

# TEAM BEHAVIOR ANALYSIS IN SPORTS USING THE POISSON EQUATION

*Cem Direkoğlu and Noel E. O'Connor*

CLARITY: Centre for Sensor Web Technologies, Dublin City University, Ireland  
cem.direkoglu, noel.oconnor@dcu.ie

## ABSTRACT

We propose a novel physics-based model for analysing team players' positions and movements on a sports playing field. The goal is to detect for each frame the region with the *highest population* of a given team's players and the region towards which the team is moving as they press for territorial advancement, termed the *region of intent*. Given the positions of team players from a plan view of the playing field at any given time, we solve a particular Poisson equation to generate a smooth distribution. The proposed distribution provides the likelihood of a point to be occupied by players so that more highly populated regions can be detected by appropriate thresholding. Computing the proposed distribution for each frame provides a sequence of distributions, which we process to detect the region of intent at any time during the game. Our model is evaluated on a field hockey dataset, and results show that the proposed approach can provide effective features that could be used to generate team statistics useful for performance evaluation or broadcasting purposes.

**Index Terms**— Feature extraction, Poisson equation, computer vision, video analysis, team sports

## 1. INTRODUCTION

Human behavior analysis is an important task in computer vision. Despite the fact that there is much research on vision-based individual behavior analysis [1], group behavior analysis remains a challenging problem. In group behavior, there are usually many people located at different positions, moving in different individual directions which makes it difficult to find effective features for higher level analysis. In this paper, we investigate team behavior in field sports. We aim to extract two important features on the field of play at any time during the game: the *highest populated region*, corresponding to the area with the highest density of the majority of players, and the *region of intent*, corresponding to the region towards which the team is moving as part of their overall strategy as they press for territorial advancement.

Detection of the highest populated regions and the regions of intent at each frame could be used to generate useful statistics about team behavior for coaches. For example, if we compute the highest populated region for each frame and take the average over the entire match, the resultant *position map* can indicate where the team spent most of its time. On the other hand, the *intention map* can similarly be computed by averaging the regions of intent obtained during the match, and it can indicate where a team was focusing its efforts during the game. For broadcasters, this information could be useful for automatic camera control to follow team activity, potentially

allowing for fewer camera operators to be present when capturing non-premium sporting events. In this paper, however, we focus only on detection of regions of highest population and regions of intent.

### 1.1. Related Work

There are few approaches which consider detecting similar features to ours in sports. Kawashima et al. [2] propose to detect group regions in broadcast soccer video sequences. However, using the broadcast camera feed is not effective for team behavior analysis, since the camera usually only captures the current region of interest (such as ball locations) and many team players may not be in that region. Using the broadcast cameras also suffers from inaccurate player localization because of occlusions, camera motion, etc. Most of the team behavior analysis methods [3, 4] use a fixed multi-camera system around the playing field to overcome the limitations of using broadcast cameras. The multi-camera system usually has a camera configuration to cover all locations on the field of play and is therefore able to capture all players simultaneously. Player detection and tracking algorithms are employed in the videos to obtain the trajectories, and then these trajectories are transformed into the top-view (plan view) of the playing field for more accurate analysis. Kim et al. [3] propose to predict where the interesting events will occur in soccer games by analysing the movements of all players (both of the teams) on the playground. They create a motion field on the top-view of the field of play by spatial and temporal interpolation of the player motions. Then, they analyze the motion field to detect the possible location of the important events. Taki and Hasegawa [4] propose “dominant regions”, where a player can arrive earlier than any other, to evaluate teamwork in soccer games. The dominant regions are formulated based on a Voronoi diagram model on the top-view image of the field of play.

### 1.2. Our Work and Contribution

We propose a new physics-based approach for team behavior analysis. We solve a particular Poisson equation on the top-view image of the playground using the team players' positions at each frame and then process the resultant distributions to determine the highest populated regions and detect the regions of highest intent at each frame. We evaluate our approach on field hockey, using a similar multi-camera capture set-up to those reported previously, where sample frames from the our field hockey dataset are shown in Figure 1. In this work, the trajectory data are collected manually using the multi-camera network. The trajectory data could also be obtained automatically using a computer vision based tracker or GPS-based wearable sensors. Regardless of how the trajectory data are obtained, the problems described in this paper must be addressed by a computer vision system designed for higher level analysis. Results show that the proposed method can extract the required features on the playground.

