Detection of Individual Ball Possession in Soccer

Martin Hoernig[1] (hoernig@in.tum.de),
Daniel Link[2] (daniel.link@tum.de),
Michael Herrmann[1](michael.herrmann@tum.de),
Bernd Radig[1](radig@in.tum.de),
Martin Lames[2](martin.lames@tum.de)

[1] Image Understanding and Knowledge-Based Systems, TUM,
[2] Department of Performance Analysis, TUM,
Arcisstraße 21, D-80333 München, Germany

Abstract. While ball possession usually is considered on team level, a model on player level brings several advantages. We calculate ball possession and control statistics for all players as well as new ball control heat maps to evaluate the players' performances. Furthermore, a basis for detecting events and tactical structure becomes available. To derive individual ball possession from spatio-temporal data, we present an automatic approach, based both on physical knowledge and machine learning techniques. Moreover, we introduce different ball possession definitions and algorithms to model various grades of ball control. When applied to flawless raw data, the algorithms show precision and recall ratios between 80 and 92 %. With approximately four percentage points less in uncorrected data, the presented algorithms are also reliable in real-world scenarios.

1 Introduction

Game analysis plays an important role in soccer coaching. Observing and analyzing tactical behavior can generate useful information that can be used for managing training processes and developing match strategies [1]. The technological innovations of recent years in the field of position tracking present new challenges in analyzing and interpreting the resulting data. The key lies in using intelligent algorithms in order to derive complex performance indicators from the raw data that add real value when it comes to game analysis [4, 5].

This paper describes and evaluates a method that enables different types of ball possession to be detected using ball and player positions. From a sports science perspective, ball possession is the most commonly investigated performance indicator [8]. Its relevance is easy to understand, since being in control of the ball is a fundamental prerequisite for being able to invade the opposing team's third of the pitch and score goals. Existing research [6, 7, 9] has been based exclusively on ball possession on the team level. Because up until now, such data has been collected by the *competition information providers (CIP)* solely on the basis of ball possession changes between teams. The reason for this

P. Chung et al. (eds.), *Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS)*, Advances in Intelligent Systems and Computing 392, DOI 10.1007/978-3-319-24560-7_13

reduction in complexity is that ball possession data is collected by human data loggers concurrent with the game, and it would be too involved and expensive to manually record data on an individual player basis.

2 Ball Possession - Models and Detection

When it comes to the definition of ball possession, there is a certain amount of flexibility. In the following we use these definitions which describe various grades of ball control:

1. *Individual Ball Possession (IBP)* begins at the moment a player is able to perform an action with the ball following an IBP of another player or a game interruption. It ends at the moment an IBP begins for another player.
2. *Individual Ball Action (IBA)* of a player begins at the moment this player is able to perform an action with the ball and had no prior IBA. It ends at the moment the player is no longer able to perform any further action with the ball or ends with the next game interruption.
3. *Individual Ball Control (IBC)* for a player begins when an IBA for this player begins and ends at the moment this particular IBA ends. In difference to the IBA, an IBC only takes place if the player is able to decide between several strategic courses of action during the IBA.

Accordingly, a successful passing involves an IBP lasting until the ball is received by another player, while IBA and IBC end when the passer is no longer able to interact with the ball. No IBC occurs if the player has only the option of this passing.

Automatic detection of ball possession involves a three-step process. In the first step, the spatio-temporal tracking data provided by CIPs are pre-processed by a Rauch-Tung-Striebel smoother [10, Ch. 3, 4, 8]. The next step is the core of the procedure: the detection of IBP and IBA start and end points, concluded by the final step, an estimation of ball control using a Bayesian network.

By using the IBP model, it is sufficient to only calculate the moments in which the IBPs start, given that the CIP provides a running flag containing a match's status (running or interrupted) for every point in time. Because of a missing z-coordinate in common CIP data, a component of the distance between the players and the ball stays unknown. Thus, a threshold based detection is prone to errors. Instead, we use local maxima of the amplitude of changes in ball velocity (ball accelerations) to detect kicks. If a kick is detected, a player is within a distance that allows a physical interaction, and if they are the player with the shortest distance to the ball, IBP is assigned to the player. We refer to this method as *kick detection*.

Following the definition, the kick detection can also be applied to calculate IBA start points. There are three possible ways an IBA for a player can end: 1) the game is interrupted, 2) another player gains IBA or 3) the player is no longer able to interact with the ball. Whereas the first case is trivial to detect because of the running flag data, and the second case can be obtained directly

Table 1: IBP and IBA detection results in % (0.6 s tolerance window)

		Recall	Precision
IBP			
	Net play time without tracking errors	80.1	86.1
	Net play time	78.0	76.9
IBA			
	Net play time without tracking errors	86.8	92.4
	Net play time	88.0	86.6

through kick detection results, only the last case requires special treatment. This involves checking whether a player will still be able to interact with the ball in the near future. Our approach makes use of the current position and velocity of the ball to give an estimate of its future location. As long as it is possible for the player currently in possession of the ball to control the ball in one second (the moment of prediction), they will retain IBA. The ability to control a ball is again checked via a distance threshold. We refer to this method as *ball prediction*.

