Visual Analytics for Cooperative and Competitive Behavior in Team Sports

Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von

Manuel Stein

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Informatik und Informationswissenschaft

Konstanz, 2020

Tag der mündlichen Prüfung: 2. März 2020

1. Referent: Prof. Dr. Daniel A. Keim, Universität Konstanz

2. Referent: Prof. Dr. Michael Grossniklaus, Universität Konstanz

3. Referent: Prof. Dr. Gennady Andrienko, Fraunhofer IAIS

Da ist das Ding.

(Oliver Kahn)

Visual Analytics for Cooperative and Competitive Behavior in Team Sports

Abstract

Automatic and interactive data analysis is instrumental in making use of increasing amounts of complex data. Owing to novel sensor modalities, analysis of data generated in professional team sport leagues such as soccer, handball, and basketball has recently become of concern, with high commercial and research interest. The analysis of team sports can serve many goals, for example, in coaching to understand the effects of strategies and tactics or to derive insights for improving performance. Also, it is often decisive for coaches and analysts to understand why a certain movement of a player or groups of players happened, and what the respective influencing factors were. We consider team sports as group movement including cooperation and competition of individuals following a specific set of rules. Analyzing team sports is a challenging problem as it involves joint understanding of heterogeneous data perspectives, including high-dimensional, video, and collective movement data, as well as considering team behavior and rules (constraints) given in the particular team sport. However, the discipline is in its infancy, largely restricted to commercial solutions developed out of necessity, while neglecting the movement context, with only a few academic contributions so far, and much room for improvement still exists. Consequently, the research in this dissertation happens at the intersection of several cutting-edge technologies, including computer vision and machine learning, data visualization, and human-computer interaction. All required research steps from data extraction and context enrichment to the visualization of cooperative and competitive behavior are covered in this thesis, enabling data acquisition and match analysis directly from existing video sources. The methods are capable of providing accurate analysis results both from a recording as well as in real time during a live match, improving and advancing the analytical possibilities of coaches and analysts in various invasive team sports. The impact of the presented methods is illustrated by highlighting how the application of proposed methods of this dissertation by the Austrian first league soccer club TSV Hartberg greatly improved their analysis process. Building on the foundations set by this dissertation will help to further revolutionize the way match analysis is being performed in the upcoming years. Ultimately, the progress enabled by research methods such as the introduced in-video visualization will not be limited to the domain of team sports analysis alone, but will have a general impact on how we visualize, see and perceive our data in the future.

Visual Analytics for Cooperative and Competitive Behavior in Team Sports

Zusammenfassung

Automatische sowie interaktive Datenanalyse ist von essentieller Bedeutung bei der Analyse von immer komplexer werdenden Datenmengen. Dank neuartiger Sensoren ist die Analyse von Daten, welche bei professionellen Mannschaftssportarten wie Fußball, Handball und Basketball gewonnen wurden, in jüngster Zeit zunehmend in den Fokus von kommerziellem wie auch wissenschaftlichem Interesse gelangt. Die Analyse von Mannschaftssportarten kann vielerlei Zwecken dienen, beispielsweise während des Trainings, um die Auswirkungen von Strategien und Taktiken zu verstehen, oder um Erkenntnisse zur Leistungssteigerung abzuleiten. Außerdem ist es für Trainer und Analysten oft entscheidend, zu verstehen, warum eine bestimmte Bewegung eines Spielers oder einer Gruppe von Spielern stattgefunden hat und was die jeweiligen Einflussfaktoren waren. Wir definieren Mannschaftssportarten als kooperative und kompetitive Gruppenbewegungen unter Berücksichtigung der spielspezifischen Regeln. Die Analyse von Mannschaftssportarten ist schwierig, da sowohl heterogene Daten einschließlich hochdimensionaler Video- und kollektiver Bewegungsdaten als auch das Mannschaftsverhalten beachtet werden müssen. Die Disziplin steckt jedoch noch in den Kinderschuhen und ist weitgehend beschränkt auf aus der Notwendigkeit heraus entwickelte kommerzielle Anwendungen, unter Vernachlässigung des Bewegungskontextes, mit nur wenigen wissenschaftlichen Beiträgen und noch viel Raum für Verbesserungen. Die Forschung in dieser Dissertation findet an der Schnittstelle modernster Technologien wie Computervision, maschinellem Lernen, Datenvisualisierung und Mensch-Computer-Interaktion statt. Alle erforderlichen Forschungsschritte, von Datenextraktion und Kontextanreicherung bis hin zur Visualisierung von kooperativem und kompetitivem Verhalten werden in der vorliegenden Arbeit behandelt und ermöglichen Datenerfassung und Spielanalysen anhand von vorhandenen Videoquellen. Die Methoden erlauben die Durchführung genauerer Analysen, sowohl aus einer Aufzeichnung heraus als auch in Echtzeit während eines Live-Spiels, und verbessern damit die analytischen Möglichkeiten von Trainern und Analysten in verschiedenen invasiven Mannschaftssportarten. Der Mehrwert der vorgestellten Verfahren wird am Beispiel des österreichischen Bundesligisten TSV Hartberg verdeutlicht, welcher durch die Anwendung der vorgestellten Methoden dieser Dissertation seinen Analyseprozess verbessern konnte und wider Erwarten den Verbleib in der Liga schaffte. Das Fundament dieser Dissertation wird dazu beitragen, die Art und Weise zu revolutionieren wie Spiele in Zukunft analysiert werden. Schließlich werden sich die Fortschritte, die durch Forschungsmethoden wie die eingeführte in-Video Visualisierung ermöglicht werden, nicht nur auf den Bereich der Mannschaftssportarten beschränken, sondern auch allgemein beeinflussen, wie wir unsere Daten in Zukunft visualisieren, sehen und wahrnehmen.

Contents

List of Figures

We want a revolution.

Christofer Clemens (DFB Chief Analyst)

1 Introduction

The history of modern soccer began in the 19th century in England [Gol07]. The Sheffield footballing fraternity published a book of their own rules in October 1858. Shortly afterwards, the London-based football association was founded in 1867. Nevertheless, public schools with their very own soccer rule sets disrespected the rules collected and distributed by the associations. Even worse, teams decided to play according to a mixture of several codes. In 1871, the Sheffield- and London-based football associations decided to combine their efforts and proposed a shared set of rules creating the Rugby Football Union. Still, soccer was a minority sport of the Victorian society in the early 1870s. This changed slowly with the establishment of a challenge cup of all clubs belonging to the football association and approximately 2000 spectators coming to the first finals. Many rules and soccer-related items have been developed and introduced since then. For instance, goal nets replaced two goal posts with a tape strung between back in 1892 and handling by the goalkeeper was restricted to the penalty area in 1912. In the beginning, there were only two referees being drawn from each team. Later in 1881, the third referee was added, and in 1891 he was given overall control of the soccer game.

Over the years, soccer evolved from a relatively localized minority sport to one with massive global appeal. Today, the Fédération Internationale de Football Association (FIFA) contains more countries than the United Nations. With ongoing attraction and mass marketing, the structures of professional soccer clubs evolved over the decades. Modern soccer clubs can be regarded as corporate entities, with the soccer team and its successful operation at the center. Many auxiliary and infrastructure departments in clubs provide supportive functions such as promotion of young players (club development), medical treatment of players (performance maintenance and optimization), and game analysis (for development and alignment of tactics and strategies). The game analysis department is directly connected to the coaching team and employs video analysts. The task of these experts is to identify strengths and weaknesses of their own team and of opponents, both in retrospect from historic matches, and in anticipation of upcoming matches. Their findings are used to adjust the training and thereby raise the team's awareness for dangerous situations, preparing for matches.

For decades, video analysts used video recordings of matches, which are manually processed, annotated, and edited for analysis and presentation. With advances in sensor technology, it has recently become possible to actively track player position and event data of soccer matches with high temporal and spatial resolution. Furthermore, video cameras installed on the stadium roof additionally allow passive tracking of players. Depending on availability and regulation (e.g., FIFA disallowed active tracking using sensors until 2015), either or both of these modalities can be used to capture match data. The now increasing availability of automatically recorded motion and event data enables the development of automatic data analysis methods to support the soccer analysis process. Many of the traditional video analyst tasks involve manual inspection and transcription of video material to identify scenes of interest and to gather descriptive data, e.g., on player performance. With motion and event data being readily available, one can ask how the traditional task of video analysts can be optimized, and which novel analyses can be supported.

However, while there is much progress, the analysis of football data continues to remain a very challenging endeavor as there are many different analysis perspectives for which different analysis workflows need to be defined. Even though it is rather straightforward to compute some measures, e.g., physical performance statistics (successful or failed passes, shots on goal, packing rate) of a player from motion data, it is much more difficult to automatically compute the impact of tactical and strategic aspects. This is partly due to the fact that there are different ways to quantify tactical and strategic aspects of a match. We define tactical behavior in the context of invasive team sports as the strategic organization, positioning and distribution of players during the match to pursue a common objective (to win against the other team) [SM12]. To reach this goal, players within a team need to work together cooperatively while being in competition with the players of the opposite team. For example, different coaching philosophies focus on different aspects, some of which are difficult to automatically detect and compute from data. Consequently, to identify and visualize tactical behavior of an opposing team, the gathered movement data of every single player needs to be put into context with each other, representing the collective (cooperative as well as competitive) movement throughout the match. *Collective movement* [CVV99], as a branch of collective behavior [Blu39], is described as the movement of individuals in close proximity with similar speed and direction. A particular form of collective movement can be observed in invasive team sports such as soccer or basketball where members of a team want to reach a collective goal. Players need to make decisions and develop strategies in cooperation with their team members as well as in competition with the players of the opposing team. Yet, to date, collective movement analysis mainly takes place within the scope of collective behavior analysis of animals, and, even in this domain, much more analysis has to be performed (see, for example, the 2018 granted excellence cluster *Centre for the Advanced Study of Collective Behaviour* of the German Research Foundation at the University of Konstanz).

We argue in this dissertation that, due to the many different analysis perspectives, no single preconfigured analysis or visualization will solve all data analysis and presentation tasks. Instead, we propose highly interactive team sport data analysis solutions focusing on the cooperative and competitive movement behavior through the integration of flexible data analysis methods with data visualization. These interactive solutions enable analysts from clubs (coaches, managers, players, ...) and media (journalists, ...) to steer the data analysis process by letting them control which computational methods get applied and which results they see in real time. As a consequence, these users provide context and dynamically guide the analysis process by exploring the data. Putting the analyst in full visual interactive control of the analysis process is the central idea of visual analytics [KMT09]. The research in this dissertation closes the aforementioned gaps for invasive team sports and advances the analytical possibilities of coaches and analysts by introducing novel visual interactive techniques for the analysis of cooperative as well as competitive behavior. The proposed analytical approach in this dissertation mainly focuses on the analysis of soccer matches. Nevertheless, the presented concepts can be applied to other (invasive) team sports. As a result, the introduced methods enable the identification as well as visualization of tactical movement behavior and lay the foundation for innumerable further and deeper analysis tasks.

1.1 Research Outline

Working with collective movement data is challenging because the information and interrelations are rather complex. In team sports, the movement of each player is restricted by a pitch and rules, driven by the predetermined objective, and influenced by the movement of players from both teams. These strong interdependencies result in movement patterns where every action causes a reaction [GMTR+16]. Observable group movement, therefore, contains either cooperative or competitive behavior or any combinations thereof. The data itself consists of a combination of hypervariate, hierarchical, relational, and temporal types. This leads to two challenges [KWB+19]. On the one side are the analytical aspects, which aim to find meaning inside the data, while on the other side is the visualization, not only of the data itself, but also of the information and knowledge that is derived from it. The latter enables us to support and enhance our understanding through meaningful representations and can give analytical insights, which help to reach the goals set out above.

A significant amount of research has been done on visualization techniques during the last decades. Meaningful progress has been made for the visualization of large multivariate, multidimensional, multitemporal, and spatial data sets. Information about and comparisons between them are available in the standard literature [Kei01, Kei02, HH09, dOL03]. However, while some of the more generic approaches like tables, diagrams, and 2D maps have been applied to visualize sports data, few approaches have been developed directly with team sports in mind. Instead, the state of the art in commercial systems is traditionally focused on video-based visualizations that can be used during a live broadcast. Video-based visualizations for soccer (and other team sports) have the advantage of augmenting information directly relevant to the actual scene. Furthermore, video-analysts employed by professional clubs are used to segment and analyze soccer matches manually based on available video recordings, which currently results in a very tedious and time-consuming process [Bia14]. Visualization superimposed on the original video recordings ideally enables the addition of relevant information in real time during the match, directing the analyst's attention to the essential aspects and events while maintaining the context of the real world. This enables a user to form an extremely robust mental map [BS16] which, in turn, enables a more robust contextual understanding, makes forming connections between actions easier, and improves memorability.

However, even when video-based visualizations are currently used in television recaps or analytical summaries, they are usually created and placed manually or semi-automatically at best, which requires effort and carries significant cost with it. Furthermore, video-based approaches are mostly limited to track players, show offside positions or player movements. Even these rather single-feature focused statistics and simple visualizations are an improvement over methods employed for the most parts of the last century, which involved simple counts (like number of successful passes, possession time, etc.) which have been read out or were later shown as overlay tables, supported by more detailed observations provided by subjective reports from experts. Recent progress in digital match tracking, computer graphics, and image manipulation has lead to increased capabilities in these regards, making advanced video-based visualizations technically feasible. However, the discipline is in its infancy, largely restricted to commercial solutions developed out of necessity, while neglecting the movement context, with only a few academic contributions so far, and much room for improvement still exists [TGM+17].

Team sport analysis now requires an interdisciplinary approach where Sport Science, Behavioral Science, and Computer Science including Data Visualization can all benefit to better using and understanding team sports data. Manual analyses are not feasible and fully automated methods can only be applied when the analyst knows the desired patterns in advance. Hence, supporting analysts focusing on the key aspects is necessary for a successful analysis but is also highly context-dependent and usually ill-defined. The methods contributed by this dissertation enable analysts to include their domain knowledge in order to select the crucial information pieces to validate their hypotheses and intuitions, while speeding up the analysis process and reducing tedious work. By seamlessly integrating video and visualization modalities, we enable analysts to draw on the advantages of both analysis forms and provide novel ways to project two-dimensional visualizations back into the original video recording. This bridges the gap between modern data analytics and traditional video analysis by preserving the field of view of coaches and analysts. The introduced techniques can be applied in match summarization and aggregation tasks when preparing for an upcoming match as well as during live analysis when analysis needs to be performed as time efficient as possible.

1.2 Target Audience

The methods presented in this dissertation are aimed at two target audiences, clubs (represented by coaches, analysts and scouts) and media (represented by journalists and correspondents). Both target audiences have a high level of interest in analyzing as well as visualizing match data for various reasons. Clubs are interested in visual interactive match analysis to improve the performance of their team. Media companies, such as television broadcasters, are interested in providing detailed insightful analysis to their viewers during the match or at the half-time break to extend their portfolio and, ultimately, increase sales and advertisement revenues.

1.3 Contributions

This dissertation strongly contributes to our understanding of cooperation and competition when analyzing collective movement. We make several application contributions, improving and advancing the analytical possibilities of coaches and analysts in various invasive team sports. In Chapter 7, we illustrate our application contribution by highlighting how the application of proposed methods of this dissertation by the Austrian first league soccer club TSV Hartberg greatly improved their analysis process. Besides these application contributions, this dissertation contains numerous technical contributions such as improving the overall data extraction process by providing various methods for reliable player tracking in real time from simple video recordings. Building on the extracted data, we make several technical contributions advancing our understanding of match situations by introducing novel methods measuring how movement context constrains the motion of players. We contribute techniques for the dynamic calculation of players interaction spaces, free spaces as well as dominant regions and cover shadows, illustrating how the influence of moving entities on each other can be measured and, therefore, provide the foundation for innumerable further analysis tasks. Building on this foundation, we contribute a novel method for computational and visual what-if analysis in soccer. Ultimately, this dissertation makes a strong visual contribution by introducing a framework which automatically augments original video recordings with complex and advanced visualizations.

1.4 Invited Experts for Evaluation

The approaches presented in this dissertation have been evaluated and assessed quantitatively as well as qualitatively through various expert studies. Experts share a need to analyze matches and communicate their findings and insights to players or television viewers. As all of the invited experts share encompassing professional experience in video analysis, we consider their feedback as highly relevant. Each of our invited experts, working as either an analyst or coach, typically watches and annotates up to three matches a day using state-of-the-art video tagging systems such as Sportscode [Hud19] or Dartfish [Dar19] to prepare their teams for upcoming matches. Briefly, these video tagging systems enable analysts to interactively define tags (*events*) being set when a specific key is pressed. These systems do not allow to work with the underlying movement data, so analysis is performed completely manually by watching videos.

Expert A

Expert A has been an active soccer player for 26 years and has been working as a coach for 13 years. He worked for the German soccer club FC Bayern München as a certified coach in the youth sector. A certified coach needs to be experienced in theory and practice of video analysis.

Expert B

Expert B has been an active soccer player for 21 years and is now serving as accredited referee. As an active soccer player, he regularly participated in briefings of his team, where video analysis was used to improve team performance.

Expert C

Expert C is a former professional international first league soccer player and is currently working as head coach of an Austrian first league (Austrian Football Bundesliga) team. In addition to being a certified first league coach (Coaching License UEFA A / UEFA Pro), he is also a certified match analyst.

Expert D

Expert D has been an active soccer player for 16 years and is working as a coach in the youth sector. Moreover, he is currently studying sport sciences analyzing the annotation of tactical movement behavior in soccer matches.

Expert E

Expert E has been working for the company VIZRT (https://www.vizrt.com/) for 4 years where he is the current leader of research & development of VizLibero (https://www.vizrt.com/products/viz_libero/). VizLibero specializes in manual immersive sports analytics and is used by television broadcasters to present interesting match scenes annotated with manual created visualizations.

1.5 Data

In order to address the research questions presented in this dissertation, several suitable data sets from various international top leagues were collected. The data is available in XML format and contains meta-information of the match, such as the competition and date, as well as detailed information about the teams involved and their players. The detailed information most relevant to our research are event data and position data of players and ball. The position data of the individual players and the ball are available in a time resolution from 33 to 100 milliseconds and provide at least information about their x- and y-positions on the soccer pitch. Additionally available event data are usually associated with the soccer ball, for example, when a shot, an offside, or a foul occurred. A detailed description of the data can be found in Chapter 2.1. Overall, the following match data have been used in this dissertation:

- 101 matches from the second half of the 2018/2019 season of the **Austrian Football Bundesliga**, the highest-ranking national league club competition in Austrian Football. The data includes 1080p video recordings.
- 34 matches (at the time of writing) of the ongoing 2019/2020 season of the **Austrian Football Bundesliga**. The data includes 4K video recordings.
- 60 matches from the 2014 season of teams in the highest tier of a professional football league in Asia. The data includes 480p video recordings for one-third of the matches.
- Three matches from the 2014/2015 season of a European primary football competition. For each match, a 480p video recording is available.
- Two matches from the **UEFA Champions League** between FC Bayern Munich and Manchester City of the 2013/2014 season. The data includes 1080p video recordings.

1.6 Thesis Structure

The research in this dissertation happens at the intersection of several most current technologies in computer science, including computer vision and machine learning, data visualization and visual analytics, as well as human-computer interaction. The overall structure of this dissertation (as depicted in Figure 1.6.1) covers all required research steps from data extraction

Figure 1.6.1: The content handled in this thesis ranges from the general research aspects of team sport data to data extraction, context enrichment as well as visual explanations and the discussion of future perspectives.

and context enrichment to the visualization of cooperative and competitive behavior and its assessment. The scientific foundation of this dissertation is introduced in Chapter 2. We discuss existing data sources in team sport analysis consisting mostly of statistics with very limited access of positional data retrieved from video or sensor data. By abstracting the data space, we define the various research challenges when working with team sports data as well as propose a research methodology consisting of data modeling, data mining, information visualization and visual analytics.

Building on the scientific foundation, we introduce several novel approaches for real-time data extraction of video recordings in team sports (Chapter 3). The presented methods are especially designed to be used with simple video recordings from a single moving camera in resolutions from 1280 x 720 to 1920 x 1080 pixels which so far have been most difficult to extract data from. The resulting methods are not limited to single camera recordings in low resolutions but can also be applied to multi-camera video streams with each camera having a resolution of up to 3840 x 2160 pixels. We provide and discuss several novel methods for camera-, player- and ball-tracking in real time.

Without further processing and analysis, however, the so far gathered data alone does not provide deeper insights into a match. Accordingly, the large amount of resulting movement data has to be presented through effective visualizations which allow an analyst to extract meaningful insights. In team sports, players within a team need to work together as well as against the players of the opposite team. Consequently, group movement is not random. Instead, the movement behavior of each player is strongly influenced by the context in which it occurs. After the initial data extraction, we, therefore, focus on enriching the obtained players' movement data with necessary contextual information of the expressed cooperative and competitive behavior, being crucial for a successful and effective analysis (Chapter 4). Detecting what influences moving entities in groups can help to semantically annotate or interpret relevant intentions of the collectively moving entities. We propose and evaluate several methods incorporating different aspects of context in our analysis techniques. Furthermore, we provide visual-interactive and data analysis support for annotating important types of soccer match elements, such as player interaction spaces, free spaces, dominant regions and cover shadows. In addition, we introduce a novel method that enables analysts to project every two-dimensional visualization back into the original video recording and, therefore, combine abstract visualizations with the context of the real world.

Using the identified context enables us to develop enhanced visualization techniques (Chapter 5) to retrieve explanations of observed collective movement patterns and to understand why, when and how specific movement behavior is expressed because of tactical behavior. Tactical behavior, in this dissertation, represents the overall effort practiced on the field to eliminate the factor of luck as much as possible. Towards a semi-automated analysis of tactical behavior, we introduce a novel methodology for explanatory storytelling in team sports covering classification, specification, explanation and alteration of match scenes. Building on this foundation as well as the previously introduced contextual measures, we present an approach for the automatic realization of the alteration step by implementing computational as well as visual what-if analyses in soccer. As our results show that experts most often agree with the suggested altered player movement (83 %), our proposed approach enhances the analytical capabilities in soccer and supports a more efficient analysis.

Every part of this dissertation is evaluated in detail in its corresponding section. In Chapter 6, we, additionally, compare the combined research in this dissertation to scientifically as well as commercially available tools, considering that most scientific approaches are not yet used in commercially available tools. The state of the art in commercial systems is instead traditionally focused on video-based visualizations that can be used during a live broadcast. We provide a comprehensive and categorized overview of the latest, non-trivial methods developed for video-based visualization of soccer matches and position the research contributed by this dissertation within the state of the art. We identify the approaches proposed by this dissertation, the systems Piero [Red18] and Viz Libero [Viz18] as well as, in the area of 3D reconstruction for soccer, True View [Int18] to be the most advanced methods currently available. In some sense, the main difference between the Piero system as well as the Viz Libero system compared to the approaches introduced within this dissertation is that the former more or less represent the state of the art in manually annotated video-based soccer analysis with advanced graphical display, while the latter is a fully automated approach.

In Chapter 7, we discuss whether we reached our initial goal of advancing team sport analysis by incorporating novel visual analytics techniques for the analysis of cooperative and competitive behavior. This includes media coverage of the presented work and invited talks as well as examples of professional clubs, where the work presented within this dissertation is already being applied to improve match performance. We demonstrate, for example, how the Austrian first league soccer club TSV Hartberg successfully applied methods of this dissertation in order to improve their match analysis processes. Furthermore, we present future research perspectives arising from this work including 3D reconstruction, skeleton analysis, measuring the influence of stress and perception, training management, match aggregation and team summarization as well as how current analysis processes can be improved in the future by building bridges between the data world and the real world, based on the methods presented in this dissertation.

1.7 Publications

Parts of this dissertation have been successfully published in journal and conference articles.

Journal Articles (in chronological order)

[SHJ+15] Manuel Stein, Johannes Häussler, Dominik Jäckle, Halldor Janetzko, Tobias Schreck, and Daniel A. Keim. Visual soccer analytics: Understanding the characteristics of collective team movement based on feature-driven analysis and abstraction. *ISPRS Int. J. Geo-Information*, 4(4):2159–2184, 2015.

This paper is a follow up of our VAST paper [JSS+14] on feature-driven visual analytics of soccer data. I took primary responsibility for this publication that provides a significant extension to our previous work by introducing an interactive machine learning approach for identification of soccer patterns, in combination with a novel workflow improving the analyses by a user specific learning stage. I did most of the programming with some help for the realization of the visual design by Johannes Häussler. Dominik Jäckle and Halldor Janetzko helped in the design of the workflow. I wrote major parts of the paper by myself and incorporated valuable feedback of all co-authors.

[SJB+16] Manuel Stein, Halldor Janetzko, Thorsten Breitkreutz, Daniel Seebacher, Tobias Schreck, Michael Grossniklaus, Iain D. Couzin, and Daniel A. Keim. Director's cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, 2016.

In this paper, we introduced visual-interactive and data analysis support for annotating important types of soccer match elements, such as player interaction spaces, free spaces, and pass options. The paper is a close collaboration between Thorsten Breitkreutz and myself as I supervised his Bachelor project and thesis. Thorsten Breitkreutz was responsible for the major implementation efforts and I was responsible for leading the project, the major ideas, and writing the paper. Halldor Janetzko, Michael Grossniklaus and Daniel Keim helped guiding the project. Tobias Schreck provided the first draft for the sidebox of related work in interactive data analysis. Daniel Seebacher supported me in evaluating the pass options. All authors commented on paper drafts and helped to improve the text. I wrote the major parts of the text and revised all sections several times.

[SJS+17] Manuel Stein, Halldor Janetzko, Daniel Seebacher, Alexander Jäger, Manuel Nagel, Jürgen Hölsch, Sven Kosub, Tobias Schreck, Daniel A Keim, and Michael Grossniklaus. How to make sense of team sport data: From acquisition to data modeling and research aspects. *Data*, 2(1):2, 2017.

This paper was realized in order to provide an overview about the important components of team sport data and to explain how to analyze team sport data in general, what challenges arise, and how computer science can help to cover these tasks. I was was leading the project in organizing the meetings and discussions, providing summarizations of results and in structuring the work. For validating the identified research aspects, I was supported by Halldor Janetzko, Manuel Nagel, Sven Kosub and Michael Grossniklaus. The major parts of the paper were written by myself and further revised several times by me. Sven Kosub wrote the section about Data Modeling. Manuel Nagel contributed a paragraph about the definition of behavior in our context. Daniel Seebacher, Alexander Jäger, Jürgen Hölsch, Tobias Schreck and Daniel Keim supported me to shape the paper and discussing the details.

[SJL+18] Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Thorsten Breitkreutz, Philip Zimmermann, Bastian Goldlücke, Tobias Schreck, Gennady L. Andrienko, Michael Grossniklaus, and Daniel A. Keim. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE Trans. Vis. Comput. Graph.*, 24(1):13–22, 2018

This paper presents a novel system which automatically displays complex and advanced 2.5D visualizations superimposed on the original video recordings. The paper is a close collaboration between Andreas Lamprecht and me. I supervised his Master project and thesis. Andreas Lamprecht was responsible for the major implementation efforts. I lead the project and developed the major ideas for the player detection, static camera generation and visualization integration. Thorsten Breitkreutz implemented the dominant region visualization under my supervision. Philip Zimmermann reimplemented a subset of time-critical computer vision methods in order to be running on GPUs. Gennady Andrienko contributed a paragraph in the related work for visualization for video analytics. Bastian Goldlücke supported me in writing the section about video movement tracking and visualization integration. Halldor Janetzko, Daniel Keim, Tobias Schreck and Michael Grossniklaus guided the project and commented on paper drafts. I wrote the major parts of the text and revised all the sections several times.

[SSM+19] Manuel Stein, Daniel Seebacher, Rui Marcelino, Tobias Schreck, Michael Grossniklaus, Daniel A. Keim, and Halldor Janetzko. Where to go: Computational and visual what-if analyses in soccer. *Journal of Sports Sciences*, 0(0):1–9, 2019

This paper presents an automatic approach for the realization of effective region-based what-if analyses in soccer. The paper covers the automatic detection of region-based faulty movement behavior, as well as the automatic suggestion of possible improved alternative movements. I took primary responsibility for this publication and developed the major ideas for the detection of faulty movement behavior as well as suggested alternative movements. Daniel Seebacher helped to improve and finalize our approach in many brainstorming sessions during the development phase. Rui Marcelino helped in shaping the introduction as well as the general outline for this non computer science journal. Tobias Schreck, Michael Grossniklaus, Daniel Keim and Halldor Janetzko commented on paper drafts and helped to revise all sections. I wrote the major parts of the text and revised all sections several times.

[SJKS19] M. Stein, H. Janetzko, D. A. Keim, and T. Schreck. Tackling similarity search for soccer match analysis: Multimodal distance measure and interactive query definition. *IEEE Computer Graphics and Applications*, pages 1–1, 2019

This paper is the journal version that has been published by IEEE Computer Graphics & Applications, based on the conference version [SJSK18] of the manuscript. The foundation of this paper is based on the Master project and thesis of Thomas Grießhaber which I supervised. We propose an enhanced similarity measure integrating spatial, player, event as well as high level context into the process of similarity search in soccer match analysis. Thomas Grießhaber was responsible for the major implementation efforts and I was responsible for leading the projects, the major ideas and writing the paper. Halldor Janetzko, Tobias Schreck and Daniel Keim helped with fruitful discussions and advices.

Conference Articles (in chronological order)

[SJL+16] Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Daniel Seebacher, Tobias Schreck, Daniel A. Keim, and Michael Grossniklaus. From game events to team tactics: Visual analysis of dangerous situations in multi-match data. In *1st International Conference on Technology and Innovation in Sports, Health and Wellbeing, TISHW 2016, Vila Real, Portugal, December 1-3, 2016*, pages 8:1–8:9, 2016

In this paper, we propose a set of effective visual-interactive methods for investigating set plays as first step towards semi-automated analysis of tactical behavior. The paper is partly based on a close collaboration of Andreas Lamprecht and me (I supervised his Master project and thesis). Andreas Lamprecht implemented the visual analysis and formation prediction component under my supervision. I initiated and lead the project. I wrote the major parts of the text and revised all sections several times. All authors commented on paper drafts and helped to improve the text.

[SBH+18] Manuel Stein, Thorsten Breitkreutz, Johannes Häussler, Daniel Seebacher, Christoph Niederberger, Tobias Schreck, Michael Grossniklaus, Daniel A. Keim, and Halldor Janetzko. Revealing the invisible: Visual analytics and explanatory storytelling for advanced team sport analysis. In *2018 International Symposium on Big Data Visual and Immersive Analytics, BDVA 2018, Konstanz, Germany, October 17-19, 2018*, pages 1–9, 2018

This paper proposes a four-step analytics conceptual workflow for an automatic selection of appropriate views for key situations in soccer matches. The paper is based on a close collaboration between Thorsten Breitkreutz and me. I supervised his Master project and thesis. Thorsten Breitkreutz was responsible for the major implementation efforts and I was responsible for leading the project, the major ideas, and writing the paper. Michael Grossniklaus, Daniel Keim and Halldor Janetzko contributed significantly in discussing and fine-tuning details regarding the methodology or regarding guidance during the analytic process. Christoph Niederberger supported me with his experience from industry and contributed a draft paragraph for the conclusion. I wrote the major parts of the text and revised all sections several times. Johannes Häussler, Daniel Seebacher and Tobias Schreck helped with fruitful discussions and advices.

[KWB+19] Matthias Kraus, Niklas Weiler, Thorsten Breitkreutz, Daniel A. Keim, and Manuel Stein. Breaking the curse of visual data exploration: Improving analyses by building bridges between data world and real world. In *10th International Conference on Information Visualization Theory and Applications*, 2019

This paper originates to an idea that I had while listening to the VDS panel of the IEEE VIS 2017. The panel speakers mentioned that data is always wrong to some extent, since performing an analysis on incomplete and noisy data cannot lead to fully complete and correct results. The purpose of this paper is to raise awareness of this discrepancy between the data world and the real world which has a high impact on the validity of analysis results in the real world. We propose two strategies which help to identify and remove specific differences between the data world and the real world. I took primary responsibility in initializing and supervising the project, and invited several Ph.D. students to join the development of the manuscript. Matthias Kraus, Niklas Weiler, Thorsten Breitkreutz and I collaborated together closely in many meetings during the writing of the paper. I wrote the first draft of the motivation and later on revised it several times in collaboration with Matthias Kraus. Thorsten Breitkreutz wrote the first draft of the related work. Matthias Kraus and Niklas Weiler provided the draft of Section 3 and Section 4 as well as provided the first draft of Figure 1. Matthias Kraus described the first strategy in Section 4.1 about reconstructing the real world while I wrote Section 4.2 about projecting results back into the real world as second strategy. Thorsten Breitkreutz provided the use case about Collective Behavior and Matthias Kraus about criminal investigation. We all commented on paper drafts and worked together to revise and improve the text.

[SSK+19] Manuel Stein, Daniel Seebacher, Tassilo Karge, Tom Polk, Michael Grossniklaus, and Daniel A. Keim. From movement to events: Improving soccer match annotations. In *MultiMedia Modeling - 25th International Conference, MMM 2019, Thessaloniki, Greece, January 8-11, 2019, Proceedings, Part I*, pages 130–142, 2019

This paper introduces a novel method for the semi-automatic definition and detection of events based entirely on movement data of players and ball. This paper is a close collaboration between Daniel Seebacher, Tassilo Karge and me. Daniel Seebacher and I supervised his Bachelor project and thesis. I was responsible for leading the project and the foundation for automated event annotation in soccer matches. Daniel Seebacher had the initial idea to incorporate Allen's interval algebra into a visual analytics system in order to enable analysts to visually define as well as search for complex, hierarchical events. In addition, Daniel Seebacher wrote the draft for the Section about detecting complex events. Tassilo Karge was responsible for the major implementation efforts. Tom Polk, Michael Grossniklaus and Daniel Keim commented on paper drafts and helped to improve the text. I wrote the major parts of the text and revised all sections several times.

[FKS19] Maximilian T. Fischer, Daniel A. Keim, and Manuel Stein. Video-based analysis of soccer matches. In *2nd International ACM Workshop on Multimedia Content Analysis in Sports (ACM MMSports'19)*, 2019

This paper provides a comprehensive overview and categorization of the methods developed for the video-based visual analysis of soccer matches. The foundation of this paper is based on an internal seminar report of Maximilian Fischer which I supervised and where I provided the initial research question. Maximilian and I worked closely together when writing the follow up paper, revising every section. Daniel Keim helped with fruitful discussions and advices.

I believe in work, in connections between the players, I think what makes football great is that it is a team sport. You can win in different ways, by being more of a team, or by having better individual players. It is the team ethic that interests me, always.

Arsene Wenger (Coach Arsenal F.C., 1996–2018)

2

Research Aspects of Team Sport Data

Contents

Recent progress in sensor development results in increasing interest in recording and analyzing movement in team sports. In this dissertation, we focus on team sports that can be classified as invasive team ball games with two opposing teams competing against each other and trying to score more points than the opponent to win a game. We have chosen this specific focus for two reasons. First, while the interaction of opposing teams in invasive team sports makes the analysis of sports data more challenging, it also opens up more opportunities for findings. Second, many of the world's most popular team sports, e.g., soccer, football, basketball, hockey, rugby, handball, etc. are invasive. Due to the popularity of these sports, the availability of corresponding data sets and the interest in their analysis are currently on the rise. Professional team sport companies invest substantial resources to analyze the own team's performance as well as the performance of future opposing teams. Various aspects and several data sources are important descriptors for the performance of a team. In practice, some of these data sets are kept confidential by respective stakeholders, e.g., when they contain exact movement trajectories. Other data sets, e.g., basic statistics, are publicly available for analysis purposes (see Section 2.1.1 for several examples). Depending on the available data, different analysis tasks can be executed. Analysts usually do not only want to have information about the *what* (e.g., "Team A won against Team B" or "Player X passed more often than player Y") but instead want to investigate the *why* behind these facts. There is a need to understand why a certain movement happened and what the influencing factors were. For example, why did a player decide to move to Point A instead of Point B and what influence did this movement decision have on members of the own and opposing teams. The results of such analyses will help, e.g., in scouting or training. However, analysis often focuses on pure statistical approaches. For decades, movement and tactical analysis has been done manually by inspecting video recordings of past matches.

In this chapter, we show how this gap can be closed by providing an overview of how to work with team sport data in the future. We introduce the various data types that are available and relevant for team sport analytics and highlight the challenges that need to be overcome when gathering and working with team sport data. We focus on the different research aspects arising as displayed in Figure 2.0.1 with respect to the set of heterogeneous data. Figure 2.0.1 shows that the arising research aspects can be grouped into specific domains. Data acquisition describes what needs to be done at first to get the data, e.g., through video processing. The context domain allows us to enrich the data (e.g., through data fusion) after the acquisition with useful additional information. After data acquisition and enrichment the analysis domain allows us to search for patterns. The resulting team sport analysis is performed on the basis of high-dimensional data that contains time series as well as trajectory data. Throughout this dissertation, we use soccer as a prime example for our proposed methodology being a highly popular team sport. Nevertheless, we present a general overview of data, methods, and tasks that are applicable to all invasive team ball games. We contribute a concise description of the enablers for data-driven sports analysis. We start with describing relevant sensors and data sources. We show which general computer science problems can be addressed while working on team sport data and propose a methodology to handle these challenges. By abstracting the data space, we reveal general research aspects being related to data modeling, data mining, information visualization and, especially, visual analytics.

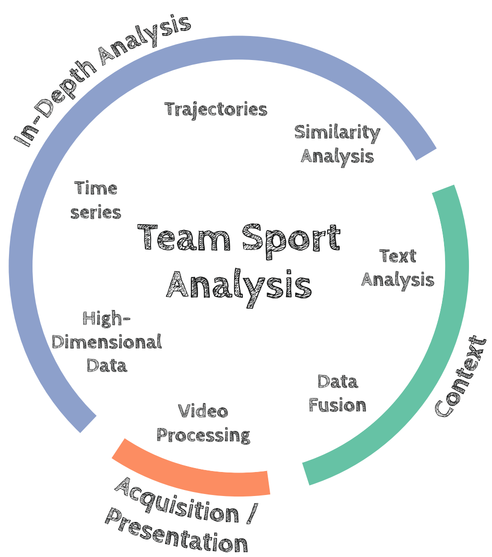


Figure 2.0.1: From acquisition (e.g., video processing) and data enrichment (e.g., data fusion) through context information to in-depth analysis tasks (e.g., trajectory analysis), many research fields are covered when analyzing team sport data.

2.1 Team Sport Data

Team sport data can either be characterized by describing different technical aspects such as the used acquisition method (e.g., optical tracking, local positioning systems, triangulation, or manual recording) or by discussing the data types arising from different data sources. In the following, we introduce various data sources and describe them in detail highlighting important technical aspects. Practically, most data in invasive team sports (like player movement, events, and descriptive statistics) are extracted from video and sensor data. Relevant information can also be obtained from live streams in social media channels, which may reveal interesting facets of the game from an audience perspective. The latter is again presented in heterogeneous data formats; depending on the social media channel, this may assume the form of text, images, or video feeds. In the following, we elaborate each data type with information about data characteristics such as size, accuracy, and resolution. Furthermore, we give an overview with respect to where such data can potentially be obtained from (i.e., which company offers which service). Whenever applicable, we introduce possible architectural requirements, e.g., information about hardware or analytics supporting databases and index structures. We believe that the most comprehensive analysis covers and combines all of the mentioned data types below.

2.1.1 Video and Sensor Data

Video recordings are ubiquitous in invasive team sports and there is an increasing demand for professional analyses performed on them. Video recordings range from television recordings from mass media with various perspectives (e.g., Sky TV has 24 cameras on the ground for soccer matches [Rya10] while for the NFL up to 70 cameras are used during a super bowl match [Gla16]) to professional recordings carried out by the teams themselves. Consequently, video data generally can be considered as the most available data source. Companies such as STATS [stab] or Opta [opt] offer the service of extracting movement, event, and statistical data based on their own recordings. Working directly with video data is much cheaper than assigning a professional company to track a team's players, events, and statistics. However, the extraction of player movement from video recordings is a non-trivial task. Nevertheless, recent publications showed the feasibility of extracting movement data based on video sources [SJL+18, SCKH97, LTL+09, PHVG02].

Player movement can also be recorded by attaching sensors directly to players or game objects (e.g., ball, sidelines, targets, etc.). Catapult [catb] is one of the well known companies in the field of such tracking devices. Practically, the applicability of this acquisition modality is depending on the legislation adopted by the sports associations. This data acquisition may be partially restricted in some invasive team sports (as it was for example in soccer until 2015) while allowed for others. For instance, the NFL [Pel14] allows active tracking by sensors placed on the player shoulders in cooperation with Zebra Technologies [zeb]. Sensors may allow real-time capturing of data via wireless data transfer. The data set of the ACM DEBS 2013 Grand Challenge is a perfect example for these kinds of sensor data [MZJ13].

