

Recognition of Household and Athletic Activities using SmartShoe

S. Ryan Edgar¹, George D. Fulk², and Edward S. Sazonov³, *Senior Member, IEEE*

Abstract— The ability to provide real time feedback concerning a person's activity level and energy expenditure can be beneficial for improving activity levels of individuals. Examples include biofeedback systems used for body weight and physical activity management and biofeedback systems for rehabilitation of stroke patients. A critical aspect of any such system is being able to accurately classify data in real-time so that active and timely feedback can be provided. In the paper we demonstrate feasibility of real-time recognition of multiple household and athletic activities on a cell phone using the data collected by a wearable sensor system consisting of SmartShoe sensor and a wrist accelerometer. The experimental data were collected for multiple household and athletic activities performed by a healthy individual. The data was used to train two neural networks, one to be used primarily for sedentary individuals and one for more active individuals. Classification of household activities including ascending stairs, descending stairs, doing the dishes, vacuuming, and folding laundry, achieved 89.62% average accuracy. Classification of athletic activities such as jumping jacks, swing dancing, and ice skating, was performed with 93.13% accuracy. As proof of real-time processing on a mobile platform the trained neural network for healthy individuals was timed and required less than 4ms to perform each feature vector construction and classification.

I. INTRODUCTION

Human activity and posture recognition serves many useful purposes. It can be used during physical therapy and for feedback on maintaining a healthy lifestyle. However, the task of monitoring a person's movement and recognizing activities is not trivial. Several solutions to this problem already exist and have been shown to increase the subject's awareness of physical activity [1]. One study developed a variety of classifiers to recognize various activities in patients with chronic obstructive pulmonary disease [2]. Fourteen gym and everyday activities were recognized using 10 wireless nodes with three accelerometers each placed on a subject's body [3]. Recognition was performed using features such as root-mean-square, range, entropy, and the correlation coefficient for 5 second samples of data. The study looked at classifiers including instance based learning, Naïve Bayes, J48 decision trees, multi-layer perceptron (MLP), random forest, and support vector machines (SVM). The best classifier was found to be a support vector machine, 12.09% error, the

second most accurate classifier was the MLP, 14.18% error.

Other studies have relied on the MLP classifier, such as the 29 subject study [4]. This study used data collected from a cell phone's accelerometers worn by healthy subjects to classify six activities. Classification was performed on 10 second intervals of data using features such as average, standard deviation, average absolute difference, average resultant acceleration, time between peaks, and binned distribution. Out of the four classifiers tested (J48, logistic regression, MLP, and straw man), the MLP was the most accurate with an error of 8.3%.

In [5] researchers attempted using Intel iMotes to stream data for activity and posture classification. They were unable to wirelessly stream the data and resorted to using a wearable PDA wired to sensors on the body. Boosting was used for feature selection and a semi-supervised model capable of adapting to a specific user's data after deployment was employed.

One study has even taken classification to the next step using a real-time processing platform [6]. The study performed with three subjects used three wireless accelerometers placed on the dominate hip, wrist, and ankle. An artificial neural network (ANN) was used to classify six activities and the results indicated a 79.76% successful classification rate.

Our previous research demonstrated accurate classification of most common postures and activities (such as sitting, standing, walking/jogging, etc) using shoe-based wearable sensors (SmartShoe [7,8]).

Classification of activities and postures in healthy individuals [7] was performed using an SVM classifier. Nine subjects wore SmartShoe equipped with pressure sensitive insole sensors and accelerometers. Subjects performed six activities which were used for classification in a fourfold validation scheme. The resulting system was able to classify data for healthy individuals with an average accuracy of 98% in the group model for which each individual subject's data was either used only in the training set or only in validation set.

Similar research conducted in people with stroke using a SVM as the classifier yielded comparable (95%) accuracy of classification [8]. SmartShoe sensors were used to collect sitting, standing, walking, ascending steps, and descending steps from eight patients who had suffered a stroke. Classification was performed by SVM with Gaussian kernel. The topic of activity and posture recognition has been explored in many studies. None of these, however, provide for the requirements of fast real-time processing on a mobile platform, high accuracy, wearability, and ability to recognize a variety of activities. Studies such as [2] and [5] are not

This research was supported by Award Number R15HD061006 from the Eunice Kennedy Shriver National Institute Of Child Health & Human Development, and the New York Physical Therapy Association.

¹ Clarkson University, Department of Electrical and Computer Engineering, Potsdam, NY, 13699, edgars@clarkson.edu

² Clarkson University, Physical Therapy Department, Potsdam, NY, 13699

³ The University of Alabama, Department of Electrical and Computer Engineering, Tuscaloosa, AL, 35487, esazonov@ieee.org, phone: 205-348-1981.

easily worn by subjects whereas [7] provides an easily worn system, but complex processing requirements. An ideal system should combine the most useful attributes and designs from the existing systems to build a system which meets the specified requirements.

