Report: Condensation Tracker

Rishabh Singh: 19-953-793

December 20, 2019

1 Implementation

This exercise deals with the objective of tracking an object using condensation tracker based on color histograms. We are considering the modelling of motion of object under two scenarios, one with a constant velocity and other with a no motion. The initial bounding box is drawn by the user across the object.

1.1 Color Histogram

We begin with a check on the coordinates of the bounding box whether they are within the bounds of the image dimensions. In cases when this doesn't happen, we clip the values to 1 or the width/height of the image accordingly. We further extract the region inside the refined bounding box and split it into its component channels. Histogram of each channel is computed individually based on the given number of bins. The final histogram is a concatenation of the above individual ones. We normalize this histogram to account for numerical consistencies.

1.2 Derive matrix A

The matrix A captures the transition from state at time $t-1$ to state at time t.

$$s_{t+dt} = As_t + w_t$$

Ignoring the stochastic part, the motion equations will be:

$$\begin{aligned} x_1 - x_0 &= v_x.dt \\ y_1 - y_0 &= v_y.dt \end{aligned}$$

where x, y are coordinates at two different time intervals separated by dt. Assuming $s_t = [x_0, y_0, v_x, v_y]$,

$$s_{t+dt} = \begin{bmatrix} x_1 \\ y_1 \\ v_x \\ v_y \end{bmatrix} = \begin{bmatrix} x_0 + v_x.dt \\ y_0 + v_y.dt \\ v_x \\ v_y \end{bmatrix} = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x_0 \\ y_0 \\ v_x \\ v_y \end{bmatrix} = A * s_t$$

In case of constant velocity model, we set $dt = 1$ in this experiment and hence, A is given by:

$$A_{constant-velocity} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In case of no motion model, $v_x = 0$ and $v_y = 0$ and thus A reduces to the first $2X2$ submatrix:

$$A_{no-motion} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

1.3 Propagation

This function implements the first equation discussed in the previous section. Matrix A is defined according to the type of model chosen. The stochastic element w is assumed to be normally distributed with zero mean and fixed variance. This variance term is used according to the model chosen i.e. $\sigma_{position}$ and/or $\sigma_{velocity}$. After updating the state, a check is employed to ensure that the new coordinates are within the bounds of the image dimensions.

1.4 Observation

This is the measurement model where we observe and update the weights of the particles. The weights are computed based on the chi-square distance between the current and target histogram, as stated in the exercise description. These weights are normalised before returning.

1.5 Estimation

The mean state of the particles is estimated as a weighted summation of individual particles at the current instance of time.

