Calculating Acceleration Vectors in American Football GPS Data

Steven Plaisance

May 10, 2023

Abstract

This paper seeks to convert scalar acceleration values from athlete GPS data to a series of vectors describing different types of athlete movement. This increased specificity of measurement provides utility both from a physiological standpoint and a sport-specific production perspective.

1. Introduction

GPS data is an increasingly popular method of monitoring athlete workload and performance indicators in collegiate and professional sports organizations. Ongoing advances in sports medicine, sports science and sports performance are supported by increasingly sophisticated means of measuring and analyzing athlete workload and performance, leading to more advanced understanding of the relationship between workload and performance among sport coaches, strength coaches, athletic trainers and dietitians.

GPS data offers unique utility due its precision of measurement, width of descriptive possibilities and relatively low invasiveness. Compared with other leading forms of athlete monitoring, GPS sensors monitor athletes directly in their sport-specific training environment (ie, on the field during practice rather than in the weight room during lifts.) The presence of GPS sensors changes very little if anything in regards to the sport-specific training environment. Often times, sensors are placed within the shoulder pads prior to practice and athletes are not explicitly aware they are being recorded.

Within American Football, GPS data has facilitated a wide range of uses in both professional and collegiate organizations, though there is little crossover between the two. Professional organizations have focused on developing sport-specific performance and production metrics to evaluate free agents, draft prospects and potential trade targets. Meanwhile collegiate organizations remain focused on physiological performance indicators as a complement to their sports medicine and strength and conditioning departments.

One major limitation of raw GPS data: velocity and acceleration are provided in scalar values. But athletes move in a variety of ways, meaning any sophisticated assessment of physiological performance would account for the various types of movement an athlete may perform. In this paper, I will convert raw scalar acceleration into 8 different vector accelerations, each describing a different type of movement.

- Forward propulsion
- Forward braking
- Backward propulsion
- Backward braking
- Left propulsion
- Left braking
- Right propulsion
- Right braking

More specific measurements of athlete movement provide utility in both the physiological performance area and the sport-specific performance area. Relevant changes in athlete performance are likely to vary across different types of movement. Athletes may also be asked to perform different movement types at varying frequencies based on their position or role within a given playcall.

2. Literature Review

At this time, there is little public research available regarding American Football GPS data from a physiological perspective. There has been a great deal of performance research using GPS conducted in other sports such as soccer and rugby. But these studies are based on summary workload data rather than raw positional data.

Meanwhile, there is a host of public, football-specific positional data research available thanks to the Big Data Bowl, an annual sports analytics contest hosted by the National Football League. The contest provides a sample set of GPS data from NFL games and offers contestants a broad prompt to answer creatively.

The winning submission of the 2021 Big Data Bowl (Peng et al. 2021). used a variable framework to assess pass coverage performance based on the defensive playcall and the positioning of nearby receivers at the time of the pass. In essence, models such as this one minimize scouting time by automating tedious charting tasks, saving organizations precious time and money and creating a potential competitive advantage.

Another Big Data Bowl submission from 2019 used a clustering-based approach to identify different route types in GPS data and went on to model optimal route combinations based on defensive personnel and alignment (Chu et al. 2019).

Ian Barnett calculated expected yardage gains based on a given field state, which "represents the positions, orientations, accelerations and speeds of every player on the field at a given moment." He then evaluated punt and kick returners based on their deviation from the expected yardage gains (Barnett 2022).

Jay Li and Ali Kasar took a more granular approach to evaluating a player's added value. They essentially turned gunner-vise interactions into zero-sum games to evaluate each inidividual (Li and Kasar 2022). Gunners are the punt coverage players responsible for moving down the field and making a play on the ball. Vises are the punt return players responsible for delaying the gunners.

Another submission calculated the optimal path for a given kick return based "given the current distribution of defenders and blockers on the field" and evaluated kick returners based on deviation from the optimal path (Gross et al. 2022).

These are all impressive and highly-specific applications of positioning data in American football. There are a number of ways that this paper could act as a bridge between the performance science side of GPS data and the football-based side.

