Using Support Vector Regression for Assessing Human Energy Expenditure Using a Triaxial Accelerometer and a Barometer

Panagiota Anastasopoulou[1,2], Sascha Härtel[2,3], Mirnes Tubic[3], and Stefan Hey[2]

[1] Institute for Information Processing Technology, Karlsruhe Institute of Technology, Karlsruhe, Germany
[2] House of Competence, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3] Institute of Sport and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany
{panagiota.anastasopoulou,sascha.haertel,stefan.hey}@kit.edu, mirnes_tuba@hotmail.com

Abstract. Physical inactivity is nowadays defined as the fourth leading risk factor for global mortality. These levels are rising worldwide with major aftereffects on the prevention of several diseases and the general health of the population. Energy expenditure (EE) is a very important parameter usually used as a dimension in physical activity assessment studies. However, the most accurate methods for the measurement of the EE are usually costly, obtrusive and most are limited by laboratory conditions. Recent technological advancements in the sensor technology along with the great progress made in algorithms have made accelerometers a powerful technique often used to assess everyday physical activity. This paper discusses the use of support vector regression (SVR) to predict EE by using a single measurement unit, equipped with a triaxial accelerometer and a barometer, attached to the subject´s hip.

Keywords: accelerometers, barometers, energy expenditure, physical activity monitoring, support vector regression.

1 Introduction

During the last decade several studies have shown the positive affect of adequate exercise on the health of people and the correlation between the lack of exercise and the risk of developing various diseases such as cardiovascular diseases, colon and breast cancers, Type 2 diabetes and osteoporosis [7]. According to the World Health Organisation (WHO), physical inactivity (lack of physical activity) has been identified as the fourth leading risk factor for global mortality (6% of deaths globally) leading to approximately 3.2 million deaths each year [10] (Fig. 1).

Energy expenditure (EE) is one of the most widely used quantitative measures in studies that try to assess people’s everyday life physical activity. Doubly labeled water and indirect calorimetry are considered to be the gold standard measures to measure EE. Nevertheless those methods are rather obtrusive and require high operating costs. Other methods like questionnaires or diaries have on the other hand

B. Godara and K.S. Nikita (Eds.): MobiHealth 2012, LNICST 61, pp. 106–113, 2013.

limited accuracy. Due to the improvements in sensor technology and algorithms, accelerometry has become the mostly used technique to assess everyday life physical activity.

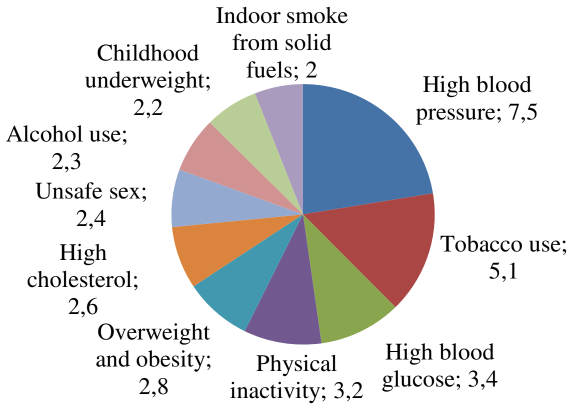

Fig. 1. 10 leading risk factors causes of death in millions per year

Jatobá et al. [3] proposed an acceleration-based EE prediction process. The activity-based linear models were developed by using the intensity of the acceleration together with some other subject-related features (such as age, weight, height).

Su et al. [8] proposed the use of SVR in order to predict EE during walking by using an acceleration sensor attached to the lower back close to the subject's center of gravity.

A well-known limitation of those devices is that they are not able to assess the increase in energy cost of walking upstairs or uphill, since the acceleration pattern remains unchanged under these conditions, although increased effort is required.

Yamazaki et al. [9] and Voleno et al. [4] introduced, along with the accelerometer, the use of an additional air pressure sensor in order to capture the movement in the vertical axis and used linear regression models for the EE estimation.

The purpose of this paper is to discuss the use of SVR to predict EE for different everyday life activities by using a single measurement unit, which consists of a triaxial accelerometer and a barometer, placed on the subject´s right side hip.

2 Measurement Setup and Data Collection

2.1 Measurement System

The Move II sensor (movisens GmbH, Karlsruhe, Germany) was used for the data collection. Move II consists of a triaxial acceleration sensor with a range of ±8 g, a resolution of 12 bit and a sampling frequency of 64 Hz. The measuring unit has an additional air pressure sensor with a sampling frequency 8 Hz and a resolution of 0.03hPa (corresponds to 15 cm at sea level). The recorded raw data were saved using the unisens format [5] and were transferred on a computer for further analysis via a USB 2.0 interface.

