Multimedia Tools and Applications
https://doi.org/10.1007/s11042-020-09409-0

Techniques and applications for soccer video analysis: A survey

Carlos Cuevas[1] · **Daniel Quilón**[1] · **Narciso García**[1]

Received: 11 July 2019 / Revised: 6 May 2020 / Accepted: 21 July 2020 / Published online: 12 August 2020

Abstract
Nowadays, soccer is the most popular sport in our society, followed by millions of people. Consequently, many video analysis applications have been developed in the last years to provide information that can be useful for viewers, referees, coaches and players. Some of these applications are focused on specific tasks, such as detecting players, segmenting the field of play, or registering the broadcast images. On the other hand, there are applications aimed at performing tasks of a higher level, such as event detection or game analysis. Here, the most meaningful techniques and applications that have been proposed throughout the last two decades to analyze soccer video sequences are surveyed. The aim of the paper is not to compare the existing techniques, but to represent a comprehensive and organized showcase for the state-of-the-art in the field: as such, it provides a thorough review of the existing types of soccer analysis applications and the techniques used in each one of them, along with the apparent recent technical trends identified from the most recent works, and discuses the challenges in soccer analysis that still remain unsolved.

Keywords Soccer · Football · Survey · Review · Overview · State-of-the-art · FIFA · Application · Method · Strategy · Event detection · Player tracking · Ball tracking · Game analysis · Team performance

1 Introduction

Currently, soccer is the most popular sport in the world [9], with more than 265 million players in more than 200 countries [141] and with the largest television audience in terms of sports. Thanks to this great popularity and to the technological advances produced in the last decade, in recent years, numerous artificial vision applications have been proposed

This work has been partially supported by the Ministerio de Ciencia, Innovación y Universidades (AEI/FEDER) of the Spanish Government under project TEC2016-75981 (IVME).

✉ Carlos Cuevas
ccr@gti.ssr.upm.es

[1] Grupo de Tratamiento de Imágenes (GTI), Information Processing and Telecommunications Center and ETSI Telecomunicación, Universidad Politécnica de Madrid, 28040, Madrid, Spain

to carry out the automatic analysis of soccer matches. Additionally, these applications are increasingly demanded by the audience, referees, coaches, and players.

These applications cover a wide range of techniques with very varied uses, such as, for example: the detection and tracking of players [2] or the ball [16], the collection of match statistics (e.g., the distance covered by players, the speed of a player during a given move, etc.) [10], or the detection of events (e.g., outside positions, goal scoring, etc.) [65].

Although many of these applications are able to provide successful results, typically they do not perform in real-time, which is an essential requirement for their use in live transmissions. In addition, they usually propose strategies that are focused on solving/analyzing very specific aspects, but they are not integrated with others so that they can be used as a final product suitable for being used by TV networks or soccer clubs. For these reasons, new applications are continually being proposed.

In spite of the growing demand for soccer video analysis applications and the increasing number of scientific publications that try to satisfy such demand, up to our knowledge, there are no recent works summarizing and comparing the soccer video analysis strategies proposed along the last years. The only existing work performing a general summary of the soccer analysis applications and the techniques they use is a survey of 2010 [35]. However, as many new soccer video analysis approaches have been proposed since then, it is outdated. More recently, some other works summarizing soccer video analysis applications have been published [1, 57, 75, 79, 104, 107, 129]. Nevertheless, they are focused on the analysis of specific subsets of the existing techniques. Therefore, they do not show a complete overview of all types of existing applications. The survey presented in [129] reviews the approaches regarding to the detection of events, the summarization of full matches based on video streams, and the application of text sources (e.g., social networks feeds) for detecting events. In [57] and [79], the video summarization approaches are also exclusively surveyed, focusing in those strategies that collect information from online platforms (e.g., Twitter) in the case of the first one. The surveys in [1] and [104] are focused on strategies for tracking the players and/or the ball. The first one [1] describes and compares the most relevant existing tracking techniques according to the specific features they use (e.g., motion, color, edges, etc.). The second one [104] analyzes the state-of-the-art corresponding to the preprocessing and processing algorithms applied in strategies for tracking players. In [107], an overview of the strategies for the analysis of player position data (tactical behavior) was presented. Finally, in [75] the strategies for tracking the ball in sports (not only in soccer, but in table tennis, tennis, etc.) are reviewed. However, in the case of soccer, all the reviewed works are previous to 2010.

This document provides a novel and complete survey of the most meaningful techniques and applications that have been proposed throughout the last two decades to analyze soccer video sequences. Its aim is not to compare the existing applications among them, nor to study their advantages and disadvantages. Its purpose is to offer an updated and complete overview of the types of existing soccer applications and their most relevant characteristics. Thanks to this general overview, researchers who begin working in this topic will have a comprehensive and organized showcase for the state-of-the-art in the field, in which the existing types of soccer analysis applications and the techniques used in each one of them are reviewed. Moreover, the apparent recent technical trends are identified from the most recent works, and the challenges in soccer analysis that still remain unsolved are discussed. In addition, although this survey contains only brief descriptions of the listed strategies, it provides valuable information for the readers to identify, at a glance, their suitability for different purposes.

The paper is structured as follows. First, the considered groups of applications and their main characteristics are introduced in Section 2. Then, the details concerning the strategies in each of these groups are presented in the following three sections: strategies for event detection in Section 3, strategies for the detection and tracking of player and/or ball in Section 4, and game analysis strategies in Section 5. Finally, the future directions and main conclusions obtained from the analysis performed in this survey are presented, respectively, in Sections 6 and 7.

2 Groups of applications

In general, depending on their objectives and needs, soccer video analysis applications can be classified into the following three groups:

– G1: Event detection (Section 3).
– G2: Detection/tracking of players and/or ball (Section 4).
– G3: Game analysis (Section 5).

The purpose of the applications in G1 is to detect the most relevant events that take place during the broadcast of the match, such as: goal scoring, fouls, shots on goal, corner kicks, penalties, etc. Typically, these applications mainly analyze audiovisual features of the broadcast TV signal: types of camera shots [5, 161], audio intensity level [135, 140], movements made by the camera [139, 158], detection of slow-motion video segments [39, 147], wipe-type shot transitions [87, 140], graphs with the statistics of the match [96, 135], etc. However, some applications not only analyze the transmitted audiovisual signal, but also the data from some other sources: additional cameras strategically placed in the stadiums [38], the content of sport press websites [32, 176], or the content published by people on social networks (e.g., Twitter) [70, 177].

The second considered group, G2, includes those applications focused on detecting and tracking the players [127, 134] and/or the ball [64, 81]. Typically, these applications make use of sophisticated techniques to solve complex analysis situations, such as occlusions between players, abrupt movements of the cameras, or sudden illumination changes [104].

Finally, the applications in G3 are focused on analyzing what is happening along the matches. First, they typically apply low-level analysis methods (e.g., detection and tracking) and, then, they use the obtained results to perform different kinds of high-level analyses about the match: distance covered by the players from each team [10], possession statistics of the ball [170], offside detection [58], team tactics [49], etc.

Typically, applications in G1 are used once the match has finished, so they are offline applications. On the other hand, detection and tracking applications (G2), and game analysis applications (G3) are typically used during the transmission of the game, either in real-time (e.g., to provide statistics) or after a few minutes to analyze certain relevant plays (e.g., a goal or an offside) [35].

In some cases, strategies corresponding to G1 are used as starting point for the applications in G2 and G3, since these applications are used exclusively in those fragments of video in which specific events have been detected. For example, in [176] it is proposed a strategy to determine the tactic followed by a team after a goal. Therefore, the detection of the goal events is a required starting point. Similarly, many game analysis applications (G3) take as input data the results provided by detection and tracking applications (G2). For example, the strategy in [94] uses ball and player positions provided by an external tracking tool to analyze the ball possession throughout the match. Consequently, as it is going to be seen in the following sections, most of the outstanding strategies proposed in the last decades correspond to groups G1 and G2, since their tasks could be considered more general.

There are also some authors who have proposed strategies that focus on specific tasks that may be useful for applications in more than one of the three previously defined groups. One of these typical specific tasks is the segmentation of the field of play [73, 125]. As it will be seen later, the segmentation of the field of play is an initial stage used in most applications for detecting and tracking players (group G2). However, it can also be useful in applications for detecting events (group G1) or analyzing the game (group G3). Another example are the works focused in the shot view classification [43]. These works are not only included in many event detection strategies (group G1), but can also be considered as event detection methods in themselves [54]. However, they can also be used in applications for tracking players (group G2) or analyzing tactics (group G3), since these applications cannot be applied to all types of shot view.

Finally, it must be mentioned that there are also some works proposing datasets that can be used to assess the quality of the results provided by soccer video analysis applications [36, 51, 169].

3 Event detection

As stated in Section 2, the purpose of this group of applications is to automatically detect the most important events taking place during the broadcast of a soccer match: goal scoring, fouls, shots on goal, corners, penalties, offside positions, etc. These applications arise from the need to automate the tasks of identifying those events that may be of interest to the viewer [79]. In this way, users acquire the ability to directly view the most interesting parts of the game. Some applications have been designed for automatically generating video summaries with the most interesting plays of the match, where users can configure the duration and the type of plays of the summary [144]. The identification of the events made by these applications is useful for many other applications with the purpose of analyzing such events (e.g., plays, trajectories of players, distances of shots to the goal, outside positions, etc.).

One of the main drawbacks of applications for detecting events is their high computational cost, since they require analyzing the full video of the game. However, all the tasks performed by these applications are typically done after the game, i.e., offline. Therefore, they do not usually have run-time or computational restrictions.

Table 1 shows a summary of the main characteristics of the most representative strategies proposed in the last two decades for event detection in soccer. According to the data in this table, such strategies can be classified attending to different criteria.

Depending on the source of information they use (second column of the table), it is possible to distinguish between strategies using:

– Broadcast signal: About 90% of the reviewed works analyze the audiovisual signal transmitted outside the stadiums (i.e., the TV signal). As stated before, event detection is typically done after the game (i.e., offline). Consequently, although not all works explicitly mention it, they use pre-recorded TV signals instead of using the live broadcast signal. In addition, it should be noted that most authors exclusively use this audiovisual signal. However, in some works it is combined with other information sources.

Table 1 Summary of the main characteristics of the the most representative strategies proposed in the last two decades for event detection in soccer

Strategy	Source	Semantic feafures	Tasks	Methods
2001-Xu [161]	Broadcast signal	F1	T1, T3	Rules with predefined thresholds
2002-Assfalg [5]	Broadcast signal	F1, F4	T1, T3	Hidden Markov Models
2003-Assfalg [4]	Broadcast signal	F1, F5, F10	T1	Finite State Machine
2003-Ekin [39]	Broadcast signal	F1, F5, F6	T1, T2	Probabilistic models Zero Crossing Measure
2004-Chen [24]	Broadcast signal	F1, F2	T1	Decision tree
2004-Leonardi [93]	Broadcast signal	F1, F2, F4	T1, T3	Controlled Markov Models
2004-Wan [155]	Broadcast signal	F1, F2, F8	T1	Decision tree
2004-Xie [158]	Broadcast signal	F1, F4	T1, T3	Hidden Markov Models
2005-Sadlier [135]	Broadcast signal	F1, F2, F3, F4, F5	T1	Support Vector Machine
2006-Huang [67]	Broadcast signal	F1, F2, F3, F4, F5, F7	T1, T2	Bayesian network
2006-Xu [159]	Broadcast signal Webcast	F1, F3, F4	T1, T4	Webcast analysis
2008-Ding [32]	Broadcast signal Webcast	F1, F4	T4	Webcast analysis
2008-Zhu [176]	Broadcast signal Webcast	F1, F4, F9, F10	T1, T4	Webcast analysis Text recognition
2009-D'orazio [38]	4 extra cameras	F10	T1	Detection and tracking
2009-Eldib [40]	Broadcast signal	F1, F3	T1, T2	Rules with predefined thresholds
2009-Halin [54]	Broadcast signal	F1	T1	Rules with predefined thresholds
2009-Liu [96]	Broadcast signal	F1, F2, F3	T1	Support Vector Machine
2010-Boyar [18]	Broadcast signal Webcast	F1	T1, T4	Hidden Markov Models Webcast analysis
2010-Tjondronegoro [150]	Broadcast signal	F1, F7	T1, T2, T3	Probabilistic models Support Vector Machine Rules with predefined thresholds
2011-Zawbaa [173]	Broadcast signal	F1, F3, F5, F7, F8	T1, T2	Support Vector Machine Neural networks
2012-Nguyen [112]	Broadcast signal	F1	T1, T2	Rules with predefined thresholds Histogram comparisons

Table 1 (continued)

Strategy	Source	Semantic feafures	Tasks	Methods
2012-Nichols [114]	Social networks	-	T4	Social network analysis
2012-Zubiaga [177]	Social networks	-	T4	Social network analysis
2013-Halin [53]	Broadcast signal Webcast	F1, F2	T1, T4	Rules with predefined thresholds Webcast analysis
2013-Hosseini [65]	Broadcast signal	F1, F2, F7, F9	T1, T2	Fuzzy logic Neural networks
2014-Esmin [41]	Social networks	-	T4	Social network analysis
2014-Jai [70]	Broadcast signal Social networks	F2, F3	T4	Social network analysis
2014-Nguyen [113]	Broadcast signal	F1, F4	T1	Rules with predefined thresholds
2014-Tavassolipour [147]	Broadcast signal	F1, F5, F6, F7, F9	T1, T2, T3	Support Vector Machine Histogram comparisonsBayesian networks Hidden Markov Models
2015-Kolekar [87]	Broadcast signal	F1, F4, F7, F9	T1, T2	Probabilistic models Bayesian networks Histogram comparisons
2015-Sigari [139]	Broadcast signal	F1, F4	T1	Support Vector Machine
2015-Sigari [140]	Broadcast signal	F1, F2, F7	T1, T2	Fuzzy Logic Decision tree
2016-Jiang [71]	Broadcast signal	F1	T1, T3	Neural network
2016-Raghuram [126]	Broadcast signal	F2	T1	Support Vector Machine Neural network
2016-Saraogi [137]	Broadcast signal	F1	T1	Probabilistic models
2017-Lee [90]	Broadcast signal	F9, F10	T1	Object tracking and motion recognition
2017-Song [142]	Broadcast signal	F1	T1	Fuzzy Logic
2017-Wang [156]	Broadcast signal Webcast	F4, F5	T1, T4	Attack-defense transition analysis Webcast analysis
2018-Hong [63]	Broadcast signal	F1	T1	Neural network
2019-Fakhar [42]	Broadcast signal	F1, F3, F4	T1, T2	Histogram comparisons Support Vector Machine Neural network
2019-Yu [168]	Broadcast signal	F1, F3	T1, T2	Neural network

Marked with F, analyzed semantic features: F1 - Shot sizes and content; F2 - audio; F3 - graphics; F4 - camera movement; F5 - field of play lines; F6 - slow motion; F7 - wipes; F8 - goal detection; F9 - player movements; F10 - ball movements

Marked with T, main tasks performed: T1 - audiovisual feature analysis; T2 - replay analysis; T3 - paused match elimination; T4 - cross-source correlation

- Extra cameras: In some strategies, the detection of events is done by using the video signal obtained with additional cameras, placed in strategic points of the stadiums (e.g., near the areas for detecting offside events, or near the goal line for detecting goal scoring), that are independent to those used to broadcast the matches.
- Webcasting: A significant percentage (about 15%) of the strategies reviewed combine the analysis of the audiovisual signal with the chronicles of the game that are made in the sports press web pages (webcasting). These strategies look for keywords (goal, fault, penalty, etc.) on the analyzed text.
- Social Networks: Similarly to the applications using webcasting data, there are some applications that analyze the text published in social networks (typically Twitter) to detect keywords.

