HARF: A Hierarchical Activity Recognition Framework Using Smartphone Sensors

Manhyung Han[1], Jae Hun Bang[1], Chris Nugent[2], Sally McClean[3], and Sungyoung Lee[1]

[1] Department of Computer Engineering, Kyung Hee University (Global Campus), Korea
{smiley,jhb,sylee}@oslab.khu.ac.kr

[2] School of Computing and Mathematics, University of Ulster, Jordanstown, U.K.
cd.nugent@ulster.ac.uk

[3] School of Computing and Information Engineering, University of Ulster, Coleraine, U.K.
si.mcclean@ulster.ac.uk

Abstract. Activity recognition for the purposes of recognizing a user's intentions using multimodal sensors is becoming a widely researched topic largely based on the prevalence of the smartphone. Previous studies have reported the difficulty in recognizing life-logs by only using a smartphone due to the challenges with activity modeling and real-time recognition. In addition, recognizing life-logs is difficult due to the absence of an established framework which enables utilizing different sources of sensor data. In this paper, we propose a smartphone based Hierarchical Activity Recognition Framework which extends the Naïve Bayes approach for the processing of activity modeling and real-time activity recognition. The proposed algorithm demonstrates higher accuracy than the Naïve Bayes approach and also enables the recognition of a user's activities within a mobile environment. The proposed algorithm has the ability to classify fifteen activities with an average classification accuracy of 92.96%.

Keywords: Activity Recognition, Smartphone, Multimodal Sensors, Naïve Bayes, Life-log.

1 Introduction

Activity Recognition (AR) solutions are capable of identifying physical actions, postures or movements using various sensors which have the ability to capture the intention and the condition of the current situation of the subject. Existing AR research has utilized wearable sensors which have been attached on the human's body or 2D/3D cameras for the purposes of capturing video based images [1,2]. A range of studies have also considered the use of mobile phones for the purposes of AR [3,4,5]. In these studies, the data is collected using a mobile phone and the AR is normally performed offline. Thus, the AR algorithms are not implemented in real-time on the phone and the classification is not performed in real-time. In [6], an AR system running purely on a smartphone is presented. The system can be trained on the mobile phone itself

G. Urzaiz et al. (Eds.): UCAmI 2013, LNCS 8276, pp. 159–166, 2013.

and also as the ability to perform the classification in real-time on the phone. The recognition is based on features calculated using a geometric template matching approach and a support vector machine (SVM) is used as the classifier.

There are multiple sensor devices embedded on a smartphone and the smartphone is able to process collected data independently. For recognizing physical movements of the user, accelerometer and gyroscope sensors are utilized. Also for gathering context data such as location, situation or environmental information, data from GPS, proximity sensor and MIC are used. A smartphone also has enough storage space, processing capability and communication modules. So we selected a smartphone as a sensing and processing platform for recognizing human's activities.

General AR is divided into two phase. Firstly a training phase is required to build the activity models and secondly a recognition phase where based on activity models the data collected is proposed. Probability-based AR algorithms for example hidden markov models (HMM), SVM or k-nearest neighbor (kNN) are difficult to apply to smartphone applications given that they requires sample data for modeling in addition to their computational complexity [11].

In this paper, a lightweight activity modeling and recognition framework defined as the Hierarchical Activity Recognition Framework (HARF) which enables the modeling and recognition the user activities on a smartphone is proposed. The proposed HARF has the ability to recognize 15 activities and uses the accelerometer, Gyroscope, Proximity sensor and GPS modules all from the smartphone.

2 Related Works

AR using multimodal sensors on a smartphone has been widely investigated due to the ability of the smartphone to collect various sensor data on a single mobile device. Previous studies have considered recognizing the activities of walking and running using accelerometer data [7,8] and moving activities by transportation using GPS data [9,10]. Nevertheless, the majority of approaches have mainly considered a single sensor and exploited multiple sensor data. To a certain extent this limits the type of activities which can be recognized in addition to the recognition accuracy. In [11], the researches utilizing multimodal sensors in a Smartphone are similar to this paper. They have, however, used HMM and GMM algorithms which require significant computing resources. In [2], the authors proved that utilizing multiple sensor data helps to improve accuracy levels through a combination of accelerometer and audio data.

