



City Research Online

City, University of London Institutional Repository

Citation: Stein, M., Janetzko, H., Lamprecht, A., Breitzkreutz, T., Zimmermann, P., Goldlucke, B., Schreck, T., Andrienko, G., Grossniklaus, M. & Keim, D. (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), pp. 13-22. doi: 10.1109/tvcg.2017.2745181

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/18380/>

Link to published version: <https://doi.org/10.1109/tvcg.2017.2745181>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Bring it to the Pitch: Combining Video and Movement Data to Enhance Team Sport Analysis

Manuel Stein, Halldor Janetzko, Andreas Lamprecht, Thorsten Breitzkreutz, Philipp Zimmermann, Bastian Goldlücke, Tobias Schreck, *Member, IEEE*, Gennady Andrienko, Michael Grossniklaus, *Member, IEEE*, and Daniel A. Keim, *Member, IEEE*

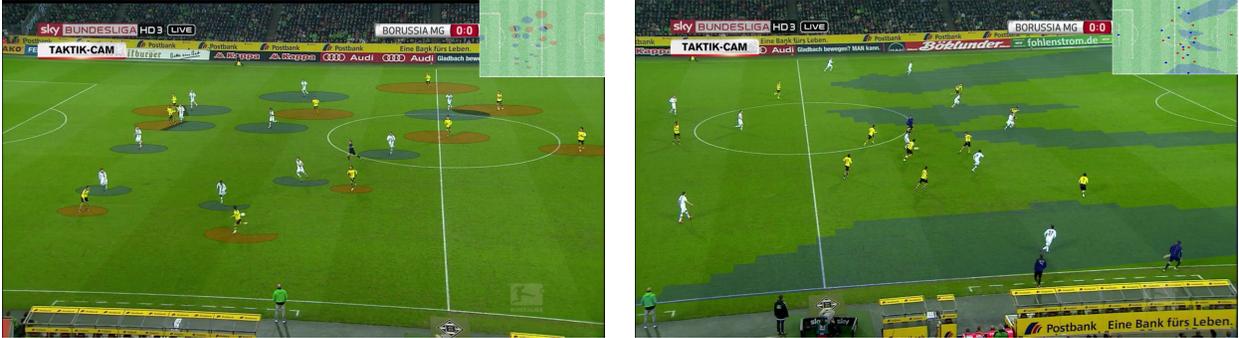


Fig. 1: Our system integrates analytical visualizations of soccer player movement data extracted from video into the same video, relying on appropriate computer vision techniques. Movement visualization techniques like interaction spaces (left) and free spaces (right) allow to analyze and explore soccer movement data in context of the original video. As our studies conducted with expert soccer analysts show, the combination of video and abstract visualization supports effective contextualized analysis and can foster user trust regarding data and analytical visualization.

Abstract—Analysts in professional team sport regularly perform analysis to gain strategic and tactical insights into player and team behavior. Goals of team sport analysis regularly include identification of weaknesses of opposing teams, or assessing performance and improvement potential of a coached team. Current analysis workflows are typically based on the analysis of team videos. Also, analysts can rely on techniques from Information Visualization, to depict e.g., player or ball trajectories. However, video analysis is typically a time-consuming process, where the analyst needs to memorize and annotate scenes. In contrast, visualization typically relies on an abstract data model, often using abstract visual mappings, and is not directly linked to the observed movement context anymore. We propose a visual analytics system that tightly integrates team sport video recordings with abstract visualization of underlying trajectory data. We apply appropriate computer vision techniques to extract trajectory data from video input. Furthermore, we apply advanced trajectory and movement analysis techniques to derive relevant team sport analytic measures for region, event and player analysis in the case of soccer analysis. Our system seamlessly integrates video and visualization modalities, enabling analysts to draw on the advantages of both analysis forms. Several expert studies conducted with team sport analysts indicate the effectiveness of our integrated approach.

Index Terms—visual analytics, sport analytics, immersive analytics

1 INTRODUCTION

Over the last years, the interest in analysis of team sport data has grown rapidly in research and practice. Among many team sport disciplines, soccer is one of the most widely practiced and also, commercially relevant kinds of team sports. The analysis of soccer match data has received much attention in the recent years by fields including sports and

behavioral science, biomedicine, and others. Computer science-based solutions help to capture soccer data, process and analyze it. Recently, several works introduced visual analysis techniques for soccer data. Soccer is a match of two opposing teams that engage in competitive and cooperative movement patterns among and between teams in space and time. Among the key goals of soccer analysis is to support understanding of patterns in this movement space on the player and team level. Also, the analysis of derived performance measures for players and teams, including predictive analysis, is of high relevance in this domain. In practice, two analysis modalities often occur. First, analysis can be done based on captured movement data, applying techniques from movement analysis. These often are used in combination with abstract movement visualization techniques. Second, in interactive video analysis, video recordings of matches are interactively inspected and annotated by experts, yielding reports and video presentations to coach players and inform coaches. Video analysts are typically used to work with video recordings and not with abstract data representations.

Several previous works investigated the usage of visually-supported methods in the soccer analysis domain. For example, the SoccerStories [37] system presented several custom-designed information visualization techniques to present and comprehend soccer match data

- Manuel Stein, Andreas Lamprecht, Thorsten Breitzkreutz, Philipp Zimmermann, Bastian Goldlücke, Michael Grossniklaus and Daniel A. Keim are with University of Konstanz. E-mail: forename.lastname@uni-konstanz.de.
- Halldor Janetzko is with University of Zürich. E-mail: Halldor.Janetzko@geo.uzh.ch.
- Tobias Schreck is with Graz University of Technology. E-mail: Tobias.Schreck@cg.v.tugraz.at.
- Gennady Andrienko is with Fraunhofer IAIS, Germany and City University London, UK. E-Mail: Gennady.Andrienko@iais.fraunhofer.de.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxx

