

Layered hidden Markov models to recognize activity with built-in sensors on Android smartphone

Young-Seol Lee¹ · Sung-Bae Cho¹

Received: 1 July 2015 / Accepted: 6 April 2016 / Published online: 20 April 2016
© Springer-Verlag London 2016

Abstract Recently, activity recognition using built-in sensors in a mobile phone becomes an important issue. It can help us to provide context-aware services: health care, suitable content recommendation for a user's activity, and user adaptive interface. This paper proposes a layered hidden Markov model (HMM) to recognize both short-term activity and long-term activity in real time. The first layer of HMMs detects short, primitive activities with acceleration, magnetic field, and orientation data, while the second layer exploits the inference of the previous layer and other sensor values to recognize goal-oriented activities of longer time period. Experimental results demonstrate the superior performance of the proposed method over the alternatives in classifying long-term activities as well as short-term activities. The performance improvement is up to 10 % in the experiments, depending on the models compared.

Keywords Activity recognition · Mobile sensors · Layered hidden Markov models · Context-aware computing

1 Introduction

Recent smartphone not only serves as the communication device, but also provides a rich set of embedded sensors, such as an accelerometer, digital compass, gyroscope, GPS

receiver, microphone, and camera. As mobile phones contain increasingly various sensors, sensing information is used for a wide variety of domains such as healthcare, social networks, safety, environmental monitoring, and transportation [1].

Because current mobile devices include built-in sensors that provide real-time data gathered from the surroundings of the devices, it is possible to infer its user's current status [2]. Activity is useful context to provide personalized services like health monitoring, user adaptive interface, and content recommendation. Because of the importance of the activity, many researchers investigated activity classification using smartphones. In most cases, machine learning techniques are used to classify activities on smartphones [3].

Recognizing human activities from data sequences is a challenging issue [4]. In order to implement practical activity-aware systems, the underlying recognition module has to handle the real world's noisy data and complexities [5]. Furthermore, the systems must consider some important constraints such as relatively insufficient memory capacity, lower CPU power (data-processing speed), and limited battery lives [6]. A lot of classification methods have been investigated. Some research incorporated the idea of simple heuristic classifiers [7]. On the other hand, the other studies used more generic methods from the machine learning techniques including decision trees, Bayesian networks [6, 8–10], support vector machines [11], neural networks [12], and Markov chains [13–15].

Probabilistic models such as Bayesian network and hidden Markov models (HMMs) are appropriate for dealing with vagueness and uncertainty of data in real life for context-aware services. However, it is difficult to apply them to mobile devices because it requires a lot of memory and CPU time. Hierarchical probabilistic model is the

✉ Sung-Bae Cho
sbcho@yonsei.ac.kr

Young-Seol Lee
tiras@sclab.yonsei.ac.kr

¹ Department of Computer Science, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, South Korea

combination of some separate classification modules. The modular structure is suitable to overcome the limitations in mobile environment. This paper proposes an activity-aware system using a hierarchical probabilistic model, layered HMMs, to recognize a user's activities in real time. The activity recognition system is developed on an Android smartphone, and has two layers of HMMs to deal with both short-term and long-term activities. To evaluate the usefulness of the system, we collected mobile sensing data from graduate students and compared the accuracy of the system with the alternative methods. Experimental results demonstrate the superior performance of the proposed method over the alternatives in classifying both long-term and short-term activities. The proposed method showed better performance up to 10 % than the alternatives.

2 Related works

2.1 Activity recognition with mobile sensors

There are many studies to recognize a user's activities with mobile sensors. Kwapisz et al. proposed activity recognition system using an accelerometer on Android phone [16]. Longstaff et al. presented a semi-supervised learning method to train a classifier for monitoring patients [3]. Maguire et al. classified human activities using k nearest neighbor and decision tree [17]. They extracted some features such as mean, standard deviation, energy, and correlation from acceleration and heart rate data. Györfi et al. developed a system to recognize a person's activities from acceleration data using a feed-forward neural network [18]. Song et al. proposed an activity recognition system for the elderly using a wearable sensor module including an accelerometer [12]. They used multi-layer perceptron to recognize activities of daily life. Zappi et al. used HMMs to recognize a user's activities from acceleration data [13]. Berchtold et al. proposed a fuzzy inference based classifier to recognize a user's activities [19]. Recently, many researchers attempt to incorporate various sensors with different places to attack the problem [20–22]. Table 1 summarizes the studies to recognize human activities using mobile sensors.

Most studies in activity recognition have been using discriminative approaches such as support vector machines (SVM) and decision trees, ignoring the time-series characteristics of sensor signals. Although these models are easy to implement, they require the use of a rich set of features, which in turn increases computational costs, in exchange for algorithm simplicity. To take advantage of this inherent characteristics of sensor signals, we propose an activity recognition system using layered HMMs to deal with multi-modal sensors efficiently.

2.2 Hierarchical probabilistic models

There are many probabilistic models, including hierarchical structure, such as hierarchical Bayesian network (HBN), hierarchical dynamic Bayesian network (HDBN), hierarchical hidden Markov model (HHMM), and layered HMMs (LHMM). The model with a hierarchical structure is an effective solution to the problems which can be divided into smaller units. For instance, to recognize human activities, the model of human is divided into smaller parts of body, head, and arms, respectively. The hierarchical models can improve accuracy and speed of activity recognition.

Min and Cho [11] proposed a method to recognize activities by combining multiple classifiers such as support vector machine (SVM) and Bayesian network (BN). Park and Aggarwal presented a method for the recognition of two-person interactions using a hierarchical Bayesian network. They divided the overall body pose into separate body parts. The pose of body parts is modeled at the low level of the BN, and the overall pose is estimated at the high level of the BN [8]. Mengistu et al. used hierarchical HMM for spoken dialogue system [27]. Wang and Tung recognized dynamic human gestures using dynamic Bayesian networks which represented multi-level probabilistic process like hierarchical HMM [9]. Du et al. proposed a dynamic Bayesian network based method to recognize activities. They divided the features for human activities into global and local features, and built a hierarchical DBN model to combine the two features [10].

Table 2 summarizes examples of hierarchical models. In most cases, hierarchical model is used to recognize an activity that consists of smaller behaviors or time-series patterns.

3 Activity recognition using layered HMM

The proposed structure consists of two steps to analyze sensor data and recognize a user's behaviors. First, after sensor data are collected from sensors on a smartphone, the data are transferred to HMMs and preprocessing units to classify a user's short-term activities. In the second phase, the second-level HMMs are used to recognize a user's long-term activities from the temporal pattern of the short-term activities and the other features. Figure 1 shows the process of the entire system.