There are several existing studies for activity modeling and recognition algorithms. In [12], the authors used multiple sensors or heterogeneous sensors for recognizing user's activities by attaching to the body. The approach required a physical connection between the sensors, however, such an approach is not suitable for long-term AR. For accelerometer data classification, several approaches to feature extraction and classification have been investigated [13]. Nevertheless, existing approaches are difficult to apply to mobile devices which have relatively less resources than computers or servers for the initial stages of training.

3 Adaptive Naïve Bayes Algorithm

In this paper the Naïve Bayes algorithm is used as a basic algorithm for recognizing a human's activities. If the activity information of users is matched to the model which has the highest possibilities among pre-constructed activities models, it is chosen by the algorithm. Generally the Naïve Bayes classifier achieves faster modeling time and less computation overheads than other machine learning algorithms. The Naïve Bayes classifier can be generate an activity model quickly, however, it has several limitations such as relatively low processing speed and it is difficult to deploy into a mobile environment which is resource constrained. In addition, one of the inherent characteristics of the Naïve Bayes classifier is that every attribute has the same priority which results in a lower accuracy of posterior probability. In the current work, an Adaptive Naïve Bayes approach is proposed in an effort to address the aforementioned issues.

Naïve Bayes is a statistical classification method which can estimate the possibility of a given sample. The Naïve Bayes probabilistic model assumes that sample data F_1 to F_n have possibilities to relating to an independent class C. The probability of C after the sample data $F_1 \dots F_n$ are collected is $p(C|F_1, \dots, F_n)$ and which is referred to as the posteriori probability. In order to calculate $p(C|F_1, \dots, F_n)$, $p(F_1, \dots, F_n)$ and $p(C)$ are required. These can be estimated from training data and are referred to as the boundary probability. By using Bayes's theorem a posteriori probability is defined as presented in equation (1):

$$p(C|F_1, \dots, F_n) = \frac{p(C)p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)} \tag{1}$$

Only considering a maximization of $p(C)p(C|F_1, \dots, F_n)$ because $p(F_1, \dots, F_n)$ has values to every class. If the boundary probability of the class is now known, only $p(F_1, \dots, F_n|C)$ may be considered. $p(F_1, \dots, F_n|C)$ is calculated by the independent assumption of Naïve Bayes. As a result, $F_1, \dots, F_n$ can be classified as the class which has the largest posteriori probability. If a sample data F_i is a classification attribute and contains one value out of several limited values, a calculation of $p(F_i|C)$ may be made according to traditional probability. However the characteristic of the training data for AR is continuous data. In this case, the distribution of probability is utilized for calculating conditional probability. And the Gaussian distribution is utilized for representing a distribution of F_i.

$$P(F_i = v|C) = \frac{1}{\sqrt{2\pi\sigma_C^2}} e^{-\frac{(v-\mu_C)^2}{2\sigma_C^2}} \tag{2}$$

In equation (2), the mean value of F_i in class C is μ_c and distribution is σ_c^2, In the current work, the proposed approach utilizes multiple sensor data gleaned from a smartphone. Taking this into consideration a lightweight modeling and recognition algorithm is therefore required due to the limited resources. In addition, given the performance issues associated with the Naïve Bayes approach a lightweight classification algorithm A-NB, which enables activity modeling and recognition in a

Smartphone, is proposed. When a system builds an activity model using Naïve Bayes, the complexity of the calculation is dependent on the number of sample data i. If features for AR are increased, it requires significant processing capabilities while calculating the mean value μ_C and the distribution σ_C^2 of data F_i. For overcoming memory overflow which may occur during real-time activity training, A-NB calculates the mean and distribution values of data F_i periodically. The repetitive approach considering memory usage and efficiency is described below:

$$\mu_N = \frac{\mu_{(N-1)} \times (N-1) + F_N}{N},\ v_N = \frac{v_{(N-1)} \times (N-1) + F_N^2}{N}, \sigma_N^2 = v_N - \mu_N^2 \tag{3}$$

In equation (3), N is the number of collected data for time t, μ_N is the mean of data N, v_N is the mean of the square of data N and σ_N^2 is the distribution. If the number of calculated mean and distribution is j, the proposed A-NB approach calculates the total mean value by combining μ_1 to μ_j, the total distribution value by mean value of σ_1^2 to σ_j^2. In order to calculate the $p'(F_i|C)$ value, equation (2) is transformed to equation (4)

$$P'(F_i = v|C) = \frac{1}{\sqrt{2\pi\mu_v}} e^{-\frac{(v-\mu_m)^2}{2\mu_v}} \tag{4}$$

μ_m is the mean value of μ_1 ... μ_j, μ_v is also the mean value of the distribution σ_1^2 ... σ_j^2. Hence the mean and distribution are calculated by data sample F_i, and by using these values the posteriori probability $P'(F_i = v|C)$ are able to can subsequently be calculated.

