Automatic Tactical Adjustment in Real-time; Modeling Adversary Formations with Radon-Cumulative Distribution Transform and Canonical Correlation Analysis

Amir M. Rahimi
HRL Laboratories
3011 Malibu Canyon Rd
Malibu, CA, 90265-4797
amrahimi@hrl.com

Soheil Kolouri
HRL Laboratories
3011 Malibu Canyon Rd
Malibu CA, 90265-4797
skolouri@hrl.com

Rajan Bhattacharyya
HRL Laboratories
3011 Malibu Canyon Rd
Malibu CA, 90265-4797
rbhattac@hrl.com

Abstract

In this paper we introduce two fundamentally different techniques for optimizing counter formations in team sports. In the first technique, we use canonical correlation analysis (CCA) to learn an "explicit" relationship between offensive and defensive formations. We then use the learned CCA components to make predictions about players' spatial position. Experimenting with the basketball dataset (NBA season 2012-2013) we are able to predict players' positions with high precision. In the second technique, we create an image-based representation of the player movements relative to the ball. The mentioned representation enables coaches to assess team formations in a glance. The recently developed Radon Cumulative Distribution Transform (RCDT) was used alongside CCA to analyze the image-based representations. With these techniques, we provide real-time feedback to optimize both players' positions and team formations.

1. Introduction

Recent advances in perception and computing power has opened up new opportunities for sport analysts to gain access to remarkable details about both individual and team behavior in sport events. Granted from the $600 billion industry [7], today the sport technology has mainly focused on changing the relationship between spectators and match officials. Some examples include 1st and ten line marking in American football, 3D ball trajectory in tennis, or tracking systems in soccer. Aside from facilitating the wide viewership and it's tremendous commercial potential, sports technology is entering an era that can essentially improve the evolution of sport itself. As a result, athletes' performances are improving and teams are behaving more intelligently. To this end, one major challenge in professional team sports is to automatically provide accurate and quantitative tactical feedback to coaches in real-time. Such high level analyses are complex and currently rely on the skill set of the tactical analysts. In this paper, we aim to automatically mimic the cognitive abstraction of expert tactical analysts in a team sport in order to improve the quality of novice teams.

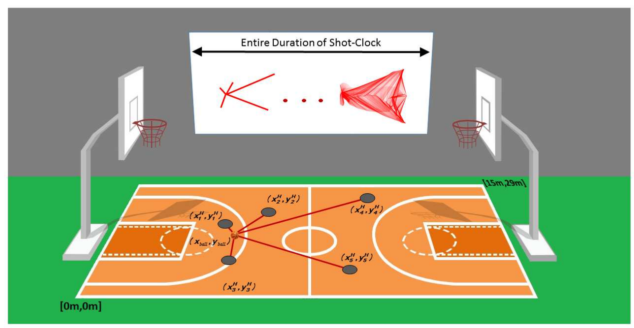

Figure 1: Player connections are shown for a team. The relative position of each player with respect to the ball is used to formulate offensive and defensive formations.

The task is to model interactions between two highly dynamic and intelligent group of individuals. Current state of the art in high-level reasoning of team behavior involves significant simplifications in order to understand the team behavior. Generally, analysis of player formation faces the problem of combinatorial permutations. Hence, recent literature have focused on projecting player movements onto a much smaller space such as "player roles" to decrease the granularity of possible interactions. For example, in basketball, researchers have decoded tactical behavior with player roles such as point-guard, shooting-guard, small-forward, power-forward, and center, or in soccer such player roles are keeper, sweeper, halfbacks, forward and striker.

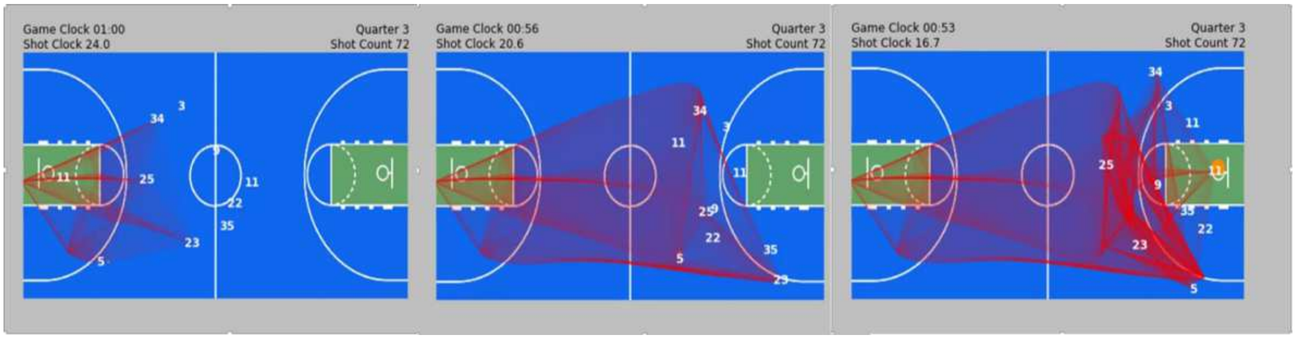

Figure 2: The progression of the players during a shot-clock and the creation of their pheromone like signature-formation.

