Field Relativity and Recognition of American Football Formations and Plays

William Johanson
Department of Computer Science and Software Engineering
University of Canterbury
Christchurch, New Zealand
wpj19@uclive.ac.nz

Professor Richard Green
Department of Computer Science and Software Engineering
University of Canterbury
Christchurch, New Zealand
richard.green@canterbury.ac.nz

***Abstract* - The power of creating sports datasets to analyse trends and patterns, of your own team and the opposition, is a large area of recent development. In a sport such as American football, with its commercial and popular success, being able to develop more and analyse data to a deeper extent can give your team the winning edge. By utilizing algorithms such as Canny edge detection, probabilistic Hough transforms, contouring, thresholding and averaging, a simple video of a game of American football can begin to be analysed by a machine. The above algorithms were successfully implemented in the following program to identify the on-field players, to identify the field markings throughout a play with 58% accuracy and to locate the line of scrimmage.**

I. Introduction

'Player tracking datasets have been a game changer for a number of sports over the past decade. Knowing the exact location of each player on the field as well as the ball is extremely powerful' [1] No sport is more involved in using analytics to gain an edge 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 utilizing the powers of computer processing [2], 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, as per the example 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).

By identifying the players positions, it can be relatively mapped 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.

II. Background

A. Existing Products

To understand the requirements of the project, two areas required investigation; existing sports data products and existing sports data research. Through analysis it can be determined where the value lies within this research project.

Firstly, are sports analytics companies. Predominantly, *Pro Football Focus* (*PFF*) [3] dominates the American football data analytics market. *PFF* analyse film taking upwards of 10 hours per game to glean critical data to allow for their analysis. For perspective of such a manual workload, there are up to 65 FBS college football games and 16 NFL games a weekend during the season to analyse. Therefore, the value in harnessing the time saving capabilities of computer processing 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 analyzing the dataset, which is where the ability arises to find trends.

Further companies have begun research into computer vision in sports and American football specifically. Utilizing computer vision is still in the development stages across all sports with companies such as *Sentio Sports Analytics* [4], which specialize 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* [5]. *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 [6] are invested in neural networks investigating 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.

B. Existing Research

An existing research paper [7] investigated automating NFL film study. The goal of the paper was to extract key player information from video of a football play. The implementation of their research would allow for efficient tactical scouting of opposing NFL teams. The paper primarily investigated field reconstruction and player recognition. The research into the two areas is an initial development of determining what the camera is looking at. By defining the frames coordinate systems, the field features could then be defined relative to a global coordinate system. The field features can be defined by the specific measurements of a football field. The field features that are pronounced on the field are the vertical yard lines and horizontal hashmarks which are primarily investigated to reconstruct the field.

The Sobel edge detector filter was utilized on a greyscale version of the frame to expose the edges on the field which is prominently the yard lines. The edges were then run through a standard Hough line transform to detect the yard lines from the edges as the yard lines are the longest prominent edge on the field.

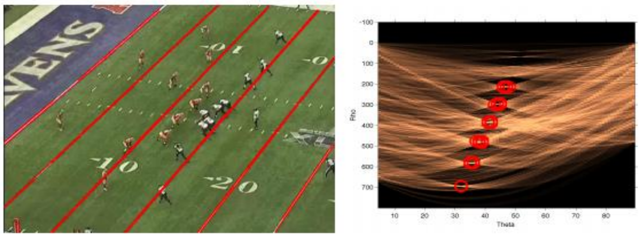

Figure 2: Identifying the field yard lines.

The hashmarks were done similarly but the color space was converted to the L*a*b* colour space. The colour space separates the colours in a way that is perceptually motivated by the human visual system. The luminance portion (L) is filtered and a blob detector extract the bright white spots on the field. The spots, which are in effect the hashmarks, are then passed to a Hough transform which joins the hashmarks together.

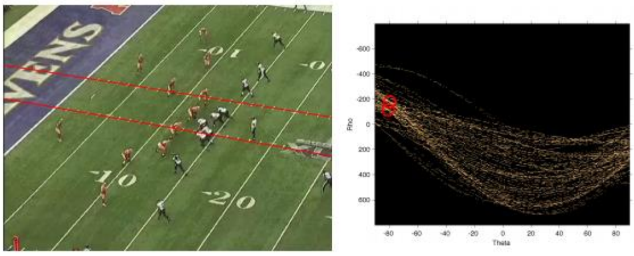

Figure 3: Identifying the field hashmarks.

The weakness of the field coordinate system is the ability to determine the position of the camera coordinates with respect to the field. The only on field coordinates predominantly in the z-direction (upwards) are the goal posts. Since the goal posts are not always visible this isn't an effective reference. The research investigated using an affine transform from the yard lines and hashmarks. However, further development defined this as an area requiring improvement.

