Human activity recognition using sensor recordings from passive infrared and microwave radar sensors

Aleksandra Bozhinoska
Ss. Cyril and Methodius University in Skopje
Faculty of Computer Science and Engineering
Skopje, Republic of North Macedonia
aleksandra.bozinoska@students.finki.ukim.mk

Dejan Gjorgjevikj
Ss. Cyril and Methodius University in Skopje
Faculty of Computer Science and Engineering
Skopje, Republic of North Macedonia
dejan.gjorgjevikj@finki.ukim.mk

Abstract*—Human activity recognition is a challenging and wide-spread research topic among the data scientists. It has ever-increasing importance in human monitoring approaches used mainly in Ambient Assisted Living (AAL) and healthcare. The most demanding task regarding this field of research is achieving high recognition accuracy while maintaining low equipment costs by using fewer simple and inexpensive sensors. An experiment has been conducted by installing three identical data collection modules in a controlled environment for human activity recording. The data collection modules were composed of Arduino microcontroller and two modified low-cost ambient sensors: microwave radar sensor and passive infrared (PIR) sensor. The sensors were modified to obtain their analog outputs that were logged on a SD card. The experiment was conducted on six subjects performing seven different activities. This has produced almost 2.5 hours of recorded sensor measurements that were then labeled with the corresponding activity producing data set of 654061 entries. Different approaches for feature extraction and preliminary tests for activity recognition were conducted. Different sliding window sizes, as well as considering only certain sensors (of the six possible) and their combinations were also examined.

***Index Terms*—human activity recognition; passive infrared sensor; microwave radar sensor; Arduino; sliding window technique**

I. Introduction

The number of elderly people is rapidly growing as a proportion of the total population in most developed countries around the world [1]. Significant number of them live alone within their own house. To be able to function independently at home, individuals have to be able to perform Activities of Daily Living (ADLs) [2] such as eating, dressing up, cooking, drinking, and taking medicine. Automating the recognition of the activities is an important step towards monitoring the functional health of a smart home resident.

Advancements in sensor technology, the declining costs of sensors and their vast availability opened up unprecedented opportunities for a wide variety of industrial, scientific, commercial, agricultural and military applications, such as home automation, health care, emergency response, smart transportation, infrastructure protection, and others [3]. Pervasive sensing technologies are becoming more and more popular offering new opportunities in smart homes such as providing health monitoring and assistance to individuals experiencing difficulties living independently at home. In order for the functional health of smart home residents to be monitored the system should recognize and track the activities that people perform at home.

Human activity recognition has attracted a lot of research activity in several fields like ambient assisted living, sports injury detection, elderly care, rehabilitation, and entertainment and surveillance in smart home environments. Most of the research was conducted using data collected by wearable sensors [4][5]. As wearable sensors smartphones, smart bands, and dedicated sensor nodes have been used. Approaches that are using unobtrusive ambient sensors were also reported [6]. Among the unobtrusive sensors use of sensor nodes equipped with acoustic sensors [7], ultrasound sensors [8], pyroelectric motion detection sensors [9] and recently microwave radar sensors [10] has been reported for human activity recognition. They are considered as essential for number of enabling technologies for independent living by the elderly such as the ambient assisted living systems (AALS).

In this paper, we present initial study of human activity recognition and identification using very cheap passive infrared (PIR) and microwave radar sensors. Three data collection modules developed around the sensors were placed in a room where volunteers were instructed to perform seven different activities in random order. The sensor measurements during the experiment were recorded by the modules and then labeled with the corresponding performed activity. Initial experiments considering different approaches for feature extraction and machine learning for automatic activity recognition were performed and the obtained results are presented and discussed.

The remainder of this paper is organized as follows: In section 2 conducting of the experiment is documented and section 3 explains the process of labeling of the gathered data. Section 4 elaborates the suggested feature extraction techniques and section 5 documents the examined machine learning approaches. Section 6 discusses and compares the used machine learning approaches and finally section 7 summarizes the results of the documented work and suggests future plans for improvement.

