Basic Human Activity Recognition Based on Sensor Fusion in Smartphones

Charlene V. San Buenaventura, Nestor Michael C. Tiglao

Electrical and Electronics Engineering Institute
Velasquez St., University of the Philippines, Diliman, Quezon City, Philippines 1101
charlene.san_buenaventura@upd.edu.ph, nestor@eee.upd.edu.ph

Abstract— **Smartphones are ubiquitous devices that offer endless possibilities for health-related applications such as Ambient Assisted Living (AAL). They are rich in sensors that can be used for Human Activity Recognition (HAR) and monitoring. The emerging problem now is the selection of optimal combinations of these sensors and existing methods to accurately and efficiently perform activity recognition in a resource and computationally constrained environment. To accomplish efficient activity recognition on mobile devices, the most discriminative features and classification algorithms must be chosen carefully. In this study, sensor fusion is employed to improve the classification results of a lightweight classifier. Furthermore, the recognition performance of accelerometer, gyroscope and magnetometer when used separately and simultaneously on a feature-level sensor fusion is examined to gain valuable knowledge that can be used in dynamic sensing and data collection. Six ambulatory activities, namely, walking, running, sitting, standing, walking upstairs and walking downstairs, are inferred from low-sensor data collected from the right trousers pocket of the subjects and feature selection is performed to further optimize resource use.**

Keywords—Human Activity Recognition, Sensor Fusion, Feature Selection, Mobile HAR

I.   INTRODUCTION

Smartphones are becoming an integral part of our daily lives. The portability and programmability of these devices promise several potential applications that seem limitless. The wide range of sensors embedded in smartphones made collection of context information, at any desired instant, possible [12].

The activity of the user is one type of context information that can be derived from smartphone sensors. Identifying the current activity of the user can be useful for health-related applications such as ambient-assisted living systems [11], physical assessment of the individual [9], and monitoring the rehabilitation progress of patients [6]. However, these applications have specific requirements such as the response time and recognition accuracy [5]. In tele-health, accurate sensing is required so that healthcare providers can assess the conditions of the individuals correctly and provide them with proper treatment. For emergency warning systems, such as fall-detection, minimum amount of latency is necessary.

Although smartphones have higher battery capacities and computing capabilities than sensor nodes, their lifetime is generally shorter since they also have to support several other services and applications [9]. In addition, memory requirements should be low to enable long-term monitoring. Therefore, efficient methods are necessary to prolong battery life and limit memory usage.

Although lightweight classifiers are more energy and computationally efficient, and are more suitable for online activity recognition in mobile devices, they give lower recognition rates compared with more sophisticated classifiers. The main goal of this study is to leverage sensor fusion to improve the classification accuracy of a lightweight classifier while minimizing resource use. A feature-level sensor fusion is made to accurately classify six ambulatory activities from three smartphone sensors. Finally, a feature selection algorithm is applied to further reduce resource costs and improve recognition results.

The rest of the paper is structured as follows. Section II summarizes relevant work on human activity recognition. Section III presents the methodology in building the system. In Section IV, experimental results are reported and analyzed to give further insights to the current HAR problem. Finally, conclusion is drawn in Section V along with future work.

II.   RELATED WORK

Accelerometers are widely used in motion sensing because of their low-power requirement and non-intrusiveness [4]. The body component of the acceleration signal is the linear acceleration that detects motion of the body itself. Hence, the gravity component is often regarded as noise [12]. However, a recent study showed that the gravity component of the accelerometer aids in discriminating between sitting and standing activities [3]. In fact, the poor performance of gyroscopes in differentiating between the two activities can be attributed to the lack of gravity [3]. The two components of the accelerometer signal actually make accelerometers suitable for detecting both body movements and postural orientations [7]. Furthermore, it was observed in a previous study that linear acceleration is in general under par or on a par with the total acceleration measurement in terms of motion sensing [3]. To save computational and storage resources, the linear acceleration attribute is omitted in this study.

