Net-Based Game Analysis by Means of the Software Tool *SOCCER*

Jürgen Perl[1] & Daniel Memmert[2]

[1]D-55124 Mainz, Germany

[2]D-50933 German Sport University Cologne, Germany

Abstract

Game analysis has become much easier by automatic position recording. However, the problem remains how to transfer the astronomic amount of available data to a selection of useful information. Our approach is based on two ideas: Data reduction and pattern recognition. In the first step, by means of SOCCER, the position data of the players of a team are reduced to those of tactical groups like offense or defence, followed by normalization, where the players' constellations on the playground are reduced to their geometric formations relative to their centroids – i.e. the playground-independent position patterns. In the second step, those patterns are learned by the self-organizing neural network DyCoN, resulting in a collection of formation clusters, each containing a variety of shapes of the corresponding formation type. Based on that information, game analysis with DyCoN and SOCCER works as follows: Along the time-axis, position data of interacting tactical groups are fed to the net, which recognizes the time-dependent corresponding formation types. A first quantitative analysis then results in frequency distributions of formation types. Recombination with the playground position information leads to a playground specific frequency distribution. And adding the time information finally allows for process and interaction oriented analyses. Moreover, SOCCER not only offers quantitative results but also qualitative ones like game animation and tactical analyses by use of additional semantic action valuation.

KEYWORDS: PATTERN ANALYSIS, FORMATION, STATISTICS, TRAJECTORY

Introduction

In the following, first results from a project are reported, which deals with net-based game analysis in soccer. The project is supported by the German Research Foundation (DFG, Project-No. ME 2678/3-2). Additional information is given in Perl (2008), Grunz, Memmert & Perl (2009), Memmert & Perl (2009-1, 2009-2) Grunz, Endler, Memmert & Perl (2011).

In the beginning of computer aided game analysis, about 35 years ago, interesting ideas were developed but could not got to work because of a lack of available data. Meanwhile, data acquisition has become much easier due to automatic position recording. The problem is to transfer the astronomic amount of available data to a selection of useful information.

Our approach is based on two ideas: Data reduction and pattern recognition.

The process is starting with position data preparation and pre-processing, which is done by means of the software tool *SOCCER**, followed by three steps of analyses:

In the first step, *SOCCER* reduces the position data of the players of a team to those of tactical groups like offense or defence, followed by normalization, where the players' constellations on the playground are reduced to their geometric formations relative to their centroids – i.e. the playground-independent position patterns.

In the second step, those patterns are learned by the self-organizing neural network *DyCoN**, resulting in a collection of formation clusters, each containing a variety of shapes of the corresponding formation type.

Based on that information, game analysis with *DyCoN* and *SOCCER* works as follows:

Along the time-axis, position data of interacting tactical groups are fed to the net, which recognizes the time-dependent corresponding formation types.

A first quantitative analysis then results in frequency distributions of formation types. Recombination with the playground position information leads to a playground specific frequency distribution. And adding the time information finally allows for process and interaction oriented analyses.

Also qualitative analyses like semantic valuation and animation of game processes are supported by *SOCCER*.

In the third step, the trajectory analysis component of *DyCoN* enables tactical analyses of the game, including interaction and phase analyses. This way, in particular long term interaction patterns as well as hidden or creative tactical activities can be recognized and analyzed regarding success.

SOCCER-Based Data Reduction

In the first processing step the position data are transformed into pieces of information that can be used for statistic, pattern and tactic analyses.

The main idea is to transform constellations of players into a pair of normalized formations and corresponding positions on the playground: As is shown in Figure 1, the position data of a group build time-dependent constellations, which can be departed into its position (i.e. centroid of the constellation) and its characteristic formation (i.e. the normalization of the constellation relative to its centroid).

