Energy Expenditure Estimation
with Wearable Accelerometers

Mitja Luštrek, Božidara Cvetković and Simon Kozina

Department of Intelligent Systems
Jožef Stefan Institute
Ljubljana, Slovenia
{mitja.lustrek, boza.cvetkovic, simon.kozina}@ijs.si

Abstract—**This paper presents a method for human activity recognition and energy expenditure estimation with two tri-axial accelerometers. Recognizing the activity of a person and measuring his/her energy expenditure is important for the management of several diseases. In the CHIRON project we aim to monitor congestive heart failure patients using wearable sensors and a smartphone. Our method uses a classifier for activity recognition constructed with machine learning. Attention was paid to the complexity of the attributes for machine learning, resulting in the omission of the most complex attributes in order to prolong the battery life. The recognized activity serves as an input to a classifier for energy expenditure estimation, which was also constructed with machine learning. The best-performing classifier turned out to be a composite of two activity-specific classifiers and a general classifier. Its mean absolute error was 0.91 metabolic equivalents of task (MET).**

I.   INTRODUCTION

The energy expenditure of a person closely corresponds to the amount of physical activity. This information is important for medical, lifestyle and sports-training purposes. Most research on the subject was probably done in the sports medicine to optimize the training and diet of athletes. However, the amount of physical activity is also critical for cardiovascular health [1], which was the focus of our work.

There are several methods for estimating the human energy expenditure. Direct calorimetry [2] measures the heat output of a person. It is the most reliable method, but can only be used in laboratory conditions. It measures the energy expenditure over medium periods of time (hours). Indirect calorimetry [3] measures the amount of inhaled and exhaled oxygen and CO_2, which is linked to the expended energy. It is fairly reliable and can be used in field conditions, even over short periods of time (minutes). However, it cannot be used in everyday life, since it requires a breathing mask. Doubly-labeled water [4], which is water labeled with deuterium and oxygen-18, can be used to measure the amount of exhaled CO_2. It is also fairly reliable and can in principle be used in everyday life, but only over long periods of time (days or weeks). It is also quite expensive. A diary of activities is very

simple, but tends to be unreliable. Finally, accelerometers and other wearable sensors are moderately reliable, inexpensive, and can be used in everyday life over short periods of time. Given the increasing ubiquity of such sensors, they seem to be the most promising tool for energy expenditure estimation.

In this paper we present a method for activity recognition and energy expenditure estimation using two wearable accelerometers. The method was developed in the CHIRON project [5], which will monitor congestive heart failure patients using wearable sensors connected to a smartphone. The activity type, energy expenditure and their relation to the heartbeat are important for the health of such patients. Two accelerometers were judged to be the right number considering the tradeoff between the reliability of the activity recognition and energy expenditure estimation on one hand, and the burden on the patients on the other hand. A heart-rate monitor and skin-temperature sensor were also considered, but experimental results showed no benefit. This is not surprising, since changes in the heart rate and skin temperature can have many causes.

The rest of the paper is structured as follows: after discussing the related work in Section II, we present the experimental setup in Section III. The development and testing of the method for activity recognition and energy expenditure estimation are described in Section IV, and the conclusions are drawn in Section V.

II.   RELATED WORK

Most methods for energy expenditure estimation using wearable sensors seek linear or nonlinear relations between the energy expenditure and the accelerometer outputs. The most basic methods use one accelerometer and one linear regression model. The estimation accuracy could be improved by multiple regression models [6] and complex attributes [7]. The regression method by Crouter et al. [8] is currently among the most accurate. It uses one accelerometer attached to the hip. In the first step it classifies a person's activity into sitting, ambulatory activity or lifestyle activity. In the second step it uses a linear regression model for the

ambulatory activity and an exponential regression model for the lifestyle activity. Sitting is always considered to have the energy expenditure of 1 metabolic equivalent of task (MET, 1 MET is the energy expended at rest).

