# Video-based Analysis of Soccer Matches

Maximilian T. Fischer
University of Konstanz
Konstanz, Germany
Max.Fischer@uni-konstanz.de

Daniel A. Keim
University of Konstanz
Konstanz, Germany
Daniel.Keim@uni-konstanz.de

Manuel Stein
University of Konstanz
Konstanz, Germany
Manuel.Stein@uni-konstanz.de

## ABSTRACT

With the increasingly detailed investigation of game play and tactics in invasive team sports such as soccer, it becomes ever more important to present causes, actions and findings in a meaningful manner. Visualizations, especially when augmenting relevant information directly inside a video recording of a match, can significantly improve and simplify soccer match preparation and tactic planning. However, while many visualization techniques for soccer have been developed in recent years, few have been directly applied to the video-based analysis of soccer matches. This paper provides a comprehensive overview and categorization of the methods developed for the video-based visual analysis of soccer matches. While identifying the advantages and disadvantages of the individual approaches, we identify and discuss open research questions, soon enabling analysts to develop winning strategies more efficiently, do rapid failure analysis or identify weaknesses in opposing teams.

## CCS CONCEPTS

• **General and reference** → **Surveys and overviews**; • **Human-centered computing** → **Visual analytics**; **Information visualization**; *Mixed / augmented reality*; *Interaction techniques*.

## KEYWORDS

visual analytics; information visualization; sports analytics; soccer analysis; immersive analytics

## 1 INTRODUCTION

The awareness of context is an important aspect for many applications incorporating and analyzing collective behavior data, of which team sports data is one rather specific but interesting example. One of the most famous team sports worldwide is soccer, with a large number of competitions held every year, many with huge financial and political investments and returns. Analyzing soccer matches to develop and implement winning strategies, do

opposition research and failure analysis is relevant from an economic as well as a reputational perspective. Therefore, matches of national teams are often politicized and matter in a symbolic way, which further acts as a catalyst for larger financing (cf. [19]). Working with collective behavior data is hard because the information and interrelations are rather complex. The data is, in general, a combination of hypervariate, hierarchical, relational, and temporal types. This leads to two challenges (cf. [17]). On the one side are the analytical aspects, which aim to find meaning inside the data, while on the other side is the visualization not only of the data itself, but also the information and knowledge that is derived from it. The latter enables us to support and enhance our understanding through meaningful representations and can give analytical insights, which help to reach the goals set out above.

A significant amount of research has been done on visualization techniques during the last decades. Meaningful progress has been made for the visualization of large multivariate, multidimensional, multitemporal, and spatial datasets. Information about and comparisons between them are available in the standard literature [10, 11, 15, 16]. However, while some of the more generic approaches like tables, diagrams, and 2D maps have been applied to visualize sports data, few approaches have been developed directly with soccer in mind. Even fewer approaches exist that visualize relevant information *in context*, i.e., directly embedded in a live stream or video recording, termed video-based visualizations. Video-based visualizations for soccer (and other team sports) have the advantage of augmenting information directly relevant to the actual scene. Furthermore, employed video-analysts by professional clubs are used to segment and analyze soccer matches manually based on available video recordings, which currently results in a very tedious and time-consuming process [7]. Visualization superimposed on the original video recordings ideally enable the addition of relevant information in real time during the match, directing the analysts attention to the essential aspects and events while maintaining the context to the real world, thus, enabling a user to form an extremely robust mental map (cf. [3]). This, in turn, enables a more robust contextual understanding, makes forming connections between actions easier, and improves memorability.

Currently, most soccer match reports contain a significant number of statistics, presented mostly in large tables and only sometimes using (very simple) static graphics like shot distributions and set play analysis to visualize results. These visualizations typically include line diagrams, bar charts, positional plots, or star glyphs. Even when video-based visualizations are used in television recaps or summaries, they are usually created and placed manually or semi-automatic at best, which requires effort and carries significant cost with it. Furthermore, video-based approaches are mostly limited to track players, show offside positions or movements. Even these rather single-feature focused statistics and simple visualizations are

an improvement of methods employed for the most parts of the last century, which involved simple counts (like number of successful passes, possession time, etc.) which have been read out or were later shown as overlay tables, with more detailed observations provided by subjective reports from experts. Recent progress in digital match tracking, computer graphics, and image manipulation has lead to increased capabilities in these regards, making advanced video-based visualizations technically feasible. However, the discipline is in its infancy, largely restricted to a commercial solution developed out of necessity, with only a few academic contributions so far, and much room for improvement still exists (cf. [29]).

In order to close this gap and improve as well as foster research in the field of video-based analysis of soccer matches, we provide a comprehensive and categorized overview of the latest, non-trivial methods developed for video-based visualization of soccer matches. We, specifically, do *not* survey solutions which only provide tagging and enhanced video libraries like Hudl Sportscode [13] or MyDartfish [9], as they do not employ video-based visualizations yet. However, plans in this regard are announced for the latter. In general, no current survey paper exists that describes the narrow field of video-based visualization of soccer matches completely. Existing surveys are either outdated [8] or incomplete [20] as they describe the whole field of sports data visualization. Other surveys like [23, 25, 29] can only be considered partial reviews (outdated, commercial solutions only, ...) or deal with a specific sub-issue. We survey the existing literature including outdated part-surveys (Section 2) and, furthermore, differentiate described video-based approaches from other visualization and analysis methods (Section 3). The existing approaches found in the literature are each presented shortly before we compare them according to eleven identified relevant features. We discuss differences and commonalities in detail, with a qualitative overview given in Table 1 for reference (both Section 4). Further we identify key issues and unsolved research challenges in the field and discuss research opportunities, before we provide a short summary of our findings (Section 5).

