Yelena S. Tarantova[1], Kirill V. Makarov[2], Alexey A. Orlov[3]   DOI: 10.25045/jpit.v10.i2.08

Murom Institute of Vladimir State University, Murom, Russia

[1]elena.tarantova@yandex.ru, [2]k-distance@rambler.ru, [3]AlexeyAlexOrlov@gmail.com

FEATURES SELECTION FOR HUMAN ACTIVITY RECOGNITION IN THE TELEREHABILITATION SYSTEM

The article researches the problem of human activity recognition in the telerehabilitation system. To recognize the patient's physical activity, the sensors of the smartphone are used: accelerometer and gyroscope. The term of telerehabilitation and an exemplary set of exercises as part of a rehabilitation event are considered. Such tasks as determining the permissible value of frequency and classification accuracy and also features select to reduce computational complexity are solved. The classification frequency value is proposed to take into account the type of activity and the patient's health group for evaluating the correctness of independent implementation of rehabilitation events by the patient. An algorithm for selecting a subset of informative features is described. An experiment is carried out to select a subset of informative features is necessary for the classification of physical activity in the telerehabilitation system, taking into account the influence of features on the classification accuracy and computational complexity in their calculation. Comparison of classification results using a feature vector and using a subset of informative features is performed.

Keywords: *feature extraction, feature selection, feature engineering, accelerometer, gyroscope, telerehabilitation, human activity recognition.*

1. Introduction

Telemedicine is one of the areas of health care, which is aimed to improving the availability of medical care by information technology. Telerehabilitation is a popular area of telemedicine, which allows to improve a patient's level of functioning in his own life environment through the independent implementation of a rehabilitation program under remote control and guidance of a rehab physician.

Telerehabilitation implies constant monitoring data on the performed physical activity of the patient. A rehab physician makes up an individual rehabilitation program for each patient. This program includes various types of physical activity and shares into four stages of intensity and duration. The duration of the exercises is determined by the patient's level of functioning and should not exceed 30 minutes per one cycle [1].

Monitoring based on the data of various sensors located on the human body is one of the options for remote monitoring of the patient's condition. In particular, it is possible to use the classification module of types of physical activity based on the readings of such smartphone sensors as a three-axis accelerometer and three-axis gyroscope to monitor the physical activity using the telerehabilitation system. Research results show that the accelerometer and gyroscope provide high classification accuracy and can be used for human activity recognition [2, 3].

This classification module allows to restore the relationship between the sensor readings of the patient's smartphone and the types of physical activity performed by the patient during the rehabilitation program. Using the feature vector obtained by calculating various characteristics from the accelerometer and gyroscope readings allows determining different classes (types) of physical activity.

2. Problem Formulation

The rehab physician needs to estimate the correctness execution of physical activity by the patient during the monitoring. a lot of research has been done as a part of solving the problem of classifying a patient's physical activity types, however, the solution of such tasks as selecting of informative features subset, determination of the permissible accuracy value and frequency of classification during telerehabilitation is still of current interest [4–9].

Selection the most informative subset of features from the feature vector plays an important role in the classification, since the selected features affect the accuracy of the classification. The more features in a subset, the higher the classification accuracy. However, increasing the number of features results in increasing computational complexity.

The classification results are used to monitor the physical activity of the patient. It is necessary to ensure such accuracy of classification, which would allow the rehab physician, considering the classification error, to estimate the correctness of rehabilitation events executed by the patient. Therefore, it is necessary to determine the valid value of the classification accuracy during telerehabilitation.

Although the doctor makes up a rehabilitation program, the patient may not perform it accurately, and sometimes not at all. Therefore, it is also necessary to define the valid value of the classification frequency to determine the patient's condition during monitoring. The frequency of classification is also important for detecting patients' critical states which may occur when performing physical activity. Patients' Critical state is understood as a situation where the patient is unable to perform a rehabilitation event due to deterioration of his condition.

Thus, it is necessary to determine the valid value of classification accuracy and frequency of its implementation for monitoring patient's physical activity as part of telerehabilitation events. It is also necessary to select such a subset of features to obtain a given classification accuracy, which would allow obtaining the required classification accuracy with low computational complexity.

