Expert Systems with Applications 41 (2014) 6067–6074

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Unsupervised learning for human activity recognition using smartphone sensors

Yongjin Kwon, Kyuchang Kang, Changseok Bae *

SW·Content Research Laboratory, Electronics and Telecommunications Research Institute, 218 Gajeong-ro, Yuseong-gu, Daejeon 305-700, Republic of Korea

ARTICLE INFO

Keywords:
Human activity recognition
Unsupervised learning
Healthcare services
Smartphone sensors
Sensor data analysis

ABSTRACT

To provide more sophisticated healthcare services, it is necessary to collect the precise information on a patient. One impressive area of study to obtain meaningful information is human activity recognition, which has proceeded through the use of supervised learning techniques in recent decades. Previous studies, however, have suffered from generating a training dataset and extending the number of activities to be recognized. In this paper, to find out a new approach that avoids these problems, we propose unsupervised learning methods for human activity recognition, with sensor data collected from smartphone sensors even when the number of activities is unknown. Experiment results show that the mixture of Gaussian exactly distinguishes those activities when the number of activities k is known, while hierarchical clustering or DBSCAN achieve above 90% accuracy by obtaining k based on Caliński–Harabasz index, or by choosing appropriate values for ε and *MinPts* when k is unknown. We believe that the results of our approach provide a way of automatically selecting an appropriate value of k at which the accuracy is maximized for activity recognition, without the generation of training datasets by hand.

1. Introduction

With an increasing interest in human health, it is necessary to obtain objective information of a patient to provide appropriate healthcare services. Although wearable sensors are convenient and useful for obtaining undistorted information from a human body, they may sometimes become an obstacle for healthcare services. Since most wearable sensors are attached directly to the user's body, some people may feel a sense of inconvenience. These disadvantages may generate incorrect results from healthcare services, making people refuse such services.

The inconvenience of wearable sensors can be resolved through the use of mobile devices (Anguita, Ghio, Oneto, Parra, & Reyes-Ortiz, 2012). Mobile devices such as smartphones and music players have recently become pervasive to the point that many people carry them at all times. Mobile devices usually incorporate various sensors, such as GPS sensors, accelerometers, or gyroscopes. For instance, most Android-powered smartphones have built-in sensors, and these sensors can be controlled by developers using Android APIs. Thus, instead of attaching a wearable sensor to the user's body, a mobile device can be kept in the pocket of the user's pants, which is less bothersome.

One of the most impressive studies on the sensors used in mobile devices is activity recognition. Human activity recognition is promising research that has been widely studied in recent decades. The sensors used in mobile devices can provide useful information for activity recognition. In particular, accelerometers, which are used as a source of fundamental information in many studies on activity recognition, are included in most smartphones (Brezmes, Gorricho, & Cotrina, 2009). If methods to recognize a user's activity with a mobile device can be achieved, it will be possible to develop many useful healthcare applications. For instance, we can monitor the activity states of a user, and can aggregate such activity states over time to obtain daily, weekly, and monthly ratios of activities. These ratios can be used to determine whether a user exercises regularly or sits for too long. Based on the estimated activity ratios, the application can recommend an appropriate activity to the user, such as walking outside or stretching. If we think beyond healthcare services, we may come up with more diverse useful applications.

Most studies on activity recognition have utilized accelerometers in either wearable sensors or smartphones. Interestingly, most such studies have considered activity recognition as a supervised learning problem. In fact, it is natural to think of activity recognition as a supervised learning problem because activity recognition

* Corresponding author. Tel.: +82 42 860 3816.
E-mail addresses: scocso@etri.re.kr (Y. Kwon), k2kang@etri.re.kr (K. Kang), csbae@etri.re.kr (C. Bae).

http://dx.doi.org/10.1016/j.eswa.2014.04.037
0957-4174/

classifies a given sensor dataset based on the activities. For the supervised learning, one of the most important things is the training dataset. The generation of a training dataset, however, is a tedious and labor-intensive work. Furthermore, the training dataset has some drawbacks. First the number of sensor records may be huge. For instance, when sensor data are recorded at a sampling rate of 50 Hz, the number of sensor records for an hour is 180,000. It is time consuming to label the whole records. Second it is difficult to remember the activities performed at a specific time. Especially, for short periods of an activity or at the boundary of consecutive activities, it is difficult to assign the correct activities. Last when the number of activities to be recognized varies, the training dataset should be regenerated. For these reasons, we need to seek new approaches of activity recognition without generating training dataset.

In this paper, we propose activity recognition using unsupervised learning assuming that the number of activities k is unknown. Although there are a few studies that have applied unsupervised learning approaches, they are inadequate to discuss the effectiveness of unsupervised learning for activity recognition, especially when k is unknown. Hence, we present experiments that examine different types of unsupervised learning algorithms to show that our approach can find an appropriate set of k at which the accuracy is maximized and can separate different activities. We first collected a series of sensor data from smartphones as the users performed five activities: walking, running, sitting, standing, and lying down. We then generated a list of feature vectors by aggregating the sensor data over sliding windows. To verify the usefulness of unsupervised learning techniques, we examined three clustering algorithms while assuming that the number of clusters is known, and observed whether they divided the vectors into five clusters precisely. We then investigated four clustering algorithms by assuming that the number of clusters is unknown to see whether they can still be applied to any series of sensor data, which are collected during an arbitrary number of activities. Hence, we observed whether unsupervised learning approaches can play an important role in activity recognition in future works.

The rest of the paper is organized as follows. In Section 2, we present the previous approaches on human activity recognition. Section 3 explains the details of the experimental setup, feature extraction, and descriptions of the sensor data for each activity. Section 4 shows the experiment results of the unsupervised learning techniques. Finally, Section 5 provides some concluding remarks regarding this research.

2. Related work

Many investigators have tried to recognize human activities using various combinations of sensors, which are included in cameras (Uddin, Thang, Kim, & Kim, 2011), wearable computers, and mobile devices. Accelerometers are common sensors for activity recognition because the accelerations measured rely on which activity the user performs (Mathie, Coster, Lovell, & Celler, 2004). Therefore, activity recognition has been studied using a number of accelerometers or with a combination of accelerometers and other types of sensors.