Some methods use a human kinematic model, from which the kinetic energy required for the movement can be computed [9]. The movement is estimated using accelerometers attached to the body. However, this requires the integration of acceleration data, which can amplify sensing errors, and the use of many accelerometers.

III. HARDWARE AND DATA

We used Shimmer accelerometers [10], which – apart from a tri-axial accelerometer – include a microcontroller, and a Bluetooth and 802.15.4 radio. The accelerometers measure the acceleration along three perpendicular axes with a configurable frequency (we chose 50 Hz). Because of the limitations of the available smartphones, we used Bluetooth to transmit the data to the smartphone or PC, where the activity recognition and energy expenditure estimation took place.

To develop and test our method, data consisting of acceleration measurements during activities with different energy expenditures were required. The data was recorded for five subjects. Each subject wore four accelerometers on the chest, right ankle, right thigh and right wrist, to study the performance of different sensor placements. Each subject performed a test scenario that served two purposes. The first was to provide the basic activity types that we wanted to recognize, and the second was to provide a range of activities with different energy expenditures (basic and complex). The activities in the scenario are shown in Table I. The basic activities to be recognized are labeled as such. The MET values for the energy expenditure were sourced from The Compendium of Physical Activities [11].

TABLE I. ACTIVITIES IN THE TEST SCENARIO

Activity	Basic	Energy expenditure (MET)
Lying	Yes	1.0
Sitting	Yes	1.0
Standing	Yes	1.2
Walking	Yes	
... normally (4.8 km/h)		3.3
... quickly (6.4 km/h)		5.0
... downstairs		3.0
... upstairs		5.0
Running	Yes	
... slowly (8 km/h)		8.0
... normally (10.8 km/h)		11.0
Cycling	Yes	
... normally (20.9 km/h)		8.0
... quickly (25.7 km/h)		11.0
... stationary normally (100 W)		7.0
... stationary quickly (150 W)		10.5
Kneeling	Yes	1.0
On all fours	Yes	1.5
Sitting doing light activities		1.5
Walking doing light chores		2.3
Scrubbing the floor		3.8
Sweeping outdoors		4.0
Sit-ups		8.0
Jumping jacks		8.0

IV. METHOD AND RESULTS

Our main task was to estimate the energy expenditure of a person, but we also wanted to recognize the type of his/her activity. The latter is both interesting in itself and a useful input to the energy expenditure estimation. Both tasks were tackled in essentially the same way. The stream of acceleration data was split into windows of length N, each window overlapping with the previous one by one half of its length. Several attributes were computed from the acceleration within each window. These attributes formed a vector, which was fed into a machine learning algorithm. The algorithm constructed a classifier which either recognized the activity within the window or estimated the energy expenditure. The former is a classification task and the latter regression, but for a lack of a better term we use "classifier" for both.

In both tasks three combinations of accelerometer placements were tested. One accelerometer was always placed on the chest, since earlier experiments showed this to be optimal for the activity recognition [12]; in addition the torso represents the largest part of the body mass, requiring the most energy to move. The other accelerometer was placed either on the ankle, thigh or wrist. The performance of the classifier for both the activity recognition and energy expenditure estimation was tested using the leave-one-person-out method. This means that the classifier was trained on four subjects and tested on the fifth. The procedure was repeated five times, using a different subject for testing each time.

A. Activity Recognition

The length of the window for the activity recognition was set to 2 s, which means that each window contained $N = 100$ acceleration measurements. Longer windows yielded higher recognition accuracies, but some short activities were missed if the length was more than 2 s. The following attributes were computed for each window. Since the method must run on a smartphone and should conserve the battery charge, we had to pay attention to the time complexity of the attribute computation; it is listed for each group of attributes.

