ECE 457: Digital Video Processing

Final Project

Player Location Map generation for the game of Badminton

- Siddharth Patki

About the sport and my motivation

- Popular indoor racquet sport
- I was playing professionally for 6 years.
  Still love to play.
- Position of player on court is very important aspect of the game. - **Tactical analysis** of the game can be done with that information.

Sample Video

INTRODUCTION

The position of the player on the court during a rally is one of the most important aspects of the game.

Analysis of the types of shots, the positions on the court they are executed from and the success rates form the basic framework of tactical analysis in badminton.

In this project, I propose a fully automated system which generates the heat maps for rallies in the game given a recorded video of international Badminton World Federation sponsored match.

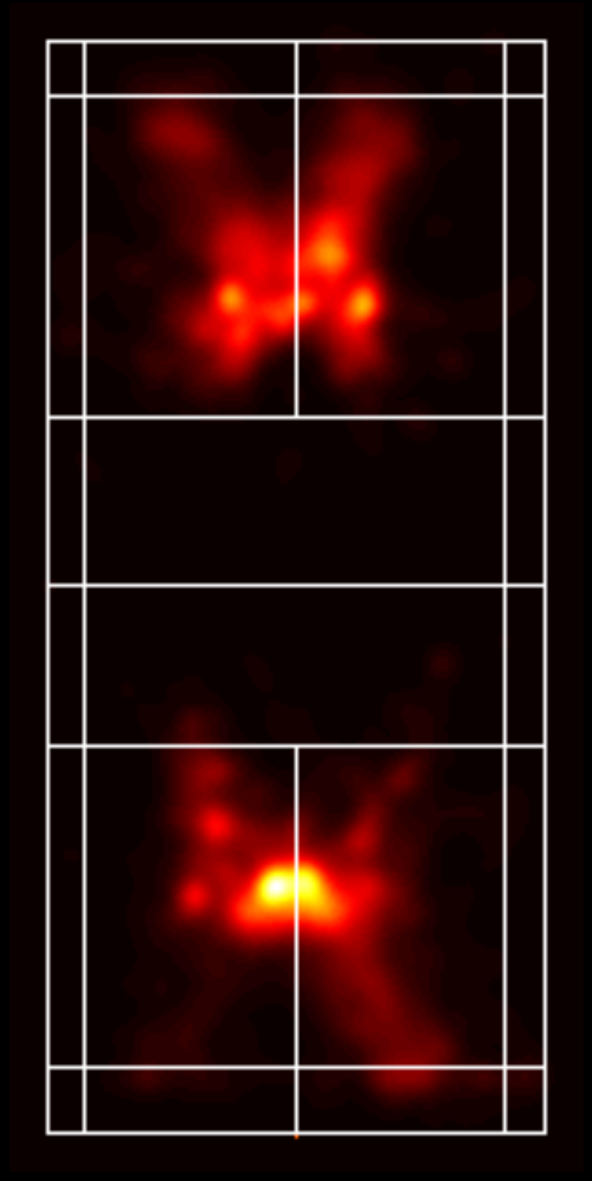
APPLICATION

Besides analyzing international matches, this information could be very useful in training facilities, where keeping watch on multiple players practicing on multiple courts is practically impossible task. By observing the generated heat map, the coach can conclude if the specified practice was done appropriately or not.

PROBLEM STATEMENT

To generate a position map (heat map visualization) of the player location on the court over rallies and complete game.

MATCH VIDEO



To do list

1. Generate approximate background image of the court.
2. Detect and track player real world player location.
3. Generate location map over rallies.

Block Diagram

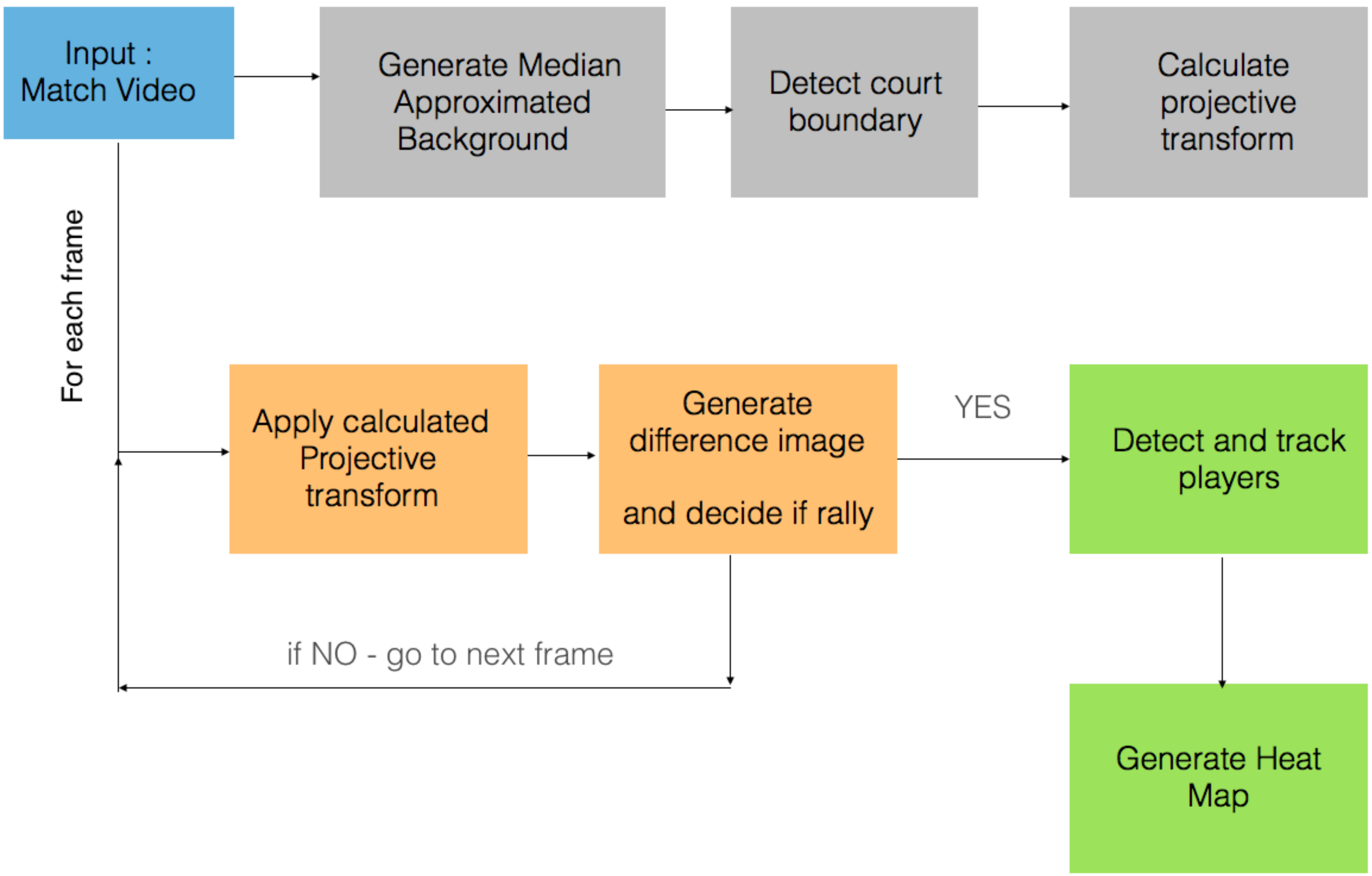
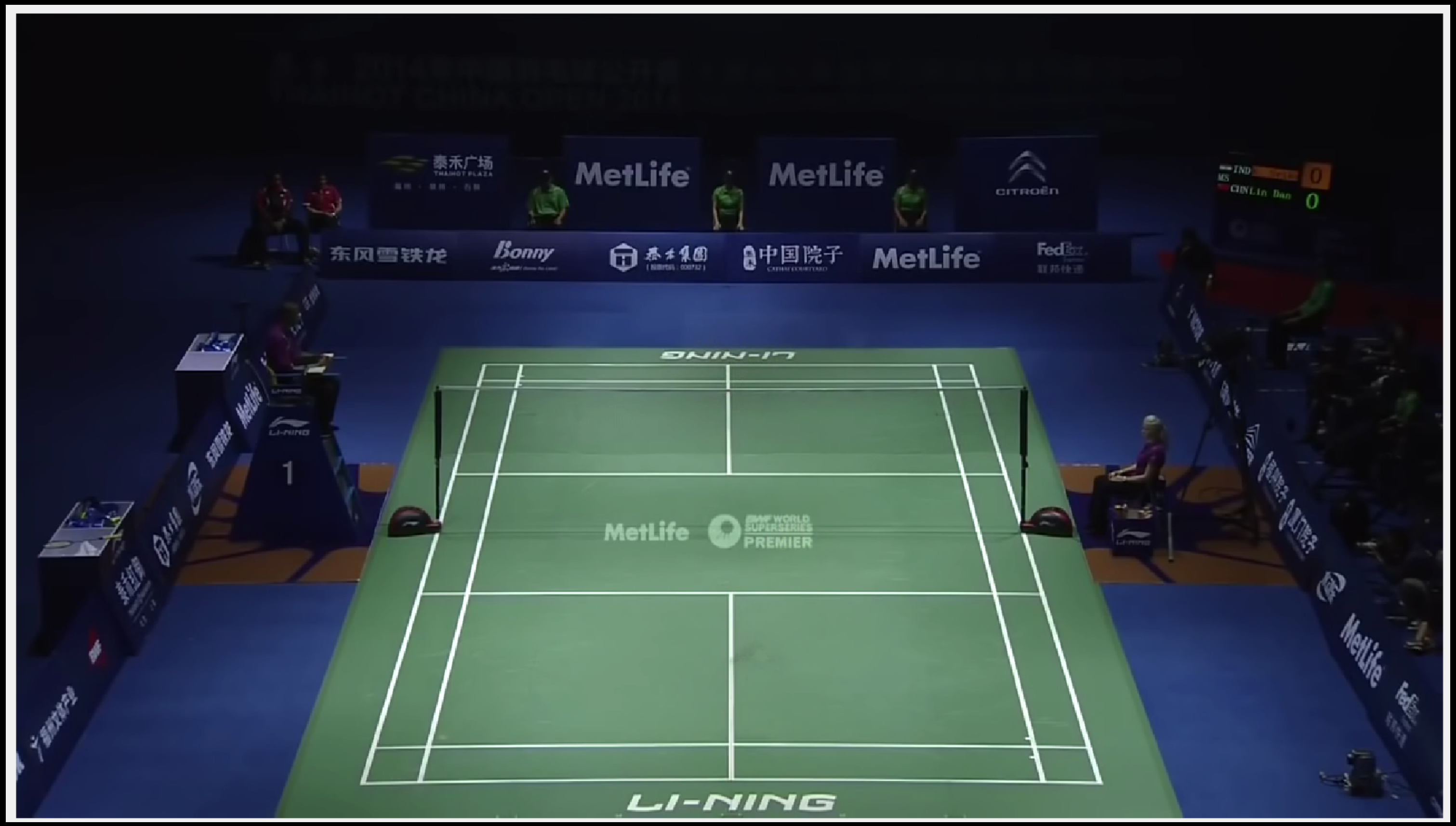


Fig. 1. Block diagram of system.

