Human Activity Recognition using Smartphone Sensors with Context Filtering

Shah Md. Shihab Hasan, Mohshi Masnad, Md. Mohiuddin Khan, Hasan Mahmud, Md. Kamrul Hasan
Systems and Software Lab (SSL), Department of Computer Science and Engineering
Islamic University of Technology (IUT), Dhaka, Bangladesh.
email: shihab11@iut-dhaka.edu, aurkoiut@gmail.com, {mohiuddin, hasan, hasank}@iut-dhaka.edu

***Abstract*—In recent times, application of Ambient Intelligent services, e.g. smart home, remote monitoring and assisted healthcare, the use of smart phones for the recognition of human activities has become a topic of high interest. Simple activities like sitting, running, walking can be recognized easily but semi-complex activities like ascending and descending stairs, slow running, jogging, fast running etc. are often difficult to recognize accurately. We aim to reduce the error rate of recognizing these kinds of activities by applying Dynamic Time Warping (DTW) algorithm and introducing context filtering. We used heart rate data and barometric pressure sensor data as elements of context filtering. We used a steady state as template and matched every activity with this steady state. To get optimum threshold values, we applied K Nearest Neighbor (KNN) algorithm on the score of DTW. After primarily classifying activities, we used the context filtering approach to further recognize activities by removing confusions. After completion of our study, we have seen that accuracy level has increased significantly for differentiating similar kinds of activities. Overall, our novel approach of applying DTW algorithm and applying context filtering shows considerable performance improvements at a low cost.**

Keywords-Human Activity Recognition (HAR), context filtering, DTW.

I. INTRODUCTION

In recent years, Human Activity Recognition (HAR) through smart phones became a well known field of research. As we have entered the era of intelligent environment, the automated detection of Activity has become a point of high interest. Intelligent environments generally exploit information gathered from users and their environments in order to produce an appropriate action [16]. In this regard, different studies have been conducted in this field. Based on these studies, we observed that basic locomotion activities like Walking, Running, Sitting, Lying on bed can be detected with good accuracy rate [15].However, similar activities, such as Going upstairs or downstairs, Slow running (Jogging), Fast Walking can not be detected perfectly [15]. It is important to detect these activities for development of systems that promote the improvement of people's Quality of Life (QoL) through the *recognition of human activities*. Especially, our focus is on differentiating and detecting similar kinds of activities: slow walking, fast running, going upstairs and going downstairs. We use DTW algorithm to recognize these activities from the inertial sensors available in smart phones. Moreover, we equipped the users with heart rate sensors and took the barometric pressure sensor data. These two factors are used to further improve the activity of the recognition. The latter factors are termed "context filters".

From the start of 21st century, HAR became a point of very high interest. In the beginning, it was confined to only surveillance using video camera but now it is incrementally used for different areas of study directly related to HAR. Ambient Intelligence (AMI) and Ambient Assisted Living (AAL) are two areas of study that are directly related to HAR.AMI is a system mainly focusing on environments that behave intelligently based on the actions of users associated with the system. These systems are unique in many ways, e.g. systems that intelligently anticipate the user needs based on information (e.g. activity patterns, past events and their solutions). These environments need seamless interaction between user and system. Smart Homes and Smart Cities are examples and concepts of such systems. Another important field is Ambient Assisted Living. This is accomplished by incorporating different systems together into a health monitoring system for elder people. Due to enormous advancement in Medical Science, people are living longer. One study based in Europe showed that by 2060 the elderly (namely people over 65 years) will be near 30% of the entire population as opposed to a 17% by 2010 [18]. In addition to the elderly, about 15% of the total population has some kind of disability (WHO 2011) [19]. So, there is a large area where AAL can bring numerous benefits. These AMI and AAL systems need continuous information from the user and from the environment. Detection of human activity can work as input to these kinds of systems. We aim to detect these activities unobtrusively by exploiting the use of smart phones as input to these kinds of systems.

