



HHS Public Access

Author manuscript

J Med Syst. Author manuscript; available in PMC 2017 December 13.

Published in final edited form as:

J Med Syst. 2015 May ; 39(5): 57. doi:10.1007/s10916-015-0239-x.

An Adaptive Hidden Markov Model for Activity Recognition Based on a Wearable Multi-Sensor Device

Zhen Li,

Department of Computer Science, Ocean University of China, Qingdao, China. Department of Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA

Zhiqiang Wei,

Department of Computer Science, Ocean University of China, Qingdao, China

Yaofeng Yue,

Department of Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA. Department of Electrical & Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA

Hao Wang,

Department of Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA. Department of Electrical & Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA

Wenyan Jia,

Department of Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA. Department of Electrical & Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA

Lora E. Burke,

Department of Health and Community Systems, University of Pittsburgh, Pittsburgh, PA, USA

Thomas Baranowski, and

Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA

Mingui Sun

Department of Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA. Department of Electrical & Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA

Abstract

Human activity recognition is important in the study of personal health, wellness and lifestyle. In order to acquire human activity information from the personal space, many wearable multi-sensor devices have been developed. In this paper, a novel technique for automatic activity recognition based on multi-sensor data is presented. In order to utilize these data efficiently and overcome the big data problem, an offline adaptive-Hidden Markov Model (HMM) is proposed. A sensor selection scheme is implemented based on an improved Viterbi algorithm. A new method is proposed that incorporates personal experience into the HMM model as a priori information. Experiments are conducted using a personal wearable computer eButton consisting of multiple sensors. Our comparative study with the standard HMM and other alternative methods in

processing the eButton data have shown that our method is more robust and efficient, providing a useful tool to evaluate human activity and lifestyle.

Keywords

Hidden Markov model; Activity recognition; Wearable device; Big data; Machine learning; Personal health

Introduction

According to the World Health Organization,¹ chronic diseases, such as certain heart diseases, respiratory conditions, cancer, diabetes, and hypertension, caused 60 % of deaths in the world population in 2002. By 2020, this percentage is expected to reach 73 %. Most chronic diseases are lifestyle related, resulting from unhealthy habits or practices such as smoking, overeating, or insufficient physical activity. Being able to monitor an individual's lifestyle automatically and objectively using advanced technology is thus extremely important in preventing and managing chronic diseases [1].

There are many methods to monitor or document an individual's lifestyle. Most of them depend on self-report. However, it is well-known that self-report based methods are inaccurate, biased and unreliable. Due to these problems, objective methods have been studied using portable electronic devices which are mostly designed to collect physical activity data. The accelerometer is one the most widely utilized sensors in these devices [2–6]. In recent years, cellphones equipped with an accelerometer chip and a camera have been used in lifestyle surveillance and physical activity studies [7, 8]. The SenseCam[9], a wearable camera with several other sensors, was used to record life experience continuously. However, this device is relatively large and heavy, thus, it must be worn using a lanyard around the neck, which is inconvenient when vigorous physical activity is performed. Although there are many forms of smart wristbands and wristwatches available for activity monitoring, they cannot identify many specific activities because of their wearing locations on the body and sensor limitations.

In order to monitor lifestyle, we have developed a wearable electronic device, eButton, as shown in Fig. 1. This device is small and light but has multiple sensors and a powerful CPU to acquire and process data. These sensors include one or two cameras, a gyroscope, an accelerometer, a GPS receiver, an electronic compass, a thermometer, a light sensor, a proximity sensor and a barometer. The eButton also has wireless channels to communicate with a smartphone, a nearby computer, or a “cloud”. Despite these powerful functions, eButton is a very convenient and unobtrusive device with an attractive and personalized appearance. It is thus a useful tool for the study of lifestyle, including diet, physical activity, and sedentary behavior.

The data collected by eButton results in a high demand for data processing. Recently, activity recognition has gained dramatic interest. Many types of visual features, including

¹http://www.who.int/chp/about/integrated_cd/en/

static and dynamic features, have been proposed for image classification. For example, the Color and Edge Directivity Descriptor (CEDD) [10] combines color and edge distributions for static image indexing. The scale-invariant feature transform (SIFT) [11] is one of the most utilized features applied to activity recognition [12–14]. There are also features specifically designed for motion sensors (accelerometer or gyroscope) for physical activity recognition such as the Signal-Magnitude Area (SMA) [3], Accumulated Height (AH), Autoregressive Coefficients (AC), Time Between Peaks (TBP) and Binned Distribution (BD) [8]. Besides these features, classification methods based on generative and discriminative approaches have been used in activity recognition. Among probabilistic methods, the Hidden Markov Model (HMM) [15], a special case of dynamic Bayesian networks, is well known. The discriminative methods include the k-Nearest Neighbors [16], Support Vector Machine (SVM) [17], Neural networks [18] and Conditional Random Field (CRF) [19].

Due to the large volume of continuously collected data by eButton, a “big data” problem must be solved to understand these data computationally. As a result, the speed of activity recognition is essential since these data are currently processed offline. The traditional algorithms for pattern recognition based on multi-sensor data cannot be used effectively, mainly due to their high computational cost which grows proportionally with the number of data channels [20]. There is a strong need to develop algorithms that make smart choices among the types of data features to be processed yet still produce reliable recognition results. Our study focuses on solving the big data problem and satisfies this need.

This paper presents the adaptive-HMM in Section 2, experiments and data analysis in Section 3, and conclusion in Section 4.

Methods

Concepts

The algorithms presented in this paper are designed based on the following new concepts: 1) Since eButton is a personalized device serving a specific individual, there is no need for it to recognize all types of human activities since this recognition is an extremely difficult task. Instead, we expect our algorithms to recognize only the activities commonly performed by an individual, which becomes a much easier task. 2) Although each individual’s lifestyle is different, they perform routine activities periodically (e.g., the weekday activities of a person with an office job). We take advantage of this important routine information, model it mathematically, and use it in our system design. 3) To avoid processing all collected data at a high computational cost, we pre-select the types of data that are determined to be relevant to the activity being tested. This approach solves the “big data” problem effectively without suffering from significant loss of accuracy.

