

Statistical Machine Learning for Automatic Assessment of Physical Activity Intensity Using Multi-axial Accelerometry and Heart Rate

Fernando García-García^{1,2}, Gema García-Sáez^{1,2}, Paloma Chausa^{1,2},
Iñaki Martínez-Sarriegui^{1,2}, Pedro José Benito³,
Enrique J. Gómez^{1,2}, and M. Elena Hernando^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, Universidad Politécnica de Madrid, Spain
{fgarcia,elena}@gbt.tfo.upm.es

² Networking Research Center on Bioengineering, Biomaterials and Nanomedicine
(CIBER-BBN), Madrid, Spain

³ Laboratory of Exercise Physiology, Facultad de Ciencias de la Actividad
Física y del Deporte (INEF), Universidad Politécnica de Madrid, Spain

Abstract. This work explores the automatic recognition of physical activity intensity patterns from multi-axial accelerometry and heart rate signals. Data collection was carried out in free-living conditions and in three controlled gymnasium circuits, for a total amount of 179.80 h of data divided into: sedentary situations (65.5%), light-to-moderate activity (17.6%) and vigorous exercise (16.9%). The proposed machine learning algorithms comprise the following steps: time-domain feature definition, standardization and PCA projection, unsupervised clustering (by k -means and GMM) and a HMM to account for long-term temporal trends. Performance was evaluated by 30 runs of a 10-fold cross-validation. Both k -means and GMM-based approaches yielded high overall accuracy (86.97% and 85.03%, respectively) and, given the imbalance of the dataset, meritorious F-measures (up to 77.88%) for non-sedentary cases. Classification errors tended to be concentrated around transients, what constrains their practical impact. Hence, we consider our proposal to be suitable for 24 h-based monitoring of physical activity in ambulatory scenarios and a first step towards intensity-specific energy expenditure estimators.

Keywords: physical activity intensity, accelerometry, heart rate, k -means, Gaussian Mixture Models, Hidden Markov Models, F-measure.

1 Introduction

The automatic assessment of physical activity facilitates ambulatory monitoring of lifestyle and may enrich traditional interventions to prevent and/or to manage chronic conditions for which regular exercise is prescribed, like cardiovascular and pulmonary diseases, obesity or diabetes.

Current gold-standard techniques to measure energy expenditure (EE), i.e. doubly-labeled water and indirect calorimetry, are costly and impractical in free-living conditions.

Thus, a frequent alternative is to use multi-axial accelerometers placed on key locations of the body (hip, wrist, etc.) and to estimate EE by linear regression from acceleration counts and personal characteristics (weight, sex, etc.). However, accelerometry-based techniques are posed to neglect any physical exercise not encompassing movements of the sensor. In addition, the accuracy of such EE estimation formulae has been reported to fluctuate substantially depending –among other factors– on whether the activity under monitoring is or not comparable in terms of type and intensity with the particular set-up used in the experiments to develop and validate the equations (Crouter et al.[1]).

Another common option is to estimate EE from heart rate (HR), since there are physiological evidences for HR exhibiting a linear relationship with oxygen uptake ratios ($\dot{V}O_2$) and energy consumptions in the case of aerobic exercise at intermediate intensity ranges (i.e. around 110–150 bpm[2]). However, such linearity can be adversely affected by a number of other factors –drugs, stress, etc.– and does not hold for anaerobic exercise[3].

Consequently, an increasing number of authors advocate for techniques combining simultaneously accelerometry and HR measurements[2, 4, 5].

On the other hand, after Crouter et al.[1] identified various limitations for the EE regression formulae, different authors have suggested activity-specific EE estimation schemes[6–8], whose first step was to recognize a particular set of activities (e.g. sitting, standing, walking, running, cycling, etc.). Machine learning (ML) techniques have been thoroughly applied in literature to such activity recognition tasks, with noticeable success and a wide range of algorithms, for example: C4.5 trees[8, 9], Naïve Bayes classifiers[8, 9], GMM[10], MLP[11] and SOM[12] neural networks, AdaBoosting[13], etc. Most approaches, like the aforementioned, were based solely on accelerometry; while some others combined it with HR measurements[7, 14].