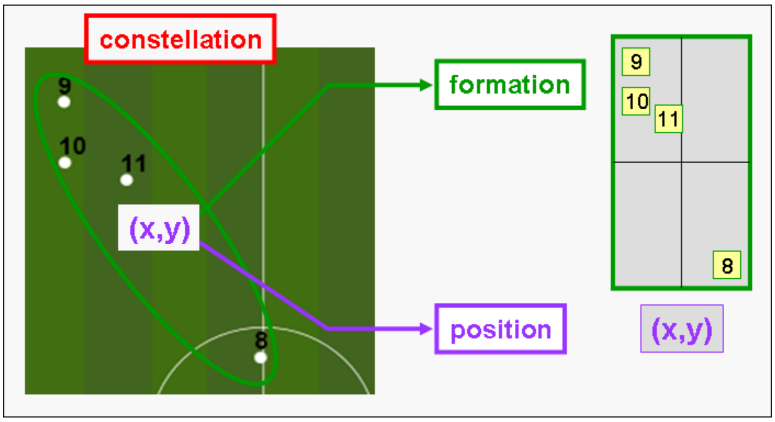

Figure 1. Departing a constellation into its position (centroid) and its characteristic formation.

To this purpose, SOCCER offers an interface for loading players' position data and selecting interesting tactical groups for normalization (see Figure 2). The normalized data are transferred to a file or a database.

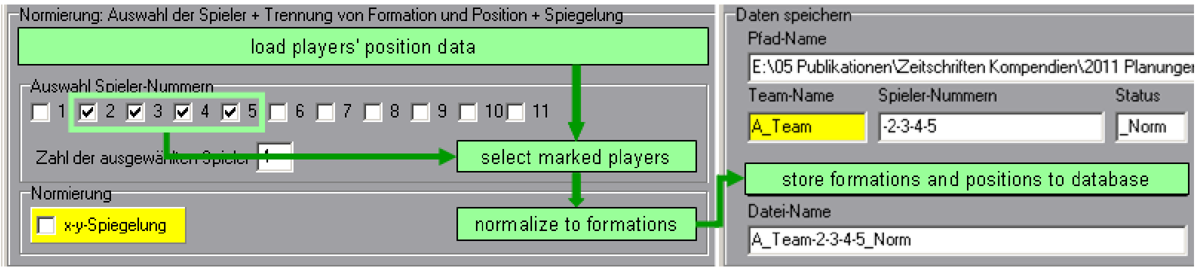

Figure 2. *SOCCER* data pre-processing interface.

Obviously, the number of characteristic formations (about 10) is much less than that of possible constellation (about 5400 per half-time). Therefore this reduction allows a much better pattern analysis of tactical group behaviour without loss of information, because the real game process can be reconstructed using the stored time-series of positions.

DyCoN-Based Pattern Analysis

By means of the self-organizing neural network *DyCoN*, the different types of formations are recognized and grouped into clusters, which in Figure 3 are presented using different colours. Within each type-representing cluster the different neurons (coloured small squares) represent different shapes of the corresponding type, as is exemplarily shown by 4 shapes of the defence formation "line".

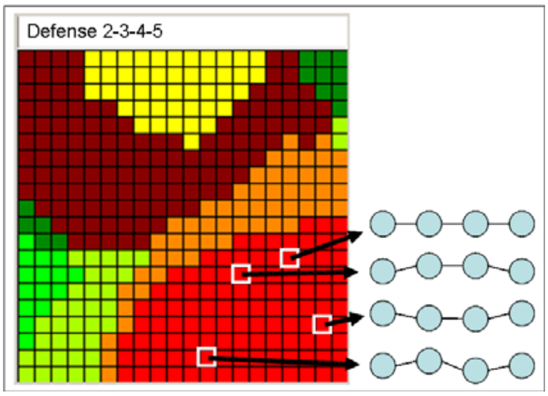

Figure 3. Trained DyCoN showing coloured clusters of formations, the neurons of which represent different shapes of the corresponding formation type.

If formation data are input to the trained net, the type (and the particular shape) is recognized by the net and now can be re-combined to the corresponding position and time data, resulting in a significantly reduced presentation of the game and its processes, nevertheless preserving all important information.

Those time-series of formation types and positions allow for a lot of analyses from space- or time-oriented distributions over success in interactions up to tactical aspects process of phases.