Generating approximate background

- Players are detected based on background subtraction.
- Median frame generation was used to estimate background image.
- In this technique, **the temporal median value for each pixel is considered.**



Detecting Court corners

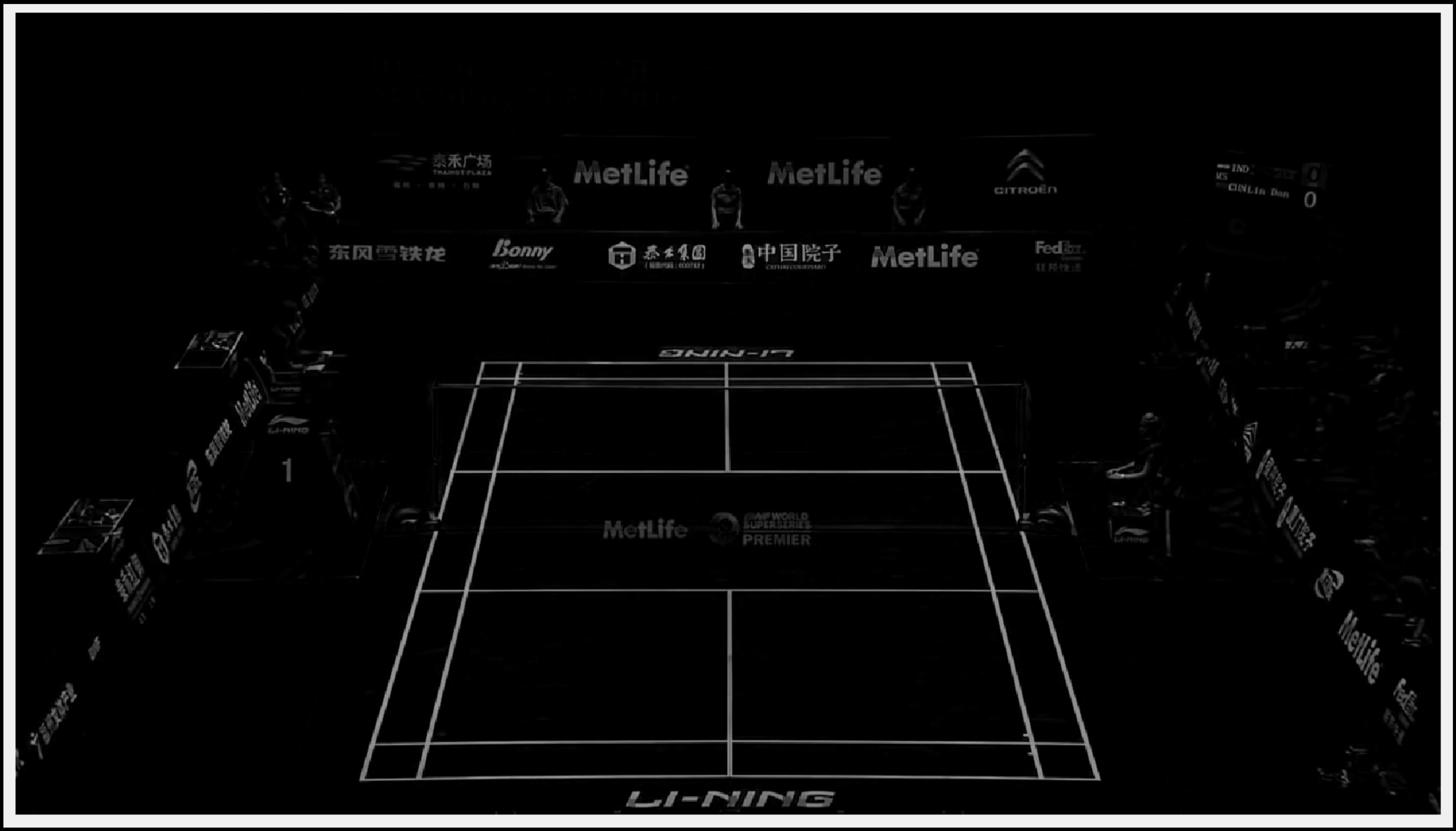
- In order to find the exact location of the player’s feet on court, the court plane needs to be transformed.
- To estimate the transform matrix, the four corner points of the court needs to be detected.
- Series of morphological operations are performed on background image in order to locate court.