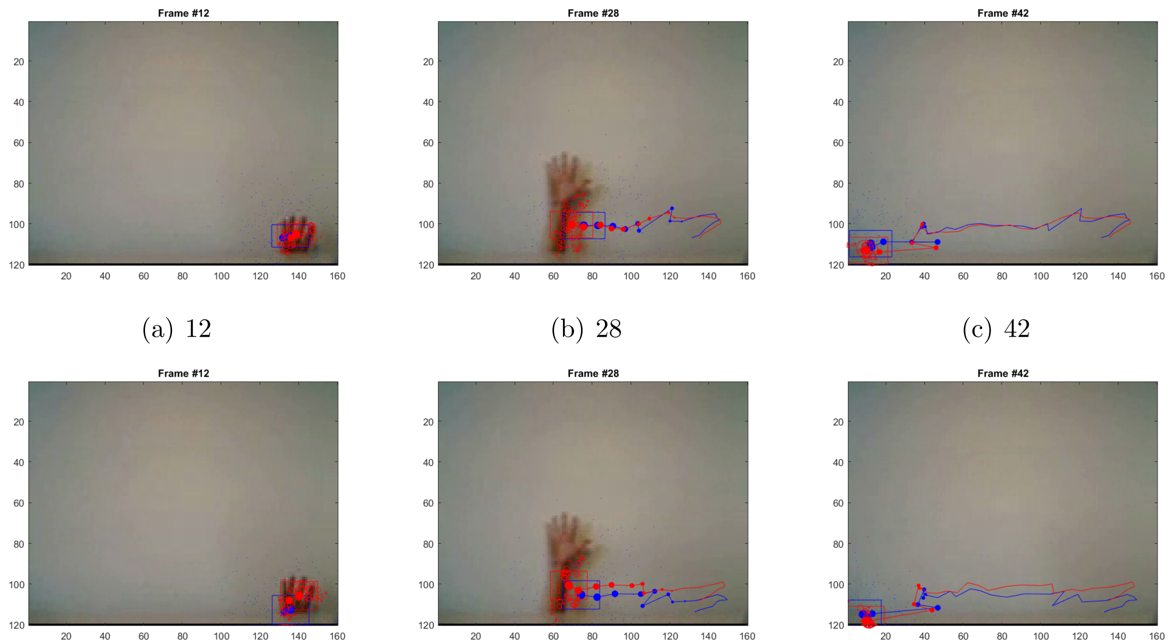
1.6 Resampling

The procedure for resampling the particles followed here is exactly as mentioned in the description of Lab-3.

2 Experiments

2.1 Video1

We have to track a moving hand in the first video. The following pictures illustrate the obtained results:



(a) 12 (b) 28 (c) 42

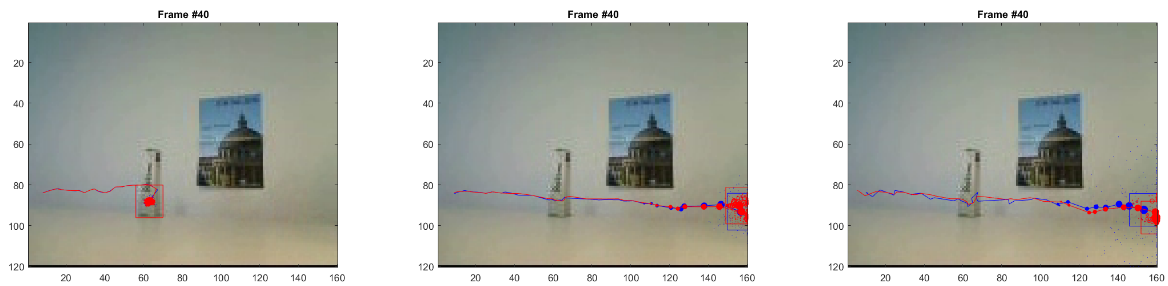
(d) 12 (e) 28 (f) 42

Figure 1: Top row shows no motion model while bottom rows show constant velocity model. The results are sufficiently good in both the cases for different frames. This may be due to uniformity in the background

The blue points are the apriori particles while the red points are the aposteriori particles. The red and blue lines correspond to the mean states respectively. The tracking results are good for the standard parameter values provided as template. Also, tracking trajectory is good for no motion model. However, note that the initial bounding box was set to cover fingers, while the objects tracked the wrist and the forearm region. One obvious reasoning which can be drawn from this observation is the presence of shadows and contrast.When the bounding box was drawn across the fingers, they were in high contrast region which later got changed as the hand started moving in the illumination regions. This was corrected by using suitable value of the α as discussed later in the dedicated section. The no motion model works pretty well here. We think it is because the hand moves very slow and smooth, so our limited noise can cover its movement area. The first attempt was successful with the parameters provided in the template, so we don't have an unsuccessful track to show.

2.2 Video2

This is a slightly more challenging problem as there is one occlusion and a background clutter present in the video frames. We first analyse the no motion model and vary the system noise.



(a) 1 (b) 5 (c) 15

Figure 2: Tracking results for no motion model under different values of $\sigma_{position}$. When the system noise is decreased, the no motion model tends to lose track and particles get stuck near occlusion.

2.2.1 Effect of system noise on no motion model

As can be seen in Fig.2, even though the hand is occluded by an object, the no motion model tracks the hand efficiently. This happens for higher system noise, here the variance in position only. This might be the case that when the hand is occluded, since the stochastic term is quite big, the particles are spread on a large surface. Therefore, when the hand reappears after the occlusion, some particles will still be on the hand and this will allow the tracking to continue. However, on decreasing the system noise to values around 1, the particles are much closer to each other and get stuck on to the occlusion as soon as the hand disappears. This problem was solved using constant velocity model as discussed below.