3.1 Feature selection for HMM state recognition

In many applications of activity recognition on a mobile device, the problem of high dimensionality of data appears. Since high dimensional data often require a large amount

Table 1 Activity recognition using accelerometers

Author	Year	Classifier	Sensor	Contribution
Gupta and Dallas [20]	2014	kNN, NB	Accelerometer	Recognize activities from sensors placed in subject's waist
Pei et al. [21]	2013	SVM	Various sensors	Construct context pyramid
Travelsi et al. [22]	2013	HMM	Accelerometer	Unsuervised approach based on HMM regression
Kwapisz et al. [16]	2010	ANN, J48	Accelerometer	Need for a little data size for recognition, fast recognition
Longstaff et al. [3]	2010	NB, DT	GPS receiver, Accelerometer	Use of co-learning, high accuracy
Berchtold et al. [19]	2010	Fuzzy logic	Accelerometer	Activity recognition service architecture on mobile phones for common users
Maguire et al. [17]	2009	kNN, J48	Accelerometer, Heart-beat rate	Fast recognition
Gyorbiri et al. [18]	2009	C4.5, NN	Accelerometer	Real time recognition, efficient battery usage
Bieber et al. [7]	2009	Heuristics (rule)	Accelerometer	Identification of physical activity in everyday life on a standard mobile phone
Lu et al. [23]	2009	Microphone	Decision tree, Markov model	Hierarchical architecture for recognizing a set of sound events
Song et al. [12]	2008	ANNs	Accelerometer	Implementation on mobile environment
Zappi et al. [13]	2008	HMM	Accelerometer	Dynamic sensor selection
Choudhury et al. [5]	2008	Accelerometer, barometer, compass, humidity, temperature	Probabilistic model	Using Mobile Sensing Platform (MSP) for activity recognition
Reddy et al. [24]	2008	GPS receiver, accelerometer	Decision tree, discrete HMM	Convenient system by not having strict position/orientation requirements
Yang et al. [25]	2007	Neuro-fuzzy	Accelerometer	Application of neuro-fuzzy
Ganti et al. [14]	2006	HMM	GPS receiver, Accelerometer	Implementation of middleware

Table 2 Related works using hierarchical model

Author	Year	Classifier	Sensor	Contribution
Min et al. [11]	2011	Bayesian network and SVMs	Accelerometer, wearable sensors	Activity recognition using classifier fusion with multi-modal sensors
Hwang et al. [6]	2009	Modular Bayesian network	GPS receiver, weather, call history, etc.	Modularization of probabilistic models
Aarno et al. [26]	2009	Layered hidden Markov models	JR3 force-torque sensor	Task segmentation and modeling using LHMM
Mengistu et al. [27]	2008	HHMM	Text data	Understanding underlying semantics of words and sentences
Wang et al. [9]	2007	HDBN	Video camera	Gesture recognition using HDBN
Du et al. [10]	2006	HDBN	Video camera	Hierarchical durational-state DBN
Park et al. [8]	2004	HDBN	Video camera	Modularization using HDBN
Oliver et al. [28]	2004	Layered hidden Markov models	Microphones, camera, keyboard, mouse	Real-time, multimodal approach to human activity recognition in an office environment

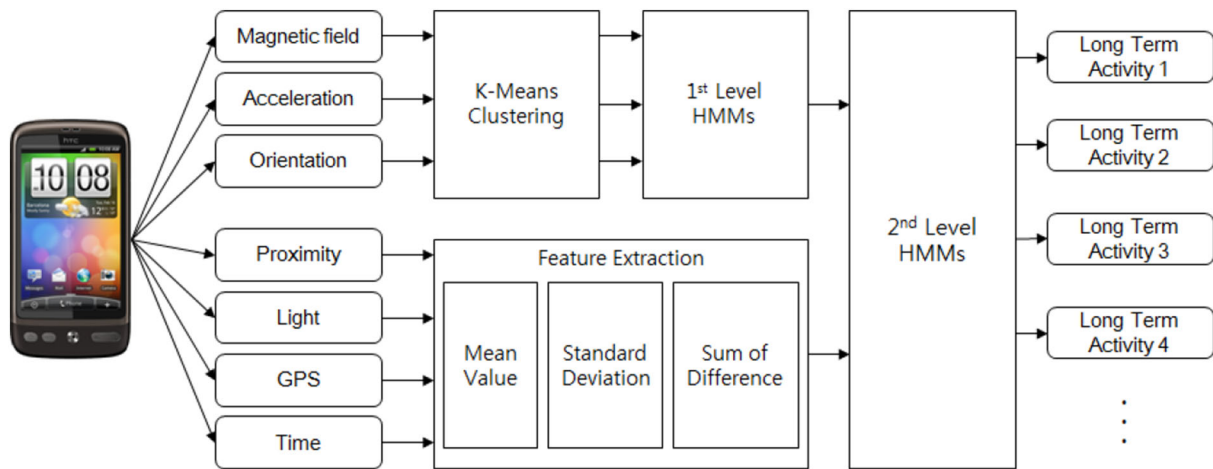


Fig. 1 Activity recognition system overview

of memory and CPU time to analyze, dimensionality reduction is essential to recognize behaviors in mobile environment. In order to reduce the dimensionality of raw data, most classifiers extract features from the data. However, it is difficult to extract good features to reduce dimensionality without any degradation to the performance for activity recognition. Usually, the feature extraction involves a function that measures the capability of the feature set to discriminate the classes [29].

This section presents the feature selection for the activity recognition system. Firstly, the window size for collecting the raw sample is reviewed to make into features. Then, the information gain value of each feature is calculated as a measure for feature selection. Finally, classification methods for features are chosen for the activity recognition with the criteria.

A window of 5 s is used for the period of activity recognition. There are various window sizes depending on sensor types and activities to recognize in a mobile environment. Hyunh and Schilele compared diverse window sizes, 0.25, 0.5, 1, 2, and 4 s, to recognize various behaviors such as ‘jogging,’ ‘walking,’ ‘skipping,’ ‘hopping,’ and ‘standing’ [30]. Bao and Intille determined each window size for activity recognition with 6.7 s [31]. Kern et al. used 1 s to detect human activity using body-worn sensors [32]. The smaller window size is not effective to consider certain long-term activities and the larger window size may include noises since multiple activities could exist. The window size of 5 s was used by our previous work [33] in classifying the classes of activities that this work targets.

To measure the “value” of each feature, the information gain (or the predictive power) of each feature is calculated [34]. Suppose that F be the set of all features and X the set of all training samples, $value(x, f)$ with $x \in X$ defines the value of a specific sample x for feature $f \in F$. E specifies the

Table 3 Information gain scores for feature selection

Sensor	Score
Orientation, pitch, roll	1.156053
Accelerometer	1.015941
Magnetic field sensor	1.013732
Light sensor	0.293353
Locations	0.267844
Proximity sensor	0.117537

entropy. The information gain for a feature $f \in F$ is defined as follows:

$$IG(X, f) = E(X) - \sum_{v \in values(a)} \frac{|\{x \in X | value(x, f) = v\}|}{|X|} \cdot E(\{x \in X | value(x, f) = v\}) \tag{1}$$

The information gain is the difference between the total entropy and the relative entropies when a specific feature value is determined. It can be used as a score to measure the power of prediction [35].

