USING UNSUPERVISED LEARNING TO CHARACTERIZE MOVEMENT PATTERNS – AN EXPLORATIVE APPROACH

Sina David[1] and Gabor Barton[2]

[1]Department of Human Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[2]Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom

The purpose of this study was to explore the ability of Kohonen neural network self-organizing maps (SOM) to visualize and characterize different movement patterns during sidestepping. The marker trajectories of 631 sidestepping trials were used to train a SOM. Out of 63731 input vectors, the SOM identified 1250 unique stick figures, determined by the markers. Visualizing the movement trajectories and adding the latent parameter time, allows for the investigation of different movement patterns. Additionally, the SOM can be used to identify zones with increased injury risk, by adding more latent parameters which opens the option to monitor athletes and give feedback. The results highlight the ability of unsupervised learning to visualize movement patterns and to give further insight into an individual athlete's status without the necessity of a-priory assumptions.

KEYWORDS: Kohonen self-organising map, neural network, sidestepping.

INTRODUCTION: Recent studies showed the impact of individual movement patterns on the injury-related structural loadings during fast sidestepping manoeuvres (David et al., 2018; Dempsey et al., 2007; Donnelly et al., 2017). The results provided a clear indication that preparatory strategies such as trunk inclination, preorientation and foot strike pattern determine the knee valgus moment which is a proxy for anterior cruciate ligament (ACL) load. However, despite this knowledge, successful prevention could not be achieved (Bahr, 2016). The nature of injuries is complex and is influenced by internal and external risk factors in addition to the inciting event with all its factors (van Mechelen et al., 1992). Modern lab technology allows researchers to investigate such complex problems from different perspectives. 3D movement analysis systems can capture accurate positions and orientations of the body and the use of inverse dynamics can combine the movement data with external forces to give insight into joint loading. Additional features can be added to further increase the amount of data and to give a bigger picture. This results in a large amount of multidimensional data which is a big advantage of modern data acquisition, but the problem arises of how to process and connect all the data. The human brain can receive, process, and remember between seven to nine items at a time (Miller, 1956). When reviewing the large datasets gathered through motion analysis, only the lower limb joint angles of the right leg in three dimensions already exceed this number of items. So, the researcher must decide on a preselection of parameters. These assumptions are dependent on the profession or experience of the investigator (Skaggs et al., 2000; Watts, 1994). In sports biomechanics, the selection is often based on the mechanical understanding of the load-tissue interaction. However, every movement is the result of the interplay of mechanics, tissue characteristics, motor control, psychology, and others. It might therefore be impossible to identify a single cause that forces an athlete into a movement pattern with an elevated risk of injury while others will end up using a different pattern in the same situation. Additionally, the transition from one movement pattern into another is continuous - as are the risk factors, which makes data analysis even more complex and the grouping of athletes to further investigate them is prohibited (Bahr, 2016).

To summarise, the state-of-the-art motion analysis approaches result in high-quality datasets. The conventional ways to approach these datasets require a reduction of the amount of data and often also force the researcher to make a-priori assumptions. This may result in a narrowed view of the dataset and eventually lead to overestimating the importance of single parameters together with loss of essential information. One solution to the raised issues are neural networks by means of Kohonen self-organizing maps (Kohonen, 2001). These unsupervised neural networks can overcome the described problems as they can process large quantities of input data without a-priory assumptions (Barton, 2006). The aim of this study was therefore to explore the ability of SOMs to visualize movement patterns during sidestepping.

METHODS: This study used the lower body marker trajectories dataset of athletes performing planned full-effort 90° sidestepping to train an unsupervised neural network by means of Kohonen self-organizing maps.

The dataset contained 631 trials of 67 athletes in total (adults: 26 male, 31 female, age 22.6 years ± 3.3, height 1.77 m ± 0.1, mass 70.9 kg ± 0.1; children: 10, age 9.8 years ± 1.0, height 1.45 m ± 0.1, mass 36.98 kg ± 6.38). All participants were free of injury or pain and gave their written informed consent to participate in the study. Ethical approval for the study was given by the University's ethical committee.

The trajectories of 28 lower body retro-reflective markers were captured by 14 infrared cameras (200 Hz, VICON, Oxford, UK) and filtered with a recursive 2nd order low pass filter and a cut-off frequency of 20 Hz.