Despite the fact that player roles reduce the problem of large permutation and also despite the fact that some 'fine' strategies can be conveyed with the use of player roles, the notion of tactics is created from a wider spectrum of interactions. Furthermore, the main purpose of tactical analysis is to adapt the team formation in order to optimize counter-attacks and maximize defense on the adversary teams. Therefore, in order to properly model each teams behavior the high-level semantics must be conditioned on the opposing team's behavior. A naive approach to this challenge is to classify team's behavior conditioned on the formation of opposing team. A better solution, however, is to find a way to relate the offensive formation to its corresponding defensive formation and perform pattern recognition in turn. This boils the problem down to modeling adversary formations relative to each other with the goal of providing an optimal response in real-time. In this paper, we focus on the game of basketball (2012-13 NBA season) with a twofold approach to this problem; We utilize canonical correlation analysis (CCA) to model the relationship between the offensive and defensive formations. In the first approach we create explicit representations of both teams' trajectories (Figure 1) and apply CCA to these explicit representations. In the second approach, we create an implicit image-based representation of the entire duration of the shot-clock (where the maximum length is 24 seconds) for corresponding pairs of offensive and defensive formations. The implicit image-based features provide rich high-level formation representations. These image-based representation are first analyzed via the recently developed Radon Cumulative Distribution Transform (RCDT) and then further analyzed by CCA to model the relationships between offensive and defensive formations.

2. Related Work

At the early stage of "machine-based" sport analytics, the main focus was to improve player detection, re-identification [5, 16], tracking [1], and activity classification and recognition [20, 3, 10]. Progress in these applications combined with recent advances in perception has paved the way for a more complex analysis of team tactics and strategies. For that purpose, the intricacy of such highly dynamic systems has led research toward simplifying assumptions such as the independence between players [18, 19, 1]. In order to better understand team behavior, Lucey et al. and Wei et al. [17, 22] proposed a role-based representation which significantly reduced high permutation in player movements. Intille et al. [9] modeled the interactions between player trajectories using a Bayesian network. Li et al. [15] used a multi-modal density function to classify different offensive plays. Li et al. [14] segmented group motion and used a spatio-temporal driving force model to identify offensive plays in American football. In soccer, Kim et al. [11] estimated the global movement of players using a dense motion field. They then looked for convergence of theses motion fields to indicate the key events. Wang et al. [21] formulated a network-flow to track all players simultaneously by considering interactions between players. Bialkowski et al. [4] used formation analysis to compare the performance of the team when the game is played at home compared to when it is played away. In most approaches the simplifying assumptions eliminate an important part of tactical behavior. In this paper, both data representations are derived from players spatial position on the court. In the first technique, we make prediction about spatial positioning of adversary players with CCA. In the second technique, we create a comprehensive image-based representation and reconstruct the offensive team's formation from the defensive team's formations. Given these representations we use CCA and RCDT+CCA, correspondingly [8, 12] to learn one teams tactical movements given the opposing team's movements.

3. Approach

Consider the problem of player position estimation in the game of basketball. In order to optimize the player positions during the game we need to have a good understanding of tactics and strategies from both teams. Before we describe