Conversely, instead of distinguishing among concrete activities from a pre-established set, in this work we will present algorithms to assess physical activity intensity (PAI) from multi-axial accelerometry and HR signals. In current practise and commercial monitoring devices, physical activity is often ranked into intensity levels based simply on ‘cut-points’ –thresholds– for accelerometer counts and/or HR, either with constant[15] or subject-dependent[16] values. Within the ML literature, some authors defined separate classes for the same exercise performed at different intensities, as in [7, 8]. In this work we will develop a combination of statistical ML techniques to determine PAI in an explicit and activity-independent manner. PAI classes will correspond to standardized intensity range definitions (Ainsworth et al.[17]): rest and sedentary situations, light-to-moderate activity and vigorous exercise.

2 Materials and Methods

2.1 Equipment

ActiTrainer Research accelerometer (ActiGraph, USA) was selected due to its function to record wireless signals from Polar Wearlink heart rate monitors (Polar Electro, Finland). ActiTrainer (51 g, 86×33×15 mm, ±3G dynamic range, 30 Hz sample frequency) registered biaxial acceleration counts (\mathbf{a}_1 , \mathbf{a}_2), number of steps –pedometer function– \mathbf{st} and heart rate \mathbf{hr} . Measurements were accumulated over 10 s periods –epochs– so that \mathbf{hr} was physiologically meaningful.

Following manufacturers’ guidelines, Polar strap was worn on the chest and ActiTrainer was tightly placed on the hip. Main and secondary axis had vertical and antero-posterior orientation, respectively.

2.2 Experiments and Participants

Two different experiments were conducted.

Exp. 1. aimed to acquire data in free-living conditions. Seven healthy subjects took part: 3 males and 4 females, aged 23–36 and with lifestyles ranging from sedentary to regular athletic training.

Volunteers received instructions on how to wear the sensors and were encouraged to report detailed written descriptions of type, duration and intensity of their physically active periods. They were requested to include: exercise at a self-selected intensity, daily life situations (e.g. sleep, rest, walk, housekeeping, office work, etc.) and the use of transportation (elevator, car, bus, metro).

An heterogeneous set of activities was obtained, including: walking at moderate speeds, dancing, jogging, vigorous endurance running, karate, football, mountain bike, etc. The elimination of periods with ambiguous annotations yielded a total of 149.35 h in 72 sequences. Data were then manually grouped according to Ainsworth et al.’s Compendium[17] into three PAI classes: *(i)* rest and sedentary situations (<3 MET, Metabolic Equivalent), *(ii)* light-to-moderate activity (3–6 MET), and *(iii)* vigorous exercise (>6 MET). See Table 1 for details.

Exp. 2. consisted of a controlled laboratory set-up where physical activity was performed in a gymnasium under researchers’ supervision. Three circuit modalities were available, comprising: upper and lower limb exercises in gymnasium machines (e.g. shoulder press), free weight training and a combined weight-aerobic (treadmill running) circuit. Each 64 min session started with a warm-up phase by 5 min of treadmill/elliptical walk and a preliminary circuit lap (7.75 min) with light load. Thereafter, 3 more circuit bouts with high load were performed, each of these bouts separated by 5 min of ‘active rests’ (i.e. walking)¹.

Nine subjects aged 20–49 years were involved (6 male, 3 female). Three were healthy active males and the remaining 6 individuals suffered overweight

¹ PRONAF Study (Trial registration ClinicalTrials.gov NCT01116856). A complete description of exercises, circuits and protocols can be found in Benito et al.[18].

Table 1. Summary of the collected dataset, divided by experiments and PAI ranges. In *Exp. 2*, 40.9% ‘REST’ data embraced pre- and post-exercise periods.