The rest of this paper is organized as follows. Section 2 gives an overview of different approaches for the activity recognition problem. It also describes the reasons of choosing DTW algorithm as the classifier to classify and recognize activity. Section 3 proposes a solution to recognize human activity with further accuracy. Section 4 includes description of the implementation of the proposed methods and also contains the evaluation of proposed methodology. Section 5 presents our result analysis and comparison with different studies. Finally, Section 6 concludes the paper.

II. BACKGROUND STUDY

In HAR, different types of studies have been conducted so far:

A. External or Environmental Sensors

From the very beginning of activity recognition research, video cameras were the first choice. Video cameras were employed in Poppe [13] for marker-less vision-based human motion analysis. Apart from video cameras, Bian et al. [14] used microphones for sound source localization in a home environment for communication activity inference. Taka et al. [3] used a Microsoft Kinect's depth sensor as an ambient sensor for position and orientation tracking for an indoor monitoring system for Parkinson's disease (PD) patients. These kinds of systems require a static infrastructure which limits their range of operation to a constrained space. So for a static environment, such as in a room it gives satisfactory performance but for dynamic environments it is less useful.

B. Wearable Sensors

Wearable sensors are commonly attached to different body parts such as the waist, wrist , chest, legs and head, as shown in Figure 1.

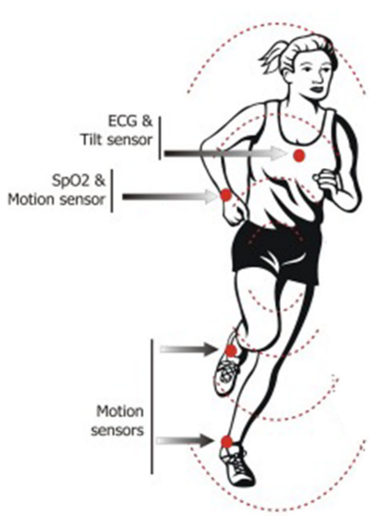

Figure 1. Wearable sensors

Skin temperature, heart rate, conductivity, Global Positioning System (GPS) location and body motion are some examples of variables that can be measured with current wearable sensor technologies [4]. All of the systems in this class are obtrusive and always conducted in a controlled environment. It is also relatively harder to synchronize the components of the system.

C. Smartphone as Wearable Sensor

Today's smart phones incorporate an array of diversified sensors within, e.g. inertial sensors (Accelerometer, Gyroscope), barometer, proximity sensor etc. and it also remains unobtrusive to users. As a result, it is now one of the main tools to recognize human activity automatically. Kwapisz et al. [5] used labeled accelerometer data from Android phones whereas Yang [6] used Nokia N95 phone. Miluzzo et al. developed CenceMe [7] using off-the-shelf, sensor-enabled mobile phones (Nokia N95) and exploited various sensors (such as a microphone, accelerometer, GPS, and camera) that are available for activity recognition. Different studies have been conducted using accelerometer embedded cell phones to detect physical activities with varying phone locations [8].

While basic activities can be recognized almost perfectly with smart phone sensors, semi–complex activities like ascending and descending stairs can be complex to recognize and differentiate between them [15].

D. Combining Different Sensors

This class of techniques involves including smart phone sensors along with other external sensors. Subramanya et al. in [9] built a model using data from a tri-axial accelerometer, two microphones, phototransistors, temperature and barometric pressure sensors, and GPS. Choudhury in [10] used multiple sensor devices consisting of seven different types of sensors to recognize activities. Cho et al. used a single tri-axial accelerometer, along with an embedded image sensor worn at the user's waist to identify nine activities [11]. Györbıróet al. [12] used "Motion Bands" attached to the dominant wrist, hip, and ankle of each subject to distinguish between six different motion patterns. Kawser et al. [1] used phone's accelerometer and another plantar pressure sensor attached in shoe to recognize activity. They mainly focused on combining data from different sensors and getting the final result based on these combined data.