The assessment of the reference for the EE was performed by using the portable indirect calorimeter Meta Max 3B (Cortex Biophysics GmbH, Leipzig, Germany). The indirect calorimeter measures the breath by breath energy consumption and transmits the measured data wirelessly to a laptop.

2.2 Subject Characteristics

Twenty healthy subjects (12 male and 8 female), all students or employees of the Karlsruhe Institute of Technology (KIT) participated in the data collection study. The data collection was performed in cooperation with the Institute of Sport and Sports Science, Karlsruhe Institute of Technology (KIT). Descriptive data of the subjects can be found in Table 1.

Table 1. Subject characteristics

Subject parameter	Males (N=12)	Females (N=8)	All subjects (N=20)
Age (yrs.)	30.1 ± 8.8	30.8 ± 8.6	30.4 ± 8.7
Height (cm)	179.8 ± 8.0	167.3 ± 4.3	174.8 ± 9.2
Weight (kg)	81.9 ± 12.9	64.9 ± 9.2	75.1 ± 14.2
BMI ($kg \cdot m^{-2}$)	25.3 ± 3.3	23.2 ± 3.1	24.5 ± 3.3

2.3 Measurement Procedure

In our study all the subjects were equipped with the measurement unit, placed over the right anterior axillary line (Fig. 2) and the indirect calorimeter. The indirect calorimeter was both used to assess the reference data for the EE and to set the markers at every transition between the different activities. The markers where later on synchronized with the acceleration and air pressure signals.

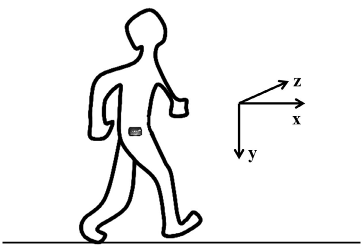

Fig. 2. Position of the sensor and direction of the acceleration axes

For the training and the evaluation of the algorithms a study with a variety of different activities (10 in total) was carried out. The data collection protocol of the study can be seen in Table 2.

Table 2. Study- Data Collection Protocol

Activity	Duration	Frequency x Distance
sitting	5 min	
standing	5 min	
walking slow		1 x 415 meter
walking fast		1 x 415 meter
jogging		2 x 415 meter
cycling	app. 5min	
walking up- /downhill		4 x130 meter
walking upstairs		3 x 84 stairs
walking downstairs		3 x 84 stairs

3 Signal Processing

3.1 Signal Preprocessing

The signal processing for the development of the EE prediction models was done with MATLAB (R2010a). Both the acceleration and the air pressure signals were preprocessed in two steps.

During the first step the acceleration signals were band-pass (0.25 Hz - 10 Hz) filtered to suppress DC-response (static acceleration due to gravity) and high frequencies that cannot arise from human movement respectively. Using the barometric formula, the air pressure signal was converted into altitude and the noise was suppressed using a Butterworth low-pass filter.

In the second step the activity was classified in intervals of 4 seconds. The activity recognition process is discussed in [6]. The classification algorithm differentiated between the following 7 activities: lying, rest (sitting/standing), cycling, up-hill/-stairs, down-hill/-stairs, level walking and jogging.

The reference for the EE was the output from the mobile indirect calorimeter. The indirect calorimeter measures the breath by breath energy consumption. Due to the non-uniform sampling frequency of the indirect calorimeter, linear interpolation was used in order to resample the EE reference to a uniform 1 Hz sampling frequency.

3.2 Parameter Extraction

The key step for the preparation of the data for the modeling is the parameter extraction. Each second of the acceleration and the air pressure data is transformed into a feature vector. The feature vector is the input for the EE modeling and estimation. The features that were extracted are as follows:

Acceleration Magnitude. The acceleration magnitude represents the intensity of the movement in each interval and is proven to be highly correlated to the EE. The acceleration magnitude was calculated using the AC part of each acceleration component as follows:

$$EEAC(i) = mean(sqrt\ (\ ax_{AC}^2(i)\ +ay_{AC}^2(i)\ +az_{AC}^2(i)\ ))\ . \quad (1)$$

Altitude Change. The altitude change corresponds to the direction of the vertical movements. Therefore the altitude change (Δh) at every segment was calculated. The altitude change was split into two separate features, the positive and the negative altitude change (Δh_{pos} and Δh_{neg} respectively). By means of this value, it is possible to estimate the intensity of the movement in the vertical axis, and therefore a better estimate of the energy consumption when walking up- or downhill.

Subject Related Data. Besides the features extracted from the acceleration and the air pressure signal, other subject-related features, which are determinant for the person's EE (body height, body weight age and sex), were used in the model development as well.

