

Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions

Uwe Maurer¹, Asim Smailagic², Daniel P. Siewiorek², Michael Deisher³,
²School of Computer Science, Carnegie Mellon University, Pittsburgh
¹Computer Science Department, Technische Universität München, Germany
³ Intel, Hillsboro, OR

Abstract

The design of an activity recognition and monitoring system based on the eWatch, multi-sensor platform worn on different body positions, is presented in this paper. The system identifies the user's activity in realtime using multiple sensors and records the classification results during a day. We compare multiple time domain feature sets and sampling rates, and analyze the tradeoff between recognition accuracy and computational complexity. The classification accuracy on different body positions used for wearing electronic devices was evaluated.

1. Introduction

The primary goal of this paper is to study the effectiveness of activity classifiers in a multi-sensor system as we vary the wearing positions of the sensors. The eWatch is used as a multi sensor platform for wearable context-aware computing. Previous feature extraction studies examining accelerometer data have shown that it is a viable input for detecting user states when it is worn on the wrist [1]. Motivated by other possible sensor platform locations, especially with mobile communication devices such as a cell phone or PDA, we designed a study to investigate the dependency of the eWatch classification accuracy on given different body positions. We investigate wearing the eWatch in the following locations: the belt, shirt pocket, trouser pocket, backpack, and necklace. The results of the study would help us decide on the best position to place such a sensor platform, and understand the nature of the trade-off between wearing position and classification performance.

In [1], the authors used multiple accelerometers worn on a person's body to recognize their physical activity. Sensor data from multiple body positions was combined for classifying the activities.

In [2], a low power sensor hardware system is presented, including accelerometer, light sensor, microphone, and wireless communication. Based on this hardware, a de-



Figure 1. eWatch sensing platform

sign method for a context recognition system is proposed. It evaluates multiple feature sets and makes the tradeoff between power consumption and recognition accuracy. A system that classifies household activities in realtime with a focus on low power consumption is presented in [3].

In [5], a system using an armband based sensor array and unsupervised machine learning algorithms was able to determine a meaningful user context model.

In Section 2, we describe the sensor platform and Section 3 explains the experimental design. Section 4 describes the activity recognition method. Section 5 presents the results of the data analysis, and Section 6 addresses the performance of our on-board activity classifier.

2. Sensing Platform

Our sensor platform, the eWatch (Figure 1), is based on the Philips LPC2106 ARM7 TDMI microcontroller, with 128kB of internal flash memory and 64kB of RAM [6]. The LPC2106 is a 32bit processor running at up to 60Mhz. eWatch contains four sensors: a dual axes accelerometer, light, temperature sensor and microphone. Sensor data can be stored in a 1MB external flash memory.

3. Experiment Design

In our study we focussed on six primary activities: *sitting, standing, walking, ascending stairs, descending stairs*

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE APR 2006		2. REPORT TYPE		3. DATES COVERED 00-00-2006 to 00-00-2006	
4. TITLE AND SUBTITLE Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Carnegie Mellon University, School of Computer Science, 5000 Forbes Ave, Pittsburgh, PA, 15213				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The design of an activity recognition and monitoring system based on the eWatch, multi-sensor platform worn on different body positions, is presented in this paper. The system identifies the user's activity in realtime using multiple sensors and records the classification results during a day. We compare multiple time domain feature sets and sampling rates, and analyze the tradeoff between recognition accuracy and computational complexity. The classification accuracy on different body positions used for wearing electronic devices was evaluated.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

and *running*. Body positions that are normally used for wearing electronic devices, such as cell phones or PDAs, were studied. We placed our sensor hardware on the left wrist, belt, necklace, in the right trouser pocket, shirt pocket, and bag. The subjects wore six eWatch devices located at these body positions during the study. The devices recorded sensor data from the accelerometer and light sensor into their flash memory. The user was asked to perform tasks that consist of the activities, such as working on the computer or walking to another building. The lead experimenter annotated the current activity and instructed the subjects on how to proceed. The annotations were done using an application running on an extra eWatch worn by the lead experimenter.

Six subjects participated in the study, each subject performed the given tasks in 45 to 50 minutes. In total we collected over 290 minutes of sensor data.

Sensor setup eWatch recorded both axes of the accelerometer and the light sensor. All sensors values were recorded with a frequency of 50Hz and with 8bit resolution. The accelerometer was calibrated so that both axes operate in a range of $\pm 2g$. Evaluation of the recorded data was done with Matlab and the WEKA software [7].

