Single-Accelerometer-Based Daily Physical Activity Classification

Xi Long, *Student Member, IEEE*, Bin Yin, and Ronald M. Aarts, *Fellow, IEEE*

Abstract—In this study, a single tri-axial accelerometer placed on the waist was used to record the acceleration data for human physical activity classification. The data collection involved 24 subjects performing daily real-life activities in a naturalistic environment without researchers' intervention. For the purpose of assessing customers' daily energy expenditure, walking, running, cycling, driving, and sports were chosen as target activities for classification. This study compared a Bayesian classification with that of a Decision Tree based approach. A Bayes classifier has the advantage to be more extensible, requiring little effort in classifier retraining and software update upon further expansion or modification of the target activities. Principal components analysis was applied to remove the correlation among features and to reduce the feature vector dimension. Experiments using leave-one-subject-out and 10-fold cross validation protocols revealed a classification accuracy of ~80%, which was comparable with that obtained by a Decision Tree classifier.

I. INTRODUCTION

Human physical activity recognition has been receiving increasing attention in recent years. Human behavior and its classification are significant for the disciplines such as medicine, behavioral sciences, physiotherapy, etc [1]. An accelerometer is an inexpensive, effective and feasible body-worn sensor which has been frequently used in daily physical activity classification [2], [3].

The majority of earlier research was based on multiple body-worn sensors placed at different positions such as chest, thigh, waist, ankle, knee, and so on [4]-[10]. The main attention of those studies was to classify a specific subset of activities for a certain application [5], [8], [11], [12]. For instance, wrist and arm sensors were employed for the classification of upper body movements [13]. For customer's lifestyle applications, however, a single sensor solution without any requirement on sensor fixation would be preferred. Fixing multiple sensors to the body may pose restrictions on physical activities performed in an everyday life context, and would be cumbersome to use. Scientific proof has been reported on the feasibility of wearing a single accelerometer in daily energy consumption assessment [14]-[16].

In this study, the acceleration data was collected in a

naturalistic, real-life environment. Previous research mostly used the supervised and semi-naturalistic approaches in a lab environment for data collection where the subject was given explicit instructions [5], [17]. The studies using only supervised data for training and only naturalistic data for validation showed limited classification accuracies [18]. Promising performance was shown in [19] by a continuous activity recognition (CAR) algorithm that used activity data collected with a single accelerometer in a real-life environment without researchers' intervention. The purpose of the study was to get the daily activity pattern of a user and in combination with energy consumption assessment help to change to a healthier lifestyle in an easy and encouraging manner. For this, five typical daily physical activity elements, i.e., walking, running, cycling, driving, and sports were chosen. In the previously developed CAR algorithm, a Decision Tree (DT) classifier was employed that, in general, showed the best performance in activity classification in previous studies [5], [20].

From the implementation point of view, however, a DT classifier has to be completely re-built if the activity set changes or/and new features are incorporated. In addition, because the tree training is normally completed on isolated activity events, it requires possibly manual tuning, thus involvement of experts especially when applied on a continuous data trace recorded in a real application. A system realizing the function of the classification algorithm can not be updated without considerable efforts. Thus, a DT-based algorithm has a low extensibility or poor forward compatibility at this aspect. An algorithm based on Bayesian classification, on the other hand, is advantageous in incorporating additional features and/or activities through expanding the dimension of the feature vector and estimating feature probability density functions (PDFs) for new activities and/or features. Also the Bayesian classification algorithm achieved comparable performance with the DT-based algorithm shown in [15]. This paper, as extension of the work in [19], [21], aims at studying a Bayes classifier combined with a Parzen window estimator [22] and compares its performance with that of a DT classifier.

II. DATA COLLECTION

Activity data was collected by the Philips NWS Activity Monitor (URL: http://www.newwellnesssolutions.com) with a built-in tri-axial accelerometer. It is a light and small-sized portable device (3×3 cm^2) which can be worn easily and

unobtrusively in an arbitrary orientation on the human body in a free-living environment. The sample rate is 20 Hz.