Movement Data

Gathered movement data describes where a player or game object is located at a specific point in time. Locations are usually measured by local coordinate systems with reference to the game pitch. These measurements contain as a minimum the x- and y-coordinates and sometimes also the z-coordinate. Depending on the acquisition technique, the positions are usually sampled around 10 to 25 times per second (Hz). In the sample data set of the ACM DEBS 2013 Grand Challenge, each player has two sensors (one in each shoe; the keeper additionally has sensors in gloves) that each transmit position reports at 200 Hz. The ball contains one sensor that transmits at 2 kHz. For each sensor, a timestamp (X, Y, Z)-position, as well as both overall and component-based velocity and acceleration is retrievable. Storing a full game results in approximately 70 MB to 10 GB of data. Body postures (player skeleton data) are not recorded in the sample data set due to the large amount of active sensors needed for robust posture recognition. We introduce a novel method for body posture detection in Chapter 3.3.1.

Event Data

Sport games can be described by an ordered sequence of events. We define events as match-relevant actions that happen during the match. Events can be derived from movement data by automatic video analysis [SSK+19, TQ01, ETM03, XCDS02, ABC+03, XZZ+08]; also manual annotation is possible and done professionally by some data providers such as Opta. From a technical perspective, events are timestamped occurrences of previously known and defined categories, optionally annotated with spatial coordinates or additional information as involved players. Most events are directly ball related and correspond to actions with the ball (for instance passes or dribbling). Other events may be time-dependent (e.g., start and end of a play period) or not directly dependent on the ball (e.g., a foul situation during a free kick). The resulting streams of semantic data are already widely used in industry and scientific communities [NWCP07, RLRM02, KTAP+95]. In practice, events might lack in accuracy, as they are usually annotated manually or as fully automatic recognition may produce false positive and negative events. As event data mostly contains information about players interacting with the ball, event data enables to conduct overall game statistics (e.g., passing networks, pass accuracy, or time between gaining the ball and shot on target). We give an example of a set of potentially relevant events for soccer analysis in Figure 2.1.1. Although arguably it may be extensible, and is not tailored towards a specific model from Sport Science, we believe it is a practical starting point for reasoning about the types of events potentially useful for analysis. The table contains the type of event, a short description of when the event is recognized, as well as a proposed categorization of events that share similar characteristics. We distinguish the following event categories:

- ***Rule-induced events*** are events that occur as a result of the match rules. For example, if the ball passes the sideline of the soccer pitch, it has to be thrown in again by the opposite team.
- Events tagged with ***prosecution*** indicate that there was a foul behavior of the related player(s) which is penalized.
- ***Player interactions with ball*** is about events that happen when a player is touching the ball. Observable, almost every event that gets tagged falls under this category besides yellow and red cards, the end of a halftime and a substitution.
- Events that interrupt the match gets marked as ***gameplay interruption***.
- If an event has a direct relation to scoring (e.g., a shot on the goal) we mark it as ***scoring related***.

Events from different categories may lend themselves to different analysis tasks. Also, they are subject to possible detection or description inaccuracies and may require parameters to work. For example, *running with ball* requires an appropriate threshold setting for distance between player to ball to be recognized. We observe that events may belong to different categories mandatory or optionally. In our table, we mark events dark blue if it always belongs to the respective category. If an event is marked in lighter blue instead, the event may falls under this category, but does not have to.

rule-induced events	prosecution	player interactions with ball	gameplay interruption	scoring related	Event Type	Description
applies always	applies always	optional	applies always	applies always	Foul Penalty	Free kick on the goal defended only by the goalkeeper
applies always	applies always	optional	applies always		Foul direct free kick	Free kick that is allowed to be directly shot into the goal
applies always	applies always	optional	applies always		Foul indirect free kick	Free kick that is not allowed to be directly shot into the goal
applies always	applies always	optional	applies always		Foul throw in	Throw in that is not correctly executed
applies always		optional			Halftime Start	First or second half starts
applies always		optional	applies always		Offside	Player is in an offside position
applies always		applies always	applies always		Out for goal kick	Ball passes the endline after an opponent touched it
applies always		applies always	applies always		Out for corner	Ball passes the endline after a player from the own team touched it
applies always		applies always	applies always		Out for throw in	Ball passes the sideline of the soccer pitch
		applies always	applies always	applies always	Goal	Awarded when the whole of the ball crosses the whole of the goal-line
		applies always	optional	applies always	Shot on target	Any shot attempt that would or does not enter the goal if left unblocked
		applies always	optional	optional	Shot not on target	Any shot attempt that would or does enter the goal if left unblocked
		applies always			Pass	Ball touch from one player with direction towards a team mate
		applies always			Reception	Ball touch made by the player after receiving it from another player
		applies always			Clearance	Hard ball touch where the player tries get the ball away from the current zone on the pitch
		applies always			Hold of ball	Play action when the keeper takes the ball with his hands without danger
		applies always			Running with ball	Used by the player to move the ball around without passing it to another player
		applies always			Cross	Hard ball touch where the executing player is positioned in the final third of the field
		applies always			Neutral contact	Characterized by ball touch which is difficult to control
		applies always			Pass assist	The last pass to a teammate in a way that leads to a goal
		applies always			Cross assist	The last cross to a teammate in a way that leads to a goal
		applies always			Catch	Keeper catches the ball and hold it in his hands on a dangerous situation
		applies always			Catch drop	Keeper does not manage to hold the ball but lets it bounce of the hands again
		applies always			Drop of ball	Keeper drops the ball after having caught it or holds it in order to play the ball
		applies always			Punch	Keeper punches the ball with his hand away
		applies always			Diving save	Keeper jumps to a side to catch the ball
		applies always			Drop kick	Kicking a ball that is dropping to the ground as it starts to bounce up
applies always	applies always		optional		Yellow card	Displayed by referee to indicate that a player has been cautioned for a foul
applies always	applies always		applies always		Red card	Displayed by referee to indicate that a player has been dismissed from the field for a foul
applies always			applies always		Halftime Ends	First or second half ends
applies always					Substitution	Replacing one player with another during a match.

applies always   optional

Figure 2.1.1: Various kinds of events categorized by their characteristics

Descriptive (Statistical) Data/Derived Data

Properties of a player or a team can be characterized by descriptive (statistical) data. Descriptive data include everything that can be counted or measured during one or several matches, for example, how often a player passes or the maximum speed and acceleration. Today, some of these descriptors can be measured automatically by tracking devices. However, most of this data is collected manually by analysts. Due to historical reasons, descriptive statistics are the most common data sources for team sport analyses, as automatic movement recording is a relatively new technical achievement. Data sets are available to access historical data such as prior match results between the teams, squad-memberships, final standings of leagues and career records. Beside several commercial providers and unstructured text websites, freely available and machine readable sources can be found online [ope, soc, foob, fooc, fooa]. Bergmann et al. [BBE+13] show how such data sources can be matched and stored.

2.1.2 External Factors

External and environmental factors can be obtained by matching game data with external data sources, such as location, date of a match or characteristics of the stadium (capacity, open/closed roof) [staa]. Historical and current weather records [nnd] can be used to heuristically estimate the quality of the playing field and possible effects on player performance due to high or low temperatures and air humidity. Furthermore, Ekin et al. [ET03, ETM03] showed that weather and stadium context can influence the effectiveness of video based tracking and needs to be taken into account.

News and Social Media

Community-generated reports about games can be gathered from social media platforms such as Twitter [twi], reddit [red], or Wikipedia [wik]. Yucesoy et al. [YB16] investigated, for example, the relation of tennis players' popularities and their performance based on visits to their Wikipedia entries. Users write before, during and after a team sport event. Many of these services provide APIs to gather at least partial data sets for matches from the past and in real time. Works from traditional news sources are collected in several projects similar to the European Media Monitor [emm]. A real-time stream of worldwide journalistic articles can be retrieved with additional metadata and event detection added by the data provider.

2.2 Abstracting the Data Space

We described the multitude of facets and aspects of data in team sports and pointed to the heterogeneity of recorded team sport events. We now investigate the nested abstraction levels depicted in Figure 2.2.1 following a bottom-up approach starting with team sport and going all the way to the basic data types, namely geospatial and temporal data.

In team sport data, two competing groups are represented that have opposed predefined objectives, meaning that (if the match did not end in a tie) only one of the two groups can achieve their objective and the other group loses the match. The challenge of analyzing team sport data is that movement is restricted by a pitch and rules, driven by the predetermined objective, and influenced by the movement of own and opposing team players. For illustration purpose, we exemplify these properties in American football: the movement of players and teams is limited to the pitch. The movement of the two opposing groups is clearly driven by a predetermined goal. The group possessing the ball wants to cross "the opposition's goal line with the ball, or catch or collect the ball in the end zone" [BBC]. The counter-objective of the opposing team is to prevent this from happening and to gain possession of the ball. American football is a very good example to illustrate how groups and individuals influence their movement mutually. Examples include the defensive line trying to block the running back (group influencing an individual and vice versa), the offensive line pushing against the defensive line (group influencing group) and corner backs covering the receivers (one individual influencing one individual). Rugby is another example for team sport being analyzed nowadays as shown by Cintia et al. [CCP16].

One abstraction level higher, team sport can be seen as a specialization of group movement. Group movement can contain either cooperative or competitive behavior or any combinations of both, being exactly what can be observed in team sport. On the highest abstraction level, movement in general is defined according to Andrienko et al. [AAB+13a] as the path of moving entities through space and time. Andrienko et al. [AAB+13b] suggest analyzing movement at two different granularities, of the individuals and of the group as a whole. Nevertheless, we also have to deal with pure time-series or time-stamped event data and with spatial topology data. There exists many works dealing with pure time series analysis mostly focusing on similarity calculation and pattern analysis. A good overview over state-of-the-art methods in temporal data mining is given by Fu [Fu11] and visualization methods for time series are discussed for instance by Aigner et al. [AMST11]. Temporal analyses alone cannot explain all the behavior observed as important spatial aspects are neglected. Nevertheless, temporal visualizations are crucial to convey temporal patterns. From a pure spatial perspective, computational geometry can be the starting point to analyze trajectories in team sports data as described by De Berg et al. [dB00]. However, pure geometric approaches do not cover the temporal aspects of movement. Consequently, spatio-temporal analysis are key for a successful analysis. For instance, Kang et al. [KHL06] represent the movement of soccer players as trajectories and propose a model which quantitatively expresses the performance of players based on the relationships between the player trajectories and the ball. Another player network based analysis with respect to performance is performed by Cintia et al. [CGP+15, CRP15]. Football strategies are investigated based on network theory analysis by Pena et al. [PT12]. A different spatial topology based approach would be the use of graph theory in sports. Bourbousson et al. [BPSS10] analyze basketball matches and Clemente et al. [CCMM15] use graph theory to analyze soccer matches.

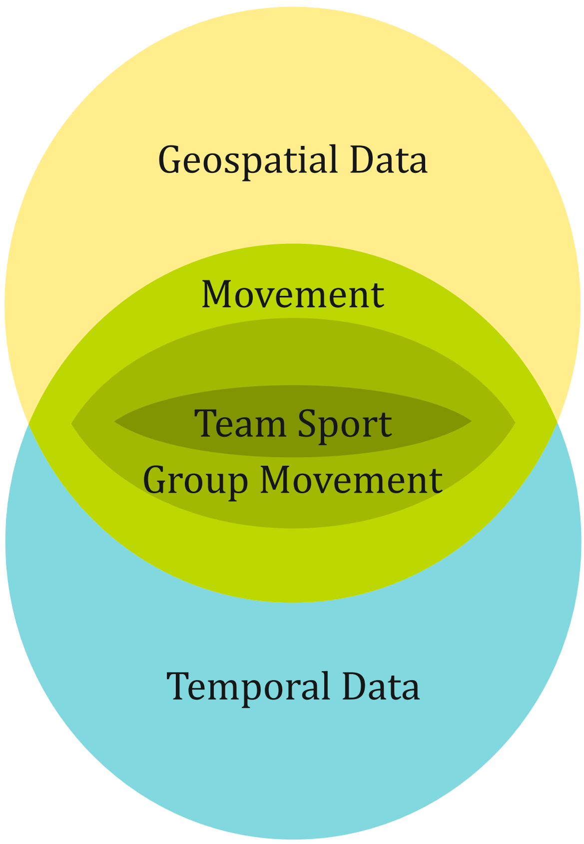


Figure 2.2.1: The abstract ingredients of team sport

Statistic	Brazil	Germany
Goals	1	7
Possession	52%	48%
Shots	18	14
Duel Quota	51%	49%
Packing	341	402
IMPECT	53	84

Table 2.3.1: Statistics of a real-world game: Brazil vs. Germany (2014 FIFA World Cup).

2.3 Research Challenges

Analysts inspecting team sport data usually pursue various goals such as identifying strengths and weaknesses of their own team and of opposing teams. The insights are used to improve training and raise the team's awareness when preparing for upcoming matches. To identify strengths and weaknesses, analysts, for example, want to understand why a team won a past match. Widespread statistical approaches, however, typically provide only a basic overview (aggregation) over the characteristics of a match. Therefore, one of the most challenging tasks for statistical analysis is the identification of expressive statistical features that help to gain more insight into factors that could influence the outcome of a match. However, in order to get a better understanding of the outcome of a match, its data needs to be analyzed on a more fine-grained level (e.g., movement of single players). One example for the inherent complexity is shown in Table 2.3.1 with a real world example. Possession, shots and duel quota, among other statistics, are usually used to statistically compare teams. They indicate that both teams are equally strong. Considering that Germany beat Brazil 7–1 this is quite surprising. A different sight on the game outcome comes with two new statistics called Packing and IMPECT, both developed by the Impect GmbH [Gmb]. Packing is the number of outplayed players and IMPECT is the number of outplayed defenders. A detailed explanation of both statistics is given by Regenhuber [Reg]. These statistics are better suited to explain the outcome of this match, with Germany having outplayed approximately 60% more defenders than Brazil.

However, even improved statistical measures cannot provide more than an indication about why a team won in a match. Features like packing in our example show that Germany did a better job in outplaying Brazilian players, but we cannot see how these attacks were played in particular. Analysts want to find out why and how something important happened in a match as well as filter out noise obscuring interesting patterns. Explaining cooperative and competitive movement, however, is very complex, as the movement of a single actor is depending on the movement of all other actors. Because of these interdependencies, almost every action causes a reaction. Adding to the complexity is that typically certain role definitions exist in a team that influence movement. Particularly, team captain and goal keeper may steer or influence the movement patterns of team players, e.g., following a tactical decision to defend or press certain players. The role of such leadership effects is prominent not only in soccer analytics, but also a question of leader-follower relationships. This complexity is added to by the fact that we do not only have to look at the interdependencies within a team but also the interdependencies between two opposing teams.

Context-aware analysis on the given spatio-temporal data is accordingly required when analyzing cooperative and competitive behavior in team sport data. Movement is more than just x- and y-coordinates: players are motivated to move by intrinsic and extrinsic contextual factors. We want to know why a decision was made. So in order for the computer to understand why someone moved from A to B, we need to detect and measure what influenced the movement, and teach the computer to see the interdependencies and the influence of the players on each other and their decisions. Summarized, context-aware analysis means that insights into group dynamic behavior are enabled by incorporating collective movement models in the analysis process. Finally, experts want to explain why, when and how specific movement behavior is expressed because of tactical behavior. Tactical behavior, in our understanding, represents the overall effort practiced on the field to eliminate the factor of luck as much as possible. Analysts want to retrieve explanations of observed cooperative movement patterns in reaction to competitive movement patterns by enhanced visualizations. Another open issue is tracking external influences as coach advices and the corresponding team reactions.

The challenges arising with team sport data are not only concerning the analysis of movement itself. Technical challenges as efficient data storage, querying, and processing arise as well. Improved tracking techniques and devices led to an increase in data volume gathered from team sports. Large data sets and in some cases vast amount of individuals tracked require new and efficient handling methods. The need for novel streaming technologies that allow real-time analysis enabling analysts and coaches to gain information even, for example, during the halftime break is increasing. Consequently, creating an encompassing view of team sport data, integrating trajectories, time series, similarity search, high dimensional data, text, and video image processing/analysis will be the major challenge promising large synergy effects and real insights.

2.4 Methodology

Analyzing team sports data requires a toolbox of methods from different domains including sport domain knowledge. We consider the following areas non-exclusively important. *Models* help in deriving patterns and features describing observed behavior in the domain. Models can stem from various perspectives. For example, in a data-driven way models can be obtained from statistical analysis and mathematical formulations; or from concepts developed in sport science. *Data mining* is the corresponding domain in computer science and uses mathematical and statistical approaches. The communication of modeling results is often achieved by methods from the *information visualization* domain. Information visualization is part of computer science and strongly connected to computer graphics and cognitive psychology. The synergetic combination of data mining and information visualization is called *visual analytics* [KKEM10]. The core idea of visual analytics is to implement steerable data mining methods and immediate visual feedback of the analysis results. Highly interactive analysis systems are the outcome supporting the integration of domain knowledge.

2.4.1 Data Modeling

Generally speaking, modeling is about giving structure to the problem of sport analysis, by prescribing which aspects are of importance to the analysis. Ultimately, this needs to be informed by the task of the analysis. Different examples are, e.g., the short-term performance analysis of a single team member, versus the long-term analysis of the team performance or its evolution over time. The modeling process will eventually identify a set of variables and/or events to observe, and a quantitative scheme to aggregate and/or compare the measures. We distinguish two main approaches to guide the modeling: Domain-specific modeling is based on theories and concepts from *Sport Science* [LMH+15, SLE14], which generalize relationships between actions and outcomes in the respective sport domain. On the other hand, *data-driven* or *explorative* modeling does typically not assume previous knowledge about the domain, but is guided by dependencies found in the data directly. In practice, both approaches often go hand-in-hand: There are expectation and theoretical models about typical dependencies in the sports events, and the analysis measures are derived to validate and quantify the expected dependencies. Besides, patterns that one may observe inconsistently in larger amounts of measurement data may lead to new domain-specific theories and hence guide the modeling approach conceptually.

Team sport is fundamentally about space-time interaction of players—understood as a combination of operation, communication, and strategy—with the most influential interaction being the interference of the opposing team which excludes in many respects absolute scales of measurement. The ongoing lack of reliable key performance indicators for soccer teams points in this direction [MC13, CWNB14]. A deeper understanding of the interaction structure and dynamics between and within teams during a match is meaningful for making progress in team sports analysis. Relative phases [BSM10a, BSM10b], couplings [FdPVL12], invasion profiles [Lin14], or centralities in passing networks [Gru12] represent exemplary concepts.

As the data collected for a match shows only one realization of a contingent situation ("what happened"), the challenge is to determine the (path-dependent) set of *possible* actions executable in that situation ("what not happened"). Operationally, the presence of the opposing team first and foremost restricts the space-time regions on the pitch accessible for meaningful action of a focal actor. On the communication level, selecting an action from the set of accessible actions ("decision") depends furthermore on the intra-team movement patterns, space cognition, and information processing. Here, coming up with realistic and feasible action models is very demanding [TH00, FS05, GW14], but worthwhile for answering to questions like "Which pass is best to execute given a set of possible passes?" Solely ball-oriented data is very limiting in respect thereof, as no information on players without the ball is available; the situation becomes much better with additional positional data of all players [GW14]. Moreover, technology is available (at least in basketball) to track head positions and head orientations of the players providing data for a feasible inference of mental maps [XWW12], e.g., which players are seen by a player. On the strategic level of interaction, expectations, style of play, tactical orientations, etc. determine a *normal* behavior of the teams. Data-based analyses of these aspects are thus required for an unbiased evaluation of action sets. However, corresponding techniques are currently less developed. Methods of game theory could be applied, e.g., building strategic games from multi-parameter interaction models based on tactical action-reaction schemes available from expert's domain knowledge.

In principle, there are two approaches to analyze the interaction of two teams in competitive dyads (matches). First, characterize matches by the behavioral characteristics shown by the teams, i.e., observed events are directly assigned to individuals, groups, or teams with positive or negative evaluation. This is the standard bottom-up approach. Second, characterize the behavior of the teams by the characteristics of matches, i.e., observed events add up to global information on the match before they are projected onto the teams. This can be classified as top-down. Next to external information (match type/competition, venue, audience, weather, etc.), typical global characteristics are parameters related to match speed. Speed can be hardly assigned to one side only. Though there is no standard definition of match speed, the growing popularity of packing rates (which relate the space occupied by a number of opposing players to time represented by pass duration) can be seen as a need in this global information.

2.4.2 Data Mining

Data mining is the automatic or semi-automatic discovery of patterns in data sets, which are too large to analyze manually. It is the analysis stage of the knowledge discovery in databases process, which also encompasses *Selection, Pre-Processing, Transformation and Interpretation* [FPS96]. Data mining itself is an umbrella term, which encompasses various methods from different computer science fields. It uses, inter alia, methods from machine learning, artificial intelligence, statistics and database systems [CEF+06]. Fayyad et al. [FPS96] group data mining techniques into six general categories. There exist already many potential uses of data mining methods in the field of team sport analysis. However, it is necessary to either adapt existing methods from various research fields or to develop new algorithms to allow an effective analysis of team sport data. In the remainder of this section, we use these categories in order to review existing data mining techniques and their possible application to sports data.

Clustering

Clustering is the grouping of similar objects into clusters. The concept of a cluster, however, is not clearly defined, which is why so many different clustering algorithms exist [Est02]. Generally speaking, a cluster is a set of objects which are more similar to each other, than to those in other clusters, however similar is defined. One could use one of the clustering algorithm by Lee et al. [LHW07] to identify common movement patterns of individuals or groups. In American Football this could be used to identify the most often used passing patterns of a team. An example where clustering is used is the work of Janetzko et al. [JSS+14]. Here, clustering is used to find common behavioral patterns of individual players. For this, the *k-means* [M+67] algorithm is used on feature vectors consisting of various features like player speed or distance to ball.

Classification

"Classification is learning a function that maps (classifies) a data item into one of several predefined classes" [FPS96]. These classes can either be defined manually or alternatively generated automatically using clustering, which is described in Section 2.4.2. This data mining method can, for example, be used—to stay with the American Football example—to automatically identify in which passing pattern a new pass can be classified.

Regression

Regression or regression analysis is the process of estimating the relationship between dependent and independent variables of an experiment [Lin90]. Its goal is the determination of parameters of a function (such as a_1, a_2 and ε in $f(x) = a_1 \times x + a_2 \times x + \varepsilon$). If a relationship exist and if a good fitting can be achieved, the resulting function can be used to predict future observations. Regression analysis can be used to expose statistical correlations between and model the behavior of players [DWW05] and teams in tournaments [Car96].

Summarization

"Summarization is a key data mining concept which involves techniques for finding a compact description of a data set" [CK07]. Methods like calculating the mean or standard deviation or dimensionality reduction are often used to analyze and visualize large and complex data sets. Clustering can also be used as a summarization method, with the centroid used as a representative for the whole cluster. This method has beneficial use to display the complex and large data sets, which are common for team sport analysis. For instance, techniques like the well-known PCA [Pea01] or t-SNE [MH08] are used to visualize high-dimensional data in lower-dimensional space. Perin et al. [PVF13] show this by proposing a visual abstraction and summarization system of, for example, certain attack paths.

Change and Deviation Detection

Change and deviation detection, which is often also called outlier detection, refers to the detection of observations, which do not correspond to the already existing patterns. On method to find such observations is Grubbs' test for outliers [Gru50]. One possible use case in the research field of team sport analysis could be the detection of players which perform extremely better or worse than all other players.

Dependency Modeling

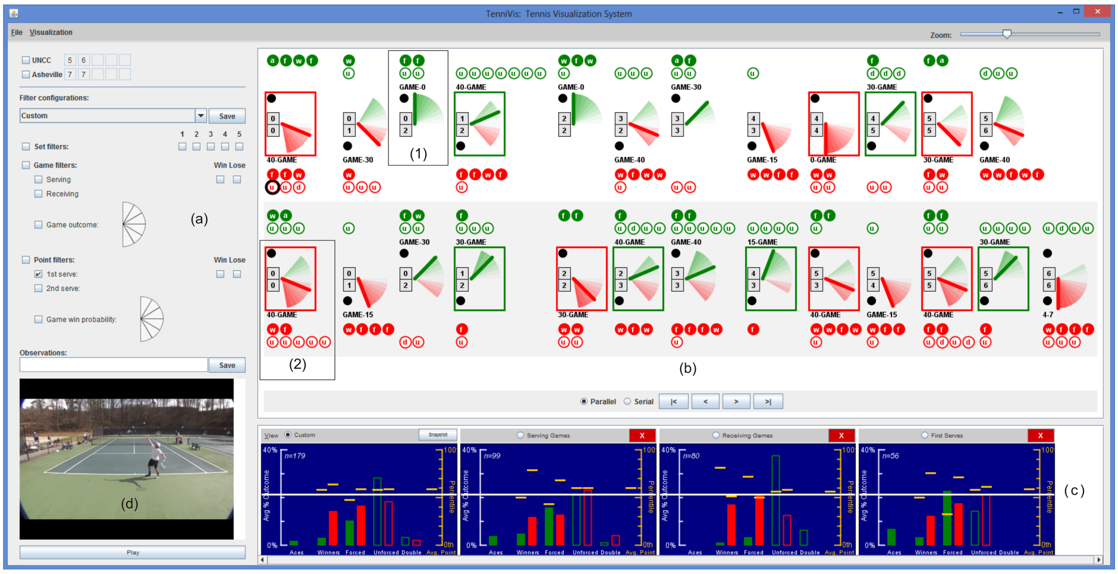
Dependency modeling or also called association rule learning is defined as the identification of significant relations between variables in the data. To identify interesting rules, the measurements of confidence and support are often used [KMR+94]. These rules are often used for market basket analysis, to identify which items are often bought together. They could also be used in the field of team sport analysis, to identify which events frequently occur if, for example, a goal was scored.

2.4.3 Information Visualization

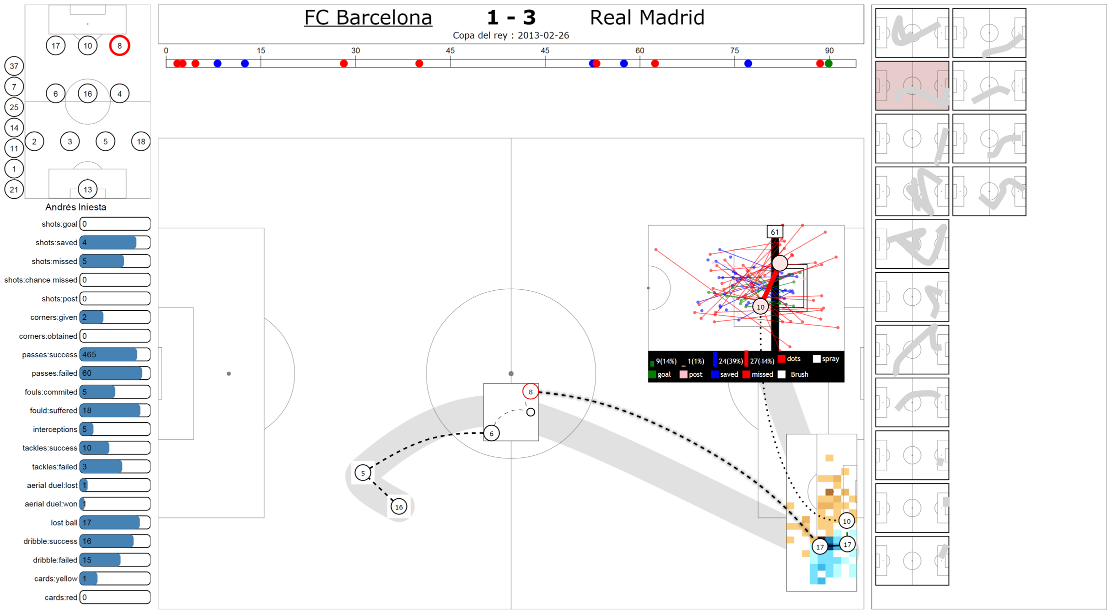
Information visualization is a growing research field with three main tasks of visualizations being directly transferable to the sports domain: exploration, hypotheses validation, and hypotheses generation. Exploration, here introduced as an example, is usually the first step when dealing with a previously unknown data set. Overview visualizations help to identify the descriptive features or to detect interesting patterns. Important visualization techniques for sports data are statistical visualizations as scatter plots or parallel coordinate plots [Ins85] as well as more specialized spatial and temporal visualizations (such as displayed in Figure 2.4.1). Spatial visualizations help to investigate distributions and individual or group movement patterns. Temporal visualizations [AMST11] show how features change over time and effect each other. Horizon graphs [R+08] are a good example for a space efficient visualization of time-dependent data. Pixel-based visualizations [Kei00] are even more space efficient, encoding data values by color-coding single pixels of the display. The challenge in pixel-based visualization is an effective layout of the data items on the screen. For instance, temporal data can be laid out hierarchically with the help of a technique called Recursive Patterns [KAK95]. However, the high complexity and multi-dimensional aspects in sports data require novel visualization techniques. Abstracting the collected data in a suitable way and pointing experts to interesting aspects is key for a successful analysis.

2.4.4 Visual Analytics

Combining data mining with information visualization (visual analytics) is creating synergistic effects of human and machine. Visual analytics enables experts to include their domain knowledge during the analysis process by interactive and steerable data mining methods and immediate visual feedback of the results. We believe that visual analytics is a very effective way to



(a)



(b)

Figure 2.4.1: Two recent systems that aim to improve the understanding of sport data by several visualization techniques. (**a**) TenniVis [PYHZ14]; (**b**) SoccerStories [PVF13].

cope with the challenging data properties in the team sport domain. The research challenges described previously showed that although the desired outcome is clear, we need to deal with ill-defined data analysis problems. For example, detecting movement patterns being of interest to the analyst requires a semi-formal description of interests of the analyst. However, this transfer from the sports domain to the data domain being understandable by machines is very difficult. Visual analytics proposes a transparent analysis process, where the manifold parameter choices of data mining algorithms are as comprehensible as possible. By interactive exploration of the data and parameter space, domain experts can get a feeling for their data. However, there is also translation and communication needed between sport experts and visual analytics experts. One possibility to enable a productive communication between experts of both domains could be the use of a *Liaison* [SMKS15].

2.5 Discussion and Conclusion

The analysis of team sport data is a very useful, yet inherently difficult problem. In this chapter, we gave a high-level overview to this fields as well as identified research aspects to contribute on. Team sport analytics involves many challenges and problems. First, the choice of data acquisition is difficult, as the type and quality of data available determines the potential analysis that can be done. Current acquisition includes video analysis, using position sensors or manual encoding. In the wider sense, data acquisition also needs to consider data modeling, integrating, and cleaning, each representing significant work steps. Being aware of data quality and comparisons of data sources is of critical importance. With regards to this work, quality observations are especially needed when it comes to the trajectories of players and balls from video image sources (Chapter 3). For meaningful analysis, domain requirements coming from experts are important and must be taken into account. Specifically, analysts nowadays are, for example, interested in the analysis, prediction and performance monitoring of their team. Each problem, however, has many sub problems. The problem of data acquisition, for example, translates into video processing, classification and annotation problems. Nevertheless, team sport analysis is a highly interesting research field where experts from many different computer science subjects (computer vision, graph theory, network analysis, visual analytics, ...) can bring in their expert knowledge.

An important problem of doing analysis relates to modeling—what are the aspects in the sports events/matches which are of importance? How can we qualitatively and quantitatively assess individual player performance or the development of tactic and strategic capabilities of a team? Single events and situations (e.g., success rate of corner kicks) can be easily statistically computed. However, it is a difficult problem to assess strategic factors to *explain* why a specific success rate is observed, or how it could be influenced. Furthermore, many analysis goals cannot be statically defined once and for all, but depend highly on the context of the analysis. Answers to these questions need to consider both data-driven approaches from Data Science, but also, models and concepts developed in Sport Science. In this dissertation, we show that both can go hand in hand by characterizing the role and influence that both fundamental approaches have and how to combine them to design novel, appropriate and adaptive analysis systems.

Whole systems which allow for automated tracking or labeling remain an open challenge.

Graham Thomas, (BBC R&D)

3

Extraction of Team Sport Data

Contents

3.1 Introduction

In common forms of tactical analysis, coaches and players rely on insights from previous matches when determining their lineups and strategies. In order to perform such an analysis, a significant amount of data needs to be acquired. Information about player movement and behavior as well as other aspects of the game are crucial for the evaluation and development of new tactics. Many sports clubs use dedicated video analysts which are tasked with watching and carefully evaluating recordings of previous matches and supplementary data. This typically involves manual tagging of important events and situations within the video using specialized video annotation software such as Dartfish or Sportscode. In practice, much of the data acquisition methods fall in one of two major categories.

Several professional providers of match data use dedicated tracking hardware such as GPS receivers and accelerometers directly worn by the players in order to track their positions and movement. A popular example is the CATAPULT [cata] system which consists of GPS-based tracking devices with a wireless transmitter. However, these systems are typically very expensive. Furthermore, a team can only use them with their own players and is, therefore, unable to gather any data from the opposing players and the ball. Complex camera setups consisting of multiple calibrated cameras placed at strategic locations inside a stadium are often used as well. These setups allow precise and continuous tracking of relevant entities and can be used to obtain three-dimensional position data. However, they are associated with high installation and operating costs and are typically only available in highly professional stadiums.

When video data is available, many employed analysts also perform manual annotation of video recordings according to their specific requirements. During this process, the analyst determines different aspects of the game such as specific events, player constellations and situations that are of interest and then uses specialized software to manually tag the respective occurrences in the video. There are typically two different use cases where manual annotation is performed. One possibility is live annotation during a running match, the other one is offline annotation performed on a video recording of a previous match. These two use cases extent to match analysis in general. Live analysis is typically used to evaluate and adjust the teams performance during a running match, for example, by providing a detailed performance report during the halftime break. Offline analysis is often used to analyze and aggregate data from previous matches in order to optimize the tactical preparation for an upcoming match. The requirements for the two use cases differ significantly. For live analysis, real-time acquisition and processing of the data is required and any latencies have to be minimized as far as possible. For offline analysis, the runtime is much less relevant and the focus lies more on achieving the best possible accuracy with a variety of analysis methods. Both of these methods, dedicated tracking setups and extensive manual annotation, require either expensive hardware and/or a significant amount of human effort and time.

Due to the inherent limitations of these methods, the area of automated video-based acquisition of match data from a single camera presents interesting opportunities. The low cost and easy installation as well as the possibility to analyze existing match recordings from TV broadcasts and other video sources makes it possible to perform match analysis for a much wider range of matches opposed to the existing, much more exclusive tracking setups. Especially interesting is the use of already existing television cameras or similar uncalibrated single-camera setups since the use of professional multi-camera arrangements is associated with significant practical hurdles such as high hardware and operating costs as well as the need for initial calibration. While single camera setups are affordable and often already available in the form of television cameras, they present additional challenges compared to calibrated setups since the position and movement of the camera itself has to be taken into account as well. Furthermore, additional difficulties arise from the fact that a single camera setup is unable to provide depth information and has to deal with frequent occlusion of target objects and numerous other challenges.

This chapter closes the aforementioned gaps by providing appropriate real-time computer vision techniques to extract player and ball positions as well as to detect the relative view-port from standard TV broadcasts, which enables capturing movement data. We contribute an improved approach for mapping between video and pitch coordinates through a temporally consistent camera tracking method. Evaluation on real-world data shows that the proposed methods are usable for many forms of video recordings and can, therefore, provide a base for more insightful video-based analysis. Quantitative evaluation is performed using existing data sets with known ground truth information.

3.2 Related Work of Object Detection and Tracking

Higher-level content analysis, e.g., for player movement, involves detection and tracking of players in the input video streams as well as mapping of the positions into a reference frame. Tracking has been studied for decades as one of the fundamental tasks in computer vision, see e.g. the work by Wu et al. [WLY13] for a more recent overview. We, therefore, restrict the following discussion to the methods used in previous work on team sports analysis.

In ***position mapping***, positions from the video are mapped to a physical reference frame, in our case the soccer pitch. A key approach to this mapping is to define and identify reference points in the video that correspond to known locations on the underlying soccer pitch. For this purpose, Seo et al. [SCKH97] define a pitch model, incorporating known positions like the center circle and center line. To find the corresponding points within a video, lines and ellipses are identified in a given image by Hough-Transformation [DH72]. Of these lines and ellipses, four positions are extracted and associated with the positions in the pitch model to estimate a coordinate transformation of the given image into the pitch model. Since it is a plane-to-plane transformation, it can be described by a projective linear map (homography) [HZ06]. A drawback of this method is, that there always have to be at least four positions associated with the soccer pitch and, most importantly, these positions have to be correct. In contrast to comparing images individually, rebuilding a larger image (e.g., panoramic views) from smaller subimages is a common technique in image vision called *stitching*. For example, Kim et al. [KH01] look for identical lines in subsequent images of the video and map related lines onto each other by a coordinate transformation. This way, a complete view of the soccer pitch is obtained image by image. Brown et al. [BL07] introduce an automated process for panorama stitching. The described approach involves the calculation of SIFT-features [Low04] within the considered images. A transformation is computed by finding matching features in image pairs. We roughly follow their approach to generate a panoramic view of the soccer pitch seen in the video recording.

The next step after position mapping is to ***detect and track objects of interest***. In case of team sport video analysis, these are typically the teams' players and the ball. In player detection, advances have been made in the last decades. Liu et al. [LTL+09] recognize players with a so-called *boosted cascade of Haar features*. This detector is constructed with carefully chosen training data (players in different poses as positive examples, background and pitch parts as negative examples). For the detection of players on a soccer pitch, Hoernig et al. [HHR13] compute a pitch mask. For that matter, they assume that the soccer pitch is a rectangle. Furthermore, the dominant color of the pitch is defined as green. The potential positions of the players are obtained by subtracting the pitch mask from the processed image. Additionally, the found contours have to meet different predefined features (for example height relations).

Perez et al. [PHVG02] follow a probabilistic approach. First, a color histogram model is created, which describes the color distribution of a given player. The detection and tracking of the player is subsequently realized with a particle filter [GRA04]. Dearden et al. [DDG06] also introduce a combined approach. In the first step, a mask of the soccer pitch is computed. For this mask, the color histogram is defined and by subtraction the possible player areas obtained. The player areas are subsequently processed with a particle filter. Seo et al. [SCKH97] use a combination of particle filter and matching of player templates. These templates are obtained by analyzing the defined player mask. The tracking of players with several stationary cameras is examined by Kang et al. [KCM03]. The modeling of players is achieved by a combination of color distribution and movement models. The tracking itself is performed by maximizing a multivariate distribution probability.

3.3 Data Extraction

Our proposed system should be capable of extracting data from standard video recordings. It should be easy to use by non-expert users, with only low-level technical requirements. The input to our system is the video stream of one camera, which zooms and rotates to capture as many players as possible. This form of recording resembles common public television recordings of team sport matches, with one main camera positioned on the side of the pitch for a tactical overview. We align the frames of the video recordings with a normalized reference pitch to associate the positions in the video and on the pitch. In particular, our system is not limited to only process data from video to normalized pitch. Instead, we can also project visualizations rendered on the normalized pitch back to the original video, which allows for further analyses by directly watching the enhanced video stream (see Chapter 4.3 for more information on the visualization projection in the original video recording). We implement our visual analytics system in *Java*, with a subset of time-critical computer vision methods in *nVidia CUDA* [Nvi11] running on GPUs. The parallel nature of the feature extraction and matching process is well suited for modern GPU architectures and allows an implementation which is several times faster than equally powerful CPU based methods. An overview the proposed system is given in the subsequent sections. See Figure 3.3.1 for an overview of the workflow.

