Multimed Tools Appl
DOI 10.1007/s11042-015-2643-0

Human activity recognition using quasiperiodic time series collected from a single tri-axial accelerometer

Andrey D. Ignatov[1] · Vadim V. Strijov[1]

Received: 16 September 2014 / Revised: 9 March 2015 / Accepted: 20 April 2015

Abstract The current generation of portable mobile devices incorporates various types of sensors that open up new areas for the analysis of human behavior. In this paper, we propose a method for human physical activity recognition using time series, collected from a single tri-axial accelerometer of a smartphone. Primarily, the method solves a problem of online time series segmentation, assuming that each meaningful segment corresponds to one fundamental period of motion. To extract the fundamental period we construct the phase trajectory matrix, applying the technique of principal component analysis. The obtained segments refer to various types of human physical activity. To recognize these activities we use the k-nearest neighbor algorithm and neural network as an alternative. We verify the accuracy of the proposed algorithms by testing them on the WISDM dataset of labeled accelerometer time series from thirteen users. The results show that our method achieves high precision, ensuring nearly 96 % recognition accuracy when using the bunch of segmentation and k-nearest neighbor algorithms.

Keywords Machine learning · k-nearest neighbor method · Neural network · Segmentation · Physical activity recognition · Singular value decomposition

✉ Andrey D. Ignatov
andrey.ignatoff@gmail.com

Vadim V. Strijov
strijov@ccas.ru

[1] Moscow Institute of Physics and Technology, Inststitutskii per. 9, 141700, Dolgoprudny, Moscow Region, Russia

1 Introduction

The current generation of portable mobile devices, such as cellular phones or music players, is becoming increasingly complex. Most of these devices incorporate various types of sensors, including accelerometers, light sensors, cameras, microphones and GPS sensors, that can be applied for analysis of everyday human behavior. One its important part is human physical activity, that reflects various aspects of health and thus is exceedingly attractive for many applications in the field of healthcare monitoring. In this paper, our purpose is to perform human physical activity recognition using data, collected from the built-in tri-axial accelerometer of a mobile phone. This data represents quasiperiodic time series corresponding to one of performed activities: walking, jogging, stair climbing, sitting or standing. For each time series we have to detect the correspondent activity type. Related problems of time series classification arise in human activity recognition on the silhouette [8], face recognition [35], gestures recognition [39], activity intensity level recognition [26] and detection of activity periods [28].

In the research of human activity recognition we generally face two related challenges. The first challenge is an online segmentation of time series into meaningful segments. The basic approach for this problem is to split each time series into equal segments [11, 15, 20]. However, in activity recognition the quality of segmentation directly influences the recognition results, and this elemental solution is actually not enough for accurate activity estimation. Thus, there should be used more sophisticated methods, based on deep analysis of the time series' structure. One natural way is to search for fundamental period of motion in these series and to consider it as a basis of partition. To extract this period we propose a technique, based on principal component analysis [19].

The second challenge is to construct an efficient classification method, that assign each obtained segment to one of the pre-determined classes. There exists a variety of different approaches, such as neural network [23, 25], logistic regression [15, 17, 18], SVM [6, 30], decision trees [23, 38], fuzzy algorithms [7] and differen heuristic algorithms [24, 28] based on spectrum analysis.

The main contributions of this paper are twofold. First, we propose an online segmentation algorithm for quasiperiodic time series, that is based on the extraction of the fundamental period. To detect this period we apply the technique of principal component analysis. Second, we propose the classification algorithm, based on the k-nearest neighbor method, which makes it possible real-time activity recognition on the smartphones. We also consider neural network as an alternative solution and compare the results, obtained by these two algorithms.

The paper is arranged as follows. Section 2 introduces some related work on human activity recognition and time series segmentation. Section 3 introduces an overview of the whole process of human activity recognition. Section 4 presents the detailed description of the algorithms of online time series segmentation, noise reduction and time series classification. Section 5 describes the experiments and provides the classification results, Section 6 summarizes our conclusions.

2 Related work

The task of human activity recognition using accelerometer has been well addressed in literature. However, the problem of time series segmentation in context of this task was not covered adequately in earlier papers. Primarily, this problem was considered as independent and was solved without verifying its results in real applications. Another lack of proposed solutions in terms of activity recognition consists in leaving out of account some particular aspects of the data, such as quasiperiodicity and homogeneity.

