: 11, 15, 21, 31, 51
- Update rate  $\alpha$ : 0.01, 0.02, 0.05, 0.1
- $\epsilon$ : 1e-6, 1e-7, 1e-8, 1e-9, 1e-10

In table I we show the 10 best performing parameter tuples. We can observe that all the results use an update rate of 0.01, which is sensible with the notion that the target object does not change significantly between frames. All the tested kernel sizes are present in the best performing parameters. This can be explained by the fact that the target moves by a small amount between frames and thus the mean shift starts near the maximum of the PDF. Because of this, kernel sizes 11 and 15 are sufficient to converge to the target. A similar observation can be made for the  $\epsilon$  parameter, where the best performing values are in the range of 1e-7 to 1e-10. For other tested parameter values the tracker quality is mostly dependant on the update rate  $\alpha$ . As we increase the update rate, the number of errors increases significantly, this would be expected as the tracker would be more sensitive to changes in the target object.

Kernel Size	Update Rate	Epsilon	Errors	FPS
11	0.01	1e-7	42	756
11	0.01	1e-9	42	747
11	0.01	1e-6	42	731
11	0.01	1e-10	42	730
11	0.01	1e-8	42	722
51	0.01	1e-7	43	774
15	0.01	1e-7	43	751
31	0.01	1e-7	43	748
31	0.01	1e-9	43	745
21	0.01	1e-6	43	743

Table I: Best performing parameters for the mean shift tracking algorithm.

## IV. CONCLUSION

The limitations of the Mean Shift tracking algorithm can be seen in the failure cases. The algorithm is not robust to fast moving objects, changes in lighting conditions and changes in object shape. To improve the tracking a method to adapt size changes and adapt for light changes would need to be added.