Field Relativity and Recognition of

American Football Formations and Plays

William Johanson

Supervisor: Richard Green

University of Canterbury

Christchurch, New Zealand

[wpj19@uclive.ac.nz](mailto:wpj19@uclive.ac.nz)

Abstract:

# Introduction:

Data analytics has become prevalent in all sports, with no sport more involved than American Football, especially at the highest levels. For teams to gain a competitive edge, being able to call upon endless amounts of historical data to determine patterns, tendencies and weaknesses of their opposition can give them the winning margin. However, creating that data repository is very time consuming when done by hand. By utilising computer vision, the amount of man hours required trawling through game film to provide the data sets will be reduced, allowing more time to be spent making sense of the data.

Using computer vision techniques, the purpose of this project is to be able to take All-22 game film, shown in Figure 1, to identify 22 players on the field, eleven for each respective team.

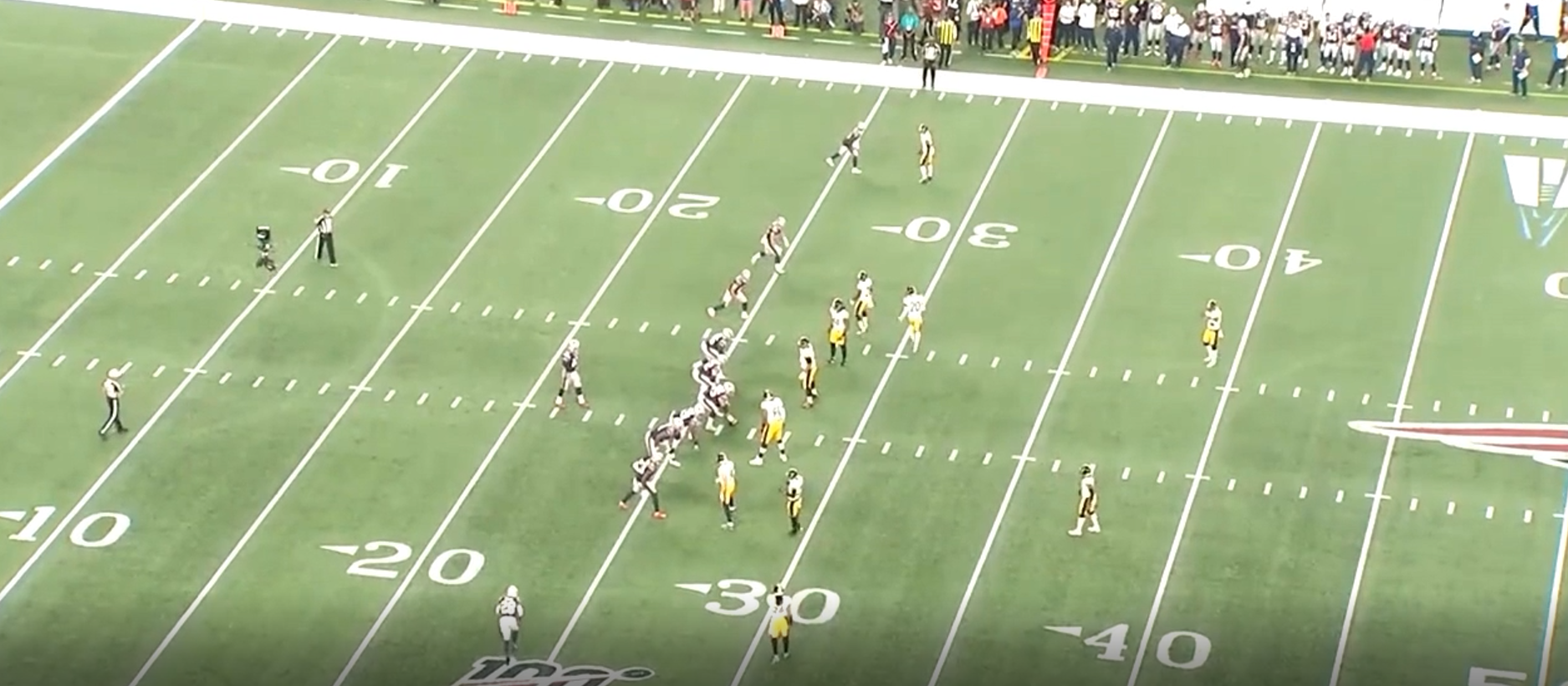


Figure 1: Screenshot from All-22 game film (courtesy of NFL Gamepass).