Detecting Court corners -

Convert to grayscale

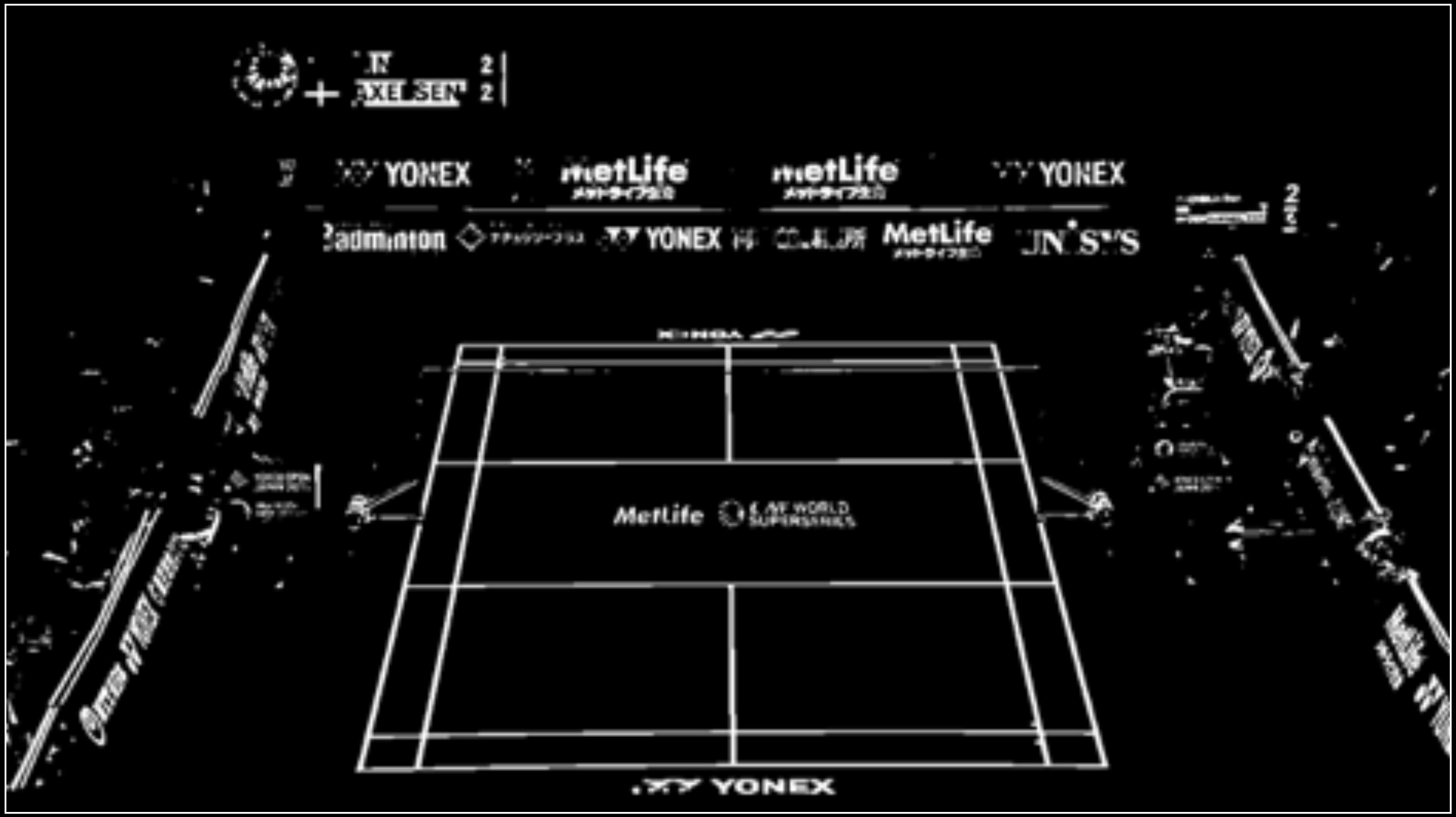
Detecting Court corners -

TopHat transform with square shaped structuring element



Detecting Court corners -

Convert to binary using low threshold

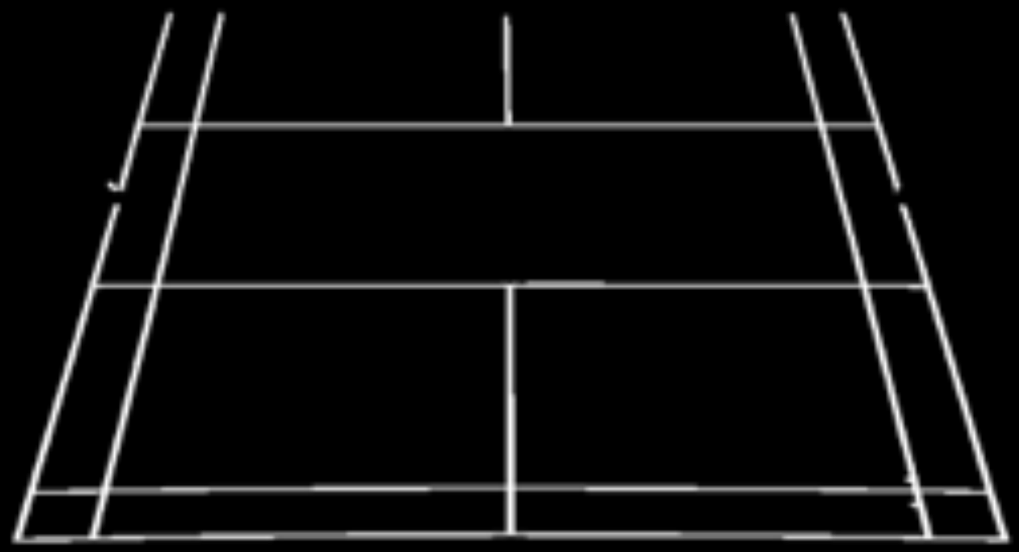


Detecting Court corners -

Removing Small Objects

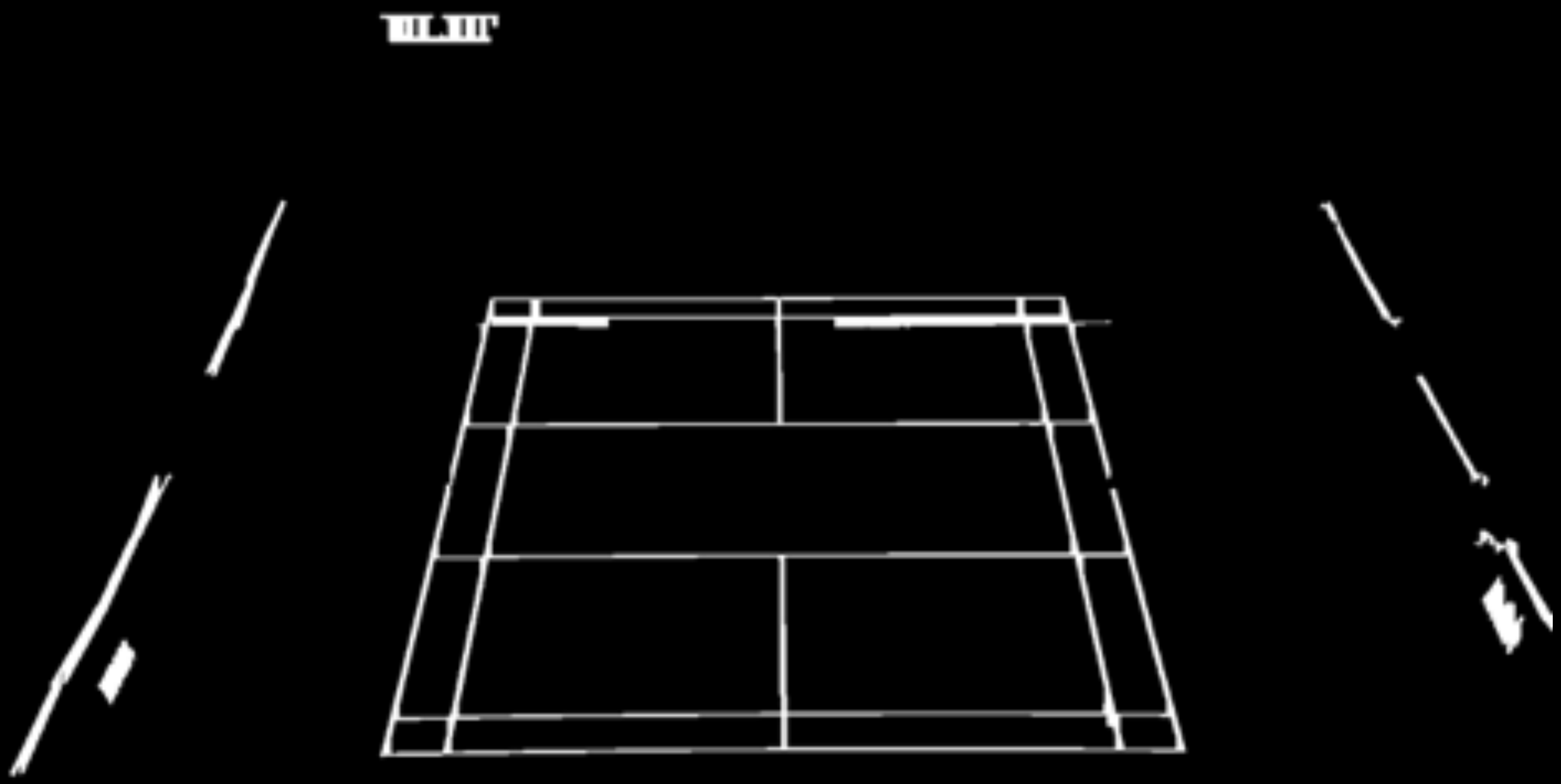
Detecting Court corners - **Problem**

- As shown in the following image, sometimes the court lines are broken due to the presence of occlusion formed by ‘net’, and they might also get filtered if the threshold is set high.



