A Comparative Evaluation of Acceleration and Jerk in Human Activity Recognition Using Machine Learning Techniques

Nurul Retno Nurwulan and Gjergji Selamaj

Abstract Acceleration data have been widely used to study human activity recognition. However, the acceleration data collected from the accelerometer do not consider the force of gravity. Thus, it has a difficulty in discriminating closely similar activities. Jerk, the derivative of acceleration, is able to describe the changes of body accelerations correctly and mutually exclusive from the sensor orientation. This study aimed to compare the performance of jerk compared to acceleration data. The dataset used to do the comparison was collected from a triaxial accelerometer built-in Samsung smartphone with a sampling frequency of 50 Hz attached to the waist of thirty subjects. The subjects performed walking on a flat surface, walking upstairs, walking downstairs, sitting, standing, and lying down. The feature data were then grouped into three categories: acceleration features, jerk features, and combined features of acceleration and jerk. The evaluation was done using k-Nearest neighbors (KNN), linear discriminant analysis (LDA), support vector machine (SVM), and random forest (RF). The results showed that jerk features performed worse than the acceleration features. However, the combined acceleration and jerk features yielded the highest accuracy with above 87% for all classifiers. The present findings show that acceleration is still better than jerk in recognizing the pattern in human activity. The present study is, therefore, increasing the understanding of acceleration data and its derivative in human activity recognition.

Keywords Human activity recognition · Accelerometer · Acceleration · Jerk · Pattern recognition

N. R. Nurwulan (✉)
Sampoerna University, Jakarta 12780, Indonesia
e-mail: nurul.nurwulan@sampoernauniversity.ac.id

G. Selamaj
Universiteti Ismail Qemali, Vlorë 9401, Albania
e-mail: gjergji.selamaj147@gmail.com

1 Introduction

Human activity recognition has been proven for its benefits in various fields such as healthcare [1–3]. Human activity recognition is often used to help the physicians to make a correct diagnosis for patients. Accelerometers are a common device in human activity recognition due to its ability to provide an objective, non-intrusive measure of activity and the high resolution of data acquisition [1]. Triaxial accelerometers in smartphones are the practical and cost-effective choice for human daily activity recognition [4].

Various studies evaluating human daily activity using the acceleration data by extracting its features either the time or frequency-domain features. High accuracy results with more than 90% accuracy have been reported in past studies [5–7]. However, the high accuracy results are only possible when the activities are simple, well-separated, and performed carefully based on researchers' instructions. If the activities are more complex, closely similar to each other, and performed in a more natural manner, human activity recognition becomes a challenging problem since there is no clear way to relate the signal data to a specific activity.

Past studies reported promising results obtained from using jerk-based features for activity recognition in animals [8]. Jerk was used because it is not possible to direct the animals with specific instructions to perform the tasks. Further, jerk could describe the sensor orientation although the sensor is loosely attached and frequently shifted [8]. Motivated by the promising results in animals, the aim of the present study was to evaluate the performance of jerk compared to the acceleration in recognizing human daily activities. This study examined whether jerk, the derivative of acceleration, could overcome the low sensitivity of acceleration data in discriminating complex and closely similar human daily activities. Machine learning techniques were used for the evaluation due to their robustness [9].

This paper is organized as follows. Section 2 describes the materials and method in obtaining the dataset and a brief description of the theoretical background behind acceleration and jerk, feature extraction, and machine learning classifiers. Section 3 presents results and discussion. Finally, the conclusion is presented in Sect. 4.

2 Materials and Method

The dataset used to evaluate the performance of acceleration and jerk was taken from the study by Anguita et al. [4]. The data was collected using the triaxial accelerometer built-in Samsung smartphone from thirty subjects aged between nineteen and forty-eight. The smartphone was attached to the waist with the sampling frequency rate was set to 50 Hz. Each subject performed walking, walking upstairs, walking downstairs, sitting, standing, and lying down. The dataset was pre-processed with a fifty percent overlapping window of 2.56 s. Then, time- and frequency-domain features from acceleration and jerk data were extracted.

The extracted features were then analyzed using supervised machine learning techniques, namely K-Nearest neighbors (KNN), Fisher's linear discriminant analysis (LDA), support vector machine (SVM), and random forest (RF). Ten-fold cross-validation using Weka environment was done to evaluate the performance of acceleration and jerk datasets.

2.1 Acceleration and Jerk

Acceleration is the rate of change of velocity in terms of speed and/or direction. The acceleration is commonly used in human activity recognition researches. However, acceleration data exclude the force of gravity and thus it is just a consequence of static load [10]. Jerk is the derivative of acceleration and it is felt like the change of accelerations. The magnitude of jerk describes the changes of accelerations independently from the sensor orientation [8]. Thus, jerk can overcome the problem of unknown sensor orientation that might happen when the smartphone on the subjects' pocket moves its position.