To track and identify players on the field, the researcher determined the most logical method would be to utilize the unique jersey colours of each team. Using the jersey colours would allow the players to be tracked in every frame throughout the play as an alternative to point clouds consisting of Harris features for each player. The issue with Harris features is that the players are always slightly different. Whether that difference be in terms of body position, relativity to the camera or different players of different sizes. The further advantage of using uniform colours is also the colour of their pants which can either be the same colour as the jersey or different. Therefore, large masks for each player occur in the binarised image by colour thresholding. A large mask allows for morphology to occur which removes the 'salt and pepper' noise in the frame mask. If the pants are a different colour to the jersey, using the difference between the different colour blobs also allows players to be identified in the image with ease. The research identified that false positives can occur with the use of this method. Notably, with the field painted logos being within the colour thresholds and when the frame includes the sideline and crowds. In contrast to the affine transforms for the field coordinates, the affine transform of the player positions proved to be capable of mapping the normalized field position for the players.

The challenges outlines within the research included the use of a single camera angle. Since the component that errors occurred in was related to struggling to determine the camera matrix, having multiple cameras or initially knowing the cameras position matrix would remove the error related to the field coordinate systems.

A second research paper [8] into the area of American Football film analysis pertained to the use of game film to identify and track the offensive players. The initial process was the same as the previous research paper, where identifying the cameras relativity to the field was done using Hough Line transforms. The paper also identified the offensive players. Their process involved identifying the line of scrimmage and using the jersey colours to identify the x and y coordinates of the players in relation to the quarterback. To improve the algorithm, a training set of 500 formation images was created. Various machine learning algorithms were then tested to identify which algorithm produced the most accurate results. The research identified the classification and regression trees algorithm to be 86.5% accurate. Using the classification and regression trees algorithm, the research recorded the algorithms accuracy at identifying different quarterback formations, with the results shown in Table 1.

TABLE I. CLASSIFICATION REPORT FOR QB POSITION

QB Position	**Precision**	**Recall**
Center	.82	.92
Shotgun	.90	.84
Pistol	.50	.12
Average	**.84**	**.85**

The limitations of the research project pertained to the limited training data for the machine learning algorithm. Since the research only investigated offensive players, their future work largely included replication of their progress for the defensive side. The data would then be able to be used to investigate relationships between offensive and defensive play calling.

Investigating game decisions was reported in a research paper which researched predicting points and valuing decisions in real time with NBA tracking data [9]. The research wanted to quantitatively evaluate the decisions made within a game. Using a factor called Expected Possession Value (EPV) to quantify each decision in a game based on the expected outcome. The observed occurrences in game can be analysed deeper than the box score statistics to create a deeper understanding of what is occurring in the game. The limitations of the research was the lack of data to be able to accurately quantify the decision, as each decision needs to be trained on similar circumstances that have occurred in games prior to allow for an accurate EPV valuation. The benefits across all sports will be the implementation of coaching choices which are statistically proven to create the winning edge.

An alternative method of investigating game decisions is to identify trends within the teams coaching, a research paper investigated automatically recognizing on-ball screens in basketball games [10]. The current method is to utilize manual labour and review excessive amounts of film to identify the trends of the oppositions play calling. The research utilized machine learning which trained on data to identify instances of on-ball screens occurring in the film. The value of such implementations is the reduced labour cost to analyse game film to identify play trends of the upcoming opposition. The limitations of the research was due to the limited game film, since every team runs their plays with slight variations, there needed to be data to train the program on for all the teams and the slight variations they run of the same play to accurately identify the occurrences.

The common error arising from previous research in sports analytics field is the lack of data to allow for accurate results. A research paper proposed a method to implement large scale sports play retrieval [11]. A form of sports data search engine along the lines of *Google* or *Yahoo*. The implementation of a search engine for specific plays or occurrences will reduce the amount of data required to be stored, therefore allowing more data to be accessed per unit memory. Once an input query arises from the program the retrieval system can return a ranked set of similar plays. The retrieval system with a deep learning data training algorithm will increase the accuracy and efficiency to analyse and produce the trained data set.

C. *Areas of Improvement*

The current colour thresholding method to find the players still produces noise due to 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 companies such as PFF is to take the position of the players, simply and robustly detected, and then develop that into a new window orientated aesthetically to 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 utilizing specific positioning to get quantifiable values for further analytics.

III. Method

A. *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 (offence and defence).
  - Line of Scrimmage (LOS) placement.
  - Outputting key data points to informative end-user plots.
- Initially develop for individual frame screenshots (*.png*) for proof of concept and methodology.
- Further develop into frame by frame (video, *.mp4*) analysis and post-snap play recognition.

B. *Field Outline*

The play frame example used for the initial development of determining the field outline is shown in Figure 4.

Figure 4: Play example used for code development.

From Figure 4, the field markings are a clear white colour in stark contrast to the green playing field turf. Based on this observation, it was determined that using the probabilistic Hough Lines detection algorithm would return a set of data for the start and end points of the field marker lines with the correct filtering applied.

