Human Activity Recognition Using Accelerometer and Photoplethysmographic Signals

Giorgio Biagetti, Paolo Crippa(✉), Laura Falaschetti, Simone Orcioni, and Claudio Turchetti

DII – Dipartimento di Ingegneria dell'Informazione,
Università Politecnica delle Marche, Via Brecce Bianche 12, 60131 Ancona, Italy
{g.biagetti,p.crippa,l.falaschetti,s.orcioni,c.turchetti}@univpm.it

Abstract. This paper presents an efficient technique for real-time recognition of human activities by using accelerometer and photoplethysmography (PPG) data. It is based on singular value decomposition (SVD) and truncated Karhunen-Loève transform (KLT) for feature extraction and reduction, and Bayesian classification for class recognition. Due to the nature of signals, and being the algorithm independent from the orientation of the inertial sensor, this technique is particularly suitable for implementation in smartwatches in order to both recognize the exercise being performed and improve the motion artifact (MA) removal from PPG signal for accurate heart rate (HR) estimation. In order to demonstrate the validity of this methodology, it has been successfully applied to a database of accelerometer and PPG data derived from four dynamic activities.

Keywords: Activity recognition · Accelerometer · Photoplethysmography (PPG) · Motion artifact reduction · Real-time · Health · Fitness · Bayesian classification · Singular Value Decomposition (SVD) · Expectation Maximization (EM) · Karhunen-Loève Transform (KLT) · Feature extraction

1 Introduction

Human activity recognition using wearable sensors, i.e. sensors that are positioned directly or indirectly on the human body, is one of the most interesting topics in the healthcare, ambient assisted living, sport and fitness research areas.

These sensors, which can be embedded into clothes, shoes, belts, sunglasses, smartwatches and smartphones, or positioned directly on the body generate signals (accelerometric, photoplethysmography (PPG) [2,5,23], electrocardiography (ECG) [3,14], surface electromyography (sEMG) [9], ...) that can be used to collect information such as body position and movement, heart rate (HR), muscle fatigue of the user performing activities [2,4,11]. In particular, exercise routines and repetitions can be counted in order to track a workout routine as well as determine the

I. Czarnowski et al. (eds.), *Intelligent Decision Technologies 2017*,
Smart Innovation, Systems and Technologies 73, DOI 10.1007/978-3-319-59424-8_6

energy expenditure of individual movements. Indeed, mobile fitness coaching has involved topics ranging from quality of performing such sports actions to detection of the specific sports activity [7].

On the one hand, among wearable sensors, accelerometers are probably the most frequently used for activity monitoring. In particular, they are effective in monitoring actions that involve repetitive body motions, such as walking, running, cycling, climbing stairs [8,13,18,20,21]. Because smartphones and smartwatches have become very popular, their accelerometer sensors can be used for providing accurate and reliable information on peoples activities and behaviors, thus ensuring a safe and sound living environment [1,15,17]. There are several techniques based on signal processing and neural networks for representing nonlinear transformations derived from stochastic systems such as the human body [22].

On the other hand, PPG is a well-known noninvasive method for monitoring the HR that shines light into the body and measures the amount of light that is reflected back to measure the blood flow. Unlike the ECG and the sEMG monitoring that need sticky metal electrodes across the body skin in order to monitor electrical activity from heart and muscles [6,10], PPG monitoring can be performed at peripheral sites on the body and needs a simpler body contact. As a result, PPG sensors are more and more used in wearable devices (smartwatches), as the preferred modality for HR monitoring in everyday activities by non-specialist users. However accurate estimation of PPG signal recorded from subject wrist, when the subject is performing various physical exercises, is often a challenging problem as the raw PPG signal is severely corrupted by motion artifacts (MAs). These are principally due to the relative movement between the PPG light source/detector and the wrist skin of the subject during motion. In order to reduce the MAs, a number of signal processing techniques based on data derived from the smartphone built-in triaxial accelerometer have been proven to be very useful [5,23].