2.2 Feature Extraction

There are two modes of features in human physiological data: time-domain features that refer to the variation of the amplitude of the signal with time; and frequency-domain features that are more robust and need signal pre-processing such as the Fast Fourier Transform (FFT) [11]. In this study, the features are mean, standard deviation, mean absolute deviation, maximum, minimum, signal magnitude area, energy, interquartile range, and spectral entropy in both the time- and frequency-domain modes.

2.3 Classifiers

In this study, we evaluated the performance of acceleration and jerk data using supervised classifiers: K-Nearest neighbors (KNN), Fisher's linear discriminant analysis (LDA), support vector machine (SVM), and random forest (RF). Weka environment was used for the machine learning analyses. The KNN is a fast and simple supervised classifier [9]. The classification is done by determining the similarity between the training set and new observation that is assigned to the most similar class based on the most votes [12]. Despite its simplicity, KNN is an effective classifier for human activity recognition [13].

The LDA is capable of ensuring the projections of samples from different classes to a line are well-separated. It can separate the classes with an obvious distance

between mean values and small variances [14]. Whereas, the SVM can distinctly classify the patterns of the data points into different classes. The difference between the SVM and LDA is that the SVM selects the hyperplane with maximum margin distance [15]. The RF constructs a large number of uncorrelated decision trees as an ensemble at training time with the output is the classification of individual trees [16]. Each individual tree has a class prediction and the class with the most votes will become the model prediction. Past studies reported high accuracy using the RF as a classifier [9, 16, 17].

3   Results

This study evaluated the performance of both acceleration and jerk in order to see whether jerk could really recognize human activity better than acceleration. The evaluation was conducted for three different groups: acceleration, jerk, and combination of both acceleration and jerk. Then, these groups were evaluated using KNN, LDA, SVM, and RF. Ten-fold cross-validation with 90% of the data used for training was selected in this study. The dataset was divided into ten folds with the same class distribution as the original dataset. Each fold was used once to test the performance of the classifier from the combined data of the remaining nine folds. As can be seen in Table 1, the RF outperformed the other classifiers in all groups with the highest accuracy is 88.30% in the combination of acceleration and jerk group.

Since the RF outperformed the other classifiers, we focus on the RF to compare the performance of acceleration and jerk. Tables 2, 3, and 4 show the confusion matrices of the RF for acceleration, jerk, and combination of acceleration and jerk, respectively. The standing, sitting, and lying are misclassified with one another. However, acceleration and combination groups classified better than jerk.

4   Discussion

We used ten-fold cross-validation because it has a low variance than a single hold-out set estimator. For the dataset in this study, the test set is very small if the single hold-out set is used because 90% of the data are used for training, and 10%

Table 1 Performance evaluation of the classifiers

Attribute	KNN (%)	LDA (%)	SVM (%)	RF (%)
Acceleration	85.06	85.84	85.79	87.15
Jerk	81.24	83.82	83.45	83.52
Combined	87.27	87.31	87.42	88.30

Table 2 Confusion matrix of RF for acceleration

Labeled	Recognized results					
	Standing	Sitting	Lying	Walking	Downstairs	Upstairs
Standing	323	174	171	0	0	0
Sitting	231	218	174	0	0	0
Lying	200	151	328	0	1	1
Walking	0	0	0	508	14	81
Downstairs	0	0	0	25	431	37
Upstairs	0	0	1	106	24	410

Table 3 Confusion matrix of RF for jerk

Labeled	Recognized results					
	Standing	Sitting	Lying	Walking	Downstairs	Upstairs
Standing	240	197	231	0	0	0
Sitting	226	198	199	0	0	0
Lying	218	196	265	0	0	2
Walking	0	0	0	369	146	88
Downstairs	0	0	0	114	329	50
Upstairs	0	0	1	96	20	424

Table 4 Confusion matrix of RF for the combination of acceleration and jerk

Labeled	Recognized results					
	Standing	Sitting	Lying	Walking	Downstairs	Upstairs
Standing	353	156	159	0	0	0
Sitting	235	224	164	0	0	0
Lying	165	126	388	0	0	2
Walking	0	0	0	572	5	26
Downstairs	0	0	0	12	431	50
Upstairs	0	0	0	31	25	485

used for testing. The findings of this study showed that RF yielded the highest accuracy. RF is an effective method to rank the importance of variables. Unlike the other decision tree classifiers, each tree in RF can only select a random subset of features that makes it possible to increase the variation among the trees in the model. Thus, the classification will result in higher accuracy considering the low correlation across trees [9, 17].

The performance of a combination of both acceleration and jerk yielded in the highest accuracy. However, when the acceleration and jerk were evaluated as individual groups, acceleration resulted in higher accuracy than jerk. The logic behind using jerk-based features is because jerk is the derivative of acceleration.

Thus, it is supposed to be able to detect slight changes in closely similar activities [8]. However, it is not the case in this current study. This could be caused by the fact that the past study [8] used jerk to evaluate the activity recognition in animals with less-to-none instructions from the researchers. On the contrary, the dataset in this study was taken from human subjects with detailed instructions for each activity. Therefore, in contrast to the past study [8], jerk did not perform well in recognizing human activities. This might happen because there were little changes in forces in human signal data during standing, sitting, and lying. Jerk is felt when there is an obvious change in force [10].