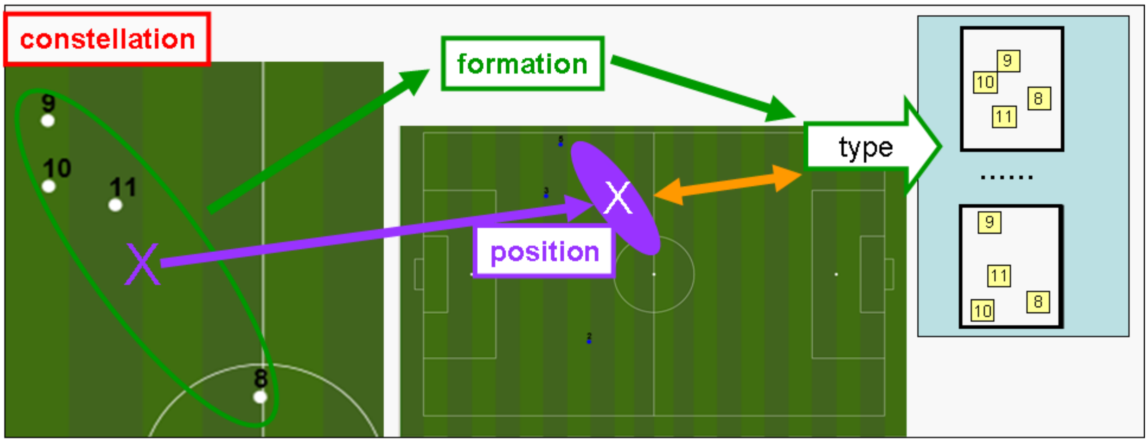

Figure 4. Net-based recognition of formation types and re-combination with position and time information.

The condensed game information from above can be used for generating a game protocol like presented in Figure 5, where time, formation type and position is completed by a manual semantic valuation.

time	type	position	action / result
...			
HT1 34:15	7	(45,23)	pass / successful
...			

Figure 5. Automatically generated game protocol manually completed by semantic valuations (yellow column).

Those protocol data then build the basis for the actual game analysis, as will be discussed in the following.

SOCCER-based Game Analysis

Game analysis in the following will roughly be distinct into three phases: Distribution analysis deals with space- and time-specific frequencies of formations as well as frequencies of corresponding interactions. Process analysis is oriented in the dynamics of the game, dealing in particular with simulation and animation of processes corresponding to formations. Tactics analysis deals with the game as a whole in order to recognize specific group behaviour and success of interaction processes.

Distribution Analyses

The distribution analysis presented in Figure 6 demonstrates a typical situation: Team A attacks in the formation of type 4, team B reacts with a defence formation of type 3. The distribution matrix shows that this particular interaction happened 523 times (i.e. at 523 seconds) in the regarding half-time.

In general, the matrix provides the distributions of formations of the teams as well as those of the respective interactions.

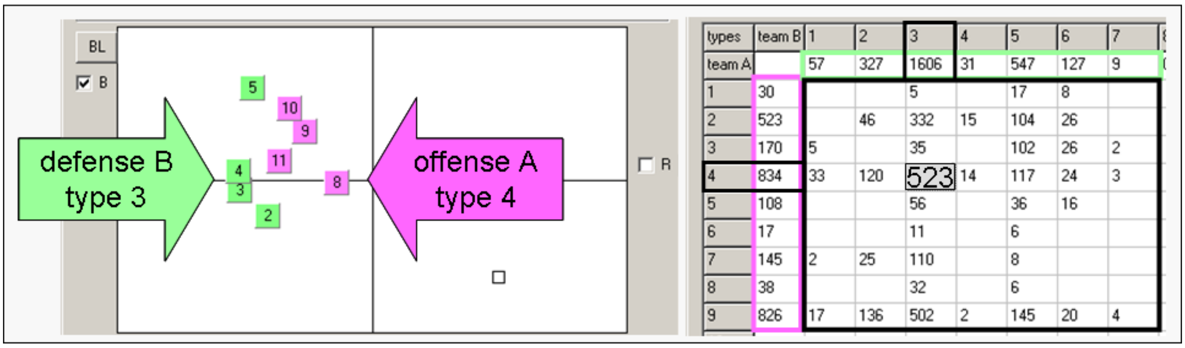

Figure 6. Distribution matrix of formation types and interactions.