According to the data in the third column of the table, the reviewed strategies can also be classified attending to the semantic features they analyze:

- F1: Shot sizes (long shot, medium shot, or close-up) and types of content in such views (stands, players, sidelines, etc.).
- F2: Level of the transmitted audio.
- F3: Graphics with statistics on the images.
- F4: Types of camera movements (e.g., fast/slow pan/tilt motion, or zoom level).
- F5: Lines of the field of play that appear in the images.
- F6: Video segments in slow motion (typical during replays).
- F7: Wipes (typical at the beginning or at the ending of a replay).
- F8: A goal, or part of it, visible in the images.
- F9: The position and/or movement of the players of each team and the referees.
- F10: The movement of the ball.

Some of the reviewed works make use of only one of the above described types of features. However, many others analyze several of such features.

As it can be seen in the fourth column in Table 1, the reviewed strategies can also be classified according to the following tasks they perform for detecting events:

- T1: Recognize any of the specific audiovisual features that typically occur during a soccer match event (e.g., fast camera movements, increasing audio level, labels with statistics on the screen, fast changes from defense positions to attack ones, camera views that show the area close to a goal, etc.).
- T2: Detect replays (since they usually show relevant events).
- T3: Detect the intervals of the video in which the game is being played (so, discarding the intervals in which the game is paused).
- T4: Relate the content of the videos with external sources of non-audiovisual information (webcasting or social networks).

In addition to the semantic features and the tasks used by the strategies that have been reviewed, the table also summarizes (last column) the methods that these strategies use for detecting events. The way in which these methods are applied by each strategy is studied more in depth throughout the following subsections. First, the methods based on the analysis of audiovisual data are analyzed in Section 3.1. Then, Section 3.2 describes the methods based on external sources of non-audiovisual information.

3.1 Methods based on audiovisual information

To identify when events are happening, the audiovisual signal is analyzed to try to detect indicators that typically appear when an event is taking place, such as wipes before and after a repetition, a slow motion video, increased audio activity, short shots on the stands or the players, graphics with statistics, etc.

Some strategies determine when an event takes place through a joint analysis of the content shown in the images, the analysis of the audio, and the operation of the camera (strategies labeled as T1 in Table 1). In events such as goal scoring or fouls, the level of the audio increases due to the shouts of the spectators, images of the stands usually appear, short camera shots are used, the players of the same team usually group in the celebration and, after the event, a digital graphic with statistics always appears on the image (e.g., goal score or number of fouls). In events corresponding to corner kicks or penalties, a specific area of the field of play is typically shown, the camera usually remains static at the time of the throw, the type of shot used is usually the same, and the players are distributed in a very characteristic way. In the case of a counterattack, the camera movements are very fast, the audio level increases, and the commentators speak with greater intensity.

During the broadcast of a match, replays are very common after outstanding plays (e.g., after scoring a goal or after committing a foul). Consequently, several authors have proposed strategies to detect replays (strategies labeled as T2 in Table 1). Typically, at the beginning and end of such replays shot transitions of type wipe appear. Additionally, these wipes usually show the logo of the broadcast competition and many times they are played back in slow motion (see Fig. 1). The main drawback of strategies for detecting replays is that they are very dependent on the broadcast of the video, not being able to detect an event if the director has decided that there is no replay associated with it. Moreover, it is also possible to classify erroneously an event if the replay is very separated in time from the corresponding event.

Typically, in all sports the game stops when a relevant event happens. Based on this fact, some of the works proposed in the literature try to detect events by identifying the stops and restarts of the game (strategies labeled as T3 in Table 1). Usually, these works consider that the game is in progress when the shots are long and the zoom level is low.

The methods that are typically used in these audiovisual-based event detection strategies are described in the following subsections.

3.1.1 SVM-based methods

A significant number of the methods included in the reviewed strategies are based on the use of Support Vector Machines (SVMs). An SVM is a supervised machine learning algorithm that, from characteristics extracted from the video and using models trained with other matches, is able to determine if an important event is taking place or not.

Fig. 1 Examples of shot transitions of type wipe

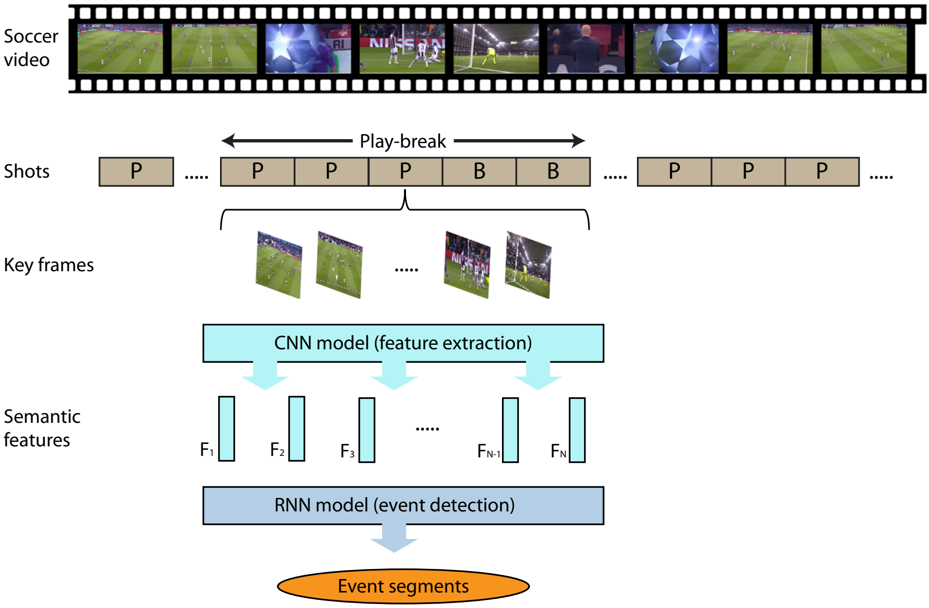


Fig. 2 Flowchart of the NN-based strategy proposed in [71]

The SVM proposed in [135] allows detecting short shots of the players, the increase in the activity of the commentators, the graphs of statistics on the screen, or the amount of movement of the camera. In [96], SVMs are used to analyze the intensity level of the commentators' voice and to look for keywords in the statistic graphics on the screen. The SVM proposed in [147] is used to classify the frames into "boundary" and "non-boundary" classes from spatio-temporal and compressed domain features. In [139], an SVM-based method for detecting counterattacks is proposed, which is based on the detection of video segments in which the camera is panning quickly. Finally, in [150], an SVM is used to search for images in which animations of about 10-20 frames occur, where an object very different to the background appears and that, additionally, follows the small-large-small appearance pattern.

3.1.2 NN-based methods

Another type of methods commonly used in the strategies for event detection are those based on Neural Networks (NNs).[1] Similarly to SVMs, they are supervised machine learning algorithms. However, they are much more flexible than SVMs because of their ability to reproduce and model nonlinear processes.

An NN trained with logos from several competitions is used in [65] a to try of detecting the replay events. The strategy proposed in [71] allows classifying some events (goals, corner kicks, and card events). First, game stop/restart temporal instants are detected. Then, a pre-trained Convolutional Neural Network (CNN) is used to extract semantic features of key frames. At last, a Recurrent Neural Network (RNN) is used to map the semantic features to the considered event types. Figure 2 illustrates the flowchart of this NN-based strategy.

[1]NNs take their name and behavior from the biological process that occurs in the brain, known as synapses.

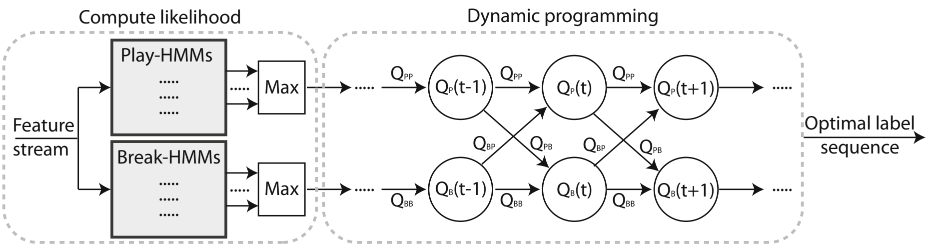


Fig. 3 Topologies of the HMMs proposed in [158], where the Qs illustrate the transition likelihoods

In [63], a strategy based on CNNs is proposed, which is able of classifying between six categories related to the possible shot sizes: corner, free-kick, long view with no events, close-up view, penalty, and goal. In [173], an automatic learning algorithm that combines SVMs with NNs is proposed, which is capable of analyzing audiovisual features and, in addition, detecting the graphic on the screen corresponding to the match score. An algorithm that also combines SVMs with NNs is proposed in [126] to analyze and segment the audio according to its characteristics (silence, only commentator's voice, only sound from the bleachers, and both the commentator's voice and bleachers sound). In [42], a strategy for detecting events (goals, cards, fouls, and corners) has recently been proposed. First, a histogram-based analysis and a linear SVM are used to detect the shot boundaries. Then, replays are detected thanks to a method based on the detection of logos. Finally, the event in each replay is recognized with an NN. Another CNN-based strategy that has been recently proposed is described in [168]. This strategy locates the replays (analyzing whether there is a logo in the images), detects the events, and finally provides video summaries.

3.1.3 HMM-based methods

A third group of supervised learning algorithms for event detection is based on Hidden Markov Models (HMMs), which are statistical models that assume that the system being modeled is a Markov process.[2]

The HMM-based approach in [5] allows detecting highlights from the camera motion and the position of the players on the field of play. The strategy in [158] first computes the dominant color ratio and the motion intensity of the videos in the compressed domain. Then, it segments the videos into two generic states ("play" and 'break") with a set of HMMs that use dynamic programming to address the transitions among states (see Fig. 3). The HMM in [147] is used for classifying the shots into the states "break" or "play", and also for classifying them according to the view types of shot. In [93], instead of detecting changes between view types, a multimodal analysis using Controlled Markov Chains (CMCs) is proposed to describe the temporal evolution of the views throughout an event. This analysis is based on the idea that the fast movements of a camera typically occur when the game is in progress, whereas the camera is typically static if the game is stopped. Additionally, to reduce the presence of false alarms, the increase of the audio signal loudness between shots is also used as a descriptor.

[2]Sequence of possible events in which the probability of each event depends only on the state attained in the previous event.

3.1.4 Probabilistic-based methods

Some of the reviewed strategies for event detection are based on Bayesian Networks (BNs). A BN is a probabilistic graphical model whose nodes represent random variables that have associated probability functions that take as input certain sets of values of the preceding nodes.

Some authors have used BNs to classify the events among several possible classes: scored goals, corner kicks, penalty kicks, and yellow and red cards in [67]; scored goals, corner kicks, yellow and red cards, and foul kicks in [147]; or scored goals, shots stopped from goalkeepers, yellow and red cards, and shots on goal in [87]. In Fig. 4, examples of BNs corresponding to the detection of different events (scored goal, corner kick, penalty kick, and card) are illustrated.

