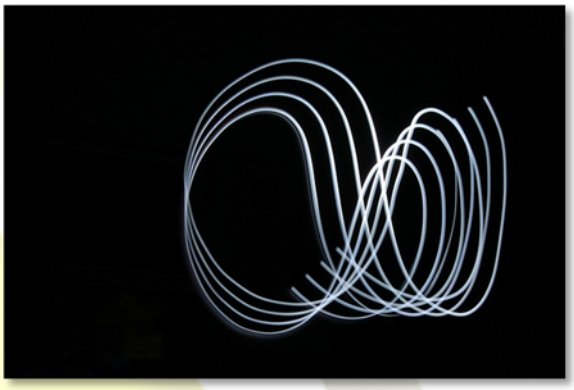
Object Tracking

Locating an object in successive frames of a video is called **tracking**. The definition sounds straight forward but in computer vision and machine learning, tracking is a very broad term that encompasses conceptually similar but technically different ideas. For example, all the following different but related ideas are generally studied under **Object Tracking**



Dense Optical flow: These algorithms help estimate the motion vector of every pixel in a video frame.

Sparse optical flow: These algorithms, like the Kanade-Lucas-Tomashi (KLT) feature tracker, track the location of a few feature points in an image.

Kalman Filtering: A very popular signal processing algorithm used to predict the location of a moving object based on prior motion information. One of the early applications of this algorithm was missile guidance! Also as mentioned here, “the on-board computer that guided the descent of the Apollo 11 lunar module to the moon had a Kalman filter”.

Meanshift and Camshift: These are algorithms for locating the maxima of a density function. They are also used for tracking.

Single object trackers: In this class of trackers, the first frame is marked using a rectangle to indicate the location of the object we want to track. The object is then tracked in subsequent frames using the tracking algorithm. In most real life applications, these trackers are used in conjunction with an object detector.

Multiple object track finding algorithms: In cases when we have a fast object detector, it makes sense to detect multiple objects in each frame and then run a track finding algorithm that identifies which rectangle in one frame corresponds to a rectangle in the next frame.

Tracking vs. Detection

Tracking is faster than Detection: Usually tracking algorithms are faster than detection algorithms. The reason is simple. When you are tracking an object that was detected in the previous frame, you know a lot about the appearance of the object. You also know the location in the previous frame and the direction and speed of its motion. So in the next frame, you can use all this information to predict the location of the object in the next frame and do a small search around the expected location of the object to accurately locate the object. A good tracking algorithm will use all information it has about the object up to that point while a detection algorithm always starts from scratch.

Tracking can help when detection fails: A good tracking algorithm, will handle some level of occlusion.

Tracking preserves identity: The output of object detection is an array of rectangles that contain the object. However, there is no identity attached to the object.