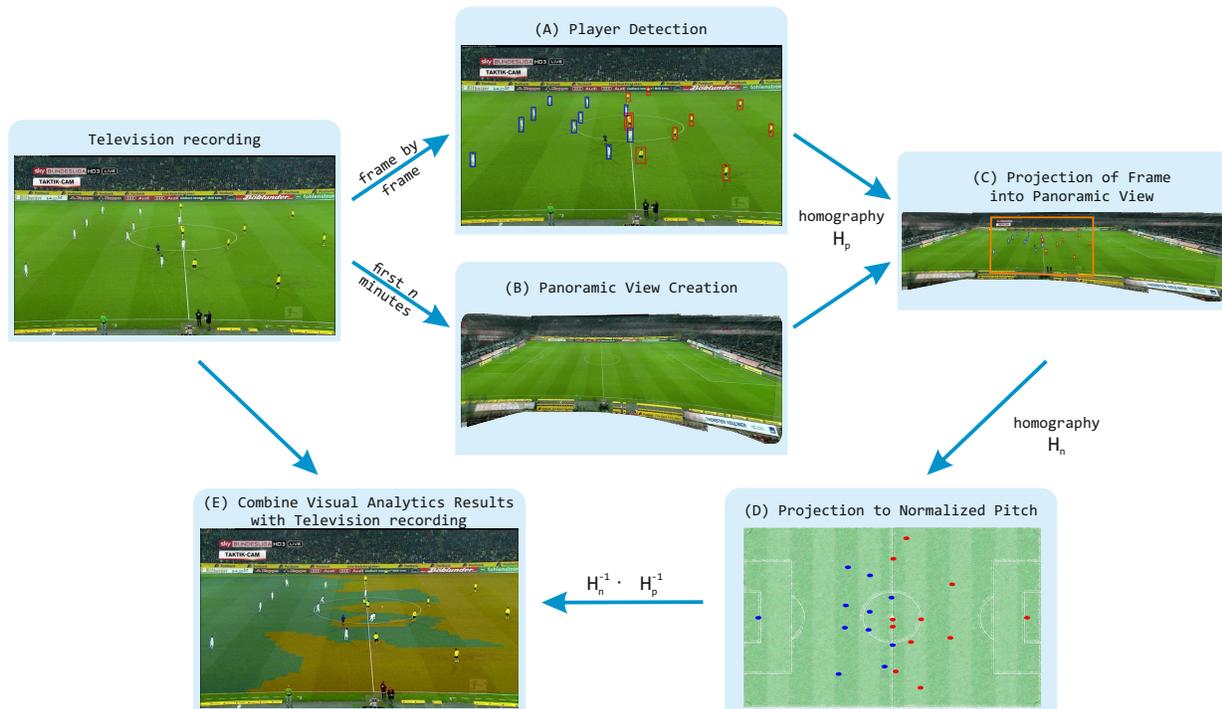


Fig. 2: Workflow of our system. A simple (TV/video) recording serves as input. We detect players on the soccer pitch (A) and, simultaneously, generate a panoramic view of the pitch (B). Afterwards, each frame gets projected into the panoramic view to extract correct player positions from a moving camera (C). The extracted player positions then can get projected into a normalized pitch (D). Based on this projection, we realize the integration of visualizations on the pitch for the video (E).

using abstract visual representations. Other works make use of trajectory visualization, hetamap-based techniques, of movement feature visualization (see Section 2 for details). In our own previous work, we have defined abstract visual representations for soccer match data and evaluated these with soccer analysts by means of interviews and case studies. While per se useful in many cases, we also observed that interpreting results obtained from abstract visual representations in context of the actual game situation required substantial mental effort by analysts. Often, experts wanted to verify observations made in the abstract visualization space in the video space, or even required the video space to interpret patterns in the visualization space. In effect, analysts we worked with wanted to bring analysis results obtained in abstract visualization space back to the pitch, as to paraphrase an expression often used by coaches during soccer practice. We hence identified the need to integrate soccer data visualization with soccer video for enhanced analysis and context provision.

We contribute an automatic method that provides effective visual analysis in the domain of team sport analysis by integrating appropriate analytical visualizations within the video context. We address the system aspect, demonstrating the applicability of a computer vision framework enabling Visual Analytics in movement context. We identify and implement an appropriate computer vision technique to extract player positions and detect the relative view-port from standard TV broadcasts, which enables capturing movement data. Furthermore, this techniques enables to map two-dimensional analytical visualizations back to the video recordings in a perspectively correct way. We additionally address a methodological aspect, raising the question of how and which analytics methods from the “2D space” can be embedded in video recordings. We give a first discussion of effects of the combination of abstract visualization and video, providing empirical evidence by domain experts to this end.

2 RELATED WORK

We first discuss related work in video analysis and data extraction (Section 2.1) followed by an overview to video analytics (Section 2.2).

We outline research in the area of sports in general and related to soccer (Section 2.3) and position our approach with the aforementioned works in Section 2.4.

2.1 Object detection and tracking

Higher-level content analysis, e.g., for player movement, involves detection and tracking of players in the input video streams as well as mapping of the positions into a reference frame. Tracking has been studied for decades as one of the fundamental tasks in computer vision, see e.g. [49] for a more recent overview. We therefore restrict the following discussion to the methods used in previous work on team sports analysis.

In *position mapping*, positions from the video are mapped to a physical reference frame, in our case the soccer pitch. A key approach to this mapping is to define and identify reference points in the video that correspond to known locations on the underlying soccer pitch. For this purpose, Choi et al. [45] define a pitch model, incorporating known positions like the center circle and center line. To find the corresponding points within a video, lines and ellipses are identified in a given image by Hough-Transformation [13]. Of these lines and ellipses, four positions are extracted and associated with the positions in the pitch model to estimate a coordinate transformation of the given image into the pitch model. Since it is a plane-to-plane transformation, it can be described by a projective linear map (homography) [16]. A drawback of this method is, that there always have to be at least four positions associated with the soccer pitch and, most importantly, these positions have to be correct. In contrast to comparing images individually, rebuilding a larger image (e.g. panoramic views) from smaller subimages is a common technique in image vision called *stitching*. For example, Kim et al. [24] look for identical lines in subsequent images of the video and map related lines onto each other by a coordinate transformation. This way, a complete view of the soccer pitch is obtained image by image. Brown et al. [7] introduce an automated process for panorama stitching. The described approach involves the calculation of SIFT-features [31] within the considered images. A transformation is computed by finding

matching features in image pairs. We roughly follow their approach to generate a panoramic view of the soccer pitch seen in the video recording.

The next step after position mapping is to *detect and track objects of interest*. In case of team sport video analysis, these are typically the teams' players and the ball. In player detection, advances have been made in the last decades. Liu et al. [29] recognize players with a so-called *boosted cascade of Haar features*. This detector is constructed with carefully chosen training data (players in different poses as positive examples, background and pitch parts as negative examples). For the detection of players on a soccer pitch, Hoernig et al. [17] compute a pitch mask. For that matter, they assume that the soccer pitch is a rectangle. Furthermore, the dominant color of the pitch is defined as green. The potential positions of the players are obtained by subtracting the pitch mask from the processed image. Additionally, the found contours have to meet different predefined features (for example height relations). Perez et al. [36] follow a probabilistic approach. First, a color histogram model is created, which describes the color distribution of a given player. The detection and tracking of the player is subsequently realized with a particle filter [15]. Dearden et al. [12] also introduce a combined approach. In the first step, a mask of the soccer pitch is computed. For this mask, the color histogram is defined and by subtraction the possible player areas obtained. The player areas are subsequently processed with a particle filter. Choi et al. [45] use a combination of particle filter and matching of player templates. These templates are obtained by analyzing the defined player mask. The tracking of players with several stationary cameras is examined by Kang et al. [22]. The modeling of players is achieved by a combination of color distribution and movement models. The tracking itself is performed by maximizing a multivariate distribution probability.

2.2 Visualization for Video Analytics

Video analysis attracted interest of visualization researches in early 2000s when millions of CCTV cameras have been installed worldwide. The human operator is not able to watch the produced footage neither in real time nor afterwards. Respectively, it was necessary to develop methods for extracting features from video, track moving objects, identify abnormal situations, and present outputs in easy-to-comprehend summarized form. Paper [11] proposed to visualize the dynamics of the detected features in so-called volume scene graphs. More recently, a series of papers from D. Weiskopf's group systematically considered challenges of video analysis, proposed a model for video visual analytics, and implemented it in a system for analysis of VAST Challenge 2009 data set [19]. Later work [18] especially addressed reconstruction of moving objects trajectories from video, clustering trajectories by similar paths and velocities, and representation of features dynamics for the clusters in a compact form. By projecting the extracted features to parallel time lines, this work provided a convenient tool for detection of unusual situations and browsing/querying the lengthy content of video.

Paper [27] describes application of interactive video analysis to sport. After extracting trajectories of players from a video of a rugby game, the proposed system supports queries based on sketches. Based on a user's sketch of a desired trajectory, the system extracts candidate fragments and characterizes the extracted trajectories in respect to multiple trajectory attributes such as curvature, tortuosity etc. The extracted features of the candidate episodes are presented on a parallel coordinates plot. Multiple representations engage users in the process of exploring the video and corresponding trajectory patterns.

2.3 Sports Analysis

The application of visual analysis methods enabling an in-depth data analysis was already introduced in many sports. Polk et al. [41] describe an analysis approach for tennis, based on meta-data as result, length of rallies, and service. Their visual method includes a comprehensive overview of the overall game and highlights strength and weaknesses of players based on pure statistics. The shot behavior of different teams in ice hockey is examined by Pileggi et al. [39]. They focus on a visual representation, realized by a circular heatmap around the goal. The shade of the color describes for different distances the possibility of

execution. Especially in team sports, the spatial area where the match takes place, is an important concept. For this reason, there exist several models to measure the utilization of the area, for example with voronoi cells [30].

