GIMA

Geographical Information Management and Applications

Geovisual football analytics

Towards the development of an interactive visual interface for football coaches, analysts and players

Universiteit Utrecht

TU Delft

University of Twente.

Wageningen UR

Printed in the Netherlands

First printing, June 2017

Lars Thomas van Hoeve

06 | 2017

Geovisual football analytics

Acknowledgments

After an intensive period of almost eight months I present to you my MSc research titled: Geovisual football analytics. This research project is commissioned by the University of Utrecht, University of Wageningen, University of Twente and the Delft University of Technology as part of the Geographical Information Management and Applications Master of Science. Completing this research would not have been possible without the assistance, encouragement and help from many people to whom I would like to express my gratitude.

Foremost, the greatest thanks goes to my supervisor Dr. Corné van Elzakker, who continuously provided me with his detailed feedback, advice and support. I would also like to express my gratitude to Prof. Dr. Menno-Jan Kraak, whose research has been an inspiration for this study. Many thanks go to my fellow students and friends Joris van den Berg, Simon Groen, Kai Passiers and Ruben van der Valk for their help and feedback. I am also very thankful to Greta Ferloni for her editorial efforts.

A special thanks have to be addressed to Anatoliy Babic and Giels Brouwer, who enthusiastically welcomed me at SciSports and opened doors to the world of football analytics I have never been able to open without them. I would also like to express my gratitude to each and every colleague from SciSports, especially for their help and support. I also like to thank Charles Perin from the University of London for providing me access to the SoccerStories interface.

Last but not least, I would like to thank all people who participated and spending their valuable time and effort to help me complete this research project.

Lars van Hoeve

Utrecht, 5 June 2017

Abstract

Due to recent technological advancements, large volumes of movement data can be collected about almost anything that moves. Professional football organizations have access to these increasing data volumes. However, literature shows that the capacity to collect data has not been matched by the ability to process it in meaningful ways. The question therefore is how to transform this data into useful information on which coaches, analysts and players dare to base their decisions. This research contributes to this development by proposing ways to make movement data more accessible for football clubs by developing a conceptual visual interface to visually explore and analyze concerted movements. The term concerted movement is used to indicate the coordinated movements of multiple objects in relation to each other and the spatiotemporal context in which they move.

Visual interfaces are tools capable of dealing with large amounts of movement data. They can yield knowledge and insight by establishing an interactive relationship between humans and machines that allows users to look at a subject from different perspectives and allow them to modify the data using different visual analytics tools. From an academic point of views, it looks like interfaces are potentially very powerful to explore, analyze and communicate movement patterns to extract relevant information from spatiotemporal data. However, the involvement of users and a successful identification of their needs is often underestimated if not overlooked altogether. This research project tries to alleviate this problem by investigating user requirements prior to the development process of a conceptual visual interface. To realize this objective, the User Centred Design (UCD) methodology is followed.

This research shows that there is a sincere need for more user research in the fields of geo-information science with a focus on geovisual football analytics. What must further be developed is the incorporation of user requirements. Therefore, it is recommended that the conceptual visual interface is turned into a high-fidelity prototype that is implemented and subjected to various iterations of usability research. Furthermore, future football analytics research should embrace a more multidisciplinary approach to support the visual exploration and analysis of spatiotemporal data.

Table of contents

1	**Introduction**	**17**
1.1	Research context	19
1.2	Problem statement	20
1.3	Research objectives	22
1.4	Research questions	23
1.5	Research methodology	24
1.6	Innovation aimed at	24
1.7	Thesis structure	25
2	**Framework and conceptual model**	**27**
2.1	Introduction	29
2.2	Fundamentals of movement	29
2.3	Concerted movements	32
	2.3.1 Movement trajectories	33
	2.3.2 Movement parameters	35
	2.3.3 Movement patterns	36
2.4	Geovisual analytics of movements	39
	2.4.1 Interactive visual interfaces	39
	2.4.2 Geovisual analytics tools	40
	2.4.3 Usability issues	45
2.5	Conceptual model	47
2.6	Conclusion	49

3 Visual analysis of football movements 51

3.1 Introduction 53

3.2 Movement observations 53

3.3 Football analytics 57

3.4 Visual analysis of concerted movements 63

3.4.1 SoccerStories 67

3.4.2 Feature-based visual analytics 68

3.4.3 FootMapp 70

3.4.4 SAP Match Insights 72

3.5 Conclusion 73

4 Methodology 75

4.1 Introduction 77

4.2 User-Centered Design 78

4.2.1 Requirement analysis 80

4.2.2 Conceptual design 83

4.2.3 Evaluation 87

4.3 Case study 88

4.4 Conclusion 89

5 Requirement analysis 91

5.1 Introduction 93

5.2 User profiles 93

5.3 Participants 96

5.4 Data collection 99

5.5 Analysis of the resulting data 103

5.5.1 Data 105

5.5.2 Analytical questions 108

5.5.3 Visualizations 109

5.5.4 Interaction functionalities 111

5.5.5 Summary 112

5.6 Data quality 113

5.7 Conclusion 114

6 Conceptual design 117

6.1 Introduction 119

6.2 User interface design 119

6.3 User experience design 124

6.4 Navigation panel 126

6.5 Focus view 126

6.5.1 Animation panel 129

6.5.2 Exploration panel 135

6.5.3 Timeline panel 136

6.6 Video and statistics panel 136

6.7 Geovisual analytics panel 139

6.8 Conclusion 144

7 Conclusion and discussion 147

7.1 Summary 149

7.2 Answering the research questions 151

7.3 Discussion 155

7.4 Recommendations 157

Appendices 159

Appendix 1: Formal interview request 159

Appendix 2: Attachment about the topic of study 160

Appendix 3: Introduction/instruction requirement analysis 161

Appendix 4: Requirement analysis interview questions 162

Appendix 5: Introduction/instructions first task-based session 163

Appendix 6: Introduction/instructions second task-based session 165

Appendix 7: Example of transcription of research materials 167

References 169

URLs 175

Curriculum Vitae 178

CH.1

Introduction

1.1. Research context

To understand the overall setting in which this research is situated it is important to provide an overview of developments in the fields of geo-information science and cartography with a focus on geovisual analytics. In this research, geovisual analytics is understood as an important analysis methodology when dealing with complex and multivariate datasets. This emerging interdisciplinary field emphasizes the importance of visual analytics, geovisualization and information visualization but focuses more on the geographic aspects dealing with spatial data (van Ho, 2013). According to MacEachren, geovisual analytics *"focuses on visual interfaces to analytical methods that support reasoning with and about geo-information to enable insights about something for which place matters"* (Kraak, 2014, p. 129). Geovisualization provides theory, methods and tools for the visual exploration, analysis, synthesis and presentation of geospatial data to stimulate visual thinking about patterns, relationships and trends (Kraak & Ormeling, 2010). It differs from cartography in the way that it uses an interactive interface for data exploration of complex datasets (Longley et al., 2011). Information visualization, which is part of computer science and strongly connected to computer graphics and cognitive psychology, has become the driving force behind the domain of visual analytics (Stein et al., 2017).

Similar to visual analytics, geovisual analytics is conceptually related to many different research areas such as data mining, geo-information science and computer science. In the domain of visual analytics, visualizations are combined with analytics to allow *"analytical reasoning facilitated by interactive visual interfaces"* (Thomas & Cook, 2005, p. 28). Data mining is the automatic or semi-automatic discovery of patterns in datasets, which are too large to analyze manually (Fayyad et al., 1996; Stein et al., 2017). The goal of visual analytics is to facilitate the analytical reasoning process through the creation of visual tools and interfaces that are capable to derive insight form large and unstructured datasets, provide understandable assessments and effectively communicate the results (van Ho, 2013). This combination enables experts to include their domain knowledge during the analysis process by interactive data mining methods and immediate visual feedback of the results.

1.2. Problem statement

Movement data is collected in large and constantly growing amounts. Many of the decisions made in the fields of science, society and business depend on proper knowledge about and a correct understanding of movement (Andrienko et al., 2013). However, current practice shows that despite this promise, the data often does not lead to insight, but to confusion. Without pre-processing the data, users may get overwhelmed by irrelevant or inappropriately processed and presented information. This problem is referred to as the information overload problem (Keim et al., 2010).

Regarding this problem, professional football organizations also have difficulties with the exploration and analysis of movement data and are laymen in the domain of geovisual analytics. Even though these clubs are in the position to collect large amounts of movement data; they do not know how to process and analyze these datasets themselves. Therefore, companies such as SciSports (URL 1) and Metrica Sports (URL 2) among others are currently filling this gab by helping football clubs to translate vast amounts of movement data into understandable information. When properly transformed, various visual tools are required to explore and analyze data in order to discover patterns. This research hopes to contribute to this development by proposing ways to make movement data more accessible for football clubs by developing a conceptual design for an interactive visual interface to the complex underlying data.

Another issue, is that currently available geovisual analytics tools to explore and analyze movement data are still weak in terms of their capability to deal with the temporal dynamics of geographic processes (Andrienko et al., 2010; Andrienko et al., 2011). In addition, existing tools mainly provide complex representation of the spatio-temporal phenomena that is not intuitive in dealing with large amounts of movement data (Dodge, 2015). These limitations are being considered as a significant hindrance. While there is currently an academic effort aimed at increasing the knowledge regarding these tools, there is still a lack in both literature and practice concerning the development of effective, efficient and satisfactory geovisual analytics tools. Therefore, this research also contributes to the resolution of this issue by suggesting a possible interface for more intuitive visual analytics capable of dealing with large amounts of data.

The third and final problem, which is at the heart of this research, is that most visual tools and interfaces are developed using a technology-centered or data-driven design approach (Roth, 2013). In these applications, the user does not hold a central part in the development and design process but only has a rather limited role as the evaluator of already developed solutions. In this way, the usability or quality of use of most products remains unsatisfying (Delikostidis, 2011). This research alleviates this problem by applying a User-Centered Design (UCD) approach.

1.3. Research objectives

The main objective of this research is *to develop a conceptual design for an interactive visual interface to movement data that helps football coaches, analysts and players to visually explore and analyze concerted movements.*

This approach is not only valid in the case of football, but, when used correctly, can also be applicable in similar domains, including movement ecology (animal tracking), transportation (vehicles, vessels and air traffic tracking) and health informatics (pedestrian movement). When accomplished, the conceptual design can be improved in further research and implemented on a broader scale.

To be able to realize the main objective stated above, this research is divided in several sub-objectives. These steps are necessary for the desired results and are formulated as follows:

1. To identify key concepts related to concerted movements and develop a conceptual framework for the visual depiction of the connections and notable relationships among these concepts.

2. To identify and evaluate currently available geovisual analytics tools related to concerted movements and their visualization.

3. To identify the context of use, relevant user requirements and specific analytical questions as starting points for the conceptual design process.

4. To develop a conceptual design of an interactive visual interface that can deal with concerted movements.

1.4. Research questions

Derived from the above discussion, the central research question that guides this research project is: *To what extent can the conceptual design for an interactive visual interface to movement data help football coaches, analysts and players to visually explore and analyze concerted movements?*

To answer this main question, several sub-questions need to be answered beforehand. The research questions (RQ) defined for this research are the following:

RQ 1: What are the relevant key concepts related to concerted movements which underpin the research project?

RQ 2: Which connections and notable relationships exist between the relevant key concepts?

RQ 3: How can an interactive visual interface support knowledge construction and insight into concerted movements?

RQ 4: Which geovisual analytics tools are currently available to visually explore and analyze concerted movements and how are they used?

RQ 5: What are the characteristics of end-users, tasks and context of use in which they operate?

RQ 6: What are the main requirements and analytical questions that end-users have when they visually explore and analyze concerted movements?

RQ 7: Which geovisual analytics tools are currently missing to meet the proposed requirements of end-users?

RQ 8: How can the conceptual design contribute to fill the gaps identified by end-users?

1.5. Research methodology

This research acknowledges that the perspective of users regarding the interface development process is often underestimated if not overlooked altogether (Delikostidis, 2011). Therefore, it adopts a User-Centered Design (UCD) methodology to produce an accessible solution for potential users. The UCD approach consists of three main phases with an associated set of research methods (Kveladze, 2015). In the framework of this research, first a requirement analysis is carried out to explore the user requirements and the analytical questions that potential users need support for. The objective of the second phase is to deliver a conceptual design of the visual interface capable of dealing with the concerted movements of football players. Finally, the evaluation phase describes how the prototype interface is tested based on its usability. However, due to time constraints this last stage of the UCD method could not be executed in this MSc research.

Understanding the behavior of objects and how they move through space over time within the complex nature of geographical and temporal spaces is a difficult challenge (Andrienko et al., 2007; Keim et al., 2008). Therefore, this research makes use of a case study, which concentrates on the coordinated positioning of defenders in space to prevent the opposing team from scoring a goal. Despite that there are comparable datasets available, this case study is chosen based on its unique characteristics. Its uniqueness is grounded in the complex way football players move through space over time in relation to their teammates, the adversary team and the ball.

1.6. Innovation aimed at

This research project is, in the context of geo-information science and cartography in general and geovisual analytics in particular, amongst the first to support football analytics by developing a conceptual design for an interactive visual interface to movement data. The novelty of this research also lies in the usage of the UCD approach to explore the analytical questions that domain experts need support for and the types of geovisual analytics tools that can be understood by them. Subsequently, this knowledge is transformed into a conceptual design of an interactive visual interface that may help professional football organizations to visually explore and analyze concerted movements.

1.7. Thesis structure

In total, this report is divided into seven chapters. In chapter 1, the topic of research is introduced and main objectives of study are provided. In chapter 2, the key concepts which underpin the research project are explored. This theoretical framework concludes with the conceptual model providing an overview of connections and notable relationships between the relevant key concepts. In chapter 3, the attention shifts from generic to specific by focusing on the visual analysis of concerted football movements. The methods used are presented in chapter 4. Chapter 5 and 6 contain a report on the implementation of the requirement analysis and the outcomes (see chapter 5) and the conceptual design that is based on these outcomes (see chapter 6). Finally, chapter 7 concludes by answering the research questions, gives a reflection on the methods used and provides recommendations for further research.

CH.2

Framework and conceptual model

2.1. Introduction

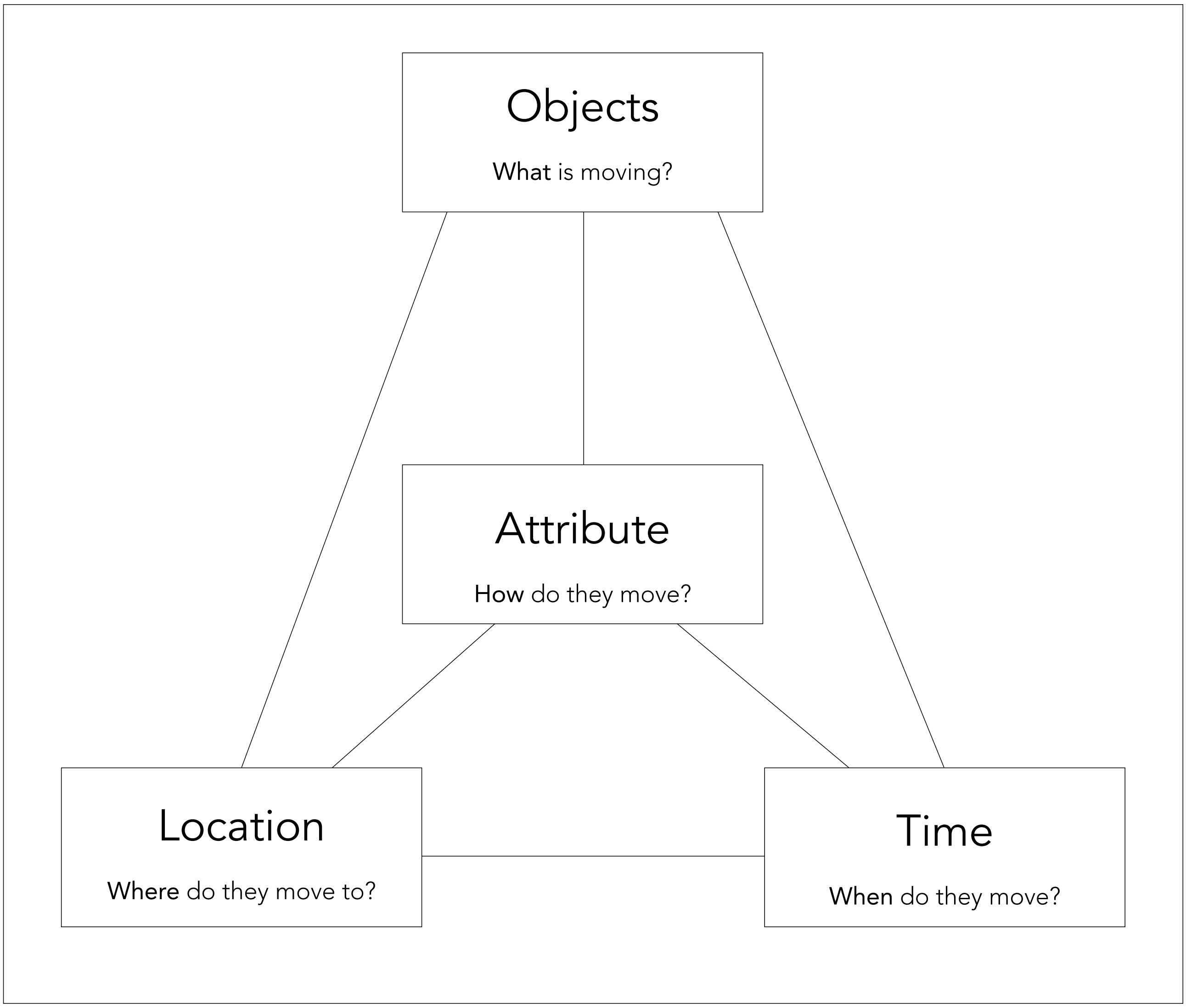
This chapter aims to outline and explore the key concepts which underpin this research project by presenting an overview of the relevant literature. These concepts, that form the framework of this research, are related to the concerted movements of objects. Together, they are used to address the research questions. In section 2.2, the foundation upon which this framework is built is discussed. Then, in section 2.3, the properties of concerted movements (i.e. trajectories, parameters and patterns) are discussed. Together, trajectories and parameters are considered as the underlying elements of movement patterns. In order to discover (by exploring, analyzing and synthesizing tracked data) and communicate these patterns geovisualization tools and geovisual analytics environments can be used. Therefore, section 2.4 further explores tools and interfaces that are needed to visualize movements. Section 2.5 summarizes which connections and notable relationships exist between the key concepts by using a conceptual model.

2.2. Fundamentals of movement

The study of movement can be considered as a young subfield of geo-information science (Dodge et al., 2016). At the center of this research domain are movement observations. These observations may be understood as spatiotemporal data (Dodge, 2015). Arising for these conceptualizations, this research defines movement data as *"spatiotemporal data describing changes of spatial positions of one or more moving objects"* (Andrienko et al., 2013, p. 380). Spatiotemporal data is a specific type of spatial data which refers to objects or phenomena with a specific location in space and linked by a certain timestamp. Spatiotemporal data refers to *"changes occurring in space over time"* (Andrienko et al., 2013, p. 382). According to Kraak (2014), these changes can be classified into three major types of change, namely: object-based (i.e. appear or disappear), attribute-based (i.e. increase or decrease) and location-based (i.e. expand, shrink or move). This research focuses on movement as change of location occurring in space and time. Movement is defined as *"the sequence of position records representing the movement of an object for the whole duration of the movement"* (Andrienko et al., 2013, p. 382).

Donna Peuquet (1994; 2002) proposes that there are three generic components in spatiotemporal data specific to movement: objects ('what?'), space ('where'?) and time ('when?'). Objects can be classified in accordance with their spatial and temporal properties (Andrienko et al., 2013). These objects can be abstracted from reality and stored in the database of a Geographic Information System (GIS) as entities (i.e. points, lines and polygons) or fields (Longley et al., 2011). The components that characterize these objects are: location, attribute and time (see figure 2.1).

Figure 2.1: The TRIAD framework, which shows the components of spatiotemporal data: location, attribute and time. Based on these components, one can ask elementary questions, like: 'what?', 'where?' and 'when?' or complex question, such as 'how long?', 'how often?', 'how fast?', or 'in what order?'. Source: Peuquet, 1994; Kraak & Ormeling, 2010



Temporal objects (i.e. spatial events) are objects with limited time of existence with respect to the period under observation. Spatial objects are objects capable to change their spatial position over time. Together, they are referred to as spatiotemporal objects. When more than one spatiotemporal object is involved, their collective movements are driven by the interactions between them and the spatiotemporal context in which they occur (Andrienko et al., 2013).

In most research projects, the factors that constrain the movements of objects are overlooked (Dodge et al., 2008). This ignores the complexity of movements of objects (or phenomena) in space as they relate to the environments in which they move and the other objects with which they interact (Dodge, 2015). Context has always been an important part of research in the domain of geo-information science. According to Buchin et al. (2014, p. 102), context refers to *"the locational circumstances of a moving agent"*. An alternative definition of context is given by Purves et al. (2014, p. 3), who defined context as *"the description of space within which movement occurs"*. This research deviates from both definitions by not only focusing on the spatial but also on the temporal context. Therefore, for the purpose of this research, spatiotemporal context is defined as: *the description of space and time within which movement occurs.*

Spatial context refers to the characteristics of the underlying landscape (e.g. characteristics of the terrain) and surrounding environment (e.g. properties of surface) in which movement takes place. In addition, the presence of other objects in the same space (e.g. interactions) and ambient attributes (e.g. atmospheric conditions) may affect movement (Kotzbek & Kainz, 2015; Andrienko et al., 2013). Together, these influence the spatial context. Temporal context refers to the temporal cycles (e.g. day or night, week or weekend) that may be relevant to the movements under study (Andrienko et al., 2013).

For instance, the challenge of analyzing football-specific spatial data is that all movements are restricted by a pitch and a set of rules, driven by the objective of the game, and influenced by the movements of own and opposing team players (Stein et al., 2017). Each game takes place in a restricted space of approximately 105 meters by 68 meters. According to Fédération Internationale de Football Association (FIFA), these are the standard international dimensions for a football field. (URL 3). In addition, there are several defined areas, like the penalty box and half way line, which are primarily relevant for the set of rules (Kotzbek & Kainz, 2014). Besides the field of play, other rules that influence the spatial context are: the number of players, the referee and assistant referees, ball in and out of play, offside, fouls, free kicks, penalty kicks, throw-ins, goal kicks and corner kicks. In this research, the referee and his or her assistants are not taken into consideration, as the level to which they influence the players' movements during the match is negligible. The temporal context is determined by two halftimes of 45 minutes each including a variable amount of additional time for time lost due to substitutions, injuries or other unforeseen events (Kotzbek & Kainz, 2014). The effect that the temporal context has on movement greatly depend on the position of objects within temporal cycles.

2.3. Concerted movements

This research uses the term concerted movement to indicate a specific type of movement. In this case, concerted movement is defined as: *the coordinated movements of multiple objects in relation to each other and the spatiotemporal context in which they move.* The underlying principles of this type of movement can be found in different fields of study, such as behavioral ecology, transportation, surveillance and athlete tracking (Demšar et al., 2015). Examples of concerted movements are particularly evident within the field that studies animal behavior. For instance, the way a flock of birds, a herd of sheep or a school of fish adapt their direction, speed and orientation to the position of the group or position of nearby predators (see figure 2.2).

Figure 2.2: Three examples of concerted movements of animals. Source: Google Images, 2017

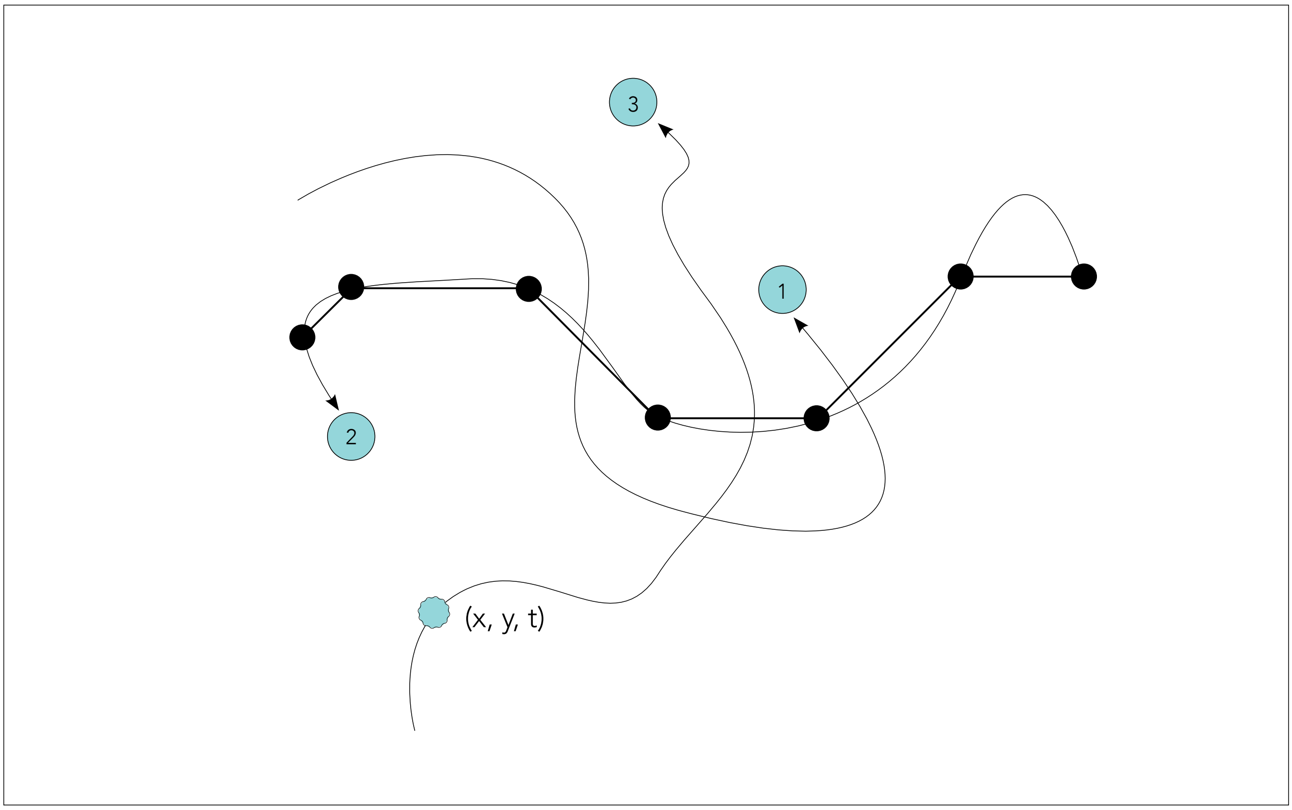
The use of these patterns does not imply that animals have understood the mathematical principles behind the usage of geometrical shapes; rather, they just follow a simple set of rules related to the movements of their neighbors (Sumpter, 2016). These simple rules adopted by animals provide a starting point for thinking about tactics and strategy in football. When interacting with their teammates, players must adopt the same rules as animals do. They must know when to accelerate, when to move into space, as well as how to use space and respond to their teammates. Apart from sports, this type of movement is extremely relevant to any sort of flow or network analysis, like traffic and logistics (Stein et al., 2016). Concerted movement occurs when multiple objects are moving through space over time in a collective manner. To understand 'when?', 'where?', 'why?' and 'how?' they are moving, the following sub-sections look at the traces, parameters and patterns which appear when an object moves.

2.3.1. Movement trajectories

When an object moves from one location to another, its path through space as a function of time is referred to as a trajectory (see figure 2.3). In search of a suitable definition, The New Oxford Dictionary of English (Stevenson & Lindberg, 2010) defines a trajectory as: *"The path followed by a projectile flying or an object moving under the action of given forces"*. To use this definition in this study, it is necessary to know: 'what are the forces that encourage an object to move?', 'why does an object moves?' and 'how does it move?'. To answer these questions, Dodge (2015) makes use of the movement ecology paradigm proposed by Nathan et al. (2008). According to this paradigm, the trajectory of a moving object is either driven by the intrinsic properties of the object or is influenced by the context. The internal state (e.g. hunger) of an object leads to a behavior (e.g. move). The state and resulting behavior influences the movement parameters (e.g. directions, speed or acceleration), which are highly influenced by the spatial context (e.g. type of terrain). These components are the forces that drive movement between two extreme positions, its origin and destination (Andrienko et al., 2013). Between their origin and destination, moving objects are influenced by various factors that impact and constrain their movements (Hägerstrand, 1970).

These paths through space and time may take different forms. The most commonly used differentiation is between continuous and discrete paths (Dodge et al., 2008). In most cases, when collected by tracking systems, trajectories are continuous paths. However, to become suitable for storage and analysis in an information system they first need to be divided into discrete parts prior to computing the movement parameters (Laube et al., 2005). In other words, discrete paths are a time-ordered sequence of steps (i.e. coordinates) between stops (Dodge, 2015). In this type of movement, the stops themselves may become more important than the movement between them (Dodge et al., 2008). This distinction is also relevant for this research, as it is not always interesting to know the exact trajectories of moving objects. Sometimes only the connection between origin and destination is valuable.

Figure 2.3: This figure shows a schematic representation of the trajectories of three moving objects. The arrows point towards the direction they are moving. Typically, moving objects are modeled as moving point objects and their location is fixed as tuples of (x, y, z). Furthermore, for data capture and implementation movement paths are typically discretized into time instants. Source: Labue et al., 2005



2.3.2. Movement parameters

Movement parameters are the derivatives of movement trajectories. They can be used to extrapolate the basic characteristics of a moving object at any given timestamp (Gudmundsson & Horton, 2016). Dodge et al. (2008) make a distinction between three major groups of parameters: primitive parameters, primary derivatives and secondary derivatives. Together, these are classified into spatial, temporal and spatiotemporal dimensions (see table 2.1).

Primitive parameters refer to the geographical coordinates of objects over time instance and time interval. From these parameters, more complex capacities of movements may be derived, such as direction, acceleration, path sinuosity and velocity. These primary and secondary derivatives consist of a combination of multiple parameters. For instance, when space and time are combined in the spatiotemporal dimension, the parameters speed and velocity are derived. Speed refers to the rate at which one or more objects move through space over time, whereas velocity refers to the speed of one or more objects moving in a certain direction (Dodge et al, 2008).

Table 2.1: Overview of movement parameters. Source: Dodge et al., 2008

Dimensions	Primitive	Primary derivatives	Secondary derivatives
Spatial	Position (x, y)	Distance	Spatial distribution
		Direction	Change of direction
		Spatial extent	Sinuosity
		Duration	Temporal distribution
Temporal	Instance	Travel time	Change of duration
	Interval	Speed	Acceleration
Spatiotemporal	-	Velocity	Approaching rate

2.3.3. Movement patterns

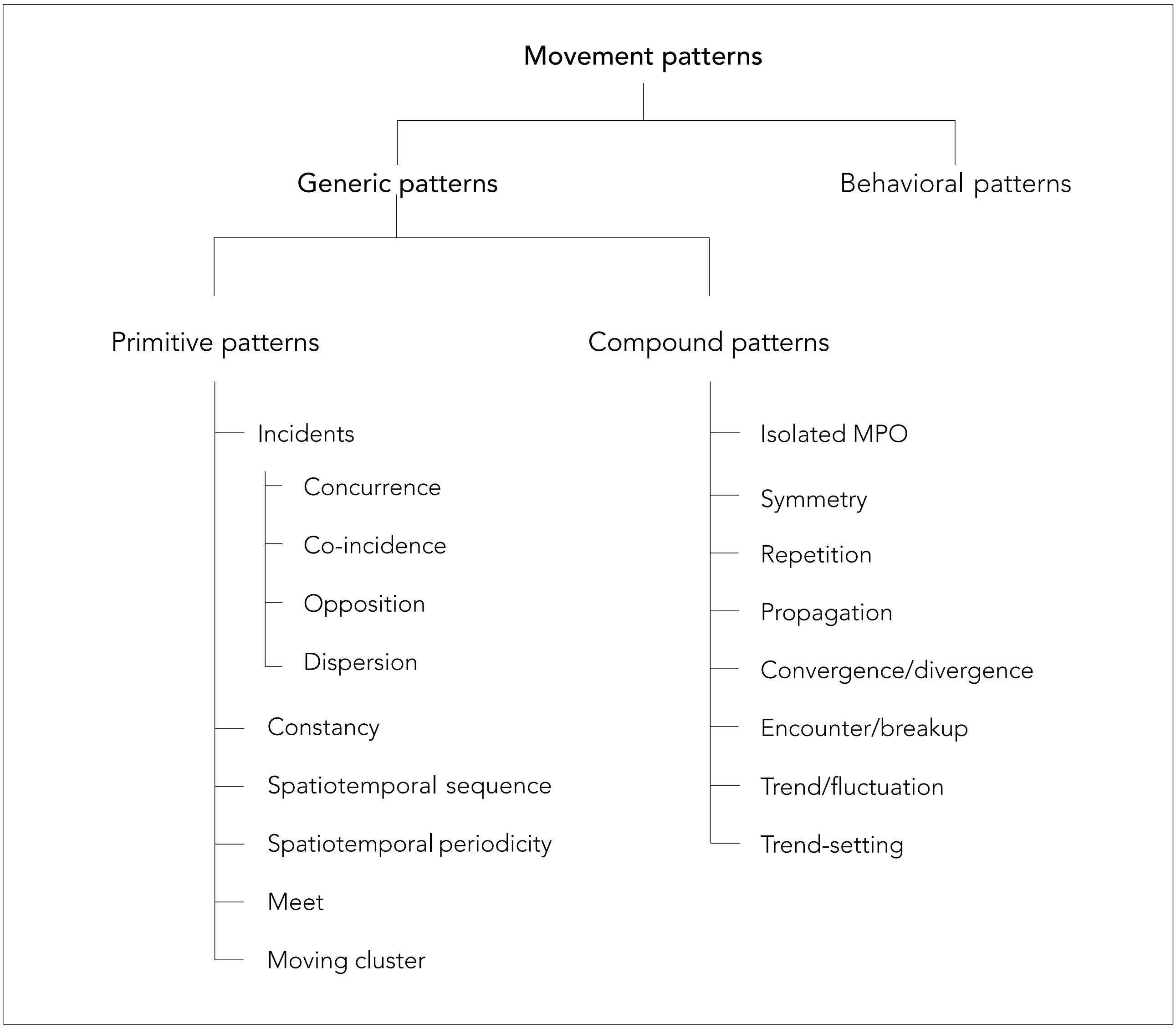
Movement patterns include any recognizable spatial and temporal trend or relationship in a set of movement data, whereas the proper definition depends on the field of research (Dodge, 2015). In this research, movement patterns are defined as *"concise descriptions of frequent behaviors, in terms of both space (i.e., the regions of space visited during movements) and time (i.e., the duration of movements)"* by Giannotti et al. (2007, p. 330). The number of involved objects is essential when dealing with movement patterns (Dodge et al., 2008). As mentioned, this research focuses on the concerted movements of moving objects. Spatiotemporal movement patterns need to reflect collective group actions or reactions to achieve a certain goal (Stein et al., 2017). When referring to multiple objects, there is a known differentiation between groups and cohorts (see table 2.2).

Table 2.2: The number of objects involved in a movement. Source: Dodge et al., 2008

Number of objects	Relationship	Examples
Individual	-	Trajectory of a person over a day
Group	Functional	Trajectories of a flock of sheep
Cohort	Statistical	Eye movement trajectories of female participants

Based upon the number of objects involved, Dodge et al. (2008) have developed a systematic classification of movement patterns relevant to the domain of geo-information science. With the main objective of study in mind, this research is not discussing all of them but only focuses on the spatiotemporal patterns related to concerted movements: concurrence, moving cluster and trend-setting (see figure 2.4).

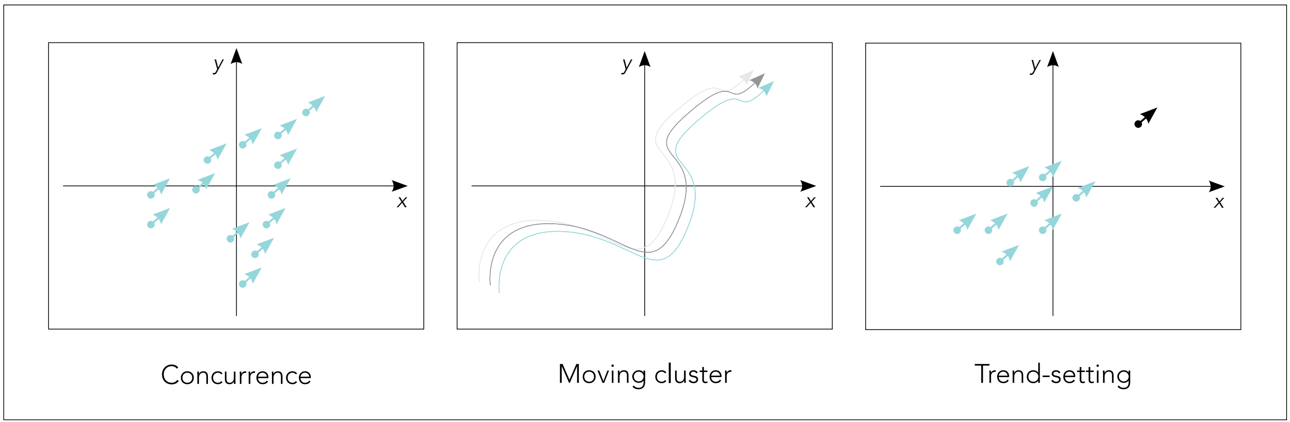
Figure 2.4: Classification of spatiotemporal movement patterns. Source: Dodge et al., 2008



The movement patterns concurrence and moving cluster fall within the group of generic primitive patterns (see figure 2.4). Generic patterns represent the building blocks used to form higher-level movement patterns. Primitive patterns are the most basic forms of movement patterns, those where only a single movement parameter varies. Concurrence is an incident (i.e. a pattern occurring among multiple objects) where a group of objects shows a coordinated movement in a certain direction for a certain duration of time (see figure 2.5). For instance, a flock of geese moving in the same direction or a group of football players chasing the ball. The term flock is often used referring to the moving cluster pattern. This pattern describes how objects within a group of objects moves in the same direction while staying near to each other (Benkert et al., 2008).

Figure 2.5: Three examples of spatiotemporal movement patterns: concurrence, moving cluster and trend-setting.

Source: Dodge et al., 2008



Another interesting pattern for this research is the trend-setting pattern, which is a generic compound pattern (see figure 2.5). Compound patterns are made up of several primitive parameters involving complex relations. This pattern refers to objects that influence the movement of others. In trend-setting, objects move in the same direction or might have other similar movement characteristics such as speed (Dodge et al., 2008). There are two variations of trend-setting: non-varying and varying. In the case of non-varying trend-setting, the number of followers does not change. In contrast, the number of followers may be changing from one-time interval to the next when referring to the varying trend-setting patterns.

In this section, an overview of the movement properties (i.e. trajectories, parameters and patterns) has been provided. Together, trajectories and parameters are considered as the underlying elements of movement patterns. In order to discover (by exploring, analyzing and synthesizing tracked data) and communicate these patterns geovisualization tools can be used. Therefore, the next section further explores tools and interfaces that are needed to visualize movements.

2.4. Geovisual analytics of movements

This chapter has provided an overview of key concepts related to the study of movement. In this section, the tools and interfaces needed to visualize movements are discussed. In sub-section 2.4.1, the basic elements of an interactive visual interface are given. Thereafter, sub-section 2.4.2 provides examples of currently available geovisual analytics tools related to the visual exploration and analysis of movement data. While there currently are quite some academic efforts aimed at increasing knowledge regarding these tools, there is still a lack in both literature and practice concerning the development of effective, efficient and satisfactory geovisual analytics tools. Therefore, sub-section 2.4.3 concludes with a discussion about usability issues.

> *''Visualization tools have the potential to provide 'windows' into the complexity of the phenomena involved"*
>
> MacEachren & Kraak, 2001, p.2

2.4.1. Interactive visual interfaces

In this research, the definition by Wassink et al. (2008, p. 2) is used to define an interactive visual interface as *"a system that not only provides different views on data, but also enables dialogues to explore and to modify the data"*. These systems are used for specific tasks and specific user groups. Therefore, it is essential to analyze what kind of visualisation techniques should be used to support the tasks at hand and what types of interaction techniques best fit the targeted users.

In a geovisual analytics environment an interactive relationship between humans and machines is achieved by means of an interactive visual interface (Thomas & Cook, 2005; Roth & MacEachren, 2014). Interactive visual interfaces support the human cognitive process by allowing analysts to look at a subject from different perspectives and at different scales and levels of detail, to link diverse pieces of information, as well as to direct and control the work of computational analytical tools. In this respect, interaction is defined as *"the dialog between a human and map, mediated through a computing device"* (Roth, 2012, p. 377). This definition acknowledges the three crucial components of interaction: human (i.e. user-centered perspective), interface (i.e. interface-centered perspective) and computing device (i.e. technology-centered perspective).