As Table 3 shows, orientation, pitch, roll, acceleration, and magnetic field show relatively high information gain score. On the other hand, the other sensors such as light sensor, location, and proximity sensor have lower information gain score. This result implies the need to analyze acceleration, magnetic field, orientation values more carefully. To consider temporal patterns delicately, hidden Markov model (HMM) is applied to the values of the three sensors of high information gain score. The other features are processed using simple rules with pre-defined thresholds.

3.2 First layer HMMs for short-term activity recognition

HMM is a probabilistic model based on Markov chains, and it is suitable to handle time series data such as speech processing and stochastic signal processing [36]. The HMM λ consists of three elements as follows:

$$\lambda = (A, B, \pi) \tag{2}$$

where λ represents a HMM model, A is a state transition probability distribution, B is an observation distribution, and π is an initial state distribution. Let us assume that we have M observation symbols and N states for this model. $A = \{a_{ij}\}$, including transition probability from state i to state j , is defined as follows:

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i), \quad 1 \leq i, j \leq N. \tag{3}$$

where $a_{ij} > 0$ for all i, j . The observation probability in state j , $B = \{b_j(k)\}$, is defined as follows:

$$b_j(k) = P(x_k \text{ at } t | q_t = S_j), \quad 1 \leq j \leq N, \quad 1 \leq k \leq M. \tag{4}$$

The initial state probability in state i , $\pi = \{\pi_i\}$, is defined as follows:

$$\pi_i = P(q_1 = S_i), \quad 1 \leq i \leq N. \tag{5}$$

The probability of the observation sequence, $X = x_1, x_2, \dots, x_T$, given the model λ , $P(X|\lambda)$ is calculated through enumerating every possible state sequence of length T (the number of observations). A fixed state sequence Q is defined as follows:

$$Q = q_1, q_2, \dots, q_T \tag{6}$$

where q_1 is the initial state. The probability of the observation sequence X for the state sequence of Eq. (6) is

$$P(X|Q, \lambda) = \prod_{t=1}^T P(X_t | q_t, \lambda) = b_{q_1}(x_1) \cdot b_{q_2}(x_2) \cdot \dots \cdot b_{q_T}(x_T). \tag{7}$$

The probability of such a state sequence Q can be written as

$$P(Q|\lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{T-1} q_T}. \tag{8}$$

The joint probability of X and Q , the probability that X and Q occur simultaneously is simply the product of the above two terms as follows.

$$P(X, Q|\lambda) = P(X|Q, \lambda)P(Q|\lambda). \tag{9}$$

The probability of X (given the model) is obtained by summing the joint probability in Eq. (9) as follows.

$$\begin{aligned} P(X|\lambda) &= \sum_{\text{all } Q} P(X|Q, \lambda)P(Q|\lambda) \\ &= \sum_{q_1, q_2, \dots, q_T} \pi_{q_1} b_{q_1}(x_1) a_{q_1 q_2} b_{q_2}(x_2) \dots a_{q_{T-1} q_T} b_{q_T}(x_T) \end{aligned} \tag{10}$$

There are two types of HMMs, continuous HMM and discrete HMM, according to data types. While the discrete HMM deal with discrete data from a categorical distribution, the continuous HMM uses a single Gaussian or a mixture of Gaussians as the continuous observation distribution. Using Gaussian mixtures for observation distributions requires evaluation of the probability densities in the mixture for each feature vector at each state in the HMM. Since the evaluations are computationally complex, they account for much of the time spent in activity recognition [37]. In order to recognize activities more efficiently, we consider to quantize the feature vectors into a finite set of symbols prior to activity recognition as shown in Fig. 2.

K-means clustering is performed to quantize the feature vectors into finite symbols. Clustering is a method to assign a set of samples into groups according to a distance metric [30]. K-means clustering aims to partition n observations into k groups in which each observation belongs to the group with the nearest mean. The algorithm uses an iterative refinement technique. Given an initial set of k means m_1, \dots, m_k , the algorithm proceeds by iterating the following two steps:

- Assignment step: assign each observation to the cluster with the closest mean

$$C_i^t = \{x_p : \|x_p - m_i^{(t)}\| \leq \|x_p - m_j^{(t)}\|, \forall 1 \leq j \leq k\}$$

where each x_p goes into exactly one $C_i^{(t)}$, even if it could go in two of x_p : an observation, m_i^t : mean for the i th cluster at time step t $C_i^{(t)}$: a set of observations assigned to the i th cluster at time step t

- Update step: calculate the new means to be the centroid of the observations in the cluster as follows

$$m_i^{(t+1)} = \frac{1}{|C_i^{(t)}|} \sum_{x_j \in C_i^{(t)}} x_j \tag{11}$$

As mentioned in the previous section, discrete HMMs are trained using orientation, acceleration, and magnetic field to analyze a user’s activity. The data are discretized into 20 states by k-means clustering with $k = 20$. Our previous work [33] used continuous HMM to recognize

