Visual Analysis of Player Interactions in Soccer Games

Ivona Ivkovic-Kihic
Graz University of Technology
University of Sarajevo
ivona.ivkovic-kihic@cgv.tugraz.at

Daniel Seebacher
University of Konstanz
daniel.seebacher@uni-konstanz.de

Manuel Stein
Subsequent GmbH
Konstanz, Germany
manuel.stein@subsequent.ai

Tobias Schreck
Graz University of Technology
tobias.schreck@cgv.tugraz.at

Reinhold Preiner
Graz University of Technology
r.preiner@cgv.tugraz.at

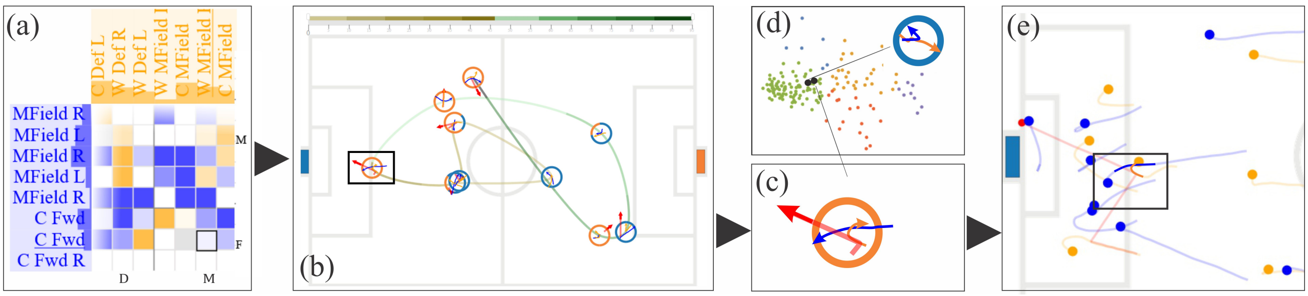


Figure 1: We develop an interactive visual analysis approach for interaction events in soccer data, which are extracted from player tracking data based on a proximity condition. An interactive framework supports explorative visual analysis based on categorical (a), spatiotemporal (b), feature-space (d), and detailed animation views (e). Based on appropriately designed visual (c) and quantitative (d) encodings of player interactions, our approach establishes suitable analysis workflows for investigating key-interactions that lead to successful outcomes – in this figure, a shot on goal (e).

ABSTRACT

Recently, visualization of sport data in general, and soccer data in particular, has received much research attention. Visual sport data exploration helps to understand behavior and performance of athletes and teams, identify possible influence factors, and changes over time, among other important tasks. In soccer match data, much of the play is determined by direct interactions between players spatially close to each other, competing for influence. We introduce a novel visual analytics system for exploring *pairwise player interactions* using a trajectory-based data representation in a highly interactive multiple view approach. Our notion of player interaction is based on proximity of pairs of players, and respective motion patterns represented as trajectories. Our approach segments player interactions from soccer match data, as the basis for linked analytical views. A matrix view allows to explore interaction frequencies between players, group of player roles, and assess overall game dominance between teams. An appropriately defined interaction glyph allows to compare interactions based on player motion, ball possession, and pitch position. We further investigate the design of a descriptor encoding the geometric configuration of interaction trajectory pairs, enabling common analytical tasks like clustering or searching for similar interactions. We demonstrate the applicability of our approach by use cases on real soccer match data, detailing the analytical tasks supported by our system.

1 INTRODUCTION

In team sports such as soccer, a fundamental task of coaches is the analysis of individual games to identify strengths and weaknesses, plan team lineups and tactics, and understand the critical key events that lead to success or failure. Besides analyzing the overall behaviour and interplay of an entire team, coaches are often interested in observing *local interactions* between individual players of competing teams, such as tackles for the ball, as the success or failure of these interactions can have a very high impact on the match outcome. The important key actions deciding the outcome of such interactions often happen within very few seconds, where the skill and training of a player can be decisive. Accordingly, coaches want their teams' tactic to be organized in a way that the strengths and weaknesses of their players regarding different interaction types perfectly outbalance the opposing team. Besides pure match analysis, a detailed assessment of a player's interaction behaviour is also important for scouting for new players that optimally complement their team.

Several previous work have investigated visual analysis approaches for interactions either based on raw motion data [21, 24], employing an interaction definition merely based on imitative motion [13], or focused on semantic aspects at the level of global team tactics [31]. However, especially within invasive team sports, considering short-time small-scale interactions and their relations is highly relevant for the analysis process. Analyzing these interactions comes with different challenges that stem from the complexity and interdependencies of their contained information. Simple pattern detection on the motion data is often not feasible for this analysis task, as the crucial motion information decisive for the interaction outcome is mostly concentrated in a very short time span, and is further dependent on additional semantic context such as the change of ball possession.

In this paper, we develop a visual analysis approach that addresses these challenges and aims at fostering an interaction-centric analysis process, allowing for investigating game-specific, player-specific and shape-based relations of individual interactions in a soccer game. To provide an accessible notion of the nature of individual interactions, we propose a suitable visual representation as glyphs, encoding both motion data and semantic context information (Section 3.2). We further investigate an appropriate quantitative encoding of its motion data, providing the analyst with a feature-related structuring of the data and giving rise to common analysis tasks like similarity assessment or clustering (Section 3.3). To foster an in-depth analysis of matches and player performances based on these interactions, a visual representation of a player-related spatiotemporal context of these Interactions Glyphs is proposed, summarizing the interaction history of individual or pairs of players (Section 4). Based on these atomic encodings, the interaction data is made accessible in an interactive analysis framework, combining means of navigation and filtering in a categorical, spatial, temporal and feature-space domain (Section 5). We demonstrate that our approach allows for quick insights into the current game situation and its relations to the performances of players within individual or groups of interactions (Section 6). In particular, our system supports a causal investigation of game outcomes, aiming for an efficient identification of key interaction events that led to a successful or unsuccessful game outcome. As a result, our approach allows for a deeper understanding of games in team sports, and gives rise to new workflows for sports-related analysis and decision-making.