2.2.2 Effect of constant velocity model

In order to address the above issues arising with occlusion, the constant velocity model was used. The initial velocity was changed from $[1, 10]$ to $[5, 0]$ as there was no motion in the vertical direction, hence the y-component of velocity was kept zero. This change brought performance improvement as seen in Fig.3. The results in Fig.3 are obtained for $\sigma_{position} = 1$, which shows that the constant velocity model works even in low system noise where no motion model fails. Thus, constant velocity model is robust to occlusion under low noise in position.

2.2.3 Effect of system noise on constant velocity model

As seen above, the constant velocity model performs well in low range variance values of position. Fig. 4 shows the tracking for $\sigma_{position} = 15$. Also, changing values of $\sigma_{velocity}$ doesn't

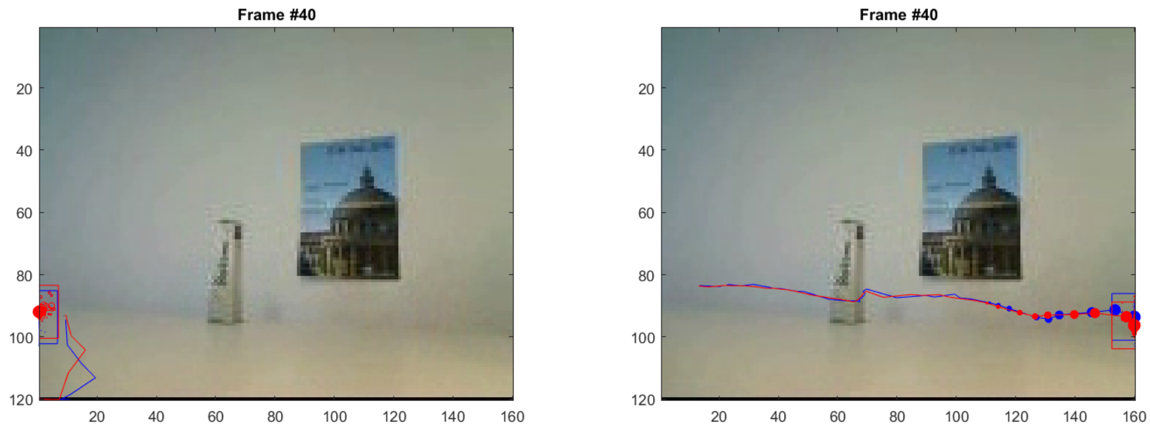
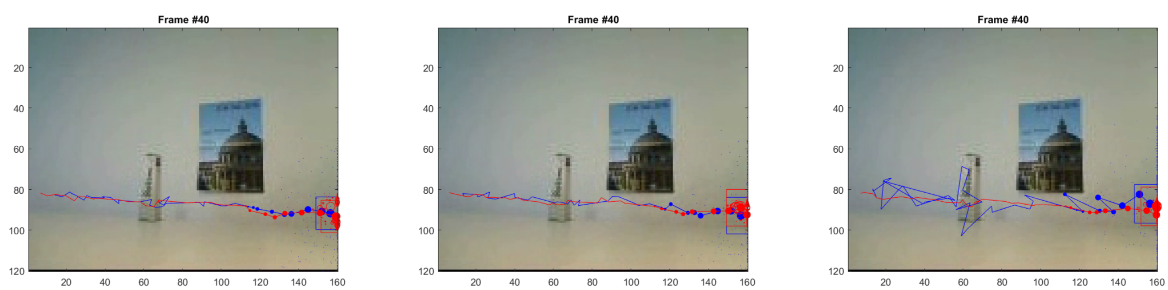


Figure 3: Tracking results for constant velocity model with $\sigma_{position} = 1$. The results are for initial velocity values $[1, 10]$(left) and $[5, 0]$ (right). The effect of occlusion under low system noise has been addressed by this model in the right image

change the tracking performance, although the mean state of apriori particles become more random for higher values. Thus, changes in system noise doesn't affect constant velocity model.

