8 Self-organizing maps and cluster analysis in elite and sub-elite athletic performance

Wolfgang Schöllhorn, Jia Yi Chow, Paul Glazier and Chris Button

This chapter examines ways in which movement patterns can be analyzed as performance contexts change or as a function of learning and development. The methods described can be used to study the effects of important factors such as fatigue, injury, learning, development or training in motor performance.

Classifying objects can be considered as one of the major tasks in science (Slife 1995). Classifying usually occurs at the beginning of the research process and involves categorizing objects by criteria specified by the investigator. Most often, these criteria are associated with implicit assumptions and rely on a certain philosophical background. In the scientific literature, two general types of classification study can be identified; those adopting a confirmatory approach or an exploratory approach, with the former being more common (Tukey 1980; Jaeger and Halliday 1998). In confirmatory approaches, the classes are given in advance and are tested for statistical significance. In exploratory approaches, by contrast, only the criteria for the classification of the objects are provided. Once certain classes are identified, the explorative approach can be followed by the confirmatory approach so as to test the identified classes for significance.

For methodological and historical reasons, the quantitative investigation of movement has typically been focused on the classification of simple movements on the basis of time-discrete amplitudes of selected variables. In biomechanical studies, these time-discrete variables (also known as performance parameters) have typically been specified by deterministic or hierarchical models and have some relationship with performance outcome. Time-series data typically have to be reduced or collapsed to single data points (maximum, minimum, average, value at specific event, etc.) and pooled so that (confirmatory) statistical analyses can be performed. In parallel, the analysis of movements has been mainly limited to the statistical comparison of single variables. In cases where multiple variables have been analyzed, potentially more useful practical knowledge can be derived.

When movement patterns are classified by means of multiple time-continuous variables, a more holistic approach is pursued. The difference between a time-discrete and time-continuous movement description can be illustrated with an analogy. If we see a known person far away standing still, it is often difficult to identify that person. Once he/she starts to walk, our visual system receives additional information that increases the likelihood of recognizing that person.

Similarly, if only the left elbow of the person walking can be viewed, the probability of correctly identifying that person is rather low. However, if the motion of other joints and segments are presented simultaneously, the probability of recognizing the person improves. Indeed, Johansson (1973) showed that perception of biological motion relies more heavily on relative motion, rather than absolute motion characteristics of limb and torso segments.

While time-discrete variables focus on instantaneous characteristics of single joint or segment motion, time-continuous variables provide the opportunity to categorize movement qualities such as types of movement (jumping, running, etc.), modes of movement (springy, tentative, rushed, etc.) and individual styles (individual expressions of types and modes) of movements. According to pattern analysis, the quantification of coordination is initiated by means of the determination of a reference pattern to which all other patterns have to be compared with. Alternatively, the similarity of all patterns to each other can also be determined.

Numerous questions can be addressed through the quantitative description of movement patterns in sport on the basis of time-continuous variables. For example, how similar are the average patterns of movement within a class of movement? How are gait patterns similar within an individual across different performance context or between different individuals? Even in the study of plants or animals and their subordination into certain classes in biology, a first step for time-continuous movement patterns is their classification on the basis of certain criteria. Because the general classification of movements already has been performed successfully by numerous scientists (e.g. Hay 1993), mathematical algorithms are first required that are relevant to everyday experience and observations. Thus, plausible criteria for the classification of objects seem to be their relative similarity or proximity in a most abstract understanding. The quantitative sorting of objects by means of linear distances led to the development of different forms of cluster analysis (Everitt *et al.* 2001), while the mathematical simulation of neuronal assemblies was associated with artificial neural nets that meanwhile provided plausible nonlinear groupings.

Without delving too deeply into the problem of interpreting similarity or proximity (Everitt *et al.* 2001), the simplest procedure is to quantify a certain quality of all objects and to determine the relative distance of these quantities (e.g. three people walking should be clustered on the basis of their minimum knee flexion during single stance phase in gait). The first step in such a procedure would be to assign a quality or variable to 'knee flexion', such as 'knee angle'. This quality is typically represented by a unit such as radians or degrees. Hence, qualities can be compared by means of their relative size or a vector distance. A commonly used measure for mathematical comparisons is the Euclidean distance, which represents the distance between two objects on a straight line. If the distance between two knee angles of two participants is smaller than the distance relative to a third participant, the first two participants would be assigned to the first cluster and the third participant to another. Applying the same procedure to movement patterns that are described by means of multiple time course variables, some vector and matrix algebra is necessary that includes similarity and proximity measures analogously. In contrast to cluster analysis, which exclusively sorts objects, self-organizing maps (SOMS; as a specific form of artificial neural network [ANN]) have the additional ability to reduce high-dimensional data to low-dimensional data. Therefore, SOMs are often used as a preparatory tool for cluster analysis.

In the following sections, we examine the application of cluster analysis and SOMs to movement analysis in sport and human movement science.