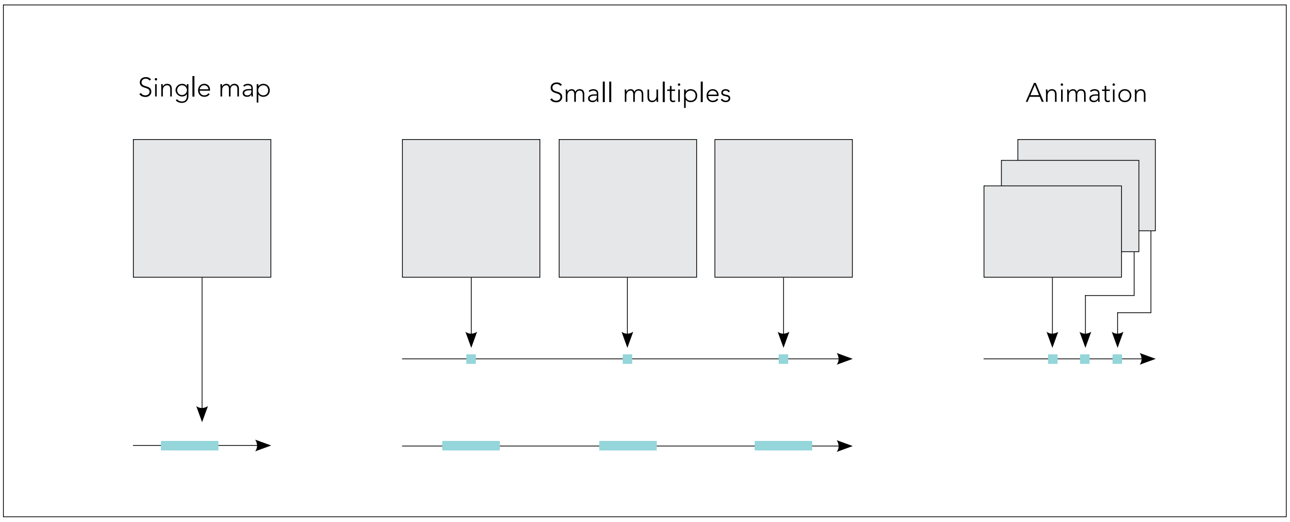
Each interface has three distinct characteristics which determine its success: the visual interaction it supports, its interface style and its interface design (Roth, 2013). Regarding the first characteristic, the stages of (inter)action model by Norman (1988) provides a framework based on the three O's of interaction. These are the objective-based (i.e. identify and compare), operator-based (i.e. brushing, focusing and linking) and operant-based approach (i.e. space, time and attribute primitives). Especially, the operator-based actions linking and brushing are of importance when developing an interface capable of dealing with large volumes of movement data. These techniques highlight selected observations across various visualizations by providing a dynamic link between maps and other types of graphical displays to communicate different aspects of the data.

2.4.2. Geovisual analytics tools

Geovisual analytics is an interdisciplinary field which integrates visual analytics with information and scientific visualization and introduces new techniques for knowledge discovery (Kveladze, 2015). In geovisual analytics environments various visual representations are used to stimulate visual thinking about spatiotemporal movement patterns. This is realized by providing several analytical tools to support the exploration of the data in several alternative ways that contain graphics of the same or related data (Kveladze, 2015). These visualizations allow reasoning among different tasks and user groups involved. Visualizations that take the form of maps and map-related graphics or diagrams are called geovisual analytics tools. These maps are used to visualize spatial data, *"that is data that refer to the location or the attributes of objects or phenomena located on Earth"* (Kraak & Ormeling, 2010, p. 1).

With the use of these tools, concerted movements can be visualized in multiple ways, using, for example, a single map, small multiples or animated maps (see figure 2.6). These three temporal cartographic depiction models are traditional ways to visualize geographical data with a temporal component (Kraak & Ormeling, 2010). Other more complex geovisual analytics tools like self-organizing maps, origin-destination (OD) matrices, parallel coordinate plots, treemaps and node-link diagrams are not taken into consideration, because it is assumed that they are difficult to understand. This research limits itself, therefore, by focusing on the three traditional ways of visualizing geographical data with a temporal component. It is more likely that these more traditional solutions will be better understood by the potential end-users.

Figure 2.6: Mapping moments and time intervals in a single map, small multiples and animation. Source: Kraak, 2014



The single map represents a snapshot about the spatial distribution of objects in space and time in which symbols and visual variables may be used to give the impression of change. Single maps are however quite limited when visualizing vast amounts of spatiotemporal data (see figure 2.7). Small multiples use multiple static maps, graphs or charts (in chronological order) to differentiate between two or more snapshots of the same place during different moments. These series of static maps are limited in their use because once the amount of time displayed becomes too large, the images become too complicated (see figure 2.8). Animated maps or animations depict change dynamically using a set of single maps (Kraak, 2014). Interactive animations refer to animations in which the user is in control. For instance, users control the speed in which images are displayed or they may pause and play the animation.

Figure 2.7: This single map originally published by Charles Joseph Minard in 1862 shows migration flows in 1858. This map is an early example of a sankey diagram, which displays information as flows. Source: Google Images & Edward Tufte, 1983

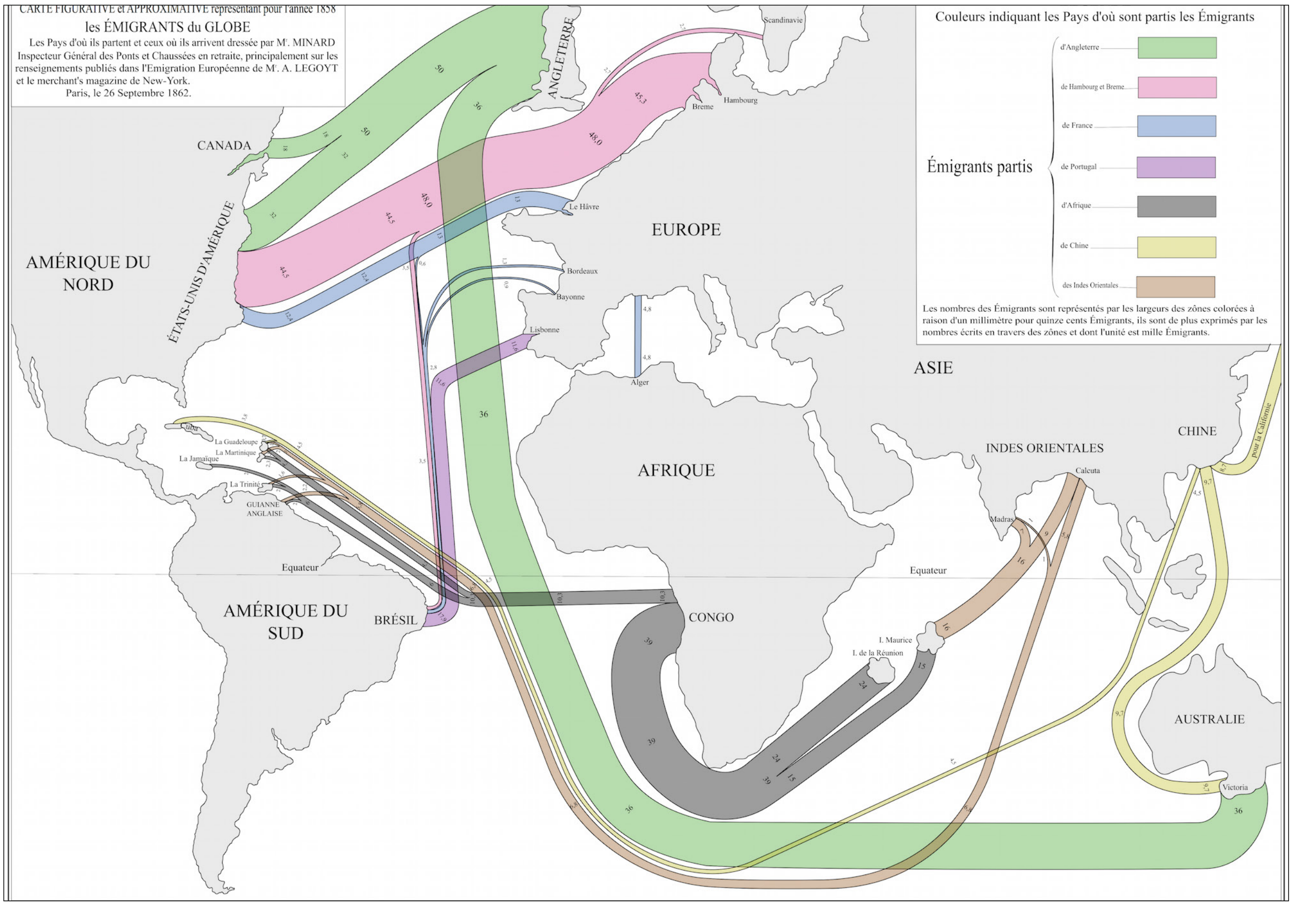
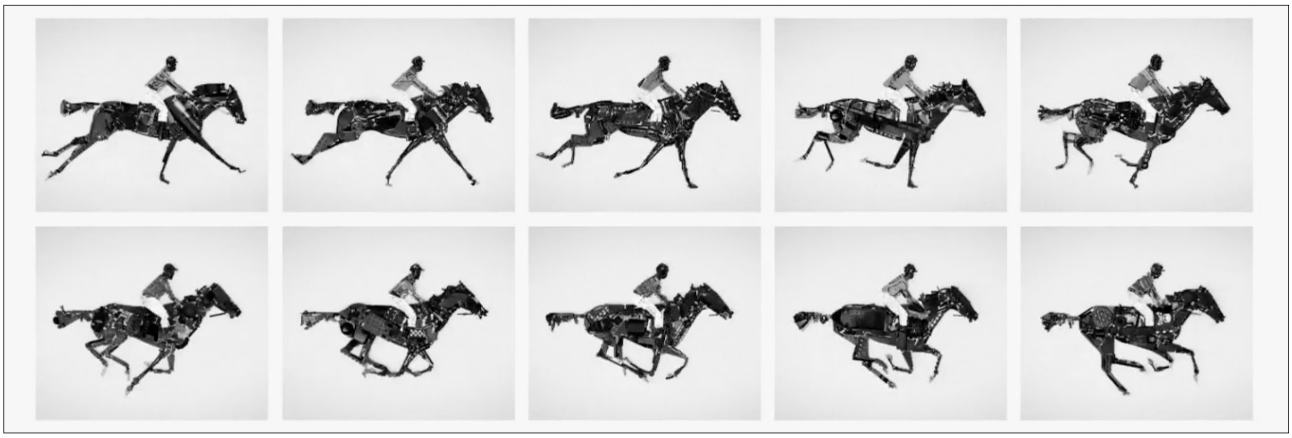


Figure 2.8: This illustration shows a modern representation of the original small multiple *Horse in Motio*n made by Eadweard Muybridge in 1886. The photos show the course traversed by the feed of the mare Sallie Gardner during a single complete stride. Source: Google Images & Edward Tufte, 1983

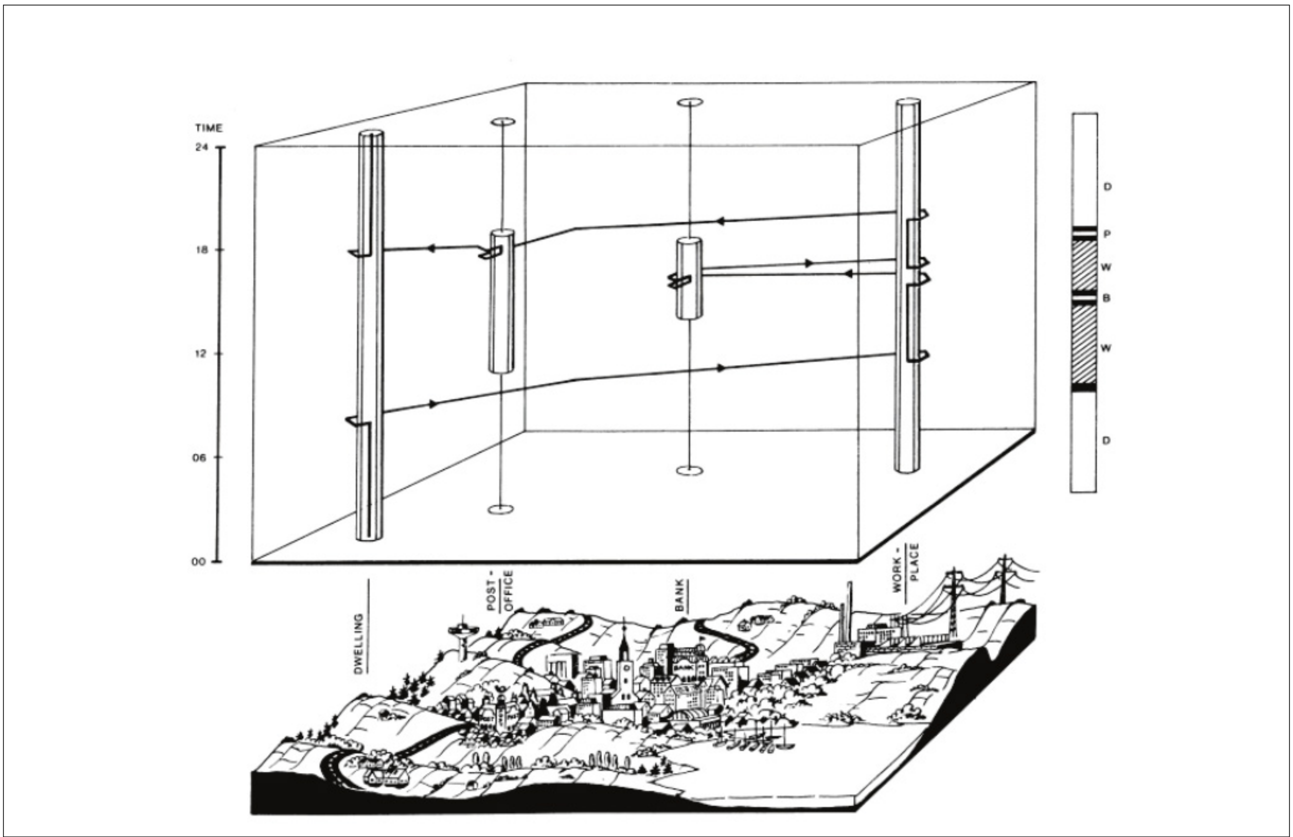


The most common type of display used to visualize spatial and spatiotemporal data is the cartographic map. As football is a spatial phenomenon it is obvious to visualize its analysis outcomes in the form of a map (Kotzbek & Kainz, 2014). This tool can visualize the spatial structure, the positions of moving objects, the trajectories of movement and flows between different locations. In this case, flow refers to the aggregated movements between locations. One possible way to aggregate trajectories into flows is by making use of flow maps (Andrienko & Andrienko, 2012). These maps make use of straight or curved line symbols to represent trajectories of moving objects (see figure 2.9). Usually, arrowheads are used to indicate their direction across a certain environment. Other visual variables like color or shadow are used when one would like to express the type of movement. The width of the arrow is used to express the amount or volume of movement (Kraak, 2014).

Figure 2.9: Based on a recent United Nations Refugees report, the New York Times published a flow map visualization, mapping the 14 million refugees who fled their country in 2014. Source: Pecahna & Wallace, 2015

Another tool capable of visualizing the trajectories of moving objects is the space-time cube. Its ability to consider space and time as two inseparable concepts is its main strength (Hägerstrand, 1970). The two-dimensional horizontal plane represents the location of an object in space and the vertical plane represents its location in time (see figure 2.10). Together, they are used to study the movements of objects in a three-dimensional cube (Kveladze, 2015). Within this cube, trajectories of moving objects between stops are visualized. However, due to the increasing availability of movement data, alternatives to the space-time cubes have been developed (Kveladze, 2015). For instance, Zhao et al. (2008) who proposed the usage of ringmaps or Demšar and Virrantaus (2010), who introduced the space–time density of trajectories. Furthermore, Tominski et al. (2012) and Andrienko et al. (2014) worked on an alternative tool called the trajectory wall. Additionally, temporal bar charts, time bars or radial histograms may be added to show temporal variation of values of a dynamic attribute within trajectories.

Figure 2.10: An early example of the Space-Time Cube. Source: Carlstein, et al., 1978 from Kveladze, 2015.



2.4.3. Usability issues

In this research, the definition by the International Organization for Standardization (ISO) is used which defines usability as: *"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"* (ISO 9241-11:1998 as cited in Kraak & Ormeling, 2010, p. 187). This definition points to the three main aspects of usability: effectiveness, efficiency and user satisfaction.

One of the first studies conducted regarding the usability of geovisualization tools, was the experiment by Koussoulakou and Kraak (1992). They investigated the capabilities of animated maps for analysis in comparison to static maps dealing with spatiotemporal data (Andrienko et al., 2002). In another experimental study, Harrower et al. (2000) assessed interactive devices for controlling map animations by testing the temporal focusing and brushing techniques suggested by Monmonier (1990). The experiment showed that most target users did not understand how to use the temporal focusing tool, and while almost all users understood the purpose of temporal brushing, very few of them actually used it. According to Andrienko et al (2002), the same is true for many other geovisualization tools, such as the ones they developed in their geovisual analytics environment CommonGIS. The novelty of such tools requires substantial training to master them and to ensure their usability. After such training, it appeared that users could use the tools more effectively (Andrienko et al., 2002).

Prior to these studies, literature on the use of geovisualization tools in general and the usability of geovisual analytics environments in particular is rather sparse (Andrienko et al., 2002). However, since the beginning of this century user research in the domain of geo-information science is thriving (URL 4). For instance, Balciunas (2013) who did a usability assessment of internet maps, Schobesberger (2012) who proposed a framework for improving the usability of web-mapping products or Opach & Nossum (2011) who evaluated the usability of cartographic animations. Even more recently, Kveladze (2015) developed a methodological framework to define the usability of the space-time cube (STC). Despite the increasing popularity of the STC, she found that usability research remains limited. Based on a UCD perspective, she used multiple qualitative and quantitative user research methods to evaluate the usability of the STC. For these experiments, two groups of users, domain experts and non-domain experts, had to execute several map use tasks to answer specific spatiotemporal questions. A comparison between both groups showed that domain experts could operate the geovisual analytical environment more effectively and efficiently due to their interest to explore their data.

So, until now, several geovisualization tools and geovisual analytics environments enabling user interaction with maps and other types of graphical displays have been developed. However, these tools and environments can only support information exploration and knowledge construction when users are able to employ them properly (Slocum et al., 2001). The involvement of the user and a successful identification of their needs determine whether tools and interfaces will become successfully used (Balciunas, 2013). Due to this reason, user research becomes the underlying trend of modern geo-information science researches and the usability of the visualizations becomes one of the basic criterion of this evaluation (Balciunas, 2013). Therefore, it also seems that there is a sincere need for more usability research in the domain of geo-information sciences dealing with large volumes of movement data.

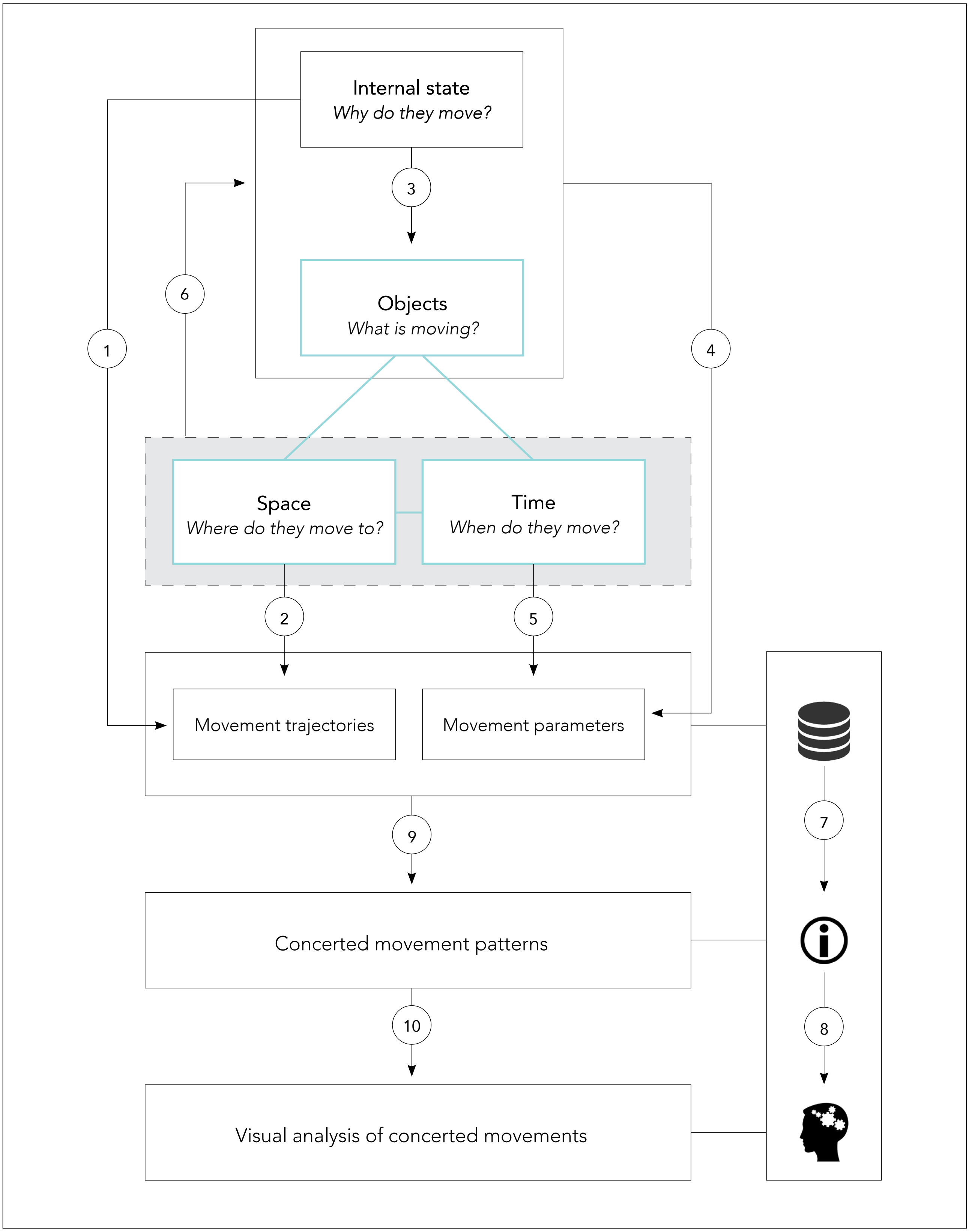
2.5. Conceptual model

This section provides a summary of the connections and notable relationships that exist between the key concepts discussed in this chapter. Together, these are visualized in a conceptual model (see figure 2.11). The TRIAD framework, which shows the generic components in spatiotemporal data specific to movement, is used to understand where and when objects move through space over time (see figure 2.1). In the conceptual model, this framework is highlighted by using a blue stroke color. When more objects are involved, their collective movements are driven by the interactions between them and the spatiotemporal context in which they occur. In the conceptual model, the combination of space and time which forms this context is visualized using a gray box.

Based upon this foundation, this research subsequently looked at the traces which appear when multiple objects move through space over time. These traces are represented in the model as movement trajectories. The conceptual model shows that the trajectory of a moving object is either driven by its internal state (1) or is influenced by its context (2). In addition, the internal state, which leads to a particular behavior (3), determines the characteristics and capacities of movement (4). Again, these movement parameters are highly influenced by the context (5). These components are the forces that drive movement between two extreme positions, its origin and destination. Between origin and destination, moving objects are influenced by various factors that impact and constrain their movements (6).

Together, trajectories and parameters may be considered as the fundamental elements of movement. Due to the recent advancements in tracking systems, increasing amounts of movement data is collected. These vast amounts of data pose a challenge on how to transform the data into information (7) and ultimately into knowledge (8). Using computational and geovisualization techniques raw movement observations can be transformed into movement patterns which include any recognizable spatial and temporal trend or relationship in a dataset (9). When properly transformed, geovisualization tools, possibly embedded in a geovisual analytics environment are potentially very powerful to explore, analyze and communicate these patterns to extract relevant information from this spatiotemporal data (10).

Figure 2.11: Conceptual model of the connections and notable relationships that exist between the key concepts related to the concerted movements of objects through space over time.



2.6. Conclusion

This chapter has identified the key concepts and developed a conceptual model for the visual depiction of the connections between these concepts and notable relationships among them (see figure 2.11). Together, the concepts are used to address the research questions. Movement data is inherently spatiotemporal. This type of data focuses on the changes occurring in space over time. The objects that change their location are classified in accordance with their spatial and temporal properties. The type of movements under investigation is referred to as concerted movement and indicates a specific type of movement defined as: *the coordinated movements of multiple objects in relation to each other and the spatiotemporal context in which they move.*

To understand concerted movements, this research looks at traces and patterns that occurs when an object decides to move from one location to another under the influences of its internal state and context. Based on these positional records, complex capacities of movement may be derived, such as distance, acceleration and velocity. These are known as movement parameters. Together, trajectories and parameters are considered as the underlying elements of movement patterns. To communicate the information extracted and stimulate visual thinking about these concerted movement patterns various visual tools and interfaces are required. In the geo-visual analytics environment, this is realized by viewing the data in several alternative ways that contain graphics of the same or related data. When combined with analytics, these visual analytics tools can be used to display concerted movement patterns in multiple ways. However, user research has showed that these tools and environments can only support information exploration and knowledge construction when users are able to employ them properly.

Based upon the scientific foundation outlined in this chapter, the next chapter specifically deals with the visual analysis of concerted movements based on movement data collected by different tracking systems.

CH.3

Visual analysis of football movements

3.1. Introduction

This chapter specifically deals with the collection, analysis and visualization of movement data in football. Section 3.2 describes how movement data is collected using different tracking systems. Section 3.3 describes how this data is used in the field of football analytics. Finally, in section 3.4, several geovisual analytics systems specifically developed for football are discussed.

3.2. Movement observations

It all starts with the collection of movement data. Nowadays there are numerous companies which provide movement data, for instance: Wyscout (URL 5), InStat (URL 6), Ortec Sports (URL 7), Opta Sports (URL 8), Scout7 (URL 9), STATS (URL 10) and Gracenote (URL 11). These data providers are using different optical and device enabled tracking systems, or manually annotated events from video recordings (Kotzbek & Kainz, 2016). Movement observations collected by tracking systems contain the positions of players and the ball and may be enriched by adding event data, such as the number of shots, fouls and passes (Gudmundsson & Horton, 2016). Occasionally, data representing the match officials is also provided (Kotzbek & Kainz, 2015). These complementary types of football-specific spatial data are distinguished as tracking and event data (Kotzbek & Kainz, 2014; 2016).

Right now, most football analysts and coaches focus on descriptive data about the physical capacities of individual players, because these data are easily accessible and interpretable to them. This must be considered as insufficient, because the only information provided is 'what?' and 'how often (it occurred)?'. Since football is all about the union of space and time, football analytics cannot disregard these two essential components any longer (Kotzbek & Kainz, 2015). Therefore, this research focuses on informing analysts and coaches about the characteristics of the concerted movements of players through space over time, information that can be transferred less easily, because of its inherent complexity. Therefore, the remainder of this section focuses on tracking data which contain the characteristics of concerted movements. Optical (i.e. passive) tracking systems make use of several fixed (high speed) cameras to capture player and ball movements with a high temporal and spatial resolution. They range from television recording to recordings carried out by the clubs themselves. Private companies such as Opta Spots and STATS extract movements, events and statistic data based on their own recordings (Stein et al., 2017). Another option is using wearable device (i.e. active) tracking systems directly attached to players or game objects. These systems consist of two distinct parts of technology: location technology and inertial sensors. Location technology refers to outdoor positioning technologies such as the Global Positioning System (GPS) or the Global Navigation Satellite System (GNSS). The other half of the technology solutions are inertial sensors, such as accelerometers, magnetometers and gyroscopes that capture detailed information about the physical capabilities of players (URL 12).

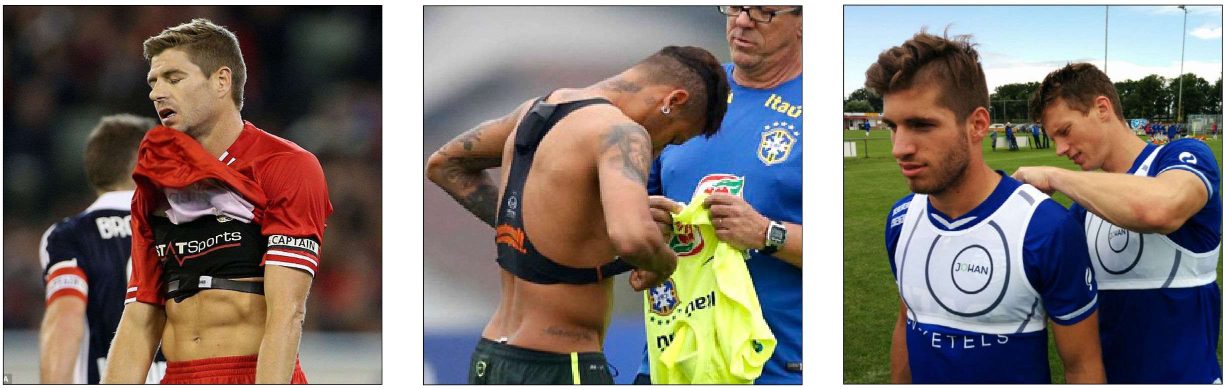
In the United States, they are already experimenting with different data collection system for several years. Since 2009, the National Basketball Association (NBA) has been a leader in this area by installing the optical tracking system SportVU in every arena (see figure 3.1). This system will be replaced at the beginning of the 2017-18 season by Second Spectrum and Sportrader to enhance NBA player tracking. Second Spectrum becomes the NBA's official optical tracking provider (URL 14). Meanwhile, Sportrader's rader360 tool makes it possible to collect data and customized data views to analyze different situations and make all data available to fans at home (URL 15). The National Hockey League (NHL), has been experimenting with chips directly embedded into pucks and jerseys since the 2015-16 season (URL 16) and the National Football League (NFL) recently adopted Radio-Frequency Identification (RFID) chips (URL 17).

In the largest and wealthiest football competitions, the English Barclays Premier League, the German Bundesliga and the Spanish La Liga, they make use of the optical tracking system TRACAB developed by ChyronHego (URL 18). The TRACAB system uses advanced image-processing technologies to determine the position of all moving objects on the pitch at 25 times per second (i.e. temporal resolution) which leads to a vast amount of spatiotemporal data (see figure 3.1). Another, quite similar system, called BallJames is currently being developed at the Dutch company SciSports (URL 19). BallJames is a fully-automated tracking system which makes use of fourteen fixed cameras. This data-engine captures accurate x, y and z coordinates of the players, the ball and the referees in real-time (see figure 3.1).

Furthermore, there are wearable systems on the market (see figure 3.2). Recently the FIFA decided to allow the usage of sensor technologies to track player positions during competitions (Stein et al., 2016). In practice, most of these wearable systems are combined with a Polar chest strap for heart rate sensing (URL 20). For example, the Viper Pod developed by STATSports (see figure 3.2). This system is regarded as the leading player performance analysis system capable of tracking speed, distance and acceleration among other parameters (URL 21). Another widely used system is called the OptimEye S5 developed by Catapult (URL 22). In addition, there is the device made by JOHAN sports which can measure location, distance and orientation as well as other variables (URL 23). A final example is the Inmotio system which uses (wireless/mobile) Location Position Measurement (LPM) technology, drone cameras and ball tracking technology based on advanced RFID technology. Besides generating real-time performance data, LPM is known for its stable and extremely accurate measurements (URL 24).

Figure 3.1: The SportVU (left), TRACAB (middle) and BallJames (right) optical tracking systems. **Sources**: URL 13; 18; URL 19

Figure 3.2: The Vipor Pod (left), OptimEye (middle) and Johan (right) wearable tracking systems in action. **Sources**: URL 21; 22 URL 23



The nature of the spatiotemporal data resulting from these tracking systems can be characterized using the so-called three V's of big-data: volume, variety and velocity (Noor et al. 2015). First, volume refers to the size of datasets. Currently, most tracking data is encoded using Extensible Markup Language (XML) and ranges in size between 86 and 300 megabytes for one match. The most common solution to provide structured access to the data is by using comma-separated values (CSV) files. However, these do not scale well to the specific nature of spatiotemporal data (Rein & Memmert, 2016). Second, variety refers to different data formats and data sources, for instance: tracking, physical, video, event and crowd data. These sources can be further distinguished into: structured, semi-structured and unstructured data. Video data is a typical example of unstructured data, whereas the current XML-files used for tracking data are considered as semi-structured (Rein & Memmert, 2016). Finally, velocity describes the speed with which the data is being generated. In the case of football the velocity varies widely between real-time streams of tracking data to manually annotated event data from video (Kotzbek & Kainz, 2016).

Tracking data corresponds to the paths through space linked with a certain timestamp and is provided as an attribute to the spatial data. With the use of this information, speed is calculated at each single point and provided as an attribute. Furthermore, the z-coordinate is also provided for the ball's path to reconstruct its trajectory (Kotzbek & Kainz, 2016). Tracking data can be considered as vector data and is spatially represented as point and line features (Longley et al., 2011). Each point feature is defined by a single pair of x, y coordinates which determines the position of the players and the ball. Subsequently, line and polygon features can be created based on the existing point data. With the use of a local coordinate system a framework is provided to accurately represent the real-world locations of the players and the ball. In a Geographical Information System (GIS) every dataset has its own coordinate system, which is used to integrate it with other spatial data layers within a common coordinate framework such as a map. In the case of football, the coordinate system used for tracking data refers to the international standardized football field. The origin of the local coordinate system is positioned at the field's center (Kotzbek & Kainz, 2016). However, the coordinate system that underlies the data can vary between data providers (Kotzbek & Kainz, 2014).

3.3. Football analytics

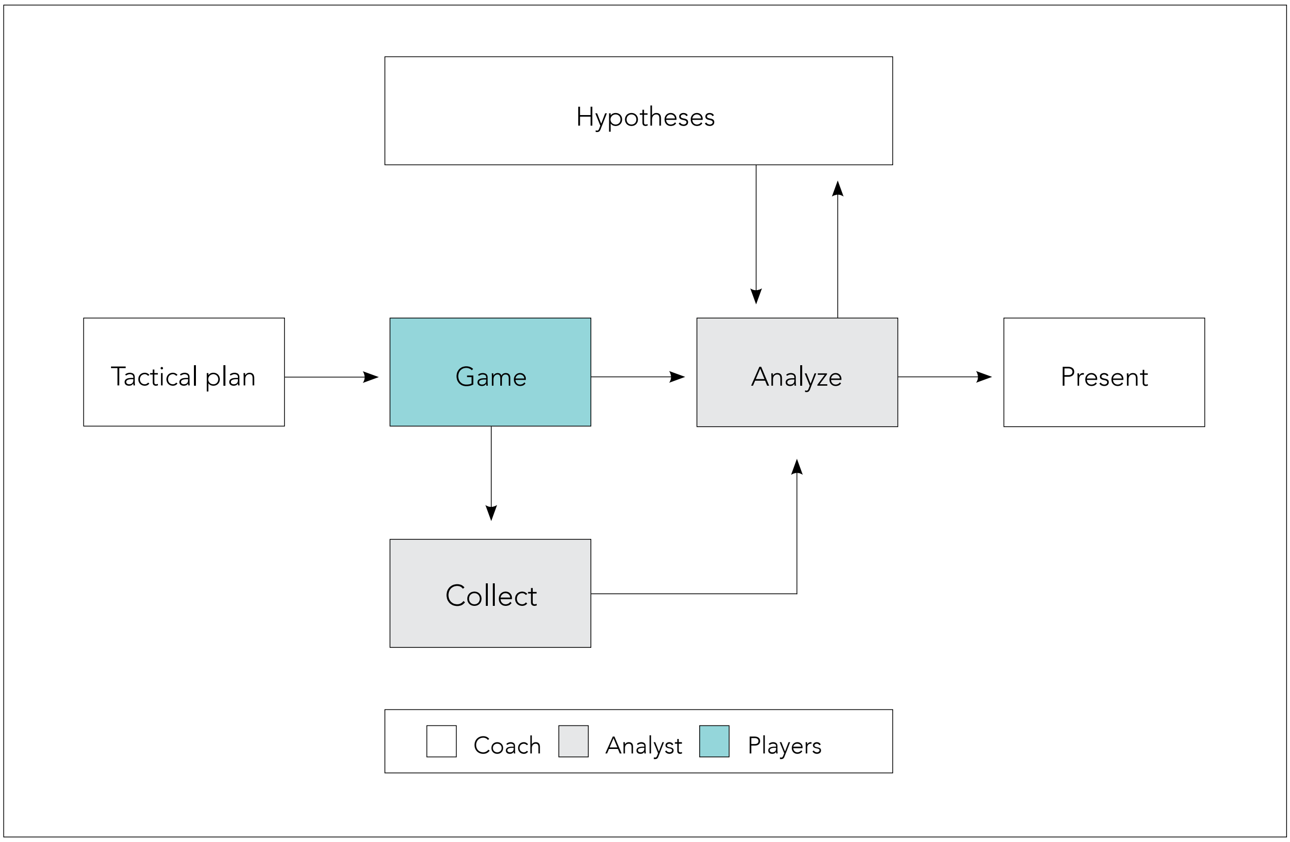
In 1977, Bill James revolutionized baseball analysis with a self-published booklet called the Baseball Abstract. His approach inspired others and soon they were referred to as sabermetrics (URL 25). In the late 1990s, the Oakland Athletics led by general manager Billy Beane began applying the sabermetric principles. In 2003, journalist Michael Lewis wrote the best-selling book Moneyball about their success (Lewis, 2003). Moneyball changed the way people thought about the use of data analytics in sports and the possible competitive advantages during the game, training, preparation or recruitment (Gudmundsson & Horton, 2016). Since then, researchers have attempted to copy the success of this approach in other sports. However, in many sports, the capacity to collect data has not been matched by an ability to process it in meaningful ways (MIT Technology Review, 2016).

The use of statistics that led to James' success is mainly applicable to static sports, such as baseball and cricket. These types sports move from one state to another and are primarily a series of one-to-one contests. On the other hand, team-based invasion sports such as basketball, ice-hockey and football consist of much more complex interactions between components involved in the game. In these sports, two teams are competing for possession of the ball or the puck in a constrained playing area, for a given period of time. Each team has the same objectives: scoring by putting the ball into the opposition's goal and defending their goal against attacks by the opposition. The team that has scored the most goals at the end of the game wins.

Until recently, most analysts and coaches used video recordings of past matches and training sessions that were manually processed, annotated and edited for analysis and presentation purposes (Stein et al., 2016). Not only are these recordings used to evaluate individual player performance and teamwork, but also tactics and scouting are potential fields of application (Kotzbek & Kainz, 2015). Although post-match analysis is crucial for most teams to maintain their competitive advantage, there are some known limitations. The main limitation of this traditional approach is that almost all contextual information is discarded (Rein & Memmert, 2016). Furthermore, these qualitative game observations are less objective, depending on the know-how of domain experts (Memmert et al., 2016).

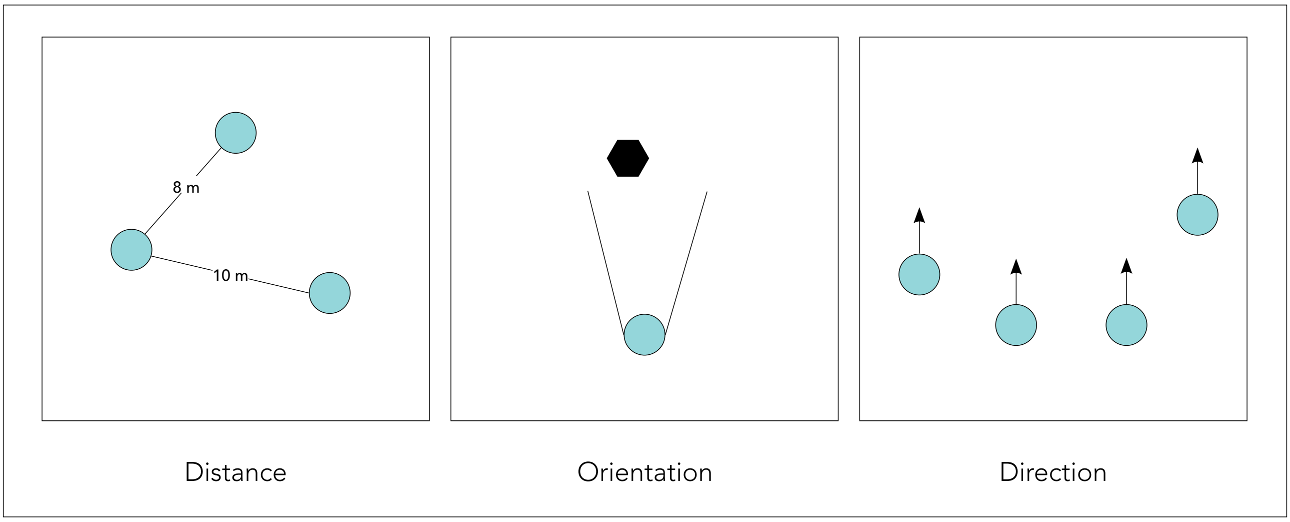
In practice, the work of the video analyst is often driven by the hypotheses of the coaching staff (see figure 3.3). Usually, the results of this process are transferred from the analysist to the coach who then communicates with the players, normally using a video presentation and sometimes with additional data visualizations (Stein et al., 2016). However, due to recent advancements of tracking systems, football clubs now have access to an increasing number of data points. For a competitive match, tracking all players and the ball in three-dimensional coordinates over 95 minutes, produces over 200 million point features (Sumpter, 2016). The availability of movement data has facilitated a variety of research efforts, across multiple disciplines, to extract spatiotemporal movement patterns from tracking data (Gudmundsson & Horton, 2016). Unfortunately, conventional methods of post-game analyses often fail to analyze and visualize the complex spatiotemporal patterns sufficiently (Kotzbek & Kainz, 2015).