	<i>Exp. 1</i>	<i>Exp. 2</i>	Total
	149.35 h (83.1%)	30.45 h (16.9%)	179.80 h (100.0%)
REST	105.40 h (70.6%)	12.45 h (40.9%)	117.85 h (65.5%)
LIGHT	26.42 h (17.7%)	5.23 h (17.2%)	31.65 h (17.6%)
VIGOR	17.53 h (11.7%)	12.77 h (41.9%)	30.30 h (16.9%)

($BMI = 28.1 \pm 1.3 \text{ kg} \cdot \text{m}^{-2}$). Informed consent was obtained in all cases. Depending on their availability for the experiment, 5 subjects completed all of the 3 circuits in different days, one subject exercised for 2 sessions and the remaining 3 participants completed only one circuit. In total, 20 sessions were registered. Data were subsequently grouped into the three available PAI ranges, following the same class criteria as for *Exp. 1*.

Combining both experiments, the total amount of time in rest/sedentary situations dominated against the two other PAI classes (see Table 1), what in practice mimics reasonably well realistic data distributions for 24 h-based monitoring scenarios.

2.3 Methodology for Automatic PAI Assessment

2.3.1 Feature Definition. Signals were divided into segments by rectangular non-overlapping windows. Note that the length of such analysis window is an important parameter for our methods, implying a trade-off between: *a*) temporal resolution –short windows necessary–, and *b*) the definition of meaningful features able to capture relevant information about the underlying phenomena –longer windows preferred–. In a preliminary stage, we tested our algorithms for different window lengths, within a range of 0.5–10 min. Evaluating the obtained overall performance in terms of accuracy and temporal resolution, we opted for 2 min windows (12 samples with 10 s epochs).

An extra magnitude, acceleration norm $\|\mathbf{a}\| = \sqrt{\mathbf{a}_1^2 + \mathbf{a}_2^2}$, was defined to combine the information from the two axis. Afterwards, statistical descriptives were computed in the time-domain for \mathbf{a}_1 , \mathbf{a}_2 , $\|\mathbf{a}\|$, \mathbf{st} and \mathbf{hr} in each windowed segment; namely: means, standard deviations, medians, maxima, minima and Pearson’s correlation coefficients r_{XY} for low-pass filtered versions of the original signals. In total, 35 features were derived.

2.3.2 Dimensionality Reduction. Principal Component Analysis (PCA) was applied to reduce the amount of redundancy present in the data. Given that features are clearly not commensurate (for example: $|r_{XY}| \in [0, 1]$ by definition, while $\|\mathbf{a}\| > 1000$ in many cases), the original feature space was standardized (i.e. subtraction of μ and division by σ in each dimension). This was done to prevent the PCA eigenvectors from being dominated by those features with largest variance, what would have masked much of the relevant information.

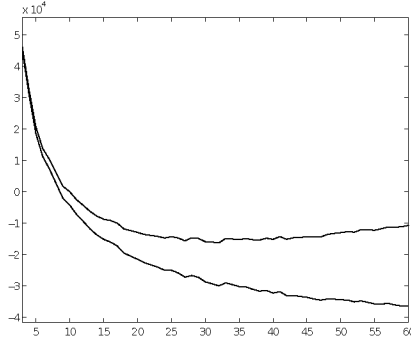


Fig. 1. Average values for Akaike (light line) and Bayes (dark line) information criteria –ordinate– with an increasing number of L mixtures in the GMM –abscissa–

We selected the $T=10$ eigenvectors with highest eigenvalues λ_i , which explained 81.5% of total variance (79.0% for $T=9$). After projection onto a new subspace formed by the eigenvectors $\{\mathbf{v}_i\}_{i=1}^T$, data were again standardized (dividing by $\sqrt{\lambda_i}$) in order to obtain commensurate magnitudes.

2.3.3 Unsupervised Clustering. Two different techniques were compared: k -means clustering and a Gaussian Mixture Model (GMM), the latter option as a particular case of Expectation-Maximization (EM) algorithm where the underlying probability density function is modeled as a mixture (i.e. linear combination) of L Gaussians.