Previous research had utilized the Sobel operator for edge detection. However, Canny is widely termed the 'optimal edge detector'. The benefits of the Canny edge detector over the Sobel detector [12] include the additional steps of non-maximum suppression and the use of hysteresis thresholding along edges. The benefits include the removal of noise in the image, thus better edge detection in noisy environments.

The first stage of filtering for the Canny edge detector implementation was to apply the Gaussian Blur filter to the original image. The Gaussian Blur reduces the amount of image noise and detail in the image to prepare the image for edge detection. Since the white lines are large edges/features, the kernel size was chosen to be relatively large (9, 9).

The image is then gray scaled. The gray scaling of the image reduces the information for each pixel by a factor of three. The information is reduced, by the gray color having equal intensity for blue, green and red. Therefore, only one value needs to be specified for each pixel as opposed to three.

An edge is where change occurs, which is measured by the derivative. Thus, the largest change has a second derivative equal to zero. The gradient of an image is:

$$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right] \qquad (1)$$

The gradient points in the direction of the most rapid change in intensity with the edge strength given by the gradient magnitude:

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \qquad (2)$$

The Canny threshold values used for hysteresis, through trial and error, were 20 and 60. The detected edges of Figure 4 are shown in Figure 5.

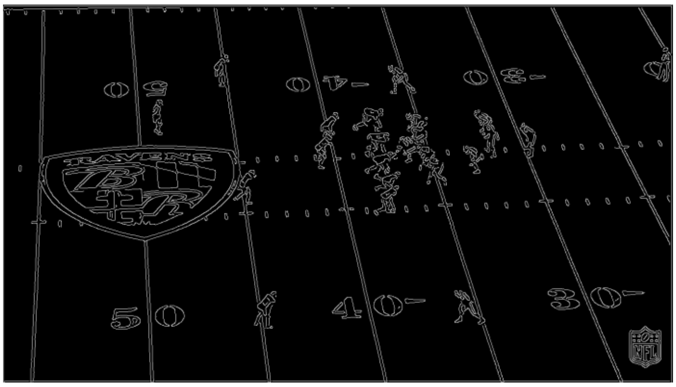

Figure 5: The Canny detected edges output.

With the edges in the image detected, a Hough transform was used on the image shown in Figure 5. The Probabilistic Hough Transform (PHT) is a mathematically correct form of the Hough Transform [13]. The PHT lines detection takes extra inputs in comparison to the normal Hough lines detection in the minimum line length and the maximum line gap. The minimum line length is useful in removing lines which are predominantly noise. The maximum line gap is an additionally useful feature for this application due to the players breaking up the edges of the field markings. The line gap allows the filter to work around the players and the probability of the edge being a line means the field markings are detected even in the presence of players in the frame. The values used for the minimum line length was 10 and the maximum line gap was set large to 1000 to ensure even groupings of players didn't affect the field marking recognition. The output from the probabilistic Hough lines transform is shown in Figure 6.

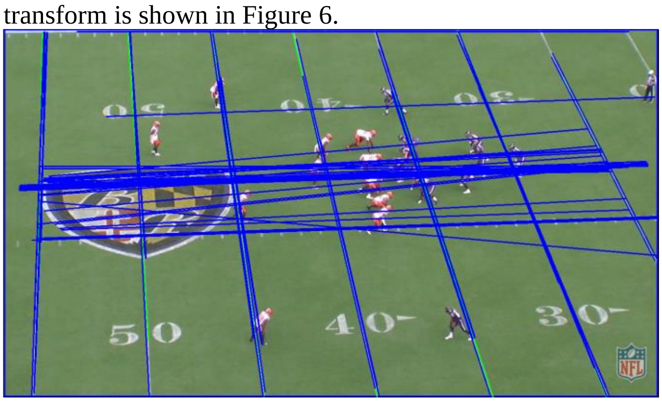

Figure 6: Hough line transform output.

The output shown in Figure 6 shows more lines than the desired field markings. Thus, orientation and frequency filtering were carried out on the lines. A further key feature of the PHT was that it returned a set of cartesian coordinates for the endpoints of each line in the form:

$$[(x_1, y_1), (x_2, y_2)] \qquad (3)$$

Slightly visible in Figure 6 are lines around the edge of the frame. The cause of these were put down to edge filtering near the edge of the frames and thus an exception was raised where if either $x_1 = x_2$ or $y_1 = y_2$, these lines were from the edge of the image and couldn't be field markings.

The field markings in the frame are closer to vertical lines than horizontal lines. To implement logic to remove the horizontal lines produced from the probabilistic Hough lines transform the array of line coordinates were filtered by the conditional statement:

$$|y_2 - y_1| > |x_2 - x_1| \qquad (4)$$

If the change in the y-direction was greater than the change in the x-direction the line was approximately vertical. The output shown in Figure 7 shows the resultant lines after the filtering.

Figure 7: Vertical field marker lines.

The noticeable error in Figure 7 is the false positive occurring at similar positions to the desired lines. The lines required further filtering to only have one line approximating each field marker.