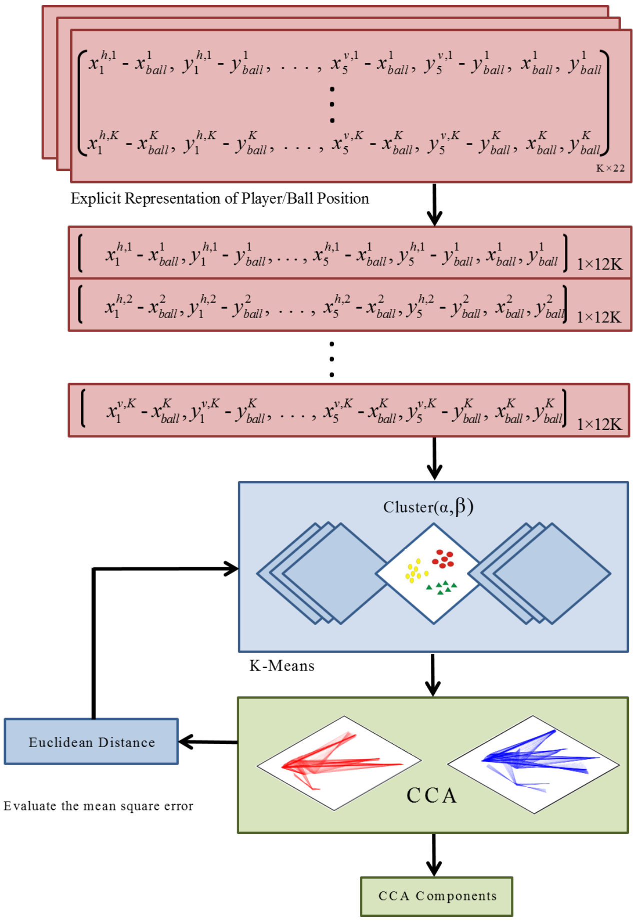

Figure 3: Explicit Method Flowchart; The explicit features are clustered and centroids are used for CCA analysis.

Algorithm 1 Explicit Tactical Analysis in Basketball

Input: Positions $x_k, y_k, x_{ball}, y_{ball}$
where $k \in \{player_1, ..., player_5\}$
Output: CCA_{comp}

for all Shot-clock Periods **do**

Generate the trajectories $hpos^n_m$, $vpos^n_m$:

$[x^h_1 - x_{\text{ball}}, y^h_1 - y_{\text{ball}}, ..., x^h_5 - x_{\text{ball}}, y^h_5 - y_{\text{ball}}, x_{\text{ball}}, y_{\text{ball}}]$
$[x^v_1 - x_{\text{ball}}, y^v_1 - y_{\text{ball}}, ..., x^v_5 - x_{\text{ball}}, y^v_5 - y_{\text{ball}}, x_{\text{ball}}, y_{\text{ball}}]$

n: Shot-clock index, $n \in \{1, ..., N\}$
m: Sample index within shot-clock, $m \in \{1, ..., M\}$

Generate pairs of tactical feature vectors $\mathbf{h}_n$,$\mathbf{v}_n$

$$H_n = \begin{bmatrix} hpos^n_1 \\ \vdots \\ hpos^n_M \end{bmatrix} \to \mathbf{h}_n = vec(H_n)$$

$$V_n = \begin{bmatrix} vpos^n_1 \\ \vdots \\ vpos^n_M \end{bmatrix} \to \mathbf{v}_n = vec(V_n)$$

end for

Calculate CCA embedding:

$$CCA_{comp} = \underset{\mathbf{u},w}{\text{argmax}} \frac{\mathbf{u}^T C_{hv} \mathbf{w}}{\sqrt{\mathbf{u}^T C_{hh} \mathbf{u}} \sqrt{\mathbf{w}^T C_{vv} \mathbf{w}}}$$

where $C_{hv} = \sum_{n=1}^N \mathbf{h}_n \mathbf{v}_n^T$, $C_{hh} = \sum_{n=1}^N \mathbf{h}_n \mathbf{h}_n^T$, and $C_{vv} = \sum_{n=1}^N \mathbf{v}_n \mathbf{v}_n^T$

Tactical Analysis with CCA:

Let $U = [\mathbf{u}_1, ..., \mathbf{u}_k]$ and $W = [\mathbf{w}_1,...,\mathbf{w}_k]$ be the top k components of CCA, then for an input tactical feature vector $\mathbf{h}$ we can predict the opposing team's reaction $\mathbf{v}$ from:

$$\hat{\mathbf{v}} = WU^T\mathbf{h}$$

our two different techniques let us briefly review the time constraints of the game. The game of Basketball consist of four quarters, the duration of each quarter is 720 seconds leading to 2880 seconds total in one NBA match. There are two processes of timekeeping in the game; First, the game clock and second, the shot-clock. Once a team has possession they have up to 24 seconds to make the shot. The shot-clock countdown resets due to various reasons including rebound, crossing over the court boundaries, or simply due to making the shot. Once possession changes, the shot-clock resets giving the opposing team a time window of 24 seconds to make their shot.

In this paper we study the relative formation of two teams during two consecutive shot-clock resets. We now describe our two techniques; the **trajectory-based** technique and the **image-based** technique.

3.1. Trajectory-Based (Explicit)

In this method, the offensive and defensive formations are captured explicitly by player positions of both teams with respect to the ball. Since the spatial formation of players is a linear representation, in the sense that a linear (or a convex) combination of these'spatial' representations could be a feasible 'spatial' representation itself, we apply CCA to model the relationship between offensive and defensive formations. In order to reduce the noise in computing the CCA components we first cluster our explicit features and then apply the CCA analysis to the centroids. The flowchart of this approach is illustrated in Figure 3.

Clustering

Given the trajectory of the team formation and the position of the ball itself, we create a $K \times 12$ matrix where each row contains the relative x and y coordinates of 5 players and the ball. We vectorize these formation matrices and cluster them using k-means [2]. The best k (for k-means) is then found with a greedy search while evaluating the performance of the CCA model.