3. Methodology

3.1. Determinining the Permissible Value of the Classification Frequency

It is necessary to consider the rehabilitation program in detail in order to learn the types of physical activity used and the duration of its execution during the day to determine the valid value of classification frequency.

A rehabilitation program can include dynamic activities such as jogging, dosed walking, gymnastic exercises, swimming, rowing, skiing and static poses, which are used for recreation and the transition between dynamic activities. Daily physical activity may also differ in patients depending on their gender, age and profession. In addition, patients with different clinical conditions may belong to different health groups.

Patients of the first health group are practically healthy people with a low risk factor occurrence of critical state in performing physical activity; they do not need constant monitoring. Patients of the second health group are people with an average risk factor occurrence of critical state in performing physical activity. Patients of the third health group are people with a high risk factor occurrence of critical state in performing physical activity; they require constant monitoring of their condition.

Thus, the rehabilitation program is made in accordance to individual features of each patient.

However, the types of physical activity and its duration for persons within a certain health group are similar. Persons aged 18 to 64 must perform physical activity for at least 150 minutes a week.

The planning of daily activities, including various types of physical activity of a person, should take place in the form of discussion with the rehab physician and should be based on the following principles [10]:

- frequency – at least 5 times a week;
- duration – 30 minutes, with an increase of up to 45-60 minutes – 2-3 times a week.

The most accessible and usual types of activity for patients with various health groups are walking, running and climbing stairs. These types of activity allow you to restore the normal mental and physiological functions of the body lost as a result of illness or injury. Walking is the most common form of dynamic physical activity, which allows improving the physical condition of the patient, provided sufficient duration of each session.

The structure of the rehabilitation event, regardless of the patient's health group, should include such components as:

1. Warm up (preparatory part) – 5-10 minutes.

2. The main part (active period) – 20-30 minutes.

3. Cool down (final part) – 5-10 minutes.

Thus, an approximate set of exercises as part of the rehabilitation event is presented in table 1.

Table 1

Approximate set of rehabilitation exercises

Stage	Type of physical activity	Duration, min
Warm up	Walking	3
Warm up	Run	1
The main part	Up and down the stairs	5
The main part	Squats	5
The main part	Biking	10
Cool down	Walking	4
Cool down	Breathing exercises in standing positions	2
The total time of the rehabilitation event, min	30-35	

The main task of the rehab physician is to monitor the patient's health. It is necessary to monitor the correctness of performance of rehabilitation events by the patient and to identify changes in his condition. To solve this problem during telerehabilitation it is necessary to classify various types of physical activity with a certain frequency.

In order to determine the classification frequency is proposed to take into account:

- type of physical activity: static or dynamic;
- patient's health group: first, second or third.

When performing dynamic types of activity, the probability of incorrect exercise is higher than when performing static types of activities, as the probability of changing from one type of activity to another increase. Accordingly, the classification of dynamic physical activity should be carried out with higher frequency than when considering static poses.

Also, with an increase in the patient's health group, more attention is required from the rehab physician, since there is a possibility of emergency situations in the performance of physical activity. Consequently, the classification of physical activity with an increase in the number of the patient's health group is required to be carried out with higher frequency.

The proposed classification frequency for the case of performing dynamic and static activity by patients of different health groups is presented in table 2.

Table 2

The classification frequency of different types of activity
performed by patient with different health groups

Health group	Classification frequency	
	Dynamic activity	Static activity
1	0.1	0.03
2	0.2	0.05
3	0.5	0.1

In order to determine the number of classifications for different periods of time following formula (1) will be used:

$$n = v \cdot t \cdot 60 \tag{1}$$

where v is the classification frequency, t is the duration of the type of physical activity, min, n is the number of classifications.

Thus, the classification frequency for different groups of health for during performing an exemplary set of exercises is presented in table 3.

Thus, the proposed value of the frequency of the classification will allow to monitor the correctness of the patient's implementation of rehabilitation events.

Table 3

The number of classifications for different duration of exercise

Stage	Type of physical activity	Duration, min	Number of classifications		
			The first group of health	The second group of health	The third group of health
Warm up	Walking	3	18	36	90
Warm up	Run	1	6	12	30
The main part	Up and down stairs	5	30	60	150
The main part	Squats	5	30	60	150
The main part	Biking	10	60	120	300
Cool down	Walking	4	24	48	120
Cool down	Breathing exercises in the standing positions	2	4	6	12

3.2. Development of an Algorithm for Selecting a Subset of Features

Feature subset selection is an important task when classifying types of physical activity, since the selected informative features affect the resulting classification accuracy and computational complexity. And often it is necessary to make a choice either in favor of improving the classification accuracy, or in favor of reducing computational complexity. To solve this problem, new features will be engineering.