The aim was to classify five activities including walking, running, cycling, driving, and sports, which generally are the main activities contributing to daily activity-related energy expenditure (AEE). The monitor was used to collect the acceleration data of 24 subjects consisting of 13 males and 11 females, ranging in age from 26 to 55 (mean 33.6, standard deviation 7.9). For each subject, the data about 10 hours was recorded without researchers' intervention during which the subject conducted physical activities as normally in his/her everyday life. During the data collection, the sensor was placed without being fixed in a certain orientation on the subject's waist. No special requirement has been announced to subjects about how to wear the sensors except for the location. Subjects were asked to annotate main activities with start and finishing times.

In order to build the ground truth, acceleration data was visually inspected with reference to the annotation sheets to determine the classes. Some abstraction was made to cluster sub-activities. For instance, walking stairs and training on a cross-trainer were categorized as walking; taking public transport such as bus and train belonged to driving; soccer, volleyball, badminton, boxing, table tennis, fitness, etc. fell into the class of sport.

III. FEATURE EXTRACTION

Nineteen features have been defined which can be separated into three categories, namely features in time domain, frequency domain and spatial domain [21]. Several examples are explained as below:

• In time domain, the "standard deviation" of the data in a frame was calculated. It has been shown that for a physical activity there is a consistent relationship between the standard deviation of the acceleration data and intensity of the movement during the activity.

• In frequency domain, "frequency-domain entropy" helped distinction of activities with a similar energy intensity by comparing their periodicities. This feature was computed as the information entropy of the normalized power spectral density (PSD) function of the input signal without including the DC component. The feature "periodicity" evaluated the periodicity of the signal that helps to distinguish cyclic and non-cyclic activities.

• In spatial domain, "orientation variation" was defined by the variation of the gravitational components at three axes of the accelerometer sensor. This feature effectively shows how severe the posture change can be during an activity. Other features looked at inertial accelerations in the vertical direction and horizontal plane, as well as their relation, which may provide with information distinct from activity to activity.

Before the feature calculation, the raw data was pre-processed to improve the signal-to-noise ratio (SNR). A low pass filter (LPF) with a cut-off of 5 Hz was applied to filter out high frequency noises. In the dataset, the acceleration data stream was divided into consecutive frames with 16 seconds each. Then a vector of 19 features was calculated over all the frames.

IV. BAYESIAN CLASSIFICATION

A. PDF Estimation

The activity classes were represented by C, which took value $c=c_k$ with $k=1,2,\ldots,5$. f_i represented the value of the feature F_i ($i=1,2,\ldots,19$). The PDF of the values of each feature F_i given a class was obtained by fitting its distribution (implemented in Matlab) based on Parzen-window method using Gaussian kernel [22]. In total, 5×19 conditional PDFs were estimated in this study.
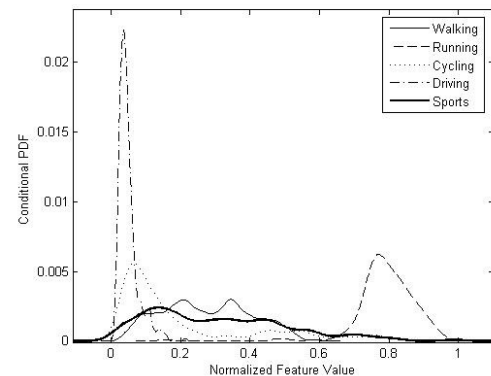

Fig. 1. Normalized training conditional PDFs of "standard deviation" of the five classes.

Taking the feature "standard deviation" as an example, Fig. 1 depicts its normalized conditional PDFs of the five classes which represent the intensity levels for each activity. It can be seen that running and cycling (or driving) are significantly separated due to the different levels of acceleration generated. Walking is in between driving and running. The acceleration power of sports is more spread because many different types of sports with both high and low intensities were included in the dataset. Hence, "standard deviation" may be a good feature for recognizing running from others. In addition, the figure suggests that the overlapping areas may generate classification errors.