In soccer, several visual data analysis methods have been proposed. Legg et al. [28] try to capture and visualize the course of a match in an abstract and summarizing manner. They employ a timeline, depicting important events like shot on goal, as well as additional information and indicators for the outcome of different events. A complete system for the analysis of soccer matches is described by Perin et al. [37]. Their system provides visualizations for many aspects of the match, like a generalized view of the course of the match or a clustered view of performed passes. Different phases of a match can also be compared in small-multiple view. Chung et al. [9] describe an approach for the multivariate sorting of event data. The sorting is aimed for the efficient classification of match recordings for subsequent analysis and evaluation of the team performance. An approach for the interactive determination and analysis of interesting events during a soccer match is proposed by Janetzko et al. [44] and extended with an application-specific visualization by Stein et al. [46]. Horton et al. [20] classify performed passes during a soccer match. Therefore, the quality of passes by means of distance, shot strength, and pressuring of the receiving player is evaluated. A general visual analysis approach for pressure in soccer matches has been introduced by Andrienko et al. [5]. Perl et al. [38] propose the analysis of tactical performance, based on a combination of pattern recognition with neural networks and analysis of ball possession.

Sport media broadcasters are integrating additional graphical cues in video and on the pitch for quite some time supported by companies such as Vizrt [3] or SAP [1]. Vizrt as one of the market leaders in this field, offers a product (*Vizrt Libero*) for virtual sports enhancements. Journalists and television stations use *Vizrt Libero*, for example, during television broadcasts to explain interesting match situations. They enable manually drawings of simple visual elements onto the pitch such as arrows or colored rectangles. Furthermore, *Vizrt Libero* allows the user to track specific players for small time frames. However, the analysis possibilities of the tracked players is limited. A tracked back four formation, for example, can be visually connected or a simple heat map can be calculated and displayed. Our system allows the user to automatically visualize complex advanced visualizations transferred from two-dimensional abstractions. Specifically, our system recognizes the positions of all visible players in a given video stream, and we use this information to compute various analytical measures which in turn can be visualized in the progressing video.

2.4 Positioning of our Work

We contribute a system for visual analysis of group movement in video data combining information visualization with video recordings. We rely on techniques for position mapping and object recognizing in video, and apply them to soccer recordings. We also rely on techniques for movement analysis enabling insights to the observed player movement patterns. Existing work in video visualization often focuses on summative views. In visual movement analysis [4], state-of-the-art techniques typically use visual abstractions from the real world. To the best of our knowledge, our system is the first to allow visual analysis of soccer matches relying on appropriately defined movement visualization techniques mapped with high precision into the match video. Hence, our approach provides analysis considering both visual abstraction of movement features and match video context, supporting effective use of both modalities in an integrated system.

In a sense, our approach also relates to the ideas of *Immersive and Augmented Analytics* which has recently again received interest by the visualization and analytics communities, e.g., as seen from the session on Immersive Analytics at IEEE InfoVis 2016 and the IEEE VR 2016 Workshop on Immersive Analytics. Our work adds to this direction specifically for team sports analysis, by bringing together abstract analytical visualization with media from the real world.



Fig. 3: The chosen points for the transformation of the panoramic (left) into the normalized (right) view must allow referencing distances in both directions of the main axes. Afterwards, a calculated homography H_n enables the transformation of panoramic view into normalized view and vice versa.

3 VIDEO MOVEMENT TRACKING AND VISUALIZATION INTEGRATION

The vital prerequisite for inspecting and analyzing team sport matches is movement data of players and the ball. A variety of companies (STATS, Opta, etc.) offer services acquiring movement, event, and statistical data based on their operating of capturing systems. However, many team sport clubs cannot take advantage from these services, for example, due to a lack of financial resources or the overly complex process of technical implementation. **Instead of professional systems, especially clubs from lower or youth leagues in soccer as well as clubs from other team sports have one static video camera recording entire matches. Representatives from these clubs have shown an increasing interest in approaches on how to analyze their performance based on this video data.** Consequently, we design our proposed system following two key requirements: the system should be capable of (a) extracting data from standard video recordings as well as (b) allow the user to overlay visualizations on the video material. Both should be easy to do for the non expert user, with only low-level technical requirements. The input to our system is the video stream of one camera, which zooms and rotates to capture as many players as possible.

This form of recording resembles common public television recordings of team sport matches, with one main camera positioned on the side of the pitch for a tactical overview. To illustrate our approach in this paper, accordingly, we use an example television match recording from the German Bundesliga being broadcasted on the Sky Sport TV channel operated by Sky UK Telecommunications [2]. All displayed figures are extracted from this match and enhanced by our proposed system. We align the frames of the video recordings with a normalized reference pitch to associate the positions in the video and on the pitch. In particular, our system is not limited to only process data from video to normalized pitch. Instead, we can also project visualizations rendered on the normalized pitch back to the original video, which allows for further analyses by directly watching the enhanced video stream. Currently, the ball detection is not done automatically but by manual annotation of events as passes, shots, etc. – similar to the ball acquisition by professional video tracking companies.

We implement our Visual Analytics system in *Java*, with a subset of time-critical computer vision methods in *nVidia CUDA* [35] running on GPUs. An overview the proposed system is given in the subsequent sections. See Figure 2 for an overview of the workflow.

3.1 Player Detection

In order to calculate exact player positions on the pitch (Figure 2 (A)) it is necessary to segment them from the background in each frame, i.e. detect all of the pixels where a player is visible. Unfortunately, we cannot define a frame-independent player size: first, the recording camera can zoom and thus focal length can be different in different frames, second, the camera is located at the side of the pitch, and thus players on the opposite side of the pitch appear smaller. Other challenging factors stem from body pose and proportions, e.g., players' shape differs based on their activities, or from adverse imaging conditions, e.g. caused by illumination changes, especially for matches with floodlight or intense rain phases. The detection should ideally also work in real-time.

Thus, we turned to low-level appearance models which require only minimal characteristics about the searched objects to be known. In

particular, we use methods based on edge detection and color based contour analysis, where we need to know dominant colors of the wanted object. During our design process, we applied a variety of image vision methods covering player classifiers [33], canny edge detection [8], contour-based player detection through pitch subtraction [17] as well as player detection through contour analysis by color histograms [36]. After evaluating the applied techniques, we found we achieve the best results by performing the player contour analysis through color histograms. A main reason is that this way, markings on the pitch do not influence player detection.

Our player detection method works in three steps. First, we create color histograms based of manually marked areas that are unique for players on the pitch. In our case, for example, player and team jerseys. Afterwards, we determine for each pixel in the image, whether the color of the pixel is contained in the color histogram with sufficient probability (i.e. a sufficient number of pixels in the marked area matches the color). To conclude exact player positions, we finally calculate the centroids of each detected area. Considering human anatomy, we use average human proportions to calculate the positions of feet, head and the corpus of the body. To obtain the extended bounding boxes enclosing the entire player, we enlarge the initially detected area by adding proportions dependent on empirically defined factors to all directions. Significantly smaller identified areas below a given threshold, for example player socks, are removed automatically. We illustrate our player detection in Figure 4.

3.2 Static Camera Generation

Previous systems typically use several static cameras, mounted at specific positions with certain angles and focal lengths, being calibrated for the sole purpose of making the player detection for the computer as efficient and effective as possible [34]. Instead, we propose to use one single, stationary camera, whose motion only consist of rotations and occasional zooms. It is well known [16] that under the additional assumption of pinhole projection, the video frames are then related to each other by a simple projective linear transformation called a homography. The advantage is that for the domain expert and final user, one single camera is set up easily without the need of longer preparations, special configurations, and calibrations.