Prior to the automatic learning phase itself, suitable values for parameters k and L had to be chosen. To do so, we assessed the complexity of our 10-dimensional feature space by fitting into the data a GMM with increasing L . For each iteration, average values of Akaike (AIC) and Bayes (BIC) information criteria were computed with a 10-fold cross-validation strategy. Figure 1 depicts an almost monotonically decreasing trend for AIC up to $L=60$ mixtures, while BIC reached a global minimum for $L=32$. Since AIC yielded a local minimum for this same number, we accepted $L=32$ as a reasonable compromise to avoid making the learning process unnecessarily complex. In addition, for k -means we selected $k=32$ as well, in order to compare both algorithms under equal circumstances (assuming none was favoured by such choice).

2.3.4 Temporal Model. In practise, when individuals exercise, their PAI does not usually fluctuate rapidly; on the contrary, changes tend to be gradual and intensity ranges are often kept almost constant for time periods that broadly exceed the length of our 2 min analysis windows. Consequently, it is common to observe long-term trends with a fixed intensity level; so that PAI for a particular moment shows strong correlation with respect to the PAI at neighbor instants. To benefit from this behaviour, we propose a multi-scale approach from a temporal point of view, where the lower scale corresponds to the span

of our analysis window and the higher scale is related to the scope of a Hidden Markov Model (HMM). While extensively used in speech and gesture recognition, different authors have also suggested HMMs to detect physical activities (e.g. falls, sit-to-stand transitions, etc.[19, 20]).

We built a single-layered HMM where: *a*) observations correspond to cluster assignments –which, being 32 possible symbols, lack of intuitive interpretation–, and where *b*) hidden states are assimilated to the targeted PAI classes. During the training phase, emission and transmission matrices (with size 3×32 and 3×3 , respectively) were estimated in a supervised manner, i.e. by direct comparison of cluster assignments and PAI labels from the ground truth. During the HMM-based classification stage, the algorithms calculated which sequence of (‘hidden’) PAI classes most likely originated a given (‘observed’) sequence of clusters.

2.3.5 Performance Evaluation. While total accuracy is the most common metric for assessing performance in classification tasks, it tends to undervalue achievements on multi-class problems or under notable class imbalance. As both circumstances apply in this work, we opted for using precision and recall metrics instead. For the i -th class:

$$\text{Precision}^{(i)} \equiv \frac{\text{TP}^{(i)}}{\text{TP}^{(i)} + \text{FP}^{(i)}} \quad \text{Recall}^{(i)} \equiv \frac{\text{TP}^{(i)}}{\text{TP}^{(i)} + \text{FN}^{(i)}} \quad (1)$$

Hence, the harmonic mean between precision and recall (often known as F-measure or F-score) is another convenient metric in our case:

$$\text{F-measure}^{(i)} \equiv 2 \frac{\text{Precision}^{(i)} \cdot \text{Recall}^{(i)}}{\text{Precision}^{(i)} + \text{Recall}^{(i)}} \quad (2)$$

Overall classification performance was estimated by 30 independent runs of a 10-fold cross-validation. Full stratification was dismissed because it would remove any temporal coherence. Instead, a partial stratification was implemented: Each of the 92 available sequences was randomly allocated into a fold. Folds were then checked to guarantee that their relative class frequencies were similar to the prior PAI distributions, and to assure that fold sizes were not highly uneven.

3 Results

The dimensionality reduction process showed HR-related features playing prominent roles in the first PCA eigenvectors. Outliers-sensitive features (e.g. maxima, minima) also exhibited a substantial discriminative power, instead of being adversely affected by noise or artifacts.

Table 2 presents a summary of results in terms of total accuracy; as well as precision, recall and F-measure for the three PAI. In general, both k -means and GMM-based algorithms yielded meritorious performance values considering the marked class imbalance in the dataset. k -means outperformed GMM in

Table 2. Classification performance of the proposed algorithms in the whole dataset

