Intelligent Statistics Computation System for soccer games

Project Report

August 2017

Table of Contents

Project Summary

Project Objectives

- Design an action recognition and automatic statistics computation system for a soccer game

Project Description

- This project involves multiple object tracking using multiple cameras, developing and applying computer algorithms for action recognition for soccer games, and providing user interfaces for operation of software for the soccer statistics company.

Project Scope

- Multiple object tracking
- Action Recognition

Project Approach

- Multiple-kernel adaptive segmentation and tracking for robust object tracking
- State transition diagram for modeling
- Model-based heuristic learning for action recognition

1. Introduction

1.0 Background

Image processing and computer vision technologies have been widely utilized in team sports analysis and statistics collection. More powerful automatic systems and tools for information retrieval of a team sports video have been developed to replace some of the manual work of statistics collection for a team sports game. Furthermore, the automatic observation system can provide real-time feedback of the game such as performance of the team and each player, which is much more timely than manual interpretation. These statistics can be further used in evaluating each player, decision making and tactics.

For soccer game analysis, building an automatic system for statistics collection is challenging because of the complex rule, player and ball occlusion and sophisticated situations during the game. One of the most important parts is to obtain each player's and the ball's position, trajectory and speed. In addition, determining the player who possesses the ball is another key point and difficulty, since we can implement action recognition based on this information, and further interpret all the statistics for the team and player.

In this project, there are four static cameras at the four corners on the soccer field and eight players for each team. The input is the videos from these four cameras recording total sixteen players as well as the ball on the soccer field. The output is the computed statistics such as short and long pass, shots, possession rate for each team and player. Furthermore, an established software and interface has been developed so that our system can be used by other people in a convenient and user-friendly way.

1.1 Project Flowchart

The flowchart of this project is shown in Figure 1 below. Section 1.2 will give a brief description for each step.

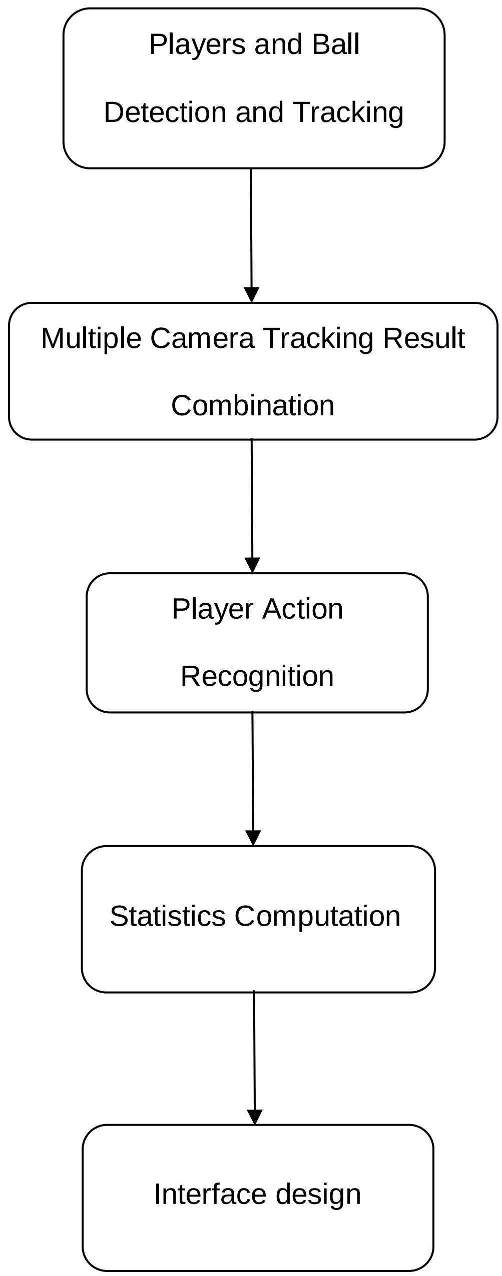


Figure 1 Project Flowchart

1.3 Project Description

This project will provide a robust detection and tracking system to all players and the ball, the tracking result will include positions, speeds, sum of distances and trajectories. Multiple-kernel adaptive segmentation and tracking(MAST)[1] is utilized in the detection and tracking phase. Multiple cameras' tracking result will then be combined to optimize and smooth out the final tracking results. A model-based method using heuristic learning, will be implemented to solve player action recognition after tracking. And a user-friendly interface will be built with all computed statistics based on player action recognition.

