Environment. Technology. Resources, Rezekne, Latvia
Proceedings of the 11th International Scientific and Practical Conference. Volume II, 147-153

The Good, the Bad and the Unrecognized: Smart Textile Signal Clustering by Self-Organizing Map

Artyom Shootov, Yuriy Chizhov, Aleksandr Oks, Alexei Katashev
Riga Technical University

Abstract. ***The present article is a series of publications dedicated to the research of smart fabric sensors integrated into socks and is also part of the project aimed at developing the measuring system based on smart fabric supplied with sensors and intellectual data processing. The aim of the article is to perform a practical study on the application of Self-Organizing Map to smart textile signal clustering. Within the framework of the research, different approaches to the organization of network training are explored. A method for encoding an input pattern is also proposed. It has been established that the network is able to recognize the signal as a good step, a bad step, and an unrecognized step. The primary classification allows further selecting specific algorithms for a detailed analysis of good steps and bad steps. The detailed analysis of bad steps is the key to solving the problem of revealing of an athlete's special type of fatigue, leading to injuries.***

Keywords: DAid® Pressure Sock System, self-organizing map, smart textile signal clustering.

I. INTRODUCTION

DAid® Pressure Sock System [1] measures and transmits a signal proportional to the pressure exerted by the foot on the sensors. In general, the system opens wide opportunities for investigating the processes of interaction of the user of smart fabric with any other objects, for example, in medicine at the stage of rehabilitation of patients or for monitoring the mobility of bedridden patients; in various sports – in equestrian sport (to control the rider's position in the saddle), dancesport, basketball (analysis of the coverage of the ball with the wrist) and in many other areas. The system transmits observations simultaneously from all sensors at predetermined intervals (for example, there are ten sensors in smart socks).

However, there is a substantial gap between the idea of signal collection (and even its physical implementation) and the operation of the system. It is due to the fact that the system must be able to recognize the signal. The signal may have noise, it is susceptible to the influence of the human body mass, tissue displacement is also not excluded, etc. Thus, one of the challenges is the analysis of the received signal. The ultimate goal of the analysis is to determine the process parameters specific to each area. For example, in case of running, as a result of analyzing the signal from the socks, it is expected to reveal such signal characteristics that would indicate the athlete's fatigue level.

As it is known, observations from clinical studies have estimated that over 60 % of running injuries could be attributed to training errors. In fact, it can be stated that all overuse running injuries are a result of training errors. An individual who has sustained an overuse running injury must have exceeded his/her limit of running distance and/or intensity in such a way that the remodeling of the injured structure predominated over the repair process due to the stresses placed on the structure [2].

On the way to the ultimate goal of identifying process parameters, there is the task of primary signal classification. Within the framework of the research, the authors investigate the use of Self-Organizing Maps (SOMs, also known as Kohonen Network) for signal clustering into the acceptable step ("the good" step that is correctly performed from the subjective point of view of the authors of the study), the unacceptable step ("the bad") and something else that cannot be called even a ugly step ("the unrecognized"). Unacceptable steps are the steps that are not executed correctly (in comparison with the "good" ones). Objective evaluation of a step is the subject of further research, once it is determined that signals of DAid® Pressure Sock System can be divided into these three classes.

The paper explores various approaches to the organization of network training. A manual way of network training is proposed. Experiments on the clustering of the signal obtained during a real run are carried out. The experiments aim at finding out whether the signal is distinguishable and determining the accuracy levels at which it is possible to classify patterns that do not participate in network training.

ISSN 1691-5402

http://dx.doi.org/10.17770/etr2017vol2.2567

The Kohonen network has been chosen owing to its simple mathematical model, which does not require significant computation and can be implemented in portable devices. The SOMs do not require a training sample. This will allow in the future, on their basis, to create systems that themselves adapt to the ordinary human step. This is a significant advantage over similar solutions.

