ADL Classification based on Autocorrelation Function of Inertial signals

Walid Gomaa[1,2], Reda Elbasiony[1], and Sara Ashry[1,3]

[1]Cyber-Physical Systems Lab.,
Egypt-Japan University of Science and Technology University, Alexandria, Egypt.
[2]Faculty of Engineering, Alexandria University, Alexandria, Egypt.
[3]Electronic Research Institute (ERI), Cairo, Egypt.
[1]walid.gomaa, reda.elbasiony, sara.ashry@ejust.edu.eg

***Abstract*—Recognition of human activities is one of the most promising research areas in artificial intelligence. This has come along with the technological advancement in sensing technologies as well as the high demand for applications that are mobile, context-aware, and real-time. In this paper, we use a smart watch to collect sensory data for 14 ADL activities. We collect three types of sensory signals: acceleration, angular velocity, and rotation displacement; each is a tri-axial signal. From each given signal we compute the autocorrelation function up to certain lag and take these computed values as representative features of the given signal. We then feed these features to a random-forest-based classifier for training and prediction. We experiment with different combinations of sensory data. The joint use of acceleration with angular velocity has achieved the best performance in prediction accuracy which reaches about 80% for the whole set of 14 activities.**

1. Introduction

Making our everyday environment intelligent is an old dream that is about to be real. Ambient intelligence becomes one of the most promising research topics [1]. Human activity modeling and recognition are the backbone of this research trend. They are fundamental topics for both smart environments and surveillance applications [2].

Activities of daily living (ADL) are the activities normal people usually learn in childhood and can do them every day without needing help. Feeding, toileting, bathing, dressing, transferring, and personal grooming are considered as the basic ADLs according to Katz et al. in [3].

Smart homes employ various types of sensors such as cameras of all their types, microphones, and motion sensors in order to recognize human activities while people perform them and try to compose them into a more complex patterns of actions [4]. Although this method can detect complex actions efficiently, it has many limitations regarding its fixed nature, for example, it is not effective if the user decided to leave the place, to be hidden from fixed sensors, or to do any activities which are not detectable by these sensors.

The other approach depends on wearable mobile-sensors instead of fixed environmental ones. The measurements are taken from mobile sensors mounted to human body parts like wrists, legs, waist, and chest. Wearable accelerometers are widely used in this area, because of being small-sized, inexpensive, and naturally embedded in most of smart mobile devices [4]. Typically, action recognition requires building computational models like Gaussian mixture modeling and regression (GMM and GMR) which are widely used.

In the current paper, we propose using the random forest classification algorithm without any preprocessing on the raw data. The input features are simple, they are extracted by computing the autocorrelation function of the input accelerometer signals. The random forests are much more stable and computationally efficient than GMM.

Table 1: The selected ADLs and related motion primitives

ADL	Motion Primitives (s)
Communication	Use telephone
Feeding	Drink from glass, Pour water into glass, Eat with knife and fork, Eat with spoon
Transferring	Climb stairs, Descend stairs, Walk, Get up from bed, Lie down on bed, Stand up from chair, Sit down on chair
Personal grooming	Brush teeth, Comb hair

From several years, many devices have been used to collect acceleration data from users. Previously, big and heavy devices have to be used to collect data such as data acquisition cards (DAQs) connected to accelerometers in order to convert analog acceleration signals into digital form and send it to a PC to record them [5]. In the recent few years, smart mobile phones have become very popular and naturally equipped with accelerometers and many other sensors [6]. Thus, it became easier to use them in this field. However, it is still uncomfortable to carry or wear them everywhere specially while sleeping or showering. Recently, smart watches have become the most ideal wearable devices to be used for the activity recognition purpose. They are small, light, and can be worn in most different situations. Figure 1 shows the smart watch.

The rest of the article is organized as follows. Section 1 is an introduction. Section 2 presents the related work. Section 3 illustrates our data collection process

0-7695-6321-X/17/31.00 ©2017 IEEE
DOI 10.1109/ICMLA.2017.00-53

approach. Experimental results are discussed in Section 4. Section 5 concludes the paper.