There are also some strategies that, instead of using BNs, propose alternative probabilistic-based methods. The strategy in [39] detects slow motion video segments using

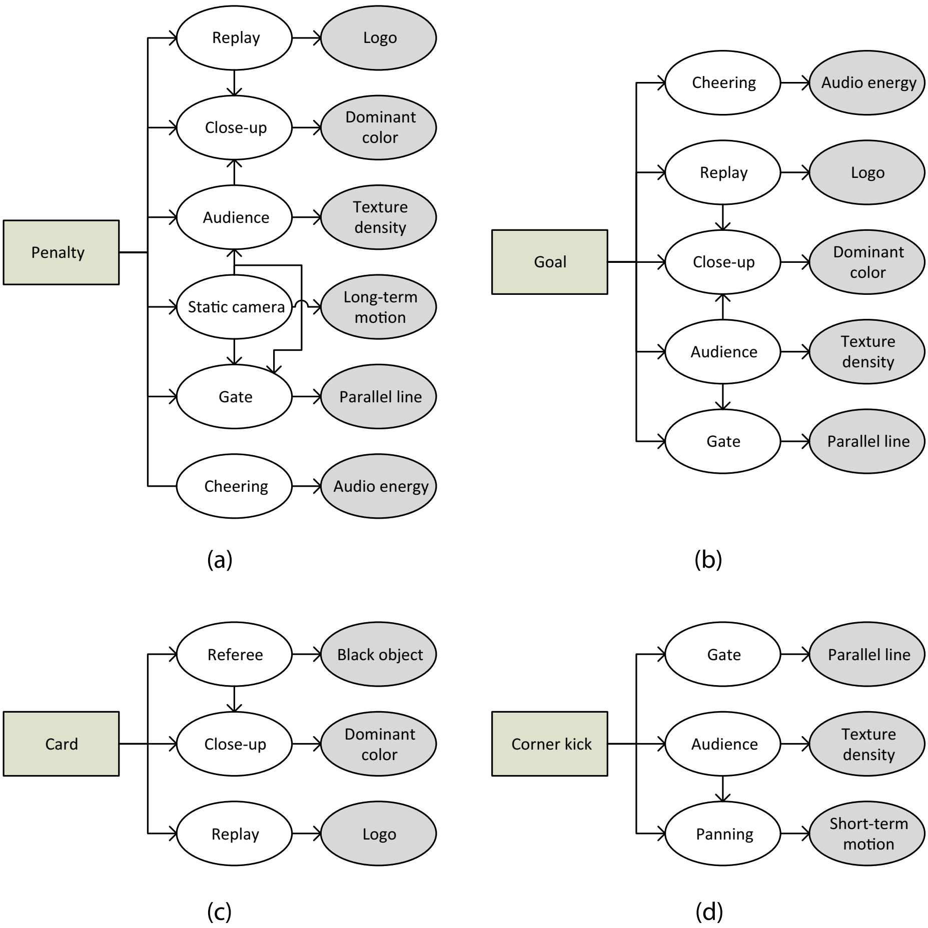


Fig. 4 Examples of BN structures in [67] for detecting four different kinds of events: **a** Goals; **b** Corner kicks; **c** Penalty kicks; **d** Cards

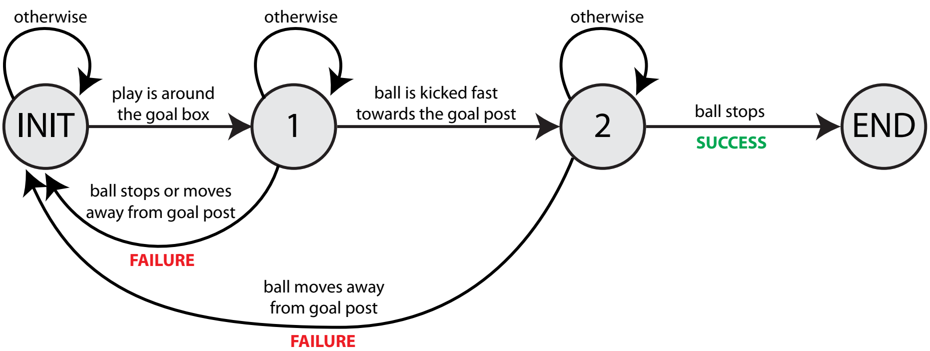


Fig. 5 FSM used in [4] for detecting the "shot on goal" highlight

the Zero Crossing Measure proposed in [120], which is based on statistic models that allow computing the probability of each video segment of being a slow motion video and, additionally, identifies the beginning and the end of the detected slow motion intervals. Other strategies [137, 150] use probabilistic models to segment the field of play. Then, by comparing the size of the segmented region with the size of the image, they determine the shot size that is being used.

3.1.5 Other audiovisual-based methods

In some works [40, 54, 112, 113, 150, 161], specific series of rules with predefined thresholds are used to analyze some audiovisual characteristics of the sequences such as, for example, the number of frames in which a goal appear, the amount of movement of the camera, the zoom level, the frequency of short shots, or the replays.

Other strategies propose detection techniques based on Decision Tree Logic (DTL) [155], which are based on applying a series of rules designed to evaluate and categorize some conditions that occur successively and correspond to a specific event (e.g., goal scoring). In [24], a decision tree model is used for goal detection. The strategy in [140], by using a decision tree to detect wipe-type shot transitions, is able of locating replay events.

In [4], as an alternative to learning-based approaches, highlights are modeled with Finite State Machines (FSMs)[3] in which the state transitions are determined by several visual descriptors. Figure 5 shows one of the FSMs used in this work.

Other works [65, 140, 142] propose making use of Fuzzy Logic (FL)[4] systems to detect events from some characteristics of the video, such as the shot size, the peaks in the audio level, or the replays.

There are also some strategies that apply detection and tracking algorithms to try of recognizing those movements of the players and/or the ball that correspond to significant events. In [38], goal detection events are detected by analyzing the ball position in images obtained with several cameras placed at strategic positions of the stadiums and independent

[3] Computation model that can be used to simulate sequential logic.

[4] Approach to variable processing that allows for multiple values to be processed through the same variable. In contrast to probability, which is a mathematical model of ignorance, fuzzy logic uses degrees of truth as a mathematical model of vagueness.

from those used in broadcasting (see Fig. 6). In [90], events are detected using a multi-object tracking algorithm to obtain the position and size of the ball, players, and referees.

Since competition logos generally appear in replay events, some strategies include histogram-based methods to detect such logos in images. In [42, 87, 112, 147], the logos are detected by analyzing changes in brightness in the histograms.

3.2 Methods based on external sources

Many of the reviewed event detection methods make use of external sources of non-audiovisual information (strategies labeled as T4 in Table 1). Typical external sources are webcasting and social networks. An important drawback of these methods is that they usually require to synchronize the external sources with the audiovisual signal. In addition, it should be considered that there is a delay between the events and the information received, which can be significantly high in the case of social networks.

Since many sport press websites offer the live match chronicles (webcasting), several authors have developed algorithms to extract information from such websites and use it to identify and classify events [18, 32, 53, 156, 159, 176]. Typically, these methods analyze the content of the chronicles in search of keywords that are related to the event to be classified (e.g., goal, corner kick, card, penalty, etc.). Then, to locate the event in the broadcast TV signal, they compare the time stamp in the text event with the overlaid video clock used to indicate the elapsed time. Finally, they determine the beginning and the end of the event by analyzing the duration of the camera shots around the moment in which the event was detected. Figure 7 shows an example of the architecture used by approaches based on webcasting analysis.

Currently, every day millions of messages are sent on social networks describing events that people are watching on TV, such as sport events. For this reason, over the last few years, some works have been proposed that try to detect and classify events based on information extracted from social networks such as Twitter [41, 70, 114, 177]. These algorithms, from the analysis of the texts of the published tweets, identify the most repeated words and relate them to the events to be detected to generate an ordered list of such events.

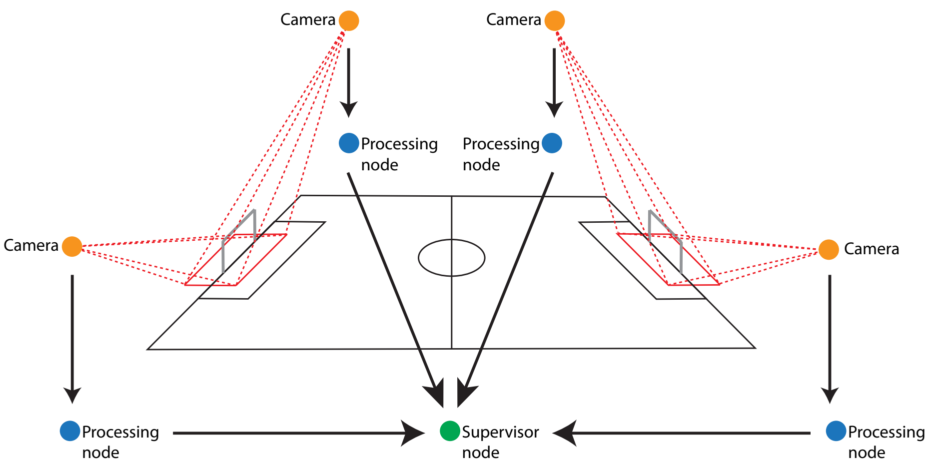


Fig. 6 Scheme of the visual system used in [38]

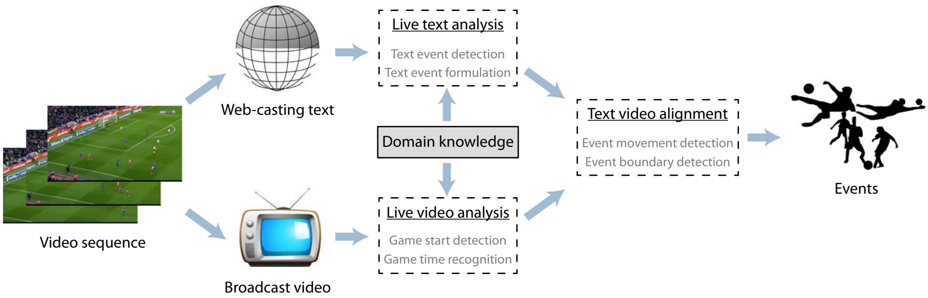


Fig. 7 Example of the webcasting-based architecture used in [159]

4 Detection and tracking

Due to the technological advances produced in the last decade in the field of image analysis, the development of object detection and tracking applications has increased in the last years [104]. The use of these applications in sport environments, such as soccer, allows tracking both the players and the ball and, consequently, determining their position throughout the matches. From these data it is possible to answer many questions that are of special interest to the teams, to the audience and even to the referees [34]: what are the strengths and weaknesses of each of the players?, what role does each player play and what are their abilities?, what are the offensive characteristics of each team?, how does a team create scoring opportunities?, etc. Moreover, the real-time analysis of soccer matches is receiving special attention from the referees, since during the game many dubious situations take place due to the perspective, occlusions, and the speed of the events.

This section describes the methods used by the most representative strategies proposed in recent years to detect and track the players and the ball. These strategies are the basis from which most game analysis applications (described in Section 5) start, and that will allow to answer the questions mentioned above. It must be noted that some of the works previously described in Section 3 use methods to detect and track players/ball as part of their strategies for detecting events. However, such algorithms are quite simple or have been taken from other approaches such as the ones described in this section, which are specifically focused on the development of algorithms for detecting and tracking the soccer players and/or the ball.

First, in Section 4.1, the lists of groups of the most typical methods used to detect and track players/ball are introduced. Then, the algorithms focused on the analysis of the players are described in Section 4.2, whereas the ones that analyze the ball are detailed in Section 4.3. Along these subsections, a distinction will be made between algorithms that only include detection methods and those that also apply tracking algorithms.

4.1 Detection and tracking algorithms

The detection methodologies used by the reviewed strategies can be classified in the following groups:

– Key-Point Detection (KPD): Methods that are based on looking for key-points in the texture of the images. Some typical key-point detectors are Harris Corner Detector

(HCD) [55], Kanade-Lucas-Tomasi feature tracker (KLT) [101], and Scale-Invariant Feature Transform (SIFT) [99].
- Background Subtraction (BGS): Methods that segment the foreground objects in the scene, by comparing each current image with a probabilistic background model [29]. Some of the most popular BGS methods are those based on Gaussian Mixture Models (GMMs) [28], Frame differencing [122], Temporal Median Filtering (TMF) [3], and Non-Parametric Modeling (NPM) [13].
- Region-Based Segmentation (RBS): Algorithms that aim at separating video frames into regions of differing characteristics, each one enclosing uniform characteristics within (e.g., color, texture, illumination, etc.). Some outstanding methods are: Mean-Shift Clustering [25], Active Contour Modeling [77], or K-Means Clustering [56].
- Supervised Learning (SL): Methods that try to locate in the analyzed images those objects that are similar to those learned in a previous training stage. These methods are typically used not only to detect moving objects but also to classify them in their teams. Some examples of SL methods are the SVMs [27], the Adaptive Boosting (AdaBoost) [48], and the CNNs [174].
- Silhouette Detection (SD): Methods that analyze the images in search of silhouettes and contours similar to a given model. The association between models and silhouettes with the contour must take into account aspects such as shape and density. Some representative SD methods are the Hough transform [68], the Wavelet transform (WT) [31], and the Hausdorff distance [167].

Regarding the strategies used to track the detected objects, the methods used by the reviewed strategies can be classified in the following three groups:

- Key-Point Tracking (KPT): Objects detected in consecutive images are related by establishing relationships between their key-points. These relationships are used to estimate the position that such objects occupy in subsequent frames. Some typical KPT algorithms are the Kalman filter [72], the Particle Filter [115], the Joint Probabilistic Data Association Filter (JPDAF) [52], and the Markov chains [118].
- Template Matching (TM): Techniques that track objects over time by analyzing the variations in their shape or appearance, whether due to translations, rotations or affine transformations. Some examples of typical TM methods are those based on Cross-correlation and Normalized cross-correlation [20], or on Sum of squares differences [124].
- Silhouette Tracking (ST): The tracking is done by estimating the region occupied by the object in each image after applying a silhouette detection. Most popular methods are Contour Tracking methods [6] and Shape Matching algorithms [7].

4.2 Players

Detecting and tracking players is essential for a better understanding of the game. However, there are several associated challenges that must be faced to obtain successful results, such as the occlusions between players, the abrupt movements of the camera, the changes of lighting, the lack of resolution in very distant players, the blurring of the players that are moving, or the players remaining static long periods of time [29].

