

# An Efficient Technique for Real-Time Human Activity Classification Using Accelerometer Data

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

**Abstract** Accurate estimation of biometric parameters recorded from subjects' wrist or waist, when the subjects are performing various physical exercises, is often a challenging problem due to the presence of motion artifacts. In order to reduce the motion artifacts, data derived from a triaxial accelerometer have been proven to be very useful. Unfortunately, wearable devices such as smartphones and smartwatches are in general differently oriented during real life activities, so the data derived from the three axes are mixed up. This paper proposes an efficient technique for real-time recognition of human activities by using accelerometer data that is based on singular value decomposition (SVD) and truncated Karhunen-Loève transform (KLT) for feature extraction and reduction, and Bayesian classification for class recognition, that is independent of the orientation of the sensor. This is particularly suitable for implementation in wearable devices. In order to demonstrate the validity of this technique, it has been successfully applied to a database of accelerometer data derived from static postures, dynamic activities, and postural transitions occurring between the static postures.

**Keywords** Activity detection · Accelerometer · Real-time · Smartphone · Health · Fitness · Bayesian classification · Singular value decomposition · SVD · Expectation maximization · EM · Feature extraction

---

G. Biagetti · P. Crippa (✉) · L. Falaschetti · S. Orcioni · C. Turchetti  
DII – Dipartimento di Ingegneria dell'Informazione,  
Università Politecnica delle Marche, Via Brecce Bianche 12, I-60131 Ancona, Italy  
e-mail: p.crippa@univpm.it

G. Biagetti  
e-mail: g.biagetti@univpm.it

L. Falaschetti  
e-mail: l.falaschetti@univpm.it

S. Orcioni  
e-mail: s.orcioni@univpm.it

C. Turchetti  
e-mail: c.turchetti@univpm.it

# 1 Introduction

Human activity recognition is one of the most interesting research topics for several areas such as pervasive and mobile computing, ambient assisted living, surveillance-based security, sport and fitness activities, healthcare.

Over the recent years sensor technologies especially high-capacity, low-power, low-cost, miniaturized and lightweight sensors, wired, wireless and hybrid communication network protocols as well as signal processing theory have greatly progressed.

Wearable sensors, i.e. sensors that are positioned directly or indirectly on the human body, generate signals (accelerometric, PPG, ECG, sEMG, ...) when the user performs activities. Therefore they can monitor features that are descriptive of the person's physiological state or movement.

These sensors can be embedded into clothes, shoes, belts, sunglasses, smartwatches and smartphones, or positioned directly on the body and can be used to collect information such as body position and movement, heart rate, muscle fatigue and skin temperature [3, 4].

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, sitting, standing, and climbing stairs.

On the one hand, activity classification using accelerometers can be obtained using one or more sensors on the body [7, 12]. However single sensor systems are more practical and in this case common location choices are the waist, upper arm, wrist, and ankle [13, 15, 16]. The waist location has been used extensively in physical activity measurements because it captures major body motions, but algorithms using waist data can underestimate overall expenditure on activities such as bicycling or arm ergometry, where the waist movement is uncorrelated to the movement of the limbs. Therefore several recent studies addressed the problem of detecting human activities from smartphones and smartwatches [1, 8, 11].

On the other hand, accurate estimation of biometric parameters recorded from subjects' wrist or waist, when the subjects are performing various physical exercises, is often a challenging problem due to the presence of motion artifacts. In order to reduce the motion artifacts, data derived from a triaxial accelerometer have been proven to be very useful [6].

Wrist-worn sensor devices can be comfortably used during activities of daily living, including sleep, and can remain active during changing of clothes and do not require special belts or clips, thus increasing the wear time. Therefore smartwatches for human activity monitoring are becoming very important tools in personal health monitoring. In particular, exercise routines and repetitions can be counted in order to track a workout routine as well as determine the 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 [5].

However, wearable devices such as smartphones and smartwatches are in general differently oriented during human activities, so the data derived from the three axes are mixed up.

This paper proposes an efficient technique for real-time recognition of human activities, and transitions between them, by using accelerometer data. 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 sensor making it particularly suitable for implementation in wearable devices such as smartphones where the orientation of the sensor can be unknown or its placement could be not always correct. In order to demonstrate the validity of this technique, it has been successfully applied to a database of accelerometer data derived from static postures, dynamic activities, and postural transitions occurring between the static postures.