As for the choice of classifier, one can use either temporal classifiers or non-temporal ones. In case of non-temporal classifiers like Decision Tree, Artificial Neural Network, k-Nearest Neighbor, K-Mean etc, one cannot provide the raw data as input directly to the classifier. First, one needs to extract some features from the raw data and then pass these features to the classifier. So there is a pre-processing on the raw data before sending it to classifier. In this case, the features that are mostly used are the following:

- Arithmetic Mean
- Standard Deviation
- Max, Min
- Median Absolute Deviation
- Frequency Signal Weighted Average

In case of temporal classifiers like AR and DTW, etc., there is no need for feature extraction. These algorithms take input of data as a time series. Moreover, one can easily plot the accelerometer and gyroscope data acquired from the smart phone into a time series. Hence, a temporal classifier is a better choice for classifying activities using the data acquired from smart phone sensors. As DTW works best on time series data, this makes DTW a good fit for our purpose.

A summarization of various HAR systems is provided in Table 1 below.

TABLE I. SUMMARY OF HAR SYSTEMS

HAR System	Sensors	Machine Learning Algorithm	Accuracy
1. Ferdaus, Sheikh, Richard (2015)	Accelerometer, Gyroscope, Planter Pressure Sensor	Decision Tree	94.37 - 99.53%
2. Davide, Alessandro, Luca, Xavier, Jorge (2012)	Accelerometer Gyroscope	Support Vector Machine	89.3%
3. Wu et al. (2012)	Accelerometer, Gyroscope	k-NN	90.2%
4. Jennifer, Gary, Samuel (2010)	Accelerometer, Gyroscope	Decision Tree, Logistic Regression, Multilayer N	91.7%
5.Maninni, Sabatini (2010)	2D-Accelerometer (Wearable)	NB, DT, kNN, ANN, GMM, cHMM	92.2 - 98.5%

From Table I, we can observe that, to obtain higher accuracy rate for recognizing Human Activities, Multimodal systems are more efficient than others.

III. OUR APPROACH

From the literature review, it is clearly observed that simpler locomotion activities can be detected easily but similar kinds of activities are difficult to differentiate. We are particularly interested in differentiating these similar activities correctly. As DTW algorithm has never been used before in HAR studies but it seems particularly suited for the purpose, we propose to apply DTW algorithm for classification purpose. We also introduce context filtering methods in HAR to filter the data for achieving a more precise result. Explanation of the context filtering is provided below.

We construct a multi-modal system that takes user's smart phone accelerometer and gyroscope value as input. It also concurrently records the change of altitude of the smart phone using its barometer sensor and the user's heart rate using a heart rate monitor during the period. Accelerometer and gyroscope values are used by the classifier to classify the data. The altitude and heart rate observation is then used as a context filter. So when classifier generates a result, it is passed through context filter and the output from the context filter is the final result.

From background study, we have shown that it is difficult for classifiers to classify similar activities like going upstairs and going downstairs, fast walking and slow running etc. The highest accuracy for detecting ascending and descending stairs is around 55-60%. The idea of context filter is particularly applicable here and it will be able to differentiate these similar activities using the observance on altitude and heart rate values and produce a more accurate result. It is often difficult to differentiate between these similar activities by only using accelerometer and gyroscope. But if we use barometer sensor here, this may improve the result. From the barometer sensor, we get the altitude. So while walking if the altitude of the person is increasing then we can assume that the person is going upstairs. If it is decreasing, then we can say the person is going downstairs. And if the altitude is stable, then we can assume the person is just walking. So, by normalizing the barometer and heart rate sensor values and then setting up rules corresponding with those, we can setup the Context Filter.

Context filter may not create an impact on the classified results instantly because the heart rate and the altitude values do not increase or decrease instantaneously. For example, when somebody starts running, his heart rate stays normal in the beginning but after a period of time, it starts increasing gradually. So our context filter will always be observing the altitude and heart rate values of the near past and will try to find a window containing noticeable fluctuation. When it will find it, it will again filter the results. This time it will filter with the new window for better optimization of the result. Figure 2 explains this scenario.