4. Activity recognition method

The sensor values recorded from the accelerometers and the light sensor are split into short time windows. These windows are then transformed into the feature space by calculating several feature functions over the individual windows.

Features Features from both accelerometer axes (X & Y), the light sensor, and a combined value of both accelerometer signals were calculated. To reduce the dependency on the orientation, both X and Y values were combined calculating the squared length of the acceleration vector. The classification accuracy with individual sensors as well as with multiple sensor combined was investigated.

Only time domain features were considered to avoid the costly computation that is required to transform the signal into the frequency domain. Table 1 shows the list of features that were considered. The functions to calculate these features were implemented on the eWatch and the required number of clock cycles per function was measured. Each function was executed 2000 times with different recorded sensor inputs, and then the average value was computed. The execution time was calculated based on the measured clock cycles and the CPU frequency at 59MHz. Table 1 shows the measured clock cycles and execution time using a four second window sampled at 20Hz (80 samples).

Figure 2 depicts the feature space after a transformation with Linear Discriminant Analysis (LDA). It shows that the

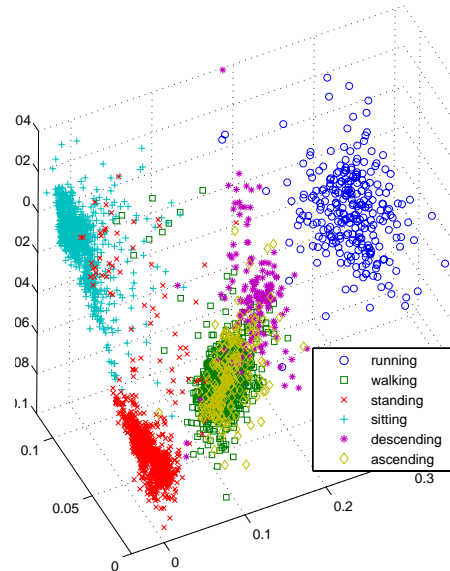


Figure 2. Feature space after LDA transformation

standing, sitting and *running* activities form separate clusters, while *walking, ascending* and *descending stairs* are closer together since these activities are very similar.

Feature Subsets To reduce the time and energy required to calculate the feature vector, several subsets of the complete feature space were evaluated. Some features are irrelevant or redundant and do not provide information to significantly improve the classification accuracy. Therefore a subset of the available features can be selected to decrease the computation time without significantly decreasing recognition accuracy.

The Correlation based Feature Selection (CFS) method from the WEKA toolkit was used to find feature sets con-

Features / Function	Name	Avg. CPU Cycles	Avg. Time in μs
Empirical Mean	<i>mean</i>	854	14.5
Root Mean Square	<i>rms</i>	1219	20.7
Standard Deviation	<i>std</i>	1139	19.3
Variance	<i>var</i>	1313	22.3
Mean Absolute Deviation	<i>mad</i>	1089	18.5
Cumulative Histogram (256 bins)	<i>hist</i>	5847	99.1
n'th Percentile ($n = 5, 10, \dots, 95$)	<i>prc</i>	142	2.4
Interquartile Range	<i>iqr</i>	289	4.9
Zero Crossing Rate	<i>zcr</i>	993	16.8
Mean Crossing Rate	<i>mcr</i>	996	16.9
Sq. Length of X,Y ($x^2 + y^2$)		1318	22.3
Decision Tree classifier (18 nodes)		138	2.3

Table 1. List of time domain features and the average clock cycles and time to calculate them on the eWatch running at 59MHz