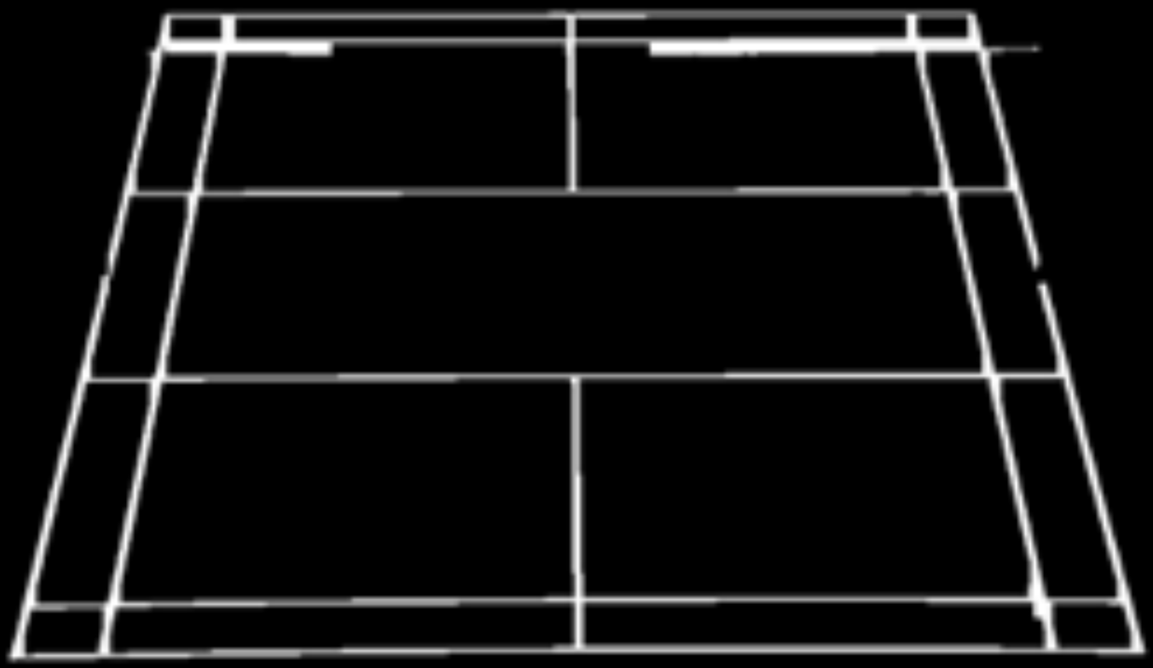
Detecting Court corners -

Solution : Closing with vertical structuring element



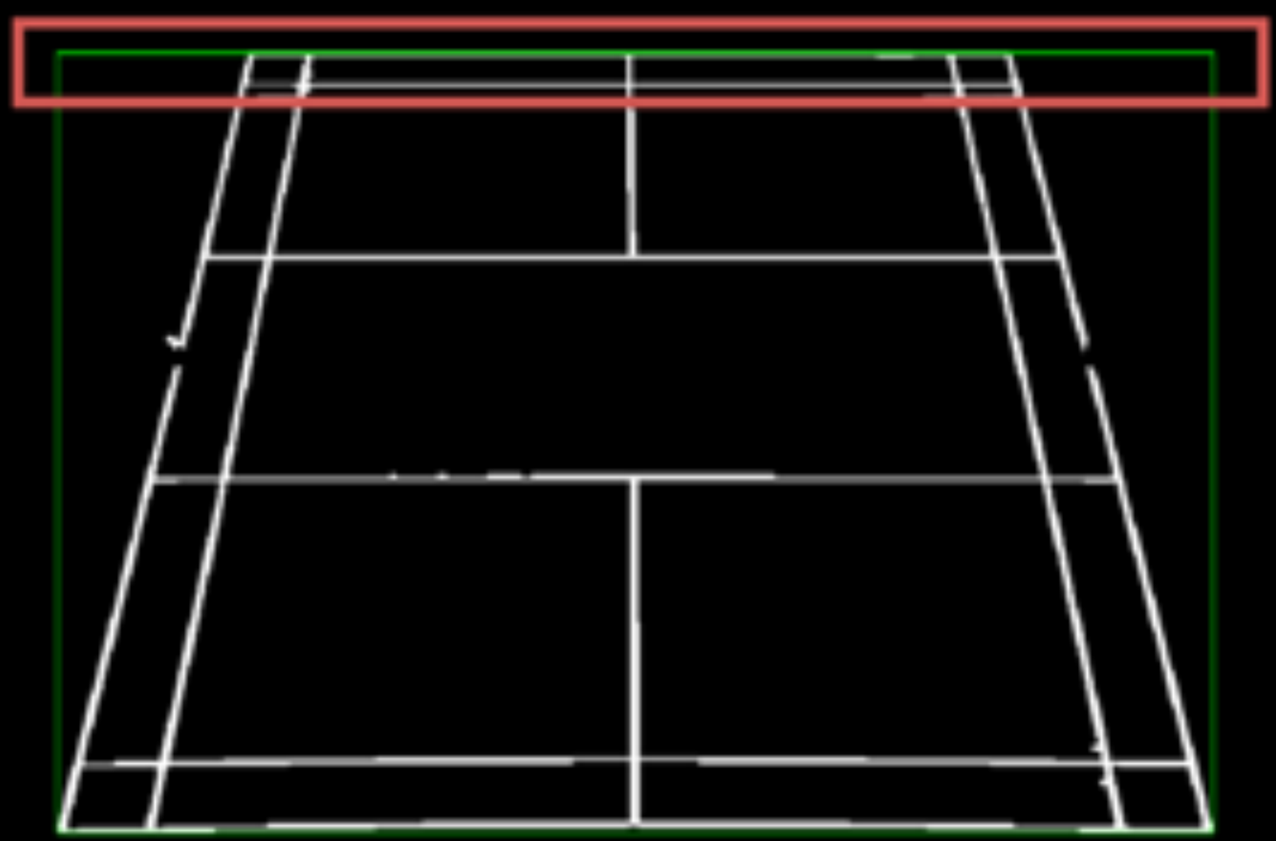
Detecting Court corners -

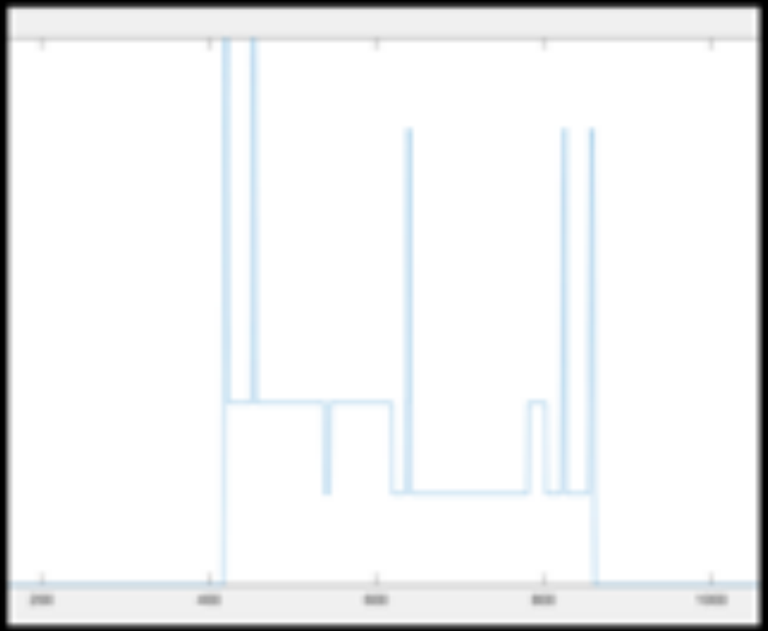
Segmented court image



Detecting Court corners -

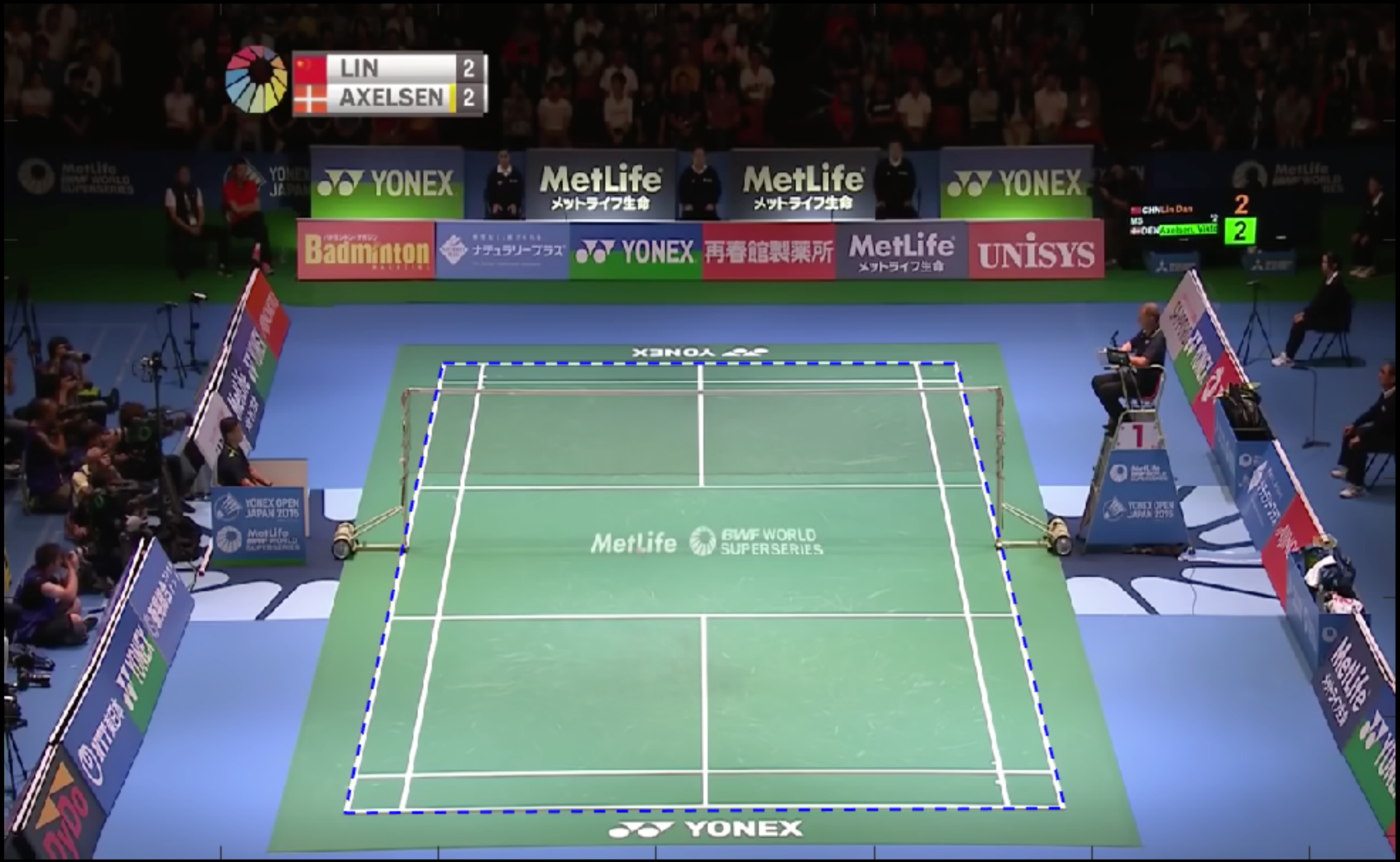
Detecting Corner points using bounding box and vertical projections





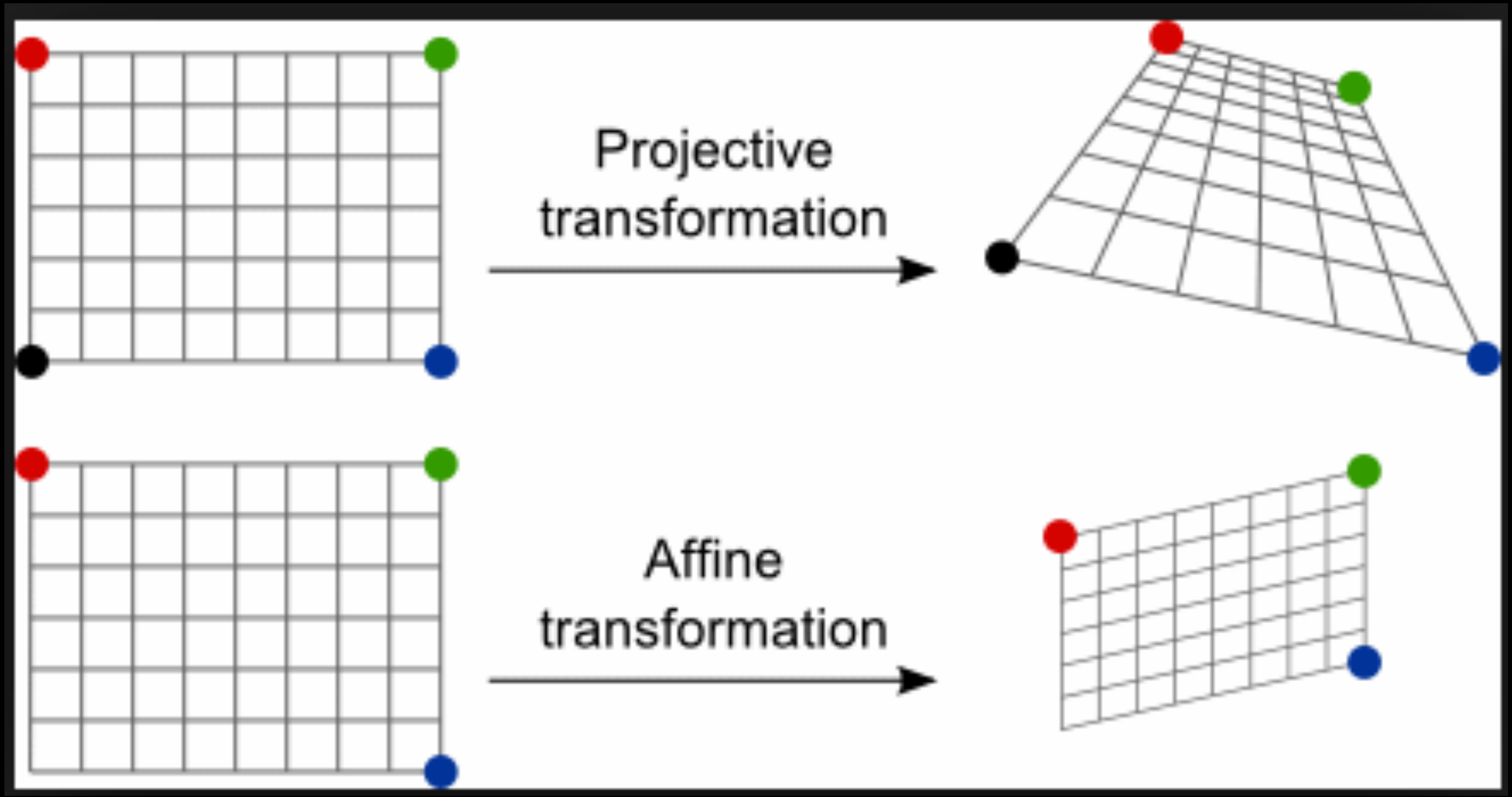
Detecting Court corners -

Finally segmented court image



Projective Correction on court

- **Projective transformation is a transformation applied to the co-ordinates of pixels.** i.e. relocating them. Projective or geometric transform is the combination of shear, rotational and translational transformation.



1. Image Source : Google

Projective Correction on court

- Based on the detected court corner points, and real world dimensions of the Badminton court, the transformation matrix is calculated, and the court plane is transformed using it.

Difference Image - Background Subtraction

Transforming current frame using calculated transform

- Background subtraction is subtracting current frame from background frame to calculate the difference.
- Since players and shuttle are the only things moving on the court, they can be seen in the difference image.
- Each video frame is transformed before subtracting from background frame.
- Following image shows an example of transformed video frame, showing players.

Difference Image - Background Subtraction

Background subtraction using RGB colorspace

- R,G and B channels of current frame and background image are subtracted from one another, and the difference is summed up to generate a difference image.

  PROBLEM : RGB colorspace is sensitive to shadows of players.

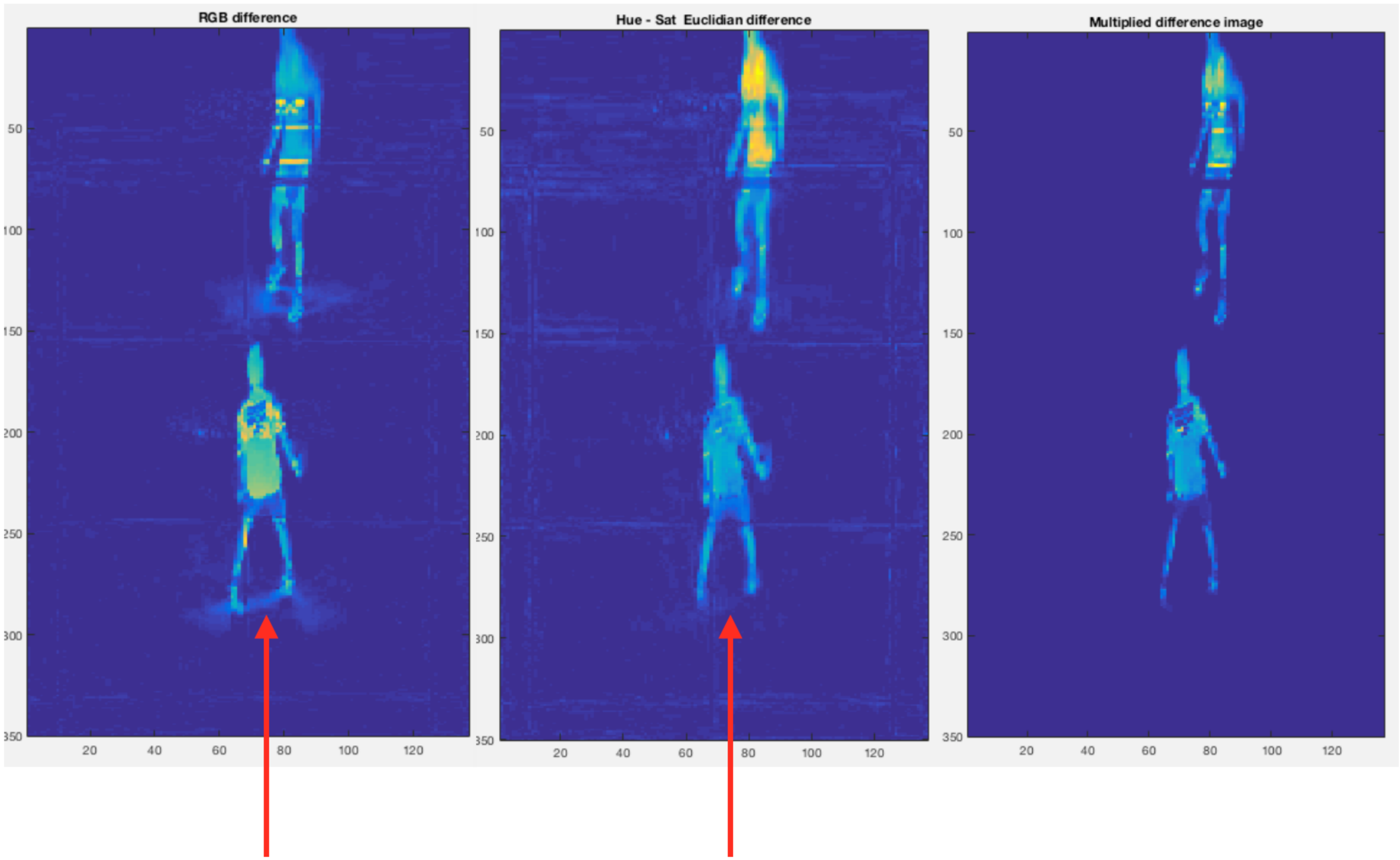
- **SOLUTION : Euclidian difference in HS(V) colorspace**

  Both Images are transformed into HSV colorspace and **euclidian difference** for H (Hue) and S (Saturation) channels is calculated. Since V (Value) accounts for darkness of shade, it is sensitive to shadow, and thus is not considered.

Both the difference images are multiplied together to generate final difference image, which is clean and more robust to shadows.