The rest of the report is organized as follows. Section2 will briefly introduce MAST detection and tracking method and the algorithm to combine multiple cameras tracking results. Feature-based and model-based action recognition methods and their simulation results will be both covered in Section3. The user interface will be presented in Section4.

2. Detection and Tracking

2.0 MAST

This project utilizes the method called Multiple-kernel adaptive segmentation and tracking(MAST) to do object detection and tracking, which dynamically controls the decision thresholds of background subtraction and shadow removal around the adaptive kernel regions based on the preliminary tracking results [1]. Players are often occluded by each other and they will have same colors if in the same team. Furthermore, the ball might also be partially occluded by players, making it ambiguous to have a robust tracking result. Multiple-kernel has been widely used in object tracking to deal with occlusion. That is the main reason why it is used in this project in the detection and tracking phase. The details of this method won't be covered in this report since the main focus of this project is on the following player action recognition part. After implementing MAST, the 2D global position of each player and ball can be obtained for each frame in videos from all four cameras.

2.1 Tracking Result Combination

There are many occlusions and ambiguity in a soccer game, both for players and the ball. For instance, the player might be far from the camera, or the player is occluded by another player, or the ball is occluded and missed. Gathering all tracking information from four cameras will significantly reduce the probability of ambiguity and missing or unclear detection and tracking. To combine the tracking results from four cameras, three important factors are considered simultaneously to determine the importance or weight of each camera.

$$W_{camera} = w_{depth} \times w_{visibility} \times w_{border} \quad (1)$$

$$w_{depth} = \frac{1}{d_{obj-camera}} \quad (2)$$

$$w_{border} = \begin{cases} 1.0 \\ \left(d_{obj-border} / \theta_{border} \right)^2 \end{cases} \quad (3)$$

As shown in (1), W_{camera} denotes the weight or importance of a camera. It is composed of three factors, where w_{depth} represents the reciprocal of the distance between the object and the camera as given in (2), $w_{visibility}$ represents the overlapped ratio of the target object's bounding box with its closest object's bounding box, and w_{border} is an index of checking whether the object is far enough to the border. When the object is close to the border, there will be more ambiguity than other cases because of the border line, so w_{border} is set to 1.0 if the object is far enough from the border, and the threshold θ_{border} is set to 0.007, if the distance between the object and border $d_{obj-border}$ is smaller than θ_{border}, then w_{border} is set to be $\left(d_{obj-border} / \theta_{border} \right)^2$.

After computing weights of all four cameras based on (1), the final position of all players can be optimized by adding the product of each camera's position data and the weight of that camera. This can significantly improve the accuracy of tracking results by eliminating those samples which have high occlusion, ambiguity and far distance.

3. Player Action Recognition

3.0 Model-based Method

The objective of this project is to realize an automatic system which can recognize the player action and compute the statistics for each team and player. For soccer games, typical player action includes shooting, short or long passing, dribbling, possessing the ball.

Using optical flow features to discriminate different actions, the feature-based method, was tried first to solve player action recognition. However, it was not very accurate and sufficient due to low image quality of the input data and ambiguous model between different actions of the player.

A model-based method, which is specifically fit for the soccer games, is then developed and implemented for this project to solve the player action recognition problem. For soccer games, the ball plays a key role in all player actions, whether it is shooting, passing or dribbling. More specifically, player action can be detected as long as the information of who possesses the ball is provided. And this information is not very hard to detect since the positions, speed of all players and the ball can be obtained from the tracking results.