Figure 3.3: Typical analysis workflow at professional football clubs. **Source**: Stein et al., 2016



While physical parameters such as distance covered, the number of high intensity sprints made or the average speed have been automatically analyzed for some years now, there are only few approaches doing this with tactical movement parameters in team sports (Memmert et al., 2016). Examples of these parameters are: the spatial distribution of players, the mutual distances between teammates, the opposition and the ball, orientation and direction of movement (see figure 3.4). These parameters describe the organizational principles of the players on the pitch and give an indication of how teams manage space, time and individual actions to win a game. This classification can be further structured depending on the number of participating players into: single player, line, team and match tactics (Rein & Memmert, 2016).

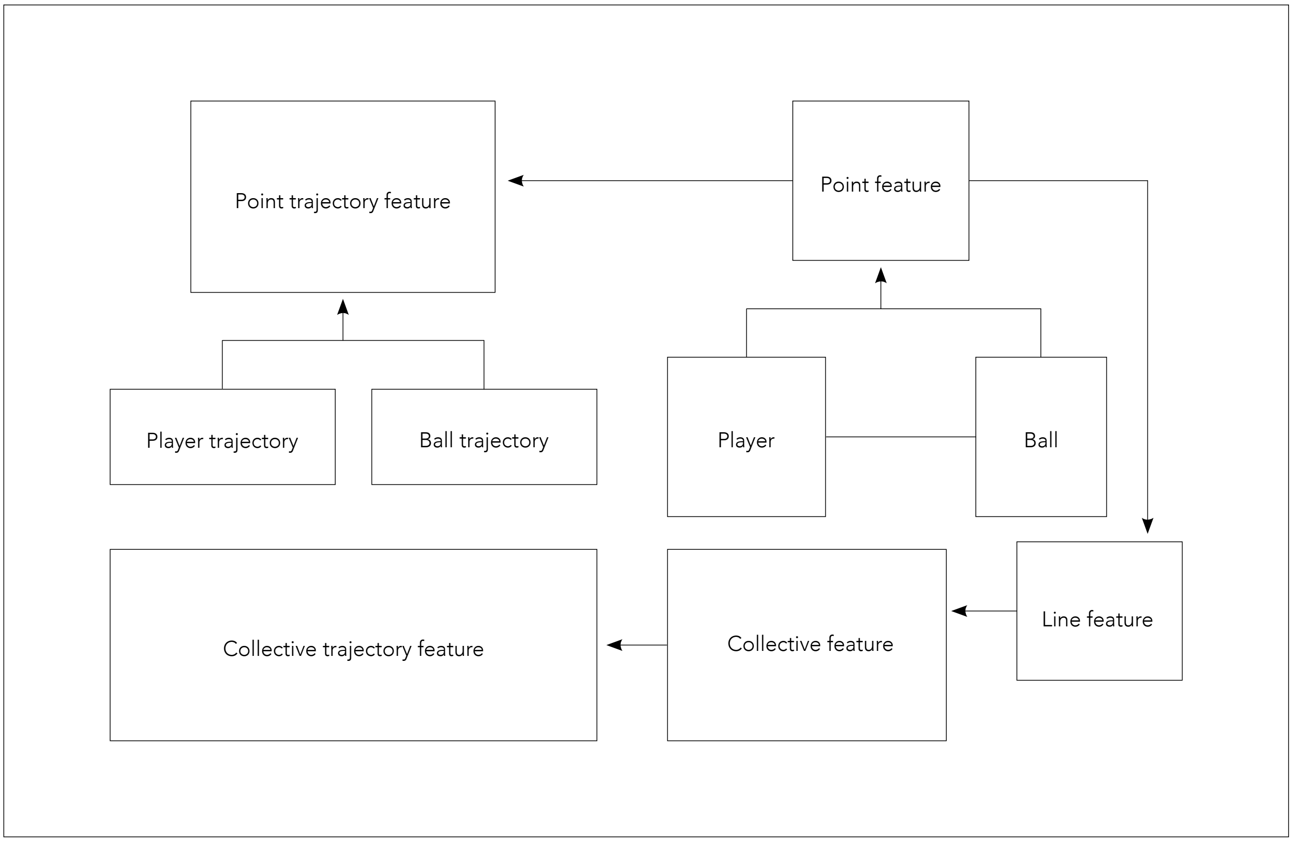
Figure 3.4: Schematic representation of the parameters: distance, orientation and direction.



Kim and colleagues (2011), developed a model of basic features that may serve as a framework to analyze these tactical patterns (see figure 3.4). In this model, point features represent the coordinates of the players and the ball at each time instance. The model also includes line features to analyze the spatial distribution of multiple players in relation to each other and the adversary team. Based on the type and shape of these lines collective tactics can be analyzed. For instance, If the depth of the defensive line is small, then it is easier to deploy an offside trap but difficult to protect the space behind the defensive line (see figure 3.5).

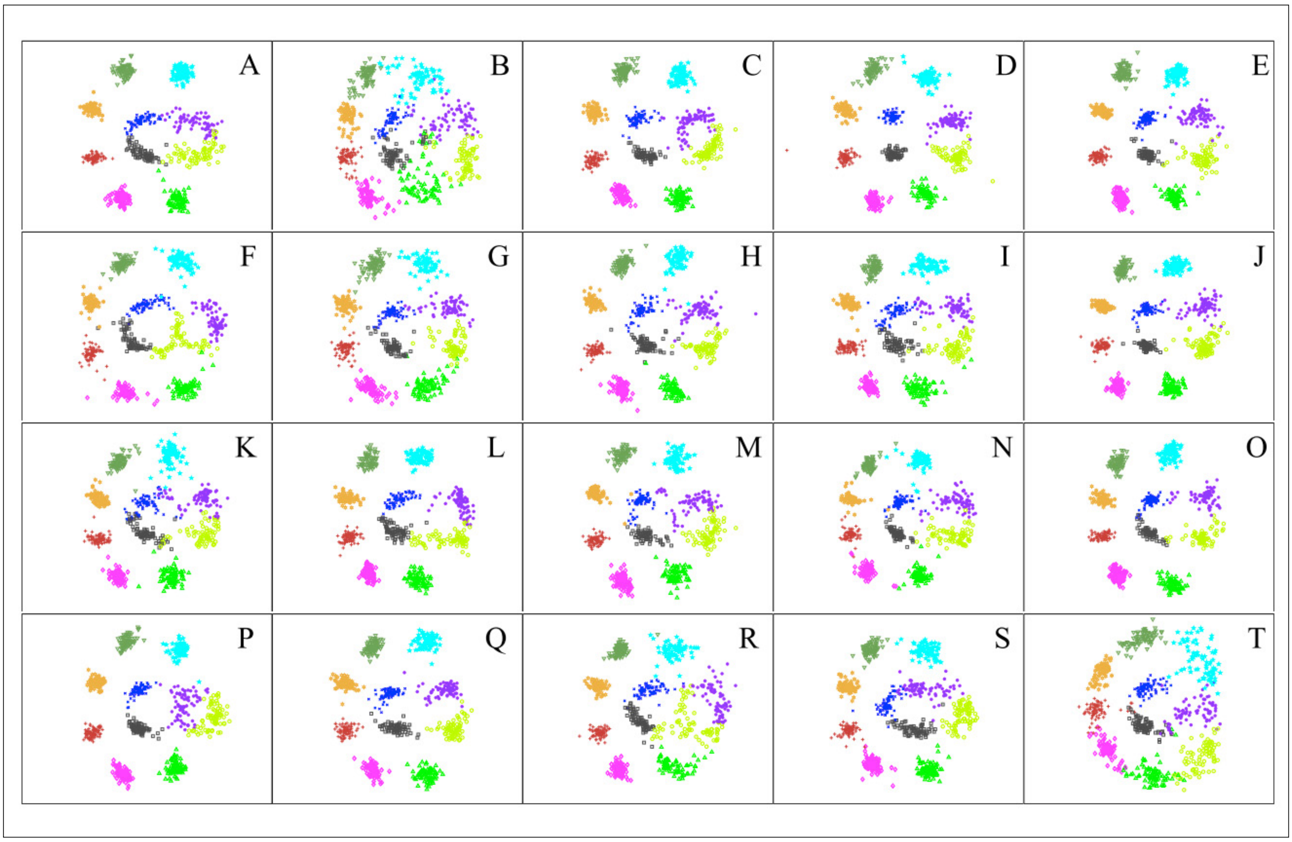
Figure 3.5: A simplification of the basic feature model that may serve as a framework to analyze tactical patterns.

Source: Kim et al., 2011



Another group of researchers who conducted research regarding the tactical analysis of football are Alina Bialkowski, Patrick Lucey and other members at Disney Research (URL 26). They used machine learning algorithms to investigate tactical decision making in professional football (Memmert et al., 2016). For instance, Lucey et al. (2013) proposed the idea of aligning player trajectories based on their role to effectively discover team formations. Additionally, Bialkowski et al. (2014b) developed a role-based representation that copes with the constant interchanging of player positions. Bialkowski et al. (2014b) extended this analysis by developing a method which can automatically detect and visualize formations using an entire season of ball and player tracking data (see figure 3.6). The results further showed that teams used more defensive formations during away games. In a follow-up article, Lucey et al. (2015) developed a method to estimate the probability of different attacking situations to result in a goal.

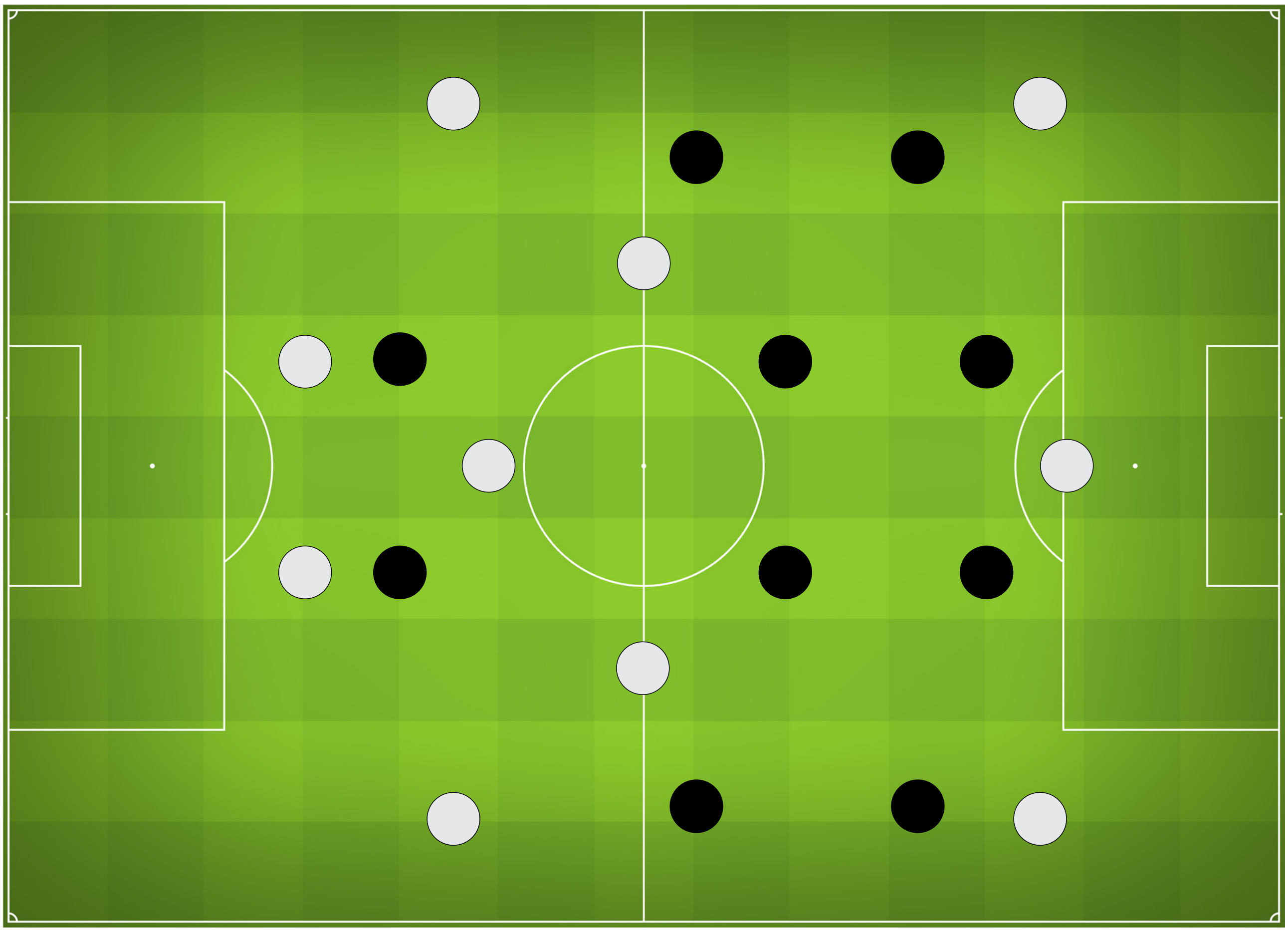
Figure 3.6: Example of formation descriptor for each team. The colors represent different roles. For visualization purposes the authors have plotted the centroid for each role for each team. Source: Bialkowski et al., 2014b



Despite the fact that Bialkowski and her colleagues made significant progress, there are still many research questions left unanswered (Gudmundsson & Horton, 2016). According to Sumpter (2016, p. 170), *"the next challenge in football analysis is to move from static descriptions of formations and positioning to dynamic analysis of player interactions*". In line with this challenge is the calculation of dominant regions, or the parts of the pitch near each player which he or she aims to control (Stein et al., 2016). Furthermore, there is the challenge to incorporate the pressure a player can put on other players and the ball by closing the free space around them (Gudmundsson & Horton, 2016). Andrienko et al. (2016) are among the first who proposed a novel approach capable of estimating, for each time frame, the pressure of the defending players upon the ball and opponents.

Arising from the discussed existing geovisual analytics methods, the primary challenge facing this research is to turn those vast amounts of tracking data into an informative picture that can be read by coaches, analysts and players. Despite the fact that football clubs increasingly begin to employ data analysts; it is not always clear how the message contained in the data carries up to the coach or out to the players. Therefore, analysts need to be able to present spatiotemporal data in terms of accurate and informative visual messages to communicate the results (Sumpter, 2016). These messages should provide a quick but clear explanation which sums up the essence of the match or training, much like tactical maps do (see figure 3.7).

Figure 3.7: Example of a tactical board mainly used during match briefings. The grey and black circles represent players of both teams. Source: URL 1



3.4. Geovisual analytics systems

Recently, heat maps appeared in the domain of the football analytics (Kotzbek & Kainz, 2016). These common approaches for visualizing aggregated information can be regarded as an improvement in respect of its consideration of space and time (Kotzbek & Kainz, 2016; Gudmundsson & Horton, 2016). Examples in the literature with respect to sport analytics are CourtVision (Goldsberry, 2012) and SnapShot (Pileggi et al., 2012). These are respectively designed for basketball and ice-hockey, and introduce specific types of heat maps tailored to ball and puck shots. CourtVision uses size and color to display the number of attempts and obtained shots visualizing the spread and range of a shooter (see figure 3.8), whereas SnapShot shows shot distances to goal and frequencies using differently colored concentric rings (see figure 3.9).

Figure 3.8: Advanced shot charts of LeBron James. Source: Goldsberry, 2012; Google Images

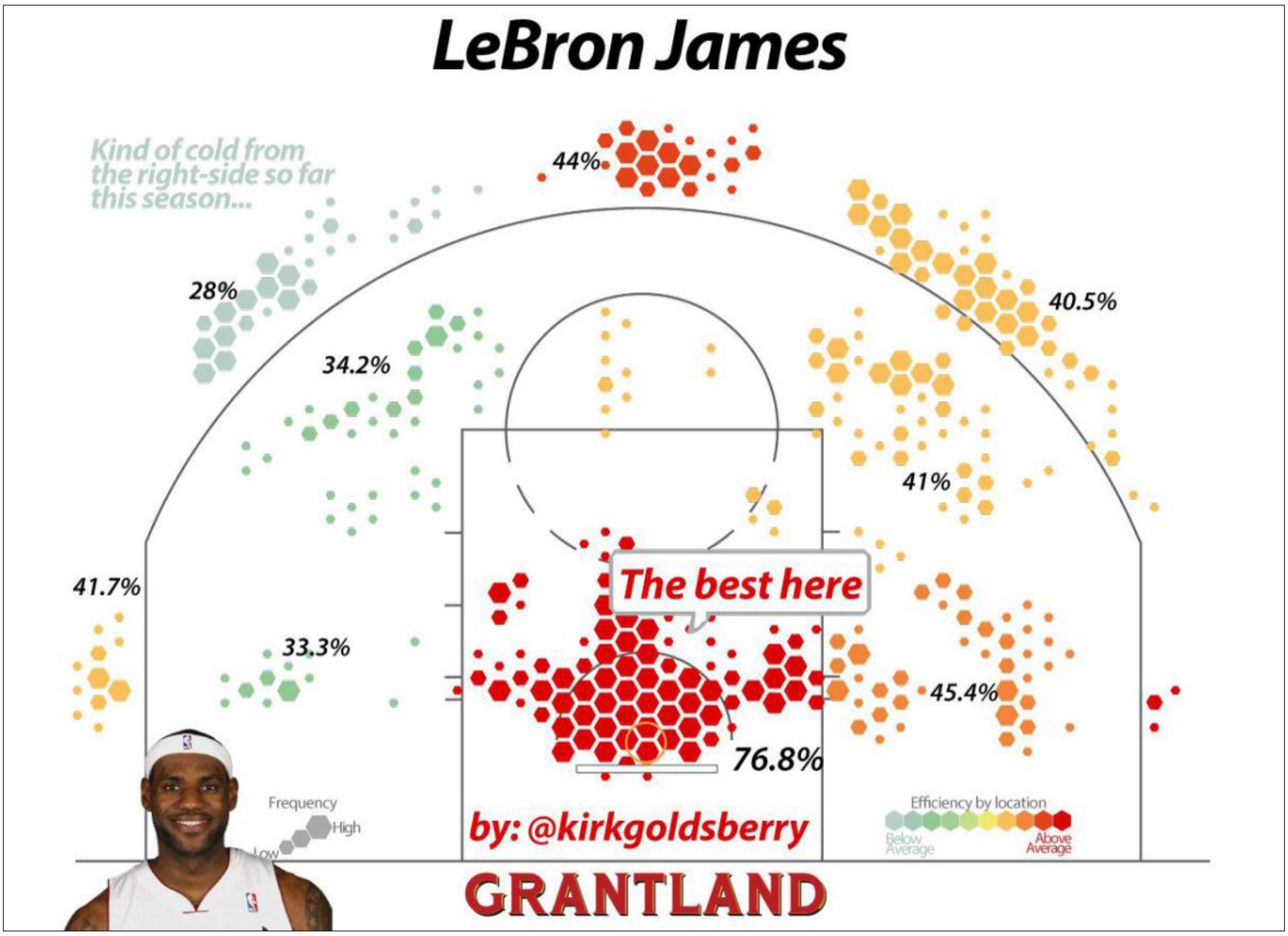
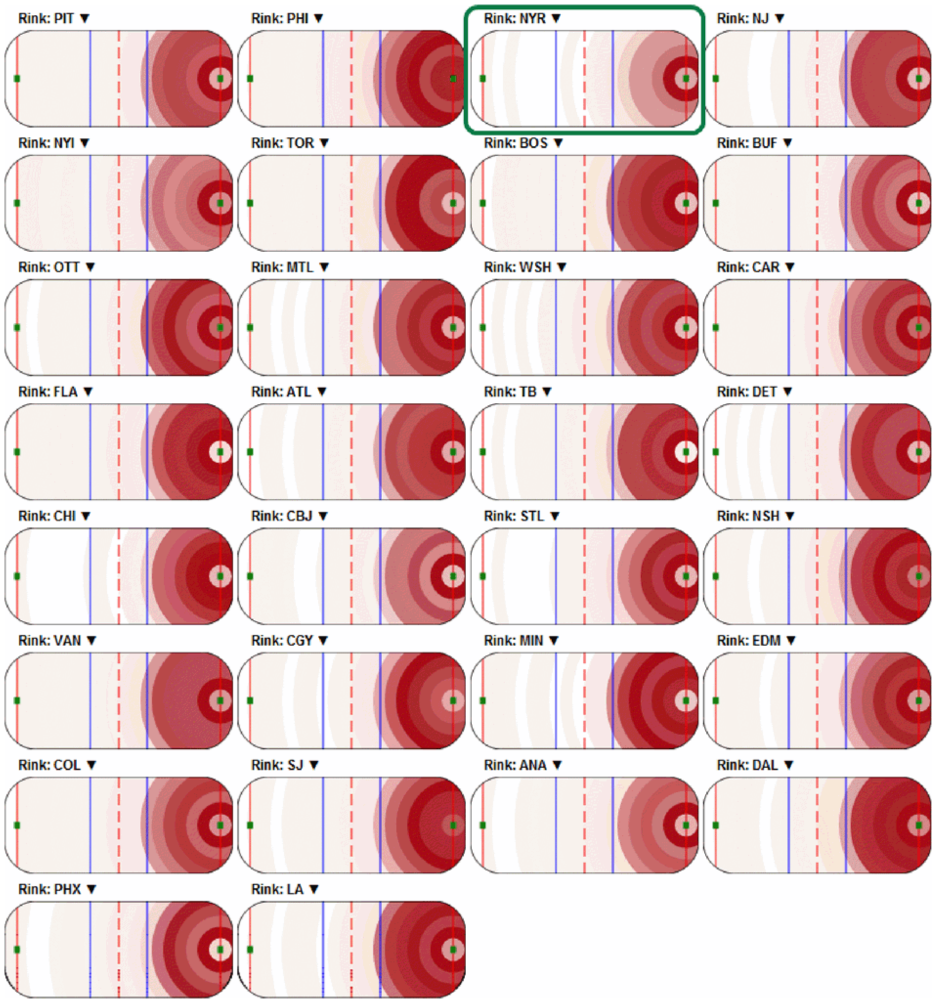


Figure 3.9: Small multiples of radial heat maps for each of the thirty NHL teams at home. The analyst's initial team of interest, the New York Rangers, is indicated by the green box. Source: Pileggi et al., 2012



SciSports also used heat maps to make so-called Goal Attempts Charts (see figure 3.10). These show the locations from where the player received the ball (blue), combined with the locations from where he attempted his shots on goal (orange). Furthermore, SciSports made use of flow maps to make Passing Maps (see figure 3.11). The line symbols (i.e. arrows) show the aggregated directions of all the passes, whereas the width of the arrow expresses the number of passes given. Another interesting visual tool made by SciSports is the Passing Chart (see figure 3.12). This tool illustrates the style of passing by dividing the passes in two categories: short and long distance. With the use of this classification, it can show the frequency, direction, accuracy and average length of each pass.

Figure 3.10: Goal Attempts Chart of Paris Saint-Germain striker Edison Cavani. **Source**: URL 1

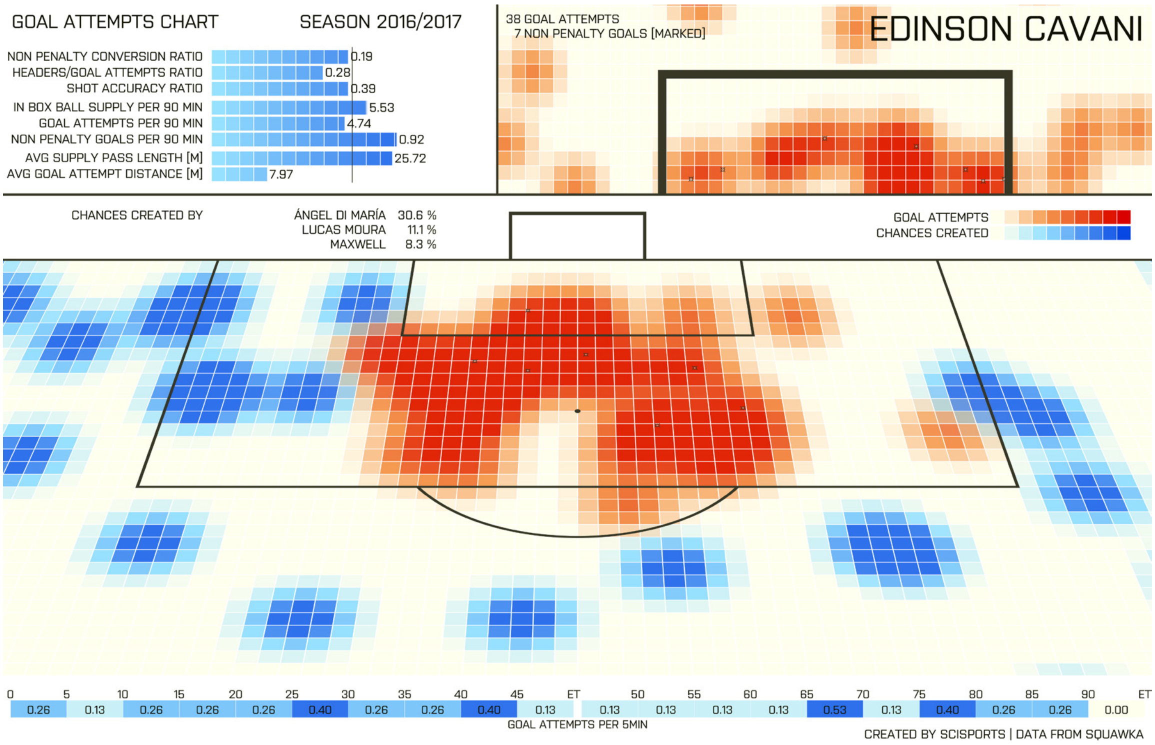


Figure 3.11: Passing Map of the Dutch football club Feyenoord. **Source**: URL 1

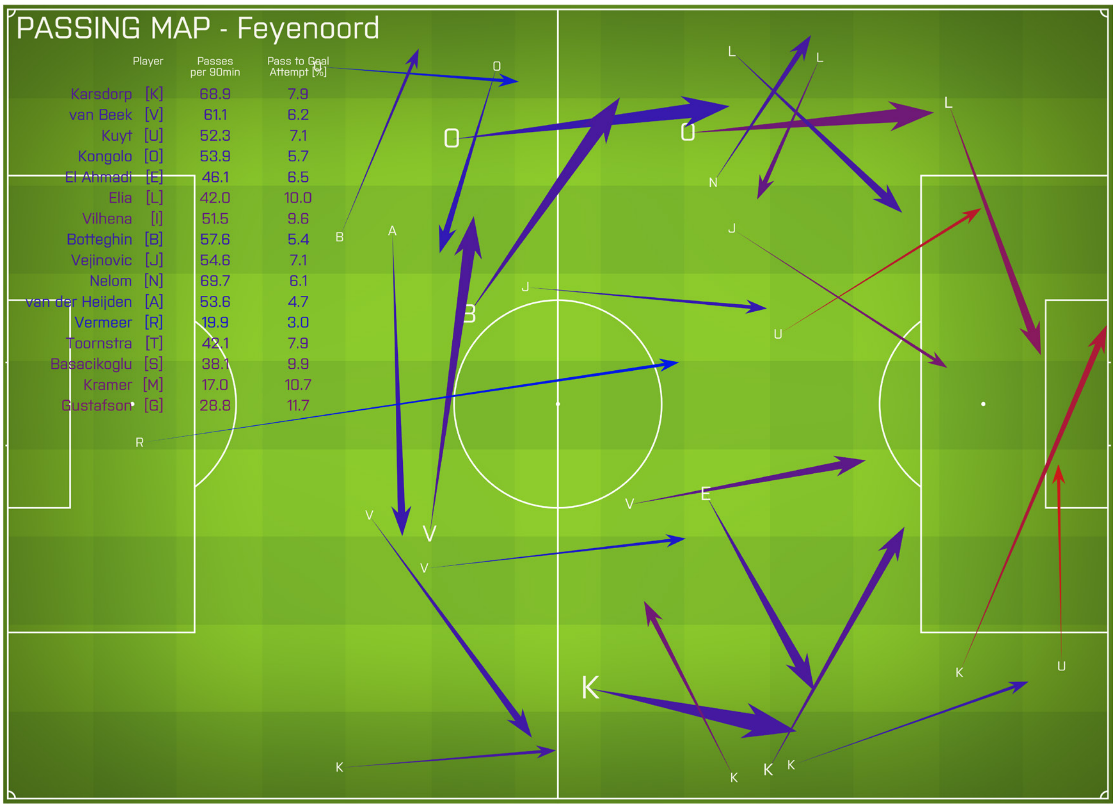
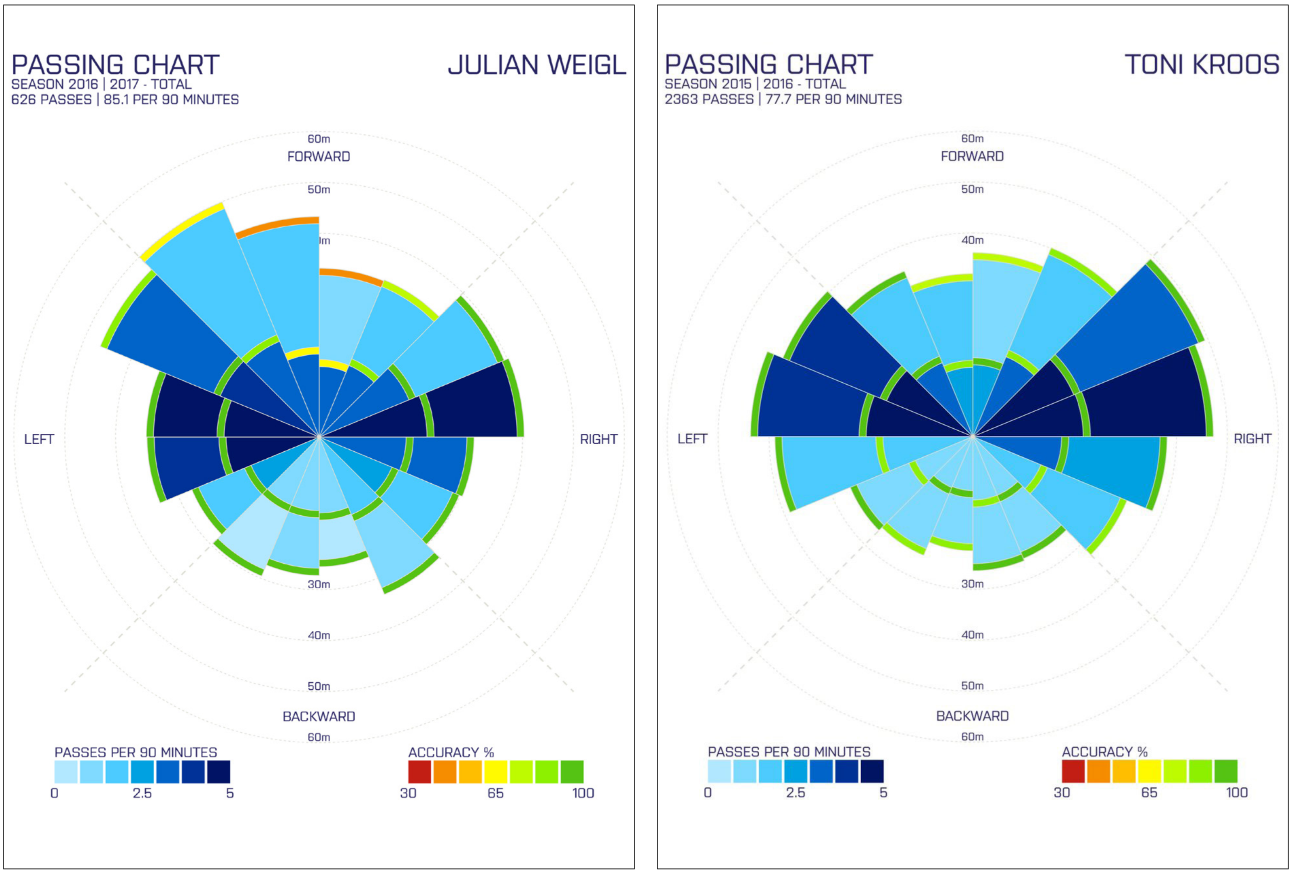


Figure 3.12: Passing Charts of Borussia Dortmund's midfielder Julian Weigl and Real Madrid's midfielder Toni Kroos.

Source: URL 1

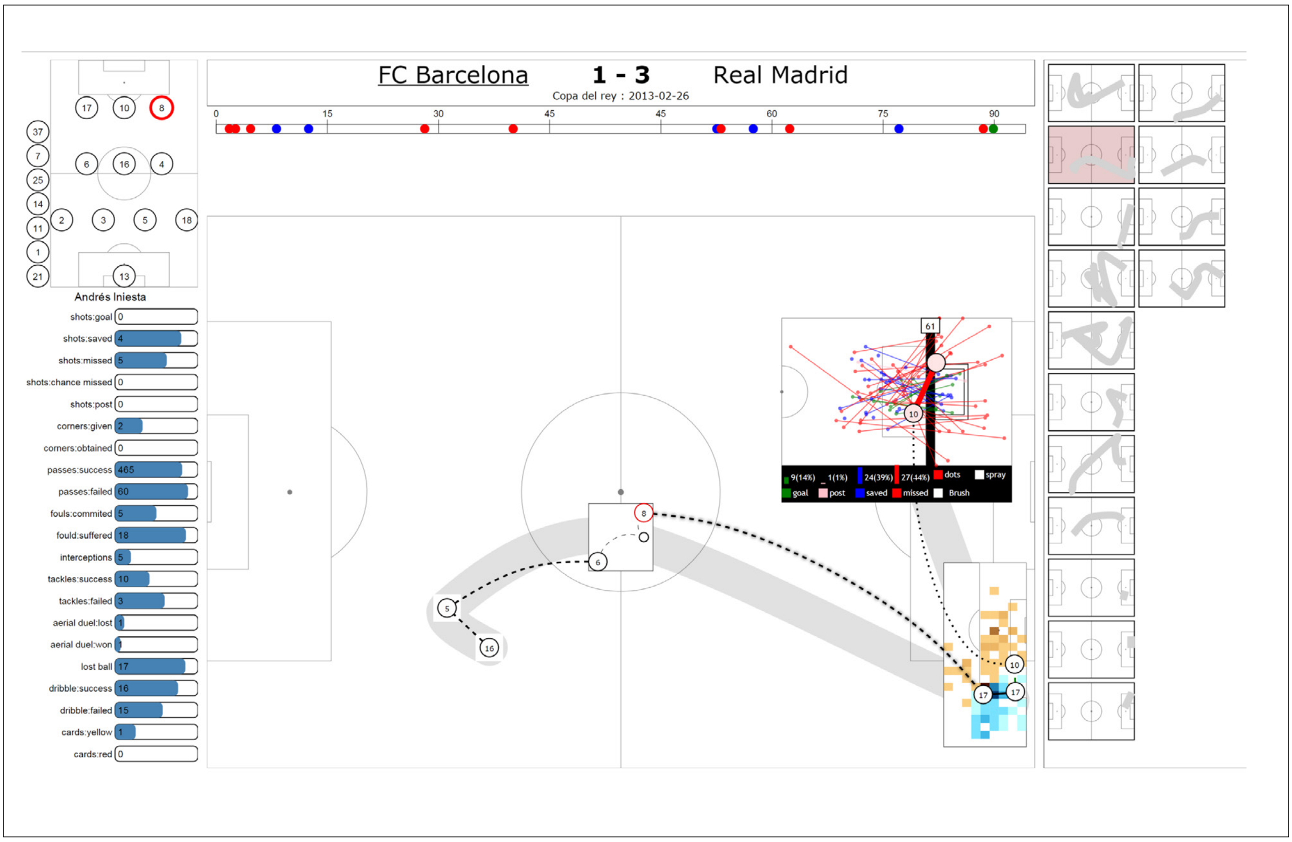


These and other applications, such as SportVis (Cox & Stasko, 2006), Soccer Scoop (Rusu et al., 2010), MatchPad (Legg et al., 2012) and TenniVis (Polk et al., 2013) acknowledge the important need and potential of visual analytics tools applied to the characteristics of (team) sports. Therefore, this section looks at four recent attempts to provide more extensive visual systems specifically made for the visual analysis of concerted movements. First, two systems made by academics are discussed. In sub-section 3.4.1, the system made by Perin et al. (2013), intended for the visual exploration of phases in football, is elaborated on. In sub-section 3.4.2, the visual analysis system for interactive recognition of patterns by Janetzko et al. (2014) and Stein et al. (2015) is discussed. Thereafter, two systems made by commercial companies are presented. In sub-section 3.4.3, the platform FootMapp developed by Metrica Sports is shown (URL 2). Finally, sub-section 3.4.4 presents the SAP Match Insight solution by SAP AG (URL 27).

3.4.1. SoccerStories

Charles Perin and his colleagues (2013), developed a user interface called SoccerStories. Their system provides an overview of game phases and detailed series of connected visualizations, where each view only focuses on one specific action. Phases are defined as a sequence of actions from one team until that team loses the ball (Perin et al., 2013). An interactive timeline and small multiples provide an overview of the game. The timeline shows each phase in chronological order and may be used to navigate to a phase of interest. The small multiples provide an overview by selecting, comparing and navigating between phases. The views are connected using various visual tools such as heat maps, hive plots and node-link diagrams (see figure 3.13). More detailed and legible figures of these visual tools can be found online (URL 33).

Figure 3.13: SoccerStories gives an overview of each match using connected visualizations of phases on a football field. Details can be found at both sides and the timeline on top is used to navigate. **Source**: Perin et al., 2013; URL 28



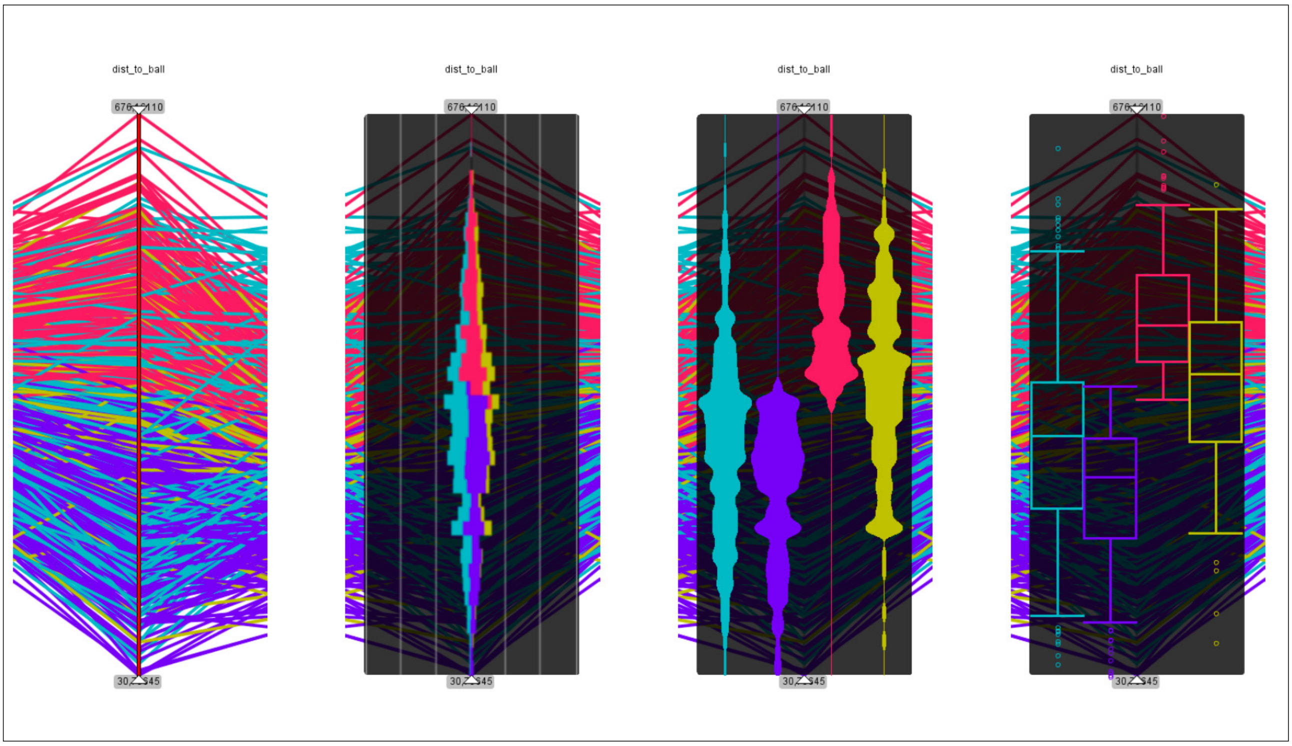
To better understand the workflow behind game analysis, Perin et al. (2013) first conducted four interviews with two online journalists, one specialist from Opta Sports and one professional coach. In addition, the authors evaluated SoccerStories afterwards based on two qualitative user studies. These showed that, with minimal training, experts where capable of finding novel insights (Perin et al., 2013). The main advantage of SoccerStories is that it provides a very intuitive overview and detailed view of phases using connected visual tools. Phases are adequate for the visual exploration and analysis of movements as they are well understood, easy to extract and communicate more information than single events do. However, the evaluation also showed that the more advanced visual tools were too complex for domain experts to understand. Therefore, this research concludes that, despite its advantages, the SoccerStories environment is still too complex for users to employ properly without any training. Nevertheless, SoccerStories can be used as a starting point for this research by exploring the kind of analytical questions that potential users need support for, the objects and processes that they want to be visualized and the functionalities the interface should support (see chapter 4).