$n = 30$ $\mu \pm \sigma$ [min max]		<i>k</i> -means + HMM			GMM + HMM		
		Accuracy	$86.97 \pm 0.67\%$	[85.58% 88.21%]		$85.03 \pm 0.82\%$	[83.60% 86.96%]
REST	Precision	$94.83 \pm 0.45\%$	[93.97% 95.73%]		$97.11 \pm 0.46\%$	[96.11% 98.11%]	
	Recall	$93.76 \pm 0.89\%$	[90.99% 95.01%]		$91.08 \pm 0.72\%$	[89.42% 93.15%]	
	F-Measure	$94.29 \pm 0.48\%$	[92.99% 94.98%]		$94.00 \pm 0.45\%$	[93.12% 95.25%]	
LIGHT	Precision	$73.60 \pm 1.99\%$	[69.39% 77.61%]		$67.40 \pm 3.27\%$	[61.18% 72.32%]	
	Recall	$63.70 \pm 1.49\%$	[60.82% 66.31%]		$65.70 \pm 1.85\%$	[61.67% 70.33%]	
	F-Measure	$68.27 \pm 1.28\%$	[65.12% 70.97%]		$66.48 \pm 1.65\%$	[63.45% 69.67%]	
VIGOR	Precision	$71.89 \pm 1.76\%$	[67.39% 74.33%]		$64.77 \pm 1.57\%$	[60.84% 67.69%]	
	Recall	$85.01 \pm 1.90\%$	[81.13% 88.96%]		$81.85 \pm 3.61\%$	[75.50% 88.63%]	
	F-Measure	$77.88 \pm 1.26\%$	[75.31% 80.10%]		$72.28 \pm 1.85\%$	[68.86% 75.47%]	

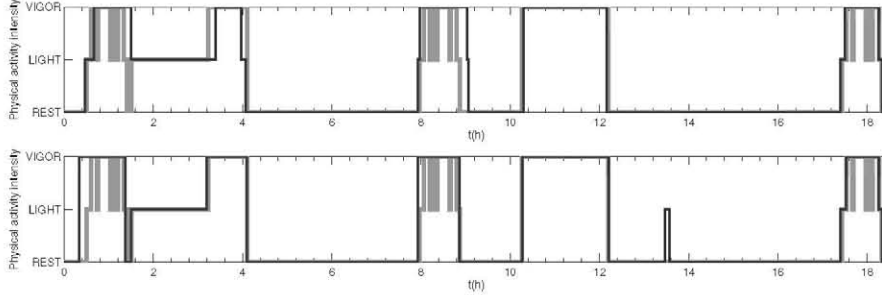


Fig. 2. Top: Example of *k*-means+HMM classification outputs (dark line, 92.45% accuracy) vs. ground truth (light line) along time. Bottom: GMM+HMM results (dark line, 94.06% accuracy) for the same ground truth sequence as above.

non-sedentary situations (average F-measures: 68.27% and 77.88% for *k*-means, versus 66.48% and 72.28% for GMM). A two-tailed unpaired t-test found significant differences between methods ($p < 0.01$ in all cases, except for F-measure in ‘REST’, where $p = 0.02$).

Figure 2 depicts an example of automatic classification results versus ground truth PAI annotations, obtained for a test subset of data kept apart from the ML training phase. In general, long-term mismatches seldom occur –none in Figure 2–. Most classification errors have a short duration and tend to be located: *a*) where the ground truth reference PAI levels oscillate rapidly (for our example, around 1.5, 8.5 and 18.0 h, where the circuit bouts from *Exp. 2* include ‘active rests’ for 5 min); or *b*) around sudden transients (especially from ‘REST’ to ‘VIGOR’ or vice versa, e.g. approx. at 0.5 and 9.0 h). Hence, on a 24 h basis the overall impact of this type of errors should be very limited in practice, with warm-up and cool-down exercising, etc.

4 Discussion

The accuracy achieved by our approach is remarkable (86.97% for k -means) and outperforms other PAI classifiers using combined accelerometry and HR, like Munguia Tapia[7] (58.40% accuracy in subject-independent classification). However, it must be noted that our proposal is explicitly intensity-oriented; while theirs was activity type-oriented (with certain types allowing several intensities), so that their number of possible classes was considerably higher.

In concordance with the conclusions by different researchers from the sport sciences background[2, 4, 5], we encountered that heart rate –and the features deriving from it– played a prominent role in PCA and classification, as they contain valuable information regarding physiological responses to exercise. In our view, this notable potential of HR has not yet been sufficiently exploited by methods from the ML background, since authors tended to focus mainly on accelerometry as their signal source.