- The average acceleration along the three axes of the accelerometer, and the average length of the acceleration vector; O (N)
- The variance of the above four accelerations, the index of dispersion (= variance / average), and the detection of movement (= index of dispersion > 0.016, determined experimentally); O (N)
- The maximum and the minimum of the four accelerations, and the difference thereof; O (N)
- The speed of change between the maximum and the minimum of the four accelerations; O (N)
- The orientation along the three axes; O (N)
- The sum of the absolute differences between the consecutive lengths of the acceleration vector; O (N)
- The frequency of the movement along the y-axis (vertical when the person is upright) computed by FFT, and the energy (= squared sum of FFT elements); O ($N \log N$)
- The correlation between all the pairs of axes; O (N^2)

The activity recognition was tested with four machine learning algorithms implemented in the Weka toolkit 0 using the default parameter values: Naïve Bayes, C4.5 decision trees, Random Forest and Support Vector Machine (SVM). The results are shown in Table II. One can see that the Random Forest was the best algorithm, while the differences between the sensor placements were small. The best placement seems chest + ankle (the best accuracy is underlined in the table). The experiment with the Random Forest algorithm and the chest + ankle placement was repeated without the two attribute groups with the above-linear complexity. The resulting accuracy was 85.1 %, which is better than the 82.3 % underlined in Table II, so these attributes were omitted.

TABLE II.   CLASSIFICATION ACCURACY OF ACTIVITY RECOGNITION (%)

	Naïve Bayes	C4.5	Random Forest	SVM
Chest + ankle	68.0	72.8	<u>82.3</u>	81.6
Chest + thigh	69.5	70.1	80.3	75.8
Chest + wrist	56.8	71.4	81.1	75.4

B. Energy Expenditure Estimation

The length of the window for the energy expenditure estimation was set to 10 s, which means that each window contained $N = 500$ acceleration measurements. Some of the attributes could not be reasonably computed for shorter windows, and we had no practical need for finer-grained estimation. The following attributes were computed for each window. Time complexities are all linear or constant, so we did not consider omitting any attributes.

- The most prevalent activity computed by the activity recognition; $O(N)$
- The average length of the acceleration vector; $O(N)$
- The area under the absolute acceleration along the three axes of the accelerometer, the area under the length of the acceleration vector, and the four areas squared; $O(N)$
- The sum of the areas and the squared areas of the acceleration along the three axes; $O(1)$
- The area under the absolute gravity-subtracted acceleration along the three axes; $O(N)$
- The number of times the length of the acceleration vector stops increasing and starts decreasing or vice versa; $O(N)$
- The change in the velocity of the accelerometer along the three axes; $O(N)$
- The integral of the change in the kinetic energy due to the movement along the three axes; $O(N)$
- The sum of the integrals of the changes in the kinetic energy along the three axes; $O(1)$

The energy expenditure estimation was tested with six machine learning algorithms implemented in the Weka toolkit 0 using the default parameter values: Linear Regression, Multi-Layer Perceptron artificial neural network, Support Vector Regression (SVR), M5P model trees, M5Rules and REPTree regression trees. The results are shown in Table III.

One can see that the REPTree is the best algorithm. However, since it constructs regression trees, it can only output a limited number of energy expenditures based on the training data, and cannot interpolate and extrapolate on new data. Therefore we considered the most suitable algorithm to be the Multi-Layer Perceptron. The best sensor placement was wrist + ankle (the error of the most suitable algorithm is underlined in the table).