The solutions for the segmentation problem can be divided into two groups. First group is based on dynamic programming [4, 16]. While these algorithms show reasonable segmentation accuracy, they are not fast enough and thus are insufficient for processing of long series. The second group represents various greedy algorithms [13, 16]. These algorithms use the specific physical properties of the measured signal and in particular cases provide solutions that are very close to the optimal ones. For instance, methods designed for walking steps detection usually use peaks in acceleration that are associated with a heel strike or other phase of walking gate [5, 33, 37]. These algorithms are derived from the threshold methods, in which the step is count every time when the signal exceeds a predefined or adaptive threshold. Among other conventional segmentation techniques are sliding window-based algorithms [14] and algorithms, based on the signal transformation into a frequency domain [3, 27]. All these algorithms rely on the specific form of the signal, and while we consider a number of different activities their applicability is limited in this case. In this paper we propose a more general approach for quasiperiodic time series segmentation. To solve the problem, we use a technique based on the analysis of the phase trajectory matrix. The method we propose is also applicable for partitioning into periods any periodical or quasiperiodical time series.

Time series classification, the second step in the human activity recognition process, has a variety of possible approaches. The commonly encountered solutions for this problem are based on the SVM [6, 30, 34], neural networks [12, 23, 25] and decision trees [2, 23, 29], while relatively little research effort has been given to the k-nearest neighbor method, particularly because the problem of segmenting and aligning the data in this case becomes especially important [21].

Among relevant research are [1, 29] that describe the use of k-nearest neighbor method to recognize human daily activities. The authors propose it as an alternative solution for the classification problem, and study [1] show promising results for this method. However, these papers, like many other earliest works, are focused on the use of multiple accelerometers. Thus, works [1, 2, 9, 23, 29, 32] deal with activity recognition using data from five accelerometers. Though proposed systems are capable of identifying a wide range of activities, they are not practical while demanding the user wear a variety of sensors.

During the last several years there was a widespread of smartphones and other commercial mobile devices. They perfectly suit for the activity recognition process, while have a variety of sensors, enough processing power and do not require any additional equipment for data collection and recognition. For this reason we have chosen an open WISDM Actitracker Dataset [36] for training and performance evaluation of our algorithm. This dataset contains accelerometer data from Android cell phones. Data was collected from thirteen different users as they performed activities such as walking, jogging, stair climbing, sitting and standing, and consists of over a million time series points measured by an accelerometer.

Another advantage of this dataset is that it was already used in several research works [11, 15, 20]. In work [15] authors divided the source data into 10-second segments and generated totally 43 features out of six basic ones for each segment. For classification of the obtained feature vectors three techniques were used: decision trees (J48), logistic regression and multilayer neural network. In [11] authors also used 10-second segments, but there were developed additional features using PCA analysis and fast-Fourier transform. For activity recognition was used SVM classifier. The detailed classification results obtained in works [11, 15] are presented in the Table 1.

In the works above basic segmentation technique was used and classification was performed using the extracted features. In this paper, we propose an advanced method for online time series segmentation, based on the analysis of the phase trajectory matrix. We verify its results in this classification problem, and use the k-nearest neighbor method to classify the obtained raw segments.

3 Problem statement

Before developing the algorithmic approach, we introduce in this section some useful notations and provide additional information about the architecture of the proposed model.

In order to collect data for the considered supervise learning problem one has to carry a smartphone while performing a specific set of activities. Let $S = \{\mathbf{x}_t\}_{t=1}^{M}$, $\mathbf{x}_t \in \mathbb{R}^3$ be the time series, obtained from the accelerometer of the phone. These data represent measurements of acceleration in three directions, and for each element $\mathbf{x}_t$, $t \in \{1, 2, \dots, M\}$ there given type of physical activity $y_t \in Y$. Define a segmentation of time series S as a sequence $\{S_i\}_{i=1}^{n}$ of n segments such that $S_1 S_2 \dots S_n = S$ and each S_i is non-empty.

The procedure of the human activity recognition consists of the three following steps:

1) segmentation of the initial time series S,
2) removal noise from each obtained segment S_i,
3) implementation of the k-nearest neighbor algorithm f_{knn} to the treated segments $\{\tilde{S}_i\}_{i=1}^{n}$.

The scheme of the whole process is presented in Fig. 1.

To evaluate the overall performance of the proposed model we use the error function G, which is ratio between the number of incorrect estimations and the total number of observations. Denote f_{knn} be our classification algorithm. Then, the accuracy of classification is:

$$G(\mathrm{X}, \mathrm{Y}, f_{\text{knn}}) = \frac{1}{N} \sum_{i=1}^{N} [y_i \neq f_{\text{knn}}(X_i)], \tag{1}$$

Table 1 Accuracy of activity recognition for methods proposed in [11, 15]

Activity type, Y	J48	Logistic regression	Multilayer perceptron	SVM
Size of the dataset	4539	4539	4539	unknown
Testing technique	10-fold CV	10-fold CV	10-fold CV	50 % for train, 50 % for test
Jogging	96.5	98.0	98.3	95.7
Walking	89.9	93.6	91.7	92.7
Upstairs	59.3	27.5	61.5	79.5
Downstairs	55.5	12.3	44.3	79.8
Sitting	95.7	92.2	95.0	87.1
Standing	93.3	87.0	91.9	93.2
Overall	85.1	78.1	91.7	92.2