Then taking those players positions and map it to their field position and pre-snap formation. Furthermore, tracking their positioning throughout the play will allow the route combinations or post snap defensive alignment to be analysed.

# Background:

## Existing Products:

There are two areas of existing products that have been analysed to determine where the value lies within this research project.

The first is analytics companies such as *Pro Football Focus* (*PFF*). *PFF* analyse film taking upwards of 10 hours per game (there are up to 65 FBS college football games and 16 NFL games a weekend during the season to analyse) to glean critical data to allow for their analysis. Therefore, the value in using computer vision to track and store information on the formation and plays being run, is a step in the direction to reducing the man hours spent creating datasets. It will allow for more time by such companies to be spent analysing the dataset, which is where the ability comes to find trends.

The second is the areas of research into computer vision in sports and American football specifically. Computer vision is still in the development stages across all sports with companies such as *Sentio Sports Analytics*, which specialise in Soccer analysis. *Sentio’s* goal is to track players in real time and then push the data produced to analytical programs on cloud. Another company, which has expanded into more sporting codes is *Sportlogiq*. *Sportlogiq* began in Ice Hockey, which is very fast paced but has less players on the playing field than in an American Football game. *Sportlogiq* offer basic tracking data to allow for ‘play by play’ calls of the game being analysed with 96% accuracy. Hence, it must be monitored and corrected manually. Their future developments are invested in 2D and 3D body pose estimation to understand position and movement of players. The benefits including a better examination of players performance and a more detailed explanation of what’s occurring in-game.

There are two research papers delving into the same area as this research paper explores. The depth of research gets to the level of determining the field boundaries and yard markers using Hough thresholds. The second task the researchers achieve is to pick out the players on the fields with varying levels of success. One of the research papers used colour thresholding, utilising the two colours of the uniform with a distance minimum between the two points, to determine both teams’ players.

NEED MORE DETAIL HERE

* Analytics companies existing and their methods.
* Research done with CV in the field.
* CV in other sports??
* What methods are useful to glean information I need

## Areas of Improvement

The current colour thresholding method to find the players still produces noise due the colours existing in other parts of the frame as well as players on the side-line which aren’t involved in the play.

A development on the existing products to create value for the likes of PFF is to take the position of the players, simply observed with a rectangle etc.. and then develop that into a new window orientated aesthetically to then map the player positions to their pre-snap formation. And post-snap, following the players to then determine the route combinations or defensive motions being run. This is a suitable next step with further developments aplenty with utilising specific positioning to get quantifiable values for further analytics.

# Method:

## Aims:

The key aims for the project are:

* Replicating the existing capabilities of other research projects but with better accuracy which includes:
  + Field outline.
  + Player detection (offense and defence).
  + Outputting key data points to informative end-user plots.
* The implementation of neural networks to learn the formations to add onto existing research.
* Having initial individual frame screenshots for proof of concept then further develop into frame by frame analysis and post-snap play recognition.

## Field Outline Methodology:

Grayscale or colour threshold to allow only the white of the field, canny edge detection, hough threshold for the longer lines, checking the angle to ensure it’s a field line as they are less than 90 from origin.

* Use tess to determine the position of the numbers and then plot those w.r.t to the lines to give relative position. *Noisy words come through though. Doesn’t show up wpords on any angle so either need to angle the whole scene and then try or find another feature to get field position.*

## Player Detection Methodology:

* Colour thresholds, primary and secondary in BGR.
* Create a mask of the both offense and defense.
* Join the masks.
* Use adaptiveThreshold with thresh gaussian and thresh binary with an 11 matrix – need to add some maths in and reasoning behind my numbers.
* Then open and close morphology with a 2,2 kernel and 3 open iterations then 2 closing.
* Find the contours.
* Find the logo contour and get its extremities points and then remove all contours within (colour threshold drawback).
* Then draw on all the contours with a rectangle with them.
* Mapping players – colour threshold. Offense/defence. Have chosen a tough play since they both have white. Plot these players against the lines and the numbers.
* Output offense and defensive formations with angle adjusted to a new window and then begin on learning formations?? Using matplotlib?? Affine Transform?

Do I do offense and defense separately?

## Neural Network Methodology:

Use neural networks to train it on what each formation key points consist of.