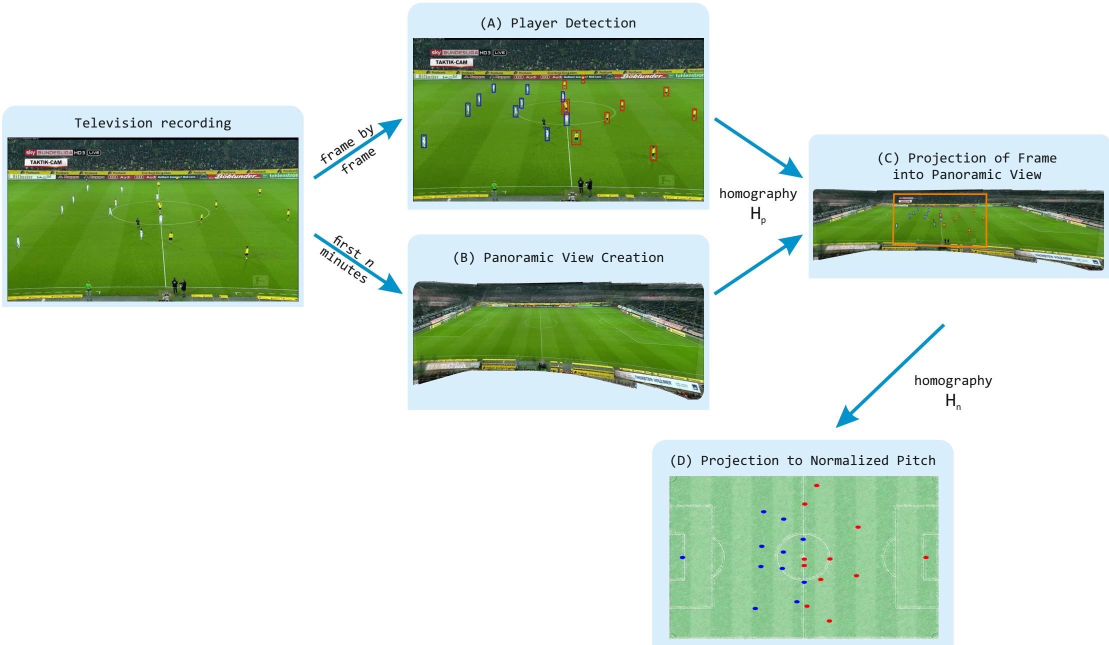
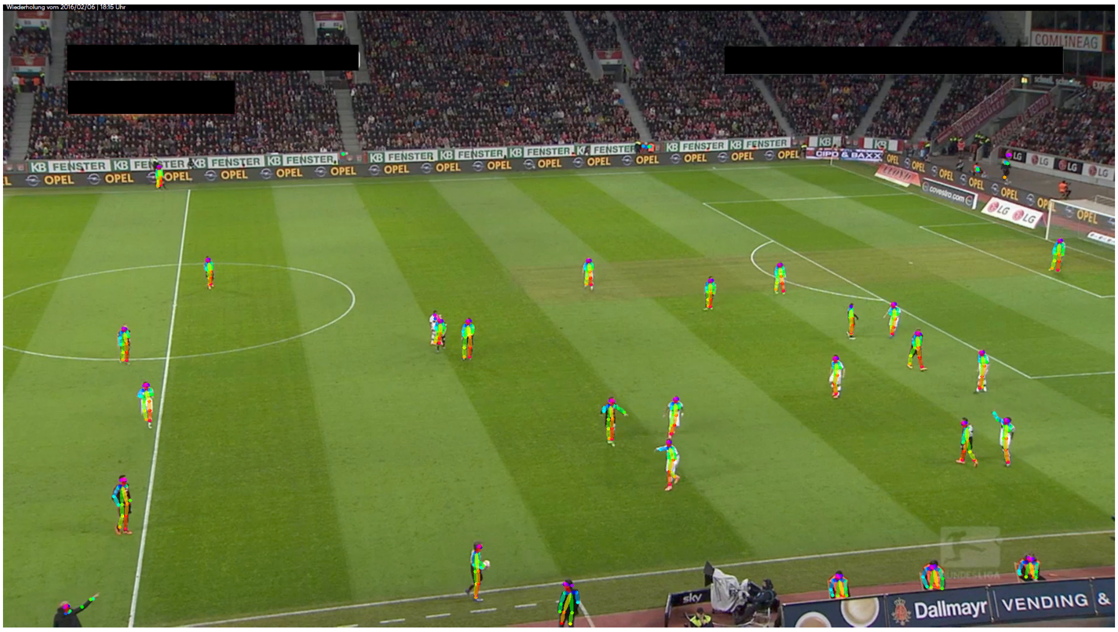


Figure 3.3.1: Workflow of our system. A simple (TV/video) recording serves as input. We detect players on the soccer pitch (A) and, simultaneously, generate a panoramic view of the pitch (B). Afterwards, each frame gets projected into the panoramic view to extract correct player positions from a moving camera (C). The extracted player positions then can get projected into a normalized pitch (D).

3.3.1 Player Detection

We propose a method to automatically extract player body poses from a single video recording of a match. Our proposed method uses a skeleton model based on a hierarchical graph structure to represent a body pose. In this graph, every node corresponds to the position of a characteristic keypoint of the skeleton such as joints, ears, eyes, etc. while each edge represents an anatomical connection between two keypoints. In total 18 different nodes are used for each skeleton which are connected to form a tree where the neck keypoints serves as the root node (see Figure 3.3.2). All keypoint locations are estimated in 2D pixel coordinates within the reference frame of the current input image. In order to obtain pose data using the previously described skeleton graph model it is required to estimate keypoint locations as well as graph edges from a given input image. The proposed approach follows a similar structure as Cao et al. [CSWS17] (also known as *OpenPose*), however, we use a custom self-developed deep learning model which has been built and trained entirely from scratch to fit our use case. Similar

(a)



(b)

Figure 3.3.2: Pose representation using skeleton graphs

Method	Output Form	Single / Multi-person	Runtime Per Frame
PIM [RMH+14]	2D skeleton	single	unknown
CPM [WRKS16]	2D skeleton	single	≈600-700ms
DeepCut [PIT+16]	2D skeleton	multi (top-down)	>500s
PAF-CPM [CSWS17]	2D skeleton	multi (bottom-up)	≈700-800ms
DensePose [GNK18]	2D surface	multi (top-down)	≈400-500ms
HMR [KBJM18]	3D mesh	single	≈40-50ms
SMPLify [BKL+16]	3D mesh	multi (top-down)	≈30-60s

Table 3.3.1: Categorization of existing works, runtime values are estimates for a single 1920x1080 frame

to Wei et al. [WRKS16], we formulate the keypoint detection and localization task as a dense regression problem. The proposed method performs this type of keypoint detection by estimating a series of dense confidence maps which encode the probability of observing a certain keypoint type at any given location in the image as well as a series of vector fields which encode edge orientation. The estimation of the confidence maps is performed using a multi-stage convolutional neural network which operates directly on the input video. The proposed model is designed to handle large scale differences in the input images and can therefore be applied to a range of possible focal lengths from zoomed-in detail views to wide-angle panoramic videos without any changes in the configuration. In contrast to Cao et al. [CSWS17] and other existing works, our proposed method is able to perform pose estimation at much smaller scales which is especially interesting for our use case since most television broadcasts of football matches contain wide-angle views where the individual players are extremely small in relation to the image size. Furthermore our model architecture as well as the inference system has been optimized for the use case of real-time low-latency video analysis and therefore exhibits dramatically increased performance in terms of latency and frame throughput when compared to similar works. Table 3.3.1 provides a characterization of existing approaches, illustrating the long runtime per frame of other methods. Figure 3.3.3 shows a runtime comparison between our proposed method and the approach by Cao et al. [CSWS17]. All runtime measurements were performed on a NVIDIA GeForce GTX 1080 as well as a NVIDIA GeForce RTX 2080ti GPU. We were unable to evaluate the approach of Cao et al. at resolutions above 1920 x 1080 due to insufficient video memory. We observe that our method requires significantly less runtime than the method proposed by Cao et al. This is especially significant at higher resolutions. At a

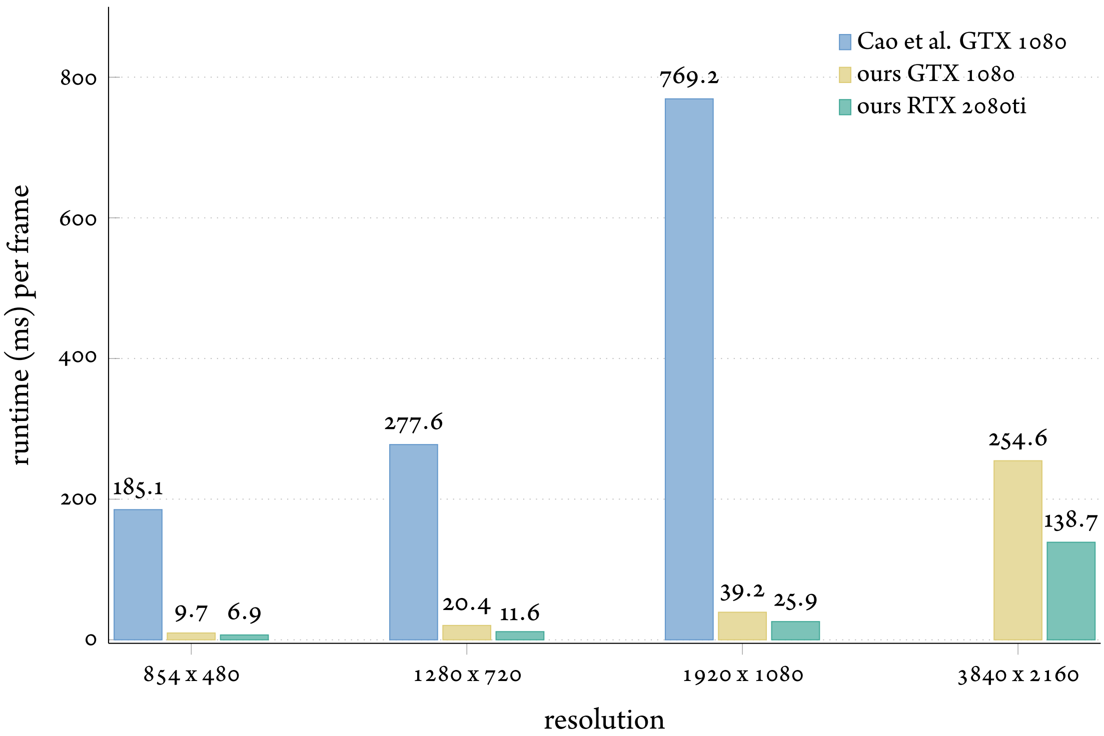


Figure 3.3.3: Total runtime for various input resolutions. The results show that our proposed approach is usable for real-time tracking even at high input resolutions. The approach of Cao et al. [CSWS17] could not be started at resolutions above 1920 x 1080 due to insufficient video memory.

resolution of 1920 x 1080 pixels, our method in development requires around 26ms on average per frame which translates to around 38 frames per second.

Evaluating the accuracy of the resulting pose data in the context of football analysis is challenging in general since no publicly available ground truth data set exists for the task of pose estimation. We performed evaluation on the COCO [LMB+14] validation set using the mean average precision and achieved results that match or exceed similar state-of-the-art methods, especially in the case of smaller relative person sizes. Comparing the mAP values in Table 3.3.2 shows that our methods performance is already reasonably close or better in comparison to the

Method	mAP	AP^{50}	AP^{75}	$\mathbf{AP^M}$	AP^L
Cao et al. [CSWS17]	61.8	84.9	67.5	**57.1**	68.2
ours	56.8	77.5	61.9	**57.3**	56.4

Table 3.3.2: Comparison using the COCO [LMB+14] evaluation metrics

one proposed by Cao et al. [CSWS17]. A small decrease is to be expected since our method is focused on the smaller person sizes present in our use case. Since we train our network for smaller input sizes than in the original COCO data set, the AP^M metric, which focuses on smaller person sizes, is of higher interest.

We also assessed the quality of our extracted data and compared our generated data with data acquired by a professional data provider with a calibrated camera system. To this end, we analyzed a match from the UEFA Champions League between FC Bayern München and Manchester City FC where we have both professionally generated position data as well as data extracted with our method from a high definition (720p) recording. To be able to compare the resulting positioning data, we synchronized the video recording (25 frames per second) and the professionally generated position data (10 positions per second). The professional system uses a set of static cameras around the pitch and proprietary software to extract the player movement. As the calibration would need manpower, these camera systems are usually permanently installed in stadiums of prime teams. We observed a maximal distance between our retrieved position and the professional data of four meters. Nevertheless, the average deviation was less than two meters with a standard deviation of roughly half a meter. We consider this as a positive result as the accuracy of professional data is not publicly known. Furthermore, the definition of a point-based location for a two-legged running person is not clear per se already causing positional differences. From a subjective perspective, our extracted locations are very close to the players in the video recording and reflect the movement very well and there are no obviously and distracting systematic errors.

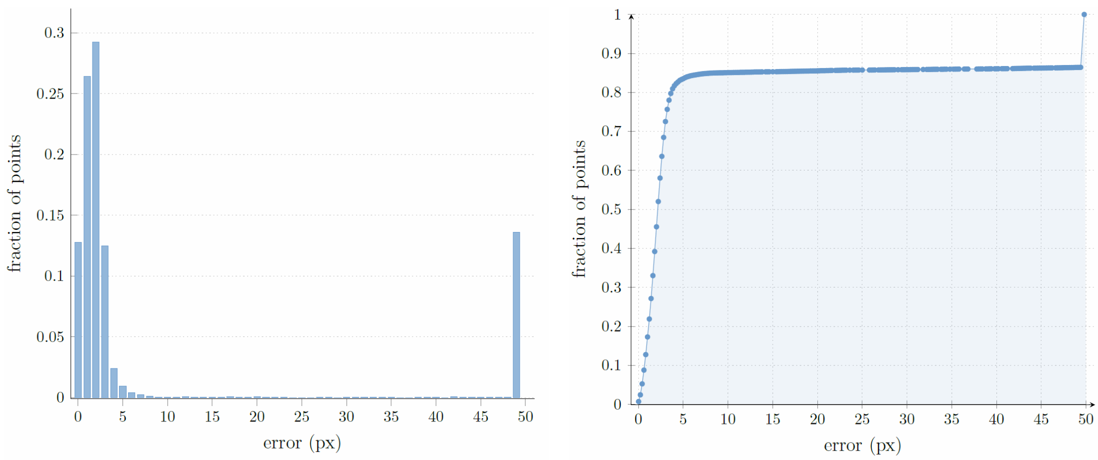
3.3.2 Detecting and Predicting Ball Movement

Methods that allow accurate and reliable extraction of ball position and movement are essential for an automated tracking system. Several previous works have presented methods for ball detection in various scenarios [CS05, MWF16]. However, robust real-time tracking of the ball using a single camera remains an unsolved problem in many cases. Ball tracking methods that are applicable to broadcast video recordings have to deal with many additional challenges such as large variations in lighting and shape, abrupt velocity changes as well as potentially long range occlusion (see the examples in Figure 3.3.4). It is, therefore, especially difficult to perform reliable tracking over longer time periods such as an entire match. Due to these challenges, it is typically not sufficient to only consider the features of a single frame for ball tracking.

(a) (b) (c)

Figure 3.3.4: Difficult ball visibility due to background color similarity, motion blur and size

Therefore, we introduce a method to perform accurate real-time ball detection from an input video recording. The proposed method uses a two-stage approach consisting of a per-frame candidate detection step as well as a temporal integration phase. In the first stage, a number of possible ball candidates are detected in each frame of the input video. The number of detected ball candidates is dynamic and depends on the image contents. The second stage then focuses on the generation of accurate ball trajectories over time from the given candidate information. The two stages are realized by a combination of a convolutional neural network for the candidate detection and a recurrent neural network for temporal filtering and integration with an additional post-processing step. Both networks are trained independently with a custom training data set. We evaluated the accuracy of the proposed method by measuring the euclidean



(a) Histogram **(b)** Cumulative distribution

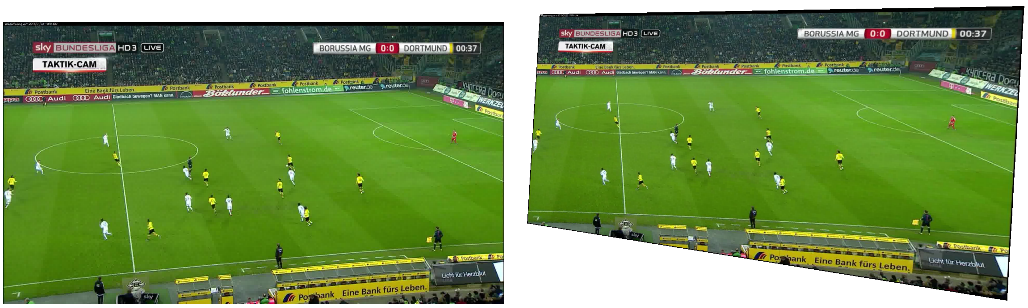
Figure 3.3.5: Distribution of the ball position error over a test set of 5 videos

distance between the predicted and ground truth ball positions within the input frame. Figure 3.3.5 (a) and Figure 3.3.5 (b) show the distribution of the ball position error over a test set of 5 videos. Figure 3.3.5 (b) shows that, with our current state, around 83 % of all ball positions are estimated with an error of less than 5 pixels, in around 13 % of all cases we observe unreliable ball positions due to occlusion or detection failure.

3.3.3 Static Camera Generation and Projection

Previous systems typically use several static cameras, mounted at specific positions with certain angles and focal lengths, being calibrated for the sole purpose of making the player detection for the computer as efficient and effective as possible [MWR09]. Instead, we propose to use one single, stationary camera, whose motion only consist of rotations and occasional zooms. It is well known [HZ06] that under the additional assumption of pinhole projection, the video frames are then related to each other by a simple projective linear transformation called a homography. The advantage is that for the domain expert and final user, one single camera is set up easily without the need of longer preparations, special configurations, and calibrations.

However, in order to have comparable results as from several cameras and relate the different frames to each other, we automatically transform the images into a single, panoramic view of the pitch (Figure 3.3.1 (B)). Our proposed approach follows Brown and Lowe [BL07] who take a set of overlapping images as input. The set of input images is aligned automatically by iterating over pairs of consecutive frames, extracting and matching SIFT features [Low04] and estimating the homography between first and subsequent frame. A homography is a 3×3 transformation matrix acting on projective image coordinates, which can be estimated by a least-squares



(a) (b)

Figure 3.3.6: Before homography (a) and after homography (b) calculation

Figure 3.3.7: The chosen points for the transformation of the panoramic (left) into the normalized (right) view must allow referencing distances in both directions of the main axes. Afterwards, a calculated homography H_n enables the transformation of panoramic view into normalized view and vice versa.

approach embedded in RANSAC to get rid of outliers [HZ06]. From the aligned frames, we select a subset of frames which spans the whole range of motion of the video camera and has sufficient overlap between frames to generate the full panoramic view. Usually, the first 2 minutes of a match are sufficient to generate a panorama reflecting (almost) the complete pitch. From the homographies computed initially, we compute the alignment of every individual frame within the panoramic view, see Figure 3.3.6 for an example transformation. For longer videos, we generate the panorama only from the beginning of the sequence, and align subsequent frames by estimating a homography directly using their and the panorama's SIFT features.

In a final pass, we generate a clean background panorama and remove foreground objects like the players. For this, we compute for every pixel in the panoramic view the median color over all frames of the selected subset. As the players are constantly moving, this should be a good approximation to the actual background color. A Graph Cut method is then used to hide objects which are only partially visible [KSE+03]. In summary, we now have a panoramic view of the background as a common reference frame, as well as for each video frame, a homography H_p which describes the coordinate transformation into the panoramic view. See Figure 3.3.1 (C) for an illustration. To obtain a normalized representation of the pitch, we map the panoramic view onto a user-supplied image, typically a standardized visualization of the pitch, see Figure 3.3.7. To define the transformation, the user selects at least four corresponding points in panoramic and normalized image manually. The four correspondences then determine a homography from the panoramic view to the reference image as before [HZ06]. Note that due to the lack of radial distortion calibration, this is only a rough approximation, which, however, is sufficient for the given tasks. Furthermore, it is a plane-to-plane homography, valid only for points on the main pitch plane. The transformation from panorama to normalized pitch is denoted H_n in the remainder of the chapter. By sequential application of the homography H_p

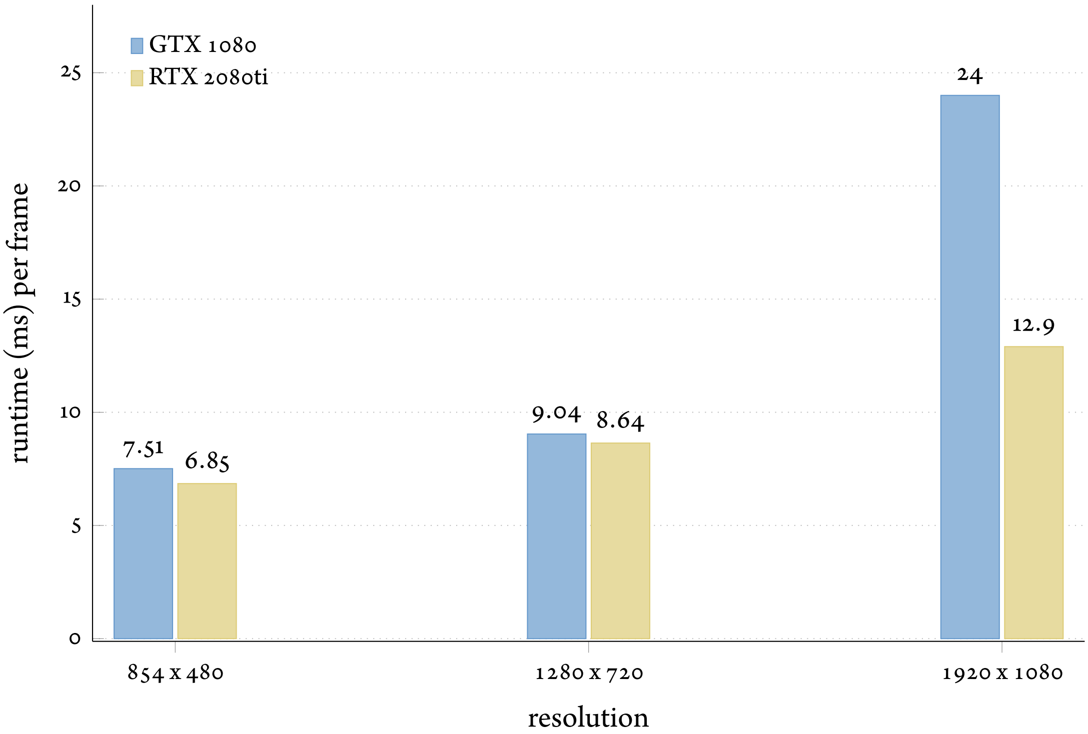


Figure 3.3.8: Total runtime of the proposed camera tracking method for various input resolutions

from frame to panorama, as well as the homography H_n from panorama to normalized frame, we can thus calculate player position coordinates on the normalized pitch, see Figure 3.3.1 (D). All positions not corresponding to a position within the normalized pitch are discarded as likely misdetections.

To investigate the resulting efficiency of our implemented methods, we calculated panoramic views and homographies for several matches on a test system using a NVIDIA GeForce GTX 1080 as well as a NVIDIA GeForce RTX 2080ti GPU. The efficiency of our automatic panoramic view generation depends on the video resolution, complexity of camera movement and the amount of moving objects in the scene. We found that we are able to generate a panoramic view in an average of 30 seconds. The efficiency of a homography calculation depends on the panorama size, input resolution and the amount of visual detail. We are able to calculate a homography for a video resolution of 1920 x 1080 pixels in around 13 milliseconds per frame. As shown in Figure 3.3.8 the runtime required for the proposed camera tracking method is well below the frametime of a 30 fps video for input resolutions up to 1920 x 1080.

3.3.4 Data Cleaning

To obtain player tracks and uniquely identify players, the detected player positions are registered by a simple location-based method, which assigns each position to the closest player in the previous frame who could realistically reach the new position within the timespan between frames. A new player is initialized for all remaining positions that could not be assigned to any existing player. This happens, for example, if the input frame did not contain all players. However, incorrectly detected player positions can lead to an initialization of a new player as well. These misdetections typically only happen over a span of few frames, so we automatically remove short tracks with only very few positions or which do not generate new positions over a longer period. Afterwards, automatic methods to further clean the data are applied, e.g., if players overlap each other we interpolate their positions by taking their current speed into account. Another problem is that the homography H_p, which maps a video frame into the panoramic view is not inevitably the optimal solution, as we compute H_p with RANSAC for a fixed number of iterations. This may result in situations where the transformed player positions are differing from the actual ones. To counteract this problem, we validate our data by applying a *moving average filter* [PM92]. In addition to the automatic systems, the user will also be given the opportunity to manually improve data gathering by interacting with the system in a variety of ways.

3.3.5 Livestream Support

One important aspect of the proposed system is the ability to work on live video data in real time. One use case in that category is the use of network cameras, which are often found in existing stadium installations. To enable real-time remote analysis we use a small single-board computer placed inside the stadium network which serves as a stream grabber. The stream grabber is able to receive multiple video streams from cameras within the stadium network. It then re-muxes the incoming streams onto a common container format. The resulting data is forwarded to an externally visible server within the university network which can then relay the video stream to a machine which runs the analysis software. Figure 3.3.9 shows the network setup used for capturing network video streams. We illustrate how this setup is used in cooperation with the Austrian first league team TSV Hartberg in Chapter 7.2.

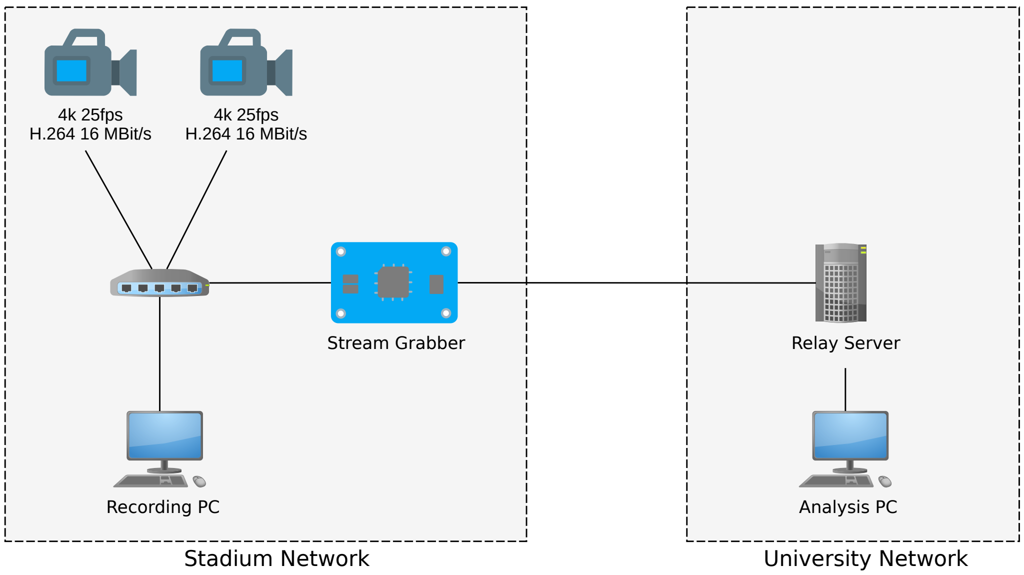


Figure 3.3.9: Network Setup for livestream analysis. The use of a small single-board computer as a stream grabber allows an easy and inexpensive integration of existing network cameras into the proposed analysis system.

3.4 Conclusion

This chapter has contributed a collection of effective computer vision methods for data acquisition from video sources in the context of real-time analysis of team sport matches. The proposed methods allows an analysis system to capture richer and more meaningful information about player behavior and can be used to evaluate additional aspects of a team sport match that could not be considered before. Evaluation on real-world data has shown that the proposed methods are usable for many forms of video recordings and can, therefore, provide a base for more insightful video-based analysis. Quantitative evaluation has been performed using existing data sets with known ground truth information. Consequently, our initially stated goal of constructing an automated reliable system for movement data extraction in team sport has been fulfilled. The resulting system lays a foundation for a large variety of possible analysis approaches that can be built on top of it.

The most exciting thing for me is to see how the transition game has developed and accelerated. There's so much happening in the eight to 10 seconds after the ball is won or lost. Those moments decide games, and a large part of our training is dedicated to what we call swarming behaviour; the synchronised movement of players.

Ralf Rangnick (Coach RB Leipzig, 2018–Present)

4 Understanding the Context of Collective Movement

Contents

4.1 Introduction

In this chapter, we focus on enriching the obtained players' movement data with necessary contextual information, being crucial for a successful and effective analysis. Group movement is not random as players within a team need to work together as well as against the players of the opposite team. Consequently, the movement behavior of each player is strongly influenced by the context in which it occurs. Without further processing and analysis, the so far gathered data alone does not provide deeper insights into a match. Detecting what influences moving entities in groups can help to semantically annotate or develop methods for the automatic interpretation of relevant intentions.

We propose several methods incorporating different aspects of context in our analysis techniques and provide visual and data analysis support for annotating important types of soccer match elements. Given the gathered movement data of a whole match (Chapter 3) or a specific match episode extracted by state-of-the-art techniques [SHJ+15], we introduce complex models illustrating how players influence each other when moving, competing and cooperating as collectives. We contribute effective automatic annotation methods addressing essential pillars of soccer match analysis, such as interaction spaces, free spaces, dominant regions, cover shadows and player reactions. Ultimately, we add the context of the *real* world represented as in the video recording of the event. By doing so, we, additionally, solve the problem of many state-of-the-art visual analysis systems in team sports being too abstract from the perspective of the analysts who are used to analyze matches from inside the stadium. In our approach, we propose a visual analytics system that tightly integrates team sport video recordings with abstract visualization of underlying trajectory data, therefore, enabling analysts to draw on the advantages of both analysis forms. The resulting system automatically displays complex and advanced 2.5D visualizations superimposed on the original video recordings. The resulting techniques serve as foundation for the effective visualization of cooperative and competitive behavior in Chapter 5.

4.2 Detecting and Annotating Movement Context

In team sport, the interactions among the players during the match contribute significantly to the overall performance. The ability of players to respond to each other, to correctly identify what other players in their team are planning, and to foresee what the opponents' next move is, are all critical contextual elements that contribute to the quality of the single player, and the success of the team as a whole. In the following, we illustrate the possibilities of context-oriented visual annotation on region control, player as well as event analysis which proofs to be relevant, coherent subsets/subareas of important, needed contextual match information during the analysis process. In the next paragraphs, we describe the proposed advanced annotation capabilities in detail. We additionally provide several videos of our results (http://files.dbvis.de/stein/Interaction_Space.mp4, http://files.dbvis.de/stein/Free_Space.mp4).

4.2.1 Interaction Spaces

Soccer players in our expert's traditional working environment are visualized as colored points on a soccer pitch. But players are characterized by more than their x- and y-location at a specific moment in time. During a soccer match, each player has a surrounding area, which he aims to control. This area is called interaction space. In the following, we introduce a model consisting of several factors determining the notion of interaction space. Interaction spaces are especially interesting when analyzing passes or their reception. Furthermore, player duels are directly related to interaction spaces. Other in-depth analyses such as the detection and assessment of pass options build on the calculation of interaction spaces.

A Model for Determining Interaction Space

The foundation for our calculation of interaction spaces was conceptualized by Grehaigne et al. [GBD97]. By performing several experiments they measured how the interaction space of a player morphs according to different velocities. Interestingly, small changes in speed are enough to drastically change the interaction space from a circle to a circular sector. We extend this discrete model and derive a continuous model by using polynomial interpolation allowing a

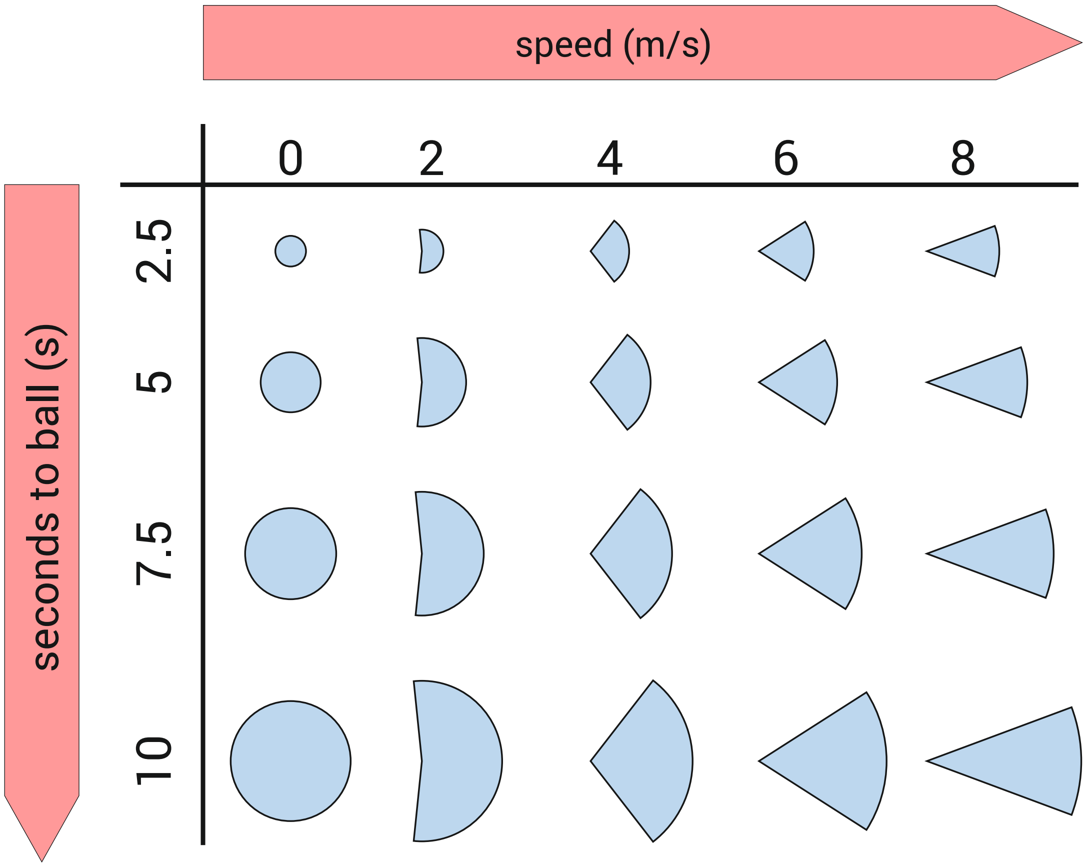


Figure 4.2.1: Speed and distance between player and ball influence the calculation of interaction spaces. The faster a player moves forward (to the right of this figure) the less possible are sudden changes in heading.

smooth transition of interaction spaces as depicted in Equation 4.1.

$$\begin{aligned} angle(v) = &-0.0038\pi \cdot v^3 + 0.0793\pi \cdot v^2 \\ &- 0.6108\pi \cdot v + 2\pi \end{aligned} \tag{4.1}$$

Furthermore, we enhance their model by integrating the distance between player and ball reflecting that a player who is further away from the ball has more time to react described by R in Equation 4.2. Figure 4.2.1 illustrates our approach with different velocities and distances between ball and player. The maximum time of ten seconds will occur when ball and player are located in opposite corners of the soccer pitch.

$$radius(v) = \begin{cases} R \cdot \bar{v} & \text{if } v \approx 0, \\ R \cdot v & \text{otherwise.} \end{cases}$$

$$\begin{aligned} \text{with } radius &: \text{radius of interaction space} \\ v &: \text{speed of player} \\ R &: \text{time until ball reaches player} \\ \bar{v} &: \text{average speed of player} \end{aligned} \tag{4.2}$$

When a player is next to players of the opposing team, his movement is restricted and he needs to pass them by. Furthermore, opposing players will try to block him. Interaction spaces have to reflect these interdependencies as shown in Figure 4.2.2. The interaction space of a player is restricted to the area that he can reach before opposing players. The intersection of interaction spaces is computed and used to visualize the corresponding restricted interaction spaces.

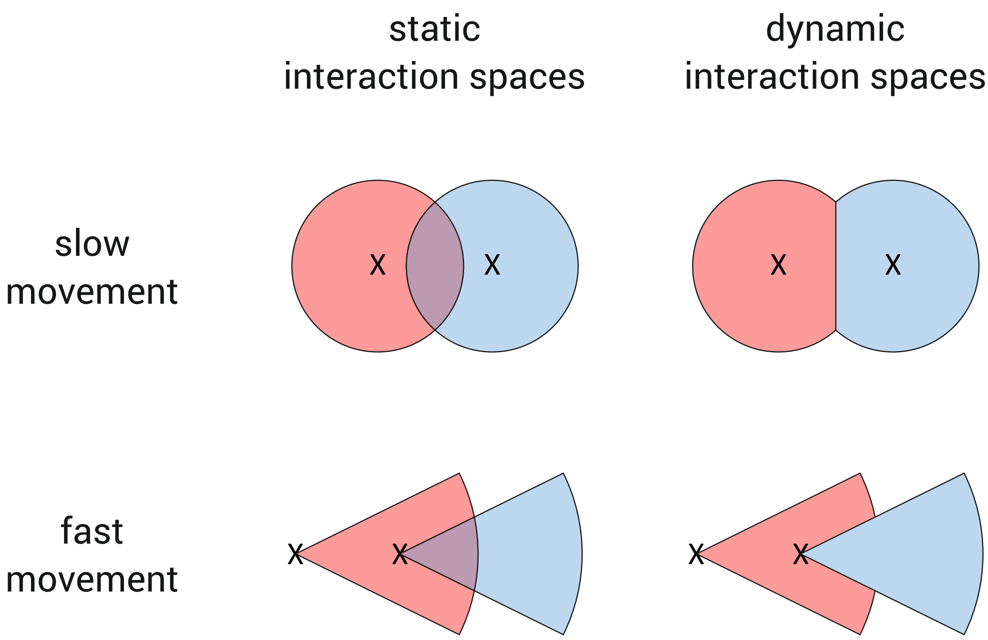
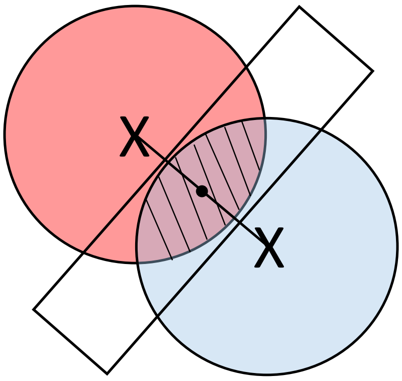


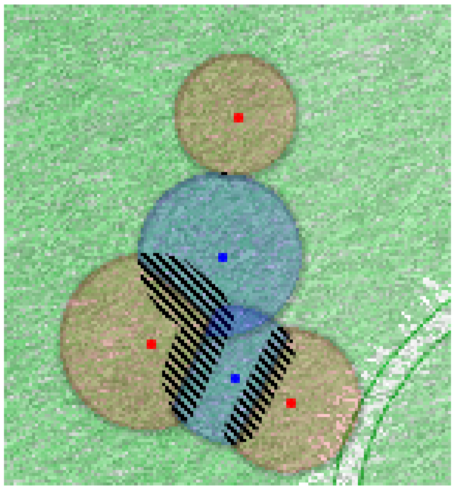
Figure 4.2.2: Interaction spaces are influenced by adjacent players

Potential Duel Areas

When two interaction spaces overlap, it is not clear who of the two players will dominate the intersection area. Individual skills, reaction times, and chance all influence ball possession or loss in these ambiguous cases. We visualize these areas by hatching as depicted in Figure 4.2.3. Hatched areas can be described as locations being potentially reached by both players simultaneously. Consequently, we first intersect both players' interaction spaces and determine the potential duel areas in the middle of both players.



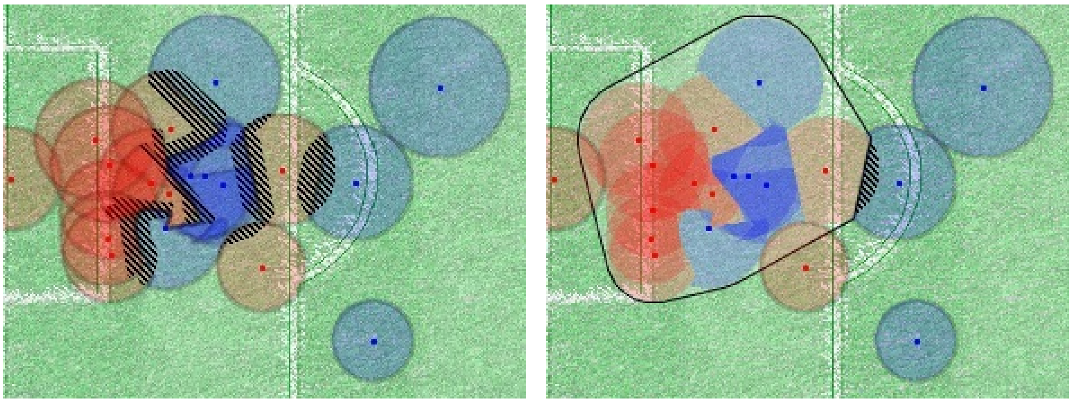
(a)



(b)

Figure 4.2.3: Potential duel area of two players visualized by hatching

In soccer, it is only natural that many players are close to each other and consequently several potential duel areas will arise. We optionally avoid visual clutter of presenting too many nearby duel areas by unifying these, showing a contour line. We show our aggregation-based approach in Figure 4.2.4. In these surrounded areas, we use transparent colors and depict the density of players. We assume that intense coloring depicts a higher possibility for one team winning the ball. We apply an adjusted density-based clustering method (DBSCAN) to detect players being close to each other. The clustering method needs as input parameters the number of players forming a cluster and a neighborhood radius to accumulate players to a cluster. In our case, the radius is dependent on the size of the respective interaction space. Found clusters are visualized by their convex hull.



(a) Many overlapping duel areas **(b)** Hull-based visualization

Figure 4.2.4: We provide an optional aggregation-based visualization avoiding visual clutter in dense regions.

4.2.2 Free Spaces

Free space is a very interesting, important contextual feature of soccer. However, there exists only an intuition among soccer experts and not a precise definition of free space, consequently, its automatic estimation is not straightforward. Based on our interviews with the subject matter experts, we identified two approaches to assess free space.

The first approach judges the notion of relevance of free space. The basic assumption is that the complete soccer pitch can be seen as free space with 22 exceptions. Employing domain knowledge, analysts manually partition the soccer pitch into several areas and decide which free spaces are relevant and irrelevant. Unfortunately, the resulting free spaces are not necessarily reproducible by asking several analysts. The second approach describes free space as a measure for how much a player is put under pressure. Pressing will hinder a player to move freely around and interact with the ball. We categorize pressing into three classes:

No pressure. The player is able to move freely around.

Weak pressure. The player is already being targeted by an opponent moving towards him. The player has less time to act.

