1 Condensation Tracker

1.1 Color Histogram

We pass over the frame and check whether we are within the bounding box, if this is the case we add one to the histogram matrix at the position x, y and z which correspond to the RGB values of the pixel we are currently checking. Once we have passed over the entire frame we normalize the histogram matrix by dividing by the number of elements we placed inside.

1.2 Derivation of matrix A

We want Ap to give us the new position after having added noise where p is the particle vector, which is why A is defined as follows:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

When we have a model with constant motion, A has the following form:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

because we want to add the displacement coming from the velocity to the position we need the ones in the off diagonal. To be able to use these matrices we need to convert the particle vector to homogeneous form.

1.3 Propagate

We first check whether we are dealing with a model in motion or without motion. If we do not have any motion we iterate over all particles and set their positions to new values r_x and r_y which were drawn using the integrated Matlab function normrnd with parameters params.sigma_position and the particles old position as the mean. In case the model is in motion we additionally call normrnd using params.sigma_velocity and the particles old velocity as the mean to generate new velocities. Finally, we check if the new particles are still inside the frame, if not we clamp them to the boundary.

1.4 Observe

For the observation we iterate over all particles, and call the color histogram function using the current particle as center of the bounding box. The resulting histogram is then compared to the target histogram using the χ^2 cost function, the resulting distance is used to compute the particle weight by computing the equation (6) given on the exercise sheet. Finally, the particle weights are normalized before being returned.

1.5 Estimate

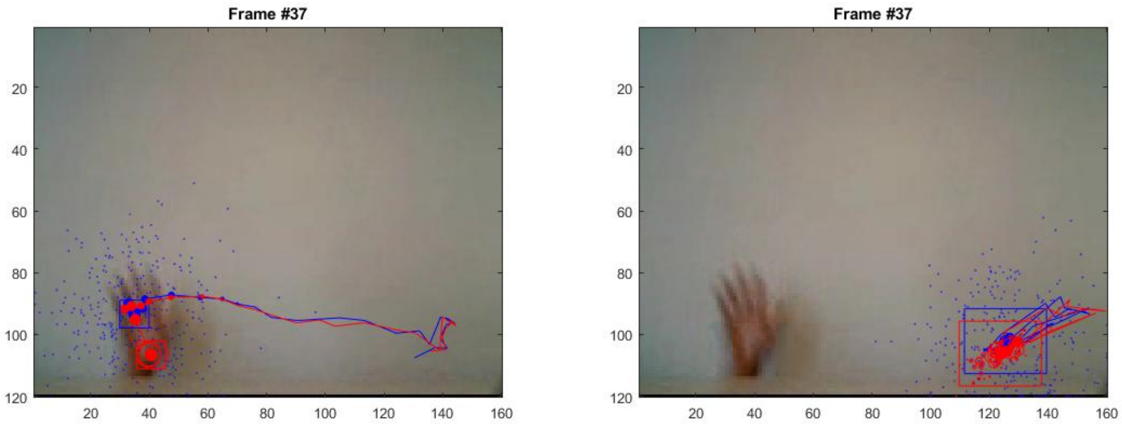
We compute the mean state which is given as the expected value of the particles and their weights, this can be done efficiently by computing the dot product of the particles and the particle weight vector.

1.6 Resample

To resample the particles and particle weights we iterate over all particles and use the integrated function datasample, this function samples particles according to their weights and returns the indices of the resampled particles and weights. The particle weights are normalized before being returned together with the new particles.