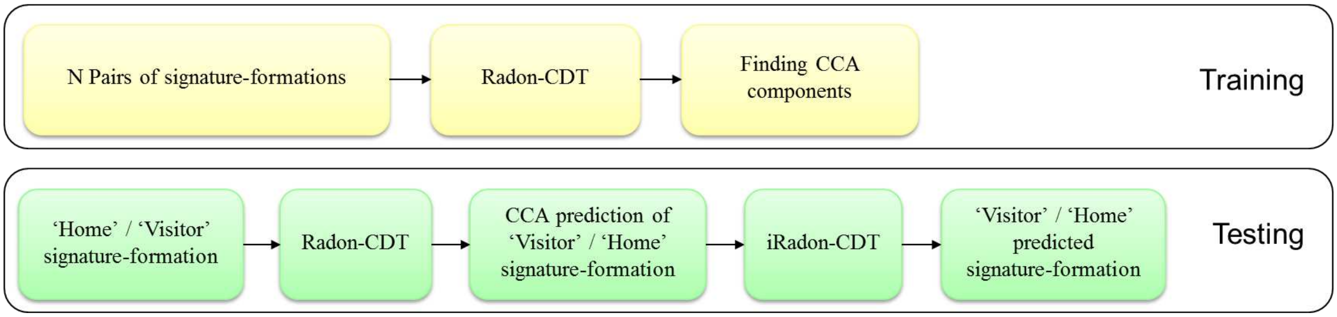

Figure 4: Information flow implicit technique.

Canonical Correlation Analysis (CCA)

In this approach, the tactical formations of the 'home' and 'visitor' teams during a shot-clock are embedded into two vectors, namely $\mathbf{h}$ and $\mathbf{v}$. Let N be the total number of tactical formations during the shot-clocks in various games such that $\mathbf{h_n} \in \mathbb{R}^{\mathbf{M}}$ and $\mathbf{v_n} \in \mathbb{R}^{\mathbf{M}}$, where $M = 12K$ is the dimension of the data observed during each shot-clock. We want to find the relationship between the 'home' and 'visitor' formations with the objective of finding a lower dimensional subspace in which the 'home' and 'visitor' formations are most correlated. In other words, the projections of 'home' formations $\mathbf{u}^T\mathbf{h}_n$ and their corresponding 'visitor' formations $\mathbf{v}_n^T\mathbf{w}$ into the shared subspace are highly correlated during each shot-clock. For this purpose we use Canonical Correlation Analysis (CCA) to encode this information. CCA seeks a shared embedding for $\mathbf{h}$ and $\mathbf{v}$ such that the embedded representations for the same shot-clock lie close to each other. In other words, CCA maximizes the following objective function:

$$
\begin{aligned}
&CCA_{comp} = \\
&\underset{\mathbf{u},w}{\operatorname{argmax}} \frac{\sum_{n=1}^{N}(\mathbf{u}^T h_n)(\mathbf{v}_n^T \mathbf{w})}{\sqrt{\sum_{n=1}^{N} \mathbf{u}^T \mathbf{h}_n \mathbf{h}_n^T \mathbf{u}}\sqrt{\sum_{n=1}^{N} \mathbf{w}^T \mathbf{v}_n \mathbf{v}_n^T \mathbf{w}}} \\
&\qquad = \underset{\mathbf{u},w}{\operatorname{argmax}} \frac{\mathbf{u}^T C_{hv} \mathbf{w}}{\sqrt{\mathbf{u}^T C_{hh} \mathbf{u}}\sqrt{\mathbf{w}^T C_{vv} \mathbf{w}}}
\end{aligned}
$$

where $\mathbf{u}$ and $\mathbf{w}$ are the CCA components which project the data onto the shared embedding and C_{hh}, C_{vv}, C_{hv} are the variance matrices. Algorithm 2 summarizes our approach for explicit tactical analysis. In subsection 4.2 we will show, through CCA, that there is a significant correlations between the offensive and the corresponding defensive formations within each shot-clock.

3.2. Image Based (Implicit)

In the second method, our goal is to predict a counter formation to that of an adversarial multi-agent system. Therefore, we propose a novel representation for the multi-agent movements over time, denoted as the "signature-formation" that captures the essence of the agent movements through time. In the following sections we present our algorithm, which takes an image-based trajectory of an adversary team and predict the 'best' response in the form of a signature-formation and vice versa. Figure 4 shows the information-flow diagram of our proposed system. The system contains a training and a testing phase. In the training phase the relationship between available signature-formation pairs are learned. The learned information from NBA players is then used to predict the most 'probable' counter signature-formation, in the testing phase.

