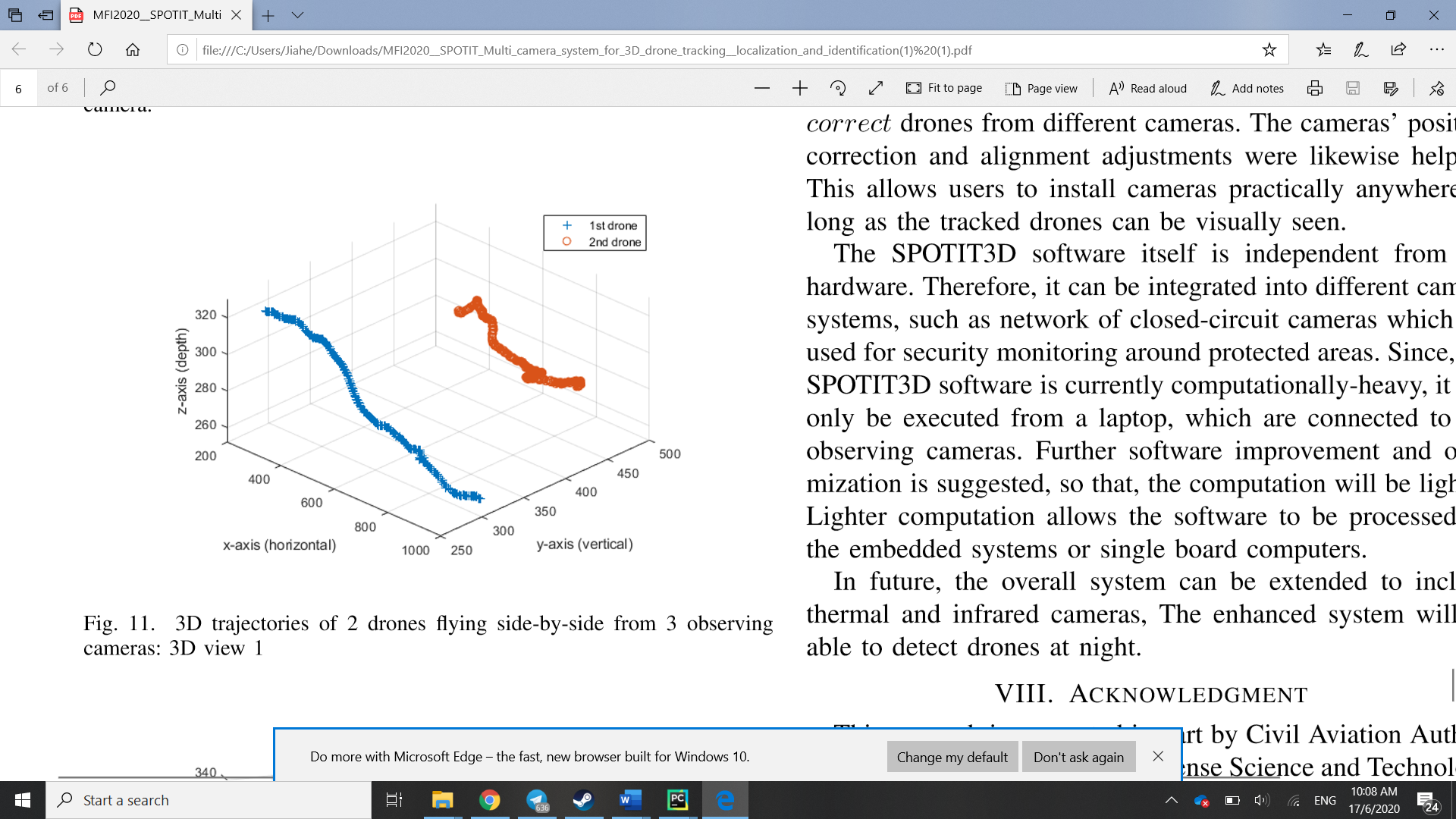
The paper “Multi-camera Multi-drone Detection, Tracking and Localization with Trajectory-based Re-identiﬁcation” suggests a method for target identification in a multiple drone, multiple camera case utilizing trajectory characteristics of each track to calculate a single track feature variable, for each track, and subsequently performing a cross-correlation on a set of from each camera to find the best match between tracks. The suggests that the actual functional form of may be arbitary, however, I believe that improvements can be made by tailoring the function used to calculate .