To implement the filtering, the mean difference between x_1 coordinates of each of the desired field markings had to be calculated. Thus, when the lines coordinates were analysed, a new line flag could be set to True. While iterating through the array of lines already drawn, if the next list in the original list array fell below the expected mean distance between the lines with any of the existing drawn lines, the flag was set to false indicating the line wasn't in a new position that hadn't been marked. The line would then be discarded from the array and not drawn on the final field markings image. Thus, there is only a singular line for each field marking.

C. Player Detection Methodology

The method chosen within previous research for identifying players was to utilize the uniform colours and implement colour thresholding. Colour thresholding is the process of binarizing the image (each pixel is either 0 or 1 relating to black or white) if the pixels BGR colour falls within a threshold range. Four sets of high and low thresholds were then created in the BGR format for the offense and defenses, primary and secondary colours. From the example in Figure 4, the secondary colour of the offense and defense was the same approximate white colour.

To further improve the players being defined within the mask, the image was also converted to the CIELAB L*a*b* colour space. By thresholding a range of luminance values within the image, an additional mask could be created which further enhanced the players in the mask. The offense and defensive mask produced from Figure 4 using just the primary uniform colour thresholds are shown in Figure 9.

Figure 8: Primary colour threshold defensive mask (top) and the primary colour threshold offensive mask (bottom).

The final mask is the sum of the primary and secondary colour thresholds for the offense and defense, as well as the L*a*b* threshold mask shown in Figure 9.

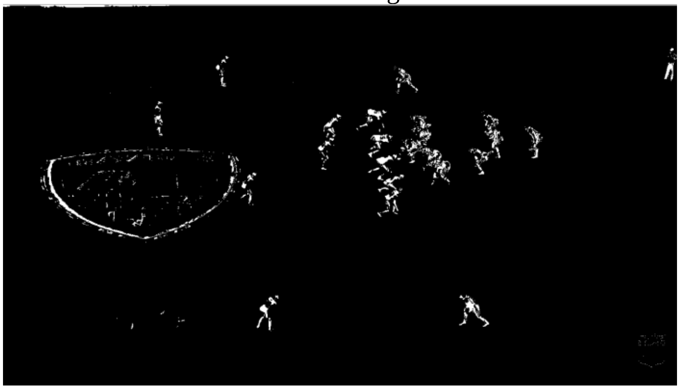

Figure 9: Final colour thresholding mask for player contouring.

The noticeable issues with the mask in Figure 10 are the visibility of the logos for both the team and the NFL in the bottom right corner, the regions of salt and pepper noise and the referee in the top right of the frame.
Before undertaking any filtering on the above image, the contours of the mask shown in Figure 9 can be produced for a reference on the impact of filtering as shown in Figure 10.

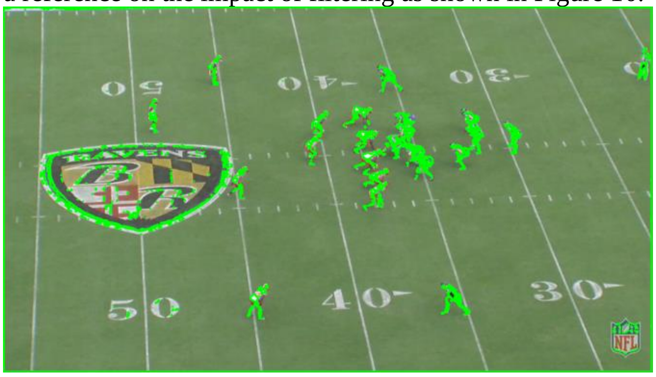

Figure 10: Contouring output of the unfiltered players mask.

The output of Figure 10 reaffirms that the mask contains false positive detections of the logos and the referee, whilst also displaying multiple contours on each player. Thus, the filtering needs to mould the contours on each player into one to create an effective ‘blob’ for each player to identify their coordinates in the frame.
The first stage of the filtering was to use adaptive thresholding. The threshold value used for the gaussian-weighted sum of the neighborhood values was 13. The large value allowed the close contours in Figure 10 to be joined together reducing the size of the set containing unique contours in the image.
Converting the thresholded image into a set of ‘blobs’ for each player required further filtering using morphology filtering. Morphology can either erode or dilate the binarised pixels in the image. The combinations of these two filtering effects are utilized primarily to reduce salt and pepper noise in the image, and additionally utilized here to reduce the detections of each player ideally into one single ‘blob’ to detect. The further advantage to the morphology is if the pixels are dilated enough, the field logo is ‘blobs’, much larger than any other detections in the image. Thus, simplifying the identification of the logo based on further contouring and easing removal. To define the parameters for the morphology filter, the kernel size as well as the opening and closing iterations must be defined. For the main image mask, the kernel size was selected to be 2×2 with 2 opening iterations followed by two closing iterations to give the output shown in Figure 11.