3.4.2. Feature-based visual analytics system

The work of Halldór Janetzko and Manuel Stein, who are part of the Data Analysis and Visualization Group at the University of Konstanz, is closely related to the work of Perin and his colleagues (2013). In two successive papers, they presented a feature-based visual analysis system capable of interactively exploring and analyzing movement patterns. Their system makes use of various visual tools to communicate the results. These tools include interactive representations and are connected via brushing and linking techniques. By using the timeline, the analyst can control the currently visualized time window which allows quick navigation of important events. The system automatically identifies complex patterns and classifies them into several parts which are visualized as ranked features. These features are displayed as rectangles and ranked based on their temporal occurrence. Each rectangle represents a descriptive part of complex and hardly interpretable moves. Interaction possibilities are added to change the ranking of a feature, to highlight or to search for similar moves.

The interface provides not only single-player and event-based visualizations, but also multi-player analytical views. When the analysist wants to investigate a single player, he or she can use colored line charts, parallel coordinate plots or small multiples. When the analysts want to compare multiple players, the system makes use of horizon graphs (Stein et al., 2015). Parallel coordinates are a common way of visualizing multivariate data. With each data point being represented by a line, overplotting may cause severe readability problems. In one of their most recent publications, Janetzko et al. (2016) proposed the usage of stacked bar charts, violin plots and box plots to improve parallel coordinate plots (see figure 3.14). These visual and analytical extensions have been motivated by the need to scale with the large volumes of movement data (Janetzko et al., 2016).

Figure 3.14: Default parallel coordinated plot, stacked bar chart, violin plot and box plot. Source: Janetzko et al., 2016



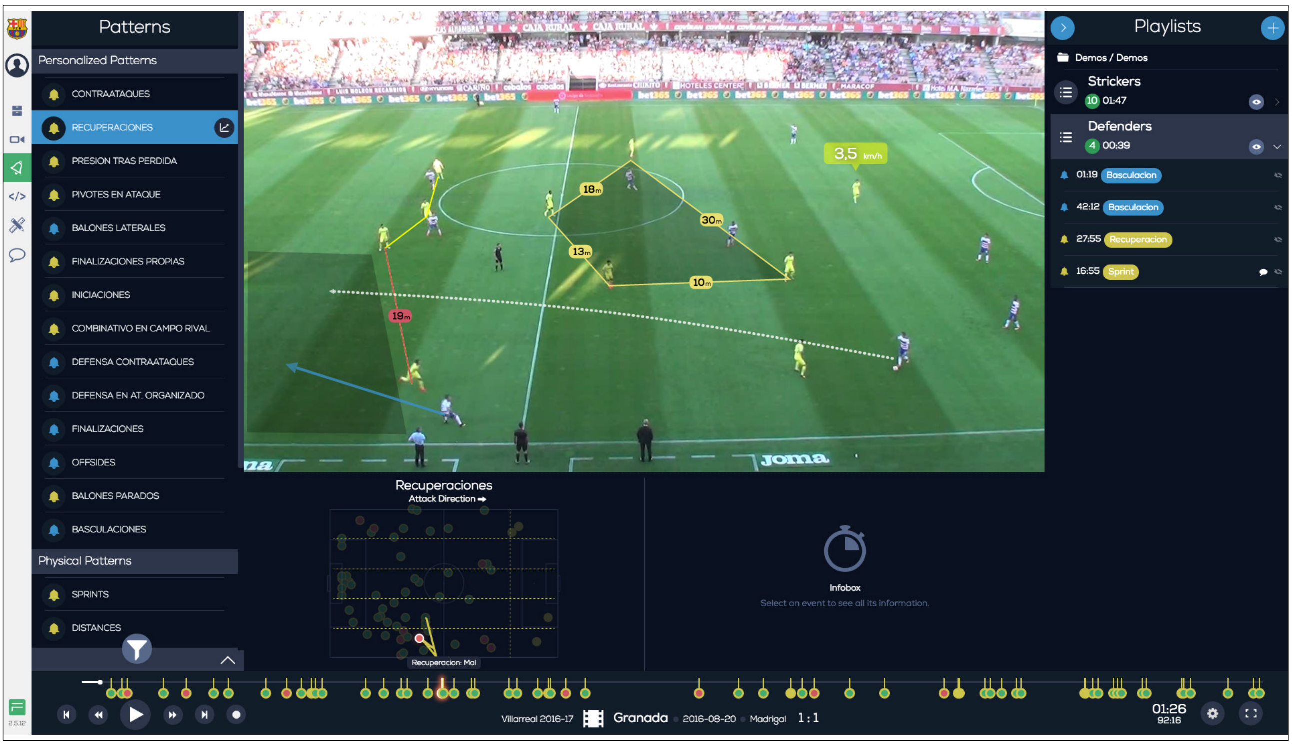
To determine the usability of their system, they made use of an approach which looks similar to the User-Centered Design (UCD) method used during this research (see chapter 4). First, they interviewed two experts (i.e. an official referee and a senior analyst at FC Bayern Munich) to get initial feedback about the semi-automatic detection of events and patterns. The second experiment consisted of a task-based user analysis. While the experts were asked to use the system, their interactions were recorded and thoughts and feelings collected using the think aloud technique. Both experts agreed that the main advantage is that the system allows analyzing more matches in less time, as compared to traditional video analysis. Furthermore, the combinations of various visual tools proved to be capable of dealing with vast amounts of movement data. However, their evaluation also showed that the visual tools used are too complex and not intuitive enough for coaches and analysts to understand (Janetzko et al., 2014; Stein et al., 2015). Another severe limitation of their evaluation is that only two experts took part in the process. For a user study, this population sample is considered to be relatively small. Therefore, my research deviates by preforming a requirement analysis prior to the conceptual design of the interface (see chapter 4). In this way, this research contributes to the area of visual interfaces by placing the end-users at the heart of the project and not the technology or the data.

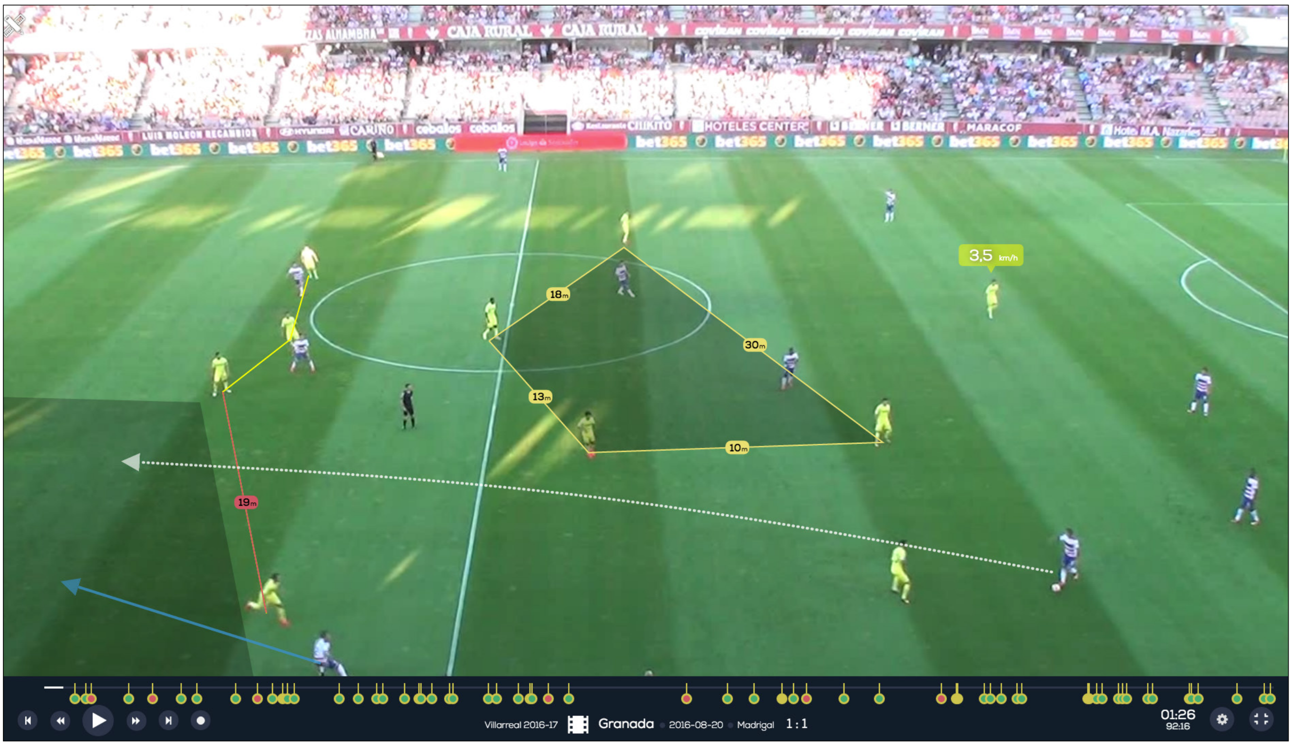
3.4.3. FootMapp

Metrica Sports also recognized that analysts and coaches are laymen in the domain of visual analytics. By developing FootMapp, they provide professional football clubs with a platform to detect, visualize and communicate found movement patterns (URL 2). FootMapp is a video editing solution based on tracking data to analyze team movement patterns and translate these into insights for the coaching staff. The system uses pattern-recognition technologies that can automatically detect key moments in the game. In addition, interactive on-video visualizations are used to visualize distances between all players, speed and acceleration (see figure 3.15). With FootMapp, coaches and analysts can analyze and improve both single player and team performance based on specific tactical patterns.

The main advantage of FootMapp is its ability to adapt its visual analytics tools to the individual needs of a team (van Zoelen, 2016). This implies that FootMapp is a custom-made tool which can automatically detect key moments in the game based on the parameters defined. Prior to the official launch of FootMapp its team of developers started by analyzing the market by talking with lots of professional football organizations (Diepstraten, 2017; Aquina, 2015). My research acknowledges the advantage of incorporating user requirements prior to the interface development process. By exploring the requirements of potential users, using a requirement analysis, my study tries to identify which types of overviews analysts and coaches need to achieve a better performance of their team.

Figure 3.15: FootMapp showing personalized and physical patterns (left-side) collected during the match Villarreal versus Granada. Furthermore, on-video visualizations are used to highlight key moments/movements. Source: URL 2

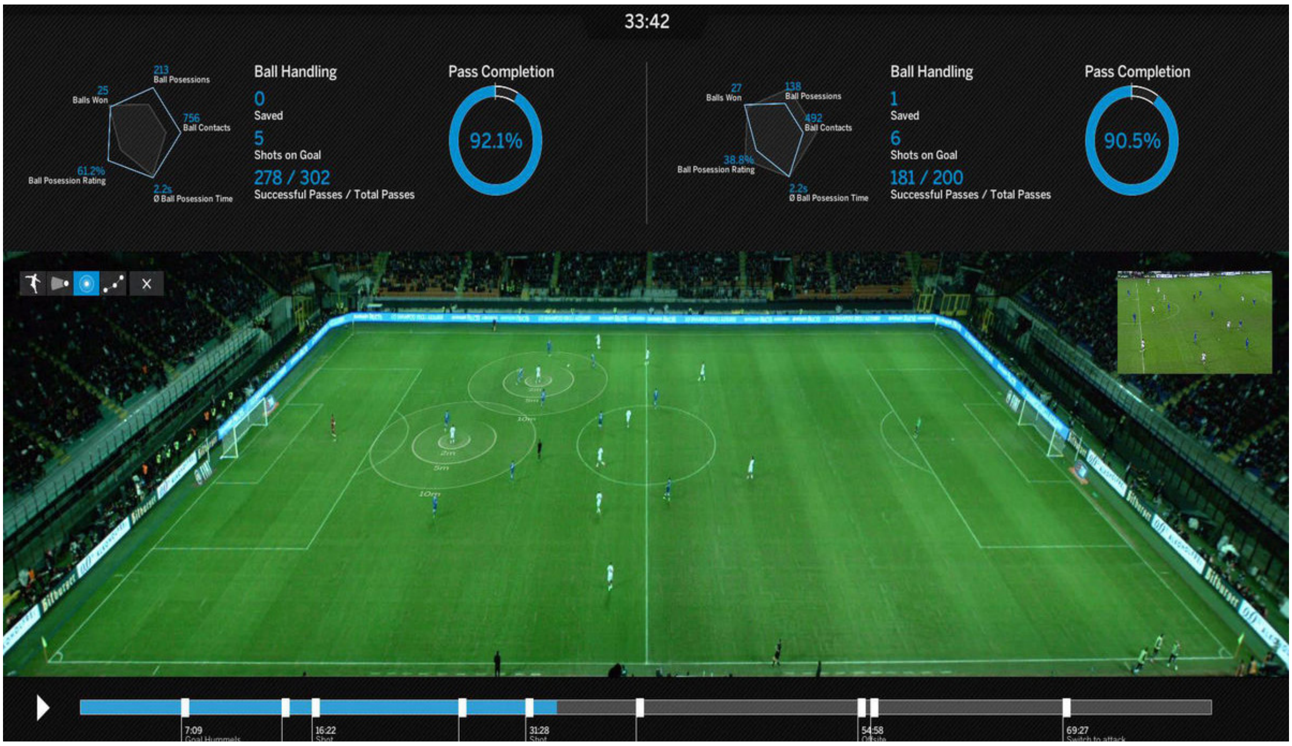




3.4.4. SAP Match Insights

SAP Match Insights is the result of a co-innovation project of the German company SAP AG and the German Football Association (DFB) prior to the FIFA World Cup of 2014. Match Insights runs on the SAP HANA cloud-based platform and is part of the integrated solution called SAP Sports One (URL 27). Part of this solution are several web and mobile applications which makes it possible for clubs to share information about player fitness, team management, scouting and match preparation by staff and players. Match Insights itself enables coaches, analysts and scouts to process vast amounts of data in order to find and assess key situations in each match to improve player and team performance. Current known customers are, among others, FC Nuernberg, TSG Hoffenheim, FC-Astoria Walldorf and the Philadelphia Union Academy. The interface combines video analysis, player statistics and tactical team information visualized on one dashboard (see figure 3.16). This relatively simple and intuitive interface may be used by players, analysts and coaches to facilitate a more interactive dialogue to debrief the team and to prepare for future matches (URL 27).

Figure 3.16: SAP Match Insights showing key statistics, movements and video footage of Philipp Lahm, (former) player of the German national team. **Source**: URL 32



3.5. Conclusion

Based upon the scientific foundation outlined in the previous chapter, this chapter has dealt with the collection, analytics and visual analysis of concerted movements in order to enable analysts, coaches and players to visually explore and analyze spatiotemporal patterns. Currently, methods of post-game analysis often fail to interpret and visualize these complex spatiotemporal patterns sufficiently. Therefore, my research aims to contribute to the visual analysis of football movements by developing a visual user interface to the underlying football-specific spatial data.

Despite the availability of movement data, most professional football organizations are still focused on statistical data and use video recordings of past matches and training sessions to analyze the individual player performance. This must be considered insufficient, because football is all about the union of space and time. With the increased availability of data, it is now possible to extract tactical movement patterns from tracking data. Even though significant progress has been made, the challenge facing this research remains turning the vast amounts of tracking data into meaningful visualizations and graphical representations that can be read by football coaches, analysts and players.

Four different attempts to provide extensive analytic systems, specifically made for the visual analysis of football movements, were discussed. Although, some user research has been executed for the development of some of these systems, that user research was rather limited and there are still concerns about the usability of these systems. Therefore, more user research is required. My research contributes to meeting this need by conducting a requirement analysis prior to the actual development process of a new conceptual design in order to understand what potential users need and require, particularly with respect to the visual exploration and analysis of concerted football movements.

In the next chapter, this research continues and contributes to this process by explaining the methodology that this research project will employ.

CH.4

Methodology

4.1. Introduction

Most geovisualization tools and geovisual analytics environments are developed using a technology-centered design approach. In this way, the usability of products remains often unsatisfying (Delikostidis, 2011). In the previous chapter, a comparison between different visual tools and interfaces was made and space for further research in this field was investigated. This chapter confirmed that most domain experts are unable to use advanced and interactive visual tools effectively when they did not get a training first. Going through different visual tools and interfaces, and inspecting the different design and development strategies followed, the selection of a methodological approach for this research was shaped.

It is essential that a consolidated procedure is being implemented throughout the project from the very start, ensuring a quality output that is reproducible and scientifically and methodologically sound. The methodological approach chosen for this research is the deductive-inductive approach. This approach implies that as a starting point, general knowledge is synthesized with the use of a structured literature review. As a follow-up to conducting this literature study which formed the basis of chapters 2 and 3, this fourth chapter describes the User-Centered Design (UCD) methodology which this research employs, outlining the procedures that lead to the results, which are discussed in chapter 5 and 6. Furthermore, the case study method is used to deal with the complexity of analyzing concerted movements.

4.2. User-Centered Design

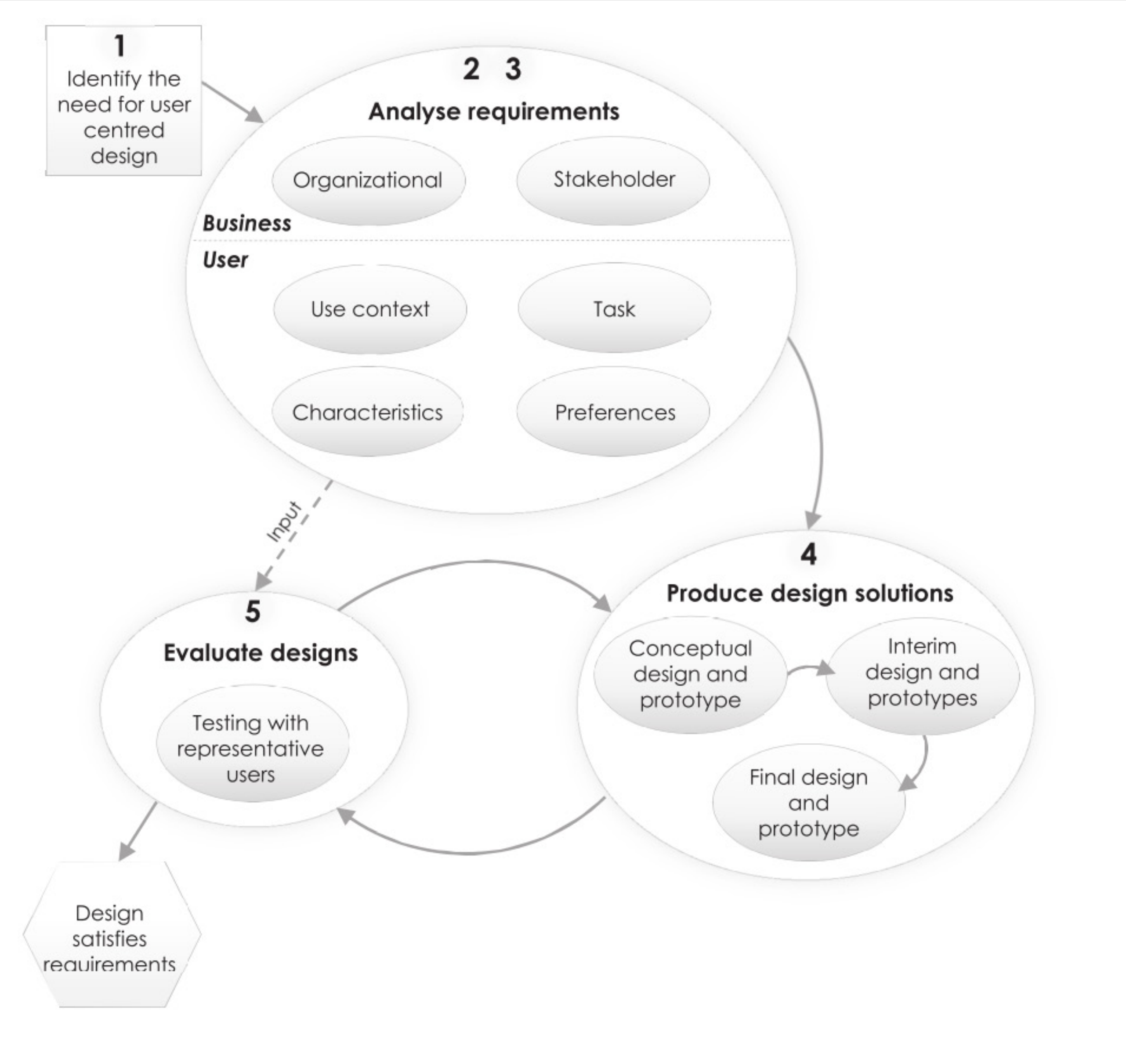
The term UCD was initially introduced by Norman and Draper in 1986 and further elaborated on by Norman in 1988. In this research, UCD is defined as *"a philosophy based on the needs and interests of the users, with an emphasis on making products usable and understandable"* (Norman, 1988, p. 188). This research acknowledges that the perspective of users regarding the interface development process is often underestimated, if not overlooked altogether (Delikostidis, 2011). Often this is the case because designers do not analyze who the end-users are, what their working environments are and what tasks they want to perform (Wassink et al., 2008). Roth et al. (2015, p. 262) proposed to mitigate these shortcomings by applying a UCD approach to ensure *"interface success by gathering input and feedback from target users throughout the design and development of the interface"*.

The overall aim of this research, as presented in chapter 1, is to develop a conceptual design of an interactive visual interface. Developing such a design with usability in mind suggests that a carefully defined UCD-based strategy should be constructed (Delikostidis, 2011). A well-structured UCD method consists of multiple phases (see figure 4.1). In the framework of this research, the three main phases that are covered are: requirement analysis, conceptual design and evaluation.

During the first phase, the current situation is being analyzed, which results in user profiles and user requirements (see sub-section 4.2.1). The qualitative data collected during this phase are used as input for the conceptual design process. During this second phase, the conceptual design is developed with the use of wireframes, which are low-fidelity representations of the actual design that focuses on structure and content without the distraction of visual elements (see sub-section 4.2.2). The conceptual design can be turned into a high-fidelity prototype that might be presented to potential users through various iterations. In the third and final phase, the usability of these high-fidelity representations could be iteratively evaluated, including an assessment of both visual representation and interaction styles (see sub-section 4.2.3). These last two steps however could not be executed due to the resources and the lack of time available for this MSc research.

Figure 4.1: A schematic representation of the main activities of the User-Centered Design process.

Source: van Elzakker and Wealands, 2007



4.2.1. Requirement analysis

User profiling is the first activity of the requirement analysis, which comprehends a description of the basic characteristics of the users who are intending to use the conceptual design and implementation thereof (Delikostidis, 2011). Different types of users often have different goals, perform different tasks and have different experience with the use of visual tools (Wassink et al., 2008). They differ, furthermore, in their perceptual abilities, purposes, roles, backgrounds and tasks. Together, all these aspects determine the needs and requirements for visualization and interaction use. Due to these differences, it is essential to perform a detailed analysis of the users, their environments and their tasks, before starting with the conceptual design of the actual interface (Wassink et al., 2008).

As said in the previous section, the requirement analysis is considered as the first phase of the UCD approach. The output of this phase consists of user profiles and user requirements. The data collected during this phase are used as input for the conceptual design and evaluation activities to reason what the valid design solutions are for the intended users. To collect this data, there are numerous different user research methods available. Most frequently used are observations, questionnaires or interviews (van Elzakker, 2004; Wassink et al., 2008).

Observations are useful when observing how users act and interact in certain situations. The main advantages of this method are threefold: it identifies actual behavior, contextual and supportive information is collected, and it is possible to conduct in multiple situations (Hennink et al., 2010). The main limitation of this method is that participants are aware that they are being observed and may get nervous, which could result in misleading observations (Bryman, 2011). Questionnaires are often used to generate quantitative data from large groups of users and may, under certain conditions, be suitable for translation into statistical information. In UCD, statistically significant outcomes are needed to validate the usability of the final design solutions. But in the requirement analysis stage the user research is more often of a qualitative nature (Kveladze, 2015).

The qualitative technique interview is more suitable when needing to obtain sensitive information about complex issues and to gain the support of the participants. The interviewer is available to explain questions and ambiguities (Wassink et al., 2008). Due to the opportunity to provide more detailed explanations, interviews form a main source of knowledge for this research. Interviews are often being used to gather in-depth information about the attitudes, preferences and needs of the targeted users (Bryman, 2011). Interviews can be structured, unstructured or semi-structured. In practice, most interviews are semi-structured, which combines features of both structured and unstructured interviews (Wassink et al., 2008). This means that the interviewer makes use of several pre-defined topics. By using a topic list, the interviewer controls the interview by asking questions useful for answering the research questions. At the same time, the interviewer should be open minded and willing to explore topics that the interviewee brings up and construct new interview questions based on the answers received (Maxwell, 2013). Interviews however are also known to be time-consuming, and therefore only a limited group of participants can be chosen (see section 5.3).

Due to the different objectives, advantages and disadvantages of these user research methods, most requirement analyses consist of a combination of several (qualitative and quantitative) methods in order to provide comprehensive and meaningful results (Nielsen, 1993). The exact combination of methods, in such a so-called "mixed methods approach", depends on the project characteristic (Kveladze, 2015). In this research, structured interviews with potential users are combined with the think aloud method and synchronized screen and audio recording. Combining these methods to pursue a single goal is also known as triangulation and provides a rich source of information for extracting user requirements (Delikostidis, 2011; Wassink et al., 2008).

Domain experts will be interviewed who are working for professional football organizations. These interviews will take place after the literature study has been finished because only then there will be sufficient knowledge about the research topic. The aim of these interviews is to understand, identify and obtain (anonymous) information regarding the collection, analysis and visualization of movement data. In addition, the interviews focus on getting a first understanding of the context of use and user requirements. For instance, 'what kind of analytical questions do users need support for?', 'what are the objects and processes that need to be visualized?' and 'what kind of insight should visualizations support?'. The content of these questions, the order and the way in which they are asked are structured and standardized. In this way, the comparability of the data is more assured (van Elzakker, 2004).

The think aloud method is used to understand the mental processes behind actions when users are completing a particular task. This method involves the analysis of audio recordings resulting from asking domain experts to speak out loud about their thoughts and feelings when executing problem-solving tasks (van Elzakker, 2004). Each participant will be given the same task to perform and asked to think aloud while their verbal and on-screen actions are recorded. The main advantage of this method is that in-depth understanding of cognitive processes is gathered. An additional advantage is that many people have no problem expressing the thoughts that come up, as they are voiced immediately. Nevertheless, there are also some limitations. Foremost, thinking aloud is a very time consuming method. It is not only the data collection that takes time, but particularly the coding and analysis afterwards. Another disadvantage is that some participants may find it difficult to verbalize their thoughts and are often tempted to say what they think is expected by the researcher (van Elzakker, 2004).

In this research, the think aloud method combined with synchronized screen and audio recording is applied during two separate task-based sessions. During the first task-based session, domain experts are asked to use the SoccerStories environment and express their thoughts and feelings. They are given the task to use SoccerStories and look for the information they would like to have during the visual analysis and exploration of phases of the football game that led to a shot on goal. This task is developed to discover the analytical questions that potential users have. Before participants were invited to execute the actual tasks, the SoccerStories environment and all necessary functions were demonstrated.

During the second task-based session, an animation of approximately one minute is shown multiple times which highlights the coordinated positioning of defenders with the use of points, lines and polygons (see section 4.3). Again, participants are given the task to look for the information they would like to obtain during the visual analysis and exploration of movement data. In addition, they are asked 'what kind of insights these visualizations should support?'. During this task, participants are encouraged to think aloud while interacting with the animation. During the actual thinking aloud, the researcher follows the comments of the participants and in case they are not clear, or if they stop talking, questions like: 'why do you think so?', 'could you please explain what do you mean?', 'what do you miss?' and 'what are you looking for?' are asked.

The interviews are audio recorded and thereafter transcribed using the free web application Otranscribe (URL 29). The main advantage of this application over other tools, like for instance QuickTime Player and VCL Media Player, is that it allows the researcher to pause, rewind and fast-forward the recordings without switching between screens and applications by using simple keyboard shortcuts. Another advantage is its capability to insert interactive timestamps which can be used to easily switch between relevant passages. The finished transcripts are encoded using NVivo (URL 30). With the aid of this data management tool, pieces of text can quickly and easily be selected and grouped systematically into themes and nodes. Through the analysis of multiple codes, it is easy to identify these across datasets. In addition, nodes can be organized and analyzed in multiple ways to make sense of and find insights in unstructured qualitative data. Another advantage is its ability to collect multimedia data from multiple devices that can be linked to the transcribed data.

Furthermore, the logging software QuickTime Player is used for synchronized audio and screen recording. Logging the manipulations on screen and capturing the thoughts and feelings is important to get full insight into the interaction with the visualizations that appear on the screen. This multimedia framework, is built-in by default on every modern Mac OS workstation and is provided with a simple interface that is easy to use (URL 31). The screen recording feature records all changes participants make including cursor manipulations. Furthermore, audio recording is supported by using the workstation's internal microphone.

4.2.2. Conceptual design

In essence, a conceptual design describes how the interface works and meets the user requirements. It is an early phase of the design process, in which the broad outlines of function and form are articulated (Saidy, 2016). The user requirements identified in the previous phase are ideally answered by the interface in a clear and understandable way (Delikostidis, 2011). At the same time, the design and interaction of the interface should follow well-defined rules regarding the presentation of information in a usable form (Kveladze, 2015).

In sub-section 2.4.1, the three approaches of cartographic interaction: the objective-based, the operator-based and the operant-based approach have been introduced. First, objective-based approaches (e.g. identify, compare and rank) emphasize the kind of tasks that the indented user may wish to complete with the cartographic interface (see figure 4.2). Second, operator-based approaches (e.g. brushing, focusing and linking) focus upon the interfaces that make manipulation of the representation possible (see figure 4.3). Finally, operand-based approaches focus on the operand, or the object with which the user is interacting with (see figure 4.4). According to Roth (2012; 2013), the operand approach focuses on the type of information (i.e. space, time or attribute) that is being represented in the map or the state in the information pipeline from data to visualization.

> *"While the outlook for interactive maps is great, ensuring they "work" successfully for the target users remains a challenge"*
>
> Roth et al., 2015, p. 263

Figure 4.2: Relationships among objective-based taxonomies and the relative frequency of the included primitives.

Source: Roth, 2012

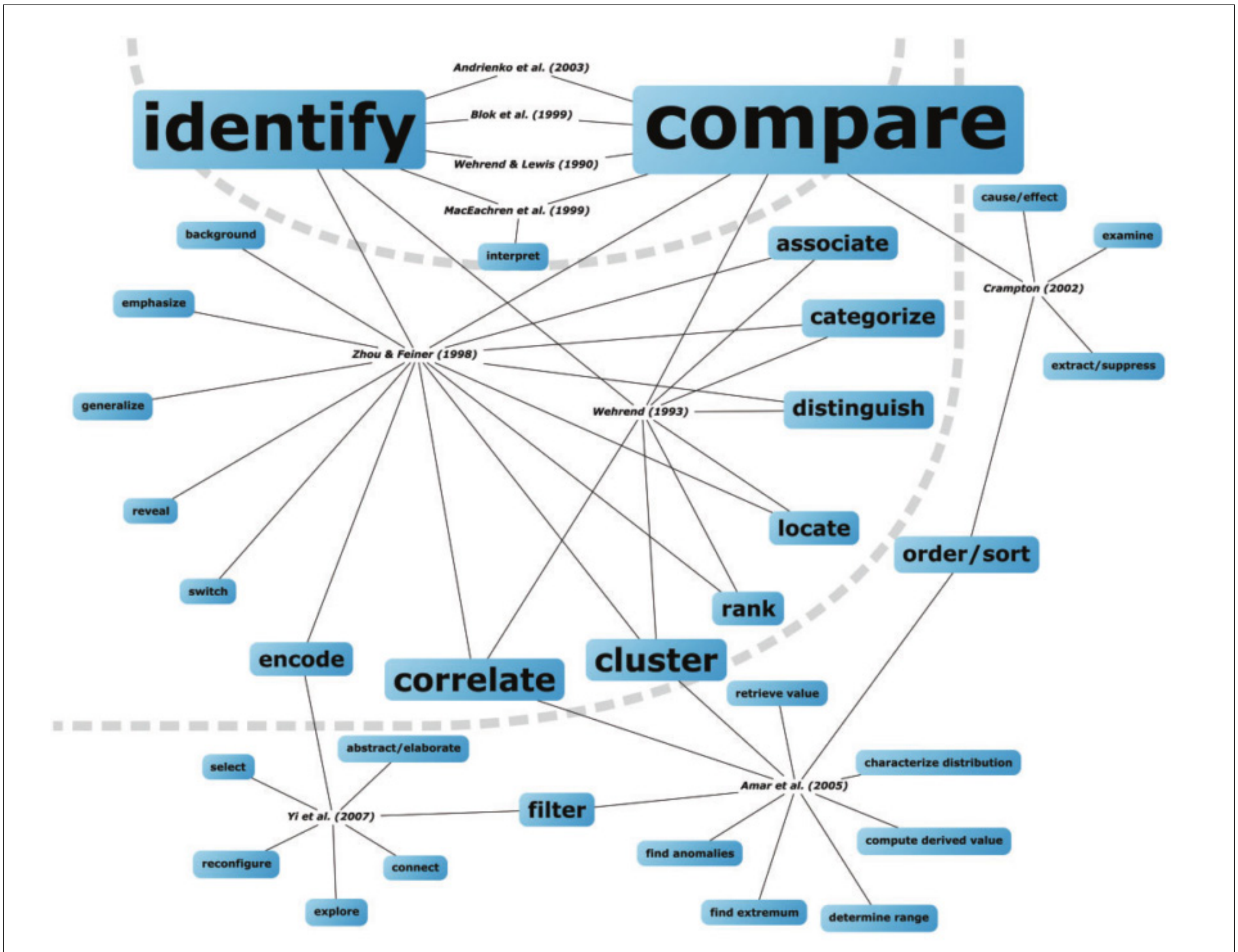


Figure 4.3: Relationships among operator-based taxonomies and the relative frequency of the included primitives.

Source: Roth, 2012

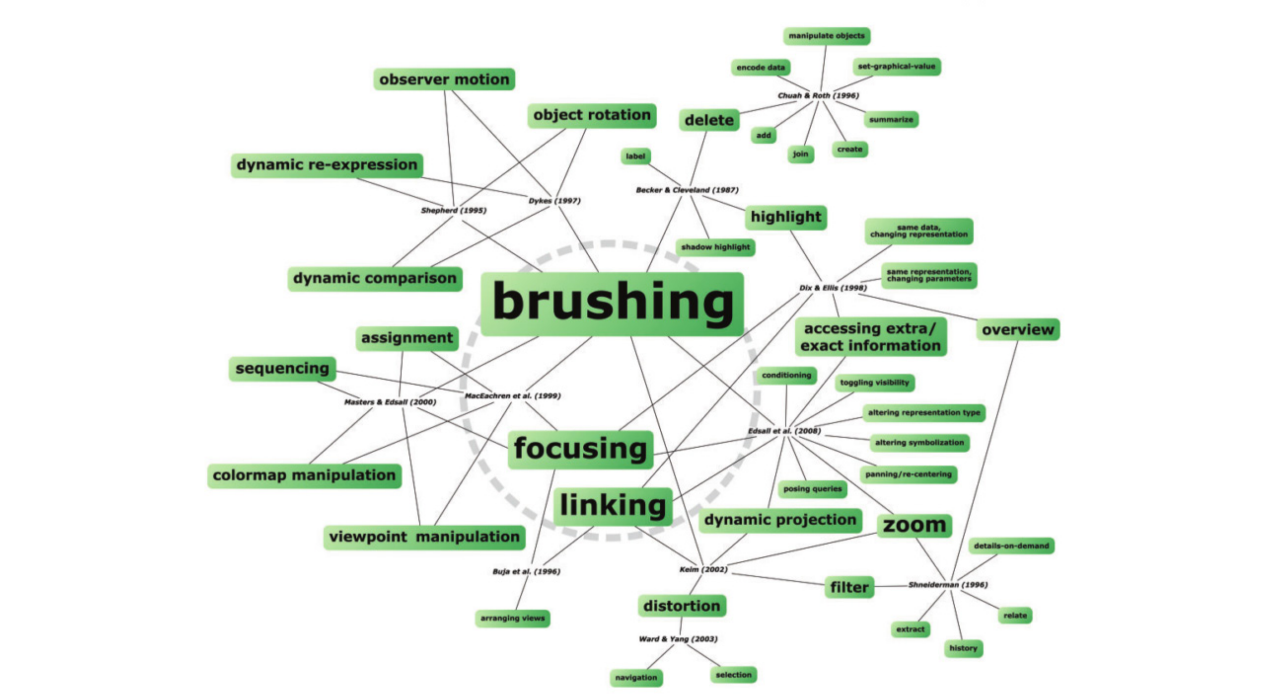
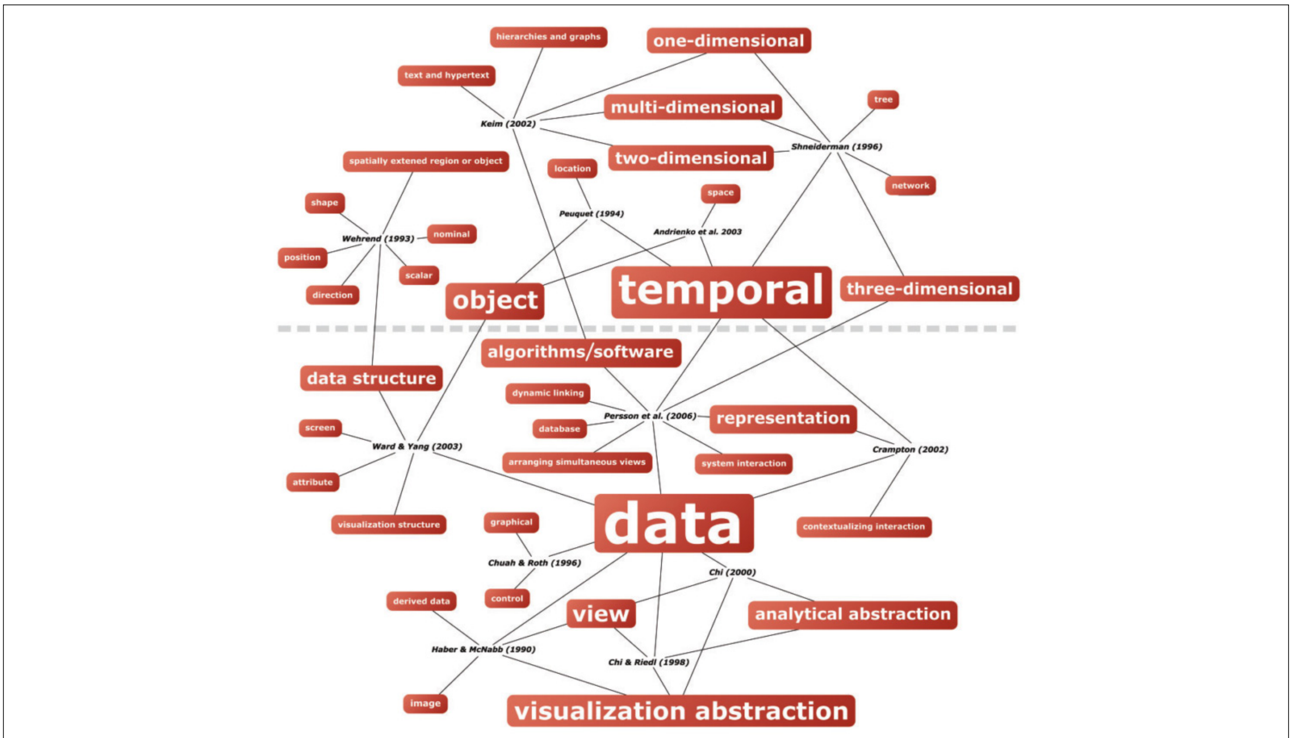


Figure 4.4: Relationships among operand-based taxonomies and the relative frequency of the included primitives.