## 2 RELATED WORK

While a high volume of literature on visualization techniques exists [10, 11, 15, 16], which are extensively used in fields such as physics, the research is decidedly more sparse on the topic of video-based visualizations and in particular on video-based approaches for soccer. The recent work by Perin et al. [20] provides a quick overview of visualizations developed for sports data, although few techniques can be regarded as video-based, in particular for soccer. For those video-based approaches, an older survey by Borgo et al. [8] provides a general overview up to the year 2012, while developing useful classification schemes. However, with the fast pace of technological development, the referenced approaches are mostly outdated. Other surveys like Rehman and Saba [23], Shih [25], Thomas et al. [29] can only be considered partial reviews or deal with a specific sub-issue. They are either outdated, not complete, or consider only commercial approaches [28]. Few solutions have been proposed in the field of video-based soccer visualization. The first relevant work can be attributed to Assfalg et al. [2] as early as 2003. It was one of the first to annotate soccer matches with calculated information, extracted from the video itself, to support

the understanding of what is shown on screen. In the paper, ways are discussed to annotate videos with simple match scene information. One of the most important contributors to the field, especially in the 2000s but also, later on, was Graham Thomas. His research as part of BBC R&D (British Broadcasting Company, Research & Development) resulted in several commercial projects [4–6], which pioneered novel solutions for sports visualization, including the Piero system [22], one of most widely used sports visualization and annotation system today and the main competitor to the Viz Libero system [30]. Both systems are developed for sports broadcasting applications and provide a visualization and special effects toolbox, tailored to supported games like soccer. Both try to automate some aspects like tracking to reduce user workload.

In general, as detailed by Thomas et al. [29], large commercial applications like Augmented Video Player [5], True View [14], Piero [22], and Viz Libero [30] for video-based soccer visualization are actually ahead of *most* of the academic research and represent the state-of-the-art, with the only notable advanced academic exceptions being Stein et al. [26, 27]. Apart from these autonomously working solution, few other approaches with the same level of sophistication exist. The work by Schlipsing et al. [24] looked initially promising, but was abandoned after the principal investigator completed his Ph.D. Some other approaches exist, like Andrade et al. [1], Liu et al. [18], Wan et al. [31], Xue et al. [32]. Nevertheless, these latter approaches remain at basic levels of visualization and surprisingly few new ideas were proposed. As described by Thomas [28], some individual and specific contributions have been made to the field of sports data analysis and visualization, however

> *whole systems which allow for automated tracking or labeling remain an open challenge (cf. [28]).*

## 3 METHODOLOGY

**Definitions and Terminology.** Soccer, official known as *Association football* and often–especially in Europe–just called *football*, is a very popular team sport played between two competing teams of eleven players each. It is played on a grass field of rectangular size and the aim is to kick a spherical ball into the opponent's goal as often as possible, generally lasting 90 minutes in two parts. More detailed rules exist but are of no concern here. Games are very popular and several leagues, cups, and competitions have formed over time, often with huge financial impact. Games are often recorded as videos or transmitted live, depending on the venue, with several thousand or sometimes even millions of spectators. As one key to winning such games is understanding the opposing team and their tactic, past games are analyzed extensively. This process is currently done either manually or with semi-automated supportive technology and very time consuming as well as expensive [7]. Here the automated match analysis comes in. As dedicated tracking devices can be expensive and more importantly would require the consent of the opponent teams, one is left with the only non-invasive technique remaining: video recordings of past and currently running matches. This leaves video-based analysis as the only viable option for assessing an arbitrary soccer match. The overall analysis process can be split into three different steps: extracting movement and event data from video data, analyzing this numeric and event data for assessments, and visualizing this information, either separately

(e.g., as a statistics table) or reintegrate it into the video itself as a form of augmented reality. In this survey, we focus mostly on the latter part, but often this is impossible without considering the first two steps. Therefore, we also describe these aspects if necessary. The central aim of video-based soccer analysis can be regarded as searching for ways to visualize relevant match information inside the actual video. This can support a direct mental map, provide contextual information, and thereby enable coaches to significantly improve and simplify soccer match preparation and tactic planning. Video-based visualizations help to more effectively develop winning strategies, perform extensive failure analysis and identify weaknesses in opponents which are hard to detect otherwise. Therefore, we provide an overview of the existing literature which answers the following research question:

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

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

## 4 SURVEY

In the following Section 4.1, we give an overview of standard approaches and the most promising novel techniques proposed in literature and industry. In Section 4.2, we propose classification methods and different criteria to assess the presented papers. We follow up with a detailed comparison of the presented approaches for each feature individually. A qualitative overview of the most important features of the approaches we consider most relevant is also given in Table 1 for reference, which is presented after the approaches and comparison criteria.

### 4.1 Overview of current approaches

In the following, we present the individual approaches. The order in which the approaches are presented in is related to their *complexity*. However, the details of this categorization can be safely ignored for now and the concept is introduced later, when relevant for the comparison in Section 4.2.