Feature engineering is the process of creating new features based on existing ones to increase classification accuracy by applying knowledge about the input data or the application area. Feature engineering will allow replacing the original informative features with new ones, maintaining the obtained classification accuracy and reducing the computational complexity.

An experiment will be carried out to select the feature set. For the experiment, the MATLAB environment will be used. Open dataset will be used as data for the experiment and objects for classification [11].The data collection was attended by people of different height and weight. They performed those types of physical activity that are used for rehabilitation purposes: dynamic movements such as walking, running, biking, walking up and down stairs, and static postures such as sitting and standing.

To carry out the experiment it is necessary:

1. To form a vector of features and to choose classification method.

2. To check the sufficiency of the training dataset size to building classification model.

3. To calculate features and to estimate the computational complexity.

4. To train and to estimate training results.

5. To evaluate the informative of the features using the feature selection algorithm.

6. To engineer new features in the case of highly informative features and high computational complexity.

7. To form a subset of features.

8. To train and to estimate the results.

9. To compare the accuracy and computational complexity when using the entire feature vector and the selected subset of informative features.

4. Experiment

4.1. Forming of the Feature Vector and Choosing of Classification Method

A literature review was carried out to form the feature vector and choosea classification method.

In many activity recognition systems using accelerometer and gyroscope data statistical values such as elements of feature vector are used. The statistical values include the average value, variance, standard deviation, asymmetry, etc. The features that are used in studies on the

recognition of human physical activity are presented in table 4.

Used features

References	Features
[4]	Range, Mean, Kurtosis, Skewness
[5]	Mean, Standard Deviation, Max, Min, Median Absolute Deviation, Frequency Signal Weighted Average
[6]	Mean, Median, Standard deviation, Peak of absolute data, Absolute integral, Correlation between axes
[7]	Mean, Standard deviation, Median absolute deviation, Max, Min, Skewness, Kurtosis, Correlation
[8]	Mean, Median, Standard deviation, Skewness
[9]	Mean, Standard deviation, Skewness, Median

Also, many studies on the recognition of physical activity compare classification methods by classification accuracy. Studies that use different classifiers and the classification accuracy are presented in table 5.

Table 5

Accuracy of classifiers used in research

References	Classification method	Accuracy
[7], [12]	Support Vector Machine (SVM)	96%
[13], [14]	Naive Bayes	90%
[15], [16]	K-Nearest Neighbors (KNN)	94%
[17], [18]	Decision Tree	92.9%

Thus, SVM will be used as a classification method in this experiment. The generated feature vector for the experiment, obtained from each axis of the accelerometer and gyroscope, is presented in table 6.

Table 6

Formed vector of features

№	Description
1-6	Mean value
7-12	Standard deviation
13-18	Skewness
19-24	Minimum element of set
25-30	Maximum element of set
31-36	Mean or median absolute deviation
37-42	Kurtosis
43-48	Interquartile range
49-54	Range of values

4.2. Checking the Sufficiency of the Training Set Size

The relationship between the features and the classified type of activity is derived from the training set. The larger the training set size, the higher the model's ability to recognize the type of physical activity correctly. Therefore, it is important to ensure that the size of the selected data set is sufficient for training.

In order to check the sufficiency of the training set size, it is necessary to construct a dependence graph of the classification accuracy versus the size of the training set.

The training set containing 12600 samples is used in experiment. From the training set, it is necessary to choose subset of different sizes, to make a model and to determine the average

classification accuracy for each subset.