However, in order to have comparable results as from several cameras and relate the different frames to each other, we automatically transform the images into a single, panoramic view of the pitch (Fig-

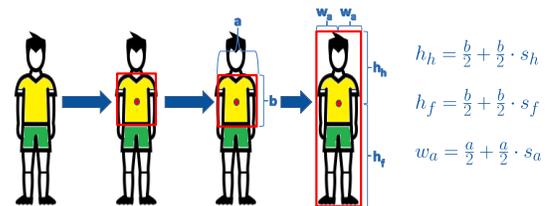


Fig. 4: Illustration of our player detection method using jersey color. Extending the centroid area of the jersey, we obtain a player enclosing rectangle. The factors s_h , s_f and s_a are empirically defined.



(a)



(b)

Fig. 5: Before homography (a) and after homography (b) calculation.

ure 2 (B)). Our proposed approach follows Lowe and Brown [7] who take a set of overlapping images as input. The set of input images is aligned automatically by iterating over pairs of consecutive frames, extracting and matching SIFT features [31] and estimating the homography between first and subsequent frame. A homography is a 3×3 transformation matrix acting on projective image coordinates, which can be estimated by a least-squares approach embedded in RANSAC to get rid of outliers [16].

From the aligned frames, we select a subset of frames which spans the whole range of motion of the video camera and has sufficient overlap between frames to generate the full panoramic view. Usually, the first 2 minutes of a match are sufficient to generate a panorama reflecting (almost) the complete pitch. From the homographies computed initially, we can compute the alignment of every individual frame within the panoramic view, see Figure 5 for an example transformation. However, for longer videos, we generate the panorama only from the beginning of the sequence, and align subsequent frames by estimating a homography directly using their and the panorama’s SIFT features.

In a final pass, we generate a clean background panorama and remove foreground objects like e.g. the players. For this, we compute for every pixel in the panoramic view the median color over all frames of the selected subset. As the players are constantly moving, this should be a good approximation to the actual background color. A Graph Cut method is then used to hide objects which are only partially visible [25]. In summary, we now have a panoramic view of the background as a common reference frame, as well as for each video frame, a homography H_p which describes the coordinate transformation into the panoramic view. See Figure 2 (C) for an illustration.

To obtain a normalized representation of the pitch, we map the panoramic view onto a user-supplied image - typically a standardized visualization of the pitch, see Figure 3. To define the transformation, the user selects at least four corresponding points in panoramic and normalized image manually. The four correspondences then determine a homography from the panoramic view to the reference image as before [16]. Note that due to the lack of radial distortion calibration, this is currently only a rough approximation, pending improvements in future work. Furthermore, it is a plane-to-plane homography, valid only for points on the main pitch plane. The transformation from panorama to normalized pitch is denoted H_n in the remainder of the paper. See Figure 2 (D) for an illustration.

3.3 Projection and Visualization Integration

From section 3.1, we have for each video frame the detections of the players. By sequential application of the homography H_p from frame to panorama, as well as the homography H_n from panorama to normalized frame, we can thus calculate player position **coordinates** on the normalized pitch, see Figure 2 (D). All positions not corresponding to a position within the normalized pitch are discarded as likely misdetections.

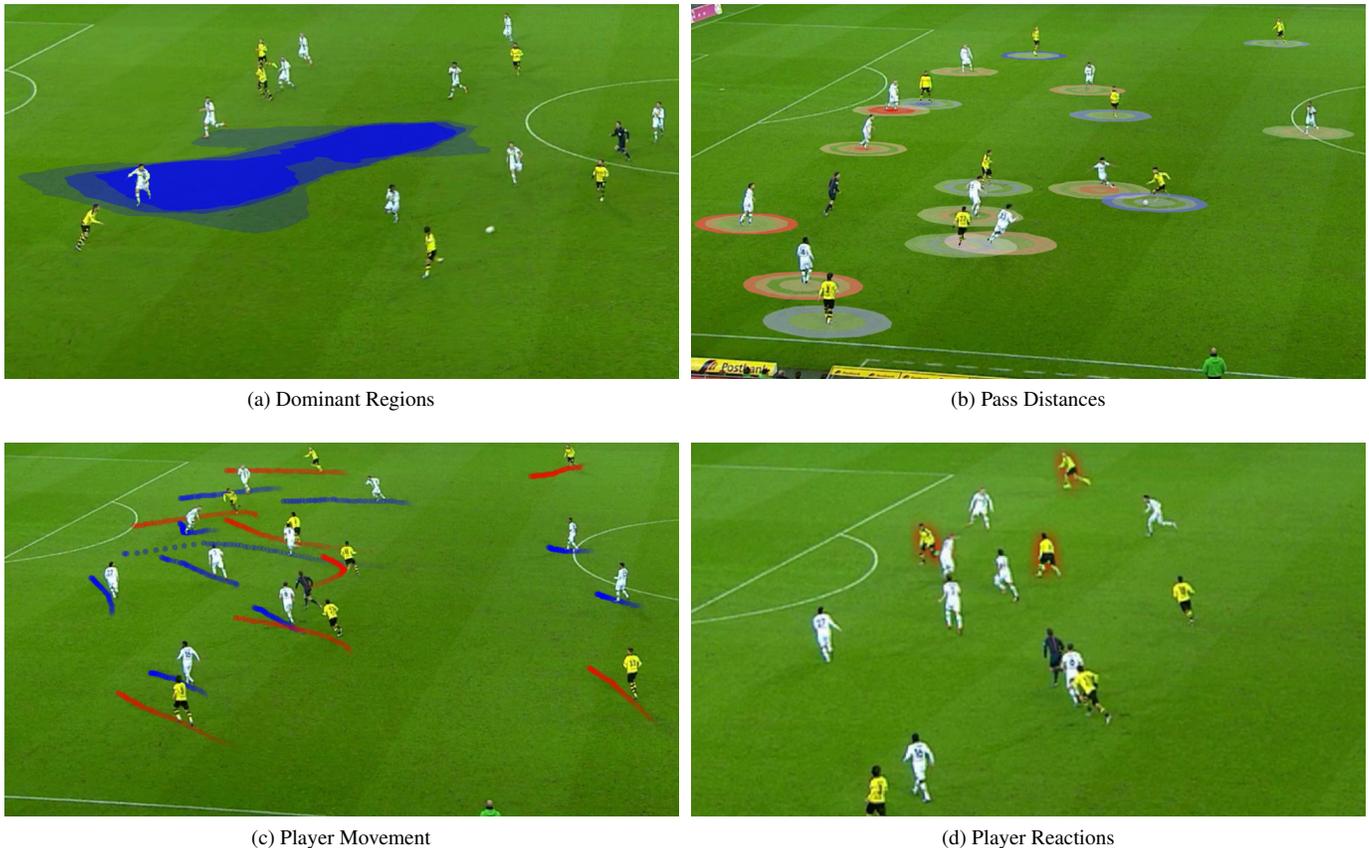
To obtain player tracks **and uniquely identify players**, the detected player positions are registered by a simple location-based method, which assigns each position to the closest player in the previous frame who could realistically reach the new position within the timespan between frames. A new player is initialized for all remaining positions that could not be assigned to any existing player. This happens, for example, if the input frame did not contain all players.

However, incorrectly detected player positions can lead to an initialization of a new player as well. These misdetections typically only happen over a span of few frames, so we automatically remove short tracks with only very few positions or which do not generate new positions over a longer period. Afterwards, automatic methods to further clean the data are applied, e.g., if players overlap each other we interpolate their positions by taking their current speed into account. Another problem is that the homography H_p , which maps a video frame into the panoramic view is not inevitably the optimal solution, as we compute H_p with RANSAC for a fixed number of iterations. This may result in situations where the transformed player positions are differing from the actual ones. To counteract this problem, we validate our data by applying a *moving average filter* [42]. In addition to the automatic systems, the user will also be given the opportunity to manually improve data gathering by interacting with the system in a variety of ways. We will evaluate the accuracy of our location detection method in Section 5.

