ANALYSIS OF LOW RESOLUTION ACCELEROMETER DATA FOR CONTINUOUS HUMAN ACTIVTY RECOGNITION

Narayanan C. Krishnan[*], Sethuraman Panchanathan*

Department of Computer Science and Engineering
Arizona State University
Tempe AZ 85281
USA

ABSTRACT

The advent of wearable sensors like accelerometers has opened a plethora of opportunities to recognize human activities from other low resolution sensory streams. In this paper we formulate recognizing activities from accelerometer data as a classification problem. In addition to the statistical and spectral features extracted from the acceleration data, we propose to extract features that characterize the variations in the first order derivative of the acceleration signal. We evaluate the performance of different state of the art discriminative classifiers like, boosted decision stumps (AdaBoost), support vector machines (SVM) and Regularized Logistic Regression(RLogReg) under three different evaluation scenarios(namely Subject Independent, Subject Adaptive and Subject Dependent). We propose a novel computationally inexpensive methodology for incorporating smoothing classification temporally, that can be coupled with any classifier with minimal training for classifying continuous sequences. While a 3% increase in the classification accuracy was observed on adding the new features, the proposed technique for continuous recognition showed a $2.5 - 3\%$ improvement in the performance.

Index Terms— Accelerometers, human activity recognition, AdaBoost, SVM

1. INTRODUCTION

Recognizing human motion patterns plays an important role in monitoring and understanding human activity. The availability of inexpensive and unobtrusive wearable sensors like accelerometers, gyroscopes, microphones, has opened up a new avenue of research for recognizing human activities, augmenting the traditional vision centric approach. The profound impact of such a system on applications in health care arena for elder care support, long term health monitoring, as a proactive assistive system for individuals with cognitive disorders, or for developing an automated daily life log system, motivates the research in this area.

The different ways in which the continuous data stream from accelerometers can be modeled has resulted in different recognition paradigms. In this paper we divide the continuous acceleration stream in to fixed length frames and classify each frame. We analyze the performance of the different discriminative classifiers on the features extracted from the frames. We propose a generic framework for incorporating temporal continuity for classification on top of the discriminative classifiers for continuous human activity recognition. In addition to the standard features extracted from the raw data, we also explore the effect, of using statistical features computed from the first order derivative of the acceleration signal, on the performance of classification.

Section 2 gives an overview of the work carried out on accelerometer based human activity recognition. In section 3, we describe the feature extraction process and the proposed methodology for classification. We discuss the results obtained in section 4 and finally conlcude and present the future work in section 5.

2. RELATED WORK

The effectiveness of using data from accelerometers placed at five different body locations for recognizing twenty different human activities (a mixture of upper body and lower body motion patterns) was first demonstrated in [1]. Decision trees, k-NN and Bayesian classifiers were trained on global features like mean, spectral energy yielding a best performance of 84%. In [2], the authors evaluated different meta classifiers for recognizing seven lower body motion patterns from a single biaxial accelerometer data and reported the best performance for boosted SVM with a subject independent accuracy of 64%.

[3], combine multiple inertial sensors (accelerometers and gyroscopes) for classifying gestures involved in eating and drinking activities. They employ a generative Gaussian HMM to recognize the individual segmented gestures. [4] compares the performance of activity classification for four different

[*]contact email: narayanan.ck@asu.edu

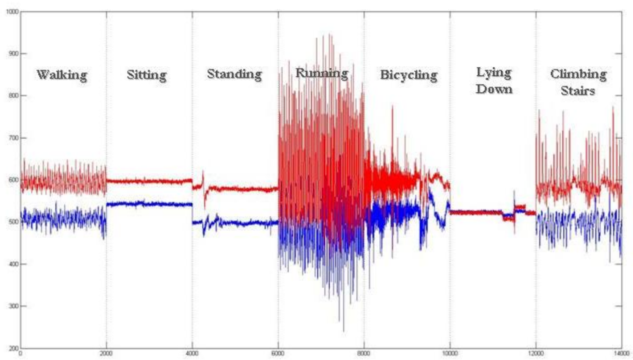

Fig. 1. Sample ankle accelerometer data from X(blue) and Y(red) axis for different activities

configurations of accelerometer placement using HMM. [5] proposes a hybrid discriminative and generative approach for modelling human activities. Features are extracted to train an ensemble of static classifiers to recognize different activities and use HMM to capture the temporal regularities and smoothness of the activities.