This work is supported by Science Foundation Ireland under grant 07/CE/I114. The authors would like to acknowledge Disney Research Pittsburgh for their help in generating the data sets used.

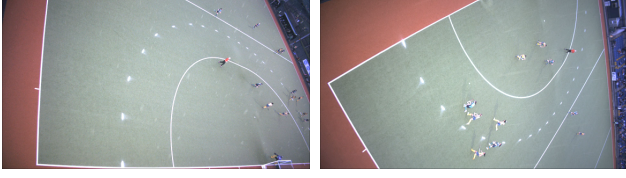


Fig. 1. Sample frames from 2 of the 8 cameras for our hockey dataset

## 2. THE POISSON EQUATION AND PROPOSED MODEL

### 2.1. Background to the Poisson Equation

In mathematics, the Poisson equation is an elliptic type partial differential equation [5] which arises usually in electrostatics, heat conduction and gravitation. The general form of the Poisson equation, in two-dimensions, is given by,

$$\nabla^2 I(\mathbf{x}) = -Q(\mathbf{x}) \quad (1)$$

where  $Q$  is a real-valued function of a space vector  $\mathbf{x} = (x, y)$  and it is known as the source term,  $I$  is the solution which is also a real-valued function and  $\nabla^2$  is the spatial Laplacian operator. Given a source term  $Q(\mathbf{x})$ , we find a solution for  $I(\mathbf{x})$  that satisfies the Poisson equation and the boundary conditions over a bounded region of interest. There are three general types of boundary conditions: Dirichlet, Neuman and Mixed. Here, we explain the Dirichlet condition, which is used in our algorithm. In the Dirichlet condition, the boundary values (solutions) are specified on the boundary. These values can be a function of space or can be constant. The Dirichlet condition is represented as  $I(\mathbf{x}) = \Phi(\mathbf{x})$ , where  $\Phi(\mathbf{x})$  is the function that defines the solution at the boundary layer.

### 2.2. The Proposed Poisson Equation and Solution

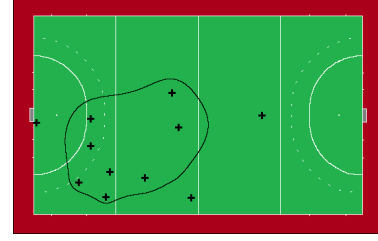
We use a particular Poisson equation for team behavior analysis in sport games. We investigate the problem in the context of field hockey, where the top-view of the hockey field of play is shown in Figure 2(a) with the team player positions (a field hockey team has 11 players). Given the positions of the team players at any time, we solve a particular Poisson equation to generate a smooth distribution.

The top-view image of the field of play is assumed to be a binary image where the player positions are one and the rest of the positions are zero at any time during the game. Although players are expected to be in the play area during the game, players sometimes can move a little outside for a variety of different reasons, such as to serve the ball, when the ball is out or in order to talk to the coach. Thus, we expand the binary image of the field of play to include the possibility that the players may move a little outside the lines. The binary image is defined to be the source term in the Poisson equation. The boundary condition is Dirichlet which has a specific solution,  $I(\mathbf{x}) = 0$ , at the boundaries of the expanded field of play. This means that there is no possibility for a player to be outside the region of interest. The proposed Poisson equation problem is,

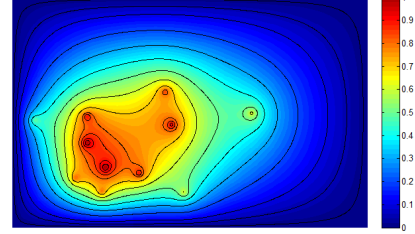
$$\nabla^2 I(x, y) = - \left( \sum_{i=1}^N \delta(x - x_i, y - y_i) \right) \quad (2)$$

$$I(x, y) = 0, \quad \text{boundary condition}$$

where  $N$  is the number of players in the team and  $(x_i, y_i)$  is the position of player  $i$ . The source function is assumed to be a linear combination of dirac-delta functions  $\delta(\cdot)$  in two dimensions. It is



(a) The top-view of the field hockey pitch with player locations



(b) The position distribution of the team

Fig. 2. Poisson equation applied to find highest populated regions

important to note that the proposed Poisson equation has a unique and steady-state solution at each frame. Therefore, when players change their position from the previous frame to the current frame, the solution also changes in the current frame.