Table 2 shows a summary of the most representative strategies published in the last years for detecting and tracking players in soccer video sequences. The second column of the table indicates the source used by the strategy developed in each work. It can be seen that most

Table 2 Summary of the main characteristics of the the most representative strategies to detect and track players in soccer video sequences

Strategy	Source	Detection	Tracking
2000-Naemura [111]	Broadcast signal	BGS	
2003-Kim [83]	Broadcast signal	RBS	KPT
2004-Iwase [69]	16 extra cameras	RBS	KPT
2005-Hayet [59]	Mobile cameras	KPD	KPT
2005-Xu [160]	8 extra cameras	BGS	KPT
2006-Beetz [12]	Broadcast signal	BGS	TM
2006-Figueroa [45]	4 extra cameras	BGS	
2006-Figueroa [46]	4 extra cameras	BGS	TM
2007-Beetz [11]	Broadcast signal	RBS	TM
2007-Spagnolo [143]	6 extra cameras	BGS, RBS	
2008-Nunez [117]	Broadcast signal	BS	
2009-Khatoonabadi [80]	Broadcast signal	BGS, RBS	KPT, TM
2009-Kim [84]	Broadcast signal	BGS	KPT
2009-Liu [97]	Broadcast signal	SL	KPT
2011-Kataoka [78]	1 extra camera	BGS	KPT
2011-Tong [151]	Broadcast signal	SL	KPT
2012-Ali [2]	Broadcast signal	BGS	
2013-Bai [8]	Broadcast signal	SL	
2014-Martín [105]	6 extra cameras	BGS	TM
2015-Gerke [50]	Broadcast signal	SL	
2015-Sabirin [134]	Broadcast signal	BGS	TM
2015-Upendra [127]	Broadcast signal	BGS	
2017-Manafifard [103]	Broadcast signal	RBS	
2017-Yang [164]	Broadcast signal	RBS	KPT
2018-Kim [86]	6 extra cameras	BGS	KPT
2018-Lu [100]	Broadcast signal	SL	
2018-Zhang [174]	Other	SL	
2019-Kim [85]	Broadcast signal	BGS	ST
2019-Komorowski [88]	6 extra cameras	SL	
2019-Sverrisson [145]	Broadcast signal	SL	KPT, other

Detection methods: KPD - Key-Point Detection; BGS - Background Subtraction; RBS - Region-Based Segmentation; SL - Supervised Learning

Tracking methods: KPT - Key-Point Tracking; TM - Template Matching

strategies use as input source the broadcast signal. However, some other works make use of one [78] or more [46, 69, 86, 88, 143, 160] additional (extra) cameras placed at strategic positions of the stadium, or additional (extra) mobile cameras [59].

The third and fourth columns of the table indicate the methods applied by each strategy to, respectively, detect and track the players. It can be observed that there are strategies that focus on exclusively detecting players [2, 8, 127], whereas most works do not only detect but track them. Figure 8 shows a flowchart with the typical steps included in the methods

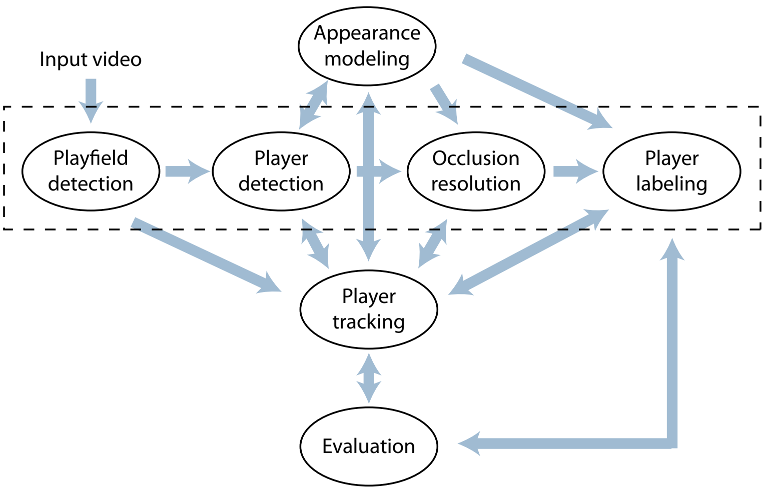


Fig. 8 Flowchart used in [104] for detecting and tracking players

to detect and track players. The way in which these methods are applied by each strategy is studied more in depth throughout the following subsections.

4.2.1 Detection

Typically, the first stage in detection strategies is a color-based foreground segmentation [125], since players always wear kits with colors that clearly differentiate them from the field of play. Some strategies analyze the histogram of the H component of the HSV color space in search of the dominant mode [78, 117, 134], which is assumed to correspond to the green color of the field of play. In other works [2, 127], the field of play, assumed to show a color distinguishable as green, is inferred as the region occupied by pixels satisfying the rule $G > R > B$, where G, R, and B are the green, red, and blue components in the RGB color space. In [160], a GMM is proposed to detect moving objects and separate them from the field of play. In [45, 111], morphological operations are applied iteratively to segment the players, the referees and the ball. The obtained results are then improved in a second stage through a Watershed Transform [132]. In [86], the player detection is performed by applying a frame differencing BGS algorithm that is improved by a second-stage analysis based on the foreground edge information [82].

Alternatively, other strategies make use of RBS methods. In [11, 83], a k-means algorithm to group pixels with the same appearance and segment the grass, the lines of the field of play, the players, the ball, and the referee was proposed. In [80], a color-based BGS method is first applied. Then, foreground pixels are grouped in regions by using an RBS algorithm. Finally, a median filtering is applied to split those players that are part of the same region. Otsu's algorithm is used in [164] to segment players. In [143], after applying a BGS, the Basic Sequential Algorithm Scheme (BSAS) [162] is used to classify the players by grouping pixels according to the color of their equipment and comparing them with known samples of such equipment. In [103], a blob-based strategy that is able of detecting and labeling multiple players is proposed. This strategy combines a two step blob detection (a color-based grass detection to detect the candidate player blobs; and a multi-scale edge-based detector to remove false positives) with a particle swarm optimization mechanism that is capable of detecting and labeling multiple players, even if some of them are partially occluded.

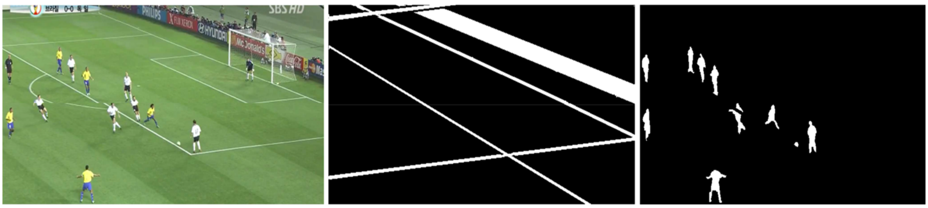
Since lines in the field of play are white, they differ significantly from the green color of the grass. Therefore, color-based segmentation methods typically require additional algorithms to remove the lines from their obtained results. Some authors propose analyzing the density and size of the segmented objects to detect and discard lines [12]. As an alternative, other works [2, 127] use the Hough transform for detecting lines and separating them from the rest of segmented objects. Figure 9 illustrates some of the results provided by the strategy in [2].

CNNs have proven to be among the most competitive methods for object detection in images [98, 146]. Consequently, several CNN-based strategies for detecting players have been recently proposed. In [100], a cascaded CNN that jointly detects players and classifies them in their teams is proposed. This strategy is able of detecting players not only in soccer, but in other sports such as basketball or ice hockey. Moreover, it is able of accurately detect players under challenging conditions such as varying illumination, camera movements, and motion blur. In [174], an end-to-end Reverse Connected Convolutional Neural Network (RC-CNN) is proposed to detect players, which is able of detecting players in multiple scales and in several challenging scenes. The CNN-based strategy that is proposed in [50] does not focus on detecting generic players on the field of play, but on automatically recognizing the numbers on the players' shirts. To this end, an initial player detection based on histogram of oriented gradients (HOG) is applied together with a linear SVM. Then, the cropped images resulting from such detection are used as the input of a CNN that has been previously trained to recognize numbers.

In addition to the strategies using CNNs, alternative SL methods have been proposed in other works to perform the detection of the players. Typically, these works not only detect players, but also classify them in their respective teams. In [97, 151], the players are classified through a Boosted Cascade Learning detector, which is a supervised learning method trained with hundreds of positive and negative images (see Fig. 10). In [8], players are detected using a method that makes use of a one-class SVM and automatically generated training data.

4.2.2 Tracking

Once the players have been detected, it is possible to track them in order to register their positions along the video sequences. Most of the reviewed strategies use KPT methods to perform such task. Some of them use the Kalman filter [59, 69, 80, 84, 160] (see Fig. 11). Alternatively, other works track the detected objects using Markov Chain Monte Carlo



(a) (b) (c)

Fig. 9 Partial result provided by the detection strategy in [2]. **a** Original image. **b** Lines detected using the Hough transform. **c** Mask of players detected

Fig. 10 Examples of positive and negative samples used in [97] for training the Boosted Cascade Learning detector

(MCMC) algorithms [97, 151], since they are able to provide better results in sequences with short-term occlusions. Finally, in [78, 164], an enhanced particle filter [116] is used to perform the tracking, which is robust to inaccurate segments (containing part of grass or lines).

To improve the quality of the results in sequences with a significant amount of occlusions between players, some of the described strategies use several cameras placed at strategical positions [46, 69, 160]. The tracking of the players is carried out independently for each camera, so if a player is occluded by another one, at least one of the cameras will be able to correctly track both players.

In other works [11, 12, 46, 134], instead of using KPT-based methods, TM techniques are proposed [105], which are typically based on the correlation of properties of the objects (color, density, size and position) between consecutive images.

As an alternative to the previous tracking algorithms, the strategy in [85] proposes a method that, after detecting the players with a BGS combined with an edge-based analysis, tracks the players by establishing relations between the topographic surfaces (i.e., 3-dimensional silhouettes) of the detected players. Another recently proposed alternative is the one described in [145], which uses a data association logic that allows players to be tracked by analyzing the position and size of bounding boxes obtained from a CNN-based detection, a set of probabilities associated with such detections, and their color histograms.

Fig. 11 Examples of the tracking results in [80]. Blue and red boxes are used for the players in each team. The goal keeper has assigned turquoise boxes. The yellow boxes show the positions of the referee. Finally, the white boxes are used to represent cases with multiple players encapsulated

Fig. 12 Some examples that illustrate the changes in the size, shape, and color of the ball in different frames of a video sequence

4.3 Ball

The most important events in soccer are determined by the positions of the ball throughout the matches. Consequently, its automatic detection is of great interest. However, such detection presents several challenges: in images with a low zoom level, the ball is represented by few pixels; it may be occluded by the players; its detection is difficult when it is on a line of the field of play; its appearance may vary in size, shape, and color (see Fig. 12); and its speed is very changing.

Table 3 shows a summary of the most representative strategies that have been proposed in the last years for detecting and tracking the ball in soccer video sequences. The second column of the table indicates the source used by the strategy developed in each work: the broadcast signal or additional (extra) cameras placed at strategic positions of the stadium. The third and fourth columns indicate the methods applied by each strategy to, respectively, detect and track the ball. It can be observed that there are strategies that focus on exclusively detecting the ball, whereas other ones also track it.

Here, only the works focused on detecting and tracking the ball in soccer matches are analyzed. However, it must be noted that there are other jobs that, despite being focused on other sports (e.g., table tennis or baseball) or being generic (multi-sport ball detection strategies), can also be used in soccer [75].

Table 3 Summary of the main characteristics of the the most representative strategies to detect and track the ball in soccer video sequences

Strategy	Source	Detection	Tracking
2002-D'orazio [33]	1 extra camera	SD	
2003-Yu [172]	Broadcast signal	BGS	TM
2004-D'orazio [34]	1 extra camera	SD	
2004-Tong [152]	Broadcast signal	BGS	ST
2008-Leo [91]	Several extra cameras	SD, KPD	
2008-Pallavi [119]	Broadcast signal	SD	
2009-Ren [131]	8 extra cameras	BGS	KPT
2011-Hossein [64]	Broadcast signal	BGS	TM
2013-Leo [92]	Several extra cameras	SD, KPD	
2016-Kia [81]	Broadcast signal	BGS, SL	
2019-Kamble [74]	6 extra cameras	BGS, SL	KPT
2019-Komorowski [88]	6 extra cameras	SL	

Detection methods: KPD - Key-Point Detection; BGS - Background Subtraction; SL - supervised Learning; SD - Silhouette Detection

Tracking methods: KPT - Key-Point Tracking; TM - Template Matching; ST - Silhouette Tracking

The main aspects of the methods proposed by the strategies in Table 3 are described below.

4.3.1 Detection

A significant amount of the reviewed works [33, 34, 92, 119] detect the silhouette of the ball by using the Circle Hough Transform (CHT) [66]. The CHT typically results in several candidate detections. Therefore, it is necessary to apply some additional stages to discard the false ones. In [34], for example, false detections are discarded using the Wavelet transform and a classifier based on an NN that has been trained to identify the pattern of the drawing of the ball.

Since the ball can take non-circular shapes in some images, instead of using the CHT, some authors propose detecting candidate balls through the analysis of the characteristics (e.g., shape, color, motion, etc.) of the objects classified as foreground after applying a BGS method [64, 81, 152, 172].

Most of the works for detecting the ball are based on SL. In [91, 92], the SIFT detector is used to compare the candidate balls with a set of training images of the ball. The strategy in [81] uses a trained Adaptive Boosting after applying a BGS-based method to detect the moving objects. The strategy in [74] combines a median filtering background modeling to detect foreground objects, with a CNN-based approach for classifying the foreground objects into three classes, i.e. ball, player, and background. In [88], an NN-based strategy for detecting not only the ball but the players is used.

4.3.2 Tracking

After detecting the ball, some strategies include algorithms to track it along the images of the video sequences. Many of them use algorithms based on KPT, typically using the Kalman filter [64, 131, 172]. However, other types of KPT-based approaches have also been proposed, such as the method based on Lucas-Kanade in [81], or the method based on the analysis of the overlapping of the ball bounding boxes in [74]. Alternatively, the algorithm proposed in [152] follows the ball by applying an ST-based method named CONDENSATION [17] that, inspired by the Particle filter, try to improve the ball tracking when the detections are noisy.