B. Classification

The corresponding Bayesian classification function C is defined as below

$$C(f_1,...,f_n) = \arg\max_c p(C = c)\prod_{i=1}^{n} p(F_i = f_i \mid C = c) \quad (1)$$

where $c=c_k$ ($k=1,2,\ldots,5$) and $n=19$ in this case. The multiplication holds when features F_i are mutually independent in (1). It also implies that the class c_k with the maximum conditional joint probability is selected.

C. Principal Components Analysis

In naive Bayesian classification, the joint conditional

probability is calculated as if the features $F_1,...,F_n$ are mutually independent. In the current study, this condition may not hold, which could deteriorate the classification performance. Principal components analysis (PCA) is an available matrix conversion approach which represents a set of vectors in a new space with usually a lower dimension. The vectors in the new space are mutually uncorrelated (equivalently independent only if the vectors are normally distributed). PCA has been widely used to identify patterns in data, and express the data so as to highlight their similarities and differences. Additionally, an important advantage of PCA is that effectively the dimension of the feature space may be reduced without much loss of information. After PCA, the naive Bayes (NB) classifier becomes less "naive". Normally the number of PCAed features is less than the original number of features as a result of redundancy removal.

V. RESULTS AND DISCUSSION

In order to evaluate the naive Bayesian classification, two cross validation protocols were used, namely the "leave-one-subject-out" cross validation (LOSO CV) and "10-fold" cross validation (10-fold CV). The latter was conducted also for the purpose of comparison with the result of the decision tree algorithm which has been verified in the previous work.

The classification performances with and without applying PCA were evaluated. The overall accuracies without PCA were 79.3% and 71.5% using LOSO CV and 10-fold CV, respectively. Fig. 2 gives the mean classification accuracies using the two cross validations with PCA versus the number of chosen PCAed features, with the selection priority on the ranking of their corresponding eigenvalues listed in a descending order. The figure shows that the performance improvement of using PCA is not significant but the NB classifier with PCA using the first 5 PCAed features already provided an optimal performance with overall accuracies of 79.5% in LOSO CV and 72.3% in 10-fold CV, meaning that there is redundancy in features and with PCA the complexity of calculation may be largely reduced without losing classification accuracy. The accuracy with more PCAed
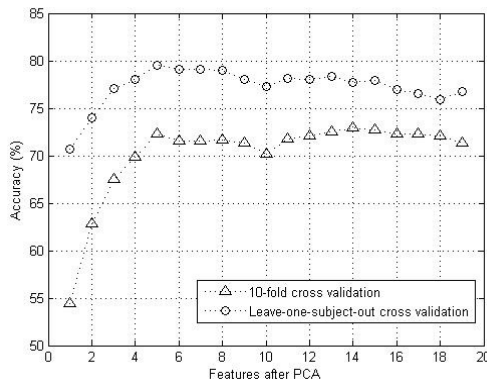

Fig. 2.   Mean classification accuracy (%) versus the number of features (1-19) after PCA for classification.

features being used did not improve further, and instead dropped to some extent, especially for LOSO CV. This may be due to that the remaining components hardly contributed to the performance because the new information carried was limited compared to the increase of the noise level.

The classification accuracies (mean ± standard deviation) using LOSO CV and 10-fold CV were 79.5% ± 11.6% and 72.3% ± 4.1%, respectively. The accuracy using 10-fold CV was lower than that using LOSO CV, simply due to that the training dataset in the former case was smaller. In addition, the standard deviation of LOSO CV was relatively large, possibly resulting from the imbalance in types and durations of activities performed by each subject, which is typical with activity data collected from everyday life.

