Computer Vision and Image Understanding 159 (2017) 19–46

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

A survey on player tracking in soccer videos

M. Manafifard[a,*], H. Ebadi[b], H. Abrishami Moghaddam[c]

[a] Dept. of Photogrammetry and Remote Sensing, K.N. Toosi University of Technology, Valieasr Street, Tehran, Iran
[b] Dept. of Photogrammetry and Remote Sensing, K. N. Toosi University of Technology, Valieasr Street, Tehran, Iran
[c] Dept. of Electrical Engineering, K. N. Toosi University of Technology, Seyedkhandan Street, Tehran, Iran

ARTICLE INFO

Article history:
Received 14 March 2016
Revised 31 January 2017
Accepted 3 February 2017
Available online 4 February 2017

Keywords:
Soccer
Playfield
Player
Detection
Labeling
Appearance
Tracking
Occlusion
Evaluation

ABSTRACT

There is a growth of demand for automatically analyzing soccer matches and tactics. Since players are the focus of attentions in soccer matches, player tracking is a fundamental element in most soccer video analysis. The aim of player tracking is to extract the trajectories of players, and its input is provided through some preprocessing steps including playfield detection, player detection, player labeling, occlusion handling and player appearance modeling. Soccer player tracking is a complex and challenging task due to difficulties such as blur, illumination change and heavy occlusions. This paper presents the state-of-the-art in preprocessing and processing methods for soccer player tracking. We categorize different approaches, analyze their strengths and weaknesses, review evaluation criteria and conclude future research directions.

1. Introduction

Automatic soccer video analysis is a mandatory response to the growing demand by sport professionals and fans for extracting semantic information. Soccer video analysis has also attracted wide applications, such as player trajectory extraction, content retrieval and indexing, summarization, highlight detection, 3D reconstruction of the soccer match, animations, virtual view generation, virtual content insertion, visualization, editorial content creation and content enhancement, content-based video compression, tactical analysis, pattern of attack or goal analysis, statistical evaluations, player action recognition, verification of referee decisions, adapting the training plan and evaluating strengths or weaknesses of a team or a player. Player detection and tracking are fundamental elements required for extracting such understanding of the game. In general, player detection and tracking are quite challenging due to many difficulties, such as similar appearance of players, complex interactions and severe occlusions, unconstrained outdoor environment, changing background, varying number of players with unpredictable movements, abrupt camera motion and zoom, calibration inaccuracy due to the low textured field and edited broadcast video, noise, lack of pixel resolution especially on small distant players, clutter and motion blur. Examples of blurred players and lines are shown in Fig. 1.

An interesting review of the state-of-the-art in tracking algorithms can be found in Yilmaz et al. (2006), and a survey on visual tracking was provided by Yang et al. (2011). Two surveys on soccer video analysis were also provided by D' Orazio and Leo (2010) and Oskouie et al. (2014). However, they were focused on various soccer video applications, while methodologies for player detection and tracking were concisely reviewed. A review of spatio-temporal analysis of team sports was also presented by Gudmundsson and Horton (2016); however, it rarely focused on soccer player tracking. The main aim of this paper is to review in detail player tracking and its preprocessing steps. Moreover, different criteria for evaluation of the performance are reviewed. Accordingly, evaluation, player tracking and its preprocessing steps (playfield detection, player detection, occlusion resolution and appearance modeling) are shown in Fig. 2. Playfield detection plays a primary role in soccer video analysis. It eliminates the spectator region and reduces false alarms and noises within the playfield, which is of much benefit to the subsequent tracking procedure. Moreover, correct player detection is essential for initializing the tracker and provides observations required by some trackers. Another important aspect is player classification into five classes corresponding to two teams, two goalkeepers and referee, namely, player labeling. On the other side, occlusion is the most challenging issue in tracking soccer players, which occurs when

* Corresponding author.
E-mail addresses: mmanafifard@mail.kntu.ac.ir (M. Manafifard), ebadi@kntu.ac.ir (H. Ebadi), moghaddam@kntu.ac.ir (H. Abrishami Moghaddam).

http://dx.doi.org/10.1016/j.cviu.2017.02.002

Fig. 1. Scenes with blurred players (in particular in feet) and blurred or duplicate lines.