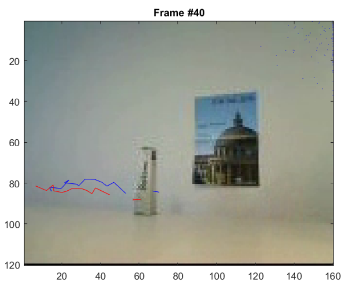


(a) 0.1 (b) 1 (c) 10

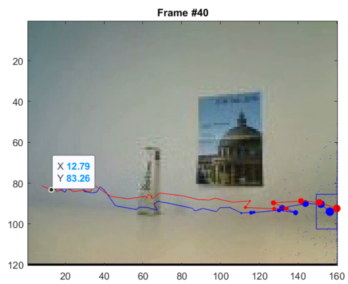
Figure 4: Tracking results for constant velocity model for different values of $\sigma_{velocity}$. These are obtained for $\sigma_{position} = 15$.

2.2.4 Effect of measurement noise

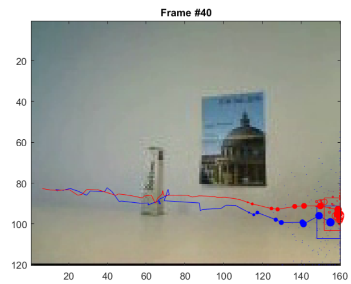
We can see the effect of measurement noise on constant velocity model in Fig.5. When we change the standard deviation of the measurement noise we decide how wide spread the weights are. For smaller values of $\sigma_{observe}$, the weight distribution is very compact. Thus, small changes can cause failure in tracking, such as the case (a) when the tracking fails on encountering the occlusion. Thus, the particles sampled from the candidates are poorly representing the target. The smaller the standard deviation the stricter is the tracker. If we go on the higher end of values, the weights are widely spread so importance is given even to particles that shouldn't be relevant. That's why the presence of background poster is able to distract the particles and they tend to track the poster instead of the hand when it disappears near occlusion. As we increase further the measurement noise (f), this phenomena is amplified: the particles stay stuck on the uniform wall and totally lose track of the hand. Similar results were observed for no motion model in Fig 6. Higher variance in measurement model tend to track the background clutter.



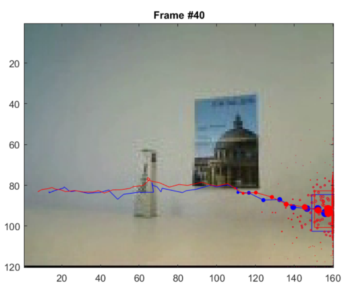
(a) 0.005



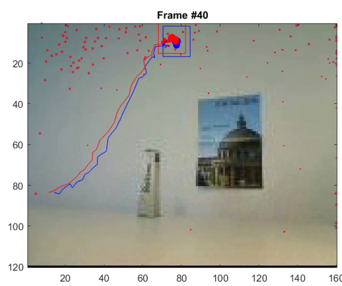
(b) 0.01



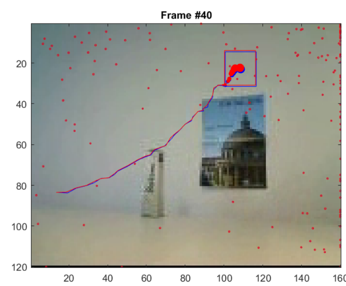
(c) 0.1



(d) 0.5



(e) 1

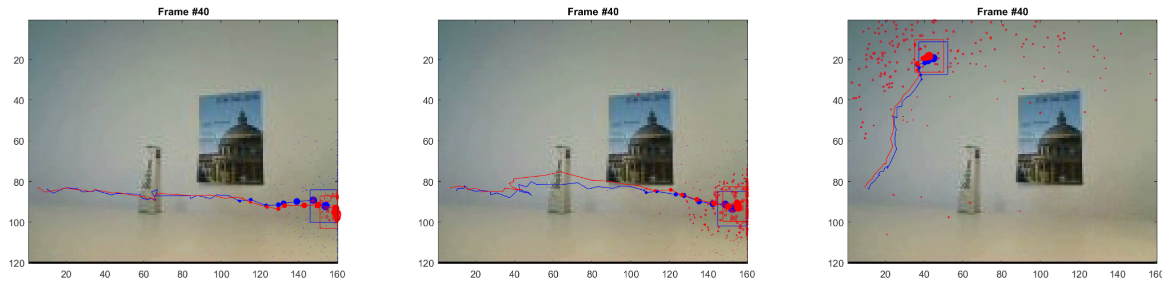


(f) 2

Figure 5: Tracking results for constant velocity model for different values of $\sigma_{observe}$. These are obtained for $\sigma_{position} = 15$, $\sigma_{velocity} = 1$ and initial velocity [5,0].

