Recognizing Movement Patterns in Automatically Identified Tactical Situations of a Football Match

Udo Feuerhake[1]

[1]Institute for Cartography and Geoinformatics, Appelstrasse 9a, 30167 Hannover, Germany
Email: feuerhake@ikg.uni-hannover.de

1. Introduction

Nowadays, there is a big variety of different football analysis methods available. A rather challenging but also promising one is the recognition of movement patterns, as those patterns contain information which can be used for different purposes. For instance, these patterns can be utilized to *predict* future movements and to *characterize* the players in general (cf. Figure 1), by revealing typical movement behaviors (Han et al. 2011). The latter can be used to also reveal the tactics of a team and thus provide valuable information for the next opponent. Further, both, the prediction and the characterization, can support the sensor-based player tracking by making it more robust against object occlusions. In this use case, either the estimated trajectory or the identified typical movement behavior can be used to determine a temporarily unknown player's identity.

Figure 1. Movement patterns can be used for predicting the future player trajectory (left) or to characterize and recognize players (right). The red trajectory is the currently observed track, whereas the blue and gray trajectories show two different typical movement behaviors.

In order to even extend the information of a movement pattern, we also consider the contextual information given in the time interval of the current play situation, which, for example, can be a complete attack or defensive situation. We assume that a player's intention also changes in different situations. For example, a defending player wants to be as close to his opponent as possible to prevent him to score, while an attacking player usually wants the opposite. Those different intensions may also be observed in different offense and defense patterns. Of course, by this we expect to find less patterns in total because we have a smaller search space for each kind of pattern. However, the resulting patterns have a stronger meaning, since they do not contain different intensions and thus are less often generated by coincidence.

In this work, movement patterns consist of repetitive similar movements (instances) which occur in the dataset. Their recognition is based on the players' trajectories, i.e. temporally ordered sets of discrete player locations. A comprehensive amount of related work in this context is presented by Gudmundsson and Wolle (2014), who propose a clustering to find repetitive trajectory segments. The segmentation is based on methods presented in Buchin et al. (2011). Niu et al. (2012) and Zhu et al. (2007) categorize attacks by their start locations and a predefined scheme in order to identify different types. In Lee et al. (2008), Liu et al. (2013), Lee et al. (2015) and Hung et al. (2015) different frameworks are introduced, which either automatically select relevant features, allow different time densities or time gaps during the clustering of the trajectories. Dodge et al. (2012) describe an approach to identify similar trajectories by applying an Edit-distance to symbolic representations of the trajectories.

In this paper we present an approach to identify time intervals containing specific situations in a football match automatically using a classification algorithm. Furthermore, we extract the players' movement patterns using the *DBSCAN* (Ester et al. 1996) in combination with the *Fréchet distance* and *optionally invariant features* to group similar trajectory segments of the previously identified time intervals. Each of the found clusters contains the instances of a particular movement pattern. Although the interpretation of the found patterns is crucial for the evaluation of the players' performances, it is not part of this work as it is only meant for recognizing those patterns.

2. Approach

The proposed approach consists of two consecutive steps, which are described in the following sections.

2.1 Identification of Situations and Segmentation of the Trajectories

The time intervals we exemplarily want to identify contain *offense* and *defense* situations. Please note, that this set can be extended by further representative situations, e.g. *corner kicks* or *free kicks*. We treat this identification as classification problem. Thus, we require a suitable classification algorithm, relevant features and training data. Since our input data only consists of the player trajectories (without the ball) as well as the field coordinates, we can only use features which can be derived from this data. In this work we use the *ID3-algorithm* (Quinlain 1986) to predict the labels for the samples. We could have also used a different algorithm, e.g. Random Forest or a Neuronal Net, as the experiments had shown that they provide very similar results. However, the reason for applying the *ID3-algorithm* is, that it helps to select the relevant features based on the provided information gain measure (see Table 1).

Table 1: The features used for the classification of the situations (ordered by their information gain).

Feature	Description: The players' …	Information gain [bit]
x_{min}	min. x-coordinate	0.519
x_{mean}	mean x-coordinate	0.502
x_{max}	max. x-coordinate	0.446
y_{min}	min. y-coordinate	0.187
y_{max}	max. y-coordinate	0.170
y_{mean}	mean y-coordinate	0.134
v_{mean}	mean velocity	0.066
h_{mean}	mean movement direction	0.049

In order to train a classification model, we have generated a training dataset by using an own implemented labelling tool, which assists the user by showing the movement trajectories and a corresponding video at the same time. With the help of this tool the timesteps of scenes containing the relevant situations have been assigned with the corresponding labels. Such a generated sample is described by a vector containing the features given in Table 1.