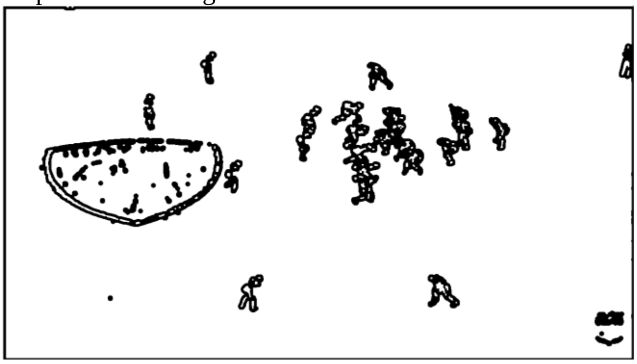

Figure 11: The morphed output of the main mask.

A larger kernel size would overlap the players in the mask which would remove the proficiency at identifying closely aligned players. However, for the NFL logo mask, which wanted to be removed the kernel was chosen to be 7×7 with 6 opening iterations and two closing iterations. The large kernel size and number of opening iterations ensured the logo became a solid blob as shown in Figure 13, which when it came to removal was fully defined as one single blob on the image.

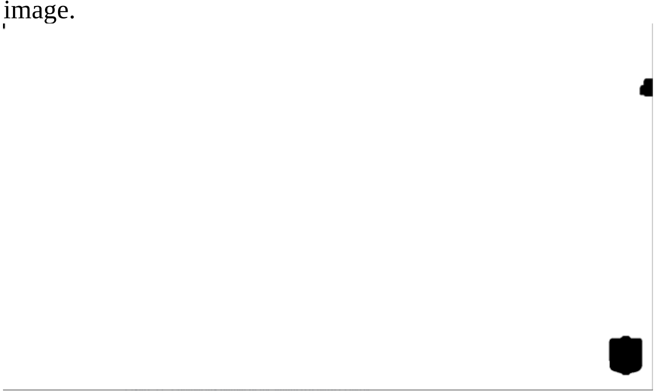

Figure 12: The morphed output of the NFL logo mask.

The contours (perimeter) of the two masks both the NFL logo and the original image were then identified. [14] The NFL logo contours were then processed first. Since it was the only set of values on the image mask, the identification is simple, and the coordinates are then saved to variables to ensure that contours are removed from the given coordinates in the main image.
The main image contours were then processed. However, their first cycle through the set of contour coordinates was to find the on-field logo. As seen in Figure 12, the logo is the largest set of connected besides the perimeter of the frame. Thus, finding the area outlined by the contours could then identify the logo contour. With identification of the logo contour, the four coordinates of the logo (a bounding box rectangle) could be saved to variables. On the second cycle through the main contour data set all the contours within the logo coordinates as well as the contour itself could be removed. The remainder of the contours could then be outputted onto the original method as perimeter contours or as rectangles to box the players identified. Another approach was to find the centroid of the contours which ideally, would be the centroid of each player to then simplify output to a plot of the players positions. This was then used in conjunctions with contours found for the primary and secondary colour masks used with the same filtering as described to create more centroid points to plot and filter.

D. Line Of Scrimmage (LOS) Determination

Identifying the LOS includes both the previous methods, line and player identification, to be able to accurately plot. Thus, the same primary colour thresholds used for player detection are defined, as well as the same probabilistic Hough lines detection parameters and methods used.
With the detected lines, the gradient of each field line must be calculated. This is made simple with the probabilistic Hough lines function as opposed to the Hough line transform as it returns the start and end point coordinates of the line. Therefore, the gradient can be calculated with:

$$M = \frac{y_2 - y_1}{x_2 - x_1} \qquad (5)$$

The set of gradients are then appended to an array. This array can easily be sorted with the python *sort()* function. The two central values of the array, if the array is even in length, or the three central values, if the array is odd, are averaged to calculate the average gradient of the field markers.
The players, using the same player detection method but with only their primary uniform colours are then detected. The LOS is the line between the two teams. Therefore, the players closest to the opposing team are also the players closest to the line of scrimmage. Hence, the coordinates of the further most right defensive player and the further most left offensive player is identified from the player contours. Using the identified coordinates and knowing the expected gradient of the LOS, two lines can be drawn from which the LOS will be in the centre of. To identify the end points for each of the line, from the given player coordinates of x_2, y_2 and the known maximum y value of 778:

$$\left[x_1 = \frac{x_2 - y_2}{M}, y_1 = 0\right] \qquad (6)$$

$$\left[x_3 = \frac{778 - y_2}{M} + x_2, y_3 = 778\right] \qquad (7)$$

With two lines now created from $(x_1, y_1) \sim (x_3, y_3)$ from the two players identified as being closest to the LOS, the actual LOS can be found. The LOS is simply in the centre of the two lines identified.