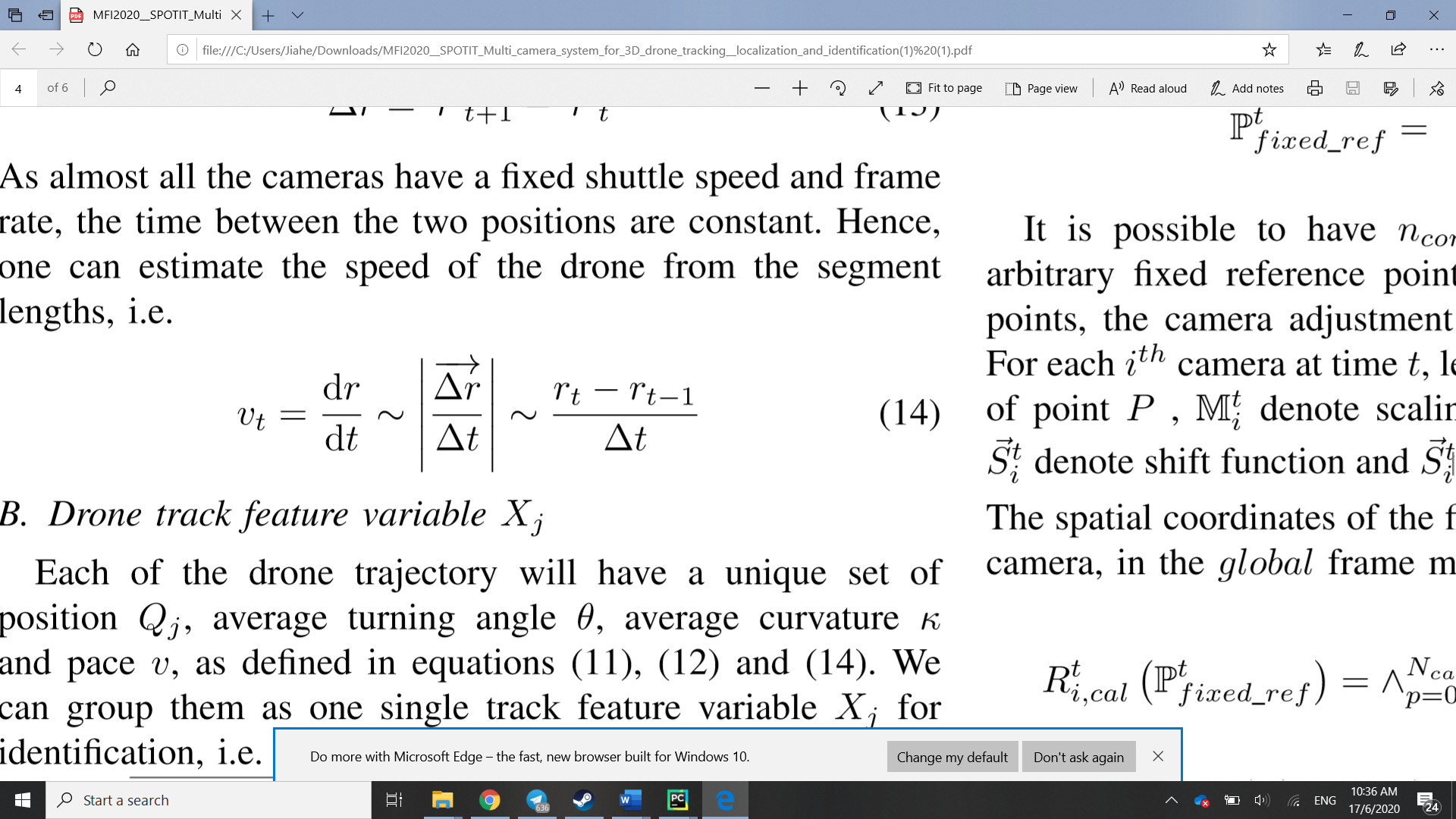


## Information only avaliable to one camera



Suppose the cameras were orthogonal to each other, which could possibly give the best results, but placed with one viewing the x-y plane and the other, the z-y plane. The camera viewing the x-y plane would observe a might higher pace than the camera viewing the z-y plane, owing to the drones motion of 600 to 800 units being much more pronounced in the horizontal direction as opposed to the maximum of 60 units in the Z-direction. If in this case, the 2nd drone possessed similar motion characteristics as the 1st but translated 90 degrees such that pace of the 2nd drone