**Basic: Textual/Numeric Table Overlay.** The most simple form of video-based data visualization is at the same time one of the oldest. Here, the information, which is typicality text or numbers, is displayed as an overlay on top of the actual video, and shown for several seconds for it to be perceived. The only relation to the



**Figure 1: A table overlay showing current match statistics, while partly occluding the gameplay. The metrics display the relative strength of both teams and aim to explain the goal scored some seconds ago. Prepared by [21].**

actual video is the timing, which can–but does not have to–match a current match situation. Information might include the current number of goals and some match statistics. An example of such a table overlay with some match statistics like goals, offsides, ball possession rate, etc. is shown in Figure 1.

**Automatic Highlight Identification.** The approach by Assfalg et al. [2] aims at automatically extracting highlights from a soccer game by analyzing the corresponding video sequences. First, playfield regions are extracted by their shape for localization. Then, player bounding boxes are found by color differentiating and adaptive template matching. To discern between different scenarios, the most likely scenario is chosen by using the normalized player position density distributions of example scenes. Then, a domain knowledge model computes the currently most likely scenario based on past predictions. The identified scenario is then reintegrated in the video in a bottom bar as event boxes for specific events, labeled as *none*, *hypothesis*, *accepted*, and *rejected*. While the detection technique was state-of-the-art at the time of publication and the accuracy was very good, the visualization was simple and did not evolve much.

**Highlights Extraction with Content Augmentation.** The idea by Wan et al. [31] differs somewhat from the previous aim by Assfalg et al. [2]. Here, the first goal is to separate the video into different segments, using audio and visual cues, but only focusing on detecting replays. The second, more relevant goal is to identify regions for content augmentation. This is done by first determining the temporal and spatial relevance of the scene, as an insertion in a highly relevant scene would be disturbing. A suitable insertion area is then found by color quantization of homogeneous regions. No specific limits are set to the content to be inserted. However, as the position is in principle arbitrary and not related to any content, the most likely use is for non-obtrusive advertisement or other non-positional information.

**Region-based Tracking and Provision of Augmented Information.** The approach by Andrade et al. [1] differs from others

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

**Virtual Content Insertion System.** The work by Liu et al. [18] is very similar to the second part of the one by Wan et al. [31]. It proposes a generic Virtual Content Insertion system, which places content inside a video when the underlying actions attract much attention, but at a position where it does not obstruct the storyline. An adequate time is found by analyzing the spatial and temporal attention and the position using the three measures of saliency, contrast, and novelty. The insertion supports affine transformations to align the content with the actual environment.

**Automatic Video Annotation System.** The system presented by Xue et al. [32] aims at extracting metadata from video sequences of sports events. Sequence boundaries are segmented and clustered according to the type of content. Players cannot be tracked automatically but must be laboriously selected manually for tracking. On-screen text is recognized using trained SVM's (support vector machines) and read using OCR (optical character recognition). The combined information is presented as a simple on-screen graphic.

**Video-based Activity Recording.** In his thesis, Schlipsing et al. [24] presents a complete recording (calibrated dual-camera setup) and analysis system workflow for sports data analysis. Players are tracked semi-automatically using segmentation. Initial assignments and re-identifications need to be made manually. In terms of in-video visualization, player positions, and numbers, player formations (between manually selected players), and compactness hulls are presented. It was one of the first approaches which featured the placement of virtual objects in a soccer video for augmentation purposes.

**Combining Video and Movement Data to Enhance Team Sport Analysis.** The approach presented by Stein et al. [27] is somewhat similar to the main idea followed by Schlipsing et al. [24], but differs in two fundamental ways in the implementation: First, it aims to use already existing video recordings–like a scouting feed or simple television recordings–as a video source, instead of a carefully calibrated dual-camera setup. Secondly, the tracking of the players aims to be autonomous and not require manual input. Then, in terms of visualization, the approach has a similar idea to enhance the gameplay and analysis by presenting visualizations in an augmented video. Here, several new ideas of features for visualization are presented: interaction spaces (the personal space that can be controlled by each player), free spaces, dominant regions, pass distances per player, and player reactions. An example visualization is shown in Figure 2.

**Explanatory Storytelling for Advanced Team Sport Analysis.** The follow-up paper to the previous approach, also by Stein et al. [26], proposes a system for automatic selection of helpful



Figure 2: Interaction space visualization by Stein et al. [27]. The augmented circles show the personal space that can easily be reached by the players and are helpful, for example, to identify suitable receivers for a pass.

visualizations in key match situations. As part of this, two further visualization ideas are presented or improved upon: The pressure imposed upon players to act (aggressiveness of the defending team to recapture the ball) and an improved version of free/interaction spaces (spaces where nobody is standing or the personal space that can be reached by each player).

The previously discussed approaches represent the state-of-the-art in the academic literature. In the following, we focus on commercial products that are deployed in the field. According to Thomas et al. [29] and Hilton et al. [12], the most widely used visualization products in the industry are *Piero* [22], developed by the BBC R&D in collaboration with Ericsson, and *True View* [14] from Intel, former FreeD by Replay Technologies. Additionally, *Viz Libero* [30] is regarded as the third large broadcasting solution. An overview of some of the systems is also given by Hilton et al. [12] and, more up to date, by Thomas et al. [29]. The commercial solutions are:

**Commercial: iView (2005-2008).** Developed by the BBC between 2005 and 2008, the iView: free-viewpoint video system [4] allows to reconstruct a 3D scene from the 2D video sequences and allows free positioning in them. In theory, arbitrary objects could be placed in the resulting 3D scene, but this was rarely done and the quality, as well as the resolution of the model, was very basic. The system was superseded by Intel True View (see below).