The learned classifier is applied to the samples of test dataset. Afterwards, consecutive and equally labelled samples are aggregated to periods representing time intervals of a specific situations. The segmentation uses the determined time intervals to cut the trajectories. The resulting meaningful segments contain the movement information of the players during the different situations.

2.2 Recognition of Movement Patterns

Since in this work a movement pattern consists of repetitive, potentially slightly deviated, motions we apply the *DBSCAN* to the previously determined segments. The resulting clusters contain the instances of each pattern. The clustering parameter ε is used for specifying the degree of similarity which is required that segments are treated to be similar. The lower ε is chosen, the more similar the instances of pattern will be. However, a small ε also means that less clusters or patterns are found. The second parameter *minIts* is used to define the minimum support count of pattern. If *minIts* is not reached, the corresponding segments will be marked as noise. The *DBSCAN* cannot be applied to the segments directly as it uses the Euclidean point distance by default. Since we have to calculate distances between trajectories, we follow Gudmundsson and Wolle (2014) and use the *Fréchet distance* as distance metric.

We are also interested in patterns which consist of movements occurring at different places on the field or moving in different directions. As shown in Figure 2 the segments can either be translated (center), translated and rotated (right) or must not be transformed at all (left) to be assigned to the same cluster. We can also add a scale invariance as an additional option. To this end, we additionally integrate invariant clustering features by replacing the standard movement representation, i.e. a trajectory consisting of a temporally ordered sequence of $[x \quad y]^T$ coordinates, by a different one. Table 2 gives an overview over the possibilities of representing the movements. The most suitable representation depends on the scenario and the allowed invariances of the segments.

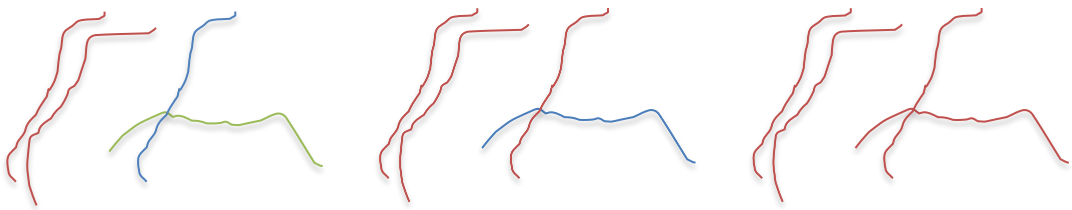

Figure 2. Using the invariant clustering features from Table 2 segments are assigned to the same cluster (same color) although being translated and rotated.

Table 2: The options to represent the movements (x,y: coordinates, ϕ: movement direction (heading), r and $\|r\|$: length and normalized length of movement vector).

Invariances	Representation
None	$[x \quad y]^T$
Translation	$[dx \quad dy]^T$ or $[\phi \quad r]^T$
Translation & rotation	$[d\phi \quad r]^T$
Translation & scale	$[\phi \quad \|r\|]^T$
Translation & rotation & scale	$[d\phi \quad \|r\|]^T$

3. Experiment and Discussion

In an experimental evaluation we apply our approach to a football dataset, which contains the player trajectories in more than 20 matches. The sampling rate is 5 Hz which leads to 297k measurements per match. The ball trajectory has not been recorded. We further have recorded several matches using cameras to have the opportunity to annotate or label the trajectories later.

After generating a training dataset, which contains 2345 labelled samples, we apply the leaned decision tree to the remaining matches. In Figure 3 we exemplarily show the resulting offense (green) and defense situation intervals (red) on the time bar of the first half of a match. These intervals are used to segment the trajectories. Three instances of the results are shown in Figure 3. Note that in the offensive situation, most of the team members are in the center or right part of the field. The evaluation of the classification results using the prelabelled data leads to precision and recall values of 0.98. The classification of the non-prelabelled data is done manually by inspecting the intervals using the corresponding trajectories and video material. It also shows that the classification works well.