Personal experience knowledge base

Each individual has his/her own lifestyle and is the best person to describe it in terms of routines in their daily lives although he/her may not remember all the particular activities performed during a specific period of time (e.g., during the last month). In our opinion,

personal experience provides the best knowledge base for the computer to recognize activities of the individual. With a relatively fixed lifestyle, the individual can roughly estimate the likelihood of each of his/her own activities given certain preconditions such as location, time or the activity performed prior to the current activity. For example, Table 1 provides a time table of an individual where column A and row T contain, respectively, m types of activities (from a^1 to a^m) and the time periods of the “average day”. Here we divide the day into consecutive 30-min intervals for simplicity so there are 48 rows in the table. The individual is asked to assign an integer value from 0 to 10 for each cell in the table which stands for the weight of activity. For the days with different lifestyles, e.g., workdays and holidays, different time tables are used. We utilize the 0–10 scale because it is more intuitive for a human to assign values. After this table is acquired, it can be converted to a probability table simply by normalization. Besides the time table, a transition table is also acquired from the individual specifying the likelihoods of consecutive activity pairs. Table 2 shows an example with three likelihood levels including Low (L), Medium (M) and High (H). Each level represents the weight of current activity A happening right after its previous activity A_p .

These tables could be obtained through machine learning based on real world data. However, since a large dataset is required to ensure the tabulated frequencies approach the true probabilities, this machine approach would be time-consuming and prone to noise and computational errors.

Adaptive hidden Markov model

We chose the HMM because the outputs of the two experience tables (Tables 1 and 2 above) provide probabilities of different activities which are relatively easily imported into the probability-based model. Besides, for each day, the sequence of activities may be modeled as a Markov chain in which each single activity is assumed to be related to its previous activity only. Clearly, this model is not strictly accurate because the second and higher order relationships are ignored. All data captured by the eButton and the entries in the experience tables are considered in the HMM model as observation values. Instead of using the traditional HMM, we propose an adaptive scheme described in Fig. 2 where A_n is assumed to be the hidden state which is obtained by segmenting the continuously recorded data based on our previous work [21]. In this segmentation, both the motion and image data are utilized to identify boundaries of unknown events according to similarities in the data. Each A_n represents a distinct activity different from its adjacent ones.

For each pair of consecutive activities, the details of the activity state transition and the observation vector are shown in Fig. 3.

The *transition probability* $P(A_n|A_{n-1})$ and *time probability* $P(A_n|O_n^1)$ are provided by the transition table and time table, respectively. Two other probabilities $P(A_n|O_n^2)$ and $P(A_n|O_n^3)$ are calculated through two trained Support Vector Machine (SVM) classifiers. The first probability $P(A_n|O_n^2)$ is defined by:

$$P(A_n|O_n^2) = \frac{\sum_j f(data_w, A_n)}{W} \quad (1)$$

With

$$f(data_w, A_n) = \begin{cases} 1 & S(data_w) = A_n \\ 0 & S(data_w) \neq A_n \end{cases} \quad (2)$$

where $data_w$ is the w^{th} segment in A_n produced by dividing the entire data equally, W is the number of segments. $S(data_w)$ is a trained SVM classifier by the GPS and motion sensor data. The output of $S(data_w)$ is the classification result. The second probability $P(A_n|O_n^3)$ is defined in the same way as $P(A_n|O_n^2)$ except that the SVM classifier is trained using the image data.

Improved Viterbi algorithm

In this model, O_n^2 and O_n^3 are adaptive nodes. Since the computation of image features is time-consuming, we process the sensor node O_n^2 before image node O_n^3 . In the case where there is enough information to identify an activity, i.e., the probability value after processing O_n^2 is sufficiently high, the O_n^3 node can be skipped. As a result, the computation of image features can be significantly shortened.

In order to select suitable features according to different activities, we improved the Viterbi algorithm, which is defined as:

$$V_n^1(j) = \max_{i \in A} (P(j|i) V_{n-1}(i)) \frac{P(j|O_n^1)}{P(j)} \quad (3)$$

where i and j are two activities, $P(j|i)$ is transition probability from activity i to activity j , n is the activity index, and $P(j)$ is the initial probability of activity j . For $q=1$ to Q , we compute the probability $V_n^q(j)$ of the most probable activity sequence for the first n activities where j is its final activity

$$\begin{aligned} & \text{if } \frac{\max_{j \in A} (V_n^q(j))}{\sum_{k \in A} V_n^q(k)} < T \text{ and } q \neq Q \\ & \quad V_n^{q+1}(j) = V_n^q(j) \frac{P(j|O_n^{q+1})}{P(j)} \\ & \text{Else } V_n(j) = V_n^q(j) \end{aligned} \quad (4)$$

where T is a predefined threshold reflecting the level of confidence in the $[0, 1]$ interval (with 0 and 1 being the least and most confidence; we set $T=0.9$), q is the number of imported nodes, Q is the total number of nodes.

Using this adaptive method, nodes are computed sequentially. If the probability of an activity dominates (e.g., at least 90 %) of all possible activities, the rest of the nodes are skipped because there is enough statistical confidence to draw a conclusion. In order to reach this point rapidly, we place the experience nodes which require little computation first, the data nodes where measurements are accurate and easy to evaluate next, and the time-consuming image nodes last. This adaptive scheme reduces computational cost significantly and represents our main strategy to solve the “big data” problem stated in the “Introduction” section.

In certain cases, the designated confidence level is never reached even when the most time-consuming data node is utilized. In this case, the final probabilities of the activities ranked in descending order are provided as the output. The result can then be judged by a computing system, the person who analyzes data, or the person who performed the activities.