2 RELATED WORK

Our work relates to several topics, including spatiotemporal visual data analysis, glyph techniques, and applications in soccer data analysis. We next discuss selected related works and how we add to it.

2.1 Visual Analysis of Spatiotemporal Data

Geospatial data arises in many areas, and to date, visual analysis of this data has received much attention. There is already a rich body of work on visual analysis of geospatial data in general [3], and movement data in particular [1]. Key analysis tasks in visual movement data include at which level of detail to describe movement, how to compare movements, and identify similarities and outliers, both for trajectories in isolation, or groups of trajectories. To date, many applications have been studied, e.g., exploration of dynamics of traffic flows [10, 23]. Also, in [25], animal movement patterns are considered. Often, functional relationships need to be considered for object movements, which may be influenced by varying environmental influences on the movement.

Besides movement in physical space, movement can also be an important factor when working with time-dependent visualizations. In [30], patterns in time-dependent scatter plot data were identified by movement analysis, e.g., allowing analysts to group similar changes and segment meaningful time intervals of the change.

In this work, we investigate a specific feature of group movement data, that is, the motion of two locally interacting entities. We analyze the interactions of such entities in terms of the geometric configuration of their trajectory intervals at the time of their encounter.

2.2 Visual Analysis with Glyphs

Glyph-based techniques are a well-known approach in visualization, often designed to give compact overviews over large amounts of data records and/or dimensions. According to [4], "Glyphs are a common form of visual design where a data set is depicted by a collection of visual objects referred to as glyphs". Different visual channels are typically used to compose glyphs, e.g., color, shape, size/height/length, orientation, texture, opacity etc. Symbolic glyphs can also be used to represent trajectories and movement [7]. Recently, Motion Glyphs [6] were introduced to show properties of large dynamic graphs. The glyph design includes an outer circle showing context of the graph, and a focal part of the graph as a node-link diagram in the center. In a case study, it was applied to sets of moving elements (fish schools), supporting analysis of leader/follower patterns among others. The Motion Rugs approach [5] is a dense visualization which provides a space-efficient overview of development of moving elements over time, supporting analysis of patterns in groups of movers.

In our work, we rely on glyphs to show trajectory interactions, using color, shape, orientation and size, as well as the outcome of an interaction in terms of change in ball possession.

2.3 Soccer Analytics

The analysis of sports data in general [9, 18], and soccer data in particular [17], has become an important application in visual data analysis. Soccer Stories [17] represents one of the first visual soccer analysis systems, giving visual designs for different soccer match situations. Interaction allows to explore soccer matches by phases and events, e.g., corner kicks, passes, dribbling etc. A large amount of work in Soccer Analytics focuses on analyzing team tactics and the global behaviour of a team [14, 16, 20, 22, 29]. Marcelino et al. [15] analyze behavior patterns of football players, measuring performance fingerprints of individuals and teams, and considering pairwise interactions to model and assess overall team performance. Similar in spirit, our work focuses on interaction pairs as the basis on which team analysis builds. In our approach we support exploration of interaction pairs by interactive cluster analysis and linked views for detail exploration. In the literature, to date many player and match features are considered for visual exploration, including free and interaction spaces [27], pressure [2], collective team movement [26], performance and event data [11], and much more. While many works consider abstract pitch and trajectory visualization, some works map derived data and visualizations onto soccer video streams, for integrated analysis. In [28], such a mapping is proposed, and shown that coupling video with visualization overlay allows for effective match context in the analysis.

For the analysis of individual player movement, an important aspect is to provide a proper visual representation and abstraction of player's trajectories [21], which also gives rise to designing suitable approaches for interactive search within the trajectory data [24]. We resort to similar abstractions for putting key interaction events between players into a spatiotemporal order. Other previous work has investigated the classification of particular match events like passes in football matches based on given spatiotemporal data [8].

In our work, interactions between players and their outcome are utilized as the key aspects of the analysis of a soccer game. We propose an appropriate design to represent these interactions as glyphs, which can be displayed in their spatial context on the soccer pitch. We are investigating a feature encoding for the trajectory footprints of two players during a mutual interaction, enabling tasks like similarity-based exploration. A proposed visual analysis system is complemented by a matrix view structuring the player interactions based on player roles, and providing an entrance point for an interactive exploration of player interactions within a game.

3 INTERACTIONS IN SOCCER GAMES

3.1 Definition and Input Data

In the context of our work, *interactions* are defined based on a set of trajectories that track the motion of players over a certain interval of time, in our case, the time frame of a soccer match. In this paper, we are using data extracted from a vision-based motion tracking technique. The data contains the position of the soccer players as well as the ball with a spatial resolution of 10 *cm* and a temporal resolution of 100 *ms*. Moreover, the data is annotated to indicate for each point in time the player that holds the ball. An interaction is defined to occur whenever the distance between two players falls below a certain proximity threshold. Starting from the point of closest distance at a reference time t_0, we extract the trajectory segments of the interacting players from the time interval $[t_0 - \Delta t, t_0 + \Delta t]$ and use these segments for visually representing the players' motion at this interaction. For the data demonstrated in this paper, we constantly use a $\Delta t = 0.75$ seconds. To obtain a robust set of non-redundant interactions, interactions with overlapping time intervals are removed, keeping the interaction centered around the minimal distance between players.

Moreover, we only extract interactions between players of opposing teams that include the ball, as these are the most important ones to affect the current and future game situation, e.g. by change of ball possession, or by preparing situations that lead to a goal. In this way, less relevant proximity interaction between players of the same team (e.g. players in a wall during a free kick scenario) are not taken into consideration. If the tracking data is labeled accordingly, we also filter out interactions that happen during non-active phases, e.g. after outs or fouls and during player substitutions. Finally, the motion data from the second half of the game is mirrored to allow a simplified spatial mapping of the player motion and their interactions in the further course of their visual analysis. Note that due to reasons of confidentiality, the datasets used in the rest of the paper are anonymized.