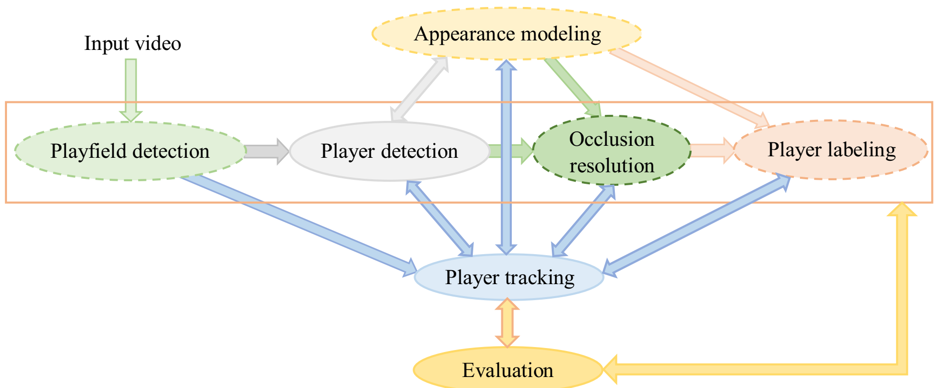

Fig. 2. The flowchart of soccer player tracking.

players hide each other either partially or completely. Occlusion is sometimes so severe that even the human observer can hardly see the occluded player, and the low quality of the images complicates more the problem. Also, more difficult situation arises due to the occlusion among teammates having similar appearances. Moreover, most of the player labeling, color-based player detection, occlusion resolution and player tracking methods require appearance representation of the players and referee while appearance modeling faces challenges such as updating models and selecting the most discriminative features. Player detection result can also be used as input to the appearance modeling step (gray bidirectional arrow) by restricting it to the detected player region. At the end of the tracking step, player tracking and its prior steps (rectangle in Fig. 2) are evaluated using some evaluation criteria. However, some prior steps are optional depending on the tracking step (dotted circles). For instance, playfield detection can be ignored, or appearance modeling is not required when color cues are ignored by a tracker. Occlusion resolution can also be performed within the tracking step (e.g. graph tracking), or player labeling can be solved during the tracking based on color cues (e.g. particle filter (PF)). Moreover, tracking results can be used as inputs to some prior steps, such as player detection, player labeling, occlusion resolution and appearance modeling (blue bidirectional arrows). For instance, probable player regions in the current frame can be predicted using the tracking result in the previous frame, or the previous appearance model can be updated using the appearance model of the tracked player. Occlusion situation can also be predicted using the distance between tracked players in the previous frame, or players' labels can be updated regarding the color-based tracking result. Moreover, all these steps can be modified regarding the evaluation results (orange bidirectional arrows).

Motion capture in commercial applications can be achieved with tracking reflective, magnetic markers or global positioning system (GPS) (Rangsee et al., 2013) on a player's body which are not always possible in sport domains, where the player movement can be affected or markers are not allowed. As a solution, computer vision techniques aim to dispose of such markers. Accordingly, some player tracking systems such as TRACAB (Capturing and visualizing large scale human action, 2016) achieved real-time and high precision tracking thanks to advanced camera setups, developed hardware and stereo vision technology.

The player detection and labeling provide observations of players; however, it is necessary to relate observations via a tracker. Moreover, most of the errors resulting from missed detections, false positives or mislabeling can be resolved by incorporating a tracker. A large number of tracking algorithms, such as Kalman filter (KF), PF, meanshift, snake and template matching, have been applied to deal with this topic. Generally, player tracking is performed during two main steps, namely, filtering and data association. Filtering is about unknown state estimation (e.g. position and velocity). However, multiple-player tracking involves the problem of data association for jointly tracking of players, since independent tracking of players tends to fail for closely spaced players. Therefore, data association works out which measurements are generated by which players. The problem gets more challenging with an increase in false alarm rate, missed observations and density of tracks. Accordingly, joint probabilistic data association filter (JPDA) and multiple hypothesis tracking (MHT) are two well-known techniques in the literature.

Although useful information was provided by previous literature surveys, the main drawback was addressing relating works without appropriate categorization of different methods or discussion about their weaknesses and strengths. Accordingly, the survey on different prior steps was superficial in the literature, and it was often addressed partially. Although an informative survey on visual tracking was provided by Yang et al. (2011), few tracking methods were reviewed. In addition, some literature surveys were focused on different soccer (D' Orazio and Leo, 2010; Oskouie et al., 2014) or sport (Gudmundsson and Horton, 2016) video analysis, and methodologies for player tracking and its prior steps were concisely reviewed. In D' Orazio and Leo, (2010) and Oskouie et al. (2014), prior steps, such as playfield detection, player detection, player labeling, appearance modeling, evaluation criteria and different camera setups were not reviewed, or few methods were addressed along with the tracking step. Moreover, different tracking methods for soccer player tracking were not categorized, and few trackers were reviewed. The main goal of this paper is to review different preprocessing steps for soccer player tracking, categorize different tracking frameworks and compare them in terms of the available evaluation criteria. Accordingly, separate sections are dedicated to different prior steps and tracking methods. It may help researchers to get familiar with the renowned and state-