Besides, the HMM introduced additional robustness into the algorithms by capturing temporal relationships among PAI labels for neighbor instants.

We observed that the recognition behaved worst for moderate intensities, with most miss-classifications occurring between ‘LIGHT’ and ‘VIGOR’ labels. A possible explanation is that, in free-living conditions, ground truth had to be based on volunteers’ self-reports, as it was the case for *Exp. 1*². Despite periods with ambiguous annotations were removed from the dataset, this dependence on subjective PAI grading poses the problem of introducing certain degree of uncertainty which cannot be eliminated from the ground truth. Nevertheless, for our research we considered essential not to restrict the analysis to laboratory situations, but to include an important amount of time in free-living conditions, to obtain datasets which assemble more realistic scenarios.

The significant class imbalance may also be a major responsible for the classifiers suffering lower performance at those underrepresented PAI levels. However, we obtained remarkable F-measures for these classes, what leads us to think that the algorithms are capable of compensating for most of the imbalance.

In current clinical and research practise (e.g. in epidemiological studies, to determine individuals’ adherence to exercise programs and their effects on health), physical activity monitoring in 24 h-based ambulatory scenarios utilizes often merely a single device: either a pedometer, accelerometer or HR monitor; thus dismissing the potential of combining information sources. In addition, manufacturers’ software for data analysis (as in the case of ActiTrainer) relies generally on simple thresholding to differentiate active periods from rest, and in linear regression for the EE computation. In contrast, our method may provide a more robust assessment of time exercised and PAI performed.

Furthermore, Crouter et al. proved in [1] that the accuracy obtained by these commercial EE estimation formulae may decrease substantially if the activity under monitoring is not comparable with the particular exercise for which the

² Such limitation was not present in *Exp. 2*, where volunteers exercised under direct supervision by the researchers.

equation was derived and validated. To solve this issue, several authors developed activity-specific EE estimation schemes[6–8]. Conversely, we intend to explore PAI-specific estimators based on the algorithms presented here.

5 Conclusions

This work presents an automatic algorithm based on statistical machine learning techniques for explicit assessment of physical activity intensity by means of simultaneous multi-axial accelerometry and heart rate signals. Our algorithms yielded up to 86.97% accuracy in PAI classification and up to 77.88% F-measure for non-sedentary situations. In addition, errors appeared mostly as brief transients. We therefore believe that our approach can be used for 24 h-based monitoring of physical activity in ambulatory scenarios and we suggest it as a first step towards the development of intensity-specific energy expenditure estimators.

Acknowledgments. The authors thank the members of the PRONAF Study Group for their assistance with data collection. This work was partially funded by a research grant of the Universidad Politécnica de Madrid and by the Spanish grants of the Ministry of Science and Innovation ‘APRIORI’ (FIS PS09/01318) and ‘PRONAF’ (DEP2008-06354-C04-01).

References

1. Crouter, S.E., Churilla, J.R., Bassett, D.R.: Estimating energy expenditure using accelerometers. *Eur. J. Appl. Physiol.* 98(6), 601–612 (2006)
2. Freedson, P.S., Miller, K.: Objective monitoring of physical activity using motion sensors and heart rate. *Res. Q Exercise Sport* 71(2 Suppl), S21 (2000)
3. Gotshalk, L.A., Berger, R.A., Kraemer, W.J.: Cardiovascular responses to a high-volume continuous circuit resistance training protocol. *J. Strength Cond. Res.* 18(4), 760 (2004)
4. Strath, S.J., Brage, S., Ekelund, U.: Integration of physiological and accelerometer data to improve physical activity assessment. *Med. Sci. Sport Exer.* 37(11), S563 (2005)
5. Plasqui, G., Westerterp, K.R.: Accelerometry and heart rate as a measure of physical fitness: Cross-validation. *Med. Sci. Sport Exer.* 38(8), 1510 (2006)
6. Bonomi, A.G., Plasqui, G., Goris, A.H.C., Westerterp, K.R.: Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *J. Appl. Physiol.* 107(3), 655 (2009)
7. Munguia Tapia, E., Intille, S.S., Haskell, W., Larson, K., Wright, J., King, A., Friedman, R.: Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In: 2007 11th IEEE International Symposium on Wearable Computers, pp. 37–40. IEEE, Los Alamitos (2007)
8. Albinali, F., Intille, S.S., Haskell, W., Rosenberger, M.: Using wearable activity type detection to improve physical activity energy expenditure estimation. In: Proceedings of the 12th ACM International Conference on Ubiquitous Computing, pp. 311–320. ACM, New York (2010)

9. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: Ferscha, A., Mattern, F. (eds.) *PERVASIVE 2004*. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
10. Allen, F.R., Ambikairajah, E., Lovell, N.H., Celler, B.G.: An adapted Gaussian Mixture Model approach to accelerometry-based movement classification using time-domain features. In: *28th Annual International Conference Engineering in Medicine and Biology Society, EMBS 2006*, pp. 3600–3603. IEEE, Los Alamitos (2008)
11. Ermes, M., Parkka, J., Mantyjarvi, J., Korhonen, I.: Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE T. Inf. Technol. B* 12(1), 20–26 (2008)
12. Krause, A., Siewiorek, D.P., Smailagic, A., Farringdon, J.: Unsupervised, dynamic identification of physiological and activity context in wearable computing. In: *Proceedings of the 7th IEEE International Symposium on Wearable Computers (ISWC 2003)*. IEEE Computer Society, Los Alamitos (2003)
13. Lester, J., Choudhury, T., Borriello, G.: A practical approach to recognizing physical activities. In: Fishkin, K.P., Schiele, B., Nixon, P., Quigley, A. (eds.) *PERVASIVE 2006*. LNCS, vol. 3968, pp. 1–16. Springer, Heidelberg (2006)
14. Duchêne, F., Garbay, C., Rialle, V.: Learning recurrent behaviors from heterogeneous multivariate time-series. *Artif. Intell. Med.* 39(1), 25–47 (2007)
15. Freedson, P.S., Melanson, E., Sirard, J.: Calibration of the Computer Science and Applications, Inc. accelerometer. *Med. Sci. Sport Exer.* 30(5), 777 (1998)
16. Brage, S., Brage, N., Franks, P.W., Ekelund, U., Wong, M.Y., Andersen, L.B., Froberg, K., Wareham, N.J.: Branched equation modeling of simultaneous accelerometry and heart rate monitoring improves estimate of directly measured physical activity energy expenditure. *J. Appl. Physiol.* 96(1), 343 (2004)
17. Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M., Swartz, A.M., Strath, S.J., O'Brien, W.L., Bassett Jr., D.R., Schmitz, K.H., Emplaincourt, P.O., et al.: Compendium of physical activities: An update of activity codes and MET intensities. *Med. Sci. Sport Exer.* 32(9 Suppl), S498 (2000)
18. Benito Peinado, P.J., Álvarez Sánchez, M., Díaz Molina, V., Peinado Lozano, A.B., Calderón Montero, F.J.: Aerobic Energy Expenditure and Intensity Prediction during a Specific Circuit Weight Training: A Pilot Study. *J. Hum. Sport Exerc.* 5(2), 134–145 (2010)
19. Lester, J., Choudhury, T., Kern, N., Borriello, G., Hannaford, B.: A hybrid discriminative/generative approach for modeling human activities. In: *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI), Citeseer* (2005)
20. He, J., Li, H., Tan, J.: Real-time daily activity classification with wireless sensor networks using Hidden Markov Model. In: *29th Annual International Conference Engineering in Medicine and Biology Society, EMBS 2007*, pp. 3192–3195. IEEE, Los Alamitos (2007)
21. García-García, F., Martínez-Sarriegui, I., Gómez, E.J., Rigla, M., Hernando, M.E.: Automatic assessment of physical activity using multi-axial accelerometry and heart rate. *J. Diab. Tech. Therap.* 13(2), 182 (2011)