This paper proposes an efficient technique for real-time recognition of human activities, by using data gathered from accelerometer and PPG sensors. The proposed technique is based on singular value decomposition (SVD) and truncated Karhunen-Loève transform (KLT) for feature extraction and reduction, and Bayesian classification for class recognition. The algorithm is independent of the orientation of the accelerometer sensor making it particularly suitable for implementation in wearable devices such as smartphones where the orientation of the sensors can be unknown or their placement could be not always correct. The algorithm could be used to both recognize the exercise being performed and improve the tracking of the PPG signal for accurate HR estimation. In order to demonstrate the validity of this technique, it has been successfully applied to a database of accelerometer and PPG data derived from four dynamic activities.

The paper is organized as follows. Section 2 provides a brief overview of the human activity recognition algorithm. Section 3 presents the experimental results carried out on a public domain data set in order to show the effectiveness of the proposed technique. Finally, the conclusions of this work are drawn in Sect. 4.

2 Recognition Algorithm

In this section both the model generation and the recognition algorithm will be presented. They exploit a Bayesian classifier based on a Gaussian mixture model (GMM) [16] of the probability density functions of the dimensionality-reduced feature vectors. The feature vectors themselves are built from the normalized singular value spectrums of both the accelerometer and the PPG Hankel data matrices. A schematic diagram of the activity detection algorithm is shown in Fig. 1.

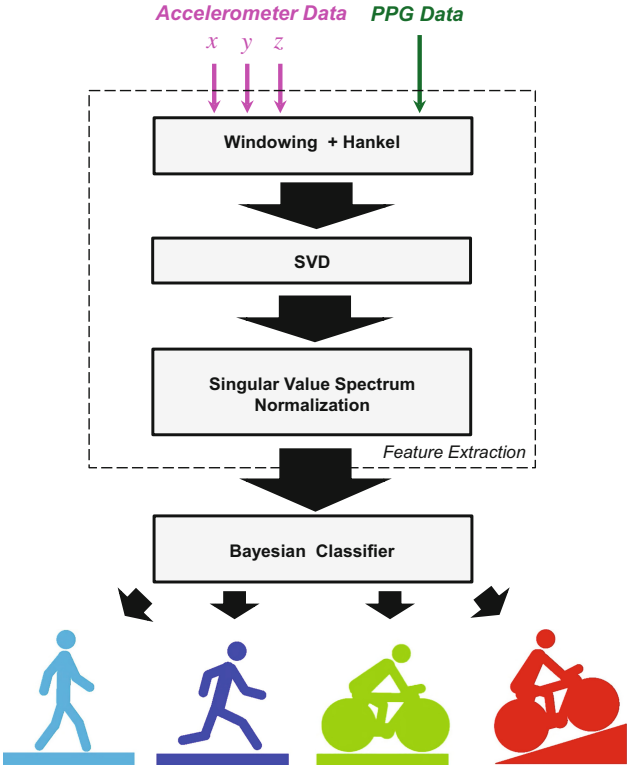

Fig. 1. Flow chart of the proposed framework for human activity classification (x, y, z are the 3-axial accelerometer signals).

In order to recognize the activity being performed at a given time instant t, the incoming signals are first windowed into relatively short windows (8 s in our sample application) wherein the activity can reasonably be considered invariant. From these slices of signals, a compact set of features ξ_t is extracted as detailed next.

Let x, y, z be the accelerometer signals and p the PPG signal, sliced into windows $N + L - 1$ samples long, indicated as $p_t = [p(t) \dots p(t + N - 1)]^T$.

Then let $P_t = [p_t^{(1)} \dots p_t^{(L)}]$, with $p_t^{(i)} = p_{t+i-1}$, be the Hankel matrix derived from the PPG signal. Analogously, let $X_t = [x_t^{(1)} \dots x_t^{(L)}]$, $Y_t = [y_t^{(1)} \dots y_t^{(L)}]$, and $Z_t = [z_t^{(1)} \dots z_t^{(L)}]$ be the Hankel data matrices for the three accelerometer signals, where $x_t^{(i)}$, $y_t^{(i)}$, $z_t^{(i)}$, $i = 1, \dots, L$, represent the observations achieved from a three-axes accelerometer, each shifted in time by i samples, just as $p_t^{(i)}$.