E. Outputting Key Data Points

To begin the journey into outputting key data and creating datasets for analysis, the findings need to be outputted in a means preferable for analysis. Whilst outputting over the original image with bounding boxes is aesthetically desirable and allows proof of methodology, the best means to output the data is by plotting the data against a set of coordinates in a consistent manner.
To output the field markings and the LOS, for reference of field location, the line coordinates can be used from the PHT output as well as the calculated LOS coordinates. The only adaptation for the purpose of consistency was to set the output from the perspective of being behind the offense. Knowing the origin of the image is in the top left-hand corner, this was a simple adjustment.
The player detections for the centroid of their primary and secondary colour thresholds can also be outputted to a graph from the same perspective. Whilst unfiltered this graph doesn't glean any critical information that can be used to identify formation. However, it is the beginning of allowing filtering on the found data points to identify the players consistently.

IV. RESULTS

With the implementation of the various elements of analysis occurring initially as independent code blocks, the methods were developed largely independently. Thus, in some cases, methods may not be compatible to produce accurate results without interference.

A. Development Environment

The project was completed in Visual Studio Code using the OpenCV 2.1 library with Python 3.8.1. The program input was in the form of *.PNG* files (1381 x 778) or *.MP4* files (1280 x 720, 30.09 frames/second). The development machine used the Windows 10 operating system with an AMD Ryzen 7 1700X eight-core processor, 3.4 GHz with 8 GB RAM memory.

B. Accuracy and Reliability

The accuracy of the field line detection method is shown in Figure 13 for a single frame.

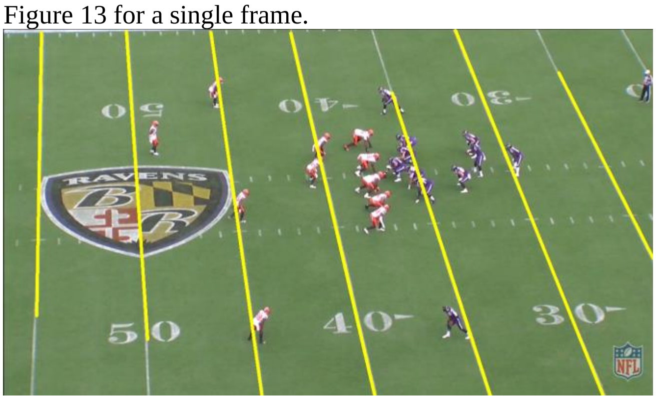

Figure 13: Final field line output for a single frame.

The output shown in Figure 13 shows an accurate positioning and only a single line for each desired field marking. When utilizing the same method over a video, the accuracy of the field markers detected compared to the field markers present is shown in Table 2.

TABLE II. FIELD MARKINGS DETECTED IN VIDEO SAMPLE SIZE OF 38 FRAMES.

Field Markings in Frame	Field Markings Detected	Percentage %
7	8	2.63
	7	57.89
	6	39.47

From Table 2, most frames detected all the field markings and not more than was in the frame. The cause of the missed field marking detection likely was due to the position of certain players breaking up a line. This could be an issue potentially solved with further identification and adjustment of the probabilistic Hough line transform filter parameters.

The player detection method could be outputted in multiple ways. For visualizing the results on the image, rectangle bounding boxes were used as in Figure 14.

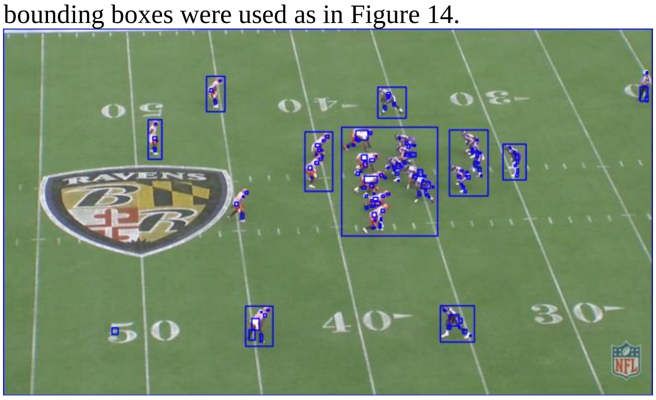

Figure 14: Players detected with their bounding boxes.

From Figure 14, the players have all been detected. Certain groupings of players also have an overarching bounding box. This could be useful for identifying formation based on select groupings of players such as the lineman around the line of scrimmage, and in the example of Figure 14, the two backs in line with the quarterback is a large indicator of the offensive formation. An alternative detection output was to use the centroid of the players detection from both their primary and secondary colour thresholds as in Figure 15.

Figure 15: Player colour threshold centroids.

Where red is the primary offensive colour centroids, green is the primary defensive colour centroids and blue is the secondary colour centroids which is the same for both sides. The output of Figure 15 shows a large set of data points for identifying players. From this output (which can easily be plotted to a graph as in Figure 16) the primary and secondary colours will be identified as certain points on the players uniforms, creating a set of expected points for each player.