An unsupervised Kohonen self-organizing map (SOM) neural network was trained with the data (Kohonen, 2001) using the SOM Matlab Toolbox (V 2.0). The input layer consisted of a matrix of 63731 input vectors [63731 x 84, 631 trials x 101 data points (631x101=63731) and 28 marker trajectories with 3 dimensions each (28x3=84)]. The Kohonen layer consisted of 1250 neurons (determined by the size of the input matrix), distributed over a 50 x 25 rectangular map with hexagonal neighbourhood topology of each neuron. The initial connection weights of the neurons are set by the first principal component of the input dataset. The following section describes the training process of the SOM:

(1) Normalization of the input vector. (2) To train the SOM, the algorithm iteratively chooses one random input vector X and calculates the Euclidian distance ε to all neurons of the Kohonen layer. The neuron with the smallest Euclidian distance ε is called the winner neuron. (3) The weight W of the winner neuron r is adjusted by using the following equation:

$$W_r^{new} = W_r^{old} + \varepsilon\,(X - W_r^{old})$$

Geometrically this means that the weight vector W of the winner neuron is moving one step of the length of ε closer towards the chosen input vector. A Mexican hat shaped function ensures that the weights of neighbouring neurons are reduced. (4) The learning procedure is stopped if no further significant changes of the weight vectors are detected.

RESULTS: The SOM is a 2D graphical representation of the multidimensional input vectors. As the input was marker trajectories, each neuron's weights can be displayed as a stick figure (Fig. 1A). Out of the 63731, the SOM reduced the input to 1250 unique stick figures. As neurons that are next to each other have a higher similarity than neurons that are far away, there is a continuous change of body postures when travelling over the SOM. One result of the SOM's training process is the location of the winning neuron that is closest to each of the input vectors. Therefore, it is possible to visualize the movement trajectory of each input trial over the SOM map and to explore the sequence of different movement patterns (Fig. 1B). Due to the connection between the SOM's neurons and the input vectors, it is also possible to label the SOM with additional, latent information. Figure 1C shows the distribution of the movement cycle (%) over the map. With this, it becomes clear, that the blue trajectory of figure 1B can be characterized with a higher number of neurons associated with the weight acceptance phase (WAP) than for example the red trajectory.

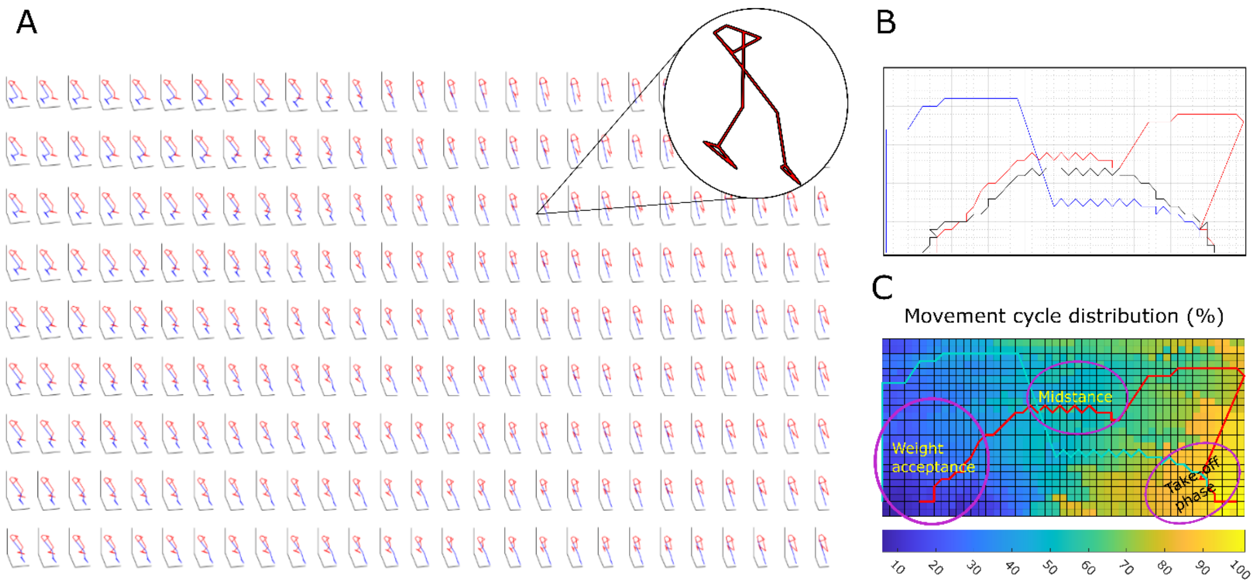

Figure 1: SOM interpretation process. A) Distribution of a reduced subset of body poses stored in the weights of the trained SOM, visualized as stick figures. B) Three selected movement trajectories travelling over the SOM map. C) SOM map after adding the latent parameter time.