Feature selection

For accelerometer and gyroscope data, we extract the following features for activity recognition: mean, standard deviation, correlation, range, root mean square, Signal-Magnitude Area (SMA) [3], Autoregressive Coefficients (AC) and Binned Distribution (BD) [8]. Only two features are extracted from the GPS data, which are the speed and center location. For image data, four features are selected from the MPEG-7 standard [22], which are the Edge Histogram (EHD), Scalable Color (SCD), Color Layer (CLD) and Color Structure (CSD) descriptors. In addition, the Spatial Color Moments (SCM) [15] and Bag of Words (BoW) of the SIFT [23] are also included as image features.

We found that, in some cases, combining all these features may produce a poorer performance than using only a part of these features due to limited specificity in the data. Therefore, a further feature selection method is needed to determine the best subset of features. A straightforward method is to calculate all combinations of candidate features. However, this method is time-consuming since, if there are d features, 2^d experimental trials of classification are required. The number of trials grow quickly as d increases. We thus used the backward elimination algorithm [24] to solve this problem. In this method, features are all utilized at the beginning of the process. Then, each feature is deleted at a time of evaluation. If the deletion improves the performance of the model, it is excluded from the feature set. The process repeats until deleting any feature further cannot improve the performance.

Experiments

Dataset

With an approval from the Institutional Review Board (IRB), ten healthy human subjects participated in an experimental study to obtain training data. Each subject wore the eButton during daytimes. As the ground truth, the sensor and image data were segmented and labeled

manually in terms of the predefined categories shown in Table 3. The collected data contained 3-axis orientation signals from the gyroscope, 3-axis motion data from the accelerometer, geographical location data from the GPS and image data from the camera. 35 segments of the same activity performed were used in each category, and each segment consisted of 100 images corresponding to 400 s in real time.

After training, the data from four subjects, between 27 and 35 years of age, who were healthy engineers following normal work hours, were evaluated. In order to reflect the personal nature of our method, we processed the data for each subject independently using the same method. All subjects wore the eButton daily for at least 10 h during the daytime. Twenty-one days of raw data were collected from each subject. There were 1,298 segments in the recorded data, the distribution of the data are shown at Table. 4. Figure 4 presents several typical segments. Only images, 3-axis accelerometer signal and GPS signal are displayed in Fig. 4. It can be seen that the accelerometer signal and image data have distinctive patterns which can be used for activity classification.

Evaluation method

To evaluate the performance of the proposed method, the F_1 measure [25] which is used in many pattern recognition problems, is implemented as below:

$$\text{Precision} = TP / (TP + FP) \quad \text{Recall} = TP / (TP + FN) \quad F_1 = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$$

where TP , FP and FN represent, respectively, true positive, false positive and false negative. F_1 is also called the harmonic mean of recall and precision.

Performance of proposed algorithm

Table 4 presents the confusion matrix of the proposed algorithm. The F_1 value (last row in Table 5) of WO, RC, CW (definitions in Table 3) are higher than other activities. We analyzed the performances and found that the GPS signal plays an important role in detecting RC and WO; activity CW is detected best using images; the performances of HW are more dependent on the motion data; and the experience information plays a critical role for ET and MT. On the other hand, activity TK has the worst performance since the features extracted from the current set of sensors are insufficient as strong indicatives of “talking”. In addition, the features extracted from these sensors overlap significantly with features of other activities. In this particular case, the experience tables do not provide much prior information to assist recognition.

In our experiment, none of the activities reached the threshold T defined in (4) after O_n^1 node were imported into the HMM. Activities meeting the threshold T after O_n^2 was imported are displayed in Table 6. Most WO activities reached the threshold, indicating that the prior experience, motion data and GPS data are powerful in detecting the WO activity. Almost 40 % of the CW and ET activities were recognized without using image features. However, the F_1 value of the TK activity is much lower for the reasons described previously.

The comparison of four subjects with respect to eight activity categories is shown in Fig. 5. In general, the results of CW, WO and RC are better than other activities. The result of subject 3 is the worst among all subjects in ET. Re-examination of data revealed that the eButton was not worn correctly. As a result, food was often missed from the recorded images. Again, the result of subject 2 in HW was inferior mainly because the illumination was insufficient in the environment where HW took place, which affected image quality and consequently the classification accuracy. Interestingly, the result of SP of subject 1 was the best because the subject shopped significantly more than others. As a result, experience tables played an important role in improving accuracy. These observations indicated that both data quality and activity types were key factors influencing activity classification results.

Evaluation of experience tables

In order to evaluate whether experience represents an important part of information for activity recognition, we tested two subsets of features without using experience tables. The first subset was obtained from the motion and GPS sensors while the second was from images. The SVM classifier was used in both cases for activity recognition. The F_1 values of different features with respect to each activity were compared with the F_1 values using experience tables. It can be observed from Fig. 6 that, in all cases, the proposed method performed better than or equivalent to the methods without using experience. The F_1 value of WO reached or was close to 1 in all three methods. However, the results of MT were poor since the features of this activity are similar to that of CW. Most MT activities were classified as CW when only motion sensor data were used. Both F_1 values of ET and SP increased significantly due to the use of experience tables, it is likely that, for the particular individual, most eating activities happened at certain times of morning, noon and evening, and most shopping activities took place at other times. The time table thus played a key role in classifying between ET and SP.

Effects of number of nodes

Since the computational complexity involving images is significantly higher than that without images, we performed an experiment comparing cases with and without images. In the first case (shown in Fig. 7 as the black line), all three nodes (time table, motion/GPS, and image nodes) shown in Fig. 3 were utilized in the adaptive hidden Markov model. While in the second case (dashed line in Fig. 7), the image node was eliminated. Although the 3-node results are better than the 2-node results, the differences in F_1 values are not drastic except for MT and TK. Therefore, if these activities are parts of the classification targets, images should be utilized despite their high computational cost.

Evaluation on feature selection

Feature selection was implemented to improve the accuracy and speed of activity recognition. In Fig. 8, the F_1 values of the proposed method with and without feature selection are compared. In the first case, many inaccurate and redundant features were removed using the method presented previously. Our results demonstrated that, in general, the method using feature selection performed better than the one without feature selection.