3 Video3

This was a challenging video given the frequent changes in velocity of the ball. The direction of motion is the key point to be considered here as unlike the other two videos, the motion is not constrained in a single direction. The parameters obtained from the observations for video 2 were used to generate the results as shown in Fig.7. Evidently, the moment ball bounces the particles keep being updated to the right and they never find the ball again.
One solution we thought was to increase the system noise. We first want to increase the noise in position of the particles with the hope that the stochasticity in location can help the particles to track the ball as shown in Fig.8. The tracking performs fairly for higher noise in position.
The above results were obtained when the variance in velocity is low i.e. 1. The results persuaded us to explore the dual scenario- low variance in position but changing noise in



(a) 0.1 (b) 0.5 (c) 1

Figure 6: Tracking results for no motion model for different values of $\sigma_{observe}$.

velocity. The results are shown in Fig. 9.
The no motion model apparently works nicely in this case as shown in Fig.10. One possible explanation is that since this models doesn't take into account the velocity component, particles are not affected by the change in velocity of the ball.
As with the case of measurement noise, higher values tend to focus more on the background clutter i.e. the furniture in case of both the models. This behavior is similar to what we observed in the case of video 2. Lower values for measurement noise tend to work better.

3.1 Effect of other parameters

3.1.1 Number of particles

For large number of particles, the tracking is more stable and accurate, but computationally expensive. For lower values, the trajectory of the mean state is very shaky, which reflects the uncertainity in the prediction of the aposterior particles. The results thus obtained are highly unreliable as shown in Fig.12.

3.1.2 Number of histogram bins

The results for larger number of bins are nicely behaving. The precision of the histogram increases for higher bins. This may come as a disadvantage too when two visually similar entities can be put in 'far-away' bins. Smaller number of bins has an averaging effect, thus the histogram is robust to noise in images to be compared. But, the performance of tracking tends to decrease, especially in the case of occlusion and background clutter. As seen in Fig 13(a), when the hand comes near the poster, the particles get stuck to the background. This may be because the histograms might be too similar.

3.1.3 Appearance model updating

We consider the first video here since it has changes in the illumination of the hand across frames. The initial frames had fingers in the higher contrast regions while later frames had forearm area in the contrast region while the fingers were in the illuminated region. As seen in

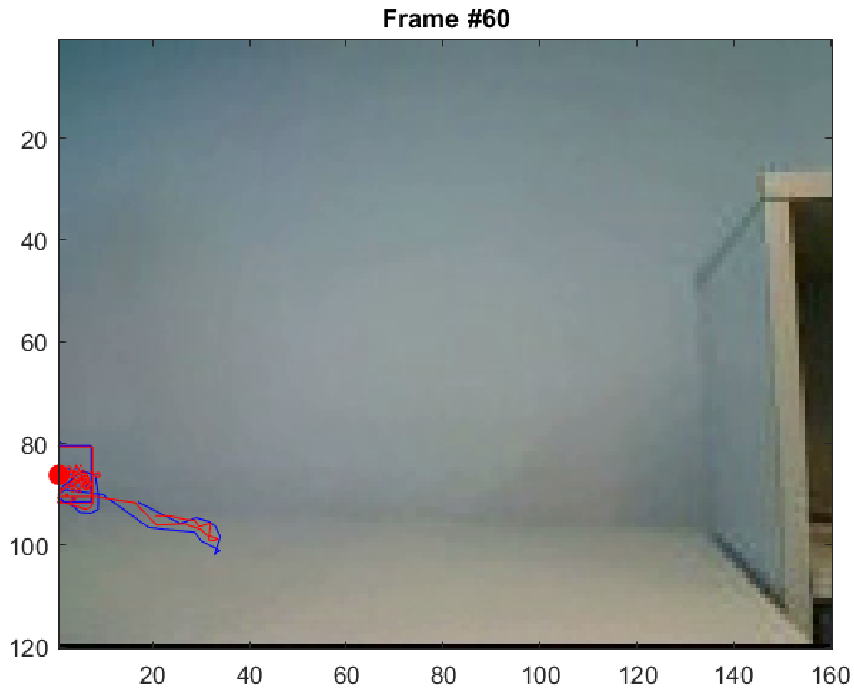
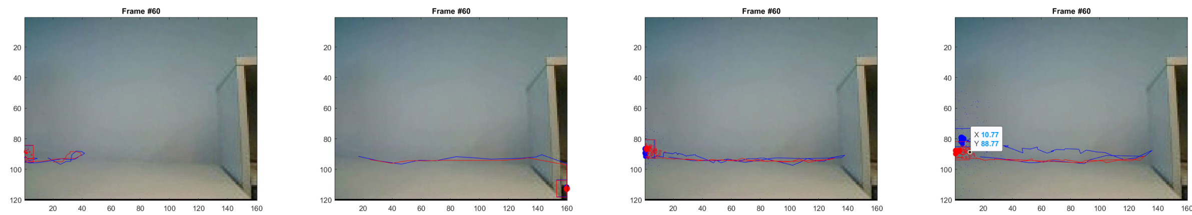