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

## 2 Recognition Algorithm

This section presents a description of the overall algorithm, based on the dimensionality-reduced singular value spectrum of the data Hankel matrix and on the Bayesian classifier, that is able to identify human activity classes from accelerometer data.

A schematic diagram of the activity detection algorithm, is shown in Fig. 1.

### 2.1 Data Preprocessing and Feature Extraction

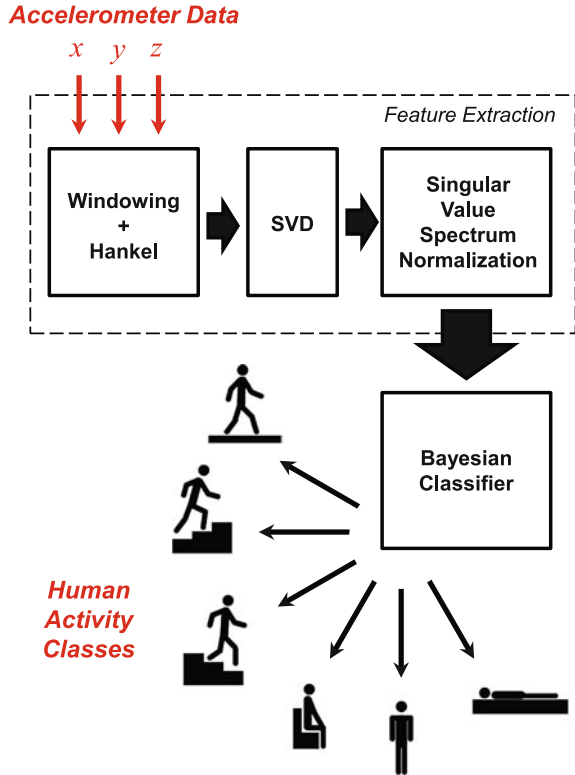
Let  $x$ ,  $y$ ,  $z$  be the accelerometer signals. After a preprocessing stage where the raw data have been windowed into windows  $N + L - 1$  samples long, the resulting accelerometer signals have been manipulated as follows.

Let  $x_t = [x(t) \dots x(t + N - 1)]^T$ ,  $x_t^{(i)} = x_{t+i-1}$ , and  $X_t = [x_t^{(1)} \dots x_t^{(L)}]$ . Analogously, let  $Y_t = [y_t^{(1)} \dots y_t^{(L)}]$  and  $Z_t = [z_t^{(1)} \dots z_t^{(L)}]$ . The matrices  $X_t$ ,  $Y_t$ , and  $Z_t$  so built are the Hankel data matrices of the three accelerometer signals, where  $x_t^{(i)}$ ,  $y_t^{(i)}$ ,  $z_t^{(i)}$ ,  $i = 1, \dots, L$ , represent the observations achieved from the three-axes accelerometer, each shifted in time by  $i$  samples.

The complete matrix of sample signals

$$H_t = [X_t Y_t Z_t] \in \mathbb{R}^{N \times 3L} \quad (1)$$

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



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

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

where, if  $N < 3L$ ,  $S_t = [s_1 \dots s_N]$ ,  $R_t = [r_1 \dots r_N]$ , with  $s_i, r_i$  being the corresponding left and right singular vectors, and  $\lambda_i$  are the singular values in decreasing order  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ .

By denoting with  $H_t \in \mathbb{R}^{N \times 3L}$  the data matrix of the accelerometer signals at each time instant  $t$ , in order to apply the human activity classification algorithm, a feature vector  $\xi_t$  has to be derived from this matrix.

We noticed that different types of activities lead to different distributions of the energy of the accelerometer signals among its eigenvectors. Thus, a suitable candidate for identifying the type of activity is the normalized spectrum of singular values  $\Lambda_t = [\lambda_1 \dots \lambda_N]$ , so as to avoid dependence on the intensity of the activity. Therefore we choose  $\xi_t = \Lambda_t / \|\Lambda_t\|$  where  $\|\cdot\|$  represents the norm of a vector.