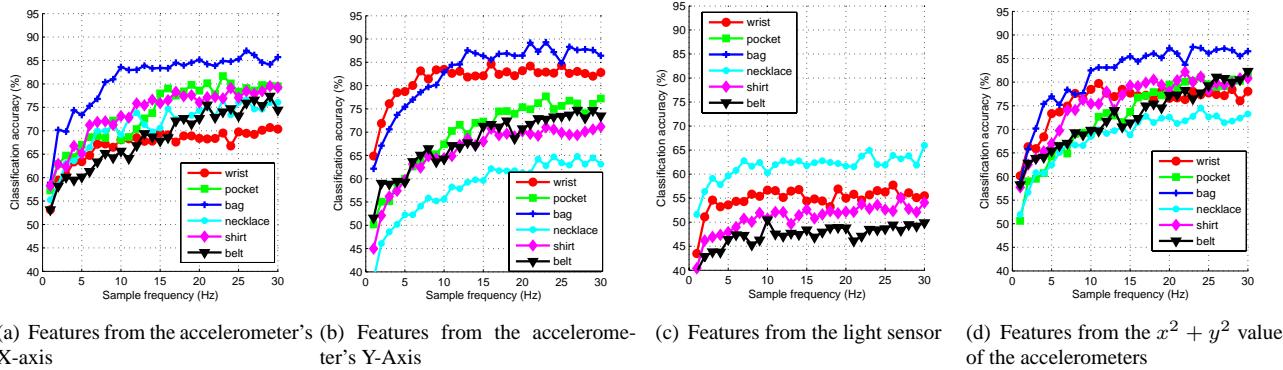


Figure 3. Recognition accuracy with different feature sets

taining features that are highly correlated within the particular class but are uncorrelated with each other. Table 2 shows the feature sets that were compared.

Classification method We evaluated and compared several classification methods, namely Decision Trees (C4.5 algorithm), k-Nearest Neighbor (k-NN), Naive-Bayes and the Bayes Net classifier. Decision Trees and Naive-Bayes were found to achieve high recognition accuracy with acceptable computational complexity. Decision Trees were used for activity classification in [1] and [2]. It was shown in [4] that the discretized version of Naive-Bayes can outperform the Decision Tree classifier for general classification problems. Finally the Decision Tree classifier was chosen as it provides a good balance between accuracy and computational complexity. For all further experiments this classifier with a 5-fold cross validation was used.

Sampling frequency During the user study the sensors were sampled with a frequency of 50Hz and later downsampled to lower frequencies. To maintain some of the high frequency components information and to reduce the computational complexity significantly, no low pass filter was used for downsampling the data. Figure 3 shows the recognition accuracy for different sample rates from 1 to 30Hz for the different body positions. The recognition accuracy was defined as the percentage of correctly classified feature vectors averaged for all six activities. The recognition accuracy increases with higher sampling rates, and with the accelerometer features the accuracy then stabilizes between 15 to 20Hz, and is only improved marginally with higher sampling rates. The accuracy with the light sensor only is lower and it stabilizes beginning with 7Hz. In Figure 3(c) the results from the belt and pocket position is not shown because the light sensor did not provide any useful classification information at these positions.

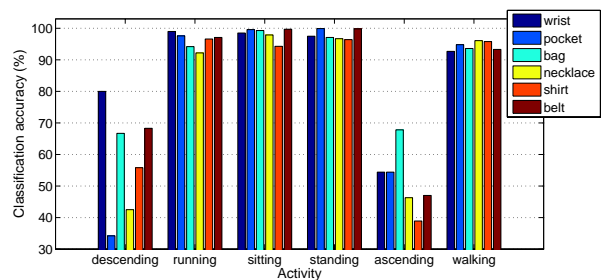


Figure 4. Recognition accuracy for the activities at different body locations

5. Data Analysis - eWatch Classification Performance for Various Wearing Positions

We calculated the classification accuracy for every activity on each of the six different body positions. The data from all subjects was combined to train a general classifier that is not specific to a person.

Table 2 shows the feature sets and the classification results for the different body positions. The features are calculated from a 20Hz signal.

Figure 4 shows the recognition accuracy for the individual activities at different body locations. For the classification the reduced feature set F6 was used. The data indicate that any of the six positions are good for detecting *walking*, *standing*, *sitting* and *running*. *Ascending* and *descending* the stairs is difficult to distinguish from *walking* in all positions, since the classifier was trained for multiple persons. The wrist performs best because the feature set was optimized for the wrist position.

6. Onboard Activity Classifier

Based on these results we implemented a decision tree classifier that runs on the eWatch. The feature set F6 was used to build the decision tree. The sensor sampling is interrupt-based, and triggers the sampling of the sensors at