TABLE I
SUMMARY OF CLASSIFICATION ACCURACIES PER CLASS BASED ON NB
CLASSIFIER AFTER PCA APPLIED WITH THE FIRST 5 FEATURES

Class	LOSO CV	10-fold CV
Walking	80.3%	77.3%
Running	92.9%	95.5%
Cycling	49.4%	49.4%
Driving	94.3%	86.9%
Sports	71.3%	57.9%

TABLE II
AGGREGATE CONFUSION MATRIX BASED ON NB CLASSIFIER WITH PCA
USING LOSO CV

Class	Walking	Running	Cycling	Driving	Sports
Walking	**237**	3	15	10	21
Running	6	**144**	2	0	3
Cycling	76	0	**125**	41	11
Driving	5	0	7	**295**	6
Sports	33	2	23	17	**186**

Table I summarizes the classification results per class using the two cross validation protocols. Because of the imbalance problem mentioned above, only the aggregate accuracies were given. The aggregate confusion matrix of the naive Bayesian classification with PCA (with the first 5 features) using LOSO CV is given by Table II. From the tables, it can be clearly seen that both running and driving achieved high accuracies, which may attribute to their distinct characteristics in generating acceleration among the five activities. Running generates very large accelerations at a certain frequency, whereas driving produces very low non-periodic acceleration signals. The classification accuracies of walking appeared moderate. For sports, the diversity and complexity of the movement should be responsible for its relatively low accuracies. Cycling gave the lowest classification accuracy of about 50%, which may be explained by the less representative acceleration generated when a sensor is placed at the waist, as well as the huge inter-subject and intra-activity variations. The false negative of cycling mainly went to walking and driving, in which a significant amount was substituted by walking (see Table II). This confirms that cycling is highly spread in feature space

TABLE III
PERFORMANCE COMPARISON OF THE THREE CLASSIFIERS USING
10-FOLD CV

Classifier	Aggregate Accuracy
DT Classifier	72.8%
NB Classifier	71.5%
NB Classifier with PCA [a]	72.3%

[a]with 5 features after PCA chosen

and thus overlaps with other activities.

The comparison between a DT classifier and the NB classifiers (both without and with PCA) using 10-fold CV is shown in Table III. It is clear that compared with the DT-based method, which has been investigated in [19], the NB classifier with PCA provides an accuracy level with a negligible drop (<1%) whereas having a more extensible algorithm structure for software implementation when new features/activities need to be incorporated.

In this study, the activity classes are semantically overlapping. One of the main ambiguities exists between walking (or running) and sports. In a real application, misclassifying walking or running as sports or vice verse (see Table II) may be accepted by users because they are not exclusive to one another to some extent. Then, studying the user perceived accuracy is interesting and valuable. After removing the misclassification between sports and walking/running, the accuracy had a significant increase (>4%).

Post-processing is crucial in activity classification in particular on continuously recorded real-life data. It aims to improve the classified results usually by using smart decision fusion techniques, for instance, smoothing or reasoning with real-life experiences. Significant improvement in classification accuracy (>5%) has been observed in [19]. Thus, a similar result is also expected with a Bayes classifier.

VI. CONCLUSION

Human daily physical activity classification using a single tri-axial accelerometer placed on the waist without fixation was studied. The data were collected in the naturalistic environments without researchers' intervention. The NB classifier with PCA provides a classification accuracy of ~80% on a variety of 5 classes, with 5 features after PCA compared to 19 before PCA. PCA was applied with a success in removing redundancy in features and thus reducing computational complexity. Compared to a DT classifier, the results showed that the Bayes classifier achieved a similar performance whereas having a more extensible algorithm structure for software implementation when new features or activities need to be incorporated.

REFERENCES

[1] J.B.J. Bussmann, *et al.*, "Measuring daily behavior using ambulatory accelerometry: The Activity Monitor," *Behav. Res. Meth. Instrum. Comput.*, vol. 33, no. 3, 2006, pp. 349-356.

[2] B.P. Clarkson. Life Patterns: Structure from Wearable Sensor. *PhD Thesis*, MIT Media Lab, 2002.

[3] J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannaford, "A Hybrid Discriminative/Generative Approach for Modeling Human Activities," in *Proc. 9th Int. Joint Conf. Artif. Intell.*, Edinburgh, Scotland, 2005, pp. 766-772.

[4] C. Randell and H. Muller, "Context Awareness by Analysing Accelerometer Data," *Proc. 4th ISWC*, 2000, pp. 175-176.