As in [8], the accelerometer signals are grouped together to avoid dependence on sensor orientation, but the PPG data is left on its own as it is not commensurable with acceleration. So, the matrices

$$H_t^{\mathrm{A}} = [X_t\, Y_t\, Z_t] \in \mathbb{R}^{N \times 3L} \qquad H_t^{\mathrm{P}} = [P_t] \in \mathbb{R}^{N \times L} \tag{1}$$

can be represented by their singular value decomposition (SVD) as

$$H_t^{\mathrm{A}} = S_t^{\mathrm{A}} \Lambda_t^{\mathrm{A}} R_t^{\mathrm{A}T} = \sum_{i=1}^{N} \lambda_i^{\mathrm{A}} s_i^{\mathrm{A}} r_i^{\mathrm{A}T}, \tag{2}$$

where, if $N < 3L$, $S_t^{\mathrm{A}} = \left[s_1^{\mathrm{A}} \dots s_N^{\mathrm{A}}\right]$, $R_t^{\mathrm{A}} = \left[r_1^{\mathrm{A}} \dots r_N^{\mathrm{A}}\right]$, with s_i^{A}, r_i^{A} being the corresponding left and right singular vectors, and λ_i^{A} are the singular values in decreasing order $\lambda_1^{\mathrm{A}} \geq \lambda_2^{\mathrm{A}} \geq \dots \geq \lambda_N^{\mathrm{A}}$. Similarly, from the PPG signal p, we compute the SVD of its Hankel matrix as

$$H_t^{\mathrm{P}} = S_t^{\mathrm{P}} \Lambda_t^{\mathrm{P}} R_t^{\mathrm{P}T} = \sum_{i=1}^{N} \lambda_i^{\mathrm{P}} s_i^{\mathrm{P}} r_i^{\mathrm{P}T}, \tag{3}$$

where λ_i^{P} are the singular values in decreasing order $\lambda_1^{\mathrm{P}} \geq \lambda_2^{\mathrm{P}} \geq \dots \geq \lambda_N^{\mathrm{P}}$.

The chosen feature vector is defined as

$$\xi_t = \begin{bmatrix} \Lambda_t^{\mathrm{A}} / ||\Lambda_t^{\mathrm{A}}|| \\ w\, \Lambda_t^{\mathrm{P}} / ||\Lambda_t^{\mathrm{P}}|| \end{bmatrix} \in \mathbb{R}^{2N} \tag{4}$$

where $||\cdot||$ represents the norm of a vector and w is a weighting factor to be determined. Since the dimension $2N$ is usually too high to directly model a GMM on it, it is reduced by means of a truncated Karhunen-Loève transformation to a vector k_{tM} of lower dimension by a linear application Ψ such that $\mathrm{k}_{tM} = \Psi\, \xi_t$ where $\xi_t \in \mathbb{R}^{2N}$, $\mathrm{k}_{tM} \in \mathbb{R}^{M}$, $\Psi \in \mathbb{R}^{M \times 2N}$, and $M \ll 2N$.

Let us refer to a frame $\mathrm{k}_{tM}[n]$, $n = 0, \dots, M-1$, containing compacted features extracted from both the accelerometer and PPG signals. For Bayesian classification, a group of Γ activities is represented by the probability density functions (pdfs) $p_\gamma(\mathrm{k}_{tM}) = p(\mathrm{k}_{tM} \mid \theta_\gamma)$, $\gamma = 1, 2, \cdots, \Gamma$, where θ_γ are the parameters to be estimated during training.

The objective of classification is to find the activity $\widehat{\gamma}$ which has the maximum *a posteriori* probability for a given frame k_{tM}, i.e.

$$\widehat{\gamma}(\mathrm{k}_{tM}) = \underset{1 \leq \gamma \leq \Gamma}{\operatorname{argmax}} \left\{ \frac{p(\mathrm{k}_{tM} \mid \theta_\gamma) p(\theta_\gamma)}{p(\mathrm{k}_{tM})} \right\} = \underset{1 \leq \gamma \leq \Gamma}{\operatorname{argmax}} \{ p(\mathrm{k}_{tM} \mid \theta_\gamma) \}, \tag{5}$$

assuming equally likely activities (i.e. $p(\theta_\gamma) = 1/\Gamma$) and noting that $p(\mathrm{k}_{tM})$ is the same for all activity models.