in the z-direction matched that of the 1st in the x direction, the pacing of the drones may lead to similar values.





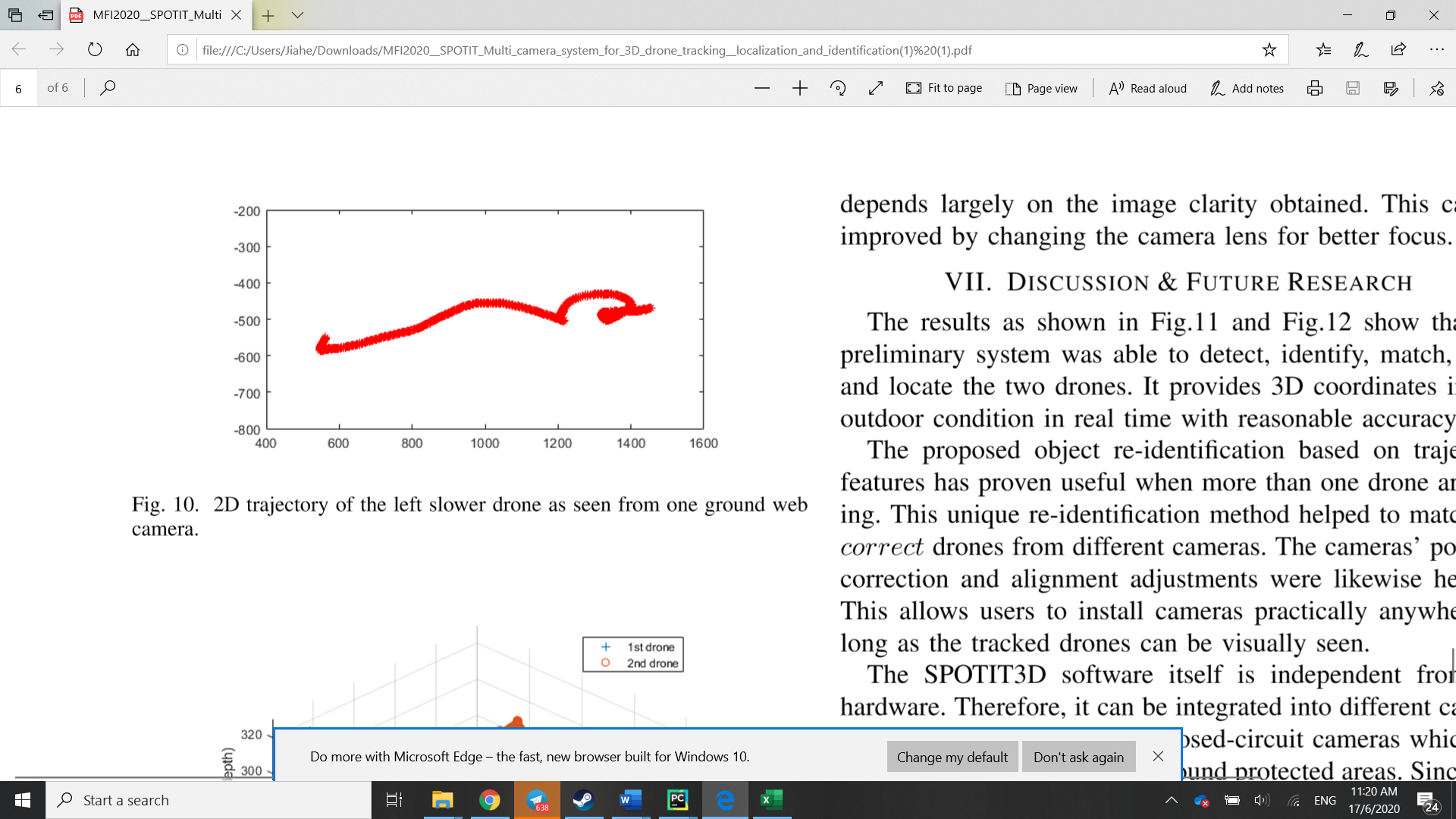
To compensate for this, I suggest that the pace could be broken down into its horizontal and vertical components in the camera frame, before being transformed by the cameras position in the camera matrix to incorporate information about which direction the drone is travelling to use for identification. In this example, information about the x-velocity of the drone is useless to the camera viewing the z-y plane and should be weighted down, allowing other trajectory characteristics to stand out more.