Our system described above has the advantage that we can perform all data analysis on the normalized reference pitch, as shown in the next section. Due to the simple homographies with respect to the video frames, we can easily embed any information visualizations targeted to the reference pitch within the original camera images, see Figure 2 (E). To further emphasize the impression that resulting visualizations are drawn directly on the pitch and players are running on top of them, we remove all pixels that are related to the pitch and pitch markings from detected players regions, i.e. the red bounding boxes illustrated in Figure 4, by background subtraction. Afterwards, the cleaned player regions are drawn on top of the visualizations. In Section 4, we show diverse examples illustrating the many possibilities for analysis.

4 VISUAL ANALYSIS OF SOCCER VIDEO

In the previous section, we described our system transforming the video space to a normalized pitch space, tracking objects, and mapping visualizations back into the video space. Although our system proposes a combined view of video and visualization, it is possible to show only video recordings or abstract visualization techniques depending on the respective analyst’s needs. We identified in our expert discussions and previous studies three main analysis areas in the soccer domain. These three areas cover some core aspects of soccer analysis, namely, analyses based on *regions*, *events*, and *players*. For the demonstration of this three types of technique, we make use of previously proposed and established visualizations to showcase how existing analytical solutions can benefit from our system. Additionally, we propose new



(a) Dominant Regions

(b) Pass Distances

(c) Player Movement

(d) Player Reactions

Fig. 6: The main Visual Analytics techniques in our approach are based on players, events and regions. When analyzing a player ((c) & (d)), analysts want to detect movement patterns while event analysis (b) aims to, e.g. describe passing preferences. Region analysis (a) tries to segment the pitch into semantically meaningful units.

visualizations for each kind of technique. We describe the techniques in the following. Note that while we here focus on soccer, our framework can easily be extended to other invasive team ball games, like handball or basketball. Such adoptions would require a slight adaptation of the tracking techniques, implementation of appropriate analytics methods, and corresponding visualizations.

4.1 Region-based Analysis

Spatial analyses in soccer mostly deal with the notion of control of regions. For example, one analysis task might be to represent which team or player was dominating which part of the pitch under specific conditions. Two recently proposed techniques include so-called interaction spaces and free spaces [47]. Interaction spaces visualize the surrounding area that each player directly aims to control. They can be computed as the area that can be reached by a player, modeled as a function of velocity and heading [47]. Intersecting interaction spaces between players of opposing teams are identified as duel areas. We transform interaction spaces using our system from previously pure two-dimensional visualization back to the video recording. An example of interaction spaces using our proposed system is presented in Figure 1 (left).

Free spaces, in contrast to interaction spaces, represent the region that can be reached by a player before any opposing players. In our approach, these regions are calculated by segmenting the pitch into grid cells assigning each cell to the player with the highest probability of arriving there first with respect to distance, speed, and heading. Resulting free spaces are assessed by their respective size, the number of opposing players, and the distance to the opposing goal. Free spaces that are most promising (exceed a specific size, small number of opposing players, close to goal) are visualized by highlighting the respective

region on an abstracted soccer pitch. With the use of our proposed system, we can now visualize free spaces on top of the original video recording. An example for the resulting free space visualization is shown in the right of Figure 1.

In addition to the previously introduced visualizations, we also propose a new visualization to illustrate *dominant regions* of a team. These regions represent areas where one team is more in control than the opposing team. We calculate dominant regions by taking into account how many players of each team can reach a dynamic region of the pitch first. We segment the pitch into regular grid cells and calculate for each cell the order in which players can reach it. To ensure that we only identify dominant regions that are present over a longer period, we use a sliding window approach and calculate the moving average over the results. Using the *marching squares*-algorithm [32], we detect the contours of the resulting dominant areas that can be reached first by at least three players of the same team. The resulting dominant regions are visualized using a dark red to dark blue bipolar colormap (with transparent colors for tie areas) representing both teams. An example for the dominant region visualization is depicted in Figure 6 (a).

4.2 Event-based Analysis

Events in team sports can be defined as match-relevant actions which occur during playtime [48]. Most events are based on the ball, e.g., like shot on goal, cross or pass. Event data is used by analysts to identify statistical patterns during a match, e.g., which passing preferences a player might have or whether a player prefers playing short passes instead of long passes. We integrated established and new approaches for event-based analysis into our system.

Ball passing in a game situation is the result of a complex decision-making process influenced by pressure of the opposing team and the

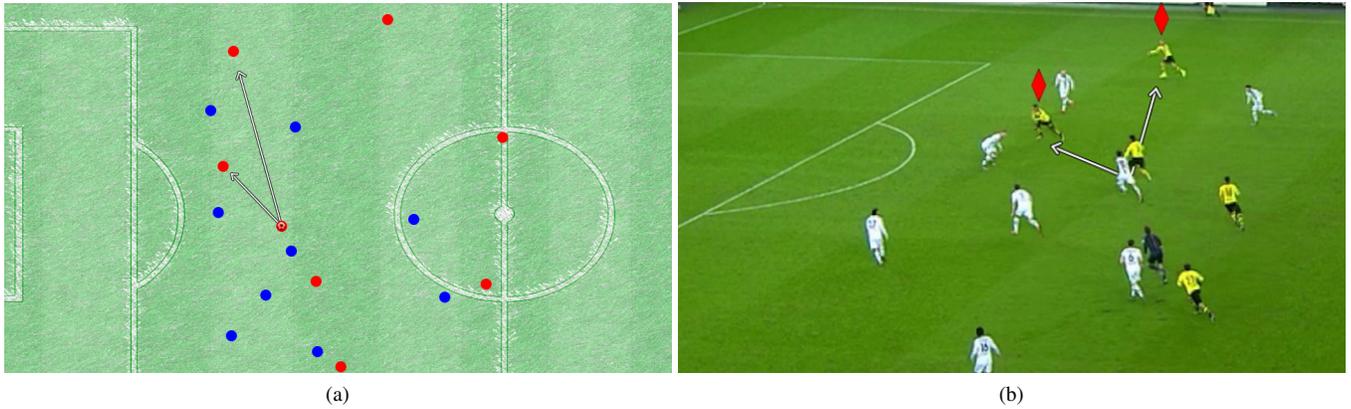


Fig. 7: Pass Options were visualized so far on an abstract pitch (a) [47]. Our proposed video visualization (b) allows us to add additional highlighting features such as a floating colored diamond-shape above the player heads in (b).

relative positions of each player. The passing behavior is usually analyzed to reveal typical tactics and to improve overall game play. Players need to take many factors into account when deciding for low-risk or more insecure passes to gain space. Recently, an analysis technique calculating pass alternatives for a player in a given game situation was introduced [47]. The overall goal was not to assess the quality of a pass but rather determine the risk involved for various passing alternatives. Resulting possible pass alternatives were visualized with arrows pointing towards potential receiving players on an abstract soccer pitch. Several criteria for the automatic assessment of passes were introduced reflecting the relative positions, distance, and direction. So far, arrows pointing towards potential receiving players on an abstract pitch are employed to visualize possible pass alternatives. While integrating the pass visualization into the video, we extended the original two-dimensional version [47] by additionally highlighting players, for example, with a floating colored diamond-shape over their heads. In Figure 7 (a), we show the previous two-dimensional visualization on an abstract rendered pitch in comparison to our enhanced version visualized on the pitch in Figure 7 (b).

Furthermore, analysts are also interested in the characteristics of player passing behavior. Understanding whether a player prefers to play short instead of long passes can provide tactical advantages for match preparation. To get an intuition about the passing preferences of a player, we developed a new visualization for player centered pass distances. Figure 6 (b) illustrates our approach. Circular rings drawn around each player visualize the aggregated normalized pass distance of the respective player. Color saturation is used to indicate whether a player usually prefers short or long distance passes. We normalize the aggregated pass number globally for all players and distance bins. Concentric rings close to the player represent short passes. This visualization allows identifying players with many passes and reveals their usual passing behavior.