In this paper we analyse the performance of discriminative classifiers like AdaBoost, SVM and RLogReg for recognizing seven different activities. An important drawback of these classifiers is that they do not consider temporal information for continuous recognition. We propose a novel framework for adding temporal information on top of the classifier. The proposed technique is generic and can be easily adopted to work with any classifier and does not require to recompute the feature vector nor requires any additional training as proposed in [5]. In addition we also propose features that capture the properties of the data corresponding to activities that have significant variations in motion.

3. EXPERIMENTAL SETUP

The data used for the experiments in this paper is a subset of the data collected by Bao and Intille in [1]. The data was collected in two different ways - supervised approach(activity), where the subject is given explicit instructions about what action to perform, and a semi-naturalistic approach(obstacle), where the subject is given instructions to perform an activity of daily life, that implicitly encodes the action patterns. The data corresponding to 10 random subjects from a pool of 20, for 7 lower body activities namely *walking, sitting, standing, running, bicycling, lying down* and *climbing stairs*, from accelerometers placed at *hip, dominant ankle, and non-dominant thigh*, for the two modes of data collection have been considered for the experiments performed in this paper. Figure 1 depicts typical samples that are obtained from the accelerometers.

3.1. Feature Extraction

The first step in the feature extraction process is to divide the acceleration stream in to frames. We divide the acceleration stream in to frames of size 512 samples, with 256 overlapping samples between successive frames, as described in [2]. For each frame, we compute statistical features like mean, variance, correlation between all the axis of all the accelerometers, along with the spectral features like energy and entropy. For activities that have a significant amount of motion like walking, running, etc the rate at which the acceleration changes is a characteristic property that distinguishes them. We propose to capture these variations by computing statistical features like mean, variance and correlation between all the axis on the first order derivative of the acceleration data in addition to the features mentioned above. The robustness of this feature was validated by computing the classification accuracy obtained by training classifiers on these features.

3.2. Classification

We analyzed the performance of AdaBoost, SVM and regularized logistic regression(RLogReg) on the features extracted from each frame. Binary classifiers were trained for each activity. Given a test sample, the class that yielded maximum margin/probability, was considered as the predicted activity. Human activity is a continuous process and though these techniques are effective in classifying an individual frame, they do not consider temporal continuity for classification. We propose a classification framework that incorporates this temporal continuity of human activity. The proposed framework does not require recomputation of the feature vector nor requires any additional training, thus remains computationally inexpensive.

3.3. Adding Temporal Continuity

The classification margin $m_c(f_t)$, for a frame f_t, belonging to a class c derived either in AdaBoost or SVM reflects the confidence of prediction. This margin can be used by the classifier to output the probability, $p_c(f_t)$, of the frame belonging to class c. A method to compute the probability directly is to fit a sigmoid function to the output of AdaBoost or SVM as described in equation 1.

$$p_c(f_t) = \frac{e^{\phi m_c(f_t)}}{1 + e^{\phi m_c(f_t)}}, \text{where } \phi \text{ is a constant.} \quad (1)$$

The probability values computed for the frame f_i at time instant i can aid in classifying successive temporally close frames. For a frame f_t, let the frames that influence its classification be f_i, where $i = t - \Delta t, \ldots t$. We weight the probability $P_c(f_i)$, for the frame at i belonging to class c, by two factors - a function of i(temporal distance between the frames) denoted by $g(i)$ and a function of the similarity between the current frame and the past frame, measured as the Euclidean distance between them denoted by $h(t-i,t)$. Thus the final probability $P_c(f_t)$ for the frame at t, is given by the equation 2, where the denominator acts as a normalizing factor.

$$P_c(f_t) = \frac{p_c(f_t) + \sum_{i=1}^{\Delta t} g(i) * h(t-i,t) * P_c(f_{t-i})}{\sum_{i=1}^{\Delta t} g(i) * h(t-i,t) + 1}, c = 1 \ldots \quad (2)$$

For the experiments conducted in this paper, we treated the function $g(i)$ as a Gaussian. This was done to ensure that frames that are farther away in time have minimal influence on each other. The function $h(t-i,t)$ was represented as $h(t-i,t) = e^{-\alpha d(f_{t-i}, f_t)}$, where $d(.)$ corresponds to the Euclidean distance between the feature vector describing the frames. This assumes that if adjacent frames are similar, then they should belong to the same class. Though in this paper, we have experimented the framework with AdaBoost and RLogLReg, as a broader impact, the proposed framework can be adopted to work with any classifier.