A proposed simple from of this implementation could be in the form:

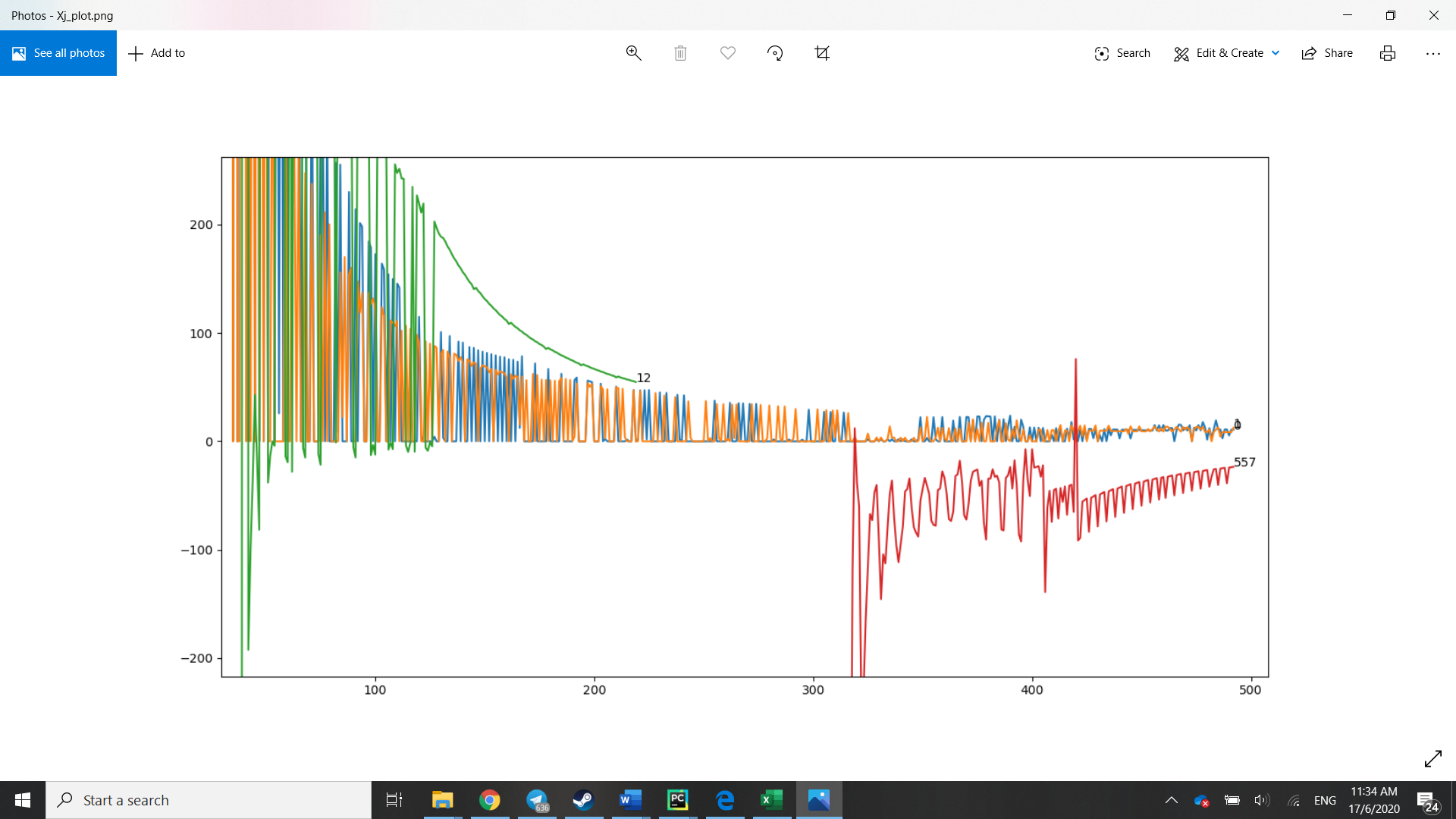
Where are the pace in the directions and is the angle between the two cameras. This proposed form assumes that both cameras are positioned on the same x-z plane, might may be extended to other or more dimensions.