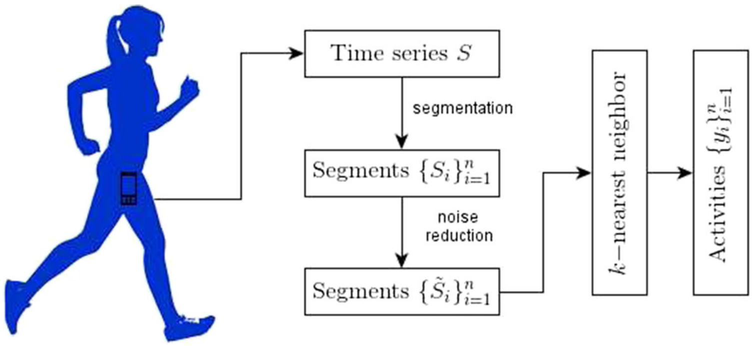

Fig. 1 Architecture of the proposed approach

where $\mathrm{X} = \{X_i\}_{i=1}^N$ is the set of test samples, $\mathrm{Y} = \{y_i\}_{i=1}^N$ is the set of answers, that corresponds to X. Here, the following notation for the indicator function is used:

$$[y \neq y'] = \begin{cases} 1, & if \quad y \neq y', \\ 0, & if \quad y = y'. \end{cases}$$

Let $\mathcal{W}$ be the set of parameters for the considered model. Thus, we treat the classification problem as the error function minimization problem:

$$\hat{\alpha} = \arg \min_{w \in \mathcal{W}} G(\mathrm{X}, \mathrm{Y}, f_{\mathrm{knn}}). \tag{2}$$

4 Algorithms

The human activity recognition procedure consists of two steps. The first step is data processing, that involves time series segmentation and noise reduction. The second step is classification of obtained segments. In this section, we provide algorithms for both of these steps, taking into account the specifics of the considered data.

4.1 Segmentation of the quasiperiodic time series

Hence, we propose an algorithm for time series segmentation under the assumption of their quasiperiodicity. This assumption is based on the nature of human motion, which consists in the repetition of movement phases. For instance, during the cycle of normal human gait, each lower limb goes through a stance phase and a swing phase and then returns to the stance phase [22, 31]. When human is inactive (e.g. standing or lying), this repetition is connected with respiration phases. Though the sequence of time series points measured by an accelerometer may change in each cycle, the characteristic form of the signal stays the same, which leads to the quasiperiodicity of time series. Based on the described physical property of human motions, we can define the fundamental period as a segment of time series, measured during one cycle of motion. From mathematical point of view quasiperiodicity of time series is equivalent to the fact that these series can be presented with high accuracy as a superposition of harmonic components with different periods. Then the harmonic with the longest period and this period are the fundamental ones. It should be noted that in the case of strict periodicity the fundamental period coincides the real period and the described method is also applicable.

Consider the one-dimensional time series $S = \{x_t\}_{t=1}^{M}$, and let $S' = \{x_t\}_{t=p}^{p+m}$ be its section, $p + 2m \leq M$. Our purpose is to extract the fundamental period from this section, assuming that its length m is much grater than the length of the period. First, we construct the Hankel matrix $\mathbf{A}$, that is a square matrix with constant skew-diagonals, which first line is S':

$$\mathbf{A} = \begin{pmatrix} x_p & x_{p+1} & x_{p+2} & \cdots & x_{p+m} \\ x_{p+1} & x_{p+2} & x_{p+3} & \cdots & x_{p+m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p+m} & x_{p+m+1} & x_{p+m+2} & \cdots & x_{p+2m-1} \end{pmatrix}.$$

Then we apply the singular value decomposition to the Hankel matrix: $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\mathsf{T}}$, where $\mathbf{\Lambda}$ is a diagonal matrix, $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrixes. Let $\{\lambda_i\}_{i=1}^{N}$ be the diagonal elements of matrix $\mathbf{\Lambda}$, and let $\mathbf{v}_k^{\mathsf{T}}$ be the k-th row of matrix $\mathbf{V}^{\mathsf{T}}$. Consider m principal components $\mathbf{y}_k = \mathbf{A}\mathbf{v}_k$ that account for 95 % of the variance.

The obtained components $\{\mathbf{y}_k\}_{k=1}^{m}$ are mapped into the m-dimensional phase space Φ_m, that consists of all possible values of these components. Thereby we get the phase trajectory $\mathcal{Y}$, every point of which is of the form $\mathcal{Y}(t) = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m)(t)$. To extract segments we should split this trajectory by the hyperplane $\mathcal{Y}_1 = \mathcal{Y}_2 = \ldots = \mathcal{Y}_m$. Assume that the hyperplane cuts the line section which connects the points $\mathcal{Y}(t)$ and $\mathcal{Y}(t+1)$. Then the time series S' should be split between x_{p+t} and x_{p+t+1}. This performs the segmentation $\left\{S_i'\right\}_{i=1}^{n}$ of S'. Figure 2 shows an example of segmentation results.