3.3 Support Vector Regression (SVR)

Support vector machine (SVM) was firstly developed in the sixties but it has been further developed during the last four decades and is nowadays one of the widely used techniques to solve classification problems. Support vector regression (SVR) is a technique based on the SVM used when the predicted values are not discrete but continuous values. SVR can be used to learn complex nonlinear relationships between predictor and predicted values.

SVR tries to find a function f(x), which given the training data $\{(x_1,y_1),..., (x_l,y_l)\}$ has at most ε deviation from the actually obtained targets y_i and at the same time is as flat as possible. SVR approximates the function in the following form:

$$f(x) = w \cdot \varphi(x) + b\ . \quad (2)$$

Where w and b are coefficients and φ maps the original data x to a high dimensional feature space. This optimization problem can be transformed by applying Lagrangian theory into:

$$f(x) = \Sigma^{l}_{i=1}\ (\alpha_i^* - \alpha_i) \cdot k(x_i,x) + b\ . \quad (3)$$

Where α_i and α_i* are Lagrange multipliers and the function $k(x_i,x)$ is called the kernel function and is defined as a linear dot product of the nonlinear mapping, i.e.,

$$k(x_i,x) = \varphi(x_i) \cdot \varphi(x) \, . \tag{4}$$

The types of kernel functions mostly used are: linear, polynomial, radial basis and sigmoid. Here the use of radial basis function was investigated and compared with the traditional linear regression (LR).

All calculations were done using the library LIBSVM [2].

3.4 Estimation of Energy Expenditure

For the estimation of the EE, the activities were divided into groups. The activity-based EE prediction included 3 groups. The first group included all passive activities (lying, sitting and standing), all walking activities (walking fast, slow, up-/downhill, up-/downstairs and jogging) in the second and cycling in the third.

For the first activity-group, the EE estimation was done using the basal metabolic rate (BMR) formulas as published in [1]. BMR is defined as the minimal rate of energy expenditure compatible with life. It is measured in the supine position under standard conditions of rest, fasting, immobility and represents the calories the body needs only for the functioning of the vital organs (e.g. heart, lungs, nervous system). The EE for the passive activities was defined as 25% above the BMR. For the other 2 groups, SVR- and LR-based models were developed. The EE estimation models were further separated into 2 groups according to the gender.

4 Results

For the validation of the EE prediction models we chose the leave-one-subject-out cross validation method. Using this method, the generalization of the models can be tested. The difference between the reference and the predicted values of the EE was computed for each subject and each activity and compared with the results of the LR.

Table 3 summarizes the results of the EE estimation, showing the mean percent error of the second-by-second estimation of EE for each activity and the mean error in kcal for the whole activity. Fig. 3 shows the measured and the estimated EE for the whole protocol for one subject.

The results suggest that the use of SVR improves the estimation of EE. The mean percent error for all the activities was reduced from -1.8 ± 12.0 %, when using LR to -1.4 ± 10.5 %, when using SVR. For some specific activities like walking slowly and fast the prediction errors were reduced significantly from -10.1 ± 14.8 % and -7.7 ± 17.5 % to -6.0 ± 14.8 % and -2.2 ± 17.1% respectively. Only during walking up-/downhill there was a slight increase in the prediction error. This might be due to the short time period. Generally the biggest errors were observed during the movement transitions; the time needed for the stabilization of the energy expenditure.

Table 3. Prediction errors (mean ± SD) for the energy expenditure

Activity	LR Mean error [kcal]	LR Percent error [%]	SVR Mean error [kcal]	SVR Percent error [%]
all activities	-0.7 ± 26.2	-1.8 ± 12.0	-1.0 ± 22.8	-1.4 ± 10.5
sitting / standing	0.6 ± 2.3	2.0 ± 12.4	0.6 ± 2.3	2.1 ± 12.4
walking slow	-2.3 ± 3.5	-10.1 ± 14.8	-1.3 ± 3.3	-6.0 ± 14.8
walking fast	-1.3 ± 3.3	-7.7 ± 17.5	-0.1 ± 3.5	-2.2 ± 17.1
jogging	3.6 ± 6.3	5.5 ± 12.1	1.7 ± 7.8	2.0 ± 13.9
cycling	0.3 ± 8.5	-1.3 ± 31.4	0.6 ± 7.9	0.5 ± 26.5
walking up- / downhill	-1.3 ± 4.7	-3.5 ± 11.5	-2.1 ± 4.1	-4.9 ± 10.5
walking upstairs	-0.6 ± 2.5	4.4 ± 15.8	-0.6 ± 2.8	4.4 ± 17.7
walking downstairs	0.4 ± 1.3	4.1 ± 13.2	0.2 ± 1.2	3.2 ± 12.1