**Commercial: Augmented video Player (2012-today).** Also developed by the BBC since 2012, the Augmented Video Player [5], allows a user to watch sports sequences and interactively overlay additional layers to augment the video with further information.

**Commercial: Intel True View (2012-today).** The Intel True view system [14] is similar to the iView system and can be considered its technological successor. Formerly developed by FreeD Replay Technologies, it enables to reconstruct high-resolution 3D scenes using several HD cameras from different angles to show matches as a 3D scene. The quality is markedly better and some simple (manually annotated) visualizations are used, like trajectories or
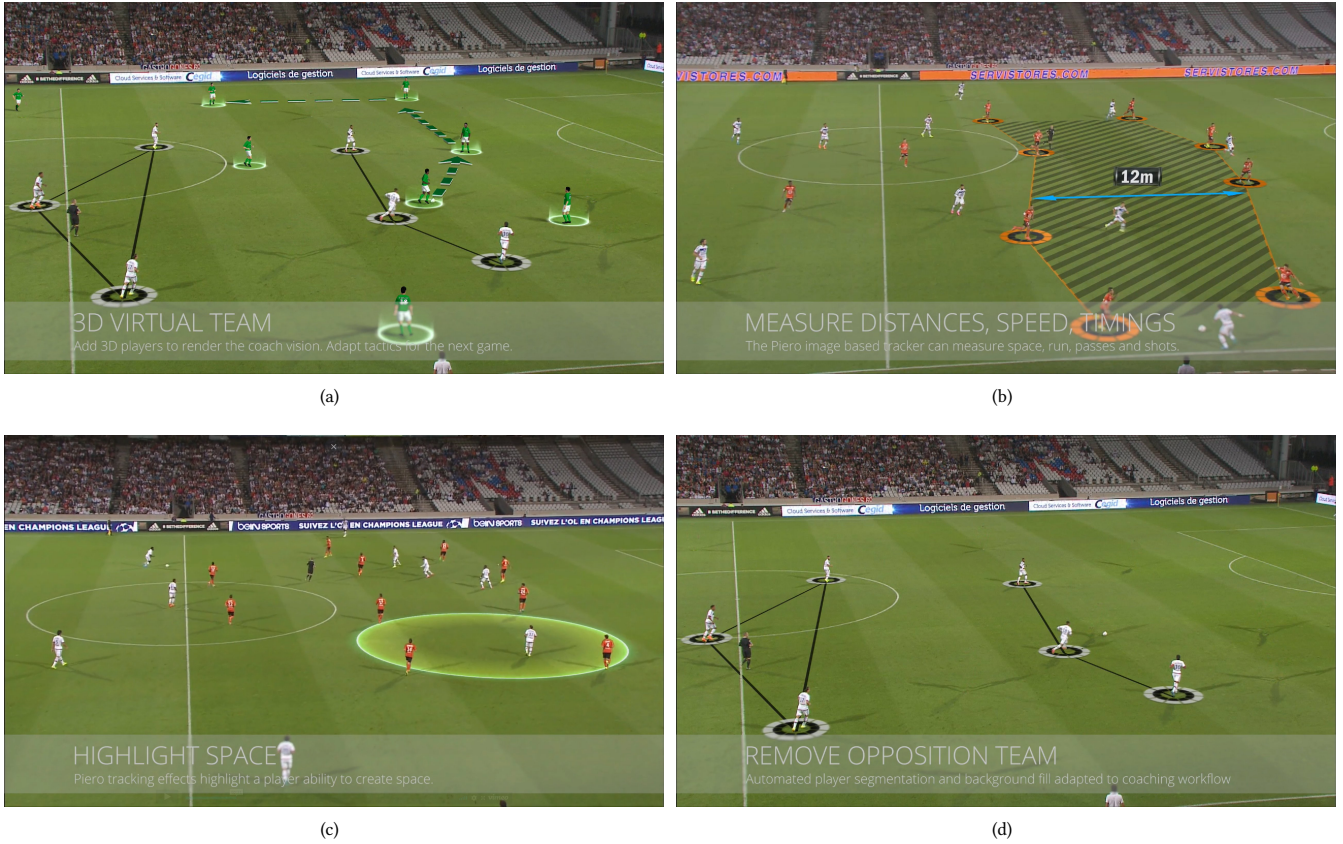
(a)



(b)



(c)



(d)

Figure 3: Four different types of visualization which are available in the Piero System [22].

highlighting.

**Commercial: Piero (2004-today).** The Piero system [6, 22] is considered one of the most advanced and widely used systems on the market. It was pioneered by the BBC, then outsourced to Redbee-media in collaboration with Ericsson. It is used in different sports, including soccer, to augment the video with additional information. This aspect makes it come very close to the actual core of the survey. Tracking of players is done semi-automatically (initial assignment and re-identification are done manually) and several automatisms exist. However, the core of the systems remains a video-effects software, specifically tailored to real-time analytics of sports games. For soccer, several manually placeable visualizations are available: They include formations, free spaces, interaction spaces, distance measurements, ball trajectories, removable players, goal area heatmaps, player trajectories, density distributions, and others. Some of these examples are demonstrated in Figure 3. However, there is still a lot of manual work involved in creating animations and overlays. The main difference between the Piero system to the approach by Stein et al. [26] is that the latter is a fully automated approach while the former represents the state-of-the-art in manually annotated video-based soccer analysis with an advanced graphical display. The visualizations by Stein et al. [26] can also be considered advanced, but, more importantly, are based

on mathematical and physiological considerations. However, sometimes they lack the clean and typographically nice style. Combining those two approaches would likely lead to enormous benefits.