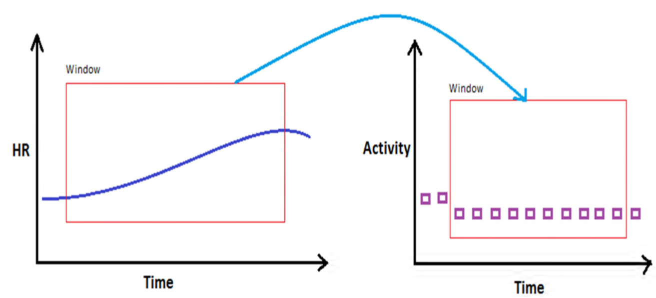

Figure 2. Context filtering Methodology

From the above figure we can see that rate of change of heart rate is gradually increasing.

IV. METHODOLOGY

Data from the Heart rate monitor and inertial sensor's data from Smartphone have been collected by an Android app. After collecting data in the Android app, it is sent to server through Wi-Fi. On the server side, all the data is processed by an offline data processing system. The classification of the data is done in the server. Figure 3 depicts the system architecture for our system.

For our study, we needed a smart phone which incorporates tri-axial Accelerometer, Gyroscope, and Barometric sensor. It must be ANT+ supported as we use Garmin Heart Rate Monitor as our sensor for obtaining heart rate data. A Samsung Galaxy S4 (I9505) smart phone was used in our experiment as it meets all the requirements.

Around 25 persons of different ages were chosen randomly for our data collection process. They were also different in physical built, height and weight.

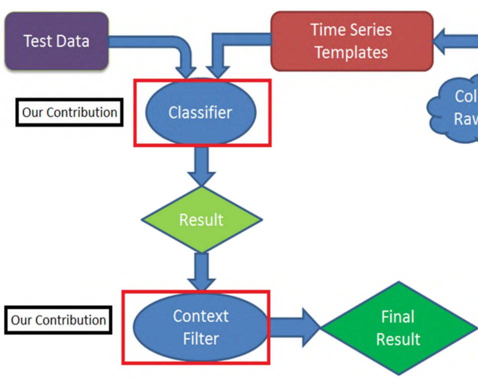

Figure 3. System Architecture

For developing the system, we built the Android app. This smart phone app was built using Android Development Tools (ADT) bundle which integrates a collection of the following programs:

- Eclipse: an integrated environment for the development of software projects with multi-language support.
- ADT plug-in: the toolset for Eclipse designed to allow the development of Android Apps.
- Android Software Development Kit (SDK): provides the API libraries and developer tools required to build apps for Android.
- ANT + SDK: provides the API libraries to use the ANT+ sensors.

From the Android App All data is passed to server via wireless medium. Then, on the server side all calculation is done. We take samples at a rate of 50 samples/second. The length of each activity sample is 8 seconds. We also calculate warp path by applying DTW algorithm. Warp path is calculated of every test data by taking the distance from the steady state. DTW works in only one dimension, but both the accelerometer and the gyroscope have three dimensions of data. So, we calculated the distance for each dimension and combined them to a single value which is our warp path using (1) and (2).

$$D_a = \sqrt{DTW(Xa)^2 + DTW(Ya)^2 + DTW(Za)^2} \quad (1)$$

$$D_g = \sqrt{DTW(Xg)^2 + DTW(Yg)^2 + DTW(Zg)^2} \quad (2)$$

Here D_a and D_g is the warp path for accelerometer and gyroscope, respectively.

For each template, each test data produced a warp distance indicating the difference from the steady state. Different kinds of activities produce different kinds of distances. So using this distance we can train a classifier which will take the warp path distance as input and then classify it to a class of activity.