Strong pressure. The opponent is close to the player and tries to get the ball. The player has nearly no opportunity to act proactively as his priority is to defend the ball. Chances for errors increase severely.

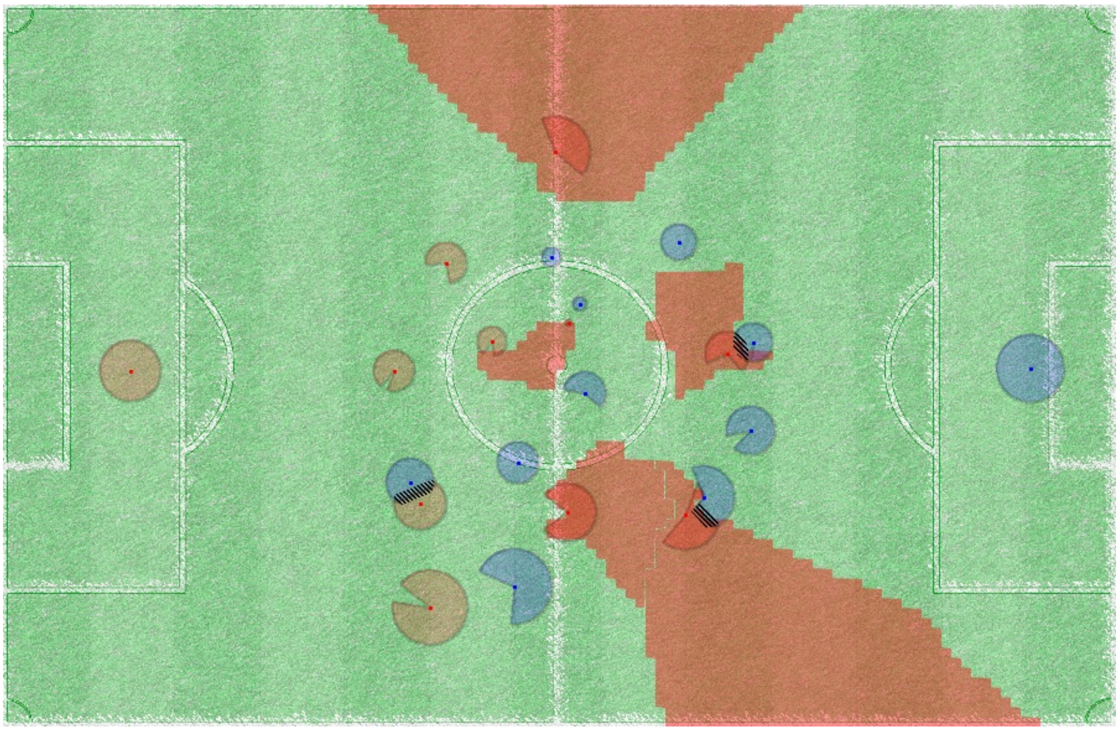


Figure 4.2.5: Grid-based free space visualization for five players of the red team.

With our approach, we follow both ideas of the free space relevance as well as the amount of pressure a player is experiencing. We assess detected free spaces by the respective size, the amount of opposing players, and the distance to the opposing goal. In detail, our method works as follows. We segment the soccer pitch into grid cells of one square meter. We assign to each cell the player with the highest probability of arriving there first with respect to distance, speed, and heading. Consequently, free space can be defined as the region a player can reach before opposing players. We visualize the resulting free space by drawing a colored grid on the soccer pitch. The analyst is able to perceive an overview about the spatial distribution of both teams. The analysis can additionally be steered into the direction of player behavior, e.g., when inspecting zonal defense. Figure 4.2.5 illustrates our free space visualization.

4.2.3 Dominant Regions

In addition to interaction spaces and free spaces, we also introduce a method to annotate *dominant regions* of a team. These regions represent areas where one team is more in control than the opposing team. We calculate dominant regions by taking into account how many players of each team can reach a dynamic region of the pitch first. Similar as for free spaces, we segment the pitch into regular grid cells and calculate for each cell the order in which players can reach it. To ensure that we only identify dominant regions that are present over a longer period, we use a sliding window approach and calculate the moving average over the results. Using the *marching squares*-algorithm [Map03], we detect the contours of the resulting dominant areas that can be reached first by at least three players of the same team. The resulting dominant regions are visualized using a dark red to dark blue bipolar colormap (with transparent colors for tie areas) representing both teams. An example for the dominant region visualization applied on an original video recording is depicted in Figure 4.3.2 (a).

4.2.4 Cover Shadows

Cover shadows are directly related to the kind of pressing each individual team plays. Annotating how players move out of the cover shadows of the opposing team allows to adapt pressing strategies to, for example, force the opposing team out of their normal behavior. A dynamic and valid calculation of cover shadows serves as foundation to this end. In the following, we provide a potential solution for the dynamic calculation of cover shadows. In our model, cover shadows are quadrilateral areas behind a defending player which cannot be accessed by a direct pass. For the calculation of the outer vertices, it is necessary to define a width of a player. For example, if a player has the width of one meter, the player can be viewed as a one meter wide, opaque wall. We base the actual width on the previously introduced interaction spaces. The foundation for our calculations can be seen in Figure 4.2.6 (a). To obtain the first two outer vertices P_1 and P_2, we connect player P and ball B by a straight line whose slope is defined by

$$m = \frac{P.y - B.y}{P.x - B.x} \tag{4.3}$$

With this, we calculate a straight line g through P which is perpendicular to h. The slope of g is defined by the negative reciprocal of m, resulting in the linear equation:

$$g(x) = -\frac{1}{m}(x - P.x) + P.y \tag{4.4}$$

To identify P_1 and P_2 on this line we define a circle c with radius $d = \frac{1}{2} * width\ of\ player\ P$

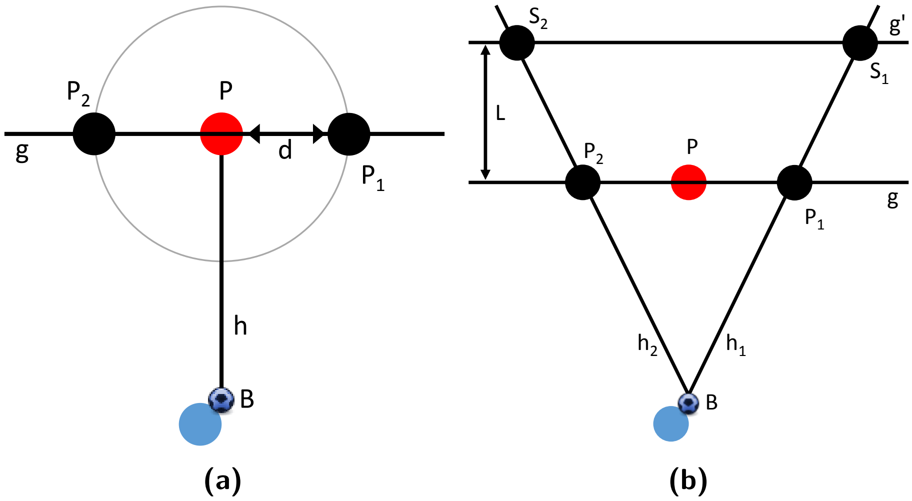


Figure 4.2.6: Cover Shadow Calculation

which satisfies the condition

$$(x - P.x)^2 + (y - P.y)^2 = d^2 \tag{4.5}$$

This ensures that all points on c have distance d to player P. Finally, intersecting c and g yields P_1 and P_2 as follows:

$$\begin{aligned} P_{1,2}.x &= P.x \pm \sqrt{\frac{d^2}{1 + (\frac{1}{m^2})}} \\ P_{1,2}.y &= g(P_{1,2}.x) \end{aligned} \tag{4.6}$$

To obtain the final two vertices, we employ a parallel shift of g to get a canvas line where the shadow will be displayed. To ensure the position of the cover shadow behind player P, the shift will be positive if the y-coordinate $P.y$ of player P is greater than the the y-coordinate $B.y$ of the ball B, otherwise negative. This parallel line is defined by

$$g'(x) = g(x) + sign(P.y - B.y) * L \tag{4.7}$$

where L is the length of the shadow. Empirical studies lead us to propose an inverse dependence to the distance b between player P and the ball B with a factor of $\frac{1}{40}$. Therefore, we set L to

$L(b) := \frac{1}{(40*b)}$. As shown in Figure 4.2.6 (b), we then calculate straight lines h_1 & h_2 through B & P_1, and B & P_2, respectively:

$$\begin{aligned} h_1 &= \frac{P_1.y - B.y}{P_1.x - B.x}(x - B.x) + B.y \\ h_2 &= \frac{P_2.y - B.y}{P_2.x - B.x}(x - B.x) + B.y \end{aligned} \tag{4.8}$$

Finally, we obtain the final two vertices S_1 and S_2 by intersecting g' with h_1 and h_2 respectively. In general, the intersection point Q of two lines $y = mx + c$ and $y' = m'x + c'$ can easily be calculated:

$$Q = \left(\frac{c' - c}{m - m'}, \frac{mc' - m'c}{m - m'}\right) \tag{4.9}$$

By using the coefficients of g', h_1 and h_2 the final intersection points can be retrieved. The cover shadow of a player is now defined by 4 points (P_1, P_2, S_1, S_2) and can be calculated for any given timestep during a match. An example illustrating the visualization of cover shadows can be seen in Figure 4.2.7. It can easily be observed how the red team is trying to avoid the cover shadows of the blue team, searching for an optimal way to position themselves.

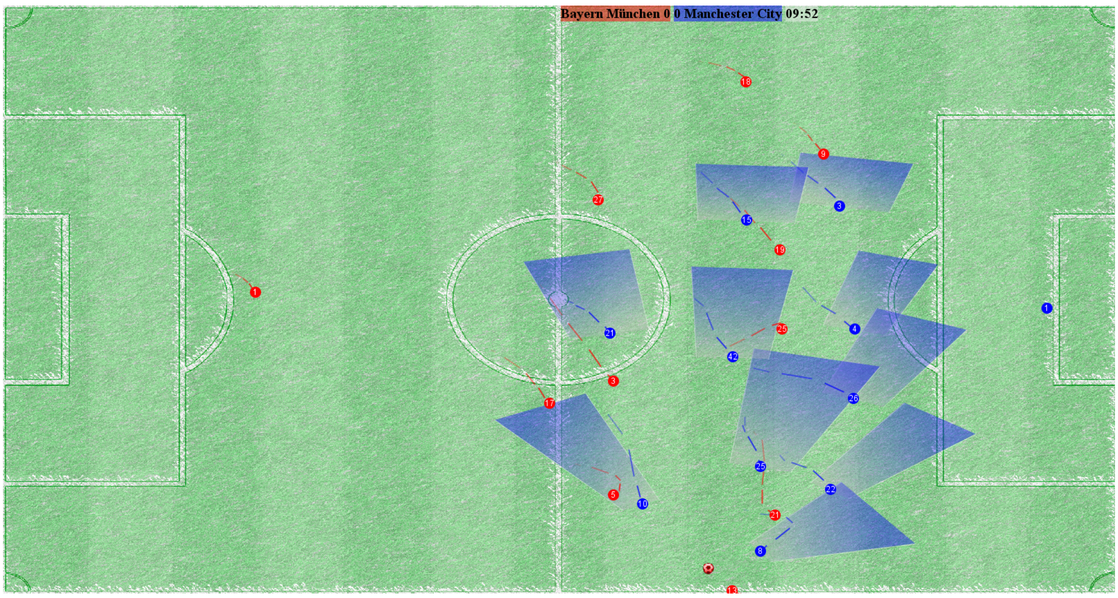


Figure 4.2.7: Example of a cover shadow calculation for a match of the UEFA Champions League between Bayern Munich and Manchester City.

4.3 Combining Video and Movement Data

Video analysts are typically used to work with video recordings and not with abstract data representations. While per se useful in many cases, we also observe that interpreting results obtained from abstract visual representations in context of the actual game situation requires substantial mental effort by analysts. Often, experts want to verify observations made in the abstract visualization space in the video space, or even require the video space to interpret patterns in the visualization space. We hence identify the need to integrate soccer data visualization with soccer video for enhanced analysis and context provision. Consequently, this section contributes an automatic method that provides effective visual analysis in the domain of team sport analysis by integrating appropriate analytical visualizations within the video context. We address the system aspect, demonstrating the applicability of a computer vision framework enabling visual analytics in movement context. We build on the foundation of our data extraction framework from Chapter 3 which enables capturing movement data by extracting player and ball positions and detecting the relative-view-port from standard TV broadcasts. We provide a technique which enables the user to map two-dimensional analytical visualizations back to the video recordings in a perspectively correct way.

To the best of our knowledge, our system is the first to allow visual analysis of soccer matches relying on appropriately defined movement visualization techniques mapped with high precision into the match video. Hence, our approach provides analysis considering both visual abstraction of movement features and match video context, supporting effective use of both modalities in an integrated system. In a sense, our approach also relates to the ideas of *Immersive and Augmented Analytics* which has recently again received interest by the visualization and analytics communities, e.g., as seen from the session on Immersive Analytics at IEEE InfoVis 2016 and the IEEE VR 2016 Workshop on Immersive Analytics. This section adds to this direction specifically for team sports analysis by bringing together abstract analytical visualization with media from the real world.

4.3.1 Integration of Visualizations in Video Recordings

Our system described in Chapter 3 has the advantage that we can perform all data analysis on the normalized reference pitch. Due to the simple homographies with respect to the video frames, we can easily embed any information visualizations targeted to the reference pitch within the original camera images, see Figure 4.3.1 (E). To further emphasize the impression that resulting visualizations are drawn directly on the pitch and players are running on top of them, we remove all pixels that are related to the pitch and pitch markings from detected players regions by background subtraction. Afterwards, the cleaned player regions are drawn on top of the visualizations. In Section 4.3.2, we show diverse examples illustrating the many possibilities for analysis.

4.3.2 Visual Analysis of Soccer Video

We identified in our expert discussions and previous studies three main analysis areas in the soccer domain. These three areas cover some core aspects of soccer analysis, namely, analyses based on ***regions***, ***events***, and ***players***. For the demonstration of this three types of technique, we make use of previously proposed and established visualizations to showcase how existing analytical solutions can benefit from our system. We describe the techniques in the following. Note that while we here focus on soccer, our framework can be extended to other invasive team ball games, like handball or basketball. Such adoptions would require a slight adaption of the tracking techniques, implementation of appropriate analytics methods, and corresponding visualizations.

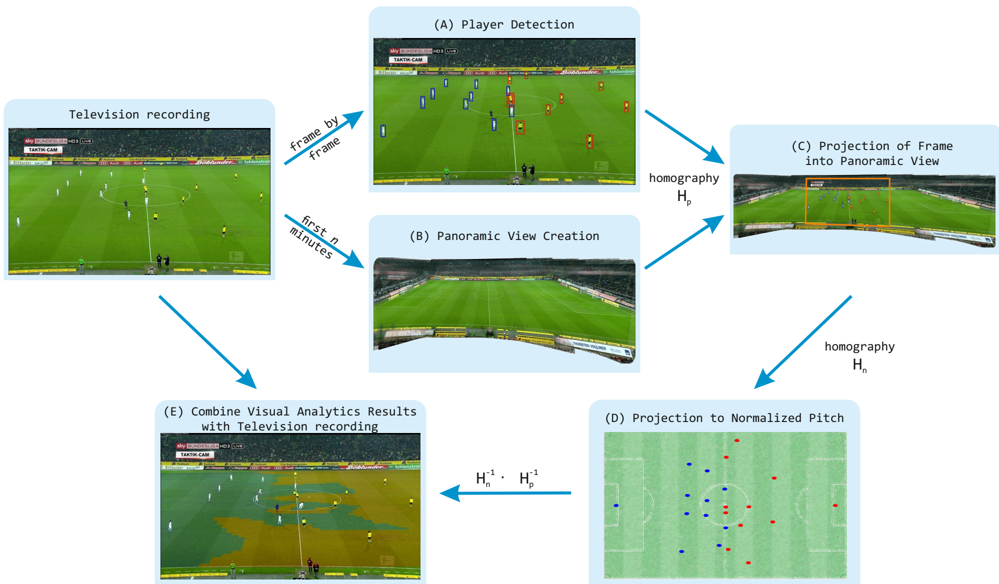
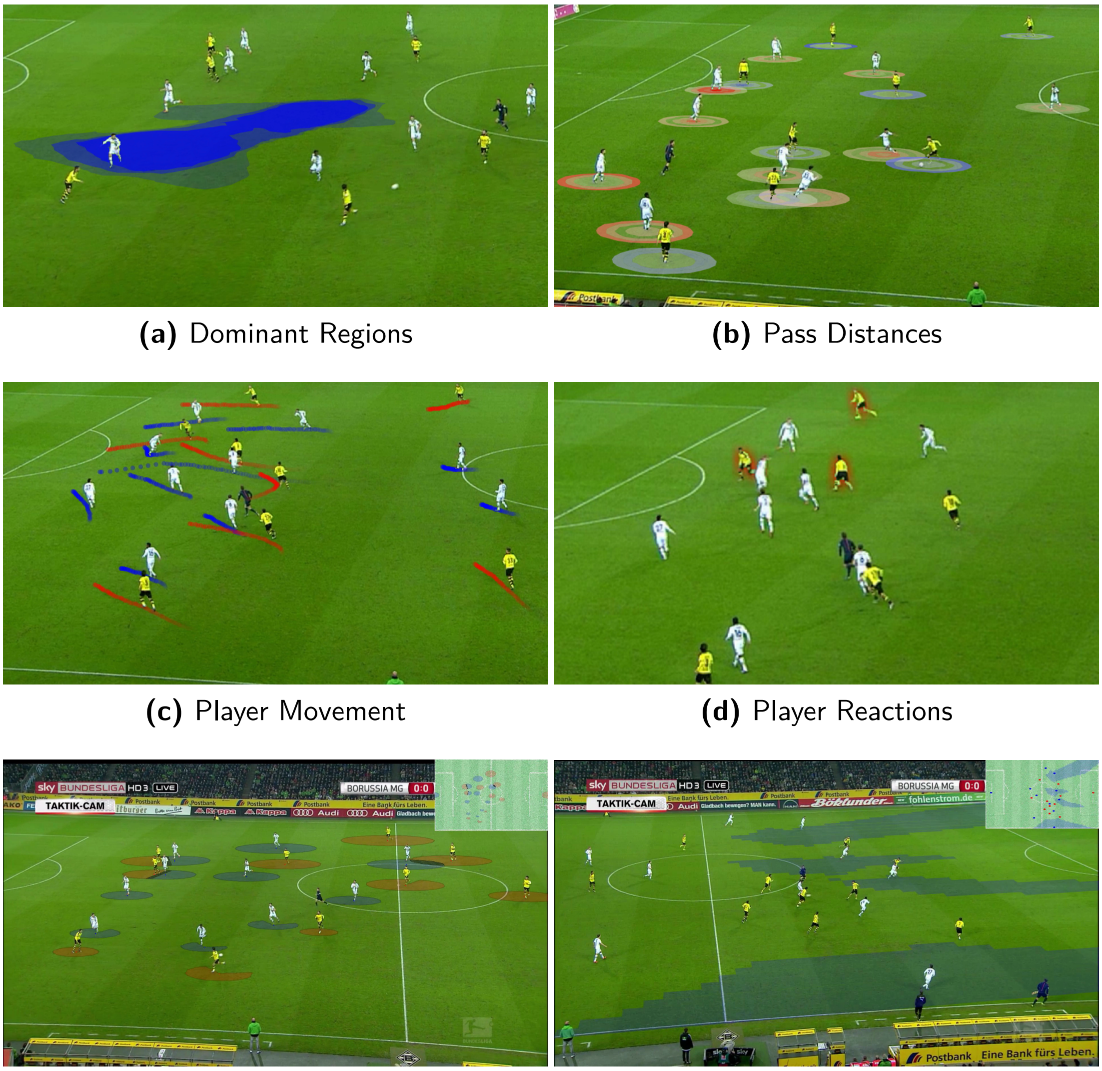


Figure 4.3.1: Extended Workflow of our data capturing process. Based on the projection step (E), we realize the integration of visualizations on the pitch for the video.



(a) Dominant Regions **(b)** Pass Distances

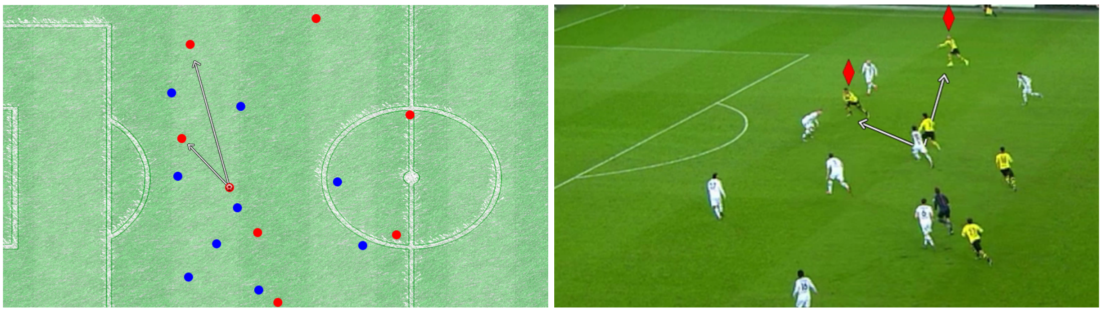
(c) Player Movement **(d)** Player Reactions

(e) Interaction Spaces **(f)** Free Spaces

Figure 4.3.2: The main visual analytics techniques in our approach are based on players, events and regions. When analyzing a player ((c) & (d)), analysts want to detect movement patterns while event analysis (b) aims to, e.g, describe passing preferences. Region analysis (a) tries to segment the pitch into semantically meaningful units.

Region-based Analysis

Spatial analyses in soccer mostly deal with the notion of control of regions, i.e., which team or player was dominating which part of the pitch under specific conditions. The previously introduced interaction spaces, free spaces as well as dominant regions (see Section 4.2) are ex-



(a) **(b)**

Figure 4.3.3: Pass options were visualized so far on an abstract pitch (a) [SJB+16]. Our proposed video visualization (b) allows us to add additional highlighting features such as a floating colored diamond-shape above the player heads in (b).

amples for this kind of analysis. With the use of our proposed system, we can now visualize interaction spaces (Figure 4.3.2 (e)), free spaces (Figure 4.3.2 (f)) as well as dominant regions (Figure 4.3.2 (a)) on top of the original video recording.

Event-based Analysis

Events in team sports can be defined as match-relevant actions which occur during playtime (Chapter 2), for example, a shot on goal, a cross or a pass. Ball passing in a game situation is the result of a complex decision-making process influenced by pressure of the opposing team and the relative positions of each player. The passing behavior is usually analyzed to reveal typical tactics and to improve overall game play. Players need to take many factors into account when deciding for low-risk or more insecure passes to gain space. In previous work [SJB+16], an analysis technique calculating pass alternatives for a player in a given game situation was introduced. So far, arrows pointing towards potential receiving players on an abstract pitch are employed to visualize possible pass alternatives. While integrating the pass visualization into the video, we extended the original two-dimensional version by additionally highlighting players, for example, with a floating colored diamond-shape over their heads. In Figure 4.3.3 (a), we show the previous two-dimensional visualization on an abstract rendered pitch in comparison to our enhanced version visualized on the pitch in Figure 4.3.3 (b).

Furthermore, analysts are also interested in the characteristics of player passing behavior. Understanding whether a player prefers to play short instead of long passes can provide tactical advantages for match preparation. To get an intuition about the passing preferences of a player, we developed a new visualization for player centered pass distances. Figure 4.3.2 (b) illustrates our approach. Circular rings drawn around each player visualize the aggregated normalized pass distance of the respective player. Color saturation is used to indicate whether a player usually prefers short or long distance passes. We normalize the aggregated pass number globally for all players and distance bins. Concentric rings close to the player represent short passes. This visualization allows identifying players with many passes and reveals their usual passing behavior.

Player-based Analysis

Besides regions and events, sport analysts need to focus on individual players and their performance. Analysis typically either concentrates on statistical performance measures (e.g., distance covered by a specific players or failed passes) or actual player movement in context as key factor to reveal and predict tactical patterns. To support the assessment whether players move or not like expected in a given situation, we enable the user of our system to display the previous and next *k* seconds as live trajectory views in the video. We show the past trajectory different to the future trajectory by using transparency. An example can be seen in Figure 4.3.2 (c).

When analyzing team sport it is, however, important to focus on more than the movement of a single entity. The collective movement, here expressed by the two opposing teams, reveals more insights about strategies and tactics. Players moving similarly or reacting on each other is important to observe as this might be an indicator of tactical patterns. Following the idea of Laube et al. [LIW05] and as an exemplary first step towards this challenging problem, we implemented a method to visualize players reacting on each other. We define a reaction as imitated behavior, for example, if one player is strongly accelerating in one direction and other players start accelerating in the same direction shortly afterwards. Using our proposed system, we increase the salience of players acting coherently. For example, we are able to give players that are to be grouped together the same jersey color. Having the same jersey color, however, may result in losing distinct teams. Our design alternative is to emphasize corresponding groups of players with distinctly colored halos instead. An example can be seen in Figure 4.3.2 (d) where two players react on the ball possessing player and accelerate very fast to position themselves in a good spot for a potential shot on goal.

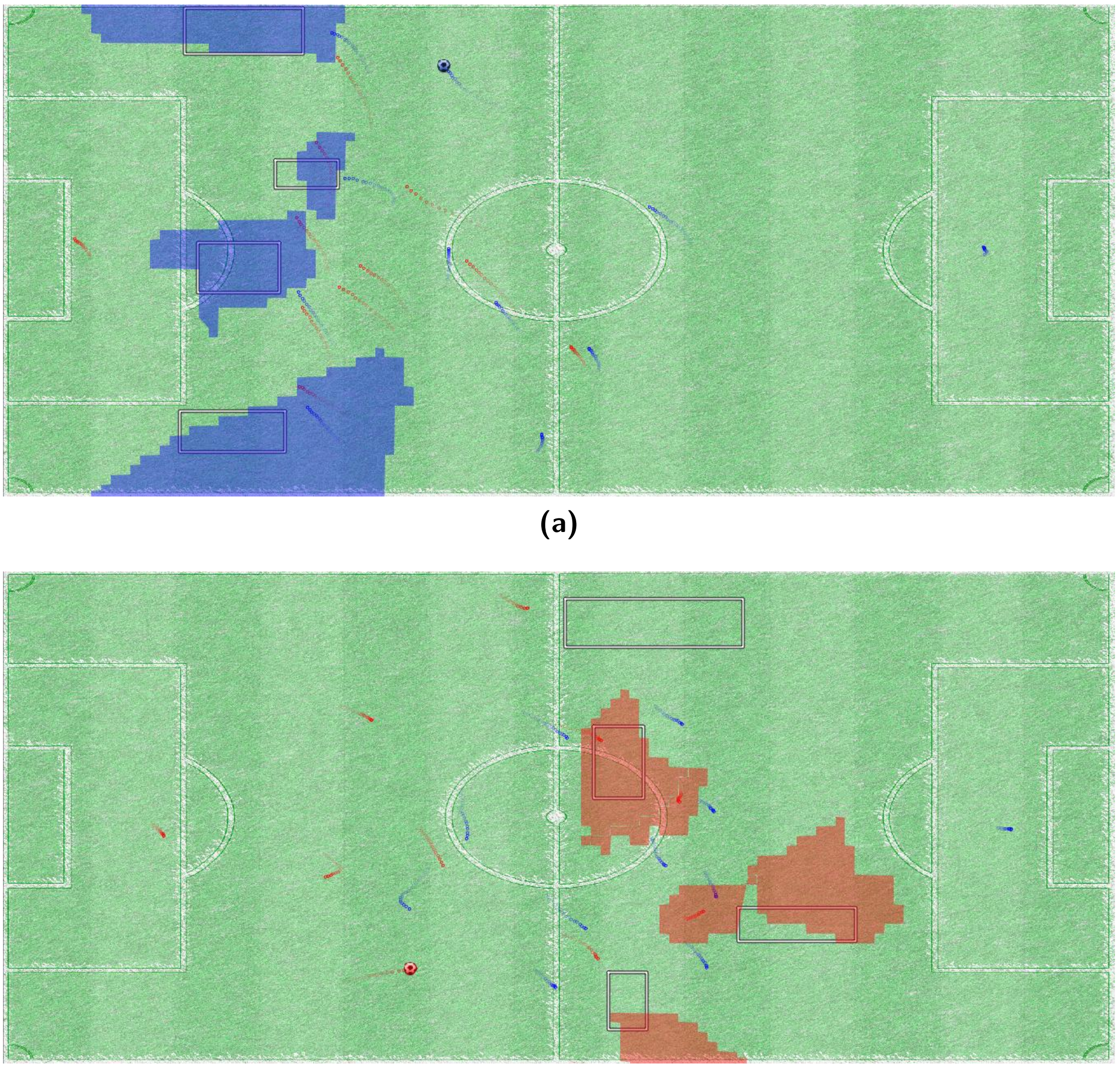
4.4 Evaluation

We demonstrate the applicability of our proposed context annotation techniques by giving insight into the results of several conducted quantitative and qualitative evaluations. At the beginning of a study, each expert was explained the annotation methods, the system and the interaction possibilities. In the following, each expert could openly navigate through the system and try out all available annotation methods and visualizations. Besides the interaction with the system, we performed open discussion rounds which we used to ask related questions. Furthermore, we encouraged the experts to express *ad hoc* comments (the thinking aloud method [ES84, BR00]). Expert interactions with the system were recorded for subsequent analysis. During the studies, we took notes of comments and recorded our observations. We concluded our evaluation with a structured interview regarding the understanding and usefulness of our proposed system. In detail, we asked for the expected impact that our introduced methods will have on the work of a professional soccer analyst and how the various techniques (interaction spaces, free spaces, dominant regions and cover shadows) were experienced. We were, additionally, highly interested in determining whether a video integrated visualization increases the experts' trust. The results of our expert studies are as follows.

For the quantitative evaluation of our introduced free spaces we conducted an evaluation, where we showed the same 52 situations containing 204 free spaces from several soccer matches separately to two of the invited experts (Expert A and Expert B). Each situation was visualized by showing the spatial locations of the 22 players and the ball. We gave our experts the task to draw the four most important free spaces for each situation. After the expert finished annotating each situation manually based on his knowledge, we automatically computed the free spaces with our method from Section 4.2.2. Consequently, we can assess the accuracy of our technique. Figure 4.4.1 shows example results for situations with manually annotated free spaces combined with an overlay of our calculated free spaces.

The results of our expert study are very promising. Even hard to detect and not obvious free spaces have been identified correctly in various sample situations. The first expert found 171 of the 204 free spaces while the second expert identified 154. In 140 of 204 cases both experts detected a free space found by our technique as well. During our evaluation study, the domain experts drew 236 distinct free spaces overall. In 167 of these 236 cases, both experts marked the same free spaces. Consequently, both experts agreed among themselves in approximately 70 % of the free spaces. Evaluating our automatic free space detection algorithm, we counted

326 matches of manually drawn free spaces with the ones proposed algorithmically. As both experts together could match their annotation with ours in 408 cases, this results in an accuracy of 79.65 %. As a consequence, our introduced method calculating and visualizing free spaces of a team can be seen as valid and valuable. Confirming our viewpoint with respect to the complex-



(a)

(b)

Figure 4.4.1: Evaluation result of two example situations where the expert had to draw in the free space (white rectangles). Afterwards, we computed the best free spaces with our technique and overlaid them (blue and red shapes). The left figure shows a complete match of manual and automatic annotation. The right figure characterizes one missed free space. In this case, someone runs from the back freely over the sides.

ity of free space detection, experts do agree among each other only in two thirds of the cases. Nevertheless, we find free spaces in around 80 % of the cases confirmed by at least one expert. Furthermore, the analysis of free spaces in video recordings is seen very important by the experts. Free spaces are seen as help to effectively identify offensive options for the team that is in possession of the ball. Visualization and visibility of free spaces are seen likewise positive. The experts, however, found the high refresh rate of free spaces partially irritating. Their high update rate (each video frame) can result in situations where small changes in the resulting free spaces are causing a perceived 'flickering' in the video visualization. A possible solution beside the limitation to, e.g., ten updates per second could be to define core areas of free spaces and cut the partially rapid changing margins. We updated our free space visualization according to the experts' comments (Figure 4.4.2).

Figure 4.4.2: The experts found the initial high refresh rate of free spaces partially irritating. As a consequence, we updated our free space visualization by defining core areas of free spaces.

Our qualitative evaluation of interaction spaces showed that the computed interaction spaces as well as their projection to video recordings correspond with the experts expectations. The experts consider interaction spaces as additionally very helpful to highlight player formations. Furthermore, our experts state that interaction spaces in video recordings support the perception of players being far away from the camera. The domain experts also approve our approach for the calculation of dominant regions. They consider the nuances in our visualization of dominant regions representing the amount of players that have control over a specific area as a gain, enabling them to segment pitch space into relevant and irrelevant regions. Furthermore, dominant regions are considered useful when identifying regions which need to be covered by players to prevent dominance of the opposing team. Cover shadows are considered "very useful and important" (exact quotation). The experts approve our calculation as correct and want to use the resulting cover shadows to investigate how players move within and out of the cover shadows of opposing players. In addition, our implementation of cover shadows enables experts to easily visualize which players should be passed to.

Pass distances were seen positive as well. The experts want to use such a visualization to detect player roles. Players might, for example, have been formally described as defenders but their playing style might be unknown. Using our annotation approach, this player behavior can now be classified. Analyzing player behavior with our complimentary visualization of player movement trajectories in past and future was interesting for the invited experts when analyzing tackles or individual tactical behavior (e.g., offside trap). Eventually, the experts approved our visualization of player reactions as beneficial information when analyzing group movement. According to our invited experts, players get trained to perform specific movement behavior, for example, to pull apart the defense lines of the opposing team and to open new free spaces. Our visualization, therefore, enables the experts to "make the invisible visible" (quoting one expert).

Ultimately, the experts unanimously appreciated the possibility to perform the analysis directly in the video. All experts consider embedded analysis tools and visualizations in video recordings as the natural (also: "perfect") procedure for in-depth analysis tasks in soccer. One of the reasons the experts mentioned is that the analysis in the video with moving real players is perceived more "natural" as in a generated view. This makes it easier for analysts to contextualize their findings from analytical visualizations. Consequently, the experts argued that their trust in the analysis strongly increases. To this end, they feel that our approach is already very advanced in terms of application in practice.

4.5 Discussion and Conclusion

This section contributed to the automatic annotation of movement context in soccer matches and offers new opportunities not yet available in professional analysis systems. We enhanced the way analysts are looking at data of soccer matches instead of a pure manual and time-consuming video editing routine. We introduced our system laying analysis foundations by providing concrete measures for the annotation of some of the most essential pillars of match analysis: interaction spaces, free spaces, dominant regions and cover shadows. Furthermore, the combination of visual analytics methods with the raw video material is a unique possibility to enhance the usability and acceptance for domain experts. To the best of our knowledge, any two-dimensional visualization that has been designed for the use on an abstract pitch can be used and transferred with our system. The combination of video and visualization bridges the gap between computer and human domain experts, allows to include context information from the video in the analytical visualization and communicates inherent uncertainties regarding pure visualization approaches. Analysts can decide for video recordings with and without visualizations as well as more abstract visualizations on a two-dimensional pitch. The usefulness of our approach is shown by expert studies with real soccer domain experts intrigued by the potential analysis capabilities. We, therefore, believe that the inverse transformation from the abstract data model to the real world enables a more insightful and effective analysis. Experts can identify and reveal uncertainties in the analysis process more easily by relating their knowledge to the visualized analysis results. Furthermore, as we experienced in our expert studies that the experts come up with their own suggestions for visualizations, the combination of video and abstract visualization seem to be a promising communication medium when designing visual analytics systems. We discuss this aspect more in detail in Chapter 7.

The experts judge the presented contextual annotation methods as very helpful for the detection, exploration, and comparison of interesting game situations. The greatest impact is expected in the area of processing and presenting findings within a match. According to the experts, our approach enables analysts to confirm or reject hypotheses by facts instead of intuitions. Furthermore, the amount of time that is needed for the analysis of a single match is expected to decrease drastically. Our experts believe professional video analysts would make extensive use of systems building on our proposed annotation methods. Our research is seen as an important step towards the efficient analysis and explanation of game situations. As next step, the experts proposed to build on our proposed methods enabling a deeper analysis by focusing on duel areas and inspecting particular movement behavior. Specifically, the experts wished for a visualization encoding where a player should move to minimize the room control of an opposing player. Following these lines, this will enable us to define hypotheses whether a player always takes advantage of existing free spaces, or instead rather focuses on distracting the opposing defenders, or in which situations the behavior would change. We propose a method fulfilling this suggestion in the following Chapter 5.3.

In football, the worst blindness is only seeing the ball.

Nelson Falcão Rodrigues (Journalist)

5

Visualizing Cooperative and Competitive Behavior

Contents

5.1 Introduction

The final challenge is to build upon the so far presented techniques with novel methods in order to retrieve explanations of observed cooperative and competitive movement patterns and to understand why, when and how specific movement behavior is expressed because of tactical behavior. A team usually plays with a certain formation, depending on many influence factors such as the match situation or the current score. Consequently, there exists a myriad of possible strategic approaches, tactics, observable behavior patterns and many more parameters that need to be analyzed and which make the analysis difficult. Visualizing the data provides the opportunity to use the human perception and cognition interpreting the data as well as human pattern recognition identifying relations and interdependencies. Consequently, the combination of analysis, visualization, and human domain knowledge is a promising approach for a successful visualization of cooperative and competitive behavior.

In the following Section 5.2, we extend the state of the art for visual match analysis by proposing a four-step analytics conceptual workflow for an automatic selection of appropriate views for key situations in soccer games. Our concept covers classification, specification, explanation, and alteration of match situations, effectively enabling the analysts to focus on important game situations and the determination of alternative moves. Ultimately, we introduce an implementation of the previously introduced alteration step by proposing an automatic approach for the realization of effective region-based what-if analyses in soccer (Section 5.3). The presented system covers the automatic detection of region-based faulty movement behavior, as well as the automatic suggestion of possible improved alternative movements. We enable domain experts to include their domain knowledge in the analysis process by interactively adjusting suggested improved movement while visualizing its implications on region control. As we show, our approach effectively supports analysts and coaches investigating matches by speeding up previously time-consuming work. We demonstrate the usefulness of our proposed approach via an expert study with three of our invited domain experts, one being head coach from the first Austrian soccer league. As our results show that experts most often agree with the suggested player

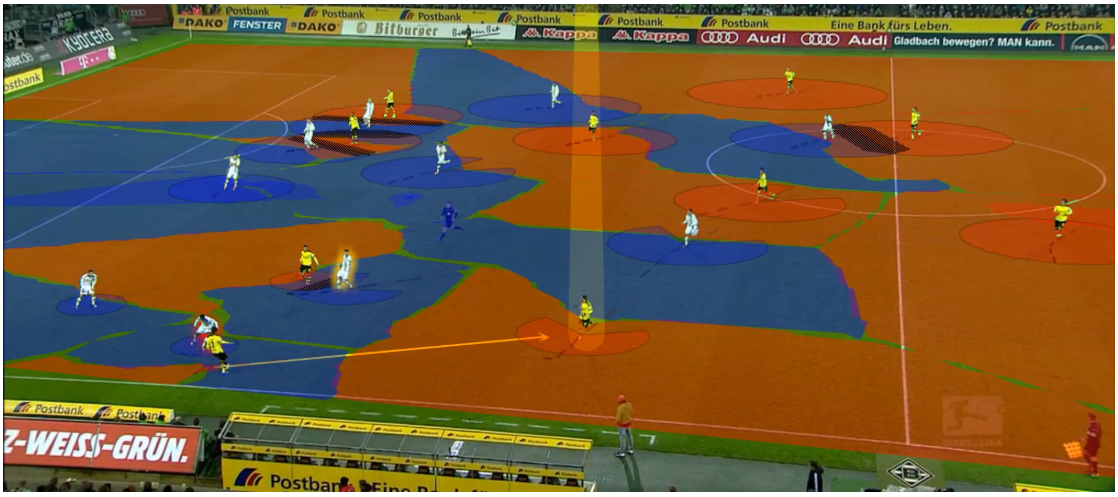


Figure 5.2.1: Various visualizations have been proposed for the analysis of team sports. It is, however, often not clear when to choose and how to combine visualizations. The figure illustrates techniques for interaction spaces, free spaces, pass alternatives and dominant regions (Chapter 4).

movement (83 %), our proposed approach enhances the analytical capabilities in soccer and supports a more efficient analysis.

5.2 Explanatory Storytelling for Open Play Situations

The goal of a video analyst is to detect weaknesses of their own as well as opposing teams, enabling them to gain advantages in the competition. Findings are typically presented to the team members, coaches, managers, and other stakeholders in form of manually annotated video clips of previous matches. The annotation is performed with the help of various symbols such as colored arrows, lines or rectangles. Afterwards, the annotated match scenes are used in the team briefing. The resulting annotations, however, are not only interesting for video analysts. Television broadcasting companies are using manual added annotations to explain interesting moves to their audience as well as to reveal alternatives. State-of-the-art industrial solutions (for example Vizrt `http://www.vizrt.com/products/viz_libero/`) support journalists by providing methods to draw single symbols, to manually highlight players as well as to aggregate player movement by a density heat map.