4.3 Player-based Analysis

Besides regions and passes, sport analysts need to focus on individual players and their performance. Analysis typically either concentrates on statistical performance measures (e.g., distance covered by a specific players or failed passes) or actual player movement in context as key factor to reveal and predict tactical patterns. To support the assessment whether players move or not like expected in a given situation, we enable the user of our system to display the previous and next k seconds as live trajectory views in the video. We show the past trajectory different to the future trajectory by using transparency. An example can be seen in Figure 6 (c).

When analyzing team sport it is, however, important to focus on more than the movement of a single entity. The collective movement – here expressed by the two opposing teams – reveals more insights about strategies and tactics. Players moving similarly or reacting on

each other is important to observe as this might be an indicator of tactical patterns. Following the idea of Laube et al. [26] and as an exemplary first step towards this challenging problem, we implemented a method to visualize players reacting on each other. We define a reaction as imitated behavior, for example, if one player is strongly accelerating in one direction and other players start accelerating in the same direction shortly afterwards. Using our proposed system, we increase the salience of players acting coherently. For example, we are able to give players that are to be grouped together the same jersey color. Having the same jersey color, however, may result in losing distinct teams. Our design alternative is to emphasize corresponding groups of players with distinctly colored halos instead. An example can be seen in Figure 6 (d) where two players react on the ball possessing player and accelerate very fast to position themselves in a good spot for a potential shot on goal.

5 DISCUSSION

We quantitatively assessed the quality (precision) of the tracked data using our system in comparison with data from commercial systems. We provide details on the performance of homography and panoramic view calculations as introduced in Section 3. Afterwards, we discuss our proposed combination of video and abstract visualization supported by insights from several expert studies. We additionally report expert feedback on the created example visualizations from Section 4.

5.1 Tracking Assessment

As our approach imposes fewer restrictions and no calibration needs to the data acquisition process than professional capturing systems, our available data quality might be lower when compared to well-calibrated, professional systems. Consequently, we assessed the quality of our extracted data and compared our generated data with data acquired by a professional data provider with a calibrated camera system. To this end, we analyzed a match from the UEFA Champions League between FC Bayern München and Manchester City FC where we have both professionally generated position data as well as data extracted with our method from a high definition (720p) recording. To be able to compare the resulting positioning data, we synchronized the video recording (25 frames per second) and the professionally generated position data (10 positions per second). The professional system uses a set of static cameras around the pitch and proprietary software to extract the player movement. As the calibration would need manpower, these camera systems are usually permanently installed in stadiums of prime teams. We observed a maximal distance between our retrieved position and the professional data of four meters. Nevertheless, the average deviation was less than two meters with a standard deviation of roughly half a meter. We consider this as a positive result as the accuracy of professional data is not publicly known. Furthermore, the definition of a point-based location for a two-legged running person

is not clear per se already causing positional differences. From a subjective perspective, our extracted locations are very close to the players in the video recording and reflect the movement very well and there are no obviously and distracting systematic errors.

One of our final goals is to analyze the acquired movement data directly in the half-time break or after the match. Consequently, we addressed the efficiency of our approach. Furthermore, it is important to compute the reverse homographies near real-time as any changes in the Visual Analytics part should be directly visible in the recordings. Therefore, the introduced homography and panoramic view calculations were implemented highly parallelized on the GPU. The parallel nature of the feature extraction and matching process is well suited for modern GPU architectures and allows an implementation which is several times faster than an equally powerful CPU based method. To investigate the resulting efficiency of our implemented methods, we calculated panoramic views and homographies for several matches on a test system (Intel Core i7-6700, 16GB RAM, NVIDIA GeForce GTX 1070). The efficiency of our automatic panoramic view generation depends on the video resolution, complexity of camera movement and the amount of moving objects in the scene. We found that we are able to generate a panoramic view in an average of 40 to 50 seconds. The efficiency of a homography calculation depends on the panorama size, input resolution and the amount of visual detail. We were able to calculate a homography in 50 to 60 milliseconds. Estimating that a match lasts approximately 95 minutes (overtime included), we need around 140 to 170 minutes for the homography calculation. Mapping the abstract visualization to the video recordings and using the inverse homography takes 4 to 10 milliseconds being real time for interactive visualization purposes.

5.2 Insights from Expert Studies

Eventually, we want to find out which image an expert or analyst is going to interpret better and trust more – a more abstract visualization or a visualization added on top of the 'real' world. Therefore, we invited two experts to test the usefulness of our system. One expert has been an active soccer player for 25 years and has been working as a coach for 11 years. Until recently, he was working for the German soccer first league club FC Bayern München as a certified coach in the youth sector. A certified coach needs to be experienced in the theory and practice of video analysis. Because of his experiences in video analysis, we consider his feedback as highly relevant for the evaluation of our system. The other expert has been an active soccer player and coach for more than 30 years. Both experts have in common that they need ways to analyze matches of their teams as well as possibilities to communicate findings and insights to the players of their team. Hence, we consider both experts as well-suited candidates to discuss our approach.

In several expert studies, we separately investigated the concept of our system as well as the example analyses with each expert. At the beginning of a study the expert was explained the system and the interaction possibilities. Afterwards, the various analysis methods and their intentions were introduced. In the following, each expert could openly navigate through the system and try out all available features and visualizations. Besides the interaction with the system, we performed open discussion rounds which we used to ask related questions. Furthermore, we encouraged the experts to express *ad hoc* comments (the thinking aloud method [6, 14]). Expert interactions with the system were recorded for subsequent analysis. During the studies, we took notes of comments and recorded our observations. We were, additionally, highly interested in determining whether a video integrated visualization increases the experts' trust. Eventually, we asked about suggestions for improvement. The results of our expert studies are as follows.

The experts unanimously appreciated the possibility to perform the analysis directly in the video. Both experts consider embedded analysis tools and visualizations in video recordings as the natural (also: "perfect") procedure for in-depth analysis tasks in soccer. One of the reasons the experts mentioned is that the analysis in the video with moving real players is perceived more "natural" as in a generated view. This makes it easier for analysts to contextualize their findings from

analytical visualizations. Consequently, the experts argued that their trust in the analysis strongly increases. To this end, they feel that our approach is already very advanced in terms of application in practice.

The analysis of free spaces in video recordings is seen very important by the experts. Free spaces help to effectively identify offensive options for the team that is in possession of the ball. Visualization and visibility of free spaces are seen likewise positive. The experts, however, found the high refresh rate of free spaces partially irritating. Their high update rate (each video frame) can result in situations where small changes in the resulting free spaces are causing a perceived 'flickering' in the video visualization. A possible solution beside the limitation to, e.g., ten updates per second could be to define core areas of free spaces and cut the partially rapid changing margins. Another straightforward option is to smooth the visual transitions over time. The transformation of interaction spaces to video recordings is seen positive, too. The experts consider interaction spaces in video recordings as additionally helpful to highlight player formations. Furthermore, our experts state that interaction spaces support the perception of players being far away from the camera.

The domain experts approve our approach for the calculation of dominant regions. They consider the nuances in our visualization of dominant regions representing the amount of players that have control over a specific area as a gain. Pass distances were seen positive as well. The experts had the hypothesis that such a visualization could in future be used for detecting player roles. Players might, for example, have been formally described as defenders but their playing style might be not known. Following this hypothesis, the experts were interested in future possibilities of classifying player behaviors and resulting improvements, for example, for the scouting process. Analyzing player behavior with our complimentary visualization of player movement trajectories in past and future was interesting for the invited experts when analyzing tackles or individual tactical behavior (e.g., offside trap). The experts mentioned that they would appreciate if players that are noticeable moving into different directions than other players get visually highlighted. Eventually, the experts approved our visualization of player reactions as beneficial information when analyzing group movement. According to our invited experts, players get trained to perform specific movement behavior, for example, to pull apart the defense lines of the opposing team and to open new free spaces. Our visualization, therefore, enables the experts to "make the invisible visible" (quoting one expert).