[5] L. Bao and S.S. Intille, "Activity Recognition from User-Annotated Acceleration Data," *Proc. 6th Int. Conf. on Perv. Comp.,* Vienna, Austria, 2004, pp. 1-17.

[6] P. Lukowicz, T.E. Starner, G. Troster, and J.A. Ward, "Activity Recognition of Assembly Tasks Using Body-Worn Microphones and Accelerometers," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1553-1567, 2006.

[7] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing Human Motion with Multiple Acceleration Sensors," in *Proc. IEEE Conf. Syst. Man. Cybern.*, Tucson, AZ, 2001, vol. 2, pp. 747-752.

[8] S.W. Lee and K. Mase, "Activity and Location Recognition Using Wearable Sensors," *IEEE Perv. Comp.*, vol. 1, pp. 24-32, 2002.

[9] N. Kern, B. Schiele, and A. Schmidt, "Multi-Sensor Activity Context Detection for Wearable Computing," in *Proc. 1st Euro. Symp. Ambient Intel.*, LNCS 2875, pp. 220-232, 2003.

[10] J. Parkka, M. Ermes, *et al.*, "Activity Classification Using Realistic Data From Wearable Sensors," *IEEE Trans. Info. Tech. Biomed.*, vol. 10, no. 1, pp. 119-128, 2006.

[11] S.M. Patterson, *et al.*, "Automated physical activity monitoring: Validation and comparison with physiological and self- report measures," *Psychophysiology*, vol. 30, no. 3, pp. 296-305, 1993.

[12] B. Najafi, K. Aminia, *et al.*, "Ambulatory System for Human Motion Analysis Using a Kinematic Sensor: Monitoring of Daily Physical Activity in the Elderly," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp.711-723, 2003.

[13] F. Foerster, M. Smeja, and J. Fahrenberg. Detection of posture and motion by accelerometry: a validation in ambulatory monitoring. *Comp. in Hum. Behav.*, vol. 15, pp. 571-583, 1999.

[14] M.J. Mathie, A.C. Coster, N.H. Lovell, and B.G. Celler, "Detection of daily physical activities using a triaxial accelerometer," *Med. & Biol. Eng. & Comp.*, vol. 41, pp. 296-301, 2003.

[15] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity Recognition From Accelerometer Data," in *Proc. 17th Innov. Appl. Artif. Intell. Conf.*, Pittsburgh, PA, 2005, pp. 1541-1546.

[16] G. Plasqui, *et al.*, "Measuring free-living energy expenditure and physical activity with triaxial accelerometry," *Obesity Research*, vol. 13, no. 8, pp. 1363-1369.

[17] N.C. Krishnan and S. Panchanathan, "Analysis of Low Resolution Accelerometer Data for Continuous Human Activity Recognition," in *Proc. IEEE Int. Conf. Acoust. Speech. Sig. Proc.*, Las Vegas, NV, 2008, pp. 3337-3340.

[18] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, "Detection of Daily Activities and Sports With Wearable Sensors in Controlled and Uncontrolled Conditions," *IEEE Trans. Info. Tech. Biomed.*, vol. 12, no. 1, pp. 20-26, 2008.

[19] B. Yin and A.H.C. Goris, "Continuous Recognition of Daily Physical Activities Using A Single Triaxial Accelerometer," *ICAMPAM*, Rotterdam, The Netherlands, 2008, p. 90.

[20] K. Aminian, *et al.*, "Physical activity monitoring based on accelerometry: validation and comparison with video observation," *Med. & Biol. Eng. & Comp.*, vol. 37, no. 3, pp. 304-308, 1999.

[21] B. Yin and A.H.C. Goris, "Detection of Sensor Wearing Positions for Accelerometry-based Daily Activity Assessment," *6th IASTED Conf. Biomed. Eng.*, Innsbruck, Austria, 2008, pp. 390-395.

[22] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.,* vol. 33, no. 3, pp. 1065-1076, 1962.