Figure 1: The smart watch

2. Related work

In regards to human activity recognition, two learning approaches are mainly used: supervised and unsupervised learning. In supervised learning, classification algorithms are used to infer special patterns for each type of human activity based on the collected labeled training data. Then, the trained classification model is used later in the activity recognition process. Many models have been used such as decision trees with c4.5 algorithm [7]; instance based learning such as K-NN [8]; support vector machines (SVM) [9]; artificial neural networks (ANN) [10]; ensembles of classifiers such as random forests [11]; deep learning such as Deep convolutional neural network (CNN) [12] by using the gyroscope, total acceleration, and linear acceleration signals.

In unsupervised learning approaches, algorithms are used to directly discover similarity between data instances. Then unlabeled groups of similar activities are constructed based on a suitable similarity measurement. Some famous algorithms are used in this type such as k-Means [13], Gaussian mixture models (GMM) [14], and hidden Markov models (HMM) [15]. Both the GMM and the HMM approaches use the expectation maximization algorithm.

Feature selection in this context has been studied also. The authors in [16] have proved that selecting the suitable features for each activity can improve both the recognition accuracy and the computational efficiency. This is because removing the useless features which confuse the classifier can improve the accuracy and reduce the number of features resulting in significant computational saving.

The count and locations of sensors are crucial design parameters in any accelerometer-based activity recognition system. With regard to the selected data and the sensors placement on the human body, the authors in [17] have proposed a technique which is tested on acceleration data collected from three sensors placed at the chest, the right thigh and the left ankle of the volunteers. However, they restrict the mobility and the freedom of the person. This drawback is solved in our approach by using only Apple Watch Series 1 at the wrist. The authors in [18] used spectrogram signals of the tri-axial accelerometer instead of the raw data to learn deep activity recognition models. Our proposed approach use the raw signals without any preprocessing and it has achieved high acceptable accuracy.

In this work we use the data of the tri-axial accelerometer and the tri-axial gyro streamed from the smart watch. The used attributes are: 3D user acceleration, 3D Rotation, and 3D angular velocity for 14 human motion primitives, which are proven to give high accuracy as shown in Figure 5.

3. Proposed approach

3.1. Human motion primitives

Our study focuses on the "Activities of daily living", which are too generic to be completely modeled and recognized. So, we selected some low level activities called "motion primitives," where each motion primitive uniquely identifies a specific ADL. Table 1 shows the selected 14 motion primitives which correspond to 4 different ADLs.

3.2. Instrumentation

We use Apple smart watch, its case weight is only about 30 gm, and water resistant. It contains Dual-Core processor, 8 GB internal storage, and 512 MB RAM. In regards to motion sensors, it contains an ST Microelectronics 6 axis sensor for acceleration and roll, pitch, and yaw (gyroscope). Thus, smart watch is ideal for motion data sensing, acquisition, and storage. It can be naturally worn while doing almost all ADLs including showering and sleeping without any uncomfortable feeling.

3.3. Data collection

To evaluate the proposed method, we collected a new dataset[1] which contains the activities shown in Table 1. Data are collected from volunteers (females and males). Each volunteer repeats the same activity about 20 times on average. Volunteers ages range from 25 to 40 years old and with weights range from 50 to 90. Every volunteer wears an Apple watch on the right wrist and records the signals of the fourteen activities. The raw signals for each activity contain information about 3D orientation (roll, pitch, and yaw), 3D angular velocity, 3D gravity components, and 3D user acceleration (x, y, z). The sampling rate is chosen to be 50 Hz. Figures 3, 4 show samples of accelerometer and gyro signals.