Cluster analysis

As discussed earlier, the application of cluster analysis in sport can be divided into exploratory and confirmatory approaches. In the exploratory approach, hierarchical cluster analyses are typically applied to holistic descriptions of movements. Starting with a single object, all other objects are clustered successively by means of a chosen metric until all objects are included. The exploration of the number and size of the clusters along the relative metric provides either previously expected groups of movements or leads to a new categorization that has to be interpreted creatively. In summary, here the least information about the structure of the objects is included in the classification process. In the second case of plausible interpretations, the confirmatory approach, can be followed. Here, the classes or at least the number of classes are given in advance and the cluster analysis procedure classifies all objects accordingly on the basis of the chosen metric. Subsequently, the identified clusters have to be validated and any cluster differences has to be tested for significance. However, irrespective of the approach, a dominant influence on the results is given by the chosen variables as well as by their pre-processing. Thus, the selection of variables is directly connected with the specificity of the hypothesis and the expectation of the investigator while the pre-processing of the data has an indirect but strong influence on the relative weighting of the variables with respect to the resulting clusters (for further details, see Everitt *et al.* 2001; Rein *et al.* 2010a).

Owing to the increased processing capacity of computers by the end of the last century, cluster analysis has since been applied increasingly to large sets of time course oriented movement data without data reduction.

Exploratory approaches

One of the first exploratory applications of cluster analysis in the context of movement pattern analysis in sports was provided by Müller (1986). Several kinematic variable time courses and muscular activities were measured during the demonstration of different skiing techniques for downhill skiers. Discrete biomechanical parameters were extracted and classified by means of a cluster analysis. It was found that irrespective of the snow and slope conditions, the techniques could be classified as similar.

An explorative cluster analysis on the basis of time-continuous variables in discus-throwing movements during practice of a single high-performance athlete provided evidence for a successful learning process that led to enduring/lasting qualitative changes of the throwing technique (Schöllhorn 1993). Eight discus-throwing trials before and after a biomechanical feedback intervention were described by means of the time courses of 40 joint angles and angular velocities during the final throwing phase. Three trials were performed within one competition before the intervention and five trials during different competitions following the intervention. For data reduction, these high-dimensional data were factor analyzed. The resulting factor-loading matrices were compared by means of a structure comparison algorithm, which led to a distance matrix. The subsequent cluster analysis (Figure 8.1) clearly separated three trials (T791–T793) before a specific training intervention from five trials (T84–T88) performed after the biomechanical feedback. Figure 8.1 displays exemplarily the history of a hierarchical cluster analysis applied to eight discus throws during a learning process. The cluster analysis begins with determining the distances between all trials and is followed by determining the two most similar (smallest distance) trials (T791 and T793). The clustered trials are then considered as a single new trial. Subsequently, the next two trials with the smallest distance are clustered together and so on until all trials are clustered together. Similar results revealed the cluster analysis of the same data when data reduction was performed by means orthogonal reference functions (Schöllhorn 1995).

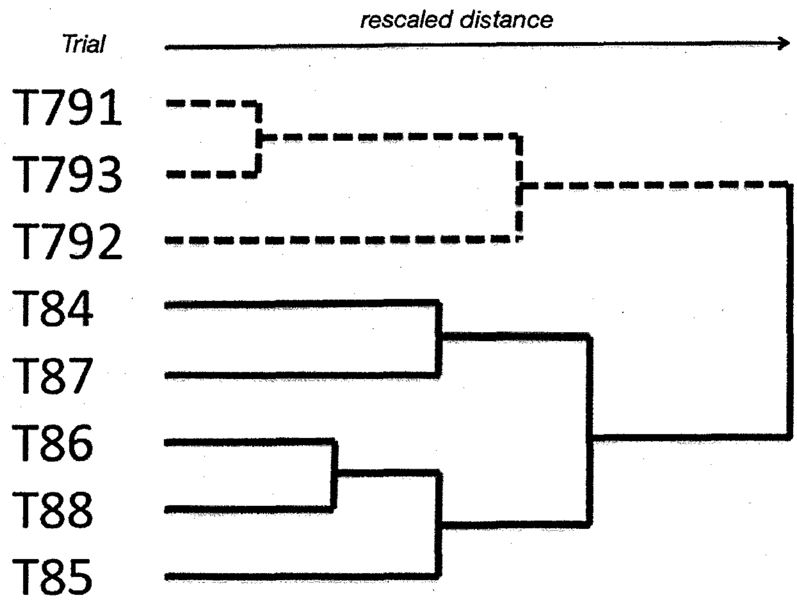

Figure 8.1 Clustering dendrogram resulting from an average linkage algorithm; horizontal lines indicate the level of the rescaled distance at which the respective movements are grouped into one cluster

Explorative cluster analysis of biomechanical data during several double steps in running provided a separation of left and right contact phase as well as of flight phase. Just by examining the ground contact phase of a single leg, 95% of the runners could be identified individually (Schöllhorn 1999). A similar application of cluster analysis to kinematic and muscular variables in long jumping allowed the identification of the athletes during take off (Jaitner *et al.* 2001a,b). Muscular variables pertaining to either the envelope of the rectified electromyography signals or their frequency spectra can be taken as input variables for the classification process for the cluster analysis. Different levels of expertise in handball players were also identified inter- and intra-individually on the basis of time-continuous, three-dimensional kinematic data of the hip, shoulder, elbow and wrist joint by Schorer *et al.* (2007). Cluster analysis yielded an assignment rate of 92% for the level of expertise.