The component due to gravity of the accelerometer output can help in orientation detection [3]. However, since accelerometers are very sensitive, they are prone to noise or

vibration. We can apply smoothing to the signal to give accurate readings, but at the expense of a slower response [2]. Gyroscopes measure angular velocities along the axes of a body and are not sensitive to gravity. They also respond fast to a change in angle but can only measure orientation indirectly through integration. Consequently, gyroscopes experience integration drift [5], inherent from the determination of the angular position or orientation from the angular velocity reading. With this, a fusion of both accelerometer and gyroscope data would compensate for the weaknesses of the individual sensors [4]. Several studies regarding accelerometer-gyroscope fusion were conducted in the past and gave promising results [17], [18].

Other sensor fusion schemes can be found in HAR literature such as fusion of acceleration and ventilation signals [16], accelerometer and air pressure sensor [1], and a wide variety of sensor combination including camera, passive infrared, acoustic, pressure and contact sensors [14]. In [10] the activity recognition performance of the fusion of accelerometer, gyroscope and magnetometer data were evaluated. In all of these cases, sensor collaboration was found to be beneficial.

In a sensor-rich environment such as the smartphone, the problem now lies in deciding which sensors to use and to deactivate at specific times, or to collect data from, to save energy and battery life. However, though human activity recognition has been an active field of research over the past decade, very few works have successfully been deployed in mobile phones. There are still several challenges in designing a smartphone-based HAR system. Constraints such as limited memory, battery life and processing power [8] do not allow an automatic adoption of the standard techniques used for conventional HAR systems. Also, smartphones can be placed anywhere on the body or held freely by the user which necessitates a classification model that is not sensor position-specific.

Similar to traditional HAR systems, Fig. 1 presents the steps in activity recognition in mobile phones. An efficient classification algorithm is more favored for real-time smartphone activity recognition. Choosing a classifier that is computationally light while maintaining an acceptable degree of accuracy is essential to prolong battery life of the mobile device. The need of a lightweight classifier as well as all sensors being on one location on the body warrants the use of sensor fusion in smartphone-based HAR systems to improve accuracy [9]. One should take note, however, that adding more sensors will result in an increase in the number of extracted features. Some of these features are likely to be highly correlated [13], thus, emphasizing the importance of feature selection. Choosing only the most discriminative set of features improves prediction performance while reducing the computational complexity of the classification. Feature selection methods can also alleviate the problem of overfitting.

Several efforts have been made to reduce energy consumption in mobile-HAR systems. Since communications take up the highest amount of energy [9], it is desirable to limit external communications and favor on-the-device processing. Sensing also expends energy largely next to radio

transmissions. Although collaboration of more sensors might yield to the greatest recognition accuracy, it is impractical to use all available sensors. Some sensors, such as gyroscopes, consume power more and thus, limiting their on-state is beneficial [9]. Lastly, more efficient algorithms are preferred to lower computational load. If chosen algorithm is computationally expensive, it will drain the battery rapidly. Moreover, computationally complex algorithms tend to exhibit higher latencies rendering them useless in applications that require swift reaction times, such as in emergency response systems. Highly complex algorithms are also more likely to cause overruns during run-time.

Overall, the success of a smartphone-based HAR system greatly relies on the choice of the sensors, features and classification algorithm to be used that meet the constraints of the device and is adaptive to the current conditions. The criteria for selecting the methods to be employed depend on the specific requirements of the application, which can include response-time, recognition accuracy and energy consumption.

III. METHODOLOGY

A. Data Collection and Pre-processing

A publicly available dataset [15] was used which consisted of accelerometer, gyroscope and magnetometer readings collected by a Samsung Galaxy S2 from four male subjects performing six physical activities. The participants performed walking, running, sitting, standing, walking upstairs and walking downstairs while the smartphone is placed in four locations on the body: right jeans pocket, belt, right arm, and right wrist. Data are collected at a sampling rate of 50 Hz, which was observed to be sufficient in recognizing physical activities in the past [11]. The sensor stream was segmented by a sliding time window of two seconds with 50% overlap. The choice of both the time window length and the amount of overlap has been shown to be effective in physical activity recognition [11]. Only sensor data from the right trousers pocket is utilized in this study.