3.4. Activities description

The chosen motion primitives can be categorized according to two factors. The first factor is the involved body limbs, where the motions 'use the telephone', 'drink from glass', 'pour water into glass', 'eat with knife and fork', 'eat with spoon', 'brush teeth', and 'comb hair' depend only on the motion of the right arm, while the other motions depend on the full body. The second factor is the recurrence of the motion, where the motions 'eat with knife and fork', 'eat with spoon', 'climb stairs', 'descend stairs', 'walk', 'brush

1. The collected dataset and its description are available at: https://www.researchgate.net/publication/318108745_Inertial_data_collected_from_iWatch. DOI: 10.13140/RG.2.2.11101.10722

teeth', and 'comb hair' are recursive motions, while the other motions are non-recursive. fig. 2 shows a volunteer while performing some motion primitives.

Figure 2: It shows the activities and data collection using smart watch

3.5. Modeling features

In our work, we use the time-domain raw signals of both the tri-axial accelerometer and gyro to calculate feature vectors. The used raw signals are *3D acceleration*, *3D rotation*, and *3D angular velocity*. For each time-series signal we calculate an estimated values for the auto-correlation function acf up to a certain lag. We have chosen that lag value that gives the best accuracy. These estimated values are taken as input features to a random forest classifier for training to predict the activities corresponding to test signals. The auto-correlation function is estimated as follows as a function of the lag h:

$$
\begin{aligned}
a\hat{c}f(h) &= \frac{\hat{\gamma(h)}}{\hat{\gamma(0)}} \\
\hat{\gamma(h)} &= \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h}-\bar{x})(x_t-\bar{x})
\end{aligned} \tag{1}
$$

where $\bar{x}$ is the sample mean. The time-series is assumed to be weakly stationary. We perform six different experiments, each differs by the type of sensory information used to train the classifier. For example, in the first experiment we only use the tri-axial accelerometer signals. For each axis we estimate the auto-correlation function up to lag $h = 10$, thus we have 11 estimated values for the auto-correlation function for one axis. So for the tri-axial signals we have a total of 33 estimated values representing a 33 dimensional feature spaces to be input to the random forest classifier.

3.6. Random forests

Random forests (RF) [19] is an ensemble learning approach. It is used for both classification and regression. It depends on constructing many decision trees and aggregating their decisions. In the training phase, each tree is trained over a randomly selected subset of the overall

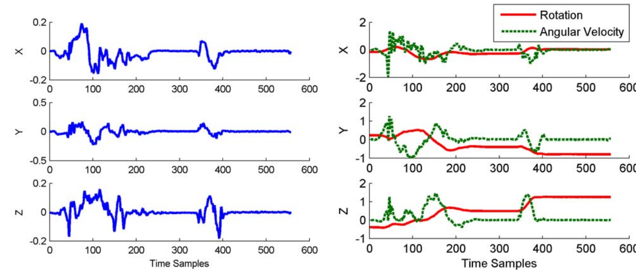

Figure 3: The left figure shows the data sample of 3D acceleration (g) signals of the 'get up from the bed' while the right figure shows the 3D rotation (rad) and angular velocity (rad/s) signals of the same activity.

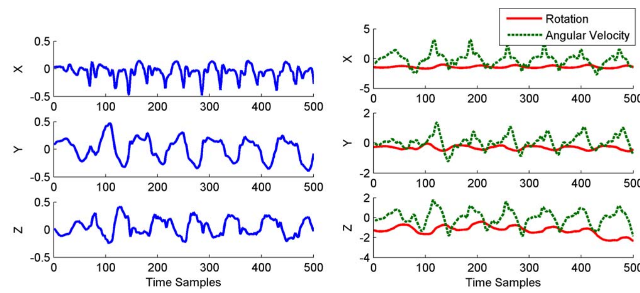

Figure 4: The left figure shows the data sample of 3D acceleration (g) signals of the 'walk' while the right figure shows the 3D rotation (rad) and angular velocity (rad/s) signals of the same activity.