3.2 Visual Encoding

In order to represent and analyze a set of player interactions throughout the game, a consistent visual representation of both motion data, as well as its semantic context in the game, is needed. To serve that purpose, glyphs combining these two types of data are used. Our proposed glyph design consists of (1) an inner part representing the trajectories of two interacting players as well as the ball trajectory, and (2) an outer ring that represents the ball possession before and after the interaction phase. Players' trajectories (Fig. 2c and Fig. 2d) are plotted based on the fixed number of points registered at the same points in time for both players, with an arrowhead pointing at the direction of the players' motion. Curves representing the movements of players are colored by the players' teams. The ball trajectory is presented by a thicker, red semi-transparent line, with the arrowhead showing direction of its movement. In general, during the interaction time interval, the ball can move larger distances than the players, e.g.,

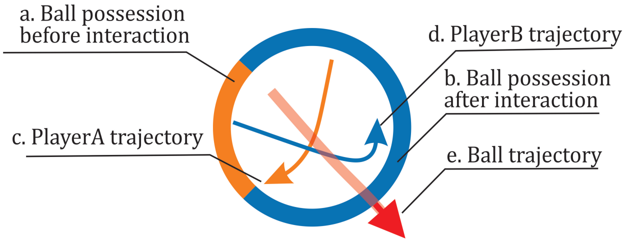


Figure 2: Interaction glyph design

when being passed or shot. In order to reduce overplotting in such cases, the ball trajectory is clipped whenever it exceeds the glyph circle, and a non-transparent arrow indicates the direction of ball movement outside the ring.

To visualize ball possession, the outer ring is divided into two parts: the first, smaller section (Fig. 2a) colored corresponding to the team holding the ball before the interaction, and the larger section (Fig. 2b) of the ring colored based on the team holding the ball right after the interaction. In this way, it can be easily determined whether the interaction caused a change in ball possession. Similarly, a uniformly colored ring indicates the continuous ball possession of the respective team. Besides the local inspection of player motion, this kind of glyph design maintains a visual overview of ball possession changes when exploring larger sets of glyphs.

3.3 Quantitative Encoding

An important aspect for analyzing the nature of interaction events between players is the ability to detect and assess similarities between different interactions, and thus allow for classifications of motion patterns whose relation to different game situations or outcomes of tackles are to be analyzed. To establish such a measure of similarity, we seek for an encoding of the common motion patterns of interacting players.

A typical approach is a quantitative encoding of motion trajectories as shape descriptors that capture the geometric features of trajectory segments [19]. However, in contrast to descriptors for single trajectories, the encoding of *pairs* of interacting trajectories in a single descriptor raises additional challenges. Besides rotation and reflection invariance, which is a commonly desirable property for shape-based descriptors, we also require it to be player and team agnostic, i.e., exchanging the motion data between involved players should lead to the same interaction descriptor. To this end, we encode the trajectory pairs of an interaction in a feature vector, capturing the most descriptive properties of the involved motion data characteristic for a tackling event: (1) the relation between the involved players' motion directions (in-sync, intercepting, or frontal approaching), and (2) the change in the scope of action for the player owning the ball, indicating how pressing the attack is. In our design, we measure these properties along 8 regular intervals during the interaction time frame, illustrated in Figure 3b. At each time step, we capture the distances d_i between player positions as well as the current motion directions $\vec{v}_i$, $\vec{w}_i$ from the i-th to the next sample point. From these data, the angles α_i between player orientations as well as the distance differentials $\Delta d_i = d_{i+1} - d_i$ are extracted and renormalized to the unit cube $[0, 1]^{16}$ based on $[-\pi, \pi]$ for angles, and on the min/max of all distance differentials in the dataset. The final normalized values are then encoded in the interaction descriptor $(\Delta d_1, \alpha_1, \ldots, \Delta d_8, \alpha_8)$.

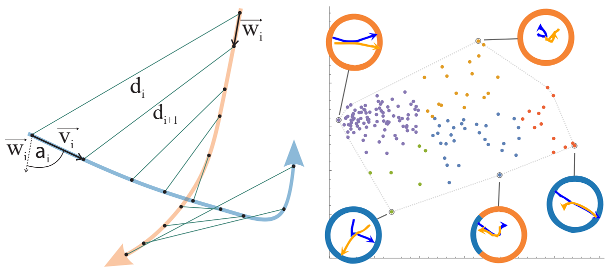


Figure 3: (a) Symmetric shape descriptor for interacting trajectory pairs. (b) Resulting feature space of interactions shown on their dominant eigenplane. The interactions exhibit a continuous distribution, smoothly varying between different motion classes shown at the outer hull.

This encoding gives rise to basic analytical tasks like clustering, performing similarity queries based on the motion characteristic, and similar. Figure 3b illustrates the interaction data extracted from a given soccer game on the dominant eigenplane of the resulting 16-dimensional feature space. The data exhibits a mainly continuous distribution of interaction data in this space, but also reveals prototypic, strongly discriminating interaction trajectory pairs around the convex hull of this subspace, as highlighted in the figure. For instance, synced parallel motion (far left) vs. frontal opposing parallel motion (far right), or cross-overs (bottom left) vs. local dribbling with U-turns (top right). Extracting such interaction *prototypes* around this convex hull enables a rough distance-based clustering (c.f. scatterplot color coding), and will later also be utilized as anchor points for an explorative visual analysis process.

4 SPATIOTEMPORAL EMBEDDING

Glyphs as described in Section 3.2 represent the two players' movements, ball movement, ball possession, and spatial placement. Glyphs as such still miss the context information like other players' movements, and the chronological ordering of the events. However, these pieces of information are crucial for understanding the context of interaction and making meaningful and comprehensive conclusions.

To this end, we design a suitable visual embedding of these glyphs into their spatiotemporal context, i.e., their position on the soccer pitch and the timeline, which at the same time establishes a connection to the involved players. Our design comprises a collection of all interaction glyphs associated with a player, or a specific pair of players, connected by a spline curve that puts them into chronological order (see Figure 4c). We therefore call this visual representation an *Interaction History Curve*. The curve segments are colored by a gradient of yellow and green for the first and second half respectively, corresponding to the timeline shown in Figure 4a and thereby establishing a visual representation of time.