5 Game analysis

Some applications focus on the analysis of team tactics and on obtaining individual and global statistics. This information is relevant, not only for the spectators, but also for the referees, coaches and players, since it allows a better understanding of the game, and the study of the tactics of the teams. In addition, it can be used to create training sessions to improve player performance [64]. Most of these applications focus on the analysis of the positions of the players and the ball along the match to provide different indicators capable of describing and improving the understanding of the game [107]: distance traveled by the players of each team, distance from the ball to the goal in a free kick, positions occupied by players throughout the game, outside detection, goal opportunities, etc.

Table 4 summarizes the characteristics of the most representative soccer game analysis strategies proposed in the last years.

Table 4 Summary of the main characteristics of the the most representative strategies for game analysis in soccer

Strategy	Source	Aim
2003-Yu [171]	Broadcast signal	BP
2005-Yu [170]	Broadcast signal	BP
2006-Hashimoto [58]	16 extra cameras	OD
2006-Kang [76]	Simulated data	PP
2007-Barros [10]	4 extra cameras	PP
2007-Zhu [175]	Broadcast signal, webcasting	TA
2008-Zhu [176]	Broadcast signal, webcasting	TA
2009-D'orazio [37]	6 extra cameras	OD
2009-Lago [89]	8 extra cameras	BP
2012-Sampaio [136]	Commercial tool	TA
2014-Bialkowski [15]	Commercial tool	TA
2016-Bialkowski [14]	Commercial tool	TA
2016-Fernandez [44]	Commercial tool	TA
2017-Gerke [49]	Broadcast signal	TA
2017-Link [94]	Commercial tool	BP
2018-McHale [106]	Commercial tool	PP
2018-Memmert [109]	Commercial tool	TA
2018-Muthuraman [110]	Simulated data	OD
2018-Ryoo [133]	Commercial tool	TA
2018-Theagarajan [149]	Broadcast signal	BP
2018-Yang [163]	Commercial tool	PP
2019-Memmert [108]	Commercial tool, 3 extra cameras	TA
2019-Pappalardo [121]	Commercial tool	PP
2019-Sarkar [138]	Broadcast signal	BP
2020-Theagarajan [148]	Broadcast signal	TA

Goals: BP - Ball possession; OD - Offside detection; PP - Player performance; TA - Tactical analysis

The second column shows the input data used by each strategy: broadcast signal, additional (extra) cameras placed at strategical positions in the stadium, webcasting, simulated data, and commercial tools. As stated before, the positions of players and ball play an important role in all game analysis applications [123], since they are typically taken as starting point. Many of the strategies reviewed in this section incorporate some of the tracking algorithms described in Section 4, whereas other ones include their own tracking algorithms (these strategies do not appear in Section 4 because their aims go beyond the tracking of players and ball). Alternatively, a third group of strategies use tracking data obtained from commercial tools developed in the last years. Some of these tools are based on semi-automatic camera systems (they require a supervisor to control and correct data) that are able of providing more accurate positions than automatic systems and, obviously, more complete results than strategies using only the broadcast signal. More recently, GPS-based tools have also been developed, which are able of providing even more accurate results. Moreover, recently, FIFA amended their rules to allow the use of GPS technology in competitive match play [60]. For these reasons, most of the strategies that have been revised, instead of incorporating their own tracking algorithms, make use of these tools. Later on, as these strategies are described, the tools they use will be mentioned.

According to their main aims, strategies for soccer gaming analysis can be classified into four groups (last column in Table 4):

- Ball possession (BP): Determine which team is in possession of the ball along the match.
- Offside detection (OD): Analyze offside events to assess whether they are correct or not.
- Player performance (PP): Evaluate the performance of players by analyzing the number of passes, their covered distance, their speed, their endurance, etc.
- Tactical analysis (TA): Analyze the disposition (tactical position) and behavior of the players throughout the game to determine the team performance, the used tactics, and the type of strategy followed by each team.

Next, the main characteristics of the strategies in each of these four groups are detailed.

5.1 Ball possession

The possession of the ball is a fundamental requirement to reach the goal of the rival team and score goals [26]. Consequently, soccer ball possession has been one of the most commonly investigated soccer performance indicators [102], resulting in several works proposing strategies to automatically determine the player and/or the team that is controlling the ball throughout the match.

To perform an automatic analysis of the ball possession, the strategy in [171] requires as input data the position of the ball and the play/break structure. It applies several methods to identify the frames in which the ball is being touched by players. In each of these frames, the player located closest to the ball is found. Finally, an SVM is applied to determine which of the teams is touching the ball. An improved version of this strategy was proposed later in [170], which integrates the algorithms for detecting and tracking the ball, and to determine the play/break structure. Additionally, it includes a view classification algorithm to determine the touching-player candidates at the touching-points.

A multiple-camera match analysis system is used in [89] to obtain the moving of the players and the ball throughout the matches. Two linear regression models are used to determine the team possession and the time that each team spends in the different zones of the field of play (defensive, middle, and attacking thirds).

The authors of [94], taking as starting point the work in [76], proposed a strategy for detecting both individual and team ball possession. To this end, they use a machine learning approach that, by analyzing not only the main characteristics of the ball motion (direction, speed, and acceleration), but also the distance between the ball and the players, allows determining how much time is the ball in the circle of influence of each player. The positional data they use as input is collected using TRACAB[5], which is a real-time optical tracking system that is able of providing the positions of ball and players using wearable devices.

A CNN-based strategy is proposed in [149] to localize the players in the field of play and identify what player is controlling the ball throughout the match. First, the CNN proposed in [128] is used to detect the soccer players. Then, the DeepSort tracking approach proposed in [157] is used to track the detected players, and a simple histogram-based matching

[5] https://chyronhego.com/products/sports-tracking/tracab-optical-tracking/

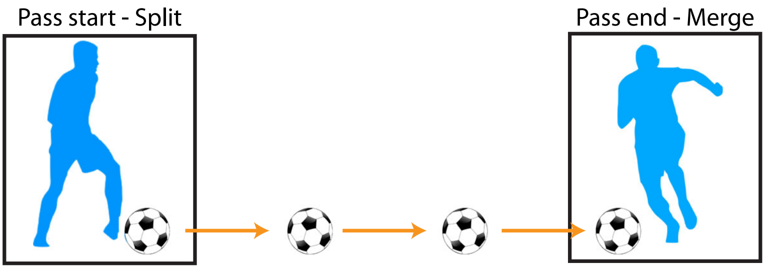


Fig. 13 Example of valid pass in [138]

approach is used to identify the team of each player. Finally, to identify the player who is in control of the ball a CNN is used, which has been trained to classify each player as “a player with the ball” or “player without the ball”.

Another strategy for determining ball possession statistics was recently proposed in [138]. First, this work segments the foreground objects and applies an SVM to classify them among four classes (ball, team A, team B, and other). Then, it uses a minimum-cost network model to detect valid passes made by each team, where the ball and the players represent the nodes of the network and split and merge events are analyzed (see Fig. 13). Finally, ball possession statistics are obtained from the amount of valid passes.

5.2 Offside detection

Offside detection is one of the most important tasks in soccer matches, since an average of 50 offside decisions are made every match and, moreover, they can drastically change the outcome of the game [110]. In addition, several offside judgments are difficult to make for referees [21]. Because of these important reasons, some works that focus on detecting offside events and assess whether they are correct or not have been proposed in the last years.

In [58], 16 fixed cameras are used to detect and track players, and to calculate their world coordinates (see Fig. 14a). The 3D trajectory of the ball and the players involved in the

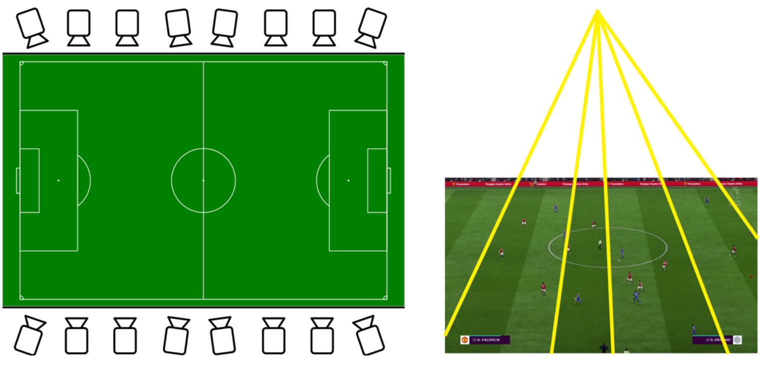


Fig. 14 **a** Setup of the 16 cameras used in [58]. **b** Example of virtual image used in [110] and lines used to determine the vanishing-point

game at each instant are analyzed to make the offside judgments. Similarly, in [37] a 6 fixed camera system is used for automatic offside detection. The position of players and ball is obtained simultaneously with all the cameras, and a multiple view analysis is carried out to determine the players who pass the ball, determining if active offside condition occurs.

In [110], instead of using extra cameras, game-play videos are used with the aim of providing a more fair approximation to real-life soccer scenarios. In this work, the players are first detected and classified in their respective teams. Then, by tracing virtual lines that cut in the vanishing-point corresponding to the transverse lines of the field of play, the offside decision is done (see Fig. 14b).

5.3 Player performance

Not only global statistics (e.g., ball possession) of the game are demanded by audiences and trainers, but also statistics related to the performance of each of the team players (e.g., running distances, number of passes, running speed, etc.). Accordingly, the number of strategies focused on the generation of these types of statistics has experienced a significant increase in recent years [130]. This increasing is largely due to the advances that have been recently made in automatic player tracking technologies, such as the wireless sensors to track player positions. In addition, in many professional soccer teams, semi-automated multi-camera systems (e.g., Prozone, which has been recently acquired by STATS[6]), which are permanently installed at the team stadium, are typically used to quantify both the physical and the technical match performance [154].

In [76], the relationship between the trajectories of the players and the ball is analyzed to evaluate the performance of the players. In this work, different regions are defined taking into account the disposition of the players in the field of play, regarding to the ball position: catchable regions, competing regions, and safe regions. The performance is measured based on the average number of times that a player is in each of these regions. For the experiments, the trajectories of players and ball are obtained from a soccer simulation game.

In [10], four fixed and partially overlapped additional cameras are used to analyze the distance covered by the players along a soccer match. Before the game, several control points are established on the field of play to calibrate the cameras. Then, player detection and tracking are performed by applying two basic automatic procedures. Finally, the spatial positions of the players are filtered using a low-pass filter.

In [106], the performance of the players is measured with the aim of identifying which players are pivotal to each team. A network analysis is used to identify which are those players that are more involved in passing actions and, consequently, are key-players for their teams. Additionally, the probability of the passes of being successful is estimated using a statistical modeling. The tracking dataset used to perform these tasks is provided by a semi-automated multi-camera system (Prozone), that includes several descriptive statistics of the players (e.g., distances covered or number of sprints).

The work proposed in [163] aims to identify the key physical and technical performance variables related to team quality. To do this, several characteristics of the game are analyzed, such as the distances sprinting, the distances covered without possession, the amount of possession in opponent's half, or the amount of entry passes in the penalty area. The input data used by this work is obtained from Amisco Sport Analysis Services [23].

[6] https://www.stats.com/football/

In [121], a strategy for evaluating the performance and ranking the players has also been proposed. This strategy extracts and analyzes features from soccer-logs provided by the commercial tool Wyscout.[7] These soccer-logs describe events using the following data: the player involved in the event, the type of event (e.g., shot, goal, or tackle), the position of the event on the field of play, and the time of the event.

It must be noted that many of the works focused on the analysis of the performance of the players, to be able of measure distances and speeds, require to register the images in a model of the field of play [61, 153]. Consequently, several strategies with this aim have also been proposed in the last years. In [165], a self-calibration method for non-fixed cameras is proposed, which is able of estimating homography matrices between a field of play model and the original acquired images. Such estimation is performed by comparing the crossing points of the field of play lines with a set of 16 possible patterns. In [151], instead of comparing the line crossing points with predefined patterns, once the detected lines have been classified among vertical and horizontal lines, the homography matrices are estimated by analyzing all the possible correspondences between image key-points and model key-points. Other works [95, 166] propose registration strategies that are able of providing successful results not only in images with several line crossing points, but also in images without or with few crossing points (e.g., images showing the central part of the field of play). To this aim, the strategy in [95] includes an ellipse detection algorithm that allows registering the central circle of the field of play. In addition, when neither crossing points, nor the central circle are detected, it is able of registering the images by using a Lucas Kanade optical flow estimator that relates the previously registered images with the current one. Similarly, in [166], a feature tracking-based algorithm is proposed to relate images without crossing points or lines with previously registered images. The strategy in [19] is focused on obtaining optimized homography maps by projecting the detected lines to the field of play model and establishing relations between line points in the original image and in the model. The authors of [62] propose a 3D registration strategy. In this strategy, first, the sets of longitudinal and transverse lines are used to determine two vanishing points, from which it is possible to perform the field of play registering. Then, an energy minimization in a Markov Random Field (MRF) is applied, which not only improves the registering process, but also allows obtaining an accurate location of both straight and circle lines. Another strategy to perform the automatic registration of the field of play has been recently proposed in [30]. This strategy is based on the detection of the line marks of the field of play (the straight lines and the central circle line). To avoid false line mark detections, it includes a pre-processing algorithm that allows discarding unwanted edge data. Then, a probabilistic decision tree is used to identify the most probable classification for the set of all detected lines. Finally, a three-step validation stage is applied to determine whether the registration is correct. Figure 15 shows some results obtained with this strategy.