In order to face the problem of dimensionality, the usual choice [10] is to reduce the vector  $\xi_t$  to a vector  $k_{tM}$  of lower dimension by a linear non-invertible transform  $\Psi$  (a rectangular matrix) such that

$$k_{tM} = \Psi \xi_t, \quad (3)$$

where  $\xi_t \in \mathbb{R}^N$ ,  $k_{tM} \in \mathbb{R}^M$ ,  $\Psi \in \mathbb{R}^{M \times N}$ , and  $M \ll N$ .

It is well known that, among the allowable linear transforms  $\Psi : \mathbb{R}^N \rightarrow \mathbb{R}^M$ , the Karhunen-Loève transform truncated to  $M < N$  orthonormal basis functions, is the one that ensures the minimum mean square error.

This normalized singular value spectrum can easily be computed immediately after having performed the SVD on the accelerometer signals, and used as input to the Bayesian classifier after a KLT-based dimensionality reduction from  $N = 96$  to  $M = 10$ .

The algorithm developed in this section follows the approach reported in [5] as it was successfully adopted in the physical exercise identification for photoplethysmography artifact reduction.

## 2.2 Bayesian Classification

Let us refer to a frame  $k_{tM}[n]$ ,  $n = 0, \dots, M - 1$ , containing features extracted from the accelerometer signals.

We assume that the observations for all human activities that need to be identified, are acquired and divided in two sets,  $\mathcal{W}$  for training and  $\mathcal{Z}$  for testing.

For Bayesian classification, a group of  $\Gamma$  activities is represented by the probability density functions (pdfs)  $p_\gamma(k_{tM}) = p(k_{tM} | \theta_\gamma)$ ,  $\gamma = 1, 2, \dots, \Gamma$ , where  $\theta_\gamma$  are the parameters to be estimated during training. Thus we can define the vector  $p = [p_1(k_{tM}), \dots, p_\Gamma(k_{tM})]^T$ .

The objective of classification is to find the model  $\theta_\gamma$  corresponding to the activity  $\gamma$  which has the maximum a posteriori probability for a given frame  $k_{tM} \in \mathcal{Z}$ . Formally:

$$\hat{\gamma}(k_{tM}) = \operatorname{argmax}_{1 \leq \gamma \leq \Gamma} \{p(\theta_\gamma | k_{tM})\} = \operatorname{argmax}_{1 \leq \gamma \leq \Gamma} \left\{ \frac{p(k_{tM} | \theta_\gamma)p(\theta_\gamma)}{p(k_{tM})} \right\}. \quad (4)$$

Assuming equally likely activities (i.e.  $p(\theta_\gamma) = 1/\Gamma$ ) and noting that  $p(k_{tM})$  is the same for all activity models, the Bayesian classification is equivalent to

$$\hat{\gamma}(k_{tM}) = \operatorname{argmax}_{1 \leq \gamma \leq \Gamma} \{p_\gamma(k_{tM})\}. \quad (5)$$

Thus Bayesian identification reduces to solving the problem stated by (5).

The most generic statistical model one can adopt for  $p(k_{iM} | \theta_\gamma)$  is the Gaussian mixture model (GMM) [14]. The GMM for the single exercise is a weighted sum of  $F$  components densities and given by the equation

$$p(k_{iM} | \theta_\gamma) = \sum_{i=1}^F \alpha_i \mathcal{N}(k_{iM} | \mu_i, C_i) \quad (6)$$

where  $\alpha_i$ ,  $i = 1, \dots, F$  are the mixing weights, and  $\mathcal{N}(k_{iM} | \mu_i, C_i)$  represents the density of a Gaussian distribution with mean  $\mu_i$  and covariance matrix  $C_i$ . It is worth noting that  $\alpha_i$  must satisfy  $0 \leq \alpha_i \leq 1$  and  $\sum_{i=1}^F \alpha_i = 1$  and  $\theta_\gamma$  is the set of parameters needed to specify the Gaussian mixture, defined as  $\theta_\gamma = \{\alpha_1, \mu_1, C_1, \dots, \alpha_F, \mu_F, C_F\}$ .