A new direction in terms of cluster analysis applications in the field of elite sports was taken by Jäger *et al.* (2003) by analyzing tactical behaviour of teams in the form of player's coordinates during volleyball matches from the 2002 World Championships. Over 250 similar situations during a tournament where six participating teams competed against each other, were video recorded at 50 Hz. Positions (x and y coordinates) of the defending teams during a defensive move in time-discrete and time-continuous form were inputted to a cluster analysis that subsequently revealed team specific behaviours. Contrary to the opinions of a coach, who believed that all teams played with the same system, the specific behaviours of each team could be identified by means of the movement of all six players during the defence movement independent of the opponent. This individuality of team dynamics has a number of potential implications for tactical training.

Confirmatory approaches

A confirmatory approach of cluster analysis with a single discrete duration parameter was examined by Lames (1992). Using ideas from dynamical systems theory and with a focus on investigating the hysteresis effects (see Kelso 1995), the duration of the driving and pitching movement of several golfers was measured for different distances. In order to initiate a hysteresis effect, the golfers performed shots from 100 metres first, then decreasing down to ten metres in five-metre steps and then increasing back to 100 metres. The subsequent cluster analysis of the movement durations led to two, three, four or five clusters from which only the three-cluster solution could be interpreted plausibly in accordance with the assumed hysteresis phenomenon.

In a more recent study by Chow *et al.* (2008), which examined coordination changes of novice participants in a football (soccer)-kicking task, cluster analysis was also effectively used to determine intra-individual differences in kicking patterns. Key kinematic data of various joint motions were captured over a four-week intervention period (40 trials per session over 12 sessions) and the kinematic data were used as input for a cluster analysis procedure. Differences in coordination patterns within individual for each session were effectively determined and the information provided valuable insights about movement pattern variability, as well as the presence of preferred kicking patterns. Such information is extremely critical in helping researchers to understand the learning processes and to investigate nonlinearity in learning evidenced by sudden transitions from one movement pattern to another, as well as the emergence of pattern variability prior to a transition.

Another confirmatory application of cluster analysis was performed by Ball and Best (2007) to determine the presence of weight transfer for two styles of golf swing. Sixty-two golfers, from professionals to high-handicap players, performed simulated drives, hitting a golfball into a net. While standing on two force plates, the centre of pressure position relative to the feet was quantified at eight swing events identified from a 200-Hz video. Cluster analysis on the basis of these time-discrete parameters revealed two major styles of golf swing: a front-foot style and a reverse style. Nevertheless, validation procedures were required.

Validation of clusters

As cluster techniques will always identify groups of data depending on the identification parameters, it is important to consider additional procedures to validate them. For supporting and providing the statistical proof of the resulting clusters, different approaches have been suggested. According to Handl *et al.* (2005), cluster validation measures can be distinguished into internal and external measures:

> External validation measures comprise all those methods that evaluate a clustering result based on the knowledge of the correct class labels... Internal validation techniques do not use additional knowledge in the form of class labels, but base their quality estimate on the information intrinsic to the data alone.
>
> (Handl *et al.* 2005: 3203)

With further extension to the work by Lames (1992), Rein *et al.* (2010b) applied an internal validation approach to their cluster analysis of basketball shooting. The phenomenon of phase transition in basketball hook shots with decreasing and increasing distances from nine metres to two metres and back to nine metres, with one-metres increments was investigated. The input variables for the cluster analysis were 12 angle variables derived from a 13-segment, rigid, three-dimensional body model. The clusters were interpreted in the terminology of systems dynamics as attractors with certain criteria. Only two of eight participants showed a clear expected phase transition behaviour. Importantly, in this preliminary study, it was possible to identify three distinctive shooting patterns with varying frequencies at different shooting distances.

As three different shooting patterns had been previously established, the external validation procedures were adopted by Rein *et al.* (2010a). Two studies in basketball served for testing the sensitivity of cluster analysis to pre-processing and for testing the phenomenon of phase transitioning in hook-shot technique. For the first experiment, four professional basketball players had to throw from three different distances with three different techniques. Owing to the impact of data normalization, the same analysis was performed with z-transformed (average = 0, standard deviation = 1) and raw data. Both sets were validated by means of bootstrapping and Hubert-Gamma method. Overall, in the first experiment, the cluster analyses led to 'entirely feasible' results and were able to reproduce a priori known differences between diverse movement patterns. In the second validation study, two basketball players were instructed to shoot baskets by means of a hook-shot technique from distances between two to nine metres. In contrast to Rein *et al.* (2010b), the task was limited by a lowered ceiling to force flight curves of the ball that are dominated by the velocity of release than by the angle of release, with the aim of causing a phase transition in the movement pattern with increasing or decreasing distance. Only one participant showed strong indications of the use of two distinct patterns whereas another participant displayed 'distinctively fewer differences' as shooting distance was manipulated. Bootstrapping and Hubert-Gamma values showed that a validation procedure is necessary for the confirmatory approach of cluster analysis.