training dataset (in-bag) and validated against the rest part of the dataset (out-of-bag). In the prediction phase, the output is the class which achieves the maximum number of trees-votes in classification, or the average of the trees-output values in regression. The advantage of multiple voting in RF makes it more robust and naturally providing distributed representation, it has much more capability of separating the activities distributions over the selected feature space, making it more robust and transparent to the particular hand-engineered feature space. In our work we have used RF for classifying the 14 activities. We have used 5-fold cross-validation repeated five times in order to optimize the hyper-parameter space. These hyper-parameters include: 'mtry' (the number of features randomly sampled as candidates at each split) and 'ntree' (the number of trees to grow). We use 65% of the data recorded for the training Phase and the other 35% for the testing phase.

4. Experimental Results

4.1. Results

We performed 6 tests to validate our approach, each test uses different set of sensory data and corresponding features as illustrated in Table 2. Each test includes 13

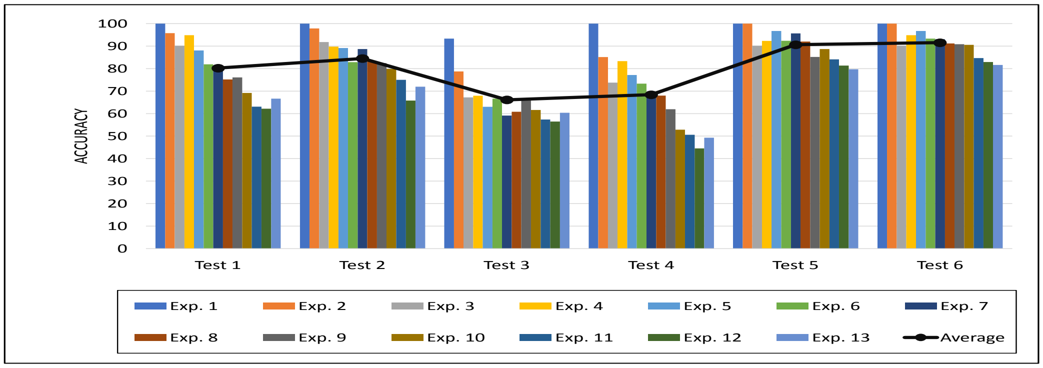

Figure 5: The accuracy of all the experiments applied on different setup categories shown in Table 2 and Table 3

Table 2: Experiments Setup

Test Number	Used Features	Lag Value for acf
Test 1	3D Acceleration	10
Test 2	3D Angular Velocity	5
Test 3	3D Rotation	15
Test 4	Magnitude of the 3D Acceleration	5
Test 5	3D Acceleration, 3D Angular Velocity	15
Test 6	3D Acceleration, 3D Angular Velocity, 3D Rotation	10

Table 3: Motion primitives in each experiment

Experiment/ Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
Climb stairs	X	X	X	X	X	X	X	X	X	X	X	X	X
Brush teeth	X	X	X	X	X	X	X	X	X	X	X	X	X
Comb hair		X	X	X	X	X	X	X	X	X	X	X	X
Descend stairs			X	X	X	X	X	X	X	X	X	X	X
Drink from glass				X	X	X	X	X	X	X	X	X	X
Eat with knife & fork					X	X	X	X	X	X	X	X	X
Eat with spoon						X	X	X	X	X	X	X	X
Get up from bed							X	X	X	X	X	X	X
Lie down on bed								X	X	X	X	X	X
Pour water into glass									X	X	X	X	X
Sit down on chair										X	X	X	X
Stand up from chair											X	X	X
Use the telephone												X	X
Walk													X

experiments, in each experiment we try to classify different number of motion primitives as illustrated in Table 3. In each experiment we add one more activity to the classification task. The choice of the added activity is done at random. Figure 5 plots the accuracy of all the experiments.