B. Feature Extraction and Selection

Features are extracted on a per time window-basis for the magnitude signals and each axis of the tri-axial accelerometer, gyroscope, and magnetometer. Time domain features such as mean, standard deviation, rms, median, variance, iqr, mad, zcr and mcr, and frequency domain features such as skewness, kurtosis and pca, were investigated in this study. The features are ranked according to their discriminatory power using the ReliefF algorithm available in the Machine Learning and Statistics Toolbox in MATLAB.
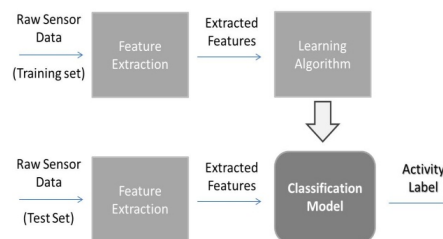

Fig. 1. General HAR Workflow

C. Classification

Two categories of machine learning algorithms that are available in the classification gallery of the ClassificationLearner App in MATLAB were explored in terms of accuracy and efficiency.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

There are several aspects in the construction of an efficient smartphone-based HAR system. One crucial step is to choose a classifier, lightweight enough to meet the constraints of the device. We begin our analysis by comparing the recognition performance of Decision Tree and Nearest Neighbor when accelerometer data alone is used. Both classifiers are popular choices for mobile-HAR. From Table 1, the Decision Tree classifier performed worse than Nearest Neighbor with slower training speed and smaller classification rate. Furthermore, Nearest Neighbor is adaptable to variable feature vector length, which can be useful for dynamic activity recognition. It is also a popular choice for user and device-orientation independent activity recognition [20]. Hence, the Nearest Neighbor classifier will be used as the lightweight classifier in this study. All evaluations are performed using 10-fold cross validation.

The individual performances of the three sensors are evaluated as well as their collaborative effects on the recognition accuracy. As seen from the summary of classification performances based on the set of sensors used in Table 2, recognition improved significantly when pairwise combinations of sensors were used. It can also be observed that accuracy was reduced slightly when magnetometer was added to the accelerometer-gyroscope combination.

Standing and sitting was not accurately discriminated from each other when using gyroscope alone. Likewise, magnetometer-only prediction greatly confused climbing upstairs and downstairs. Moreover, the activity climbing stairs was often confused with walking, and vice versa. Nevertheless, the percentage of false positives was reduced significantly when sensor fusion was utilized for classification. Fusion the three sensors can be used for improving the distinction between climbing upstairs and downstairs.

From Table 3, it can be seen that addition of frequency domain features to time domain features did not show any improvement in the classification. When only mean feature was used, the activity 'Laying' was well discriminated from the other activities. By analyzing the results further, it was also observed that the skewness feature of the accelerometer data was useful in distinguishing between upstairs and downstairs.

The selection of feature subsets is based on the ranking provided by the ReliefF, which is a filter method for selecting features and is independent of the classification algorithm used. Fig. 2 plots the accuracy vs. number of features used. It can be seen that accuracy increases almost linearly as more features are added but only up to a certain point. This point is when eight features are selected and also where the maximum accuracy is attained. Beyond this point, accuracy oscillates irregularly until it saturates to a much lower value. This can be due to the presence of redundant features. This also verifies that the Nearest Neighbor classifier scales poorly in high

dimensions [19] and justifies the need for feature selection. The top eight features are shown in Fig. 3.

TABLE I. COMPARISON OF CLASSIFIERS FOR ACCELEROMETER-ONLY ACTIVITY RECOGNITION

Classifier	Classification Rate	Training Time	Memory Used
Decision Tree	74.0%	46.692 s	83.1818 MB
Nearest Neighbor	96.8%	6.832 s	115.2081 MB

TABLE II. CLASSIFICATION RATES BASED ON SENSOR/S USED

	Type of Sensor/s	Classification Rate
Single Sensor	Accelerometer	96.9%
	Gyroscope	94.8%
	Magnetometer	91.5%
Two Sensors	Accelerometer & Gyroscope	98.7%
	Accelerometer & Magnetometer	97.7%
	Gyroscope & Magnetometer	97.7%
Three Sensors	Accelerometer, Gyroscope & Magnetometer	98.1%