5.4 Tactical analysis

Choosing a suitable tactic is crucial when preparing the game. The tactic specifies the way in which teams manage space (where on the pitch actions to win take place, or which areas of the field of play tend to be occupied by each team), time (frequency and duration of events, such as ball possession), and individual actions (kinds of actions performed by each player, such as dribbles, passes, and crosses) to win a game [47]. The analysis of the tactics

[7] https://wyscout.com/

Fig. 15 Some results obtained with the strategy in [30]. **a** Original images and line marks detected. **b** Projection of the images to a field of play model and points used to perform the registration

can be carried out at the individual level (i.e., only one player), at the level of the group of players, or at the level of the whole team [130]. On the one hand, the individual tactics are those describing the one-on-one events throughout offensive and defensive play, both with the ball and without it. On the other hand, the group tactics describe the cooperation between sub groups within a team. Finally, the team tactics refer to the playing philosophy of each team. The main characteristics of the most relevant works that have proposed tactical analysis strategies in the last years are described below.

In [175, 176], a strategy to extract tactic information from goal events was proposed. First, the goal events are detected by analyzing the broadcast video and webcasting data. Then, the trajectories of players and ball are obtained by employing a multi-object detection and tracking algorithm. Finally, the spatio-temporal interaction among the players and the ball is analyzed.

In [44], a study to define and categorize different styles of play in elite soccer using factor analysis was presented. To this aim, the multicamera match analysis system provided by Amisco Pro [22] is used to monitor matches from the Spanish and English leagues. The factor analysis is performed from a total of nineteen performance indicators (e.g., crosses, possession of the ball, and shots): fourteen of these indicators are related to attacking play and the remaining five are focused on describing aspects of defensive play.

More recently, in [109], a model to connect Big Data and match analysis from new performance indicators was introduced. All calculations are based solely on the player and ball positions that are automatically obtained with the software tool SOCCER [123]. The main key performance indicators that are used in this work are: space control (the space of the field of play that is controlled by each team is quantified by using Voronoi diagrams), outplayed opponents (effectiveness of the vertical passes of each team is measured), and pressing index (the average speed of all players with respect to the ball is used to measure the pressing behavior of the teams).

A system for determining the performance of the teams was proposed in [133], which is based on analyzing several player data and comparing the team performance changes before and after player transfers.

Among works for soccer tactical analysis, there are some ones focused on the analysis of the identification of the roles of each player throughout the matches. To this aim, a strategy

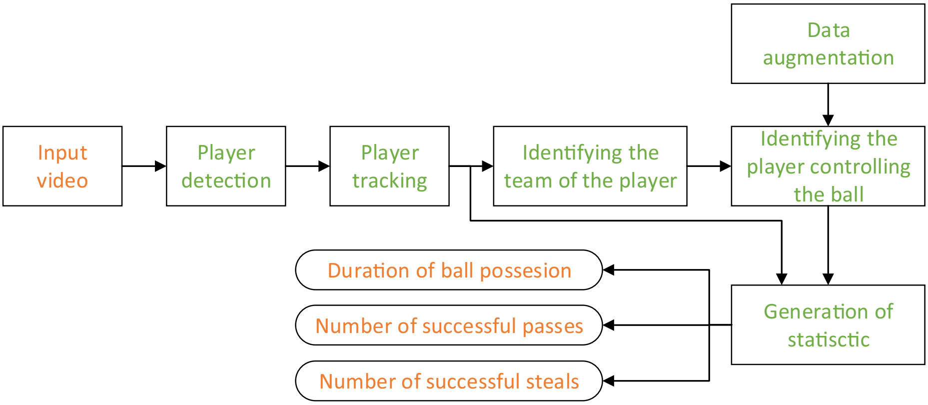


Fig. 16 Overall architecture of the system proposed in [148] for generating tactical analysis statistics. In orange the input and output data. In green the states of the system

that is able of identifying roles from spatial constellations was proposed in [49]. This strategy models such constellations as histograms over the spatial positions of all the players of a team. Additionally, it includes a convolutional NN-based method for recognizing the jersey numbers and mapping the players with their corresponding numbers. In contrast to other strategies, this one is able of avoiding to categorically associate each player with one single tactical role (central defender, winger, forward, etc.). Therefore, this strategy not only allows analyzing tactics, but it also determines the role of each player. Other works that also analyze individual roles and team tactics are those proposed in [14, 15, 136]. The strategies in these works are based on the idea that strategically, the players of each team are typically distributed in a way that covers the field of play with the minimum overlap between them. To identify the roles of the players and discover the strategies used by both teams, using the positions of the players along the match as input data, they learn the role of each player by applying minimum entropy data partitioning methods. These three works use as starting data the positions provided by commercial tracking tools. In [136], the GPSSports[8] is used, which is a GPS-based tool. On the other hand, in [14] and [15], Prozone is used.

The system proposed in [148] allows generating three tactical statistics for each player: the duration of the ball possession, the amount of successful passes, and the amount of successful steals. First, it detects and tracks the players using CNNs. Then, a Triplet-CNN is used to extract features from the detected players. These features are used for predicting the team of each player and to find out when each player controls the ball. Figure 16 illustrates the overall architecture of this system.

Finally, the strategy proposed in [108] allows the evaluation of the performance of the teams when different formations are used (e.g., 4-2-3-1 or 3-5-2). The performance indicators used in this strategy are based on the analysis of the following dynamical positioning variables: effective player space, player length width ratio, team separateness, space control gain, and pressure passing efficiency. In this work, the input data is obtained by combining a commercial player tracking system (Kinexon[9]) and three video cameras placed around the pitch.

[8] http://gpsports.com/football/
[9] https://kinexon.com/

6 Future directions

The aim of this section is to provide the reader with some ideas on possible future directions in the techniques and applications for video soccer analysis.

As it has been seen throughout the paper, the integration of the strategies to detect events in the strategies for analyzing the game is increasingly common. Consequently, since many game analysis applications require real-time performance, and despite the fact that in Section 3 it was said that event detection is generally done offline, the observed trend suggests that interest in applications that can detect events in real time will increase.

It has also been observed that there is a marked trend in the development of applications using NN-based learning methods, both in event detection strategies and in game analysis applications. This is because thanks to the recent advances in computing (e.g., greater parallel processing capacity), NNs not only provide successful results in many cases, but they also perform close to real-time.

The observed trends also suggest that there is a growing interest in the development of applications based on game analysis and, more specifically, on strategies that focus on tactical analysis. Most such strategies are very dependent on the position of players throughout the matches, which are required as input data. Since it is very complex to deal with typical player detection and tracking challenges (e.g., occlusions) with only the video cameras used to broadcast the match, the trend is to obtain positioning data through commercial tools such as those mentioned in Section 5, which are based on several extra video cameras (e.g., Prozone or Amisco Pro) or wearable devices (e.g., TRACAB or GPSSport). However, these commercial tools are typically expensive and only high-budget leagues can afford them. On the other hand, the development of strategies based on the cameras used to broadcast the matches continue been of interest in the case of leagues with a more reduced budget. Therefore, it seems reasonable that detection and tracking strategies based on the broadcast signal (such as most of those described in Section 4) continue to be proposed.

Regarding the strategies for detecting and tracking the ball, manufacturers are often reluctant to incorporate sensors. Therefore vision-based methods will continue to be necessary.

7 Conclusions

In this paper, the most meaningful techniques and applications proposed throughout the last two decades to analyze soccer video sequences have been surveyed. The existing strategies have been grouped into three categories according to their objectives and needs: strategies for event detection, strategies for detecting and tracking players and/or ball, and strategies for the analysis of the game. The main characteristics of the strategies in each group have been analyzed, detailing the data they use as starting point, their specific objectives, the analysis techniques they apply to achieve such objectives, their strengths, and their weaknesses.

In the strategies for detecting events, four main tasks have been identified: recognize specific audiovisual features (e.g., camera movements or increasing audio level); detect replays; detect the intervals in which the game is being played; and relate the content of the videos with external sources of non-audiovisual information (e.g., social networks). All these tasks are typically done after the game (i.e., offline). Therefore, they do not usually have run-time of computational restrictions.

Regarding detection and tracking strategies, it is possible to find several works that focus on detecting players, but also some works that try to detect the ball. Furthermore, it has been found that a significant number of works are not limited to detecting the players or the ball, but also tracking the detected objects. The main challenges identified in the case of strategies for detecting/tracking players are the following: occlusions between players; abrupt movements of the camera; illumination changes; lack of resolution in very distant players; and blurring of players that are moving. In the case of the strategies for detecting/tracking the ball, the identified challenges are: images where the ball is represented by very few pixels; occlusions due to players; misdetections when the ball is on a line mark of the field of play; and fast changes in the size, shape, color, and speed of the ball.

Finally, concerning the game analysis applications, it has been found that they are focused on the analysis of tactics used by teams and in obtaining individual and global statistics that can be relevant for spectators, referees, coaches and players. To perform such analyses and obtain statistics, most of the reviewed strategies require the position of the players and the ball along the match. So, they either include their own detection and tracking strategies or use commercial tools. According to their aims, these strategies have been classified into four groups: applications focused on the analysis of the possession of the ball; algorithms to analyze offside events; strategies to evaluate the performance of the players by analyzing, for example, their covered distance, their speed, or their endurance; and algorithms to analyze the tactical position and behavior of the players throughout the game to determine the performance of the team.

References

1. Al-Ali A, Almaadeed S (2017) A review on soccer player tracking techniques based on extracted features. In: IEEE International conference on information and communication technology and accessibility (ICTA), IEEE, pp 1–6
2. Ali MN, Abdullah-alWadud M, Lee SL (2012) An efficient algorithm for detection of soccer ball and players. In: Proc 16th ASTL Control and Networking, pp 16
3. Arce GR (2005) Nonlinear signal processing: a statistical approach. John Wiley & Sons
4. Assfalg J, Bertini M, Colombo C, Del Bimbo A, Nunziati W (2003) Semantic annotation of soccer videos: automatic highlights identification. Comput Vis Image Underst 92(2):285–305
5. Assfalg J, Bertini M, Del Bimbo A, Nunziati W, Pala P (2002) Soccer highlights detection and recognition using hmms. In: IEEE International conference on multimedia and expo, vol 1, IEEE, pp 825–828
6. Athanesious J, Suresh P (2013) Implementation and comparison of kernel and silhouette based object tracking. International Journal of Advanced Research in Computer Engineering & Technology, pp 1298–1303
7. Athanesious JJ, Suresh P (2012) Systematic survey on object tracking methods in video. Int J Adv Res Comput Eng Technol (IJARCET) 1(8):242–247
8. Bai X, Zhang T, Wang C, El-Latif AAA, Niu X (2013) A fully automatic player detection method based on one-class svm. IEICE Trans Inf Syst 96(2):387–391
9. Bandyopadhyay K (2017) Legacies of great men in world soccer: heroes, Icons, Legends. Routledge
10. Barros RM, Misuta MS, Menezes RP, Figueroa PJ, Moura FA, Cunha SA, Anido R, Leite NJ (2007) Analysis of the distances covered by first division brazilian soccer players obtained with an automatic tracking method. J Sports Sci Medic 6(2):233
11. Beetz M, Gedikli S, Bandouch J, Kirchlechner B, von Hoyningen-Huene N, Perzylo AC (2007) Visually tracking football games based on tv broadcasts. In: International joint conference on artificial intelligence (IJCAI), pp 2066–2071
12. Beetz M, Hoyningen-huene NV, Bandouch J, Kirchlechner B, Gedikli S, Maldonado A (2006) Camera-based observation of football games for analyzing multi-agent activities. In: International joint conference on autonomous agents and multiagent systems, ACM, pp 42–49