4.2. Discussion

As shown in figure 5, we calculated the average accuracy for the 13 experiments of each test. The lowest average accuracy is achieved by Tests 3 and 4 (66.1% and 68.4% respectively). Tests 1 and 2 achieve moderate accuracy average (80.2% and 84.5% respectively). Finally tests 5 and 6 achieved the highest accuracy average (90.6% and 91.5% respectively). Thus, this proves that the output accuracy is highly affected by the sensory information used for the classification. Generally, as more sensory data are used the better the achieved accuracy.

When we depend only on one signal to perform the classification process, 3D angular velocity is considered as the best candidate by achieving the highest discrimination between classes with the best average accuracy of 84.5% (Test 2), then acceleration is the second with average accuracy of 80.2% (Test 1), then the 3D Rotation is the lowest 66.1% (Test 3). However, using combinations of sensory signals generally improves the accuracy. This is noticed in Test 5, where both the 3D acceleration and 3D angular velocity are used together to achieve average accuracy 90.6%. Finally, although the use of the 3D rotation signal alone achieved the worst accuracy, using it in combination with the other 2 signals (3D acceleration and 3D angular velocity) has given the best accuracy across all tests with an average accuracy of 91.5%.

To compare our results to the previous work, we use the sensitivity and specificity as an evaluation metric as shown in table 4. Authors in [20] evaluated their proposed approach using a set of 8 motion primitives which are very similar to our proposed motion primitives. They achieved average sensitivity of 66.5% and average specificity of 84.5%. Compared to our results in Test 6, as shown in Table 4, the smallest average sensitivity and specificity achieved by our method are 81% and 98% respectively which are much higher than the results in [20]

5. Conclusion and Future Work

In this paper we have developed a simple, yet very effective, classification scheme for activities of daily living. We have collected several streams of sensory inertial data from a wearable smart watch. It is a very natural easy, non-intrusive way of data collection. The streams are collected for a set of 14 activities. From the collected timer series signals we estimate the auto-correlation function up to a certain (relatively low) lag, and use these values as input features to a random forest classifier. We measured the performance in terms of the overall accuracy as well as

Table 4: Sensitivity of all the experiments in Test 6

Sensitivity	Climb stairs	Brush teeth	Comb hair	Descend stairs	Drink (glass)	Eat (fork)	Eat (spoon)	Get up from bed	Lie down on bed	Pour water	Sit down on chair	Stand up from chair	Use the telephone	Walk	AVG
Exp.1	1	1													1
Exp.2	1	1	1												1
Exp.3	0.9286	1	1	0.6429											0.89
Exp.4	0.7857	1	0.9412	1	1										0.95
Exp.5	0.9286	0.9375	1	0.9286	1	1									0.96
Exp.6	0.7857	1	1	0.9286	0.9412	0.8571	1								0.93
Exp.7	0.78571	0.9375	1	1	0.9412	1	0.84615	0.9							0.92
Exp.8	0.7143	1	0.9412	0.7857	1	1	0.9231	0.9	0.9						0.9
Exp.9	0.85714	1	0.9412	0.92857	0.8824	0.92857	0.92308	0.9	1	0.76471					0.91
Exp.10	0.71429	1	1	0.92857	0.82353	1	1	0.7	0.8	0.9412	0.9412				0.89
Exp.11	0.85714	1	1	0.78571	0.82353	0.92857	1	0.8	0.8	0.88235	0.64706	0.64706			0.85
Exp.12	1	1	1	0.78571	0.76471	1	0.76923	1	0.7	0.88235	0.88235	0.64706	0.41176		0.83
Exp.13	0.57143	1	1	0.78571	0.58824	0.92857	1	0.8	0.6	0.88235	0.64706	0.76471	0.82353	1	0.81

the sensitivity and specificity of each activity. Our worst performance in all our experiments exceed 80%, indicating that the auto-correlation function is a good discriminator for such activities.

In the future, we plan to improve the trained models by adding new motion primitives to the training dataset. Also, we aim at detecting and recognizing motion primitives in a continuous streams of such primitives employing methods such as Markov regime switching. We also plan to implement our approach on smart phones and watches for real-time classification.