II. OVERVIEW OF CONTEMPORARY RESEARCH IN THE FIELD

Many articles are devoted to the analysis of the signal coming from clothes, shoes, wearable equipment. The current research is close to studies [3] and [4], in which the authors discuss how to detect the gait phases and abnormalities of the phases. The source of the signal is ground contact force (GCF) sensors, which use air bladders and air pressure sensors. The authors use predefined rules to determine the appropriate gait phase. It should be noted that the integration of sensors in the shoes aligns the relief of the foot. At the same time, the use of predefined rules narrows the scope of the method.

It is worth mentioning the study [5], in which to determine the appropriate gait phase at each gait moment the authors use a threshold-based detection algorithm employing state transition theory. To obtain the signal, the smart shoe is used, combining two types of sensors: force sensitive resistors (FSRs) and a gyroscope. The data of smart shoes are used in another study of these authors [6], which recognizes gait phases. The disadvantage of using force sensitive resistors is that the FSR does not adequately reflect the actual foot pressure due to its small sensing area and limited sensing range [7]. In its turn, the study [7] uses the GCF sensors. For the analysis of the gait phases in the gait motions, the authors apply a hidden Markov model.

Summarizing the existing studies in the field of smart textile signal processing, it is possible to formulate several approaches that allow drawing conclusions about the parameters of running or walking:

- To “disassemble” the original signal into components and explore geometric shapes, its properties, etc.;
- To cluster the signal and then denote each cluster in one way or another depending on the parameters of the signals it is made up;
- On the basis of expert knowledge, to define (a) the norms of time intervals between the extremes of sensor signals, as well as (b) other indicators characterizing the steps. Then, it is necessary to divide the signal into steps, and for each step to calculate the values of the indicators;
- As in the previous approach, it is necessary to calculate the values of step indicators and then build a time series of changes in indicators in order to analyze the trend of change in indicators. The approach allows forecasting the development of various parameters of the race.

Each approach has its own advantages in determining what happens to the athlete in the running process. Solving this problem, the greatest effect can be achieved through synergistic approaches. The present research is based on the approach related to signal clustering and denoting each cluster.

III. THE EXPLOITED SYSTEM FOR PLANTAR PRESSURE ANALYSIS

In the present research, the wireless DAid® Pressure Sock System [1] (Fig. 1) has been used to measure temporal gait parameters and plantar pressure control. The proposed system consists of the array of sensors distributed over the sole part of socks, connected by conductive lines and custom designed connector with electronic devices that collect and transmit data from sensors to the data processing device [1].

Fig. 1. Allocation of left foot sensors of Daid® Pressure Sock System.

In Fig. 1 (right side), the labels ADC0, ADC1, ADC2, ADC3, ADC4 denote pressure sensors. White lines (see Fig. 1, left image) are conductive lines, which deliver the signal from sensor to sock connectors. The transmitting device connected to each sock uses Bluetooth to deliver the signal to the host computer.

The designed Daid® Pressure Sock System provides an opportunity to control relative pressure distribution, temporal gait features; it can potentially provide recognition of walk/run patterns and their long-time alterations that could help avoid possible injuries due to foot overload. The developed sensors are inexpensive, do not disturb plantar pressure distribution and may be easily customized following recommendation of physician [1]. In Fig. 2, the example of one sensor signal is depicted, where V^{-1} denotes inverted sensor output.

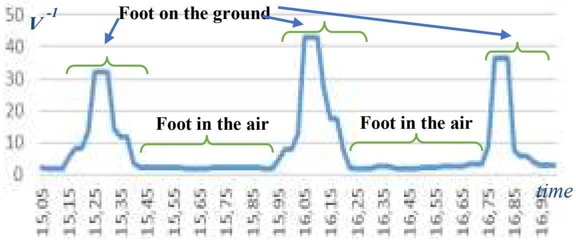

Fig. 2. A signal fragment produced by one sensor.

IV. SOM IN SIGNAL CLUSTERING TASK

SOMs are neural network based clustering methods used for the analysis and visualization of high dimensional data [8]. SOMs have an important characteristic for this study – they do not require a training sample. The only thing that is required is expert knowledge of the approximate number of classes that are expected to be received. SOMs can yield satisfactory results with a comparatively small training set and can be significantly faster than conventional MLPs for exploratory classification problems [8].