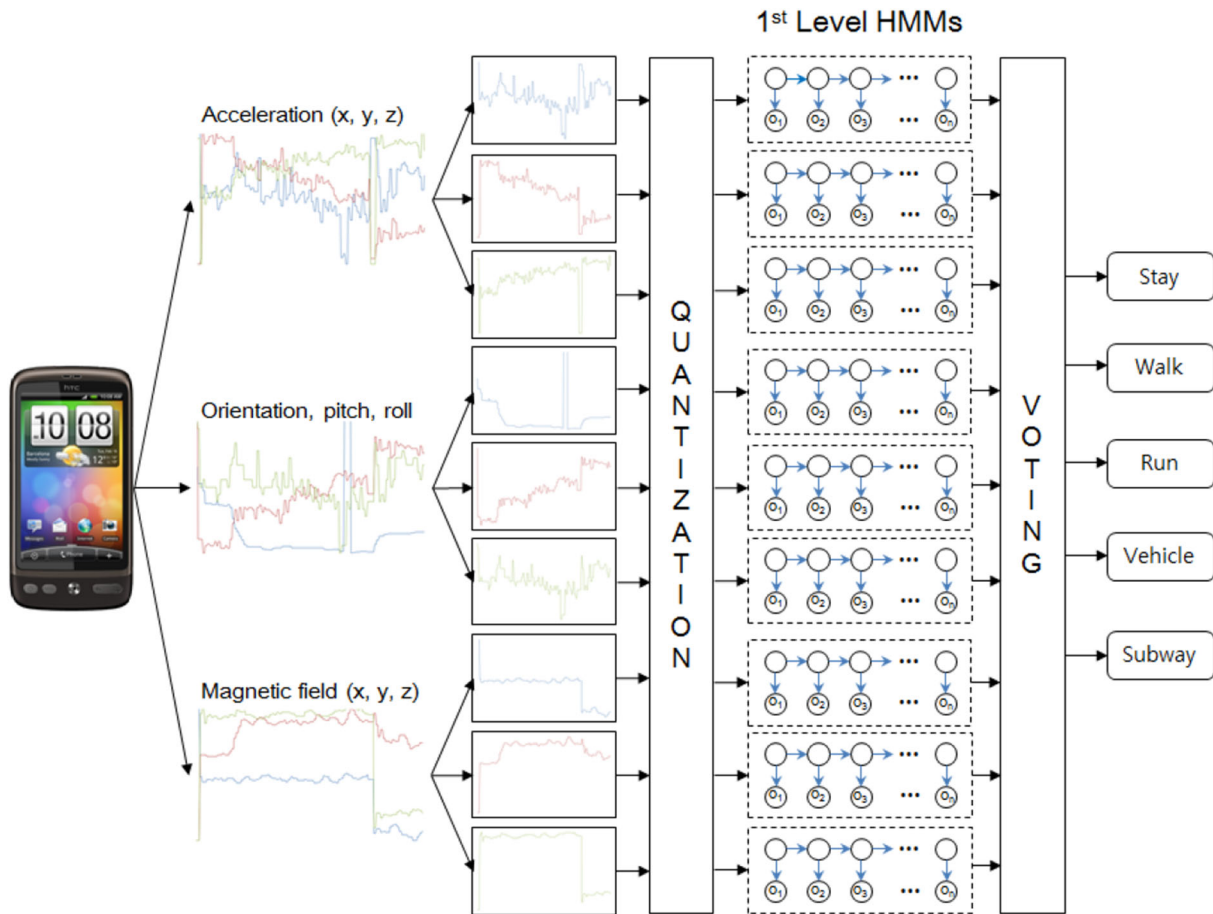


Fig. 2 First layer HMMs for short-term activity

activities from acceleration, where the observations were assumed to follow a Gaussian distribution. Although the acceleration data follow a Gaussian distribution, continuous HMMs for magnetic field and orientation, not following Gaussian distribution, cannot show good performance in the experiments. The HMMs at the first layer recognize short-term activities: stay, walk, run, vehicle, and subway, and they have five hidden states.

3.3 Second layer HMMs for long-term activity recognition

After the first layer HMMs, we can get the inference results of the short-term activities among stay, walk, run, vehicle and subway. In addition, the features of the light, proximity, and time, and the locations visited at that time constitute a feature vector, which is passed to the next layer of activity recognition as shown in Fig. 3. The models at this level are also discrete HMMs, with one HMM per long-term activity to classify. This layer of HMMs gets the sequence of the feature vectors for about 5 min to handle the concepts that have longer temporal patterns. Long-term

activities recognized by the system include: relaxing, moving, working, and eating.

The final goal of the system is to decompose in real-time the temporal sequence obtained from the sensors into concepts at different levels of abstraction or temporal granularity. At each level, we use the forward-backward algorithm to compute the likelihood of a sequence given a particular model. The HMM model with the highest likelihood is selected to perform inference with LHMMs. In the approach, the most probable short-term activity is used as an input to the HMMs at the next level [28].

Let us suppose that we train K HMMs at level L of the hierarchy, λ_k^L , with $k = 1, \dots, K$. The log-likelihood of the observed sequence $X_{1:T}$ for model λ_k^L , $L(k)_T^L$ is defined as follows:

$$L(k)_T^L = \log(P(X_{1:T}|\lambda_k^L)) = \log \sum_i \alpha_T(i; \lambda_k^L) \quad (12)$$

where $\alpha_T(i; \lambda_k^L)$ is the alpha variable of the standard Baum-Welch algorithm at time T , state i and for model λ_k^L . Equation (13) shows a recursive function $\alpha_T(i; \lambda_k^L)$.