#	Features	CPU Cycles	Time in μs	Classification Accuracy for Body Position					
				wrist	pocket	bag	necklace	shirt	belt
F1	All features, all sensors	56242	953.6	87.1%	85.2%	92.8%	86.8%	89.5%	87.0%
F2	All features from accelerometer X	14731	249.8	68.4%	78.6%	85.2%	74.3%	76.1%	72.6%
F3	All features from accelerometer Y	14731	249.8	83.2%	75.4%	86.4%	61.3%	70.1%	70.8%
F4	All features from light	14731	249.8	55.0%	16.7%	18.0%	61.7%	52.2%	48.8%
F5	All features from accelerometer XY ($x^2 + y^2$)	15049	255.2	76.6%	79.5%	87.2%	72.6%	78.0%	77.2%
F6	$prc_y(3), rms_{xy}, prc_y(20), prc_y(97), rms_{light}, mad_x, mean_y, prc_y(10)$	12114	205.4	87.0%	80.1%	86.5%	78.6%	79.6%	84.2%
F8	$prc_y(3), iqr_y, prc_y(10), prc_y(97), mad_x$	7651	129.7	82.0%	62.4%	68.9%	56.6%	69.8%	71.7%
F9	$rms_{xy}, qrt_x, rms_x, mad_{xy}, mean_{xy}$	10746	182.2	77.3%	78.2%	80.9%	72.3%	75.4%	76.5%

Table 2. Feature sub sets and classification accuracy for body positions

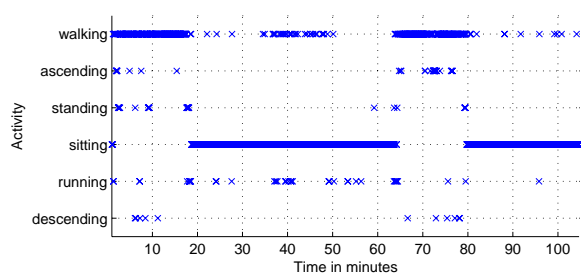


Figure 5. Activity classification recorded over 100 minutes

20Hz. The sensor value is stored in a buffer with the size of the sliding window. The activity is classified every 0.5 seconds based on the sensor data from the 4 second buffer. The classification results are stored into flash memory and are downloaded to a computer later for further processing and analysis. They can also be transferred in realtime over the Bluetooth connection. In order to save energy, the system remains idle between servicing interrupts.

A subject wore the eWatch with the built-in activity classifier on the wrist during the day. The system classified the activity in realtime and recorded the classification results to flash memory. Figure 5 shows 100 minutes of activity classification, as the user walked to a restaurant, sat down, ate lunch, went back to the office and sat down to continue working. The classification results match well with the actual activities; eating lunch was partially interpreted as *walking* or *running* activity due to arm movements.

7. Conclusions and Future Work

The activity recognition and monitoring system that can identify and record the user's activity in realtime using multiple sensors is presented. We compared multiple feature sets and sampling rates to find an optimized classification method, and showed how well they perform on different body locations that are commonly used for wearing elec-

tronic devices.

We will extend our activity classifier to other activities and investigate how the activity classification can support the recognition of the user's location.

7.1. Acknowledgments

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. NBCHD030010, the National Science Foundation under Grant Nos. 0205266 and 0203448, a grant from Intel Corporation, and PA Infrastructure (PITA).

References

- [1] L. Bao and S. S. Intille. Activity Recognition from User-Annotated Acceleration Data. In A. Ferscha and F. Mattern, editors, *Pervasive*, volume 3001 of *Lecture Notes in Computer Science*, pages 1–17. Springer, 2004.
- [2] N. B. Bharatula, M. Stäger, P. Lukowicz, and G. Tröster. Empirical Study of Design Choices in Multi-Sensor Context Recognition Systems. In *IFAWC: 2nd International Forum on Applied Wearable Computing*, pages 79–93, Mar. 2005.
- [3] N. B. Bharatula, M. Stäger, P. Lukowicz, and G. Tröster. Power and Size Optimized Multisensor Context Recognition Platform. In *ISWC 2005: Proceedings of the 9th IEEE International Symposium on Wearable Computers*, pages 194–195, Oct. 2005.
- [4] J. Dougherty, R. Kohavi, and M. Sahami. Supervised and unsupervised discretization of continuous features. In *International Conference on Machine Learning*, pages 194–202, 1995.
- [5] A. Krause, D. P. Siewiorek, A. Smailagic, and J. Farrington. Unsupervised, dynamic identification of physiological and activity context in wearable computing. In *Seventh IEEE International Symposium on Wearable Computers (ISWC'03)*, pages 88–97, 2003.
- [6] U. Maurer, A. Rowe, A. Smailagic, and D. P. Siewiorek. eWatch: A Wearable Sensor and Notification Platform. In *Body Sensor Networks Workshop, submitted*, 2006.
- [7] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 2nd edition, 2005.