After the values of classification accuracy were obtained using a training set of different sizes, it is possible to construct a dependence graph of the accuracy versus the size of the training set (see figure 1).
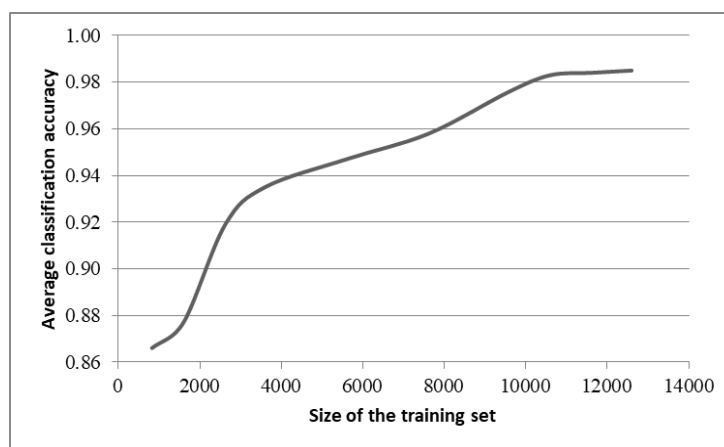

Figure 1. The dependence graph of the accuracy versus the size of the training set

As can be seen from the graph, for small training set sizes, the accuracy curve continues to increase, and after 10,000 it is aligned. This means that the classification accuracy will not significantly change with a further increasing of training set size. Accordingly, the existing size of the training set is sufficient for further experimentation.

4.3. Calculation of Features and Evaluation of the Computational Complexity of Its Calculation

To calculate the generated feature vector, it is possible to use the MATLAB functions or to calculate the features using the formulas that are presented in table 7.

Table 7

Functions and formulas for calculating features

№	Description	Function	Formula
1-6	Mean value	mean(x)	$\bar{x} = \dfrac{1}{N}\sum_{i=1}^{N} x_i$
7-12	Standard deviation	std(x)	$\sigma = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$
13-18	Skewness	skewness(x)	$E\left[\left(\dfrac{x-\bar{x}}{\sigma}\right)^3\right]$
19-24	Minimum element of set	min(x)	$\min_i(x_i)$
25-30	Maximum element of set	max(x)	$\max_i(x_i)$
31-36	Mean or median absolute deviation	mad(x)	$median_i\left(\left|x_i - median_j(x_j)\right|\right)$
37-42	Kurtosis	kurtosis(x)	$\dfrac{E\left[(x-\bar{x})^4\right]}{E\left[(x-\bar{x})^2\right]^2}$
43-48	Interquartile range	iqr(x)	$Q3(x) - Q1(x)$
49-54	Range of values	range(x)	$\max(x) - \min(x)$

The time of calculation of each feature was also determined. Results presented in table 8.

Table 8

Feature calculation time

Features	Time, s
min	0.007
max	0.007
mean	0.008
range	0.017
mad	0.064
std	0.082
iqr	0.309
kurtosis	0.566
skewness	0.569

The table shows that the calculation of such features as the interquartile range, skewness and kurtosis takes more time than the other features.

Next it is needed to train using formed vector of features and the SVM classifier. After the training, a classification accuracy of 98.5% was obtained. The confusion matrix for each type of activity with taken combination of features is presented in figure 2.

Figure 2. The confusion matrix

4.4. Evaluation the Informative of the Features

In order to reduce the computational complexity when using the generated feature vector, it is necessary to select a subset of informative features. To select the features, the ReliefF algorithm will be used. ReliefF is a widely used filtering method that ranks objects by weighing based on quality.

The algorithm finds the nearest hit (a data point from the same class) and the nearest misses (data points from different classes) for each object. Then differences of values of the corresponding features are calculated by (2).

$$\omega_i = \sum_{j=1}^{N} \left(x_i^j - near\_miss(x^j)_i \right)^2 - \left(x_i^j - near\_hit(x^j)_i \right)^2 \qquad (2)$$

where ω_i is the weight of the ith feature;
x_i^j is the value of the ith feature for point x^j;
$near\_miss(x^j)$ *is* the nearest data point to x^j in the different classes;

near_hit(x^j) is the nearest data point to x^j in the same classes.

The weights of the features are determined by the difference in values and the significance of the features is associated with its weights [19]. The weights of features are presented in table 9.