The obtained segments have different lengths. Since the basic classification algorithms, including k-nearest neighbor method, cannot be applied directly to this segments, they should be primarily reduced to an equal length. Assume that the segments should be rescaled to the predefined size L. For this reason, each segment S_i' is approximated by the polyline $\tilde{S}_i'$ of length $l_i = |S_i'|$, so that the k-th vertex of this polyline is the k-th term of the time series S_i'. Then we construct the time series X_i of length L, defined as:

$$X_i(j) = \tilde{S}_i'\left(\frac{j \times l_i}{L}\right),$$

that corresponds to the time series S_i'.

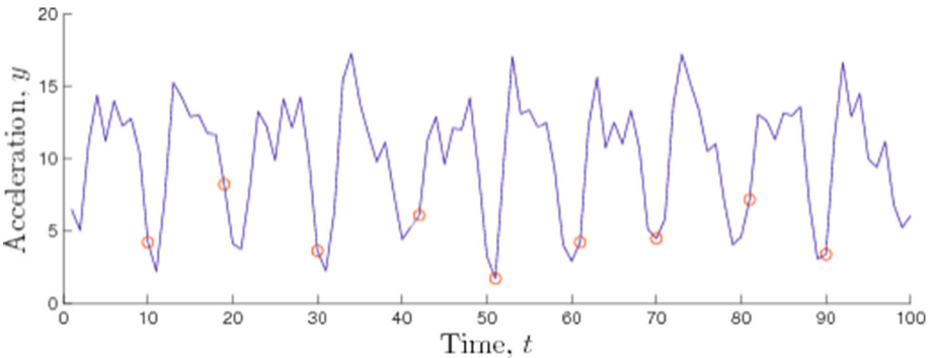

Fig. 2 Segmentation results. Red points mark the limits of segments

4.2 Noise reduction

We assume that the time series S contains normally distributed Gaussian noise with parameter $\mathcal{N}\left(0, \sigma^2\right)$. It introduces an additional error in the time series classification process, and in order to improve the classification accuracy we apply noise reduction. To remove noise we apply an algorithm, based on the singular value decomposition technique.

Here we use the notation introduced in the paragraph above. Let S' be the section of the time series S, and let $\mathbf{A}$ be the Hankel matrix corresponding to this section. Consider the singular value decomposition of matrix $\mathbf{A}$. To reduce the noise level we compute m principal components $\mathbf{y}_k = \mathbf{A}\mathbf{v}_k$ that account for 95 % of the variance. Then we construct matrix $\mathbf{Y}$, where k-th row is the k-th principal component $\mathbf{y}_k$. Define series $\tilde{S}'$, which i-th element is a skew-diagonal sum of matrix $\mathbf{Y}$. The obtained series $\tilde{S}'$ has less noise variance $\sigma'^2 < \sigma^2$ than the source series S'.

4.3 K-nearest neighbor method

To solve the time series classification problem in this paper we use the k-nearest neighbor method. The brief description of the method is provided below. Let $\mathrm{X} = \{X_i\}_{i=1}^{N}$ be the learning sample and let $\mathrm{Y} = \{y_i\}_{i=1}^{N}$ be the set of answers, that corresponds to X. Suppose we need to classify the object $X' \notin \mathrm{X}$. First, we arrange objects from the learning sample X in the distance ascending order to X':

$$\rho\left(X', X_{i_1}\right) \le \rho\left(X', X_{i_2}\right) \le \dots \le \rho\left(X', X_{i_k}\right). \tag{3}$$

Here, X_{i_n} is the n-th nearest neighbor for X' and ρ is the standard Minkowsky metric. Denote by y_{i_n} the class of the n-th neighbor. Then the class y' of X' is defined as:

$$y' = \arg\max_{y \in \mathrm{Y}} \sum_{n=1}^{k} \left[y_{i_n} = y\right] w(n, X_{i_n}), \tag{4}$$

where w is a weight function, that evaluates the significance of n-th neighbor for the classification of object X'. In this paper we consider four different weight functions, presented in the Table 2.

5 Experiments and evaluation

To evaluate the performance of the proposed model and compare it with existing alternatives we carried out a set of experiments, described in this section. Experiments were performed on the WISDM dataset [36] that represents time series from the accelerometer of a cell phone. These data were collected from 13 different people while they were performing a

Table 2 Weight functions w

Function	Value
$w_1(n, X_{i_n})$	1
$w_2(n, X_{i_n})$	$1 - \frac{n}{k+1}$
$w_3(n, X_{i_n})$	$\frac{1}{\rho(X', X_{i_n})}$
$w_4(n, X_{i_n})$	$\frac{1}{\rho(X', X_{i_n})^2}$

specific set of six activities: walking, jogging, ascending stairs, descending stairs, sitting and standing.