Comparison between the proposed method and standard HMM method

The differences between our method and standard HMM method is mainly in the use of the adaptive observation nodes. The adaptive method is advantageous in its high computational efficiency. Figure 9 presents the accuracy rates of the adaptive-HMM and standard HMM methods with respect to 21 days of data recorded from 4 subjects. It can be seen that these two cases are similar. In fact, by assigning score 1 to the method with better performance and 0 to the other method (if the performance of two method were the same, assigning 1 to both), the accumulated score (75) obtained by the standard HMM method was essentially the same as the score (74) of the proposed method. It is understandable that the proposed method could not overperform the standard HMM because the former method observed less amounts of data than the later method in making classification decisions. The slight difference in performance indicated that the data ignored by the proposed method were mostly redundant or irrelevant. As a result, the proposed method was much more computationally efficient than the standard HMM. Figure 10 presents the computational time of the two methods for processing all 21 days, 4-subject data using the Matlab code running on a laptop computer with a 3.3 GHz Core i3 CPU. The adaptive-HMM took 3.25 h to process a full day data on average, which was 25.4 % faster than standard HMM. Although over 3 h of computing is still considerably long, we expect it to be reduced tremendously by re-written our current un-optimized Matlab code into the optimized C code.

Conclusions

An adaptive Hidden Markov Model has been proposed for activity recognition. This model consists of a sequence of states corresponding to the sequence of activities performed. Its novel design features with a utilization of personal experience in the HMM model. Our experimental results demonstrated that personal experience plays an important role in activity recognition with a significant increase in classification accuracy. In addition, the algorithm becomes more stable and is capable of handling many real life events effectively which are otherwise difficult to recognize using the existing methods.

In order to increase the speed of the algorithm, an adaptive Viterbi algorithm is proposed to select features from motion sensor, GPS sensor and image data. Unlike conventional methods which include data from all sensors in the computational process, the observation node in our HMM model is expanded to multiple nodes corresponding to different sensors. In the recognition process, the first node is imported to the proposed model. If data are sufficiently strong to recognize the activity, other nodes are skipped. Otherwise, more nodes are imported sequentially until the termination condition is reached. With this method, most accurate and computationally efficient sensor data are placed in the front of the node list, and more computationally expensive nodes, such as those corresponding to image features, are placed in the back of the list. This approach speeds up the classification process significantly.

Our experimental results demonstrated that the proposed algorithm performs better than other methods without personal experience. The adaptive HMM saved 25.4 % of computation compared to the standard HMM with essentially the same activity recognition accuracy.

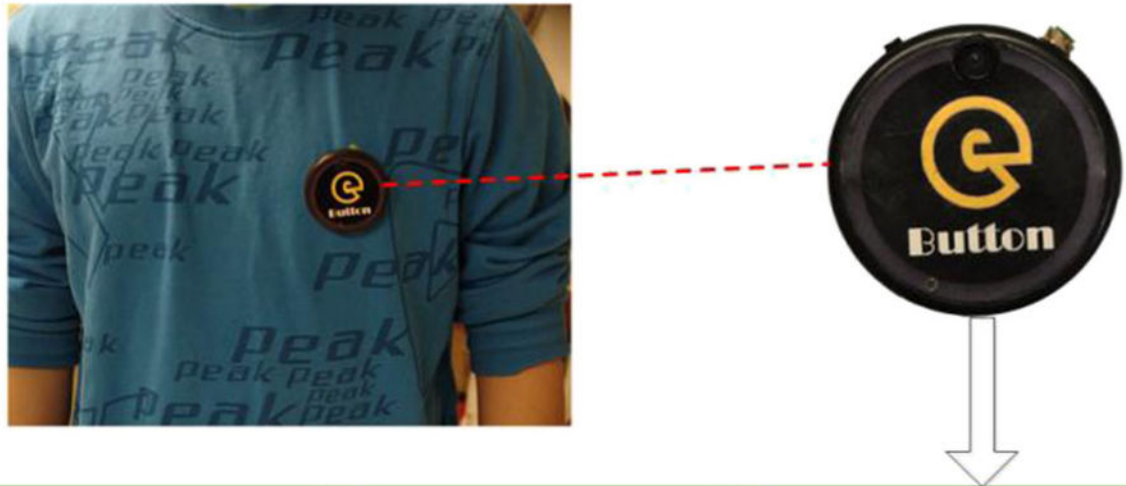
Acknowledgments

This work was supported by National Institutes of Health Grants No. R01CA165255, R21CA172864, and P30 AG024827 of the United States, and the National Natural Science Foundation No. 61402428 of China.