II. Conducting the experiment

For the needs of human activity recognition experiment, a controlled environment was created for the realization of the following seven activities: sitting on the bed, sitting at the table, walking out of the room, walking into the room, walking around the room, eating at the table and laying on the bed. For that purpose, an isolated room was prepared and supplied with only the needed furniture elements. The floorplan of the prepared room for activity recording is given in Fig.1. Three identical data collection modules were placed on three distant locations in the room, namely the two side walls and the ceiling. The modules consisted of Arduino microcontroller board, real-time clock (RTC) microSD card adapter and two motion detection sensors: passive infrared sensor HC-SR501 (PIR) and microwave radar sensor RCWL-0516 [11]. The sensors were modified to obtain the analog signal that was digitized to 10-bit precision and logged on the SD card at rate of 25 samples per second. The clocks of the sensors were synchronized before the experiment to guarantee alignment in the sensor measurements taken by the separate data collection modules.

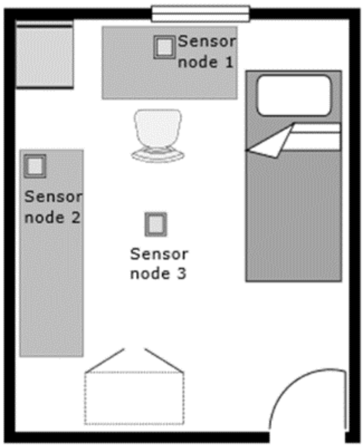

Fig. 1. Floorplan of the room where the sensors were placed

Six volunteers were instructed to perform the activities in the room one by one while the data streams from the 3 pairs of sensors were recorded. Each of the subjects was instructed on the next activity to perform by continuous delivery of messages to his mobile phone using previously developed simple mobile application. The activities were provided in random order taking into consideration the same activity not to be repeated twice in a row. The subjects were also continuously filmed by a dashboard camera for the purpose of ground truth labeling of the data. The experiment produced about two and a half hours of sensor recordings that after the cleansing resulted in 654061 entries.

The recording was performed on micro SD card in each of the modules, in digital compressed format that contains precise timestamp of each sample and also readings from some other sensors (ambient temperature, lighting and noise level) that were not used in this experiment. The data from the three modules was combined in time aligned fashion and only the readings from the PIR and the microwave radar sensors from each of the modules were considered.

III. Data fusion and labeling

The ground truth labeling was performed manually by following the activities of the subjects in the filmed videos. Nine different labels were used for labeling the entire entry set of sensor recordings, namely the labels were: SB (sitting on the bed), ST (sitting at the table), WO (walking out of the room), WI (walking into the room), WA (walking around the room), ET (eating at the table), LB (laying on the bed), SD (standing in one place) and one more used to denote subject absence or empty room, denoted as no activity - (NA). Although the last two were not part of the initial set of activities, they were added after being recognized as a potential source of confusion by the classifiers in the phase of activity recognition. The labeled experiment data consisted of 553 non-overlapping consecutive activities undertaken by the subjects with average duration time of 15.89 seconds. Statistics about the label distribution are presented in Table 1.

Table 1. Class distributions among the labeled samples

Target class	Number of samples	Percent of total number of samples
SB	31271	14,34
ST	28032	12,86
WO	4046	1,86
WI	3846	1,76
WA	37587	17,24
ET	17831	8,18
LB	30870	14,16
SD	13535	6,21
NA	51002	23,39

IV. Data processing and feature extraction approaches

A sliding window technique was exploited in the feature extraction approaches in order to capture recent history context of measured values for each of the sensors. By experimenting with different sliding window lengths varying from 2 to 20 seconds, the optimal values of 3, 4 and 5 were identified. In all the feature extraction approaches, new generated data sets were produced by applying computations on the series of values for the individual sensors in the current sliding window and then combining the new features. The feature extraction approaches were more efficient when applied to the combined data set of entries from all three modules, generated by joining the corresponding entries from each module in one entry (corresponding are the entries with the corresponding serial numbers). The following grouping operations were applied to the series of values for the 6 different sensors in the combined data set as different feature extraction approaches:

A. *Averaging (computing cumulative average, weighted moving average and exponential weighted moving average)*

Three different averaging approaches were examined as possible feature extraction techniques. First evaluated averaging approach was the cumulative average, or computing the standard average of all sensor measurements in the current window for each sensor. This approach gives equal importance to all the sampled values in the previous n-seconds. The number of entries in the sliding window was computed as $n * 25 + 1$ (with n being the window size in seconds), considering the sampling rate. The complete equation is defined as:

$$CMA_{n,si} = \frac{v_{1,si} + v_{2,si} + \cdots + v_{k,si}}{k} \tag{1}$$

where $k = n * 25 + 1$ and $1 \le i \le 6$, with si being the i-th sensor and n the sliding window size in seconds.