D_a or D_g = DTW (Test Data)
Output = Classifier (D_a or D_g)

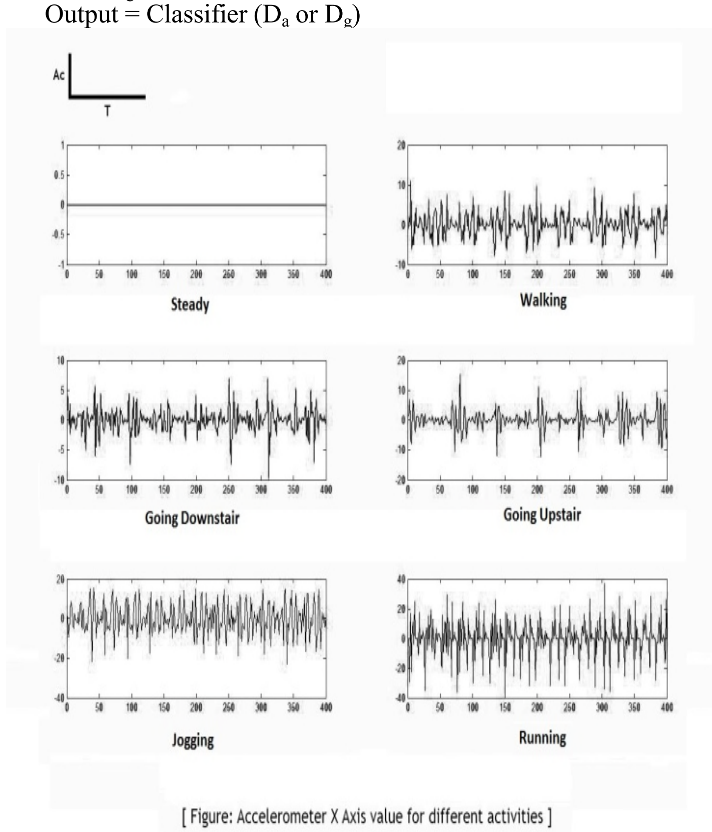

Figure 4. Accelerometer X axis value for different activities

If we observe the accelerometer x-axis values for different activities from Figure 4, then we will see that each of the activities has a specific pattern of acceleration amplitude levels during the activity. For example, in case of walking it is between 10 to -10, for jogging it is 20 to -20 and for running it is 40 to -40. However, the range is quite the same for walking, going upstairs and downstairs. All these characteristics remain consistent for the other accelerometer axis also. So if we compare these time series of the different activities with respect to a steady state then the output will be quite similar for walking, going downstairs and going upstairs but different for jogging and running.

There are three axis in accelerometer and DTW can be applied to only one time series at a time. So the warp path

 ISBN: 978-1-61208-468-8

distance for each axis with respect to steady state has to be measured separately and then combined to get a single value. So we have calculated the DTW warp path distance for each axis and combined them to a single value using the approach illustrated in Figure 5.

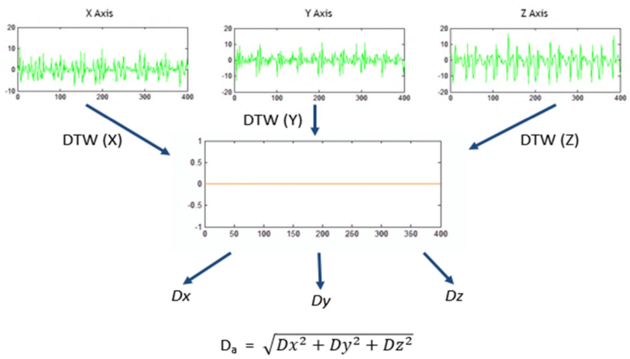

Figure 5. Applying DTW with respect to Steady State

We have already seen that the amplitude of acceleration differs for different kinds of activities. But, for walking, going downstairs and upstairs, the values are quite similar. That is why the total warp path distance (D_a) measured using DTW algorithm with respect to the Steady State is quite the same for these activities but different for jogging and running. Figure 6 shows the various warp path distance values for five different activities.