13. Berjón D, Cuevas C, Moran F, García N (2018) Real-time nonparametric background subtraction with tracking-based foreground update. Pattern Recogn 74:156–170
14. Bialkowski A, Lucey P, Carr P, Matthews I, Sridharan S, Fookes C (2016) Discovering team structures in soccer from spatiotemporal data. IEEE Trans Knowl Data Eng 28(10):2596–2605
15. Bialkowski A, Lucey P, Carr P, Yue Y, Sridharan S, Matthews I (2014) Identifying team style in soccer using formations learned from spatiotemporal tracking data. In: IEEE International conference on data mining workshop (ICDMW), IEEE, pp 9–14
16. Bialkowski A, Lucey P, Carr P, Yue Y, Sridharan S, Matthews I (2014) Large-scale analysis of soccer matches using spatiotemporal tracking data. In: IEEE International conference on data mining (ICDM), IEEE, pp 725–730
17. Blake A, Isard M (1997) The condensation algorithm-conditional density propagation and applications to visual tracking. In: Advances in neural information processing systems, pp 361–367
18. Boyar M, Alan Ö, Akpinar S, Sabuncu O, Çiçekli NK, Alpaslan FN (2010) Event boundary detection using audio-visual features and web-casting texts with imprecise time information. In: IEEE International conference on multimedia and expo (ICME), IEEE, pp 578–583
19. Bozorgpour A, Fotouhi M, Kasaei S (2015) Robust homography optimization in soccer scenes. In: The 23rd iranian conference on electrical engineering, ICEE, pp 787–792
20. Bracewell RN, Bracewell RN (1986) The Fourier transform and its applications, vol 31999. McGraw-Hill, New York
21. del Campo VL, Fariñas AC, Márquez FJD, Martín JM (2018) The influence of refereeing experiences judging offside actions in football. Psychology of Sport and Exercise 37:139–145
22. Carling C, Bloomfield J, Nelsen L, Reilly T (2008) The role of motion analysis in elite soccer. Sports Medicine 38(10):839–862
23. Castellano J, Alvarez-Pastor D, Bradley PS (2014) Evaluation of research using computerised tracking systems (amisco®and prozone®) to analyse physical performance in elite soccer: a systematic review. Sports Medicine 44(5):701–712
24. Chen SC, Shyu ML, Chen M, Zhang C (2004) A decision tree-based multimodal data mining framework for soccer goal detection. In: IEEE International conference on multimedia and expo, vol 1, IEEE, pp 265–268
25. Cheng Y (1995) Mean shift, mode seeking, and clustering. IEEE Trans Patt Anal Mach Intell 17(8):790–799
26. Collet C (2013) The possession game? a comparative analysis of ball retention and team success in european and international football, 2007–2010. J Sports Scie 31(2):123–136
27. Cortes C, Vapnik V (1995) Support vector machine. Mach Learn 20(3):273–297
28. Cuevas C, García N, Salgado L (2008) A new strategy based on adaptive mixture of gaussians for real-time moving objects segmentation. In: Real-time image processing 2008, vol 6811, International society for optics and photonics, pp 681111
29. Cuevas C, Martínez R, García N (2016) Detection of stationary foreground objects: A survey. Comput Vis Image Understanding 152:41–57
30. Cuevas C, Quilón D, García N (2020) Automatic soccer field of play registration. Pattern Recogn 103(107):278
31. Daubechies I (1990) The wavelet transform, time-frequency localization and signal analysis. IEEE Trans Inf Theory 36(5):961–1005
32. Ding XF, Miao YJ, Bu F, Sun LF, Yang SQ (2008) Highlight detection in soccer video using web-casting text. In: IEEE 10Th workshop on multimedia signal processing, IEEE, pp 604–609
33. D'Orazio T, Ancona N, Cicirelli G, Nitti M (2002) A ball detection algorithm for real soccer image sequence. In: IEEE International conference on pattern recognition, vol 1, IEEE, pp 210–213
34. D'Orazio T, Guaragnella C, Leo M, Distante A (2004) A new algorithm for ball recognition using circle hough transform and neural classifier. Patt Recogn 37(3):393–408
35. D'Orazio T, Leo M (2010) A review of vision-based systems for soccer video analysis. Patt Recogn 43(8):2911–2926
36. D'Orazio T, Leo M, Mosca N, Spagnolo P, Mazzeo PL (2009) A semi-automatic system for ground truth generation of soccer video sequences. In: IEEE International conference on advanced video and signal based surveillance, IEEE, pp 559–564
37. D'Orazio T, Leo M, Spagnolo P, Mazzeo PL, Mosca N, Nitti M, Distante A (2009) An investigation into the feasibility of real-time soccer offside detection from a multiple camera system. IEEE Trans Circ Syst Vid Technol 19(12):1804–1818
38. D'Orazio T, Leo M, Spagnolo P, Nitti M, Mosca N, Distante A (2009) A visual system for real time detection of goal events during soccer matches. Comput Vis Image Underst 113(5):622–632

39. Ekin A, Tekalp AM, Mehrotra R (2003) Automatic soccer video analysis and summarization. IEEE Trans Image process 12(7):796–807
40. Eldib MY, Zaid BSA, Zawbaa HM, El-Zahar M, El-Saban M (2009) Soccer video summarization using enhanced logo detection. In: IEEE International conference on image processing (ICIP), IEEE, pp 4345–4348
41. Esmin AA, Júnior RS, Santos WS, Botaro CO, Nobre TP (2014) Real-time summarization of scheduled soccer games from twitter stream. In: International conference on applications of natural language to data bases/information systems, Springer, pp 220–223
42. Fakhar B, Kanan HR, Behrad A (2019) Event detection in soccer videos using unsupervised learning of spatio-temporal features based on pooled spatial pyramid model. Multimedia Tools and Applications 78(12):16,995–17,025
43. Fani M, Yazdi M, Clausi DA, Wong A (2017) Soccer video structure analysis by parallel feature fusion network and hidden-to-observable transferring markov model. IEEE Access 5:27,322–27,336
44. Fernandez-Navarro J, Fradua L, Zubillaga A, Ford PR, McRobert AP (2016) Attacking and defensive styles of play in soccer: analysis of spanish and english elite teams. J Sports Sci 34(24):2195–2204
45. Figueroa PJ, Leite NJ, Barros RM (2006) Background recovering in outdoor image sequences: an example of soccer players segmentation. Image Vis Comput 24(4):363–374
46. Figueroa PJ, Leite NJ, Barros RM (2006) Tracking soccer players aiming their kinematical motion analysis. Comput Vis Image Underst 101(2):122–135
47. Fradua L, Zubillaga A, Caro Ó, Iván Fernández-garcía Á, Ruiz-Ruiz C, Tenga A (2013) Designing small-sided games for training tactical aspects in soccer: Extrapolating pitch sizes from full-size professional matches. J Sports Sci 31(6):573–581
48. Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 55(1):119–139
49. Gerke S, Linnemann A, Müller K (2017) Soccer player recognition using spatial constellation features and jersey number recognition. Comput Vis Image Underst 159:105–115
50. Gerke S, Muller K, Schafer R (2015) Soccer jersey number recognition using convolutional neural networks. In: IEEE International conference on computer vision workshops, pp 17–24
51. Giancola S, Amine M, Dghaily T, Ghanem B (2018) Soccernet:, A scalable dataset for action spotting in soccer videos. arXiv:1804.04527
52. Habtemariam B, Tharmarasa R, Thayaparan T, Mallick M, Kirubarajan T (2013) A multiple-detection joint probabilistic data association filter. IEEE J Select Top Signal Process 7(3):461–471
53. Halin AA, Rajeswari M, Abbasnejad M et al (2013) Soccer event detection via collaborative multimodal feature analysis and candidate ranking. Int Arab J Inf Technol 10(5):493–502
54. Halin AA, Rajeswari M, Ramachandram D (2009) Shot view classification for playfield-based sports video. In: 2009 IEEE International conference on signal and image processing applications, IEEE, pp 410–414
55. Harris C, Stephens M (1988) A combined corner and edge detector. In: Alvey vision conference, Citeseer, pp 147–151
56. Hartigan JA, Wong MA (1979) Algorithm as 136: a k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics) 28(1):100–108
57. Hasan M, Orgun MA, Schwitter R (2018) A survey on real-time event detection from the twitter data stream. J Inf Sci 44(4):443–463
58. Hashimoto S, Ozawa S (2006) A system for automatic judgment of offsides in soccer games. In: IEEE International conference on multimedia and expo, IEEE, pp 1889–1892
59. Hayet JB, Mathes T, Czyz J, Piater J, Verly J, Macq B (2005) A modular multi-camera framework for team sports tracking. In: IEEE Conference on advanced video and signal based surveillance, IEEE, pp 493–498
60. Hennessy L, Jeffreys I (2018) The current use of gps, its potential, and limitations in soccer. Strength and Conditioning Journal 40(3):83–94
61. Homayounfar N, Fidler S, Urtasun R (2016) Soccer field localization from a single image. arXiv:1604.02715
62. Homayounfar N, Fidler S, Urtasun R (2017) Sports field localization via deep structured models. In: IEEE Conference on computer vision and pattern recognition (CVPR), IEEE, pp 4012–4020
63. Hong Y, Ling C, Ye Z (2018) End-to-end soccer video scene and event classification with deep transfer learning. In: International conference on intelligent systems and computer vision (ISCV), IEEE, pp 1–4
64. Hossein-Khani J, Soltanian-Zadeh H, Kamarei M, Staadt O (2011) Ball detection with the aim of corner event detection in soccer video. In: IEEE International symposium on parallel and distributed processing with applications workshops (ISPAW), IEEE, pp 147–152

65. Hosseini MS, Eftekhari-Moghadam AM (2013) Fuzzy rule-based reasoning approach for event detection and annotation of broadcast soccer video. Appl Soft Comput 13(2):846–866
66. Hough PV (1962) Method and means for recognizing complex patterns. US Patent 3,069,654
67. Huang CL, Shih HC, Chao CY (2006) Semantic analysis of soccer video using dynamic bayesian network. IEEE Trans Multimedia 8(4):749–760
68. Illingworth J, Kittler J (1988) A survey of the hough transform. Comput Vis Graph Image process 44(1):87–116
69. Iwase S, Saito H (2004) Parallel tracking of all soccer players by integrating detected positions in multiple view images. In: IEEE International conference on pattern recognition, vol 4, IEEE, pp 751–754
70. Jai-Andaloussi S, Mohamed A, Madrane N, Sekkaki A (2014) Soccer video summarization using video content analysis and social media streams. In: IEEE/ACM International symposium on big data computing, IEEE, pp 1–7
71. Jiang H, Lu Y, Xue J (2016) Automatic soccer video event detection based on a deep neural network combined cnn and rnn. In: IEEE 28Th international conference on tools with artificial intelligence (ICTAI), IEEE, pp 490–494
72. Julier SJ, Uhlmann JK (1997) New extension of the kalman filter to nonlinear systems. In: Signal processing, sensor fusion, and target recognition VI, vol 3068, International Society for Optics and Photonics, pp 182–194
73. Kaewbuadee P, Chomphuwiset P (2016) Playfield detection in soccer images using prior information. In: International joint conference on computer science and software engineering (JCSSE), IEEE, pp 1–6
74. Kamble P, Keskar A, Bhurchandi K (2019) A deep learning ball tracking system in soccer videos. Opto-Electronics Rev 27(1):58–69
75. Kamble PR, Keskar AG, Bhurchandi KM (2017) Ball tracking in sports: a survey. Artif Intell Rev, pp 1–51
76. Kang CH, Hwang JR, Li KJ (2006) Trajectory analysis for soccer players. In; IEEE International conference on data mining workshops, IEEE, pp 377–381
77. Kass M, Witkin A, Terzopoulos D (1988) Snakes: Active contour models. Int J Comput Vis 1(4):321–331
78. Kataoka H, Hashimoto K, Aoki Y (2011) Player position estimation by monocular camera for soccer video analysis. In: IEEE SICE annual conference, IEEE, pp 1985–1990
79. Khan YS, Pawar S (2015) Video summarization: survey on event detection and summarization in soccer videos. International Journal of Advanced Computer Science and Applications 6(11):256–259
80. Khatoonabadi SH, Rahmati M (2009) Automatic soccer players tracking in goal scenes by camera motion elimination. Image Vis Comput 27(4):469–479
81. Kia M (2016) Ball automatic detection and tracking in long shot views. Int J Comput Sci Netw Secur (IJCSNS) 16(6):1
82. Kim C, Hwang JN (2002) Fast and automatic video object segmentation and tracking for content-based applications. IEEE Trans Circ Syst Vid Technol 12(2):122–129
83. Kim H, Nam S, Kim J (2003) Player segmentation evaluation for trajectory estimation in soccer games. In: Conference on image and vision computing, Massey university, pp 159–162
84. Kim JY, Kim TY (2009) Soccer ball tracking using dynamic kalman filter with velocity control. In: IEEE International conference on computer graphics, imaging and visualization, IEEE, pp 367–374
85. Kim W (2019) Multiple object tracking in soccer videos using topographic surface analysis. J Vis Commun Image Represent 65(102):683
86. Kim W, Moon SW, Lee J, Nam DW, Jung C (2018) Multiple player tracking in soccer videos: an adaptive multiscale sampling approach. Multimedia Systems, pp 1–13
87. Kolekar MH, Sengupta S (2015) Bayesian network-based customized highlight generation for broadcast soccer videos. IEEE Trans Broadcast 61(2):195–209
88. Komorowski J, Kurzejamski G, Sarwas G (2019) Footandball:, Integrated player and ball detector. arXiv:1912.05445
89. Lago C (2009) The influence of match location, quality of opposition, and match status on possession strategies in professional association football. Journal of Sports Sciences 27(13):1463–1469
90. Lee J, Nam DW, Moon S, Lee J, Yoo W (2017) Soccer event recognition technique based on pattern matching. In: IEEE Federated conference on computer science and information systems, IEEE, pp 643–646
91. Leo M, D’Orazio T, Spagnolo P, Mazzeo P, Distante A (2008) Sift based ball recognition in soccer images. Image and Signal Processing, pp 263–272
92. Leo M, Mazzeo PL, Nitti M, Spagnolo P (2013) Accurate ball detection in soccer images using probabilistic analysis of salient regions. Machine Vision and Applications 24(8):1561–1574