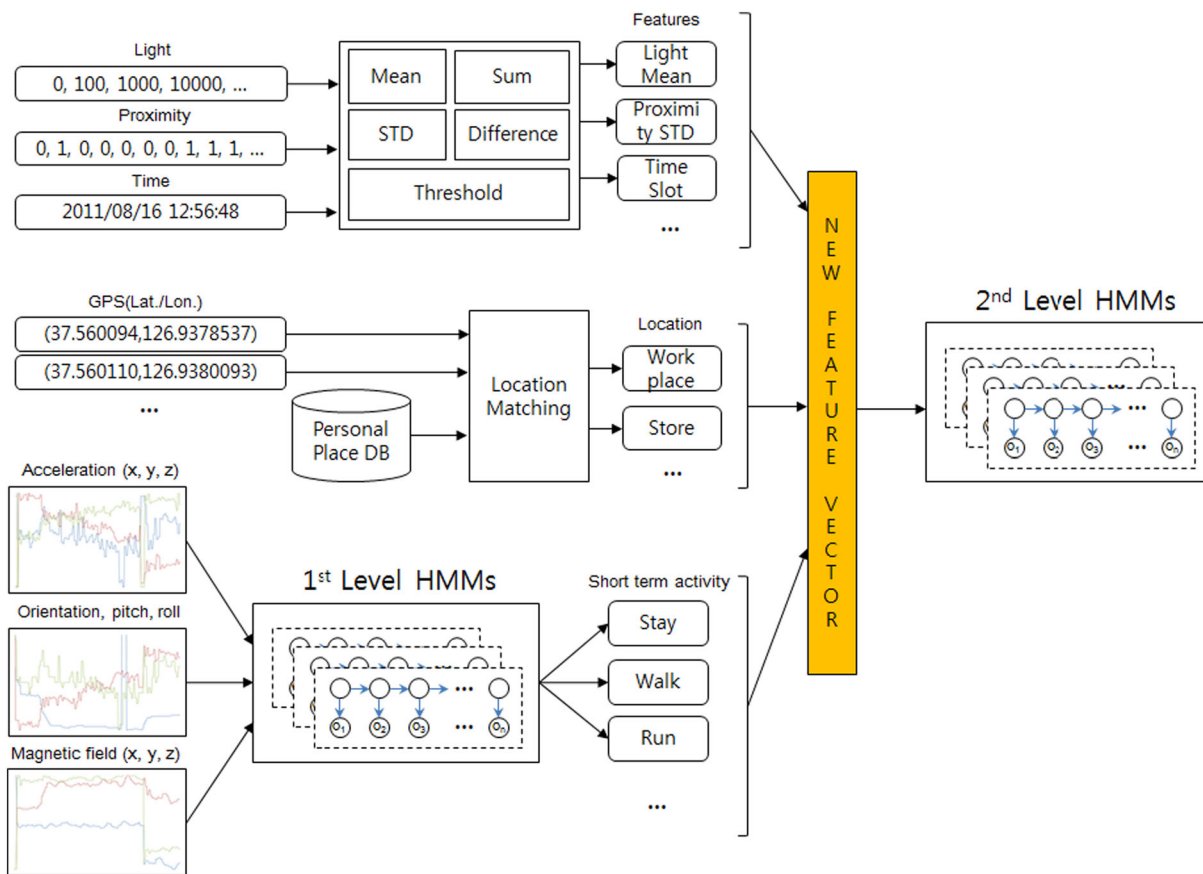


Fig. 3 Second layer HMMs for long-term activity

$$\alpha_{T+1}(j; \lambda_k^L) = \sum_{i=1}^N \alpha_T(i; \lambda_k^L) a_{ji}^{\lambda_k^L} b_j(x_T; \lambda_k^L)$$

$a_{ji}^{\lambda_k^L}$: the transition probability from state j to state i from model λ_k^L
 $b_j(x_T; \lambda_k^L)$: the probability for state j in model λ_k^L of observing x_T

(13)

At that level, we classify the observations by declaring class $\text{Class}(T)^L$ as follows.

$$\text{Class}(T)^L = \arg \max_k L(k)_t^L, \quad k = 1, \dots, K \tag{14}$$

The window size varies with the granularity of each level. At the first level of the hierarchy, the samples of the time window are extracted from the raw sensor data. At the second level of the hierarchy, the inference outputs of the previous level are used as part of samples. The other sensors except acceleration, orientation, and magnetic field are analyzed to extract three features in Eqs. (15), (16), and (17).

$$\text{sum}_X = \sum_{i=1}^S \sqrt{(x_i - x_{i-1})^2} \tag{15}$$

$$\text{mean}_X = \frac{\sum_{i=1}^S \sqrt{(x_{i+1} - x_i)^2}}{S} \tag{16}$$

$$\text{std}_X = \sqrt{\left(\sum \sqrt{(x_{i+1} - x_i)^2} - \text{mean}_X \right)^2} \tag{17}$$

where x_i is the value from a specific sensor at time step i , and S is the total number of samples in a window.

3.4 Mobile interface for visualization

The recognized activity is displayed by using mobile applications. We develop two applications to manage personal information such as visited places, short-term activity, and long-term activity. Figure 4 shows an application for short-term activity visualization. It has three interfaces: activity representation by text, activity representation by graph, and place representation on a map. In addition, the other program summarizes long-term activities in a day as shown in Fig. 5.

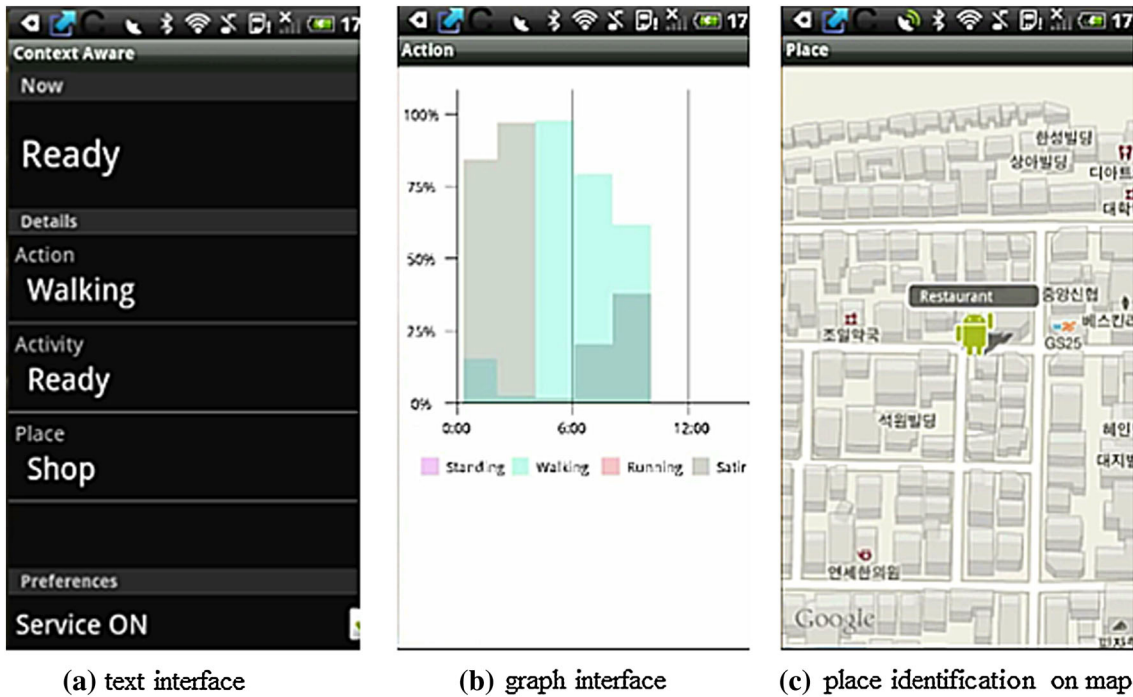
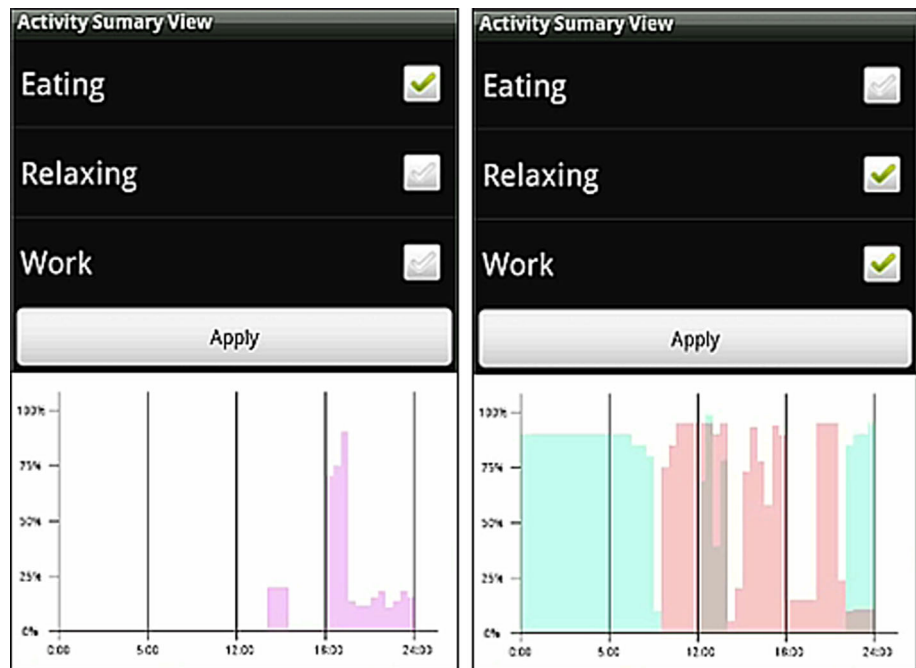


Fig. 4 Short-term activity visualization on a mobile phone

Fig. 5 Long-term activity visualization on a mobile phone



4 Experiments

4.1 Data collection

Mobile sensor data were collected from four graduate students who are 26–32 years old for over a week. Figure 6

shows a part of the logs, which illustrates the correlations between the temporal pattern and the short-term activities. In this paper, attending a lecture and studying were regarded as a work, one of long-term activities, since the users were students. Table 4 summarizes the details of collected data from a mobile phone.