In the following, we are proposing an interactive approach that integrates the visual representations developed so far into an visual analysis tool allowing for a task-oriented exploration and analysis of soccer interaction data.

5 PLAYER INTERACTION ANALYSIS

In this section, we present an interactive system that combines different views utilizing the visual and quantitative encoding of the interaction data (Section 3) as well as their spatiotemporal embedding (Section 4) enabling an explorative visual analysis workflow. In order to provide the user with a suitable overview on the data, we require different views on the interactions present in a game, providing a categorical overview, showing the frequencies of mutual interactions between players of opposing teams and their success in keeping or stealing the ball, a spatial overview to show where these interactions happened, and a temporal structure to indicate when the interactions happened. At the same time, a user might want to explore data through multiple levels of details, including: (1) an overall view of the entire spectrum of interaction data present in a single game, as well as (2) their spatial distributions and concentrations, (3) the quantitative distribution of interactions over the participating players and player pairs, (4) groups of interactions based on our feature-based cluster analysis, as well as (5) detailed analysis of interactions (details on demand).

Based on the above requirements, we design a user interface that provides these views and interactively links them to allow for an encompassing visual analysis and exploration of the data.

5.1 Interaction Matrix

The Interaction Matrix (Figure 4b) shows data at a global level and at a high level of abstraction, and is used as a starting point for player-based analysis.

Cells of the upper triangular matrix represent the overall view of the number of interactions of each player and each pair of players, respectively. Color intensity of non-diagonal cells is determined based on the number of interactions that occurred between the two players, where darker colored cells indicate that more interactions happened between the two players. Diagonal cells represent the number of interactions that one player had with all other players during the game, where more intense color means more interactions were registered, and the color is determined by the team color of the player.

In contrast, the lower triangular matrix represents the players performances in the means of ball possession changes during the interactions. Each cell is colored based on the team color of the player that had more positive outcomes. In this context, a positive outcome for playerA means that (a) playerA's team kept the ball, or (b) playerA's team stole the ball from playerB's team. A higher color saturation represents a higher percentage of interactions with a positive outcome for the respective team. The matrix is complemented by bars behind the player's names indicating the total ball possession time for each player. Combining this information with the number of interactions allows for a more distinctive assessment of the player's performance. Players in the Matrix are ordered first by the team, then by their playing positions, and different roles are distinguished by line separators and labeled accordingly. This particular ordering allows for both a player-based performance analysis as well as an overall role-based assessment between teams.

When hovering over the cells on the upper part of the matrix, corresponding cells in the lover part are highlighted, and vice versa. In this way, users can easily explore both interaction numbers and possession ratio for the selected pair of players at the same time. Finally, selecting a non-zero cell in the matrix, all

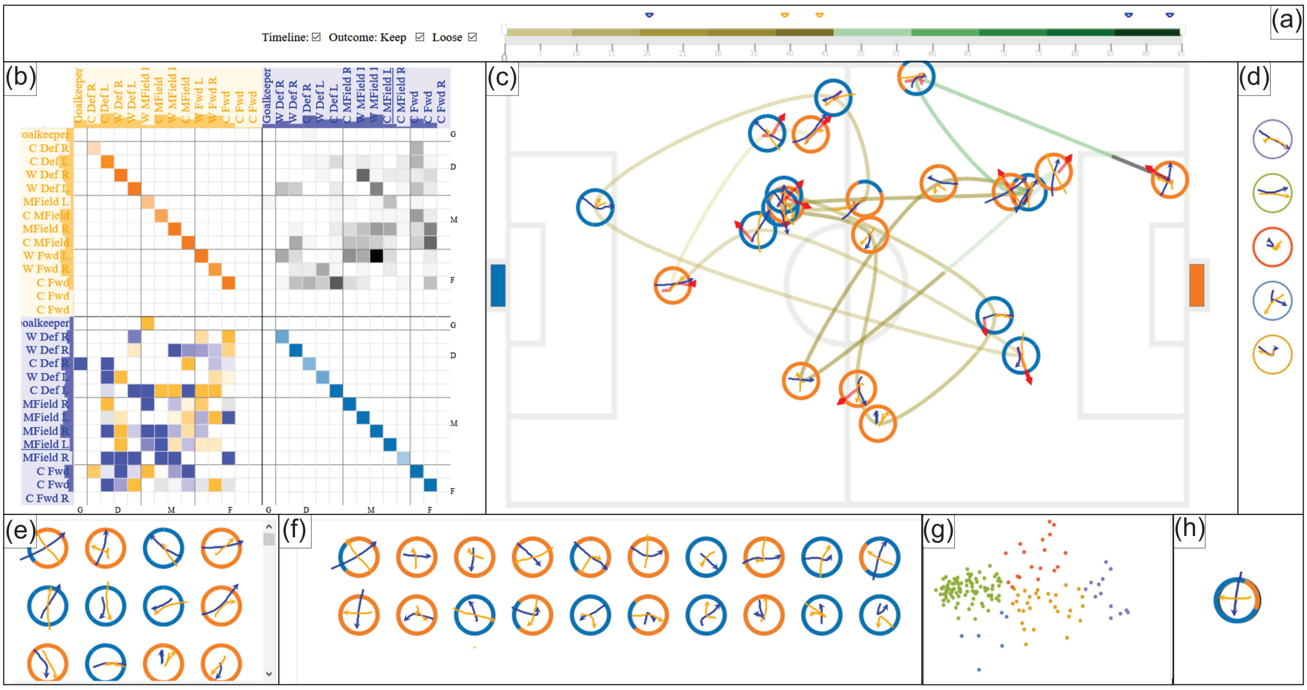


Figure 4: Overview of our visual analysis interface for soccer player interactions for a specific game and selected pair of players. (a) Timeline and Toolbox. (b) Interaction Matrix. (c) Soccer Field View. (d) Interaction Prototypes. (e) Selection grid. (f) Similarity Search Grid (g) Interactive scatter plot. (h) Zoom panel.

interactions of the corresponding player or pair of players are shown on the Soccer Field View and Similarity Search Panel.