Signature-Formation

We exploit the tactical patterns in the image domain via signature-formations. The signature-formation is essentially the temporal integration of the agents movement (without tracking agents) in each shot-clock period. In other words, the signature-formation captures a pheromone like effect of the agents' movements. Figure 2, demonstrates several snapshots during a shot-clock and the development of a signature-formation (the red pattern). In our dataset, we have signature-formations of 'home' and 'visitor' teams for nearly 10,000 shot-clocks from NBA games. Next we describe our system to predict the counter signature-formation for a given formation.

Radon-CDT

The signature-formations are two-dimensional images. These images lay on a nonlinear manifold, in the sense that the linear (or convex) combination of two images does not necessarily belong to the same set of images (see Figures 5 and 6. Therefore, a nonlinear method is needed to analyze these images. Deep convolutional neural networks and deep convolutional encoder/decoder could be used to model the nonlinearities of the signature formations. However, such methods often require millions of training samples to achieve an acceptable performance. On the other hand, recently Radon-CDT, which is a nonlinear and invertible image transformation, was introduced in the image pro-

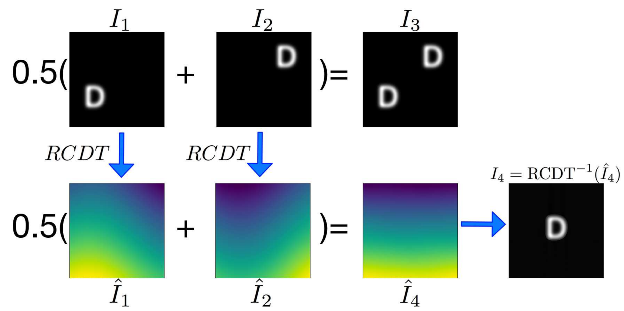

Figure 5: Linear combination of two sample images in the image space and in the Radon CDT space.

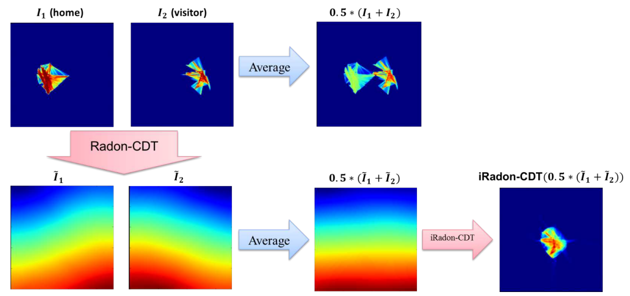

Figure 6: Averaging two signature-formations in the image space, and in the Radon-CDT space.

cessing community. Radon-CDT is a powerful and mathematically rigorous tool, which enables one to model various nonlinearities in sets of images. In order to be able to define Radon-CDT we first need to review the Radon transform [13]. For a two-dimensional image, $I : R^2 \to (0, 1]$, its Radon transform can be written as,

$$\hat{I}(t, \theta) = \int_{-\infty}^{\infty} I(x, y)\delta(t - x\cos(\theta) - y\sin(\theta))dxdy$$

where $\delta(.)$ is the Dirac function, and θ is the projection angle. Let a given signature-formation, $I : R^2 \to (0, 1]$, be normalized, such that,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x, y)dxdy = 1$$

then the Radon-CDT with respect to a normalized template signature-formation, I_0, is defined as:

$$\tilde{I}(., \theta) = (f(., \theta) - id(.))\sqrt{\hat{I}_0(., \theta)}$$

where *id* is the identity function, θ is the projection angle, $\hat{I}_0$ is the Radon transform of the template, and $f(., \theta)$ is a transport map that satisfies the following equation:

$$\int_{-\infty}^{f(t,\theta)} \hat{I}(\tau, \theta)d\tau = \int_{-\infty}^{t} \hat{I}_0(\tau, \theta)d\tau$$

Note that, since the right hand side of the above equation is a monotonically increasing function in t and the left hand side is monotonically increasing function in $f(t, \theta)$, there is a unique solution to above equation and $f(t, \theta)$ has a closed form solution for a fixed projection angle [13]. More importantly, Radon-CDT is invertible and the inverse Radon-CDT (iRadon-CDT) is defined through:

$$I = \mathfrak{R}^{-1}(det(Dg)\hat{I}_0(g)) \tag{1}$$

where $\mathfrak{R}^{-1}(.)$ is the inverse Radon transform, and $g(t, \theta) = [f^{-1}(t, \theta), \theta]^T$. For a more detailed explanation of the transform please refer to Kolouri et al [13].