In some researches, attempts at using multiple accelerometers attached to different locations have been progressively conducted. The authors in Veltink, Bussmann, de Vries, Martens, and van Lummel (1996) performed a number of experiments that used two or three uniaxial accelerometers to distinguish several activities, including standing, sitting, lying down, walking, ascending stairs, descending stairs, and cycling. The researchers in Aminian et al. (1999) studied whether activities (lying down, sitting, standing, and walking) can be recognized using two accelerometers, one attached to the chest and the other to the rear of the thigh. In Foerster and Fahrenberg (2000), three uniaxial accelerometers were strapped to the sternum, and two uniaxial accelerometers were located on the left and right thighs to detect four basic activities (sitting, standing, lying down, and moving). Using only two accelerometers, the authors in Laerhoven and Cakmakci (2000) identified seven activities, sitting, standing, walking, running, climbing stairs, descending stairs, and riding a bicycle. The researchers in Bussmann et al. (2001) provided a technical description of an activity monitor in which four uniaxial accelerometers and one biaxial accelerometer were used to recognize activities such as standing, sitting, walking, climbing up, climbing down, cycling, driving, running, and laying down. The authors in Mantyjarvi, Himberg, and& Seppanen (2001) tried to recognize different moving activities, such as walking on a level surface, walking downstairs, walking upstairs, and not walking using two sets of accelerometers. In Bao and Intille (2004), sensor data were collected from 20 individuals wearing five biaxial accelerometers while doing twenty activities to show that a decision-tree classifier can recognize such activities with reasonable accuracy. The researchers in Krishnan, Colbry, Juillard, and Panchanathan (2008) examined five activities, sitting, standing, walking, running, and lying down, using two accelerometers. In Krishnan and Panchanathan (2008), the authors collected data from ten subjects wearing three accelerometers to identify seven activities, walking, sitting, standing, running, bicycling, lying down, and climbing stairs. In Mannini and Sabatini (2010), some experiments were performed that are similar to those in Bao and Intille (2004) in that they also used five biaxial accelerometers to collect the sensor data from 20 individuals, but applied classifiers based on Hidden Markov Models. The authors in Banos, Damas, Pomares, Prieto, and Rojas (2012) measured accelerations using a set of accelerometers placed on the hip, wrist, arm, ankle, and thigh to recognize four activities, walking, sitting, standing, and running. In Zhang, Liu, Zhu, and Zhu (2012), the authors varied the number of accelerometers, with different settings, and examined eight activities, standing, walking, running, jumping, lying, sitting, tooth brushing, and eating.

Some studies have combined accelerometers with other sensors, such as gyro sensors, microphones, and digital compasses. The researchers in Foerster, Smeja, and Fahrenberg (1999) tried to recognize nine activities, sitting, standing, lying, sitting and talking, sitting and operating a computer keyboard, walking, going up stairs, going down stairs, and cycling, using four accelerometers and some additional channels such as a microphone and an electrocardiogram. The authors in Lee and Masc (2002) created a system that uses a biaxial accelerometer, a gyroscope, and a digital compass to identify the user's location and activities, such as sitting, standing, walking on level ground, and going up and down a stairway. In Najafi et al. (2003), the authors utilized two accelerometers and one gyroscope on the chest to identify whether elderly persons were standing, sitting, walking, or lying down. In Parkka et al. (2006), the authors built a system that measures two accelerations using two accelerometers (one on the chest and the other on the wrist) and 16 different quantities with 20 additional sensors to recognize such activities as lying down, sitting, standing, walking, Nordic walking, running, rowing, and cycling. The authors in Subramanya, Raj, Bilmes, and Fox (2006) addressed similar activities by building a model using data from a triaxial accelerometer, two microphones, phototransistors, temperature and barometric pressure sensors, and GPS to distinguish between a stationary state, walking, jogging, driving a vehicle, and climbing up and down stairs. In Tapia et al. (2007), five accelerometers and a heart rate monitor were incorporated to automatically recognize activities with different intensities (lying down, standing, sitting, walking, running, etc.). The authors in Banos, Galvez, Damas, Pomares, and Rojas (2014) evaluated the significance of signal segmentation. They did experiments that used nine inertial sensors on different parts of the body while varying the sliding window sizes ranging 0.25–7 s. With supervised learning methods, they showed that the most accurate results are obtained for very small windows (0.25–0.5 s). The authors in Gao, Bourke, and Nelson (2014) compared single-sensor wearable systems to multi-sensor systems with five recognizers to identify six activities, including lying, sitting, standing, level walking, and climbing up/down. They concluded that the single-sensor systems did not beat the multi-sensor systems on the recognition accuracy regardless of a higher sampling rate, more complicated features, and more dedicated classifier. The authors in Guiry, van de Ven, Nelson, Warmerdam, and Riper (2014) compared a number of activity recognition algorithms, including their own classifier, to identify six key activities using two accelerometers, one in the trouser pocket, and the other in a chest strap.

Some other researchers have used a single accelerometer for activity recognition. The authors in Randell and Muller (2000) used a biaxial accelerometer to identify six activities, walking, running, sitting, walking upstairs, walking downstairs, and standing. The authors in Sekine, Tamura, Togawa, and Fukui (2000) and Sekine et al. (2002) attempted to distinguish the states of walking (level walking, walking upstairs, and walking down stairs) using a triaxial accelerometer attached to the waist. In Lee, Park, Hong, Lee, and Kim (2003), a triaxial accelerometer was placed on the vertebra of the subjects. The acceleration measurements were used to recognize five static and four dynamic activities. The authors in Allen, Ambikairajah, Lovell, and Celler (2006) concentrated on elderly people. Acceleration data were collected from six elderly subjects with a triaxial accelerometer placed on the waist to identify eight activities. In Lee et al. (2009), five activities, walking, running, sitting, standing, and lying down, were identified using a triaxial accelerometer attached to the left waist. The researchers in Long, Yin, and Aarts (2009) also used a triaxial accelerometer placed on the waist to distinguish five different activities.