5.2 Soccer Field View

This view (Figure 4c) shows interaction glyphs as described in Section 3.2 plotted at the spatial location they occurred on the pitch. Glyphs also serve as trigger points for the situation animation. When clicked, they show an animation of movements of all players and the ball at the time of the interaction, providing the user with a detailed visualization of the game situation. Moreover, a marker on the timeline appears to clearly indicate when the interaction takes place in the game. When activating the corresponding setting in the Toolbox, Interaction History Curves are displayed connecting all the interactions in chronological order. Glyph rings representing the ball possession are oriented according to the tangent of the Interaction History Curve at that point.

5.3 Similarity Search Panel

Selection Grid and Similarity Search Grid. A selection grid, shown in Figure 4e, represents the same set of interactions as the Soccer Field View, ordered by their time stamp. They serve as trigger points for a query search that finds the most similar interactions according to the descriptor introduced in Section 3.3. Similarity is measured by the Euclidean distance between the interaction descriptor. After selecting a query interaction, the closest interactions are plotted in the Similarity Search Grid (shown in Figure 4f). Using the query search, a user can explore similar interactions to the interesting one and search for the common behaviors in the game.

5.4 Feature Space and Interaction Prototypes

Feature Space View. Finally, we add another view on the data, by structuring them in the visualization of their feature space as established in Section 3.3. The view consists of two parts: an interactive scatter plot (Figure 4g) showing the feature space of the data, and the zoom panel showing selected interaction while the user hovers over the scatter plot (Figure 4h). This panel should provide a more elaborate view on the data records.

Interaction Prototypes. This panel, shown in Figure 4d, lists the cluster prototypes extracted around the convex hull as described in Section 3.3. These prototypical interactions can serve as a starting point for clusters exploration. By selecting a cluster prototype, the cluster members are being plotted on the Soccer Field View as well as the interaction selection grid to allow for further visual analysis. As a result, users can search for patterns in similar interactions from one cluster.

5.5 Timeline and Toolbox

The Timeline on top of the Soccer Field View represents the time of individual interactions and it is colored with the gradient of yellow and green, similar to the Interaction History Curve. The Timeline can be used to filter the time span of interactions shown on the Soccer Field View, which is also useful for reducing the plot density on this spatial view. On top of the Timeline, *goal indicators* are shown as clickable markers that trigger an animation replaying the last few seconds before the respected goal. These can be used for detail inspection of interactions that lead to a goal.

Finally, a toolbox, shown in Figure 4a is a simple set of filters and user settings made for the purpose of easier analysis of different situations. The Interaction History Curve can be switched off to reduce the visual load on demand. Outcomes can be filtered in order to show only interactions that led to a change or no change in ball possession This is used to analyze the success rate of players of teams in general.

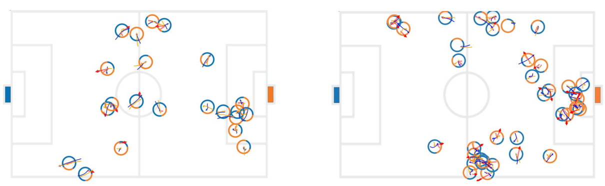
5.6 Implementation

We implemented our system as a web-based JavaScript application using D3 for the interactive matrix plot, the interaction curves, and other responsive elements. Interaction information is extracted from tracked soccer game motion datasets in an offline preprocess based on a predefined proximity range. For each interaction between two players from opposing teams that involves the ball, we store the player's trajectory segments for a fixed time interval around the point of closest proximity. Based on these trajectory pairs, we then compute the corresponding interaction descriptors and extract interaction clusters from the resulting feature points as described in Section 3.3. The resulting trajectory pairs, interaction descriptors, and cluster assignments are then loaded from the precomputed files into the framework for interactive analysis.

6 USE CASES AND RESULTS

In order to demonstrate the usefulness of our approach, we identify and discuss several typical analysis use cases. For this example, we analyzed an anonymized match from a well-known European club competition.

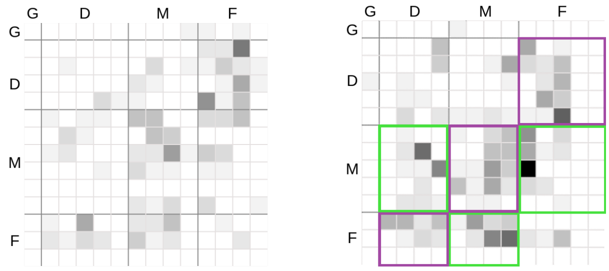
Assessing Game Dominance. An interesting finding we immediately discovered with our approach during the analysis of this match was the spatio-temporal distribution of interaction glyphs depicting transition phases in the first and second half. We compared the distribution of the glyphs in the first half of the match (see Figure 5a) and the second half (see Figure 5b), using the time range filter in the timeline interface. In the first half, a relatively even distribution of transition interacting glyphs can be observed. However, in the second half of the match, the majority of possession changes occurred in the orange team's half of the pitch. This difference in the distribution of the glyphs shows that the blue team was way more dominant during the second half. This assumption is also reflected in the outcome of the match. After a 1:2 in the first half, the blue team was able to score two additional goals in the seconds half, earning them a 3:2 score.



(a) Half 1 (b) Half 2

Figure 5: The contrast in the distribution of changes in the ball possession clearly shows that the blue team was more dominant in the second half of the game.