Additionally, we observed that the experts came up with new ideas inspecting integrated visualizations in combination with the real movement on the pitch. According to the experts, they became more aware of a visualization's limitations and possibilities for improvement in the future. As, for example, soccer players were not represented by moving dots on an abstract pitch anymore but with the real persons, the experts noticed that the body pose is currently not always reflected correctly in the calculation of interaction spaces. If a player is running forwards or backwards, the resulting interaction spaces are identical. This exemplary problem could not attract attention outside of the video visualization as no data about the body poses are collected. To this end, novel methods [21, 40] from the field of human pose estimation could be used to gather data about body poses and take these into account, e.g., while calculating interaction spaces. Another idea the experts had was to extend the visualizations with respect to a players viewing direction. The viewing direction could, for example, be used to determine whether single players are able to perceive particular free spaces. Free spaces that can not be detected could be removed from the calculation. We believe it is an important and interesting finding of our evaluation to see experts being inspired by the analysis transformed back to the observed real world behavior.

6 IMPLICATIONS

Our work addresses two main aspects: a system aspect integrating computer vision methods with interactive visualization for video movement analysis. The second aspect involves questions of usability and analytical effectiveness prompted by the integration of analytical visualizations with the underlying video. **We will outline and describe the**

implications for future work in this section.

6.1 System implications

As opposed to professional stadium-based video-based tracking systems, we are relying with non-calibrated single-sensor recordings of arbitrary focal lengths and lens systems. Consequently, our data acquisition misses movement players being not visible in the recording, e.g., goal keepers usually will be only visible if the ball is nearby. Another issue with single-camera tracking is that jumps of players lead to inaccuracies. We plan to improve the player detection by more sophisticated and novel methods tackling for example the radial distortion by camera lenses (visible in the right Figure 1). A first starting point in this direction is the technique introduced by Kim et al. [23] detecting straight lines over the video frames. Furthermore, we plan to extend the computer vision processing, such as to retrieve additional features like viewing direction, body postures, or even the facial expression like confidence or surprise, which may be valuable features for analysis and able to introduce more semantics to the analysis beyond spatiotemporal properties. Additionally, our next steps involve automatic ball detection and, consequently, automatically deriving events (e.g., passes, dribbling, etc.) based on ball movement.

Another future direction is the real-time analysis of soccer games with a live stream and live processing of the video recordings. Beside hardware requirements for real-time processing, the challenges are the proper presentations and suitable analysis methods for first ad-hoc explorative analyses.

6.2 Integration implications

Our approach offers new opportunities not yet available in professional tracking systems. First of all, the combination of Visual Analytics methods with the raw video material is a unique possibility to enhance the usability and acceptance for domain experts. The combination of video and visualization bridges the gap between computer and human domain experts being crucial for Visual Analytics. We therefore believe that the inverse transformation from the abstract data model to the real world enables a more insightful and effective analysis. Experts can identify and reveal uncertainties in the analysis process more easily by relating their knowledge to the visualized analysis results. **Furthermore, as we experienced in our expert studies that the experts come up with their own suggestions for visualizations, the combination of video and abstract visualization seem to be a promising communication medium when designing Visual Analytics systems.**

To the best of our knowledge, any two-dimensional visualization that has been designed for the use on an abstract pitch can be used and transferred with our system. However, some additional preparations might be beneficial for visualizations over time, for example from Sacha et al. [43], such as presenting aggregated trajectories on a pitch without players. We will test this in our future work.

By merging the raw recording with two-dimensional visualizations, a potential problem may be overplotting, contrast effects, or distraction caused by non-match visual information in the video. Furthermore, the perspective of video recording at hand might not match the most interesting area. In our used recordings the upper side of the pitch is always smaller as it is further away, leading to less image space for the visualization in this area. Hence, we offer both the undistorted (plain two-dimensional) visual result representation and the real-world movement enriched by distorted visualizations. Detailed experiments should be done to assess which analysis tasks are more effectively done in either the raw video, in an abstract analytical visualization, or in the video with integrated visualization. Eventually, from such studies, an adaptive approach could be designed that chooses the most appropriate view based on task, user and situation in the match video.

We point out that of course other forms of integration of video and visualization are possible. For example, side-by-side approaches make use of explicit or implicit links, e.g., based on arcs or color-coding. One may exemplarily investigate how approaches like VisLinks [10] could be adapted to our case. Finally, in this work we have not specifically addressed user operations regarding interaction with the display. Future work may include using, e.g., sketch-based interfaces to search directly

in the video recordings for events or scenes of interest. This could be an effective way to specify the context for a search, e.g., player formations or neighboring players directly from the video.

7 CONCLUSION

We presented a system for visual analysis of team sport data integrating video with analytical trajectory visualizations. Based on single camera video recordings of soccer matches, we derive and analyze movement data as well as transform the resulting analytical visualizations back to the original video. The integrated visualizations allow to include context information from the video in the analytical visualization and can bridge inherent uncertainties regarding pure visualization approaches. Our proposed workflow enables domain experts to investigate the data in their desired and needed way. Analysts can decide for video recordings with and without visualizations as well as more abstract visualizations on a two-dimensional pitch. The usefulness of our approach is shown by expert studies with real soccer domain experts intrigued by the potential analysis capabilities. Interesting future work besides technical challenges on the computer vision side include an in-depth analysis of advantages and disadvantages of video, visualization as well as integrated video and visualization cases. Eventually, systems may be designed that can adaptively choose the best representation depending on task and data.

ACKNOWLEDGMENTS

This work was supported by the ERC Starting Grant “Light Field Imaging and Analysis” and the SFB Transregio 161 “Quantitative Methods for Visual Computing”.

REFERENCES

- [1] Sap sports one. <https://www.sap.com/products/sports-one.html>. (Accessed on June 21, 2017). 3
- [2] Sky go - moenchengladbach vs dortmund. <http://www.skygo.sky.de/>. (Accessed on January 23, 2016). 4
- [3] Vizrt. <http://www.vizrt.com/>. (Accessed on June 21, 2017). 3
- [4] G. Andrienko, N. Andrienko, P. Bak, D. Keim, and S. Wrobel. *Visual Analytics of Movement*. Springer, 2013. doi: 10.1007/978-3-642-37583-5 3
- [5] G. Andrienko, N. Andrienko, G. Budziak, J. Dykes, G. Fuchs, T. von Landesberger, and H. Weber. Visual analysis of pressure in football. *Data Mining and Knowledge Discovery*, pp. 1–47, 2017. 3
- [6] T. Boren and J. Ramey. Thinking aloud: Reconciling theory and practice. *IEEE transactions on professional communication*, 43(3):261–278, 2000. 8
- [7] M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. *International journal of computer vision*, 74(1):59–73, 2007. 2, 5
- [8] J. Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986. 4
- [9] D. Chung, P. Legg, M. Parry, I. Griffiths, R. Brown, R. Laramee, and M. Chen. Visual analytics for multivariate sorting of sport event data. In *Workshop on Sports Data Visualization*, 2013. 3
- [10] C. Collins and S. Carpendale. Vislink: Revealing relationships amongst visualizations. *IEEE Trans. on Visualization and Computer Graphics (Proc. of the IEEE Conf. on Information Visualization (InfoVis))*, 13(6), Nov./Dec. 2007. doi: 10.1109/TVCG.2007.70521 9
- [11] G. Daniel and M. Chen. Video visualization. In *Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, VIS '03, pp. 54–. IEEE Computer Society, Washington, DC, USA, 2003. doi: 10.1109/VISUAL.2003.1250401 3
- [12] A. Dearden, Y. Demiris, and O. Grau. Tracking football player movement from a single moving camera using particle filters. pp. 29–37. IET, 2006. doi: 10.1049/ep:20061968 3
- [13] R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972. 2
- [14] K. A. Ericsson and H. A. Simon. *Protocol analysis*. MIT press Cambridge, MA, 1993. 8
- [15] N. Gordon, B. Ristic, and S. Arulampalam. Beyond the kalman filter: Particle filters for tracking applications. *Artech House, London*, 2004. 3