4 Proposed Hierarchical Activity Recognition Framework

Although the AR using multimodal sensors can increase recognizable activities and enable the recognition various situations, it lowers the accuracy of the overall recognition result given that the classifier is required to consider more factors from input data. In order to overcome this issue, HARF which recognizes activities in hierarchical approach has been proposed. The approach includes the ability to not only recognize a simple act, however, also to consider the spatial location of the user and their activity within a given context. The approach also considers that there may be different meanings associated with the user's location. The activities are categorized into 3 types as presented in Table 1.

Figure 1 depicts the proposed HARF architecture for real-time AR processing based on the A-NB algorithm. If the A-NB is applied to AR, classification is performed firstly using location information and the heuristic approach is applied as described in Table 1. Once recognition is performing, a system recognizes the location first for differentiating indoor (Home and Office) and outdoor. If the user is at unregistered location, the system tries to recognize the current activity among outdoor activities (Walking, Sitting, Standing, Jogging and Riding a car) with heuristic-based approach. However if the user is at registered location, the system firstly look up the location-based activity list. If the user is at the home or office, the system tries to recognize the activity using the proposed approach with multimodal sensor data.

Table 1. Activity categorization for hierarchical activity recognition

Type	Area	Activity	Sensors
Location & multimodal sensor based activity recognition	Home	Walking Sitting Standing	Accelerometer, Gyroscope, Proximity and GPS
	Office	Walking Sitting Standing	
	Outdoor	Walking Sitting Standing Jogging	
Location based activity recognition	Outdoor	Waiting for bus at bus stop Having a meal at cafeteria Exercising at gym Visiting a park	GPS
Heuristic based activity recognition	Outdoor	Riding a car	Accelerometer, Gyroscope, Proximity, GPS and Heuristic

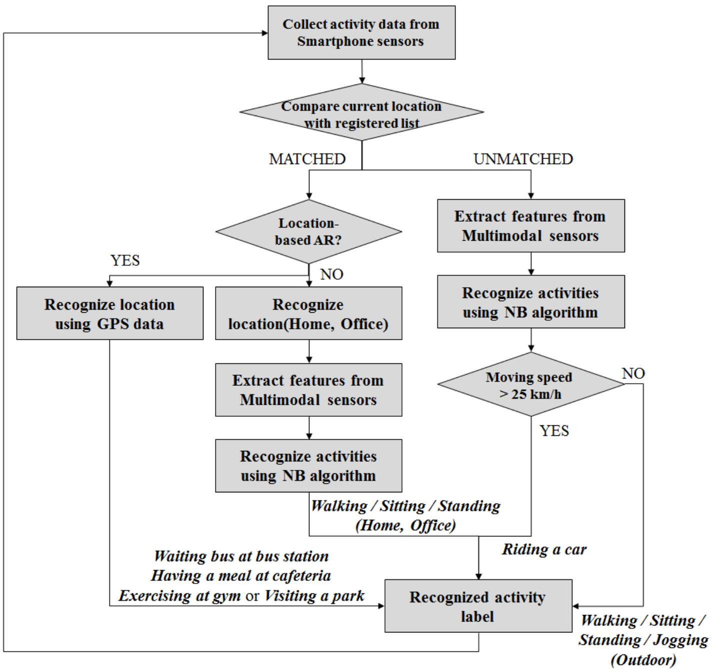

Fig. 1. Hierarchical AR framework for enhancing recognition accuracy

5 Performance Evaluation

For the verification of the A-NB algorithms and the HARF framework, a real-time AR system has been implemented in the form of a Smartphone application using the Android OS. Table 2 presents the results of the AR using the developed Smartphone application. The experiments were conducted on 15 activities including 4 location based activities (Waiting for bus at bus stop, Having a meal at cafeteria, Exercising at gym, Visiting a park). However, the results of recognizing 4 location based activities are not presented. There was only 1 misrecognized case out of 200 testing samples. Nevertheless, if the GPS on the Smartphone is guaranteed to work well, location based activities are well recognized in the proposed system. Therefore, the experimental results in Table 2 present the accuracy of 11 activities without location-based activities.