For example, the route identification could be used in tandem with acceleration vectors to assess different movement types based on the route run. Curl routes necessitate a strong forward propulsion followed by forward braking, while crossing routes require a smooth transition from forward propulsion to left/right propulsion depending on the direction of the route.

Defensively, athletes with greater backpedal propulsion would make better cornerbacks, while athletes with better forward braking and propulsion would make for better safeties. A more specific approach could cause a defensive coordinator to choose his man-to-man matchups based on the individual movement profiles of each receiver and defender.

3. Data

The raw GPS data includes the following variables:

- x: field x coordinates (meters)
- y: field y coordinates (meters)
- ts: POSIX time in seconds since the start of the epoch
- cs: Observation time offset in centiseconds
- face: the magnetic facing of the unit
- v: velocity (meters per second)
- a: acceleration (meters per second per second)

The data is sampled at 10Hz, meaning there are 10 observations per athlete per second.

Due to data size limitations, the data provided in this paper is specific to one athlete from one practice.

4. Empirical Methods

A handful of new variables must be calulcated in order to arrive at acceleration vectors.

First, dx and dy represent the instantaneous change in positioning from the last observation in the x and y planes. These are used to calculate alpha (the direction the athlete is moving)

$$\alpha = atan2(dy, dx)$$

The key variable used to calculate acceleration vectors is designed as "theta" and describes the difference, in degrees, between the direction that the athlete is facing (face) and the direction that the athlete is moving (alpha)

$$\theta = \alpha - face$$

Next, we calculate four proportion variables, one for each direction of movement (F for forward, B for backward, L for left and R for right). These variables take a range from 0 to 1 and describe the proportion of scalar movement that is in the designated direction.

$$p_F = \begin{cases} 1 - (|\theta|/90), \ \theta < 90 \ \& \ \theta > -90 \\ 0, \ Otherwise \end{cases}$$

$$p_B = \begin{cases} (|\theta| - 90)/90, \ \theta > 90 \ OR \ \theta < -90 \\ 0, \ Otherwise \end{cases}$$

$$p_L = \begin{cases} 1, \ \theta = -90 \\ 0, \ \theta = -180 \\ (90 - (-\theta \mod 90))/90, \ \theta < -90 \\ -\theta/90, \ \theta < 0 \\ 0, \ Otherwise \end{cases}$$

$$p_R = \begin{cases} 1, \ \theta = 90 \\ 0, \ \theta = -180 \\ (90 - (\theta \mod 90))/90, \ \theta > 90 \\ \theta/90, \ \theta > 0 \\ 0, \ Otherwise \end{cases}$$

From there, the proportion variables are multiplied by the scalar distance variable to find vector distance in each direction. Vector velocities are calculated from changes in vector distance, and vector accelerations calculated from changes in vector velocity.

5. Research Findings

The theta variable is critical to all subsequent calculations, so it was first necessary to ensure that theta was functioning properly. Theta is the difference of two variables, face and alpha, which represent global coordinates. Conbining the two into theta is the key conversion from global coordinates to the athlete's individual axes.

Plot 1.1 shows the distribution of all values of theta throughout the session. As expected, values tend to center around 0 and distribute around either side equally, though there does appear to be some bias toward the athlete's left side in this case.

We know that football players spend the majority of their time moving forward during a practice, so plot 1.1 confirms the expected behavior.

We should also expect that the highest values of scalar velocity and scalar acceleration correlate with values of theta that are near 0.

Plot 1.2 and 1.3 investigate the relationships between theta and scalar velocity and acceleration respectively. As expected, there is a clear relationship between the highest values of each and values of theta near 0, though there seems to be more noise than expected, particularly in the case of acceleration.

Moving on to the acceleration vectors themselves, table 1.4 displays summary statistics for each scalar acceleration and each acceleration vector.

Again, these values are largely in-line with what we would expect, though the bias toward the athlete's left side appears once again.

Additionally, the difference between the maximum values of scalar acceleration and forward propulsion is much larger than would be reasonable. 5 m/s^2 is extremely fast acceleration for an athlete and would not be possible under any circumstance other than forward propulsion.