**Commercial: Viz Libero (2010-today).** The Viz Libero system [30] is considered, together with the previously presented Piero system, the most advanced and widely used system on the market. It is also used for different sports, including soccer and augments the video with additional information. Available broadcast streams or video recordings are used as input streams, which means that no dedicated camera setup is required. Tracking is also semi-automatic and again several automatisms exist. The available visualizations are similar to the ones offered by Piero. One striking difference between Piero and Viz Libero is that the latter uses homography estimations and therefore allows limited reconstruction from single video sources. This means the camera viewpoint can be freely (within limits) be re-positioned within the scene to get the view from another perspective. One example of an augmented scene in Viz Libero is shown in Figure 4.

## 4.2 Categorization and Comparison

We first propose possible classifications and criteria to compare the different approaches. We establish the following nomenclature to
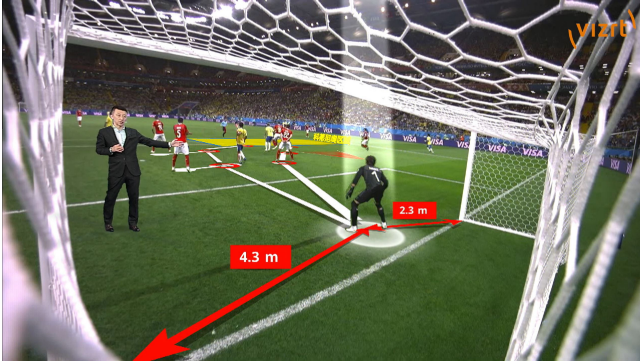
**Figure 4: Annotated game scene with added distance measurements, highlighting and perspectively correct content placement by Viz Libero [30].**

classify the techniques into different categories. First, we differentiate between **scientific research** and **commercial applications**. For the latter, it is often the case that no paper exists that describes the system in detail, so one is left with product descriptions of the producer, manuals, as well as demo videos. Secondly, we differentiate between the **complexity** of the approaches, ranging from **low** over **medium** to **high**. A special case is a simple data table overlay, which we assign to the class **very low**, and is provided for reference as the simplest technique available. To judge the amount of manual work required or how far an approach can work autonomous, we classify the proposals according to their **automation level**, also ranging from **(very) low** over **medium** to **high**. In addition, we distinguish between the approaches according to several key concepts: Are they **insertion** or **reconstruction** techniques? Is augmented information contextually **embedded** in the actual scene or merely overlayed and positioned? In this context, we further differentiate between static and support for **dynamic** content as well as the support to **interact** with the visualization, for example, by clicking or settings different parameters. Further options for comparison are if the content is **placed** at fixed positions or positioned intelligently and if the system can in principle work in **realtime** or not. To assess the impact of the research we further differentiate the papers and approaches by the presence or absence of a usability or case **study**, either in the paper itself, as a follow-up paper, or in form of industrial applications. Lastly, it is relevant to note if the technique is still actively researched or even **used** in the field. A qualitative comparison of the important features from the presented approaches is given in Table 1. In this table, the individual features are grouped semantically, serving as an overview of the main findings of this survey. Sometimes the labels given in the table are over-simplified and do not give the whole aspect credit. The full evaluation is given in text-form in the following sections for each feature individually.

**Complexity.** The most basic approach, with very low complexity, is just displaying a data table (or other types of simple visualization without location information) as an overlay on top of the video during the relevant period. The approaches with low complexity

are all insertion techniques which employ non-sophisticated positioning inside the video and, with the exception of Liu et al. [18], all operate in the screen space instead of the world space (embedding techniques). Several approaches of Assfalg et al. [2], BBC R&D (Augmented Video Player) [5], Andrade et al. [1], Wan et al. [31], Liu et al. [18], Xue et al. [32] belong in this category. More sophisticated ideas with medium complexity are introduced by BBC R&D (iView) [4] and Schlipsing et al. [24]. The information is visualized as part of the scene itself. The approaches Piero [6, 22], Viz Libero [30], True View [14], Stein et al. [26, 27], with high complexity are those which employ advanced, embedded visualizations or reconstruction techniques, which require significant work to develop.

**Automation.** The automation level indicates how much user input is necessary for the analysis and to produce the visualization. Again, the simple approach of just displaying a data table contains **very low automation**, as variables like placement, size, and duration all have to be determined manually. The vast amount of approaches like Assfalg et al. [2], Wan et al. [31], Andrade et al. [1], iView [4], Liu et al. [18], Augmented Video Player [5], Xue et al. [32] offer **low automation** and often several manual steps are necessary. The approach by Schlipsing et al. [24] offers a higher degree of automatic visualizations, leading to **medium automaton**. However, the initial selection of players for tracking and then the re-identification is still solved manually. The same is true for the systems True View [14], Piero [22], and Viz Libero [30]. These approaches all offer automatic support to reduce repeating and laborious steps. Nevertheless, the type of visualization and initial positions have to be specified manually. More often than not, even the actual content of the visualization is created manually by combining different building blocks. A **high automation** level is only reached by the systems presented by Stein et al. [26, 27]. The visualizations are created automatically and without any user input, when using the default parameters.