Figure 7: Tracking results for constant velocity model for video3 with parameters for video2 $\sigma_{velocity} = 1$, $initial_{velocity} = [5, 0]$, $\sigma_{position} = 1$, $\sigma_{observe} = 0.1$

Fig.14a, the model tracks the forearm region. But was the value to α is increased, the region for which bounding box was defined initially gets tracked successfully. One explanation is that a non-zero value to α allows for the histogram to adapt to the color or illumination changes across frames.

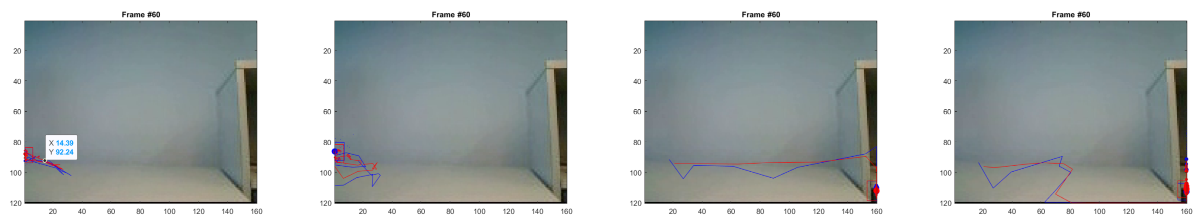
4 Discussion

- By increasing the system noise we spread our apriori estimate more. The tracking thus, is more smooth and stable.
- A suitable value of system noise helps avoiding occlusion. Lower values fail to track the object as the aposteriori particles tend to cluster around occlusion.
- A suitable value of measurement noise helps avoiding background clutter.
- By assuming constant motion we can move our apriori 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, especially in the cases when system noise is low.



(a) 1 (b) 3 (c) 7 (d) 15

Figure 8: Tracking results for constant velocity model for different values of $\sigma_{position}$. The results improve for higher values.



(a) 3 (b) 5 (c) 7 (d) 10

Figure 9: Tracking results for constant velocity model for different values of $\sigma_{velocity}$ when $\sigma_{position} = 1$. The results are not good for low noise in position, irrespective of the variance in velocity.

- Increasing the number of particles means a more accurate mean state and therefore a smoother motion and better quality of the tracking.
- Less number of bins of color histogram decrease the precision of the histograms to be compared, and too strong average can destroy the recognizability of the object.
- Appearance model updating improves tracking in cases where the object undergoes small changes in appearance or changes in illumination.
- The performance of this tracker is limited by the subjectivity in choice of bounding box chosen initially by the user.