Difference Image - Background Subtraction

Difference images



Shadow in RGB difference image

No shadow in HS Euclidian difference

Multiplication of two
- Cleaner difference image

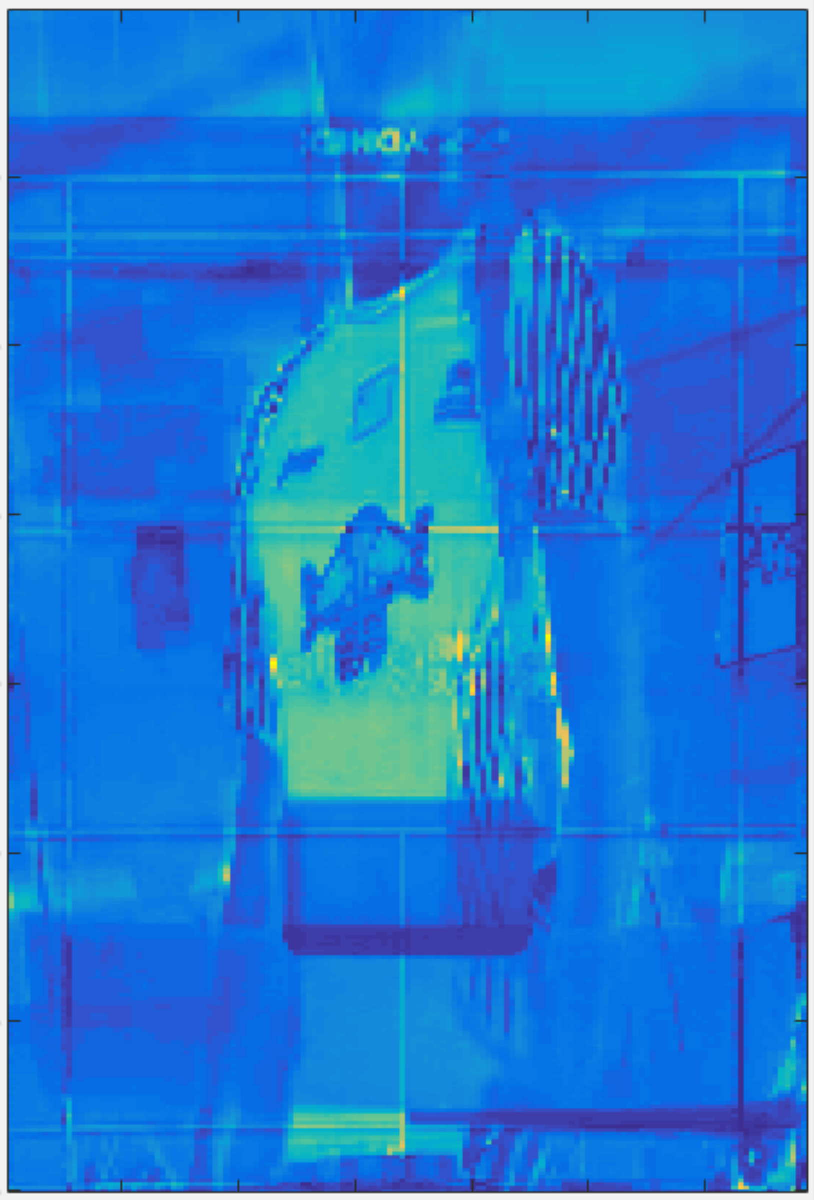
Deciding if the current frame is rally or not

- It was observed that in all matches, the court view (such as in our median approximated background scene) was only shown when the players were playing.

- If the court view is not shown in the current frame, the **sum of pixels in difference image will be way higher than usual.**

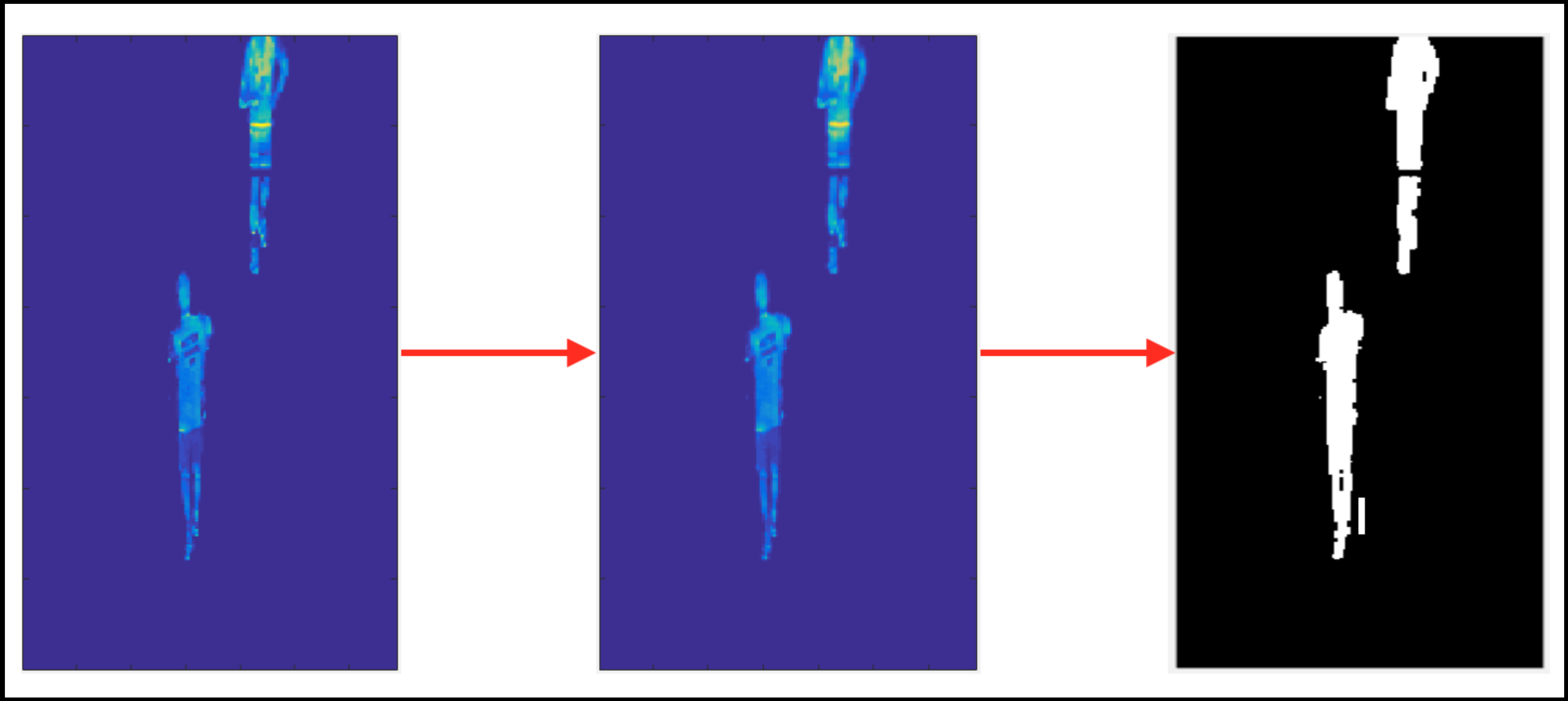
- If the frame belongs to rally, it is passed to the foreground mask detection block.

  Entire video of the match was automatically divided into small videos of rallies based on this idea, and data was generated.