**Insertion.** Techniques are considered insertion techniques when content information is placed inside the video, during a specific time and at a specific location. Most of the presented approaches belong to this category. The exceptions are iView [4] and True View [14], which belong to the reconstructive techniques.

**Embedded.** This subcategory of insertion techniques contains approaches which actually embed the content into the 3D scene and respect relevant aspects like camera distortion and perspective projections. Approximately half of the approaches can be considered of this type, namely the ones by Liu et al. [18], Schlipsing et al. [24], both from Stein et al. [26, 27], Piero [22], and Viz Libero [30]. The four latter ones are the most sophisticated, leading to aesthetically pleasing results.

**Reconstruction.** Reconstructive techniques are not (primarily) based on inserting content into the video to help with the analysis, but reconstruct the scene as a 3D object to allow for arbitrary positioning. This can support match analysis by enabling different views. While the field of 3D reconstruction and structure from motion, part of image analysis and computer graphics, is relatively

Table 1: Comparison overview of the visualization techniques for video-based soccer match analysis we reviewed. Interesting to note is the low amount of current academic research and the relatively large amount of commercial solutions. For a detailed discussion, see Section 4.2.

| | Table Overlay | Assfalg [2] | Wan [31] | Andrade [1] | iView* [4] | Liu [18] | AVP*,3 [5] | Schlipsing [24] | Xue [32] | Stein4 [27] | Stein4 [26] | True View* [14] | Piero* [6, 22] | Viz Libero* [30] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *(year)* | 1960+ | 2003 | 2004 | 2005 | 2005 | 2008 | 2012 | 2014 | 2017 | 2018 | 2018 | now | now | now |
| **Complexity[1]** | very low | low | low | low | medium | low | low | medium | low | high | high | high | high | high |
| **Insertion** | partly | present | present | present | | present | present | present | present | present | present | | present | present |
| **Embedded** | | | | | | | present | | present | present | present | | present | present |
| **Reconstruction** | | | | | present | | | | | | | present | | partly |
| **Dynamic** | | | present | | | | present | | | present | present | | present | present |
| **Interactive** | | | | | present | | present | | | | | present | | partly |
| **Auto-Placement** | | | present | present | present | present | present | present | present | present | present | | present | present |
| **Real-time** | | | | | | | | | partly | present | present | partly | partly | partly |
| **User Study** | | partly | | | | | present | present | | present | present | partly | partly | partly |
| **Active/Used** | present | | | | | | | | | present | present | present | present | present |
| **Automation[1]** | very low | very low | very low | very low | very low | very low | very low | medium | low | high | high | low | low | low |

Legend:
[1] very low · low · medium · high
[2] not present · partly · present
[3] Augmented Video Player
[4] real-time capability in current, yet to be published versions
* Commercial applications

large, it has not primarily been applied to soccer, due to the unique challenges with tiny, moving targets. The exceptions are the commercial systems iView [4] and True View [14], which follow this reconstructive approach. True View is the more advanced.

**Dynamic Content.** Depending on the type of content that can be inserted, the techniques can be classified to only support static, fixed size content like text and images, or they support dynamic content, which can be animated or morphed. From the presented approaches, only Wan et al. [31], Stein et al. [26, 27], BBC R&D [5], Piero [22], and Viz Libero [30] support this.

**Interactive.** Most approaches output a generated augmented video file, which can not be interacted with. We consider a system to be interactive when, during playback, different parameters to modify the visualization or actually interact with it can be set by keyboard or mouse/touch. Of the presented techniques, only the reconstructive ones (iView [4], True View [14]), and Augmented Video Player [5] are truly interactive. However, Viz Libero [30] can be considered to be mostly interactive and, to a lesser degree, this is also true for Piero [22]. However, it should be said, that the boundaries between static and interactive are not strict and overlapping can exist. Also, some of the techniques can be adapted to provide interactivity.

**Auto-Placement.** The aspect of auto-placement is somewhat related to embedding, as it refers to the ability to place the content intelligently at an appropriate position. While this is a requirement for embedding, for overlaying techniques this is not strictly necessary, albeit often desired. Several of the presented systems are capable of automatically placing the content at the desired position, also different levels of accuracy and difficulty exist. While this is one of the main aspects for reconstructive techniques (merge different images for a 3D reconstruction), the approaches by Wan et al. [31], Andrade et al. [1], and Liu et al. [18] follow a simple positioning and only the approaches by Schlipsing et al. [24] and

Stein et al. [26, 27] are more advanced, with the latter doing a better job at the embedding.

**Real-time.** Of the presented approaches, only a few support real-time analysis and augmentation. For the purpose of this study, we consider any approach to be real time if the delay is below 50ms, a typical order up from when humans notice a lag. The reconstructive approaches can be considered near-real time with a delay in the order of seconds. This is achieved by not doing full 3D reconstruction but some forms of approximation yielding good enough results. For Piero [22], the response time depends on the required human operations. The time needed to generate visualizations is often between several seconds to some minutes, with no upper limit in principle. Viz Libero [30] instead takes several minutes to generate non-trivial visualizations. Only the systems by Stein et al. [26, 27] support real time in their currently existing version.