To analyze the segmentation accuracy we performed two sets of experiments. In the first set we used three conventional segmentation techniques: temporally-fixed segmentation, sliding window [14], and threshold-based algorithm [5, 10]. The second set was conducted using the proposed segmentation algorithm based on the fundamental period extraction. In both experiments we used k-nearest neighbor method to classify the obtained segments.

The classification technique plays the crucial role in the activity recognition problem. To compare the results obtained by the proposed model, we determined parameters and applied to the same dataset a classification method based on the neural network.

The classification accuracy for all methods was verified using 10-fold cross validation, the experiments we conducted on the set of 5600 samples.

5.1 Conventional segmentation techniques

To compare the proposed segmentation algorithm with alternatives, we applied three different online segmentation techniques to the same dataset. The first one is an elemental solution that splits the time series into segments of equal predefined length N. The second technique is based on the sliding window algorithm [14]. In this algorithm, a segment is grown until it exceeds a predefined error bound and then the process is repeated with the first data point not included in the newly obtained segment. Also, we considered a threshold-based algorithm [5, 10], in which a new segment is created every time the signal exceeds preliminary specified threshold. The value of the threshold, the length of segments N, the error bound and other parameters of these algorithms were determined in the computational experiment.

The obtained segments were classified using the k-nearest neighbor method. We considered different types of the weight function w (Table 2), and the number of neighbors k. Table 3 resumes the classification results obtained using the determined optimal parameters.

5.2 Segmentation based on the fundamental period extraction

In this section we describe experiment that was conducted to determine parameters and to estimate accuracy of the segmentation algorithm, proposed in this paper.

Considered algorithm had two parameters. First we had to select the threshold M, that cuts off segments of length less then M. The second parameter was length L, to which all the remained segments must be rescaled. In this experiment both M and L were over

Table 3 Classification results for the conventional segmentation techniques

Activity type, Y	Size of the test sample	Temporally-fixed	Threshold-based	Sliding window
Jogging	1600	90.6	91.3	94.3
Walking	1600	98.8	96.5	97.2
Upstairs	800	91.3	89.4	92.4
Downstairs	800	90.0	86.5	91.6
Sitting	400	100	87.5	100
Standing	400	100	75.2	97.5
Overall	5600	94.29	90.41	95.11
Processing time, s	–	3.1	3.9	161.7

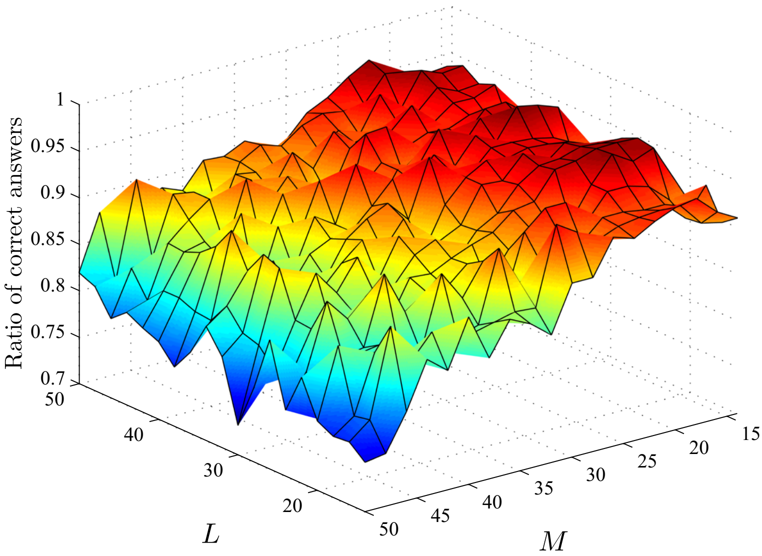

Fig. 3 Dependence between the ratio of correct answers and parameters M, L

the range 14 to 50 with a two-step and the segmentation algorithm was implemented for all possible combinations of these parameters. Obtained segments were classified using the k-nearest neighbor algorithm with the linear weight function w_2 and the value of $k = 1$. Figure 3 shows the classification results of the experiment.

The figure illustrates that the best classification results correspond to the values M that are between 20 and 25. There is also a slight increase of classification accuracy for coincident values of parameters M and L. Using parameters $M = 22$, $N = 22$ we obtained final classification results, presented in the Table 4. In comparison to the temporally-fixed segmentation algorithm, the overall accuracy was improved by more than 2 % to 96.61 %. In terms of individual class accuracy, the results for descending stairs remained the same, while the accuracy of recognition jogging and ascending stairs outperforms the previous results by almost 5 %. Comparing with the sliding window algorithm, the overall accuracy was also improved by 1.5 % while the processing time was about 4 times smaller.