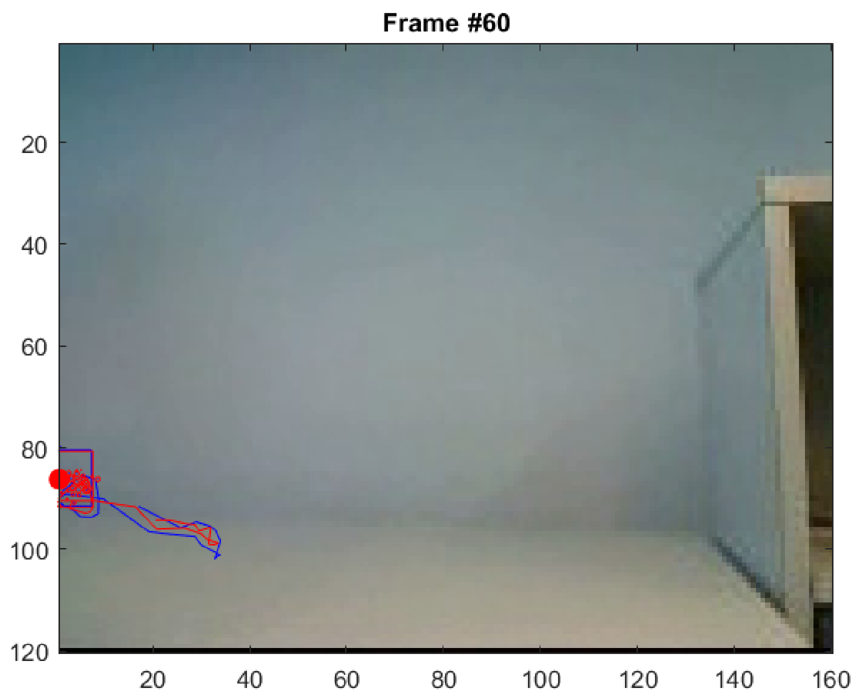
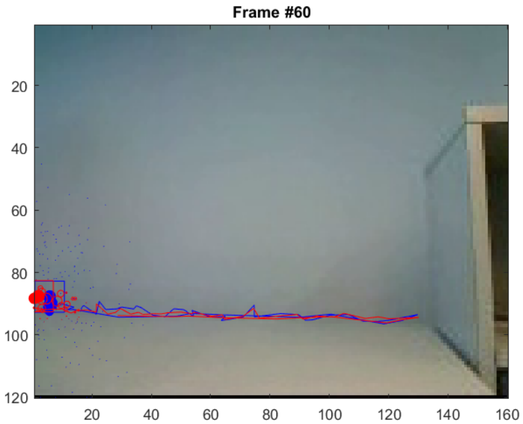
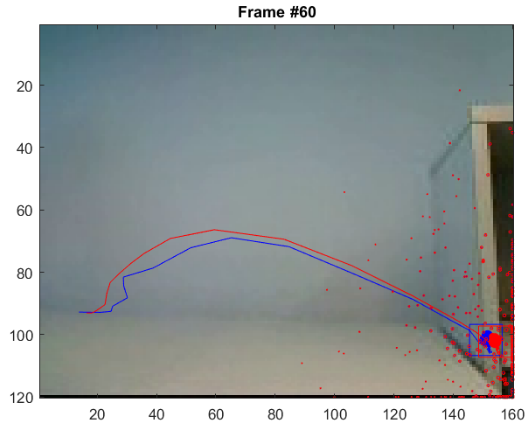


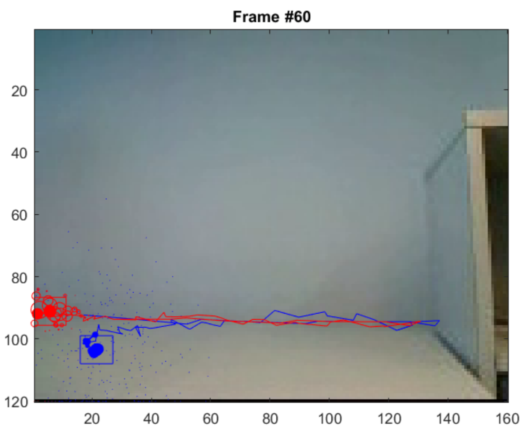
Figure 10: Tracking results for no motion model for video3 with parameters for $initial_{velocity} = [5, 0]$, $\sigma_{position} = 15$, $\sigma_{observe} = 0.1$



