Neural Network-Based Process Analysis in Sport

Juergen Perl
University of Mainz, Germany

INTRODUCTION

Processes in sport like motions or games are influenced by communication, interaction, adaptation, and spontaneous decisions. Therefore, on the one hand, those processes are often fuzzy and unpredictable and so have not extensively been dealt with, yet. On the other hand, most of those processes structurally are roughly determined by intention, rules, and context conditions and so can be classified by means of information patterns deduced from data models of the processes.

Self organizing neural networks of type Kohonen Feature Map (KFM) help for classifying information patterns – either by mapping whole processes to corresponding neurons (see Perl & Lames, 2000; McGarry & Perl, 2004) or by mapping process steps to neurons, which then can be connected by trajectories that can be taken as process patterns for further analyses (see examples below). In any case, the dimension of the original data (i.e. the number of contained attributes) is reduced to the dimension of the representing neuron (normally 2 or 3), which makes it much easier to deal with.

Additionally, extensions of the KFM-approach are introduced, which are able to flexibly adjust the net to dynamically changing training situations. Moreover, those extensions allow for simulating adaptation processes like learning or tactical behaviour.

Finally, a current project is introduced, where tactical processes in soccer are analysed under the aspect of simulation-based optimization.

BACKGROUND

A major problem in analysing complex processes in sport like motions or games often is the reduction of available data to useful information. Two examples shall make plain what the particular problems in sport are:

In Motor Analysis, a lot of data regarding positions, angels, speed, or acceleration of articulations can be recorded automatically by means of markers and high speed digital cameras. The problem is that those recorded data show a high degree of redundancy and inherent correlation: A leg consisting of thigh, lower leg, foot, and the articulations hip joint, knee, and ankle obviously has only a comparably small range of possible movements due to natural restrictions. Therefore the quota of characteristic motion data is comparably small as well. Classification can help for deducing that relevant information from recorded data by mapping them to representative types or patterns.

In Game Analysis, during the last about 5 years an increasing number of approaches have been developed which enable for automatic recording of position data. Based on the video time precision of 25 frames per second, 9.315.000 x-y-z-coordinate data from 22 players and the ball can be taken from a 90-minutes soccer game. Obviously, the amount of data has to be reduced and to be focused to the major tactical patterns of the teams. Similar to what coaches are doing, the collection of players' positions can be reduced to constellations of tactical groups which interact like super-players and therefore enable for a computer-aided game analysis based on pattern analysis.

As is demonstrated in the following, neural network-based pattern analysis can support the handling of those problems.

MAIN FOCUS OF THE CHAPTER

Artificial Neural Networks

Current developments in the fields of Soft Computing and/or Computational Intelligence demonstrate how information patterns can be taken from data collections by means of fuzziness, similarity and learning, which the approach of Artificial Neural Networks gives an impressive example for. In particular self organizing neural networks of type KFM (Kohonen Feature Map) play an important role in aggregating input data to clusters or types by means of a self organized similarity analysis (Kohonen, 1995).

Net-Based Process Analysis

Processes can be mapped to attribute vectors – in a game, for example, by recording the positions of the players – which then can be learned by neurons. There is, of course, a certain loss of precision if replacing an attribute vector by a representing neuron, the entry of which is similar but normally not identical to that attribute vector. Nevertheless, there are two major advantages of the way a KFM maps input data to corresponding neurons:

1. The number of objects is dramatically reduced if using the representing neurons instead of the original attribute vectors: a 2-dimensional 20×20-neuron-matrix contains 400 neurons, while a 10-dimensional vector space with only 10 different values per attribute already contains 10^{10} = 10.000.000.000 vectors.
2. The dimension of input data is reduced to the dimension of the network (i.e. normally 2 or at most 3). This for example enables for mapping time-series of high-dimensional attribute vectors to trajectories of neurons that can easily be presented graphically.

There are three ways of gaining information from data by means of Artificial Neural Networks of KFM-type:

1. Neurons represent classes of similar data and so define types of information patterns.
2. Clusters of neurons represent time-static classes of similar information patterns and so build structures of information patterns.
3. Trajectories of neurons represent time-dynamic sequences of information patterns and so build 2-dimensional mappings of time-dependent processes. Moreover, trajectories themselves build patterns and therefore can be input to a network for classifying their similarities – which is extremely helpful not least in motor analysis or in game analysis.

There are a large number of successful applications that demonstrate how those neural networks can be used for that pattern analysis (see Perl & Dauscher, 2006).

Example "Gait Analysis": Reduction of Redundancy and Dimensionality

In gait analysis, data from articulations like for example hip-joint, knee and ankle can automatically be recorded using markers and so build a time series of n-dimensional attribute vectors which can be trained to a net. The result is that each of those n-dimensional vectors is mapped to a 2-dimensional neuron of the net – i.e. the dimension is reduced from n to 2. Corresponding to the original time series the neurons can be connected by a

Figure 1. Two trajectories of the same gait process, using 20 attribute values (left) and 10 attribute values (right), respectively. The high degree of similarity suggests that the missing 10 values are redundant and can be neglected.