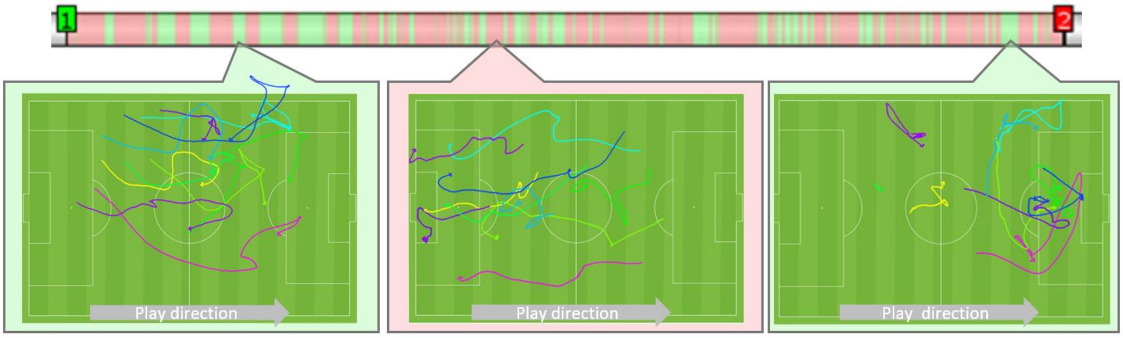

Figure 3. Three examples for identified offense (green) and defense situations (red) as well as the player trajectories of the analyzed team (in different colors). Please note the given direction of team.

To evaluate the resulting patterns, we apply this recognition approach to a single match consisting of the trajectories of one team with in total about 152k points. After the segmentation the trajectories are split into 741 offense trajectories. Table 3 shows the resulting number of patterns and the average number of instances per pattern depending on the given invariances and clustering parameters.

Table 3: The number of patterns and the number of their instances of one match changes depending on the allowed invariances.

Invariances	Parameters	# patterns	Avg. # instances
None	$\varepsilon = 10, minIts = 2$	15	2.6
Translation	$\varepsilon = 0.3, minIts = 2$	7	3.9
Translation & rotation	$\varepsilon = 0.3, minIts = 2$	15	6.1
Translation & scale	$\varepsilon = 2, minIts = 2$	15	4.5
Translation & rotation & scale	$\varepsilon = 2, minIts = 2$	10	6

Figure 4 shows some examples for recognized offense patterns of two teams allowing *no invariances*. These patterns give a hint on their different offensive strategies. While team A (left) seems to principally attack over the wings, team B's patterns are more distributed over the complete width of the field (right). This kind of information could be used to **predict** a team's next attack.

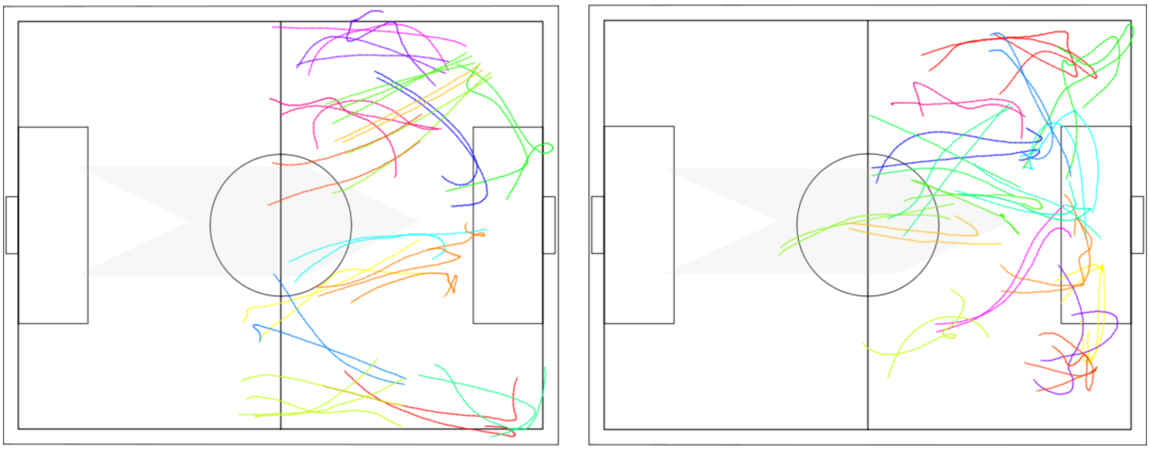

Figure 4. The recognized patterns (different colors) of two teams reveal their different offensive strategies.

A similar analysis can also be done to compare two individual players with different roles. This time the patterns are *translation invariant*. Figure 5 presents the patterns of a wing player (left) and central midfielder (right). The movements of the wing player are straighter and located on the sides of the field, whereas the central midfielder generally stays in the center. His patterns are curvier which indicates a more flexible movement. In this case the patterns **characterize** both players as they obviously reveal different movement behaviors.