In the present research, the input vector transmits the signal of a step obtained from five sensors of smart sock (see Fig. 1). To find the start and end points of a step in the streaming signal of the race, a simple comparison of the sum signal across all sensors with a threshold value is used. If the sum is below a given threshold (see Fig. 3), the closest (right-hand) minimum of this sum will be the cut point of next step.

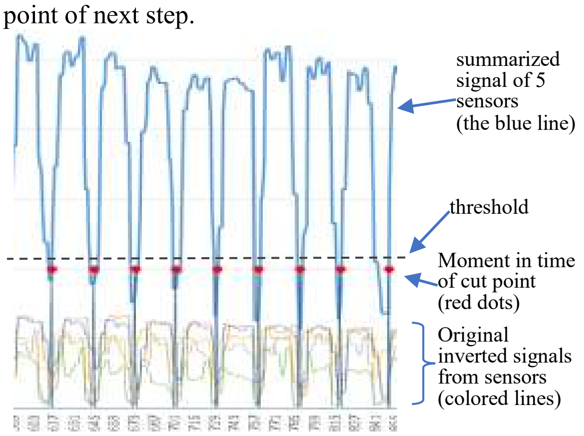

Fig. 3. Visual representation of weight vectors of the three neurons.

It is also possible to use the method of detection of the so-called reference points, based on the transformation of dynamic data into static images and the application of the multilayer perceptron (see [9] for detailed explanation). Any approach allows dividing the signal into steps. The five-channel signal of each step forms one input vector of the Kohonen network.

As the number of signal measurements of a person's step can vary, to determine the size of the input vector, the case with the largest number of measurements of the step (23 measurements) has been identified among all the steps of the two experimental races R_1 and R2. Two more formal measurements have been added to this number in order to ensure that all measurements are made. However, this is an optional activity. If the duration of any step is less than 25 measurements, for example, 18, then measurements from 19 to 25 are filled with zeros. In total, the input vector is formed by 125 measurements – 25 for each of the five sensors as shown in Fig. 4.

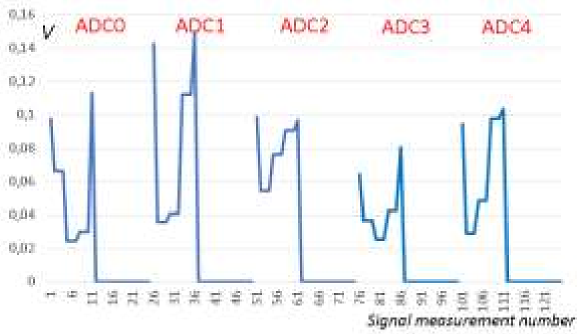

Fig. 4. Visual representation of the composite input vector of a step.

In the present paper, as a metric for determining the winning neuron of the input vector, the Euclidean metric

$$E(a,b)=\sum_{i=1}^{N}\left|a_i-b_i\right| \quad (1)$$

and shape-based criteria are used:

$$S(a,b)=\sum_{i=1}^{N-1}\left|(a_i-a_{i+1})-(b_i-b_{i+1})\right| \quad (2)$$

The Kohonen network consists of neurons organized into a rectangular structure. As a function of calculating the proximity of neighboring neurons, the Gaussian function is used:

$$\eta_i(d,k)=e^{\frac{d_i}{2\sigma(k)}}. \quad (3)$$

The study uses the coefficient of learning rate η equal to 0.1 and the standard deviation of the distribution σ equal to 0.5.

Depending on the goal of the experiment, the network consists of two, three or six neurons. In fact, their number indicates the number of clusters (types of steps), into which the Self-Organizing Map divides the "cut" signals of step. Accordingly, the more neurons are used in the Kohonen network, the more details are recognized in the signal of step.

To provide full control over all parameters of the Kohonen network, use arbitrary metrics for calculating the closeness of vectors and manually set the values of the weights, the authors have created their own SOM program in the Java programming language. In future, this development can be transferred to the Android platform, which will allow for real-time monitoring of the race parameters calculated on the side of this device.