The external validation approach pursues the probability of ending up with certain clusters relative to arbitrary or accidental data. The internal approach compares the variation within a cluster relative to the variation between clusters (Bauer and Schöllhorn 1997). Both approaches demonstrate the problem and importance of variable selection and preparation that is known in the context of analyzing complex self-organizing systems (Haken *et al.* 1995). According to Haken *et al.* (1995), the selection of collective variables or order parameters is highly dependent on the investigator's intuition. The problem seems to become even bigger with increasing complexity of the movement task and the possibilities of compensations. Yet there are no general rules for the choice and selection of variables as well as for the preparation of the data before cluster analysis.

In summary, cluster analysis provides a powerful dimension reduction tool which can serve different purposes depending on the nature (i.e. explorative or confirmatory) of the research undertaken. In contrast to cluster analysis, ANNs have the ability to separate two classes nonlinearly (Figure 8.2) and therefore lead to higher recognition rates and potentially more flexibility in their use (Haykin 1994; Schöllhorn and Jäger 2006). As we discuss in the next section, ANNs can be administered independently or in combination with cluster analysis.

Self-organizing maps

SOMs, also known as Kohonen maps, are a specific type of ANN that can be used to mathematically model specific characteristics of neuronal cell assemblies. In contrast to supervised learning ANNs such as multi-layer-perceptrons (MLPs), SOMs are trained using unsupervised learning to typically produce a two-dimensional discrete representation of the input space of the training samples, called a map. A big advantage of SOMs is the way in which it captures

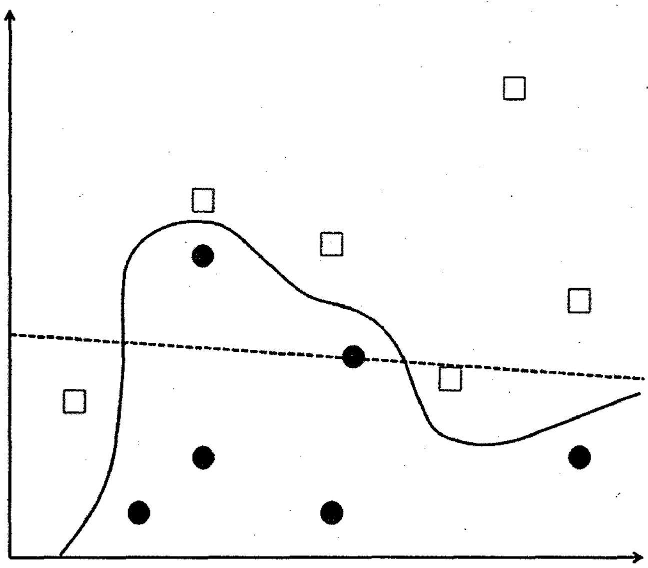

Figure 8.2 Linear versus nonlinear separation

low-dimensional views of high-dimensional data, akin to multidimensional scaling. SOMs are different from other ANNs because of their usage of a neighbourhood function that preserves the topological properties of the input space. Similar to most ANNs, SOMs operate in two phases. In the first phase, the map is trained using input examples. The second phase, called mapping, automatically classifies a new input vector. The competitive working process is also called vector quantization. Components of a SOM are called nodes or neurons that are associated with a weight vector of the same dimension as the input data vectors and a position in the map space. The usual arrangement of nodes is a hexagonal of rectangular grid. The procedure for placing a vector from input data space on to the map is to first find the node with the closest weight vector to the vector taken from input data space. Once the closest neuron is located, it is assigned the values from the vectors taken from the input data space. Mathematically, SOMs are sometimes associated with nonlinear forms of principal component analysis.

Owing to their capacity to map high-dimensional data to a low-dimensional map whilst preserving the topological characteristics of the original complex movement data, in principle, four different modes of application in the analysis of complex movement patterns can be distinguished (Figure 8.3A–D). In all approaches, an explorative strategy is pursued.

(A) In relation to describing a single movement by means of multiple time-discrete data, each movement is represented by a single vector that is mapped to a single node in the map space. By mapping several movements to the map space, SOMs function as classifiers that group similar movements to each other, owing to the characteristics of neighbourhood preservation.