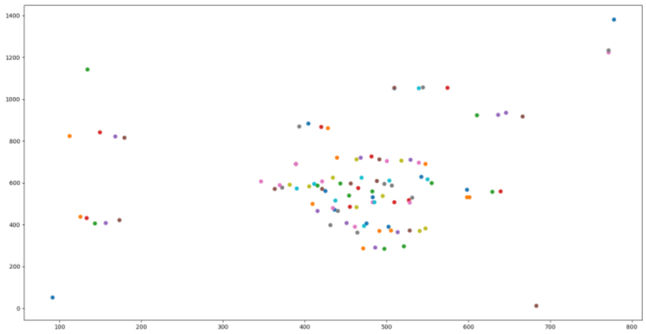

Figure 16: The player centroids (unfiltered) plotted against the pixel location.

Ideally, this could be used to map each player. The issue with this is the movement and obscuring of parts of players means identification of each expected point does not happen in every frame for every player.

The logo removal, for both the on field and NFL logo, were successfully removed from detection. The NFL logo was a removal that can be replicated no matter what the frame/team or field is. The field logo removal for the example frame shown in Figure 15 was successful. However, it is impossible to determine whether the method will be able to be replicated in all situations.

The line of scrimmage detection, while largely limited by successful player detections still has a successful output over the example frame used as shown in Figure 17.

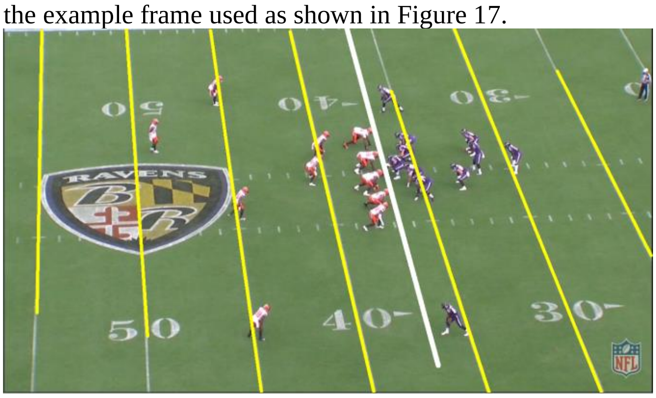

Figure 17: The Line of Scrimmage (white line) successfully placed in the expected position.

The improvements on the previous research in the field, particularly when it comes to the field line markers was in the ability to track the lines throughout the play. The previous research simply found a desirable frame to use as an example from which the lines could be identified. However, the development within this project showed with 58% accuracy, all the field markers within the frame could be identified throughout the play.

Of the two existing research documents into the area of American football data sets with the aid of computer vision, neither achieved both the field markers and the line of scrimmage detection for the same frame, simply each had one of the occurrences detected.

The output shown in Figure 17 plotted is shown in Figure 18.

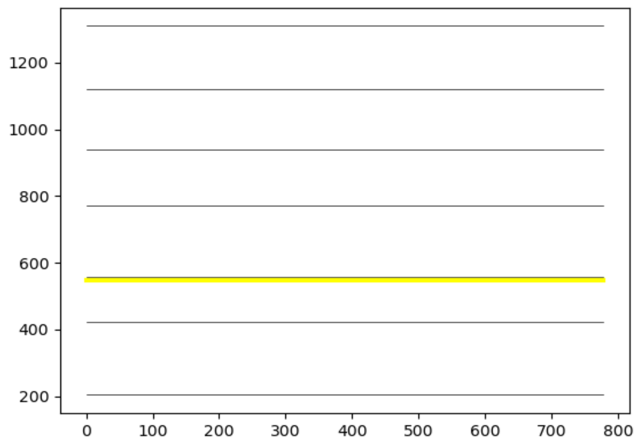

Figure 18: LOS (yellow) and field marker (dark) graphed output against the pixel position.

C. Limitations

The largest limitation of the LOS detection was the ability to usefully detect players to find the players closest to the opposition for each of the respective sides. False detections were attempted to be minimized using mask filtering with methods such as thresholding and morphology. However, false detections could still occur which would result in the LOS being offset from the true LOS in the frame. A potential solution for events down the line would be to only utilize the first frame of the play for field markings and LOS detection which is then stored for the rest of the play whilst only player detection occurs during the remaining frames. This will also add the further benefit of reducing processing time for each frame.

A limitation of the line detection method is the probabilistic Hough line transform returns lines that don't consistently give the coordinates of the very end points of the field markers within the frame. This results in lines that don't cover the full length of the field marker with no suitable explanation as to why the lines comes up short. Some lines re visibly interrupted by players breaking the edge of a line. However, from Figure 17, the leftmost line can be seen to not have any interruptions yet still come up short of the end of the visible field marker.

The final limitation was the inability to develop a neural network. A neural network is useful in training towards learning certain inclinations and traits of teams and/or players [16]. For this project it would be able to identify groupings of players and map the groupings to a certain formation. Ideally, this would have been expanded to not only identifying the formations but also the subsequent routes and movements run during the remainder of the play.