1.7 Experiments

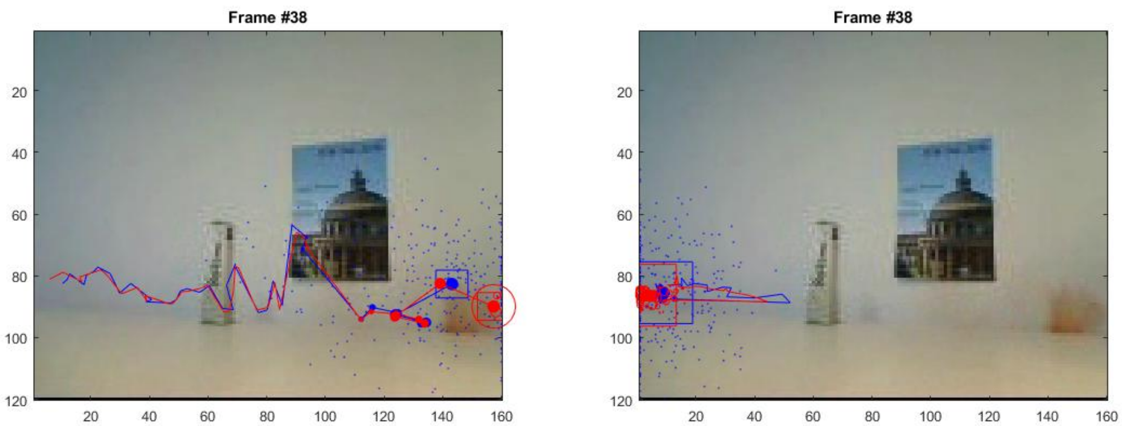
The experiment with using different amounts of histogram bins showed that using substantially more or less bins often results in tracking loss, especially in the second video after the hand has past the poster in the background. This is because we are either considering too much or too little color variation. For the first video we only reliably lose tracking if we choose a bad starting bounding box, altering the parameters did not guarantee failure of tracking for a very tight bounding box. The following images display a successful and unsuccessful tracking in the first video.



(a) successful tracking (b) unsuccessful tracking

Figure 1: successful and unsuccessful tracking of hand in video1.

For the second video we used the motion model since we have a constant velocity. We set the initial velocity to $v_x = 1$ and $v_y = -0.5$ since the hand moves to the right and slightly upwards. The tracking in this video is susceptible to the background poster as it gives a very different color histogram than the rest of the background. By increasing the observation noise we were able to achieve slightly smoother results with less jumps. Figure 2 shows a noisy tracking with the default observation noise 0.1, using 0.5 as our observation noise we were able to achieve a much smoother tracking displayed in Figure 3.



(a) successful tracking (b) unsuccessful tracking

Figure 2: successful and unsuccessful tracking of hand in video2.

- What is the effect of using a constant velocity motion model?
  By assuming constant motion we can move our a priori estimate into the direction of the motion, without increasing the uncertainty of the objects position. This is especially useful when an object becomes occluded since we have a guess on how and to where the object is moving.