4. RESULTS AND ANALYSIS

4.1. Feature Analysis

We compared the effect of the features on classification performance of Adaboost. Separate Adaboost classifiers were trained with the standard set of features, statistical features of the first derivative of the acceleration data and a combination of both. The accuracies for the three scenarios were $89.82\%, 81.94\%$ and 92.81% respectively. It is evident that the standard features perform significantly better than the proposed features when it is not combined. However, there was a 3% increase in the accuracy when both the features were combined. Figure 2 gives the class-wise accuracy for the three scenarios. It can be noticed that the proposed features are able to distinguish accurately activities characterised by distinctive motion patterns like walking, running etc(1, 4, 5,and7). The accuracies for these classes are on par with that of the standard features. This indicates that features extracted from the first derivative of acceleration data are able to capture the subtleties in the motion data. Though the features proposed are very rudimentary, the results indicate a scope for improvement through more sophisticated features extracted from the first order derivative of the raw acceleration data.

4.2. Static Classification

We experiment AdaBoost, RLogReg and Linear SVM on the combined feature set for classifying the seven activities. Three different evaluation scenarios were considered for the analysis. For the subject independent scenario, activity data from nine subjects were considered as training samples and the obstacle data from the remaining one subject was the test data. The activity data of all the ten subjects were considered as the training set and the obstacle data from each of the subject formed the test set, for the subject adaptive scenario. The activity and obstacle data from only a single subject formed the training and test set for the subject dependent evaluation.

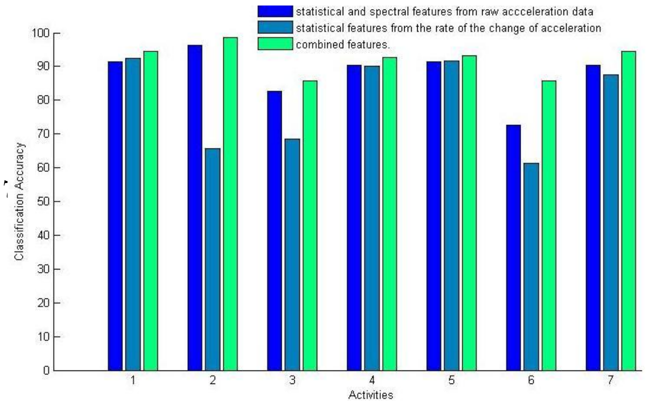

Fig. 2. Class-wise accuracies using AdaBoost trained on the three features.[1 - Walking, 2 - Sitting, 3 -Standing, 4 - Running, 5 - Bicycling, 6 - Lying Down, 7 - Climbing Stairs].

Classifiers	Subject Independent	Subject Adaptive	Subject Dependent
AdaBoost	92.81	93.96	47.88
RLogReg	86.55	88.14	74.56
Linear SVM	82.28	83.60	72.64

Table 1. Subject Independent, Adaptive and Dependent Classification Accuracies.

The results, summarized in table 1 show that AdaBoost performed best in both subject independent and adaptive scenario, while RlogReg had the highest accuracy in subject dependent case. The 90% reduction in the size of the training data for the subject dependent scenario, was the cause for the poor peformance of Adaboost. We did not experiment with kernels for SVM due to the high computational costs associated with them. The confusion matrix for classification aggregated over the 10 subjects for subject independent scenario using AdaBoost is presented in table 2. The misclassification of walking samples as climbing stairs and vice versa, suggests that the motion patterns involved in them are similar. There were also misclassifications occurring between activities that do not involve any quantitative motion in them, probably indicating that necessity of data from other parts of the body.

