IDENTIFICATION OF EMG FREQUENCY PATTERNS IN RUNNING BY WAVELET ANALYSIS AND SUPPORT VECTOR MACHINES

Thomas Jaitner[1], Daniel Janssen[2], Ronald Burger[2] and Uwe Wenzel[3]

Institute for Sports Science, University of Augsburg, Augsburg, Germany[1]
Institute for Sports Science, Johannes Gutenberg University, Mainz, Germany[2]
Faculty of Sports Science, University of Leipzig, Leipzig, Germany[3]

The purpose of this study was to identify EMG pattern of running at different speed and incline based on a trial-to-trial analysis. Eight subjects performed treadmill running at five different conditions (4, 5 and 6 m/s, 5m/s at 5° incline, 5m/s at 2° decline). EMG data of eight leg muscles were recorded and transformed by a wavelet analysis (van Tscharner, 2000). Ten subsequent steps of each subject and condition were classified by support vector machines. Between 93 and 100% of all EMG patterns were assigned correctly to the individual. According to the different running conditions recognition rates ranged between 78 and 88%. Hence, support vector machines can be considered as powerful nonlinear tool for the classification of dynamic EMG patterns.

KEYWORDS: pattern recognition, support vector machines, electromyography, running.

INTRODUCTION: The electromyographic activity of a single muscle is considered as a complex stochastic signal that results from the superposition of the electrical activity of several motor units and therefore shows a high trial-to-trial variability. In dynamic movements such as running, the coordination of muscles with similar function (e.g. extension of the knee) can enhance this variability since the same movement outcome might be produced by different activities of single muscles. The most common approaches for the analyses of the surface electromyogram (EMG) aim on the extraction of the essential contents of the signal by averaging over time and trials (De Luca, 1997; Hermens et al 1999). In running, such techniques have been applied to analyze the EMG of the leg muscles at different speed (e.g. Kyröläinen et al. 2005, Gazendam & Hof 2007) and incline (e.g. Swanson & Caldwell 2000). However, there are two major disadvantages of these techniques. Calculating the signal mean over time (e.g. by root mean square) allows quantifying the overall intensity but does not provide any process-related information about the changes of EMG intensity during performance. If the EMG is averaged over trials, signal variations from trial to trial are primarily considered as noise and therefore neglected. It is implicitly assumed that differences between groups of trials (e.g. according to varying running conditions) must reflect in the average activity of a single muscle. Variations of the EMG that result from compensatory activities of different muscles cannot be obtained and hence the interplay between muscles cannot be analyzed.
In this paper a different approach was chosen that takes benefit of the variability of the EMG signal and considers the interplay in muscular activity between different muscle groups. The main objective was to identify EMG patterns for running at different speed and incline.

METHOD: Eight track and field athletes (age 18.6 years ±2.4; 5 male/3 female) participated in this study. All subjects were free from recent lower extremity injury or pain and trained regularly for at least 2 years. After a five minute warm up the subjects were asked to run five times 200 m on a treadmill at different conditions (4, 5 and 6 m/s, 5m/s at 5° incline, 5m/s at 2° decline). Recovery periods between the trials were chosen by the subjects individually and lasted normally approximately one minute. EMG of eight muscles of the right limb were recorded at 2400 Hz using bipolar surface electrodes (AMBU 720 00-S) and single differential amplifiers (BIOVISION). The following muscles were considered: M. gastrocnemius medialis, M. lateralis, M. soleus, M. tibialis anterior, M. biceps femoris, M. rectus femoris, M. vastus medialis and M. vastus lateralis. The raw signals of each EMG channel were amplified by a factor of 2000. A band-pass filter with a bandwidth from 10 to 700 Hz was applied. An accelerometer was fixed on the subjects' right shoe and was used to determine the time of heel strike by a rapid change in acceleration. All data were collected synchronously and stored on a PDA that the subjects carried on their back (fig. 1). For each condition, the EMG of ten consecutive double steps were cut off and analyzed step by step.