Two more averaging techniques were utilized in order to impose higher importance to the most recently sampled values while computing the averaged sensor values. Therefore, weighted moving average technique was applied to the values in the window, defining weights as consecutive integers in the interval [1, number of values in the window] and assigning them to the ordered sensor values. The complete equation for the weighted moving average is defined as:

$$WMA_{n,si} = \frac{w_1 * v_{1,si} + w_2 * v_{2,si} + \cdots + w_k * v_{k,si}}{w_1 + w_2 + \cdots + w_k} \tag{2}$$

where $w_j = j,\ k = n * 25 + 1$ and $1 \le i \le 6$, with si being the i-th sensor and n the sliding window size in seconds.

The last considered averaging technique was exponential weighted moving average with weights decreasing exponentially from latest to the earliest sampled sensor value in the window. The computation formula is defined as:

$$EWMA_{n,si} = \frac{w_k * v_{1,si} + w_{k-1} * v_{2,si} + \cdots + w_1 * v_{k,si}}{w_1 + w_2 + \cdots + w_k} \quad (3)$$

where $w_j = (1-\propto)^{j-1}, \propto = \frac{2}{k+1}, k = n * 25 + 1$ and $1 \le i \le 6$, with si being the i-th sensor and n the sliding window size in seconds. Fig. 2 visualizes the different rates of weight decrease in weighted moving average and exponential weighted moving average approaches.

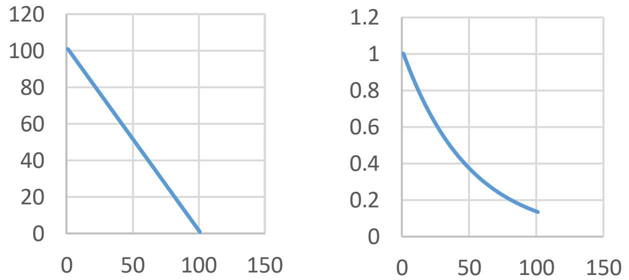

Fig. 2. Decrease of weights in WMA approach (left) and EWMA approach (right) for assumed 4 seconds sliding window size

B. *Computing number of zero crossings*

One proposed feature for computing the rate of change of the sensor values in the window was modified zero crossings approach. Namely this feature represented the count of occurrences when two consecutive measured values in the sequence appeared in different subintervals of the interval of all possible values [0-1023]. The first subinterval is the interval containing values less than 512 and the second one is the interval containing greater values that 512.

C. *Computing number of local extremes*

Another examined feature extraction approach was computing the number of local extremes considering the discrete values in the sliding window for each of the sensors. The feature was extracted by first computing the slope of the function for each pair of consecutive discrete values in the sliding window and then counting the number of pairs of consecutive slopes in the computed sequence with opposite signs. The used computation formula is defined as:

$$slope_{j,si} = \frac{v_{j,si} - v_{j-1,si}}{t_j - t_{j-1}} \quad (4)$$

where $2 \le j \le k, k = n * 25 + 1$ and $1 \le i \le 6$, with si being the i-th sensor and n the sliding window size in seconds.

D. *Integrating*

Integration was used as a feature extraction approach in order to capture the magnitude of measured values for each of the sensors in the sliding window. Integrals were computed on the function represented by pairs of transformed measured sensor value and time of measurement. Measured values were transformed by first subtracting 512 (the median from all possible digitized values) and then computing absolute value. Time of measurement in the interval was represented as sequence of discrete values from 0 to 40 * window-size- in-entries with step of 40 (because of sampling rates of the values). Simpson's rule for approximating integral of function represented by discrete pairs of values was used by calling its implementation in scipy.integrate library in Python [12].