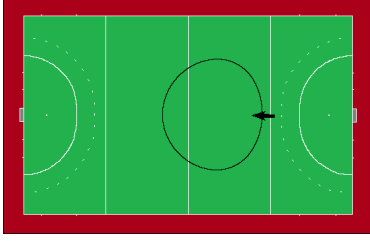
There exist both direct and iterative numerical solution methods of the Poisson equation. In [6], Simchony et al. pointed out that direct methods are more efficient than multigrid-based iterative methods for solving the Poisson equation on a rectangular domain, since direct methods can be implemented using the Fast Fourier Transform (FFT). In our work, since the field of play is rectangular, we employ FFT-based direct methods to solve the proposed Poisson equation. The proposed equation has a Dirichlet boundary condition that needs discrete sine transforms (using FFT) to achieve an exact solution, where the detailed description of the solution method is given in [6]. The solution to the proposed equation forms peaks at the player positions. To smooth these peaks, we apply Gauss-Seidel iterations (8 iterations), as a post-processing stage, to relax the surface while maintaining the boundary condition ( $I(\mathbf{x}) = 0$ ) outside the region of interest.

## 3. TEAM BEHAVIOR ANALYSIS

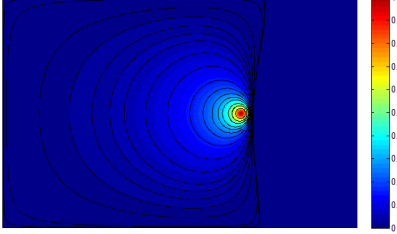
We use the resultant distributions at each frame to extract two important features related to team behavior: the highest populated regions and the regions of intent.

### 3.1. Highest Populated Regions

Since the resultant distribution provides the likelihood of a position to be occupied by players at that time, the distribution is called the *position distribution* of the team. Figure 2(b) shows the normalized position distribution (divided by the maximum value in the distribution) for the given example. The position distribution has higher values in a region where the players are close to each other. So, the highest populated region can be detected by applying an appropriate threshold to the distribution. If there is more than one region above the threshold, the biggest region is selected since our aim is to detect the region with the majority of the players. We assume that there is always a highest populated region at each frame, because the majority of the team players move together in the same direction during



(a) The player movement and the region of intent



(b) The intention distribution of the player

**Fig. 3.** Computing the intention distribution of a player

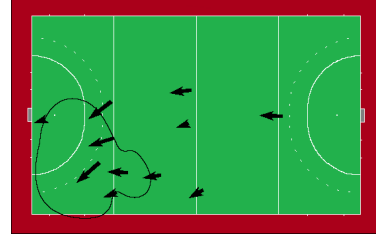
the game to follow the ball so that they are close to each other and occupy the same region. The level sets of the position distribution, which is shown in Figure 2(b), can also be considered as different threshold levels. We use 85 highest percentiles of the distribution to detect the highest populated region at each frame, where the boundary of the highest populated region, for the given example, is shown in Figure 2(a) with a black contour. It can be observed that we obtain a smooth shaped region of highest population. Three players are excluded (the goalkeeper and the forward players) in this region, since they are away from the majority of the players.

### 3.2. Region of Intent

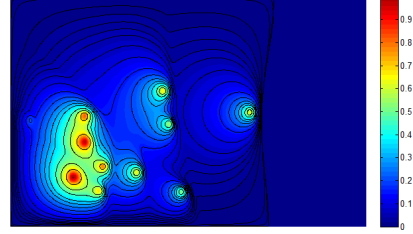
We aim to detect the region of intent at any time during the game. The region of intent can be understood as a region to where most of the team players are moving. We propose a new distribution, termed the *intention distribution*, which is obtained by processing the position distributions of the players. First we explain how we compute the intention distribution of a single player, and then how to compute the intention distribution of the team.