**Active/Used.** Promising or successful systems are typically developed further or actively used. For the presented systems, to the best of our knowledge only the simple Table Overlay, the commercial systems True View [14], Piero [22], and Viz Libero [30] are used in the field. Further, the approaches by Stein et al. [26, 27] are actively developed and improved.

**Study/Evaluation.** Table Overlays have been in use for decades in the industry. These representations are extremely common and provide the basis for other, more detailed visualizations. Studies discussing the other approaches only exists rudimentary for Wan et al. [31], and as evaluations for Liu et al. [18], Schlipsing et al. [24], as well as both paper by Stein et al. [26, 27]. For the industry applications True View [14], Piero [22], and Viz Libero [30] the usefulness is proven in form of continued usage and viewer feedback.

Concluding the comparison, we consider Piero [22] and Viz Libero [30] the most advanced approach employed in the field in terms of visualization, while the research by Stein et al. [26, 27] represents the state-of-the-art in terms of automation of the whole process, generating the visualizations without user input and augmenting the video with relevant and useful information.

## 5 DISCUSSION AND CONCLUSION

Visualization aims to provide an abstract representation to enhance and speed-up the understanding of complex data. This is especially true for video-based soccer visualization, where the collective behavior of two groups of persons, trying to follow opposing tactics, leads to complex relational and hypervariate data, which is hard to visualize in context. An efficient and useful representation often is not known a priori, and more often than not depends largely on the specific aspects under consideration. This survey paper provided a comprehensive overview of the latest methods developed for the video-based analysis of soccer matches. The field is underdeveloped and few significant publications exist. The existing literature has been surveyed and relevant approaches have been identified. Apart from traditional ones like data tables, charts, and 2D maps, these include more advanced visualizations of formations, free spaces, interaction spaces, distance measurements, ball trajectories, selecting

visible actors, player trajectories, and density distributions. The papers under consideration have been compared according to eleven criteria in the five areas **complexity**, **visualization type**, **capabilities**, **assessment**, and **automation level**. For all approaches, we have cited accompanying studies that show how each approach is expected to significantly improve and simplify soccer analysis, with some example scenarios given. We identify the approaches by Stein et al [26, 27], the systems Piero [22] and Viz Libero [30] as well as, in the area of 3D reconstruction for soccer, True View [14] to be the most advanced methods currently available. In some sense, the main difference between the Piero system as well as the Viz Libero system compared to the approach by Stein et al. [27] is that the former more or less represent the state-of-the-art in manually annotated video-based soccer analysis with advanced graphical display, while the latter is a fully automated approach. Combining those two would likely lead to significant benefits. Almost all other academic approaches are in their infancy. This is surprising, given the technological advances in recent years and the commercial and political benefits which could be leveraged.

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

## REFERENCES

[1] E. L. Andrade, J. C. Woods, E. Khan, and M. Ghanbari. 2005. Region-based analysis and retrieval for tracking of semantic objects and provision of augmented information in interactive sport scenes. *Trans. Multi.* 7, 6 (2005), 1084–1096. https://doi.org/10.1109/TMM.2005.854417

[2] Jürgen Assfalg, Marco Bertini, Carlo Colombo, Alberto Del Bimbo, and Walter Nunziati. 2003. Semantic annotation of soccer videos: automatic highlights

identification. *Computer Vision and Image Understanding* 92, 2-3 (2003), 285–305. https://doi.org/10.1016/j.cviu.2003.06.004

[3] Rahul C. Basole and Dietmar Saupe. 2016. Sports Data Visualization [Guest editors' introduction]. *IEEE Computer Graphics and Applications* 36, 5 (2016), 24–26. https://doi.org/10.1109/MCG.2016.85

[4] BBC R&D. 2005. iview: free-viewpoint video. https://www.bbc.co.uk/rd/projects/iview

[5] BBC R&D. 2012. Augmented Video Player. https://www.bbc.co.uk/rd/projects/augmented-video-player

[6] BBC R&D and Ericsson. 2004. Piero sports graphics system. https://www.bbc.co.uk/rd/projects/piero

[7] Carl Bialik. 10.06.2014. The People Tracking Every Touch, Pass And Tackle in the World Cup. *FiveThirtyEight* (10.06.2014). https://fivethirtyeight.com/features/the-people-tracking-every-touch-pass-and-tackle-in-the-world-cup/

[8] R. Borgo, M. Chen, B. Daubney, E. Grundy, G. Heidemann, B. Höferlin, M. Höferlin, H. Leitte, D. Weiskopf, and X. Xie. 2012. State of the Art Report on Video-Based Graphics and Video Visualization. *Computer Graphics Forum* 31, 8 (2012), 2450–2477. https://doi.org/10.1111/j.1467-8659.2012.03158.x

[9] Dartfish. 2019. MyDartfish. https://www.dartfish.com/football

[10] M.C.F. de Oliveira and H. Levkowitz. 2003. From visual data exploration to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphics* 9, 3 (2003), 378–394. https://doi.org/10.1109/TVCG.2003.1207445

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

[12] Adrian Hilton, Jean-Yves Guillemaut, Joe Kilner, Oliver Grau, and Graham Thomas. 2011. 3D-TV Production From Conventional Cameras for Sports Broadcast. *IEEE Transactions on Broadcasting* 57, 2 (2011), 462–476. https://doi.org/10.1109/TBC.2011.2131870