Source: Roth, 2012



Research on the science of interaction furthermore recognizes a distinction between interactions, the action-response sequence between a human and a machine, and interfaces, the tools developed to support the interaction in a digital environment (Roth 2012; 2013). In both academy and industry, it is increasingly common to refer to this distinction as user experience (UX) design versus user interface (UI) design (Roth & MacEachren, 2014). Design is a creative activity and the results are proposed solutions. To analyze these solutions and to communicate them to the end-users at an early stage, one needs to present these solutions in the form of scenarios, mock-ups, wireframes or interactive prototypes. These presentations can be used for two purposes. First, they can function as proof-of-concept to test whether the design is acceptable and whether it can be expected to meet the needs it aims to fulfill. Secondly, the conceptual design acts as means of communication between the designer and end-users (Wassink et al., 2008).

There are many different approaches available to build the foundation upon which the actual interface can be developed. These may range from simple paper prototyping to fully interactive prototypes based on code. Two types of design are generally being distinguished: low-fidelity and high-fidelity (Wassink et al., 2008). Fidelity refers to the detail of the design. Low-fidelity designs, such as sketches and storyboards, support creativity during the design process by showing different states of the interaction design. Real interaction is, however, impossible. High-fidelity designs will look and behave like the expected end-product (Wassink et al., 2008). Whatever the status of the conceptual design, whether it is a simulation or real interactive version of the intended system, its goal is to assess design decisions and possibly to reconsider them (Wassink et al., 2008).

Common artefacts of conceptual design are concept sketches and models. This study, given the resources and time available, makes use of Adobe Experience Design CC (Adobe XD) to develop wireframes (URL 32). Adobe XD is used because it switches easily and fast between wireframing, high-fidelity designs and prototyping, all in one application. Another advantage of this tool is that there are multiple free user interface kits (e.g. wireframe templates and icon sets) and resources (e.g. tutorials and design related articles) available on the web (URL 33).

A wireframe is a low-fidelity visual representation of an interface. The purpose of a wireframe is to determine the main groups of contents, the structure of information, functionality and relationship between the different elements of the interface (Cao, 2016). Wireframes serve as a blue print and are created before any design work and coding is started. In this way, the focus is on the structure and the content of the interface without the distraction of color, font choices, logos or other visual elements (Paste Interactive, 2009). Wireframes use placeholders, such as labelled boxes, to outline the specific size and placement of elements, features and navigation controls. Creating wireframes pushes usability issues to the forefront by exhibiting the layouts at their core (Wassink et al., 2008).

4.2.3. Evaluation

During the final phase, the conceptual design and implementation should be evaluated to check whether it meets the user requirements found during phase one. Evaluation of any visualization tool or interface must include an assessment of both visual representation and interaction styles (Wassink et al., 2008). This can be done by observing the potential of the prototype interface in the context of a geovisual analytics environment (see section 2.4). This environment allows an interaction with the locational, attribute and time components of the underlying spatiotemporal data (Kveladze, 2015; Kraak & Ormeling, 2010). End-users are expected to execute different tasks that focus on these characteristics of the data. The observation of this process leads to the evaluation of the visual interface based on its effectiveness, efficiency and user satisfaction. However, due to time constraints this last stage of the UCD method could not be executed in this MSc research.

4.3. Case study

To deal with the complexity of visually exploring and analyzing concerted movements, this research restricts itself by concentrating on the coordinated positioning of defenders in space to prevent the adversary team from scoring a goal. This case is used during the second task-based session in order to apply focus and structure.

Defensive dynamics are part of the tactical dynamics that underlie how space is used. The primary tactical dynamics that underlie how space is used are threefold: tactical shape, player movements and ball movements (URL 34). The tactical shape refers to the position of a player in relation to his or her teammates. This coordinated positioning provides a team with a basic structure that enables it to more effectively occupy space in different areas of the pitch. Player movements enable a team to dynamically occupy space and help the defense to restrict space and protect space whilst retaining compactness (URL 34). Ball movement is related to restricting space, interceptions and possession. Based on the shape (i.e. length, width, depth and angles) and concerted movements of the defensive line, collective tactics can be analyzed (Kim et al., 2011).

To visually explore and analyze these defensive dynamics, this research makes use of a single set of tracking data. This dataset contains the positions of players and the ball in x- (parallel to sidelines) and y-coordinates, which amounts to approximately 1 million data points. These point features represent a chronological sequence of players and ball movements within its spatiotemporal context. With the use of this dataset, an animation is made that shows the coordinated positioning of defenders in space. More detailed information about the datasets used and animation made are provided in the next chapter (see section 5.4).

4.4. Conclusion

Most geovisualization tools and geovisual analytics environments often fail to work because the usability of these products remains unsatisfying. This research alleviates this problem by applying a User-Centered Design (UCD) approach. Developing an interface with quality of use in mind suggests that a carefully defined UCD-based strategy should be constructed. The UCD is a design and evaluation methodology that consists of three main phases with an associated set of research techniques and methods.

During the first phase, structured user interviews are combined with the think aloud method and synchronized screen and audio recording to collect qualitative data about user profiles and requirements. By combining interviews with the think aloud method during two separate task-based sessions more in-depth information of the mental processes behind actions when completing a particular task is gathered. During the second task-based session, a case study is used to deal with the complexity of visually exploring and analyzing concerted movements by applying more focus and structure. In the second phase, the requirements identified are translated into a conceptual design of the interactive visual interface by using low-fidelity wireframes. The conceptual design determines the main groups of contents, the structure of information and describes how the interface works and meets the user requirements. This low-fidelity design can be turned into a high-fidelity prototype that may be presented to and evaluated by potential users through various iterations. However, given the resources and time available for this MSc research, these final two steps could not be executed.

In the next chapter, the implementation and results of the first phase, the requirement analysis, are discussed. The goal of this analysis is to explore and analyze user profiles and requirements.

CH.5

Requirement analysis

5.1. Introduction

In the previous chapter the methodological approach for this research was shaped. This fifth chapter contains a report on the implementation of the requirement analysis and the outcomes thereof. Section 5.2 provides background information about the user profiles of the anticipated users. In section 5.3, the representativeness of these user groups is discussed. Section 5.4 provides information about the implementation of the research methodology involved. In section 5.5, the results of the requirement analysis interviews and task-based sessions are provided. In section 5.6, the quality of the collected data is being discussed.

5.2. User profiles

This research focuses on developing a conceptual design of an interactive visual interface that helps football coaches, analysts and players to visually explore and analyze concerted movements. This conceptual design may help these different user groups to deal with increasingly large volumes of movement data and provide the basis for more advanced analyses. These intended user groups differ in their abilities, capabilities, needs and requirements. This section, therefore, provides information about their basic characteristics, tasks, responsibilities and the conditions in which they will probably use the interface.

The first user profile is that of the coach, also referred to as manager. He or she is head of the coaching staff and is responsible for the overall team performance. Generally, depending on the resources of a club, the coaching staff consists of a head coach, one or more assistant managers and one or more goalkeeping coaches. Furthermore, most professional football organizations have their own medical department (e.g. fitness coach, doctor, physician and physiotherapist), scouting department and youth academy. Many clubs have several performance analysts and video analysts on the payroll, but only some have an entire analysis department. The coach and his team of analysts are working full-time to improve the performance of individual players, enhance team performance and maintain their competitive edge or improve their ranking and status.

The second defined user profile is that of the analysts. Their main responsibilities and tasks lie with providing the coaching staff, medical staff, scouting department and players with pre-and post-match video analyses and player statistics in relation to team performance, training analysis, opposition analysis and player recruitment. Most analysts currenlty work with video. In a normal day to day working situation, these analysts manually process and annotate video recordings of matches and/or training sessions for analysis and presentation purposes. Video analysts are often highly skilled in the use of encoding software. The main limitations of these qualitative game observations are that they are less objective depending on the know-how of domain experts and are known to be time consuming (see section 3.3).

Besides video recordings, most professional football organizations also collect physical parameters through numerous pieces of technology. For instance, the total distance covered and the number of high intensity runs made. These statistics are collected by the performance analyst, sport scientist or exercise physiologist and used by the medical staff to prevent injury and improve strength and stamina. Most of these so-called performance analysts have an educational background in the fields of (human) movement science, exercise physiology or sport and exercise science. On match days, video and performance analysts are involved in the live analysis, communicating with the coaching staff to review key moments and preparing footage to be relayed at half-time and full-time. After collection, most clubs store their data in external databases such as Microsoft Excel, Access or SQL in order to manage, analyze and monitor. Printed reports including graphs and diagrams are used to provide feedback for the coaching staff and players. Besides video footage and physical parameters, analysts and coaches obtain additional data through detailed performance, match, player and scouting reports.

The third and final user profile defined are the players. The number of players within the first team varies approximately between twenty and thirty. This number also relates to the number of injuries, suspensions and the number of players on loan. The players are not only referred to by their names but also very often by using their current position and role on the pitch. Generally, a distinction is made between goalkeepers, defenders, midfielders and forwards. Depending on their position and role(s) these players may vary in their technical, physical, physiological and tactical abilities. Most professional players are characterized by their spatial intelligence. For instance, elite passing midfielders like Andrea Pirlo and Xavi Hernandez are not so much defined by their passing abilities but more by their awareness of position and ability to move into space with or without the ball (Anderson & Sally, 2013). Finally, we must assume that most players don't have any experience using maps and should be considered laymen in the domain of geovisual analytics. This means that for them graphics of the data should be very clear and easy to understand.

When following UCD approaches it is also important to carefully consider the context in which the application designed will or may be used. For this research project, the application is designed to be used during pre- and post-match briefings preferably in a classroom setting, and not, for instance, on or near the pitch. In this setting, the interface is used by the coaching staff to provide, for example, pre-match presentations of upcoming opponents to be delivered to the players to assist with match preparation. This means that there is no need for a scalable system with a responsive design that also can be used on tablets, smartphones, smart watches or other other devices. When using the interface, it is important that the graphics and maps displayed are legible. The graphics displayed by the interface are too detailed for small screens and simply scaling down is inefficient, as it makes the maps difficult or even impossible to read. The application is designed to be used in a desktop environment on a personal computer or laptop. A desktop-based application is preferred above a browser-based web application because desktop applications are stand-alone in nature and do not face any hindrances resulting from internet connectivity and are protected from various security vulnerabilities (Bychkov, 2013).

5.3. Participants

The previous section showed that the main objective of this study targets three user profiles: coaches, analysts and players. To get in touch with representatives of these groups without having any connections with professional football organizations, proved to be very difficult. Fortunately the access to these user groups within the given amount of time for this MSc research project, could be facilitated by SciSports. This data intelligence company is increasingly finding its way into professional football. Besides working with clubs in the Netherlands and Belgium, SciSports is also active in England, Germany and Italy (URL 1). They provided the contact details of scouts, agents and analysts working for clubs in the Netherlands, Belgium and England. Due to the limited monetary resources, the researcher decided to approach only those working in the Netherlands and Belgium. These coaches, analysts and players have been formally approached by sending an interview request through an e-mail account of SciSports (see appendix 1). In this way, no contact details were shared with other parties and all details were kept confidential. In addition to this request, a short description of the research was attached to inform the participants (see appendix 2).

In total sixteen interview requests have been sent to representatives of each user group. Thirteen of them reacted and were willing to participate on short notice. Eventually ten interviews have been conducted (see table 5.1). Each participant has been given an identification number (P1, P2 and so on) in order to address the correct interviewee in an anonymous manner. Thereafter, based on their current function, each participant has been assigned to one of the three user groups previously described (see section 5.2). Participant 3, who works as a player agent, did not fit into one of the three user profiles. Therefore, the outcomes of the interview with him were not taken into consideration. Additionally, table 5.1 provides information about the experience that participants have working with tracking systems and football data.

Table 5.1: Overview of participants.

ID	Function	User profile	Tracking system	Data
P1	Football player	Player	STATS and Polar	Video and statistics
P2	Head of recruitment	Analyst	Catapult and Polar	Video and physical data
P3	Player agent	-	-	-
P4	Sport scientist	Analyst	Catapult and Polar	Physical data and reports
P5	Video analyst	Analyst	Catapult and Polar	Video and physical data
P6	Sport scientist	Analyst	Inmotio and SportVU	Movement data
P7	Performance analyst	Analyst	Catapult and Polar	Physical data
P8	Video analyst	Analyst	Catapult and Polar	Video and reports
P9	Performance analyst	Analyst	Inmotio and SportVU	Movement data
P10	Head coach	Coach	JOHAN Sports	Video and physical data

Most analysts (P2, P4, P5, P7 and P8) indicated that they are accustomed to work with the wearable tracking device made by Catapult (URL 22) which is combined with the Polar chest strap for heart rate sensing (URL 20). The interviews revealed that these analysts mainly used these systems to collect statistics about the physical workload of players in order to enhance performance and mitigate injury risk. This means they don't use internal GNSS-enabled device to collect positional data that can be used for more advanced tactical analyses. Furthermore, they are accustomed to use paper-based reports and simple graphs and diagrams to communicate the results found with the coaching staff, medical staff and players. Therefore, they are considered as novices in this field.

In contrast, participants 6 and 9 indicated to work with accurate movement data collected with the Inmotio system which is using the advanced LPM technology (URL 24) combined with the optical tracking system SportVU (URL 13). They also indicated to have experience with the use of different and more advanced data visualizations and interactive data visualization products like Tableau (URL 35). Both participants therefore are considered being expert users. During the analysis and interpretation of the collected data the differences between the backgrounds and characteristics of these users compared to novice user have been looked at. The answers and thoughts of these expert users have not been crucial to the development process, because the aim of this research is to develop an interface to movement data that can also be used by less experienced users.

Clearly the intended user groups are not equally represented in the sample whereas the group of analysts seems to be overrepresented (see table 5.1). This is considered as a shortcoming of this research and has several underlying reasons. Foremost almost all contact details provided by SciSports were those of analysts. Their digital address book didn't contain any representatives of the user groups coach and player. Coaches and players proved to be very difficult to reach due to the high position jobs they hold and the busy work schedules these entail. The second reason, in line with the first, is that analysts are relatively more experienced working with football data and are more familiar with the use of visualizations than coaches and players (see section 5.2). Usually these potential users are put forward by their clubs as the designated users of tracking systems. The third and final reason is that not all clubs were open and willing to participate, because they are reluctant and skeptical towards innovations and not yet convinced of the added value of interactive visual interfaces. One of the objectives of this research in particular is to remove this skepticism and make football data more accessible to coaches, players and other less experienced users.

5.4. Data collection

The qualitative user research methodology involves structured user requirement interviews and thinking aloud with synchronized audio and screen recording (see sub-section 4.2.1). The combination of these methods allows to investigate thoughts and actions, and keeps, at the same time, records of all the activities. Participants were first asked to answer several interview questions aimed at gathering information regarding the collection, analysis and visualization of movement data. These questions can be found in appendix 4. As this appendix is in Dutch, more details about the nature of the questions can be found in table 5.2 and the paragraph below.

Table 5.2: English translation of interview questions.

ID	Question
1	How would you describe your role based on your tasks and responsibilities?
2	Have you ever worked with (the results of) movement data?
3	With the help of which tracking system(s) does your club works to collect movement data?
4	How are these systems used, how often and by whom?
5	With the help of which systems is movement data analyzed and by whom?
6	Are there any problems you encounter during the analysis of movement data?
7	What kind of information can you currently get working with (the results of) movement data?
8	What kind of information can you currently not get working with (the results of) movement data?
9	Do you use visualizations during the exploration and analysis of movement data?
10	Are there analytical tasks in which you would like to be supported by visualizations?
11	What kind of insights should visualizations support?

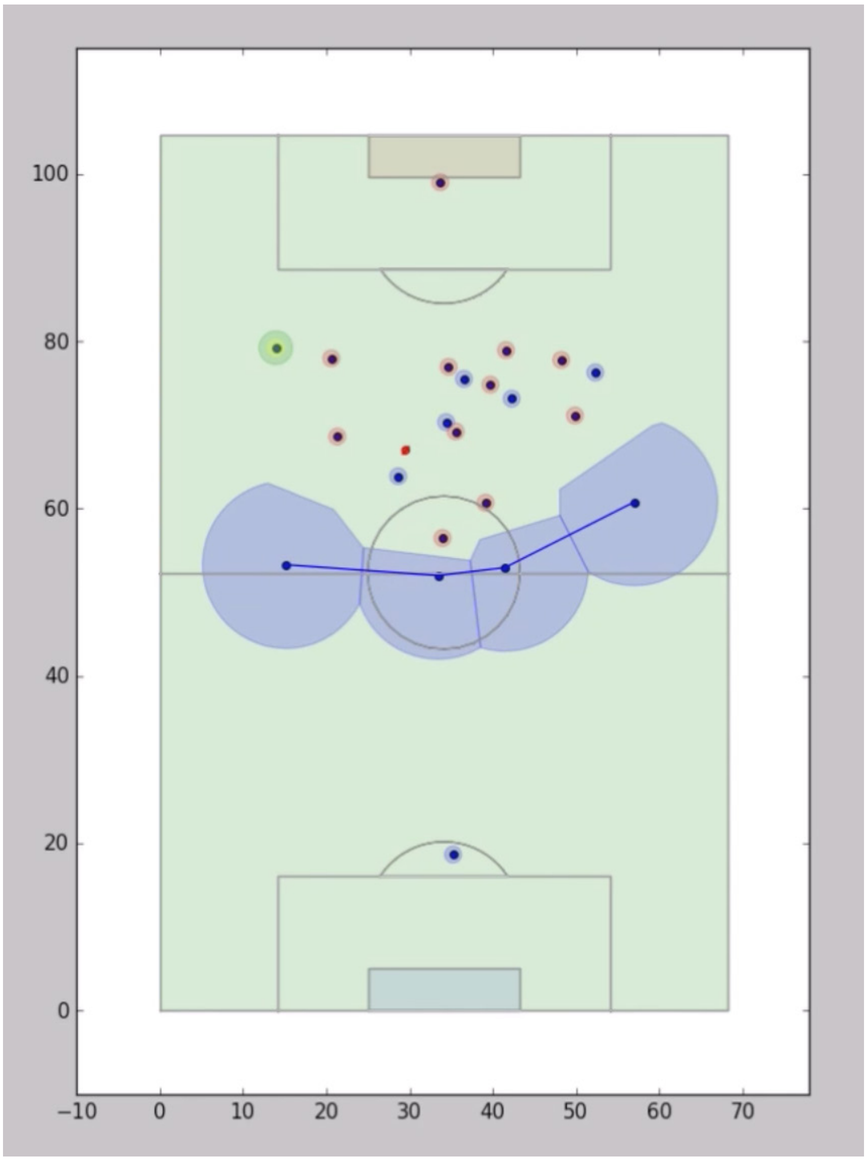
The first question aims to get familiar with the tasks and responsibilities of domain experts in their current working environment. Subsequently they are asked about their experience working with (the results of) movement data. Then two questions follow that focus on the use (i.e. how, how often and by whom) of different tracking systems to collect movement data. Questions 5 and 6 consider the analysis of movement data and the analytical problems domain experts experience working with this type of data. Furthermore, domain experts are asked about the type of information they can or cannot currently get from (the results of) movement data. Finally, domain experts are asked about the usage of visualizations during the analysis of movement data, the tasks in which they would like to be supported by visual tools and the insights these should support. At the end of each interview domain experts are encouraged to discuss additional comments, suggestions and ideas.

During the first task-based session, participants were asked to use the SoccerStories working environment and speak out loud about their thoughts and feelings (see appendix 5). SoccerStories is an interactive visual interface which provides an overview of game phases and a detailed series of connected visualizations, in which each view only focusses on one specific action (see sub-section 3.4.1). The event dataset used for this visual analytics system has been provided by Opta Sports (Perin et al., 2013). This dataset contains events (i.e. corner kicks, crosses, the distribution of shots, long runs and passes) collected during eight matches in which either FC Barcelona, Real Madrid CF or both teams played during the 2012-2013 season.

During the second task-based session an animation of approximately one minute was shown multiple times which highlights the coordinated positioning of defenders by using points, lines and polygons (see section 4.3). Participants were again given the task to look for the information they would like to obtain during the visual analysis and exploration of movement data (see appendix 6). To develop this animation a single set of movement data, obtained via SciSports, has been used. The positions of players and ball are defined by a single pair of geographical coordinates captured with the use of the Inmotio tracking system. Speed has been calculated at each point and provided as an attribute to the data. Furthermore, the names of each player including identification numbers have been provided. The data have been collected during the 2015-2016 season.

After collection, the data have been stored as a CSV-file. The animation displays the concerted movements of first team players during a 11 versus 11 training match of a professional football club. The ball is shown as a relatively small red dot and the players are visualized as black dots surrounded by either a red or a blue circle to indicate the team they play in. There is one player however playing in the blue team who is surrounded by a green circle. This was a mistake in the visual output of the data (see figure 5.1). The four defenders of the blue team are connected by straight lines that highlight their collective motion. The space they aim to control is indicated by means of a Voronoi diagram. The changing shape and size of these polygons have been calculated based on the points (teammates, opposition and ball) closest to a player.

Figure 5.1: Screenshot of the animation shown during the second task-based session.



The sessions with potential users took place between the 24th of January and the 15th of February 2017. Almost all sessions took place at the current working location of the interviewees in order to make them feel more comfortable. Two sessions were conducted at another location suggested by the participants themselves. At the beginning of each session, each participant was given a printed document containing a brief introduction to the topic and explanation about the purpose of the session, the procedure and user research ethics (see appendix 3). In addition, each participant was informed about the user research methods to be applied. The same document was also sent as a pdf-file prior to the interview by e-mail to give each participant the chance to prepare. The average duration of each session was an hour.

For the task-based sessions, a uniform working environment has been created by using one laptop which runs a Mac OS Sierra workstation with a built-in retina display of 13,3 inch at a 2560 by 1600-pixel resolution, and a wireless mouse for a more flexible manipulation. While thinking out loud, user activities and cursor manipulations on the screen were recorded using the logging software QuickTime Player. The sessions were audio recorded, capturing the think aloud actions of the intended users and the interviews with them. This is important to minimize the influence of interpretation and confirm the real meaning of their pronunciations. For these audio recordings, an IPhone 5S has been used.

The collected audio and screen recordings have been transcribed with use of the web application Otranscribe (URL 29). The finished transcripts have been downloaded as text files and imported into Microsoft Word. With the use of this text editing software, each transcript has been given the same lay-out. These final transcripts were saved and sent to the participants in order to ask for their approval. In this way, participants had the opportunity to give their feedback. If they did not agree with the translation of the interview they were given the opportunity to adjust. In total, five out of ten participants responded and only one of them made some relatively small adjustments.

5.5. Analysis of the resulting data

The answers to the interview questions and the information derived from the thinking aloud were brought together into written verbal and action protocols generated with the help of Otranscribe (URL 29). These transcripts do not contain literally everything said and done during the one-hour sessions, but are limited to those parts of the session directly related to the objectives of this study. For every participant, a separate document has been made and was about 10 pages long (see appendix 7). To discover relevant connections and relations within this unstructured data and quickly find new valuable insights the encoding software NVivo has been used. With the use of this tool, pieces of text can be selected and grouped into nodes, also known as codes. Coding is the iterative process of organizing and sorting qualitative data.

For this study, the same structure applied as in chapter 3 has been used to categorize the resulting unstructured data into three main themes: collect, analyze and visualize. In addition, the main code 'interaction functionality' has been added to the codebook after reading and analyzing the transcripts a second time. It appeared that during both task-based sessions participants had valuable thoughts and ideas about the way the interface should function and interact. Therefore, this code has been added to the coding scheme in order to incorporate this as a user requirement (see table 5.3). In addition, the key concepts discussed in chapter 2 combined with own knowledge and expectations about the topic have led to seven sub-themes (see table 5.3). Based on this classification, all transcripts have been read thoroughly, analyzed and coded in order to find distinct concepts and categories in the unstructured data.

Table 5.3: Coding scheme used to analyze the resulting unstructured qualitative data.

ID	Code	Description (passages in transcripts that refers to…)
1.0	Data	Different types of football-specific spatial data used by domain experts. For instance, movement data collected with the use of tracking systems.
1.1	Movement parameters	The basic characteristics of a moving object at any given timestamp. For example: distance, direction and acceleration.
1.2	Playing system	The influence of the system (or style) of play and vision of the coaching staff on the interpretation of movement parameters.
1.3	Context	The environments in which players and ball move and the other objects with which they interact.
2.0	Analytical questions	Questions of domain experts that the system should be able to answer in an intuitive way.
2.1	Spatial	Questions in which space matters: where?
2.2	Temporal	Question in which time matters: when?
3.0	Visualization	Graphics and diagrams that support the visual exploration and analysis of football-specific data.
3.1	Usage	The use of visual tools and interfaces to derive insight from and support the visual exploration and analysis of movement data.
3.2	Insight	Insights that domain experts currently get from working with movement data.
4.0	Interaction functionality	The functionalities the system should incorporate to allow users to interact with the data at hand.

The remainder of this section is divided into four sub-sections, each focusing on one of the four main themes and corresponding sub-themes shown in table 5.3. In sub-section 5.5.5, an overview of the resulting user requirements is provided.

5.5.1. Data

As previously stated it all starts with the collection of data by professional football organizations themselves or by commercial data providers (see section 3.2). By now most clubs mainly explore and analyze event data, video recordings and physical parameters. Despite their lack of experience working with movement data, participants P1, P4, P7, P8 and P10 made some interesting remarks about the interpretation of these data. They indicated that the way movement data are interpreted is crucial for the success of the system. One participant expressed this as follows:

> *"If you have a lot of information at your disposal, the way you interpret these data is very important. I think they [data] can contribute a lot if they are made easily accessible for coaches, analysts and players" (P8)*

But how can movement data be made more accessible for the intended user groups? To answer this question participant 4 stressed the importance of critically evaluating the collected movement parameters. In sub-section 2.3.2 these parameters have been defined as *"the derivatives of movement trajectories that can be used to deduce the basic characteristics of a moving object at any given timestamp*" (Gudmundsson & Horton, 2016). Indispensable parameters to be included into the system are quite numerous: mutual distances to all other moving objects, spatial distribution, current direction, change of direction, current speed, potential speed, orientation, acceleration, ball position and interaction spaces (P1, P4, P6, P7 and P10). Concerning acceleration and ball position, participant 6 pointed out that measuring those parameters is still difficult because the data, collected with the optical tracking system SportVU (URL 13), is simply not accurate enough. Besides participants 3 and 6 stated that the problem with calculating interaction spaces is that differences in maximum speed between players are not sufficiently considered.

According to participant 4 the way these parameters are assessed and weighted is crucial to the way movement data are interpreted. Due to the lack of consensus between domain experts about the actual definition of these parameters and their applicability however, it remains problematic to translate event data into measurable units in an objective way. When the amount of movement data and the number of different and more detailed parameters keeps growing at this pace it may become problematic, participant 4 concludes.

"The combination of different parameters and their applicability. That is probably a black box that will never be unlocked". (P4)

Furthermore, multiple participants (P4, P6, P8, P9 and P10) mentioned that the assessment, definition and weighting of movement parameters is strongly influenced by the vison of the coach and depends on the playing system he or she propagates. it is almost impossible to visually explore and analyze concerted movements, not knowing the tasks and responsibilities of the players and the instructions they are given (P10). Therefore, it is required that the system includes a function that makes it possible to incorporate the vision of the coach and the playing system in order to link these to the movement observations collected. This important system requirement has been well articulated by participants 9 and 6:

"We can all generate data, but data are no information yet. The playing system should be the interpretation mechanism and nothing else (…) it determines the value you attach to the various parameters" (P9)

"The playing system is going to have so much influence on the information you search for. (…) this influences the way you build such tool or visualization eventually. It makes it very difficult". (P6)

Another relevant problem indicated by almost all interviewees, is the lack of contextual data. This type of data refers to spatiotemporal data defined as the description of space and time within movement occurs. Spatial context refers to the characteristics of the landscape, the surrounding environment, the presence of other objects and ambient attributes that may affect movement (Kotzbek & Kainz, 2015; Andrienko et al., 2013). Whereas temporal context refers to the temporal cycles that may be relevant to the concerted movements under investigation (see section 2.2). Without knowledge of the context the interface cannot provide a good representation of reality. During the first task-based session multiple participants (P2, P4, P5, P6, P7 and P8) confirmed this problem by stating that the visualizations of the SoccerStories environment lacked contextual data. SoccerStories provides an overview of game phases where each view only focusses on one specific action (see sub-section 3.4.1).

"I am more interested in how this opportunity came about. The lines [passes] are therefore more important than the actual result" (P8)

According to participant 2 not the number of shots on target is crucial, but the considerations behind the shot. In addition, participant 7 noticed that SoccerStories only provides a single part of the complex system. Tactics and strategy in football are much more complicated because they depend on the concerted movements of all other moving objects through space over time. Without this understanding it is nearly impossible to objectively determine the quality of a player and answer complex spatiotemporal questions (P2 and P10).

"In this situation, I don't see the relationship with the other players. For me, it is always about the context" (P2)

The above statement supports the idea of producing a two-dimensional animation participants were confronted with during the second task-based session. Since football is all about the union of space and time, movement data seem suited to add a part of this missing context by providing answers to the 'when?' and 'where?' questions. According to participant 7 movement data give a more objective overview of the situation compared to video recordings. Participant 9, experienced in working with movement data, agreed, and pointed out that movement data are necessary in football analytics. He added there is also a sincere need for more and accurate 'off the ball data'. Off the ball data are parameters focusing on the way a team is playing when the opposition is in possession (P9). Currently, data providers such as Opta Sports do not own their own tracking system which records every movement. If a player does not interact with the ball (i.e. event), nothing is registered.

"On the ball data don't say anything about the player in relation to the style and playing system. In this case position data are extremely important" (P9)

Finally, participants (P4, P5, P6 and P7) added that they would like to have a system that combines video recordings and physical parameters with movement data into an integrated visual analytics environment. They indicated that this link is currently absent which leads to a lack of knowledge:

> *"Currently I can only say something about the physical condition of a player compared to others and our expectations. I still miss movement data to see whether his performance was good or bad". (P7)*

5.5.2. Analytical questions

To visually explore and analyze concerted movements the interface should be able to provide answers to the spatial and temporal questions coaches, analysts and players need support for. These analytical questions should, therefore, be translated into system requirements and answered in a simple and intuitive way with the use of (geo)visual analytics tools. Analytical questions regarding the spatial component focus on the coordinated change of location (i.e. movement) of players and ball at a certain timestamp. For instance, participant 10 wanted to know which specific area in front of the goal his team should protect during the defensive phases in order to prevent the adversary team from scoring a goal. Another example has been given by participant 2 during the second task-based session. During the visual analysis and exploration of the animation he wanted to find out from where and when to the central defender had sent his passes.

> *"For me this [the animation] is fantastic. It shows in one glance what happens during the game. Based on this the coach can ask question like: where do you stand, why do you stand there, where is the ball and what do you need to do next?" (P2)*

Besides these relatively simple spatial questions many participants (P1, P2, P4, P6, P8 and P9) wanted to know 'how players collectively move and orientate themselves in relation to the ball, the opposition and each other?'. During the interviews, they referred to this type movement as tactical movements. In this research, these tactical movements are defined as concerted movements. These are the coordinated movements of multiple objects in relation to each other and the spatiotemporal context in which they move (see section 2.3). This means that the system must be able to visualize these movements in an effective and efficient way.

"I am interested in the way I move with respect to my teammates in both the offensive and defensive phases" (P1)

Another equally interesting topic is that participants (P2, P6 and P8) indicated they would like to know whether a certain event or movement happened just once or if it is a reoccurring pattern. To be able to recognize these movement patterns it is necessary to incorporate the spatiotemporal dimension into the system. As mentioned in sub-section 2.3.3, spatiotemporal movement patterns in football need to reflect collective group actions or reactions to achieve a certain goal. For instance, 'how long does it take before the adversary team has been switching sides (P8)?', 'how often did we lose or win the ball during the ten times we gave high pressure forward (P4)?' and 'what is the density of certain variables [parameters] to occur and how do they develop over time (P9)?'. Finally, participant 4 added that it would be helpful to see those moments back in one visualization instead of several separate video clips. The challenge, therefore, is to develop an interactive visual interface that incorporates and answers these spatial and temporal questions in an understandable way.

'How are we going to visualize those things [patterns] over multiple games?' Those are questions very difficult to answer and visualize" (P6)

5.5.3. Visualization

The previous sub-section concluded with emphasizing the important role visualizations play in the exploration and analysis of spatiotemporal movement patterns. At present, most analysts and coaches prefer the usage of video and statistics during their pre-match and post-match analysis because these data are easily accessible and interpretable to them. Relatively standard graphs furthermore are used to monitor heart rate, show recovery and compare the potential of new players to those who have reached their potential. Only participants 6 and 9 indicated to have experience working with movement data and visualizations. With the use of software provided by Inmotio, video recordings can be combined with two-dimensional and even three-dimensional animations, statistics and graphs. The two-dimensional animations show, for instance, the constantly changing mutual distances between players connected by lines (P9). Players can use this kind of animations to visually explore important moments during the game.

"It is beautiful to visually compare these moments with each other, maybe there is a pattern to detect. (…) when you find one, that is something you should share with the coach." (P8)

Participant 6 however remained skeptical during the entire interview about the usability of these two- and three-dimensional visualizations and stressed that the actual usage of these tools is still quite limited. In addition, participant 6 noticed that almost all data are stored into a SQL-database and is semi-automatically transformed into graphical reports with the use of Tableau (URL 35) or Qlik View (URL 36). These and other programs, such as SAS Visual Analytics (URL 37), Spotfire (URL 38) and Domo (URL 39), are intended to help businesses to visualize and understand their data by building dashboards/interfaces. Participant 6 indicated that the reports could easily be transformed into web-based or desktop applications. The coaches, even the younger ones, nevertheless, still want to have printed reports on their desk to write and draw on with a pen, participant 6 stated.

"Even though we can make attractive dashboards, they [coaches] do not use it. Until then, it is just a waste of time". (P6)

Another recurring theme is the desired incorporation of the playing system into the interface. According to participant 2 visualizations should always support the playing system and the vision of the coach. The instructions the coach gives to his players you want to see in a clear and understandable visualization (P2). Furthermore, participant 6 said that visual information is mostly above all used by the coach. Off course it may also be used by the performance analyst or other staff members, but most issues take place on a tactical level. This is the level in which the coach and his assistant managers act (P6). Depending on the playing system the kind of things the coaches want to know from there analytical staff are always different. Therefore, participant 6 thinks a visual analytical system should always be custom-made. Especially clubs that have their own data department or a relatively large team of analysts are not eagerly waiting for a tool generally applicable. These clubs prefer making their own tools and visualizations because this works much faster, participant 6 said.

In one fragment [two-dimensional animations] players can see their mutual distances. They know their tasks within the playing style (…) look what happens here". (P9)

5.5.4. Interaction functionalities

To be able to work with the data, the geovisual analytics environment needs a proper functionality. It should support a smooth and consistent transition between components when the user makes changes for visualization and exploration purposes (Kveladze, 2015). Participant 8 indicates that it would be helpful to have multiple different windows instead of a single detailed display showing too much information at once. This concept is in line with the overview and detail interface of game phases proposed by Perin et al. (2013). Additionally, during the second task-based session, participant 1 pointed out that it would be interesting to combine the events shown in the two-dimensional visualization with the actual video recordings. When an object is selected in one view (e.g. video) it should automatically be highlighted in the other views as well (P1, P2, P5 and P8). This process is called interactive (dynamic) linking and aims to highlight an object in multiple coordinated views. Another technique, called brushing, aims to identify subsets of data through selection (Kveladze, 2015). Both techniques highlight the selected observations across various visualizations by providing a dynamic link between maps and other types of graphical displays to communicate different aspects of the data.

"We could also create links between other players (…) you could link the entire defense for example. (…) in this way, we could see how we were moving in relation to each other". (P1)

Besides linking and brushing, the system should also be able to execute other tasks: identify objects, select objects by clicking, filter events and draw points, lines, arrows and other basic shapes such as a rectangle, ellipse and polygon (P1, P2, P8). Additionally, participants P2, P5 and P8 indicated that the use of colors and shapes could help to make player roles and playing system more observable and recognizable. The adversary team could for instance be distinguished by using a different shape such as a triangle instead of an ellipse (P2). Another example is the usage of color to indicate the different roles players possess during the game (P8). Also, colors (e.g. red and green) might be used to indicate how many times the team did what it was supposed to do, or what the team should not have done (P8). Furthermore participants 5 and 8 indicated they liked the way the timeline in the SoccerStories environment could be used by selecting events that led to a shot on goal (see sub-section 2.4.1). Therefore, the interface should also include an interactive timeline in which key moments can be selected by clicking on them.

5.5.5. Summary

In this sub-section, an overview of all previously discussed user requirements is provided (see table 5.4). In the next chapter, these requirements will be translated into interface content and interaction functionalities.

Table 5.4: Overview of user requirements that will be incorporated into the conceptual design

ID	Requirement	Description (the interface should …)
1	Off the ball data	Display parameters that focus on the way a team is playing when the opposition has possession.
1.1	Indispensable parameters	Display the parameters: mutual distances to all other moving objects, spatial distribution, current direction, change of direction, current speed, potential speed, orientation, acceleration, ball position and interaction spaces.
1.2	Playing system	Allow the user to incorporate the playing system and link these to the movement observations collected.
1.3	Contextual data	Display knowledge about the context in order to give a better representation of reality.
2	Geovisual analytics tools	Provide answers to analytical questions in a simple and intuitive way with the use of geovisual analytics tools.
2.1	Concerted movements	Display concerted movements in an effective and efficient way.
2.2	Movement patterns	Incorporate the spatiotemporal dimension to recognize reoccurring movement patterns.
3	Integrated system	Integrate video recordings, statistics and movement data into one integrated visual analytics environment.
3.1	Printable	Allow the user to print and share reports and visualizations made with help of the interface.
3.2	Custom-made	Adopt to the playing system and incorporate tools and visualizations the clubs prefer.
4	Linking and brushing	Allow the user to use linking and brushing techniques to highlight and identify an object in multiple coordinated views.
4	Exploration functions	Allow the user to use identify, select and drawing tools to explore the movement data at hand.

5.6. Data quality

During the last decades, researchers have become more aware of data quality issues when dealing with qualitative research. The quality of qualitative research refers to the reliability and validity of the research design (Maxwell, 2013; Bryman, 2012). Reliability refers to the degree in which other researchers performing similar observations would produce comparable results. In addition, reliability depends on the consistency within the employed analytical process (Noble & Smith, 2015). Validity refers to the application of the methods undertaken and evaluates the credibility whether or not an interpretation accurately reflects the collected data. External validity (i.e. generalizability) refers to the degree to which research findings are transferable to other specific settings than those directly studied (Baksh, 2012; Noyes et al., 2011).

In this research, multiple strategies to enhance the reliability and validity have been applied. First, a structured and standardized procedure has been implemented throughout the project to ensure a quality output that is reproducible and methodologically sound (see chapter 4). This includes that each participant has been approached in a similar way (see appendix 1), was given the same introduction to the topic (see appendix 2; 3), confronted with a consistent and standard interview topic list (see appendix 4) and given the same instructions during both task-based sessions (see appendix 5; 6). Second, the researcher explained what (and how and why) steps were taken and which choices were made during the research process in order to be as transparent as possible (see section 5.4). Finally, even though novice and expert users were given the same interview questions and the same tasks to perform, this research also considered the influence of their different characteristics on the research its outcomes. This means that while interpreting the collected data their backgrounds have been considered to develop an interface that all of them can use (see chapter 6).

Interviews have furthermore been combined with the think aloud method and synchronized screen and audio recording. The combination of different methods to observe the same topic from different perspectives is known as (method) triangulation and leads to greater credibility (Baksh, 2012). Another approach to improve the validity of this research is known as member checks. Finished transcripts were presented to the participants with the request to check the correctness of the collected data.

5.7. Conclusion

In this chapter, the implementation of the requirement analysis and the outcomes have been discussed. For this research, three potential user profiles have been selected and defined: coaches, analysts and players. However, representatives of these groups are not equally represented in the sample whereas the group of analysts seems to be overrepresented. The qualitative user research methodology involved structured user requirement interviews and thinking aloud with audio and screen recording. This combination of different methods to observe the same topic from different perspectives is known as triangulation and leads to a greater credibility. During all sessions, a uniform working environment was created by using the same laptop. Furthermore, a structured and standardized procedure has been implemented throughout the project to ensure a quality output that is reproducible and methodologically sound.

To discover relevant connections in the unstructured data and quickly find new valuable insights the encoding software NVivo has been used. Based on the code scheme made, the answers that participants have given are divided into main and sub themes of analysis (see table 5.3). In sub-section 5.5.5, a summary of the most important user requirements can be found (see table 5.4). By utilizing this knowledge, the next chapter aims to develop a conceptual design of an interactive visual interface that may help coaches, analysts and players to deal with increasingly large volumes of movement data and provide the basis for more advanced analyses.