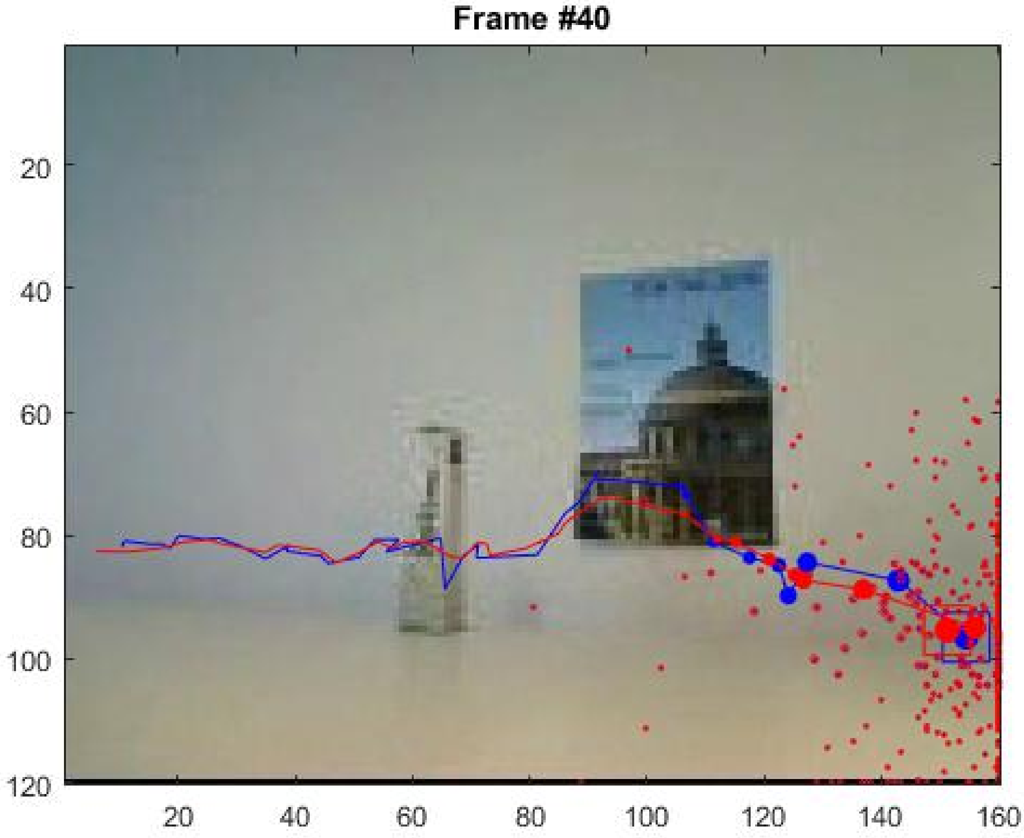


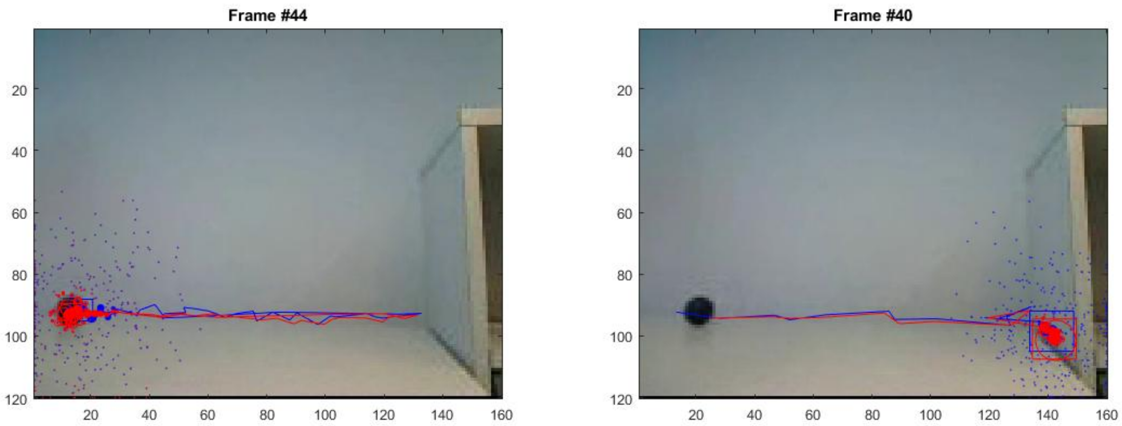
Figure 3: successful and smooth tracking using sigma_observe = 0.5.

- What is the effect of assuming decreased/increased system noise?
  By increasing the system noise we spread our a priori estimate more, this results in a more stable and smooth tracking compared to a decreased system noise. If the system noise is decreased too much we will most likely lose the object when it moves swiftly from one frame to the next since we cannot probe the location it has moved to as the a priori particles are all clustered on one spot.
- What is the effect of assuming decreased/increased measurement noise?
  As we have seen in the second video, increasing the measurement noise increases the smoothness of the tracking since we are more likely to find the same location of the corresponding object. A low measurement noise is acceptable if we trust the measurements as for example in the first video, the measurement in the second video however is very rough when using a low measurment noise as can be seen in Figure 2.