TABLE III. CLASSIFICATION RATES BASED ON TYPE OF FEATURES USED

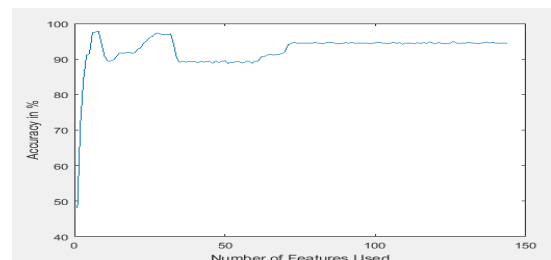| Type of Feature	Classification Rate
Time domain features only	98.3%
Frequency domain features only	95.8%
Time and frequency domain features	98.1%

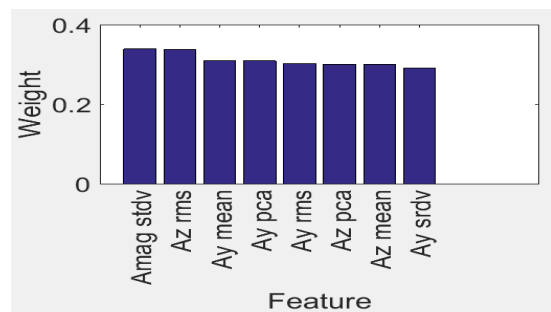Fig. 2. Recognition Accuracy Vs. Number of Features Used

Fig. 3. Top eight features ranked by ReliefF Algorithm

V. Conclusion and Future Work

Smartphones can be an unobtrusive means of gaining contextual information from the user. In this study, a feature-level sensor fusion was implemented to classify six ambulation activities from sensor data collected from a smartphone.

Nearest Neighbor was chosen as the lightweight classifier for this study and its overall accuracy was boosted through sensor fusion. It was found that accelerometer-gyroscope combination gave the best result while a subset of eight features was sufficient for recognition. Any addition of features hereafter results to model degradation. The optimal number of features was determined by plotting the accuracy versus the number of features used, as ranked by the ReliefF feature selection algorithm. It was also seen that activating or collecting data from all the sensors at all times is not necessary. Hence, a more efficient sensing scheme and data collection can be implemented to lower resource-use and enable long-term activity monitoring.

It was also shown that the addition of the selected frequency domain features did not improve classification. Thus, limiting the use of the more expensive frequency domain features can help minimize memory and computational expenditure. Confusion between certain activities was also alleviated using sensor fusion techniques.

Magnetometer has been shown to be ineffective in feature-level sensor fusion for improving the recognition performance. For future work, soft sensors that can be derived by fusing magnetometer with other sensors on a sensor-level fusion can be added to improve recognition. Physical features and additional statistical features in the frequency domain can also be incorporated to further improve the accuracy of the classification. Furthermore, other feature selection techniques that are more data-driven will be exploited for mobile-HAR. Other classifiers that are adaptable to variable feature vector length must be explored for efficient online mobile activity recognition. Energy consumption, memory usage, and latency should also be closely monitored on the actual device. Generally, the choice of proper methods can accomplish the correct balance between accuracy and meeting the constraints of the device.

References

[1] K. Sagawa, T. Ishihara, A. Ina, and H. Inooka. Classification of human moving patterns using air pressure and acceleration. Industrial Electronics Society, 1998. IECON '98. Proceedings of the 24th Annual Conference of the IEEE, 2:1214 – 1219, 1998.

[2] Accelerometer and gyro integration. Retrieved from http://www.hobbytronics.co.uk/accelerometer-gyro

[3] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. M. Havinga, "Preprocessing techniques for context recognition from accelerometer data", Personal and Ubiquitous Computing, vol.14, no.7, pp.645–662, 2010.

[4] O. Politi, I. Mporas, and V. Megalooikonomou, "Human motion detection in daily activity tasks using wearable sensors," Rion-Patras, Greece, 2011.