The advent of mobile devices has initiated activity recognition studies using smartphones. The authors in Miluzzo et al. (2008) introduced the CenceMe application, which includes an activity classification engine that collects acceleration data from a Nokia N95 phone to distinguish sitting, standing, walking, and running activities. In Brezmes et al. (2009), a real-time classification system for some basic movements, including walking, climbing upstairs, climbing down stairs, sitting, standing, and falling, was developed using an accelerometer in a Nokia N95 phone. The authors in Yang (2009) also collected acceleration data from a Nokia N95 phone and examined whether several classifiers can distinguish six activities, sitting, standing, walking, running, driving, and cycling. The researchers in Kwapisz, Weiss, and Moore (2010) built an activity recognition system that uses accelerometers embedded in smartphones. The system can recognize six activities with a number of classification algorithms. In Anguita, et al. (2012), the authors proposed a multiclass HF-SVM model that adapted a support vector machine to address the limitations of mobile phones. They used an accelerometer and a gyroscope in a Samsung Galaxy S2 phone to show that their model can distinguish six basic activities. The authors in Cho, Kim, and Kim (2012) used a triaxial accelerometer, gyroscope, and magnetometer in a smartphone to extract the acceleration and orientation of the device. They showed that five activities, walking, going up stairs, going downstairs, running, and not moving, can be recognized through a linear discriminant analysis and use of a support vector machine. In Kwon, Heo, Kang, and Bae (2013), the authors introduced an adaptive boundary correction approach to enhance the particle swarm optimization. They examined whether their new method can identify three basic activities, sitting, standing, and walking. The authors in Chianga, Yanga, and Tub (2013) proposed a smartphone-based activity pattern recording system that recognizes daily activity patterns using four traditional classifiers. In Anjum and Ilyas (2013), the authors implemented a smartphone application that recognized seven activities (walking, running, climbing/descending stairs, driving, cycling, and inactive) using smartphone accelerometer and gyroscope. Their application used C4.5 decision tree for activity recognition because it showed the best results.

Since we have adopted unsupervised learning methods, it is reasonable to look up the previous studies that used unsupervised learning approaches for activity recognition. There have also been some studies considering unsupervised learning methods. The authors in Wyatt, Philipose, and Choudhury (2005) believed that activities can be differentiated by the objects used. Given the names of these objects, they built models of activities from the Web, and showed that the models can distinguish 26 activities. In Huynh (2008), the authors introduced an unsupervised learning method that handles multiple time scales, based on multiple eigenspaces. They tried to recognize six different activities with four inertial sensors. The researchers in Li and Dustdar (2011) analyzed the feasibility of incorporating unsupervised learning methods in activity recognition. They proposed a hybrid process combining activity classification with subspace clustering to handle the curse of dimensionality. In Vandewiele and Motamed (2011), a smart home environment with a network of sensors that produce redundant information was assumed. The authors stated that an unsupervised approach is useful for activity recognition in such an environment. To our best knowledge, the most similar study to our work is Trabelsi, Mohammed, Chamroukhi, Oukhellou, and Amirat (2013). The authors in Trabelsi et al. (2013) proposed an unsupervised method for activity recognition with three accelerometers, attached at the chest, right thigh, and left ankle. They took the Hidden Markov Model Regression and showed that their method outperformed other unsupervised learning methods and was comparable to supervised learning methods. Their algorithm, however, assumed that the number of activities is known.

3. Methodology

3.1. Experimental setup

The experiments exploit a single smartphone kept in a pants pocket to collect the sensor data during five common physical activities, each of which was performed for ten minutes. Since we want to examine the effectiveness of unsupervised learning for human activity recognition, we selected the five most common activities during a daily routine, namely, walking, running, sitting, standing, and lying down. The sensor data were collected from a Galaxy Nexus smartphone, as it has an embedded accelerometer and gyroscope for triaxial linear acceleration and angular velocity, respectively, at a 50 Hz sampling rate. A simple low-pass filter was applied to eliminate noise from the sensor data. Each record was labeled with the corresponding activity to evaluate the quality of the clustering results.

3.2. Feature extraction

The features were computed on fixed-width windows of 64 sensor datasets with a 50% overlap between consecutive windows. We chose a smaller window length than in other researches because a window with a large size does not seem to represent the "moment" in which an activity is performed. For each window, some statistical properties, including average and standard deviations, were used as features in both the time and frequency domains. To retrieve the frequency components, a Fast Fourier Transform (Cooley & Tukey, 1965) was applied. Therefore, the sample space is represented by 24 features: 12 features from time domain (six from acceleration and six from angular velocities) and 12 features from frequency domain (six from acceleration and six from angular velocities). After feature extraction, the feature values were normalized to within [0,1].

3.3. Activities

In these experiments, we choose five activities: walking, running, sitting, standing, and lying down, which are performed frequently by most people in their daily lives. It is therefore valuable to see how different the patterns of acceleration and angular velocity are for these five activities.

Fig. 1 plots the acceleration for all three axes and each of the five activities. Since walking and running are energetic activities, the *x*, *y*, and *z* values vary significantly. In particular, the sensor values from running vary more considerably than those of walking, because running is a more vibrant activity. Sitting, standing, and lying down, on the other hand, are dormant activities, and thus their *x*, *y*, and *z* values are almost constant. One interesting point is that for walking, running, and standing, the *y* values have the lowest accelerations, and for sitting and lying down, the *z* values have the largest accelerations. This is because the force of gravity influences the entire acceleration in the direction of the center of the Earth. For walking, running, and standing, the direction of gravity mainly corresponds to the *y* axis, whereas for sitting and lying, it mainly corresponds to the *z* axis.

Fig. 2 plots the angular velocity for all three axes and each of the five activities. For walking and running, the *x*, *y*, and *z* values significantly change, as shown in the acceleration. For sitting, standing, and lying down, on the other hand, the *x*, *y*, and *z* values are much smaller (<0.25 rad/s). Note that the *x*, *y*, and *z* values for standing vary considerably because standing is a relatively unstable activity compared with sitting and lying. However, the angular velocity does not seem to be useful for distinguishing these activities since they show a similar pattern and their magnitudes are very small.