3.0.1 Problem Modeling

To model the soccer game, a state transition diagram of the ball is designed. There are three ball states in the soccer game: Pass, Possession and OutOfField. The detailed definition of these three states is as follows:

- The ball is in **Pass** state when it is not within the reachable position of a player and is in the field or it is within the reachable position of a player but has a high velocity

- The ball is in **Possession** state when it is very close to one player or it is within the reachable position of a player and has a relatively similar velocity with the closest player
- The ball is **OutOfField** state when the ball is kicked outside the field

After all the ball states is defined, the soccer game can be modeled into a state transition diagram as given in Figure 2.

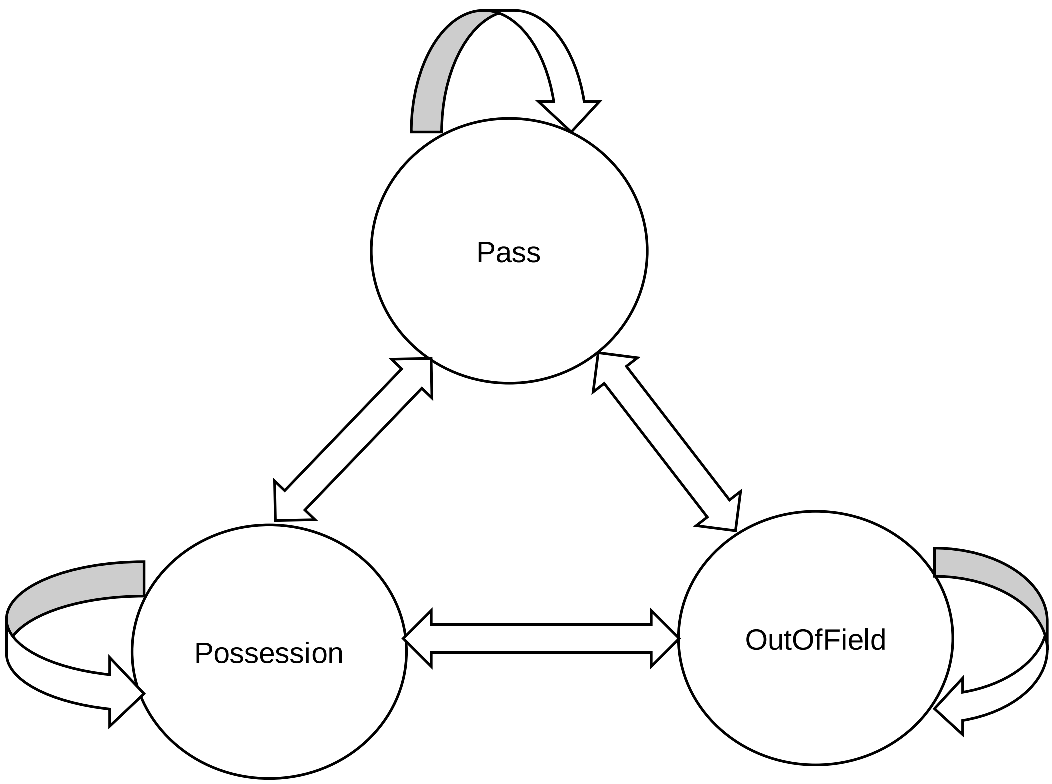


Figure 2 Ball State Transition Diagram

Player action can be recognized if the ball state changes. For instance, if the ball state changes from Pass to Possession, the player who is closest to the ball should possess that ball. More cases will be covered in Section 3.0.3.