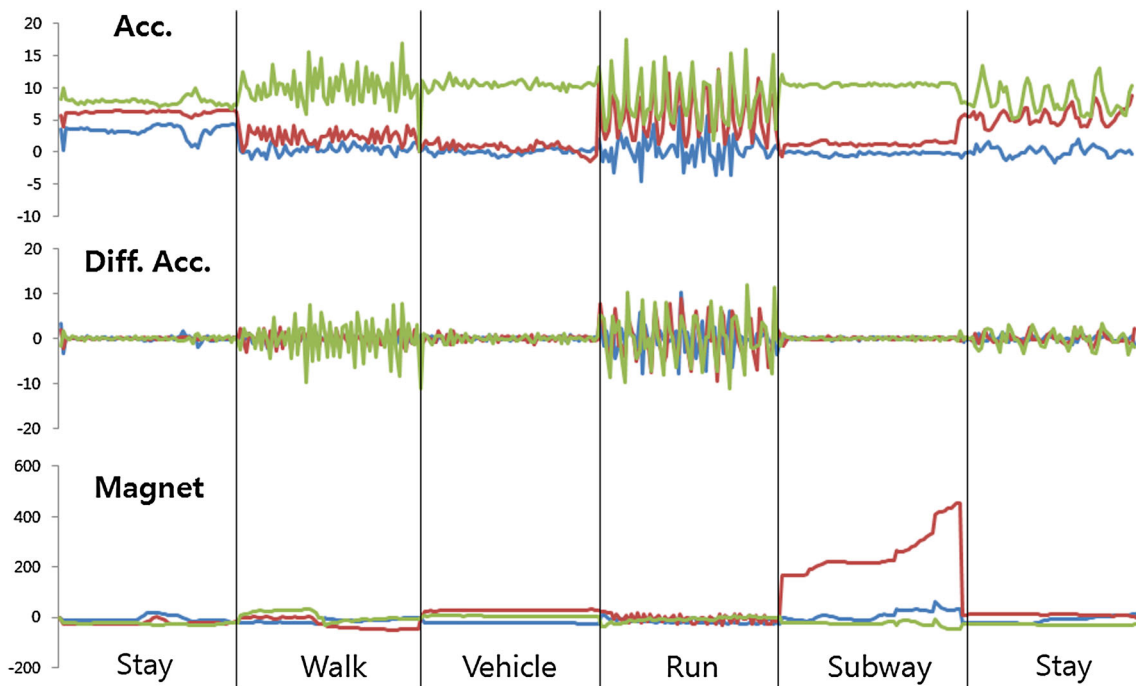


Fig. 6 Acceleration and magnetic field data for each short-term activity

Table 4 A summary of the data sets used

Type	Format	Description
Time	YYYYMMDD hh:mm:ss	Date and time
Location	Location name	Matched or labeled location
Position	Hand/pocket	Attached position
Acceleration	Real number (<i>X</i> axis)	−2–2 g
	Real number (<i>Y</i> axis)	−2–2 g
	Real number (<i>Z</i> axis)	−2–2 g
Orientation	Real number	0–359
Pitch	Real number	−180–180
Roll	Real number	−90–90
Magnetic field	Real number (<i>X</i> axis)	Digital compass
	Real number (<i>Y</i> axis)	Digital compass
	Real number (<i>Z</i> axis)	Digital compass
Light	Integer number	Light intensity
Proximity	1 or 0	Contact or not contact

4.2 Evaluation of HMMs for short-term activity recognition

To compare the proposed k-means clustering-based HMM (KMC + HMM) with other classification methods, we conducted an experiment using the collected data set. In the experiment, we applied naïve Bayes (NB), multi-layer perceptron (MLP), j48 decision tree (J48), support vector machine (SVM), and Bayesian network (BN) as well as a single layer HMM with quantization by k-means clustering. Those models were from the Weka (<http://www.cs.waikato.ac.nz/~ml/weka>).

Some features such as mean, standard deviation, and summation in a window are used to recognize short-term activities (stay, walk, run, vehicle, and subway) by the classification methods because most of them cannot handle time series data. Figure 7 shows a part of the short-term activities.

We tested the precision and the recall of the recognized short-term activities by comparing them with the actual labels acquired from the users, and achieved about 80 % for the short-term activities as shown in Table 5. Here, the proposed first-level HMM shows the comparable

Fig. 7 Short-term activities to be recognized by KMC + HMM



Table 5 Performance comparison of classification methods for short-term activity

	KMC + DHMM	NB	MLP	SVM	J48	BN
Avg. precision	0.79	0.73	0.65	0.80	0.76	0.73
Avg. recall	0.77	0.66	0.66	0.69	0.75	0.80

Table 6 Comparison of precision for short-term activity

	KMC + DHMM	NB	MLP	SVM	J48	BN
Staying	0.63	0.62	0.73	0.70	0.67	0.59
Walking	0.86	0.67	0.81	0.90	0.79	0.46
Running	1.00	1.00	0.00	0.79	0.58	0.79
Vehicle	0.79	0.73	0.81	0.67	0.79	0.78
Subway	0.61	0.57	0.92	0.95	0.99	0.97

performance with other classification methods, even though it is not the best for both criteria. However, we can see the proposed model produces the results consistently. Tables 6 and 7 show the performance for each short-term activity. There are some subjects who used subway during data collection. Although location information is not used at this level, using subway is recognized well because of the magnetic field sensor.