Manual in-depth annotation and explanation of soccer matches, however, is not feasible because of high-frequent matches during seasons. An additional factor is the duration until suitable annotations must be provided. This affects, e.g., (sports-)TV broadcaster that aim to give first insights and reasonable stories already after a short commercial break. There is a natural trade-off between invested resources (e.g., time, persons, other resources) and quality of the analysis. Automatic player and ball tracking (and directly derived data such as speed and acceleration) have so far been mainly used for quantitative comparisons. Qualitative analyses as the automatic selection and combination of visual analysis techniques which are helpful to identify and understand patterns of movement and player performance is a challenging problem. On the one hand, human analysts are able to intuitively highlight interesting aspects by selecting the proper tool in their toolbox (status quo) but are limited with respect to time and possibly, also cognitive biases. On the other hand, automatic annotation algorithms have a set of visualization algorithms but lack the human ability of assessing when to choose and how to combine visual annotations as illustrated in Figure 5.2.1. The inherent combinatorial problem leads to an exponential search space not feasible for interactive inspection.

Instead, this section contributes a conceptual workflow for finding a valid and useful solution to identify suitable visualizations in this large search space. Our approach includes a rating-based approach for the semi-automatic selection of visualizations and creation of match reports. Our concept includes four consecutive steps: The **classification** of a given situation rating and verifying whether a certain visualization is interesting or not as a first step, followed by a **specification** to take the context as well as the characteristics of a situation into account. Using the results of classification and specification, user specified selections of visualizations can **explain** a given situation. Our conceptual workflow concludes by the assessment of **alterations** that would have been possible during a respective match scene. We propose to employ our concept in the original input video recordings via the previously presented techniques to improve intuitive understanding. This will extend the analytical possibilities, e.g., during television coverage of team sport events by providing novel ways to extract and present interesting stories with the help of visual analytics. We give several showcases in the soccer domain, focusing on player interactions, player behavior as well as match events. Additionally, we provide initial expert feedback to our proposed approach. Our overall workflow can be combined with existing approaches and is largely orthogonal to existing work. To this end, we introduce an implementation of the alteration step in Section 5.3.

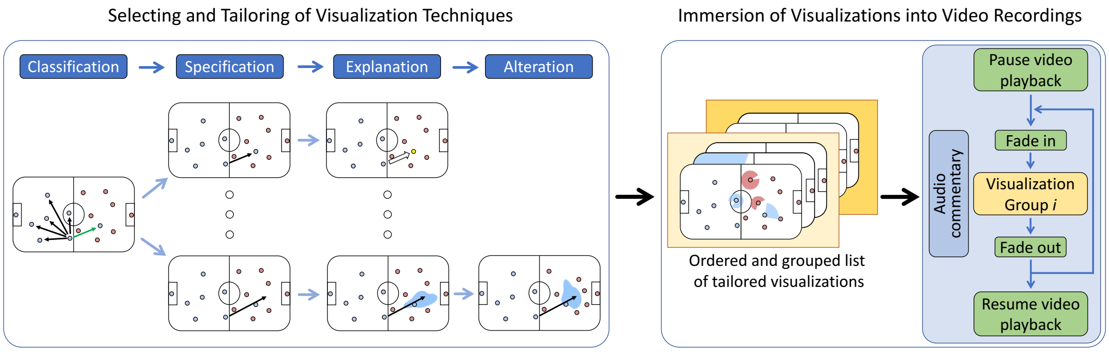


Figure 5.2.2: Enriching a single situation: Our proposed process consists of two steps. Selecting and tailoring of visualizations and combining these visualizations with the video recording. The situation of interest is first classified and afterwards contextualized in the specification step. Fine-tuning visualization parameters helps to explain the situation further. An optional alteration step should allow *what-if* explorations. The resulting list of visualizations is the input for the combination of video recordings and visualizations.

5.2.1 Explanatory Storytelling in Soccer

Lee et al. [LRIC15] describe the structure of a story as a sequence of visualizations where each visualization describes a specific characteristic of the overall story. One of the resulting challenges is to identify appropriate visualizations for given characteristics as well as to determine their order of appearance. Our semi-automatic annotation system is aimed to create stories in the form of annotated video recordings. Following the categorization of Segel and Heer [SH10], our generated story elements are a combination of the types *annotated chart* and *video animation*, in which we blend video footage with elements from trajectory and region-based visualization. Our concept for the semi automatic-annotation of match scenes is depicted in Figure 5.2.2. In this concept, we tackle the large search space by assessing the predefined, parameterized visualizations and, consequently, support the creation of a narrative of events in a soccer match. When applied, we envision our concept as particularly suited for summative preservation of key events of a match, e.g., in a TV sports broadcast. We structure our conceptual workflow according to the following very general questions for a given situation. Please note that these questions are very common in the visualization domain and are by no means restricted to soccer.

A. **Is the visualization relevant?** With this question we cover the transformation from data to visualization.

B. **What were the specifics of the situation?** We slightly adjust parameters of the visualization reflecting the situation context best.

C. **How did the situation evolve?** We bridge the gap between visualization and human with visual explanations.

D. **What are alternative situations or movements?** The iterative loop combining human and computational power with visual analytics enables to derive new insights.

Classification: Assess Available Visualizations

The first step of our proposed conceptual workflow is the **classification**. During the classification, each available visualization needs to be assessed based on the respective relevance for the given situation as well as on the task and analyst's preference. For example, the visualized free space of a goalkeeper might not be of large relevance during an offensive counter attack. Consequently, we need to identify and classify observed movement patterns and determine the suitability of the available visualizations. In order for our proposed concept to be generalizable, the classification will assess different features depending on the respective visualization. Still, some visualizations might be useful in all situations within a specific game phase (e.g., defending phase). These game-dependent meta-features will influence the classification of game situations as well.

Specification: Enrich with Situation Context

If the result of the classification (the assessment) matches a certain threshold, the conceptual workflow continues with the **specification**. The goal of the specification is to use the situation context to categorize and set up a visualization. A pass, for example, can be displayed in several ways focusing on different aspects like pressure, safety, or free spaces. Another game-specific meta-feature employed for differentiation could be whether it is a pass in the own or in the opposing half of the pitch. Features reflecting collective movement patterns are used as well for differentiation. In the case of passes, these features could be for example whether the pass is directed at a player or into his or her free space.

Explanation: Tell the Story

Once the observed situation has been enriched with context information, the existing fine-tuned visualization techniques can be used for the **explanation** of the movement. The results of the previous specification step are very important here as the added context enables to reflect different semantically meaningful situations. Depending on the context, different visual representations will be more suitable than others. In order to enrich the story presentation explaining the details behind a situation, we propose to automatically generate textual as well as audio descriptions. These descriptions assist the user with following the identified storyline. They ensure that the system and the user have the same contextual basis such that misinterpretations can be minimized. We intentionally propose to convey the explanation via the audible information transmission channel. On the one hand, this enables a clear distinction between following a game situation (audio commentary) and the scope for analytical reasoning (visualization). Associating the audio commentary with the well-known metaphor of a soccer commentator substantiates this distinction. Thus, users are able to focus on the audible explanation of the game situation or the visual details according to their needs. On the other hand, the audio description increases the effectiveness of conveying the story without the usage of additional screen space.

Alteration: What-If Analysis

A location change can happen for various reasons. A coach, for example, might want to predict possible changes or alternatives for the current situation as well as where a player should have positioned itself to prevent following actions. The **alteration** step, consequently, enables focusing on these *what-if* scenarios. As the implications of such location changes needs to be analyzable, each visualization needs to get recalculated during position changes so that experts are enabled to compare, confirm, or reject their hypotheses with respect to the displayed visualizations.

Further investigating what-if scenarios, however, is very complex, as each change might have unforeseen implications. Simply moving players closer to the ball will ultimately result in a potentially worse situation as, for example, more free spaces can open up. Furthermore, a player should theoretically be able to reach the proposed position respecting physics. We consider this as an optimization problem assuming that players are constantly trying to reach their optimum location with respect to all other player positions and respective goals. In order to detect possible alternatives, we need to define whether players are losing their optimum location and identify when they are moving less than optimal. Accordingly, one option could be to visually annotate where these players should have moved instead and why they should have moved there. We introduce an example implementation of the alteration step in Section 5.3.

Immersion: Improve Intuitive Understanding

After the four-step concept selecting and tailoring visualizations, we need to create the **immersion** of abstract visualizations and video recordings. A special challenge in the creation process is to display the information at the right time for the proper amount of time. In order to allow the analyst to focus on the provided visualizations, we propose to stop the current situation and fade in the visualizations right after another (right part of Figure 5.2.2). This helps to steer the attention towards the annotation as well as prevents a cognitive overload and change blindness. A visualization itself should be shown as long as the provided audio commentator is speaking. Afterwards, the video continues to the next relevant scene.

5.2.2 Proof of Concept

In order to showcase the usefulness of our proposed workflow covering the complete range of match analysis tasks, we provide a proof of concept in collaboration with the domain experts introduced in Chapter 1.4. We explained our proposed concept during several interviews and asked the invited domain experts to use our concept for the creation of automatic annotation workflows for passes (see Section 5.2.2) and passes into free spaces (see Section 5.2.2). For the proof of concept, we used a television match recording from the German Bundesliga being broadcasted on the Sky Sport TV channel operated by Sky UK Telecommunications [sky]. All figures are extracted from this match and visually enhanced by our conceptual workflow presented above.

As a prerequisite for each presented proof of concept, the analyst has the possibility to provide an interesting video sequence serving as input for the story creation. We intentionally allow to make an initial manual selection in order to limit the created video annotations to what is interesting for the user. Nevertheless, previously proposed systems [SHJ+15] can be used to automatically detect interesting situations for visual annotation. In order for the visualizations to be superimposed on the original video recording, we make use of our previously presented technology (Chapter 4.3). The audio commentary is realized by making use of Microsoft's *Text-*

To-Speech API [mic]. A set of predefined textual comments is provided by the system and can be extended by domain experts anytime during their analysis. Depending on the classification and specification steps, the most suitable comments are chosen for the identified characteristics. Additionally, they are enriched with situation- and game-specific information, such as player names, directions of movement and locations on the pitch.

Passes

Passes are among the most important means in a soccer match to play around opposing players as well as to gain space. We apply our proposed concept in collaboration with the invited experts focusing on passes on the ground. The resulting annotation workflow that can be applied automatically to any given input video sequence can be seen in Figure 5.2.3. The first step in our conceptual workflow (**classification**) is to define when a pass or a pass alternative is relevant for the understanding of a situation. For this purpose, the experts want that each possible pass is assessed by defined and established criteria [SJB+16], e.g., pass distance, intersection interaction spaces, or pressure. If the rating of a pass exceeds a user-defined threshold, we will consider the pass or the pass alternative as relevant for the given situation. For the **specification**, we need to define in which variants a pass can be played. Our experts mostly differentiated between passes in the own and the opposing half of the pitch. According to our experts, it can be assumed

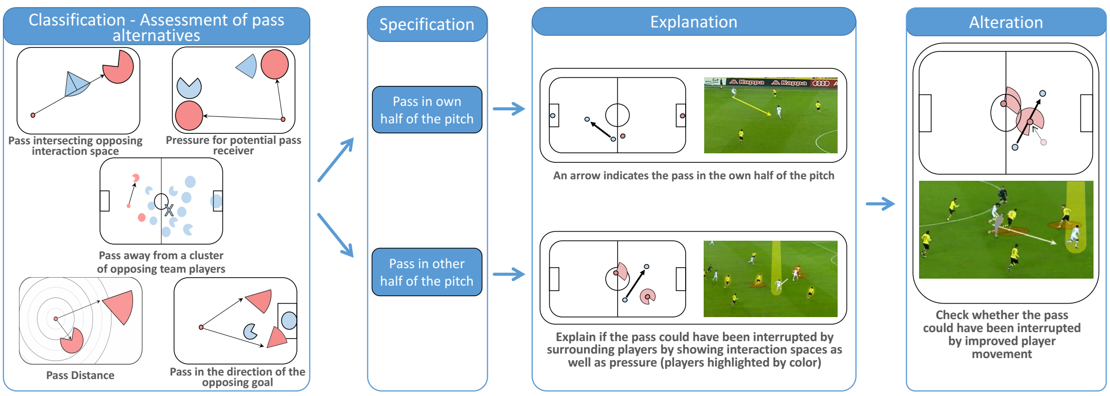


Figure 5.2.3: Exemplified simplification by a common match situation. In this user-defined scenario, we automatically analyze and visually annotate passes, depending whether they have been played in the own or in the opposing half of the pitch. During the explanation, we focus on contextual aspects such as whether a pass could have been interrupted through opposing players interaction spaces. Eventually, we verify whether the player movement could have been better.

that passes in the own half of the pitch, e.g., back passes to team members, are safer [SJB+16] since the probability to lose the ball is lower while passes in the opposing half of the pitch are considered more dangerous as usually more opposing players will try to interrupt it.

In the **explanation** step, we specify the visual representation of the respective pass variants. Our invited experts decided that passes in the own half are sufficiently indicated by arrows on the pitch. Additional visualizations are not required as the probability of a turnover is considered to be low. However, a pass in the opposing half of the pitch can be intercepted more easily by near opposing players. Consequently, the experts verify whether a pass can be intercepted computing and visualizing the corresponding interaction spaces (Chapter 4.2) of nearby opposing players. Another desired information is the pressure a player is experiencing as this increases the probability of losing the ball. Various models to calculate player pressure have been introduced in the past, e.g., by Andrienko et al. [AAB+17]. We visualize player pressure by a colored halo surrounding the pressing player using a color scale from yellow (medium pressure) to red (high pressure). Additionally, the invited domain experts decided that they want to highlight involved players such as pass receivers by a spotlight visualization. In the last phase of our applied conceptual workflow (**alteration**), the experts want to find out whether the pass could have been intercepted by improved player movement. Accordingly, we automatically detect close players which would have been able to gain possession of the ball if they altered their movement slightly based on their corresponding interaction spaces. Those players can be automatically moved to their new position.

Passes into Free Spaces

A more complex form of passing is the pass into free spaces. Here, it is crucial to analyze the observed cooperative and competitive group movement resulting in the spaces players are able to control. The resulting annotation workflow created in collaboration with the invited experts can be seen in Figure 5.2.4. During the **classification** step, our experts defined two main requirements for the creation of a story. The first requirement focuses on the free space of the players, defined as the space a player is able to reach before any opposing player. Several criteria have been proposed in previous works (Chapter 4.2) to assess how good and relevant a free space is in terms of offensiveness (size of the respective free space, its distance to the opposing goal, how much opposing players can be outplayed). A positively evaluated free space without a realistic chance of passing the ball to this position is not beneficial for the analysis. Consequently,

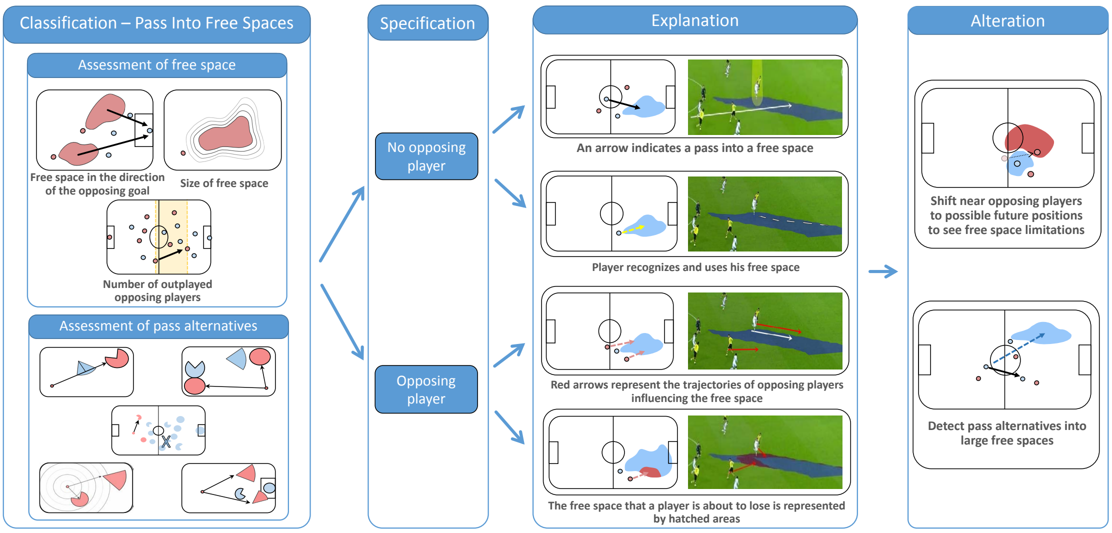


Figure 5.2.4: Interesting passes into free spaces are annotated by incorporating cooperative and competitive group movement context aspects. We distinguish between situations where no opposing player is directly influencing an arising free space and situations where opposing players have a strong influence. The alteration step enables experts to be aware of not realized pass alternatives.

the experts add as the second requirement whether the ball could be played into the identified free spaces. We check if a ball can be played to a certain proposed free space by assessing the possibility to play a pass from the player, who is in possession of the ball, towards the center of the proposed free space. The pass assessment is done based on our defined criteria from the first proof of concept containing regular passing behavior. As we assess the risk of losing the ball, the result is a relevant free space with a high change of not losing the ball possession.

The **specification** and visual **explanation** are focused mainly on the movement behavior of the involved players. The experts distinguish between two main types of possible match situations: the first case describes situations in which no opposing player is limiting the player's path to the opposing goal using a respective free space. In this case, the movement of players of the own team is important, according to our experts. The experts want to find out whether players are able to detect and use their own free spaces. Consequently, we visualize the free space as well as the possible pass options and whether the movement of a player follows the created free space in the next few seconds. The second kind of situations occurs when opposing players strongly influence the arising free spaces, forcing the player who is in possession of the ball to react to their exact movement. To cope with this specific situations, we need to analyze how free spaces are influencing and limiting each other. Consequently, we calculate and visualize the intersection of each free space with every free space of the opposing team for a user-defined timespan, e.g., the next two seconds. This allows detecting which and how opposing players are reacting on a player's free space. Additionally, we visualize the future movement of nearby opposing players, giving us the advantage to verify whether opposing players detected and moved towards a dangerous free space.

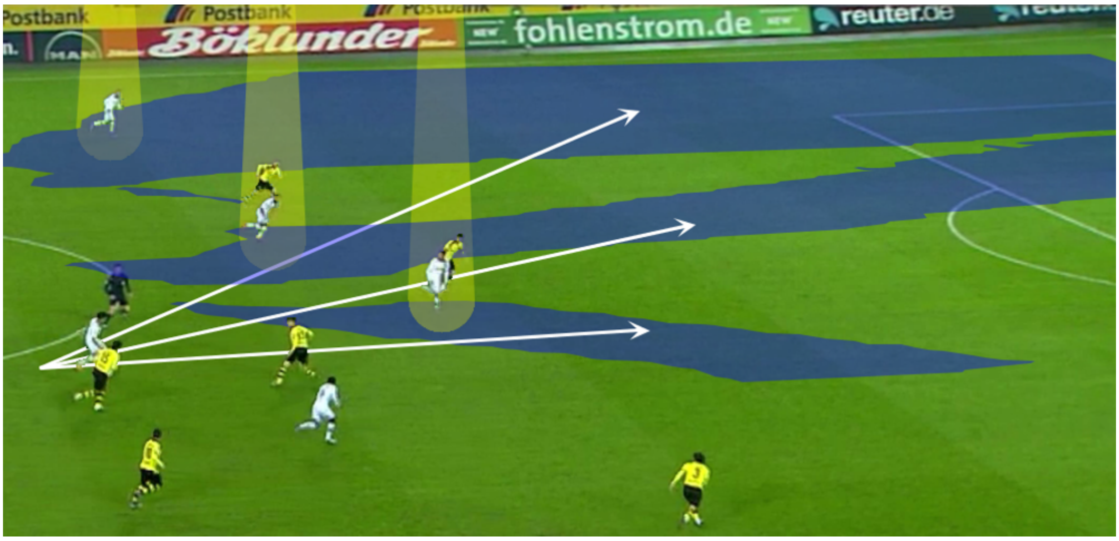
In the **alteration** step, the domain experts are, for example, interested in detecting possible pass alternatives into relevant free spaces that have not been realized during a match. Therefore, we calculate at each timestep whether big offensive free spaces arise for the attacking team and assess the chances of a successful pass based on the criteria described above. If such a pass alternative in an open free space gets detected, we highlight the identified players automatically with a spotlight.

5.2.3 Initial Expert Feedback

To validate whether our invited domain experts agree with the automatic created annotations of our proof of concept, we performed several experiments with each expert. At the beginning of a study, unmodified video recordings of match segments with a time span of approximately 20 seconds were shown to each expert. Afterwards, each expert was asked to replay the scene and to pause the recording whenever he wants to annotate something on the pitch. Experts were allowed to interactively draw on the pitch, using various symbols to highlight players and regions. After the entire situation has been annotated manually, we provided the expert with the automated annotation of the given video sequence. Together with our experts, we then compared and discussed the final annotations with and without automated support to identify strengths and weaknesses of our proposed concept. Additionally, we recorded their interactions for subsequent analysis. Furthermore, we were particularly interested whether our proposed video annotations strengthen the trust of experts into their analysis.

The initial expert feedback revealed that all experts consider our concept as highly valuable for the annotation of soccer matches as well as other invasive team sports. The provided proof of concept is, according to our experts, illustrating the usefulness of our approach. For example, our implemented conceptual workflow for passes into free spaces is already perceived very similar to the annotation our experts would manually propose, as can be seen in Figure 5.2.5. Figure 5.2.5 (a) shows the manually added expert annotations while Figure 5.2.5 (b) shows the

(a) Manually proposed pass alternatives annotated by domain expert



(b) Automatically annotated pass alternatives proposed by our system

Figure 5.2.5: To validate the annotations proposed by our system, we asked domain experts to interactively annotate their findings on the pitch. Blue areas depict free spaces, arrows depict potential pass options while players highlighted with a spotlight indicate potentially involved players. The high similarity of both annotations suggests our concept is capable of supporting domain experts revealing invisible movement explanations.

annotations proposed by our system. The invited experts state that our proposed conceptual workflow is helping in filtering out probably unnecessary visualizations while still allowing the user to intervene in the analysis process when desired. This allows focusing on the important aspects of a situation. The single steps of our conceptual workflow are considered believable as well. The specification step was especially mentioned by one expert enabling to express, detect and visualize the different characteristics of a situation. The shown specifications of our proof of concept already reflect "*up to 90 percent of all possible cases*" (quoting one expert). Consequently, we believe our concept can serve as solid foundation for the creation of match annotations.

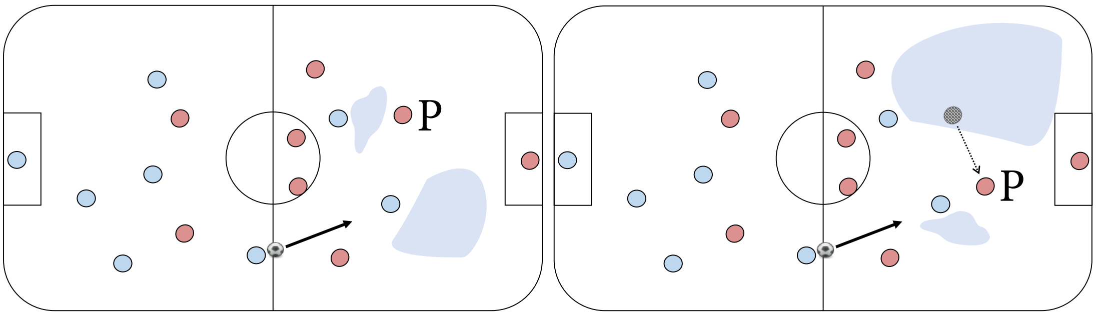
The experts also found the possibility to let the system annotate alternatives moves very useful while the user can still interact and decide whether he wants an annotation in the final story. One expert argued that he would have felt overwhelmed if he could not decide on his own which alternatives to show. The way annotations are displayed in our system is also seen positively. The experts liked that involved players are highlighted within the original video recordings, e.g., by a colored halo or a spotlight. Furthermore, they approved our way of displaying the visualizations one after another and exactly synchronized when the audio commentator is explaining what can be seen. According to our experts, this allows to explain a complete situation in one playback. To this end, our invited experts praised the possibility to dynamically add visualizations to players and regions using a simple drop-down menu and by right-clicking on the respective object.

5.3 Computational and Visual What-If Analyses

In this section, we introduce a novel interactive visual analysis method supporting positional *what-if* analysis in soccer, realizing and implementing the previously introduced **alteration** step (Section 5.2). The resulting visual analysis enables experts to automatically detect and explain suboptimal movement behavior as well as to indicate where players could have moved to improve the situation. To make this possible, our presented approach supports the experts in several parts of the analysis process. Initially, suggestions are made for situations of interest, which should be investigated in more detail. In these situations, our algorithm can then perform the what-if analysis and optimize the players' positions. Additionally, we provide visualizations for assessment that help users to understand the expected effects of the repositioning of the players. Finally, we show that our technique enables domain experts to create realistic and expressive analyses of alternative movement options through several expert studies, including feedback from coaches and analysts of professional soccer clubs. We contribute to the state of the art by enabling analysts to validate hypotheses automatically as well as interactively by manually repositioning players via drag-and-drop on a visual pitch. The impact of the change is re-analyzed and visualized, e.g., in form of free spaces (see Chapter 4.2) and helps analysts to identify improvement options to use during coaching. This combined automatic and interactive approach enables new forms of interactive video analysis, and is especially suited also for discussion and presentation within a team.

5.3.1 Foundations

The effectiveness of what-if analyses strongly depends on the solutions to two major challenges: (1) the identification of a point in time at which the movement outcome can be positively influenced and (2) the prediction ability for alternative movement behavior. Identifying a point in time where we can positively influence movement behavior is inherently difficult due to two aspects: the identified point of time to optimize the movement behavior is either too close to the begin of the observed period of time (i.e., there is not enough time left to optimize the movement) or the begin point of time is identified too far in the past and the predicted movement is not reliable anymore (i.e., all players have too much time to react resulting in a high variability). In between both temporal extremes is an intermediate sweet spot allowing enough time to optimize the movement outcome, and still do stable outcome prediction. This sweet spot is unfortunately not always the same, preventing methods for a fully automatic identification so far.



(a) Observed Movement **(b)** Wrongly suggested alteration

Figure 5.3.1: The automatic assessment of how a player should have moved to receive a better outcome of a situation is difficult. An alteration of real movement can even result in a worse situation, as e.g., shrinking the lower free space (a) by simply moving the red player **P** results in enlarging the upper free space (b). Please note that we reduced the number of players in this figure for legibility reasons.

In the following, we propose an approach to detect the most promising point in time as a starting point for an alternative movement prediction. This serves to improve the detected arguable movement behavior, additionally allowing the analyst to explore the temporal neighborhood as well. The prediction quality relies on the context inclusion in the movement prediction algorithm. Currently, there exist no way to validate the hypotheses of what-if scenarios based on real player movement data causing tedious and error-prone manual analyses. Experts have to rely on their own experience while looking for suboptimal movement behavior as well as alternative movement possibilities.

Computational assessment of alternative movement and potential outcome of a situation are challenging problems. An alternative location of players can easily result in a worse situation as depicted in Figure 5.3.1. Consequently, both domain experts and algorithms have to take the situation context into account, assessing and changing the positions of players. Therefore, a proper algorithmic solution needs to make sure that proposed local alterations do not result in a worse situation. Thus, we incorporate the collective movement of both teams into the assessment and the proposal of alternatives, e.g., by evaluating the creation and limitation of free spaces. Processing large amounts of player movement data, our method proposes how players could have altered their movement in respective situations based on real underlying data. Based on these considerations, and on requirements obtained by interviews with several coaches and analysts from international first league soccer clubs, we state the following research hypotheses:

1. The combination of automatic and interactive visual analysis is beneficial for the detection when and where suboptimal or faulty movement behavior occurs, incorporating all relevant data;

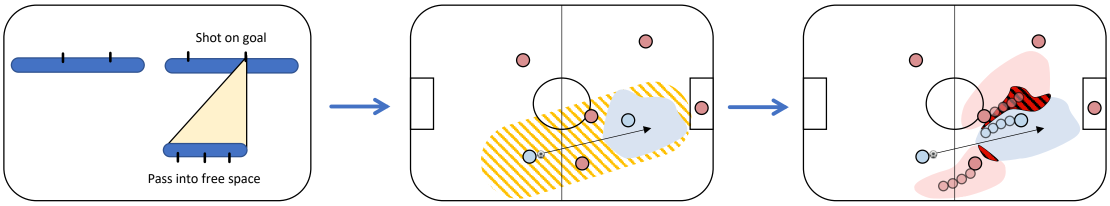
2. Faulty movement detection and generation of what-if scenarios can be algorithmically modeled and computed efficiently;

3. The synergetic automatic and visual-interactive approach effectively supports and improves the work of coaches and video analysts.

5.3.2 Procedures

We consider match situations—including significant or match deciding events involving faults—that lead to a major disadvantage of one team as basis to perform what-if analyses. For these situations, analysts are usually highly interested in **(a)** finding faulty movement behavior as well as **(b)** providing alternatives to the team in order to avoid these faults. While **(b)** reflects the what-if analyses, **(a)** is a necessary input to perform such analyses. Consequently, we consider the detection of faulty movement behavior as an essential first step for a successful what-if analysis. We focus in this work on defensive behavior only, as a good defense is crucial for a successful team. Ball possession is highly depending on a productive defending behavior and faulty positioning in defending phases can be fatal.

To support analysts in investigating matches, our designed technique has to be transparent, traceable, as well as adjustable to the needs of the analyst. Therefore, we rely on criteria describing soccer movement and high-level features such as interaction spaces and free spaces (Chapter 4.2). Furthermore, we present a transparent rule-based approach that produces explainable and interpretable results. This is an intentional design decision to avoid often used black-box techniques such as deep learning methods. While widely used black-box machine learning techniques can achieve remarkable results, they do so at the expense of explainability and interpretability. There are some first efforts to make such models comprehensible, but some important scientific questions remain unanswered, which can lead to practical, ethical and trust issues [GMR+19]. A video describing and showing our system in use can be found online (`https://files.dbvis.de/stein/SystemExamplesWhereToGo.mp4`).



(a) Detect time steps where a space gain happens

(b) Recognize opposing players who could be responsible for the large free space

(c) Identify loss of free space from the defending team

Figure 5.3.2: Detecting faulty movement behavior is non-trivial. For our approach, we first identify time steps when a pass is being played into a large free space. Afterwards, we identify opposing players who were close to the ball trajectory as possible responsible candidates for the control of this free space. These candidates are further analyzed by extracting the area lost to the attacking player. The larger the loss of free space, the more likely it is that this opposing player has moved incorrectly.

Detection of Faulty Movement Behavior

There exist a large variety of possible faulty movement behaviors ranging from lost header duels, to players positioning themselves wrongly on the pitch. However, as our data as previously described consists mainly of two-dimensional x- and y-coordinates, we restrict ourselves on the detection of movement- and region-based faulty behavior. Free space analysis has proven to be a powerful concept to estimate the regions that certain players or whole teams are potentially controlling. An overview of our resulting approach for the detection of faulty movement behavior can be seen in Figure 5.3.2. Our approach can either be applied to the whole match, or to previously chosen situations, e.g., by enabling analysts to define important events that have to be included such as shot on goal events. For each given interesting situation, we afterwards need to verify whether this situation could have been caused by faulty movement behavior. Consequently, we iterate backwards starting from the end of the situation, inspecting the collective movement of all individual players at each point in time. We calculate all free spaces for each time step in a predefined time period prior to the interesting situation (Figure 5.3.2 (a)), specifically looking at passes in large open free spaces. Consequently, we assess which opposing players lose the most free space to the player of the attacking team who receives the pass in the created free space (Figure 5.3.2 (b) and Figure 5.3.2 (c)). These defending players are our candidates for potential faulty movement behavior and also serve as the starting point for the what-if analysis. In the following, we describe the detection of faulty movement behavior in detail. All processing steps can be configured using several parameters if desired.

1. Detection of free spaces. (Figure 5.3.2 (a)) We calculate the free spaces of each player for each time step in a predefined time period before the occurrence of given interesting situations, in our case *shot on goal* events. The duration of the timespan can be set by the analyst, for example, the 30 seconds before a shot on goal event. Within this selected time period, we subsequently identify all passes into large free spaces as these passes are used to overcome opposing players.

2. Detection of involved players. (Figure 5.3.2 (b)) After the detection of free spaces, we identify the involved players which may have shown false movement behavior. To detect these players, we check which players are within a certain distance to the ball positions during the pass as displayed in Figure 5.3.2 (b). Realistically, only players near the ball would have been able to intercept the ball or reduce the free space of the attacking players. The hatched area in Figure 5.3.2 (b) represents the distance around the ball which is used to identify involved defending players. After discussion with experts, as well as after reviewing the first preliminary results, we experimentally set this value to 8.75m. Alterations of this value affect which players are considered for the identification of faulty movement behavior. Too high values include too many players, which may result in players being analyzed who did not participate in the respective game situation. On the other hand, if the value is set too low, important involved players may not be considered.

3. Identification of faulty movement behavior. (Figure 5.3.2 (c)) In the last step, we examine which players may have performed faulty movement behavior in the given match situation. Consequently, we take the players of the defending team that we identified in the previous step and calculate, for any time step between two seconds before a pass and its reception, how much free space each of them lost to the ball receiving player of the attacking team. The identified difference corresponds to the region the attacker gained control over by her or his movement behavior. Conversely, we recognize when a defending player potentially shows a faulty movement behavior based on a large loss of space. Additionally, as free spaces depend on all moving players on the field and can thus change quickly, we calculate the duration the defending player loses free space to the attacking player. Afterwards, the identified candidates for faulty movement behavior are assessed based on duration, area of free space lost, as well as distance to the time step where the pass was received by the attacking player.

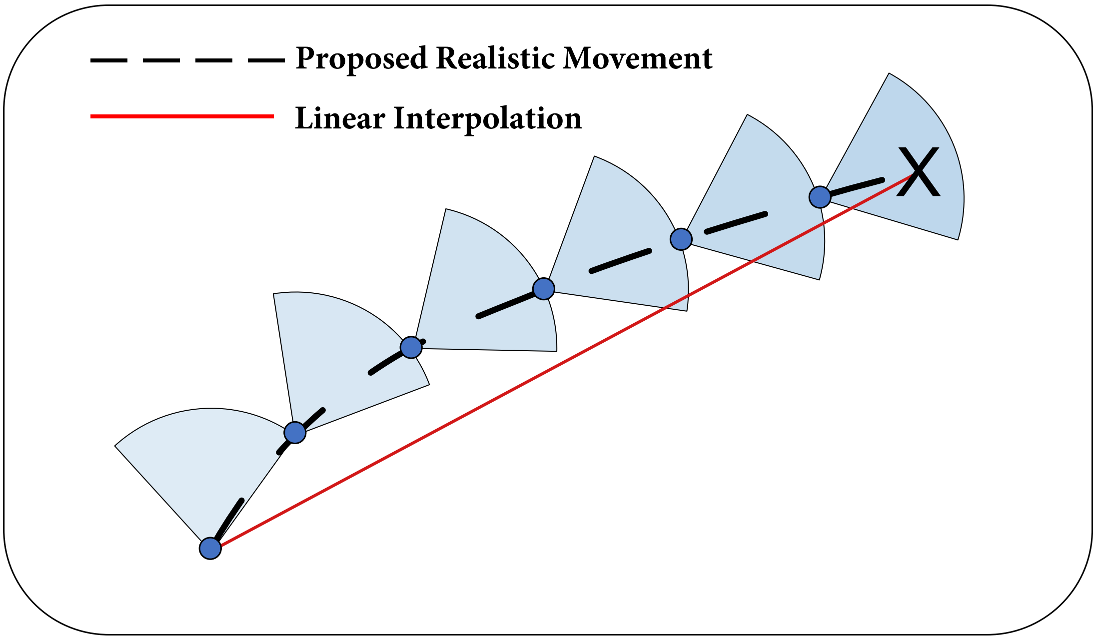


Figure 5.3.3: To enable domain experts to properly assess the proposed optimal position, we provide a realistic player trajectory. It is representing how the player could have moved to the identified optimal position based on the player's interaction space (dashed line), based on an appropriate movement and interaction model (Chapter 4.2).

Region-Based What-If Analyses

After identifying situations containing potentially faulty movement behavior, we focus on enabling an effective what-if analysis for these cases. It is important for coaches and analysts to recognize and present faulty movement behavior and movement alternatives to the team to improve team performance. Using our proposed what-if analysis, we show how a physically realistic position change of a player could have affected the current situation. Here, our goal is to maximize important regions which are under the control of a team, for example, near the own defenders by shifting the player who makes the mistake to a realistic and improved position. Exploring what-if scenarios, however, is very complex, as each change might have unforeseen implications. Simply moving players closer to the ball or into opponents' free spaces could result in an overall loss of the free space of their team which is potentially worse. Furthermore, a player should theoretically be able to reach a proposed position respecting physics. We consider this as an optimization problem assuming that players are constantly trying to reach their optimum location concerning all other player positions and respective goals.

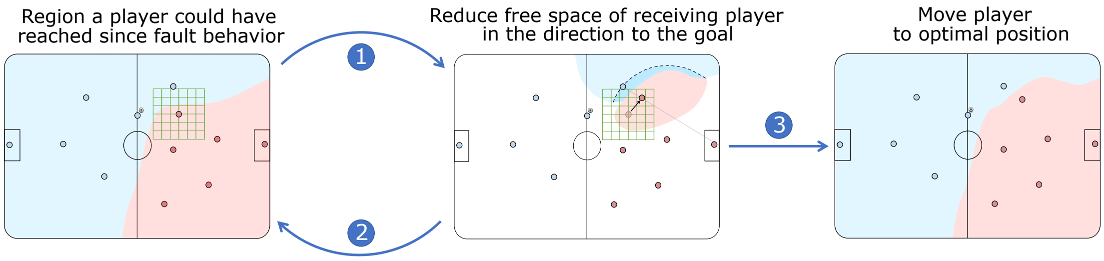


Figure 5.3.4: After we detected which defending players were causing a dangerous opposing free space by losing important room control, we calculate optimized player positions. We calculate the region each player could have reached since the initially faulty movement behavior (left). Afterwards, we asses each possible position determining the optimal player position based on how much the free space of the attacking opposing player can be reduced towards the own goal (middle). Eventually, each detected player is shifted towards the region with the best result (right). In this schematic picture, only six players of each team are depicted for visual clarity. Our algorithm considers all players from both teams.

We distinguish between an user-driven, interactive approach as well as an automatic approach, presenting previously unknown alternatives to perform what-if analyses. For the first approach, we allow the user to interactively inspect the implications of possible changes or alternative positions for the current situation by a *drag-and-drop* interaction with any player. Once a player has been selected by the user (mouse clicking), our system recalculates free spaces during every position change (mouse moving) and calculates optimized, realistic player trajectories for defending players on their path to the identified optimal position (Figure 5.3.3), thus, enabling experts to compare, confirm, or reject their hypotheses with respect to the displayed results. For our automatic approach, we calculate an improved position for a defending player with faulty movement behavior as displayed in Figure 5.3.4. Afterwards, our system proposes optimized, realistic player trajectories for defending players on their path to the identified optimal position (Figure 5.3.3). Eventually, a complete overview of the workflow of our implemented what-if analysis can be seen in Figure 5.3.5. Below, we describe every step of our automatic approach in detail.

1. Calculate improved player movement. The starting point of our automatic what-if analysis is the first time step of the identified faulty movement behavior, calculated as previously described in Section 5.3.2. We now define which region a defending player could have reached between faulty movement behavior and the pass into the arising free space by placing an invisible, dynamic grid with the defending player at the center (Figure 5.3.4 left) on the soccer pitch.

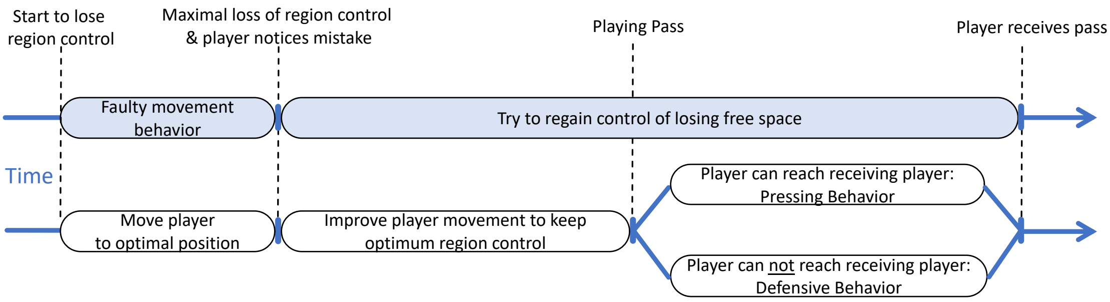


Figure 5.3.5: Every step of our workflow for the implemented what-if analysis. Our what-if analysis starts at the moment an identified defending player starts to lose room control. Afterwards, we constantly optimize the defending player's movement trajectory to reduce the free space of the ball receiving attacking player towards the own goal. Eventually, during the pass to the attacking player, our system either proposes pressing behavior or a more defensive behavior, based on whether the attacking player can be reached in time.