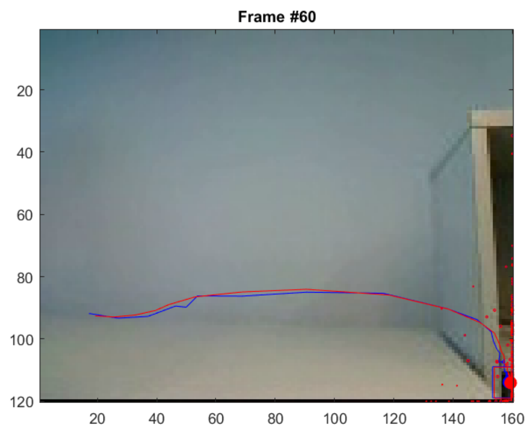
(a) 0.1



(b) 1

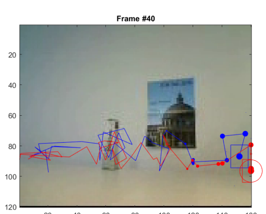


(c) 0.1

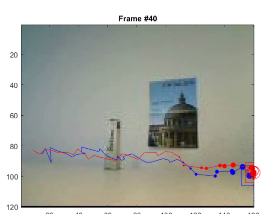


(d) 1

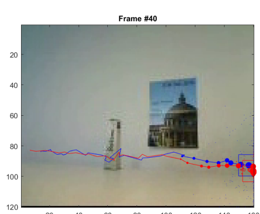
Figure 11: Tracking results for varying measurement noise for no motion (top) and constant velocity (bottom). High measurement noise tracks background clutter.



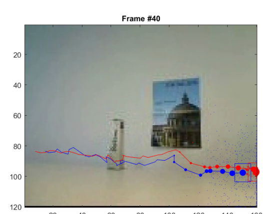
(a) 10



(b) 100

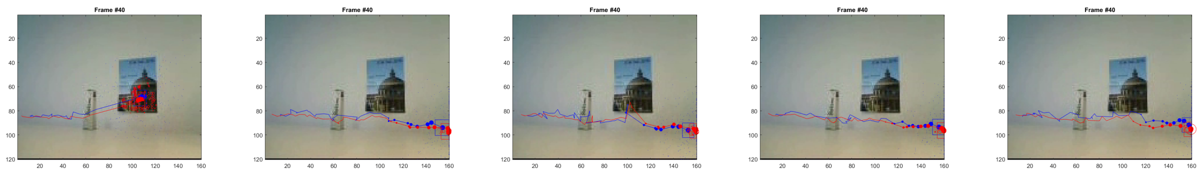


(c) 300



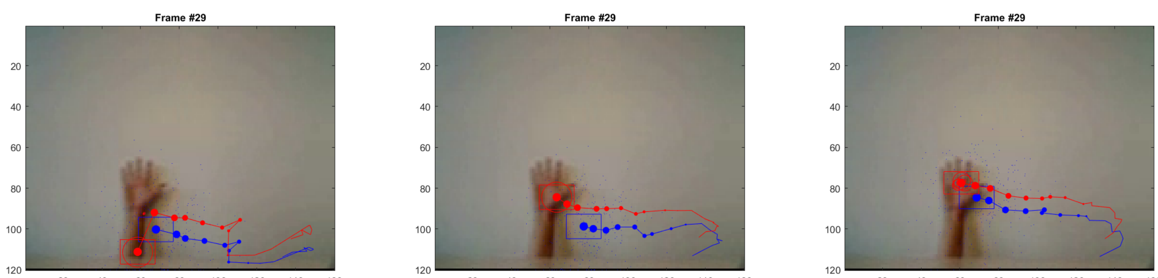
(d) 700

Figure 12: Tracking results for constant velocity model for different number of particles . Lower number of particles produce unreliable results.



(a) 2 (b) 16 (c) 30 (d) 50 (e) 100

Figure 13: Tracking results for constant velocity model for different number of bins of color histogram . Lower number of bins get affected by background.



(a) 0 (b) 0.5 (c) 0.8

Figure 14: Tracking results for constant velocity model for different values of α. Higher values tend to be more robust to illumination changes.