The non-linearity and invertibility of the Radon-CDT enables one to apply the well-established linear modeling techniques in the transform space, and then invert the results back to the image space. To demonstrate non-linearity and invertibility of the Radon-CDT, we take the linear combination of two images in the image space and in the Radon-CDT transform space. We then invert the linear combination of transformed images back to the image space. Figure 5 shows this linear combination and demonstrates the nonlinear nature of the Radon-CDT. The first row in Figure 5 represents the image space, the second row represents the Radon-CDT space. Furthermore, we also apply the Radon-CDT to two sample signature-formations to demonstrate its applicability to more complex images (see Figure 6). Note that, the process of averaging is only used to demonstrate a linear operator (i.e. linear combination of images). All signature-formations for home and visitor teams are first transformed to the Radon-CDT space. Then the representations are vectorized and processed via canonical correlation analysis as described below.

RCDT + CCA

In our approach, the signature-formations of the 'home' and 'visitor teams during a shot-clock are first normalized (to sum to one) and processed through the Radon-CDT and then embedded into two vectors, namely h and v. Let N be the total number of tactical formations during the shot-clocks in various games such that $h_n \in R^M$ and $v_n \in R^M$, where M is the length of the vectorized Radon-CDT presentation of signature-formations. Our goal is to find the relationship between the 'home' and 'visitor signature-formations. Formally, for a given formation of the home team, h, we would like to find the most probable formation of the adversary, v. This can be achieved via CCA that seeks a shared embedding for and such that the embedded representations for the same shot-clock lay close to each other. In the training phase, similar objection function was used as 3.1. Where u and v are the CCA components that project the data onto the shared embedding and C_{hh}, C_{vv}, and C_{hv} are the covariance matrices. Let $U = [u_1, , u_K] \in R^{MK}$

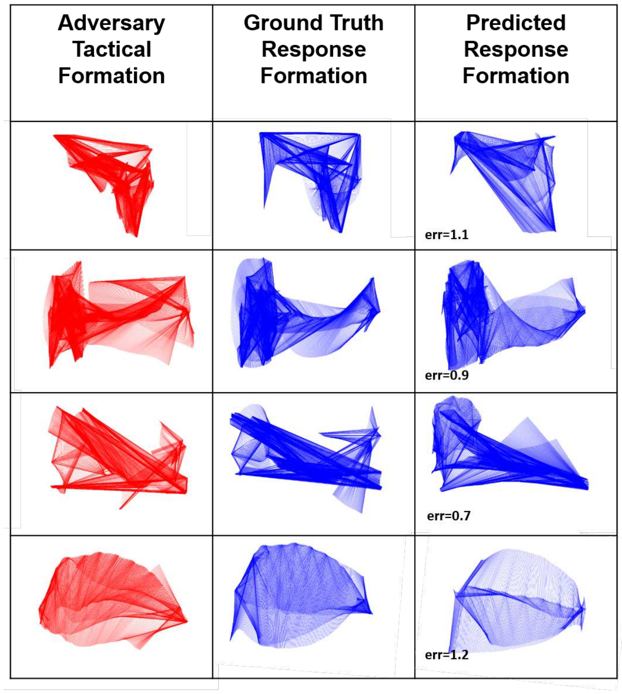

Figure 7: Prediction of signature-formations using CCA method. The first column is the input, the second column is the ground truth counter signature-formation, and the third column is our prediction.

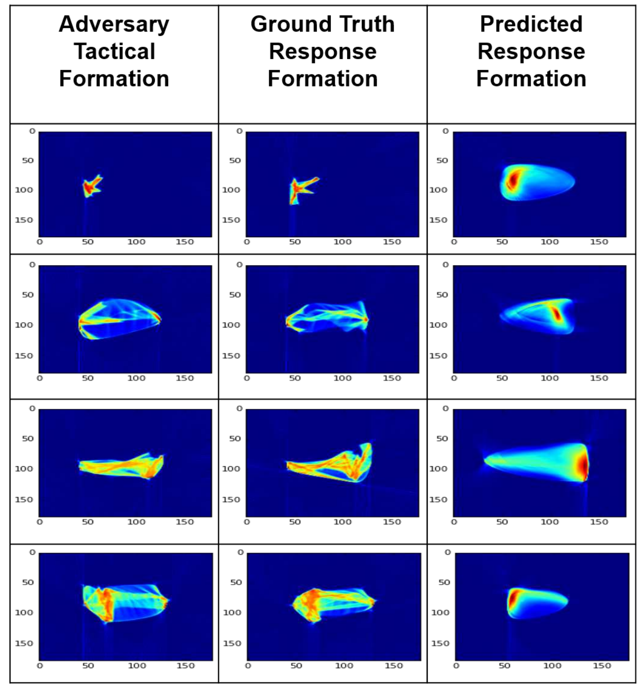

Figure 8: Prediction of signature-formations using RCDT method. The first column is the input, the second column is the ground truth counter signature-formation, and the third column is our prediction.

and $W = [w_1, , w_K] \in R^{MK}$ be the canonical component matrices, containing the top K canonical correlation components learned based on the training data (i.e. the covariance matrices where calculated based on the training data).