The statistical model we adopted for $p(\mathrm{k}_{tM} \mid \theta_\gamma)$ of the γ-th exercise is the GMM [19] given by the equation

$$p(\mathrm{k}_{tM} \mid \theta_\gamma) = \sum_{i=1}^{F} \alpha_i \mathcal{N}(\mathrm{k}_{tM} \mid \mu_i, \mathrm{C}_i) \tag{6}$$

where α_i, $i = 1, \ldots, F$ are the mixing weights, and $\mathcal{N}(\mathrm{k}_{tM} \mid \mu_i, \mathrm{C}_i)$ represents the density of a Gaussian distribution with mean μ_i and covariance matrix C_i, and θ_γ is the set of parameters defined as $\theta_\gamma = \{\alpha_1, \mu_1, \mathrm{C}_1, \ldots, \alpha_F, \mu_F, \mathrm{C}_F\}$.

To obtain an estimate of the mixture parameters we used a variant of the expectation maximization (EM) algorithm [16], which integrates both model estimation and component selection, i.e. the ability of choosing the best number of mixture components F according to a predefined minimization criterion, in a single framework.

3 Experimental Results

PPG recordings were collected from 8 subjects, for approximately 5 min each, as they undertook a range of different physical activities on a treadmill and on an exercise bike. The activities performed by each person are detailed in Table 1. All of the signals (including a simultaneous ECG for gold standard heart rate) belong to the Physionet database in https://physionet.org/works/WristPPGduringexercise/ [12]. We follow the original Authors' naming of the subjects, i.e., s1, ..., s6, s8, s9, and we splitted the available signals between a training set and a testing set, so as to include exactly two signals for each activity type in the training, as detailed in the same table.

Table 1. Distribution of the activities performed by the various subjects. The database was also split in a training set (*), comprising subjects s1, s2, s4, and s8 (8 signals in total), and a testing set (+), comprising subjects s3, s5, s6, and s9 (10 signals in total, all of the remaining).

Activities	s1	s2	s3	s4	s5	s6	s8	s9
High resistance bike (H)	*	*	+					
Low resistance bike (L)	*	*	+		+	+		
Run (R)			+	*	+	+	*	
Walk (W)	*		+			+	*	+

The available data was sampled at 256 Hz, and it was sliced in overlapping windows of 2048 samples (8 s long), each window being shifted by 512 samples (2 s). These windows were used to build four $N \times L$ Hankel matrices, one for each acceleration direction and one for the PPG signal, with $N = 400$ and $L = 1649$. The acceleration-derived matrices are then fed together to the SVD, to remove

sensor orientation effects, and the PPG-derived matrix is fed to an independent SVD block. The resulting normalized singular values are concatenated, with a weight w applied to the PPG-derived values, and are used as the feature vectors for the classifier, after KLT-based dimensionality reduction to $M = 10$ principal components.

As an example, the average of these feature vectors, within each activity, is shown in Fig. 2. It is apparent that different types of motion (bike, run, walk) produce differing distribution of the singular values, making recognition of the class of activity quite easy. Unfortunately, there is almost no way to differentiate the resistance level on the bike.

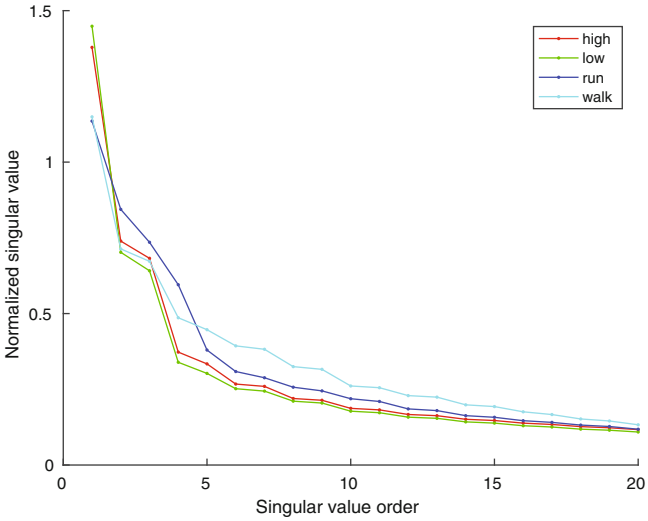

Fig. 2. Per-class average of a portion (extracted from the accelerometer data) of the feature vectors.