Role-based Assessment. The matrix-based visualization of interactions provides an overview of the respective interactions of all players in a game. The ordering of the columns and rows by team and player roles allows the analysis of substructures in the player interactions, which would be difficult or impossible to inspect with classical matrix reordering techniques. An



(a) Match 1 (b) Match 2

Figure 6: Comparing the player interactions between the first and second match shows a clear increase in the overall frequency of interactions, as well as a shift away from the anti-diagonal.

example is given in Figure 6, which compares the first match and the return match between the same teams. At first glance, it is noticeable that the number of interactions between the first and second game differs significantly. However, the particular matrix ordering also immediately reveals for which combinations of player roles these interactions have increased in particular, which allows to draw direct conclusions about the game. In a balanced game, the majority of interactions would be expected to occur along the anti-diagonal of the player-role matrix (highlighted in purple), where defenders face forwards and midfielders encounter midfielders. In contrast, the second game also shows a noticeable increase in the frequency of interactions between midfielders and forwards of both teams (green). Overall, these are indicators that the second game was more aggressive since not only the frequency of interactions increased, but the player combinations within which they happened also shifted away from the usual anti-diagonal. This might be explained by the fact that one of the teams would have went out of the competition if they would have lost this return match in the group phase.

Player-Centric Assessment. Another important analysis task is the investigation of individual player performances. For the game data investigated in this paper, the matrix plot indicates a particularly strong involvement of the orange Wide Forward Left with the Blue Wide Midfield Right player (Figure 7). A look on their common interaction history curve provides interesting insights into the interaction history between these two players. First, most interactions happened on the side of the orange team,

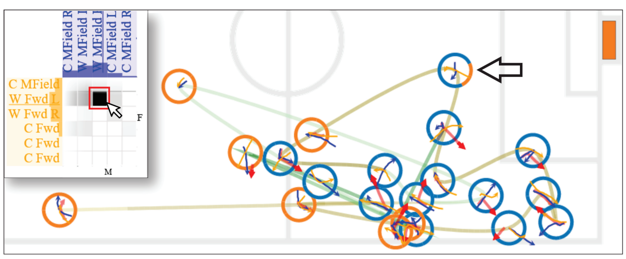


Figure 7: Interaction History curve between the orange Wide Forward Left and the blue Wide Midfield Right player over the course of the game, showing a dominance of the blue midfield player and a large number of unsuccessful tackle attempts of the orange player.

where the orange forward player was forced into a typical defender role far back in his own side of the field. Here he attempted several tackles on the pressing blue midfield player. However, as the ring colors of the interaction glyphs indicate, none of these attempts led to a successful gain of ball possession. In contrast, in a single interaction between these players, the blue midfield player even obtained ball possession from the orange forward player, in a relatively close distance to the orange goal (see black arrow). This reflects the above insights into the the game dominance in the second half of the game, on a player-focused level of detail, showing that the orange forward player was heavily outgunned by the blue midfield player.

Identifying Key Interactions. A further interesting situation we discovered during the analysis of the interaction glyphs is depicted in Figure 8. The glyph shows that the orange team lost possession of the ball in the midfield. While investigating the replay animation of this scene, we noticed that the orange player shown in the glyph first receives a pass from a player of his own team. However, the blue player in the glyph is immediately putting pressure on the orange player. The orange player's subsequent pass was a miss pass, which may have been caused by the applied pressure. After this miss pass, the blue team was able to bring the ball directly to the strikers, which created a very dangerous situation with two blue players in ball possession and a free path towards the opposing goal.

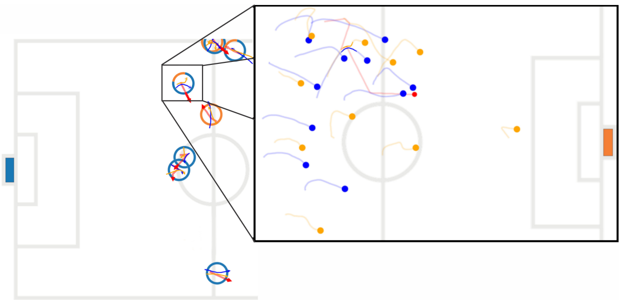
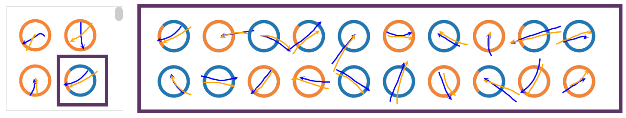
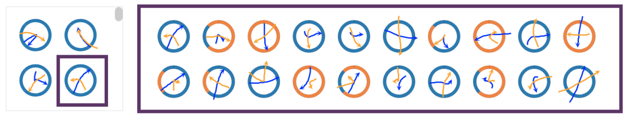


Figure 8: Loss of ball possession by the orange team results in a dangerous game situation, as there are now two blue players in ball possession with a free path to the goal.

Shape-based Interaction Exploration. In many cases, the type and interpretation of an interaction between players is directly reflected by their trajectory segments shown in the glyph. Based on a selected interaction glyph of interest, our system allows the user to investigate additional interactions of the same type. Selecting a glyph in the Selection Grid (Fig. 4e) issues a search for similar interactions throughout the game, utilizing the interaction shape descriptor developed in Section 3.3. The 20 most similar interactions are then presented in the Similarity Search Grid (Fig. 4f). These can then be further investigated by clicking them, inspecting their geographic context on the soccer pitch (Figure 4c), and inspecting them in detail using our system's animation capabilities. Figure 9 shows the retrieval results of the for two different interaction queries. In Fig. 9a, the user searches for glyphs similar to a parallel run of two players. The result set shows 20 configurations of high geometric similarity, and proofs the rotation invariance of the proposed descriptor. Moreover, the result set indicates that interactions of these type rarely lead to a change in ball possession. The query in Fig. 9b denotes an interaction of players with crossing trajectories. Again, most similarity



(a) Parallel Runs



(b) Cross-Overs

Figure 9: Similarity queries on selected glyphs yield interactions with most similar trajectory configurations, enabling further exploration of similar interactions.

results exhibit a similar cross-over shape. Comparing the sets of parallel runs and cross-over glyphs, we can also observe overall longer trajectory path lengths in the parallel runs. As trajectory segments correspond to constant time intervals (1.5 seconds in our examples), this indicates that these type of interactions are generally much more fast-paced than the cross-overs.