4.3 Evaluation of LHMMs for long-term activity recognition

We attempt to recognize four kinds of long-term activities (move, eating, relaxing, and working) as shown in Fig. 8. It is assumed that the activities have different temporal patterns correlated to sensing values. For instance, when a user is moving at a department store, he or she is going to repeat walking and standing regularly. On the other hand, if a user works at the office, the acceleration pattern may be similar to ‘staying’ for sitting in a seat, and his location is

Table 7 Comparison of recall for short-term activity

	KMC + DHMM	NB	MLP	SVM	J48	BN
Staying	0.78	0.30	0.69	0.40	0.66	0.46
Walking	0.82	0.79	0.86	0.66	0.66	0.88
Running	0.75	0.75	0.00	1.00	0.75	1.00
Vehicle	0.58	0.81	0.85	0.94	0.84	0.82
Subway	0.90	0.80	0.96	0.60	0.90	0.82

fixed to ‘office.’ The long term activities can be recognized by considering location, and time as well as a sequential pattern of short-term activities. Especially, location is important information to estimate a user’s activities. Figure 9 shows the locations for each long-term activity.

In Table 8, the precision and the recall of the proposed layered HMM are compared with other classification methods such as BN, NB, MLP, SVM, and J48. The layered HMM shows better performance in average than the other methods. For precision, Relax is the only long-term activity that the proposed method was not perfect, and for recall, the proposed method was perfect except the activity of Eat. It also shows that naïve Bayes classifier has worse performance than all the other methods. The difference between NB and BN implies that hierarchical structure is suitable to recognize the long-term activities.

Tables 9 and 10 depict the performance of the classifiers for each long-term activity. The experiment was done with tenfold cross validation. J48 and LHMM show better performance for all activities. However, it is difficult to classify ‘relax’ and ‘eat’ activities sometimes because they have similar short-term activity patterns, location, and time.

5 Concluding remarks

In this paper, we attempt to recognize short-term/long-term activities in real time using mobile sensors on the Android platform. The layered HMM structure is used to model the temporal patterns with multi-dimensional data. As HMM is

Fig. 8 Long-term activities to be recognized by HHMM

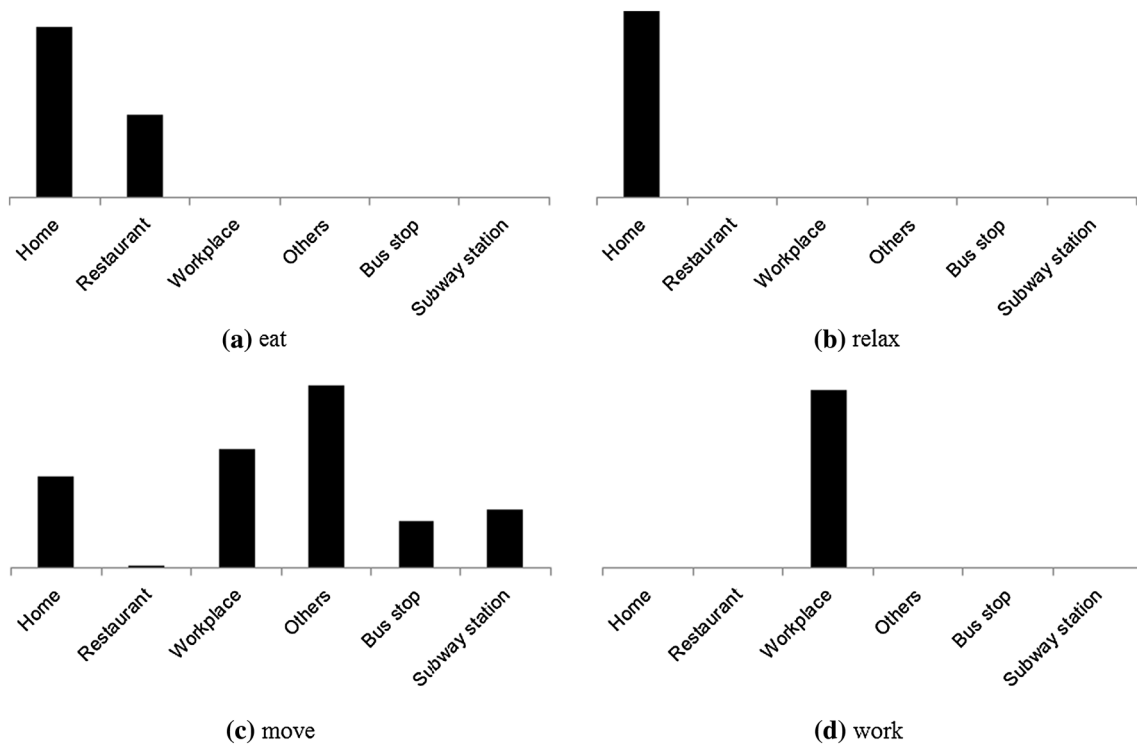


Fig. 9 Locations for each long-term activity

a Markov chain with both hidden and unhidden stochastic processes, for activity recognition, the unhidden or observable components are the sensor signals, while the hidden element is the user’s activity. Looking at the results

of short-term activity recognition, it shows comparable performance with other classification methods. LHMM has the best accuracy among other classification methods for recognizing the long-term activities.

Table 8 Performance comparison of classification methods for long-term activity

	BN	NB	MLP	SVM	J48	LHMM
Precision	0.86	0.76	0.94	0.84	0.96	0.97
Recall	0.84	0.58	0.94	0.83	0.96	0.97

Table 9 Comparison of precision for long-term activity

	BN	NB	MLP	SVM	J48	LHMM
Eat	0.83	0.85	0.97	0.54	0.95	1.00
Move	0.86	0.83	0.94	0.90	0.96	1.00
Relax	0.71	0.37	0.82	0.64	0.92	0.80
Work	0.95	0.84	0.99	0.96	0.99	1.00

Table 10 Comparison of recall for long-term activity

	BN	NB	MLP	SVM	J48	LHMM
Eat	0.40	0.14	0.83	0.74	0.95	0.80
Move	0.86	0.26	0.98	0.93	0.96	1.00
Relax	0.85	0.90	0.82	0.34	0.88	1.00
Work	0.95	0.97	0.98	0.96	0.98	1.00

There are still many problems to be solved for the activity recognition on a mobile phone. In our experiment, ‘relax’ and ‘eat’ activities have some difficult patterns to classify, and we need to use more wearable sensors. It is necessary to consider modeling a user’s variations for personalized services as well. Moreover, the comparison with more diverse classification methods such as dynamic time warping (DTW), hierarchical HMM (HHMM), and hierarchical dynamic Bayesian network (HDBN) is also a very crucial issue to be considered as a future work.

Acknowledgments This work was supported by the industrial strategic technology development program, 10044828, Development of augmenting multisensory technology for enhancing significant effect on service industry, funded by the Ministry of Trade, Industry & Energy (MI, Korea).