Figure 3(a) shows the direction of movement of the player from the previous frame (50 frames before) to the current frame, where the starting point of the arrow represents the position of the player at the previous frame and the end point represents the position of the player at the current frame. We compute the position distribution for that player at the previous and at the current frames. Since the player moves from the previous position to the current position, the player creates higher position distribution values in the direction of motion. To obtain the intention distribution at the current frame, we apply change detection by simply subtracting the previous distribution from the current distribution and keep the positive values while setting the negative values to zero, i.e.  $(I_i^n - I_i^{n-m}) > 0$ , where  $I_i^n$  represents the position distribution of the player  $i$  at frame number  $n$  and  $m$  is the number of frames between the current and the previous frame. Figure 3(b) shows the normalized intention distribution (divided by the maximum value) of the player. The black contour in Figure 3(a) represents the boundary of the region of intent, which is detected using 90 highest percentiles of the distribution.

The intention distribution of the team at the current frame can be computed by simply summing the intention distributions of the



(a) Team players movements and the region of intent



(b) The intention distribution of the team

**Fig. 4.** Computing the intention distribution of team.

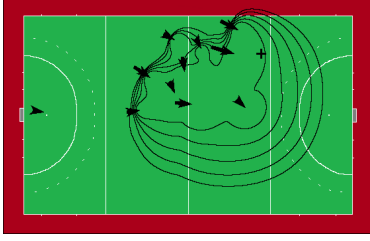
individual players:

$$ID^n = \sum_{i=1}^N (I_i^n - I_i^{n-m}) \text{ where } \forall (I_i^n - I_i^{n-m}) > 0 \quad (3)$$

where  $ID^n$  is the intention distribution of the team. As a post-processing stage, Gauss-Seidel iterations (8 iterations) are applied to relax the surface while maintaining the boundary values at zero. The directions of players movements are shown in Figure 4(a), where the starting points of the arrows represent players' positions at the previous frame (50 frame before) and the arrow magnitudes are proportional to the movement magnitude. The normalized intention distribution (divided by the maximum value) of the team is shown in Figure 4(b) with the level sets that are considered as different threshold levels. The region of intent can be extracted by applying an appropriate threshold to the distribution. As before, in case of multiple regions above the threshold, we choose the largest. We use 85 highest percentiles of the distribution to detect the highest intention region of the team at each frame, where the boundary of the highest intention region is illustrated in Figure 4(a) with a black contour. It can be observed that we detect a smooth shaped region of intent.

## 4. EVALUATION

The proposed model is validated on a field hockey dataset, which is an outdoor team sport with field of play of  $91.4 \times 55$  meters. We collected a dataset by recording an international match between Ireland and Australia (adult ladies). We use eight fixed (and synchronized) cameras around the field in order to cover the entire field of play and, each camera (Prosilica type) is mounted on pole 20 meters high. The top-view field of play coordinates of the Australian team players were then extracted as follows: First players' positions were manually labeled for each camera view, then we use homographic transformation for each camera view to transform the players' positions from the image domain to the top-view image domain. If the same player is identified by more than one camera, the center of mass of the position candidates on the top-view is computed after the transformations. We prepared a dataset (team players' positions) consisting of 2000 frames (a period of one minute and twenty



**Fig. 5.** Region of intent detection (team) with different thresholds.

seconds), which is all that could be generated to date given the extremely labour-intensive nature of generating this data.

Quantitative evaluation of these kind of approaches is a difficult task, since the highest populated region or the regions of intent can be subjective, and there is no ground truth for defining these regions. Kim et al. [3] evaluated their approach, which predicts the locations of where the interesting event will happen, by comparing with the location of the ball. Since players focus on the ball during the game and they move towards the ball position, the ball position can be considered to be the ground truth to evaluate the detection of the region of intent of the team. Following this approach, the ball position is manually annotated to indicate the ground truth position on the field of play for our approach. Thus, we extract a region of intent at each frame, and compare with the ball position. If the ball position is included in the region of intent, it is assumed that the detection is correct. We believe that this is a reasonable validation technique which can provide quantitative results for the accuracy of the region of intent detection. We only evaluate during game play, since if the game stops such as for free throw, corner or penalty, players do not run after the ball. The ball is on the playground in 1457 frames in our dataset. Therefore, we evaluate our method, for the accuracy, using 1457 frames and compute the correct detection rate (CDR%) with respect to different threshold levels (70, 75, 80, 85, 90 highest percentiles) as shown in Table 1.

