**Technical Assessment Instructions – Engineer**

Objective

We want to create a metric (or up to 2 metrics) that summarizes a given pitcher’s arm path during his delivery.

We know we have a good metric, or two metrics, when they can show that there are at least 2 different arm paths in the sample data that has been sent over.

Baseball Terminology

I do believe that a key part of success in this position is being able to have an understanding of the game of baseball. However, I think that aspect can be learned and not needed for the purposes of this assessment. So here is some information that may be helpful if you unfamiliar with specifics of pitching.

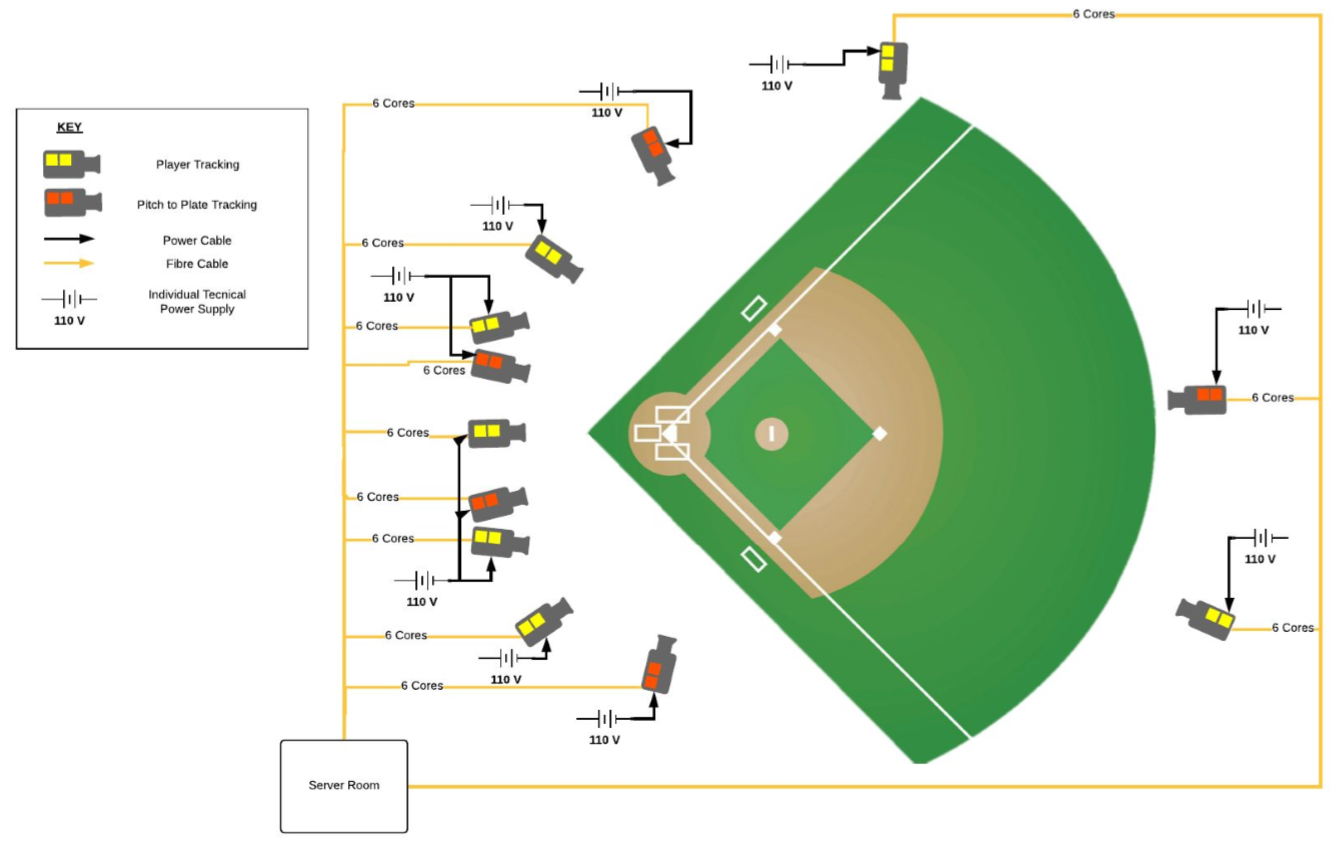
Below are visuals of the different landmarks in the pitching delivery. Each one of them can be quantified, although some are harder than others. In this exercise, we are primarily interested in Leg Lift to Ball release. However, it does not have to start with leg lift. It is possible you find more signal if you start at hand break, or drive, or something else. I can just say for certain that we do not care about how a pitcher comes set or his follow through.

A baseball player running with a baseball glove

Description automatically generated with medium confidence

For this exercise, we are only interested in what the throwing arm is doing. This consists of the shoulder, elbow, and wrist. Although, it does not have to contain all 3 of those. You may have success just looking at the wrist, or wrist + elbow.

About the Data

* Columns
  + sched\_id : integer to denote a unique game
  + joint\_type\_id : to be joined with the joint\_type.csv file. There are no joints with ids greater than 18 because they are not useful
  + time : 0 = ball release. < 0 is before ball release. Units are in seconds.
  + x,y,z : global coordinates of the specified joint in feet
  + astros\_pitch\_id : integer to denote unique pitches within a given sched\_id (game)
  + pitcher\_id : integer to denote a specific pitcher
  + bats : the handedness of the batter that is currently hitting (right or left)
  + throws : the hand the pitcher throws with (use this to identify the throwing arm)
* Coordinate System
  + Z axis : Vertical Axis, pointing to the sky
  + X axis : Lateral Axis, pointing to First Base
  + Y axis : Anterior / Posterior Axis, pointing to centerfield
  + Origin: Home Plate
* 

Origin

* Other Useful Information
  + This data is raw. There will certainly be bad readings in the timeseries data. Your knowledge of signal processing is one aspect of this assessment, but not the primary aspect. If you are strong in signal processing and can quickly improve the data then that is a huge plus. But if you are not familiar at all, then I would just throw out the bad data and focus more on developing a useful metric.
  + The wrists and the hips are going to be the most noisy. The wrists get a little noisy around release, when the arm gets moving very fast. The hips are noisy throughout the delivery because the CV model has a hard time with most uniforms since the pants and shirts are the same color.