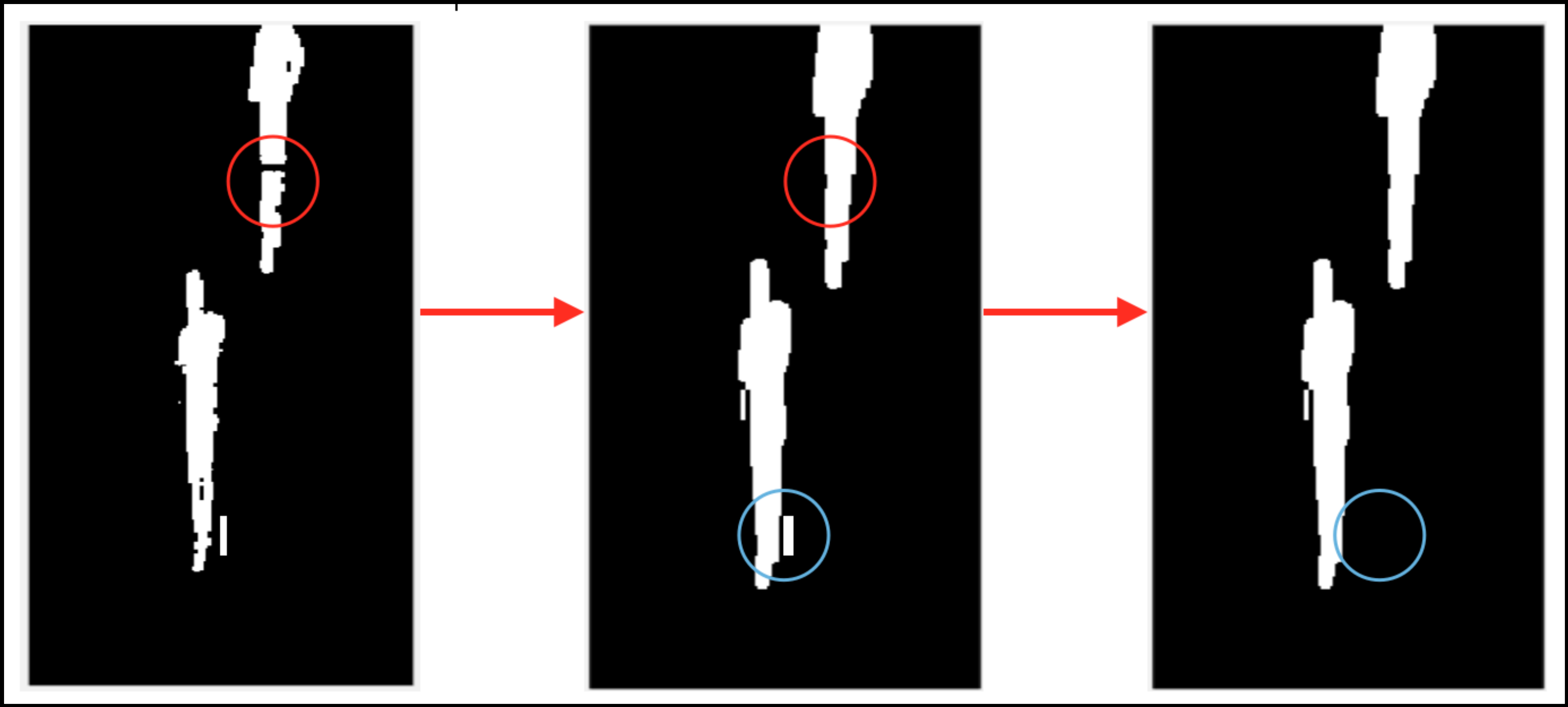
Foreground Mask Generation

Gaussian filter + Thresholding



Foreground Mask Generation

Dilation with vertical structuring element + removing small objects



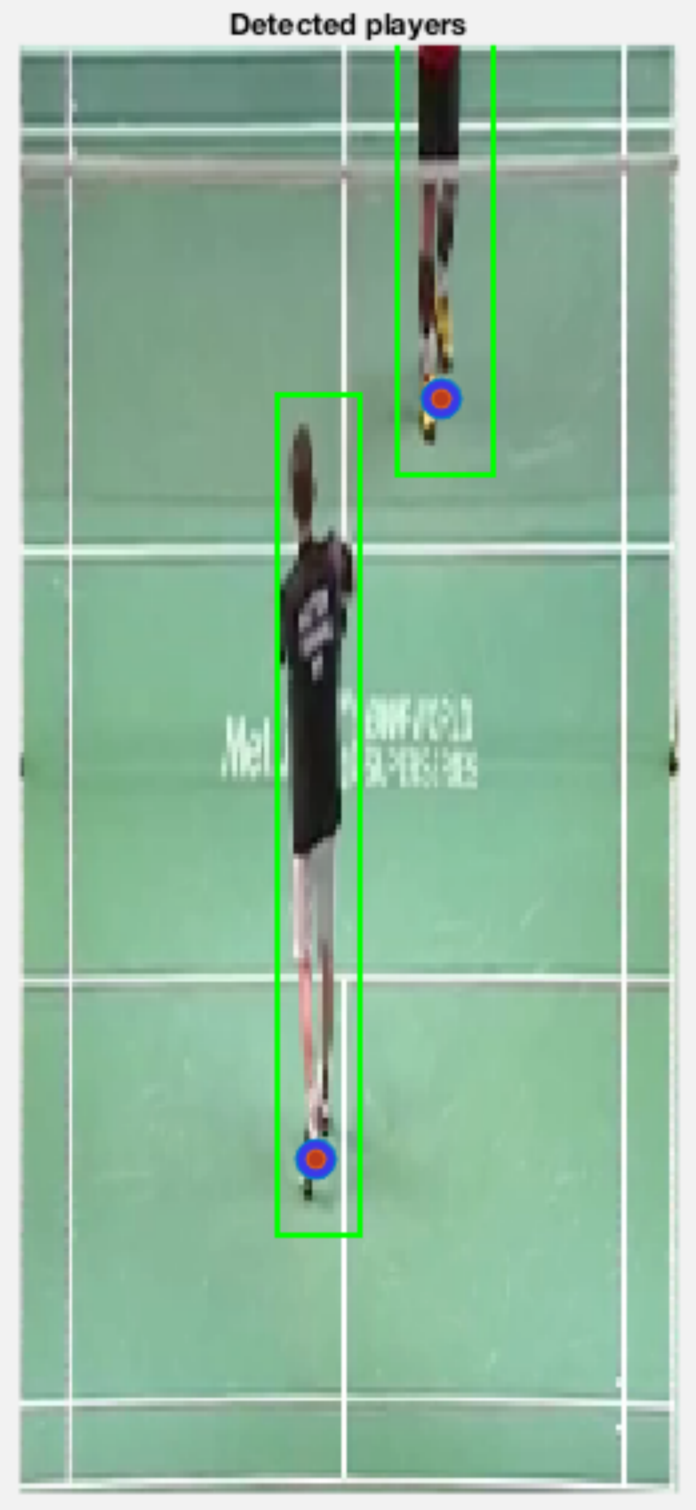
Detecting player's approximate location

- Since the player's feet are closest to the court plane, its location in the transformed space is good approximation for the players on court location in real world.

  - **Centroids and bounding boxes are found for the top 2 objects with highest areas, and having vertically oriented major axis.**

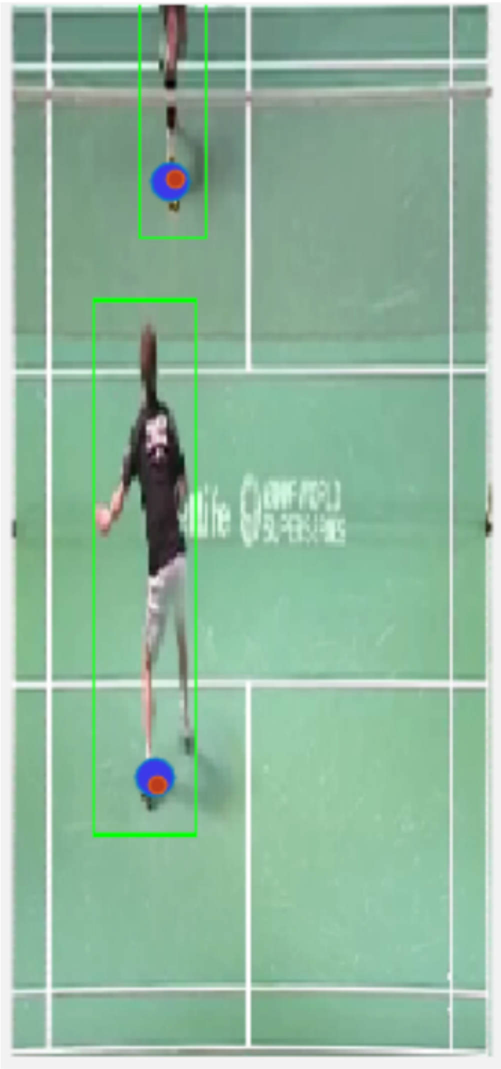
  - **The y coordinate of centroid is then dropped down to the end of bounding box to reach the feet location.**

  Note: The y- coordinate is finally adjusted to compensate for dilation operation performed using vertical structuring element.



WHY USE CENTROID :

Centroid of the object depends on the distribution of the mass of the object. i.e. even if the player raises his arm to one of the sides, since the mass of arm (area - in case of images) is less as compared to the whole body, the centroid won't shift much. Whereas the width of bounding box will drastically change.



Kalman Filter

- **NEED:**
  Since the players move fast on the court, and foreground extraction or shadow removal techniques are not highly robust, there is noise found in detections.
- Also, when both players come forward to the net, occlusion takes place - generating single object in masked image. This situation leads to missing detection for one of the players.

  To deal with this noise, Kalman filter is applied to the detected location.

Kalman Filter - Overview

Kaman Filter Overview:

Kalman filter predicts the location of the object given the past information and new data. In simple terms, the Kalman filter is a linear system of equations. It basically calculates following 3 terms.

1. State Prediction :

$$X_{predicted\,(t)} = A * X_{predicted\,(t-1)} + B * \mu_{(t)} + error_{(gaussian)t}$$

2. Sensor Prediction :

$$Z_{predicted\,(t)} = C * X_{predicted\,(t)} + error_{(gaussian)t}$$

3. State Estimation :

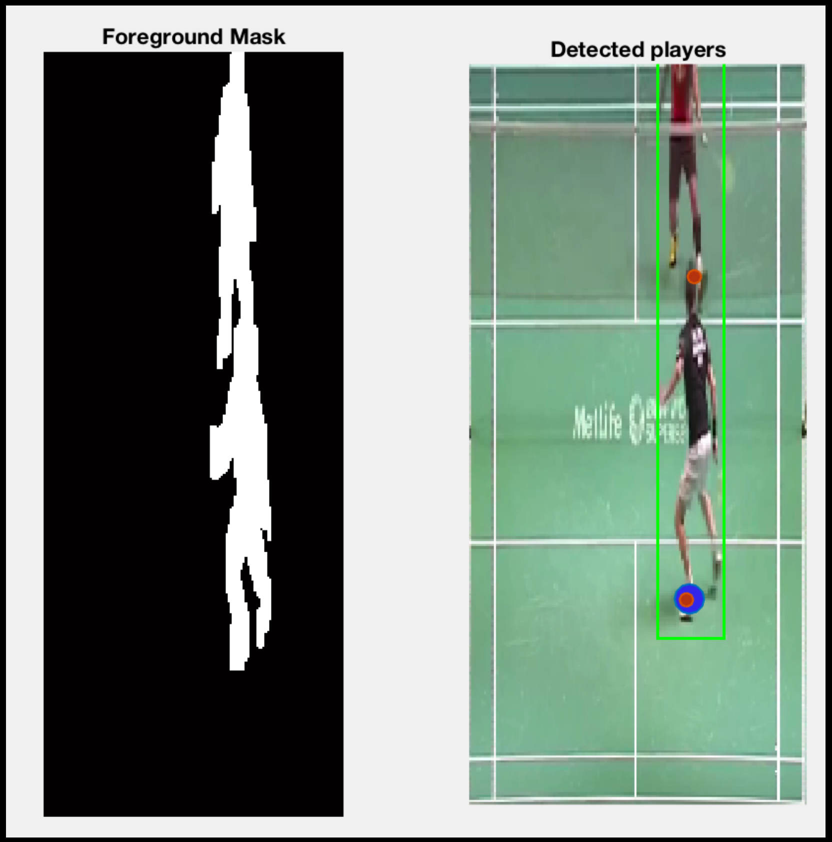
$$X_{Estimation} = X_{predicted\,(t)} + K_{(gain)} * (Z_{predicted\,(t)} - Z_{measured\,(t)})$$

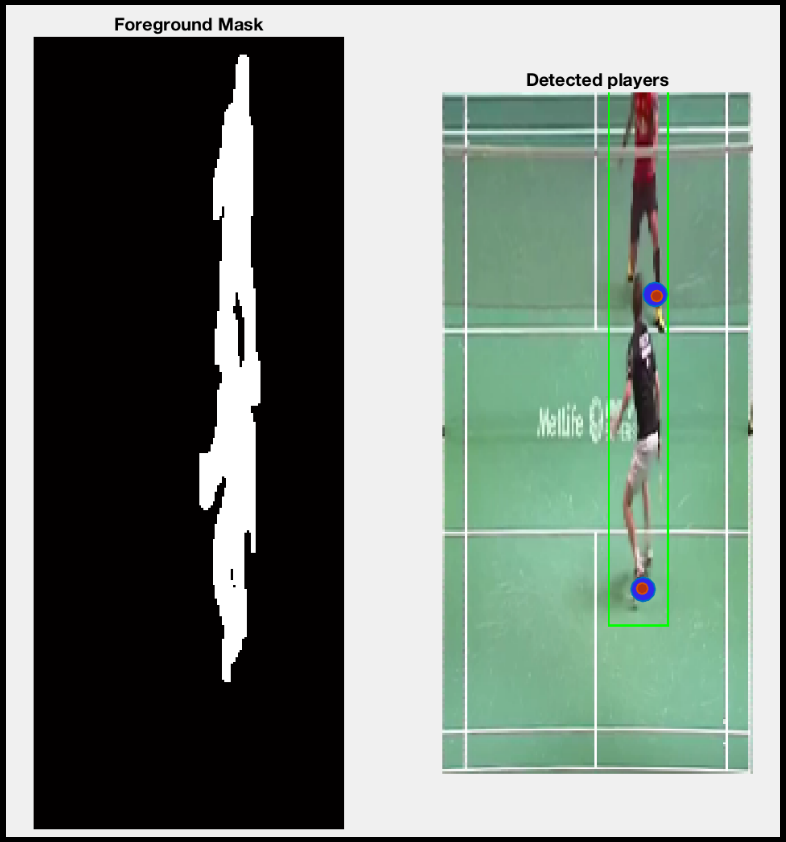
- The whole idea is to incorporate the predictions of state and sensor information, with the real sensor information we get to get best estimate.

Kalman Filter - Detections

- **These** images show the detected feet location (in Blue) and kaman filtered location (in Red) respectively.

- In the first image we can see that the blobs are connected, and the detection (blue) for one of the players is lost.
- Kaman filter predicts the correct location in such cases.
- Also, in such overlap cases, the location is not updated after certain number of frames to avoid the drift of location according to the physical model assumed by kaman filter.