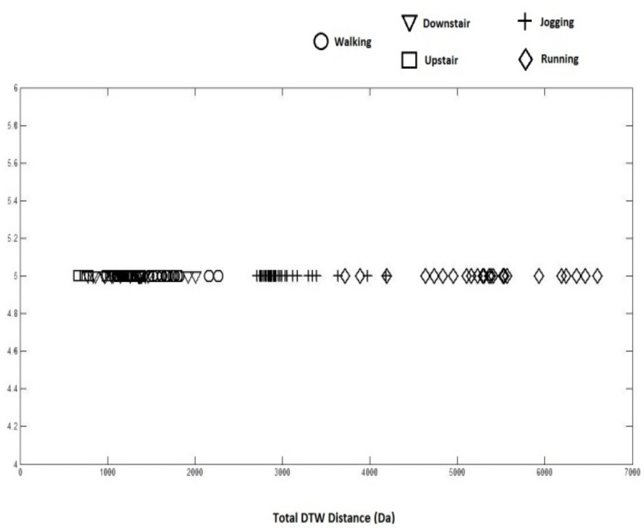

Figure 6. Total DTW distance in amplitude

It is clear from the figure that jogging and running has its separate zone but walking, going downstairs and going downstairs are in the same zone. So if we consider walking, going downstairs and going upstairs as a single class (Combined Class) then using the total warp path distance of an activity with a classifier, we can differentiate between the combined class, jogging and running. Using this characteristic, we have trained a k-NN classifier which takes the total warp path distance (D_a) of an activity as input and classifies into the combined class or jogging and running

Though the basic classification can be accomplished by the aforementioned method, still the confusion remains as we do not know the true class of an activity when it is classified to the Combined Class. So here we are applying our context filtering approach which uses barometer sensor values to differentiate them. When a data is classified as Combined Class, we are again filtering the decision using Context Filtering.

If we take the differences of the altitude of the smart phone from starting to ending point of combined class activities, i.e.

dA = Altitude of ending point – Altitude of starting point

and plot them on a graph, it will look as depicted in Figure 7.

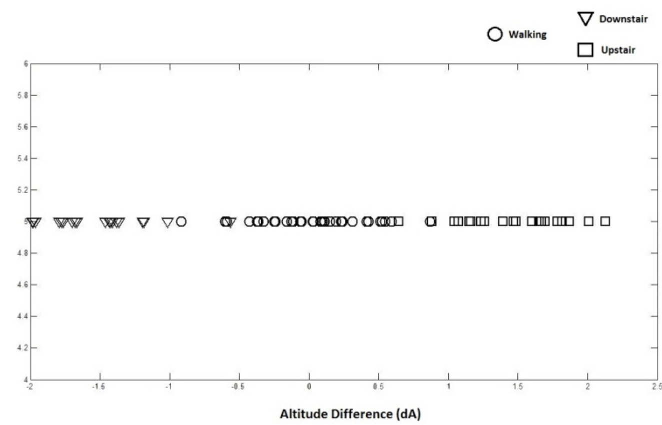

Figure 7. Altitude difference after applying KNN algorithm to find optimal threshold

So if we use ***dA*** of a combined class activity, we can easily differentiate them as each of walking, going downstairs and going upstairs has their own separate area and it is easily classifiable. We have used another k-NN classifier here which uses dA of a combined class activity to identify the true class. This is how an activity can be classified into one of the five classes by reducing the confusions.

V. RESULTS

Figures 8, 9 and 10 show the accuracy of detecting the five activities, respectively, using only accelerometer, only gyroscope and a combination of both sensors. On the basis of the different accuracy level of the different cases, it is clear that using only accelerometer in DTW produces better result in our approach. The computational complexity is also decreased as only the accelerometer is enough to detect the basic activity.

More importantly, the accuracy level for detecting similar activities like going downstairs and going upstairs has also been increased with respect to previous studies. For these two particular cases, we obtained accuracy of 92.85% and 100% respectively whereas in previous studies [5], the accuracy was 59.3% and 55.5%.