The distribution matrix of *SOCCER* is interactive – i.e. clicking one of the entries shows the distribution of the corresponding events on the grids of the playground (which can be arranged arbitrarily). Figure 7 shows how the 132 interactions 'defence type 3 vs. offense type 1' (green highlight) are distributed on the playground. Clicking one of the grid fields (blue highlight) shows the distribution of the remaining entries in that area.

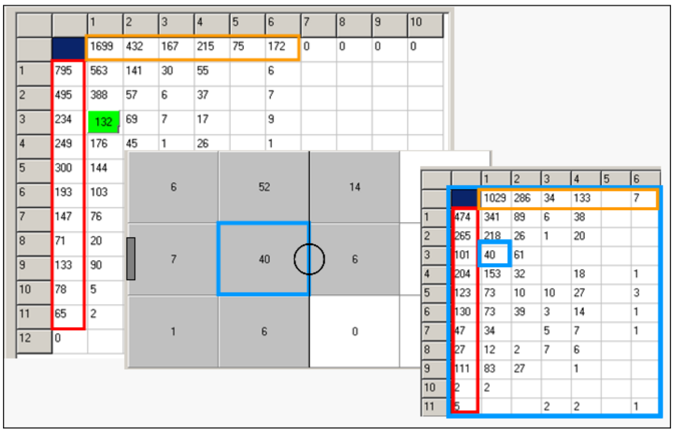

Figure 7. Interactive presentation shows distributions of formation interactions.

Statistical analysis is helpful for a first recognition of normal and of seldom or striking situations. In order to recognize the role they play in the game process, statistical analysis can be combined with animated process analysis:

Process Analysis

Clicking one of the entries of the distribution matrix also shows a list of all processes where that interaction took part. In Figure 8 the clicked interaction (green highlight) activates a list of 14 processes, the 8th of which took part in minute 23 from second 1374 to 1384 (violet highlight). Also this sequence can be activated resulting in an animation of the motion of the teams' centroids on the ground.

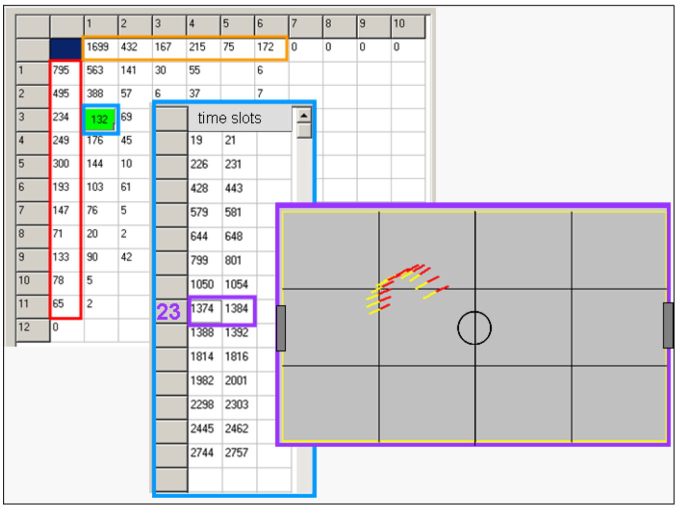

Figure 8. Interactive presentation: Number of interactions (green highlight) → list of processes with selected process (violet highlight) → sequence of teams' centroids.

Combined Quantitative and Qualitative Analysis

As has been mentioned above (see Figure 5), the formation data can be completed by semantic ones like technical or tactical aspects and success. In the following two examples demonstrate the valuation of success by means of *SOCCER*:

In the first example the success of a single player, depending on his tactical position and his action in the context of the current formation, is valuated. In Figure 9 the blue highlights mark the selected team (ITA), player (number 10) and formation (number 2, which is shown in the left most graphic together with the moving direction) as well as the selected action (number 6 of 13) and the selected tactical position (number 9 of 9). The graphic on the right hand side shows that there were 27 events in that situation, 24 of which were successful.