The choice for obtaining an estimate of the mixture parameters is an unsupervised algorithm for learning a finite mixture model from multivariate data, that overcomes the main lacks of the standard expectation maximization (EM) algorithm, i.e. sensitiveness to initialization and selection of number  $F$  of components [9]. This algorithm 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

We used the ‘‘Human Activity Recognition Using Smartphones’’ data set [2]. This data set includes data recorded from experiments made by a group of 30 volunteers, each of which performed six different activities, three belonging to the ‘‘static’’ class, i.e., standing, sitting, and lying, and three belonging to the ‘‘cyclic’’ class, i.e., walking, climbing stairs, descending stairs. Data recorded during the transitions occurring between static postures were labeled accordingly as ‘‘transitions’’. 3-axial linear acceleration was recorded at a 50 Hz sampling rate and the experiments were video-recorded to allow accurate manual data labeling.

The signals so gathered were split in 2.56 s long windows with 50 % overlap. Windows containing unlabeled portions of signal had been discarded, as were windows containing more than 25 % of signal with inconsistent labeling (i.e., a label that differs from that of the majority of the data points within the window). This yielded a total of 10991 windows to process, 7808 of which were used for training and the remaining 3183 for testing. Separation was done so that data recorded by any given person never occurred both in the training and testing sets.

Data was pre-processed by using these windows to build three  $N \times L$  Hankel matrices, one for each acceleration direction, with  $N = 96$  and  $L = 33$ . These are then fed together to the SVD, so as to remove the effect of sensor orientation, and all the ensuing normalized singular values used, after dimensionality reduction to  $M = 10$  principal components, as the feature vectors for the classifier.

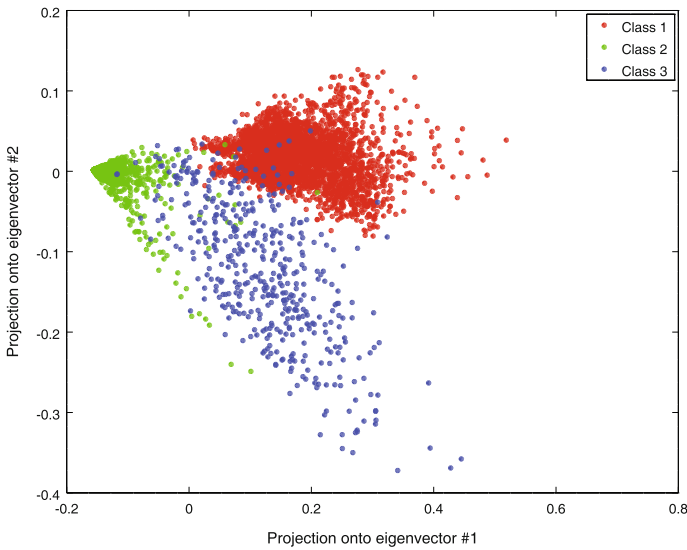
**Table 1** Confusion matrix of the exercise type identifier evaluated on the whole testing set

	Standing	Sitting	Lying	Walking	W. down- stairs	W. up- stairs	Stand- to-sit	Sit-to- stand	Sit-to-lie	Lie-to-sit	Stand- to-lie	Lie-to- stand
Standing	420	27	26	0	0	0	1	0	0	0	1	0
Sitting	98	273	64	0	0	0	1	0	2	0	12	0
Lying	66	63	266	0	0	0	0	0	0	0	0	0
Walking	0	0	0	292	125	102	2	3	0	0	0	0
W. downstairs	0	0	0	228	178	151	5	1	0	0	0	0
W. upstairs	0	0	0	269	128	158	1	3	0	1	0	0
Stand-to-sit	0	0	0	2	2	0	9	6	4	3	5	3
Sit-to-stand	0	0	0	0	2	0	5	4	0	1	3	1
Sit-to-lie	0	0	0	0	0	0	5	4	12	5	11	4
Lie-to-sit	0	0	0	0	0	0	2	3	1	10	7	12
Stand-to-lie	0	0	0	2	0	0	6	4	3	6	28	10
Lie-to-stand	0	1	0	0	0	0	2	2	3	6	7	10

As a first experiment, the performance of the classifier in recognizing the exact activity being performed was assessed. Of course, this method was never intended to be able to discriminate between all these activities, as by mixing the axes of the accelerometer output is clearly high impossible to discriminate, e.g., between the static postures. This is clearly shown if Table 1, where the confusion matrix of this classification experiment is reported. The resulting performance is reported in Table 2, yielding an overall accuracy of 52.15 % with an F1-score of 43.22 %.