Table 9

The weights of features

Features	Accelerometer			Gyroscope		
	x	y	Z	x	y	z
mean	0.0249	0.0428	0.0372	0.0636	0.0822	0.0501
std	0.1091	0.1628	0.1045	0.1562	0.1277	0.1013
skewness	0.0291	0.0359	0.0257	0.0347	0.0263	0.0363
min	0.0961	0.1516	0.1073	0.1185	0.1279	0.0553
max	0.0994	0.1109	0.1014	0.1139	0.0989	0.0583
mad	0.0992	0.1520	0.0973	0.1592	0.1235	0.1086
kurtosis	0.0136	0.0211	0.0201	0.0148	0.0123	0.0163
iqr	0.0770	0.1236	0.0866	0.1522	0.1114	0.0952
range	0.1126	0.1428	0.1229	0.1270	0.1160	0.0728

As can be seen from the table, such features as mean, skewness and kurtosis were not informative, so they will not be included in the subset of features. A feature like iqr is informative, but it takes a lot of time to calculate it, so it needs to be replaced with a new one, created from the original feature.

4.5. Feature Engineering

Interquartile range is the distance between the first and third quartiles (the difference between 75th and 25th percentiles). This feature is proposed to be replaced by the second percentile (the value of X, to which 2% of the distribution is located). Denote the second percentile as iqr*.

On the calculation iqr on average spent 0,309 s, and in the calculation of the iqr* – 0,180. The computational complexity has been reduced. Now it is necessary to make sure that the weight of the iqr* is not less than the iqr.

Weights of iqr and iqr* for each axis of the accelerometer and gyroscope are presented in table 10.

Table 10

Weights of iqr and iqr*

Features	Accelerometer			Gyroscope		
	x	y	z	X	y	z
iqr	0.0770	0.1236	0.0866	0.1522	0.1114	0.0952
iqr*	0.0885	0.1555	0.1003	0.1215	0.1211	0.0619

Thus, the selected features for the subset of features are presented in Table 11 and are indicated by the "+"sign.

Now it is need to train using formed subset of features. After the training, the classification accuracy was 97.6%. The classification confusion matrix for each type of activity using a subset of features is presented in figure 3.

Table 11

The selected features for the subset of features

Features	Accelerometer			Gyroscope		
	x	y	z	X	y	z
mean	–	–	–	–	–	–
std	+	+	+	+	+	+

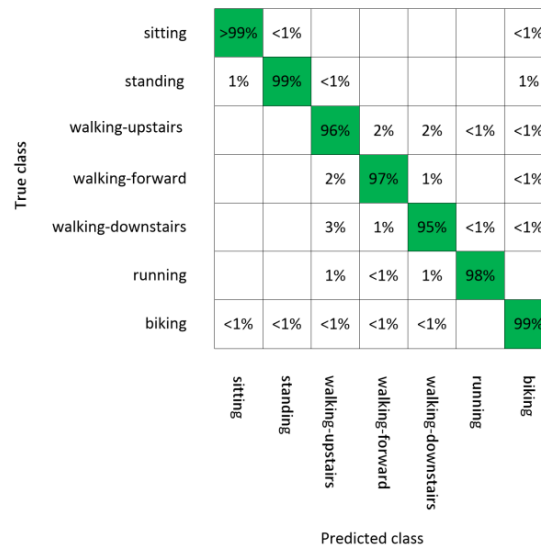| skewness	–	–	–	–	–	–
min	–	+	+	+	+	–
max	–	+	+	+	–	–
mad	–	+	–	+	+	+
kurtosis	–	–	–	–	–	–
iqr*	–	+	+	+	+	–
range	+	+	+	+	+	–

Figure 3. The confusion matrix

4.6. Determination of the Permissible Value of the Classification Accuracy

During the experiment, the accuracy of the classification of physical activity types was obtained, using the selected subset of features, equal to 97.6%. It is necessary to determine if a given value of classification accuracy is permissible during telerehabilitation.

The permissible value of classification accuracy is proposed to be determined by the number of classification errors per minute. The number of errors in the classification of physical activity for a given period of time is determined by a rehab physician.

An experiment to determine the number of errors per minute during classification of physical activity using a training data set was carried out. Experimental results show that one classification error occurs on average per minute with a classification accuracy of 97.6%.

Thus, the obtained value of the classification accuracy in the course of the experiment is permissible during telerehabilitation.