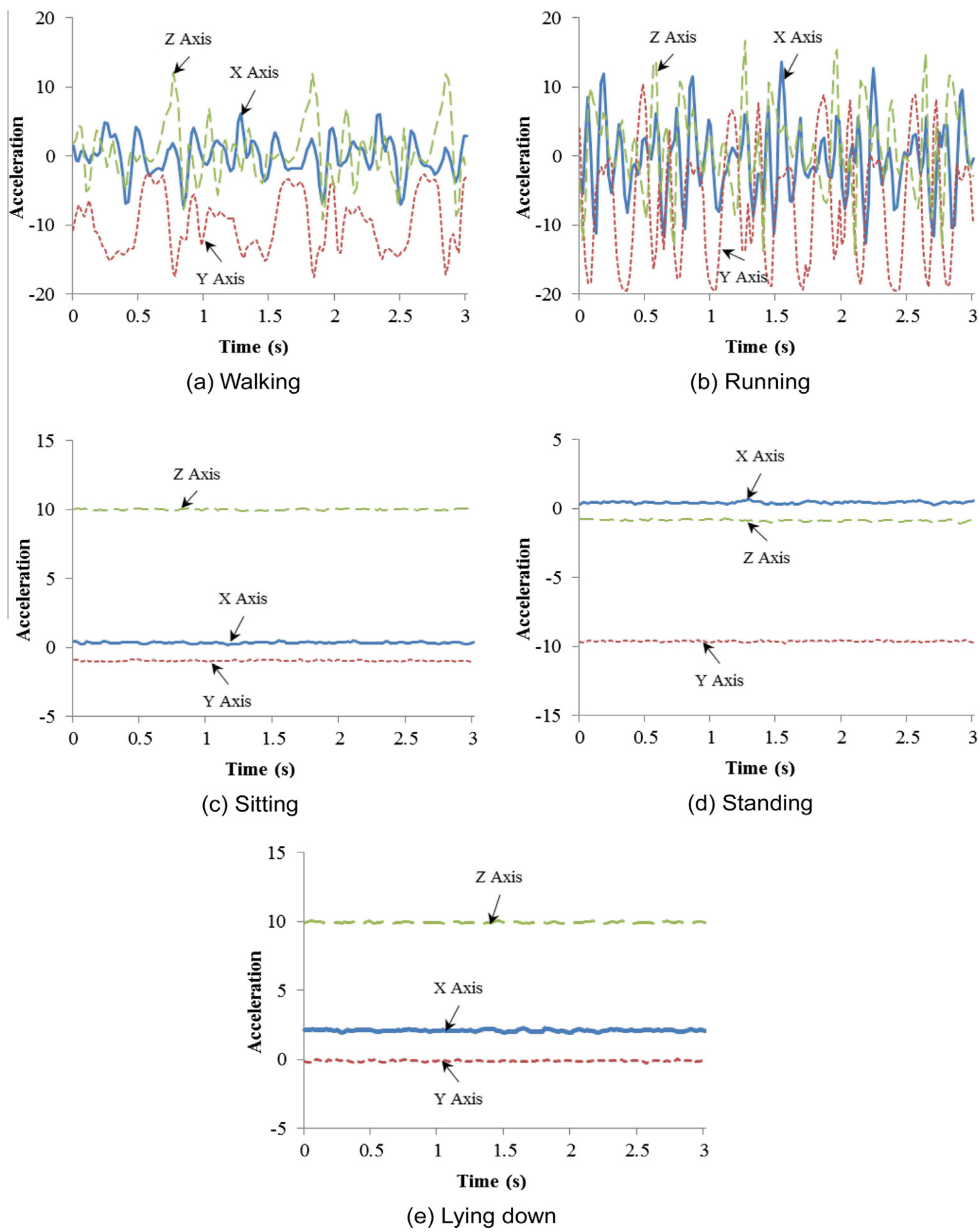

Fig. 1. Acceleration for the five activities.

