Ensembles of Multiple Sensors for Human Energy Expenditure Estimation

Hristijan Gjoreski[1,2], Boštjan Kaluža[1,2], Matjaž Gams[1,2], Radoje Milić[3], Mitja Luštrek[1,2]

[1]Jožef Stefan Institute,
[2]Jožef Stefan International Postgraduate School
hristijan.gjoreski@ijs.si

[3]University of Ljubljana, Faculty of Sport -
Institute of Sport
radoje.milic@fsp.uni-lj.si

ABSTRACT

Monitoring human energy expenditure is important in many health and sport applications, since the energy expenditure directly reflects the level of physical activity. The actual energy expenditure is unpractical to measure; hence, the field aims at estimating it by measuring the physical activity with accelerometers and other sensors. Current advanced estimators use a context-dependent approach in which a different regression model is invoked for different activities of the user. In this paper, we go a step further and use multiple contexts corresponding to multiple sensors, resulting in an ensemble of models for energy expenditure estimation. This provides a multi-view perspective, which leads to a better estimation of the energy. The proposed method was experimentally evaluated on a comprehensive set of activities where it outperformed the current state-of-the-art.

Author Keywords

Human energy expenditure, regression, ensembles, multiple sensors, wearable sensors, context.

ACM Classification Keywords

I.2.6 Artificial Intelligence: Learning.

General Terms

Algorithms; Human Factors; Measurement.

INTRODUCTION

Human energy expenditure (EE) directly reflects the level of physical activity, which makes it important for sports training, weight control, management of metabolic disorders (e.g., diabetes) and other health goals. True EE is difficult to measure by traditional means. Direct calorimetry measures the total heat output of a person in an accurate way, but is only usable in laboratory conditions. The

slightly less accurate indirect calorimetry analyzes the respiratory gases, which requires a breathing mask. Doubly labeled water is both accurate and convenient, but can measure only long-term EE. Finally, self-reporting is highly unreliable.

With the increasing accessibility and miniaturization of sensors and microprocessors, ubiquitous monitoring systems are becoming a practical solution for measuring EE. Such systems primarily measure the physical activity with accelerometers, but can include additional sensors that indirectly measure the metabolic activity, such as a heart rate monitor or thermometer. The main challenge is how to estimate the EE from sensor outputs accurately, irrespectively of the user's activity, ambient conditions and context in general.

Early approaches to EE estimation ignored the context and used simple regression on accelerometer outputs. As the field matured, it became clear that a context-dependent approach in which a different regression model is used for different activities of the user improves the performance. In this paper we go a step further and use multiple sensors and therefore multiple context components, resulting in an ensemble of models for EE estimation. We show experimentally that such an ensemble outperforms the current state-of-the-art.

RELATED WORK

The first methods for EE estimation with wearable sensors used linear regression to map the accelerometer output to EE. The accelerometer output was often expressed in "counts", an aggregate acceleration measure reported by devices such as Actigraph [4, 6, 10]. More recent approaches used a richer representation of the accelerometer output consisting of multiple features, as well as non-linear regression methods such as artificial neural networks and model trees [9, 11]. These approaches were experimentally shown to substantially improve upon the early Actigraph-based work.

Researchers soon realized that different contexts require different energy-expenditure equations [6, 10]. This realization resulted in the state-of-the-art Actigraph-based approach by Crouter et al. [2], which assigned 1 MET (Metabolic Equivalent of Task) to inactivity and used two

different (non-linear) regression models for light and intense activity. Lester et al. [7] used three predefined EE estimation formulas for rest, walking and running. They also considered GPS and barometer information to estimate the slope of walking/running. The reported results are promising, but the range of activities was quite limited. The advances in the accelerometer-based activity recognition allowed finer-grained activities as the context for EE estimation [1, 8, 11]. The current state-of-the-art is probably the commercial SenseWear device by BodyMedia [12], which uses the activity as the context, and combines multiple context-dependent regression models according to the probabilities for their respective contexts. It also uses multiple sensors: an accelerometer, two thermometers, galvanic-skin-response and heat-flux sensors.

Our work uses not only the activity as the context, but utilizes multiple context components from multiple sensors, so that each measurement can be placed in multiple contexts simultaneously (e.g., activity = running, heart rate = high, breath rate = moderate, etc.). This gives us the benefit of multiple "viewpoints" from which one can estimate the EE, and later combine them in order to achieve greater accuracy than a single context component.