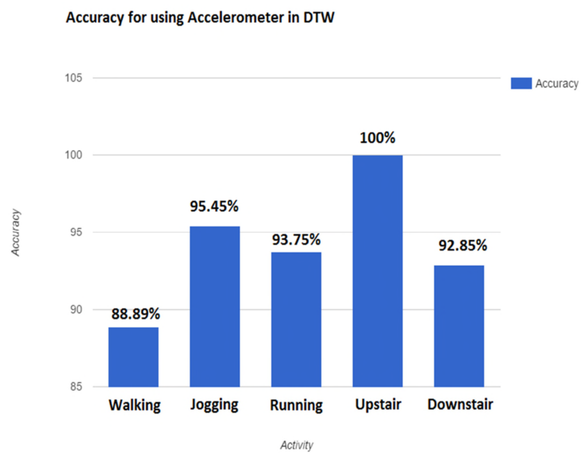

Figure 8. Accuracy using only Accelerometer

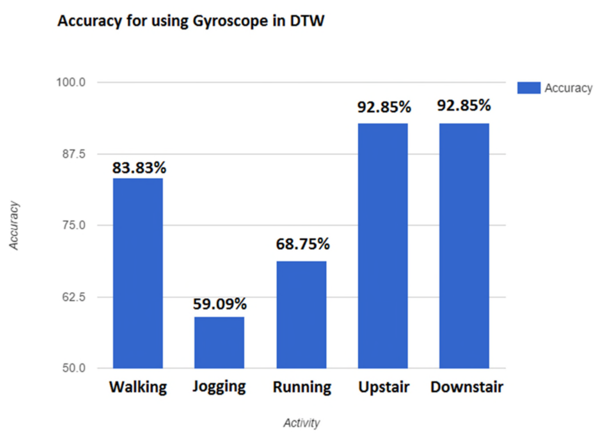

Figure 9. Accuracy for Using Only Gyroscope Value

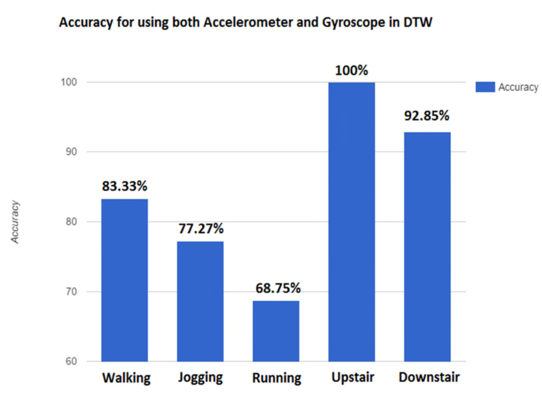

Figure 10. Accuracy for using Accelerometer & Gyroscope in DTW

Activity recognition success rate can again be increased for people of different category if we use only the data of that category as training samples. As people of different categories were chosen for our data collection process, the result we have acquired is quite for a general case.

VI. CONCLUSIONS

In this paper, we address some critical challenges of Activity Recognition with Mobile phones. DTW is an expensive algorithm with respect to time. May be this is the reason why this algorithm has not yet been used in this research field before, thinking that it may not be suitable for real time recognition. But rather using DTW for traditional template matching with respect to standard templates of each kind of activity, we are using it only once with respect to the steady state, reducing the sample number. We have also differentiated between similar activities like going upstairs and going downstairs. We all know that mobile phones have limited power capacity. So we have run our activity recognition classifier on our server instead on mobile phones to reduce the use of mobile battery. In the previous studies of human activity recognition, we have seen that the orientation and location of the smart phone was fixed to a certain part of a human body, mostly the waist. But we have placed the phone in the right pocket of the pant. So, it is more user friendly than the previous ones.