7 DISCUSSION AND FUTURE WORK

In the previous section, we have shown a set of basic but important analysis use cases relevant to soccer coaches and analysts. These range from an overall assessment of a team's performance, to investigating the role, importance and performance of individual players, up to detailed analyzes of key interactions and their influence on the game outcome. Our system provides several different views on the interaction data present in a game, and allows for temporal, player-related or shape-based filtering. The latter is enabled by defining appropriate shape descriptors for pairs of trajectory segments that constitute an interaction event.

Limitations. Targeting at the visual analysis of potentially complex player interactions over a whole game, our approach currently exhibits several limitations. One issue is the problem of visual clutter when many interaction glyphs are superimposed in the spatial embedding, as seen in Fig. 7. These need to be addressed using appropriate means of visual reduction, like density based scaling or visual simplification, and combined with suitable interactive detail-on-demand techniques.

Moreover, our current way of extracting semantic interactions from the raw trajectory data enforces the assumption that any interaction only involves two players. However, in general dribbling events or close-range interactions inside the penalty box, e.g. after a corner kick, more than two players can be close to or in physical contact to the ball. Our current approach would break these down to a set of two-player interactions, which is not able to visually express the complexity of the interaction in a single glyph, or encode it in a descriptor.

Another aspect is the usability of abstract views on the data provided by our system. While shape-based similarity search and data- or feature-based views are typical elements used by visual analysis experts, they might be not as intuitive to domain users like soccer coaches. However, work from other domains indicate that such elements can indeed be useful for domain experts [12].

Future Directions. Other interesting future work directions involves finding a taxonomy of interaction patterns, e.g. by enhancing the cluster analysis of interactions by visual cluster analysis tools. For instance, allowing users to label interaction examples from the PCA view and using this data to train an interaction classifier would further improve its usability.

Conceptually, besides single matches, we may also look at the differences in interaction patterns between matches of the same teams, and compare them with coaching strategies. An obvious extension of the player interaction matrix is to code it for types of interaction patterns, e.g., to observe if same players more frequently interact in similar patterns, or whether their patterns change or evolve over time. We presume that sequence mining methods can be applicable to this problem as well.

Finally, there are several possibilities to enhance our shape descriptor and glyph representation. These currently focus on the interacting trajectories between two players and the ball and are agnostic to the surrounding context such as the positioning of nearby players, which however can influence individual interactions to some degree. Including this information could therefore add to the expressiveness and suitability to capture and explain individual interactions in a game.

Ultimately, we are going to evaluate the utilization and usability of our system and its individual components, i.e., views on the data, together with end users such has soccer coaches.

8 CONCLUSION

We have shown that the visual analysis of soccer games with a focus on the interactions between the players is a promising yet barely researched approach to game exploration. Although we demonstrated our approach only on the example of soccer games, it should generalize to other team sports as well. We believe that this approach can have a great potential in helping coaches improve the performance of their teams based on the deeper insight on their weaknesses and strengths.

REFERENCES

[1] Gennady L. Andrienko, Natalia V. Andrienko, Peter Bak, Daniel A. Keim, and Stefan Wrobel. 2013. *Visual Analytics of Movement.* Springer. https://doi.org/10.1007/978-3-642-37583-5

[2] Gennady L. Andrienko, Natalia V. Andrienko, Guido Budziak, Tatiana von Landesberger, and Hendrik Weber. 2018. Exploring pressure in football. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces, AVI 2018, Castiglione della Pescaia, Italy, May 29 - June 01, 2018.* 54:1–54:3. https://doi.org/10.1145/3206505.3206558

[3] Natalia V. Andrienko and Gennady L. Andrienko. 2006. *Exploratory analysis of spatial and temporal data - a systematic approach.* Springer. https://doi.org/10.1007/3-540-31190-4

[4] Rita Borgo, Johannes Kehrer, David H. S. Chung, Eamonn Maguire, Robert S. Laramee, Helwig Hauser, Matthew Ward, and Min Chen. 2013. Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications. In *34th Annual Conference of the European Association for Computer Graphics, Eurographics 2013 - State of the Art Reports, Girona, Spain, May 6-10, 2013.* 39–63. https://doi.org/10.2312/conf/EG2013/stars/039-063

[5] Juri Buchmüller, Dominik Jäckle, Eren Cakmak, Ulrik Brandes, and Daniel A. Keim. 2019. MotionRugs: Visualizing Collective Trends in Space and Time. *IEEE Trans. Vis. Comput. Graph.* 25, 1 (2019), 76–86. https://doi.org/10.1109/TVCG.2018.2865049

[6] Eren Cakmak, Hanna Schäfer, Juri Buchmüller, Johannes Fuchs, Tobias Schreck, Alex Jordan, and D Keim. 2020. MotionGlyphs: Visual Abstraction of Spatio-Temporal Networks in Collective Animal Behavior. In *Computer Graphics Forum*, Vol. 39. https://doi.org/10.1109/TVCG.2013.192

[7] Somayeh Dodge, Patrick Laube, and Robert Weibel. 2012. Movement similarity assessment using symbolic representation of trajectories. *Int. J. Geogr. Inf. Sci.* 26, 9 (2012), 1563–1588. https://doi.org/10.1080/13658816.2011.630003

[8] Joel Estephan, Joachim Gudmundsson, Michael Horton, and Sanjay Chawla. 2017. Classification of Passes in Football Matches Using Spatiotemporal Data. (2017).

[9] Joachim Gudmundsson and Michael Horton. 2017. Spatio-temporal analysis of team sports. *ACM Computing Surveys (CSUR)* 50, 2 (2017), 1–34.

[10] Hanqi Guo, Zuchao Wang, Bowen Yu, Huijing Zhao, and Xiaoru Yuan. 2011. Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In *2011 IEEE Pacific Visualization Symposium.* IEEE, 163–170.