For the third image we switched back to the model without motion as the motion of the ball is not constant. Since the results were more smooth and robust with an observation noise of 0.5 we keep this value as well. The tracking of the ball is very strongly dependent on the initial bounding box size, as long as the ball covers the bounding box the tracking is very accurate. If we choose a bounding box which exceeds the balls size we either fail to track the ball from the start or we fail to track the ball after colliding with the wall.

- What is the effect of using more or fewer particles?
  Apart from the computational cost using fewer or more particles has an influence on how the tracker behaves. Using more particles gives us a better spread over a greater area, meaning we can cover more ground and are less likely to lose track of the object, using less particles causes the opposite.

- What is the effect of using more or fewer bins in the histogram color model?
  The more bins we use the more number of distinct colors we will consider, this is a double edged knife as we can on the one hand better track a very distinct object but on the other hand are also more susceptible to color noise from the background. Similar problems apply when using less bins, by only having a few bins we might get a similar color histogram for many frames as using few bins can be seen as a kind of average of the colors in the bounding box, it is easy to imagine that a too strong average can destroy the recognizability of the object.
- What is the advantage/disadvantage of allowing appearance model updating?
  The advantage is that we can track an object which is changing in appearance, for example a face turning sideways or by having our object move over a different background since the histogram allows for these slight changes. The disadvantage is for the same reason, as we have seen in class a slow moving hand over a face tracker can trick the algorithm into believing the histogram of the hand represents the face, which causes the tracker to now follow the hand.



(a) successful tracking (b) unsuccessful tracking

Figure 4: successful and unsuccessful tracking of ball in video3.

1.8 Own Video

We recorded a stationary plush piggy with a handheld camera moving around it. The initial frame can be seen in Figure 5. Initially we used the default parameters provided in the exercise and tried to track the right eye and forehead of the toy. Similar to the second video we noticed a very rough tracking when using a low observation noise, increasing it to 0.5 resulted in a much smoother tracking, the differences can be seen in Figure 6. Both trackers never lost the object but occasionally moved to the neck and back area, this was to be expected as the histogram for all these areas is very similar. We experimented with different histogram bin sizes, when using 32 instead of 16 the tracker moved to the lower back region and converged on the neck area, when using a very low number of bins the tracker lost the object within the first 10 frames. We left the alpha parameter at zero since we do not want to update the histogram. As a measure of performance we say a tracker is optimal if it was able to track the head region for most of the time, this can be seen in Figure 8 as the tracker only moved to the back for around 5 frames before convering back to the head region. Next we tried to track the object using a bounding box encasing the entire toy, the initial setup and final trace can be seen in Figure 7. The tracking was still fairly accurate but there was a strong performance impact as the bounding box now contained roughly 1600 pixels compared to the previous 100 pixels. We did not use the motion model as the motion is not constant.

One improvement would be to allow for not only rectangular but more flexible bounding boxes, as can be seen in Figure 7 (b) the rectangular grid defined in the beginning is not a good representation any more of the space the body occupies. The other improvement would be to use a more optimized variant of the color histogram which is currently the bottleneck in performance when scaling up the bounding box.