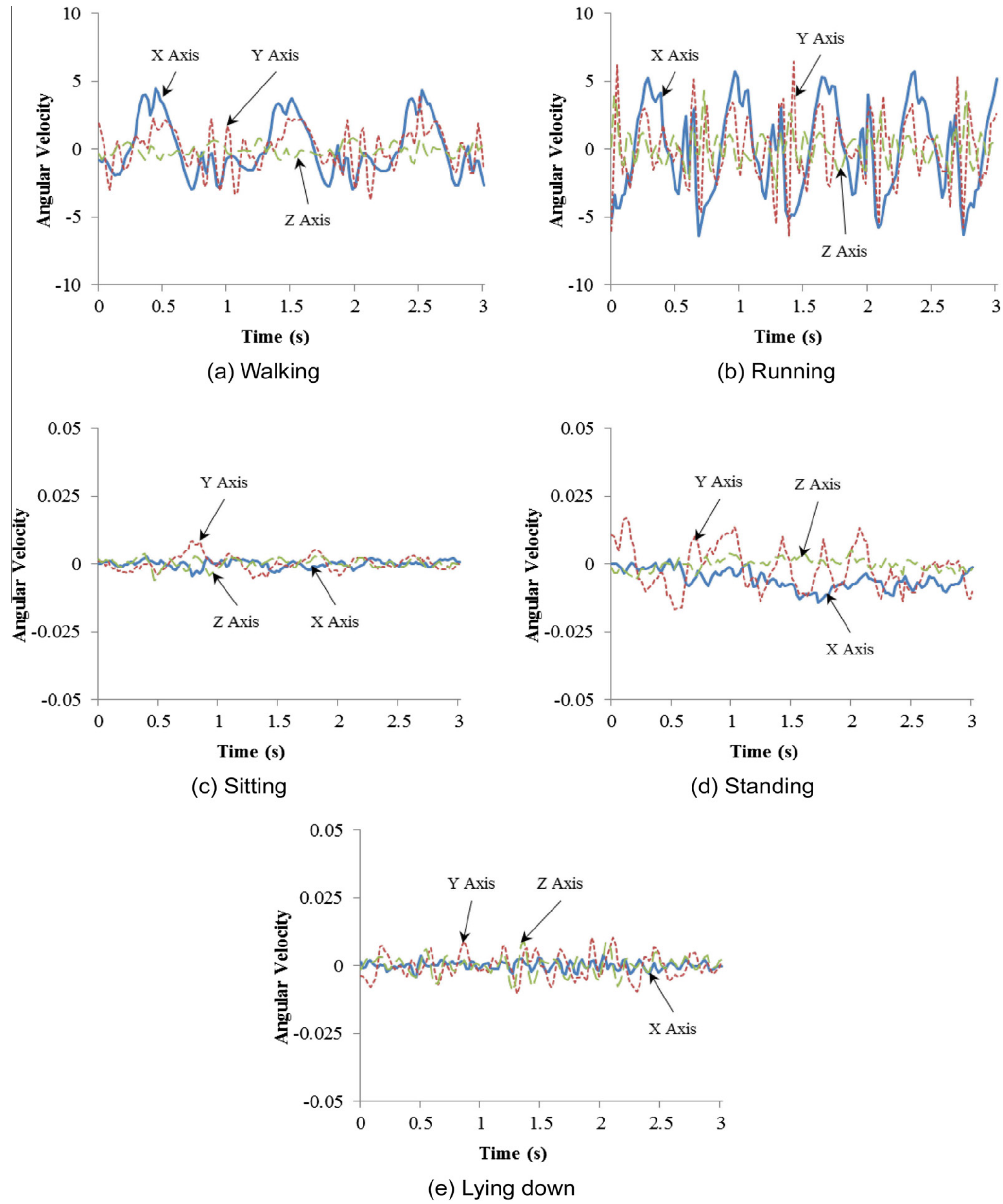

Fig. 2. Angular velocity signals for the five activities.

Since each activity has its own properties in terms of acceleration and angular velocity, the features should represent the properties definitively and clearly. We believe that the average and standard deviations can distinctly express the properties of the activities. For instance, sitting, standing, and lying down have their own intervals of *x*, *y*, and *z* values for acceleration. In this case, the average value will identify each interval without difficulty. Another example is a comparison of running to walking. As shown in Figs. 1 and 2, running shows larger signal amplitudes than walking. Since the standard deviation indicates how distributed the values are from the average, it can be a key feature to distinguish running from walking.

4. Experiments

To evaluate the performance of unsupervised learning methods for activity recognition, two kinds of experiments were carried out based on the knowledge of the number of activities (k). First, we verified whether unsupervised learning is a reasonable way for activity recognition when k is known, and second, showed whether unsupervised learning can still be useful when k is unknown, indicating its potential to distinguish an arbitrary number of activities.

4.1. Evaluation measures

To evaluate the quality of the clustering algorithms, we used two measures: accuracy and normalized mutual information (*NMI*) (Strehl & Ghosh, 2002). Both measures require a labeled dataset. The accuracy is computed using a confusion matrix. The *NMI* is a measure ranging from 0 to 1, and *NMI* = 1 when each cluster corresponds exactly to a class. The *NMI* is computed as

$$NMI = \frac{\sum_{i=1}^{r}\sum_{j=1}^{s} n_{i,j}\log\left(\frac{n \cdot n_{i,j}}{n_i \cdot n_j}\right)}{\sqrt{\sum_{i=1}^{r} n_i \log\frac{n_i}{n}\sum_{j=1}^{s} n_j \log\frac{n_j}{n}}}$$

where r is the number of clusters, s is the number of classes, $n_{i,j}$ is the number of instances in cluster i and class j, n_i is the number

Table 1
Accuracy and NMI when $k = 5$.

	Accuracy	NMI
k-Means	0.7198	0.8670
GMM	1.0000	1.0000
HIER	0.7998	0.9092

of instances in cluster i, n_j is the number of instances in class j, and n is the number of instances.

4.2. When k is known

In this experiment, we chose five activities ($k = 5$): walking, running, sitting, standing, and lying down. We chose three clustering algorithms, k-means clustering, mixture of Gaussian (GMM), and average-linkage hierarchical agglomerative clustering (HIER), from the WEKA data mining suite (Hall et al., 2009) to see whether they can distinguish these activities. Since they are different types of clustering approaches, we can compare the methods from one another.

The experiment results are presented in Table 1. This table specifies the overall accuracy and *NMI* for each clustering algorithm. GMM shows a perfect recognition of all activities. In contrast, k-means and HIER show lower accuracy because they fail to distinguish walking from running, although they successfully identify sitting, standing, and lying. Actually, the ranges of the signals while walking and running are quite overlapped. Therefore, the average features hinder a precise recognition of walking and running for these two algorithms. GMM, nonetheless, succeeds in identifying all activities, because it is based on a robust Gaussian distribution model.

4.3. When k is unknown

In reality, individuals may perform a different number of activities. It is therefore necessary to check the usefulness of unsupervised learning methods even without the number of clusters. We chose four clustering algorithms: k-means clustering, mixture of Gaussian (GMM), average-linkage hierarchical agglomerative clustering (HIER), and DBSCAN (Ester, Kriegel, Sander, & Xu, 1996). For k-means, GMM, and HIER, we chose appropriate values of k based on local maxima of the Caliński–Harabasz index (CH) (Caliński & Harabasz, 1974); given a set $X = \{x_1, x_2, \ldots, x_n\}$ of n instances having a total center c, and a partition of these instances into k mutually exclusive clusters, each of which has n_i instances and center c_i, the value CH can be computed as

$$CH = \frac{SSB/(k-1)}{SSW/(n-k)}$$

where

$$SSB = \sum_{i=1}^{k} n_i \| c_i - c \|^2, \quad SSW = \sum_{i=1}^{k} \sum_{j=1}^{n} \| x_j - c_i \|^2$$

Since the number of clusters may be different from that of activities, the evaluation of accuracy should be adjusted. For each activity, we pick as a correct one a cluster that has the largest portion of the true cluster. If two or more activities share the cluster, then the cluster is assigned to the activity that has more points of it, and for the other activities, we pick the second largest cluster. Note that *NMI* does not

