Analysis and clustering of soccer game situations in order to populate virtual environments

Matthieu Marbac-Lourdelle, ENSAI-CREST, France.
Brigitte Gelein, ENSAI-IRMAR, France.
Anthony Sorel, Univ. Rennes, Inria, M2S, France.

Soccer and data context:

The world of sports is no exception to the exponential inflation of the amount of data available. For a few years, vision systems enriched by machine learning algorithms have made possible the capture of the 2D movements of the ball and soccer players at the scale of the entire field during matches with very good spatial (< 10 cm) and temporal (> 10 Hz) resolution. These data allow for statistical descriptions of the match, offering increasingly fine-grained possibilities of game analysis for coaches [1], but also for researchers in STAPS who use these data to better understand the mechanisms that make a collective organization effective [2].

One of the questions raised lies in the individual perception that each player has of the collective movements around him, which guides his decision on his future action - an action that will contribute, in turn, to the evolution of collective movements in a permanent perception-action loop. This interweaving of the perception-action couplings of the different actors of the game can be questioned thanks to virtual reality. By immersing athletes in virtual environments simulating game situations in which it is possible to manipulate sensory afferences (sources of information), the M2S laboratory questions these information and decision making processes [3, 4].

To generate these contextualized virtual situations, with realistic player movements, M2S has a database of game situations from real soccer matches. In order to efficiently exploit this type of database, which is bound to grow, a powerful information processing tool is necessary. It must allow for an automatic classification of the different situations into similar varieties.

Professional clubs, such as SRFC, EAG or Paris FC, which already collaborate with M2S, as well as the French Football Federation, which currently structures its research-performance unit, are also very interested in these immersive technologies.

Methodological and technical approaches:

We chose to represent a game situation, at a given time t, by pairs of heatmaps that reflect the pressure exerted by the two teams on the different areas of the field. The use of Voronoi surfaces has the disadvantage of not taking into account the dynamics of the players. A model integrating the speed and direction of the players seems to be a more promising measure of the pressure. A descriptor, following this idea, has to be built. This descriptor has to translate the capacity of a player i to reach rapidly a certain point M of the field at time t. The data to be classified are therefore spatio-temporal data.

The degree of accessibility was initially inspired by Manuel Stein [5], but the PhD work we propose aims at refining its definition. The degree of accessibility of a point of the field can be the object of a

complex modeling taking into account of the individual characteristics of the players and the interactions exerted between them. Optimal transport can be used to define a distance between heatmaps. To go further, unbalanced optimal transport could be used in order to take into account the degrees of accessibility without normalization. After the choice of the distance, the clustering method of the game situations has to be chosen.

Among the approaches foreseen we can mention deep neural networks [6], U-Nets [7], hidden Markov chains [8], optimal transport [9], long short-term memories [10] or Generative Adversarial Networks [11]. This last approach could also be **promising** to add variability in each situation extracted from the data basis [12].

Duration and location of the position: 3 years at Ensai, Campus de Ker-Lann, 51 Rue Blaise Pascal, BP 37203 – 35172 BRUZ Cedex FRANCE

Expected profile of the candidate: The candidate will have a background in applied mathematics, data science and programming in R and/or Python with a strong taste for applications in sport. The taste for theoretical developments in statistics is also expected.

How to apply : Before June 18th, send a CV, a letter of motivation and two letters of recommendation to *bgelein@ensai.fr*

References:

[1] : Memmert, D., Lemmink, K. A., & Sampaio, J. Current approaches to tactical performance analyses in soccer using position data. Sports Medicine, 47(1), 1-10. 2017.

[2] : Low, B.; Coutinho, D.; Gonçalves, B.; Rein, R.; Memmert, D. & Sampaio, J. A Systematic Review of Collective Tactical Behaviours in Football Using Positional Data. Sports Medicine. 2019.

[3] : Bideau B, Kulpa R, Vignais N, Brault S, Multon F & Craig C. Using virtual reality to analyze sports performance. IEEE Computer Graphics and Applications. 2010.

[4] : Brault S, Bideau B, Kulpa R, Craig CM. Detecting deception in movement: the case of the side-step in rugby. PLoS One. 2012.

[5] Stein, M.. Visual Analytics for Cooperative and Competitive Behavior in Team Sports, PhD thesis, Universität Konstanz. 2020.

[6] Jing, L.; Tian, Y. Self-supervised Visual Feature Learning with Deep Neural Networks: A Survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

[7] Ronneberger, O.; Fischer, P.; Brox, P., "U-Net: Convolutional Networks for Biomedical Image Segmentation", Medical Image Computing and Computer-Assisted Intervention, 2015.

[8] : J. Alon, S. Sclaroff, G. Kollios and V. Pavlovic, "Discovering clusters in motion time-series data," 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., Madison, WI, USA. 2003.

[9] : H. Janati, M. Cuturi, A. Gramfort. "Spatio-Temporal Alignments: Optimal transport through space and time," AISTATS 2020

[10] Tsunoda, T.; Komori, Y.; Matsugu, M. & Harada, T. Football Action Recognition Using Hierarchical LSTM. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017.

[11] Cui, Q.; Sun, H.; Kong, Y.; Zhang, X. & Li, Y. Efficient human motion prediction using temporal convolutional generative adversarial network. Information Sciences, 2021, 545, pp. 427-447

[12] Shokrollahi, O., Rohani, B., & Nobakhti, A.. Predicting the outcome of team movements - Player time series analysis using fuzzy and deep methods for representation learning. ArXiv, abs/2109.07570. 2021.