6. Conclusion

While the goal of this paper remains an attainable and insightful endeavor, I conclude that this iteration of acceleration vectors is too unreliable for any sustained use in an athletics or healthcare setting. Instead, these formulas should remaiun experimental until a method for calculating more precise acceleration vectors becomes possible.

After studying the data more closely, the athlete's 'face' value will often shift significantly toward the left or right in the midst of a highly-explosive forward propulsion. Interestingly, this phenomenon seemed more likely to occur in the top-speed phase of the movement rather than the acceleration phase (first ~5 seconds).

The most likely explanation, given that this athlete plays wide receiver, is that the athlete is turning to face the passer and catch the ball during his route.

The GPS sensors are located between the athlete's shoulder blades, which in theory provides an accurate indicator of which direction the athlete's torso is facing. Unfortunately, the premise of this paper would be more easily accomplished with an accurate indicator of which direction the athlete's hips are facing, not his torso.

As for next steps, there remains an abundance of possible insights to gain from positional data, both in the performance science and the football science space.

Performance Science Research Areas - Injury Rehabilitation - Return-to-play Protocol - Workload Monitoring - Performance Assesment

Football Science Research Areas - Roster Construction - Draft Analytics - In-game Decision-making - Play signal processing

In terms of this specific premise, experimentation with GPS sensors located at the hips may yield more precise results and solidify acceleration vectors for use in both performance-related and football-related applications.

References

Barnett, Ian. 2022. "Introducing COYOTE," NFL big data bowl,.

Chu, Dani, Matthew Reyers, James Thomson, and Lucas Wu. 2019. "Route Identification in the NFL," NFL big data bowl,.

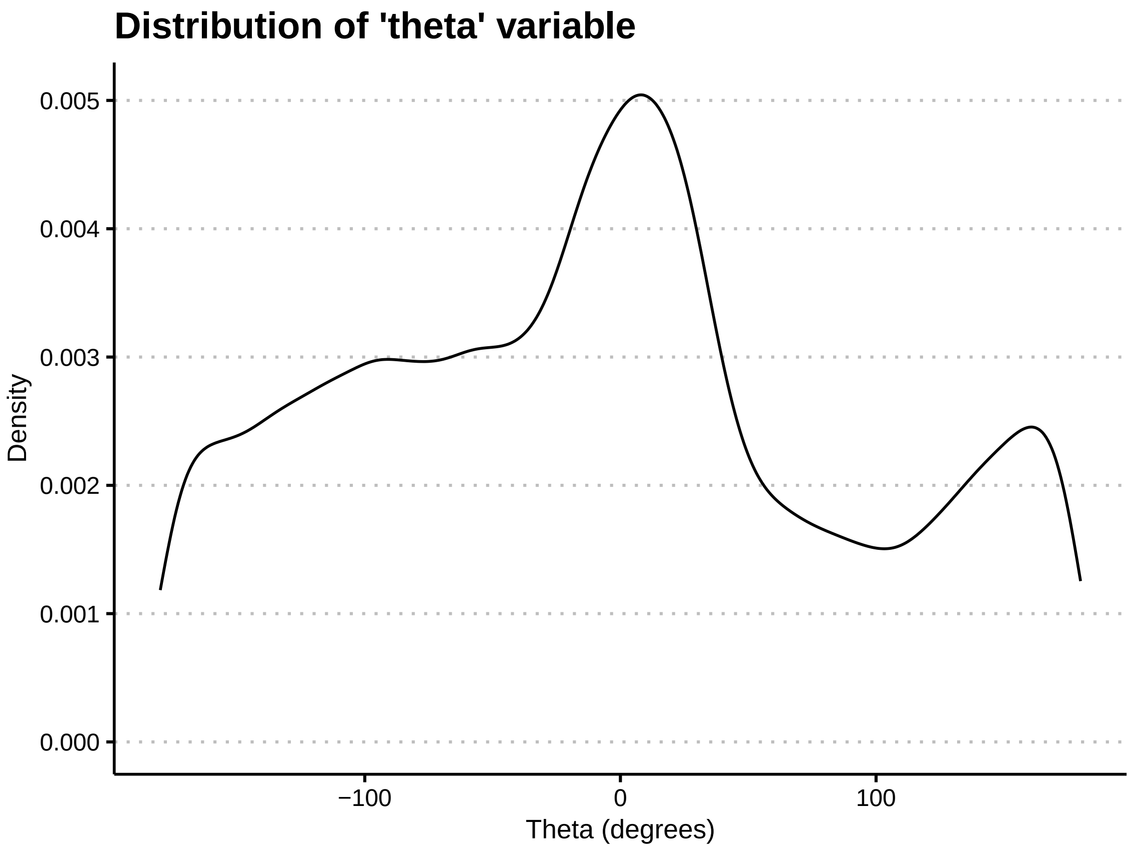
Gross, Ryan, Joseph Rudoler, Tai Nguyen, and Ryan Brill. 2022. "Optimal Run Path for Kick Returners," NFL big data bowl,.