This is confirmed by a recognition experiment, whose results are reported in Table 2. Here the test was performed disregarding the PPG signal (i.e., by setting $w = 0$), and the first two classes are often confused. The PPG signal can help differentiate between these two cases, that differs only in the effort and not in the type of motion.

Unfortunately, as can be seen in Table 3, which reports the results of an experiment made disregarding the accelerometer signal, PPG alone is not a good candidate to recognize the physical activity. To see if it can help in discriminating the effort it must be combined with the accelerometer signal.

In order to appropriately combine the information coming from the accelerometer and from the PPG for identification purposes, a 4-fold cross-validation was performed on the training set, leaving out one of the subjects at a time and averaging the overall recognition accuracy on the three remaining

Table 2. Confusion matrix obtained using only the accelerometer data as the feature vector. The resulting overall accuracy is 65.7%.

Input	Recognized			
	H	L	R	W
H	127	9	0	1
L	220	170	16	9
R	8	30	408	1
W	95	63	32	220

Table 3. Confusion matrix obtained using only the PPG data as the feature vector. The resulting overall accuracy is 44.7%.

Input	Recognized			
	H	L	R	W
H	20	117	0	0
L	5	140	172	98
R	0	0	157	290
W	0	0	97	313

Table 4. Confusion matrix obtained using both the PPG and the accelerometer data as the feature vector, with the PPG features being optimally scaled with $w = 0.155$. The resulting overall accuracy is 78.0%.

Input	Recognized			
	H	L	R	W
H	120	17	0	0
L	65	232	49	69
R	0	0	435	12
W	0	9	89	312

Table 5. Summary of the performance obtained using both the PPG and the accelerometer data as the feature vector, with the PPG features being optimally scaled with $w = 0.155$.

Activities	Sensitivity [%]	Precision [%]	F1-score [%]
H	87.59	64.86	74.53
L	55.90	89.92	68.95
R	97.32	75.92	85.29
W	76.10	79.39	77.71