ENSEMBLES OF MULTIPLE SENSORS

In general, the context is any information that characterizes the circumstances in which an event occurs [3]. In our approach, the context consists of eight context components that mostly correspond to the sensors used: activity (A), acceleration peaks count (AP), heart rate (HR), breath rate (BR), chest skin temperature (CST), galvanic skin response (GSR), arm skin temperature (AST) and ambient temperature (AT). An ensemble of multiple sensors consists of multiple regression models for EE estimation, which correspond to the context components. For each possible value of a context component, a different model is built (e.g., one model for activity = running, another for activity = lying). Therefore, for each data sample a custom ensemble is assembled from the models that correspond to the contexts of that sample (e.g., if the sample contains running with a high heart rate, the models for activity = running and heart rate = high are included in the ensemble).

To explain the ensembles, let us consider a simple version with only three context components: activity (A), heart rate (HR) and breath rate (BR). Since HR and BR have numerical values, we first discretize their values into M and L intervals, respectively. The activity feature is already discrete and contains N activities. The context components are shown in the first row of Figure 1 and their discrete values (intervals) in the second row.

In the next, training phase, for each discrete value (e.g., A_1) of each context component (e.g., A), a regression model (e.g., M_{A_1}) is trained. The training dataset for the model is a subset of the whole training dataset. It contains only the data samples which have the corresponding component value; thus, the model for the first activity M_{A_1}

is trained only on the data samples that contain the activity A_1. Once the training data is selected, the model is trained using an arbitrary regression method.
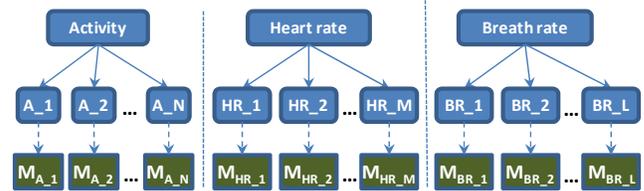

Figure 1. Multiple sensors ensemble scheme.

In the final, estimation phase, the EE of a data sample (feature vector) is estimated by an ensemble consisting of a subset of previously trained models. The models included in the ensemble are invoked according to the context component values. The final EE estimation is provided by combining the outputs of all invoked models. For example, consider the following scenario: a user is running with the heart rate of 140 min^{-1} and breath rate of 35 min^{-1}. Let us assume that running is the second of the N modeled activities (A_2), the heart-rate value falls into the second heart-rate interval (HR_2), and the breath-rate value into the first breath-rate interval (BR_1). The data sample will thus be evaluated by the models M_{A_2}, M_{HR_2} and M_{BR_1}, whose outputs will be combined (e.g., by averaging, choosing the median, etc.) to estimate the final EE.

EXPERIMENTS

Experimental Setup
The sensor equipment consisted of two Shimmer three-axis accelerometer[1], a Zephyr BioHarness sensor[2], a BodyMedia SenseWear sensor[3], and a Cosmed indirect calorimeter[4]. Each of the sensors used in this study provided different information about the user's EE. For the final evaluations we used the features shown in Table 1.

The two accelerometers were placed on the chest and the right thigh because our previous tests showed that those placements perform best for activity recognition [5]. The Zephyr sensor – a commercial sports strap – was worn on the chest, and measured the user's heart rate, breath rate, and chest skin temperature. The BodyMedia SenseWear sensor – a state-of-the-art commercial device for EE estimation – was worn on the left upper arm. It served as a benchmark, and additionally measured the user's galvanic skin response, ambient temperature and arm skin temperature. Finally, the true EE was measured on a breath-by-breath basis with the Cosmed indirect calorimeter. The Cosmed equipment consists of a sensing unit, a wireless transmitter unit, a mask and a shoulder-belt system. The output from this sensor in MET was used as the target value for training and evaluating the EE estimation models.

---

[1] http://www.shimmer-research.com/
[2] http://www.zephyr-technology.com/products/bioharness-3
[3] http://sensewear.bodymedia.com/
[4] http://www.cosmed.com/quarkcpet

Sensor	Features		
Shimmer accelerometers	Activity, acceleration peaks count	Context components	
Zephyr	Heart rate, breath rate, chest skin temperature		
BodyMedia SenseWear	Galvanic skin response, arm skin temperature, ambient temperature		
Cosmed	Reference METs	MET	

Table 1. Features per sensor.

Almost all of the features were provided directly from the sensors, except for the activity and the acceleration peaks count. These two features were extracted from the acceleration data. For the activity recognition, we used our previous experience and developed a method based on machine learning [5]. The method uses the data from the two accelerometers (chest and thigh) and applies a Random forest classification model to recognize the user activity. The acceleration peaks count is the number of times the length of the acceleration vector stops increasing and starts decreasing or vice versa.