**Table 1.** CDR% with respect to different threshold levels.

Threshold	>70%	>75%	>80%	>85%	>90%
CDR %	97.0	95.6	90.6	83.9	67.4

It is observed that using 70 highest percentiles of the distribution can achieve 97% for ball position detection, and as the threshold decreases the CDR% also decreases, where using 90 highest percentiles of the distribution achieves 67.4%. Our evaluation shows that we generate a high quality intention distribution since we can estimate the region of the ball in most of the frames. Figure 5 shows an example of team movement, and the region of intent detection with respect to different threshold levels which we use in the evaluation. Black contours, outermost to innermost, represent the boundary of the intention region that is detected using the 70, 75, 80, 85 and 90 highest percentiles of the distribution respectively. The black sign (+) represents the ball position and it is seen that the ball is included inside the region of intent at all threshold levels.

It should be noted that we cannot compare our approach with that of Kim et al. [3], because they estimate where the region of intent will be in a future frame (not in the current frame as in our approach) by analysing the two teams players positions. Given the players positions (two teams) at the current frame and at the few previous frames, they estimate the region of intent in the future (up to 120 frames later from the current frame). On the other hand, in our work, given the players position (one team) at the current frame and in the previous frame (50 frame before), we estimate the region

of intent for the current frame.

The computational time for each part of the proposed approach are given in Table 2. Since the computation is highly dependent on the resolution of the top-view playground, we present time evaluations for two different resolutions ( $360 \times 558$  and  $240 \times 372$  pixels). Note that average times for the position distribution and the intention distribution generation, as well as for the threshold selection and application are obtained using the 2000 frames.

**Table 2.** Average times for the position distribution and the intention distribution generation for the team, as well as for the threshold selection and application to the distributions. Results are obtained using Matlab 7 on a Windows 7 Operating System with Intel Core i7-870, 2.93GHz and 8MB RAM.

Playground Resolution (pixels)	Position Distribution (ms)	Intention Distribution (ms)	Threshold Selection and Application (ms)
$360 \times 558$	388.7	6283.2	21.6
$240 \times 372$	135.5	2213.3	9.7

## 5. CONCLUSIONS & FUTURE WORK

We have presented a novel physics-based approach for team behavior analysis. Given the team players' positions, we solve a particular Poisson equation on the top-view image of the playfield at each frame and then process the resultant distributions to detect the highest populated region and the region of intent at each frame. Results show that the proposed approach can extract these features successfully on the playground. In the future, we plan to equip athletes with GPS-based wearable sensors in order to obtain player locations. This will facilitate testing on significantly greater volumes of data. Our approach can be applied across different team-based sports. We have performed a preliminary validation of this on a European handball. We have also verified that the proposed Poisson equation can be used to recognize the team activities such as different types of offense and defense in a European handball dataset, which will be presented in the future.

## 6. REFERENCES

- [1] J. K. Aggarwal and M. S. Ryoo. "Human Activity Analysis: A Review", *ACM Computing Surveys*, 43(3):16, 2011.
- [2] T. Kawashima, K. Yoshino and Y. Aoki. "Qualitative Image Analysis of Group Behaviour", *IEEE Conference on CVPR*, pages:690-693, 1994.
- [3] K. Kim, M. Grundmann, A. Shamir, I. Matthews, J. Hodgins and I. Essa. "Motion Fields to Predict Play Evolution in Dynamic Sport Scenes", *IEEE Conference on CVPR*, pp. 840-847, 2010.
- [4] T. Taki and J. Hasegawa. "Visualization of Dominant Region in Team Games and Its Application to Teamwork Analysis". *Computer Graphics International*, pages:227-238, 2000.
- [5] M. Braun. *Differential Equations and Their Applications*, 4th Edition, Springer-Verlag, 1993.
- [6] T. Simchony, R. Chellappa, and M. Shao. "Direct Analytical Methods for Solving Poisson Equations in Computer Vision Problems". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):435-446, 1990.