For the analysis of dynamic contractions, the non-stationarity of the EMG signals must be considered as this might cause errors in the time as well as in the frequency domain. Therefore, the EMG data were preprocessed by a wavelet analysis (van Tscharner, 2000). The wavelet-transformation is a suitable method to analyse non-stationarity bio-signals simultaneously in time- and frequency-domain. The wavelet transformation of an EMG Signal is performed as convolution of the signal with the wavelet. A filter bank of 11 non-linearly scaled wavelets was used that has been especially developed for EMG application (van Tscharner, 2000, 2002).

All-in-all 378 movement patterns were analyzed. Every single movement pattern is represented by a *n x D* matrix, where *n* is the number of acquired data vectors during the stride length and *D* is the dimension of the data vectors. Each data vector consists of 88 features, as every EMG of the 8 muscles was transformed into 11 wavelets. In order to cope with the huge amount of data, dimension reduction was necessary. Dimension reduction tries to eliminate redundancy from the data by so called feature extraction. The dimension *D* is mapped to a lower dimension *d* while trying to retain the geometry of the data as much as possible. Several linear and nonlinear methods exist for this purpose. Two possibilities can be applied to reduce the matrix in which a single movement pattern is stored. The first possibility is to reduce the dimension of the features. This will lead to a matrix with *n* acquired data vectors that transport some encapsulated information on the whole movement. In the other case the dimension reduction is conducted over the time. This technique is for example used in gait analysis with kinematic position data (Troje, 2002). For example, if a Principal Component Analysis is applied to the data, the movement is transferred into a low-dimensional space spanned by the first (few) so-called eigenpostures of the walker. Similar to this approach multidimensional scaling (MDS; Cox & Cox, 1994; Kruskal, 1964) was used as a nonlinear reduction technique. MDS tries to retain the pair wise distances between the data points as much as possible during the mapping of the data to a lower dimension. The quality of the dimension reduction is expressed as a so-called stress function that is a measure of the error between the pair wise distances in the low-dimensional and high-dimensional data space. The main goal of the mapping process is hereby the minimization of the stress function (Cox & Cox, 1994). In several tests an intrinsic dimension of 4 was estimated as optimal using Matlab and a proper toolbox (Van der Maaten, 2007). Hence, the data matrices were reduced to 4 x 88 matrices. A second (so called two-fold) reduction was omitted in first instance.

After dimension reduction support vector machines (SVM; Vapnik, 1995; Chang & Lin, 2001) were used for the classification of the movement patterns. SVMs are supervised machine learning methods used for classification and regression, dealing successfully with small datasets and finding global minima (Bennett & Campbell, 2000). After compulsory amplitude normalization of the data, SVMs were trained with the bigger part of the data linked with the associated class memberships (i.e. person; running speed), and tested with the remaining data in order to calculate rates of how well those patterns were linked with the correct classes, that were excluded from the training process. This was conducted using cross validation (Jain, Duin & Mao, 2000), a standard technique to ensure more precise recognition rates and to avoid overtraining.

RESULTS: Table 1 shows the recognition rates for EMG patterns according to the individual subject. Overall, the recognition rates ranged between 92,9% and 100%. Best results were found for level running at different speeds, where all EMG pattern were assigned correctly to each subject. Similar recognition rates were found if only the two different incline conditions were considered.

Table 1. Recognition of Individual EMG Patterns

Condition	Recognition rate
Level running at 4, 5 and 6m/s (238 trials)	100%
Running at 5m/s [+5°/±0°/-2°] (220 trials)	97.7%
Slope running [+5°/-2°] (140 trials)	99.3%
All trials (378 trials)	92.9%

Recognition rates observed for the running speed and incline conditions are listed in Table 2. About 88% of all EMG patterns were classified correctly if all trials at a running speed of 5 m/s were analyzed. This includes level running as well as runs at an incline of 5° and a decline of -2°. This sample shows an even better recognition rate as for the incline conditions only (82,1%). Lowest rates were achieved for level running. Here, about 78% of all trials were assigned to the correct running speed.