# Results:

# Further Developments:

* Adapting to all colours of uniforms home/away.
* Training more formations offense/defense.
* Outputting player data e.g. stats/physical data e.g. distance run.

# Conclusions:

# References:

<https://sportlogiq.com/en/about-us>

<https://youtu.be/1VMt05K7xPY>

<https://cvgl.stanford.edu/teaching/cs231a_winter1415/prev/projects/LeeTimothy.pdf>

<https://www.sentiosports.com/>

To view:

<https://siliconangle.com/2019/12/05/nfl-teams-aws-simulate-predict-player-injuries/>

Sportvision, Description of Football Technologies, www.sportvision.com/football.

STATS LLC, “SportVU Player Tracking Technology,” www.stats.com/sportvu/sportvu.asp.

D. Cervone, A. D’Amour, L. Bornn, and K. Goldsberry, “POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data,” 8th Annual MIT Sloan Sports Analytics Conference, February 28 – March 1, 2014.

A. McQueen, J. Wiens, and J. Guttag, “Automatically Recognizing On-Ball Screens,” 8th Annual MIT Sloan Sports Analytics Conference, February 28 – March 1, 2014.

National Football League, “NFL Rules Digest: Field,” www.nfl.com/rulebook/field. Timothy Lee – leete@stanford.edu – CS231A – Final Project Progress Report 9 / 9

R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2004.

R. Szeliski, Computer Vision: Algorithms and Applications, Springer, 2011.

Disclaimer

The All-22 film utilized in this investigation is property of the National Football League. Utilization of this data is strictly done in an academic and non-commercial manner, and such use is congruent with the Terms & Conditions of the National Football League. Out of respect for the copyright of the National Football League, the investigation dataset is not available for public distribution.

**I am after two**(or more) **pages of background research on your project.**

Note that this is so that I can give you feedback and your text/etc can be used in your final paper.

Please submit with the file name format of your: first name space last name.

**As well as following the checklist and tips below (including your updated title),**

**please detail your OS, IDE, language, device(PC/smartphone/etc), camera(resolution/frame-rate/etc), OpenCV version, test data, etc.**

**Checklist/guide for writing a research paper for publication**(red items are most important)**:**

**title** with your name followed by name of supervisor (me in this case)  
**introduction**  
\* goal / problem statement (what is your goal/what problem are you solving)  
**background** (focused specifically on concepts/papers related directly to your work - not too general)  
\* cite publications **critically** (be critical of prior research => mention limits/constraints/etc)  
\* equations (you may be able to find many of these in other papers/text books)  
\* diagrams (you may be able to find many of these in other papers/text books)  
\* images (you may be able to find many of these in other papers/text books)  
\* summarise limitations of prior research (the entire reason for doing this paper and also shows that you can read critically)  
**solution/method**

\* how your solution overcomes limitations of prior research  
\* equations (may be old equation with slight tweak)  
\* diagrams  
\* images

Checklist/guide for writing a research paper for publication (where red items are most important)

* title
  + with your name followed by name of supervisor (me in this case)
* abstract
  + is an summary/overview of the entire paper (independent of the rest of the paper)
* introduction
  + goal / problem statement (what is your goal/what problem are you solving)
* background
  + (focused specifically on concepts/papers related directly to your work - not too general)
  + cite publications critically (be critical of prior research => mention limits/constraints/etc)
  + equations (you may be able to find many of these in other papers/text books)
  + diagrams (you may be able to find many of these in other papers/text books)
  + images (you may be able to find many of these in other papers/text books)
  + summarise limitations of prior research (the entire reason for doing this paper and also shows that you can read critically)
* solution/method
  + how your solution overcomes limitations of prior research

equations (may be old equation with slight tweak)

* + diagrams
  + images
* results
  + must be quantified - i.e. numbers, not just images (this is science/engineering research, not arts research)
  + graphs and/or tables
  + images
  + limitations of your research (gives more credibility)
* conclusion
  + results (very brief i.e. one sentence summing up results)

how your research improved on (or compares with) prior research (quantify and cite the prior research you are comparing with)

* + limitations of your research (very brief)
  + future research to overcome these limitations
* references (around say 15-25 references)
  + recent (at least a quarter in the last few years - shows you did a recent literature review)
  + quality (at least half should be journals/conferences – preferred over websites)
  + if websites are referenced, date (month/year) the last time you accessed it
  + reviewers are really fussy about consistency (such as does the year always come last, etc)