The side length of the individual cells is experimentally set to 0.5 meters, as this is approximately the space that one person occupies. The side length can be set differently, however, there is a trade-off between computational complexity and accuracy. For each cell, we calculate the free spaces if the defending player would have been at this position covering different speeds and running directions and, therefore, update the controlled regions of all players. We asses the quality of a position change by calculating how much the free space of the attacking player towards the goal is being reduced as depicted in the middle of Figure 5.3.4. After assessing each cell, the defending player is moved towards the identified optimal position corresponding to the position with the best rating (Figure 5.3.4 right).

To enable domain experts to properly assess the proposed optimal position, we also focus on providing a realistic player trajectory representing how the player could have moved to the identified optimal position. As players are bound by the laws of physics and, consequently, speed and direction influence their running behavior (see Interaction Spaces of Chapter 4.2), we provide an interpolation based on the possible movement of a player instead of, for example a simple linear interpolation between past and optimal player position. Figure 5.3.3 demonstrates how we calculate the movement to a determined optimal player position. The red line in Figure 5.3.3 shows a linear interpolation between player and calculated optimal position, proposing unrealistic movement as humans in the real world, bound to the laws of physics, cannot change their running direction immediately. To obtain a more realistic player trajectory, we introduce a movement calculation based on the player's interaction space. As soon as the identified defending player starts to lose room control towards the attacking opposing player (so right at the beginning of the faulty movement behavior), we visualize how the defending player moves to the optimal position by calculating the player's interaction space, searching for the nearest point between interaction space and optimal position. Next, we determine the new running direction of the player by calculating the angle between the current and the detected positions. These steps are repeated until the optimal position is reached. The detected player positions are represented as blue dots in Figure 5.3.3.

2. Calculate further movement behavior until the pass is played. So far, we proposed a realistic player trajectory aiming at the calculated optimal position which mostly reduces the free space and, therefore, the room control of the attacking player at the point in time when the originally detected faulty movement behavior occurred. Still, the situation might not have finished yet as the attacking team possibly did not yet try to pass in the, initially large, free space of the attacking player. To demonstrate how the defending player could have stayed in control of the important space between attacking player and own goal until the pass is played, we again move the player to the optimal position at the time of the pass, calculated as previously described.

3. Which behavior should the moved player show during the pass? The last step is to determine the movement behavior of the defending player, right at the moment of the pass towards the attacking player. The defending player can either stay defensive, or try to attack the opposing player. Increasing the pressure on the attacking, ball receiving player can result in a higher probability of gaining possession of the ball. A more defensive behavior allows to keep room control towards the own goal and might be useful, e.g., when a technically high-skilled attacking player is currently in ball possession. To determine whether the system should suggest pressing or a more defensive play style, we calculate whether the defending player could reach the attacking player, at least at the moment of ball reception. If this is the case, our system proposes the defending player to put pressure on the attacking, ball receiving player. Otherwise, our system proposes a more defensive behavior.

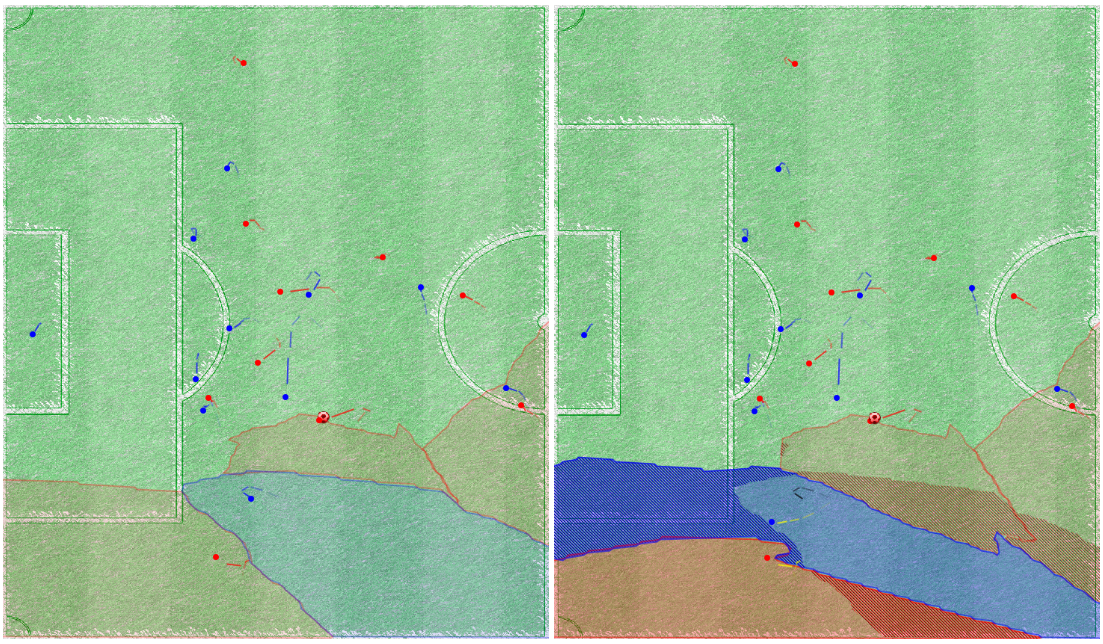
3a. Pressing Behavior. A defending player is putting pressure on an attacking opposing player if the defending player is moving towards the attacking player to disturb or stop the ball reception. In our system, we let the defending player run towards the attacking player who is about to receive the ball. We reuse the previously introduced calculation from Figure 5.3.3 to compute a realistic single player movement. Additionally, we pay special attention to avoid players running directly onto the other player, but instead moving slightly in front of the attacking player, constantly blocking the angle and path to the own goal.

3b. Defensive Behavior. If the defending player is expected to show a more defensive behavior, our system keeps the defending player moving on an optimal path which is recalculated and updated every time step. Consequently, the attacking player's free space and movement is further being reduced. We realize this by reusing our previously introduced approach from Figure 5.3.4 with the difference that we constantly use the spatial coordinates of the attacking player who is about to receive the ball instead of the ball.

5.3.3 Evaluation

For our evaluation, we separately invited three of our experts (see Chapter 1.4 for the expert introduction) to use our system for the detection of faulty movement behavior as well as our proposed resulting region-based what-if analyses. All capabilities of our proposed system were explained to each expert. Afterwards, we loaded our proposed approach using a match from a European international soccer club competition. For the given match, our system proposed 14 situations, where an assumed faulty movement behavior lead to a successful shot on goal event. For each assumed faulty movement behavior, we asked each expert separately to verify whether the classification as faulty movement behavior was correct as well as whether our automatically what-if analysis proposed an improved (ideal) solution. Every expert was allowed to spend as much time as desired with each identified faulty movement behavior as well as express ad-hoc comments. Afterwards, we performed an interview with each expert in order to gather feedback about the usefulness, completeness as well as applicability of our implementation of the **alteration** step of our conceptual workflow.

The results of both expert studies are very promising. The experts identified faulty movement behavior in every proposed situation. An exemplary situation the experts approved can be seen in Figure 5.3.6. Furthermore, all three domain experts agreed with our suggested player movement in 35 out of 42 cases (83.33 %). Interestingly, the experts did not always agree among each other when manually classifying the suggested player movement (5 out of 14 situations). We believe this level of disagreement is to be expected given individual experience and coaching strategies of involved experts. Overall, we rate it as a promising result. Table 5.3.1 shows the



(a) Observed Movement **(b)** Suggested player movement

Figure 5.3.6: Showing an attacking situation of the red team from a professional match of an European international soccer club competition. The free spaces of both teams are colored in the respective team color. Our system identified faulty movement behavior of the bottom defending player of the blue team (a), directly before the pass towards the bottom attacking player of the red team. Our system proposes an improved movement trajectory for the defending player (b) which would have allowed him to keep control of the important free space towards his own goal. Each hatched area indicates free spaces that a player is now controlling after the improved movement.

experts' assessment for each of the single situation. The experts highly appreciated that the system detects possible faulty movement behavior and suggests improved player movement. The focus on defensive behavior was seen appropriate as the analysis of defensive behavior takes an important role during match preparation. Overall, each evaluation including interviews and discussion lasted around two to three hours. According to our experts, the suggestions are very useful as starting point while they approved the possibility to integrate their expert knowledge interactively in the system by moving the players around via *drag-and-drop*. As our proposed system detects region-based faulty movement behavior, all experts see great potential for an improved training process, to increase a team's performance when directly working with the identified mistakes. One expert especially mentioned the *intuitive setup* (exact quote) of our

Scene	Expert C	Expert A	Expert D
1 (Shot 1, 2 Passes Before Shot)	✓	✓	✓
2 (Shot 1, 3 Passes Before Shot)	✓	✓	✗
3 (Shot 5, 1 Pass Before Shot)	✓	✓	✓
4 (Shot 6, 1 Pass Before Shot)	✓	✓	✓
5 (Shot 8, 1 Pass Before Shot)	✓	✗	✗
6 (Shot 9, 5 Passes Before Shot)	✓	✓	✓
7 (Shot 9, 6 Passes Before Shot)	✓	✓	✓
8 (Shot 9, 8 Passes Before Shot)	✓	✓	✗
9 (Shot 11, 6 Passes Before Shot)	✗	✓	✓
10 (Shot 11, 7 Passes Before Shot)	✓	✓	✓
11 (Shot 13, 2 Passes Before Shot)	✗	✓	✗
12 (Shot 20, 3 Passes Before Shot)	✓	✓	✓
13 (Shot 22, 1 Pass Before Shot)	✓	✓	✓
14 (Shot 22, 3 Passes Before Shot)	✓	✓	✓
Accuracy	**85.71 %**	**92.85 %**	**71.42 %**

Table 5.3.1: We invited three domain experts in order to test our proposed what-if analysis with real match data from an European international soccer club competition. Our system identified 14 situations, where an assumed faulty movement behavior lead to a successful shot on goal event. All experts agreed with the 14 situations being classified as faulty movement behavior. As further shown in the table, the invited experts mainly agreed with our following suggested improved player movement.

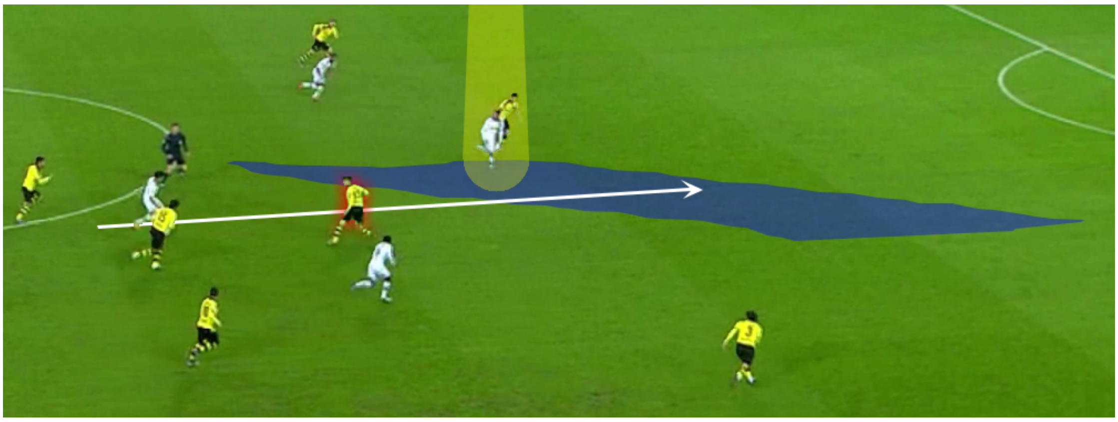
visual storytelling as being very beneficial for coaches and analysts. The other experts agree and added that they consider the way we communicate which player could have altered his or her movement is clearly understandable. Besides, the experts mentioned the big potential of using our approach to aggregate identified faulty movement behavior over several matches, for example, to detect players that are prone to certain mistakes.

5.4 Discussion and Conclusion

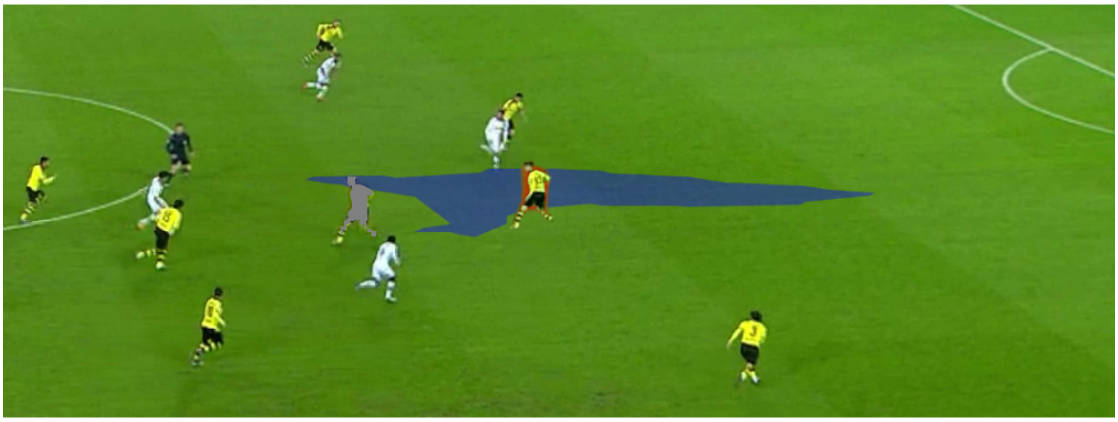
From both, a media production and a soccer analysis perspective, the presented system provides a significant step towards automated storytelling in sports analysis. Based on discussions with invited domain experts, we integrated a number of methods to detect relevant soccer event types and determine corresponding visualizations that can represent these by overlaying visualization on the video. The detection of situations and types of situations is a difficult problem, as it has to deal with segmentation in time and space. Main problems for a realized system include the *completeness* and the *precision* of the detection. In our proof of concept, we implemented a number of detection methods representing a set of important, frequently occurring events in soccer matches. Since the alteration is a complex optimization problem, we presented an approach for effective what-if analyses in soccer covering (1) the detection of faulty movement in space and time as described in Section 5.3.2, (2) the computation and prediction of realistic movement alternatives proposed in Section 5.3.2 as well as (3) the improvement of the everyday work of coaches and video analysts by realistic what-if analyses evaluated in Section 5.3.3.

We proofed the effectiveness of our detection of relevant scenes as well as suggested player repositioning. Our approach is based on a realistic movement model, and our prediction is easily explainable and hence, expected to show acceptance by analysts and players. We performed our what-if analysis on our described data sets where the initial detection of faulty movement behavior as well as the suggestion of improved player movement averaged on 27 minutes per match. Once calculated, loading every detected scene in a match takes less than a second. Furthermore, our system is expected to significantly accelerate the analysis process by automatically identifying relevant situations instead of a tedious manual analysis of every scene. Thanks to the techniques presented in Chapter 3 and Chapter 4.3, our approach can also be used interactively within original video recordings. An example showing the inline video editing revealing faulty movement and proposed repositionings can be seen in Figure 5.4.1.

Gathered expert feedback showed that experts were convinced by the applicability and effectiveness of our conceptual workflow. The possibility to select one particular story out of a few valuable options including the automatic visualization elements and explanations can simplify the in-game analysis and reduce turn-around time to a few seconds. With that, mid-game analysis of journalists can be enriched with interesting and captivating content. The audio commentary as an additional communication channel between system and expert enhances the understanding being crucial because of the editorial content of the created story. Our presented concept is seen applicable to other team sports but needs manual parameter settings, which can be set semi-automatically. Ultimately, this chapter serves as the foundation towards a better transformation from gameplay to strategic improvements via visual predictive analytics. We envision coaches being driven by data-based decisions in their training and tactics supported by visual data analytics. Consequently, the discussions on team performance will be based on data insights and not on pure intuitions.



(a) Observed Movement



(b) Suggested player position

Figure 5.4.1: Example outcome of our proposed alteration technique used in a real video recording. In this situation, a player (highlighted in red color) noticed his wrong location too late. If the player had positioned himself better, the free space of the opposing player would be much less severe.

Finding the right approach to successfully use Big Data will be of central importance for future success.

Daniel Memmert (German Sport University Cologne)

6

Evaluation on Video-Based Analysis of Soccer Matches

Contents

6.1 Introduction

In this chapter, we assess the combined research in this dissertation in comparison to scientifically as well as commercially available tools, considering that most scientific approaches are not yet used in commercially available tools. We provide a comprehensive and categorized overview of the latest, non-trivial methods developed for video-based visualization of soccer matches. We, specifically, do *not* survey solutions which only provide tagging and enhanced video libraries like Sportscode or Dartfish, as they do not employ video-based visualizations yet. However, plans in this regard are announced for the latter. In general, no current survey paper exists that describes the narrow field of video-based visualization of soccer matches completely. Existing surveys are either outdated [BCD+12] or incomplete [PVS+18] as they describe the whole field of sports data visualization. Other surveys [RS14, Shi18, TGM+17] can only be considered partial reviews (outdated, commercial solutions only, ...) or deal with a specific sub-issue. We survey the existing literature including outdated part-surveys (Section 6.2) and, furthermore, differentiate described video-based approaches from other visualization and analysis methods (Section 6.3). The existing approaches found in the literature are each presented shortly before we compare them according to eleven identified relevant features. We discuss differences and commonalities in detail, with a qualitative overview given in Table 6.3.1 for reference (both Section 6.3). Further, we identify key issues and unsolved research challenges in the field and discuss research opportunities, before we provide a short summary of our findings (Section 6.4).

6.2 Related Work

While a high volume of literature on visualization techniques exists [Kei01, Kei02, HH09, dOL03], which are extensively used in fields such as physics, the research is decidedly more sparse on the topic of video-based visualizations and in particular on video-based approaches for soccer. The recent work by Perin et al. [PVS+18] provides a quick overview of visualizations developed for sports data, although few techniques can be regarded as video-based, in particular for soccer. For those video-based approaches, an older survey by Borgo et al. [BCD+12] provides a general overview up to the year 2012, while developing useful classification schemes.

However, with the fast pace of technological development, the referenced approaches are mostly outdated. Other surveys [RS14, Shi18, TGM+17] can only be considered partial reviews or deal with a specific sub-issue. They are either outdated, not complete, or consider only commercial approaches [Tho07]. Few solutions have been proposed in the field of video-based soccer visualization. The first relevant work can be attributed to Assfalg et al. [ABC+03] as early as 2003. It was one of the first to annotate soccer matches with calculated information, extracted from the video itself, to support the understanding of what is shown on screen. In the paper, ways are discussed to annotate videos with simple match scene information.

One of the most important contributors to the field, especially in the 2000s but also, later on, was Graham Thomas. His research as part of BBC R&D (British Broadcasting Company, Research & Development) resulted in several commercial projects [BBC05, BBC12, BE04], which pioneered novel solutions for sports visualization, including the Piero system [Red18], one of most widely used sports visualization and annotation system today and the main competitor to the Viz Libero system [Viz18]. Both systems are developed for sports broadcasting applications and provide a visualization and special effects toolbox, tailored to supported games like soccer. Both try to automate some aspects like tracking to reduce user workload.

In general, as detailed by Thomas et al. [TGM+17], large commercial applications like Augmented Video Player [BBC12], True View [Int18], Piero [Red18], and Viz Libero [Viz18] for video-based soccer visualization are actually ahead of *most* of the academic research and represent the state of the art, with only a few notable advanced academic exceptions from the methods proposed by this dissertation. Apart from this autonomously working solution, few other approaches with the same level of sophistication exist. The work by Schlipsing [SSTI14] looked initially promising, but was abandoned after the principal investigator completed his Ph.D. Some other approaches exist [WWXT04, LJHX08, XSL+17, AWKG05]. Nevertheless, these latter approaches remain at basic levels of visualization and surprisingly few new ideas were proposed. As described by Thomas et al. [Tho07], some individual and specific contributions have been made to the field of sports data analysis and visualization, however *whole systems which allow for automated tracking or labeling remain an open challenge* (exact quotation, cf. [Tho07]). Therefore, we provide an overview of the existing literature which answers the following research question:

> *Which techniques exist and are currently employed in video-based analysis of soccer matches?*

6.3 Survey

Papers are categorized into three groups: those that present a variety of approaches (mostly survey papers), those who present a single, new idea and those approaches which are described academically only rudimentary, as they are commercially available. For most cases, the approaches were either tested in reality (commercial solutions when developed and deployed) or, in case of academic research, have been evaluated as part of an expert study, to asses the practical significance of the new visualization. Interestingly, for most technical papers, the video-based visualization is only a by-product, with a larger part dedicated to analysis aspects. Therefore, less than a handful of papers describe in more detail the actual techniques employed and rationale behind the presented visualizations. Comparison between different video-based approaches, specifically for soccer, is virtually non-existent.

In the following Section 6.3.1, we give an overview of standard approaches and the most promising novel techniques proposed in literature and industry. In Section 6.3.2, we propose classification methods and different criteria to assess the presented papers. We follow up with a detailed comparison of the presented approaches for each feature individually. A qualitative overview of the most important features of the approaches we consider most relevant is also given in Table 6.3.1 for reference, which is presented after the approaches and comparison criteria.

6.3.1 Overview of current approaches

In the following, we present the individual approaches. The order in which the approaches are presented in is related to their *complexity*. However, the details of this categorization can be safely ignored for now and the concept is introduced later, when relevant for the comparison in Section 6.3.2.

Basic: Textual/Numeric Table Overlay. The most simple form of video-based data visualization is at the same time one of the oldest. Here, the information, which is typically text or numbers, is displayed as an overlay on top of the actual video, and shown for several seconds for it to be perceived. The only relation to the actual video is the timing, which can - but does not have to - match a current match situation. Information might include the current number of goals and some match statistics. An example of such a table overlay with some match statistics like goals, offsides, ball possession rate, etc. is shown in Figure 6.3.1.

Figure 6.3.1: A table overlay showing current match statistics, while partly occluding the gameplay. The metrics display the relative strength of both teams and aim to explain the goal scored some seconds ago (Source: Simple Thought Productions).

Automatic Highlight Identification. The approach by Assfalg et al. [ABC+03] aims at automatically extracting highlights from a soccer game by analyzing corresponding video sequences. First, playfield regions are extracted by their shape for localization. Then, player bounding boxes are found by color differentiating and adaptive template matching. To discern between different scenarios, the most likely scenario is chosen by using the normalized player position density distributions of example scenes. Then, a domain knowledge model computes the currently most likely scenario based on past predictions. The identified scenario is then reintegrated in the video in a bottom bar as event boxes for specific events, labeled as *none*, *hypothesis*, *accepted*, and *rejected*. While the detection technique was state of the art at the time of publication and the accuracy was good, the visualization was simple and did not evolve much.

Highlights Extraction with Content Augmentation. The idea by Want et al. [WWXT04] differs somewhat from the previous aim by Assfalg et al. [ABC+03]. Here, the first goal is to separate the video into different segments, using audio and visual cues, but only focusing on detecting replays. The second, more relevant goal is to identify regions for content augmentation. This is done by first determining the temporal and spatial relevance of the scene, as an insertion in a highly relevant scene would be disturbing. A suitable insertion area is then found by color quantization of homogeneous regions. No specific limits are set to the content to be inserted. However, as the position is in principle arbitrary and not related to any content, the most likely use is for non-obtrusive advertisement or other non-positional information.

Region-based Tracking and Provision of Augmented Information. The approach by Andrade et al. [AWKG05] differs from others by being primarily about extracting and tracking objects based on regions instead of single pixels. This region-based approach allows for a hierarchy of regions, represented as a graph structure. In turn, searching and tracking between different scenes are made more robust. Detected regions and features can the be augmented inside the video, by user-supplied graphics, although it should be noted that the primary focus of the work lies on the region-based tracking and the detection of semantic objects.

Virtual Content Insertion System. The work by Liu et al. [LJHX08] is very similar to the second part of the one by Wan et al. [WWXT04]. It proposes a generic virtual content insertion system, which places content inside a video when the underlying actions attract much attention, but at a position where it does not obstruct the storyline. An adequate time is found by analyzing the spatial and temporal attention and the position using the three measures of saliency, contrast, and novelty. The insertion supports affine transformations to align the content with the actual environment.

Automatic Video Annotation System. The system presented by Xue et al. [XSL+17] aims at extracting metadata from video sequences of sports events. Sequence boundaries are segmented and clustered according to the type of content. Players cannot be tracked automatically but must be laboriously selected manually for tracking. On-screen text is recognized using trained SVM's (support vector machines) and read using OCR (optical character recognition). The combined information is presented as a simple on-screen graphic.

Video-based Activity Recording. In his thesis, Schlipsing [SSTI14] presents a complete recording (calibrated dual-camera setup) and analysis system workflow for sports data analysis. Players are tracked semi-automatically using segmentation. Initial assignments and re-identifications need to be made manually. In terms of in-video visualization, player positions,

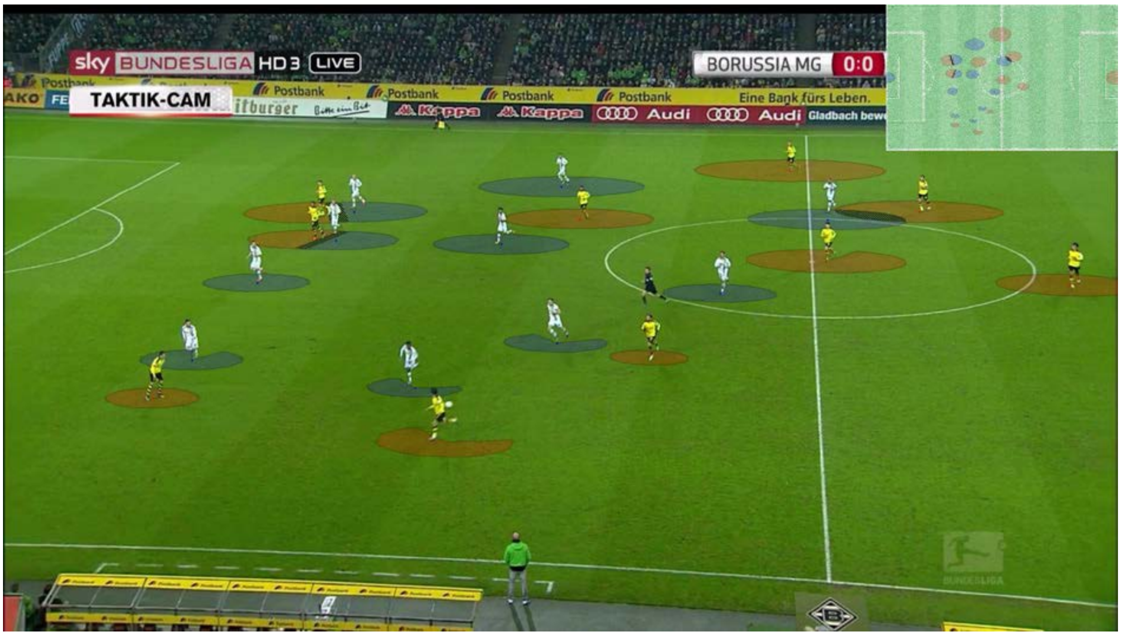


Figure 6.3.2: Interaction space visualization by this dissertation (Chapter 4.3). The augmented circles show the personal space that can easily be reached by the players and are helpful, for example, to identify suitable receivers for a pass.

and numbers, player formations (between manually selected players), and compactness hulls are presented. It was one of the first approaches which featured the placement of virtual objects in a soccer video for augmentation purposes.

Combining Video and Movement Data to Enhance Team Sport Analysis. The approach presented by this dissertation is somewhat similar to the main idea followed by Schlipsing [SSTI14], but differs in two fundamental ways in the implementation: First, it aims to use already existing video recordings - like a scouting feed or simple television recordings capturing all moving players - as a video source, instead of a carefully calibrated dual-camera setup. Secondly, the tracking of the players aims to be autonomous and not require manual input. Then, in terms of visualization, the approach has a similar idea to enhance the gameplay and analysis by presenting visualizations in an augmented video. Here, several new ideas of features for visualization are presented such as interaction spaces (the personal space that can be controlled by each player), free spaces, dominant regions, pass distances per player, and player reactions. An example visualization is shown in Figure 6.3.2. Building on this foundation, this dissertation proposes a system for automatic annotation as well as selection of helpful visualizations in key match situations. As part of this, further visualization ideas are presented or improved upon such as the pressure imposed upon players to act (aggressiveness of the defending team to recapture the ball).

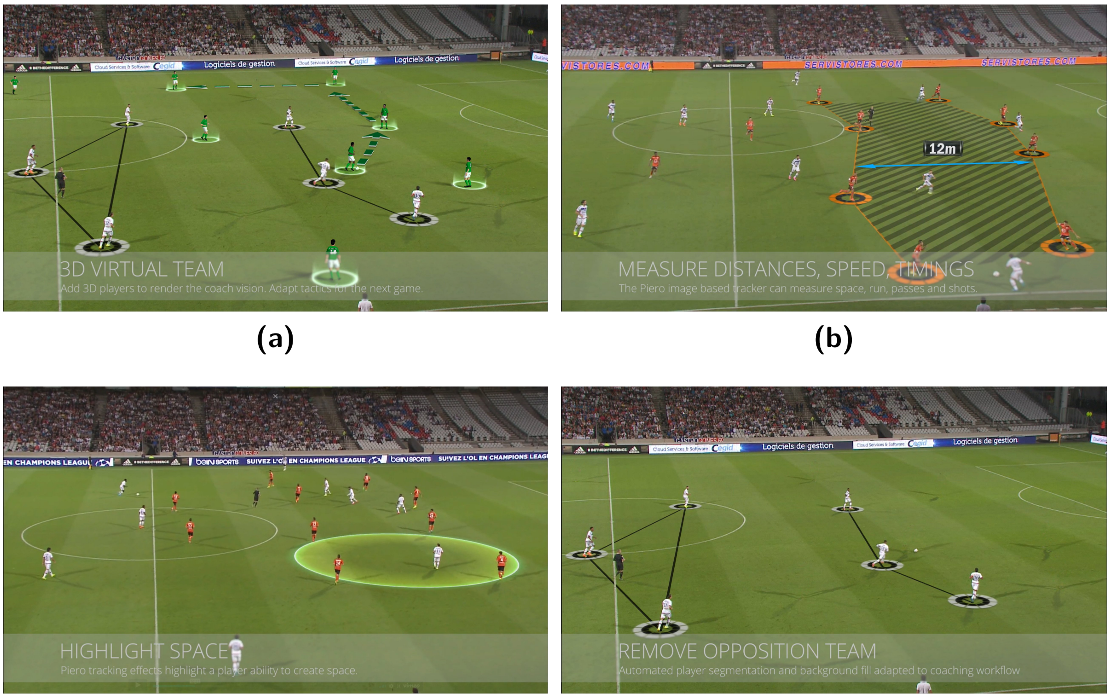
The previously discussed approaches represent the state of the art in the academic literature. In the following, we focus on commercial products that are deployed in the field. According to Thomas et al. [TGM+17] and Hilton et al. [HGK+11], the most widely used visualization products in the industry are *Piero* [Red18], developed by the BBC R&D in collaboration with Ericsson, and *True View* [Int18] from Intel, former FreeD by Replay Technologies. Additionally, *Viz Libero* [Viz18] is regarded as the third large broadcasting solution. An overview of some of the systems is also given by Hilton et al. [HGK+11] and, more up to date, by Thomas et al. [TGM+17]. The commercial solutions are:

Commercial: iView (2005-2008). Developed by the BBC between 2005 and 2008, the iView: free-viewpoint video system [BBC05] allows to reconstruct a 3D scene from the 2D video sequences and allows free positioning in them. In theory, arbitrary objects could be placed in the resulting 3D scene, but this was rarely done and the quality, as well as the resolution of the model, was very basic. The system was superseded by Intel True View (see below).

Commercial: Augmented Video Player (2012-today). Also developed by the BBC since 2012, the Augmented Video Player [BBC12], allows a user to watch sports sequences and interactively overlay additional layers to augment the video with further information.

Commercial: Intel True View (2012-today). The Intel True view system [Int18] is similar to the iView system and can be considered its technological successor. Formerly developed by FreeD Replay Technologies, it enables to reconstruct high-resolution 3D scenes using several HD cameras from different angles to show matches as a 3D scene. The quality is markedly better and some simple (manually annotated) visualizations are used, like trajectories or highlighting.

Commercial: Piero (2004-today). The Piero system [BE04, Red18] is considered one of the most advanced and widely used systems on the market. It was pioneered by the BBC, then outsourced to Redbeemedia in collaboration with Ericsson. It is used in different sports, including



(a) (b) (c) (d)

Figure 6.3.3: Four different types of visualization which are available in the Piero system [Red18].

soccer, to augment the video with additional information. This aspect makes it come very close to the actual core of this survey. Tracking of players is done semi-automatically (initial assignment and re-identification are done manually) and several automatisms exist. However, the core of the systems remains a video-effects software, specifically tailored to real-time analytics of sports games. For soccer, several manually placeable visualizations are available: They can be used, for example, for formations, free spaces, interaction spaces, distance measurements, ball trajectories, removable players, goal area heatmaps, player trajectories, density distributions, and others. Some of these examples are demonstrated in Figure 6.3.3. However, there is still a lot of manual work involved in creating animations and overlays. The main difference between the Piero system to the approach presented within this dissertation is that the latter is a fully automated approach while the former represents the state of the art in manually annotated video-based soccer analysis with an advanced graphical display. The visualizations presented in this dissertation can also be considered advanced, but, more importantly, are based on mathematical and physiological considerations. However, sometimes they lack the clean and

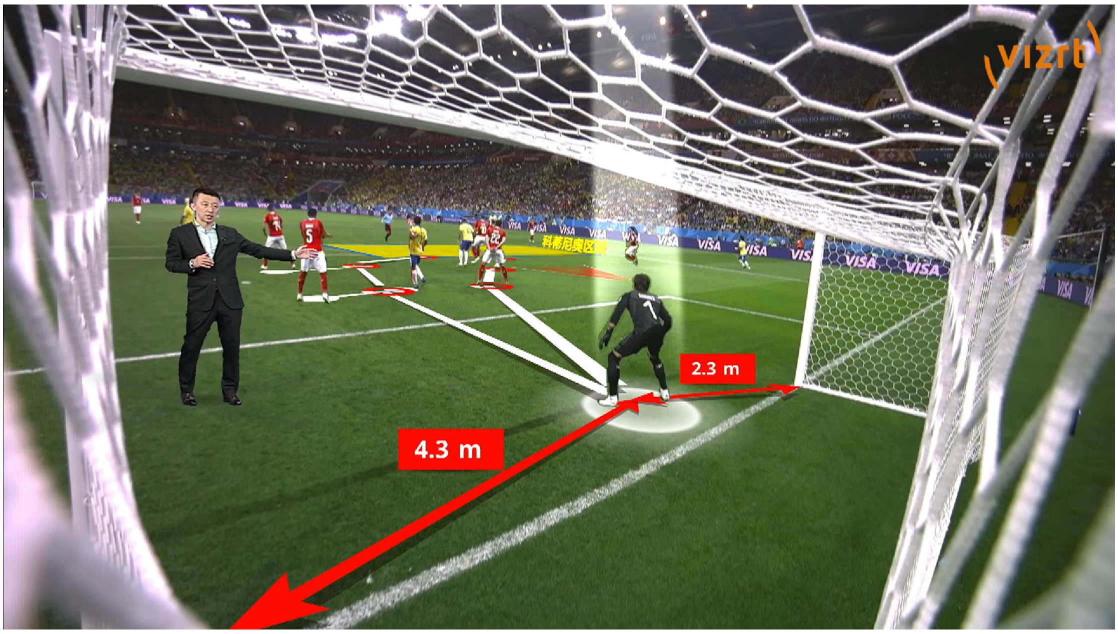


Figure 6.3.4: Annotated game scene with added distance measurements, highlighting and perspectively correct content placement by Viz Libero [Viz18].

typographically nice style. Combining those two approaches would likely lead to enormous benefits.

Commercial: Viz Libero (2010-today). The Viz Libero system [Viz18] is considered, together with the previously presented Piero system, the most advanced and widely used system on the market. It is also used for different sports, including soccer and augments the video with additional information. Available broadcast streams or video recordings are used as input streams, which means that no dedicated camera setup is required. Tracking is also semi-automatic and again several automatisms exist. The available visualizations are similar to the ones offered by Piero. One striking difference between Piero and Viz Libero is that the latter uses homography estimations and therefore allows limited reconstruction from single video sources. This means the camera viewpoint can be freely (within limits) be re-positioned within the scene to get the view from another perspective. One example of an augmented scene in Viz Libero is shown in Figure 6.3.4.

6.3.2 Categorization and Comparison

We first propose possible classifications and criteria to compare the different approaches. We establish the following nomenclature to classify the techniques into different categories. First, we differentiate between ***scientific research*** and ***commercial applications***. For the latter, it is often the case that no paper exists that describes the system in detail, so one is left with product descriptions of the producer, manuals, as well as demo videos. Secondly, we differentiate between the ***complexity*** of the approaches, ranging from ***low*** over ***medium*** to ***high***. A special case is a simple data table overlay, which we assign to the class ***very low***, and is provided for reference as the simplest technique available. To judge the amount of manual work required or how far an approach can work autonomous, we classify the proposals according to their ***automation level***, also ranging from ***(very) low*** over ***medium*** to ***high***. In addition, we distinguish between the approaches according to several key concepts: Are they ***insertion*** or ***reconstruction*** techniques? Is augmented information contextually ***embedded*** in the actual scene or merely overlayed and positioned? In this context, we further differentiate between static and support for ***dynamic*** content as well as the support to ***interact*** with the visualization, for example, by clicking or settings different parameters. Further options for comparison are if the content is ***placed*** at fixed positions or positioned intelligently and if the system can in principle work in ***real time*** or not. To assess the impact of the research we further differentiate the papers and approaches by the presence or absence of a usability or case ***study***, either in the paper itself, as a follow-up paper, or in form of industrial applications. Lastly, it is relevant to note if the technique is still actively researched or even ***used*** in the field. A qualitative comparison of the important features from the presented approaches is given in Table 6.3.1. In this table, the individual features are grouped semantically, serving as an overview of the main findings of this survey. Sometimes the labels given in the table are over-simplified and do not give the whole aspect credit. The full evaluation is given in text-form in the following sections for each feature individually.

Complexity. The most basic approach, with very low complexity, is just displaying a data table (or other types of simple visualization without location information) as an overlay on top of the video during the relevant period. The approaches with low complexity are all insertion techniques which employ non-sophisticated positioning inside the video and, with the exception of Liu et al. [LJHX08], all operate in the screen space instead of the world space (embedding techniques). Several approaches of Assfalg et al. [ABC+03], BBC R&D (Augmented Video

	1960+	2003	2004	2005	2005	2008	2012	2014	2017	now	now	now	now
	Table Overlay	Assfalg [ABC+03]	Wan [WWXT04]	Andrade [AWKG05]	iView* [BBC05]	Liu [LJHX08]	AVP*,3 [BBC12]	Schlipsing [SSTI14]	Xue [XSL+17]	True View* [Int18]	Piero* [BE04, Red18]	Viz Libero* [Viz18]	This Dissertation
Complexity[1]	very low	low	low	low	medium	low	low	medium	low	high	high	high	high
Insertion[2]	partly	present	present	present	not present	present	present	present	present	not present	present	present	present
Embedded[2]	not present	not present	not present	not present	not present	present	not present	present	not present	not present	present	present	present
Reconstruction[2]	not present	not present	not present	not present	present	not present	not present	not present	not present	present	not present	partly	not present
Dynamic[2]	not present	not present	present	not present	not present	not present	present	not present	not present	not present	present	present	present
Interactive[2]	not present	not present	not present	not present	present	not present	present	not present	not present	present	partly	present	present
Auto-Placement[2]	not present	not present	present	present	present	present	not present	present	not present	present	partly	partly	present
Real Time[2]	not present	not present	not present	not present	partly	not present	not present	not present	not present	partly	partly	partly	present
User Study[2]	not present	not present	partly	not present	not present	present	not present	present	not present	partly	partly	partly	present
Active/Used[2]	present	not present	not present	not present	not present	not present	not present	not present	not present	present	present	present	present
Automation[1]	very low	low	low	low	low	low	low	medium	low	medium	medium	medium	high

[1] very low, low, medium, high
[2] not present, partly, present
[3] Augmented Video Player
* Commercial applications

Table 6.3.1: Comparison overview of the visualization techniques for video-based soccer match analysis we reviewed. Interesting to note is the low amount of current academic research and the relatively large amount of commercial solutions. For a detailed discussion, see Section 6.3.2.

Player) [BBC12], Andrade et al. [AWKG05], Wan et al. [WWXT04], Liu et al. [LJHX08] and Xue et al. [XSL+17] belong in this category. More sophisticated ideas with medium complexity are introduced by BBC R&D (iView) [BBC05] and Schlipsing et al. [SSTI14]. The information is visualized as part of the scene itself. The approaches Piero [BE04, Red18], Viz Libero [Viz18], True View [Int18], as well as the ones from this dissertation with high complexity are those which employ advanced, embedded visualizations or reconstruction techniques, which require significant work to develop.