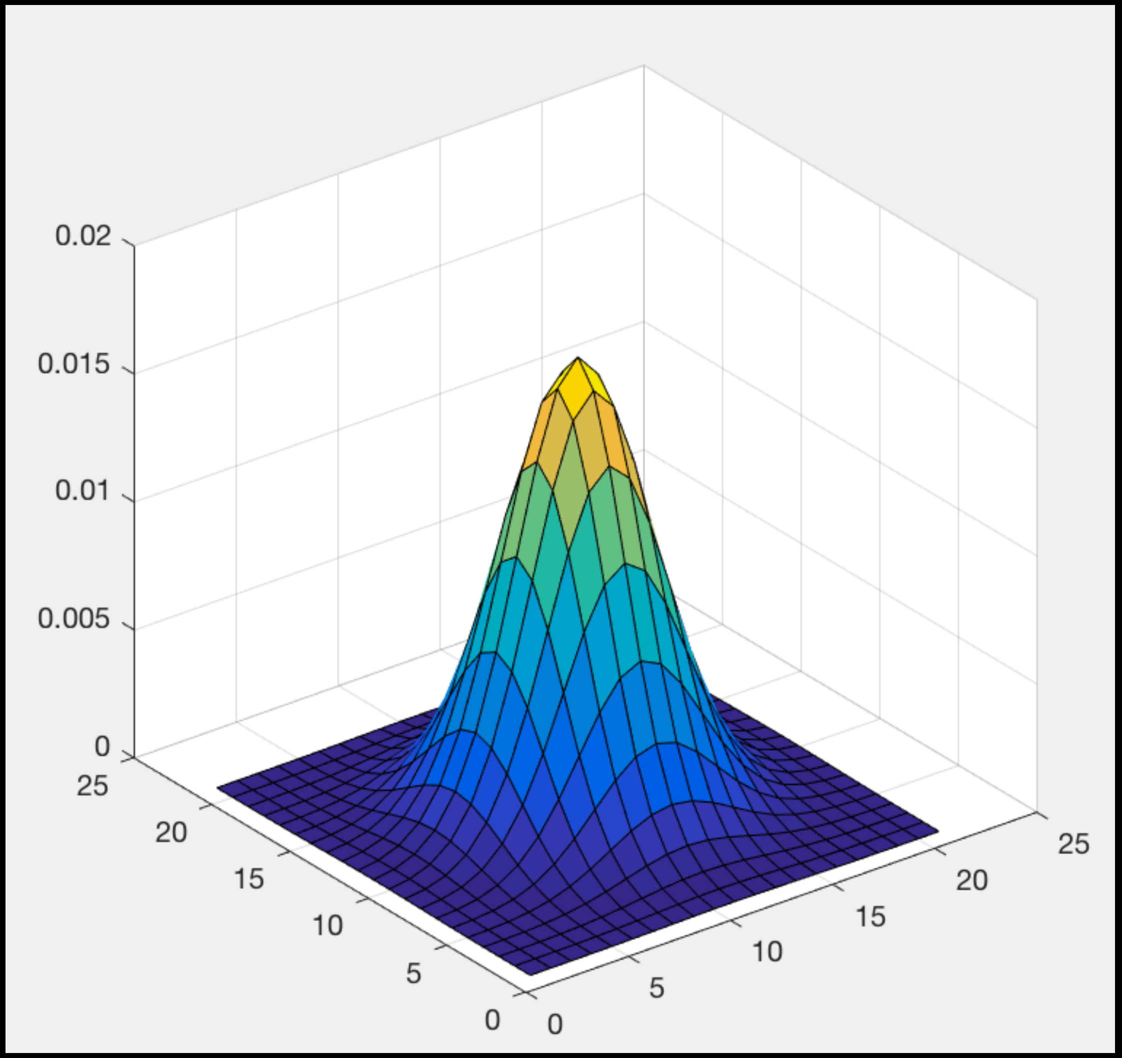




Location Map Generation

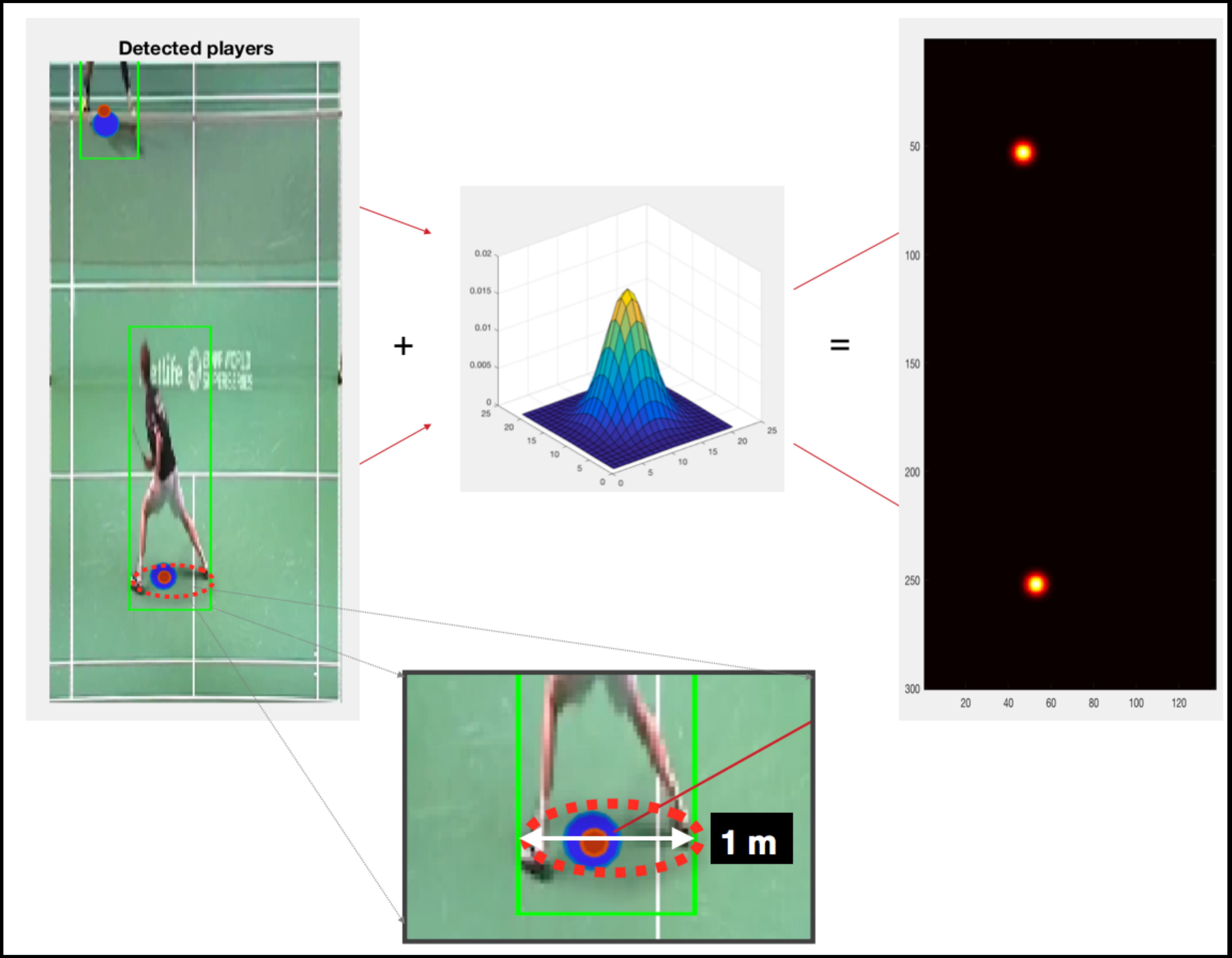
- The detected location is always an approximation of the player's real world location, and thus we can assign a normalized gaussian distribution to the court space around the detected location to indicate the probability of player being at that location.

- This also leads to an interesting visualization of the players location over single rally or complete match.



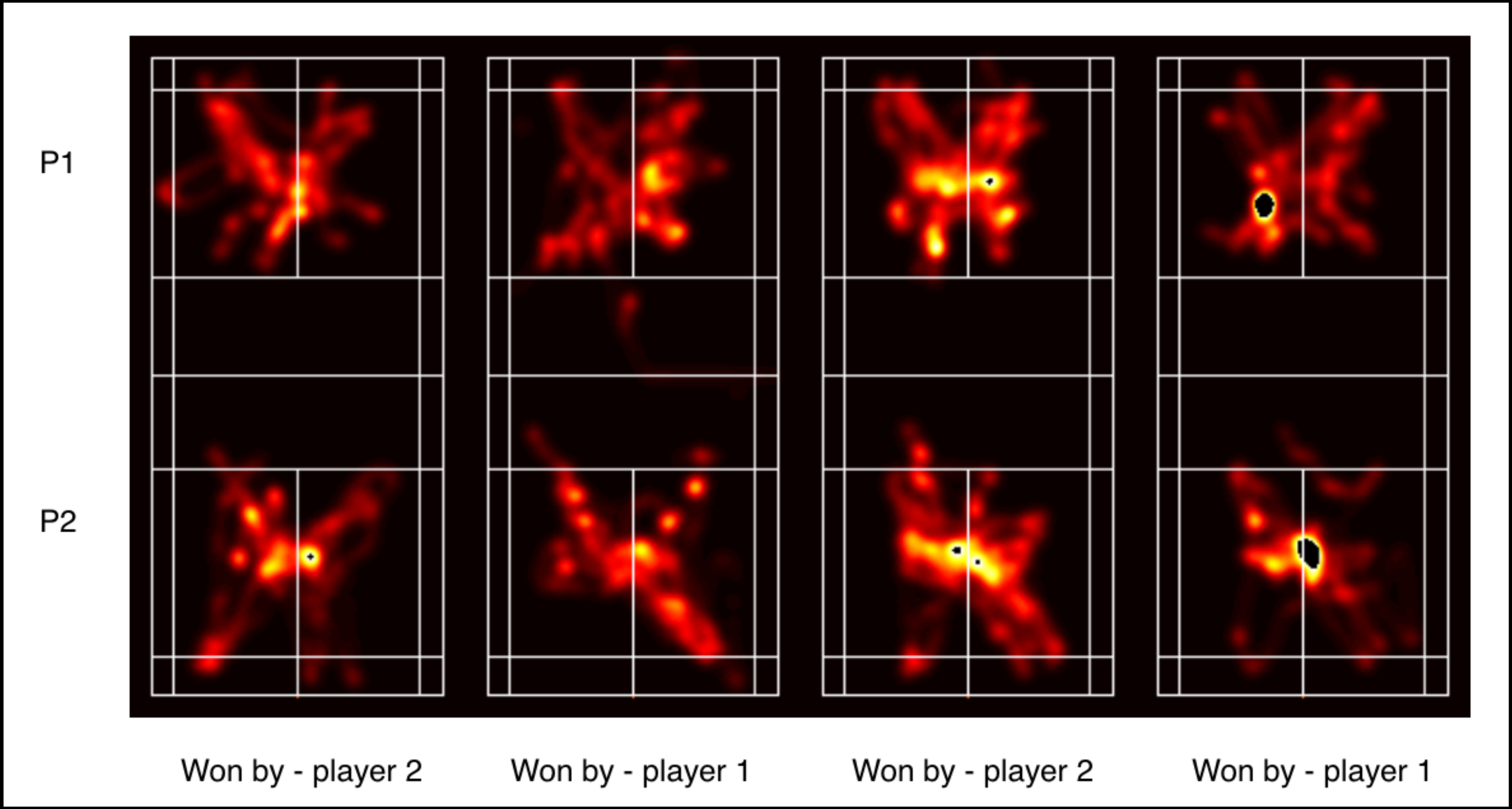
Location Map Generation

- 2D gaussian is accumulated at each detected location.