The first goal of this paper is to expand the range of activities recognized by SmartShoe to include various household chores and sports by using a hand acceleration sensor to add information on use of upper extremities. The second goal is to demonstrate feasibility of real-time classification on a mobile platform.

II. METHODOLOGY

A. Sensors

SmartShoe integrates a flexible insole with three pressure sensors, seen in Fig. 1a, with a tri-axial accelerometer and Bluetooth wireless electronics integrated into a clip-on easily attachable to the side of a shoe (Fig. 1b). In addition to SmartShoe a wrist node containing a tri-axial accelerometer (Fig. 1c) was used to collect motion data from the dominant arm. The signals from each of the 9 sensors were digitized and read as 12bit unsigned integers.

The data collection platform is based on SWiSH: A Smartphone based Wireless System for Human Activity Monitoring, a portable wireless wearable data collection system, documented in [9]. This system allows for a flexible number of wireless nodes, sensor configurations and sampling frequencies for data collection and testing. Data is collected using a series of wireless nodes which are controlled by a smartphone, in this case the Samsung Omnia I8000. Each wireless node contains three accelerometer channels and allows for external sensors. Nodes are small, discrete and do not prohibit the wearer's movements. Time synchronous data collection with a synchronization of less than 10ms on average is provided between distinct nodes at a frequency of 250Hz. SWiSH is a minimal data loss system and provides adequate features for testing different wearable sensor configurations.

B. Subjects and Protocol

The ability to recognize different household and athletic activities will help to provide a more accurate picture of energy expenditure and activities throughout the day. The activities were divided into three categories: basic, household and athletic. Basic activities included sitting, standing and walking. Household activities included ascending and descending steps, doing the dishes, folding laundry, and vacuuming. Athletic activities included jumping jacks, east coast six count swing dancing, and ice skating. Data was collected for several minutes for each activity at a frequency of 100Hz. All data was eventually down sampled to 10Hz, the intended target frequency. The data collection platform provides a mechanism for marking data as invalid should a problem occur. This was used to remove bad data resulting in 3 minutes of useable data for each activity for a total of 33 minutes.

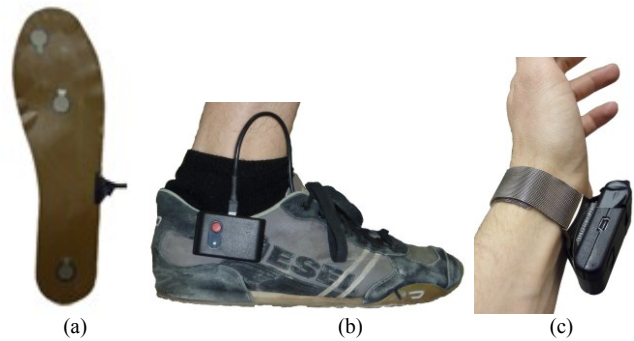


Fig 1. Nodes were located on the inside of the wrist (a) and the foot (b). The foot node connected to an insole with three pressure sensors (c).

One subject was used to determine if a network could recognize the activities for a particular individual. The subject was a 22 year old male with an amateur skill level in both ice skating and swing dance.

C. Signal processing

Classification was performed using 2 second intervals, or epochs, of raw sensor data. Raw data was sampled at 10Hz in order to provide fewer inputs to the ANN. This resulted in 10 sets of 20 data points (one every 100ms) for each of the 9 sensors per 2 second time interval, or 10 features vectors of 180 sensor readings each. Using this technique, there will be 10 different feature vectors representing any given 2 second time interval. Instead of passing only raw sensor data into the ANN in a way similar to [6], 39 features were added to increase classification accuracy. The additional features were based off of those specified in [9] during the initial design and development of the system. The first additional set of features contained the standard deviation of each sensor over the sampled time period, 9 features. The second set contained the number of times the signal from each sensor oscillated about the mean with amplitude greater than 150, 9 features. The value 150 was selected based on the input range of 0 to 4098. Third, the sum of the three pressure sensors for each reading was included, 20 features. This sum is representative of the force applied to the foot. Finally, a single logical feature, true or false, was included which indicated whether the maximum pressure reading was greater than a predefined threshold.