A 90-minute (excluding breaks) scenario was designed to capture a wide range of everyday activities and exercises with various EEs. The scenario is shown in Table 2, together with the average EEs for each activity as measured with the Cosmed sensor. It was recorded by ten volunteers aged 24–33, mean 27.3; BMI 20–28.9, mean 24.1. The resulting dataset consisted of approximately 1,000,000 raw-data samples per volunteer.

The proposed approach was implemented in Java using the WEKA machine learning toolkit [14]. First, the data from each sensor was synchronized using the starting timestamps. A sliding window of ten seconds was applied to segment the sensor data. Such length of the window was empirically determined as reasonable for the estimation of the EE. Because each sensor provides data with different sampling frequencies, the value within each window was averaged. For the non-numeric features such as the activity, the majority value was used. Additionally, because the data is user-specific, it was normalized using the first lying interval in the test scenario (shown in Table 2). The average value for each sensor data from this resting interval was computed. It was used as a calibration (reference) value and was subtracted from each sensor value. Once the feature vectors were formed, each numeric feature was discretized into four intervals using the Yong split criterion [13] implemented in WEKA. The discretization procedure was used only to define the ensemble structure as shown in Figure 1. Afterwards, the EE of each testing data sample was estimated depending on the values it contained for each context component. To train the regression models, four methods were compared: linear regression, Gaussian processes, multilayer perceptron (artificial neural network) and SMOReg (support vector regression). Finally, the median of the outputs of the invoked models was used as the final result.

Activity	EE (MET)
Lying	1.19 MET
Sitting	1.26 MET
Standing	1.26 MET
Walking slowly on a treadmill (4 km/h)	3.50 MET
Walking quickly on a treadmill (6 km/h)	5.03 MET
Running slowly on a treadmill (8 km/h)	7.80 MET
Stationary cycling lightly (1 W/kg)	4.91 MET
Stationary cycling vigorously (2 W/kg)	7.22 MET
Kneeling	1.30 MET
On all fours	1.77 MET
Lying doing light exercise	1.28 MET
Sitting doing light activities	2.30 MET
Walking doing light chores	2.30 MET
Scrubbing the floor	2.65 MET
Shovelling snow, digging	3.40 MET

Table 2. Test scenario.

The method evaluation was performed with the leave-one-person-out cross-validation technique; that is, models were trained on the data of nine people and tested on the remaining person. The same was done with the activity recognition classifier, whose output was used as a feature the in EE estimation. This procedure was repeated ten times, for each person. As the evaluation metric, we chose the mean absolute percentage error (MAPE; the mean absolute error divided by the true value), since it is the most common metric in the EE estimation domain [12].

Results

To evaluate our ensemble approach, we compared the previously mentioned four standard regression methods. In addition, as a baseline for comparison, we evaluated the same regression methods without the ensemble scheme, i.e., single regression models were constructed over the whole feature set. Also the results were compared to the MET output of the BodyMedia SenseWear commercial sensor. It should be noted that the SenseWear averages over 1 minute while our methods over 10-second intervals.

The results presented in Figure 2 show that for each method, the ensembles significantly outperform the baseline, as well as the SenseWear sensor. The multilayer perceptron method achieves the best 27.5% MAPE, which is by 7 percentage points better than the SenseWear. Thus, the further analysis is performed only for this method.
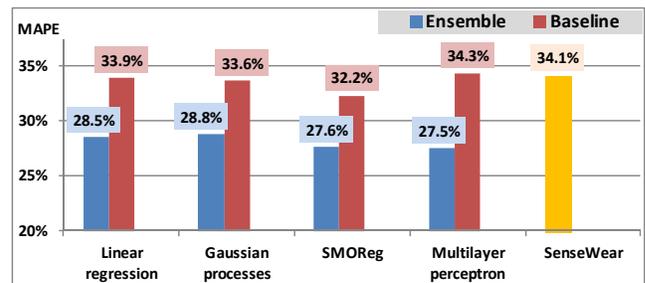

Figure 2. Comparison of the MAPEs for different methods.

Since the ensembles consist of eight context components, we additionally show the results achieved by each of the

context models used individually. In Figure 3 one can see that the MAPE of the final method is better than the MAPEs of the individual models. This shows that even by combining the models using a simple median of their outputs, the ensemble outperformed the individual models.
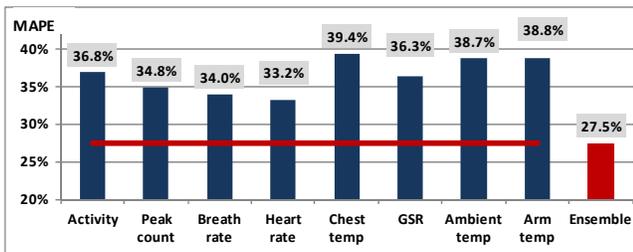

Figure 3. Comparison of the context models used individually.