V. IMPLEMENTATION OF EXPERIMENTS

The aim of the series of experiments is to investigate how successfully Self-Organizing Maps cluster the signal received from the DAid® Pressure Sock System sensors. For simplicity, the experiments use sensor signals from only one sock (from the left foot). In total, two experimental races were made (see Table 1). The first part of the path of each race was run normally by the test persor (in the opinion of the tester), and the second half – deliberately worse, performing the so-called unacceptable steps. Under the unacceptable ("the bad") step, the following sequence of foot-to-ground contact is implied:

1) contact of the sock with the ground at a large angle (the plane of the foot at this moment is practically perpendicular to the surface);
2) contact of the heel with the ground.

The number of steps accomplished by the test person during normal and bad races is given in Table 1.

Table 1
Experimental Race

Race identifier	Total step number	Good* step count	Bad* step count
R_1	105	71	34
R_2	64	36	28

** in the test person's opinion*

It should be noted that due to the human factor, within the normal style of running, various deviations from the normal style could be tolerated (the same can be attributed to the style of bad running – some "bad" steps could be performed as quite good steps or, conversely, as "not steps" at all). The comparison of the test person's opinion with the "opinion" of the artificial neural network is also one of the important points of the present research.

Apart from races, there were also two series of jumps on the left foot: 11 jumps for the training sample and 10 jumps for the test sample. The obtained jumps are used contrary to the steps and are referred to as the "unrecognized" signals.

A. The First Experiment: Clustering Steps into Two Clusters and Choosing the Best Metric

In the first experiment, the Self-Organizing Map should cluster the R_1 race steps (see Table 1) into two clusters of steps, using two metrics: the Euclidean metric (1) and contour-based metric (2). As a result of applying the Euclidean metric, one step out of the 71 steps of the normal running style is recognized as the "bad" one by the SOM. In turn, 4 steps out of the 34 steps taken in bad running style are attributed by the SOM to the cluster of "good" ones, which makes the clustering accuracy equal to 95 %. It cannot be stated that these five steps were incorrectly clustered. It is assumed that the test person unintentionally (and not realizing) took one bad and four good steps at the wrong time. In turn, the Self-Organizing Map has its own "opinion".

Fig. 5 demonstrates the weight vectors of the SOM neurons obtained during training (clustering).

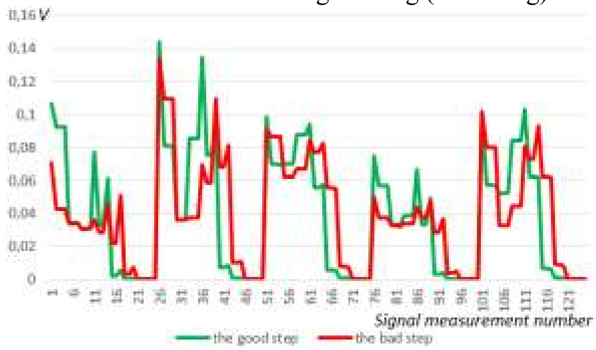

Fig. 5. Visual representation of weight vectors of two neurons.

Since the weights of the neurons contain approximately average values of all the elements of the cluster, the lines in Fig. 5 actually reflect the generalized image of a "good" step and that of a "bad" step.

In case of using the counter-based metric, the network recognizes the steps in a way similar to that of the Euclidean metric, but clusters the other seven steps in a different way. However, one cannot say that the network "made a mistake" as another metric was used, which otherwise evaluated the similarity of time series. In subsequent experiments, this trend remains unchanged – using the Euclidean metric, the clustering of steps practically coincides with the test person's opinion. Therefore, in subsequent experiments it was decided to use only the Euclidean metric. The model obtained in the first experiment is denoted as SOM_{R1}^{2}(two clusters, the first sample).