[5] Ortiz, J. L. (2015). Smartphone-Based Human Activity Recognition. Springer Theses. doi:10.1007/978-3-319-14274-6.

[6] Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). Training Computationally Efficient Smartphone–Based Human Activity Recognition Models. Artificial Neural Networks and Machine Learning – ICANN 2013 Lecture Notes in Computer Science, 426-433. doi:10.1007/978-3-642-40728-4_54, 2013.

[7] Hong, Yu-Jin, Ig-Jae Kim, Sang Chul Ahn, and Hyoung-Gon Kim. "Mobile health monitoring system based on activity recognition using accelerometer." Simulation Modelling Practice and Theory 18.4 (2010): 446-55. Web.

[8] Pires, I., Garcia, N., Pombo, N., & Flórez-Revuelta, F. (2016). From data acquisition to data fusion: a comprehensive review and a roadmap for the identification of activities of daily living using mobile devices. Sensors, 16(2), 184. doi:10.3390/s16020184.

[9] Rault, T., Bouabdallah, A., Challal, Y., & Marin, F. (2016). A survey of energy-efficient context recognition systems using wearable sensors for healthcare applications. Pervasive and Mobile Computing. doi:10.1016/j.pmcj.2016.08.003.

[10] Shoaib, M., Bosch, S., Incel, O., Scholten, H., & Havinga, P. (2014). Fusion of smartphone motion sensors for physical activity recognition. sensors, 14(6), 10146-10176. doi:10.3390/s140610146.

[11] Wen, Jiahui, Loke, Seng W., Indulska, Jadwiga and Zhong, Mingyang (2015). Sensor-based activity recognition with dynamically added context. In: Mihaela Ulieru and Valeriy Vyatkin, 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services MOBIQUITOUS 2015. International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, Coimbra, Portugal, (e4.1-e4.10). 22 – 24 July 2015. doi:10.4108/eai.22-7-2015.2260164.

[12] Khan, A. M., Tufail, A., Khattak, A. M., & Laine, T. H. (2014). Activity recognition on smartphones via sensor-fusion and KDA-based SVMs. International Journal of Distributed Sensor Networks, 2014, 1-14. doi:10.1155/2014/503291.

[13] M. Zhang, and A. Sawchuk, "Manifold learning and recognition of human activity using body-area sensors", in IEEE 10th International Conference on Machine Learning and Applications, 2011.

[14] Wichit, N. (2014). Multisensor data fusion model for activity detection. 2014 Twelfth International Conference on ICT and Knowledge Engineering. doi:10.1109/ictke.2014.7001535.

[15] Shoaib, M. and Scholten, J. and Havinga, P.J.M, "Towards physical activity recognition using smartphone sensors", in 10th IEEE International Conference on Ubiquitous Intelligence and Computing, UIC 2013, 18-20, Vietri sul Mare, Italy. pp. 80-87. IEEE Computer Society. ISBN 978-1-4799-2481-3 December 2013.

[16] Liu, S., Gao, R. X., John, D., Staudenmayer, J., & Freedson, P. S. (2011). SVM-based multi-sensor fusion for free-living physical activity assessment. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. doi:10.1109/iembs.2011.6090868.

[17] Wang, A., Chen, G., Yang, J., Zhao, S., & Chang, C. (2016). A comparative study on human activity recognition using inertial sensors in a smartphone. IEEE Sensors Journal, 16(11), 4566-4578. doi:10.1109/jsen.2016.2545708.

[18] Shen, C., Chen, Y., & Yang, G. (2016). On motion-sensor behavior analysis for human-activity recognition via smartphones. 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA). doi:10.1109/isba.2016.7477231.

[19] A. Wang, N. An, G. Chen, L. Li, and G. Alterovitz, "Accelerating wrapper-based feature selection with K-nearest-neighbor,"Knowl.-Based Syst., vol. 83, pp. 81–91, Jul. 2015.

[20] Ustev, Y. E., Incel, O. D., & Ersoy, C. (2013). User, device and orientation independent human activity recognition on mobile phones. Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication - UbiComp '13 Adjunct. doi:10.1145/2494091.249603.