In this manner, information that is irrelevant to the other camera of the pair is downscaled to reduce divergence of resultant from information captured by only one camera.

## Dilution over time



Suppose we have two cameras, one of which captured the drone at an earlier time, resulting in the longer track captured by the larger box, and another camera which only captured the drone at a later time, resulting in a shorter track. Using the mean turning angle and curvature, it is likely that the value from the first camera will be lower than the second as there are many samples diluting the mean.



The sensitivity of the value is also affected by the number of samples used to calculate it. While the long term value might be used to distinguish objects with different flight behaviours, such as the birds indicated by 12 and 557, and the drones indicated by 0 and 1, if it is not sensitive to changes it might not be able to tell different objects with the same flight characteristics apart. Suppose if one of the drones where to perform a manoeuvre, the magnitude of would not only be dependent on the manoeuvre itself, but also how long the drone has been tracked beforehand, as longer tracks during a flight phase with lower values would bring down the average.

A quick solution could be to adjust the number of recent samples used to calculate the mean, so as to have a more sensitive value which could be used to identify drones which the same long-term behaviour but in different phases of flight, e.g. hovering, steady level flight and acrobatic manoeuvring. A long-term running average could still be kept to distinguish the type of object being tracked.

## Discarding information

Having as a scalar value may be beneficial if simplicity and less computations are desired. However, they only posses a single degree of freedom and might be more susceptible to overlapping values.

If computational complexity is not a significant consideration, it may be beneficial to perform cross-correlation on the individual trajectory characteristics, before subsequently matching tracks with matches in more areas.

For example, performing cross-correlation on a single velocity component, or the velocity in vector form would allow directional information to be captured as well, allowing objects with the same speed, traveling up to be distinguished from objects travelling down or to the left or right, for example. It is at this stage that a simple weighted sum of the correlation values could be performed to match objects with similar velocities, curvatures and turning angles. As these values can change with time, it would also be important to use a smaller history window for correlation.

The best matching would be a higher value for

Where i and j are the indicies of the cameras and are gains for adjustment. I.e. the smaller the absolute difference between individual terms, the higher the value of .