[11] Halldór Janetzko, Dominik Sacha, Manuel Stein, Tobias Schreck, Daniel A. Keim, and Oliver Deussen. 2014. Feature-driven visual analytics of soccer data. In *2014 IEEE Conference on Visual Analytics Science and Technology, VAST 2014, Paris, France, October 25-31, 2014.* 13–22. https://doi.org/10.1109/VAST.2014.7042477

[12] Wolfgang Jentner, Dominik Sacha, Florian Stoffel, Geoffrey Ellis, Leishi Zhang, and Daniel A. Keim. 2018. Making Machine Intelligence Less Scary for Criminal Analysts: Reflections on Designing a Visual Comparative Case Analysis Tool. *The Visual Computer Journal* (2018). https://doi.org/10.1007/s00371-018-1483-0

[13] Hoang M Le, Peter Carr, Yisong Yue, and Patrick Lucey. 2017. Data-driven ghosting using deep imitation learning. In *MIT Sloan Sports Analytics Conference.* https://resolver.caltech.edu/CaltechAUTHORS:20170316-121646643

[14] Jose Luis Sotomayor Malqui, Noemí Maritza Lapa Romero, Rafael Garcia, Hande Alemdar, and João LD Comba. 2019. How do soccer teams coordinate consecutive passes? A visual analytics system for analysing the complexity of passing sequences using soccer flow motifs. *Computers & Graphics* 84 (2019), 122–133.

[15] Rui Marcelino, Jaime Sampaio, Guy Amichay, Bruno Gonçalves, Iain D Couzin, and Máté Nagy. 2020. Collective movement analysis reveals coordination tactics of team players in football matches. *Chaos, Solitons & Fractals* 138 (2020), 109831.

[16] Daniel Memmert, Koen APM Lemmink, and Jaime Sampaio. 2017. Current approaches to tactical performance analyses in soccer using position data. *Sports Medicine* 47, 1 (2017), 1–10.

[17] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. 2013. SoccerStories: A Kick-off for Visual Soccer Analysis. *IEEE Trans. Vis. Comput. Graph.* 19, 12 (2013), 2506–2515. https://doi.org/10.1109/TVCG.2013.192

[18] Charles Perin, Romain Vuillemot, Charles D. Stolper, John T. Stasko, Jo Wood, and Sheelagh Carpendale. 2018. State of the Art of Sports Data Visualization. *Comput. Graph. Forum* 37, 3 (2018), 663–686. https://doi.org/10.1111/cgf.13447

[19] Telmo J. P. Pires. and Mário A. T. Figueiredo . 2017. Shape-based Trajectory Clustering. In *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM,.* INSTICC, SciTePress, 71–81. https://doi.org/10.5220/0006117400710081

[20] Ángel Ric Diez, Carlota Torrents Martín, Bruno Gonçalves, Jaime Sampaio, and Robert Hristovski. 2016. Soft-assembled multilevel dynamics of tactical behaviors in soccer. *Frontiers in Psychology, 2016, vol. 7, núm. 1513, p. 1-12* (2016).

[21] Dominik Sacha, Feeras Al-Masoudi, Manuel Stein, Tobias Schreck, Daniel A Keim, Gennady Andrienko, and Halldór Janetzko. 2017. Dynamic visual abstraction of soccer movement. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 305–315.

[22] Jaime Sampaio and Vitor Maçãs. 2012. Measuring tactical behaviour in football. *International journal of sports medicine* 33, 05 (2012), 395–401.

[23] Roeland Scheepens, Christophe Hurter, Huub van de Wetering, and Jarke J. van Wijk. 2016. Visualization, Selection, and Analysis of Traffic Flows. *IEEE Trans. Vis. Comput. Graph.* 22, 1 (2016), 379–388. https://doi.org/10.1109/TVCG.2015.2467112

[24] Lin Shao, Dominik Sacha, Benjamin Neldner, Manuel Stein, and Tobias Schreck. 2016. Visual-interactive search for soccer trajectories to identify interesting game situations. *Electronic Imaging* 2016, 1 (2016), 1–10.

[25] Aidan Slingsby and E. Emiel van Loon. 2016. Exploratory Visual Analysis for Animal Movement Ecology. *Comput. Graph. Forum* 35, 3 (2016), 471–480. https://doi.org/10.1111/cgf.12923

[26] Manuel Stein, Johannes Häußler, Dominik Jäckle, Halldór Janetzko, Tobias Schreck, and Daniel A Keim. 2015. Visual soccer analytics: Understanding the characteristics of collective team movement based on feature-driven analysis and abstraction. *ISPRS International Journal of Geo-Information* 4, 4 (2015), 2159–2184.

[27] Manuel Stein, Halldor Janetzko, Thorsten Breitkreutz, Daniel Seebacher, Tobias Schreck, Michael Grossniklaus, Iain D. Couzin, and Daniel A. Keim. 2016. Director's Cut: Analysis and Annotation of Soccer Matches. *IEEE Computer Graphics and Applications* 36, 5 (2016), 50–60. https://doi.org/10.1109/MCG.2016.102

[28] Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Thorsten Breitkreutz, Philipp Zimmermann, Bastian Goldlücke, Tobias Schreck, Gennady Andrienko, Michael Grossniklaus, and Daniel A Keim. 2017. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 13–22.

[29] Manuel Stein, Halldór Janetzko, Andreas Lamprecht, Daniel Seebacher, Tobias Schreck, Daniel Keim, and Michael Grossniklaus. 2016. From game events to team tactics: Visual analysis of dangerous situations in multi-match data. In *2016 1st International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW).* IEEE, 1–9.

[30] Tatiana von Landesberger, Sebastian Bremm, Tobias Schreck, and Dieter W. Fellner. 2014. Feature-based automatic identification of interesting data segments in group movement data. *Information Visualization* 13, 3 (2014), 190–212. https://doi.org/10.1177/1473871613477851

[31] Yingcai Wu, Xiao Xie, Jiachen Wang, Dazhen Deng, Hongye Liang, Hui Zhang, Shoubin Cheng, and Wei Chen. 2019. ForVizor: Visualizing Spatio-Temporal Team Formations in Soccer. *IEEE Trans. Vis. Comput. Graph.* 25, 1 (2019), 65–75. https://doi.org/10.1109/TVCG.2018.2865041