3.0.2 Architecture

The architecture of this model-based method for player action recognition can be summarized as Figure 3:

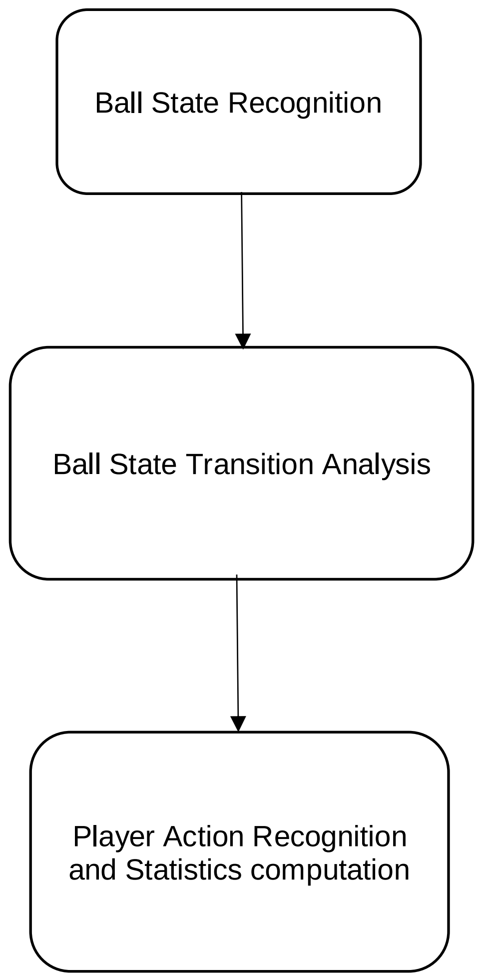


Figure 3 Architecture of Player Action Recognition

The detail implementation of each step in this architecture will be covered in Section 3.0.3.

3.0.3 Implementation

3.0.3.1 Ball State Recognition

A decision tree is used to detect the ball state, as given in Figure 4:

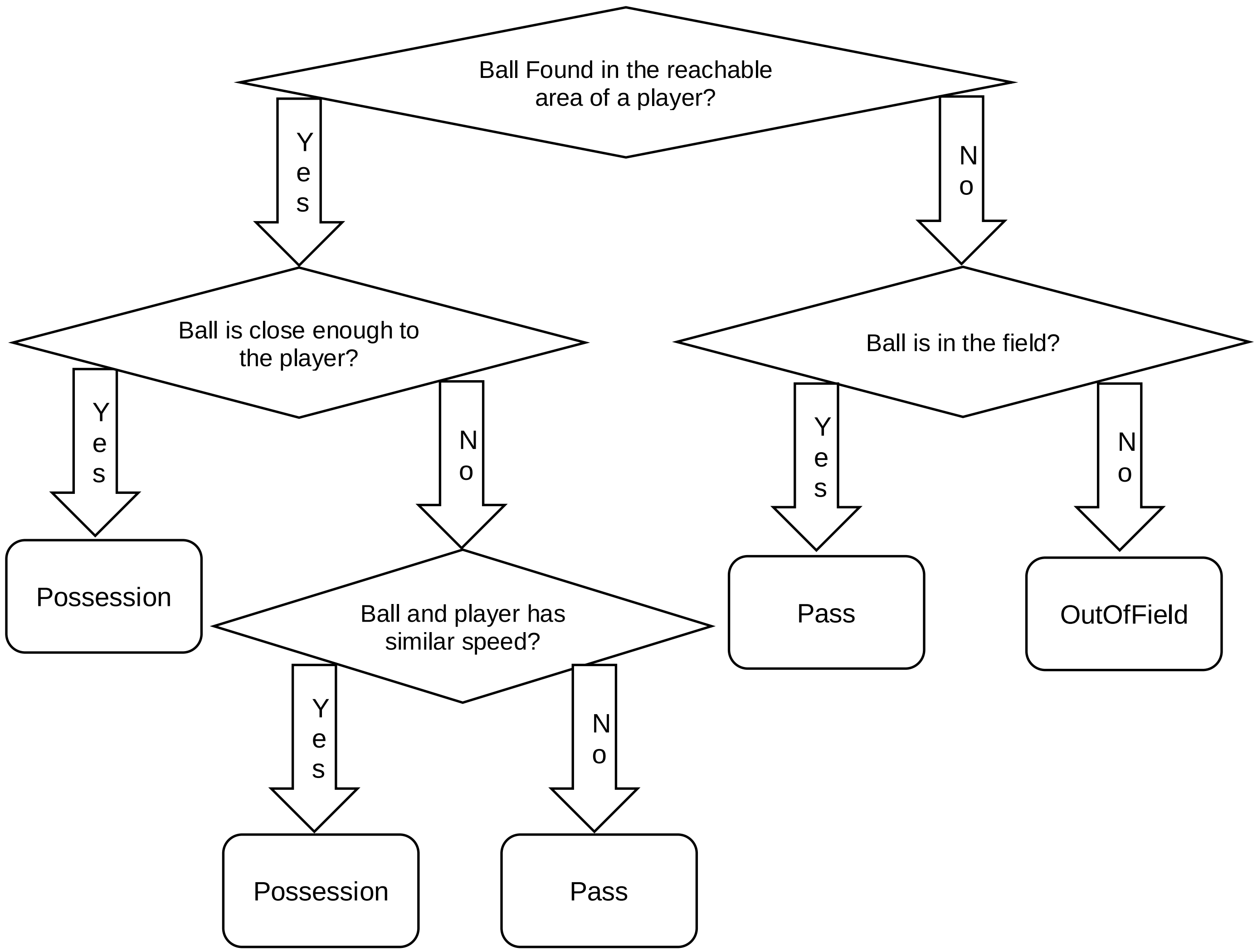


Figure 4 Ball State Recognition Decision Tree

All the thresholds in each node of the decision tree are learned by manually testing. After the ball state in each frame being detected, player action and statistics can be obtained based on the information from ball state transition.

3.0.3.2 <u>Statistics Computation</u>

- Possession rate for each team

  Whenever a Possession state for the ball is detected, add one frame time slot to the player's team who are possessing the ball.

- Short and Long Pass for each team and player

  To detect a pass, the state transition of the ball should be Pass to Possession. The player's ID and position who previously possessed the ball is recorded.

  If a new Possession state for the ball is detected, meaning that if the current and previous possessor are different, then:

  - Subcase1: If the current and previous player are in the same team, and the distance between them are large enough, add one success long pass to the team and the previous player
  - Subcase2: If the current and previous player are in the same team, and the distance between them are not large enough, add one success short pass to the team and the previous player
  - Subcase3: If the current and previous player are not in the same team, and the distance between them are large enough, add one failed long pass to the team and the previous player
  - Subcase4: If the current and previous player are not in the same team, and the distance between them are not large

enough, add one failed short pass to the team and the previous player

- Shots

  Shots can be considered as a special passing, and in most cases shots happen around the goal area and the ball's velocity is relatively higher than passing. The player's ID and position who previously possess the ball is recorded, the average velocity of the ball is also recorded.

  If the previous player who possessed the ball is around the goal area, the average speed of the ball during passing state is large enough and the ball is possessed by the goal keeper, then add one shot to the opposite team of the goal keeper's.

- Dribbling

  To detect a dribbling, the state transition of the ball should be Possession to Possession.

  If the ball is close enough to the player and player has relatively similar speed with the ball, that player is detected to be dribbling.

- Goals, Free Kick, Corner Kick, Penalty Kick

  These statistics can be computed using the same method because the ball will be put on a specific position for a certain period of time when these happen in the soccer game. The equation (4) is utilized to compute all these statistics:

$$S_i = \begin{cases} 0 \\ S_{i-1} + 1 \end{cases} \quad (4)$$

S_i denotes the consecutive frames at frame i where the ball's position changes within a threshold $\theta_{same-position}$, which is determined by manually testing. S_i is set to be $S_{i-1} + 1$ if the ball's position change is smaller than the threshold $\theta_{same-position}$, and S_i is set to be 0 if else.