Automation. The automation level indicates how much user input is necessary for the analysis and to produce the visualization. Again, the simple approach of just displaying a data table contains *very low automation*, as variables like placement, size, and duration all have to be determined manually. The vast amount of approaches like Assfalg et al. [ABC+03], Wan et al. [WWXT04], Andrade et al. [AWKG05], iView [BBC05], Liu et al. [LJHX08], Augmented Video Player [BBC12], Xue et al. [XSL+17] offer *low automation* and often several manual steps are necessary. The approach by Schlipsing et al. [SSTI14] offers a higher degree of automatic visualizations, leading to *medium automaton*. However, the initial selection of players for tracking and then the re-identification is still solved manually. The same is true for the systems True View [Int18], Piero [Red18], and Viz Libero [Viz18]. These approaches all offer automatic support to reduce repeating and laborious steps. Nevertheless, the type of visualization and initial positions have to be specified manually. More often than not, even the actual content of the visualization is created manually by combining different building blocks. A *high automation* level is only reached by the systems presented in this dissertation. The visualizations are created automatically and without any user input, when using the default parameters.

Insertion. Techniques are considered insertion techniques when content information is placed inside the video, during a specific time and at a specific location. Most of the presented approaches belong to this category. The exceptions are iView [BBC05] and True View [Int18], which belong to the reconstructive techniques.

Embedded. This subcategory of insertion techniques contains approaches which actually embed the content into the 3D scene and respect relevant aspects like camera distortion and perspective projections. Approximately half of the approaches can be considered of this type, namely the ones by Liu et al. [LJHX08], Schlipsing et al. [SSTI14], the ones from this dissertation, Piero [Red18], and Viz Libero [Viz18]. The four latter ones are the most sophisticated, leading to aesthetically pleasing results.

Reconstruction. Reconstructive techniques are not (primarily) based on inserting content into the video to help with the analysis, but reconstruct the scene as a 3D object to allow for arbitrary positioning. This can support match analysis by enabling different views. While the field of 3D reconstruction and structure from motion, part of image analysis and computer graphics, is relatively large, it has not primarily been applied to soccer, due to the unique challenges with tiny, moving targets. The exceptions are the commercial systems iView [BBC05] and True View [Int18], which follow this reconstructive approach. True View is the more advanced.

Dynamic Content. Depending on the type of content that can be inserted, the techniques can be classified to only support static, fixed size content like text and images, or they support dynamic content, which can be animated or morphed. From the presented approaches, only Wan et al. [WWXT04], the approaches of this dissertation, Augmented Video Player [BBC12], Piero [Red18], and Viz Libero [Viz18] support this.

Interactive. Most approaches output a generated augmented video file, which can not be interacted with. We consider a system to be interactive when, during playback, different parameters to modify the visualization or actually interact with it can be set by keyboard or mouse/touch. Of the presented techniques, only the reconstructive ones (iView [BBC05], True View [Int18]), and Augmented Video Player [BBC12] and the approach by this dissertation are truly interactive. Viz Libero [Viz18] can be considered to be mostly interactive and, to a lesser degree, this is also true for Piero [Red18]. However, it should be said, that the boundaries between static and interactive are not strict and overlapping can exist. Also, some of the techniques can be adapted to provide interactivity.

Auto-Placement. The aspect of auto-placement is related to embedding, as it refers to the ability to place the content intelligently at an appropriate position. While this is a requirement for embedding, for overlaying techniques this is not strictly necessary, albeit often desired. Several of the presented systems are capable of automatically placing the content at desired position, also different levels of accuracy and difficulty exist. While this is one of the main aspects for reconstructive techniques (merge different images for a 3D reconstruction), the approaches by Wan et al. [WWXT04], Andrade et al. [AWKG05], and Liu et al. [LJHX08] follow a simple positioning and only the approaches by Schlipsing et al. [SSTI14] and the approaches from this dissertation are more advanced, with the latter doing a better job at the embedding.

Real Time. Of the presented approaches, only a few support real-time analysis and augmentation. For the purpose of this study, we consider any approach to be real time if the delay is below 50ms, a typical order up from when humans notice a lag. The reconstructive approaches can be considered near real-time with a delay in the order of seconds. This is achieved by not doing full 3D reconstruction but some forms of approximation yielding good enough results. For Piero [Red18], the response time depends on the required human operations. The time needed to generate visualizations is often between several seconds to some minutes, with no upper limit in principle. Viz Libero [Viz18] instead takes several minutes to generate non-trivial visualizations. Only the systems in this dissertation support real time analysis in their currently existing version.

Active/Used. Promising or successful systems are typically developed further or actively used. For the presented systems, to the best of our knowledge only the simple Table Overlay, the commercial systems True View [Int18], Piero [Red18], and Viz Libero [Viz18] are used in the field. Further, the approaches of this dissertation are actively developed and improved as well as used in the field (see Chapter 7.2.2), for example, in the first Austrian soccer league.

Study/Evaluation. Table Overlays have been in use for decades in the industry. These representations are extremely common and provide the basis for other, more detailed visualizations. Studies discussing the other approaches only exists rudimentary for Wan et al. [WWXT04], and as evaluations for Liu et al. [LJHX08], Schlipsing et al. [SSTI14], as well as the approaches from this dissertation. For the industry applications True View [Int18], Piero [Red18], and Viz Libero [Viz18] the usefulness is proven in form of continued usage and viewer feedback.

Concluding the comparison, we consider Piero [Red18] and Viz Libero [Viz18] the most advanced approaches employed in the field in terms of manually annotated visualizations, while the research within this dissertation represents the state of the art in terms of automation of the whole process, generating the visualizations without user input and augmenting the video with relevant and useful information. As can be seen in Table 6.3.1, the approach proposed by this dissertation currently only lacks of reconstruction possibilities. When discussing the future work of this dissertation, we present our current experiments on novel possibilities for 3D reconstruction, based on single camera video recordings, which will close this gap in the future (Chapter 7.3.1).

6.4 Discussion and Conclusion

Visualization aims to provide an abstract representation to enhance and speed-up the understanding of complex data. This is especially true for video-based soccer visualization, where the cooperative and competitive behavior of two groups of persons, trying to follow opposing tactics, leads to complex relational and hypervariate data, which is hard to visualize in context. An efficient and useful representation often is not known a priori, and more often than not depends largely on the specific aspects under consideration. The survey in this chapter provided a comprehensive overview of the latest methods developed for the video-based analysis of soccer matches. The field is underdeveloped and few significant publications exist. The existing literature has been surveyed and relevant approaches have been identified. Apart from traditional ones like data tables, charts, and 2D maps, these include more advanced visualizations of formations, free spaces, interaction spaces, distance measurements, ball trajectories, selecting visible actors, player trajectories, and density distributions. The works under consideration have been compared according to eleven criteria in the five areas **complexity**, **visualization type**, **capabilities**, **assessment**, and **automation level**. For all approaches, we have cited accompanying studies that show how each approach is expected to significantly improve and simplify soccer analysis, with some example scenarios given. We identify the approaches proposed by this dissertation, the systems Piero [Red18] and Viz Libero [Viz18] as well as, in the area of 3D reconstruction for soccer, True View [Int18] to be the most advanced methods currently available. In some sense, the main difference between the Piero system as well as the Viz Libero system compared to the approaches introduced within this dissertation is that the former more or less represent the state of the art in manually annotated video-based soccer analysis with advanced graphical display, while the latter is a fully automated approach. Combining those two would likely lead to significant benefits. Almost all other academic approaches are in their infancy. This is surprising, given the technological advances in recent years and the commercial and political benefits which could be leveraged.

The process of analyzing soccer match performance and then representing the results in an understandable and meaningful manner, using video-based approaches, is a relevant topic which has not been adequately explored yet. The lack of approaches could be attributed either to a lack of research in this area, which might be considered more of an engineering than a scientific problem. However, this view falls short to respect that the academic literature has produced relevant publications of new types of visualization and the underlying mathematical definitions thereof, for example, for interaction spaces. Further, several open research questions exist, like aggregation and high-level representation of a game. It is a sign of significance that many of the presented systems are commercial solutions rather than research papers. This indicates that some works remain unpublished and are directly integrated into or developed for commercial products. A significant obstacle for scientific research is data unavailability, either due to legal or economic aspects. Nevertheless, given novel methods for data extraction (Chapter 3), the ascent of big data analysis and machine learning in the last decade, as well as the increased awareness for performance analysis by the teams themselves, we expect research to intensify in the mid- and long-term, likely lead by private and company/club-internal research, which later will transpire to the academic community. Utilizing existing and proposed visualization techniques will provide great value to coaches in the areas of match analysis, game preparation, and opposition research. Simultaneously this will enable them to spot key issues otherwise gone unnoticed. With the emergence of increasingly detailed analysis of gameplay and tactics, visualization and, especially, in-video augmenting plays an increasingly important role. They help to understand capabilities and limits of a given team, identify and resolve specific issues and optimize the overall gameplay to adapt to new and complex challenges.

The Cost Of Bad Data Is The Illusion Of Knowledge.

Tomasz Tunguz

7 Discussion and Future Perspectives

Contents

7.1 Introduction

Due to ever increasing competition pressure, it is important for a team to look for every possible opportunity to improve their performance. Evaluating the effectiveness of existing tactics as well as the development of new strategies is, therefore, an essential task for a successful team. In order to evaluate team tactics and gain strategical insights, both players and analysts are typically tasked with the tedious process of manually analyzing video recordings of previous matches. Such task involves time-consuming operations like the manual tagging and annotation of numerous match recordings. The availability of an automated data acquisition and match analysis pipeline as contributed by this dissertation significantly speeds up the match analysis process by reducing the need for tedious manual work.

The ability to extract and analyze match data in real time during a live match provides a great advantage over previous analysis approaches. With our real-time match analysis, coaches can now enrich their decision-making process with additional facts from the analysis system concerning both the performance of their own team as well as possible changes in the tactical behavior of the opposing team, instead of solely relying on their own intuition to judge the situation. Furthermore, the availability of real-time match data also allows a coach to rapidly analyze the structure of common situations in a match, such as set play situations, in order to react to them as quickly as possible. The availability of accurate match data and related analysis results open up new possibilities for more detailed and effective training of players since many additional aspects of player behavior can be captured, including small details and nuances which would otherwise be unnoticeable for the human eye alone. Analysis results from the systems presented in this dissertation can also be used outside of the training/match preparation field. One example that we have shown is medial presentation of match results. Here, the proposed in-video visualizations can be used to convey tactical concepts to the audience in, e.g., a TV broadcast.

In conclusion, we have proposed a collection of novel and effective methods from fields such as computer vision, machine learning and visual analytics which enable data acquisition and match analysis directly from existing video sources. These methods are capable of providing accurate analysis results both from a recording as well as in real time during a live match. Building on the foundations set by this dissertation will help to further revolutionize the way match analysis is being performed in the upcoming years. Ultimately, the progress enabled by research methods such as in-video visualization will not be limited to the domain of team sports alone, but will have a general impact on how we visualize, see and perceive our data in the future.

7.2 Implications

The research work in this dissertation not only contributes to the theoretical body of knowledge, but also provides real-world benefits, for example, to soccer organizations who have utilized applications stemming from this work and plan to continue doing so in the foreseeable future. Already during the writing of this dissertation, the attention of different media was drawn to the introduced research work (Section 7.2.1). The attention generated by media coverage, press releases as well as invited talks resulted in various established cooperations with international first league clubs (Section 7.2.2).

7.2.1 Media Coverage and Invited Talks

Our work about the detection and annotation of movement context including interaction spaces and free spaces (Chapter 4.2) was content of the blog post *How can clubs use detailed tracking data to evaluate player's decision-making?* (2017-03-03) of David Sumpter, well-known author of books about soccer and collective behavior. Our work about combing video and movement data to generate in-video visualizations (Chapter 4.3 of this dissertation) was featured by the IEEE Xplore Innovation Spotlight in the article *A Winning Combination: Pairing Video and Movement Data to Enhance Sports Data Analysis*. The journal article on which Chapter 2 of this dissertation is partially based on (*How to make sense of team sport data: From acquisition to data modeling and research aspects* [SJS+17]) gathered attention in social media right after publication with 93 tweets from 50 individual users covering an upper bound of 105,858 followers. Tweets came from professionals with different backgrounds such as Mathieu Lacome (Head of R&D at Paris Saint-Germain, current champion of the Ligue 1 in France), Darian Wilken (Coach and performance analyst of the Orlando Pirates, a first league team of South Africa) and Paul Neilson (Director of market development at HUDL analysis, one of the market leaders in the domain of data extraction covering, e.g., the CATAPULT sensor tracking system). In addition, several interviews (*Football and Big Data: A search engine for movements* from the Graz University of Technology), press releases (*Football through the eyes of a computer* from the University of Konstanz) as well as radio broadcasts (*Digitalisation in sport: training data in real time* by Deutschlandfunk) covered the presented work in this dissertation.

Based on the media coverage, the works in this dissertation were presented at several invited talks, for example, at the keynote presentation *Visual Analytics for Soccer* at the 11th International Symposium on Computer Science in Sport (IACSS 2017). Following the aforementioned press release of the University of Konstanz, an invited talk (*Football through the eyes of a computer*) was given at the leading international trade fair MEDICA Medicine + Sports Conference in Düsseldorf. Additionally, the contents of this dissertation were presented as ceremonial speech for the award ceremony of the Airbus Research Award 'Claude Dornier'. Ultimately, parts of this dissertation are featured at the Barça Innovation Hub in the invited Section *Identifying the Match Plan: Context-oriented Interactive Visual Analysis of Football Matches* of the football analytics guide of the Spanish professional soccer club FC Barcelona. Our section in this football analytics guide represents the only work, out of eleven invited sections, with a focus on visual data analytics.

7.2.2 Application by Sports Clubs

Through the previously described media coverage, contact to several professional clubs, interested to apply the presented methods of this dissertation in order to improve their performance, was established. The Austrian first league soccer club TSV Hartberg with its coach Markus Schopp, who is a former professional soccer player, is applying the methods presented within this dissertation for the second season already (at the time of writing). According to Markus Schopp, he is, in his work as a coach of a first league team, always looking for information to improve and advance his performance as a coach. To achieve this, he wants to know as much as possible about an opposing team. He needs possibilities to effectively analyze matches in detail automatically as well as visually to be able to recognize tactical patterns nobody else can see (see Figure 7.2.1). For example, it is very important for him to use our proposed free space detection method (Chapter 4.2) in order to quantify, detect and visualize important free spaces. Besides the initial detection, the visualizations presented in this dissertation help him to communicate his findings with his players (see Figure 7.2.2). An example and representative everyday situation where he analyzed the performance of his team after a recent match can be seen in Figure 7.2.3. His team (TSV Hartberg) is represented by the blue dots. In the given

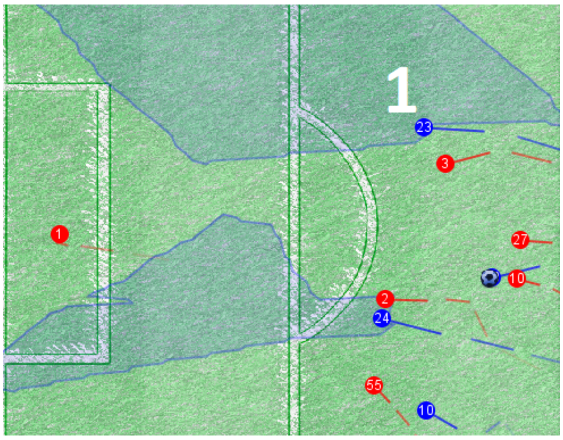
Figure 7.2.1: Markus Schopp (left) using our systems during match preparation and discussing obtained findings with his assistant coach Alexander Kontra (right).

Figure 7.2.2: Markus Schopp (middle) giving tactical advices to one of his players during a match.

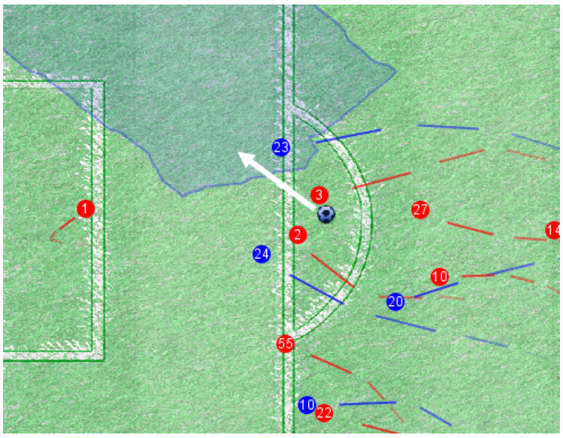
scenario, he inspects a situation where his team scored a goal against the opposing team. First (Figure 7.2.3 (a)), he uses the free space method of this dissertation in order to validate whether the player who did the pass to the striker anticipated the free space of his team mate correctly and if there would have been any alternative players to pass to. Afterwards, in Figure 7.2.3 (b), he verifies if the ball receiving player is moving correctly, based on his movement trajectory and his calculated free space. Ultimately, he analyzes whether his striker decided for the right moment in time to shoot on the opposing goal, so that he is as close as possible but not so close that defending players or the goalkeeper are able to tackle him. For this kind of analysis, he makes extensive use of our proposed interaction space visualization. After studying the situation with the proposed visualization methods, he uses the animated visualizations to coach his players individually, one by one or in subgroups of the team (e.g., defenders, midfielders, attackers).

Besides the important value of our proposed analytical methods for his everyday work, this dissertation advanced his methods as a coach also on the data extraction as well as data annotation side. Here, Markus Schopp is highly interested in our works on real-time data extraction as presented in Chapter 3. He argues that it is very important for him and his team to not be limited to analysis after a match, but ideally be able to detect tactical patterns in real time in order for them to react as fast as possible during the match to identified strategies. The presented skeleton detection (Chapter 3.3) is of additional interest for him. According to Markus Schopp, our skeleton detection offers great potential in the future for more detailed analysis, especially, in the domains of movement and body posture analysis. Skeleton detection will help to include body posture data into the analysis process and, consequently, improve analysis by adding additional information such as whether a player is facing the opposing goal or not. Furthermore, developing novel methods for assessing the perception (field of view, see Section 7.3.3) of players will provide him and his team completely new possibilities during match analysis. Concluding, Markus Schopp as coach of the first league Austrian soccer club TSV Hartberg reports to be very happy about being able to apply the research works proposed in this dissertation for his team and states that our systems contributed to the successful last season in which they managed to keep the league, although every coach presumed the TSV Hartberg to be relegated.

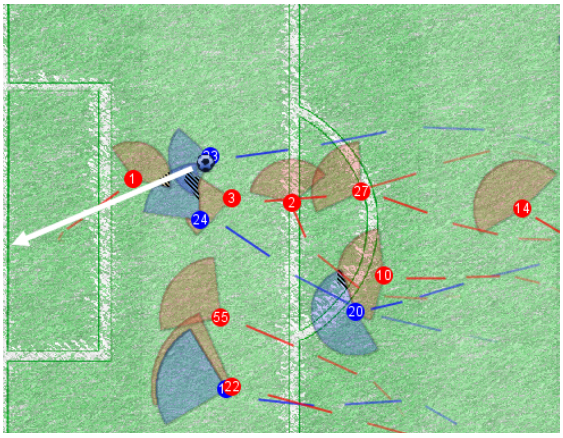
Besides the application by TSV Hartberg, further pilot studies are currently performed with the German first league women's soccer club FF USV Jena. Christoph Schliewe, managing director of the FF USV Jena, expresses high interested in using our systems for performance analysis of their first team as well as youth teams that also play in their respective first leagues. Women soccer clubs in the first league of Germany do not get provided with positional match tracking



(a) Anticipating the free space of player 1



(b) Passing towards the free space



(c) After the pass and before the shot on goal

Figure 7.2.3: Goal scene of TSV Hartberg (blue team) during match analysis of Markus Schopp (coach TSV Hartberg). Markus Schopp analyzed that the player of his team correctly anticipated the free space of his team mate (a) and made a good pass towards him (b). Once the pass reaches the striker (c), the defending team could not put him under enough pressure to stop him from scoring the goal as visualized by the interaction spaces.

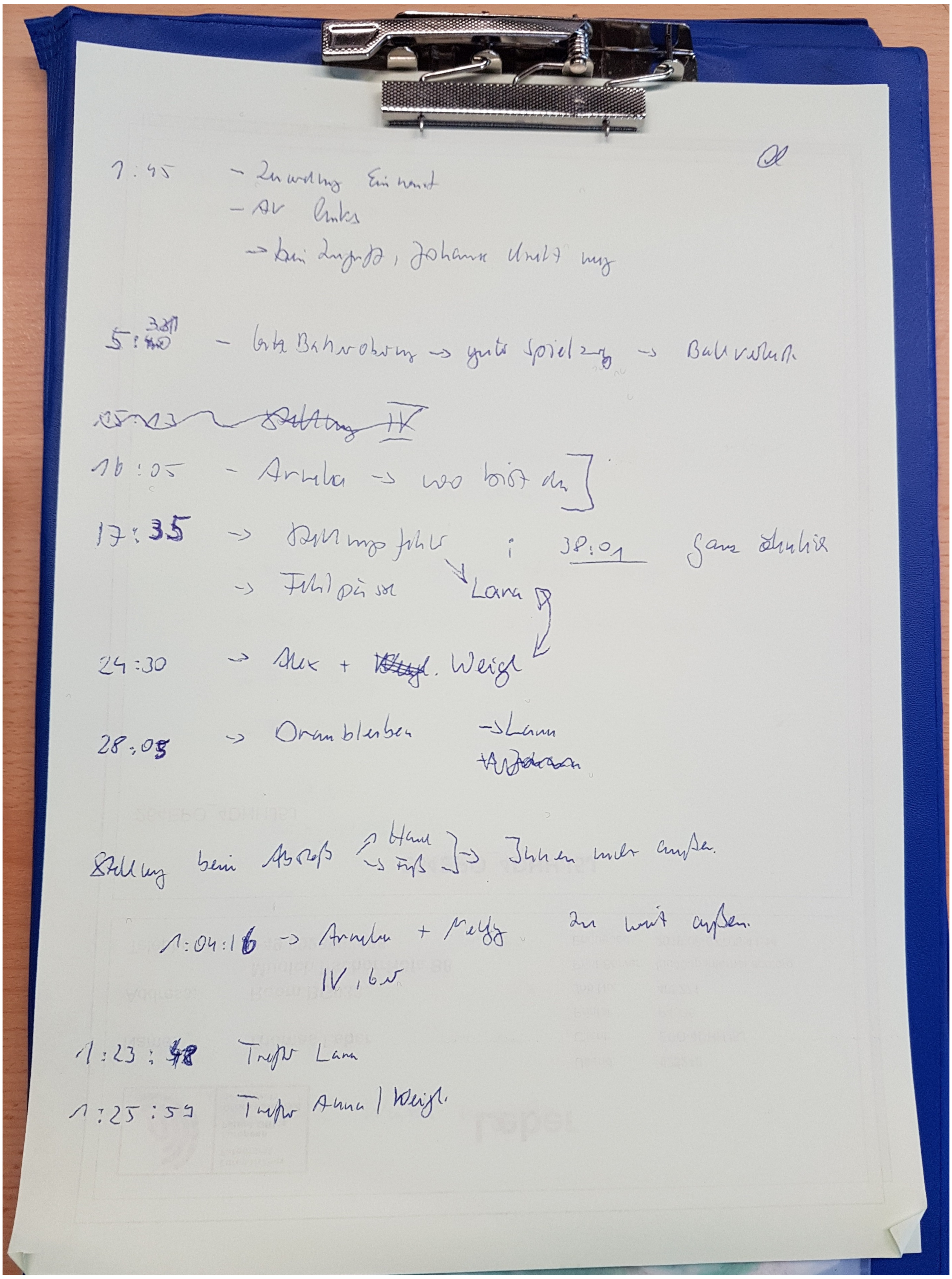


Figure 7.2.4: Until today, due to high costs of data extraction and resulting data unavailability as well as lack of established movement analysis methods, coaches and analysts still have to fall back on manual (handwritten) notes when analyzing a match

Figure 7.2.5: Coaches of USV Jena (first women's German Bundesliga) providing guidance to their players during the half-time break. So far, they have no data analytics solution to analyze their performance.

data. Instead, they still have to rely on manual (handwritten) notes during match analysis (Figure 7.2.4 and Figure 7.2.5). Our approaches for data extraction, automatic context enrichment as well as visual analysis are considered very innovative and are currently tested for full application. Similar to the FF USV Jena, the Austrian first league handball team and record champion Bregenz Handball with their coach Markus Burger are currently installing high resolution cameras in their arena to apply the contributed methods to handball. During several meetings, Björn Tyrner, the managing director of Bregenz Handball, emphasized the innovative character as well as the potential of our proposed approach. Ultimately, meetings with Pascal Bauer (manager for data analysis and machine learning at the German Football Association DFB) are scheduled to discuss how the research introduced in this dissertation can be applied at the DFB.

7.3 Future Work

Based on the research of this dissertation, a large variety of new and interesting research challenges in many domains arises. In the following, we provide brief discussions of some of the most interesting research fields in the upcoming years. For some of these areas, we already performed first experiments.

7.3.1 3D Reconstruction

One interesting future use case is the possibility to use the tracking information from a match analysis system, which includes player body poses, ball movement as well as many other characteristics to reconstruct a recorded match situation in a 3D environment. A player could then use a Virtual Reality Headset to place himself directly into a virtual playback of the recorded match. This type of immersive presentation allows the player to observe the situation in a more realistic way compared to a simple video recording and can result in a more effective learning experience since many psychological aspects (see Section 7.3.3) which affect the player in a match situation can be stimulated, which would not be possible with a simple video playback. In the state of the art, several systems in this direction have been proposed, however existing methods are not able to reconstruct an existing match situation from a recorded match. Especially reconstructions from video recordings of single moving cameras are challenging. Using our detected player skeletons from Chapter 3, we performed first experiments with promising results building on top of established methods for 3D human pose estimation [HL18, WZX17, BKL+16, MHRL17, MSS+17, ZMS18]. The intermediate result can be seen in Figure 7.3.1 (a)-(c). On the left side, respectively, is the detected 2D player skeleton by our approach. On the right side is the corresponding interactive 3D reconstruction of the player.

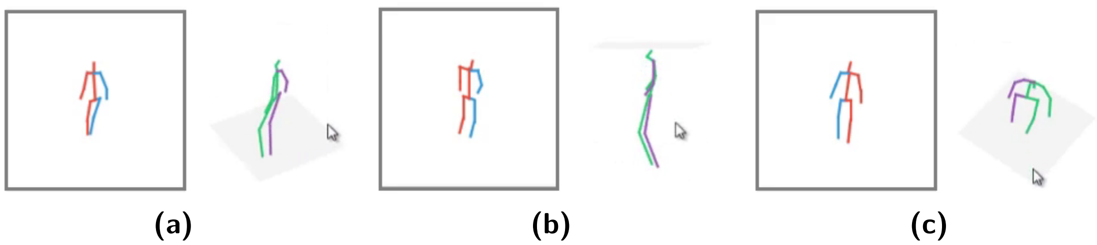


Figure 7.3.1: 3D player reconstruction based on 2D pose data. The mouse can be used to interactively explore the 3D body pose.

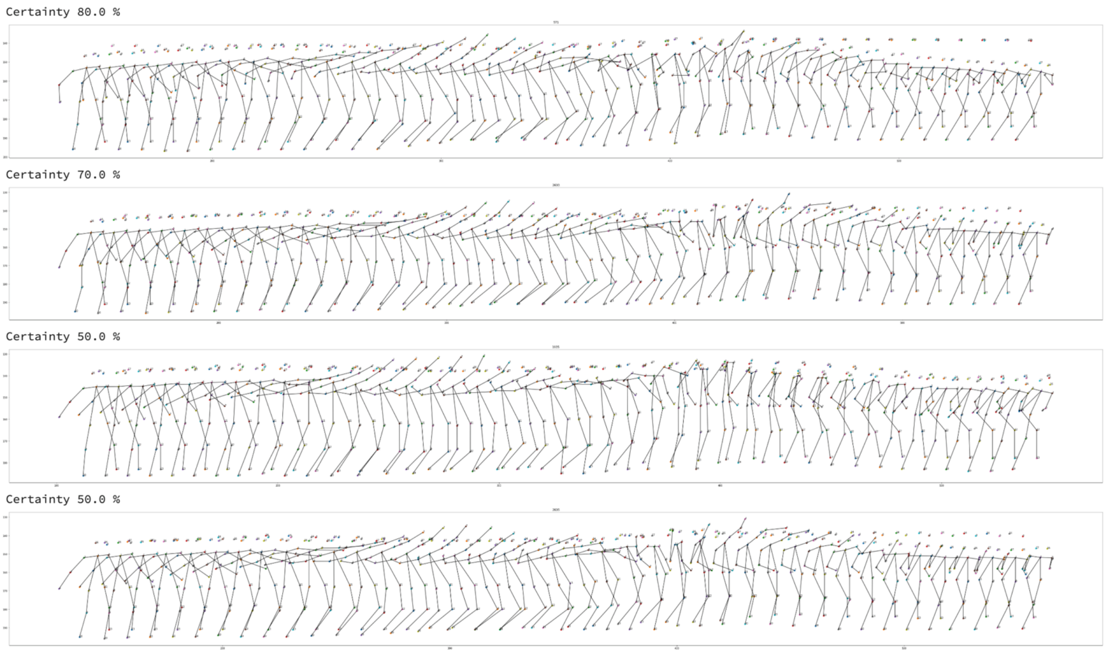
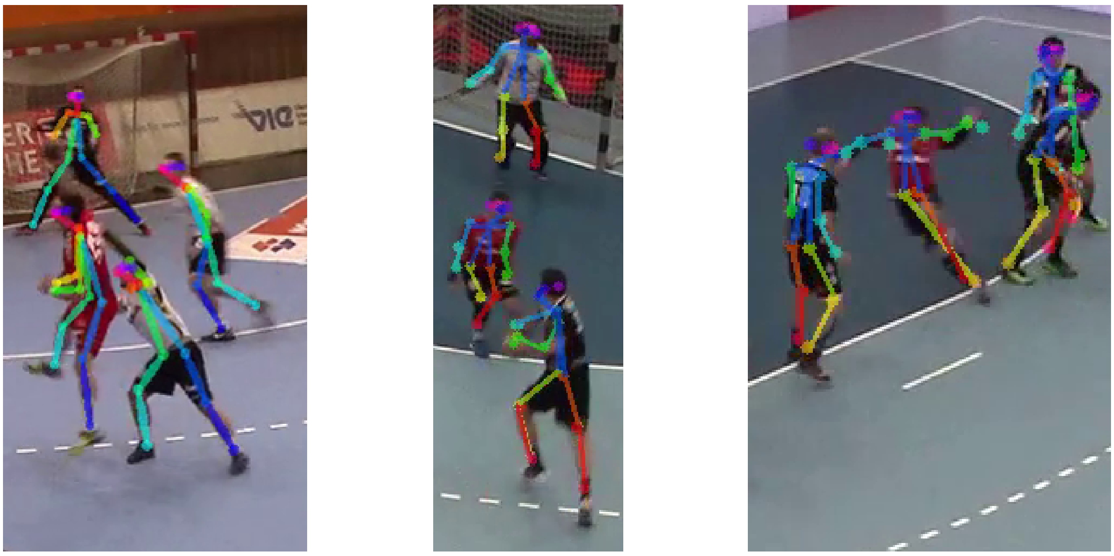


Figure 7.3.2: Automatically detecting and visualizing serves in a tennis match. Analyzing body posture data provides novel opportunities to gain insights in player performance analysis. It can be used to train machine learning classifiers in order to find repeating skeleton movement patterns or even to generate representing *fingerprints* of a moving person.

7.3.2 Skeleton Analysis

Another interesting direction for future research is the analysis of body postures and skeleton motion data for performance analysis. As the proposed skeleton detection provides more fine-grained data of human skeletons including specific body parts, more detailed analysis such as technique training (throwing and shooting motions) become possible. For example, the exact motion of a serving tennis player (see Figure 7.3.2) can be automatically detected and, afterwards, compared in order to identify possibilities for improvement. It would also be feasible to calculate *fingerprints* of players based solely on skeleton movement features, similar to proposed approaches for literature fingerprinting [KO07] in the past. The resulting fingerprints can then be used in, for example, youth sports to improve teaching and the possibilities to provide feedback. Visual skeleton analysis will also be very helpful for live injury prophylaxis as forces that have affected a player in a duel as well as physical deficits and malpositions can be recognized and visualized (see Figure 7.3.3). Ultimately, skeleton analysis enables the implementation of novel techniques incorporating the body orientation into established analysis processes.



(a) (b) (c)

Figure 7.3.3: Sport Medicine and Injury Prophylaxis can be improved using novel algorithms for the analysis of player skeleton data. Displayed are several examples of automatically detected player skeletons during a handball match.

7.3.3 Evaluating the Influence of Stress and Perception

There are many factors located within the individual player that contribute to the success of the team as a whole. Besides obvious aspects like physical strength, bodily fitness, quick wit and endurance, there are a number of psychological factors that also affect the performance of the individual player, and contribute to the physical factors mentioned above. One such factor is the individual psychological arousal of the player, which is often referred to as stress. Stress can be acute or chronic, and can have many causes. Tension among the members of the team can lead to individual members being more stressed. In addition or parallel to these are personal factors, e.g., family disputes, unresolved conflicts, financial troubles, which in isolation or combined can further contribute to elevated individual stress levels. Efforts are currently underway to determine whether stress can be detected from video recordings of individual players. As part of a collaborative project between the Departments of Computer Science and Psychology at the University of Konstanz, we are currently planning to execute a project where we record individual subjects in various physiological states and apply machine learning to investigate whether we can identify the levels of stress from body motion analysis using our previously described skeleton detection method. Additionally, we see great potential in developing a novel measure for dynamic field of view calculation incorporating stress factors based on player skeletons. Existing works in the fields of psychology [Wil88, Dir83, Wil95, Ros16] provide the first step towards this direction.

7.3.4 Analyzing Training Data

Professional clubs have usually at least one match per week. In order to prepare for the match day, a team has regularly around four to five training courses between matches. The analysis of training performance, therefore, takes an important role as coaches decide which players to nominate for the line up based on the performance of each individual player during training. However, the state of the art in training analysis is so far mainly limited to the assessment of physiological factors such as training load [VNRD17], injury prevention [ALVM+19] and the rating of perceived exertion [HSD+17]. Movement data is only used for the calculation of simple statistics such as the number of sprints a player performed or the distance run. In addition, the mental well-being of a player is currently often not included in the analysis despite its importance (see the previous section).

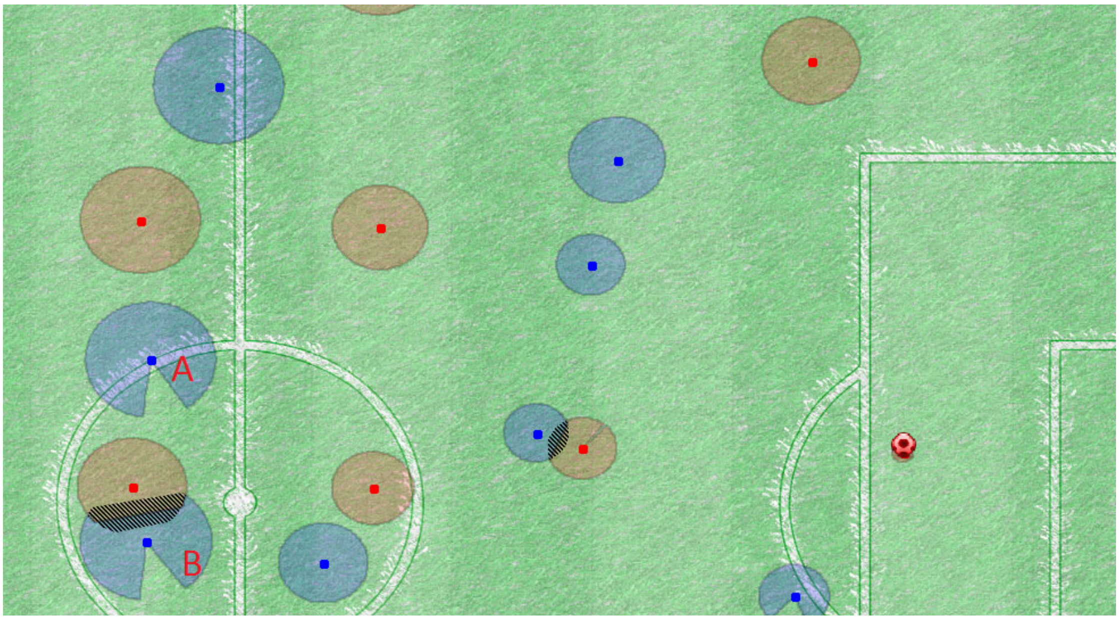
Consequently, many interesting research questions in the analysis of training performance will arise in future work. To this end, it will be very interesting to analyze how the mental state of players influences their performance, how training data can be integrated to improve the possibilities of match analysis and how novel measures for training performance can be used to predict match performance. Furthermore, using novel machine learning approaches for automatically segmenting and assessing different exercises is expected to decrease the time required for the analysis of training data. To tackle these research questions, we are currently performing iterative design studies in direct cooperation with the aforementioned professional clubs to create methods for easy data acquisition and analysis. Here, one goal is to develop novel visual analytics methods to summarize the performance of players over several complete training courses as well as individual exercises.

7.3.5 Match Aggregation and Team Summarization

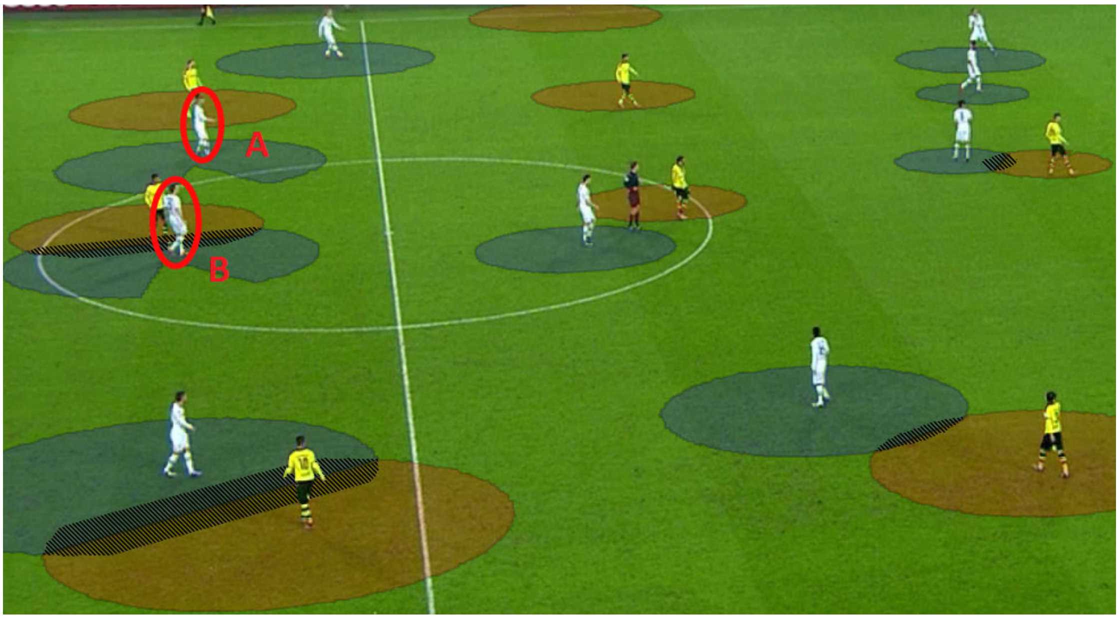
The visual analysis of team sports is in most cases still limited to individual games. While multi-match statistics are manually provided in a static way as match reports, only few automatic high-level analysis of the movement behavior are performed [AAA+19]. Instead, match reports contain basic visualizations such as bar charts, heatmaps or positional plots for events and star glyphs for performance metrics. Additionally, these static match reports often contain raw data tables without the possibility to export, link or further process the data. Therefore, future work will provide great opportunities to contribute novel visual analytics techniques for the summarization of matches as well as the assessment of teams. Interactive reports focusing on the collective movement behavior and steerable by the analyst could, for example, provide match overviews quickly and enable analysts to gain significant insights on the characteristics of a team such as player role analysis and formation behavior. Additionally, the analysis of player and team development over a whole season will be of high interest.

7.3.6 Transfer Possibilities of In-Video Visualizations

Ultimately, the in-video visualizations presented in this dissertation are useful in many application areas besides sports analysis. In fact, we argue that by projecting the analysis results into a space that is closer to the real world, users are enabled to reveal contradictions that would go unnoticed otherwise. Several of our invited experts repeatedly reported that they became more aware of a visualization's limitations and possibilities for improvement in the future using in-video visualizations as, for example, soccer players were not represented by moving dots on an abstract pitch anymore but with the real persons. An example based on the proposed interaction spaces can be seen in Figure 7.3.4. Automatically determined interaction spaces of players were initially calculated based on players' speed, distance to the ball and running direction representing the body orientation. Afterwards, interaction spaces are visualized as circles or circular sectors on an abstracted soccer pitch as can be seen in Figure 7.3.4 (a). However, a players movement direction does not necessarily reflect his or her body orientation. If the same visualization is projected into a video of the real world soccer match, as shown in Figure 7.3.4 (b), we notice that the annotated players **A** and **B** are not running forwards, but sideways which was at first not reflected in our initial model of the interaction spaces. This initial finding by our experts enabled us to adjust our design using our proposed skeleton detection approach to incorporate the body orientation of a player. This exemplary problem could not attract attention outside of the in-video visualization providing real world context. We, consequently, believe that the field of in-video visualizations will play an important role in the future of visual data analysis where it has the potential to lead to enormous benefits in many application fields.