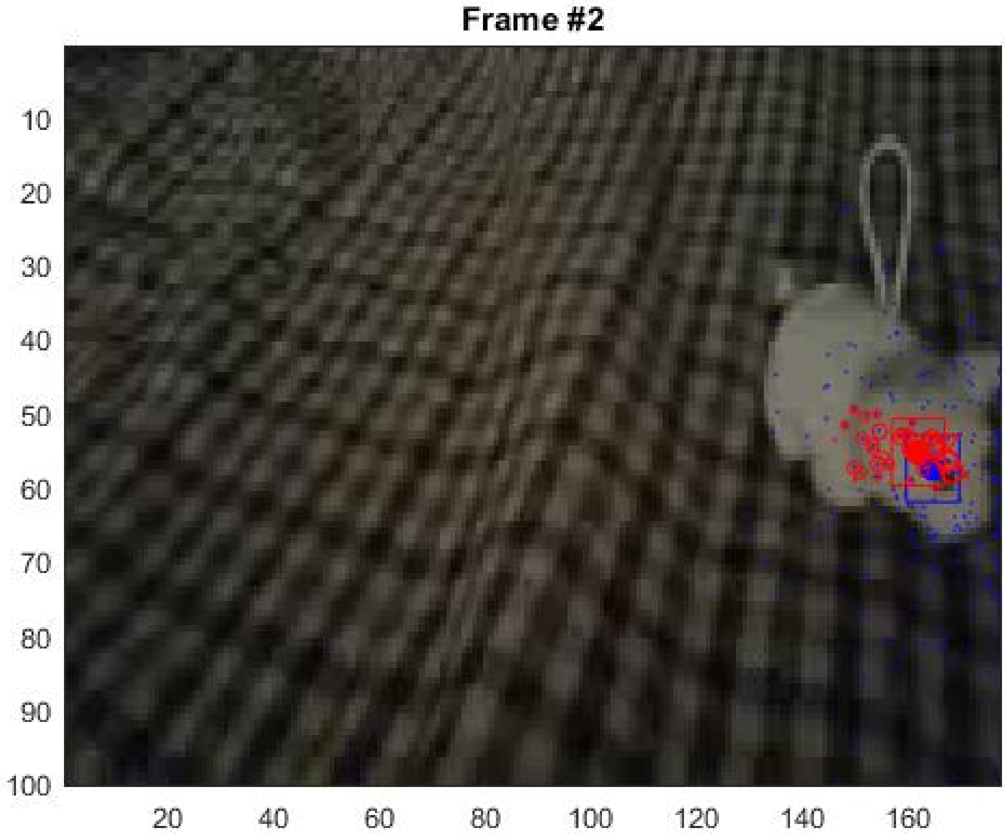
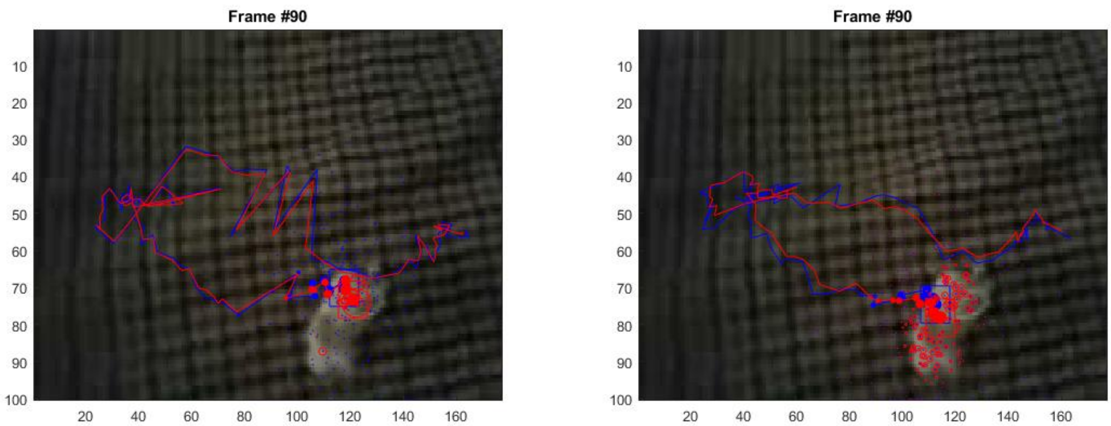


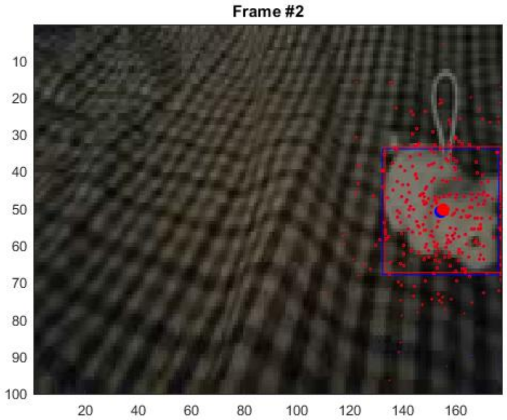
Figure 5: initial bounding box for forehead tracking using 0.1 observation noise.