TABLE III.   MEAN ABSOLUTE ERROR OF ENERGY EXPENDITURE ESTIMATION (MET) USING THE GENERAL CLASSIFIER

	Linear Regression	Multi-Layer Perceptron	SVR	M5P	M5Rules	REPTree
Chest + ankle	5.09	1.65	3.29	2.18	1.56	1.41
Chest + thigh	6.75	1.66	3.68	2.38	5.97	1.58
Chest + wrist	6.57	<u>1.39</u>	3.94	4.95	4.69	1.30

Fig. 1 shows the estimated vs. true energy expenditure using the Multi-Layer Perceptron. The largest errors occurred at large energy expenditures, which correspond to running and cycling. The reason for this is probably that most activities were not so energetic, which caused the classifier to adapt more strongly to the less energetic ones. We tackled this problem by constructing separate classifiers for running and cycling. The classifiers were trained on the recordings of these two activities only, and used to classify the activities recognized as running and cycling.
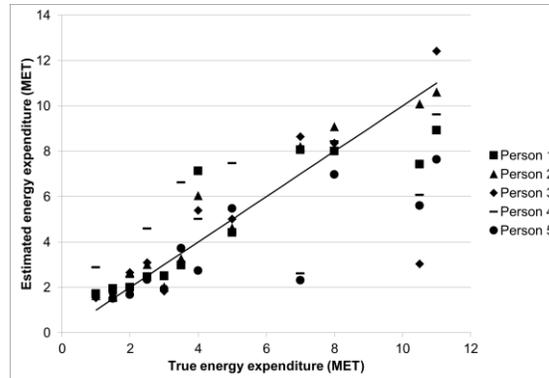
Figure 1.   Estimated vs. true energy expenditure using the general classifier.

Besides the two classifiers for running and cycling, we also trained a classifier on all the activities except running and cycling. These three activity-specific classifiers were tested with the same six machine learning algorithms as the general classifier (trained on all the activities) using the chest + wrist sensor placement. We compared the performance of the activity-specific classifiers with the performance of the general classifier when tested on running, cycling and all the other activities separately. The results are shown in Table IV. One can see that the activity-specific classifiers for running and cycling outperformed the general classifier on the activities recognized as running and cycling, which was to be expected. Surprisingly, the classifier trained on the other activities only did not outperform the general classifier on the

activities not recognized as running and cycling. The reason may be that running and cycling were sometimes not recognized correctly: in such cases the classifier trained on the other activities was used to estimate the energy expenditure, and it naturally did worse than the general classifier.

TABLE IV.   MEAN ABSOLUTE ERROR OF ENERGY EXPENDITURE ESTIMATION (MET) USING THE GENERAL AND ACTIVITY-SPECIFIC CLASSIFIERS

	Linear Regression	Multi-Layer Perceptron	SVR	M5P	M5Rules	REPTree
General						
... tested on running	3.28	1.56	3.08	1.83	1.52	1.40
... tested on cycling	7.92	3.39	5.96	4.14	6.17	2.84
... tested on the others	5.88	<u>0.87</u>	1.83	2.36	3.88	1.48
Activity-specific						
... tested on running	3.12	<u>1.48</u>	2.90	1.48	1.49	1.20
... tested on cycling	7.09	3.04	<u>2.71</u>	3.84	5.35	2.75
... tested on the others	6.33	2.19	6.48	6.59	4.46	1.44

Based on the results in Table IV we constructed the final composite classifier, which consisted of two activity-specific classifiers and the general classifier. The first activity-specific classifier was used to estimate the energy of the activities recognized as running, and was constructed by the Multi-Layer Perceptron (its error is underlined in the table, as are the errors of the following two classifiers). The second activity-specific classifier was used for cycling, and was constructed by the SVR. All the other activities used the general classifier constructed by the Multi-Layer Perceptron. Fig. 2 shows the estimated vs. true energy expenditure for the composite classifier. Its mean absolute error was 0.91 MET.
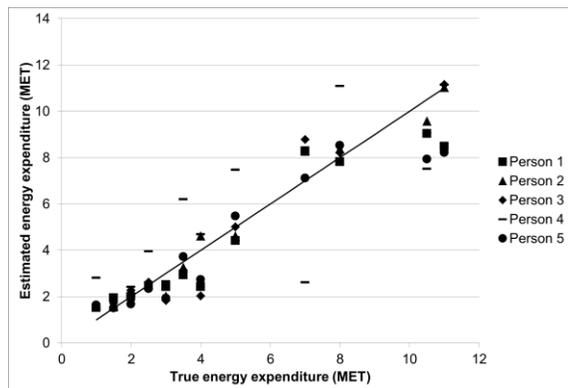
Figure 2.   Estimated vs. true energy expenditure using the final composite classifier.