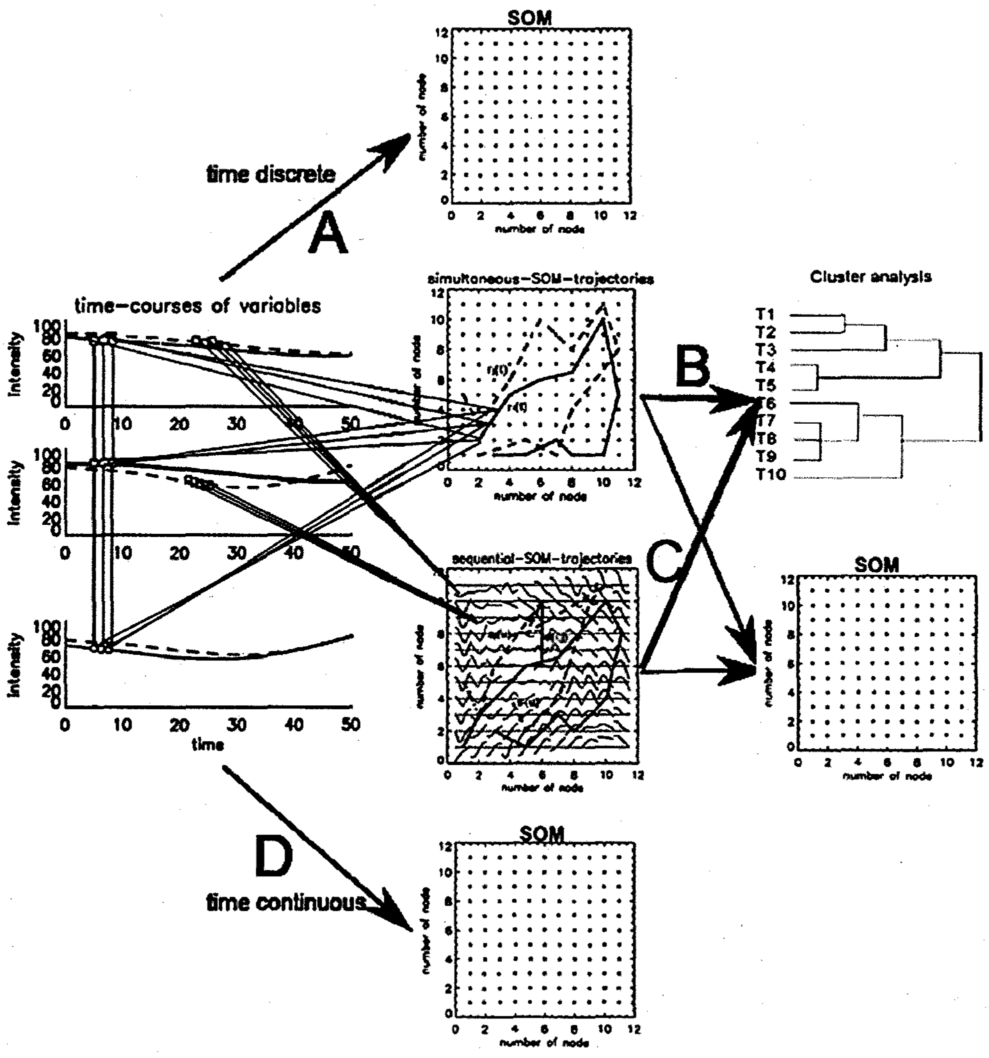

Figure 8.3 Variations of data processing by means of self-organizing maps

When the complex sport movement is described by means of time course data, three different approaches can be adopted. In all approaches, a single movement is described by means of several variable time courses.

(B) In the first time course oriented approach, the input vectors are constructed from all variable intensities at single instants. The number of vectors of a single movement is given by the number of instants the movement was measured with. Here, each feature vector is mapped to a single node in the map space and the whole movement is characterized by a two-dimensional trajectory in the map space (Figure 8.3).

(C) The second time course oriented approach considers each variable time-course as a single vector. Consequently, the number of input vectors per movement is given by the number of variables the movement was described with. The SOM maps each variable time course characteristics to a single node. Each resulting two-dimensional trajectory in the map space links the different time course characteristics of a single movement (Figure 8.3).

Once the movements are mapped to the node space (Figure 8.3B,C), the coordinates of the trajectories form new vectors, which can be either fed into another SOM or given into a cluster analysis to classify all movements by their relative similarity. When the trajectories are put into another SOM, the trajectories are mapped to different single nodes. The distribution of the nodes over the whole grid then displays the relative similarity of the movements in a two-dimensional space. However, these two-dimensional coordinates can be mapped with a cluster analysis to a linear grouping. When the coordinate vectors of the trajectories are directly put into a cluster analysis, the movements are grouped according to their one-dimensional similarity. In both cases (Figure 8.3B,C) when the SOMs are applied in a first-order form to the map, the movements on a trajectory in the low-dimensional node space the SOMs are taken for nonlinear data reduction comparable to linear analogies such as principal component analysis or Karhunen–Loève transformation. When SOMs are applied in the second-order form for the mapping of the trajectories, we can consider them as nonlinear classification procedures comparable to linear analogies such as cluster analysis.

(D) Instead of modelling the movement qualities by means of two SOMs or one SOM and a cluster analysis, a fourth approach is suggested. Here, all variable time courses of a single movement are put in a single, rather long, vector, which forms the input vector for the SOM. In this case, the SOM directly provides the two-dimensional map where each node corresponds to the model of a movement and the whole node grid provides the relative similarities of several movements.

In the following, the SOM research that is related to the area of movement execution in particular will be followed by the sport game related applications of SOMs.

To date, most applications of SOMs in elite sports can be found in the form of (B) or (C). An explorative combination of SOM and cluster analysis in form of (B) was applied by Bauer and Schöllhorn (1997). Two high-performance discus throwers were analyzed during different competition and training events over one year. The athletes' movement during the final throwing phase were described using a 14-segment, rigid body model by means of 13 joint angles and 13 angular velocities plus the trunk orientation angle and angular velocity, to achieve a physically complete description. Feature vectors of the variable time courses were mapped on a 10 × 10-node SOM. The resulting trajectories were subsequently placed into a cluster analysis. The results revealed, firstly, a clear differentiation of the two athletes. Secondly, they identified a successful intervention that resulted in a distinct separation of three trials before and five trials after the intervention period and, thirdly, sessions of training could be clearly separated from each other and provided evidence for day-dependent throwing strategies. However, the identification of signature movements for each discus thrower as well as continuous variations called traditional model- and repetition-oriented approaches into question.