(a) Interaction Spaces visualized on an abstract soccer pitch



(b) Interaction Spaces superimposed on the original video recording

Figure 7.3.4: Calculating interaction spaces for the same scene in a soccer match once visualized on an abstracted soccer pitch (top) and once superimposed on the original video recording (bottom). As interaction spaces are used to calculate the region each player is able to control until the ball reaches him or her, a player's orientation is important during calculation. In this scene, the annotated players **A** and **B** are moving upwards which influences their respective interaction spaces. By projecting the same visualization into the original video recording (bottom), we notice that the players are not running forwards, but sideways which was initially not reflected in a first (outdated) model of the interaction spaces.

Danksagung

Ich erinnere mich gerne daran, wie ich vor vielen Jahren als angehender Master-Student durch die Gänge der Universität Konstanz gelaufen bin und über Fußballanalyse als mögliche Ausrichtung für meine Masterarbeit nachgedacht habe. Ich hätte mir niemals erträumen können, was für ein Abenteuer vor mir lag. Ich möchte mich von ganzem Herzen bei allen Menschen bedanken, die mich auf diesem Weg begleitet haben und von denen ich lernen durfte.

Zuallererst möchte ich mich bei meinem Erstgutachter, Prof. Dr. Daniel A. Keim, bedanken, der mir die Möglichkeit gegeben hat, einem Spaßprojekt und meiner Leidenschaft für Bewegungsdaten zu folgen. Ich würde sagen dieses Spaßprojekt hat seinen Sinn erfüllt und viel Freude bereitet. Vielen Dank Daniel, für deine Unterstützung in den letzten 11 Jahren! Großer Dank gilt ebenfalls meinem Zweitgutachter, Prof. Dr. Michael Grossniklaus. Danke Michael, dass du bei jedem Anliegen stets ein Ohr sowie Rat für mich gefunden hast. Bedanken möchte ich mich ebenfalls bei meinem Drittgutachter, Prof. Dr. Gennady Andrienko, der meine Leidenschaft zur Fußballanalyse stets geteilt und insbesondere in den Arbeiten zur Abstraktion von Bewegungsdaten sowie zur in-Video Visualisierung großen Anteil hat.

Neben meinen Gutachtern möchte ich mich bei allen ehemaligen und momentanen Mitarbeitern des Lehrstuhls für Datenanalyse und Visualisierung an der Universität Konstanz bedanken. Hervorheben möchte ich hier insbesondere meinen ehemaligen Masterbetreuer Halldór Janetzko sowie meinen Kollegen Daniel Seebacher. Danke Halldór, dass du mir beigebracht hast zu meinen Schwächen zu stehen und sie zu meinen Stärken zu machen. Lieber Daniel, danke für jedes kalte Getränk, jedes gute Lied und jeden gemeinsamen Besuch einer Kirschessigfliegenkonferenz.

Großen Dank möchte ich ebenfalls meinen Co-Autoren aussprechen: Daniel Keim (25 Paper), Tobias Schreck (19 Paper), Halldór Janetzko (17 Paper), Daniel Seebacher (14 Paper), Michael Grossniklaus (8 Paper), Thorsten Breitkreutz (4 Paper), Johannes Häußler

(4 Paper), Dominik Sacha (4 Paper), Philip Zimmermann (2 Paper), Gennady Andrienko (2 Paper), Matthias Kraus (2 Paper), Andreas Lamprecht (2 Paper), Niklas Weiler (2 Paper), Feeras Al-Masoudi (2 Paper), Markus Schopp, Julian Bruns, Juri Buchmüller, Iain Couzin, Oliver Deussen, Mennatallah El-Assady, Ulrich Engelke, Max Fischer, Matthias Frank, Bastian Goldlücke, Jürgen Hölsch, Dominik Jäckle, Alexander Jäger, Tassilo Karge, Sven Kosub, Matthias Miller, Manuel Nagel, Benjamin Neldner, Christoph Niederberger, Jens Nimis, Tom Polk, Lin Shao, Lyubka Sharalieva, Viliam Simko, David Spretke und Patrick Wiener. Danke für euren Einsatz während zahlreicher Iterationen und Feedbackrunden!

Zusätzlich möchte ich mich bei allen Studenten bedanken, die ich während dieser Doktorarbeit in ihren Abschlussarbeiten betreuen durfte: Thorsten Breitkreutz (Bachelor und Master), Andreas Lamprecht, Johannes Häußler, Philip Zimmermann (Bachelor und Master), Sebastian Strumbelj, Tassilo Karge, Alexej Gluschkow, Dagmar Sorg, Daniel Hafner, Julia Klein, Max Fischer und Thomas Griesshaber. Danke! Es hat mir unglaublichen Spaß gemacht mit euch zusammen zu arbeiten!

Herzlichen Dank auch an die vielen genannten und ungenannten Experten, die durch ihre wiederholten Teilnahmen an zahllosen Studien und Feedbackrunden großen Anteil am Erfolg dieser Arbeit haben. Hier möchte ich mich insbesondere bei Markus Schopp für die Zusammenarbeit in den letzten zwei Jahren bedanken. Danke Markus! Unser Austausch motiviert mich stets noch einen Schritt weiter zu gehen. Ich möchte mich außerdem bei allen genannten und ungenannten Korrekturlesern bedanken. Bitte verzeiht mir jeden Kommafehler in dieser Danksagung!

Schließlich möchte ich mich bei meiner Familie bedanken. Danke Mama und Papa, dass ihr mich zu dem Menschen erzogen habt, der ich heute bin und es mir ermöglicht habt, meinen Träumen zu folgen, egal ob beim Flossenschwimmen oder im Studium. Lieber Papa, du fehlst mir. Ich habe während dieser Doktorarbeit oft an dich gedacht. Danke auch an meinen Bruder (mein Bruder!) Johannes, der stets mit einer guten Offensive an meiner Seite steht.

Zuallerletzt. Danke Dilanie, dass du seit meinem Master mit mir über die volle Distanz gehst. Ich liebe dich von ganzem Herzen und weiß, dass deine Unterstützung in den vergangenen Jahren nicht selbstverständlich war. Ohne dich hätte ich das nicht geschafft. Lieber Lio, auch bei dir möchte ich mich bedanken (auch wenn du diesen Text erst in einigen Jahren lesen wirst). Danke, dass du mein Leben jeden Tag aufs Neue bereicherst, seit du als Baby in meinen Armen gelegen bist während ich meinen Vortrag zum Proposal-Talk geübt habe. Jetzt bin ich wirklich Dr. Papa.

Bibliography

[AAA+19] G. Andrienko, N. Andrienko, G. Anzer, P. Bauer, G. Budziak, G. Fuchs, D. Hecker, H. Weber, and S. Wrobel. Constructing spaces and times for tactical analysis in football. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, 2019.

[AAB+13a] Gennady L. Andrienko, Natalia V. Andrienko, Peter Bak, Daniel A. Keim, and Stefan Wrobel. *Visual Analytics of Movement*. Springer, 2013.

[AAB+13b] Natalia V. Andrienko, Gennady L. Andrienko, Louise Barrett, Marcus Dostie, and S. Peter Henzi. Space transformation for understanding group movement. *IEEE Trans. Vis. Comput. Graph.*, 19(12):2169–2178, 2013.

[AAB+17] Gennady L. Andrienko, Natalia V. Andrienko, Guido Budziak, Jason Dykes, Georg Fuchs, Tatiana von Landesberger, and Hendrik Weber. Visual analysis of pressure in football. *Data Min. Knowl. Discov.*, 31(6):1793–1839, 2017.

[ABC+03] Jürgen Assfalg, Marco Bertini, Carlo Colombo, Alberto Del Bimbo, and Walter Nunziati. Semantic annotation of soccer videos: automatic highlights identification. *Computer Vision and Image Understanding*, 92(2-3):285–305, 2003.

[ALVM+19] Francisco Ayala, Alejandro López-Valenciano, Jose Antonio Gámez Martín, Mark De Ste Croix, Francisco J Vera-Garcia, Maria del Pilar García-Vaquero, Iñaki Ruiz-Pérez, and Gregory D Myer. A preventive model for hamstring injuries in professional soccer: Learning algorithms. *International journal of sports medicine*, 40(05):344–353, 2019.

[AMST11] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. *Visualization of Time-Oriented Data*. Human-Computer Interaction Series. Springer, 2011.

[AWKG05] E. L. Andrade, J. C. Woods, E. Khan, and M. Ghanbari. Region-based analysis and retrieval for tracking of semantic objects and provision of augmented information in interactive sport scenes. *Trans. Multi.*, 7(6):1084–1096, 2005.

[BBC] Bbc sport - american football - nfl in a nutshell. `http://news.bbc.co.uk/sport2/hi/other_sports/american_football/3192002.stm`. (Accessed on October 6, 2019).

[BBC05] BBC R&D. iview: free-viewpoint video, 2005.

[BBC12] BBC R&D. Augmented Video Player, 2012.

[BBE+13] Tanja Bergmann, Stefan Bunk, Johannes Eschrig, Christian Hentschel, Magnus Knuth, Harald Sack, and Ricarda Schüler. Linked soccer data. In *Proceedings of the I-SEMANTICS 2013 Posters & Demonstrations Track, Graz, Austria, September 4-6, 2013*, pages 25–29, 2013.

[BCD+12] R. Borgo, M. Chen, B. Daubney, E. Grundy, G. Heidemann, B. Höferlin, M. Höferlin, H. Leitte, D. Weiskopf, and X. Xie. State of the Art Report on Video-Based Graphics and Video Visualization. *Computer Graphics Forum*, 31(8):2450–2477, 2012.

[BE04] BBC R&D and Ericsson. Piero sports graphics system, 2004.

[Bia14] Carl Bialik. The People Tracking Every Touch, Pass And Tackle in the World Cup. *FiveThirtyEight*, 10.06.2014.

[BKL+16] Federica Bogo, Angjoo Kanazawa, Christoph Lassner, Peter V. Gehler, Javier Romero, and Michael J. Black. Keep it SMPL: automatic estimation of 3d human pose and shape from a single image. In *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V*, pages 561–578, 2016.

[BL07] Matthew Brown and David G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 74(1):59–73, 2007.

[Blu39] Herbert Blumer. Collective behavior. In Robert E. Park, editor, *An Outline of Principles of Sociology*, pages 219–280. Barnes & Noble, New York, 1939.

[BPSS10] Jerome Bourbousson, Germain Poizat, Jacques Saury, and Carole Seve. Team coordination in basketball: Description of the cognitive connections among teammates. *Journal of Applied Sport Psychology*, 22(2):150–166, 2010.

[BR00] T. Boren and J. Ramey. Thinking aloud: reconciling theory and practice. *IEEE Transactions on Professional Communication*, 43(3):261–278, Sep. 2000.

[BS16] Rahul C. Basole and Dietmar Saupe. Sports Data Visualization [Guest editors' introduction]. *IEEE Computer Graphics and Applications*, 36(5):24–26, 2016.

[BSM10a] Jerome Bourbosson, Carole Seve, and Tim McGarry. Space-time coordination dynamics in basketball: Part 1. intra- and inter-couplings among player dyads. *Journal of Sports Sciences*, 28(3):339–347, 2010.

[BSM10b] Jerome Bourbosson, Carole Seve, and Tim McGarry. Space-time coordination dynamics in basketball: Part 2. the interaction between the two teams. *Journal of Sports Sciences*, 28(3):349–358, 2010.

[Car96] Bradley P Carlin. Improved ncaa basketball tournament modeling via point spread and team strength information. *The American Statistician*, 50(1):39–43, 1996.

[cata] CATAPULT. https://www.catapultsports.com/. (Accessed on October 6, 2019).

[catb] Soccerstats.us. https://www.catapultsports.com/. (Accessed on October 6, 2019).

[CCMM15] Filipe Manuel Clemente, Micael Santos Couceiro, Fernando Manuel Lourenço Martins, and Rui Sousa Mendes. Using network metrics in soccer: A macro-analysis. *Journal of human kinetics*, 45(1):123–134, 2015.

[CCP16] Paolo Cintia, Michele Coscia, and Luca Pappalardo. The haka network: Evaluating rugby team performance with dynamic graph analysis. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016, San Francisco, CA, USA, August 18-21, 2016*, pages 1095–1102, 2016.

[CEF+06] Soumen Chakrabarti, Martin Ester, Usama Fayyad, Johannes Gehrke, Jiawei Han, Shinichi Morishita, Gregory Piatetsky-Shapiro, and Wei Wang. Data mining curriculum: A proposal (version 1.0). *Intensive Working Group of ACM SIGKDD Curriculum Committee*, page 140, 2006.

[CGP+15] Paolo Cintia, Fosca Giannotti, Luca Pappalardo, Dino Pedreschi, and Marco Malvaldi. The harsh rule of the goals: Data-driven performance indicators for football teams. In *2015 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2015, Campus des Cordeliers, Paris, France, October 19-21, 2015*, pages 1–10, 2015.

[CK07] Varun Chandola and Vipin Kumar. Summarization - compressing data into an informative representation. *Knowl. Inf. Syst.*, 12(3):355–378, 2007.

[CRP15] Paolo Cintia, Salvatore Rinzivillo, and Luca Pappalardo. A network-based approach to evaluate the performance of football teams. In *Machine Learning and Data Mining for Sports Analytics Workshop, Porto, Portugal*, 2015.

[CS05] Kyuhyoung Choi and Yongduek Seo. Tracking soccer ball in TV broadcast video. In *Image Analysis and Processing - ICIAP 2005, 13th International Conference, Cagliari, Italy, September 6-8, 2005, Proceedings*, pages 661–668, 2005.

[CSWS17] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 1302–1310, 2017.

[CVV99] András Czirók, Mária Vicsek, and Tamás Vicsek. Collective motion of organisms in three dimensions. *Physica A: Statistical Mechanics and its Applications*, 264(1):299–304, 1999.

[CWNB14] Christopher Carling, Craig Wright, Lee John Nelson, and Paul S. Bradley. Comment on 'performance analysis in football: A critical review and implications for future research'. *Journal of Sports Sciences*, 32(1):2–7, 2014.

[Dar19] Dartfish. MyDartfish, 2019.

[dB00] Mark de Berg. Computational geometry: algorithms and applications, 2nd edition. Springer, 2000.

[DDG06] A Dearden, Y Demiris, and O Grau. Tracking football player movement from a single moving camera using particle filters. pages 29–37. IET, 2006.

[DH72] Richard O. Duda and Peter E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15, 1972.

[Dir83] GR Dirkin. Cognitive tunneling: Use of visual information under stress. *Perceptual and Motor Skills*, 56(1):191–198, 1983.

[dOL03] M.C.F. de Oliveira and H. Levkowitz. From visual data exploration to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 9(3):378–394, 2003.

[DWW05] Matthias Dick, Oliver Wellnitz, and Lars C. Wolf. Analysis of factors affecting players' performance and perception in multiplayer games. In *Proceedings of the 4th Workshop on Network and System Support for Games, NETGAMES 2005, Hawthorne, New York, USA, October 10-11, 2005*, pages 1–7, 2005.

[emm] European media monitor. https://emm.newsbrief.eu. (Accessed on October 6, 2019).

[ES84] K Anders Ericsson and Herbert Alexander Simon. *Protocol analysis*. MIT-press, 1984.

[Est02] Vladimir Estivill-Castro. Why so many clustering algorithms: a position paper. *SIGKDD Explorations*, 4(1):65–75, 2002.

[ET03] Ahmet Ekin and A. Murat Tekalp. Robust dominant color region detection and color-based applications for sports video. In *Proceedings of the 2003 International Conference on Image Processing, ICIP 2003, Barcelona, Catalonia, Spain, September 14-18, 2003*, pages 21–24, 2003.

[ETM03] Ahmet Ekin, A. Murat Tekalp, and Rajiv Mehrotra. Automatic soccer video analysis and summarization. *IEEE Trans. Image Processing*, 12(7):796–807, 2003.

[FdPVL12] Wouter Frencken, Harjo de Poel, Chris Visscher, and Koen Lemmink. Variablitiy of inter-team distances associated with match events in elite-standard soccer. *Journal of Sports Science*, 30(12):1207–1213, 2012.

[FKS19] Maximilian T. Fischer, Daniel A. Keim, and Manuel Stein. Video-based analysis of soccer matches. In *2nd International ACM Workshop on Multimedia Content Analysis in Sports (ACM MMSports'19)*, 2019.

[fooa] Football data dump from football-data.co.uk. https://github.com/jokecamp/FootballData/tree/master/football-data.co.uk. (Accessed on October 6, 2019).

[foob] Football-data.org - restful football data. http://api.football-data.org/index. (Accessed on October 6, 2019).

[fooc] Footballsquads. http://www.footballsquads.co.uk/. (Accessed on October 6, 2019).

[FPS96] Usama M. Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37–54, 1996.

[FS05] Akira Fujimura and Kokichi Sugihara. Geometric analysis and quantitative evaluation of sport teamwork. *Systems and Computers in Japan*, 36(6):49–58, 2005.

[Fu11] Tak-Chung Fu. A review on time series data mining. *Eng. Appl. of AI*, 24(1):164–181, 2011.

[GBD97] Jean-Francis Grehaigne, Daniel Bouthier, and Bernard David. Dynamic-system analysis of opponent relationships in collective actions in soccer. *Journal of Sports Sciences*, 15(2):137–149, 1997.

[Gla16] April Glaser. The Cameras That'll Make the Super Bowl Way More Interesting This Year. http://www.wired.com/2016/01/the-cameras-thatll-make-the-super-bowl-way-more-interesting-this-year/, January 2016. Accessed on October 6, 2019.

[Gmb] Impect GmbH. Impect. http://www.impect.com. (Accessed on October 6, 2019).

[GMR+19] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM Comput. Surv.*, 51(5):93:1–93:42, 2019.

[GMTR+16] B Gonçalves, R Marcelino, L Torres-Ronda, C Torrents, and J Sampaio. Effects of emphasising opposition and cooperation on collective movement behaviour during football small-sided games. *Journal of sports sciences*, 34(14):1346–1354, 2016.

[GNK18] Riza Alp Güler, Natalia Neverova, and Iasonas Kokkinos. Densepose: Dense human pose estimation in the wild. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 7297–7306, 2018.

[Gol07] David Goldblatt. *The ball is round: a global history of football*. Penguin UK, 2007.

[GRA04] Neil Gordon, B Ristic, and S Arulampalam. Beyond the kalman filter: Particle filters for tracking applications. *Artech House, London*, 2004.

[Gru50] Frank E Grubbs. Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, pages 27–58, 1950.

[Gru12] Thomas U. Grund. Network structure and team performance: The case of english premier league soccer teams. *Social Networks*, 34(4):682–690, 2012.

[GW14] Joachim Gudmundsson and Thomas Wolle. Football analysis using spatio-temporal tools. *Computers, Environment and Urban Systems*, 47:16–27, 2014.

[HGK+11] Adrian Hilton, Jean-Yves Guillemaut, Joe Kilner, Oliver Grau, and Graham Thomas. 3D-TV Production From Conventional Cameras for Sports Broadcast. *IEEE Transactions on Broadcasting*, 57(2):462–476, 2011.

[HH09] H. L. Jin and H. J. Liu. Research on Visualization Techniques in Data Mining. In *International Conference on Computational Intelligence and Software Engineering 2009*, 2009.

[HHR13] Martin Hoernig, Michael Herrmann, and Bernd Radig. Real time soccer field analysis from monocular TV video data. In *11th International Conference on Pattern Recognition and Image Analysis (PRIA-11-2013)*, volume 2, pages 567–570, Samara, September 2013. The Russian Academy of Sciences.

[HL18] Mir Rayat Imtiaz Hossain and James J. Little. Exploiting temporal information for 3d human pose estimation. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part X*, pages 69–86, 2018.

[HSD+17] Monoem Haddad, Georgios Stylianides, Leo Djaoui, Alexandre Dellal, and Karim Chamari. Session-rpe method for training load monitoring: validity, ecological usefulness, and influencing factors. *Frontiers in neuroscience*, 11:612, 2017.

[Hud19] Hudl. Sportscode, 2019.

[HZ06] Andrew Harltey and Andrew Zisserman. *Multiple view geometry in computer vision (2. ed.)*. Cambridge University Press, 2006.

[Ins85] Alfred Inselberg. The plane with parallel coordinates. *The Visual Computer*, 1(2):69–91, 1985.

[Int18] Intel. True View, 2018.

[JSS+14] Halldór Janetzko, Dominik Sacha, Manuel Stein, Tobias Schreck, Daniel A. Keim, and Oliver Deussen. 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*, pages 13–22, 2014.

[KAK95] Daniel A. Keim, Mihael Ankerst, and Hans-Peter Kriegel. Recursive pattern: A technique for visualizing very large amounts of data. In *IEEE Visualization '95, Proceedings, Atlanta, Georgia, USA, October 29 - November 3, 1995.*, pages 279–286, 1995.

[KBJM18] Angjoo Kanazawa, Michael J. Black, David W. Jacobs, and Jitendra Malik. End-to-end recovery of human shape and pose. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 7122–7131, 2018.

[KCM03] Jinman Kang, Isaac Cohen, and Gerard Medioni. Soccer player tracking across uncalibrated camera streams. In *In: Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS) In Conjunction with ICCV*, pages 172–179, 2003.

[Kei00] Daniel A. Keim. Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Trans. Vis. Comput. Graph.*, 6(1):59–78, 2000.

[Kei01] Daniel A. Keim. Visual exploration of large data sets. *Communications of the ACM*, 44(8):38–44, 2001.

[Kei02] D. A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002.

[KH01] Hyunwoon Kim and Ki-Sang Hong. Robust image mosaicing of soccer videos using self-calibration and line tracking. *Pattern Anal. Appl.*, 4(1):9–19, 2001.

[KHL06] Chan-Hyun Kang, Jung-Rae Hwang, and Ki-Joune Li. Trajectory analysis for soccer players. In *Workshops Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), 18-22 December 2006, Hong Kong, China*, pages 377–381, 2006.

[KKEM10] Daniel A. Keim, Jörn Kohlhammer, Geoffrey P. Ellis, and Florian Mansmann. *Mastering the Information Age - Solving Problems with Visual Analytics*. Eurographics Association, 2010.

[KMR+94] Mika Klemettinen, Heikki Mannila, Pirjo Ronkainen, Hannu Toivonen, and A. Inkeri Verkamo. Finding interesting rules from large sets of discovered association rules. In *Proceedings of the Third International Conference on Information and Knowledge Management (CIKM'94), Gaithersburg, Maryland, USA, November 29 - December 2, 1994*, pages 401–407, 1994.

[KMT09] Daniel A. Keim, Florian Mansmann, and Jim Thomas. Visual analytics: how much visualization and how much analytics? *SIGKDD Explorations*, 11(2):5–8, 2009.

[KO07] Daniel A. Keim and Daniela Oelke. Literature fingerprinting: A new method for visual literary analysis. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology, IEEE VAST 2007, Sacramento, California, USA, October 30-November 1, 2007*, pages 115–122, 2007.

[KSE+03] Vivek Kwatra, Arno Schödl, Irfan A. Essa, Greg Turk, and Aaron F. Bobick. Graphcut textures: image and video synthesis using graph cuts. *ACM Trans. Graph.*, 22(3):277–286, 2003.

[KTAP+95] Urho M Kujala, Simo Taimela, Ilkka Antti-Poika, Sakari Orava, Risto Tuominen, and Pertti Myllynen. Acute injuries in soccer, ice hockey, volleyball, basketball, judo, and karate: analysis of national registry data. *Bmj*, 311(7018):1465–1468, 1995.

[KWB+19] Matthias Kraus, Niklas Weiler, Thorsten Breitkreutz, Daniel A. Keim, and Manuel Stein. Breaking the curse of visual data exploration: Improving analyses by building bridges between data world and real world. In *10th International Conference on Information Visualization Theory and Applications*, 2019.

[LHW07] Jae-Gil Lee, Jiawei Han, and Kyu-Young Whang. Trajectory clustering: a partition-and-group framework. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, Beijing, China, June 12-14, 2007*, pages 593–604, 2007.

[Lin90] DV Lindley. Regression and correlation analysis. In *Time Series and Statistics*, pages 237–243. Springer, 1990.

[Lin14] Daniel Link. Using of invasion profiles as a performance indicator in soccer. In *Proceedings of the International Association of Computer Science in Sports Conference*, 2014.

[LIW05] Patrick Laube, Stephan Imfeld, and Robert Weibel. Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science*, 19(6):639–668, 2005.

[LJHX08] Huiying Liu, Shuqiang Jiang, Qingming Huang, and Changsheng Xu. A generic virtual content insertion system based on visual attention analysis. In Abdulmotaleb EL Saddik, Son Vuong, Carsten Griwodz, Alberto Del Bimbo, K. Selcuk Candan, and Alejandro Jaimes, editors, *Proceedings of the 16th ACM international conference on Multimedia*, page 379, New York, NY, 2008. ACM.

[LMB+14] Tsung-Yi Lin, Michael Maire, Serge J. Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: common objects in context. In *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V*, pages 740–755, 2014.

[LMH+15] R. Leser, B. Moser, T. Hoch, J. Stoegerer, G. Kellermayr, S. Reinsch, and A. Baca. Expert-oriented modelling of a 1vs1-situation in football. *International Journal of Performance Analysis in Sport*, 15(12):949–966, 2015.

[Low04] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[LRIC15] Bongshin Lee, Nathalie Henry Riche, Petra Isenberg, and Sheelagh Carpendale. More than telling a story: Transforming data into visually shared stories. *IEEE Computer Graphics and Applications*, 35(5):84–90, 2015.

[LTL+09] Jia Liu, Xiaofeng Tong, Wenlong Li, Tao Wang, Yimin Zhang, and Hongqi Wang. Automatic player detection, labeling and tracking in broadcast soccer video. *Pattern Recognition Letters*, 30(2):103–113, 2009.

[M+67] James MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA., 1967.

[Map03] C. Maple. Geometric design and space planning using the marching squares and marching cube algorithms. In *2003 International Conference on Geometric Modeling and Graphics, 2003. Proceedings*, pages 90–95, July 2003.

[MC13] Rob MacKenzie and Chris Cushion. Performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences*, 31(6):639–676, 2013.

[MH08] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(Nov):2579–2605, 2008.

[MHRL17] Julieta Martinez, Rayat Hossain, Javier Romero, and James J. Little. A simple yet effective baseline for 3d human pose estimation. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 2659–2668, 2017.

[mic] Bing text to speech api. https://docs.microsoft.com/en-us/azure/cognitive-services/speech/api-reference-rest/bingvoiceoutput. (Accessed on October 6, 2019).

[MSS+17] Dushyant Mehta, Srinath Sridhar, Oleksandr Sotnychenko, Helge Rhodin, Mohammad Shafiei, Hans-Peter Seidel, Weipeng Xu, Dan Casas, and Christian Theobalt. Vnect: real-time 3d human pose estimation with a single RGB camera. *ACM Trans. Graph.*, 36(4):44:1–44:14, 2017.

[MWF16] Andrii Maksai, Xinchao Wang, and Pascal Fua. What players do with the ball: A physically constrained interaction modeling. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 972–981, 2016.

[MWR09] Emad Monier, Per Wilhelm, and Ulrich Rückert. A computer vision based tracking system for indoor team sports. In *The fourth International Conference on Intelligent Computing and Information Systems*, 2009.

[MZJ13] Christopher Mutschler, Holger Ziekow, and Zbigniew Jerzak. The DEBS 2013 grand challenge. In *The 7th ACM International Conference on Distributed Event-Based Systems, DEBS '13, Arlington, TX, USA - June 29 - July 03, 2013*, pages 289–294, 2013.

[nnd] Nndc - climate data online. http://www7.ncdc.noaa.gov/CDO/cdo. (Accessed on October 6, 2019).

[Nvi11] Nvidia, CUDA. C Programming Guide Version 4.0. *Nvidia Corporation*, 2011.

[NWCP07] Catherine D Newell, Mark D Wood, Kathleen M Costello, and Robert B Poetker. Automatic story creation using semantic classifiers for images and associated meta data, June 5 2007. US Patent App. 11/758,358.

[ope] Football.db - free open public domain football data. http://openfootball.github.io/. (Accessed on October 6, 2019).

[opt] Opta. http://www.optasports.com/. (Accessed on October 6, 2019).

[Pea01] Karl Pearson. Liii. on lines and planes of closest fit to systems of points in space, 1901.

[Pel14] Tom Pelissero. Player-tracking system will let NFL fans go deeper than ever. http://www.usatoday.com/story/sports/nfl/2014/07/30/metrics-sensor-shoulder-pads-zebra-speed-tracking/13382443/, 2014. Accessed on October 6, 2019.

[PHVG02] Patrick Pérez, Carine Hue, Jaco Vermaak, and Michel Gangnet. Color-based probabilistic tracking. In *Computer Vision - ECCV 2002, 7th European Conference on Computer Vision, Copenhagen, Denmark, May 28-31, 2002, Proceedings, Part I*, pages 661–675, 2002.

[PIT+16] Leonid Pishchulin, Eldar Insafutdinov, Siyu Tang, Bjoern Andres, Mykhaylo Andriluka, Peter V. Gehler, and Bernt Schiele. Deepcut: Joint subset partition and labeling for multi person pose estimation. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 4929–4937, 2016.

[PM92] John G. Proakis and Dimitris G. Manolakis. *Digital signal processing - principles, algorithms and applications (2. ed.)*. Macmillan, 1992.

[PT12] Javier López Pena and Hugo Touchette. A network theory analysis of football strategies. *arXiv preprint arXiv:1206.6904*, 2012.

[PVF13] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. Soccerstories: A kick-off for visual soccer analysis. *IEEE Trans. Vis. Comput. Graph.*, 19(12):2506–2515, 2013.

[PVS+18] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale. State of the Art of Sports Data Visualization. *Computer Graphics Forum*, 37(3):663–686, 2018.

[PYHZ14] Tom Polk, Jing Yang, Yueqi Hu, and Ye Zhao. Tennivis: Visualization for tennis match analysis. *IEEE Trans. Vis. Comput. Graph.*, 20(12):2339–2348, 2014.

[R+08] Hannes Reijner et al. The development of the horizon graph. 2008.

[red] Reddit api documentation. https://www.reddit.com/dev/api/. (Accessed on October 6, 2019).

[Red18] Redbeemedia. Piero Sports Graphics, 2018.

[Reg] Max Regenhuber. Impect & packing: the future of football analytics is here. http://bundesligafanatic.com/impect-packing-the-future-of-football-analytics-is-here/. (Accessed on October 6, 2019).

[RLRM02] Marirose A Radelet, Scott M Lephart, Elaine N Rubinstein, and Joseph B Myers. Survey of the injury rate for children in community sports. *Pediatrics*, 110(3):e28–e28, 2002.

[RMH+14] Varun Ramakrishna, Daniel Munoz, Martial Hebert, James Andrew Bagnell, and Yaser Sheikh. Pose machines: Articulated pose estimation via inference machines. In *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part II*, pages 33–47, 2014.

[Ros16] Ruth Rosenholtz. Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, 2:437–457, 2016.

[RS14] Amjad Rehman and Tanzila Saba. Features extraction for soccer video semantic analysis: current achievements and remaining issues. *Artificial Intelligence Review*, 41(3):451–461, 2014.

[Rya10] Mark Ryan. The impossible job: Sky TV have 24 cameras but referees can only see so much. http://www.dailymail.co.uk/sport/football/article-1301170/The-impossible-job-Sky-TV-24-cameras-referees-much.html, August 2010. Accessed on October 6, 2019.

[SBH+18] Manuel Stein, Thorsten Breitkreutz, Johannes Häussler, Daniel Seebacher, Christoph Niederberger, Tobias Schreck, Michael Grossniklaus, Daniel A. Keim, and Halldor Janetzko. Revealing the invisible: Visual analytics and explanatory storytelling for advanced team sport analysis. In *2018 International Symposium on Big Data Visual and Immersive Analytics, BDVA 2018, Konstanz, Germany, October 17-19, 2018*, pages 1–9, 2018.

[SCKH97] Yongduek Seo, Sunghoon Choi, Hyunwoo Kim, and Ki-Sang Hong. Where are the ball and players? soccer game analysis with color based tracking and image mosaick. In *Image Analysis and Processing, 9th International Conference, ICIAP '97, Florence, Italy, September 17-19, 1997, Proceedings, Volume II*, pages 196–203, 1997.

[SH10] Edward Segel and Jeffrey Heer. Narrative visualization: Telling stories with data. *IEEE Trans. Vis. Comput. Graph.*, 16(6):1139–1148, 2010.

[Shi18] Huang-Chia Shih. A Survey of Content-Aware Video Analysis for Sports. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(5):1212–1231, 2018.

[SHJ+15] Manuel Stein, Johannes Häussler, Dominik Jäckle, Halldor Janetzko, Tobias Schreck, and Daniel A. Keim. Visual soccer analytics: Understanding the characteristics of collective team movement based on feature-driven analysis and abstraction. *ISPRS Int. J. Geo-Information*, 4(4):2159–2184, 2015.

[SJB+16] Manuel Stein, Halldor Janetzko, Thorsten Breitkreutz, Daniel Seebacher, Tobias Schreck, Michael Grossniklaus, Iain D. Couzin, and Daniel A. Keim. Director's cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, 2016.

[SJKS19] M. Stein, H. Janetzko, D. A. Keim, and T. Schreck. Tackling similarity search for soccer match analysis: Multimodal distance measure and interactive query definition. *IEEE Computer Graphics and Applications*, pages 1–1, 2019.

[SJL+16] Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Daniel Seebacher, Tobias Schreck, Daniel A. Keim, and Michael Grossniklaus. From game events to team tactics: Visual analysis of dangerous situations in multi-match data. In *1st International Conference on Technology and Innovation in Sports, Health and Wellbeing, TISHW 2016, Vila Real, Portugal, December 1-3, 2016*, pages 8:1–8:9, 2016.

[SJL+18] Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Thorsten Breitkreutz, Philip Zimmermann, Bastian Goldlücke, Tobias Schreck, Gennady L. Andrienko, Michael Grossniklaus, and Daniel A. Keim. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE Trans. Vis. Comput. Graph.*, 24(1):13–22, 2018.

[SJS+17] Manuel Stein, Halldor Janetzko, Daniel Seebacher, Alexander Jäger, Manuel Nagel, Jürgen Hölsch, Sven Kosub, Tobias Schreck, Daniel A Keim, and Michael Grossniklaus. How to make sense of team sport data: From acquisition to data modeling and research aspects. *Data*, 2(1):2, 2017.

[SJSK18] M Stein, H Janetzko, T Schreck, and D Keim. Tackling similarity search for soccer match analysis: Multimodal distance measure and interactive query definition. In *Proc. 4th Symposium on Visualization in Data Science*, 2018.

[sky] Sky go - moenchengladbach vs dortmund. http://www.skygo.sky.de/. (Accessed on October 6, 2019).

[SLE14] S. Schmidhofer, R. Leser, and M. Ebert. A comparison between the structure in elite tennis and kids tennis on scaled courts (tennis 10s). *International Journal of Performance Analysis in Sport*, 14(12):829–840, 2014.

[SM12] Jaime Sampaio and Vitor Maçãs. Measuring tactical behaviour in football. *International journal of sports medicine*, 33(05):395–401, 2012.

[SMKS15] Svenja Simon, Sebastian Mittelstädt, Daniel A. Keim, and Michael Sedlmair. Bridging the gap of domain and visualization experts with a liaison. In *Eurographics Conference on Visualization, EuroVis 2015, Short Papers, Cagliari, Sardinia, Italy, May 25-29, 2015*, pages 127–131, 2015.

[soc] Soccerstats.us. http://soccerstats.us/. (Accessed on October 6, 2019).

[SSK+19] Manuel Stein, Daniel Seebacher, Tassilo Karge, Tom Polk, Michael Grossniklaus, and Daniel A. Keim. From movement to events: Improving soccer match annotations. In *MultiMedia Modeling - 25th International Conference, MMM 2019, Thessaloniki, Greece, January 8-11, 2019, Proceedings, Part I*, pages 130–142, 2019.

[SSM+19] Manuel Stein, Daniel Seebacher, Rui Marcelino, Tobias Schreck, Michael Grossniklaus, Daniel A. Keim, and Halldor Janetzko. Where to go: Computational and visual what-if analyses in soccer. *Journal of Sports Sciences*, 0(0):1–9, 2019.

[SSTI14] Marc Schlipsing, Jan Salmen, Marc Tschentscher, and Christian Igel. Adaptive pattern recognition in real-time video-based soccer analysis. *Journal of Real-Time Image Processing*, 13(2):345–361, 2014.

[staa] Stadiumdb - stadium database. http://stadiumdb.com. (Accessed on October 6, 2019).

[stab] Stats. http://www.stats.com/. (Accessed on October 6, 2019).

[TGM+17] Graham Thomas, Rikke Gade, Thomas B. Moeslund, Peter Carr, and Adrian Hilton. Computer vision for sports: Current applications and research topics. *Computer Vision and Image Understanding*, 159:3–18, 2017.

[TH00] Tsuyoshi Taki and Jun-ichi Hasegawa. Visualization of dominant region in team games and its application to teamwork analysis. In *Computer Graphics International Conference, CGI 2000, Geneva, Switzerland, June 19-24, 2000*, pages 227–235, 2000.

[Tho07] Graham Thomas. Real-time camera tracking using sports pitch markings. *Journal of Real-Time Image Processing*, 2(2-3):117–132, 2007.

[TQ01] Vasanth Tovinkere and Richard J. Qian. Detecting semantic events in soccer games: Towards A complete solution. In *Proceedings of the 2001 IEEE International Conference on Multimedia and Expo, ICME 2001, August 22-25, 2001, Tokyo, Japan*, 2001.

[twi] Twitter developers. https://dev.twitter.com/. (Accessed on October 6, 2019).

[Viz18] Viz Libero. Viz Libero, 2018.

[VNRD17] Jos Vanrenterghem, Niels Jensby Nedergaard, Mark A Robinson, and Barry Drust. Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways. *Sports Medicine*, 47(11):2135–2142, 2017.

[wik] Wikipedia. https://www.wikipedia.org/. (Accessed on October 6, 2019).

[Wil88] Leonard J Williams. Tunnel vision or general interference? cognitive load and attentional bias are both important. *The American journal of psychology*, pages 171–191, 1988.

[Wil95] Leonard J Williams. Peripheral target recognition and visual field narrowing in aviators and nonaviators. *The International Journal of Aviation Psychology*, 5(2):215–232, 1995.

[WLY13] Yi Wu, Jongwoo Lim, and Ming-Hsuan Yang. Online object tracking: A benchmark. In *2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013*, pages 2411–2418, 2013.

[WRKS16] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. Convolutional pose machines. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 4724–4732, 2016.

[WWXT04] Kongwah Wan, Jinjun Wang, Changsheng Xu, and Qi Tian. Automatic Sports Highlights Extraction with Content Augmentation. In Kiyoharu Aizawa, Yūichi Nakamura, and Shin'ichi Satoh, editors, *Advances in multimedia information processing - PCM 2004*, volume 3332 of *Lecture notes in computer science, 0302-9743*, pages 19–26. Springer, Berlin and Great Britain, 2004.

[WZX17] Qingfu Wan, Wei Zhang, and Xiangyang Xue. Deepskeleton: Skeleton map for 3d human pose regression. *CoRR*, abs/1711.10796, 2017.

[XCDS02] Lexing Xie, Shih-Fu Chang, Ajay Divakaran, and Huifang Sun. Structure analysis of soccer video with hidden markov models. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2002, May 13-17 2002, Orlando, Florida, USA*, pages 4096–4099, 2002.

[XSL+17] Yuanyi Xue, Yilin Song, Chenge Li, An-Ti Chiang, and Xiaoran Ning. Automatic Video Annotation System for Archival Sports Video. In LSSB, editor, *2017 IEEE Winter Conference on Applications of Computer Vision workshops*, pages 23–28, Piscataway, NJ, 2017. IEEE.

[XWW12] Li-ming Xia, Qian Wang, and Lian-shi Wu. Vision-based behavior prediction of ball carrier in basketball matches. *Journal of Central South University*, 19(11):2142–2151, 2012.

[XZZ+08] Changsheng Xu, Yifan Zhang, Guangyu Zhu, Yong Rui, Hanqing Lu, and Qingming Huang. Using webcast text for semantic event detection in broadcast sports video. *IEEE Trans. Multimedia*, 10(7):1342–1355, 2008.

[YB16] Burcu Yucesoy and Albert-László Barabási. Untangling performance from success. *EPJ Data Sci.*, 5(1):17, 2016.

[zeb] Zebra technologies. https://www.zebra.com/us/en/nfl.html. (Accessed on October 6, 2019).

[ZMS18] Andrei Zanfir, Elisabeta Marinoiu, and Cristian Sminchisescu. Monocular 3d pose and shape estimation of multiple people in natural scenes - the importance of multiple scene constraints. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 2148–2157, 2018.