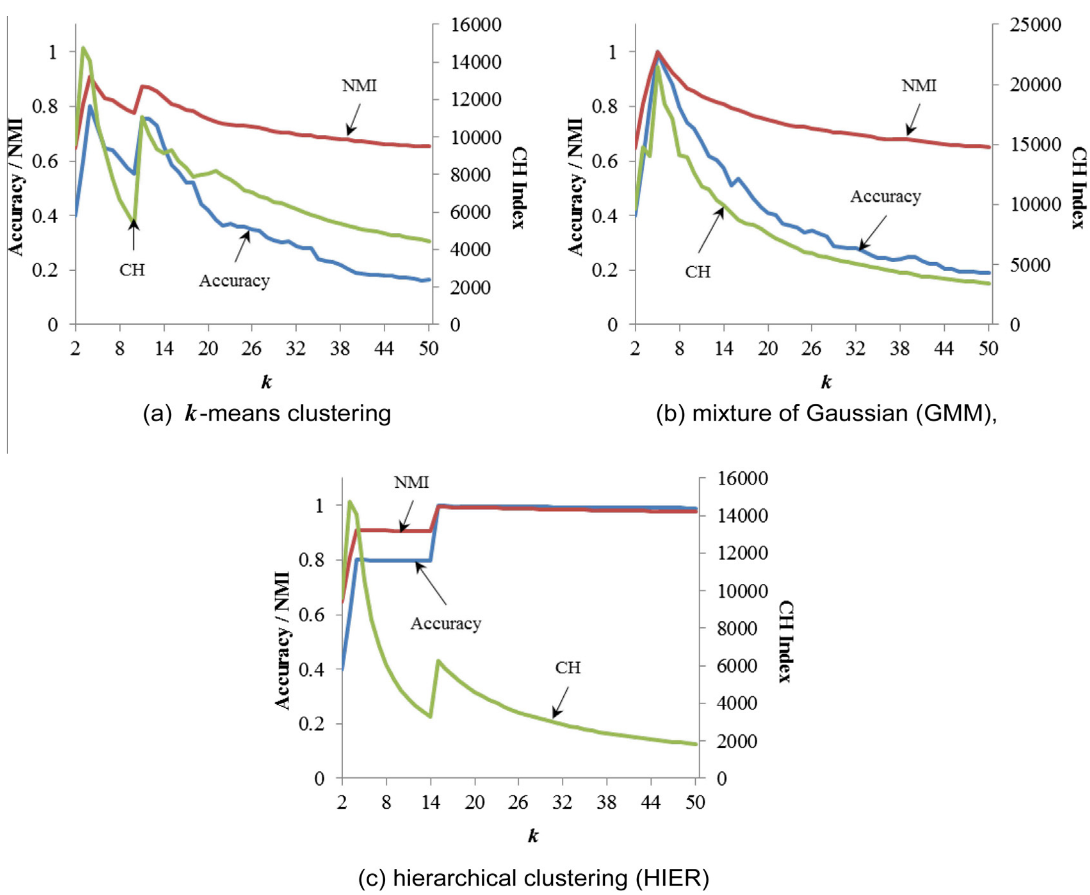

Fig. 3. Experiment results for clustering algorithms.

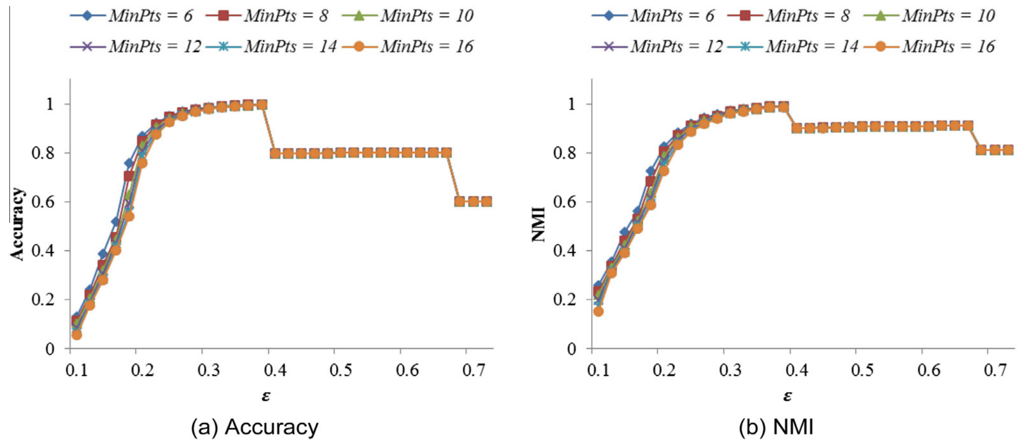

Fig. 4. Experiment results for DBSCAN.

require the condition that the number of activities is the same as that of clusters.

Fig. 3 shows the experiment results of k-means, GMM, and HIER. The k-means algorithm shows a relatively lower accuracy than others. The accuracy never exceeds 0.8 for every k. In addition, k-means clustering tends to find spherical clusters of a similar size. The sensor data, however, are not of a spherical shape, and are noisy. Hence, k-means clustering has difficulty showing a precise separation. The accuracy and other measures significantly descend when $k = 10$ because the main clusters begin to merge. More specifically, the two clusters corresponding primarily to walking and running merge at $k = 10$.

While GMM shows the exact recognition when k is known ($k = 5$), the accuracy of the other measures decrease dramatically as k increases. The reason for this can be found in the basic GMM model. GMM is based on a robust Gaussian distribution model. GMM can separate the activities when $k = 5$, but it divides normal clusters into several clusters when $k > 5$. For instance, a walking cluster is divided into two smaller clusters when $k = 6$, and a sitting cluster is also separated into two smaller clusters when $k = 7$. Therefore, GMM needs to find an appropriate k with intrinsic measures such as the CH index.

For HIER, the clusters are distinctively recognizable when k is large. Since HIER first merges closer points, the five main clusters are quickly generated. HIER therefore shows high accuracy for a large k. The accuracy and other measures are suddenly dropped when $k = 14$ because HIER begins to merge the main clusters. More specifically, the two clusters corresponding primarily to walking and running are merged at $k = 14$. As a result, HIER is inaccurate when $k = 5$.