CH.6

Conceptual design

6.1. Introduction

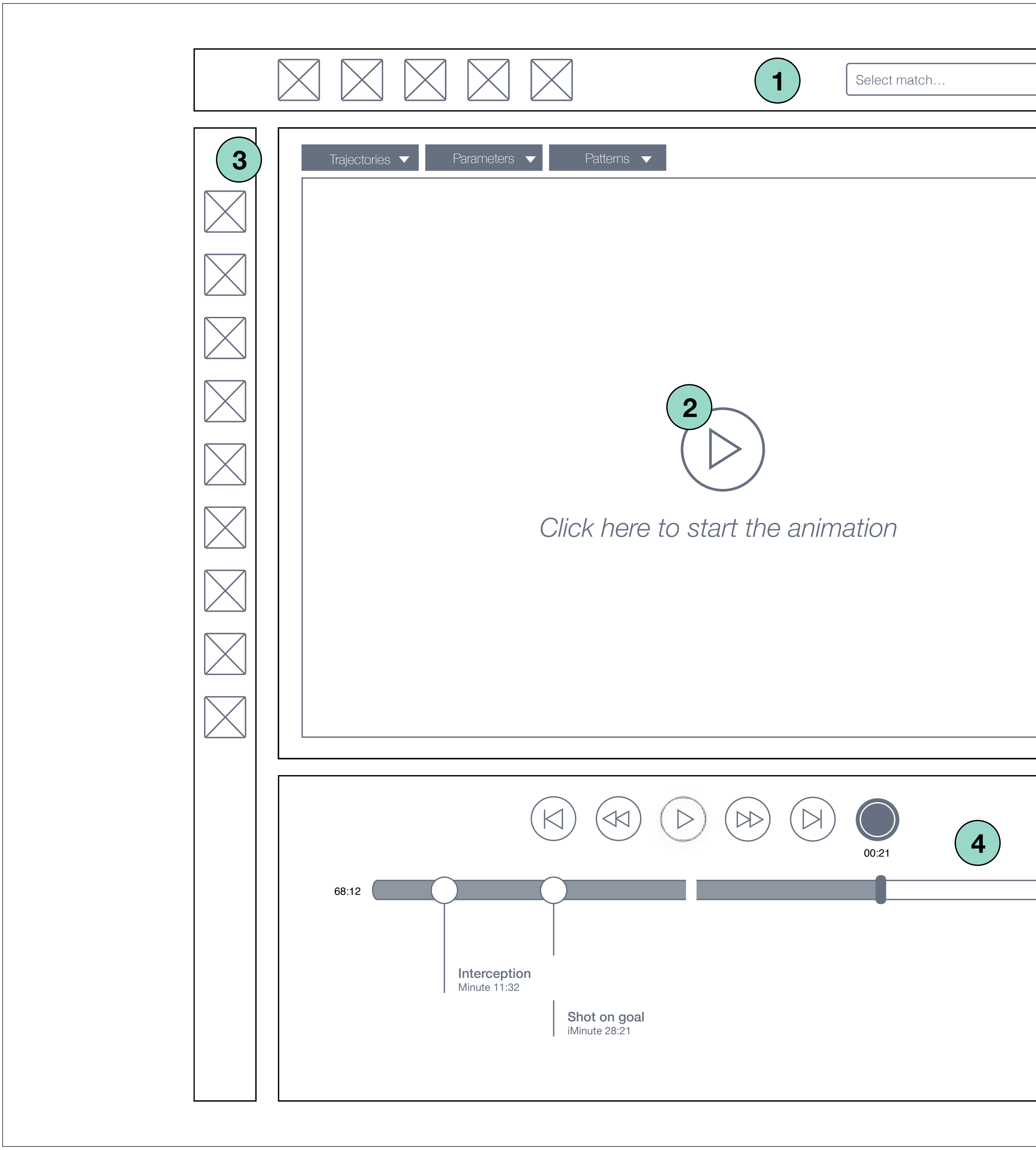
In the previous chapter a report on the implementation of the requirement analysis and the outcomes thereof was provided. The user requirements collected during this phase are used as input for the conceptual design and should be answered by the application in an accessible and understandable way. This sixth chapter contains a description of the development process that leads to the visual interface and the interaction functionalities it supports. The conceptual design provides the required information on the key concepts outlined in chapter 2, and provide answers to the analytical questions and user requirements discussed in chapter 5.

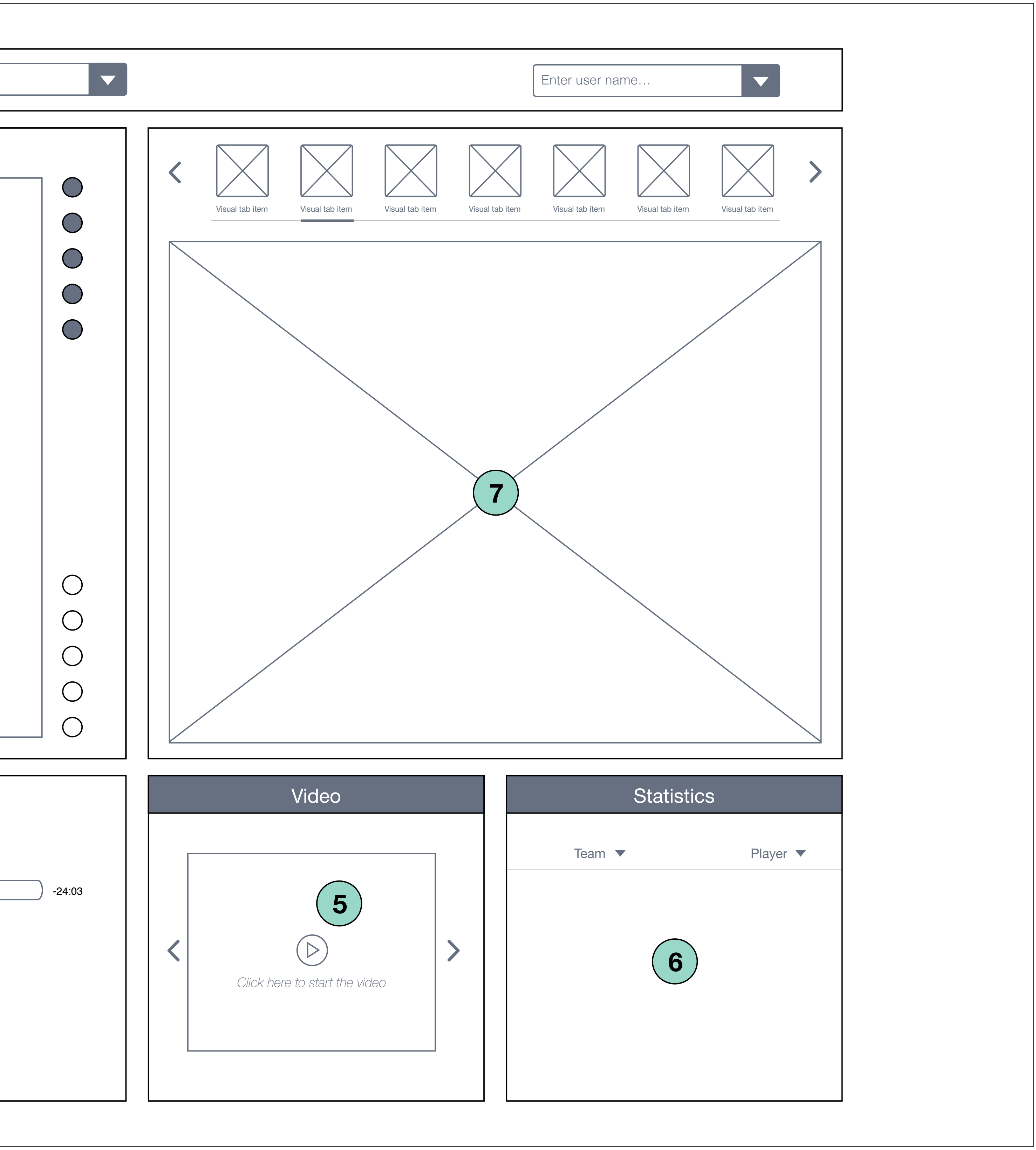
This chapter starts with a description of the distinction between user interface (UI) design (see section 6.2) and user experience (UX) design (see section 6.3). In section 6.2, a low-fidelity wireframe is presented that shows the interface developed for this research project. Section 6.3 discusses how the different content areas are linked to each other and work together. In the sections 6.4 till 6.7 subsequently more detailed design recommendations and interaction functionalities regarding the interface are discussed per panel.

6.2. User interface design

This section presents the visual interface developed for this research project consisting of seven separate panels: (1) navigation, (2) animation, (3) exploration, (4) timeline, (5) video, (6) statistics and (7) geovisual analytics panel (see figure 6.1). Besides multiple blocks of content, the wireframes made for this study also contain more detailed elements such as buttons, time sliders, video placeholders and dropdown lists. Placeholders are used to mark the location of images, video or other fixed content (URL 40; URL 41). The elements that are included in the visual interface are enumerated in table 6.1.

Figure 6.1: Overview of the visual interface developed for this study. This low-fidelity wireframe provides a visual guide suggesting the layout of interface content. Elements per content area are enumerated in table 6.1.




Enter user name…
Visual tab item
Visual tab item
Visual tab item
Visual tab item
Visual tab item
Visual tab item
Visual tab item
7
-24:03
Video
5
Click here to start the video
Statistics
Team
Player
6

The requirement analysis made clear that currently most target users are laymen in the domain of geovisual football analytics. The content and the elements included in the visual interface should therefore speak for themselves and provide feedback when necessary. Being explicit reduce some of the initial learning curve (Cao et al., n.d.). Wireframes provide a basic visual guide that suggest the layout and structure of fundamental user interface elements. These force the reviewers to debate the arrangement of components, rather than focusing on how they may eventually look. They make the transition to high-fidelity designs and eventually interactive prototypes based on code easier, saving valuable time and effort later in the development process (Bank & Zuberi, n.d.).

Figure 6.1 (see full screen view on the previous page): Overview of the visual interface developed for this study. This low-fidelity wireframe provides a visual guide suggesting the layout of interface content. Elements per content area are enumerated in table 6.1.

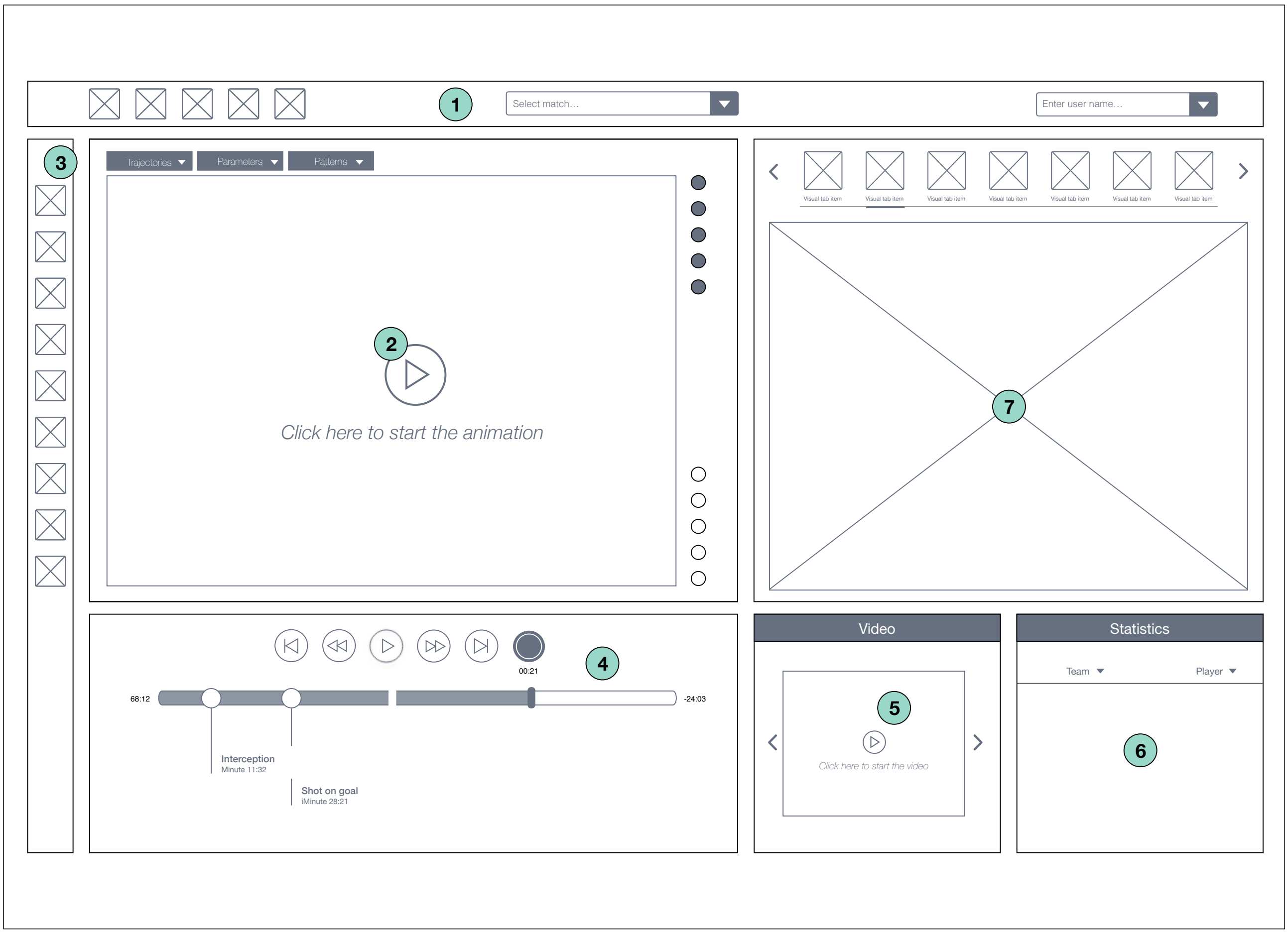


Table 6.1: Content inventory of the user interface elements. The visual interface developed for this research consists of seven panels (see figure 6.1). Each panel contains one or more separate elements as shown in this table.

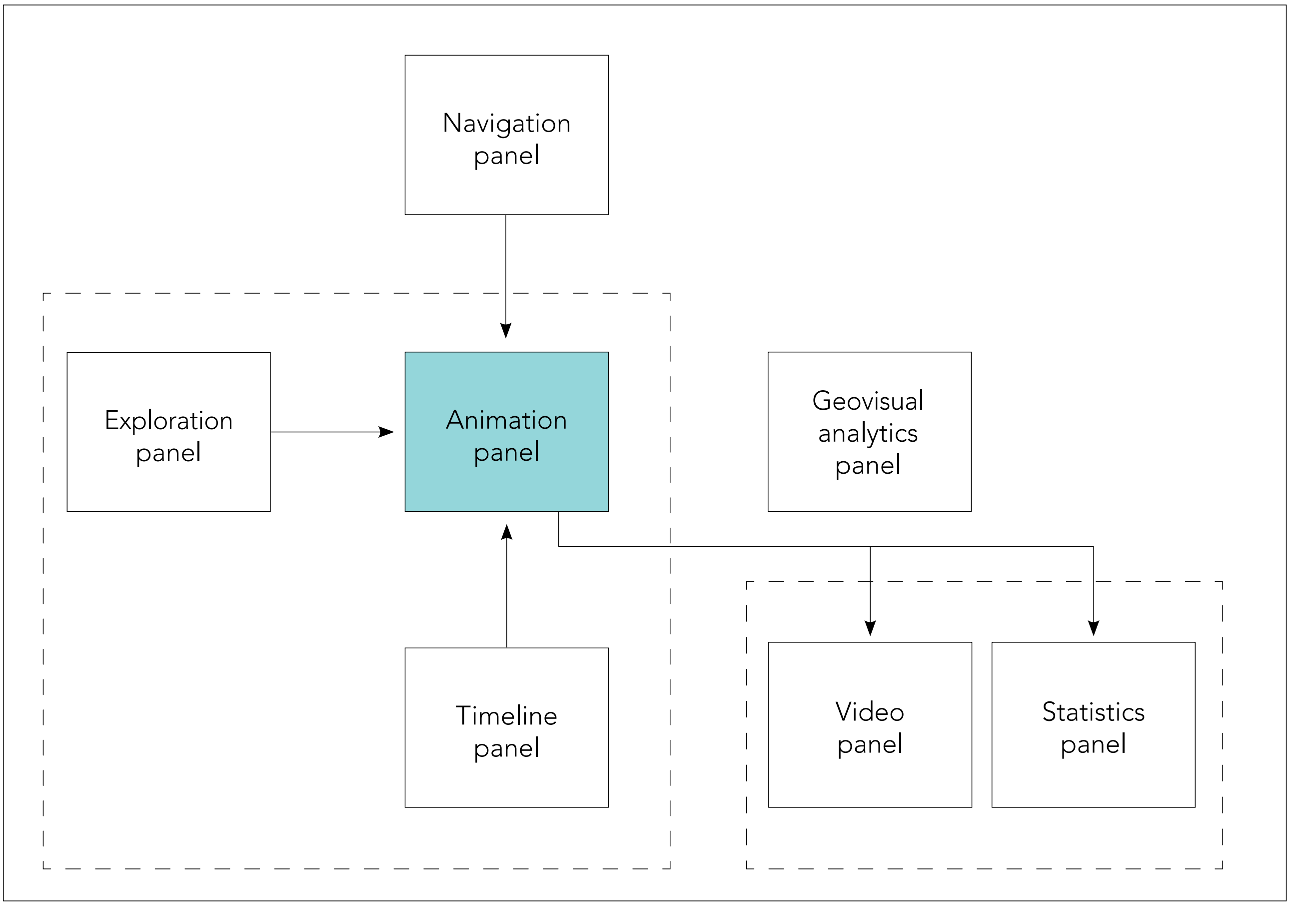
ID	Interface content	Interface elements
1.0	Navigation panel	Header
1.1	Application controls	Buttons
1.2	Match selector	Dropdown list and search field
1.3	Setting controls	Dropdown list and text field
2.0	Animation panel	Content area
2.1	Animation placeholder	Window
2.2	Trajectories	Tab and dropdown button
2.3	Parameters	Tab and dropdown button
2.4	Patterns	Tab and dropdown button
3.0	Exploration panel	Sidebar
3.1	Exploration tools	Buttons
4.0	Timeline panel	Content area
4.1	Playback controls	Buttons
4.2	Timeline	Progress bar and slider
4.3	Annotation fields	Text fields
5.0	Geovisual analytics panel	Content area
5.1	Visual analytic tools	Carousel with buttons
5.2	Visual analytic placeholder	Window
6.0	Video panel	Content area
6.1	Title bar	Heading
6.2	Video placeholder	Window
6.3	Video recordings	Video carousel
7.0	Statistics panel	Content area
7.1	Title bar	Heading
7.2	Team statistics	Tab and dropdown button
7.3	Player statistics	Tab and dropdown button

6.3. User experience design

As shown in figure 6.1 and table 6.1, the visual interface consists of seven separate panels that are linked together using the operator-based approaches brushing and linking (see sub-section 4.2.2). These techniques make manipulation of the data possible and allow users to explore and analyze different aspects of the data. This type of environment benefits from the idea that insight is formed through interaction (Roth, 2013).

With help of the navigation panel users start the analysis by importing a new dataset, connect to a database or opening an existing project. After the movement data have been imported into the system and a certain match has been selected, the visual exploration process continues in the animation panel. This panel is the focus view of the application and is linked to the timeline and exploration panels. These panels together form the primary focus point from where the visual exploration and analysis of movement data starts. In addition, the animation panel is connected to the video and statistic panels to highlight and identify an object in multiple coordinated views. After the exploration phase, users can turn to the geovisual analytics panel where visual analytics tools can be used to obtain more detailed information about the movement patterns displayed in the animation panel and statistics shown in the statistics panel. A schematic overview how these panels are linked together is provided in figure 6.2.

Figure 6.2: Overview of interaction functionalities between the seven panels shown in figure 6.1. The animation panel is placed at the center of the application highlighted by a blue box. The arrows indicate the direction of interaction. Together the animation, exploration and timeline panels are primary focus point from where the visual exploration and analysis of movement data starts.



6.4. Navigation panel

At the top of the visual interface, occupying the full-width of the screen, the navigation panel is placed (see figure 6.3). This panel has been shaped like a bar and is placed as a header to control, search for and evoke actions across the application. This position is chosen because most users first scan the horizontal line on top of the page. Designs following the 'reading pattern' of users improve the readability and learnability of the interface (Cao et al., n.d.). The navigation panel is always visible and there are no controls to hide it from view.

Figure 6.3: Overview of the navigation panel. This panel consists of three separate elements. From left to right: application controls, match selector and setting controls (see table 6.1).

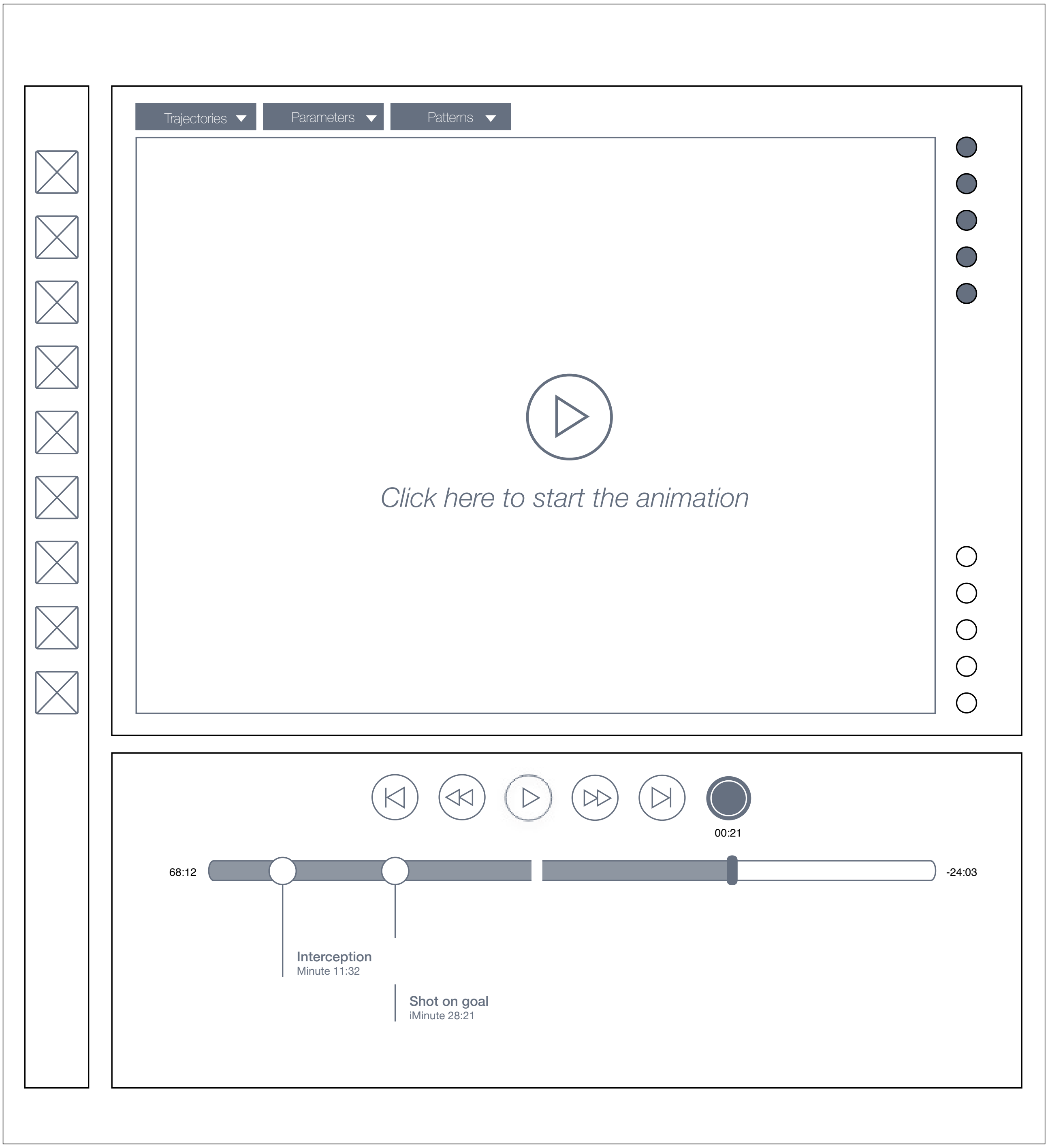
The navigation panel consists of three separate elements: application controls, match selector and setting controls (see table 6.1). With the application controls users can import new datasets or open existing ones, save and share files and print screenshots. In the wireframe of figure 6.4, these controls are visualized by placeholders representing buttons that indicate an action when clicked upon. For this conceptual visual design, the buttons are shown as icons. When clicked upon they request to open a model window. These pop-up windows require users to interact before they can return to the main view. With the match selector users can select available datasets representing matches previously played and imported into the system. The user can interact with the match selector in two ways. They can enter a keyword or phrase in the search field (left side) to search the index, or they can use the dropdown button (right side) to open the dropdown list to select other available matches (see figure 6.4). With the stetting controls users can view and alter their account/profile settings and shut down or restart the interface.

Figure 6.4: More detailed overview of the navigation panel. The three separate elements are visualized to give an impression of how they may look when users interact with them.

6.5. Focus view

The animation, exploration and timeline panels are the so-called focus view from where the visual exploration and analysis of movement data starts (see figure 6.5). To draw the attention towards these content areas they are made to look bigger than the other elements and are located at high attention areas (Cao et al., n.d.). The animation and timeline panels are linked together and do not work without one another (see figure 6.2). These three panels together form the first step in the exploration and analysis process and are called focus view. For more in-depth analyses, users can turn to the geovisual analytics panel where visual analytics tools can be used to obtain information about the concerted movement patterns displayed in the animation panel (see section 6.7).

Figure 6.5: Overview of the focus view. This wireframe provides an overview of their size and placement and shows all the elements included per panel (see figure 6.1 and table 6.1). These content areas are linked to each other with the use of operator-based approaches (see figure 6.2).



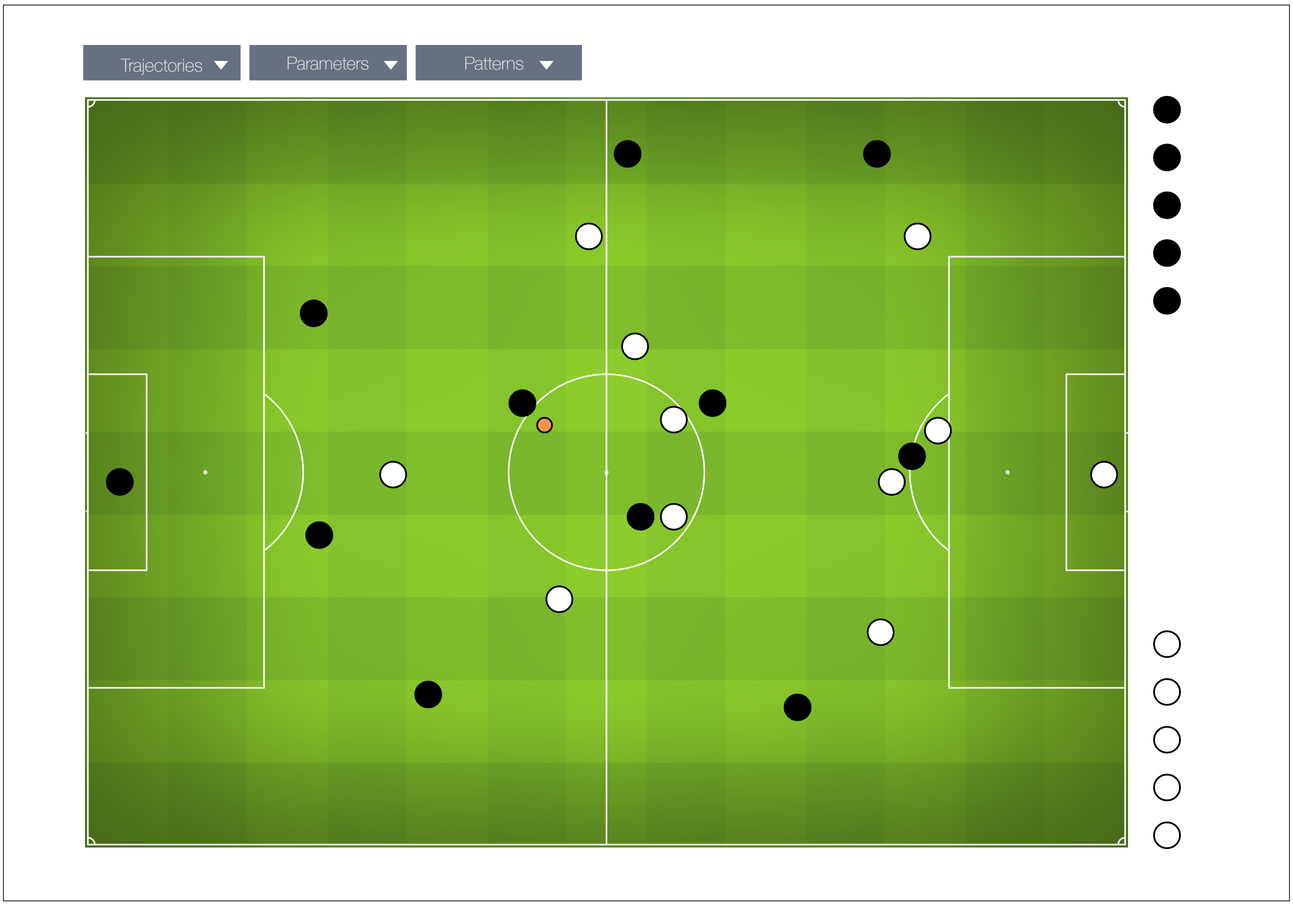
6.5.1. Animation panel

The visual interface should be able to answer the analytical questions coaches, analysts and players have (see chapter 5). Therefore, the challenge is to develop an application that incorporates and answers these spatio-temporal questions in an understandable way. Analytical questions regarding the spatiotemporal component focus on the coordinated movements of players and ball in relation to each other and the context in which they move. This means that the system must be able to visualize these concerted movements in an effective and efficient way.

Video recordings are currently the most used data source of professional football organizations to extract football-specific spatial data from movement observations. Even though this conventional method looks very realistic, it often fails to provide sufficient insights into concerted movement patterns (see section 3.3). Therefore, this study proposes another way of visualizing geographical data with a temporal component, using single maps, small multiples or animated maps (see section 2.4). Maps are an abstraction and simplification of geographic reality. In these models the cartographer compiles an abstraction by concealing real world information using different filters (Longley et al., 2011). Maps are furthermore better capable of visualizing concerted movement patterns and are more likely to be understood by novice users.

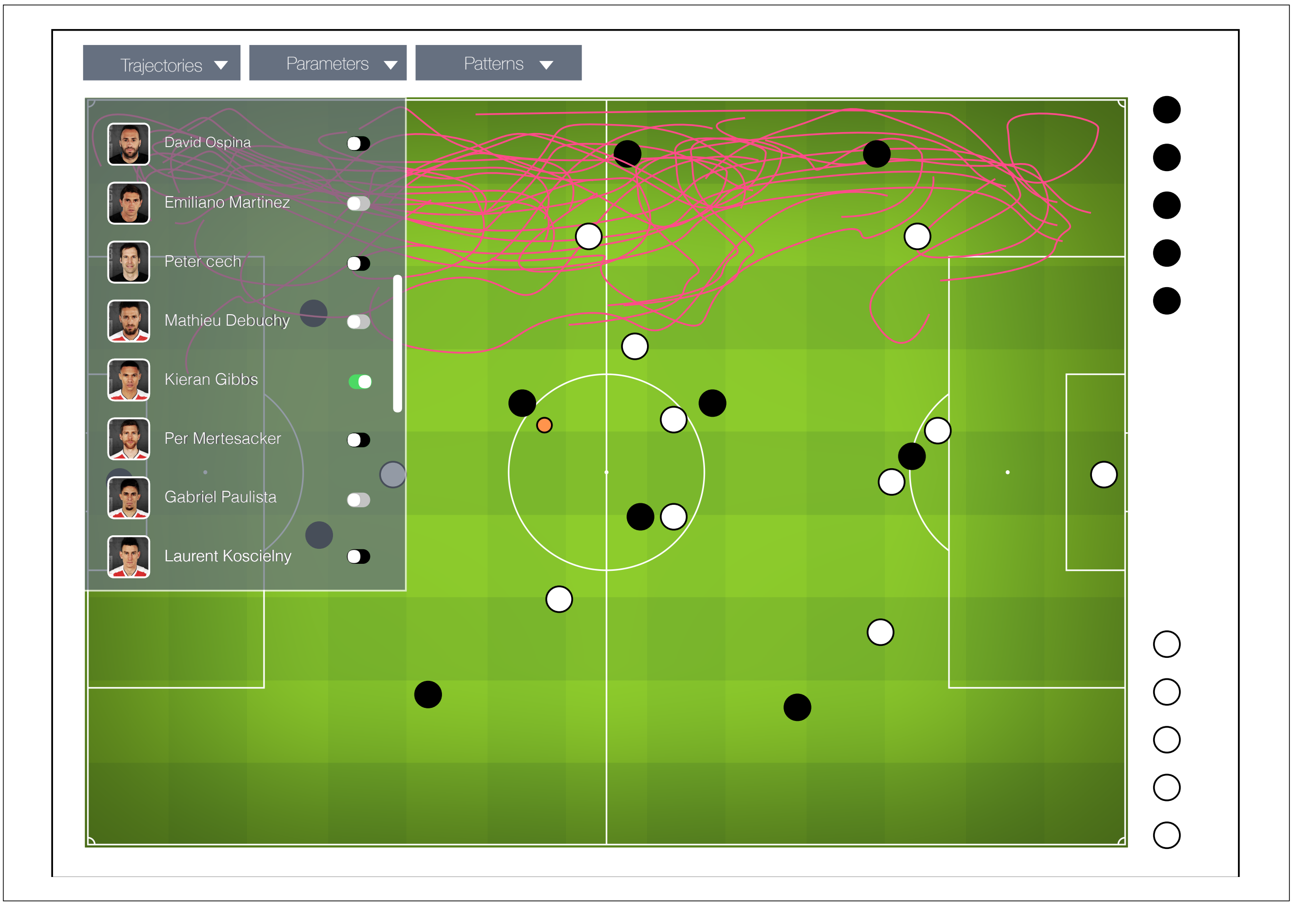
The above paragraph supports the idea of incorporating a two-dimensional animation in the main panel of the visual interface based on the one participants were confronted with during the second task-based session (see figure 5.1). The animation is incorporated in the main panel and displays the formation, player and ball positions on top of a base map that displays the main parts of a football pitch (see figure 6.6). This provides a quick and clear explanation that sums up the essence of the match or training and gives a more objective overview of the situation compared to video recordings. In addition, geometric aspects of movement parameters such as distance and direction can be retrieved, trajectories displayed, patterns revealed and relations understood and quantified. For more in-depth analyses users can connect directly with the three tab items shown at the top of the animation panel just above the actual animation (see figure 6.6). By selecting the dropdown button at the right side of the tab name or clicking directly on the tab name, a dropdown menu slides in from the top.

Figure 6.6: Overview of the animation panel. The base map shows a football pitch with penalty areas, center circle and touchlines using white lines. On top of this base map the positions of all 22 players and the ball are visualized using dots. The substitutes are shown on the right side of the pitch. In this illustration, the orange dot represents the ball. With the use of the dropdown buttons users can open or close the corresponding dropdown menu.



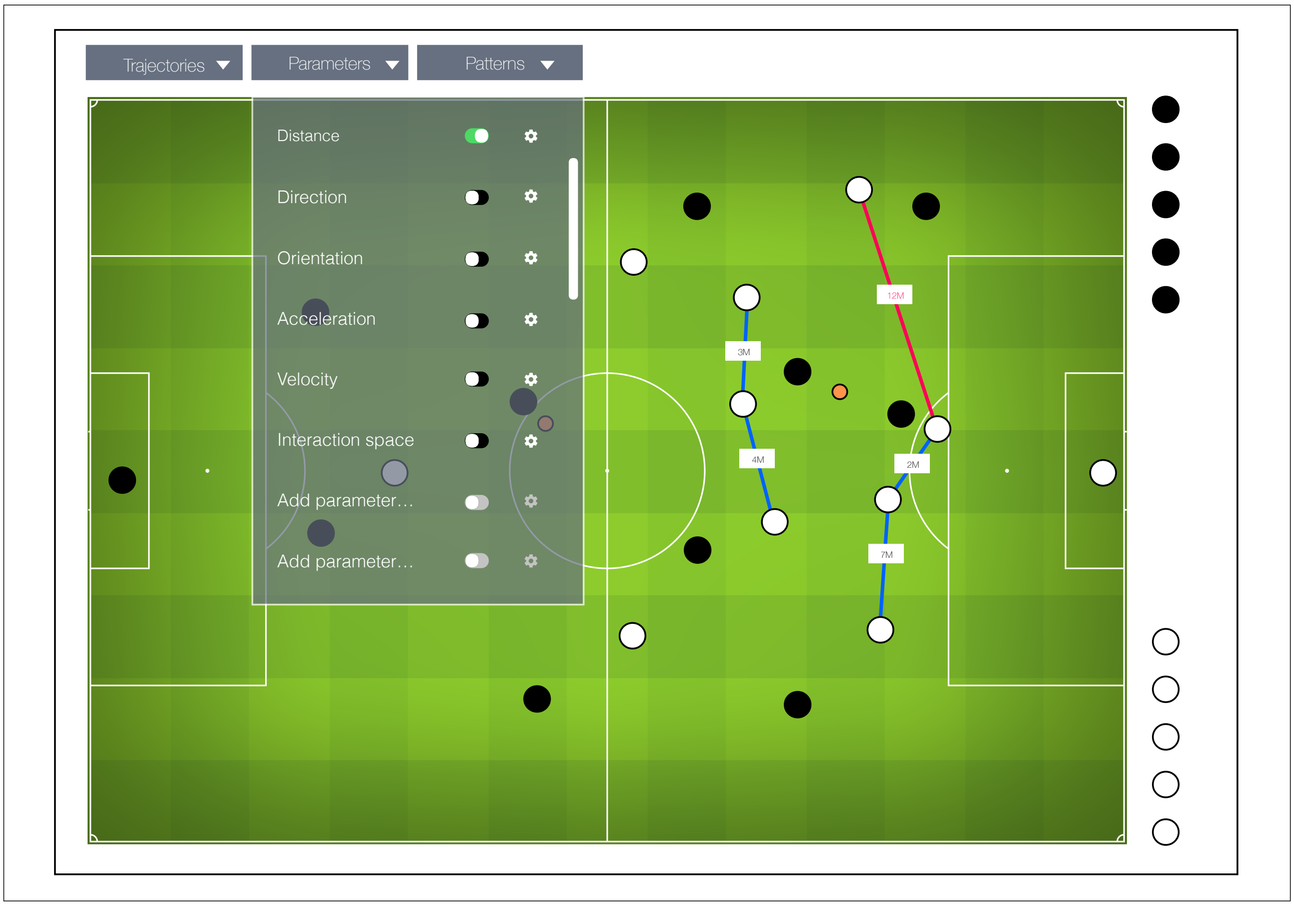
Behind these tabs more detailed information about trajectories, parameters and patterns are concealed. First, using the dropdown button, users can open the trajectories dropdown menu from where they can toggle trajectories on or off per player with the use of switch icons (see figure 6.7). A switch icon allows the user to change a setting between two states. By default, each switch is turned off. In this case, the switch remains black. When the switches are active they are shown at the right side and the icons turns green. Alternatively, when there are currently no data available the switch is shown at the left side and turns gray. Furthermore, the scroll bar at the right side of the menu supports a longer list of options if needed (see figure 6.7).

Figure 6.7: Overview of the trajectories tab showing each individual player alphabetically. When the trajectory switch is turned on this element turns green. When there are no data available this element turns grey. By default, each switch is turned off. In this case, the switch remains black. In this example, the trajectories of Arsenal player Kieran Gibbs are highlighted using a red stroke color.



With the second dropdown button users can open the parameters dropdown menu from where they can switch movement parameters on or off, add new parameters, and change their settings (see figure 6.8). During the requirement analysis, participants indicated that parameters to be included into the system are quite numerous (see section 5.5). Furthermore, the assessment of these parameters is strongly influenced by the playing system. The playing system determines which value users attach to the different parameters. In the visual interface this function is incorporated in the animation panel. For instance, when users want to explore if, when or how often the distance between the center backs exceeds the agreed threshold, they can enter this threshold using the settings icon and turn the distance switch on (see figure 6.8). The animation panel automatically updates its view and shows the constantly changing mutual distances between both players. When the players exceed the agreed threshold the line between them gets a different color, for instance red. In this way, players can see in an instance what they were supposed to do or what they should not have done.