Table 2. Activity recognition accuracy table of 11 activities for validating proposed HARF

Location		**Home**			**Office**			**Outdoor**				
		Standing	Walking	Sitting	Standing	Walking	Sitting	Standing	Walking	Sitting	Jogging	Car
Home	Standing	**90.32**	-	9.68	-	-	-	-	-	-	-	-
	Walking	10.43	**83.47**	6.1	-	-	-	-	-	-	-	-
	Sitting	2.56	-	**98.44**	-	-	-	-	-	-	-	-
Office	Standing	-	-	-	**95.2**	-	4.8	-	-	-	-	-
	Walking	-	-	-	4.84	**94.35**	0.81	-	-	-	-	-
	Sitting	-	-	-	1.2	0.61	**98.19**	-	-	-	-	-
Outdoor	Standing	-	-	-	-	-	-	**94.34**	-	5.66	-	-
	Walking	-	-	-	-	-	-	12.77	**80.85**	6.38	-	-
	Sitting	-	-	-	-	-	-	2.5	-	**97.5**	-	-
	Jogging	-	-	-	-	-	-	2.17	10.86	1.47	**85.5**	-
	Car	-	-	-	-	-	-	16.25	6.25	1.25	-	**76.25**

The recognition result of 15 activities shows an accuracy of 92.96% and the result of 11 activities, without activities based on only location, is 90.4%. There are several cases which show different accuracy on the same activities. This indicates that the activity can be recognized differently on where the activities took place. For example, walking activities in the home or outdoors are seldom recognized as a standing activity given that the user is frequently stopped for changing a direction. The recognition accuracy of both sitting and standing activities are relatively higher than others because of their static characteristic. In the case of jogging and car, there are some misrecognition results because a jogging activity is similar to walking and a car is frequently stopped or driven slowly.

A performance comparison of the HARF and the Naïve Bayes algorithm is presented in Figure 2. The result of the Naive Bayes and the proposed HARF are 81.17% and 89.88% respectively. In the case of recognizing Car Driving, HARF showed 76% and but the Naïve Bayes showed a lower accuracy of around 50%.

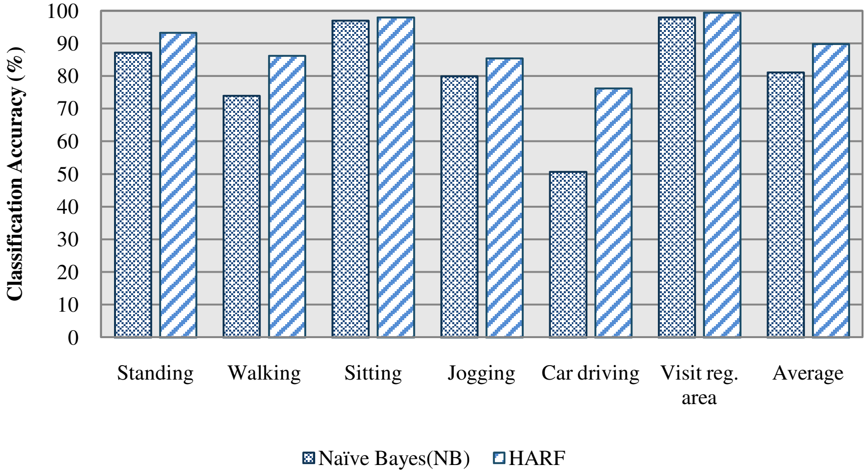

Fig. 2. Accuracy comparison between Naïve Bayes and HARF for 15 activities

6 Conclusion and Future Work

In this paper, we proposed a personalized activity modeling and real-time activity recognition framework for understanding a human's intention or requirements based on multimodal sensors in a Smartphone. A Hierarchical Activity Recognition Framework for modeling and recognizing user activities on resource restricted smartphone was proposed. For compensating a memory overflow error which may occur within the constrained resources of the smartphone when considering data from multimodal sensors, we proposed a hierarchical activity modeling and recognition approach. The testing results show that the proposed system can recognize 15 activities with an accuracy of 92.96%. This is 10.73% higher than using a conventional Naïve Bayes approach.