93. Leonardi R, Migliorati P, Prandini M (2004) Semantic indexing of soccer audio-visual sequences: a multimodal approach based on controlled markov chains. IEEE Trans Circ Syst Video Technol 14(5):634–643
94. Link D, Hoernig M (2017) Individual ball possession in soccer. PloS One 12(7):e0179,953
95. Linnemann A, Gerke S, Kriener S, Ndjiki-Nya P (2013) Temporally consistent soccer field registration. In: IEEE International conference on image processing (ICIP), IEEE, pp 1316–1320
96. Liu HY, He T (2009) Integrating multiple feature fusion for semantic event detection in soccer video. In: IEEE International joint conference on artificial intelligence, IEEE, pp 128–131
97. Liu J, Tong X, Li W, Wang T, Zhang Y, Wang H (2009) Automatic player detection, labeling and tracking in broadcast soccer video. Pattern Recogn Lett 30(2):103–113
98. Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE (2017) A survey of deep neural network architectures and their applications. Neurocomputing 234:11–26
99. Lowe DG (1999) Object recognition from local scale-invariant features. In: IEEE International conference on computer vision, vol 2, IEEE, pp 1150–1157
100. Lu K, Chen J, Little JJ, He H (2018) Lightweight convolutional neural networks for player detection and classification. Computer Vision and Image Understanding
101. Lucas BD, Kanade T (1981) An iterative image registration technique with an application to stereo vision. In: International joint conference on artificial intelligence, pp 674–679
102. Mackenzie R, Cushion C (2013) Performance analysis in football: a critical review and implications for future research. Journal of Sports Sciences 31(6):639–676
103. Manafifard M, Ebadi H, Moghaddam HA (2017) Multi-player detection in soccer broadcast videos using a blob-guided particle swarm optimization method. Multimedia Tools and Applications 76(10):12,251–12,280
104. Manafifard M, Ebadi H, Moghaddam HA (2017) A survey on player tracking in soccer videos. Comput Vis Image Underst 159:19–46
105. Martín R, Martínez JM (2014) A semi-supervised system for players detection and tracking in multi-camera soccer videos. Multimedia tools and applications 73(3):1617–1642
106. McHale IG, Relton SD (2018) Identifying key players in soccer teams using network analysis and pass difficulty. Eur J Oper Res 268(1):339–347
107. Memmert D, Lemmink KA, Sampaio J (2017) Current approaches to tactical performance analyses in soccer using position data. Sports Med 47(1):1–10
108. Memmert D, Raabe D, Schwab S, Rein R (2019) A tactical comparison of the 4-2-3-1 and 3-5-2 formation in soccer: A theory-oriented, experimental approach based on positional data in an 11 vs. 11 game set-up. PloS One 14(1) e0210191. https://doi.org/10.1371/journal.pone.0210191
109. Memmert D, Rein R (2018) Match analysis, big data and tactics: Current trends in elite soccer. Deutsche Zeitschrift für Sportmedizin 69(3):65–71
110. Muthuraman K, Joshi P, Raman SK (2018) Vision based dynamic offside line marker for soccer games. arXiv:1804.06438
111. Naemura M, Fukuda A, Mizutani Y, Izumi Y, Tanaka Y, Enami K (2000) Morphological segmentation of sport scenes using color information. IEEE Trans Broadcast 46(3):181–188
112. Nguyen N, Yoshitaka A (2012) Shot type and replay detection for soccer video parsing. In: 2012 IEEE International symposium on multimedia, IEEE, pp 344–347
113. Nguyen N, Yoshitaka A (2014) Soccer video summarization based on cinematography and motion analysis. In: IEEE 16Th international workshop on multimedia signal processing (MMSP), IEEE, pp 1–6
114. Nichols J, Mahmud J, Drews C (2012) Summarizing sporting events using twitter. In: ACM International conference on intelligent user interfaces, ACM, pp 189–198
115. Nieto M, Cuevas C, Salgado L (2009) Measurement-based reclustering for multiple object tracking with particle filters. In: 2009 16Th IEEE international conference on image processing (ICIP), IEEE, pp 4097–4100
116. Nummiaro K, Koller-Meier E, Van Gool L (2003) An adaptive color-based particle filter. Image and Vision Computing 21(1):99–110
117. Nunez JR, Facon J, de Souza Brito A (2008) Soccer video segmentation: referee and player detection. In: IEEE International conference on systems, signals and image processing, IEEE, pp 279–282
118. Oh S, Russell S, Sastry S (2009) Markov chain monte carlo data association for multi-target tracking. IEEE Trans Autom Control 54(3):481–497
119. Pallavi V, Mukherjee J, Majumdar AK, Sural S (2008) Ball detection from broadcast soccer videos using static and dynamic features. J Vis Commun Image Represent 19(7):426–436

120. Pan H, Van Beek P, Sezan MI (2001) Detection of slow-motion replay segments in sports video for highlights generation. In: IEEE International conference on acoustics, speech, and signal processing, vol 3, IEEE, pp 1649–1652
121. Pappalardo L, Cintia P, Ferragina P, Massucco E, Pedreschi D, Giannotti F (2019) Playerank: data-driven performance evaluation and player ranking in soccer via a machine learning approach. ACM Trans Intel Syst Technol (TIST) 10(5):1–27
122. Patel B, Patel N (2012) Motion detection based on multi frame video under surveillance system. Int J Comput Sci Netw Secur (IJCSNS) 12(3):100
123. Perl J, Grunz A, Memmert D (2013) Tactics analysis in soccer–an advanced approach. Int J Comput Sci Sport 12(1):33–44
124. Powell M (1965) A method for minimizing a sum of squares of non-linear functions without calculating derivatives. The Comput J 7(4):303–307
125. Quilón D, Mohedano R, Cuevas C, García N (2015) Unsupervised high-quality soccer field segmentation. In: IEEE International symposium on consumer electronics (ISCE), IEEE, pp 1–2
126. Raghuram M, Chavan NR, Koolagudi SG, Ramteke PB (2016) Efficient audio segmentation in soccer videos. In: IEEE Canadian conference on electrical and computer engineering (CCECE), IEEE, pp 1–4
127. Rao U, Pati UC (2015) A novel algorithm for detection of soccer ball and player. In: IEEE International conference on communications and signal processing (ICCSP), IEEE, pp 0344–0348
128. Redmon J, Farhadi A (2017) Yolo9000: Better, faster, stronger. In: IEEE Conference on computer vision and pattern recognition (CVPR), IEEE, pp 6517–6525
129. Rehman A, Saba T (2014) Features extraction for soccer video semantic analysis: current achievements and remaining issues. Artif Intell Rev, pp 1–11
130. Rein R, Memmert D (2016) Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. SpringerPlus 5(1):1410
131. Ren J, Orwell J, Jones GA, Xu M (2009) Tracking the soccer ball using multiple fixed cameras. Comput Vis Image Underst 113(5):633–642
132. Roerdink JB, Meijster A (2000) The watershed transform: Definitions, algorithms and parallelization strategies. Fundamenta Informaticae 41(1, 2):187–228
133. Ryoo M, Kim N, Park K (2018) Visual analysis of soccer players and a team. Multimedia Tools and Applications 77(12):15,603–15,623
134. Sabirin H, Sankoh H, Naito S (2015) Automatic soccer player tracking in single camera with robust occlusion handling using attribute matching. IEICE Trans Inform Syst 98(8):1580–1588
135. Sadlier DA, O'Connor NE (2005) Event detection in field sports video using audio-visual features and a support vector machine. IEEE Transactions on Circuits and Systems for Video Technology 15(10):1225–1233
136. Sampaio J, Maçãs V (2012) Measuring tactical behaviour in football. Int J Sports Med 33(05):395–401
137. Saraogi H, Sharma RA, Kumar V (2016) Event recognition in broadcast soccer videos. In: Indian conference on computer vision, graphics and image processing, ACM, pp 14
138. Sarkar S, Chakrabarti A, Prasad Mukherjee D (2019) Generation of ball possession statistics in soccer using minimum-cost flow network. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 0–0
139. Sigari MH, Soltanian-Zadeh H, Kiani V, Pourreza AR (2015) Counterattack detection in broadcast soccer videos using camera motion estimation. in: IEEE International symposium on artificial intelligence and signal processing (AISP), IEEE, pp 101–106
140. Sigari MH, Soltanianzadeh H, Pourreza HR (2015) Fast highlight detection and scoring for broadcast soccer video summarization using on-demand feature extraction and fuzzy inference. Int J Compute Graph 6(1):13–36
141. da Silva R, Dahmen SR (2014) Universality in the distance between two teams in a football tournament. Physica A:, Statistical Mechanics and its Applications 398:56–64
142. Song W, Hagras H (2017) A type-2 fuzzy logic system for event detection in soccer videos. In: 2017 IEEE international conference on fuzzy systems (FUZZ-IEEE), IEEE, pp 1–6
143. Spagnolo P, Mazzeo PL, Leo M, D'Orazio T (2007) Unsupervised algorithms for segmentation and clustering applied to soccer players classification. In: International conference on signal processing and multimedia applications, pp 129–134
144. Sukhwani M, Kothari R (2017) A parameterized approach to personalized variable length summarization of soccer matches. arXiv:1706.09193
145. Sverrisson S, Grancharov V, Pobloth H (2019) Real-time tracking-by-detection in broadcast sports videos. In: Scandinavian conference on image analysis, Springer, pp 399–411
146. Sze V, Chen YH, Yang TJ, Emer JS (2017) Efficient processing of deep neural networks: a tutorial and survey. Proc IEEE 105(12):2295–2329

147. Tavassolipour M, Karimian M, Kasaei S (2014) Event detection and summarization in soccer videos using bayesian network and copula. IEEE Trans Circ Syst Video Technol 24(2):291–304
148. Theagarajan R, Bhanu B (2020) An automated system for generating tactical performance statistics for individual soccer players from videos. IEEE Transactions on Circuits and Systems for Video Technology
149. Theagarajan R, Pala F, Zhang X, Bhanu B (2018) Soccer: Who has the ball? generating visual analytics and player statistics. In: IEEE Conference on computer vision and pattern recognition workshops, pp 1749–1757
150. Tjondronegoro DW, Chen YPP (2010) Knowledge-discounted event detection in sports video. IEEE Trans Syst, Man, and Cybern-Part A:, Syst Humans 40(5):1009–1024
151. Tong X, Liu J, Wang T, Zhang Y (2011) Automatic player labeling, tracking and field registration and trajectory mapping in broadcast soccer video. ACM Trans Intel Syst Technol (TIST) 2(2):15
152. Tong XF, Lu HQ, Liu QS (2004) An effective and fast soccer ball detection and tracking method. In: IEEE International conference on pattern recognition, vol 4, IEEE, pp 795–798
153. Tsai PS, Meijome T, Austin PG (2007) Scout: a game speed analysis and tracking system. Mach Vis Appl 18(5):289–299
154. Varley MC, Gregson W, McMillan K, Bonanno D, Stafford K, Modonutti M, Di Salvo V (2017) Physical and technical performance of elite youth soccer players during international tournaments: Influence of playing position and team success and opponent quality. Science and Medicine in Football 1(1):18–29
155. Wan K, Xu C (2004) Efficient multimodal features for automatic soccer highlight generation. In: IEEE International conference on pattern recognition, vol 3, IEEE, pp 973–976
156. Wang Z, Yu J, He Y (2017) Soccer video event annotation by synchronization of attack–defense clips and match reports with coarse-grained time information. IEEE Trans Circ Syst Video Technol 27(5):1104–1117
157. Wojke N, Bewley A, Paulus D (2017) Simple online and realtime tracking with a deep association metric. In: IEEE International conference on image processing (ICIP), IEEE, pp 3645–3649
158. Xie L, Xu P, Chang SF, Divakaran A, Sun H (2004) Structure analysis of soccer video with domain knowledge and hidden markov models. Pattern Recogn Lett 25(7):767–775
159. Xu C, Wang J, Wan K, Li Y, Duan L (2006) Live sports event detection based on broadcast video and web-casting text. In: ACM International conference on multimedia, ACM, pp 221–230
160. Xu M, Orwell J, Lowey L, Thirde D (2005) Architecture and algorithms for tracking football players with multiple cameras. IEE Proceedings-Vision, Image and Signal Process 152(2):232–241
161. Xu P, Xie L, Chang SF, Divakaran A, Vetro A, Sun H (2001) Algorithms and system for segmentation and structure analysis in soccer video. In: IEEE International conference on multimedia and expo, vol 1, IEEE, pp 928–931
162. Xu R, Wunsch D (2005) Survey of clustering algorithms. IEEE Trans Neural Netw 16(3):645–678
163. Yang G, Leicht AS, Lago C, Gómez MÁ (2018) Key team physical and technical performance indicators indicative of team quality in the soccer chinese super league. Res Sports Med 26(2):158–167
164. Yang Y, Li D (2017) Robust player detection and tracking in broadcast soccer video based on enhanced particle filter. J Vis Commun Image Represent 46:81–94
165. Yao Q, Kubota A, Kawakita K, Nonaka K, Sankoh H, Naito S (2017) Fast camera self-calibration for synthesizing free viewpoint soccer video. In: IEEE International conference on acoustics, speech and signal processing (ICASSP), IEEE, pp 1612–1616
166. Yao Q, Nonaka K, Sankoh H, Naito S (2016) Robust moving camera calibration for synthesizing free viewpoint soccer video. In: IEEE International conference on image processing (ICIP), IEEE, pp 1185–1189
167. Yilmaz A, Javed O, Shah M (2006) Object tracking: a survey. ACM Computing Surveys (CSUR) 38(4):1–45
168. Yu J, Lei A, Hu Y (2019) Soccer video event detection based on deep learning. In: International conference on multimedia modeling, Springer, pp 377–389
169. Yu J, Lei A, Song Z, Wang T, Cai H, Feng N (2018) Comprehensive dataset of broadcast soccer videos. In: IEEE Conference on multimedia information processing and retrieval (MIPR), IEEE
170. Yu X, Hay TS, Yan X, Chng E (2005) A player-possession acquisition system for broadcast soccer video. In: IEEE International conference on multimedia and expo, IEEE, pp 522–525
171. Yu X, Leong HW, Lim JH, Tian Q, Jiang Z (2003) Team possession analysis for broadcast soccer video based on ball trajectory. In: IEEE International conference on information, communications and signal processing, vol 3, IEEE, pp 1811–1815

172. Yu X, Xu C, Leong HW, Tian Q, Tang Q, Wan KW (2003) Trajectory-based ball detection and tracking with applications to semantic analysis of broadcast soccer video. In: ACM International conference on multimedia, ACM, pp 11–20
173. Zawbaa HM, El-Bendary N, Hassanien AE, Kim Th (2011) Machine learning-based soccer video summarization system. In: Multimedia, computer graphics and broadcasting, Springer, pp 19–28
174. Zhang L, Lu Y, Song G, Zheng H (2018) Rc-cnn: Reverse connected convolutional neural network for accurate player detection. In: Pacific rim international conference on artificial intelligence, Springer, pp 438–446
175. Zhu G, Huang Q, Xu C, Rui Y, Jiang S, Gao W, Yao H (2007) Trajectory based event tactics analysis in broadcast sports video. In: ACM International conference on multimedia, ACM, pp 58–67
176. Zhu G, Xu C, Zhang Y, Huang Q, Lu H (2008) Event tactic analysis based on player and ball trajectory in broadcast video. In: Proceedings of the 2008 international conference on Content-based image and video retrieval, ACM, pp 515–524
177. Zubiaga A, Spina D, Amigó E, Gonzalo J (2012) Towards real-time summarization of scheduled events from twitter streams. In: ACM Conference on hypertext and social media, ACM, pp 319–320