V.   CONCLUSION

We developed a method for recognizing the activity of a person and estimating his/her energy expenditure using two accelerometers placed on the chest and wrist. For the energy expenditure estimation we investigated many attributes and several machine learning algorithms. We also considered ways to combine multiple classifiers, resulting in the final composite classifier with the mean absolute error of 0.91 MET, which is a substantial improvement over the single-classifier approach with the error of 1.39 MET. One can see in Fig. 2 that the composite classifier estimated the energy expenditure very accurately for most subjects; it only had difficulties with Person 4, who apparently has unusual movement patterns. Cardiologists judged such accuracy adequate for the monitoring of congestive heart failure patients.

In the future we will mostly focus on an efficient implementation of our method. Since the Shimmer sensors are capable of some limited processing, we will attempt to find a way to distribute the computation of the attributes for machine learning between the sensors and the smartphone, that maximizes the battery life. We will also pay attention to an energy-efficient transmission of data between the sensors and the smartphone. Finally, we will consider implementing the activity recognition and energy expenditure estimation in dedicated hardware. In order to do that, the classifier for the activity recognition will have to be trained with the SVM algorithm instead of Random Forest, since the latter constructs highly complex classifiers, but otherwise our method seems well-suited to hardware implementation.

REFERENCES

[1] J. Myers, "Exercise and cardiovascular health", Circulation, vol. 107, pp. e2–e5, 2003.

[2] P. Webb, J. F. Annis, and S. J. Troutman Jr., "Energy balance in man measured by direct and indirect calorimetry," American Journal of Clinical Nutrition, vol. 33, pp. 1287–1298, 1980.

[3] J. A. Levine, "Measurement of Energy Expenditure," Public Health Nutrition, vol. 8(7A), pp. 1123–1132, 2005.

[4] J. Speakman, "Doubly labelled water: Theory and practice," Springer, 1997.

[5] "Chiron project," http://www.chiron-project.eu/.

[6] D. P. Heil, "Predicting activity energy expenditure using the actical activity monitor," Research quarterly for exercise and sport, vol. 77, pp. 64–80, 2006.

[7] C. V. Bouten, K. R. Westerterp, M. Verduin, and J. D. Janssen, "Assessment of energy expenditure for physical activity using a triaxial accelerometer," Medicine and science in sports and exercise, vol. 26, pp. 1516–1523, 1994.

[8] S. E. Crouter, K. G. Clowers, and D. R. Bassett, "A novel method for using accelerometer data to predict energy expenditure," Journal of applied physiology, vol. 100, pp. 1324–1331, 2006.

[9] Y. Jang, M. Jung, J. Kang, and H. K. Chan, "An Wearable Energy Expenditure Analysis System based on the 15-channel Whole-body Segment Acceleration Measurement," in IEEE-EMBS 2005, 27th Annual International Conference of the IEEE, pp. 3834–3836, 2005.

[10] "Shimmer Research," http://www.shimmer-research.com/.

[11] B. E. Ainsworth, W. L. Haskell, M. C. Whitt, M. L. Irwin, A. M. Swartz, et al., "Compendium of physical activities: An update of activity codes and MET intensities," Medicine and Science in Sports and Exercise, vol. 32 (suppl.), pp. 498–516, 2000.

[12] H. Gjoreski, "Adaptive human activity/posture recognition and fall detection using wearable sensors," Master thesis, Jožef Stefan International Postgraduate School, 2011.

[13] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," in SIGKDD Explorations, vol. 11(1), pp. 10–18, 2009.