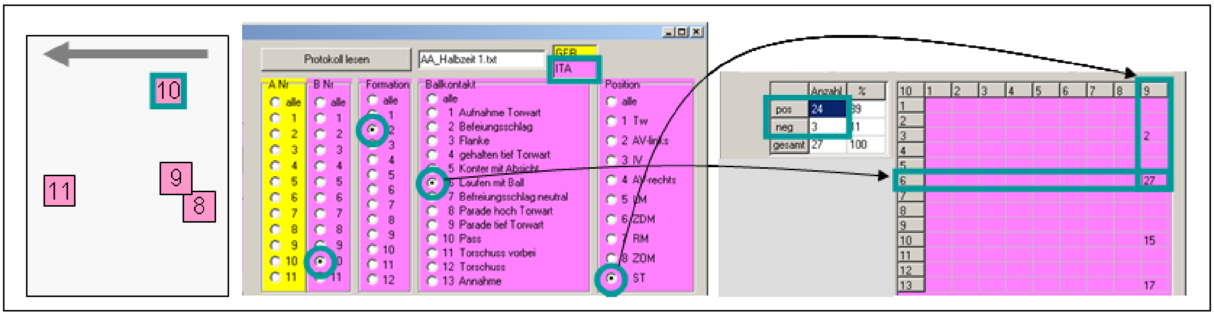

Figure 9. Tables for selecting team, formation, player, tactical position, and action with resulting information about player's success.

The second example deals with valuating the success of the team in a given formation interaction. Figure 10 shows from left to right the number of valuated interactions of a team, followed by the negative ones in absolute numbers and as percentages. Concentrating on the right graphic it seems that team A has serious problem in the interaction of formation 3 vs. formation 3. However, the absolute numbers are very small, reducing the importance significantly. Also '5 vs. 5' is negative but does not seem very important, while '5 vs. 2' seems to be a significant weakness, although the percentage of negative results is only 16.

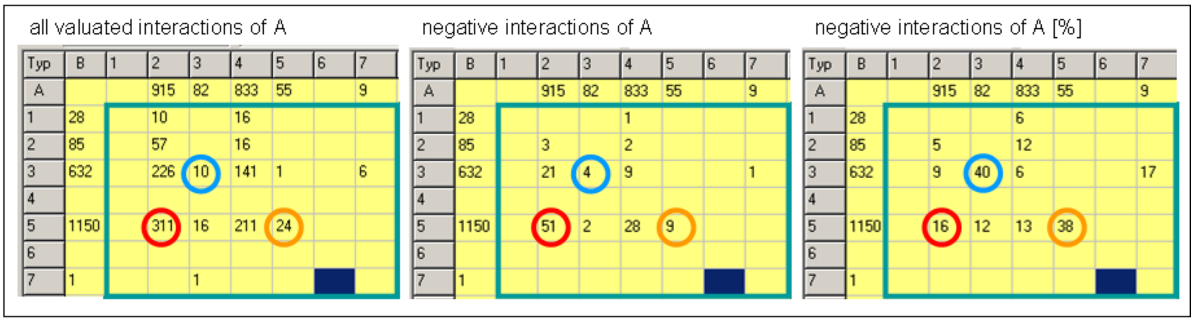

Figure 10: Matrices of valuated team success in the context of formation interaction.

Note, that the presented kinds of analysis are just examples, which can arbitrarily be completed if the valuation data are once available.

Net-based Tactics Analysis

A last kind of analysis that is provided by *SOCCER* is the net-based trajectory analysis. The idea is that at each point in time the formations of tactical groups are identified and therefore can be encoded by a corresponding number respectively colour. In Figure 11 the net has identified 10 offense formations, which are coloured from violet (1) to dark grey (10). In the left graphics it is shown, how team A and team B change their formations during the 60 seconds of the 16th minute – i.e. the process trajectories of teams A and B in minute 16. In the same way the trajectories can be shown for the whole game or, as has been done in Figure 11, for a particularly interesting part of one half-time (here minute 15 to 30): The selected part shows that the trajectory behaviours of A and B are quite different, possibly reflecting different tactical concepts.