B. The Second Experiment: Clustering Steps Into Three Clusters: The "Good", the "Bad" and the "Unrecognized"

In the second experiment, the Kohonen network is formed by three neurons, ensuring that the signal is clustered into three classes of steps. To the selection of R_1 race signals, 11 signals of jumps were added from the toe (hereinafter, the "unrecognized" signals). As a result of training, one step (and the same as in the first experiment) out of the 71 steps of the normal running style was recognized by the Kohonen network as a bad step. Eight out of the 34 steps of the bad running style were attributed by the network to a cluster of good steps. Finally, six "unrecognized" signals out of 11 signals were recognized as bad steps. Fig. 6 shows the values of weights of all three neurons.

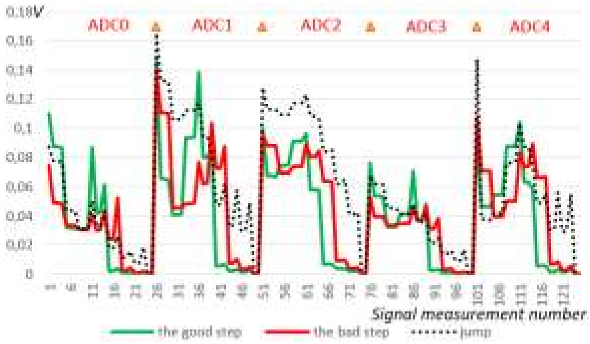

Fig. 6. The values of the weights of the neurons, the “good” step, the “bad” step and jump (unrecognized).

Fig. 6 demonstrates that the “unrecognized” signal (dotted line) significantly differs from the signal of good (green line) and bad steps, and it is also longer. This is clearly seen in the measurements 15–25 of ADC4 sensor (Fig. 6, signals 115–125): although the steps are already completed, there is still the “unrecognized” signal (dotted line). The model obtained in the second experiment is denoted as SOM_{R1}^{3}.

C. The Third Experiment: Classification of Test Steps on the Trained Network

In the third experiment, the steps from the R_2 race are classified on the models SOM_{R1}^{2} and SOM_{R1}^{3} obtained in the previous experiments (see Table 1). The model SOM_{R1}^{2} allowed classifying the signals of the race R_2 at 100 % accuracy. In turn, the model SOM_{R1}^{3} also classifies good and bad steps at 100 % accuracy, but the three “unrecognized” signals out of the 10 test “unrecognized” signals are attributed to bad steps. Fig. 7 compares signals of the three “unrecognized” signals with the values of weights of the neuron of bad steps and that of the “unrecognized” signal.

As seen in Fig. 7, duration of “unrecognized” signals (recognized by the SOM as bad steps) is 18 measurements, which is typical of bad steps.

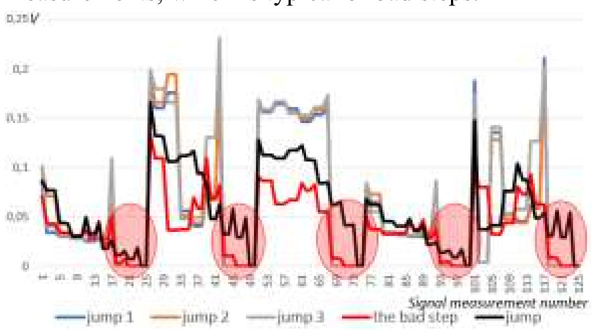

Fig. 7. Comparison of duration of the “unrecognized” signal to that of the signal of bad steps.

Indeed, repulsions from the surface when performing jumps on the left toe had similar features with landings on the toe when performing bad steps.

D. The Fourth Experiment: Clustering of the Signal of Steps by a Network with Six Neurons

Within the framework of the fourth experiment, all the steps of R1 and R2 races were used. The aim of the experiment was to investigate the clustering results of the signal of steps by a network with two neurons (the obtained model was denoted as $SOM_{R1,R2}^{2}$), and by a network with six neurons (the obtained model was denoted as $SOM_{R1,R2}^{6}$). In the case of $SOM_{R1,R2}^{2}$, the four steps out of the 107 steps made during the normal running style were recognized by the SOM as bad steps, and the eight steps out of the 62 steps made during the unacceptable running style were recognized by the SOM as acceptable steps.

It should be noted again that this does not indicate an error in the operation of the SOM – the test person could unconsciously take steps of the type not required of him.