- [16] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second ed., 2004. 2, 4, 5
- [17] M. Hoernig, M. Herrmann, and B. Radig. Real time soccer field analysis from monocular TV video data. In *11th International Conference on Pattern Recognition and Image Analysis (PRIA-11-2013)*, vol. 2, pp. 567–570. The Russian Academy of Sciences, Samara, Sept. 2013. 3, 4
- [18] B. Höferlin, M. Höferlin, G. Heidemann, and D. Weiskopf. Scalable video visual analytics. *Information Visualization*, 14(1):10–26, 2015. doi: 10.1177/1473871613488571 3
- [19] M. Höferlin, B. Höferlin, G. Heidemann, and D. Weiskopf. Interactive schematic summaries for faceted exploration of surveillance video. *IEEE Trans. Multimedia*, 15(4):908–920, 2013. doi: 10.1109/TMM.2013.2238521 3
- [20] M. Horton, J. Gudmundsson, S. Chawla, and J. Estephan. Automated classification of passing in football. In T. Cao, E.-P. Lim, Z.-H. Zhou, T.-B. Ho, D. W.-L. Cheung, and H. Motoda, eds., *PAKDD (2)*, vol. 9078 of *Lecture Notes in Computer Science*, pp. 319–330. Springer, 2015. 3
- [21] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, and B. Schiele. Deepcut: A deeper, stronger, and faster multi-person pose estimation model. In *European Conference on Computer Vision (ECCV)*, May 2016. 8
- [22] J. Kang, I. Cohen, and G. Medioni. Soccer player tracking across uncalibrated camera streams. In *Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS) In Conjunction with ICCV*, pp. 172–179, 2003. 3
- [23] D. Kim, H. Shin, J. Oh, and K. Sohn. Automatic radial distortion correction in zoom lens video camera. *Journal of Electronic Imaging*, 19(4):043010–043010, 2010. 9
- [24] H. Kim and K.-S. Hong. Robust image mosaicing of soccer videos using self-calibration and line tracking. *Pattern Anal. Appl*, 4(1):9–19, 2001. 2
- [25] V. Kwatra, A. Schdl, I. Essa, G. Turk, and A. Bobick. Graphcut textures: Image and video synthesis using graph cuts. *ACM Transactions on Graphics, SIGGRAPH 2003*, 22(3):277–286, July 2003. 5
- [26] P. Laube, S. Imfeld, and R. Weibel. Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science*, 19(6):639–668, 2005. 7
- [27] P. A. Legg, D. H. S. Chung, M. L. Parry, R. Bown, M. W. Jones, I. W. Griffiths, and M. Chen. Transformation of an uncertain video search pipeline to a sketch-based visual analytics loop. *IEEE Trans. Vis. Comput. Graph.*, 19(12):2109–2118, 2013. doi: 10.1109/TVCG.2013.207 3
- [28] P. A. Legg, D. H. S. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths, and M. Chen. Matchpad: Interactive glyph-based visualization for real-time sports performance analysis. *Comput. Graph. Forum*, 31(3):1255–1264, 2012. 3
- [29] J. Liu, X. F. Tong, W. L. Li, T. Wang, Y. M. Zhang, and H. Q. Wang. Automatic player detection, labeling and tracking in broadcast soccer video. *Pattern Recognition Letters*, 30(2):103–113, Jan. 2009. 3
- [30] A. Lopes, S. Fonseca, R. Leser, and A. Baca. Using voronoi diagrams to describe tactical behaviour in invasive team sports: an application in basketball. *Cuadernos de Psicología del Deporte*, 1:123–130, 2015. 3
- [31] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. 2, 5
- [32] C. Maple. Geometric design and space planning using the marching squares and marching cube algorithms. In *2003 International Conference on Geometric Modeling and Graphics, 2003. Proceedings*, pp. 90–95, July 2003. doi: 10.1109/GMAG.2003.1219671 6
- [33] MATLAB and Computer Vision System Toolbox Release 2015b. The MathWorks Inc., Natick, Massachusetts, United States. 4
- [34] E. Monier, P. Wilhelm, and U. Rückert. A computer vision based tracking system for indoor team sports. In *The fourth International Conference on Intelligent Computing and Information Systems*, 2009. 4
- [35] Nvidia. CUDA. C Programming Guide Version 4.0. *Nvidia Corporation*, 2011. 4
- [36] P. Perez, C. Hue, J. Vermaak, and M. Gangnet. Color-based probabilistic tracking. In *ECCV*, pp. I: 661–675, 2002. 3, 4
- [37] C. Perin, R. Vuillemot, and J.-D. Fekete. Soccerstories: A kick-off for visual soccer analysis. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2506–2515, 2013. 1, 3
- [38] J. Perl, A. Grunz, and D. Memmert. Tactics analysis in soccer—an advanced approach. *International Journal of Computer Science in Sport*, 12(1):33–44, 2013. 3
- [39] H. Pileggi, C. D. Stolper, J. M. Boyle, and J. T. Stasko. Snapshot: Visualization to propel ice hockey analytics. *IEEE Trans. Vis. Comput. Graph.* 18(12):2819–2828, 2012. 3
- [40] L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. Gehler, and B. Schiele. Deepcut: Joint subset partition and labeling for multi person pose estimation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 8
- [41] T. Polk, J. Yang, Y. Hu, and Y. Zhao. Tennis: Visualization for tennis match analysis. *IEEE Trans. Vis. Comput. Graph.*, 20(12):2339–2348, 2014. doi: 10.1109/TVCG.2014.2346445 3
- [42] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: Principles Algorithms and Applications*. Prentice Hall, 2001. 5
- [43] D. Sacha, F. Al-Masoudi, M. Stein, T. Schreck, D. A. Keim, G. Andrienko, and H. Janetzko. Dynamic Visual Abstraction of Soccer Movement. In *Eurographics Conference on Visualization (EuroVis 2017)*, 2017. 9
- [44] D. Sacha, M. Stein, T. Schreck, D. A. Keim, O. Deussen, et al. Feature-driven visual analytics of soccer data. In *Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on*, pp. 13–22. IEEE, 2014. 3
- [45] Y. Seo, S. Choi, H. Kim, and K.-S. Hong. Where are the ball and players? soccer game analysis with color-based tracking and image mosaick. In *International Conference on Image Analysis and Processing*, pp. 196–203. Springer, 1997. 2, 3
- [46] M. Stein, J. Häubler, D. Jäckle, H. Janetzko, T. Schreck, and D. A. Keim. 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, 2015. 3
- [47] M. Stein, H. Janetzko, T. Breitzkreutz, D. Seebacher, T. Schreck, M. Grossniklaus, I. D. Couzin, and D. A. Keim. Director’s cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, Sept 2016. doi: 10.1109/MCG.2016.102 6, 7
- [48] M. Stein, H. Janetzko, D. Seebacher, A. Jäger, M. Nagel, J. Hölsch, S. Kosub, T. Schreck, D. A. Keim, and M. Grossniklaus. How to make sense of team sport data: From acquisition to data modeling and research aspects. *Data*, 2(1):2, 2017. 6
- [49] Y. Wu, J. Lim, and M.-H. Yang. Online object tracking: A benchmark. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013. 2