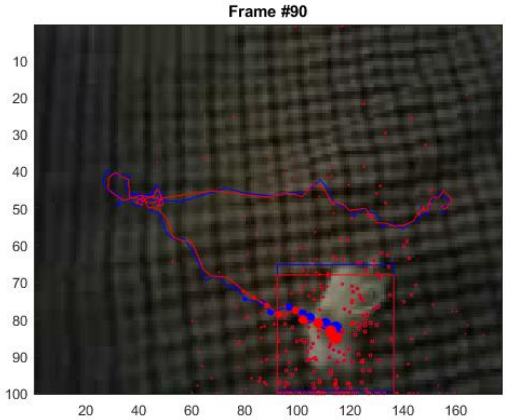
(a) successful but rough tracking

(b) successful smooth tracking

Figure 6: tracking of plush piggy with 0.1 and 0.5 observation noise.



(a) initial bounding box for entire body tracking using 0.5 observation noise.



(b) successful tracking using large bounding box

Figure 7: tracking of plush piggy whole-body bounding box.

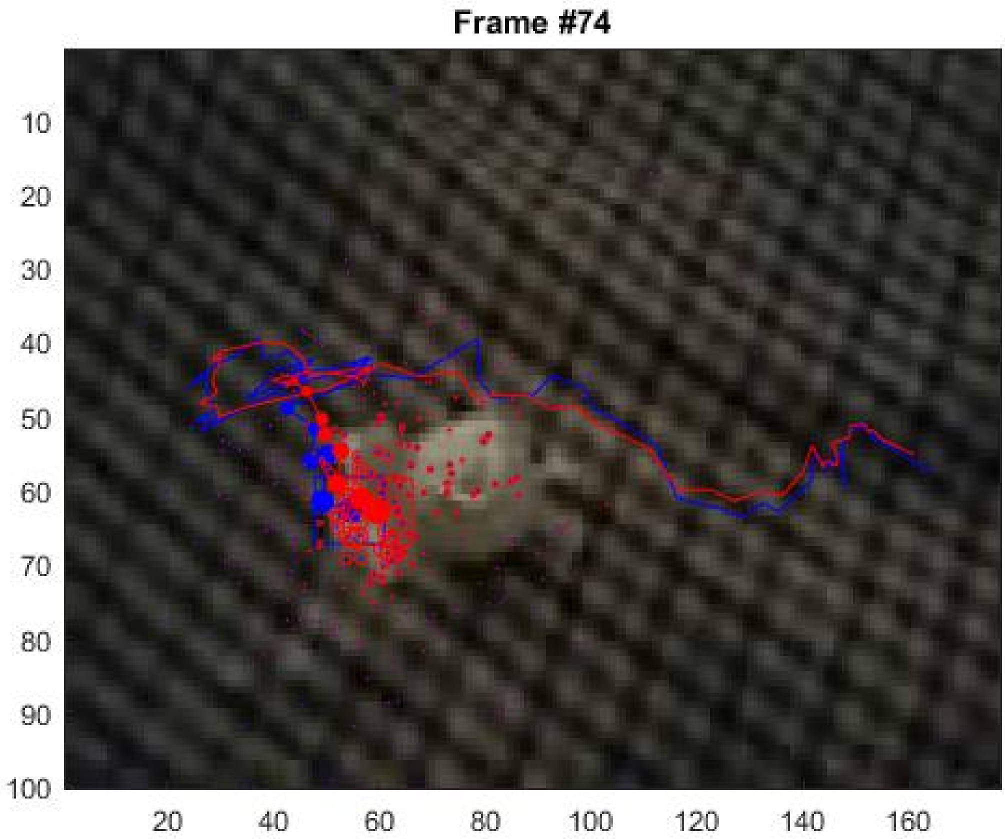


Figure 8: smooth tracking on head region.