The proposed HARF exhibited relatively high accuracy when recognizing simple or fixed patterns of activities (jogging, standing). In the case of activities such as riding a car or walking, which may have various patterns, it showed lower accuracy. In order to compensate for these problems, other studies using not only smartphone, however, external sensor devices or utilizing environmental sound are required. A compound approach utilizing multimodal sensor data and external sensor data is expected to enhance the accuracy of activity modeling and recognition.

Acknowledgements. This research was supported by the MSIP (Ministry of Science, ICT & Future Planning), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2013-(H0301-13-2001)).

This work was supported by the Industrial Strategic Technology Development Program (10035348, Development of a Cognitive Planning and Learning Model for Mobile Platforms) funded by the Ministry of Knowledge Economy (MKE, Korea).

References

1. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity Recognition using Cell Phone Accelerometers. ACM SIGKDD Explorations Newsletter 12, 74–82 (2010)
2. Ward, J.A., Lukowicz, P., Troster, G., Starner, T.: Activity Recognition of Assembly Tasks using Body-Worn Microphones and Accelerometers. IEEE Transactions on Pattern Anal. Mach. Intell. 28, 1553–1567 (2006)
3. Brezmes, T., Gorricho, J.L., Cotrina, J.: Activity recognition from accelerometer data on a mobile phone. In: Omatu, S., Rocha, M.P., Bravo, J., Fernández, F., Corchado, E., Bustillo, A., Corchado, J.M. (eds.) IWANN 2009, Part II. LNCS, vol. 5518, pp. 796–799. Springer, Heidelberg (2009)
4. Lu, H., Yang, J., Liu, Z., Lane, N.D., Choudhury, T., Campbell, A.T.: The Jigsaw continuous sensing engine for mobile phone applications. In: Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems, SenSys 2010, pp. 71–84 (2010)
5. Peebles, D., Lu, H., Lane, N.D., Choudhury, T., Campbell, A.T.: Community-guided learning: Exploiting mobile sensor users to model human behavior. In: Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15 (2010)
6. Frank, J., Mannor, S., Precup, D.: Activity recognition with mobile phones. In: Gunopulos, D., Hofmann, T., Malerba, D., Vazirgiannis, M. (eds.) ECML PKDD 2011, Part III. LNCS, vol. 6913, pp. 630–633. Springer, Heidelberg (2011)
7. Khan, A.M., Lee, Y.K., Lee, S.Y., Kim, T.S.: A Triaxial Accelerometer-Based Physical-Activity Recognition Via Augmented-Signal Features and a Hierarchical Recognizer. IEEE Transactions on Information Technology in Biomedicine 14(5), 1166–1172 (2010)
8. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity Recognition from Accelerometer Data. In: Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence, pp. 1541–1549 (2005)
9. Liao, L., Fox, D., Kautz, H.: Extracting Places and Activities from GPS Traces using Hierarchical Conditional Random Fields. Int. J. Rob. Res. 26, 119–134 (2007)
10. Vinh, L.T., Lee, S.Y., Park, Y.T., d'Auriol, B.: A Novel Feature Selection Method Based on Normalized Mutual Information. Appl. Intell. 37, 100–120 (2012)
11. Han, M., Vinh, L.T., Lee, Y.-K., Lee, S.: Comprehensive Context Recognizer Based on Multimodal Sensors in a Smartphone. Journal of Sensors 12(9), 12588–12605 (2012)
12. Minnen, D., Starner, T., Ward, J., Lukowicz, P., Troester, G.: Recognizing and discovering human actions from on-body sensor data. In: Proc. IEEE Int. Conf. Multimedia Expo, pp. 1545–1548 (2005)
13. Shen, K.Q., Ong, C.J., Li, X.P., Wilder-Smith, E.P.V.: Novel Multi-Class Feature Selection Methods using Sensitivity Analysis of Posterior Probabilities. In: Proc. of the IEEE International Conference on Systems, Man and Cybernetics, pp. 1116–1121 (2008)