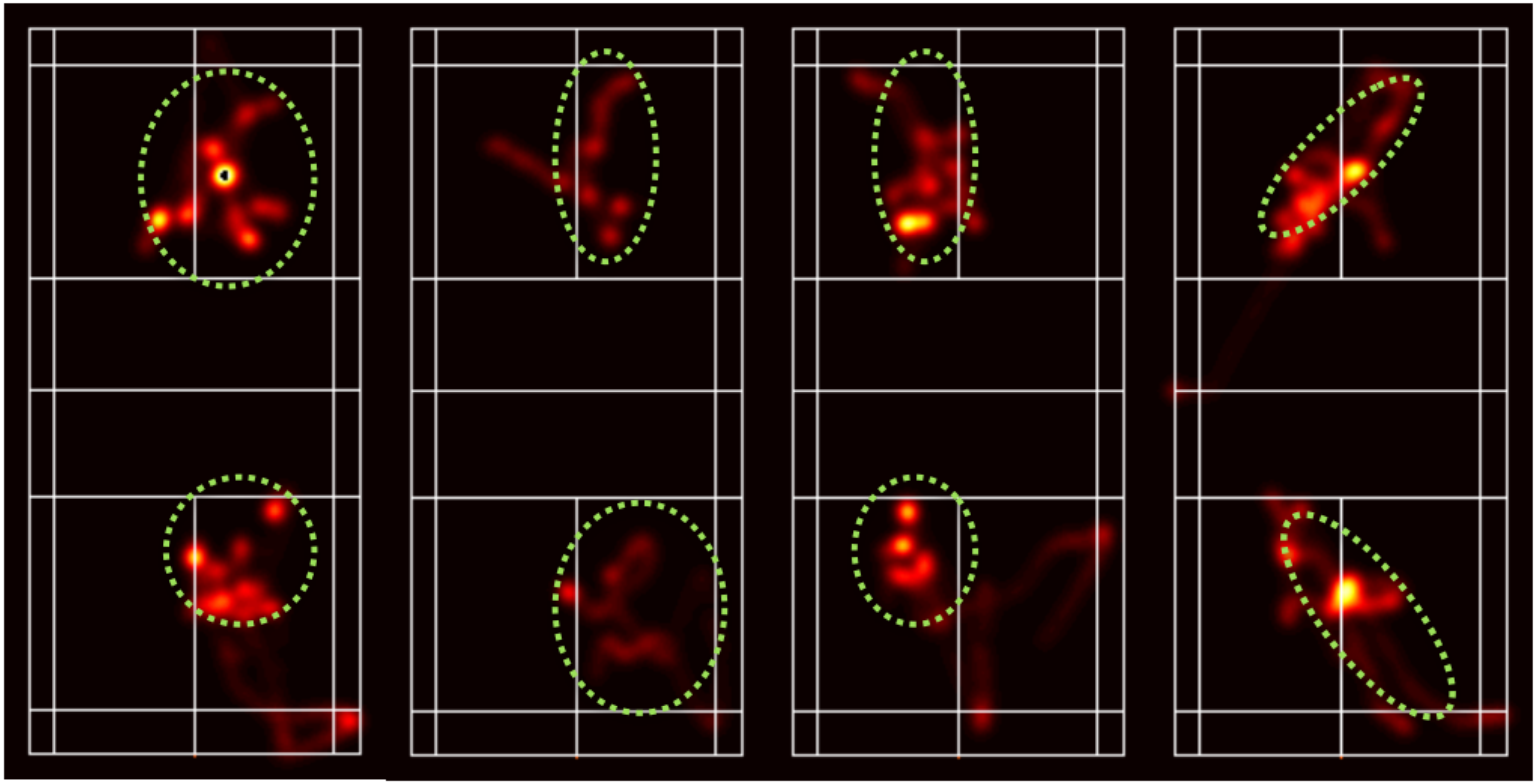
Results - long rallies

- Following images show the heat maps generated for few of the biggest rallies in a match



Results - moderate length rallies

Following images show the heat maps for moderate length rallies. **It can be easily observed that some of the rallies were biased on one side or part of the court.**



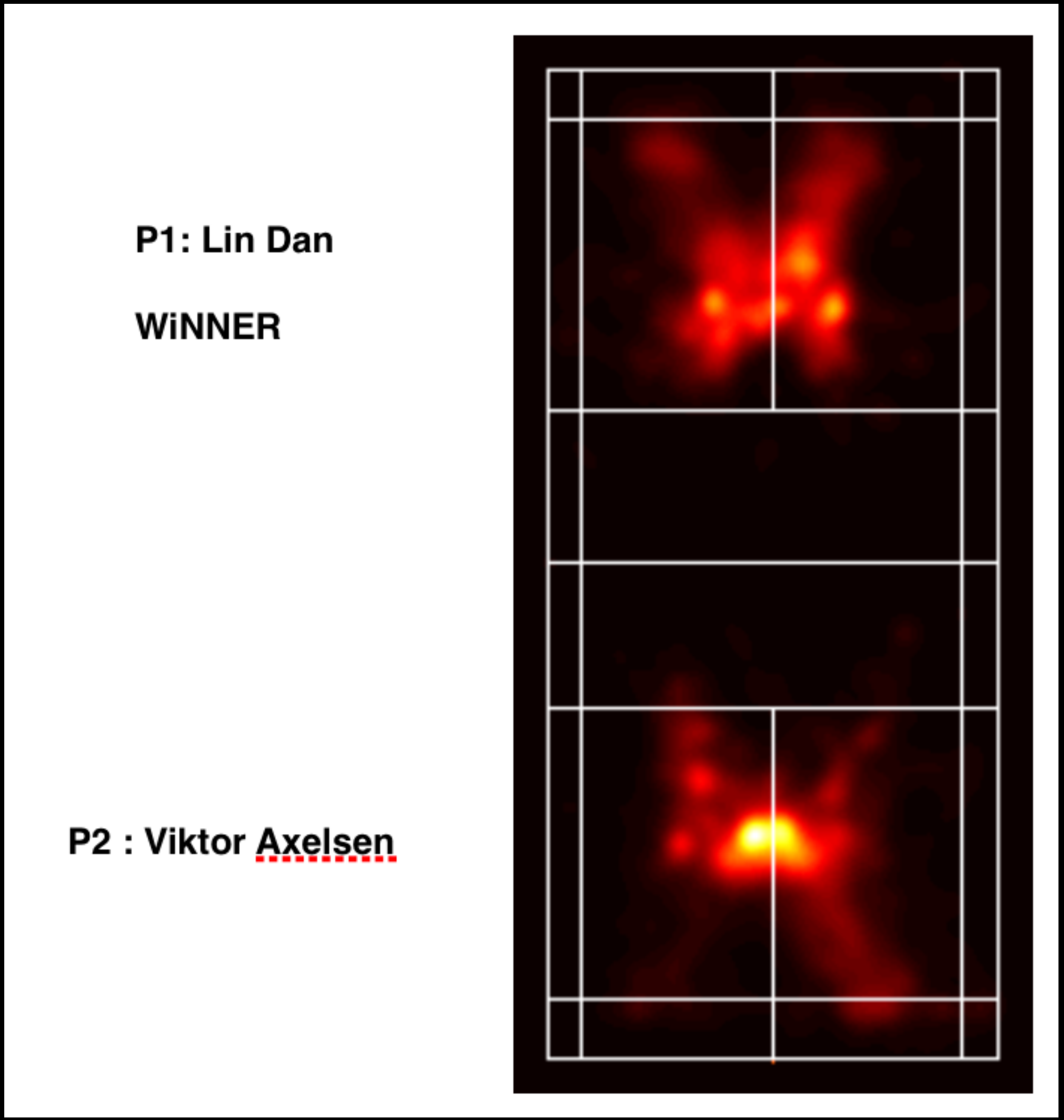
- This data can be further analyzed, and a model may be learned to predict which player won rally based on the heat map data.

Results - Full Match

- **OBSERVATIONS:**

  1. We can see a prominent spot at the center for player 2, but multiple less prominent spots around center for player 1 (which is Lin Dan). We can say that Lin dan tries to position himself on either side of court when at center based on gameplay at that time, but Viktor tried to stand at exact center location.

  2. It can also be observed that Player 2 (Viktor) was not made to play many shots on his backhand toss side (he was right handed).
  he detected location is always an approximation of the player's real world location, and thus we can assign a normalized gaussian distribution to the court space around the detected location to indicate the probability of player being at that location.



Experimentation

- Due to the time constraints the experiments were carried on only one match

Future Work

1.Data would be generated for multiple matches played by one player in a hope to find some trends.

2. Data should be ground truth to estimate accuracy of the system,

3. Movement of two players can be correlated based on the temporal location data of individual players, and shot annotation can be attempted.

4. Based on the temporal location data for multiple matches of the same player and shot annotation, a probabilistic model can be learnt which will then assign prior probability for each shot at each time instant.

5. Since temporal position data is available, the velocity and acceleration can be calculated for each player. Energy spent is directly proportional to change in acceleration - thus an energy estimate can be evaluated.

6**.** Better method for foreground extraction, occlusion handling can be tried.
indoor racquet sport

References

[1] Computer Vision in sports. Editors- Thomas B. Moeslund,Graham Thomas, Adrian Hilton. Advances inAdvances in Computer Vision and Pattern Recognition. **Workshop CVPR 2013.**

[2] Application of computer vision to automate notation for tactical analysis of badminton.Kokum Weeratunga, Khoo Boon How, Anuja Dharmaratne, Chris Messom.2014 13th International Conference on Control, Automation, Robotics & Vision, 2014.

[3] Object Detection and Tracking based on Trajectory in Broadcast Tennis Video. M. Archana, M. Kalaisevi Geetha.Second International Symposium on Computer Vision and the Internet (VisionNet'15).

[4] C. Bingqi and W. Zhiqiang, "A statistical method for analysis of technical data of a badminton match based on 2-D seriate images," Tsinghua Sci. Technol., vol. 12, no. 5, pp. 594–601, 2007.

[5] J. Han, D. Farin, and P. H. N. de With, "Broadcast court-net sports video analysis using fast 3-D camera modeling," IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 11, pp. 1628–1638, Nov. 2008.

References

[6] D. Zhong and S.-F. Chang, “Real-time view recognition and event detection for sports video,” J. Vis. Commun. Image Represent., vol. 15, no. 3, pp. 330–347, Sep. 2004.

[7] J. Han, D. Farin, P. H. N. De With, S. Member, and W. Lao, “Real- time video content analysis tool for consumer media storage system,” IEEE Trans. Consum. Electron., vol. 52, no. 3, pp. 870–878, 2006.

[8] D. Connaghan, K. Moran, and N. E. O’Connor, “An automatic visual analysis system for tennis,” Proceedings. Part P J. Sport. Eng. Technol., pp. 1–16, Feb. 2013.

[9] W.-T. Chu and W.-H. Tsai, “Modeling spatiotemporal relationships between moving objects for event tactics analysis in tennis videos,” Multimed. Tools Appl., vol. 50, no. 1, pp. 149–171, Sep. 2009.

[10] S. L. Teng and R. Paramesran, “Detection of service activity in a badminton game,” in TENCON 2011 - 2011 IEEE Region 10 Conference, 2011, no. rule 3, pp. 312–315.
A. Borrie, G. K. Jonsson, and M. S. Magnusson, “Temporal pattern analysis and its applicability in sport: an explanation and exemplar data,” J. Sports Sci., vol. 20, no. 10, pp. 845–852, Oct. 2002.

Thank you !