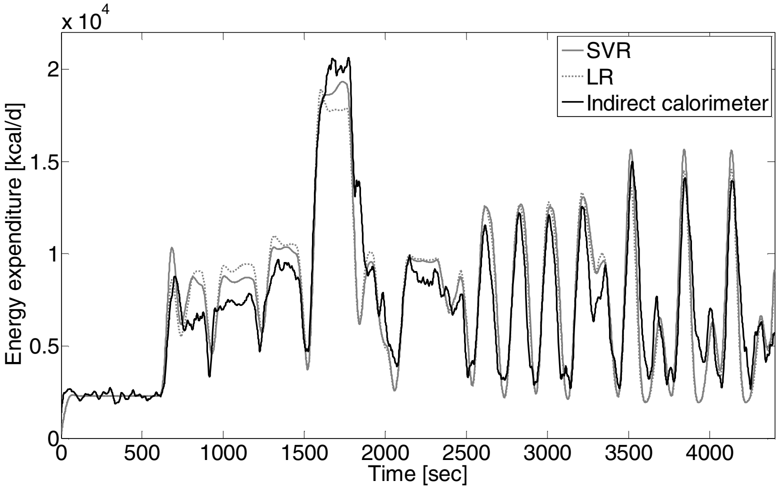

Fig. 3. Energy expenditure. The gray line is the estimated EE using the SVR, the gray dashed line is the estimated EE using the LR and the black solid line is the smoothed gold standard measure obtained by the portable indirect calorimeter.

5 Conclusion

A novel method for the estimation of human everyday activity EE has been presented and evaluated. The algorithm uses the acceleration and the air pressure signals assessed at the subject's hip and with the developed SVR-based models, it predicts the EE with high accuracy. The biggest errors were observed during the activities-transitions. This is due to the fact that the time to arrive at the steady state depends mostly on the fitness index of the subject, which was not included in the model parameters.

Future work will focus on the evaluation of the method in other population groups such as the elderly or people suffering from obesity, where physical activity monitoring is needed.

Acknowledgments. The authors would like to thank the company movisens GmbH and all the people that supported the data collection study as well as all the subjects that voluntarily participated in it.

References

1. Energy and protein requirements: Report of a joint FAO/WHO/UNU expert consultation. WHO Technical Report Series No. 724. Geneva (1985)
2. Chang, C.-C., Lin, C.-J.: LIBSVM: A Library for Support Vector Machines. ACM Transactions on Intelligent Systems and Technology 2(3), Article 27 (2011)
3. Jatobá, L.C., Großmann, U., Ottenbacher, J., Härtel, S., Haaren, B., Stork, W., Müller-Glaser, K.D., Bös, K.: Obtaining Energy Expenditure and Physical Activity from Acceleration Signals for Context-aware Evaluation of Cardiovascular Parameters. In: CLAIB 2007. IFMBE Proceedings, vol. 18, pp. 475–479 (2007)
4. Voleno, M., Redmond, S.J., Cerutti, S., Lovell, N.H.: Energy Expenditure Estimation Using Triaxial Accelerometry and Barometric Pressure Measurement. In: 32th Annual International Conference of the Engineering in Medicine and Biology Society, pp. 5185–5188 (2010)
5. Krist, M., Ottenbacher, J.: Open Source Data Format unisens (universal data format for multi sensor data), `http://www.unisens.org`
6. Anastasopoulou, P., Tansella, M., Stumpp, J., Shammas, L., Hey, S.: Classification of Human Physical Activity and Energy Expenditure Estimation by Accelerometry and Barometry. In: 34th Annual International Conference of the Engineering in Medicine and Biology Society, pp. 6451–6454 (2012)
7. Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee Report, Washington, DC, USA: Department of Health and Human Services (2008)
8. Su, S.W., Wang, L., Celler, B.C., Ambikairajah, E., Savkin, A.V.: Estimation of Walking Energy Expenditure by Using Support Vector Regression. In: 27th Annual International Conference of the Engineering in Medicine and Biology Society, pp. 3526–3529 (2005)
9. Yamazaki, T., Gen-No, H., Kamijo, Y.-I., Okazaki, K., Masuki, S., Nose, H.: A New Device to Estimate $\dot{V}O2$ during Incline Walking by Accelerometry and Barometry. Medicine & Science in Sports & Exercise 12, 2213–2219 (2009)
10. World Health Organisation. Global health risks: Mortality and burden of disease attributable to selected major risks (2009)