Table 2
Distribution of Steps by Clusters

Type of a step	Number of steps (for two races)	$SOM_{R1,R2}^{2}$		$SOM_{R1,R2}^{4of6}$			
		Cluster “Good”	Cluster “Bad”	Cluster A	Cluster B	Cluster C	Cluster D
Good	107	104	3	2	18	43	44
Bad	62	8	54	54	8	0	0

It is worth noting that in the case of six neurons in the Kohonen network ($SOM_{R1,R2}^{6}$), two neurons turned out to be “dead” – they were not involved. The eight steps attributed by the network $SOM_{R1,R2}^{2}$ to good steps were classified into the cluster B. The remaining 54 steps of the unacceptable style were classified by the SOM into the cluster A. As a result, it can be concluded that this is the cluster of bad steps. Those 103 steps out of 107 steps (acceptable style), which were previously attributed to good steps, were classified into the clusters C and D, while the remaining four steps were grouped into the clusters A and B, each containing two steps. These data are provided in

Table 2.

On the basis of the weight vectors of the four neurons, the matrix of distances was created, which resulted in a dendrogram shown in Fig. 5.

Fig. 8. Dendrogram of the similarity of clusters.

As shown in the dendrogram, all the steps can be divided into three types: the “bad” cluster A, the “good” clusters B and D. The question arises: what steps were attributed by the network to the cluster C?

To answer this question, let us take a look at the matrix of cluster distances (see Fig. 9).

	a	b	c	d
a	0	2,73944	2,80785	2,92033
b		0	1,96216	0,66857
c			0	1,77845
d				0

Fig. 9. The matrix of cluster distances.

The cluster C is closer to the cluster D, which is a subset of good steps. Thus, clusters B, C and D are types of good steps with varying degrees of acceptability. In particular, the cluster C holds the steps from the sample of good steps, which were unintentionally not so well made. In fact, this group is an indication of the deviations in the test person’s race from normal running style.

E. The Fifth Experiment: Comparison of Manual Clustering with SOM Clustering

The final experiment performed manual clustering (the obtained model was denoted as $Manual^2_{R1}$) and clustering by the Self-Organizing Map (SOM^2_{R1}). As the test person believed that he ran well the first half of the race and the second half – worse, then the average values of “good” and “bad” steps were taken. In Fig. 10, Fig. 11 and Fig. 12, these values are compared to the values of weights of the network SOM^2_{R1} of the first experiment.

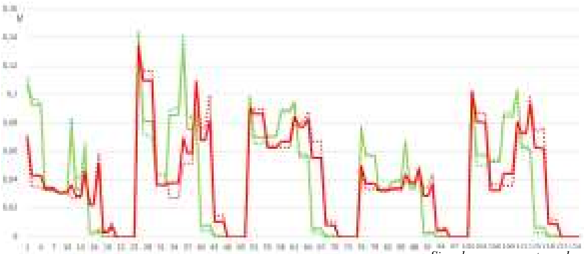

Fig. 10. The average values of steps obtained manually and by the network SOM^2_{R1}.

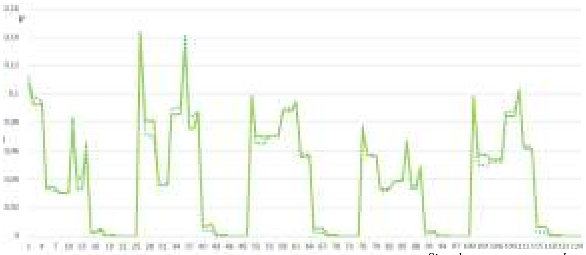

Fig. 11. The average values of “good” steps obtained manually and by the network SOM^2_{R1}.

Fig. 12. The average values of “bad” steps obtained manually and by the network SOM^2_{R1}.

As can be seen in the figures above, manual clustering practically coincides with the clustering of the Self-Organizing Map: the graph of the average value of acceptable steps coincides with the graph of the weights of the corresponding neuron by 98.5 %, and for unacceptable steps the coincidence is 97 %. This is an important result, showing that the “opinion” of the network basically coincides with that of the test person. There are also minor “disagreements” described in the previous experiments.