Predicting the Signature-Formations

In the testing phase, for an input signature-formation of the 'visitor' team (i.e. adversary), $J : R^2 \to (0, 1]$, we first calculate its Radon-CDT, $\tilde{J}$. Then the Radon-CDT representation is vectorized, $v = vec(\tilde{J})$. Next, CCA is used to predict the corresponding transformed and vectorized signature-formation, $\mathbf{h}$, as follows:

$$\mathbf{h} = UW^T\mathbf{v} \tag{2}$$

The predicted transformed and vectorized signature-formation is reshaped and then the iRadon-CDT 1 is applied to it to obtain the predicted signature-formation for the home team, I=iRadon-CDT(*reshape*($\mathbf{h}$)). This process can also be done in the other direction to predict adversarys signature-formation for a given home signature-formation. In our Reduction to Practice section we will show, through CCA, that there is a significant correlations between the offensive and the corresponding defensive formations within each shot-clock.

4. Experiments and Results

In this section we describe the details of our implementation and show our results for both techniques. We use the same train, test and validation set in both techniques to make direct comparison of the two tactical predictions. First, we describe our dataset. Next, we go over our feature representation for both techniques followed by the discussion of our results on both techniques.

4.1. Dataset

Our dataset is obtained from STATS SportsVU tracking data for the 2012-2013 NBA season. SportVU dataset includes visual data collected from six cameras installed on top of the basketball arenas. The available information contains players position, ball position, team IDs and player IDs, game clock, shot-clock, and quarter indication for 663 games across 13 NBA teams with the frequency of 25 frames per second. In our experiments we use player and ball 2D positions and shot-clocks. In the future we plan improve our methods with dependency analysis based on score and player IDs.

4.2. Feature Representation

Given our high level objective of tactical analysis, the feature extraction plays a critical role. With a simplifying assumption that a team's tactic is revealed over the entire duration of a shot-clock (which is a maximum of 24 second), and that players contribute equally to the team, our feature representation contains the player and the ball movements for the entire duration between two consecutive shot-clock resets. Given the importance of ball position and its relative distance and orientation to each player we construct our features such that in both implicit and explicit approaches our feature representations contain 'relative' distance and orientation of each player with respect to the ball. This convention provides a richer representation compared to absolute spatial positions as the tactical maneuver is highly dependant on the position of the ball at each instance.

Trajectory-Based Features

For explicit tactical analysis we create a 2 dimensional matrix that, contrary to implicit features, we explicitly encode with the relative position and orientation of players with respect to the ball during the shot-clock period. The number of rows correspond to the number of shot-clock periods and the columns of the matrix correspond to relative position of the players with respect to the ball such that x, y which are respectively the horizontal and the vertical distance at each instance (Figure 3).

Image-based Features

In implicit tactical analysis we draw a line connecting each player to the ball to create a star shaped (Figure 1) for each snapshot of the data. Next, by overlaying the updated pattern throughout the entire duration of a shot-clock we create an image for each pair of offensive and defensive formations that correspond to each shot-clock. An example of such images are shown in Figure 7.

4.3. Tactical Analysis

Trajectory-Based Tactical Analysis

Referring back to our feature representation for trajectory-based tactical analysis (Figure 3), notice that the formation of each team (offensive or defensive) was expressed with a spatio-temporal information of players thought each shot-clock. Each tactical data point is then expressed with a set of M vectors each with a size of $[1 \times 12]$ i.e. $q = [x_1, y_1, ...x_5, y_5, x_{ball}, y_{ball}]$. The parameter M was set to 10 in our experiments, which means that despite the length of consecutive shot-clock resets we take 10 equally spaced samples to encode the formation. Therefore, for each shot-clock period (complete duration of a shot-clock) we obtain a pair of offensive vectors with their corresponding defensive vectors. The results of our prediction for the trajectory-based tactical formations are shown in Figure 7.

Image-Based Tactical Analysis

This technique consists of three phases. In the first phase, our system receives the adversary's signature formation in the form of a two-dimensional heat map and applies the Radon Cumulative Distribution Transform (Radon-CDT) to the input. Radon-CDT is a nonlinear and invertible transformation that enables linear modeling of two-dimensional signature heat maps. In the second phase, canonical correlation analysis is used to predict the corresponding counter signature-formation in the Radon-CDT space. In the final phase, inverse Radon-CDT (iRadon-CDT) is used to invert the predicted signature-formation from the Radon-CDT space to the image space, and display it to the user. The result of our predictions is shown in Figure 8.