V. CONCLUSION

A. Summary

The results of the project showed automation of american football game footage data collection is feasible. The intital developments investigated and implemented in this project included the detection of the field markers throughout a play even with the motion of the camera. Using the players primary and secondary uniform colours aswell as the L*a*b* luminance, a mask was created of the players which could be filtered and turned into a series of points on the player or bounding boxes around the players. Utilising the player detection as well as averaging the field markers detected, the LOS could be calculated and implemented into the output.

B. Improvements Upon Prior Research

The improvements upon prior developments in this field revolved largely around a more robust implementation of the key concepts to begin dataset creation. Previous examples used distinct uniform disparities to be able to detect players simply with the colour thresholds. Even with the ideal uniform colours there was still detections which missed players or doubled up the detections. The findings of ths

paper showed improvement that each player could be identified even with issues in the frame such as logos and the camera angle causing player overlap.

The field marker detection was also very robust and didn't require any certain camera angles to be able to identify the field markers. The field markings could be detected throughout the play even with the camera motion, which included panning and zooming.

C. Limitations

The limitations of the project were the LOS and the field markers implementation. The field markers, while detected with reasonable accuracy and consistency, often did not stretch the length of the field marking in the output. The LOS implementation was limited by the false positives in the player detection. The false positives would mean when the LOS was implemented from the averaging technique, it would not be in the expected position and thus falsely implemented.

D. Future Improvements

The future improvements are largely due to expanding the capabilities of the detections. From the above example, for player detection a single uniform colour set was used for each team. However, of the 32 teams, each team will have multiple uniform combinations. Thus, all the colour thresholds of all possible combinations need to be stored to allow player detection with any teams playing.

An issue with the colour thresholding is the on-field logo. Whilst the NFL logo can be easily removed and the players being in its position on the frame being very unlikely, it is very likely a player will move into the on-field logo. As it currently stands, that player would effectively be removed from the detection. This is not an ideal occurrence and comes down to either a limitation of the colour thresholding method or an alternative means of filtering must be identified.

The final improvement from this project would be to implement a neural network and begin training the program towards learning all the possible formations f plays of each of the teams. From there specific data and trend analysis can occur which is the desired result of this project and where the possible industry benefits lie.

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.

REFERENCES

[1] N. Johnson, "Extracting Player Tracking Data from Video Using Non-Stationary Cameras and a Combination of Computer Vision Techniques," ESPN Analytics, 14th Annual MIT Sloan Sports Analytics Conference March 6-7, 2020.

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

[3] Pro Football Focus, Available: https://www.pff.com/

[4] Sentio Sports, https://www.sentiosports.com/

[5] SportsLogiq, https://sportlogiq.com/en/about-us

[6] M. Horton, "Learning Feature Representations from Football Tracking," Sportlogiq, 14th Annual MIT Sloan Sports Analytics Conference March 6-7, 2020.

[7] Timothy E. Lee, "Automating NFL Film Study: Using Computer Vision to Analyze All-22 NFL Film," pp. 1–9, Sourced https://cvgl.stanford.edu/teaching/cs231a_winter1415/prev/projects/LeeTimothy.pdf. Last accessed: 5/05/2020.

[8] O. Ajmeri and A.Shah, "Using Computer Vision and Machine Learning to Automatically Classify NFL Game Film and Develop a Player Tracking System" MIT Sloan Sports Analytics Conference, 2018 pp. 1-9 Sourced http://www.sloansportsconference.com/wp-content/uploads/2018/02/2012.pdf. Last accessed: 5/05/2020

[9] D. Cervone, A. D'Amour, L. Bornn and K. Goldsberry, " 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.

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

[11] M. Di, D. Klabjan, L. Sha and P. Lucey, "Large-Scale Adversial Sports Play Retrieval with Learning to Rank," ACM Transactions on Knowledge Discovery from Data, Vol. 12, No. 6, Article 69. Publication date: August 2018.

[12] D. Kim, "Sobel Operator and Canny Edge Detector," Fall 2013 pp. 1-10. Sourced: https://www.egr.msu.edu/classes/ece480/capstone/fall13/group04/docs/danapp.pdf Last accessed: 13/05/2020.

[13] R.S. Stephens, "A Probabilistic Approach to the Hough Transform," Sourced: http://www.bmva.org/bmvc/1990/bmvc-90-012.pdf Computing Devices Eastbourne Ltd, East Sussex. 1990, Last accessed: 13/05/2020.

[14] R. Szeliski, Computer Vision: Algorithms and Applications, Springer, 2011 pp. 270-284

[15] W. Gurpinar-Morgan, D. Dinsdale, J. Gallagher, A. Cherukumudi and P. Lucey, " You Cannot Do That Ben Stokes: Dynamically Predicting Shot Type in Cricket Using a Personalized Deep Neural Network," 14th Annual MIT Sloan Sports Analytics Conference March 6-7, 2020.