Further evidence for the individuality of movement patterns in running patterns was found by Schöllhorn and Bauer (1997). By analyzing three to five double steps of 20 runners during a 2000-metre run by means of a SOM and cluster analysis, on the basis of the same variables as in Bauer and Schöllhorn (1997, it was possible to separate automatically contact phases of the right/left foot and the flight phase, as well as to identify the athletes by means of their kinematic data during a single ground contact phase. Identification was based on a 91 per cent recognition rate for the left foot and 96 per cent for the right foot but only 75 per cent for the flight phase.

The same approach with SOM and cluster analysis was pursued by Schöllhorn and Bauer (1998) in an investigation of world-class javelin throwers from different nations over several years in the search for individuality in elite sports. The same variables and data processing approach as in Bauer and Schöllhorn (1997) were used for describing and mapping the movements over a SOM towards a dendrogram. Most intriguingly, the results revealed the recognition of individual movement patterns of two female javelin throwers over several years. Obviously, the fingerprint-like throwing patterns were identifiable over several years. Furthermore, evidence was provided for a clear distinction of male and female throwing patterns as well as for nation-typical throwing patterns.

Two consecutive SOMs were applied to establish whether the order of movement executions affected variability over repeated trials (Janssen *et al.* 2010). Five participants performed ten first and ten second tennis serves in blocked and alternating (serial) series. The first serves were mainly accomplished with higher serving speed but less accuracy, whereas the second serves were mainly characterized by higher precision, lower serving speed. As the input dimension of the kinematic data (15 angles) was high, a self-implemented network called 2-SOM (because of its structure, which consisted of two series-connected SOMs) was chosen for analysis purposes. Within the 2-SOM, the first SOM was used to reduce data dimension, whereas the second SOM was used to classify the serve patterns. The 2-SOM revealed a person recognition rate of 100 per cent using all serves of all participants for only the blocked or serial condition. Within the 'person clusters', a network that was trained either with all individuals' data from the blocked or serial condition was able to distinguish between an individual's first and second serve with an accuracy of 99 per cent (in the blocked condition) and 90 per cent (in the serial condition), respectively. In general, movement variability was greater for the serial condition compared with the blocked condition over all participants.

Lees *et al.* (2003) reported the results of a study that applied SOMs with 12 × 8 neurons to analyze instep kicks by two soccer players for distance or accuracy.

The output patterns were repeatable for the same task for one player. The authors claimed that the trajectories in the output map were related to characteristics of the movement technique, although these characteristics were not determined. In a subsequent study, Lees and Barton (2005) used a similar approach to model several kicks by six soccer players. Fourteen joint angles were obtained from three-dimensional coordinates for 80 equispaced time instants from take off for the last stride to the end of the follow through of the kick. The output maps distinguished between right and left-footed groups. Intra-player differences were small.

The analysis of 60 golf chip shots of four low-handicap male golfers over distances of 4–24 metres in four-metre intervals, randomly assigned, was the objective of the investigation of Lamb *et al.* (2011). From a 16-segment model, a 24-dimensional input dataset consisting of joint angle time courses was derived. Because of high heterogeneity in the training data between the golfers, separate SOMs were trained on each player's respective kinematic data. Subsequently, a second SOM was trained using the best-matching nodes from the previous SOM for each respective player. The first SOMs varied between 24 × 16 and 25 × 15 nodes, the second SOMs had in general 32 × 1 nodes. A U-matrix representation for each player allowed a phase specific comparison within each player. The attractor layout diagrams were presented as evidence of nonlinear phase transitions for three of the four shot distances.

A confirmatory approach was pursued as well when data were reanalyzed from Bartlett *et al.* (2006) with respect to differences in treadmill running with marker and no-marker conditions (Lamb *et al.* 2008). Kinematic time course data were analyzed using a 29 × 23 hexagonal lattice. In general, the findings of Bartlett *et al.* (2006) could be replicated. However, SOMs do require more work to be applied properly. The ability to represent high-dimensional coordination patterns on visualizable low-dimensional map hold great potential (Bartlett *et al.* 2006).

In another study by this group, Lamb *et al.* (2010) showed how SOMs could be used to objectively distinguish between three types of basketball shots. Two SOMs were used in this investigation. Kinematic data from right and left ankles, knees, hips and shoulder were processed and input to the first SOM. Each variable was range normalized to maximum and minimum values of +1 and −1. The map size of the first SOM was 42 × 13 nodes and was used to analyze different phases of the movements, while the second SOM had 9 × 6 nodes and was used for the classification of the complete trials. There remains some debate over the best way to represent data (i.e. complete or broken down into key phases). In particular, the trial SOM was considered more relevant in providing a macro representation (see Lamb *et al.* 2010).