Fig. 4 shows the experiment results of DBSCAN according to the two parameters, ε and *MinPts*. Since DBSCAN does not require the number of clusters, it is a suitable algorithm for this case. The graphs show that the accuracy is larger than 0.9 with $0.24 \leqslant \varepsilon < 0.4$, and the measures become a little smaller with a larger *MinPts*. Using the appropriate parameters, DBSCAN can recognize the activities precisely without k. Note that both the accuracy and *NMI* significantly decrease when $\varepsilon = 0.4$ since the two clusters corresponding primarily to walking and running are merged into a single cluster. Similarly the measures decrease when $\varepsilon = 0.68$ owing to the combination of the two clusters corresponding to sitting and lying down.

5. Conclusions

Although most of previous studies have adopted supervised learning methods, they have suffered from some difficulties, for instance, in generating a huge number of training data by hand, in making the training dataset flawless, and in extending the number of activities to be recognized. Hence, it is necessary to seek new approaches to resolve those problems. In this paper, we proposed unsupervised learning algorithms, instead of supervised learning ones, for human activity recognition using smartphone sensors, even when the number of activities is unknown. We have evaluated the effectiveness of unsupervised learning methods for activity recognition with sensor data during five common activities: walking, running, sitting, standing, and lying down. Our results show that when the number of activities k is known, the mixture of Gaussian method, a solid model-based algorithm, exactly distinguished those activities. Even though k is unknown, unsupervised learning methods can achieve the high recognition accuracy by selecting k based on CH index or by selecting appropriate parameters. Especially, hierarchical clustering or DBSCAN attained above 90% accuracy by selecting k based on CH index, or by selecting appropriate values for ε and *MinPts*. Through our experiments, we believe that our approach can be used to reveal a new way of activity recognition.

As a future work, we plan to develop a sophisticated algorithm for activity recognition with unsupervised approaches. In the experiments, we found some cases that the results become more accurate when the number of activities k exceeds the actual value. This is because when k equals the actual value, main clusters that primarily correspond to each activity are merged, although small clusters are alive. For instance, the accuracy of hierarchical clustering is steeply dropped at $k = 14$ because two main clusters, corresponding primarily to walking and running, are merged while ten small clusters are alive. Since the small clusters are actually outliers, the sophisticated algorithm should detect and handle the outliers to improve the recognition accuracies.

We also plan to develop a novel method for recognizing more complicated activities, such as working, resting, eating, and sleeping, based on the routines of different individuals. Thus far, activity recognition studies have concentrated on basic activities, such as walking, sitting, or ascending stairs. Smartphone sensors, however, provide little information about the users. To surpass previous studies, we are concerning spatio-temporal data, as well as unsupervised learning approaches. We believe that the time and location data will provide additional semantics to specify basic activities that can be recognized by sensors. For instance, sitting may imply the user is "working" at the office on weekdays from 9 to 6, or "watching TV" at home in the evening. To separate the basic activities into more specific activities, it will be necessary to apply clustering methods with respect to both sensor values and spatio-temporal data. We also believe that unsupervised learning methods are essentially necessary since an extensive number of activities can be targeted. Hence, it is believed that through unsupervised learning methods and spatio-temporal data, activity recognition methods can be enhanced to identify more specific activities.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIP) (No. 2010-0028631).

References

Allen, F. R., Ambikairajah, E., Lovell, N. H., & Celler, B. G. (2006). Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models. *Physiological Measurement, 27*(10), 935–951.

Aminian, K., Robert, Ph., Buchser, E. E., Rutschmann, B., Hayoz, D., & Depairon, M. (1999). Physical activity monitoring based on accelerometry: Validation and comparison with video observation. *Medical & Biological Engineering & Computing, 37*(3), 304–308.

Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2012). Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *Proceedings of the 4th international conference on Ambient assisted living and home care* (pp. 216–223).

Anjum, A., & Ilyas, M. U. (2013). Activity recognition using smartphone sensors. In *First workshop on people centric sensing and communications*.

Banos, O., Damas, M., Pomares, H., Prieto, A., & Rojas, I. (2012). Daily living activity recognition based on statistical feature quality group selection. *Expert Systems with Applications: An International Journal, 39*(9), 8013–8021.

Banos, O., Galvez, J.-M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. *Sensors, 14*(4), 6474–6499.

Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. *Lecture Notes in Computer Science, 3001*, 1–17.

Brezmes, T., Gorricho, J.-L., Cotrina, J. (2009). Activity recognition from accelerometer data on a mobile phone. In *Proceedings of the 10th international work-conference on artificial neural networks: PART II: Distributed computing, artificial intelligence, bioinformatics, soft computing, and ambient assisted, living* (pp. 796–799).

Bussmann, J. B. J., Martens, W. L. J., Tulen, J. H. M., Schasfoort, F. C., van den Berg-Emons, H. J. G., & Stam, H. J. (2001). Measuring daily behavior using ambulatory accelerometry: The activity monitor. *Behavior Research Methods, Instruments, & Computers, 33*(3), 349–356.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics, 3*(1), 1–27.

Chianga, J.-H., Yanga, P.-C., & Tub, H. (2013). Pattern analysis in daily physical activity data for personal health management. *Pervasive and Mobile Computing*. http://dx.doi.org/10.1016/j.pmcj.2013.12.003. in press.

Cho, J., Kim, J., & Kim, T. (2012). Smart phone-based human activity classification and energy expenditure generation in building environments. In *Proceedings of the 7th international symposium on sustainable healthy buildings* (pp. 97–105).

Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation, 19*(90), 297–301.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the second international conference on knowledge discovery and data mining* (pp. 226–231).

Foerster, F., & Fahrenberg, J. (2000). Motion pattern and posture: Correctly assessed by calibrated accelerometers. *Behavior Research Methods, Instruments, & Computers, 32*(3), 450–457.

Foerster, F., Smeja, M., & Fahrenberg, J. (1999). Detection of posture and motion by accelerometry: A validation study in ambulatory monitoring. *Computers in Human Behavior, 15*(1), 571–583.

Gao, L., Bourke, A. K., & Nelson, J. (2014). Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems. *Medical Engineering & Physics*. http://dx.doi.org/10.1016/j.medengphy.2014.02.012. in press.

Guiry, J. J., van de Ven, P., Nelson, J., Warmerdam, L., & Riper, N. (2014). Activity recognition with smartphone support. *Medical Engineering & Physics*. http://dx.doi.org/10.1016/j.medengphy.2014.02.009. in press.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter, 11*(1), 10–18.