One ANN classifier was trained to classify basic and household activities and a second was trained to classify basic and athletic activities. The classifiers are intended to operate separately, one for use on primarily sedentary individuals and the second for more active individuals. The training set was the first two minutes of data and the validation set was the last one minute of data. Thus, there was no overlap in the signal used between data sets; an important point as 10 features vectors represent the same 2 seconds of time. Sets of 10 features vectors representing the same period of time were kept within the same data set. There were 600 and 300 feature vectors per activity, for training and validation respectively. Training and validation sets for each ANN were composed of 4800 and 2400 feature vectors respectively; 600 and 300 vectors for each of the 8

activities. MATLAB was used for all data processing and training data was represented in the standard floating point notation while validation data was represented in fixed point. Validation data in fixed point format was required because the trained ANN should be usable on a smartphone which may not be efficient at floating point calculations.

D. MLP ANN Training and validation

Both ANNs were trained in a way such that a particular combination of features on the network’s inputs activated an output representing a single activity or posture. The output node with the highest value indicated the posture. A pre-existing framework, Fast Artificial Neural Network (FANN) [10], was used to speed development time. When training an ANN, the activation function, error function, learning algorithm, and learning rate must be selected. A symmetrical sigmoid function was used for the activation function. Mean squared error was used for error evaluation during training. A standard batch training algorithm was used with a learning rate of 0.7.

A network was trained using the household activity data with 219 inputs, two hidden layers with 60 and 20 neurons, and 8 output neurons. It was trained to a mean square error of 0.01. Similarly a network was trained for athletic activities with 219 inputs, a hidden layer with 25 neurons, and 8 output neurons to a mean square error of 0.02. Each was run against its respective validation set. The configurations listed above were the result of trial and error testing until approximately less than 10% error was obtained for the classification of combined training and validation data.

E. MLP ANN Benchmarking

The SWiSH platform provides a library of functions to

interact with the nodes. A custom user interface with the ability to construct feature vectors and classify them using an ANN was developed on top of the SWiSH library. It uses a fixed point neural network to process incoming sensor data and provides a graphical representation of the activity being performed.

Neural networks are fast, but increasing the number of input features and hidden nodes can significantly impact performance. It is necessary to determine if the neural network will be able to process the incoming sensor data in a continuously running live processing application. The ANN must be able to construct and classify 10 feature vectors within a 2 second time period while allowing ample idle time for the processor. The application can then determine the most prevalently recognized activity and assume the user is performing that activity.

Network performance was analyzed by profiling the network through a data set five times for the largest network configuration from the trials above. In performing this test, the purpose was to determine the speed at which classification of raw data occurred, not the accuracy. The ANN used had 219 input neurons, two hidden layers of 140 and 60, and 8 output neurons. The data set consisted of 16000 vectors including training and validation data. The test was performed on the Omnia I8000 smartphone which was running Windows Mobile 6.5 with an 800Mhz CPU and 256MB RAM [11]. The raw sensor data, 180 readings for each interval, was loaded into the applications memory. Each feature vector was created with 219 features and classified by the ANN. Results were written to a CSV file to indicate the classification assigned to each raw data input vector. Timing started after the raw data was read from memory, and it was stopped before writing the results.

TABLE I
ANN HOUSEHOLD ACTIVITIES VALIDATION RESULTS

		Predicted								
Actual	Sit	290	8	0	0	0	2	0	0	0.9667
	Stand	0	270	0	0	0	0	20	10	0.9
	Walk	0	0	300	0	0	0	0	0	1
	Up Stairs	0	0	0	233	67	0	0	0	0.7767
	Down Stairs	0	0	0	13	287	0	0	0	0.9567
	Dishes	0	0	0	0	0	290	0	10	0.9667
	Vacuum	0	0	0	0	0	0	295	5	0.9833
	Laundry	0	11	0	0	0	93	10	186	0.62
	Precision	1	0.9343	1	0.9472	0.8107	0.7532	0.9077	0.8815	0.8962

TABLE II
ANN ATHLETIC ACTIVITIES VALIDATION RESULTS

		Predicted								
Actual	Sit	290	0	0	0	10	0	0	0	0.9667
	Stand	0	300	0	0	0	0	0	0	1
	Walk	0	0	300	0	0	0	0	0	1
	Jumping Jacks	0	0	0	300	0	0	0	0	1
	Skate Forward	0	0	19	0	237	44	0	0	0.79
	Skate Backward	0	10	0	0	0	290	0	0	0.9667
	Swing Lead	0	0	0	0	0	0	228	72	0.76
	Swing Follow	0	0	0	0	0	0	10	290	0.9667
	Precision	1	0.9677	0.9404	1	0.9595	0.8683	0.958	0.8011	0.9313

III. RESULTS

For the household activities set, training of the network was successful and resulted in a successful classification of 99.33% for the training data and 89.62% for the validation data. The validation results are summarized by the confusion matrix in Table I.

Results for the classification of athletic activities indicated the ANN could successfully recognize the activities. Results showed a 93.13% successful classification rate. See Table II for the confusion matrix.