**Table 2** Performance (sensitivity, precision, and F1-score) of the exercise type identifier evaluated on the whole testing set

Activities	Sensitivity (%)	Precision (%)	F1-score (%)
Standing	88.42	71.92	79.32
Sitting	60.67	75.00	67.08
Lying	67.34	74.72	70.84
Walking	55.73	36.82	44.34
Walking downstairs	31.62	40.92	35.67
Walking upstairs	28.21	38.44	32.54
Stand-to-sit	26.47	23.08	24.66
Sit-to-stand	25.00	13.33	17.39
Sit-to-lie	29.27	48.00	36.36
Lie-to-sit	28.57	31.25	29.85
Stand-to-lie	47.46	37.84	42.11
Lie-to-stand	32.26	25.00	28.17



**Fig. 2** Eigenvector projections



**Table 3** Confusion matrix of the human activity class identifier evaluated on the whole testing set

	Static	Cyclic	Transitional
Static	1303	0	17
Cyclic	0	1629	18
Transitional	1	7	208

**Table 4** Performance (sensitivity, precision, and F1-score) of the activity class identifier evaluated on the whole testing set

Activities	Sensitivity (%)	Precision (%)	F1-score [%]
Static	98.71	99.92	99.31
Cyclic	98.91	99.57	99.24
Transitional	96.30	85.60	90.63

Despite these results, it is quite clear from observation of Table 1 that the recognizer very seldom makes mistakes between the three classes of exercises (static postures, cyclic movements, transitions). A scatter plot of the first two features, shown in Fig. 2, also confirms this idea, as in just two dimensions the three classes are very well separated.

This is better assessed by the second experiment, where only these three classes were considered for training the models. The confusion matrix is shown in Table 3, and the performance in Table 4. The overall accuracy in this context is 98.65 % and the F1-score is 96.48 %, denoting the high reliability of the chosen features to discriminate between these classes. It should also be recalled that a number of windows, 5.8 % to be exact, contain up to 25 % of “noise”, i.e., data belonging to different classes, so some classification error is to be expected.

## 4 Conclusion

In this paper we present a feature set, based on the dimensionality-reduced singular value spectrum of the data Hankel matrix, that is suitable to identify human activity classes from accelerometer data. Since the singular value spectrum is inherently invariant with respect to rotation matrices, classification accuracy is independent of the orientation of the sensor, freeing the user from having to worry about its correct placement.

Experimental results conducted on a public domain data set show the effectiveness of the proposed approach.

## 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. Anguita, D., Ghio, A., Oneto, L., Parra, X., Reyes-Ortiz, J.L.: A public domain dataset for human activity recognition using smartphones. In: *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2013)*, pp. 437–442 (2013)
3. 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 (EUSIPCO 2015)*, pp. 2696–2700 (2015)
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.: A rule based framework for smart training using sEMG signal. In: *Intelligent Decision Technologies: Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015)*, pp. 89–99. Sorrento, Italy (2015)
6. 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 (ICPRAM 2016)*, pp. 467–474. Rome, Italy (2016)
7. Catal, C., Tufekci, S., Pirmitt, E., Kocabag, G.: On the use of ensemble of classifiers for accelerometer-based activity recognition. *Appl. Soft Comput.* **37**, 1018–1022 (2015)
8. 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 (IE)*, pp. 214–221 (2012)
9. Figueiredo, M.A.F., Jain, A.K.: Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(3), 381–396 (2002)
10. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: a review. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(1), 4–37 (2000)
11. 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 (FutureTech)*, pp. 1–6 (2010)
12. Mannini, A., Sabatini, A.M.: Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* **10**(2), 1154–1175 (2010)
13. 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)
14. 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)
15. Rodríguez-Martin, D., Samà, A., Perez-Lopez, C., Català, A., Cabestany, J., Rodríguez-Molinero, A.: SVM-based posture identification with a single waist-located triaxial accelerometer. *Expert Syst. Appl.* **40**(18), 7203–7211 (2013)
16. 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 (LASCAS)*, pp. 1–4 (2015)