DISCUSSION: The visualization of the marker trajectories during sidestepping using a Kohonen self-organising map allows for a new way to investigate individual movement patterns and the associated research questions. The SOM itself can be used to understand the distribution of the different postures along the individual movement trajectories. The option to add other latent parameters such as time gives insight into the proportion of time taken within the different phases of the movement. This does not mean that an athlete spends more time for the weight acceptance phase, but the resolution of the SOM for this athlete is higher in this phase (indicated by the number of neurons in this phase), meaning that the differences between the single input vectors are higher for this athlete than for the other examples chosen for figure 1C. This is directly linked to the movement pattern. By reconstructing the stick figures (Fig 1A) along the three trajectories in figure 1B, it can be observed that the athlete represented by the blue curve shows more knee flexion of the execution leg than the other two athletes, suggesting that a higher number of different postures were achieved. Adding the lower limb joint angles as latent variables as done with time, could be one way of getting deeper insight into these individual movement patterns. It has to be mentioned that time or joint angles were not included in the input vectors, thus the smooth distribution of the timing (Fig. 1C) shows the quality of the learning process.

Taking this one step further, labels that are associated with injury risk such as joint moments can also be added as latent variables. With this, it is likely to identify zones on the map, where several risk factors fall together, such as a high knee valgus moment during the weight acceptance phase. A movement trajectory that is travelling through this zone could identify an athlete with a higher risk of injury. This approach is completely different from the screening tools that are commonly used. There is no need to define thresholds for joint loading or other risk factors, which was critically discussed recently (Bahr, 2016). An athlete that is closer to the centre of a risk zone will be more likely to be exposed to injury relevant load. Also, there is no need to group athletes according to any hypotheses, which always contains the risk of missing effects or choosing a grouping variable that is not discriminative. Therefore, SOMs are a sensitive tool, to highlight individual movement patterns and given their mode of operation offer the identification of risk.

The huge advantage of the proposed method is the gradual change that is distributed across the map. With this, the effect of small changes in the movement on the overall movement pattern can be visualized. Another big benefit of the proposed method is its flexibility. If the research question focuses on the ankle instead of knee injuries, the same map can be used but labelled with ankle relevant latent parameters. Also, the SOM can be used to monitor athletes. The advances in pose estimation would allow feeding the network with other features such as joint angles or even parameterised images. For example, this could be used for direct feedback to the athlete in form of warnings if risky postures are detected repeatedly.

CONCLUSION: Kohonen self-organizing maps are a useful method to investigate the outcome of different movement patterns without the need for assumptions or grouping of athletes according to global parameters. They can highlight how small adaptations in the movement pattern influence the movement pattern. In combination with the advances in posture estimation, they can be used for feedback training or athlete evaluation.

REFERENCES

Bagwell, J. J., Fukuda, T. Y., & Powers, C. M. (2016). Sagittal plane pelvis motion influences transverse plane motion of the femur: Kinematic coupling at the hip joint. *Gait & Posture*, *43*. https://doi.org/10.1016/j.gaitpost.2015.09.010

Bahr, R. (2016). Why screening tests to predict injury do not work-and probably never will…: a critical review. *British Journal of Sports Medicine*, *50*(13), 776–780. https://doi.org/10.1136/bjsports-2016-096256

Barton, G.J., Lees, A., Lisboa, P., Attfield, S. (2006) Visualisation of gait data with Kohonen self organising neural maps. Gait and Posture. 24/1: 46-53.

David, S., Mundt, M., Komnik, I., & Potthast, W. (2018). Understanding cutting maneuvers – The mechanical consequence of preparatory strategies and foot strike pattern. *Human Movement Science*, *62*, 202–210. https://doi.org/10.1016/j.humov.2018.10.005

Dempsey, A. R., Lloyd, D. G., Elliott, B. C., Steele, J. R., Munro, B. J., & Russo, K. A. (2007). The effect of technique change on knee loads during sidestep cutting. *Medicine & Science in Sports & Exercise*, *39*(10), 1765–1773. https://doi.org/10.1249/mss.0b013e31812f56d1

Donnelly, C. J., Chinnasee, C., Weir, G., Sasimontonkul, S., & Alderson, J. (2017). Joint dynamics of rear- and fore-foot unplanned sidestepping. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *20*(1), 32–37. https://doi.org/10.1016/j.jsams.2016.06.002

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2). https://doi.org/10.1037/h0043158

Skaggs, D. L., Rethlefsen, S. A., Kay, R. M., Dennis, S. W., Reynolds, R. A. K., & Tolo, V. T. (2000). Variability in gait analysis interpretation. *Journal of Pediatric Orthopaedics*, *20*(6), 759–764.

van Mechelen, W., Hlobil, H., & Kemper, H. C. G. (1992). Incidence, Severity, Aetiology and Prevention of Sports Injuries. *Sports Medicine*, *14*(2). https://doi.org/10.2165/00007256-199214020-00002

Watts, H. G. (1994). Gait laboratory analysis for preoperative decision making in spastic cerebral palsy: is it all it's cracked up to be? *Journal of Pediatric Orthopedics*, *14*(6), 703–704.