References

- Lane ND, Miluzzo E, Peebles D, Choudhury T, Campbell AT (2010) A survey of mobile phone sensing. *IEEE Commun Mag* 48(9):140–150
- Lara OD, Labrador MA (2013) A survey on human activity recognition using wearable sensors. *IEEE Commun Surv Tutor* 15(3):1192–1209
- Bulling A, Blanke U, Schiele B (2014) A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput Surv* 46(3):1–33
- Chen L, Hoey J, Nugent C, Cook D, Yu Z (2012) Sensor-based activity recognition. *IEEE Trans Syst Man Cybern Part C Appl Rev* 42(6):790–808
- Choudhury T, Consolvo S, Harrison B, Hightower J, LaMarca A, LeGrand L, Rahimi A, Rea A, Borriello G, Hemingway B, Klasnja PP, Koscher K, Landay JA, Lester J, Wyatt D, Haehnel D (2008) The mobile sensing platform: an embedded activity recognition system. *IEEE Pervasive Comput* 7(2):32–41
- Hwang K-S, Cho S-B (2009) Landmark detection from mobile life log using a modular Bayesian network model. *Expert Syst Appl* 36(10):12065–12076
- Bieber G, Voskamp J, Urban B (2009) Activity recognition for everyday life on mobile phones. In: *Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction, part II: intelligent and ubiquitous interaction environments* 289–296
- Park S, Aggarwal KJ (2004) A hierarchical Bayesian network for event recognition of human actions and interactions. *Multimedia Syst* 10(9):164–179
- Wang AW-H, Tung C-L (2008) Dynamic gesture recognition based on dynamic Bayesian networks. *WSEAS Trans Bus Econ* 4(11):168–173
- Du Y, Chen F, Xu W, Zhang W (2006) Interacting activity recognition using hierarchical durational-state dynamic Bayesian network. *Lect Notes Comput Sci* 4261:185–192
- Min J-K, Hong J-H, Cho S-B (2015) Combining localized fusion and dynamic selection for high-performance SVM. *Expert Syst Appl* 42:9–20
- Song SK, Jang J, Park S (2008) A phone for human activity recognition using triaxial acceleration sensor. *IEEE Int Conf Consum Electron* 1–2. doi:10.1109/ICCE.2008.4587903
- Zappi P, Lombriser C, Stiefmeier T, Farella E, Roggen D, Benini L, Tröster G (2008) Activity recognition from on-body sensors: accuracy-power trade-off by dynamic sensor selection. *Lect Notes Comput Sci* 4913:17–43
- Ganti RK, Jayachandran P, Abdelzaher TF, Stankovic JA (2006) Satire: a software architecture for smart attire. In: *Proceedings of the 4th International Conference on Mobile Systems, Applications and Services* 110–123
- Khan AM, Lee YK, Lee SY, Kim TS (2010) Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. In: *5th International Conference on Future Information Technology* 1–6
- Kwapisz JR, Weiss GM, Moore SA (2010) Activity recognition using cell phone accelerometers. In: *4th ACM SIGKDD International Workshop on Knowledge Discovery from Sensor Data*
- Maguire D, Frisby R (2009) Comparison of feature classification algorithm for activity recognition based on accelerometer and heart rate data. In: *9th IT & T Conference*
- Györfi N, Fábán Á, Hományi G (2009) An activity recognition system for mobile phone. *Mob Netw Appl* 14(1):82–91
- Berchtold M, Budde M, Gordon D, Schmidtke H, Beigl M (2010) ActiServ: activity recognition service for mobile phones. In: *2010 International Symposium on Wearable Computers* 1–8
- Gupta P, Dallas T (2014) Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Trans Biomed Eng* 61(6):1780–1786
- Pei L, Guinness R, Chen R, Liu J, Kuusniemi H, Chen Y, Chen L, Kaistinen J (2013) Human behavior cognition using smartphone sensors. *Sensors* 13(2):1402–1424
- Travelsi D, Mohammed S, Chamroukhi F, Oukhellou L, Amirat Y (2013) An unsupervised approach for automatic activity

- recognition based on hidden Markov model regression. *IEEE Trans Autom Sci Eng* 10(3):829–835
23. Lu H, Pan W, Lane ND, Choudhury T, Campbell AT (2009) SoundSense: Scalable sound sensing for people-centric applications on mobile phones. In: *Proceedings of the 7th International Conference on Mobile Systems* 165–178
 24. Reddy S, Burke J, Estrin D, Hansen M, Srivastava M (2008) Determining transportation mode on mobile phones. In: *Proceedings of 12th IEEE International Symposium on Wearable Computers* 25–28
 25. Yang J-Y, Chen Y-P, Lee G-Y, Liou S-N, Wang J-S (2007) Activity recognition using one triaxial accelerometer: a neuro-fuzzy classifier with feature reduction. *Lect Notes Comput Sci* 4740:395–400
 26. Aarno D, Kragic D (2006) Layered HMM for motion intention recognition. *IEEE RSJ Int Conf Intell Robot Syst* 5130–5135. doi:[10.1109/IROS.2006.282606](https://doi.org/10.1109/IROS.2006.282606)
 27. Mengistu KT, Hannemann M, Baum T, Wendemuth A (2008) Hierarchical HMM-based semantic concept labeling model. *IEEE Spoken Language Technology Workshop, IEEE, Goa*, pp 57–60. doi:[10.1109/SLT.2008.4777839](https://doi.org/10.1109/SLT.2008.4777839)
 28. Oliver N, Garg A, Horvitz E (2004) Layered representations for learning and inferring office activity from multiple sensory channels. *Comput Vision Image Underst* 96(2):163–180
 29. Michalak K, Kwaśnicka H (2006) Correlation-based feature selection strategy in classification problems. *Appl Math Comput Sci* 16(4):503–511
 30. Huynh T, Schiele B (2005) Analyzing features for activity recognition. In: *Proceedings of the Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies* 159–164
 31. Bao L, Intille SS (2004) Activity recognition from user-annotated acceleration data. *Lect Notes Comput Sci* 3001:1–17
 32. Kern N, Schiele B, Schmidt A (2003) Multi-sensor activity context detection for wearable computing. *Lect Notes Comput Sci* 2875:220–232
 33. Lee Y-S, Cho S-B (2011) Activity recognition using hierarchical hidden Markov models on a smartphone with 3D accelerometer. *Lect Notes Artif Intell* 6678:460–467
 34. Reddy S, Mun M, Burke J, Estrin D, Hansen M, Srivastava M (2010) Using mobile phones to determine transportation modes. *ACM Trans Sens Netw* 6(2) article 13. doi:[10.1145/1689239.1689243](https://doi.org/10.1145/1689239.1689243)
 35. Peirola R (2011) Information gain as a score for probabilistic forecasts. *Meteorol Appl* 18:9–17
 36. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE* 77(2):257–286
 37. Chen FR, Wilcox LD, Bloomberg DS (1995) A comparison of discrete and continuous hidden Markov models for phrase spotting in text images. In: *Proceedings of the Third International Conference on Document Analysis and Recognition* 398–402