If S_i is detected to be larger than a time threshold $\theta_{same-time}$, which is also determined by testing, we can compute goals, corner kick, penalty kick and free kick according to the positions the ball is located on.

3.0.4 Simulation Results

After tuning the parameters and thresholds in each step of player action recognition, the precision and recall of detecting actions and statistics computation has been significantly improved as given in Figure 5, the input is a two minutes soccer game video with 3000 * 4 frames from the four cameras.

Actions	Total	Detected	Precision	Recall
Short Pass	35	27	77.1%	87.1%
Long Pass	7	6	85.7%	100.0%
Shots	2	2	100.0%	100.0%

Figure 5 Precision and Recall Results

For other statistics such as goals, corner kicks, penalty kicks, the sample video lacks these cases so the accuracy of the model-based method cannot be tested. But these statistics will be detected with high accuracy as long as the tracking results are not too worse.

4. User Interface

After tracking, action recognition and statistics computation, a user-friendly interface is developed to include all parts together so that it can used by other people in the company.

4.0 Team Statistics

As Figure 6 shows, statistics such as Goal, Penalty Kick, Shots, Possession Rate, Passing, Passing Rate, Free Kick and Corner Kick is collected for both team. The length of color bar for each team represents the value of each statistics, and the larger value will have longer bar.

Figure 6 Team Statistics User Interface

4.1 Individual Statistics

4.1.1 Individual General Data

General data for each player includes Total Running Distance, Max Speed, Rushing Times and Possession Time, as given in Figure 7. Users can switch and check all players in different team by clicking the "Team1" and "Team2" button.

Player	Total Running Distance(m)	Max Speed(m/s)	Rushing Times	Possession Time(s)
1	52.424644	2.395862	0	0.760000
2	148.589813	4.863122	3	0.040000
3	139.414078	4.284173	2	2.399999
4	195.447769	5.982901	8	2.279999
5	202.008148	4.487658	9	0.680000
6	164.766006	4.682640	6	0.920000
7	149.984695	5.383352	4	4.599997
8	213.454559	4.261814	9	2.119999

Figure 7 Individual General Data

4.1.2 Individual Passing Data

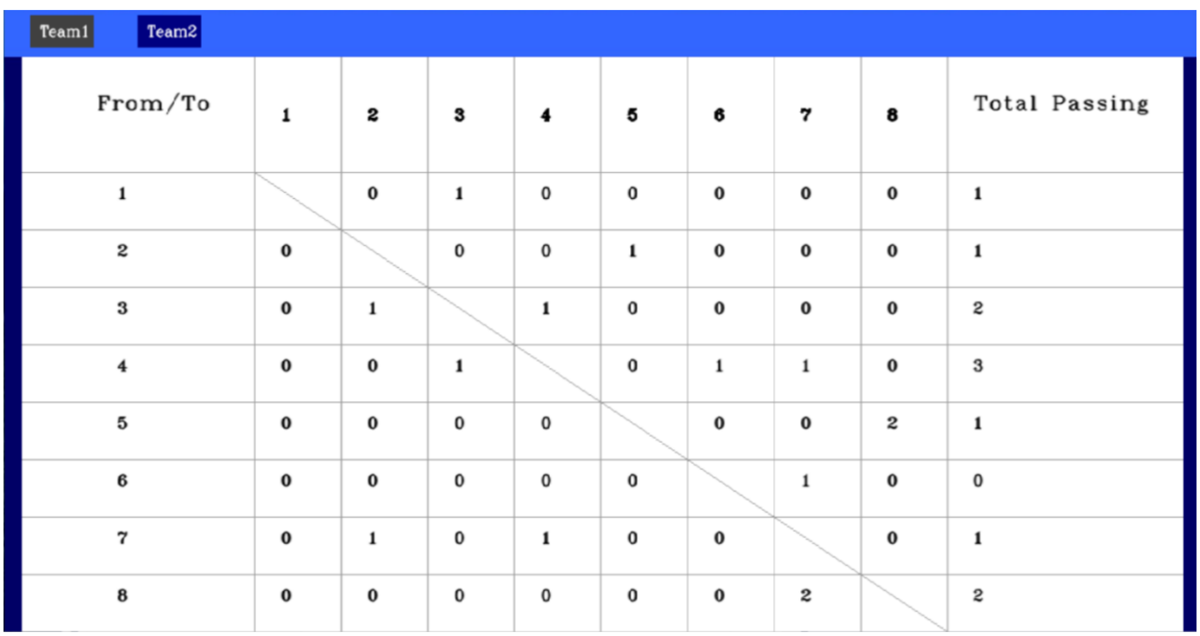
Passing data for each player includes Total Pass Times, Total Pass Success, Short Pass Times, Short Pass Success, Long Pass Times and Long Pass Success, as given in Figure 8. Users can switch and check all players in different team by clicking the 'Team1' and 'Team2' button.