V. Experiments and Results

Different data sets were produced by applying one or more of the proposed feature extraction approaches in the previous chapter to the combined data set of all six sensors. Each of the approaches was used to compute the target feature for each of the 6 sensors so the number of the features in the generated data set was number-of-combined-approaches * 6. Two alternatives for the target label of the computed features were tested, namely taking the label of the last test entry or taking the label of the middle test entry in the current window as the target class label. The experiments regarding generating the data sets and then applying machine learning algorithms were conducted using Jupyter Notebook with Python 3.7.0 and scikit-learn library version 0.19.2. The models were tested on three different classifiers: Random forest classifier, Artificial neural network and Support vector machine. The random forest classifier included 500 estimators, ANN consisted of two hidden layers of 200 and 100 neurons, and the kernel of SVM classifier was Radial Basis Function with $\gamma = \frac{1}{number\ of\ features}$ and $C = 1.0$. Overlapping and non-overlapping sliding windows were examined as well as different sliding window sizes. The overlapping window approach showed to be troublesome for proper testing because of the unequal class label distribution over the data set and also sequential dependency of the consecutive entries (in order to produce valid values for each of the features). The non-overlapping approach on the other side, was trained on the 80% of the samples in the data set and then evaluated on the remaining 20% (considering equal target class distributions after shuffling all the samples). It was examined on the data sets generated by applying only the individual feature extraction approaches as well as the paired averaging techniques (cumulative average, weighted moving average and exponential weighted moving average) with zero crossings, local extremes and integrals. Two more combinations of three different feature extraction approaches were also examined, namely: Cumulative average + Local extremes + Integrals and Exponential weighted moving average + Local extremes + Integrals. The evaluation results are presented in Table 2 and Table 3. Table 2 summarizes the recognition accuracy by feature extraction approach and by classifier using the last test label variant and Table 3 presents the same information for the middle test label approach.

VI. Discussion

From the provided recognition accuracies in Table 2 And Table 3 we can conclude that Random forest classifier generally provided the highest recognition accuracies compared to the other two classifiers, although all the results produced by all classifiers are pretty close and similar. The

Table 2. Recognition accuracy by feature extraction approach and by classifier using the last test label approach given in percent

Classifier	**Random Forest**		**ANN**		**SVM**	
Window size	**3s**	**4s**	**3s**	**4s**	**3s**	**4s**
Cumulative average (CA)	21,95	19,44	15,51	20,37	24,22	21,06
Weighted moving average (WMA)	23,00	20,83	17,60	14,81	23,00	25,00
Exponential weighted moving average (EWMA)	22,82	20,37	18,12	15,97	20,73	24,54
Zero crossings	23,00	24,31	24,74	24,31	25,44	25,46
Local extremes	24,74	24,31	24,04	23,61	25,61	23,84
Integrals	22,47	25,93	16,72	15,28	21,43	22,92
CA + Zero crossings	24,74	25,93	15,85	19,44	24,74	23,38
CA + Local extremes	25,78	24,54	15,68	13,89	22,30	24,07
CA + Integrals	25,09	23,84	19,34	20,83	24,39	23,38
WMA + Zero crossings	26,13	29,17	23,34	18,98	24,39	23,61
WMA + Local extremes	27,00	25,46	15,67	20,37	24,39	26,16
WMA + Integrals	25,96	26,16	16,72	14,35	21,95	22,92
EWMA + Zero crossings	26,83	23,84	19,86	18,75	24,56	27,78
EWMA + Local extremes	25,44	23,38	13,24	18,98	22,65	24,54
EWMA + Integrals	25,26	25,69	15,33	22,22	24,91	22,22
CA + Local extremes + Integrals	26,48	28,94	17,60	20,37	23,00	23,61
EWMA + Local extremes + Integrals	27,70	25,46	20,91	17,82	24,22	21,30

Table 3. Recognition accuracy by feature extraction approach and by classifier using the middle test label approach given in percent

Classifier	**Random Forest**		**ANN**		**SVM**	
Window size	**3s**	**4s**	**3s**	**4s**	**3s**	**4s**
Cumulative average (CA)	23,17	19,68	19,34	14,58	22,65	21,06
Weighted moving average (WMA)	20,38	22,22	16,20	16,90	22,82	22,45
Exponential weighted moving average (EWMA)	18,99	20,60	10,98	16,20	22,30	21,30
Zero crossings	21,78	22,92	23,87	22,92	24,39	25,93
Local extremes	22,47	21,30	21,78	21,53	24,56	21,76
Integrals	23,34	22,22	21,25	12,27	23,00	22,92
CA + Zero crossings	23,52	28,47	19,51	15,51	22,65	26,16
CA + Local extremes	26,31	26,62	14,29	17,82	24,22	22,69
CA + Integrals	22,47	28,94	13,76	20,83	21,25	25,23
WMA + Zero crossings	26,66	27,78	15,33	22,92	26,31	24,07
WMA + Local extremes	24,39	26,62	22,13	16,44	23,52	25,00
WMA + Integrals	27,18	28,24	14,29	15,51	24,39	25,23
EWMA + Zero crossings	24,74	26,16	18,12	11,34	22,65	23,38
EWMA + Local extremes	24,22	23,61	21,08	11,81	23,17	23,15
EWMA + Integrals	26,48	25,46	21,60	17,36	26,13	23,61
CA + Local extremes + Integrals	28,22	32,00	20,91	16,67	25,78	24,77
EWMA + Local extremes + Integrals	26,31	29,86	16,20	15,05	24,56	23,84