[13] Hudl. 2019. Sportscode. https://www.hudl.com/products/sportscode

[14] Intel. 2018. True View. https://www.intel.de/content/www/de/de/sports/technology/true-view.html

[15] Daniel A. Keim. 2001. Visual exploration of large data sets. *Communications of the ACM* 44, 8 (2001), 38–44. https://doi.org/10.1145/381641.381656

[16] D. A. Keim. 2002. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics* 8, 1 (2002), 1–8. https://doi.org/10.1109/2945.981847

[17] Matthias Kraus, Niklas Weiler, Thorsten Breitkreutz, Daniel A. Keim, and Manuel Stein. 2019. Breaking the Curse of Visual Data Exploration: Improving Analyses by Building Bridges between Data World and Real World. In *10th International Conference on Information Visualization Theory and Applications*.

[18] Huiying Liu, Shuqiang Jiang, Qingming Huang, and Changsheng Xu. 2008. A generic virtual content insertion system based on visual attention analysis. In *Proceedings of the 16th ACM international conference on Multimedia*, Abdulmotaleb EL Saddik, Son Vuong, Carsten Griwodz, Alberto Del Bimbo, K. Selcuk Candan, and Alejandro Jaimes (Eds.). ACM, New York, NY, 379. https://doi.org/10.1145/1459359.1459410

[19] Vinicius Machado, Roger Leite, Felipe Moura, Sergio Cunha, Filip Sadlo, and João L.D. Comba. 2017. Visual soccer match analysis using spatiotemporal

positions of players. *Computers & Graphics* 68 (2017), 84–95. https://doi.org/10.1016/j.cag.2017.08.006

[20] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale. 2018. State of the Art of Sports Data Visualization. *Computer Graphics Forum* 37, 3 (2018), 663–686. https://doi.org/10.1111/cgf.13447

[21] Simple Thought Productions. 2015. Production Photos. https://www.simplethoughtproductions.com

[22] Redbeemedia. 2018. Piero Sports Graphics. https://www.redbeemedia.com/portfolio/piero-sports-graphics/

[23] Amjad Rehman and Tanzila Saba. 2014. Features extraction for soccer video semantic analysis: current achievements and remaining issues. *Artificial Intelligence Review* 41, 3 (2014), 451–461. https://doi.org/10.1007/s10462-012-9319-1

[24] Marc Schlipsing, Jan Salmen, Marc Tschentscher, and Christian Igel. 2014. Adaptive pattern recognition in real-time video-based soccer analysis. *Journal of Real-Time Image Processing* 13, 2 (2014), 345–361. https://doi.org/10.1007/s11554-014-0406-1

[25] Huang-Chia Shih. 2018. A Survey of Content-Aware Video Analysis for Sports. *IEEE Transactions on Circuits and Systems for Video Technology* 28, 5 (2018), 1212–1231. https://doi.org/10.1109/TCSVT.2017.2655624

[26] Manuel Stein, Thorsten Breitkreutz, Johannes Haussler, Daniel Seebacher, Christoph Niederberger, Tobias Schreck, Michael Grossniklaus, Daniel Keim, and Halldor Janetzko. 17.10.2018 - 19.10.2018. Revealing the Invisible: Visual Analytics and Explanatory Storytelling for Advanced Team Sport Analysis. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*. IEEE, 1–9. https://doi.org/10.1109/BDVA.2018.8534022

[27] Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Thorsten Breitkreutz, Philipp Zimmermann, Bastian Goldlucke, Tobias Schreck, Gennady Andrienko, Michael Grossniklaus, and Daniel A. Keim. 2018. Bring It to the Pitch: Combining Video and Movement Data to Enhance Team Sport Analysis. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 13–22. https://doi.org/10.1109/TVCG.2017.2745181

[28] Graham Thomas. 2007. Real-time camera tracking using sports pitch markings. *Journal of Real-Time Image Processing* 2, 2-3 (2007), 117–132. https://doi.org/10.1007/s11554-007-0041-1

[29] Graham Thomas, Rikke Gade, Thomas B. Moeslund, Peter Carr, and Adrian Hilton. 2017. Computer vision for sports: Current applications and research topics. *Computer Vision and Image Understanding* 159 (2017), 3–18. https://doi.org/10.1016/j.cviu.2017.04.011

[30] Viz Libero. 2018. Viz Libero. https://www.vizrt.com/products/viz-libero

[31] Kongwah Wan, Jinjun Wang, Changsheng Xu, and Qi Tian. 2004. Automatic Sports Highlights Extraction with Content Augmentation. In *Advances in multimedia information processing - PCM 2004*, Kiyoharu Aizawa, Yūichi Nakamura, and Shin'ichi Satoh (Eds.). Lecture notes in computer science, 0302-9743, Vol. 3332. Springer, Berlin and Great Britain, 19–26. https://doi.org/10.1007/978-3-540-30542-2{_}3

[32] Yuanyi Xue, Yilin Song, Chenge Li, An-Ti Chiang, and Xiaoran Ning. 2017. Automatic Video Annotation System for Archival Sports Video. In *2017 IEEE Winter Conference on Applications of Computer Vision workshops*, LSSB (Ed.). IEEE, Piscataway, NJ, 23–28. https://doi.org/10.1109/WACVW.2017.11