Once IBA start and end points are known, the IBA intervals can be derived. We distinguish ball control IBAs from IBAs without courses of strategic interaction by categorizing them into IBCs and non-IBCs. For this reason, we decided to train a Bayesian network [3] to classify ball control based on a set of variables like duration of ball action, average ball velocity and acceleration, variance of ball velocity and acceleration, average distance between ball and the player in ball possession, and number of opposing players within certain distances.

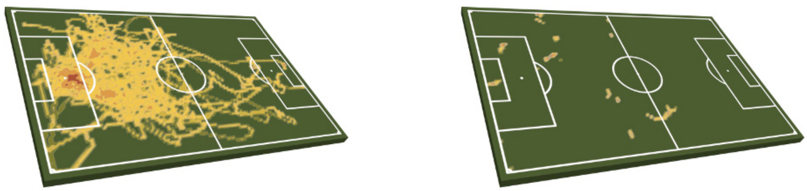
3 Evaluation

A match in a top European league served as a test sample in order to evaluate the quality of IBP, IBA, and IBC detection as well as a basis to calibrate the presented thresholds. The ball possession data were manually annotated by a trained, independent observer after the game to form a ground truth. Table 1 shows detection rates for sequences without errors which were caused by the underlying tracking system. These are compared to the results logged throughout the complete net play time. Precision is the ratio of correctly identified possession changes to the total number of changes in the measurements, recall is the ratio between the correctly identified changes to the total number of changes in the ground truth. By way of comparison, the CIP's team ball possession achieved a precision of 52.2 % to the given ground truth.

The IBC recognition rates were also determined by a ground truth comparison. The degree of consistency according to Cohen [2] is $\kappa = 0.38$ along with 92 % accordance to the ground truth. However, with only 25 non-IBC intervals, the used training set may not be sufficient for the proposed machine learning approach.

Fig. 1: Heat maps of a center-forward based on all positions during the game (a) compared to positions during IBC (b). Their team played from right to left.

(a) Movement (b) Ball control



4 Applications

IBP can be used to describe the characteristic of performance during a game. The analyzed match had a duration of 90:12 min and a net game duration (excluding stoppages) of 57:56 min. IBC was present in 1,291 phases totaling 29:48 min. The home team had a greater share of IBP, especially in the first half, whereas the away team controlled the ball for only 2:57 min. The average IBC duration, i.e. the time interval in which a player in possession of the ball could make and execute a tactical decision, was 1:27 min for field players who played for the entire match. There are, however, large differences between the players covering from 0:22 up to 3:38 min.

Fig. 1 shows a "traditional" heat map for a center-forward in contrast to a heat map based only on the time periods where the player had IBC. This player had the shortest average IBC intervals (0.9 s) in his team, which is not surprising for a center-forward.

5 Conclusion

Nowadays, the quality of tracking allows for individual ball possession to be reliably detected. Using the proposed methods on uncorrected data results in precision and recall ratios of around 80 %. This outcome indicates a wide variety of potential applications. Among statistical analyses and visual representations basic event detectors can be built easily based on individual possession data. Events such as passes, tackles, or shots on goal can be deduced directly. Also, being able to detect ball possession is a fundamental prerequisite for being able to discern higher value tactical structures like availability, pressing strategies, or marking tactics. The capability to recognize ball possession types holds considerable potential for improving the quality of match analysis in professional soccer. Further research has to be done in order for its importance to be assessed. Automation enables an additional performance-relevant parameter to be detected and evaluated in large data samples without the need for additional resources.

References

1. Carling, C., Reilly, T., Williams, A.M.: Performance assessment for field sports: physiological, and match notational assessment in practice. Routledge (2009)
2. Cohen, J.: A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20, 37–46 (1960)
3. Cooper, G.F., Herskovits, E.: A bayesian method for the induction of probabilistic networks from data. Machine learning 9(4), 309–347 (1992)
4. Grunz, A., Memmert, D., Perl, J.: Tactical pattern recognition in soccer games by means of special self-organizing maps. Human movement science 31(2), 334–343 (2012)
5. Gudmundsson, J., Wolle, T.: Football analysis using spatio-temporal tools. Computers, Environment and Urban Systems 47, 16–27 (2014)
6. Hughes, M., Franks, I.: Analysis of passing sequences, shots and goals in soccer. Journal of Sports Sciences 23(5), 509–514 (2005)
7. Jones, P., James, N., Mellalieu, S.D.: Possession as a performance indicator in soccer. International Journal of Performance Analysis in Sport 4(1), 98–102 (2004)
8. Mackenzie, R., Cushion, C.: Performance analysis in football: A critical review and implications for future research. Journal of sports sciences 31(6), 639–676 (2013)
9. Pratas, J., Volossovitch, A., Ferreira, A.: The effect of situational variables on teams performance in offensive sequences ending in a shot on goal. a case study. Open Sports Sciences Journal 5, 193–199 (2012)
10. Särkkä, S.: Bayesian filtering and smoothing, vol. 3. Cambridge University Press (2013)