Table 2. Recognition of Speed and Incline

Condition	Recognition rate
Level running at 4, 5 and 6m/s (238 trials)	78.6%
Running at 5m/s [+5°/±0°/-2°] (220 trials)	88.2%
Slope running [+5°/-2°] (140 trials)	82.1%

DISCUSSION: The single EMG patterns embody highly individual characteristics that remain stable for each subject more or less independent from speed and incline. The individual recognition rates are far beyond chance level and reach values that have been reported in previous studies for less variable kinematic data (e.g. Jaitner et. al. 2001). Moreover, specific patterns for different running speeds and inclines can be identified with high probability. This is even more remarkable since the incline differs only slightly from level running. It is therefore assumed that the muscular activity of the leg muscles during running adapts very sensitively to environmental changes.

From a methodological view, two specific aspects of the muscular activity can be addressed within this analysis, that at the first sight seem contrarily: a high individuality of muscular activity pattern that seems widely independent from changes in speed and incline and on the other hand specific pattern that remain stable for certain conditions. This highlights some critical aspects of traditional approaches in EMG analysis described in the introduction. Overall, support vector machines can be considered as powerful nonlinear tool for the classification of dynamic EMG patterns.

If the EMG patterns for a specific running condition (e.g. running at a speed of 5m/s) remain stable within the subject but differ substantially between subjects this has considerable impact on the interpretation of EMG data. Practical implications that result from the comparison of different subjects might be misleading and should be drawn with particular care. Hence, approaches that focus on the analysis of multiple trials of the same subjects might be more reliable.

An emphasis of this study was on the complex interplay between different muscles in running. The results indicate that compensatory muscular activities could be a key factor for the overall stability of the EMG patterns. The analysis of the interaction between various leg muscles therefore might provide a more detailed insight in the mechanism of running coordination. However, further research is needed to allow a better understanding of the intermuscular coordination in complex movement patterns.

REFERENCES:

Bennett, K.P. & Campbell, C. (2000). Support vector machines: hype or hallelujah? *SIGKDD Explorations*, 2(2), 1-13.

Chang, C.C., & Lin, C.J. (2001). *LIBSVM: a library for support vector machines*. Software.

Cox, T. & Cox, M. (1994). *Multidimensional scaling*. London: Chapman & Hall.

De Luca, C.J. (1997) The use of surface EMG in biomechanics. *Journal of Applied Biomechanics*, 13, 135-163

Gazendam, M.G.J., Hof, A.L. (2007) Averaged EMG profiles in jogging and running at different speeds. *Gait & Posture*, 25, 604-614

Hermens, H.J., Freriks, B., Merletti, R., Stegemann, D., Blok, J., Rau, G. et al. (1999) *European Recommendations for Surface Electromygraphy – Results of the SENIAM Projects*. Enschede: Roessingh Research Developement

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4-37.

Jaitner, T., Mendoza, L., Schöllhorn, W.I. (2001) Analysis of the Long Jump Technique in the Transition From Approach to Takeoff Based on Time-Continuous Kinematic Data. *European Journal of Sports Science*, 1(5), 1-12

Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness offit to a nonmetric hypothesis. *Psychometrika*, 29, 1-27.

Kyröläinen, H., Avela, J., Komi, P.V. (2005) Changes in muscle activity with increasing running speed. Journal of Sports Sciences, 23 (10), 1101-1109

Lau, H., Tong, K., & Zhu, H. (2009). Support vector machine for classification of walking conditions of persons after stroke with dropped foot. *Human Movement Science*, 28(4), 504–514.

Swanson, S. C., Caldwell, G. E. (2000) An integrated biomechanical analysis of high speed incline and level treadmill running. Medicine and Science in Sports and Exercise, 32 (6), 1146-1155.

Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2, 371-387.

Van der Maaten, L. J. P (2007). *The nature of statistical learning theory: An introduction to dimensionality reduction using matlab. Technical Report MICC-IKAT 07-06*. Maastricht University, Maastricht, The Netherlands.

von Tscharner, V. (2000). Intensity analysis in time-frequency space of surface myoelectric signals by wavelets of specified resolution. *Journal of Electromyography and Kinesiology*, 10(6), 433-445.

von Tscharner, V. (2002). Time-frequency and principal-component methods for the analysis of EMGs recorded during a mildly fatiguing exercise on a cycle ergometer. Journal of Electromyography and Kinesiology, 12(6), 479-492.

Vapnik, V. (1995). *The nature of statistical learning*. New York, NY: Springer.