As for walking on a flat surface, walking downstairs, and walking upstairs, jerk is also the worst in classifying the activities correctly. Both acceleration and combined groups are able to distinguish the walking activities. This might also happen because the subjects in this study were young adults with good postural control [18]. They could compensate with both increasing and decreasing forces that happened during walking downstairs and upstairs. Based on the results in this study, it can be said that jerk does not perform better than acceleration in human activity recognition when the change of force in the activities is minimum. This finding increases the understanding of the usage of acceleration data and its derivative in human activity recognition. Further studies with more variation of activities including activities with an obvious change of force, to evaluate the performance of jerk and other derivatives of acceleration are needed. In addition, future studies should also evaluate the performance of jerk in other age brackets such as older adults since the other age brackets might not have a good postural control like the young adults.

5   Conclusion and Future Work

This paper has presented the performance evaluation of jerk compared to acceleration using machine learning techniques, namely the k-Nearest neighbors (KNN), Fisher's linear discriminant analysis (LDA), support vector machine (SVM), and random forest (RF). Based on the evaluation results, it can be concluded that the RF outperformed the other machine learning techniques in terms of accuracy and jerk did not perform better than acceleration in recognizing human activities. This could happen because there were only slight force changes in acceleration in the activities being evaluated. The results of the current study increase the understanding of the usage of acceleration data and its derivative in human activity recognition. Further studies evaluating activities with an obvious change of force are important to confirm whether jerk is more sensitive than acceleration in activities with a bigger change of force.

References

1. Nurwulan NR, Jiang BC (2016) Possibility of using entropy method to evaluate the distracting effect of mobile phones on pedestrian. Entropy 18:390
2. Ogbuabor G, Labs R (2018) Human activity recognition for healthcare using smartphones. In: Proceedings of the 10th international conference on machine learning and computing, Macau, pp 41–46
3. Yan W, Shuang C, Yu H (2019) A survey on wearable sensor modality centred human activity recognition in health care. Expert Syst Appl 137:167–190
4. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL (2013) A public domain dataset for human activity recognition using smartphones. In: Proceedings of the 21th European symposium on artificial neural networks, computational intelligence and machine learning, ESANN, Bruges, Belgium
5. Walse KH, Dharaskar RV, Thakare VM (2016) PCA based optimal ANN classifiers for human activity recognition using mobile sensors data. In: Proceedings of 1st international conference on information and communication technology for intelligent systems, Ahmedabad, India, pp 429–436
6. Fullerton E, Heller B, Munoz-Organero M (2017) Recognizing human activity in free-living using body-worn accelerometers. IEEE Sens J 17(16):5290–5297
7. Pavey TG, Gilson ND, Gomersall SJ, Clark B, Trost SG (2017) Field evaluation of random forest activity classifier for wrist-worn accelerometer data. J Sci Med Sport 20(1):75–80
8. Hämäläinen W, Järvinen M, Martiskainen P, Mononen J (2011) Jerk-based feature extraction for robust activity recognition from acceleration data. In: Proceedings of the 11th international conference on intelligent systems design and applications (ISDA 2011). IEEE computer society, pp 831–836
9. Nurwulan NR (2020) Performance evaluation of decomposition methods in perturbed walking. Revue d'Intelligence Artificielle 34(1):45–50
10. Eager D, Pendrill AM, Reistad N (2016) Beyond velocity and acceleration: jerk, snap and higher derivatives. Eur J Phys 37:065008
11. Nurwulan N, Jiang BC (2020) Window selection impact in human activity recognition. Int J Innovative Technol Interdisc Sci 3(1):381–394
12. Suto J, Oniga S, Pop Sitar P (2016) Feature analysis to human activity recognition. Int J Comput Commun Control 12(1):116–130
13. Subasi A, Radhwan M, Kurdi R, Khateeb K (2018) IoT based mobile healthcare system for human activity recognition. In: Proceedings of 15th learning and technology conference, Jeddah, Saudi Arabia
14. Hassan MM, Uddin MZ, Mohamed A, Almogren A (2018) A robust human activity recognition system using smartphone sensors and deep learning. Future Gener Comput Syst 81:307–313
15. Zeng H, Li K, Tian X, Wei N, Song R, Zhou L (2018) Classification of hand motions using linear discriminant analysis and support vector machine. Chinese Autom Congr, Jinan, China
16. Nunes UM, Faria DR, Peixoto P (2017) A human activity recognition framework using max-min features and key poses with differential evolution random forests classifier. Pattern Recogn Lett 99:21–31
17. Mehrang S, Pietila J, Korhonen I (2018) An activity recognition framework deploying the random forest classifier and a single optical heart rate monitoring and triaxial accelerometer wrist-band. Sensors 18(2):613
18. Selamaj G (2020) Impacts of mobile phone distractions on walking performance. Indonesian J Comput Eng Design 2(1):32–37