References

1. Touati F, Tabish R. U-healthcare system: state-of-the-art review and challenges. *J Med Syst.* 2013; 37:1–20.
2. Gyllensten IC, Bonomi AG. Identifying types of physical activity with a single accelerometer: evaluating laboratory-trained algorithms in daily life. *IEEE Trans Biomed Eng.* 2011; 58:2656–2663. [PubMed: 21712150]
3. Khan AM, Lee Y-KLY-K, Lee SY, Kim T-SKT-S. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE Trans Inf Technol Biomed.* 2010; 14:1166–1172. [PubMed: 20529753]
4. Gao, L., Bourke, AK., Nelson, J. Activity recognition using dynamic multiple sensor fusion in body sensor networks. 2012 Annu Int Conf IEEE Eng Med Biol Soc; 2012. p. 1077-1080.
5. Chang SY, Lai CF, Chao HC, et al. An environmental-adaptive fall detection system on mobile device. *J Med Syst.* 2011; 35:1299–1312. [PubMed: 21424848]
6. Jin GH, Lee SB, Lee TS. Context awareness of human motion states using accelerometer. *J Med Syst.* 2008; 32:93–100. [PubMed: 18461812]
7. Arab L, Winter A. Automated camera-phone experience with the frequency of imaging necessary to capture diet. *J Am Diet Assoc.* 2010; 110:1238–1241. [PubMed: 20656101]
8. Kwapisz JR, Weiss GM, Moore SA. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explor Newsl.* 2011; 12:74–82.
9. Doherty, AR., Ó Conaire, C., Blighe, M., et al. Combining image descriptors to effectively retrieve events from visual lifelogs. *Proc. 1st ACM Int. Conf. Multimed. Inf. Retr;* 2008. p. 10-17.
10. Chatzichristofis, SA., Boutalis, YS. CEDD: color and edge directivity descriptor. A compact descriptor for image indexing and retrieval. *Proc. 6th Int. Conf. Comput. Vis. Syst;* 2008. p. 312-322.
11. Lowe, DG. Object recognition from local scale-invariant features. *Proc Seventh IEEE Int Conf Comput Vis;* 1999. p. 1150-1157.
12. Duan L, Xu D, Tsang IWH, Luo J. Visual event recognition in videos by learning from Web data. *IEEE Trans Pattern Anal Mach Intell.* 2012; 34:1667–1680. [PubMed: 22201057]
13. Bebars, AA., Hemayed, EE. Comparative study for feature detectors in human activity recognition. 2013 9th Int. Comput. Eng. Conf; 2013. p. 19-24.
14. Hassan, SM., Al-Sadek, AF., Hemayed, EE. Rule-based approach for enhancing the motion trajectories in human activity recognition. 2010 10th Int. Conf. Intell. Syst. Des. Appl; 2010. p. 829-834.
15. Boutell M, Brown C. Pictures are not taken in a vacuum - an overview of exploiting context for semantic scene content understanding. *IEEE Signal Process Mag.* 2006; 23:101–114.
16. Arif M, Bilal M, Kattan A, Ahamed SI. Better physical activity classification using smartphone acceleration sensor. *J Med Syst.* 2014; 38:1–10. [PubMed: 24395031]
17. Yin J, Yang Q, Pan JJ. Sensor-based abnormal human-activity detection. *IEEE Trans Knowl Data Eng.* 2008; 20:1082–1090.
18. Giansanti D, Macellari V, Maccioni G. New neural network classifier of fall-risk based on the Mahalanobis distance and kinematic parameters assessed by a wearable device. *Physiol Meas.* 2008; 29:N11–N19. [PubMed: 18367804]
19. Cao LCL, Luo JLJ, Kautz H, Huang TS. Image annotation within the context of personal photo collections using hierarchical event and scene models. *IEEE Trans Multimed.* 2009; 11:208–219.
20. Hughes G. On the mean accuracy of statistical pattern recognizers. *IEEE Trans Inf Theory.* 1968; 14:55–63.

21. Zhang, W., Jia, W., Sun, M. Segmentation for efficient browsing of chronological video recorded by a wearable device. Proc 2010 I.E. 36th Annu. Northeast. Bioeng. Conf; 2010. p. 1-2.
22. Manjunath BS, Ohm JR, Vasudevan VV, Yamada A. Color and texture descriptors. IEEE Trans Circ Syst Video Technol. 2001; 11:703–715.
23. Csurka, G., Dance, C., Fan, L., et al. Visual categorization with bags of keypoints. Work. Stat. Learn. Comput. Vision ECCV; 2004. p. 1-2.
24. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 2005; 27:1226–1238. [PubMed: 16119262]
25. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. J Mach Learn Technol. 2011; 2:37–63.



Camera Data	Accelerometer Data	Gyroscope Data
GPS Data	Electronic compass	Light Data
Proximity Data	Barometer Data	Thermometer Data

Fig. 1.
eButton and its functional components

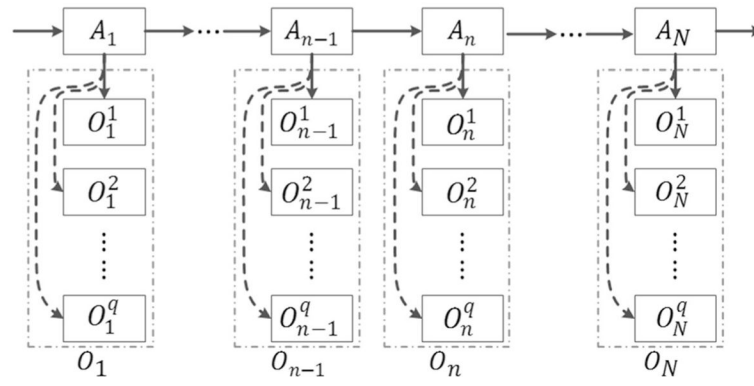


Fig. 2. Adaptive HMM. A_n is the n^{th} of total N activity states to be identified and O_n is the observation vector. Index q in O_n^q represents the specific type of sensor or experience

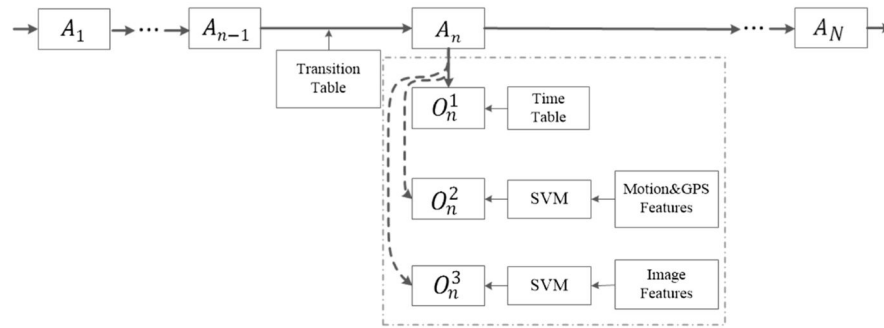


Fig. 3. Details of activity state transition and observation vectors in Fig. 2. The time, motion/GPS sensor, and image elements of the observation vector O_n are denoted by O_n^1 , O_n^2 and O_n^3 , respectively