In contrast to this confirmatory approach, Beckmann *et al.* (2012) employed a more exploratory approach for the possibility of identifying individual-specific movement characteristics in different athletic events. Specifically, five decathletes were recorded using two high-speed video cameras and three-dimensionally reconstructed during the final throwing phases of shot put, discus and javelin movements while competing in the same decathlon. On the basis of a rigid 14-segment body model, time-normalized angle and angular velocity time courses formed the input vectors for the SOM and was compared with support vector machines. By inclusion of all variables, the disciplines could be distinguished at a 100 per cent level. When the throwing or shot put arm was excluded, no individual could be identified by means of the SOM but an individual recognition rate over all throwing disciplines by means of the support vector machines achieved at 98.5 per cent.

By coding the rallies as a series of zones, from where the players were hitting the volleyball, an intriguing approach was suggested by Perl and Lames (2000) that transferred the modelling by means of dynamic SOMs to game sports. Approximately 5000 rallies of men and women's volleyball matches of first and second German division formed the training set for the 20 × 20 node grid. However, the only plausible rallies that could be identified were those in which single nodes corresponded to characteristic events in a single rally; e.g. serve or offence smash. A similar approach has also been applied to squash (Perl 2002) and table tennis (Perl and Baca 2003).

Summary

Overall, cluster analysis and SOMs enjoy increasing popularity among movement scientists, owing to their capacity to explore and validate different qualities in movement science and match analysis. Both methods of analysis offer a fruitful basis for characterizing and interpreting high-dimensional datasets. The need to balance exploratory and confirmatory approaches in combination with time-continuous and time-discrete approaches is one of the biggest challenges for the coming years. The willingness to apply most recent methodological developments from related disciplines is growing and offers a promising wide new field of research.

References

Ball, K. A. and Best, R. J. (2007) Different centre of pressure patterns within the golf stroke I: Cluster analysis. *Journal of Sports Science*, 25: 757–70.

Bartlett, R., Bussey, M. and Flyger, N. (2006) Movement variability cannot be determined reliably from no-marker conditions. *Journal of Biomechanics*, 39: 3076–9.

Bauer, H. U. and Schöllhorn, W. (1997) Self-organizing maps for the analysis of complex movement patterns. *Neural Processing Letters*, 5: 193–9.

Beckmann, H., Janssen, D. and Schöllhorn, W. I. (2012) *Identifikation individueller disziplinübergreifender Bewegungsstile* [Identification of individual and discipline independent movement styles], in Bundesinstitut für Sportwissenschaft (ed.) *BISp-Jahrbuch Forschungsförderung* [Federal Institute of Sport Science (ed.) *Annual Report of Governmental Funded Research Projects*]. Bonn, Germany.

Chow, J. Y., Davids, K., Button, C. and Rein, R. (2008) Dynamics of movement patterning in learning a discrete multiarticular action. *Motor Control*, 12: 219–40.

Everitt, B. S., Landau, S. and Leese, M. (2001) *Cluster Analysis*, 4th edn. New York: Arnold.

Haken, H., Wunderlin, A. and Yigitbasi. A (1995) An introduction to synergetics. *Open Systems and Information Dynamics*, 3 (1): 97–130.

Handl, J., Knowles, J. and Kell, D. B. (2005) Computational cluster validation in post-genomic data analysis. *Bioinformatics*, 21 (15): 3201–12.

Hay, J. G. (1993) *The Biomechanics of Sports Technique*, 4th edn. Englewood Cliffs, NJ: Prentice Hall.

Haykin, S. (1994) *Neural Networks: A Comprehensive Foundation.* New York: Macmillan.

Jaeger, R. G. and Halliday, T. R. (1998) On confirmatory versus exploratory research. *Herpetologica*, 54: S64–6.

Jäger, J., Schöllhorn, W. I. and Schwerdfeger, B. (2003) A pattern recognition approach for an opponent specific classification of tactical moves in team sports, in E. Müller, H. Schwameder, G. Zallinger and V. Fastenbauer (eds) *Proceedings of the 8th Annual congress of the European College of Sport Science*. Salzburg: Institute of Sport Science, pp. 370.

Jaitner, T., Mendoza, L. and Schöllhorn, W. I. (2001a) Analysis of the long jump technique in the transition from approach to takeoff based on time-continuous kinematic data. *European Journal of Sport Science*, 1: 1–12.

Jaitner, T., Ernst, H., Mendoza, L., Schöllhorn, W. I. (2001b) Changes of EMG patterns during motor learning of ballistic movements, in: H. Gerber and R. Müller, R. (eds) Proceedings of the XVIIIth Congress of the International Society of Biomechanics (ISB). Zürich: ETH.

Janssen, D., Gebkenjans, F., Beckmann, H. and Schöllhorn, W. I. (2010) Analyzing learning approaches by means of complex movement pattern analysis. *International Journal of Sport Psychology*, 41 (Special Issue): 18–21.

Johansson, G. (1973) Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14: 201–11.

Kelso, J. A. S. (1995) *Dynamic Patterns: The Self-Organization of Brain and Behavior.* Cambridge, MA: MIT Press.