We also collected heart rate data during an activity. However, heart rate is mostly person dependent. Each person has a specific pattern of heart rate characteristic during an activity. So it is quite difficult to extract a universal feature from the heart rate values of a group of people. That is why we could not use the heart rate data to improve the classification result. However, as heart rate is person dependent, in case of user dependent classification, it may have some impact. This would be our future study to use heart rate for user dependent activity recognition learning. Recognition of semi-complex and complex activities like cooking, dancing, and travelling in bus etc. still remains a challenge. So along with locomotion activities, we will also try to recognize these complex and semi complex activities combining smart phone sensors and the knowledge we gain from our current work.

REFERENCES

[1] F. Kawsar, S. Ahamed, and R. Love, ”Remote Monitoring Using Smartphone Based Plantar Pressure Sensors: Unimodal and Multimodal Activity Detection”. In *Smart Homes and Health Telematics* (pp. 138-146). Springer International Publishing 2015.

[2] Jorge Luis Reyes Ortiz, “*Smartphone-based human activity recognition”*. Springer, 2015.

[3] Boris Takač, "Position and orientation tracking in a ubiquitous monitoring system for parkinson disease patients with freezing of gait symptom." *JMIR mHealth and uHealth* 1.2 (2013)

[4] Yang, Guang-Zhong, and MagdiYacoub, "Body sensor networks." (2006): 500.

[5] Kwapisz,R. Jennifer, M. Gary Weiss, and A. Samuel Moore., "Activity recognition using cell phone accelerometers." *ACM SigKDD Explorations Newsletter* 12.2 (2011): 74-82.

[6] Yang and Jun, "Toward physical activity diary: motion recognition using simple acceleration features with mobile phones." *Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics*. ACM, 2009.

[7] Miluzzo and Emiliano, "Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application." *Proceedings of the 6th ACM conference on Embedded network sensor systems*. ACM, 2008.

[8] Sun and Lin, "Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations." *Ubiquitous intelligence and computing*. Springer Berlin Heidelberg, 2010. 548-562.

[9] Subramanya and Amarnag,"Recognizing activities and spatial context using wearable sensors." *arXiv preprint arXiv:1206.6869* (2012).

[10] Choudhury, Tonmoy, Sunny Consolvo, Brent Harrison, Jeffrey Hightower, Antonio Lamarca, Louis LeGrand, AzarRahimi et al. "The mobile sensing platform: An embedded activity recognition system." *Pervasive Computing, IEEE* 7, no. 2 (2008): 32-41.

[11] Cho and Yongwon, "SmartBuckle: human activity recognition using a 3-axis accelerometer and a wearable camera." *Proceedings of the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments*. ACM, 2008.

[12] Győrbíró, Norbert, ÁkosFábián, and GergelyHományi, "An activity recognition system for mobile phones." *Mobile Networks and Applications*14.1 (2009): 82-91

[13] Poppe and Ronald, "Vision-based human motion analysis: An overview."Computer vision and image understanding 108.1 (2007): 4-18.

[14] Bian, Xuehai, D. Gregory Abowd, and M. James Rehg, "Using sound source localization in a home environment." *Pervasive Computing*. Springer Berlin Heidelberg, 2005. 19-36.

[15] Incel, Ozlem Durmaz, Mustafa Kose, and CemErsoy, "A review and taxonomy of activity recognition on mobile phones." *BioNanoScience* 3.2 (2013): 145-171.

[16] Campbell and T. Andrew, "The rise of people-centric sensing." *Internet Computing, IEEE* 12.4 (2008): 12-21.

[17] Salvador, Stan, and Philip Chan, "FastDTW: Toward accurate dynamic time warping in linear time and space." *KDD workshop on mining temporal and sequential data*. 2004.

[18] European Commission. Directorate-General for Economic and Financial Affairs. *The 2012 Ageing Report: Underlying Assumptions and Projection Methodologies*. Office for Official Publications of the European Communities, 2011.

[19] WHO, *World Report on Disability: Summary*. Technical report, (UN World Health Organization, 2011)