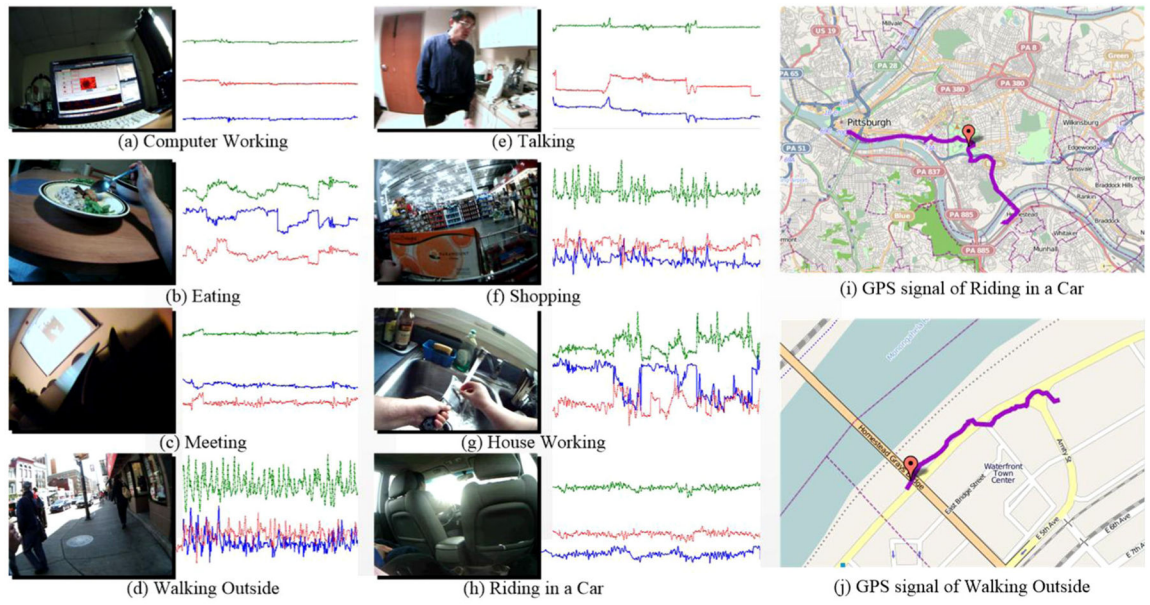


Fig. 4. Data captured by eButton: (a)–(h): image data and accelerometer signal of eight activities; (i) and (j): GPS signals of two outdoor activities