Figure 4 shows a scatter plot comparing the true and estimated MET values for different activities. The results show that in general, the estimations by the ensembles compared to the other two approaches better match the actual Cosmed MET values (the diagonal line in Figure 4). The SenseWear sensor has better estimations for the more dynamic activities (running, cycling and walking) compared to the everyday activities, which is probably because the device is intended for physically active users.
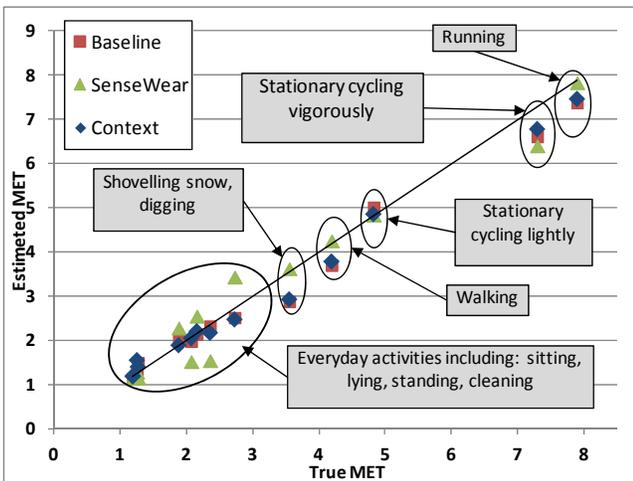

Figure 4. Comparison between the true and predicted MET values for different activities.

CONCLUSION

The paper presented a novel approach to human EE estimation using ensembles of multiple sensors. The key idea is to use not only a single regression model at a time, but an ensemble of models, each of which is constructed for a particular context corresponding to a different sensor. This provides a multi-view perspective, which leads to a better estimation of the EE. This can benefit all who are interested in precisely managing their caloric output because of their health condition, sports activity etc.

The proposed method was experimentally evaluated on a special comprehensive scenario and compared to state-of-the-art approaches. The results showed that the proposed ensembles significantly outperformed the competing approaches. In the future we first plan to focus on more sophisticated methods for combining the outputs of the models in an ensemble.

REFERENCES

1. Albinali, F., Intille, S.S., Haskell, W., and Rosenberger, M. Using wearable activity type detection to improve physical activity energy expenditure estimation. In *Proc. Ubicomp '10*, ACM Press (2010), 211-320.

2. Crouter, S.E., Clowers, K.G., and Bassett, D.R. A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol 100* (2006), 1324-1331.

3. Dey, A., Salber, D., Abowd G., and Futakawa, M. The conference assistant: Combining context awareness with wearable computing. In *Proc ISWC*, (1999).

4. Freedson, P.S., Melanson, E., and Sirard, J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc*, 5 (1998), 777-781.

5. Gjoreski, H., Luštrek, M., and Gams, M. Accelerometer Placement for Posture Recognition and Fall Detection. In *Proc IE' 2011*, (2011), 47–54.

6. Hendelman, D., Miller, K., Bagget, C., Debold, E., and Freedson, P.S. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc 32*, 9 (2000), 442-449.

7. Lester, J., Hartung, C., Pina, L., Libby, R., Borriello, G., & Duncan, G.(2009, September). Validated caloric expenditure estimation using a single body-worn sensor. In *Proc. Ubicomp '09*, ACM Press (2009), 225-234.

8. Luštrek, M., Cvetković, B., and Kozina, S. Energy expenditure estimation with wearable accelerometers. In *Proc. ISCAS 2012*, IEEE (2012), 5-8.

9. Rothney, M.P., Neumann, M., Béziat, A., and Chen, K.Y. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. *J Appl Physiol 103* (2007), 1419-1427.

10. Swartz, A., Strath, S.J., Bassett, D.J., O'Brien, W., King, G.A., and Ainsworth, B.E. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc 32*, 9 (2000), 450-456.

11. Tapia, E.M. Using machine learning for real-time activity recognition and estimation of energy expenditure. *PhD thesis*, MIT, 2008.

12. Vyas, N., Farringdon, J., Andre, D., and Stivoric, J. Machine learning and sensor fusion for estimating continuous energy expenditure. In *Proc. 23rd IAAI*, AAAI (2012), 1613-1620.

13. Wang, Y., and Witten, I. Induction of model trees for predicting continuous classes. In *Proc*. ECML (1997),128-137.

14. Witten, I., and Frank, E. Data Mining: Practical machine learning tools and techniques. *Morgan Kaufmann*. (2005).