Table 4 Classification results for algorithm, based on the fundamental periods extraction

Activity type, Y	Size of the test sample	Correct answers, %
Jogging	1600	96.9
Walking	1600	99.4
Upstairs	800	95.0
Downstairs	800	90.0
Sitting	400	100
Standing	400	97.5
Overall	5600	96.61
Processing time, s	–	32.3

Table 5 Classification results for the neural network

Activity type, Y	Size of the test sample	Correct answers, %
Jogging	1600	96.3
Walking	1600	100
Upstairs	800	68.8
Downstairs	800	85.0
Sitting	400	100
Standing	400	100
Overall	5600	92.32

5.3 Neural network-based approach

This paragraph is devoted to the approach based on the neural network. Here we consider two-layer perceptron with sigmoidal activation function, trained with the backpropagation method.

Our first experiments showed that the performance of this method on the set of raw segments is rather low. Thus, we extracted a feature vector for each segment and used it in the further classification process. We calculated 15 features for each 3-dimensional segment, 5 for each direction: the mean value, the mean absolute value, the difference between maximum and minimum value, the total value of absolute differences and the mean value of the contained fundamental period.

The main parameter of the neural network was the number of neurons H in the hidden unit. To estimate optimal value of this parameter we implemented neural network for H over the range 8 to 32 and computed the classification accuracy and the root-mean-square error. The best results were for value of H near 24, and Table 5 presents classification results for this parameter. The overall number of correct answers was 517 or 92.32 %. For four common activities, jogging, walking, sitting and standing, we archived perfect accuracy, while the most complicated task for the classifier was to distinguish ascending and descending stairs.

5.4 Processing time

In this section we estimate the velocity of the described KNN-based method and compare it with the velocity of the methods previously proposed in papers [11, 15]. There was no information about the processing time in these papers, so we implemented the algorithms exactly as they were described there to test them under the same conditions. As in original works, we used algorithms from the WEKA machine learning library for the classification of the segments. All methods were launched on the same ten datasets of different lengths ranging between 1400 and 14000 on an Intel Core-i3 330M 2 CPU 2.13 GHz machine. The processing time for each method is presented on the Fig. 4. The slowest methods are based on logistic regression and neural network [15], they require far more processing time than the rest ones. Much faster are methods based on SVM [11] and J48 decision tree [15], they show similar velocity. The shortest time is achieved by the KNN-based method. Though KNN is not generally a fast algorithm, in this case due to the small length of segments (22 points) this method shows the best performance.

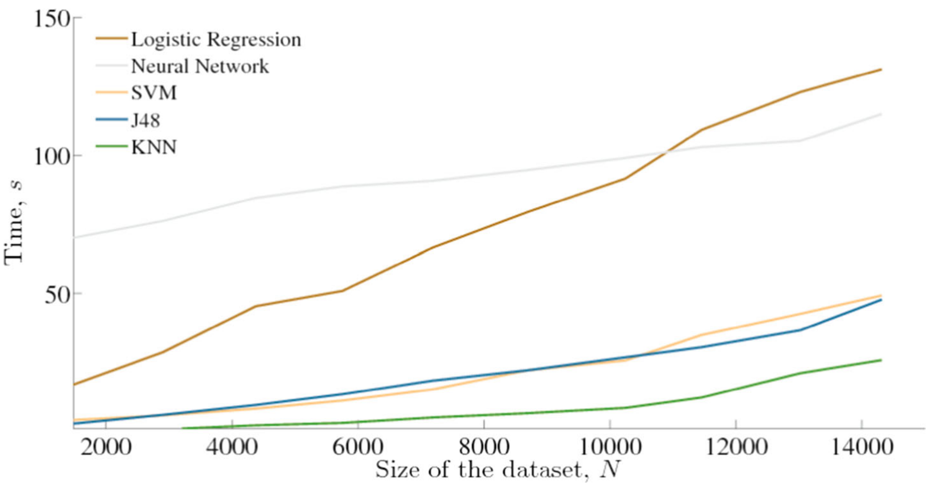

Fig. 4 Dependence between the length of the dataset and the processing time for five activity recognition methods

6 Conclusion and future work

In this paper, we proposed a method for human activity recognition using time series, collected from an accelerometer of a cell phone. We introduced an algorithm for online time series segmentation and proposed the k-nearest neighbor method for the classification of obtained segments. The accuracy of activity recognition was verified in three computational experiments on the WISDM dataset. The experiment, in which we used a combination of feature extraction and neural network has shown almost the same overall accuracy as a multilayer perceptron in the work [15]. The method, based on the basic time series segmentation and the k-nearest neighbor algorithm has improved the results by about 2 %. The best results, over 96 % of correct answers, was shown by a combination of the advanced segmentation and the k-nearest neighbor algorithms. This method has outperformed the rest in terms of overall and individual accuracy, especially in the distinguishing between ascending and descending stairs. More important is that it requires the segments to be of the length 20-30 or 1.0-1.5 seconds, that is about 7 times shorter than in works [11, 15]. This reduces computational costs and together with a low classification time of each segment (much less than a second) makes it possible real-time activity recognition.