5. Discussion

In this paper we described two fundamentally different techniques to predict the multi-agent adversary movements. What we essentially learned was to model professional teams' behaviors, specifically in terms of how they react to each other's tactical movement. For the feature representation we constructed two novel representations, which contain the team interactions in time. In the first technique, the extracted features lay in a linear space, in the sense that the linear combination of two feature vectors is also a feasible feature vector. Hence, considering the linearity of features, we exploited canonical correlation analysis (CCA) to predict relative formation of the adversary teams. Experimenting with the basketball dataset (2012-2013 NBA season) we are able to predict adversary team player positions at each instance between the shot-clock resets with high precision. In the second technique we predicted multi-agent adversary movements in scenarios for which a perfect tracking of each agent at each time step is not known but the overall formation of the adversary is known. With this technique we predict the suitable counter formation given a signature-formation of the adversary team through a two-dimensional heat map.

To the best of our knowledge this work is the first attempt to exploit such high level semantics in sports. In future work, we aim to include weights in our CCA analysis such that we are able to enrich our model with the skill set of each individual player [6].

6. Acknowledgment

We would like to acknowledge R.C. Buford from San Antonio Spears and Andy Walshe from Red Bull for their constructive feedback during this research.

References

[1] S. Ali and M. Shah. Floor fields for tracking in high density crowd scenes. In *European conference on computer vision*, pages 1–14. Springer, 2008. 2

[2] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007. 3

[3] A. Bialkowski, P. Lucey, P. Carr, S. Denman, I. Matthews, and S. Sridharan. Recognising team activities from noisy data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 984–990, 2013. 2

[4] A. Bialkowski, P. Lucey, P. Carr, Y. Yue, and I. Matthews. Win at home and draw away: automatic formation analysis highlighting the differences in home and away team behaviors. In *Proceedings of 8th Annual MIT Sloan Sports Analytics Conference*. Citeseer, 2014. 2

[5] A. Bialkowski, P. Lucey, X. Wei, and S. Sridharan. Person re-identification using group information. In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*, pages 1–6. IEEE, 2013. 2

[6] F. De la Torre. A least-squares framework for component analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(6):1041–1055, 2012. 7

[7] F. K. Fuss, A. Subic, and S. Ujihashi. *The impact of technology on sport II*. CRC Press, 2007. 1

[8] D. R. Hardoon, S. Szedmak, and J. Shawe-Taylor. Canonical correlation analysis: An overview with application to learning methods. *Neural computation*, 16(12):2639–2664, 2004. 2

[9] S. S. Intille and A. F. Bobick. A framework for recognizing multi-agent action from visual evidence. *AAAI/IAAI*, 99:518–525, 1999. 2

[10] M. Kempe, A. Grunz, and D. Memmert. Detecting tactical patterns in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks. *European journal of sport science*, 15(4):249–255, 2015. 2

[11] K. Kim, M. Grundmann, A. Shamir, I. Matthews, J. Hodgins, and I. Essa. Motion fields to predict play evolution in dynamic sport scenes. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 840–847. IEEE, 2010. 2

[12] T.-K. Kim, S.-F. Wong, and R. Cipolla. Tensor canonical correlation analysis for action classification. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007. 2

[13] S. Kolouri, S. R. Park, and G. K. Rohde. The radon cumulative distribution transform and its application to image classification. *IEEE transactions on image processing*, 25(2):920–934, 2016. 5

[14] R. Li and R. Chellappa. Group motion segmentation using a spatio-temporal driving force model. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2038–2045. IEEE, 2010. 2

[15] R. Li, R. Chellappa, and S. K. Zhou. Learning multimodal densities on discriminative temporal interaction manifold for group activity recognition. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2450–2457. IEEE, 2009. 2

[16] J. Liu, X. Tong, W. Li, T. Wang, Y. Zhang, and H. Wang. Automatic player detection, labeling and tracking in broadcast soccer video. *Pattern Recognition Letters*, 30(2):103–113, 2009. 2

[17] P. Lucey, A. Bialkowski, P. Carr, S. Morgan, I. Matthews, and Y. Sheikh. Representing and discovering adversarial team behaviors using player roles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2706–2713, 2013. 2

[18] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In *2009 IEEE 12th International Conference on Computer Vision*, pages 261–268. IEEE, 2009. 2

[19] D. Tran and J. Yuan. Optimal spatio-temporal path discovery for video event detection. In *Proceedings of the IEEE Conference on Computer Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3321 – 3328. IEEE, 2011. 2

[20] H. Wang, A. Kläser, C. Schmid, and C.-L. Liu. Action recognition by dense trajectories. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3169–3176. IEEE, 2011. 2

[21] X. Wang, E. Türetken, F. Fleuret, and P. Fua. Tracking interacting objects optimally using integer programming. In *European Conference on Computer Vision*, pages 17–32. Springer, 2014. 2

[22] X. Wei, L. Sha, P. Lucey, S. Morgan, and S. Sridharan. Large-scale analysis of formations in soccer. In *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on*, pages 1–8. IEEE, 2013. 2