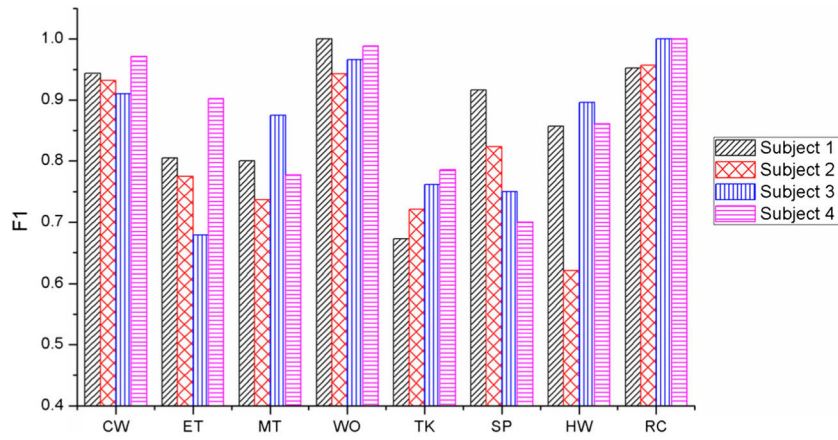


Fig. 5.
 F_1 comparison among subjects

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

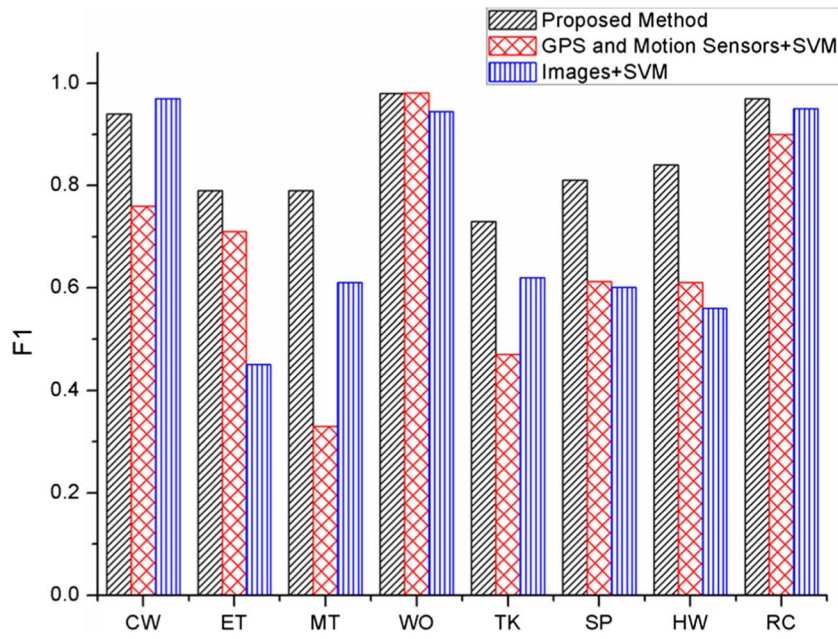


Fig. 6. Performance on methods with experience and without experience

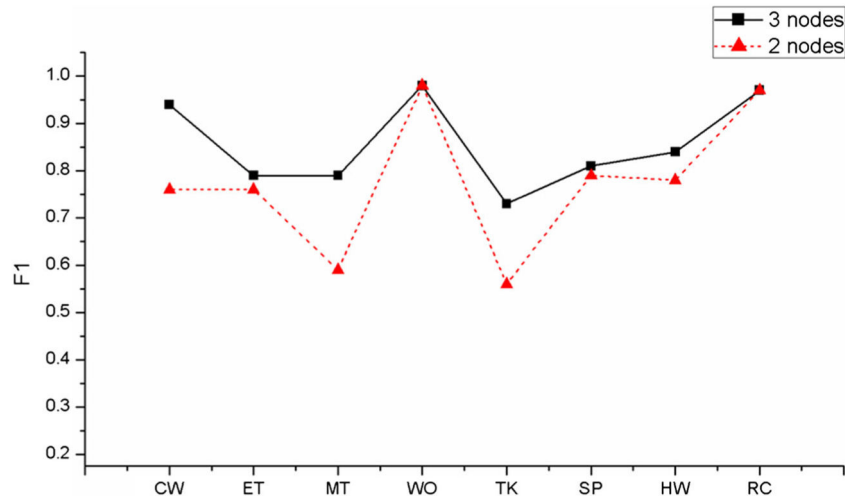


Fig. 7.
Performance on methods with different node numbers

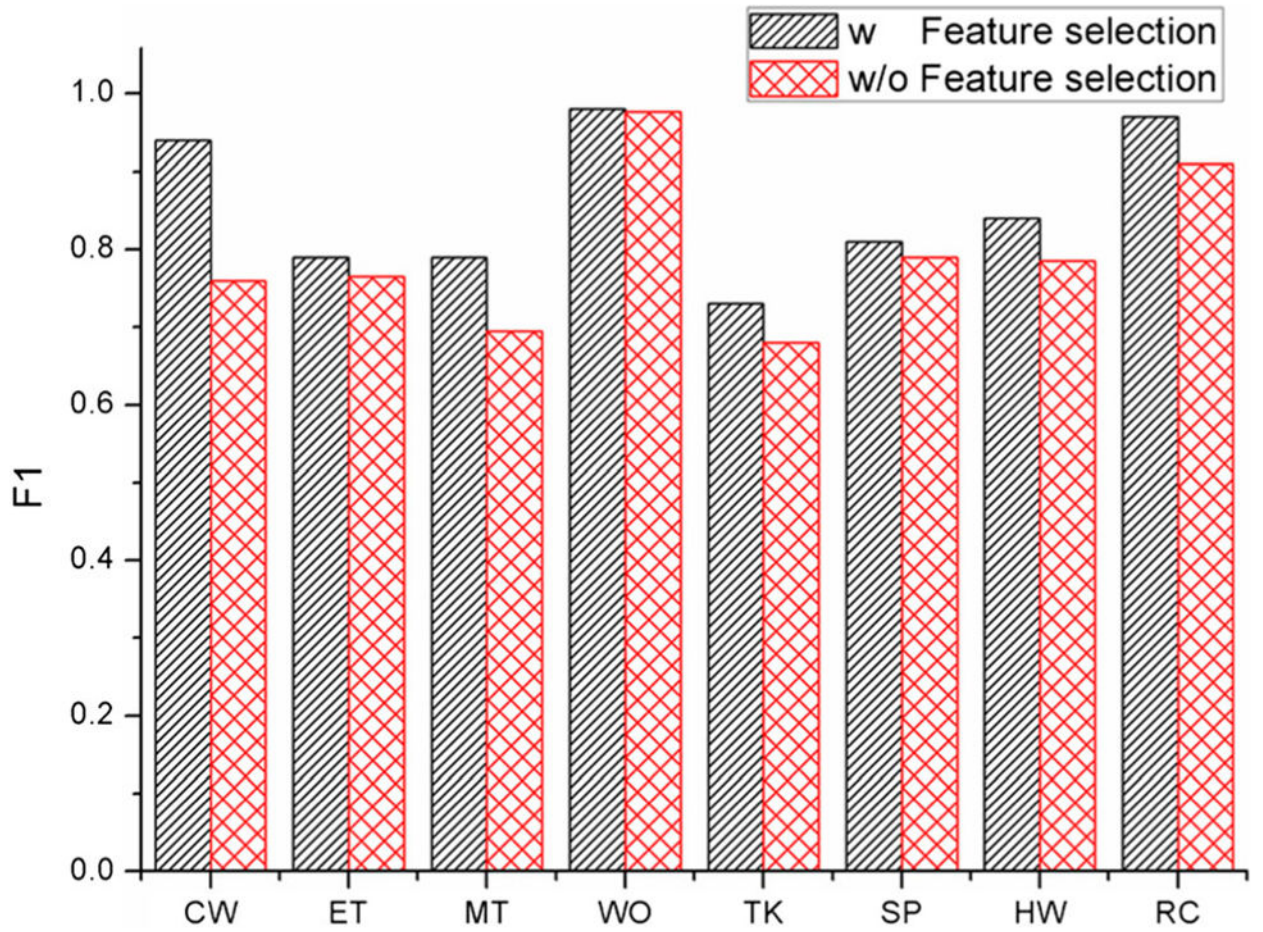


Fig. 8.
Comparison of F_1 values with and without feature selection

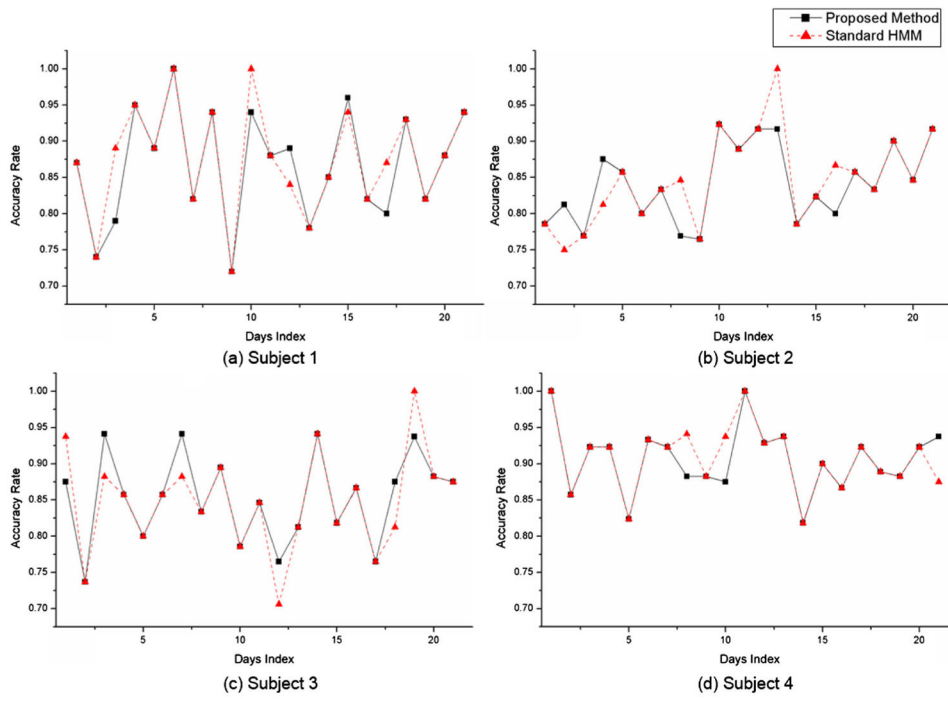


Fig. 9. Performance comparison between proposed method and standard HMM

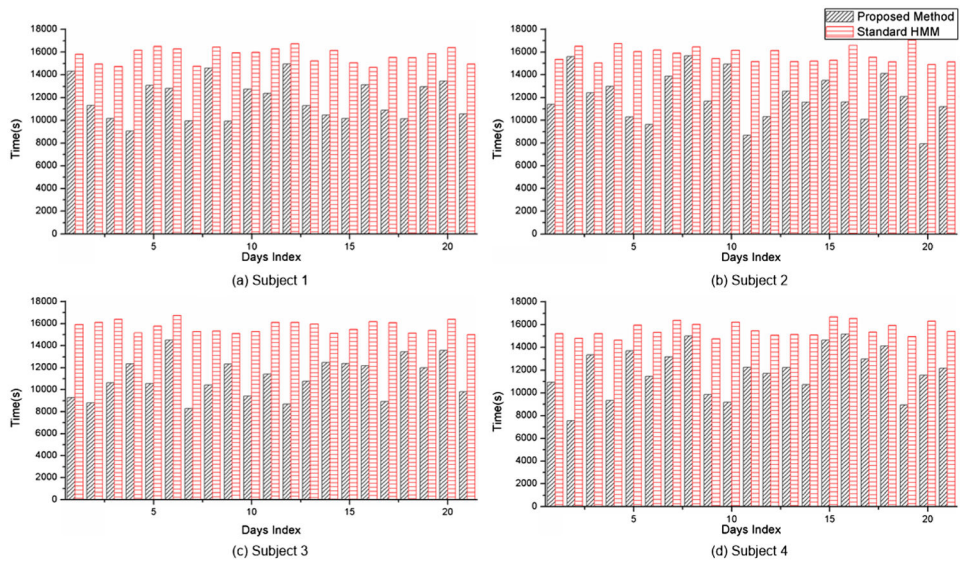


Fig. 10. Computational time consumption of the proposed method and standard HMM

Table 1

Time table

<i>T</i>	<i>A</i>				
	a^1	a^2	...	a^{m-1}	a^m
1 (0:00–0:30)	9	1	...	2	1
2 (0:30–1:00)	9	1	...	2	1
...
20 (9:00–9:30)	4	6	...	2	3
21 (9:30–10:00)	7	8	...	3	4
...
47 (23:00–23:30)	9	1	...	3	1
48 (23:30–24:00)	9	1	...	3	1

Table 2

Transition table

A_p	A	a^1	a^2	...	a^{m-1}	a^m
a^1	L	H	L	...	M	M
a^2	H	L	...	H	M	M
...
a^{m-1}	M	H	...	L	M	M
a^m	M	M	...	L	L	L

L=low transition probability, M=medium transition probability, H=high transition probability

Table 3