We plan to improve our activity recognition system in several ways. First, we intend to teach our system to recognize additional activities, such as driving, jumping and riding the bicycle. For this purpose we need to extend the existing dataset by adding new training samples. Second, we plan to reduce the size of the training set by removing similar time series. This will lower the computational load and the memory occupied by the program. Finally, we hope to develop an application for mobile devices that performs real-time activity recognition and distribute the code as an open library for Java applications.

Acknowledgments This work is supported by the Russian Foundation for Basic Research (RFBR) under grant number 14-07-31326 and by the Ministry of Education and Science of the Russian Federation, RFMEFI60414X0041.

References

1. Altun K, Barshan B (2010) Human activity recognition using inertial/magnetic sensor units. In: Salah A, Gevers T, Sebe N, Vinciarelli A (eds) Human behavior understanding, volume 6219 of lecture notes in computer science. Springer, Berlin Heidelberg, pp 38–51
2. Bao L, Intille SS (2004) Activity recognition from user-annotated acceleration data. In: Ferscha A, Mattern F (eds) Pervasive computing, volume 3001 of lecture notes in computer science. Springer, Berlin Heidelberg, pp 1–17
3. Barralon P, Vuillerme N, Noury N (2006) Walk detection with a kinematic sensor: frequency and wavelet comparison. Proceedings of the 28th IEEE EMBS annual international conference 1:1711–1714
4. Bellman R (1961) On the approximation of curves by line segments using dynamic programming. Commun ACM 4(6):284
5. Brajdic A, Harle R (2013) Walk detection and step counting on unconstrained smartphones. UbiComp 2013 1:225–234
6. Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc 2(2):121–167
7. Chiang S-Y, Kan Y-C, Tu Y-C, Lin H-C (2012) Activity recognition by fuzzy logic system in wireless sensor network for physical therapy. In: Watada J, Watanabe T, Phillips-Wren G, Howlett RJ, Jain LC (eds) Intelligent Decision Technologies, volume 16 of Smart Innovation, Systems and Technologies. Springer, Berlin Heidelberg, pp 191–200
8. Chin T-J, Wang L, Schindler K, Suter D (2007) Extrapolating learned manifolds for human activity recognition. 2007 IEEE International Conference on Image Processing 1:383–384
9. Foerster F, Fahrenberg J (2000) Motion pattern and posture: Correctly assessed by calibrated accelerometers. Behav Res Methods Instrum Comput 32(3):450–457
10. Goyal P, Ribeiro VJ, Saran H, Kumar A (2011) Strap-down pedestrian dead-reckoning system. In: 2011 International conference on indoor positioning and indoor navigation (IPIN), pp 1–7
11. Kai K, Tony P (2013) Classification of mobile device accelerometer data for unique activity identificationunique activity identification. http://cs229.stanford.edu/proj2013/PratkanisKuspa-ClassificationOfMobileDeviceAccelerometerDataforUniqueActivityIdentification.pdf
12. Khan AM, Lee YK, Kim TS (2008) Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets. In: Engineering in medicine and biology society, 2008, EMBS 2008, 30th annual international conference of the IEEE, pp 5172–5175
13. Keogh E, Chu S, Hart D, Pazzani M (2001) An online algorithm for segmenting time series. In: Data Mining, 2001 Proceedings IEEE International Conference on ICDM 2001, pp 289–296
14. Keogh E, Chu S, Hart D, Pazzani M (1993) Segmenting time series: A survey and novel approach. In: In an edited volume, data mining in time series databases, published by world scientific. Publishing Company, pp 1–22
15. Kwapisz JR, Weiss GM, Moore SA (2011) Activity recognition using cell phone accelerometers. SIGKDD Explor Newsl 12(2):74–82
16. lai Chung F, Tak chung F, Ng V, Luk RWP (2004) An evolutionary approach to pattern-based time series segmentation. IEEE Trans Evol Comput 8(5):471–489
17. Liu W, Li Y, Lin X, Tao D, Wang Y (2014) Hessian-regularized co-training for social activity recognition. PLoS ONE 9:e108474
18. Liu W, Liu H, Tao D, Wang Y, Ke L (2014) Multiview hessian regularized logistic regression for action recognition. In: Signal Processing
19. Liu W, Zhang H, Tao D, Wang Y, Lu K (2014) Large-scale paralleled sparse principal component analysis
20. Lockhart J (2011) Mobile sensor data mining. Fordham Undergraduate Research Journal 1:67–68
21. Lockhart JW, Weiss GM (2014) Limitations with activity recognition methodology & #38; data sets. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct Publication, UbiComp ’14 adjunct. ACM, New York, pp 747–756
22. Loudon J (2008) The clinical orthopedic assessment guide. Human Kinetics 2:395–408
23. Mannini A, Sabatini AM (2010) Machine learning methods for classifying human physical activity from on-body accelerometers. Sensors 10(2):1154–1175
24. Mi Z, Sawchuk AA (2013) Human daily activity recognition with sparse representation using wearable sensors. IEEE Journal of Biomedical and Health Informatics 17(3):553–560
25. Mi Z, Alexander AS (2011) Manifold learning and recognition of human activity using body-area sensors. 10th International Conference 2:7–13