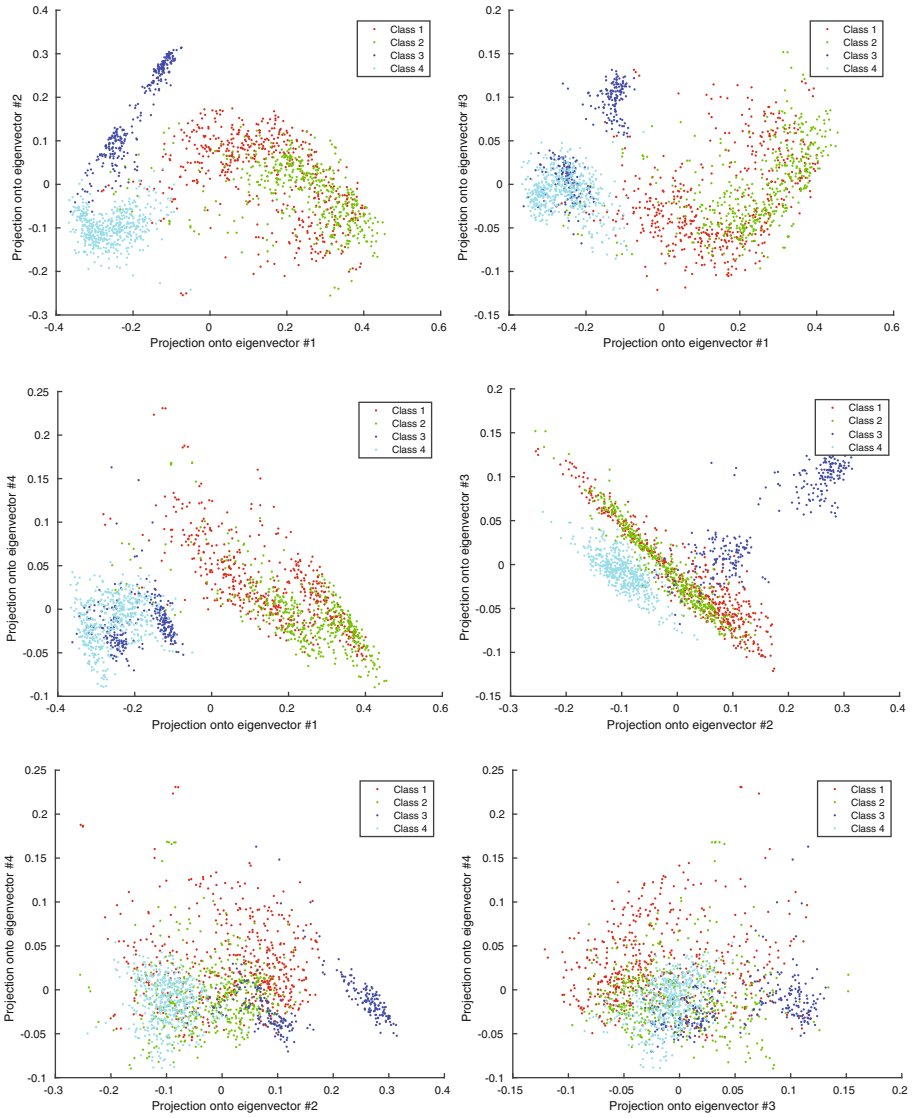

Fig. 3. Projections of the training set over the first four eigenvectors.

subjects, while varying the weight w used to combine the two signals. The optimum value was found to be $w = 0.155$. The results of this experiment are shown in Table 4, and it clearly improved the overall accuracy from 65.7% to 78.0%. For a better view of the performance, a few indices are also reported in Table 5.

Finally, Fig. 3 shows projections of the feature vectors over the first few eigenvectors. As can be seen, on the strongest projections the first two classes (bike)

have a significant overlap, but they can be separated by higher-order projection thanks to the introduction of the PPG signal.

4 Conclusion

In this paper we extended the results presented in [8] to also take into account PPG data, besides acceleration data, to help in human activity identification. In fact, there are sometimes different activities that can usefully be differentiated but that involves essentially the same movements, differing only in the amount of effort exerted, like pedaling on a bike at different resistance levels. These activities are difficult to identify with acceleration alone, as the results here presented show. By adding information from a PPG sensor, recognition accuracy can be considerably improved provided the two types of signals are appropriately combined. To this end, a cross-validation approach was employed, resulting in an improvement from an initial 65.7% to 78.0% overall accuracy on experiments conducted on a publicly-available data set.

References

1. Anguita, D., Ghio, A., Oneto, L., Parra, X., Reyes-Ortiz, J.L.: Energy efficient smartphone-based activity recognition using fixed-point arithmetic. J. Univ. Comput. Sci. **19**(9), 1295–1314 (2013)
2. Bacà, A., Biagetti, G., Camilletti, M., Crippa, P., Falaschetti, L., Orcioni, S., Rossini, L., Tonelli, D., Turchetti, C.: CARMA: a robust motion artifact reduction algorithm for heart rate monitoring from PPG signals. In: 23rd European Signal Processing Conference, pp. 2696–2700, September 2015
3. Biagetti, G., Crippa, P., Curzi, A., Orcioni, S., Turchetti, C.: A multi-class ECG beat classifier based on the truncated KLT representation. In: UKSim-AMSS 8th European Modelling Symposium on Computer Modelling and Simulation (EMS 2014), pp. 93–98, October 2014
4. Biagetti, G., Crippa, P., Curzi, A., Orcioni, S., Turchetti, C.: Analysis of the EMG signal during cyclic movements using multicomponent AM-FM decomposition. IEEE J. Biomed. Health Inf. **19**(5), 1672–1681 (2015)
5. Biagetti, G., Crippa, P., Falaschetti, L., Orcioni, S., Turchetti, C.: Artifact reduction in photoplethysmography using Bayesian classification for physical exercise identification. In: Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods, Rome, Italy, pp. 467–474, February 2016
6. Biagetti, G., Crippa, P., Falaschetti, L., Orcioni, S., Turchetti, C.: Wireless surface electromyograph and electrocardiograph system on 802.15.4. IEEE Trans. Consum. Electron. **62**(3), 258–266 (2016)
7. Biagetti, G., Crippa, P., Falaschetti, L., Orcioni, S., Turchetti, C.: A rule based framework for smart training using sEMG signal. In: Neves-Silva, R., Jain, L.C., Howlett, R.J. (eds.) Intelligent Decision Technologies, Smart Innovation, Systems and Technologies, vol. 39, pp. 89–99. Springer, Cham (2015)
8. Biagetti, G., Crippa, P., Falaschetti, L., Orcioni, S., Turchetti, C.: An efficient technique for real-time human activity classification using accelerometer data. In: Czarnowski, I., Caballero, A.M., Howlett, R.J., Jain, L.C. (eds.) Intelligent Decision Technologies 2016: Proceedings of the 8th KES International Conference on Intelligent Decision Technologies - Part I, pp. 425–434. Springer, Cham (2016)