Player	Total Pass Times	Total Pass Success%	Short Pass Times	Short Pass Success%	Long Pass Times	Long Pass Success%
1	1	100.00%	1	100.00%	0	-
2	1	100.00%	0	-	1	100.00%
3	2	100.00%	1	100.00%	1	100.00%
4	3	100.00%	3	100.00%	0	-
5	4	25.00%	4	25.00%	0	-
6	1	0.00%	1	0.00%	0	-
7	2	50.00%	1	100.00%	1	0.00%
8	2	100.00%	2	100.00%	0	-

Figure 8 Individual Passing Data

4.2 Passing Analysis

Users can check the passing analysis between each player in both teams in this page as shown in Figure 9. Click the 'Team1' and 'Team2' button to switch the team.



From/To	1	2	3	4	5	6	7	8	Total Passing
1		0	1	0	0	0	0	0	1
2	0		0	0	1	0	0	0	1
3	0	1		1	0	0	0	0	2
4	0	0	1		0	1	1	0	3
5	0	0	0	0		0	0	2	1
6	0	0	0	0	0		1	0	0
7	0	1	0	1	0	0		0	1
8	0	0	0	0	0	0	2		2

Figure 9   Passing Analysis

5. Conclusion

5.0 Summary

This project builds an automatic software and tool to compute soccer game statistics from video recorded by multiple static cameras. There are three important steps thought the process, detection and tracking, player action recognition, interface design.

Multiple-kernel adaptive segmentation and tracking(MAST) method is utilized in detection and tracking to solve occlusion problem, the tracking results from 4 cameras are then combined using weighted factors based on distance, visibility and ambiguity.

A model-based method is implemented in player action recognition, the game is modeled into a ball state transition diagram, all the player actions and statistics are detected and computed based on ball's state transition. The simulation result indicates that this method has a high accuracy both on precision and recall.

Finally, a user-friendly interface is developed to show all obtained statistics for further analysis.

5.1 Future Works

There are many more statistics for a soccer game such as shooting by head, robbing. These statistics can be easily added to the existing system and tested by future demo videos.

The model-based player action recognition method is highly depended the accuracy and robustness of the previous tracking results. If the image quality is enhanced in the future, then the feature-based method such as using optical flow to help discriminating different player actions can be used to further improve the accuracy of player action recognition and statistics computation.

6. References

[1] Zheng Tang, Jenq-Neng Hwang, Yen-Shuo Lin, “Multiple-kernel adaptive segmentation and tracking (MAST) for robust object tracking, Acoustics”, *Speech and Signal Processing (ICASSP), 2016 IEEE International Conference,* no 16021221, 2016.