Huynh, D. T. G. (2008). *Human activity recognition with wearable sensors* (Ph.D. thesis). Germany: TU Darmstadt.

Krishnan, N. C., & Panchanathan, S. (2008). Analysis of low resolution accelerometer data for continuous human activity recognition. In *Proceedings of the international conference on acoustic speech and signal processing* (pp. 3337–3340).

Krishnan, N. C., Colbry, D., Juillard, C., & Panchanathan, S. (2008). Real time human activity recognition using tri-axial accelerometers. In *Sensors signals and information processing workshop*.

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2010). Activity recognition using cell phone accelerometers. *SIGKDD Explorations Newsletter, 12*(2), 74–82.

Kwon, Y., Heo, S., Kang, K., & Bae, C. (2013). Particle swarm optimization using adaptive boundary correction for human activity recognition. In *Proceedings of the international conference on internet* (pp. 147–155).

Laerhoven, K. V., & Cakmakci, O. (2000). What shall we teach our pants? In *Proceedings of the 4th IEEE international symposium on wearable computers* (pp. 77–83).

Lee, S. H., Park, H. D., Hong, S. Y., Lee, K. J., & Kim, Y. H. (2003). A study on the activity classification using a triaxial accelerometer. In *Proceedings of the 25th annual international conference of the IEEE engineering in medicine and biology society* (pp. 2941–2943).

Lee, M., Kim, J., Kim, K., Lee, I., Jee, S., & Yoo, S. (2009). Physical activity recognition using a single tri-axis accelerometer. In *Proceedings of the world congress on engineering and computer science* (pp. 14–17).

Lee, S., & Masc, K. (2002). Activity and location recognition using wearable sensors. *IEEE Pervasive Computing, 1*(3), 24–32.

Li, F., & Dustdar, S. (2011). Incorporating unsupervised learning in activity recognition. In *AAAI workshops at the 25th AAAI conference on artificial intelligence*.

Long, X., Yin, B., & Aarts, R. M. (2009). Single-accelerometer-based daily physical activity classification. In *Proceedings of 31st annual international conference of the IEEE engineering in medicine and biology society* (pp. 6107–6110).

Mannini, A., & Sabatini, A. M. (2010). Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors, 10*(2), 1154–1175.

Mantyjarvi, J., Himberg, J., & Seppanen, T. (2001). Recognizing human motion with multiple acceleration sensors. In *Proceedings of the international conference on systems, man, and cybernetics* (pp. 747–752).

Mathie, M. J., Coster, A. C. F., Lovell, N. H., & Celler, B. G. (2004). Accelerometry: Providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement, 25*(2). R1–20.

Miluzzo, E., Lane, N. D., Fodor, K., Peterson, R., Lu, H., & Musolesi, M., et al. (2008). Sensing meets mobile social networks: The design, implementation and evaluation of the CenceMe application. In *Proceedings of the 6th ACM conference on embedded network sensor systems* (pp. 337–350).

Najafi, B., Aminian, K., Paraschiv-Ionescu, A., Loew, F., Büla, C. J., & Robert, P. (2003). Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly. *IEEE Transactions on Biomedical Engineering, 50*(6), 711–723.

Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., & Korhonen, I. (2006). Activity classification using realistic data from wearable sensors. *IEEE Transactions on Information Technology in Biomedicine, 10*(1), 119–128.

Randell, C., & Muller, H. (2000). Context awareness by analysing accelerometer data. In *Proceedings of the fourth international symposium on wearable computers* (pp. 175–176).

Sekine, M., Tamura, T., Akay, M., Fujimoto, T., Togawa, T., & Fukui, Y. (2002). Discrimination of walking patterns using wavelet-based fractal analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 10*(3), 188–196.

Sekine, M., Tamura, T., Togawa, T., & Fukui, Y. (2000). Classification of waist-acceleration signals in a continuous walking record. *Medical Engineering & Physics, 22*, 285–291.

Strehl, A., & Ghosh, J. (2002). Cluster ensembles a knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research, 3*, 583–617.

Subramanya, A., Raj, A., Bilmes, J., & Fox, D. (2006). Recognizing activities and spatial context using wearable sensors. In *Proceedings of the 22nd conference on uncertainty in, artificial intelligence* (pp. 494–502).

Tapia, E. M., Intille, S. S., Haskell, W., Larson, K., Wright, J., & King, A., et al. (2007). Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In *Proceedings of the 11th international symposium on wearable computers* (pp. 37–40).

Trabelsi, D., Mohammed, S., Chamroukhi, F., Oukhellou, L., & Amirat, Y. (2013). An unsupervised approach for automatic activity recognition based on hidden Markov model regression. *IEEE Transactions on Automation Science and Engineering, 10*(3), 829–835.

Uddin, Md. Z., Thang, N. D., Kim, J., & Kim, T. (2011). Human activity recognition using body joint-angle features and hidden Markov model. *ETRI Journal, 33*(4), 569–579.

Vandewiele, F., & Motamed, C. (2011). An unsupervised learning method for human activity recognition based on a temporal qualitative model. In *International workshop on behaviour analysis and video understanding*.

Veltink, P. H., Bussmann, H. B. J., de Vries, W., Martens, W. L. J., & van Lummel, R. C. (1996). Detection of static and dynamic activities using uniaxial accelerometers. *IEEE Transactions on Rehabilitation Engineering, 4*(4), 375–385.

Wyatt, D., Philipose, M., & Choudhury, T. (2005). Unsupervised activity recognition using automatically mined common sense. In *Proceedings of the 20th national conference on artificial intelligence* (pp. 21–27).

Yang, J. (2009). Toward physical activity diary: Motion recognition using simple acceleration features with mobile phones. In *Proceedings of the 1st international workshop on interactive multimedia for consumer electronics* (pp. 1–10).

Zhang, L., Liu, T., Zhu, S., & Zhu, Z. (2012). Human activity recognition based on triaxial accelerometer. In *Proceedings of the 7th international conference on computer sciences and convergence information technology* (pp. 261–266).