VI. CONCLUSIONS

The authors of the present paper have investigated the application of the Self-Organizing Map to the clustering of the signal of steps obtained from sensors of DAid® Pressure Sock System. The experiments have been carried out with the aim to use the trained SOM in order to classify steps that do not participate in training.

As a result of the first experiment, clustering of the steps into two clusters has been achieved at 95 % accuracy. The error of five percent is explained by subjectivity when the test person attributed steps to a particular class. To increase the accuracy, expert knowledge in the field of physiology is required. It has been found that for a given signal type the Euclidean metric gives a result that is closer to the opinion of the test person than the contour-based metric does.

In the second experiment, signals of the third type (actually not being steps) have been added, and clustering into three groups has been performed. The coincidence with the opinion of the test person has been 87 %.

In the third experiment, the models obtained during the training in the first and second experiments (SOM^2_{R1} and SOM^3_{R1}) have been used. The coincidence of the classification of steps (previously not involved in training) with the opinion of the test person has been 100 % for the two types of steps, and 96 % for the steps and “unrecognized” signals.

In the fourth experiment, the authors have compared the clustering of acceptable and unacceptable steps into two and six clusters. In case of six neurons in the Kohonen network, two have not been not involved. One cluster contains unacceptable steps, three clusters – acceptable. As a result, the gradation of good steps has been obtained. It should be noted that bad steps have not been divided into subgroups. It is also worth mentioning that even with such a small amount of data, it has been possible to demonstrate the clustering approach with further assigning of a name to each class in accordance with the steps included in it.

Finally, in the fifth experiment, the results of manual clustering and clustering by the SOM network have been compared. It has been found that the results of manual clustering and that of the SOM network practically coincide.

Within the framework of the present research, a practical study of the application of Self-Organizing Map to smart textile signal clustering has been carried out. Based on the results of the experiments, it can be concluded that the network is able to recognize the signal as the "good" step, "bad" step, and not a step at all. Such a primary (rough) classification allows further selecting specific algorithms for a detailed analysis of "good" steps and "bad" steps. The detailed analysis of bad steps is the key to solving the problem of revealing of an athlete's special type of fatigue, leading to injuries. This is also the area of the further research.

REFERENCES

[1] Oks A., Katashev A., Zadinans M., Rancans M., Litvak J. Development of Smart Sock System for Gate Analysis and Foot Pressure Control. XIV Mediterranean Conference on Medical and Biological Engineering and Computing. 2016, pp. 472-475.

[2] Hreljac A. Impact and overuse injuries in runners. Med Sci Sports Exerc. Vol. 36, No. 5, 2004, pp. 845-849.

[3] Kong K., Tomizuka M. Smooth and Continuous Human Gait Phase Detection Based on Foot Pressure Patterns. IEEE International Conference on Robotics and Automation, 2008. pp. 3678-3683.

[4] Kong K., Bae J., Tomizuka M. Detection of Abnormalities in a Human Gait Using Smart Shoes. SPIE Smart Structures/NDE, Health Monitoring. 2008.

[5] Srivises W., Nilkhamhang, I., Tungpimolrut, K. Design of a Smart Shoe for Reliable Gait Analysis Using State Transition Theory. In Proceedings of 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. 2012, pp. 1-4.

[6] Srivises W., Nilkhamhang, I., Tungpimolrut, K. Design of a Smart Shoe for Reliable Gait Analysis Using Fuzzy Logic. Proceedings of SICE Annual Conference (SICE) 2012. 2012, pp. 834-838.

[7] J.Bae, M.Tomizuka. Gait Phase Analysis Based on a Hidden Markov Model. Mechatronics. Vol. 21, no. 6, 2011, pp. 961-970.

[8] Kohonen T., Honkela T. Kohonen Network. 2007, Scholarpedia, 2(1):1568.

[9] Shootov A., Chizhov Y., Aleksejeva L., Oks A., Katashev A. Artificial Neural Network Based Approach for Control Points Detection in Smart Textile Signals. Procedia Computer Science. Vol. 104, 2017, pp. 548-555.