best recognition accuracies were achieved with 4-second-sized window and using the middle test label approach. Because of the very limited number of subjects and therefore number of entries in the data set, the highest individual sample recognition accuracy was 32% using the combination of cumulative average, local extremes and integrals as feature extraction techniques. Probably the results produced by ANN and SVM classifiers could be improved by fitting the classifier characteristics and also the overall recognition score by enlarging the testing data set.

VII. CONCLUSION

In this work, various approaches to feature extraction for human activity recognition on a custom data set were presented. The data set was created over captured sensor measurements in controlled environment in which six volunteers were performing seven different activities. Different feature extraction techniques using sliding window approach, as well as different combination of features and considered sensors were investigated. Numbers of experiments of automatic activity recognition using several classifiers were conducted. Considering the very limited data set and the variety of activities and the way they can be performed, an individual entry classification rate of 32% was obtained. Collecting a bigger data set and considering alternative feature extraction and feature fusion approaches are planned for further research.

ACKNOWLEDGMENT

This research was partially funded by the Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje.

REFERENCES

[1] "World Population Ageing 2015", (ST/ESA/SER.A/390), United Nations, Department of Economic and Social Affairs, Population Division (2015).

[2] V. Wadley, O. Okonkwo, M. Crowe and L.A. Ross-Meadows, "Mild Cognitive Impairment and everyday function: Evidence of reduced speed in performing instrumental activities of daily living", American Journal of Geriatric Psychiatry 16 (2007), 416424.

[3] O. Kanoun and H. Trankler, "Sensor Technology Advances and Future Trends", IEEE Transactions on Instrumentation and Measurement, vol. 53, no. 6, pp. 1497-1501, 2004.

[4] O. Lara and M. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors", IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1192-1209, 2013.

[5] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou and Y. Amirat, "Physical Human Activity Recognition Using Wearable Sensors", Sensors, vol. 15, no. 12, pp. 31314-31338, 2015.

[6] E. Salomons, P. Havinga and H. van Leeuwen, "Inferring Human Activity Recognition with Ambient Sound on Wireless Sensor Nodes", Sensors, vol. 16, no. 10, p. 1586, 2016.

[7] J. Sim, Y. Lee and O. Kwon, "Acoustic Sensor Based Recognition of Human Activity in Everyday Life for Smart Home Services", International Journal of Distributed Sensor Networks, vol. 11, no. 9, p. 679123, 2015.

[8] J. Sim, Y. Lee and O. Kwon, "Acoustic Sensor Based Recognition of Human Activity in Everyday Life for Smart Home Services", International Journal of Distributed Sensor Networks, vol. 11, no. 9, p. 679123, 2015.

[9] X. Luo, Q. Guan, H. Tan, L. Gao, Z. Wang and X. Luo, "Simultaneous Indoor Tracking and Activity Recognition Using Pyroelectric Infrared Sensors", Sensors, vol. 17, no. 8, p. 1738, 2017.

[10] G. Diraco, A. Leone and P. Siciliano, "A Radar-Based Smart Sensor for Unobtrusive Elderly Monitoring in Ambient Assisted Living Applica- tions", Biosensors, vol. 7, no. 4, p. 55, 2017.

[11] D. Gjorgjevikj, Gj. Madjarov, "Data Collection Module for Human Activity Recognition", in Miroslav Kotevski (edt.), Proceedings of 15th International Conference, ETAI 2018, pp. ETAI1-1, Ohrid, Republic of North Macedonia, 20-23 September 2018.

[12] Scipy.integrate v1.1.0, https://docs.scipy.org/doc/scipy-1.1.0/reference/index.htm.