4.3. Continuous Recognition

We considered the continuous acceleration stream from the obstacle dataset as a sequence of overlapping frames. Each frame was classified using the proposed methodology. The number of past frames considered for classifying the current

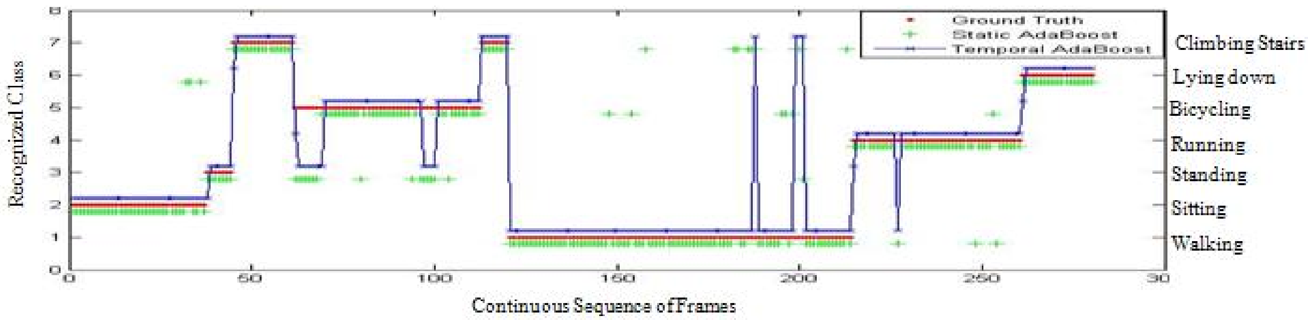

Fig. 3. Output of Static AdaBoost and Temporal AdaBoost compared against the ground truth

Labelled as	A	B	C	D	E	F	G
A (total = 889)	840	0	18	2	9	0	20
B (total = 300)	0	296	0	0	0	4	0
C (total = 150)	0	13	128	0	8	0	1
D (total = 494)	10	0	6	458	11	0	9
E (total = 435)	0	0	32	0	443	0	0
F (total = 400)	0	57	0	0	0	343	0
G (total = 342)	13	0	4	1	1	0	323

Table 2. The aggregate confusion matrix obtained from subject independent 10 fold cross validation using AdaBoost trained on combined features. The total corresponds to total number of frames belonging to that activity.[A - Walking, B - Sitting, C -Standing, D - Running, E - Bicycling, F - Lying Down, G - Climbing Stairs]

frames was varied. The optimal performance was achieved when three past frames were considered for classifying the current frame. We experimented the framework on Adaboost and RLogReg. While adding temporal information to static adaboost resulted in an average 10 fold cross validation accuracy of 95.35%, RLogReg resulted in 89.63%. For both the algorithms, an improvement of about $2.5 - 3\%$ was observed. Figure 3, illustrates the effect of adding the temporal component to the static AdaBoost classifier for one subject. There is a reduction in the number of misclassifications by the blue line that corresponds to the classification result of adaboost with temporal component added to it.

5. CONCLUSIONS

We have demonstrated the effectiveness of different discriminative classifiers for human activity recognition from low resolution accelerometer data. The experiemnts we have conducted have shown that complementing the standard features with rudimentary statistical features extracted from the first order derivative of the accelerometer data enhances the classification performance. The superiority of AdaBoost for subject independent classification was observed. The proposed technique for adding temporal continuity to the classification yielded promising results with about $2.5 - 3\%$ improvement in accuracy.

In future, we would like to experiment extensively with new features, that would capture the characteristics of motion patterns across different activities. While, only lower body activities was the focus in the paper, we plan to conduct experiments involving wide range of upper body activities and possibly simultaneous occurrences of both.

6. REFERENCES

[1] Ling Bao and Stephen S. Intille, "Activity recognition from user-annotated acceleration data.," in *Proceedings, Pervasive Computing, Second International Conference*, 2004, pp. 1–17.

[2] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L. Littman, "Activity recognition from accelerometer data.," in *Proceedings, The Seventeenth Innovative Applications of Artificial Intelligence Conference*, 2005, pp. 1541–1546.

[3] Oliver Amft, Holger Junker, and Gerhard Tröster, "Detection of eating and drinking arm gestures using inertial body-worn sensors.," in *Proceedings, Ninth IEEE International Symposium on Wearable Computers (ISWC)*, 2005, pp. 160–163.

[4] Daniel Olgun Olgun and Alex (Sandy) Pentland, "Human activity recognition: Accuracy across common locations for wearable sensors.," in *Proceedings, IEEE 10th International Symposium on Wearable Computing (Student Colloquium Proceedings)*, 2006.

[5] Jonathan Lester, Tanzeem Choudhury, Nicky Kern, Gaetano Borriello, and Blake Hannaford, "A hybrid discriminative/generative approach for modeling human activities.," in *Proceedings, Nineteenth International Joint Conference on Artificial Intelligence*, 2005, pp. 766–772.