Li, Jay, and Rahul Kasar. 2022. "FireTime: Evaluating Gunners and Vises," NFL big data bowl,.

Peng, Wei, Marc Richards, Sam Walczak, and Jack Werner. 2021. "A Defensive Player Coverage Evaluation Framework," NFL big data bowl,.

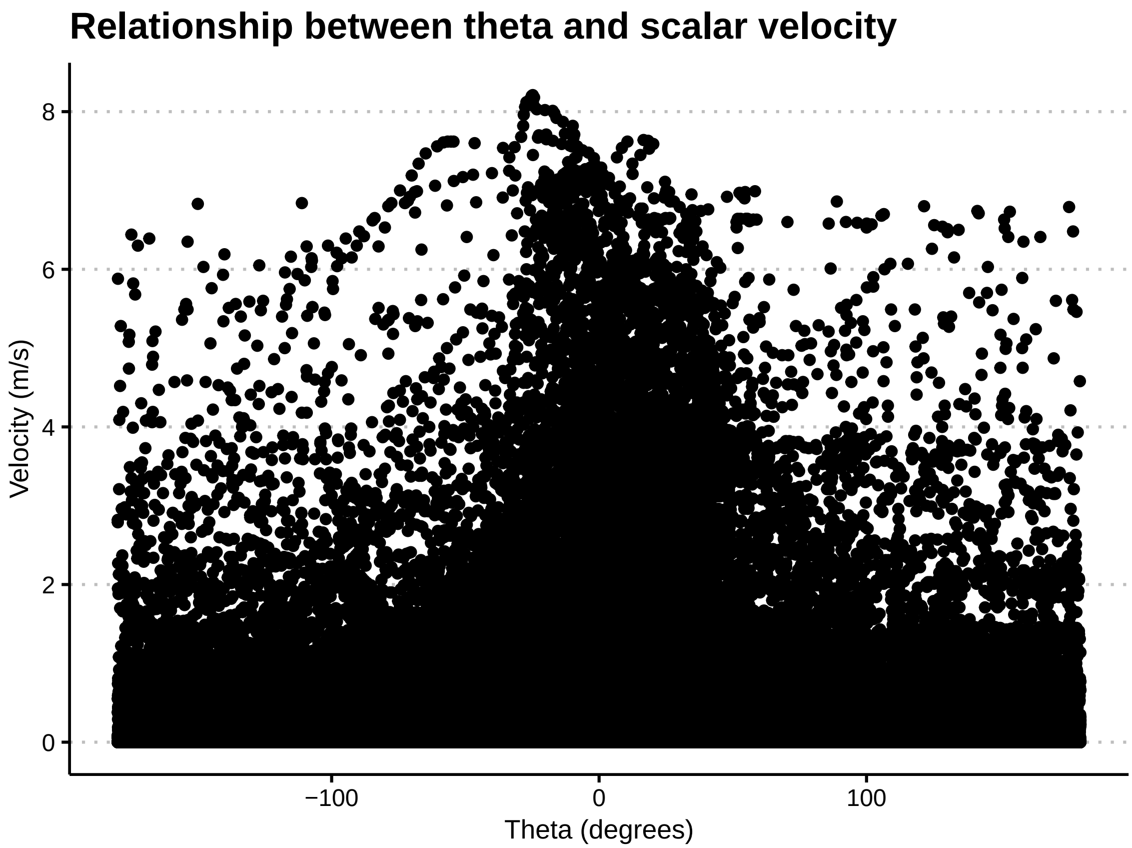
Appendix

1.1



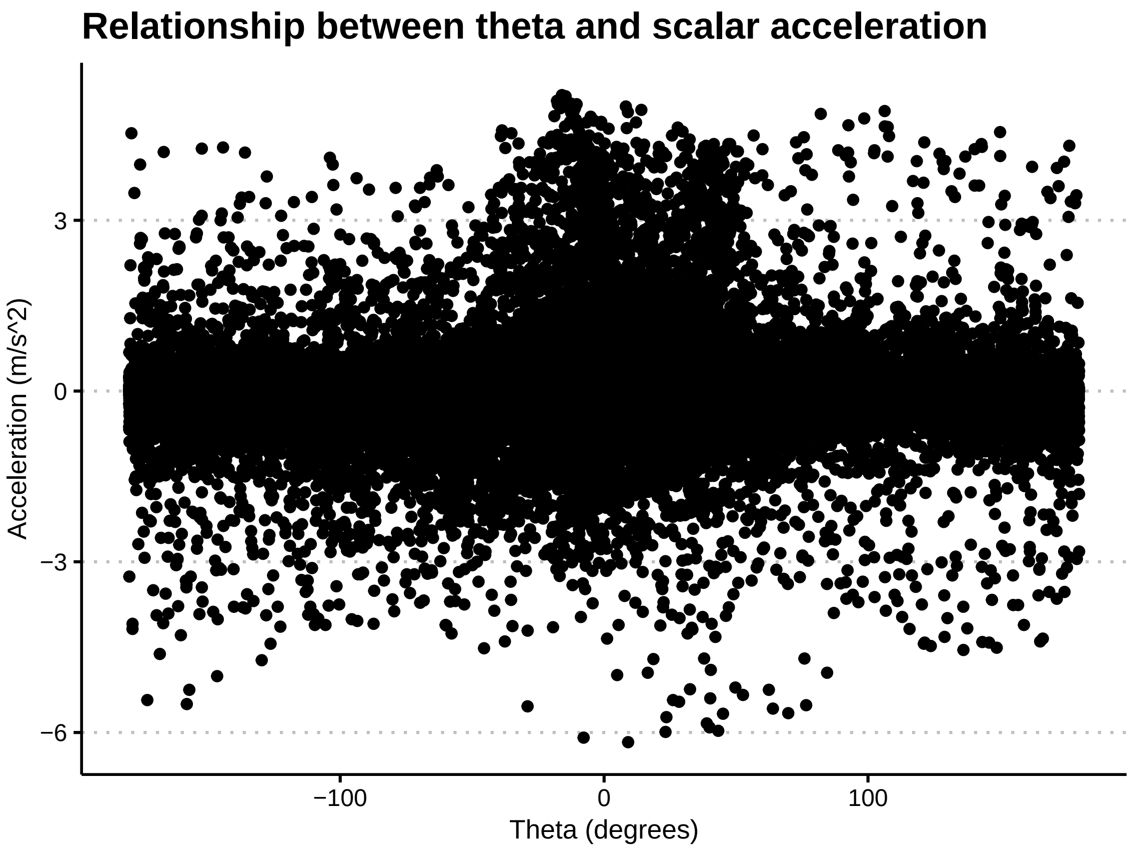
Distribution of 'theta' variable
0.005
0.004
0.003
0.002
0.001
0.000
Density
−100
0
100
Theta (degrees)

1.2



Relationship between theta and scalar velocity
Velocity (m/s)
8
6
4
2
0
−100
0
100
Theta (degrees)

1.3



1.4

Summary Statistics of Scalar and Vector Accelerations

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Scalar Acceleration	910	0	0.0	0.5	-6.2	0.0	5.2
Forward Acceleration	81219	0	-0.0	0.4	-4.6	0.0	4.3
Backward Acceleration	64838	0	-0.0	0.2	-3.5	0.0	3.4
Left Acceleration	82138	0	-0.0	0.2	-4.0	0.0	4.0
Right Acceleration	73880	0	-0.0	0.2	-3.3	0.0	3.6