9. Biagetti, G., Crippa, P., Falaschetti, L., Orcioni, S., Turchetti, C.: Homomorphic deconvolution for MUAP estimation from surface EMG signals. IEEE J. Biomed. Health Inf. **21**(2), 328–338 (2017)
10. Biagetti, G., Crippa, P., Orcioni, S., Turchetti, C.: An analog front-end for combined EMG/ECG wireless sensors. In: Conti, M., Martínez Madrid, N., Seepold, R., Orcioni, S. (eds.) Mobile Networks for Biometric Data Analysis, pp. 215–224. Springer, Cham (2016)
11. Biagetti, G., Crippa, P., Orcioni, S., Turchetti, C.: Surface EMG fatigue analysis by means of homomorphic deconvolution. In: Conti, M., Martínez Madrid, N., Seepold, R., Orcioni, S. (eds.) Mobile Networks for Biometric Data Analysis, pp. 173–188. Springer, Cham (2016)
12. Casson, A.J., Galvez, A.V., Jarchi, D.: Gyroscope vs. accelerometer measurements of motion from wrist PPG during physical exercise. ICT Expr. **2**(4), 175–179 (2016)
13. Catal, C., Tufekci, S., Pirmit, E., Kocabag, G.: On the use of ensemble of classifiers for accelerometer-based activity recognition. Appl. Soft. Comput. **37**, 1018–1022 (2015)
14. Crippa, P., Curzi, A., Falaschetti, L., Turchetti, C.: Multi-class ECG beat classification based on a Gaussian mixture model of Karhunen-Loève transform. Int. J. Simul. Syst. Sci. Technol. **16**(1), 2.1–2.10 (2015)
15. Dernbach, S., Das, B., Krishnan, N.C., Thomas, B.L., Cook, D.J.: Simple and complex activity recognition through smart phones. In: 8th International Conference on Intelligent Environments, pp. 214–221, June 2012
16. Figueiredo, M.A.F., Jain, A.K.: Unsupervised learning of finite mixture models. IEEE Trans. Pattern Anal. Mach. Intell. **24**(3), 381–396 (2002)
17. Khan, A., Lee, Y.K., Lee, S., Kim, T.S.: Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. In: 2010 5th International Conference on Future Information Technology, pp. 1–6, May 2010
18. Mannini, A., Intille, S.S., Rosenberger, M., Sabatini, A.M., Haskell, W.: Activity recognition using a single accelerometer placed at the wrist or ankle. Med. Sci. Sports Exerc. **45**(11), 2193–2203 (2013)
19. Reynolds, D.A., Rose, R.C.: Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Trans. Speech Audio Process. **3**(1), 72–83 (1995)
20. Rodriguez-Martin, D., Samà, A., Perez-Lopez, C., Català, A., Cabestany, J., Rodriguez-Molinero, A.: SVM-based posture identification with a single waist-located triaxial accelerometer. Expert Syst. Appl. **40**(18), 7203–7211 (2013)
21. Torres-Huitzil, C., Nuno-Maganda, M.: Robust smartphone-based human activity recognition using a tri-axial accelerometer. In: 2015 IEEE 6th Latin American Symposium on Circuits Systems, pp. 1–4, February 2015
22. Turchetti, C., Crippa, P., Pirani, M., Biagetti, G.: Representation of nonlinear random transformations by non-Gaussian stochastic neural networks. IEEE Trans. Neural Networks **19**(6), 1033–1060 (2008)
23. Zhang, Z., Pi, Z., Liu, B.: TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Trans. Biomed. Eng. **62**(2), 522–531 (2015)