Figure 6.8: Overview of the parameters tab which shows the movement parameters available. In this example, the distance parameter turned on using the switch icon. With the settings icon, just to the right of the switch icon, users can set the agreed minimum and maximum threshold distance. As a result, distance is shown between the players selected by the user using a blue line. When the threshold is exceeded this line automatically turns red. The actual distance in meters is also shown using a white textbox on top of each line.



Participants also stressed that they want to know whether a certain event happened just once or whether it is a recurring pattern worth further investigation. With the third and final dropdown button, users can open the patterns tab. With this tab, they can distill typical patterns and compare observed patterns with ideal patterns (see figure 6.9). This tab is directly linked to the animation panel. When users click on one of the collected or recorded patterns, the animation panel automatically updates its view and shows a replay of the selected pattern.

Figure 6.9: Overview of the patterns tab showing observed (left side) and ideal patterns (right side). Ideal patterns can be manually added by the users with use of the settings icon within the patterns tab. Observed patterns can are automatically discovered using machine learning algorithms that provide computers with the ability to learn without being programmed and make data-driven predictions.

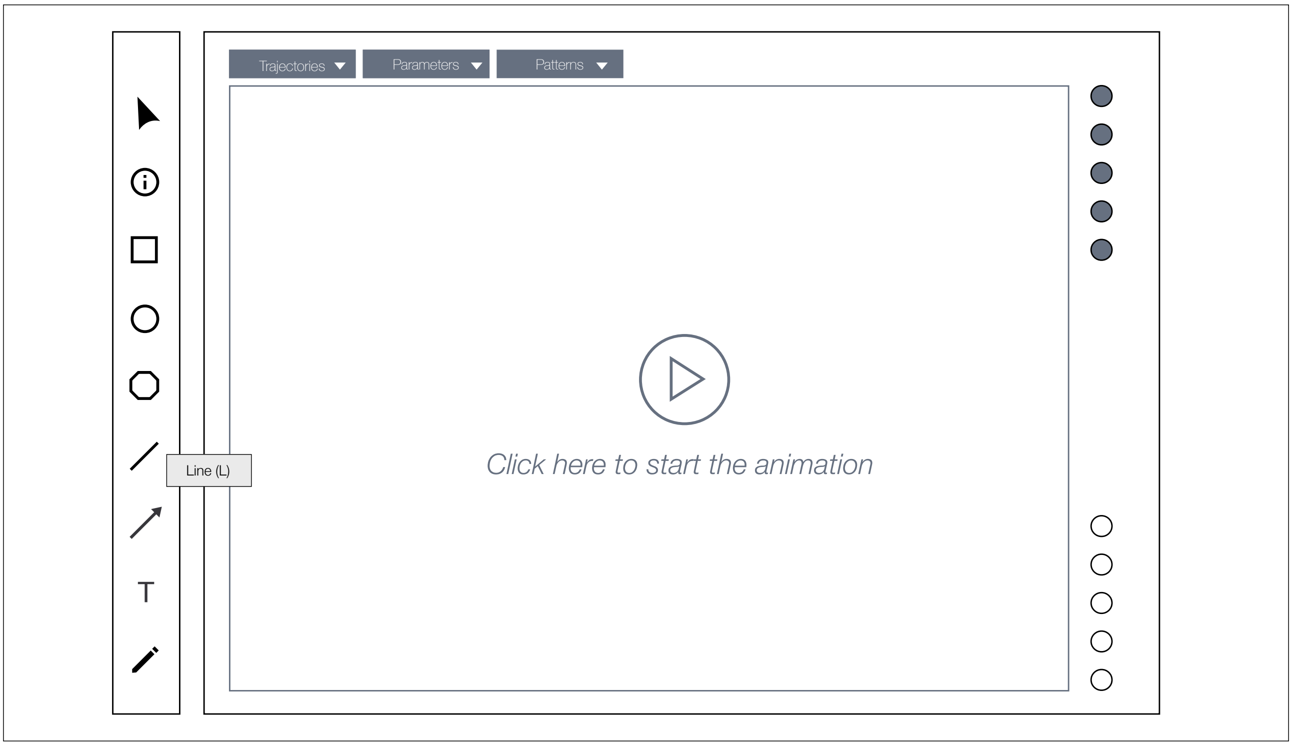
With the help of clustering techniques, movement data can be turned into summaries of similar movement patterns and typical patterns can be sorted out (see sub-section 2.3.3). When properly transformed, these observed patterns can be compared with ideal patterns in order to examine, for example, what the players' ideal position would have been. Observed patterns include any recognizable spatial and temporal trend or relationship in a set of movement data. Ideal patterns, in contrast, depend on the playing system. When interacting with their teammates, players must adopt a set of rules corresponding to the tactics following the playing system. They must know, for instance, when to accelerate, when to move into space, as well as how to use space and respond to their teammates (Sumpter, 2016).

To visually explore and analyze these patterns, it is recommended to link and translate the playing system into a spatiotemporal model. This model incorporates the vision of the coach as an interpretation mechanism and can be validated with the use of the movement data collected. Nevertheless, it remains very difficult to distill ideal patterns and compare them with observed patterns because the actual playing style is mostly unknown. Furthermore, pattern recognition based on movement data is still in an early experimental phase. This makes it hard to determine where a player goes next given his current location, speed, direction and acceleration. More research is required to establish the viability of a spatiotemporal model using movement data (see section 7.4).

6.5.2. Exploration panel

To control the information displayed in the animation panel a sidebar is added. Sidebars can be used in many ways. In this application, the sidebar contains button placeholders in which different (visual) tools are placed that emphasize the kind of tasks the user may wish to execute with the user interface (see figure 6.10). During the task-based sessions in the requirement analysis, participants indicated that the application should be able to identify and select players, filter events and draw points, lines, arrows and other basic shapes directly on screen (see section 5.5). These simple but effective features are incorporated into the exploration panel to provide both novice and expert users with a basic toolbox which can assist them during pre- and post-match briefings. The user can select, for instance the rectangle tool or the arrow tool drawing free spaces, passing options or other tactical information directly on-screen during match briefings. Besides direct interaction with the interface, by clicking on the icon, the system also provides indirect manipulation using keyboard shortcuts.

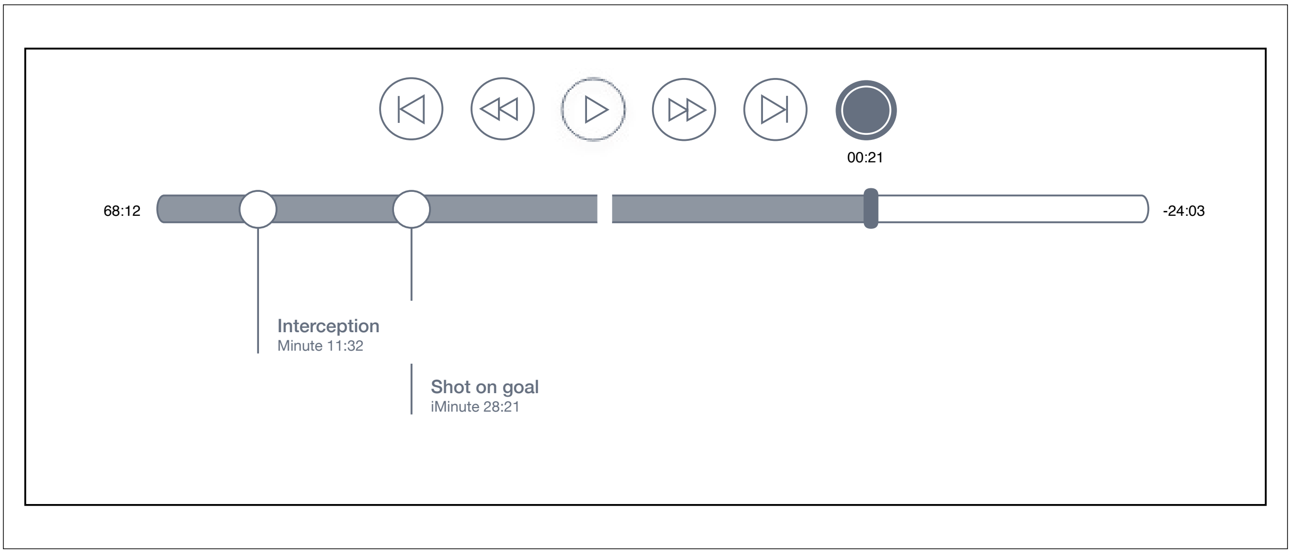
Figure 6.10: Overview of the exploration panel (left side). The tools placed in this sidebar can be used to visually explore the movement data shown in the animation panel (right side). Tools are visualized using icons. When the mouse pointer moves over a tool a message box appears showing the name of the tool and its keyboard shortcut, see for instance the line tool.



6.5.3. Timeline panel

With the incorporation of the timeline panel, the visual interface provides an intuitive overview of both space and time within which movements occur (see figure 6.11). The timeline panel consists of three basic elements: playback controls, timeline and annotation fields (see table 6.1). When users interact with the playing controls, timeline or annotations fields embedded in the timeline panel, this automatically evokes a change in the animation panel (see figure 6.2). The playback controls (i.e. slow-motion, previous, play/pause, next, fast forward and record) are used to control the two-dimensional animation. For instance, with the slow-motion button users can control the speed with which the animation is displayed. With the timeline users can control the animation by dragging the slider thumb towards a specific moment. This thumb is visualized by a gray bar. In addition, half time is shown using a whitespace/break in the middle of the timeline (see figure 6.11). Below the timeline users can manually add annotations with a maximum of thirty characters in the designated text fields. These annotations refer to a specific movement or event. When users add new annotations by entering text in the designated text fields, a node automatically appears on the timeline (see figure 6.11). When clicking such a node tagged to the timeline, users request the system to respond by changing the view (i.e. viewpoint manipulation) in the animation panel.

Figure 6.11: Overview of the timeline panel. Most elements shown in this illustration are interactive and linked to each other and to the animation panel described in the previous paragraphs. Together, they are used to incorporate both space (animation panel) and time (timeline panel) in the domain of football analytics.



6.6. Video and statistics panels

From chapter 3 it became clear that most analysts manually process and annotate video recordings to evaluate both player and team performances. Besides video recording, most professional football organizations also collect descriptive data about the physical workload of players with the use of wearable tracking devices. This study assumes that video and statistics alone do not grasp the full potential of football analytics. The video and statistics panels are not considered to be the main focal point of the visual interface and, therefore, have been placed at the bottom right corner of the application (see figure 6.1). Both are complementary content areas and are connected via brushing and linking techniques to the animation panel. When users, for instance, select a player in the animation panel by clicking or dragging, both video and statistics panels automatically update their views and show all video recordings and statistics available for that particular player or team.

Because both panels are relatively small, their presence is highlighted using a title bar at the top of each panel (see figure 6.12). The video panel contains a video placeholder that is situated in the center of a video carousel. Using the control buttons on either side, users can browse through a series of video recordings that are available for the player or moment selected (see figure 6.12). Within the video panel, the selected video is being played in small screen mode by default. However, when users double click the video image, it expands and becomes a drag and drop window that can be positioned by clicking on it and dragged anywhere within the application viewport. With the dropdown buttons in the statistics panel users can open the team and player dropdown menus from where they can obtain more detailed statistics (see figure 6.13).

Figure 6.12: Overview of the video and statistics panels. Both panels are linked to the animation panel as shown in figure 6.2.

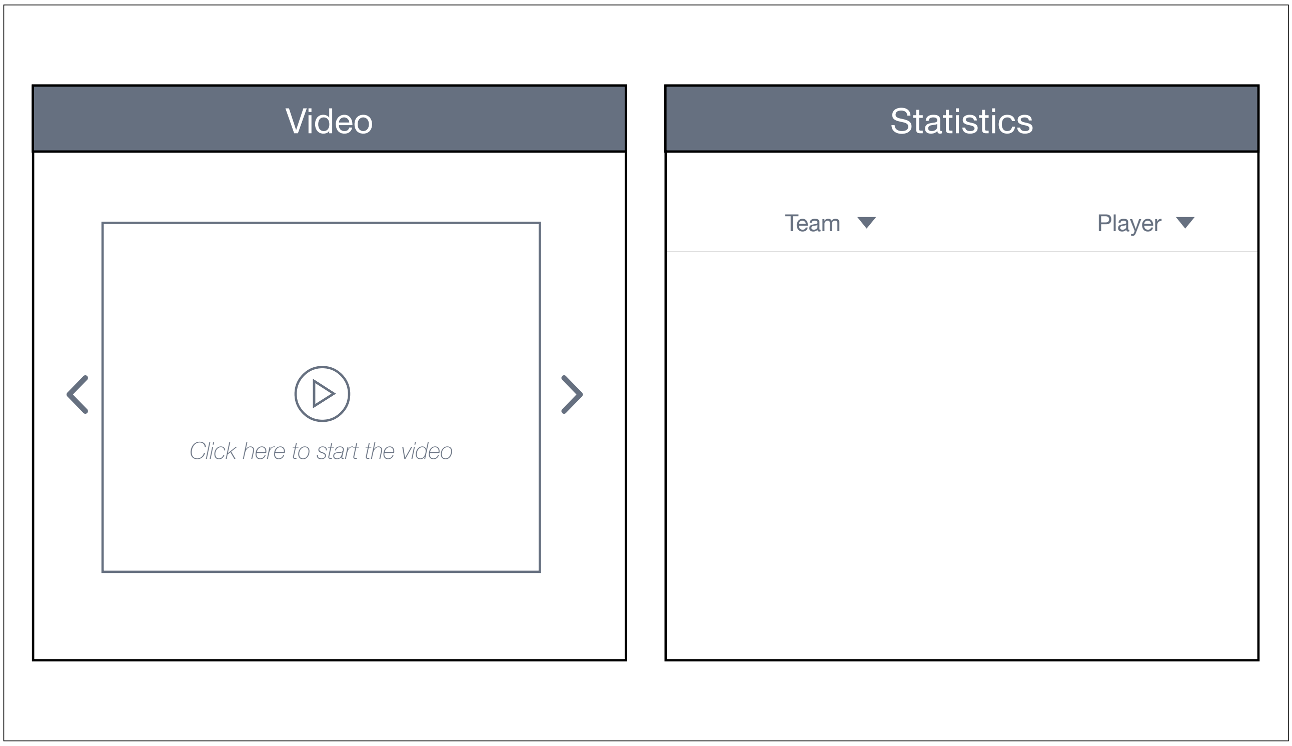


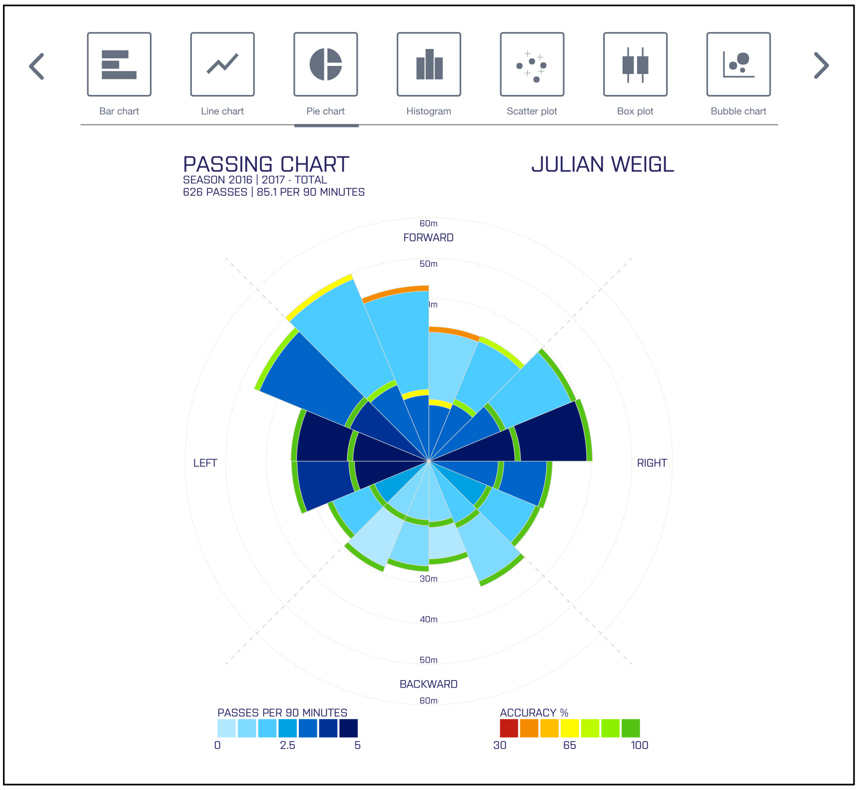
Figure 6.13: Example of more detailed player statistics. Source: URL 42

RANK	PLAYER	PLAYED		SHOT POSITION			
		Games	Mins	Inside Area	Outside Area	Total Shots	Shot Accuracy %
1	Alexis Sánchez Forward - Arsenal	33	2804	67	42	109	57
2	Theo Walcott Forward - Arsenal	26	1903	51	10	61	58
3	Mesut Özil Midfielder - Arsenal	28	2401	27	9	36	56
4	Alex Iwobi Midfielder - Arsenal	24	1414	27	8	35	45
5	Olivier Giroud Forward - Arsenal	25	1021	29	4	33	54
6	Aaron Ramsey Midfielder - Arsenal	18	884	20	12	32	32
7	Granit Xhaka Midfielder - Arsenal	27	2077	1	26	27	26
8	Oxlade-Chamberlain Midfielder - Arsenal	27	1455	18	8	26	41
9	Shkodran Mustafi Defender - Arsenal	23	2004	16	2	18	23
10	Héctor Bellerín Defender - Arsenal	28	2183	12	5	17	31

6.7. Geovisual analytics panel

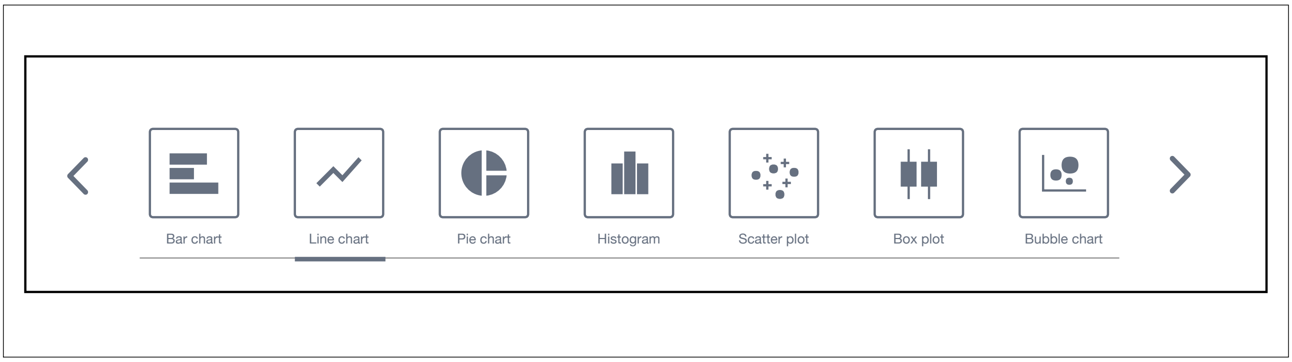
The first task-based session of the user requirement analysis showed that participants had difficulties executing the task given while using the SoccerStories environment (see chapter 5). While thinking out loud, participants indicated that they were not impressed with the interaction functionalities of the visual analytics tools provided by the interface. In addition, the screen recordings showed that participants hardly used any of these tools during the visual exploration. Furthermore, their interaction capabilities were not explained by the interface and could not easily be understood without any training or guidance. These results learned that the visual analytics tools provided by SoccerStories are too complex and not intuitive for most domain experts to be considered as useful. This visual interface therefore only includes visual tools that are accessible and easy to use without any training or guidance (see figure 6.14).

Figure 6.14: Overview of the geovisual analytics panel. This illustration provides a first glance on the visual analytics tools that are shown by default. When users select, for instance, the Pie chart tool, a Passing chart may automatically be constructed. Source: URL 1



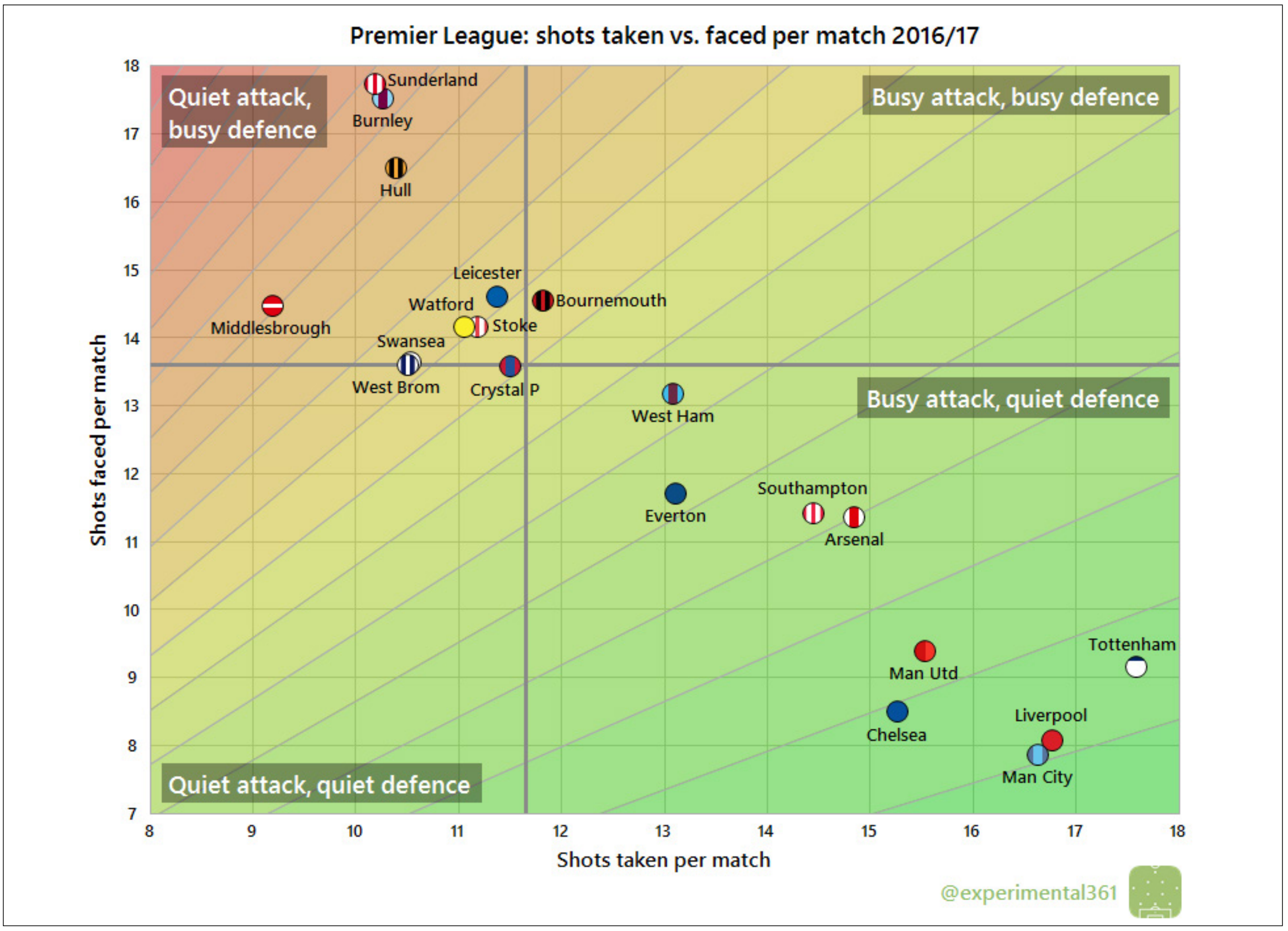
The geovisual analytics panel consists of two connected elements: visual analytics tools and the visual analytics placeholder (see figure 6.1 and table 6.1). In the wireframe of figure 6.1, the visual analytics tools are visualized by placeholders representing buttons that indicate an action when clicked upon. For this conceptual visual design, these buttons are shown as icons (see figure 6.15). From the carousel, users can select the visual analytics tool they prefer by clicking on an icon. The carousel can be controlled using the arrow buttons on either side allowing users to browse through the available set of visual tools. The visual tools that are shown by default are relatively straightforward and accessible for both novice and expert users. Novice users are expected not to use more advanced capabilities because they are unfamiliar with complex visual tools and are laymen in the field of geovisual analytics. The tools displayed are the following: Bar chart, Line chart, Pie chart, Histogram, Scatter plot, Box (and whisker) plot and Bubble chart (see figure 6. 15).

Figure 6.15: Overview of the visual analytics tools placed in the carousel.



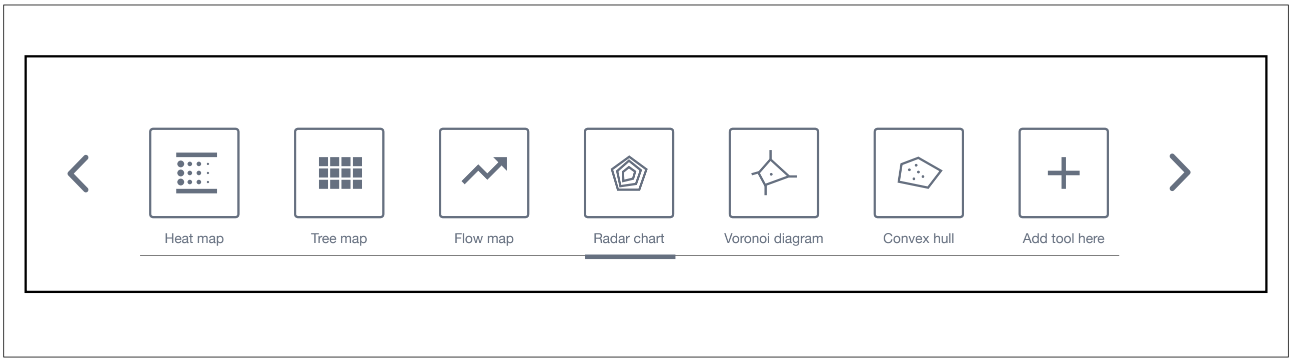
Bar charts can be plotted horizontally or stacked. They present categorical data using rectangular bars with lengths proportional to the values that they represent. A vertical Bar chart or Line chart presents a series of data points connected by straight line segments. Line charts are often used to visualize a trend, for instance, heart rate data collected with the use of a Polar chest strap (URL 20). Pie charts are circular and divided into slices that display numerical proportions. This chart can for instance be used to illustrate the style of passing by dividing the passes in two categories: short and long distance (see figure 6.14). Histograms are a good option if users want to see the frequencies/distribution of scores for a single variable. Scatter plots are graphs that plot a collection of points to look at the relationships between two variables (see figure 6.16). Box plots or box-whisker diagrams divide data into quartiles. Whiskers are used to indicate upper and lower quartiles or the most and least extreme scores respectively. Finally, the Bubble chart tool can be used to display three dimensions of data.

Figure 6.16: Scatter plot of the number of shots taken by each Premier League club compared with those face in return. The average number of shots taken per match is on the horizontal and the average number faced is on the vertical. The stripes are like contours: the greener the stripe, the better the performance (and vice versa for red). Source: URL 43



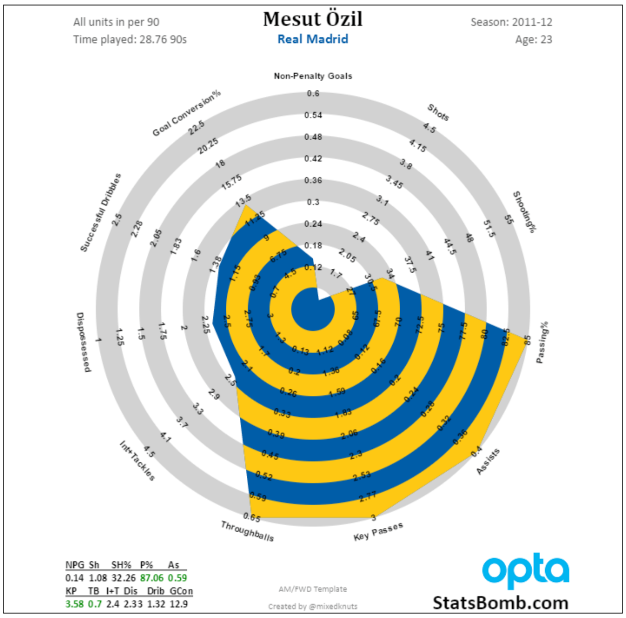
Expert users differ in their abilities, capabilities, requirements and conditions under which they will use the application (see chapter 5). For them, more complex analytics tools are available. The tools included in this conceptual visual design are: Heat map, Tree map, Flow map, Radar chart, Voronoi diagram/tessellation and the Convex hull (see figure 6.17).

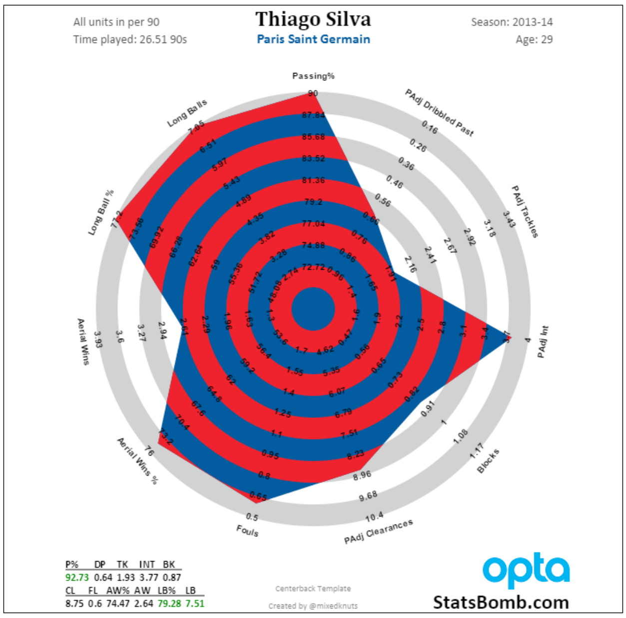
Figure 6.17: Overview of the more complex geovisual analytics tools placed in the carousel.



Heat maps are used to visualize aggregated information and indicate spatial patterns mostly using a red to green gradient. Tree maps display hierarchical data with the use of nested rectangles. Flow maps make use of line symbols that show the movement of objects between locations. Arrowheads are used to indicate their direction across a certain environment. The thickness of the arrow is used to express the amount or volume of movement. Radar charts display multivariate data in the form of a two-dimensional chart. These can for instance be used to compare player statistics in one visualization (see figure 6.18). Other tools included in the application are Voronoi diagrams and Convex hulls. Voronoi diagrams divide the pitch into regions based on the distance between all players. The larger the area, the greater space they have around them. A convex hull visually groups nodes and takes the outermost nodes of a vector to create a polygon of the minimum sized area. In the case of football, the larger the convex hull, the greater the area on the pitch covered by the team's players.

Figure 6.18: Two examples of Radar charts. They help to evaluate players based on statistics about their shooting percentage, key passes and successful dribbles, among others. Source: Knutson, 2016





6.8. Conclusion

In this chapter, the development process that lead to the visual interface and the interaction functionalities it supports has been discussed. The requirement analysis made clear that most target users are laymen in the domain of geovisual analytics and have little experience working with actual movement data. Therefore, the content and the elements included in the interface should immediately make clear their purpose and functionality.

The interface consists of seven separate panels that are linked together using the operator-based approaches brushing and linking. After movement data has been imported into the system using the navigation panel, the visual exploration and analysis starts at the animation panel. Together with the exploration and timeline panels, these panels are considered as the primary focus point. The animation panel provides a quick and clear explanation which sums up the essence of the match or training, much like tactical maps do. They display formation(s), players and ball positions on top of a base map. Furthermore, geometric aspects of movement parameters can be retrieved, trajectories displayed, patterns revealed and relations understood and quantified using the three tab items. After the exploration phase, users can turn to the geovisual analytics panel in which (geo)visual analytics tools can be used to obtain more detailed information. This panel only includes visual tools that are accessible and easy to use without any training or guidance because novice users are not expected to use more advanced capabilities.

The next chapter concludes by answering the research questions, gives a reflection on the methods used and provides recommendations for further research.

CH.7

Conclusion and discussion

7.1. Summary

Interactive visual interfaces are potential tools dealing with large amounts of movement data by providing users the opportunity to look at the subject from different perspectives. Also, they allow them to modify the data using different visual analytics tools. The involvement of users and a successful identification of their needs however is often underestimated if not overlooked altogether. This research project tries to alleviate this problem, investigating user requirements prior to the development process of a conceptual visual interface. To realize this objective, the User Centred Design (UCD) methodology is followed.

In chapter 2 an overview of literature is provided aiming to outline the key concepts that form the focus of this study. The resulting framework shows that the vast amounts of movement data pose a challenge on how to transform these into movement patterns. Visual tools and interfaces are potentially very effective to explore, analyze and communicate these patterns. These tools however can only support information exploration and knowledge construction when users are able to employ them properly. This chapter has shown that there is still a sincere need for more use, user and usability research in the domain of geo-information sciences to develop effective and efficient geovisual analytics tools that can deal with large volumes of movement data. The lack of such knowledge is the main driver of this research.

Chapter 3 provides an overview of literature focusing on the visual analysis of the concerted movements of football players based on movement data collected by different tracking systems. The challenge this chapter faced was to turn these vast amounts of data into an intuitive visualization that is accessible for coaches, analysts and players. The comparison between different visual tools and interfaces shows that the usability of visual tools remains often unsatisfying. This research project, in the context of geo-information science and cartography in general and geovisual analytics in particular, is to be considered as one of the first attempts to incorporate user requirements and transforming them into a usable conceptual design of a geovisual football analytics interface.

Chapter 4 contains a description of the UCD-based methodology used, outlining the elements and consecutive phases leading to the desired results. The first phase deals with the use and user requirements, while the second phase focuses on the conceptual design of a more usable interface to address the requirements obtained during the first phase. Being of a qualitative nature, structured user requirement interviews are combined with the think aloud method. All the resulting research materials are analyzed and turned into verbal and action protocols.

Chapter 5 describes the implementation of the research methodology involved and the outcomes thereof. The results of the requirement analysis interviews and task-based sessions obtained during this phase were used as input for the conceptual design and are to be answered by the application in an accessible and understandable way. Besides the analytical questions and user requirements identified, chapter five also provides information about the basic characteristics, tasks, responsibilities of and conditions under which the intended user groups will probably use the interface. Finally, the reliability and validity of the research design is discussed.

Chapter 6 embodies the actual conceptual design of the interactive visual interface. In this chapter, a description is provided of the development process leading to the conceptual visual interface and the interaction functionalities it supports. With the use of wireframes, low-fidelity visual representations of the interface have been made. These illustrate the main groups of contents, the structure of information, the functionality and the relationship between the different elements. The conceptual visual interface consists of seven individual panels which are linked using brushing and linking techniques. The animation panel and geovisual analytics panel are considered as main content areas. These provide information on the key concepts outlined in chapter 2, and provide answers to the analytical questions and user requirements that are discussed in chapter 5.

7.2. Answering the research questions

The answers to each of the research questions set in section 1.4 will be summarized as follows in this section:

RQ 1: What are the relevant key concepts related to concerted movements which underpin the research project?

The key concepts that form the focus of this research project are: movement trajectories, movement parameters and movement patterns. Trajectories are paths through space and time under the action of given forces. Parameters are the derivatives of movement trajectories. Together, trajectories and parameters are the underlying elements of movement patterns. Patterns describe frequent behaviors of moving objects, in terms of both space and time.

RQ 2: Which connections and notable relationships exist between the relevant key concepts?

The connections and relationships that exist between key concepts are all related to the concept of movement. Movement is defined as the change of location of one or more objects occurring in space and time. Spatiotemporal data refer to these location-based changes. According to the TRIAD framework, objects, space and time are the generic components in spatiotemporal data specific to movement which is related to the internal state of an object, its properties and the context in which it moves. This relationship results in trajectories and parameters that can be observed and stored into an information system. These movement observations can be transformed, using, for example, data mining techniques, into movement patterns connected to the field of geovisual analytics which focusses on visual interfaces to enable insight from spatiotemporal data.

RQ 3: How can an interactive visual interface support knowledge construction and insight into concerted movements?

An interface can yield knowledge and insight by establishing an interactive relationship between humans and machines that allows users to look at a subject from different perspectives. Also, diverse pieces of information can be linked and controlled by geovisual analytics tools. From an academic point of view, it looks like this is promising. There is however still a shortcoming concerning the development of geovisual analytics tools, because the research on usability is still limited. Geovisual tools and environments can only support information exploration and knowledge construction when users are able to employ them properly. There is a sincere need for more user research.

RQ 4: Which geovisual analytics tools are currently available to visually explore and analyze concerted movements and how are they used?

Maps and map-related graphics or diagrams are called geovisual analytics tools. These tools cover a wide range. It varies from a single map, small multiples or animated maps to complex and interactive tools such as parallel coordinate plots, treemaps, and node-link diagrams. Most targeted users indicate that they are accustomed to use simple graphs and diagrams to communicate the results found with the coaching staff, medical staff and players. Complex geovisual analytics tools are assumed to be difficult to understand by the intended end-users.

From a cartographic point of view, the most common type of display used to visualize spatial data is the cartographic map. As football is a spatial phenomenon it is most likely to visualize the positions and trajectories of the players and the ball on a map representing a football pitch. Trajectories may be visualized using a flow map. This tool makes use of line symbols and arrowheads to indicate the direction. Width is used to express the volume of movement. Another tool capable of visualizing the trajectories of moving objects between stops is the space-time cube. This three-dimensional cube is used to study the movements of objects through both space and time.

RQ 5: What are the characteristics of end-users, tasks and context of use in which they operate?

There are clear differences between novices and experts in the visual exploration and analysis of movement data. Most coaches, analysts and players are considered to be novice users. Despite being characterized by their spatial intelligence, they lack experience working with movement data and (complex) visualizations. In contrast, expert users do have these experiences and execute more higher-level tasks than novices do. Experts and novices both operate in an office setting, and not, for instance on or near the pitch because the graphics displayed are too detailed and simply scaling down is inefficient.

RQ 6: What are the main requirements and analytical questions that end-users have when they visually explore and analyze concerted movements?

There are at least four essential requirements end-users expect of the interface. It should display both on and off the ball (movement) parameters. It should provide answers to both spatial and temporal questions in a simple and intuitive way. It should integrate multiple coordinated views into one visual analytics environment. And it should allow users to interact with the data to be able to explore and analyze patterns and include the possibility to incorporate the playing system. Most analytical questions focus on the coordinated movements of players and the context in which they move, for example: 'how players collectively move and orientate themselves in relation to the ball, the opposition and each other?'.

RQ 7: Which geovisual analytics tools are currently missing to meet the proposed requirements of end-users?

According to the results of the requirement analysis, video recordings are the most used data source by professional football organizations. Watching the game in replay is a method that comes near to reality. However, these video presentations do not meet the proposed requirements. Therefore, the conceptual design attempts to translate movement observations into movement patterns visualized by a two-dimensional animation. With the use of this animation users can distill typical patterns and compare these observed patterns. In this way, they can find out where the players should have been on the pitch ideally. This abstract presentation of reality is currently missing to visually explore and analyze movement patterns and could improve the quality of football game analyses by making it more transparent and accessible.

RQ 8: How can the conceptual design contribute to fill the gaps identified by end-users?

The gaps identified by end-users can be filled by providing the required information about the key concepts outlined in chapter 2, and providing answers to the analytical questions and user requirements discussed in chapter 5. The user requirements collected are deployed as input for the conceptual design and should be met by the application in an accessible and understandable way. By using computational techniques, movement observations can be transformed into movement patterns. By providing insight into movement patterns using an interactive visual interface incorporating both space and time, the dynamic nature of football is made understandable. Geovisualization tools, possibly embedded in a user interface environment, are potentially very powerful to explore, analyze and communicate these patterns to extract relevant information from this spatio-temporal data.

The main conclusion of this research project is that technology-driven solutions offer professional football organizations instruments for developing visual tools and interfaces that create a new flow of information. Information that helps to get a more complete picture of both individual players and teams. However, many clubs find it difficult to incorporate these resources into the daily working routine. The question is how to transform the enormous amount of data into useful information and ultimately into knowledge. Therefore, what must be developed is the incorporation of user requirements and user centered design. Otherwise, the usability of these geovisual tools and interfaces remains unsatisfying and will be of little use for coaches, analysts and players dealing with large amounts of movement data. This research tried to solve this problem by proposing ways to make movement data more accessible for clubs by developing a conceptual visual interface using a User-Centered Design approach.

7.3. Discussion

For this research, it was decided to combine two qualitative research techniques that should lead to a greater validity of the obtained results. The chosen methods included structured interviews and the thinking aloud / observation during task execution. To ensure valid results, prevent bias and ensure a reproducible and methodologically sound output, a structured and standardized procedure has been implemented throughout the whole project. In addition, participants were made to feel comfortable by conducting the one-hour sessions at their current working location. The same conditions were created during the task-based sessions by using a uniform working environment. The results obtained from these sessions are however not transferable to other settings because the sample used is not considered to be representative for the whole population of users. This does not mean that the results are worthless. On the contrary, they provide a source of valuable information regarding the development of a conceptual visual interface that helps coaches, analysts and players to visually explore and analyze movement data.