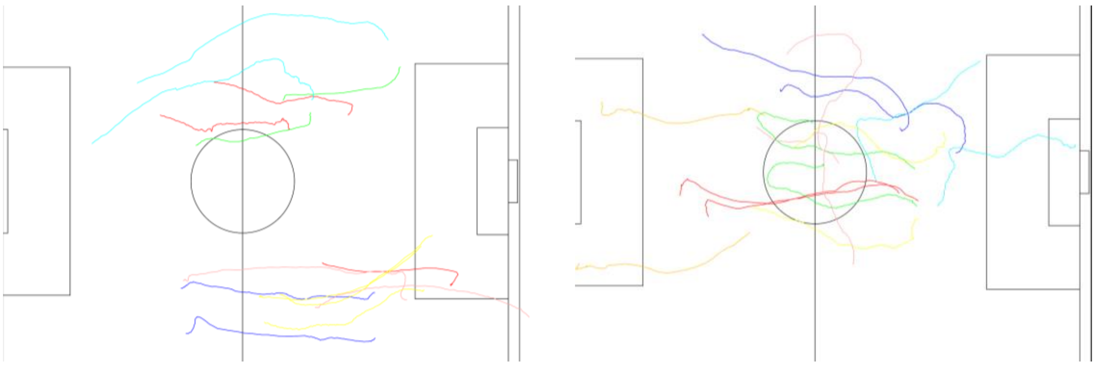

Figure 5. The patterns (different colors) of players with different roles (wing player, central midfielder) differ.

4. Conclusion and Outlook

In this paper we propose an approach to recognize situation-dependent movement patterns, which therefore are especially meaningful as their contextual background is considered. The performed experiment shows the automatic extraction of offensive and defensive patterns. In future work, this approach can be extended by identifying additional situations, e.g. corner, free or goal kicks, successful attacks which lead to a goal, etc. To this end, the requirement for the ball trajectory has to be examined, too.

Further, we plan to investigate the replacement of the clustering-based pattern recognition step by a sequence analysis-based one, as proposed in Feuerhake (2016), because it is able to detect patterns, which do not extend over the entire segment. We finally plan to apply this approach to datasets from other domains, e.g. from the observation of animals, to verify its portability.

References

Buchin, K., Buchin, M., Van Kreveld, M., Luo, J., 2011, Finding long and similar parts of trajectories. *Computational Geometry* 44 (9), S. 465-476.

Dodge, S., Laube, P., Weibel, R., 2012. Movement similarity assessment using symbolic representation of trajectories. *International Journal of Geographical Information Science* 26 (9), S. 1563-1588.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996, A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Kdd*. Bd. 96. S. 226-231.

Feuerhake, U., 2016, Recognition of Repetitive Movement Patterns - The Case of Football Analysis. *ISPRS International Journal of Geo-Information* 5 (11), S. 208.

Gudmundsson, J., Wolle, T., 2014, Football analysis using spatio-temporal tools. *Computers, Environment and Urban Systems* 47, S. 16-27.

Han, J., Kamber, M., Pei, J., 2011, Data mining: concepts and techniques. Elsevier.

Hung, C.-C., Peng, W.-C., Lee, W.-C., 2015, Clustering and aggregating clues of trajectories for mining trajectory patterns and routes. *The VLDB Journal* 24 (2), S. 169-192.

Lee, J. G., Han, J., Li, X., 2015, A unifying framework of mining trajectory patterns of various temporal tightness. *IEEE Transactions on Knowledge and Data Engineering* 27 (6), S. 1478-1490.

Lee, J.-G., Han, J., Li, X., Gonzalez, H., 2008, TraClass: Trajectory Classification Using Hierarchical Region-based and Trajectory-based Clustering. *Proc. VLDB Endow*. 1 (1), S. 1081-1094.

Liu, S., Wang, S., Jayarajah, K., Misra, A., Krishnan, R., 2013, TODMIS: mining communities from trajectories. In: *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management. CIKM '13*. ACM, New York, NY, USA, S. 2109-2118.

Niu, Z., Gao, X., Tian, Q., 2012, Tactic analysis based on real-world ball trajectory in soccer video. *Pattern Recognition* 45 (5), S. 1937-1947.

Quinlan, J. R., 1986, Induction of decision trees. Machine Learning 1 (1), S. 81-106.

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: *Proceedings of the 15th ACM international conference on Multimedia. ACM*, S. 58-67.