Lamb, P., Bartlett, R. M., Robins, A. and Kennedy, G. (2008) Self-organizing maps as a tool to analyze movement variability. *International Journal of Computer Science in Sport*, 7: 28–39.

Lamb, P., Bartlett, R. and Robins, A. (2010) Self-organising maps: an objective method for clustering complex human movement. *International Journal of Computer Science in Sport*, 9: 20–9.

Lamb, P. F., Bartlett, R. M. and Robins, A. (2011) Artificial neural networks for analyzing inter-limb coordination: the golf chip shot. *Human Movement Science*, 30: 1129–43.

Lames, M. (1992) Synergetik als Konzept in der Sportmotorik. [Synergetics as a concept for sport movements]. *Sportpsychologie*, 6 (3): 12–18.

Lees, A., Barton, G. and Kershaw, L. (2003) The use of Kohonen neural network analysis to qualitatively characterize technique in soccer kicking. *Journal of Sports Sciences*, 21: 243–4.

Lees, A. and Barton, G. (2005) A characterisation of technique in the soccer kick using Kohonen neural network analysis, in T. Reilly, J. Cabri and D. Araújo (eds) *Science and Football V: The Proceedings of the Fifth World Congress on Science and Football.* London: Routledge, pp. 83–8.

Müller, E. (1986) *Biomechanische Analyse alpine Schilauftechniken*. [Biomechanical analysis of skiing techniques.] Innsbruck: Inn-Verlag.

Perl, J. (2002) Game analysis and control by means of continuously learning networks. *International Journal of Performance Analysis of Sport*, 2: 21–35.

Perl, J. and Baca, A. (2003) Application of neural networks to analyze performance in sports, in E. Müller, H. Schwameder, G. Zallinger, V. Fastenbauer (eds) *Book of*

Abstracts of the 8th Annual Congress of the European College of Sport Science in Salzburg, Austria, 9–12 July 2003. Available online at www.informatik.uni-mainz.de/dycon/ABS_2003__Perl_Baca_Appl_NN.pdf (accessed 7 June 2013).

Perl, J. and Lames, M. (2000) Identifikation von Ballwechselverlaufstypen mit Neuronalen Netzen am Beispiel Volleyball. [Identification of rallies in Volleyball by means of SOMs], in W. Schmidt and A. Knollenberg (eds) *Sport – Spiel – Forschung: Gestern. Heute. Morgen*. Hamburg: Czwalina, pp. 211–16.

Rein, R., Button C., Davids, K. and Summer, J. (2010a) Cluster analysis of movement patterns in multiarticular actions: a tutorial. *Motor Control*, 14: 211–39.

Rein, R., Davids, K. and Button, C. (2010b) Adaptive and phase transition behavior in performance of discrete multi-articular actions by degenerate neurobiological systems. *Experimental Brain Research*, 201: 307–22.

Schöllhorn, W. I. (1993) *Biomechanische Einzelfallanalyse im Diskuswurf*. [Biomechanical single case study of discus throwing]. Frankfurt: Harri Deutsch.

Schöllhorn, W. I. (1995) Time course oriented analysis of biomechanical movement patterns by means of orthogonal reference functions. Paper presented at the XVth Congress of the International Society of Biomechanics (ISB) 2–6 July, Jyväskylä.

Schöllhorn, W. I. (1999) Complex individual movement styles identified by means of a simple pattern recognition method, in P. Parisi, F. Pigozzi and G. Prinzi (eds) *Book of Abstracts of the 4th Annual Congress of the European College of Sport Science, Rome, Italy, 14–17 July 1999*. Cologne: European College of Sport Science.

Schöllhorn, W. I. and Bauer, H. U. (1997) Linear–nonlinear classification of complex time course patterns, in J. Bangsbo, B. Saltin, H. Bonde, Y. Hellsten, B. Ibsen, M. Kjaer and G. Sjogaard (eds) *Conference Proceedings of the 2nd European College of Sport Science*. Copenhagen: University of Copenhagen, pp. 308–9.

Schöllhorn, W. I. and Bauer, H. U. (1998) Identifying individual movement styles in high performance sports by means of self-organizing Kohonen maps, in H. Riehle, and M. Vieten (eds) *XVI International Symposium on Biomechanics in Sports*. Konstanz: Universitätsverlag, pp. 574–77.

Schöllhorn, W.I. and Jäger, J. M. (2006) A survey on various applications of artificial neural networks in selected fields of healthcare, in R. Begg, J. Kamruzzaman and R. A. Sarker (eds) *Neural Networks in Healthcare: Potentials and Challenges*. Hershey, PA: Idea Group Inc., pp. 20–58.

Schorer, J. Baker, J. Fath, F. and Jaitner, T. (2007) Identification of interindividual and intraindividual movement patterns in handball players of varying expertise levels. *Journal of Motor Behavior*, 39: 409–21.

Slife, B. D. and Williams, R. N. (1995) *What's Behind the Research? Discovering Hidden Assumptions in the Behavioral Sciences*. Thousand Oaks, CA: Sage Publications.

Tukey, J. W. (1980) We need both exploratory and confirmatory. *American Statistician*, 34 (1): 23–5.