5. Conclusion

Thus, to determine the frequency and accuracy of classification, the types of physical activity and their duration in rehabilitation events were reviewed. During the experiment, the classification frequency values were proposed, which depend on the patient's health group and the type of physical activity. To reduce the computational complexity of the classification, an algorithm was developed and an experiment was carried out to select a subset of informative features. During the experiment, the calculation time and the weight of each feature were determined. From such a feature as the interquartile range, a new one was engineered, because the original feature was informative, however its calculation was time consuming.

Training was also performed and the accuracy of classification physical activity was obtained using the feature vector and a subset of informative features. The difference in

classification accuracy when using the feature vector and the feature subset was less than 1%, and the computational complexity was reduced almost by three times.

Thus, the value of classification frequency and a subset of informative features for the classification of physical activity were proposed, which allowed for the required classification accuracy and reduced computational complexity.

References

1. Vladzimirskij A.V., Lebedev G.S. Telemedicina, M.: GJeOTAR, Media, 2018, 576 s.
2. Lara O.D., Labrador M.A. A Survey on Human Activity Recognition using Wearable Sensors // IEEE Commun. Surv. Tutorials, 2013, vol.15, no.3, pp.1192–1209.
3. Walse K., Dharaskar R.V. A Survey on Human Activity Recognition using Smartphone // Int. J. Adv. Res. Comput. Sci. Manag. Stud, 2017, vol.5, no.3, pp.118–125.
4. Capela N.A., Lemaire E.D., Baddour N. Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients // PLoS One. Public Library of Science, 2015, vol.10, no.4.
5. Hasan S.S. et al. Human Activity Recognition using Smartphone Sensors with Context Filtering / ACHI 2016 Ninth Int. Conf. Adv. Comput. Interact. Hum., 2016, vol.571–572, pp. 1019–1029.
6. Attila Reiss (2014). Personalized mobile physical activity monitoring for everyday life. (Doctoral Thesis, Technical University of Kaiserslautern). https://kluedo.ub.uni-kl.de/files/ 3681/ _PhDThesis_AttilaReiss.pdf
7. Jorge Luis Reyes Ortiz (2015).Smartphone-based human activity recognition (Doctoral Thesis, Universitat Politècnica de Catalunya). https://link.springer.com/book/10. 1007%2F978-3-319-14274-6.
8. He Y., Li Y. Physical activity recognition utilizing the built-in Kinematic sensors of a smartphone // Int. J. Distrib. Sens. Networks. SAGE PublicationsSage UK: London, England, 2013, vol.2013, no.4, pp.481–580.
9. Miao F. et al. Identifying typical physical activity on smartphone with varying positions and orientations // Biomed. Eng. Online. BioMed Central, 2015, vol.14, no.1, pp.32–46.
10. Bubnova M.G., Aronov D.M., Bojcov S.A. Obespechenie fizicheskoj aktivnosti u grazhdan, imejushhih ogranichenija v sostojanii zdorov'ja: Metodicheskie rekomendacii. Moskva: Federal'noe gosudarstvennoe bjudzhetnoe Uchrezhdenie «Gosudarstvennyj nauchno-issledovatel'skij centr profilakticheskoj mediciny», 2015, 95 s.
11. Shoaib M. et al. Fusion of smartphone motion sensors for physical activity recognition // Sensors (Switzerland), vol.14, pp.10146–10176, 2014.
12. Morillo L. et al. Low Energy Physical Activity Recognition System on Smartphones // Sensors, 2015, vol.15, no.3, pp.5163–5196.
13. Suryanarayana D. A Comparative Study of Random Forest & K – Nearest Neighbors on HAR dataset Using Caret // Int. J. Innov. Res. Technol, 2017, vol.3, no.9, pp.6–9.
14. Lara Ó.D. et al. Centinela: A human activity recognition system based on acceleration and vital sign data // Pervasive and Mobile Computing. Elsevier, 2012, vol.8, no.5. pp.717–729.
15. Bugdol M.D. et al. Human Activity Recognition Using Smartphone Sensors // International Research Journal of Engineering and Technology, 2016, pp.41–47.
16. Marinho L.B., de Souza Junior A.H., Rebouças Filho P.P. A New Approach to Human Activity Recognition Using Machine Learning Techniques / International conference on Intelligent Systems Design and Applications, Springer, Cham, 2017, pp.529–538.
17. Fan L., Wang Z., Wang H. Human Activity Recognition Model Based on Decision Tree / 2013 International Conference on Advanced Cloud and Big Data. IEEE, 2013, pp.64–68.