The answers to the interview questions and the information derived from the thinking aloud were brought together into protocols. These showed that most representatives of the selected user profiles lack experience working with movement data and visualizations and are therefore considered as novice users. This result is not in line with the expectation that most professional football organizations are familiar with football analytics. Literature has shown that the capacity to collect vast amounts of movement data has not been matched by the ability to process them in meaningful ways. Video recordings moreover still seem to be the main source of data collection because these data are easily accessible to and interpretable for domain experts.

In general, there were no major problems with respect to the UCD method employed. There are however a few issues that should be considered in future similar studies. Foremost, the reader should bear in mind that the results of this study are based on data collected from only nine participants. These participants, working at professional football organizations, do not represent the targeted user profiles as intended, whereas the group of analysts is overrepresented. This is an important and useful issue for future research, and for other researchers to consider when developing their own research plan. Given the limitations of this research, this study was furthermore unable to perform continuous iteration cycles of design and development with users. Further research is therefore required to develop and implement a high-fidelity prototype and evaluate its quality of use. In this case, the outcomes of this study (i.e. the recordings and protocols) may be used as basis for the development of an actual working tool. Finally, this study provided some examples of currently available geovisual analytics tools. Combining traditional data with location data and bringing geographical context into the analyses is becoming more popular with the use of software such as Tableau (URL 35) and SAS (URL 37). This interactive data visualization software helps users to examine and understand their data by developing interactive reports and dashboards with advanced visual tools without programming. These platforms are also becoming more popular in the field of visual analytics and the domain of geovisualization. Further research is required to determine whether these platforms could be used for the development of geovisual analytics tools.

7.4. Recommendations

This research project should be considered as a first step into a new application area for geo-information scientists and cartographers. The issues that were encountered therefore lead to the following recommendations for future research:

- This research has shown that there is a sincere need for more user research in the fields of geo-information science with a focus on geovisual football analytics. What must further be developed is the incorporation of user requirements and design for the needs of specific users. Otherwise, the usability of these interfaces remains unsatisfying and will be of little help for coaches, analysts and players dealing with large amounts of movement data. Therefore, it is recommended that the conceptual design is turned into a high-fidelity prototype that is implemented and subjected to various iterations of usability research.

- Conventional models are currently used to translate movement observations into an impression of the game close to reality. However, to visually analyze the complex concerted movement patterns sufficiently it is recommended to link and translate the playing system into spatiotemporal models. These spatiotemporal models incorporate the vision of the coach as an interpretation mechanism and can be validated using the movement data collected. From the perspective of geo-information science this could be considered as a contribution to the domain of football analytics. More research is required to establish the viability of a spatiotemporal model using the movement data.

- Geovisual analytics is conceptually related to various fields of research such as data mining, geovisualization, visual analytics, statistics and machine learning. Therefore, it is recommended that future football analytics research embraces a more multidisciplinary approach to make sense of the increasing volumes and complexities of datasets and support the visual exploration and analysis of spatiotemporal data. In addition, as more movement data and novel methods to analyze the data are becoming available it is suggested that domain experts should work closer together with scientists to ban skepticism and comfortably rely upon knowledge derived from both domains.

Appendices

1 Appendix 1: Formal interview request

Beste (..),

Mijn naam is Lars van Hoeve en momenteel ben ik bezig met mijn afstudeeronderzoek van de opleiding Geographical Information Management & Applications (msc-gima.nl). Ik doe dit onderzoek in samenwerking met SciSports. Het doel van mijn onderzoek is het ontwikkelen van een visuele interface voor trainers, analisten en spelers die data uit wedstrijden en trainingen toegankelijker moet maken. Het toegankelijker presenteren kan jullie helpen om meer informatie te halen uit de grote hoeveelheden data die verzameld worden. Op deze manier kunnen mogelijke problemen eerder worden opgespoord en aangepakt waardoor een beter resultaat kan worden behaald. In de bijlage van deze e-mail een korte toelichting over mijn onderzoek.

Ik zou graag op korte termijn met je in contact komen voor een interview (duur: 60 minuten) op een tijdstip nader te bepalen in de periode van 23 januari tot 17 februari. Kan je in een reply op deze e-mail laten weten of je daarvoor open staat? Dan neem ik vervolgens spoedig contact met je op.

Het spreekt vanzelf dat de resultaten van het onderzoek na afloop met je worden gedeeld.

Alvast bedankt!

Met vriendelijke groet,

Lars van Hoeve

Geovisual football analytics

Lars van Hoeve

Until recently, analysts used video recordings that were manually processed, annotated and edited for analysis and presentation purposes. However, due to recent advancements of tracking systems, football clubs now have access to increasing amounts of movement data. Observations collected by these systems contain the positions of players (and the ball) and may be enriched by adding event data, such as the number of shots, fouls and passes. The complexity and volume (of data) pose a challenge on how to transform this data into information and ultimately into knowledge.

Data science is an emerging area of work concerned with the task of extracting insight from large collections of data. Data science methods and approaches address all stages of the transition from data to knowledge and action. Interactive visual interfaces support the human cognitive process by allowing analysts to look at a subject from different perspectives and at different scales and levels of detail, link diverse pieces of information, as well as direct and control the work of computational analytical tools. Visualization and visual analytics approaches play an essential role in data science. Therefore, the challenge facing this research is to turn vast amounts of movement data into an informative picture that can be understood by football coaches and analysts. These visual messages should provide a quick but clear explanation which sums up the essence of the match or training, much like tactical maps do.

Furthermore, this research acknowledges that the perspective of users regarding the development of these visual tools is often underestimated if not overlooked altogether. Often, this is because designers do not analyze who the users are, which tasks they want to perform using the visualizations, and what their working environments are. This research alleviates this problem by applying a user-centered design (UCD) method as a valuable method to explore the analytical functionality needed by experts and the types of visual tools that can be understood by them.

3 Appendix 3: Introduction and instruction to the requirement analysis sessions

Vooral als gevolg van technologische ontwikkelingen is het gebruik van data een snelgroeiend en belangrijk onderdeel van de sport geworden. Ook in het voetbal wordt sinds enkele jaren de revolutie der data-analyse verkondigd. Deze revolutie, overgewaaid uit de Verenigde Staten, zou nieuwe manieren moeten opleveren om prestaties van teams en spelers te waarderen, evalueren en uiteindelijk te verbeteren. Dit onderzoek richt zich op de gecoördineerde/collectieve bewegingen van spelers in relatie tot elkaar, de bal, de tegenstander en de omgeving waarin ze bewegen. Dit type data noemen we bewegingsdata ook wel bekend als positiedata.

Door technologische vooruitgang in zowel sensor- als videosystemen worden bewegingsdata verzameld in grote en constant groeiende hoeveelheden. Maar terwijl technieken om bewegingsdata te verzamelen steeds geavanceerder worden, blijkt het voor veel clubs moeilijk om optimaal bruikbare informatie en kennis te genereren uit de data. De hoeveelheid, diversiteit en complexiteit van de beschikbare bewegingsdata is een groot probleem. Dit onderzoek erkent dit probleem en probeert bij te dragen aan een oplossing door het ontwerpen van een interactieve visuele interface voor trainers, analisten, scouts en spelers die bewegingsdata toegankelijker zou moeten maken. Het toegankelijker presenteren door middel van visualisaties moet clubs helpen om meer informatie te halen uit de grote hoeveelheden bewegingsdata.

Instructies

Deze sessie duurt ongeveer 60 minuten en bestaat uit drie onderdelen. Eerst volgt een vragenlijst die als doel heeft het verzamelen van informatie over het gebruik van bewegingsdata en visualisaties binnen uw organisatie. Na het beantwoorden van deze vragen, gaat u zelf aan de slag met een reeds bestaande visuele interface. Tot slot, laat ik u een animatie zien die zich richt op het toegankelijk presenteren van de gecoördineerde bewegingen en positie inname van verdedigers. Het spreekt voor zich dat persoonlijke informatie op geen enkele manier wordt gedeeld met derde. Deelnemers worden geanonimiseerd door middel van een identificatienummer (P1, P2, etc.) en toestemming zal worden gevraagd om tijdens de sessie audio- en schermopnamen te maken. De resultaten van het onderzoek worden na afloop met u gedeeld.

Vragenlijst

Het doel van deze vragenlijst is te weten te komen op welke vragen u graag een antwoord zou willen en op welke vragen kunt u op dit moment nog geen antwoord verkrijgen? Daarnaast wordt informatie verzameld over het gebruik van visualisaties tijdens het analyse proces.

Vragen:

1. Hoe zou u uw rol beschrijven op basis van uw taken en verantwoordelijkheden?
2. Heeft u weleens met (de resultaten van) bewegingsdata gewerkt?
3. Met behulp van welke systemen worden bewegingsdata verzameld?
4. Hoe worden deze systemen gebruikt, hoe vaak en door wie?
5. Met behulp van welke systemen wordt bewegingsdata geanalyseerd en door wie?
6. Zijn er taken waarbij u problemen ondervindt tijdens de analyse van bewegingsdata?
7. Welke informatie kunt u momenteel uit (de resultaten van) bewegingsdata halen?
8. Welke informatie kunt u nog niet uit (de resultaten van) bewegingsdata halen maar zal u er wel graag uit willen halen?
9. Maakt u gebruik van visualisaties tijdens de analyse van bewegingsdata?
10. Zijn er taken waarbij u ondersteund zou willen worden door middel van visualisaties?
11. Wat voor soort inzichten zouden visualisaties volgens u moeten ondersteunen?

SoccerStories

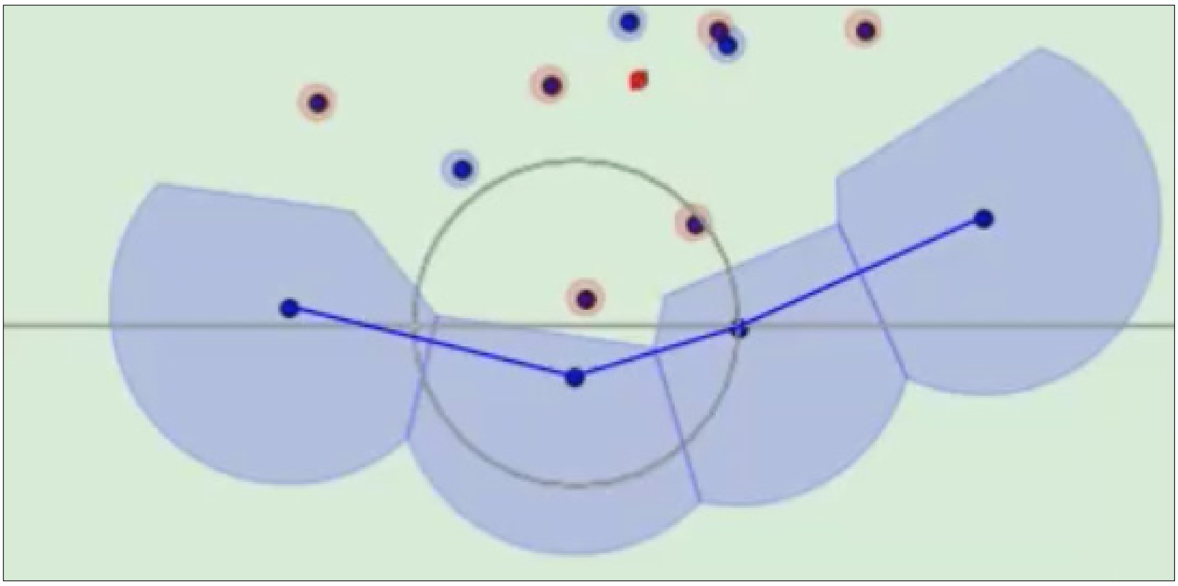
SoccerStories is een interactieve visuele interface. De interface maakt gebruik van fases om een overzicht te geven van het verloop van een wedstrijd. Iedere fase bestaat uit een serie van acties (pass, lange bal, lopen met de bal, voorzet, etc.) en eindigt met een schot richting goal. De interface bestaat uit meerderen componenten (zie afbeelding):

1. Het meest linker paneel geeft informatie over de opstelling. Door te klikken op individuele spelers worden statistieken per speler zichtbaar.
2. Het paneel midden bovenaan is de tijdlijn. De gekleurden bolletjes geven aan wanneer een bepaalde fase voorkwam. De kleur van bolletjes staat voor de uitkomst van iedere fase: rood is een gemiste kans, blauw is een schot gered en groen is een doelpunt. De tijdlijn kan worden gebruikt om te navigeren naar een fase van interesse.
3. Rechts van het voetbalveld is een overzicht van iedere fase weergegeven. Dit overzicht is gekoppeld aan de tijdlijn en het veld. Dit overzicht kan gebruikt worden om verschillende fases te vergelijken en te navigeren naar een fase van interesse.
4. Het centrale paneel geeft een voetbalveld weer. Hier kan iedere fase in detail worden bestudeerd door middel van verschillende visualisaties.
5. Het meest rechter paneel geeft annotaties/tekst weer die automatische worden gegenereerd op basis van de weergeven data.
6. Midden/rechts onderaan vindt u een besturingspaneel met een aantal knoppen en keuzevelden. Door in het besturingspaneel op de clusterize button te klikken begint de visuele verkenning en analyse van de data.
7. Spelers worden gevisualiseerd als een cirkel. Een pass wordt weergegeven door een doorbroken lijn. Bewegingen met de bal worden weergegeven door een kromme lijn.

Instructies: U krijgt nu 10 minuten de tijd om de interface zelf te gebruiken en te ontdekken welke informatie (door de visualisaties) wordt overgebracht. Ik zou u willen vragen op zoek te gaan naar de informatie die u graag zou willen hebben tijdens de analyse van de fases die hebben geleid tot een schot op goal. Tijdens het gebruik van de interface wil ik u vragen hardop te vertellen wat u doet, ziet en wat hierbij uw gedachten zijn. Daarnaast kunt u gebruik maken van de muis om zaken op het scherm aan te wijzen/klikken tijdens het praten. Tijdens het hardop denken volg ik uw opmerkingen en in het geval deze niet duidelijk zijn of als u stopt met praten zal ik u vragen stellen als: Zou u kunnen uitleggen waarom u dat denkt? Waar bent u naar aan het zoeken?

Gezamenlijk verdedigen

Om te kunnen omgaan met de complexiteit van de gecoördineerde bewegingen van spelers in relatie tot elkaar, de bal, de tegenstander en de omgeving waarin ze bewegen richt dit onderdeel op de gecoördineerde bewegingen en positie inname van verdedigers. Om dit inzichtelijk te maken laat ik u de komende 10 minuten naar een animatie van ongeveer 1 minuut kijken. Deze visualisatie is gebaseerd op bewegingsdata verzameld tijdens een trainingswedstrijd van een professionele Nederlandse voetbalclub. De visualisatie bevat de bewegingen van twee teams van elf spelers en de bal (rode stip). De vier verdedigers van het blauwe team zijn verbonden door een blauwe lijn. Daarnaast is de ruimte die zij controleren ten opzichte van de andere bewegende objecten (medespelers, tegenstanders en de bal) weergeven door middel van een blauwe geometrische vorm (zie onderstaande afbeelding).



Instructie: U krijgt nu 10 minuten de tijd om de animatie te bekijken en te ontdekken welke informatie (door de visualisatie) wordt overgebracht. Ik zou u willen vragen op zoek te gaan naar de informatie die u graag zou willen hebben tijdens de analyse van bewegingsdata. Oftewel, wat voor soort inzichten zouden visualisaties volgens u moeten ondersteunen?

Tijdens het bekijken van de animatie wil ik u weer vragen hardop te vertellen wat u ziet en wat hierbij uw gedachten zijn. Daarnaast kunt u gebruik maken van de muis om zaken op het scherm aan te wijzen/klikken tijdens het praten. U bent vrij om vooruit en achteruit te spoelen. Tijdens het hardop denken volg ik uw opmerkingen en in het geval deze niet duidelijk zijn of als u stopt met praten zal ik u vragen stellen als: Zou u kunnen uitleggen waarom u dat denkt? Waar bent u naar aan het zoeken?

7

Appendix 7: Example of transcription of research materials for one test person

01:03 R10: Kijk ook hier heb je volgens mij altijd een eerste vraag die je beantwoord wilt hebben waarbij statistieken wel kunnen helpen voor een groot gedeelte. Stel nou dat uit al die jaren en al die wedstrijden waar een feit uit voorkomt dat je weet dat de kans heel erg laag is als de speler van afstand gaat schieten. Dus hoeverre moet je als verdediger dan al agressief uit stappen met de kans dat je uitgekapt wordt en de kans eigenlijk veel groter wordt dat er uiteindelijk wel een doelpunt komt.

02:59 R10: Kijk als ik hier iets zinnigs over wil zeggen dan zou ik eerst moeten weten welk gebied moeten wij beschermen als verdedigers. Dat is waarschijnlijk wel iets wat je kan vangen in gegevens, wat ze ook in andere sporten hebben gedaan.

05:00 R10: Op het moment dat hij naar voren stapt dan kun je zeggen van kijk het verdedigt agressief naar voren op de bal, maar misschien blijkt uit al die statistieken en feiten van afgelopen jaren die er nu denk ik wel al zijn dat dit totaal niet gevaarlijk is als hij hier de bal heeft. Sterker nog wat ze nu doen is waarschijnlijk erg gevaarlijk, want ze gaan allemaal naar die bal toe.

07:24 R10: Wat ze volgens mij in andere sporten al veel meer doen is gewoon de gevaarlijkste zones, waar je punten of doelpunten kunt tegen krijgen, die ga je zo goed mogelijk verdedigen. En dat stukje kun je denk ik wel in statistieken vormgeven en dan tegelijkertijd terugkoppelen in die visualisatie hier.

07:53 R10: Alleen het begint dus wel dat de feiten kloppend moeten zijn. Dus als we ervan uit gaan dat 1 op de 90 voorzetten een succes is dan wil je dat de bal naar de zijkanten gaan en dat ze voorzetten geven, want de verdedigen we en dat wordt nooit een doelpunt.

12:31 R10: Alleen hierbij zou ik ook meer gegevens willen hebben, de pure feiten, over op welke manier nou de meeste goals worden gescoord.

23:03 I: Maar dan zou zoiets, als je van ten voren bedenkt welke gegevens je hebt, wel ondersteunend kunnen werken?

References

A

Adiseshiah, E. G., (2016, December 22). Minimalist web design: Dreaming of white (space) Christmas. Retrieved April 2017, from: www.justinmind.com blogminimalism-in-interactive-design-dreaming-of-a-white-space-xmas/

Anderson, C. & Sally, D. (2013). The Numbers Game: Why Everything You Know About Football is Wrong. London: Penguin books.

Andrienko, G., Andrienko, N., Budziak, G., Landesberger, T. Von, & Weber, H. (2016). Exploring Pressure in Football.

Andrienko, G., Andrienko, N., Dykes, J., Kraak, M.-J., & Schumann, H. (2014). GeoVisual analytics, time to focus on time. Information Visualization, 13(3), 187–189.

Andrienko, G., Andrienko, N., Bak, P., Keim, D., & Wrobel, S. (2013). Visual Analytics of Movement. Berlin/Heidelberg: Springer-Verlag.

Andrienko, N., & Andrienko, G. (2012). Visual analytics of movement: An overview of methods, tools and procedures. Information Visualization, 12(1), 3–24.

Andrienko, G., Andrienko, N., Keim, D., MacEachren, A. M., & Wrobel, S. (2011). Challenging problems of geospatial visual analytics. Journal of Visual Languages and Computing, 22, 251–256.

Andrienko, G., Andrienko, N., Dykes, J., Kraak, M.-J., & Schumann, H. (2010). GeoVA(t) – Geospatial visual analytics: focus on time. Journal of Location Based Services, 4(3–4), 141–146.

Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M., MacEachren, A., & Wrobel, S. (2007). Geovisual analytics for spatial decision support: Setting the research agenda. International Journal of Geographical Information Science, 21(8), 839–857.

Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J., & Kretchmer, U. (2002). Testing the Usability of Interactive Maps in CommonGIS. Cartography and Geographic Information Science, 29(4), 325–342.

Aquina, L. (2015, May 7). Metrica Sports: voetbal data voor Mourinho en Guardiola. Retrieved February 2017, from: http://www.sportknowhowxl.nl

B

Balciunas, A. (2013). User-Driven Usability Assessment of Internet Maps.

Bank, C. (2016, November 18). Understanding Web UI Visual Hierarchy. Retrieved April 2017, from: https://www.awwwards.com/understanding-web-ui-visual-hierarchy.html

Bank, C. & Zuberi, W. (n.d.). The Guide to Wireframing: For Designers, PMs, Engineers and Anyone Who Touches Product. Retrieved April 2017 from: https://www.uxpin.com/studio/ebooks/guide-to-wireframing/

Benkert, M., Gudmundsson, J., Hübner, F., & Wolle, T. (2008). Reporting flock patterns. Computational Geometry: Theory and Applications, 41(3), 111–125.

Bevan, N. (1995). Measuring usability as quality of use. Software Quality Journal, 4(2), 115–130.

Bialkowski, A., Lucey, P., Carr, P., Yue, Y., & Matthews, I. (2014a). "Win at Home and Draw Away": Automatic Formation Analysis Highlighting the Differences in Home and Away Team Behaviors. MIT Sloan Sports Analytics Conference.

Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S., & Matthews, I. (2014b). Identifying Team Style in Soccer using Formations from Spatiotemporal Tracking Data.

Bialkowski, A., Lucey, P., Carr, P., Denman, S., Matthews, I., & Sridharan, S. (2013). Recognising team activities from noisy data. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 984–990.

Bryman, A. (2012). Social Research Methods. New York: Oxford University Press.

Buchin, M., Dodge, S., & Speckmann, B. (2014). Similarity of trajectories taking into account geographic context. Journal of Spatial Information Science, 9(2014), 101–124.

Bychkov, D. (2013, June 7). Desktop vs. Web Applications: A Deeper Look and Comparison. Retrieved April 2017, from: http://www.seguetech.com/desktop-vs-web-applications/

C

Cao, J. (2016, April 22). How to Design a Useful Wireframe. Retrieved March 2017, from: http://blog.teamtreehouse.com/how-to-design-a-useful-wireframe.

Cao, J., Zieba, K., & Ellis, M. (n.d.). Interaction Design Best Practices: Mastering Words, Visuals, Space. Retrieved April 2017 from: https://www.uxpin.com/studio/ebooks/interaction-design-best-practices-tangibles/

Cox, A., & Stasko, J. (2006). SportVis : Discovering Meaning in Sports Statistics Through Information Visualization.

D

Delikostidis, I. (2011). Improving the usability of pedestrian navigation systems. ITC Dissertation, University of Twente Faculty of Geo-Information and Earth Observation ITC, Enschede.

Demšar, U., Buchin, K., Cagnacci, F., Safi, K., Speckmann, B., Van de Weghe, N., & Weibel, R. (2015). Analysis and visualisation of movement: an interdisciplinary review. Movement Ecology, 3(5), 1-24.

Demšar, U., & Virrantaus, K. (2010). Space-time density of trajectories: exploring spatio-temporal patterns in movement data. International Journal of Geographical Information Science, 24(10), 1527–1542.

Diepstraten, J. (2017, January 27). Retrieved February 2017, from: http://siliconcanals.nl/news/startups/football-fairytale-amsterdam-based-startup-metrica-sports/

Dodge, S., Weibel, R., Ahearn, S. C., Buchin, M., & Miller, J. A. (2016). Analysis of movement data. International Journal of Geographical Information Science, 30(5), 825–834.

Dodge, S. (2015). From Observation to Prediction: The Trajectory of Movement Research in GIScience. In Advancing Geographic Information Science: Chapter 9. Colorado: Department of Geography and Environmental Studies University of Colorado.

Dodge, S., Weibel, R., & Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. Information Visualization, 7(2008), 240–252.

Dykes, J., MacEachren, A. M., & Kraak, M.-J. (2005). Exploring geovisulization (First). Atlanta: Elsevier Ltd.

E

van **E**lzakker, C. P. J. M., & Wealands, K. (2007). Use and users of multimedia cartography. in Multimedia Cartography. ed. by Cartwright, W., Peterson, M. and Gartner, G., 487–504.

van **E**lzakker, C. P. J. M. (2004). The use of maps in the exploration of geographic data, Netherlands Geographical Studies. Borchert JG (ed). Koninklijk Nederlands Aardrijkskundig Genootschap / Faculteit Geowetenschappen, Universiteit Utrecht / International Institute for Geo-Information Science and Earth Obser- vation, Utrecht / Enschede.

F

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). Knowledge Discovery and Data Mining: Towards a Unifying Framework, 2, 82–88.

G

Giannotti, F., Nanni, M., Pedreschi, D., & Pinelli, F. (2007). Trajectory pattern mining. Discovery and Data Mining, 339.

Goldsberry, K. (2012). Courtvision: New visual and spatial analytics for the NBA. In Proc. 6th Annual MIT Sloan Sports Analytics Conference.

Goodchild, M. F. (2010). Twenty years of progress: GIScience in 2010. Journal of Spatial Information Science, 1(1), 3–20.

Gudmundsson, J., & Horton, M. (2016). Spatio-Temporal Analysis of Team Sports - A Survey. arXiv, (May), 184–193

Gudmundsson, J., & Wolle, T. (2014). Football analysis using spatio-temporal tools. Computers, Environment and Urban Systems, 47, 16–27.

Gudmundsson, J., Laube, P., & Wolle, T. (2012). Computational Movement Analsis.

H

Hägerstrand, T. (1970). What About People In Regional Science?, 'Papers in Regional Science', 24(1), 6-21.

Harrower, M., MacEachren, A. M., & Griffin, A. L. (2000). Developing a Geographic Visualization Tool to Support Earth Science Learning. Cartography and Geographic Information Science, 27(March 2015), 279–293.

Hennink, M., Hutter, I., and Bailey, A. (2010). Qualitative Research Methods. London: SAGE Publications Ltd.

Ho, Q. V. (2013). Architecture and Applications of a Geovisual Analytics Framework.

J

Janetzko, H., Stein, M., Sacha, D., & Schreck, T. (2016). Enhancing Parallel Coordinates: Statistical Visualizations for Analyzing Soccer Data. Electronic Imaging, 2016(1), 1–8.

Janetzko, H., Sacha, D., Stein, M., Schreck, T., Keim, D., & Deussen, O. (2014). Feature-Driven Visual Analytics of Soccer Data. In Visual Analytics Science and Technology (VAST), 2014 IEEE Conference.

K

Keim, D., Jörn, K., Ellis, G., & Mansmann, F. (2010). Mastering the Information Age Solving Problems with Visual Analytics. Goslar: Eurographics Association.

Keim, D., Andrienko, G., Fekete, J. D., Gorg, C., Kohlhammer, J., & Melancon, G. (2008). Visual analytics: Definition, process, and challenges. Lecture Notes in Computer Science. Department of Computer and Information Science, University of Konstanz.

Knutson, T. (2016). Understanding Football Raders for Muggs and Muggles. Retrieved April 2017 from: http://statsbomb.com/2016/04/understand-football-radars-for-mugs-and-muggles/#prettyPhoto

Kotzbek, G., & Kainz, W. (2016). Towards Automated GIS-based Analysis of Scoring Attempt Patterns in Association Football. AGILE 2006.

Kotzbek, G., & Kainz, W. (2015). GIS-Based Football Game Analysis – A Brief In- troduction to the Applied Data Base and a Guideline on How to Utilise It.

Kotzbek, G., & Kainz, W. (2014). Football Game Analysis: A New Application Area for Cartographers and GI-Scientists? 5th International Conference on Carthography and GIS, 299-306.

Koussoulakou, A., & Kraak, M. J. (1992). Spatia-temporal maps and cartographic communication. Cartographic Journal, The, 29(2), 101–108.

Kraak, M.-J. (2014). Mapping Time: illustrated by Minard's map of Napoleon's Russion campaign of 1812 (First). New York: Esri Press.

Kraak, M.-J., & Ormeling, F. (2010). Cartography: Visualiztion of Spatial Data (Thrid). Harlow: Pearson Education Limited.

Kraak, M.-J. (2008). From Geovisualisation Towards Geovisual Analytics. The Cartographic Journal, 45(2), 163–164.

Kveladze, I. (2015). Space - time cube design and usability. ITC Dissertation, University of Twente Faculty of Geo-Information and Earth Observation ITC, Enschede.

Kveladze, I., Kraak, M.-J., & van Elzakker, C. P. J. M. (2013). A Methodological Framework for Researching the Usability of the Space-Time Cube. The Cartographic Journal, 50(3), 201–210.

L

Laube, P., & Purves, R. S. (2006). An approach to evaluating motion pattern detection techniques in spatio-temporal data. Computers, Environment and Urban Systems, 30(3), 347–374.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. International Journal of Geographical Information Science, 19(6), 639–668.

Legg, P. A., Chung, D. H. S., Parry, M. L., Jones, M. W., Long, R., Griffiths, I. W., & Chen, M. (2012). MatchPad: Interactive Glyph-Based Visualization for Real-Time Sports Performance Analysis. Computer Graphics Forum, 31(3), 1255–1264.

Lewis, M. (2003). Moneyball: The Art of Winning an Unfair Game. New York: Writers House. New York: Writers House.

Long, J. a., & Nelson, T. a. (2012). A review of quantitative methods for movement data. International Journal of Geographical Information Science, 27(2), 1–27.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2011). Geographic Information Systems & Science (Third). Hoboken: John Wiley & Sons Inc

Lucey, P., Bialkowski, A., Monfort, M., Carr, P., & Matthews, I. (2015). "Quality vs Quantity": Improved Shot Prediction in Soccer using Strategic Features from Spatiotemporal Data. Proc. 8th Annual MIT Sloan Sports Analytics Conference, 1–9.

M

MacEachren, A. M., & Kraak, M.-J. (2001). Research Challenges in Geovisualization. Cartography and Geographic Information Science, 28(1), 3–12.

MIT Technology Review (2016, March 7). Big Data Analysis Is Changing the Nature of Sports Science. Retrieved January 2017, from: https://www.technologyreview.com/big-data-analysis-is-changing-the-nature-of-sports

Monmonier, M. (1990). Strategies for the visualization of geographic time-series data. Cartographica 27(1): 30-45.

Muller, G. (2012, November 6). Whitespace in Web Design: What it is and Why You Should use it. Retrieved April 2017, from: http://blog.teamtreehouse/white-space-in-web-design-what-it-is-and-why-you-should-use-it

N

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. Proceedings of the National Academy of Sciences, 105(49), 19052–19059.

Nielsen, J. (1993). Usability Engineering. San Francisco: Morgan Kaufmann.

Noor, A. M., Holmberg, L., Gillett, C., & Grigoriadis, A. (2015). Big Data: the challenge for small research groups in the era of cancer genomics. British Journal of Cancer, 113, 1–8.

Norman, D. (1988). The Design of Everyday Things. (Basic Books, Ed.). New York: Doubleday.

O

Opach, T., & Nossum, A. (2011). Evaluating the usability of cartographic animations with eye-movement analysis. 25th International Cartographic Conference, (1), 11.

P

Blog by **P**aste Interactive (2009, September 09). Content Wireframing: Duck meet row. Retrieved March 2017, from: http://gluue.com/2009/01/content-wireframing-ducks-meet-row/

Pecahna, S. & Wallace, T. (2015, June 20). The Flight of Refugees Around the Globe. New York Times. Retrieved April 2017 from: www.nytimes.com/world/map-flow-desperate-migration-refugee-crisis.html

Perin, C., Vuillemot, R., & Fekete, J.-D. (2013). SoccerStories : A Kick-off for Visual Soccer Analysis. In IEEE Transactions on Visualization and Computer Graphics, 19(12), 2506-2515.

Peuquet, D. J. (2002). Representations of space and time. New York: Guilford.

Peuquet, D. J. (1994). It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. Annals of the Association of American Geographers, 84(3), 441–461.

Pileggi, H., Stolper, C. D., Boyle, J. M., & John, T. (2012). SnapShot: Visualization to Propel Ice Hockey Analytics. IEEE Transactions on Visualization and Computer Graphics, 18(12), 2819-2828

Purves, R. S., Laube, P., Buchin, M., & Speckmann, B. (2014). Moving beyond the point: An agenda for research in movement analysis with real data. Computers, Environment and Urban Systems, 47, 1–4.

R

Roth, R., Ross, K., & MacEachren, A. (2015). User-Centered Design for Interactive Maps: A Case Study in Crime Analysis. ISPRS International Journal of Geo-Information, 4(1), 262–301.

Roth, R. E., & Maceachren, A. M. (2014). Geovisual Analytics & the Science of Interaction : A Case Study.

Roth, R. E. (2013). Interactive mapping: What we know and what we need to know. Journal of Spatial Information Science, 6(6), 59–115.

Roth, R. E. (2012). Cartographic Interaction Primitives: Framework and Synthesis. The Cartographic Journal, 49(4), 376–395.

Rusu, A., Stoica, D., Burns, E., Hample, B., McGarry, K., & Russell, R. (2010). Dynamic visualizations for soccer statistical analysis. Proceedings of the International Conference on Information Visualisation, 207–212.

S

Saidy, N. (2016, April 18). Getting started with Wireframes. Retrieved March 2017, from: https://blog.prototypr.io/getting-started-with-wireframes-8aff9b92a4c0

Schobesberger, D. (2012). Towards a Framework for Improving the Usability of Web-mapping Products, 379.

Sharp, E. (2015, June 2). 5 Ways Wireframing Helps You Plan for Better Content. Retrieved April 2017, from: http://www.protofuse.com/blog/website-wireframing-helps-plan-better-content/

Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrmann, S., & Hedley, N. R. (2001). Cognitive and Usability Issues in Geovisualization. Cartography and Geographic Information Science, 28(1), 61–75.

Stein, M., Janetzko, H., Seebacher, D., Jäger, A., Nagel, M., Hölsch, J., Grossniklaus, M. (2017). How to Make Sense of Team Sport Data: From Acquisition to Data Modeling and Research Aspects. Data, 2(1), 2.

Stein, M., Janetzko, H., Breitkreutz, T., Seebacher, D., Schreck, T., Grossniklaus, M., Keim, D. A. (2016). Director's Cut: Analysis and Annotation of Soccer Matches. IEEE Computer Graphics and Applications (CG&A).

Stein, M., Häußler, J., Jäckle, D., Janetzko, H., Schreck, T., & Keim, D. A. (2015). Visual Soccer Analytics: Understanding the Characteristics of Collective Team Movement Based on Feature-Driven Analysis and Abstraction. ISPRS International Journal of Geo-Information, 4(4), 2159–2184.

Stevenson, A., & Lindberg, C. A. (2010). New Oxford Dictionary of English (Third). New York: Oxford University Press, Inc.

Sumpter, D. (2016). Soccermatics: Mathematical Adventures in the Beautiful Game. London: Bloomsbury Publishing Plc.

T

Thomas, J. J., & Cook, K. A. (2005). Illuminating the path: The research and development agenda for visual analytics. IEEE Computer Society, 184.

Tominski, C., Schumann, H., Andrienko, G., & Andrienko, N. (2012). Stacking-based visualization of trajectory attribute data. IEEE Transactions on Visualization and Computer Graphics, 18(12), 2565–2574.

Tufte, E. R. (1983). The Visual Display of Quantitative Information. Cheshire: Graphics Press.

W

Wassink, I., Kulyk, O., Van Dijk, B., van der Veer, G., & van der Vet, P. (2008). Applying a user-centered approach to interactive visualisation design. Advanced Information and Knowledge Processing, 36, 175–199.

Z

Zhao, J., Forer, P., & Harvey, A. S. (2008). Activities, ringmaps and geovisualization of large human movement fields. Information Visualization, 7, 198–209.

van **Z**oelen, B. (2016, December 2016). Retrieved January 2017, from: http://www.parool.nl/amsterdam/amsterdamse-start-up-metrica-sports-ziet-meer-dan-voetbalanalist~a4432052/

URLs

1. Website SciSports. Retrieved September 2016, from: http://www.scisports.com
2. Website Metrica Sports. Retrieved October 2016, from: http://metrica-sports.com
3. Website official documents FIFA. Retrieved December 2016, from: http://www.fifa.com/about-fifa/official-documents/
4. Website use, User and Usability Issues bibliography ICA. Retrieved November 2016, from: http://use.icaci.org/bibliography/
5. Website Wyscout. Retrieved January 2017, from: https://wyscout.com/
6. Website Instat. Retrieved January 2017, from: http://instatfootball.com/
7. Website Ortec. Retrieved January 2017, from: http://ortecsports.com/
8. Website Opta Sports. Retrieved January 2017, from: http://www.optasports.com/
9. Website Scout7. Retrieved January 2017, from: http://info.scout7.com/en/
10. Website STATS. Retrieved January 2017, from: https://www.stats.com/
11. Website Gracenote. Retrieved January 2017, from: http://www.gracenote.com/
12. Website webinar Catapult Sports. Retrieved January 2017, from: http://www.catapultsports.com/uk/education/webinars/
13. Website SportVU player tracking system. Retrieved January 2017, from: https://www.stats.com/sportvu-basketball/
14. Website Second Spectrum. Retrieved January 2017, from: https://www.secondspectrum.com/
15. Webinar Sport Radar. Retrieved January 2017, from: https://www.sportradar.com/
16. Website Player and puck tracking NHL. Retrieved January 2017, from: http://www.sporttechie.com
17. Website Zebra Sports. Retrieved January 2017, from: https://www.zebra.com/us/en/solutions/location-solutions
18. Website TRACAB optical tracking. Retrieved January 2017, from: http://chyronhego.com/sports-data/tracab
19. Website BallJames optical tracking. Retrieved January 2017, from: http://www.balljames.scisports.com/
20. Website polar. Retrieved January 2017, from: https://www.polar.com/uk-en/b2b_products/team_sports/team_pro
21. Website Viper Pop, STAT Sports. Retrieved January 2017, from: http://statsports.com/technology/viper-pod/
22. Website Catapult Sports. Retrieved January 2017, from: http://www.catapultsports.com/uk/system/outdoor/
23. Website Johan Sports. Retrieved January 2017, from: http://www.johan-sports.com/
24. Website Inmotio. Retrieved January 2017, from: http://inmotio.eu/nl-NL/34/voetbal.html
25. Website Society for American baseball research. Retrieved January 2017, from: http://sabr.org/

26. Website Disney Research. Retrieved November 2016,

from: https://www.disneyresearch.com/project/strategy-analysis-in-soccer/

27. Website SAP. Retrieved February 2017,

from: http://www.sap.com/solution/industry/sports-entertainment/team-management/sports-one.html#

28. Website Aviz visual analytics project SoccerStories. Retrieved January 2017, from: http://www.aviz.fr/soccer

29. Website Otranscribe. Retrieved February 2017, from: http://otranscribe.com/

30. Website NVivo. Retrieved February 2017, from: http://www.qsrinternational.com/what-is-nvivo

31. Website QuickTime Player. Retrieved February 2017, from: https://support.apple.com/quicktime

32. Website Adobe Experience Design CC (Beta). Retrieved March 2017,

from: https://www.adobe.com/nl/products/experience-design.html

33. Website adobe XD UI kits and resources. Retrieved March 2017, from: http://www.xdguru.com/adobe-xd-ui-kits/

34. Website on football tactics. Retrieved December 2016, from: http://www.guidetofootball.com/tactics/tactical-dynamics w

35. Website Tableau. Retrieved January 2017, from: https://www.tableau.com/

36. Website Qlik. Retrieved January 2017, from: http://www.qlik.com/nl-nl/

37. Website SAS. Retrieved January 2017, from: https://www.sas.com/nl_nl/software/business-intelligence/visual-analytics.html

38. Website TIBCO Spotfire. Retrieved April 2017, from: http://spotfire.tibco.com/

39. Website Domo. Retrieved April 2017, from: https://www.domo.com/

40. Website user interface elements. Retrieved January 2017,

from: https://www.usability.gov/how-to-and-tools/methods/user-interface-elements.html

41. Website interface flow patterns. Retrieved April 2017,

from: http://uxmovement.com/products/flow-patterns-make-site-flows-in-fine-visual-detail/

42. Website squawka. Retrieved May 2017, from: http://www.squawka.com/news/

43. Website Experimental361. Retrieved May 2017, from: https://experimental361.com

44. Website Veldacademie Rotterdam. Retrieved April 2014, from: http://www.veldacademie.nl/

projecten/gedeeld-verleden-gezamenlijke-toekomst

Curriculum Vitae

Lars van Hoeve was born on the 25th of August 1991 in Goes, the Netherlands. From 2004 until 2010 he followed his pre-university education at the Ostrea Lyceum in Goes. In 2010, he embarked on his studies in Social Geography and Urban Planning at the University of Utrecht. Urban planning was the main subject, with cartography, cultural history, Geographical Information Science and planning as subsidiary subjects. In 2013, he obtained a minor in Architecture, Urbanism and Building Science at the Delft University of Technology. In association with the Veldacademie in Rotterdam he wrote his Bachelor Thesis titled: *common past, shared future* (URL 44). In 2014, he was awarded the Bachelor's Degree. In 2016, he started with the Geographical Information Management and Applications Master of Science. After this research project, he will continue his interests in geovisual football analytics during his internship at SciSports.