26. Mo L, Liu S, Gao RX, John D, Staudenmayer JW, Freedson PS (2012) Wireless design of a multisensor system for physical activity monitoring. IEEE Trans Biomed Eng 59(11):3230–3237
27. Nyan M, Tay F, Seah K, Sitoh Y (2006) Classification of gait patterns in the time-frequency domain. J Biomech 39:2647–2656
28. Olivares A, Ramirez J, Gorriz JM, Olivares G, Damas M (2012) Detection of (in)activity periods in human body motion using inertial sensors: A comparative study. Sensors 12(5): 5791–5814
29. Ravi N, Dandekar N, Mysore P, Littman ML (2005) Activity recognition from accelerometer data. In: Proceedings of the 17th conference on innovative applications of artificial intelligence - Volume 3, IAAI'05. AAAI Press, pp 1541–1546
30. Schuldt C, Laptev I, Caputo B (2004) Recognizing human actions: A local svm approach. In: Proceedings of the pattern recognition, 17th international conference on (ICPR'04) Volume 3 - Volume 03, ICPR '04. IEEE Computer Society, Washington, pp 32–36
31. Shultz SJ, Houglum PA, Perrin DH (2005) Examination of muskoskeletal injuries. Human Kinetics 2:55–60
32. Tapia EM, Intille SS, Haskell W, Larson K, Wright J, King A, Friedman R (2007) Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In: 2007 11th IEEE international symposium on wearable computers, pp 37–40
33. Thuer G, Verwimp T (2008) Step detection algorithms for accelerometers. Master's thesis, Paper of the E-LAB Master Thesis. http://nitarc.be/map/paper/AMBIT_ThuerVerwimp.pdf
34. Wang S, Yang J, Chen N, Chen X, Zhang Q (2005) Human activity recognition with user-free accelerometers in the sensor networks. In: Neural networks and brain, 2005, international conference on ICNN B '05, vol 2, pp 1212–1217
35. Wen-Ming C, Li N (2010) Face recognition based on manifold learning and renyi entropy. Nat Sci 2:49–53
36. Wisdm's activity prediction dataset (2012). http://www.cis.fordham.edu/wisdm/dataset.php
37. Ying H, Silex C, Schnitzer A, Leonhardt S, Schiek M (2007) Automatic step detection in the accelerometer signal. In: Leonhardt S, Falck T, Mähönen P (eds) 4th International workshop on wearable and implantable body sensor networks (BSN 2007), volume 13 of IFMBE proceedings. Springer, Berlin Heidelberg, pp 80–85
38. Yurur O, Liu C-H, Moreno W (2013) Unsupervised posture detection by smartphone accelerometer. Electron Lett 49(8):562–564
39. Zhou Y, Cheng Z, Jing L (2014) Threshold selection and adjustment for online segmentation of one-stroke finger gestures using single tri-axial accelerometer. Multimedia Tools and Applications 70:1–20

Andrey D. Ignatov is a student of the Moscow Institute of Physics and Technology (MIPT). Currently, he is a member of the ECG Informational Processing Research group, an assistant in the Laboratory of the Experimental Economics in the MIPT University and the author and co-author of several research publications in the field of machine learning applications for medical purposes. His current interests include computational intelligence, data analysis, bio-signal analysis and intelligent healthcare systems.

Vadim V. Strijov obtained Ph.D. in Physics and Mathematics from the Computing Centre of the Russian Academy of Sciences in 2002 with a thesis on Mathematical Modelling and Data Analysis and M.S. in System Analysis and Computer-Aided Design from Ufa Aircraft Institute in 1992. He is the author of the Expert estimation concordance theory. Field of research: Machine Learning, Data Analysis, Model Selection, Structure Learning, Preference Learning. Current duties: Principal Investigator at the Computing Center of the Russian Academy of Sciences, Associate Professor at the Moscow Institute of Physics and Technology, Editor in chief of the Journal of Machine Learning and Data Analysis.