The same data set of 16000 raw data vectors was classified five times by the test application. The timing began after loading all vectors were loaded into memory and was stopped once the final vector was classified. On average, it took 57800ms to classify the data set. This is 3.6125ms per feature vector.

IV. DISCUSSION

Considering the two networks, results were generally positive and showed approximately 90% accuracy. The lowest accuracy or precision for any one activity was 75%. This is a high level of accuracy considering the limited number of sensors, nine, the short sampling period, 2 seconds, and the number of activities.

Overall, results demonstrate that a high level of accurate classification is possible. For a single individual the network is able to learn and recognize basic movement patterns and trends within the feature vectors. An area of possible confusion for the matrix was activities performed while standing: scrubbing dishes, vacuuming, and doing the laundry. These activities were occasionally confused with each other and with standing itself. Each activity is performed while standing and the only significant difference is the movement of the hands. This must be picked up by the wrist sensor.

As for the network trained to recognize athletic activities, the greatest confusion was for skating backward and swing dancing follow. The confusion between skating forward and skating backward is attributed to the subject skating forward for short periods during the backward skating session in order to work up enough speed to turn around easily. Similarly for Swing dancing lead versus follow, due to the subject normally being a lead, his foot work sometimes reverted to the foot work of the lead during the follow data collection session. Additionally, the activities are only subtly different in motion patterns.

Given the repetitive nature of these activities, it is expected an ANN could easily be trained for multiple subjects and generalized to subjects not seen during training. The rhythm of ice skating and the basic step of swing dancing are hypothesized to be the main features used for recognition.

Given the time required for feature vector construction and classification, the smartphone can classify 10 feature vectors within a 2 second time period. Classification will not interfere with data collection or processing using SWiSH

DLL as the classification routine uses a small percentage of overall CPU resources.

V. CONCLUSION

This research is the foundation for a real time activity monitoring system to increase activity levels in individuals post-stroke. It successfully shows the ability to recognize complex activities for a single subject on a mobile platform.

Neural Networks are a powerful tool for real time classification of postures and activities. Using the SWiSH Activity Monitoring system data can be processed in real-time on a smartphone and provide approximately 90% accuracy for recognizing different postures and activities.

As a next step in this research, a small group study should be conducted with several individuals using the same activities. This would provide further evidence of the SWiSH platform and a neural network to classify household and athletic activities for more accurate energy expenditure and activity recognition.

Evaluation of the largest network trained showed the ability of the fixed point neural network to run in fast speeds on a Samsung Omnia I8000 smartphone. The feature vector construction and classification was performed in less than 4ms. This shows the system's ability to operate in a mobile environment in real time.

REFERENCES

- [1] S. Consolvo, K. Everitt, I. Smith, and J. A. Landay, "Design requirements for technologies that encourage physical activity," presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems, 2006, pp. 457-466.
- [2] S. Patel, C. Mancinelli, J. Healey, M. Moy, and P. Bonato, "Using wearable sensors to monitor physical activities of patients with copd: A comparison of classifier performance," *2009 Body Sensor Networks*, pp. 234-239, 2009.
- [3] A. Burns et al., "SHIMMER™—A Wireless Sensor Platform for Noninvasive Biomedical Research," *Sensors Journal, IEEE*, vol. 10, no. 9, pp. 1527-1534, 2010.
- [4] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," presented at the Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data, 2010, pp. 10-18.
- [5] T. Choudhury et al., "The mobile sensing platform: An embedded activity recognition system," *IEEE Pervasive Computing*, pp. 32-41, 2008.
- [6] N. Gyrbiró, Á. Fábrián, and G. Hományi, "An activity recognition system for mobile phones," *Mobile Networks and Applications*, vol. 14, no. 1, pp. 82-91, 2009.
- [7] E. S. Sazonov, G. Fulk, J. Hill, Y. Schutz, and R. Browning, "Monitoring of Posture Allocations and Activities by a Shoe-Based Wearable Sensor," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 4, pp. 983-990, 2011.
- [8] G. D. Fulk and E. Sazonov, "Using Sensors to Measure Activity in People with Stroke," *Topics in Stroke Rehabilitation*. 2011;18(6): 746-757.
- [9] S. R. Edgar, "SWiSH: A Smartphone based Wireless System for Human Activity Monitoring," Clarkson University, Potsdam, NY, 2011.
- [10] S. Nissen, "Fast Artificial Neural Network Library (FANN)." [Online]. Available: <http://leenissen.dk/fann/wp/>. [Accessed: 08-Apr-2011].
- [11] "Samsung I8000 Omnia II - Full phone specifications." [Online]. Available: http://www.gsmarena.com/samsung_i8000_omnia_ii-2836.php. [Accessed: 30-Jan-2012].