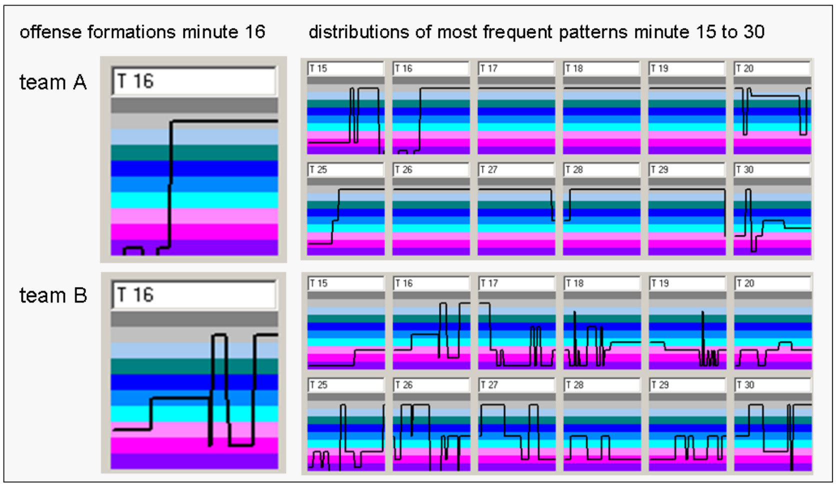

Figure 11. Process trajectories based on the time-series of formation types.

If and how much such trajectories really tell about tactics obviously depends on the tactical quality of the teams and has to be judged by experts like trainers.

Conclusion and Outlook

The aim of the distribution was to demonstrate what kinds of game analyses are possible if combining net-based pattern analysis with conventional statistic methods, as is symbolized in Figure 12:

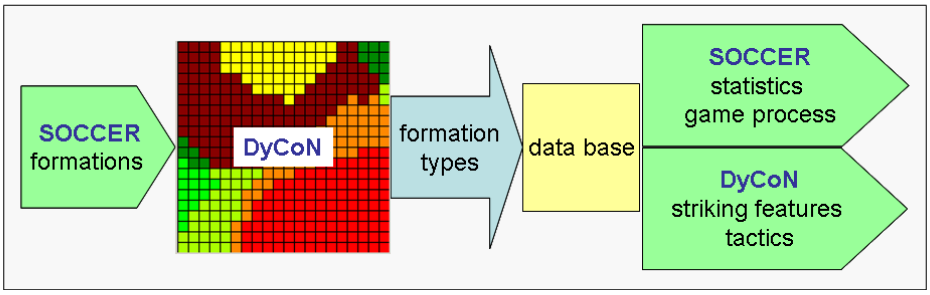

Figure 12. The concept of SOCCER: Combining net-based pattern recognition with statistical methods.

The results are encouraging and promise to be helpful in order to run successfully the above mentioned project as well as in getting new ideas and results for further and improved game analysis approaches.

References

Perl, J. (2008). Modelling. In P. Dabnichki & A. Baca (Eds.), Computers in Sport (pp. 121–160). Wit Press.

Grunz A., Memmert D. & Perl J. (2009). Analysis and Simulation of Actions in Games by Means of Special Self-Organizing Maps. In International Journal of Computer Science in Sport, *8*(1), 22-37.

Memmert, D. & Perl, J. (2009-1). Game Creativity Analysis by Means of Neural Networks. Journal of Sport Science, *27*, 139–149.

Memmert, D. & Perl, J. (2009-2). Analysis and simulation of creativity learning by means of artificial neural networks. *Human Movement Science*, *28*, 263-282.

Grunz, A., Endler, S. Memmert, D. & Perl, J. (to appear 2011). Netz-gestützte Konstellations-Analyse im Fußball [Net-based constellation analysis in soccer]. Darmstadt, Workshop on Computer Science in Sport 2010.