References

[1] N.-Y. Chong and F. Mastrogiovanni, *Handbook of Research on Ambient Intelligence and Smart Environments: Trends and Perspective*. Information Science Reference, 2011.

[2] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, “Activity recognition from accelerometer data,” in *Aaai*, vol. 5, no. 2005, 2005, pp. 1541–1546.

[3] S. Katz, A. Chinn, and L. Cordrey, “Multidisciplinary studies of illness in aged persons ii: a new classification of functional status in activities of daily living,” *J Chronic Dis*, vol. 9, no. 1, pp. 55–62, 1959.

[4] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, “Human motion modelling and recognition: A computational approach,” in *Automation Science and Engineering (CASE), 2012 IEEE International Conference on*. IEEE, 2012, pp. 156–161.

[5] J. Mantyjarvi, J. Himberg, and T. Seppanen, “Recognizing human motion with multiple acceleration sensors,” in *Systems, Man, and Cybernetics, 2001 IEEE International Conference on*, vol. 2. IEEE, 2001, pp. 747–752.

[6] S. Zhang, P. McCullagh, C. Nugent, and H. Zheng, “Activity monitoring using a smart phone’s accelerometer with hierarchical classification,” in *Intelligent Environments (IE), 2010 Sixth International Conference on*. IEEE, 2010, pp. 158–163.

[7] O. D. Lara and M. A. Labrador, “A mobile platform for real-time human activity recognition,” in *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*. IEEE, 2012, pp. 667–671.

[8] C.-L. Liu, C.-H. Lee, and P.-M. Lin, “A fall detection system using k-nearest neighbor classifier,” *Expert systems with applications*, vol. 37, no. 10, pp. 7174–7181, 2010.

[9] S. Weng, L. Xiang, W. Tang, H. Yang, L. Zheng, H. Lu, and H. Zheng, “A low power and high accuracy mems sensor based activity recognition algorithm,” in *Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on*. IEEE, 2014, pp. 33–38.

[10] H. Xu, J. Liu, H. Hu, and Y. Zhang, “Wearable sensor-based human activity recognition method with multi-features extracted from hilbert-huang transform,” *Sensors*, vol. 16, no. 12, p. 2048, 2016.

[11] H. Qian, Y. Mao, W. Xiang, and Z. Wang, “Recognition of human activities using svm multi-class classifier,” *Pattern Recognition Letters*, vol. 31, no. 2, pp. 100–111, 2010.

[12] W. Jiang and Z. Yin, “Human activity recognition using wearable sensors by deep convolutional neural networks,” in *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 2015, pp. 1307–1310.

[13] R. Duda, P. Hart, and D. Stork, “Pattern clasification, johon wiley & sons,” *Inc., New York*, 2000.

[14] A. Mannini and A. M. Sabatini, “Machine learning methods for classifying human physical activity from on-body accelerometers,” *Sensors*, vol. 10, no. 2, pp. 1154–1175, 2010.

[15] J. F.-S. Lin and D. Kulic, “Automatic human motion segmentation and identification using feature guided hmm for physical rehabilitation exercises,” in *Robotics for Neurology and Rehabilitation, Workshop at IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011.

[16] T. Huynh and B. Schiele, “Analyzing features for activity recognition,” in *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*. ACM, 2005, pp. 159–163.

[17] D. Trabelsi, S. Mohammed, F. Chamroukhi, L. Oukhellou, and Y. Amirat, “An unsupervised approach for automatic activity recognition based on hidden markov model regression,” *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 3, pp. 829–835, 2013.

[18] M. A. Alsheikh, A. Selim, D. Niyato, L. Doyle, S. Lin, and H.-P. Tan, “Deep activity recognition models with triaxial accelerometers.” in *AAAI Workshop: Artificial Intelligence Applied to Assistive Technologies and Smart Environments*, 2016.

[19] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[20] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, “Analysis of human behavior recognition algorithms based on acceleration data,” in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1602–1607.