Activities for classification

ID	Abbreviation	Activities
1	CW	Computer Working
2	ET	Eating
3	MT	Meeting
4	WO	Walking Outside
5	TK	Talking
6	SP	Shopping
7	HW	Housework(cooking, food preparing)
8	RC	Riding in a Car

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Distribution of test data

Subject	CW	ET	MT	WO	TK	SP	HW	RC	Total
1	111	43	8	57	54	11	73	10	367
2	107	49	10	46	31	7	30	12	292
3	95	60	8	43	55	8	57	8	334
4	101	51	8	41	49	11	36	8	305
Total	414	203	34	187	189	37	196	38	1298

Table 5

Confusion matrix of adaptive-HMM

Classified as	Activity									
	CW	ET	MT	WO	TK	SP	HW	RC		
CW	403	17	5	0	19	0	0	0	0	0
ET	6	154	1	3	19	0	4	0	0	0
MT	0	1	27	0	3	0	3	0	0	0
WO	0	0	0	181	0	3	0	0	0	0
TK	4	14	0	0	130	0	17	0	0	0
SP	0	0	0	0	4	31	4	1	0	0
HW	0	17	1	3	14	3	168	0	0	0
RC	1	0	0	0	0	0	0	37	0	0
F ₁	0.94	0.79	0.79	0.98	0.73	0.81	0.84	0.97	0.81	0.84

Table 6

Activities classified based on experience tables and motion GPS sensor

Classified as	Activity								
	CW	ET	MT	WO	TK	SP	HW	RC	
CW	174	0	0	0	3	0	0	0	0
ET	0	61	0	0	7	0	0	0	0
MT	0	0	0	0	0	0	0	0	0
WO	0	0	0	170	0	0	0	0	0
TK	1	1	0	0	10	0	1	0	0
SP	0	0	0	0	0	7	0	0	0
HW	0	0	0	3	0	0	15	0	0
RC	0	0	0	0	0	0	0	9	0
F ₁	0.99	0.94	N/A	0.99	0.61	1.00	0.88	1.00	1.00