NFL Health & Safety - Helmet Assignment

Project 1

*An Introduction to video tracking*

*Cao Tianwei*

**1 .Introduction**

* 1. **Source of the topic**

This topic is proposed by National Football League (NFL), who aims to monitor and prevent injury in football matches. In this case, what NFL wants to do is to assign specific players to each helmet, which would help accurately identify each player's “exposures” throughout a football play.

To be specific, we are required to identify and assign football players’ helmets from video footage. We are expected to build a algorithm capable of detect and track moving object in a video.

It is of great significance to achieve this goal. To start with, the success of this algorithm can help protect players from injury. Secondly, by implementing this algorithm, NFL no longer need to manually label each exposure, this would dramatically increase the speed of research and analyze related to helmet impact.

Also, since we are expected to build algorithm without mapping, our project is also budget-saving.

* 1. **Video tracking**

Tracking object is actually an improvement on detecting an object in static image, it requires us to merge the knowledge of object detection and object tracking so as to predict it’s moving trajectory. According to ‘Video Tracking’ in Wikipedia,the objective of video tracking is to associate target objects in consecutive video frames. The association can be especially difficult when the objects are moving fast relative to the [frame rate](https://en.wikipedia.org/wiki/Frame_rate). Another situation that increases the complexity of the problem is when the tracked object changes orientation over time. For these situations video tracking systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object.

* 1. **Other Applications**

Apart from identify and track helmets in a football game, video tracking can do more.

According to ‘Small Animal Video Tracking for Activity and Path Analysis Using a Novel Open-Source Multi-Platform Application(AnimApp)’ by Srinivasa R. Rao’s team. Video tracking can track experimental biological model system outcomes such as altered animal movement capability or behaviors that are difficult to quantify manually. In their case, they build a video tracking app based on OpenCV to track movement of little wild animals such as rabbits and rodents, which turns out to be effective.

Video tracking can be used in aviation field too. In ‘Energy-based video tracking using joint target density processing with an application to unmanned aerial vehicle surveillance’, Li. Hong and his team build an algorithm from the continuous wavelet transform (CWT) to estimating target trajectories and sizes of flying objects. His work is excellent in velocity selection and is applied to a UAV surveillance application where multiple vehicles move closely to each other on a multi-lane road.

**2. Object detection and Object tracking**

Video tracking can be divided into object detection and object tracking.

Object detection is to identify an object in a frame, put a bounding box or a mask around it and classify the object(this process is able to detect multiple objects in a single frame).

Object tracking is what we do after identification of an object, we track the object across the entire frame. In this case, we have to predict the trajectory of an object in subsequent frames.

1. **Challenge in video tracking**

Video tracking has following challenges.

To start with, we have to deal with occlusion in video tracking. Occlusion happens when a certain object can be detected in one frame but goes undetected in the next frame. In this case, we have to figure out how to detect a same object in current frame as well as a much later frame associates with his older track.

Secondly, we have to overcome the variation in viewpoint. This happens because our target keeps moving in the frame, it may rotates and shift from time to time. In this case, there may be significant variation in viewpoint, which may hamper the ability of our object tracking material.

Thirdly, the non-stationary camera can add difficulty to our project. It is because our camera sometimes undergoes vibrations in the progress of recording, this happens frequently in sports games especially football. In this case, we must build a robust tracker to overcome this problem.

Also, it sometimes difficult to get good training data, it is because we require a set of data which can represent the trajectory of an object. In this case, we can not randomly choose unconnected images as training data(what we may do in an object detector).

1. **Traditional methods for video tracking** 
   1. **Meanshift**

The goal of this algorithm is to find all the modes in the given data distribution.

This algorithm is a modification of traditional K-Means in which simple centroid technique of calculating cluster centers is replaced by a weighted average that gives importance to points that are closer to the mean.

* 1. **Optical flow**

When building this algorithm, we do not need to extra features from object. Instead, we make use of spatio-temporal *image brightness variations* at a pixel level to do video tracking. To be specific, we hope to obtain a displacement vector in order to track the object in a video. This algorithm relies on these assumptions:

1. The brightness of an certain area changes little in two nearby frames.
2. The motion of two nearby points on a same surface act similar in two nearby frames.
3. Motions of a patch changes gradually.
4. The manner of points of an object do not move in a haphazard way.

**5. Deep SORT(Simple Real time Tracker)**

Deep SORT is a efficient algorithm which is very popular in certain area. This is an modified version of SORT (Simple Real time Tracker) algorithm. Deep SORT is based on Kalman filter, in which we hope to use detections from the past to arrive a best guess for the future. In this process, the appearance of a few errors is acceptable

Using Kalman filter, we can achieve some predicted Kalman states and we have to associate them with newly arrived measurements. To do so, we build an assignment problem and use Hungarian algorithm to solve it.

So at this point, we have built an object detector and a Kalman filter, also the associate problem has been solved, what we have to do next is to build an distance metric based on the ‘appearance’ of the object called ‘the appearance feature vector’. This vector is used to offset the short come(like occlusion and other things we have mentioned) of Kalman filter when dealing with real world application. This step makes use of deep learning strategy.

**6. References:**

1. ‘DeepSORT: Deep Learning to Track Custom Objects in a Video’ by [Shishira R Maiya](https://nanonets.com/blog/author/shishira/). Link: https://nanonets.com/blog/object-tracking-deepsort/

2. ‘SIMPLE ONLINE AND REALTIME TRACKING WITH A DEEP ASSOCIATION METRIC’ by Nicolai Wojke†, Alex Bewley, Dietrich Paulus.

3. ‘Small Animal Video Tracking for Activity and Path Analysis Using a Novel Open-Source Multi-Platform Application (AnimApp)’ by Srinivasa R. Rao, Sam W. Z. Olechnowicz, Patrick Krätschmer, James E. C. Jepson, Claire M. Edwards & James R. Edwards.

4. ‘Energy-based video tracking using joint target density processing with an application to unmanned aerial vehicle surveillance’ by L. Hong, Y. Ruan, W. Li, D. Wicker, J. Layne.