18. Zheng Y. Human Activity Recognition Based on the Hierarchical Feature Selection and Classification Framework // J. Electr. Comput. Eng. Hindawi Publishing Corp., 2015, vol. 2015, pp.1–9.
19. Kononenko I. ReliefF for estimation and discretization of attributes in classification, regression, and ILP problems // Artif. Intell: methodology, systems, applications, 1996, pp.1–15.

UOT 004.93'11

Tarantova Yelena S. [1], Makarov Kiril V. [2], Orlov Aleksey A. [3]

Vladimir Dövlət Universitetinin Murom İnstitutu, Murom, Rusiya

[1]elena.tarantova@yandex.ru, [2]k-distance@rambler.ru, [3]AlexeyAlexOrlov@gmail.com

Telereabilitasiya sistemi çərçivəsində insanın fiziki fəallığının tanınması üçün əlamətlərin seçilməsi

Məqalədə telereabilitasiya çərçivəsində insanın fiziki fəallığının tanınması probleminə baxılır. Pasiyentin fiziki fəallığının tanınması üçün smartfonun tənzimləyiciləri istifadə olunur: akselerometr və giroskop. Telereabilitasiya anlayışı və reabilitasiya hadisəsi çərçivəsində tapşırıqların nümunəvi dəstinə baxılır. Klassifikasiyanın tezliyi və dəqiqliyinin mümkün qiymətlərinin müəyyən olunması, həmçinin hesablama xərclərinin ixtisarı üçün əlamətlərin seçimi kimi məsələlər həll edilir. Fəallığın növünü və pasiyentin sağlamlıq qruplarını nəzərə alaraq pasiyent tərəfindən tapşırıqların icrasının düzgünlüyünün qiymətləndirilməsi üçün klassifikasiya tezliyinin qiyməti təklif edilir. İnformativ əlamətlərin alt çoxluğunun seçimi alqoritmi təsvir edilir. Əlamətlərin hesablanması zamanı onun klassifikasiya dəqiqliyinə və hesablama çətinliyinə təsirlərini nəzərə alaraq, telereabilitasiya sistemində fiziki fəallığın klassifikasiyası üçün lazım olan informativ əlamətlər alt çoxluqlarının seçilməsi üçün sınaqlar keçirilir. Əlamətlər vektorundan və informativ əlamətlərin alt çoxluğundan istifadə ilə klassifikasiya nəticələrinin müqayisəsi aparılır.

Açar sözlər: əlamətlərin çıxarılması, əlamətlərin seçimi, əlamətlərin layihələndirilməsi, akselerometr, giroskop, telereabilitasiya, insanın fiziki fəallığının tanınması.

УДК 004.93'11

Тарантова Елена С.[1], Макаров Кирил В.[2], Орлов Алексей А.[3]

Муромский институт Владимирского государственного университета, Муром, Россия

[1]elena.tarantova@yandex.ru, [2]k-distance@rambler.ru, [3]AlexeyAlexOrlov@gmail.com

Выбор признаков для распознавания физической активности человека в рамках телереабилитационной системы

В статье рассматривается проблема распознавания физической активности человека в рамках телереабилитации. Для распознавания физической активности пациента используются датчики смартфона: акселерометр и гироскоп. Рассматриваются понятие телереабилитации и примерный набор упражнений в рамках реабилитационного мероприятия. Решаются такие задачи, как определение допустимых значений частоты и точности классификации, а также выбор признаков для сокращения вычислительных затрат. Предлагается значение частоты классификации для оценивания корректности выполнения упражнений пациентом с учетом вида активности и группы здоровья пациента. Описан алгоритм выбора подмножества информативных признаков. Проводится эксперимент по выбору подмножества информативных признаков, необходимых для классификации физической активности в телереабилитационной системе, с учетом влияния признаков на точность классификации и вычислительную сложность при их расчете. Производится сравнение результатов классификации с использованием вектора признаков и с использованием подмножества информативных признаков.

Ключевые слова: извлечение признаков, выбор признаков, проектирование признаков, акселерометр, гироскоп, телереабилитация, распознавание физической активности человека.