# HUMAN MOTION DETECTION IN DAILY ACTIVITY TASKS USING WEARABLE SENSORS

*Olga Politi[1], Iosif Mporas[2], Vasileios Megalooikonomou[1]*
Multidimensional Data and Knowledge Discovery Laboratory
Dept. of Computer Engineering and Informatics, University of Patras
26500 Rion-Patras, Greece
[1]{politi, vasilis}@ceid.upatras.gr, [2]imporas@upatras.gr

## ABSTRACT

In this article we present a human motion detection framework, based on data derived from a single tri-axial accelerometer. The frameworkuses a set of different preprocessing methods that produce data representations which are respectively parameterized by statistical and physical features. These features are then concatenated and classified using well-known classification algorithms for the problem of motion recognition. Experimental evaluation was carried out according to a subject-dependent scenario, meaning that the classification is performed for each subject separately using their own data and the average accuracy for all individuals is computed. The best achieved detection performance for 14 everyday human motion activities, using the USC-HAD database, was approximately 95%. The results compare favorably are competitive to the best reported performance of 93.1% for the same database.

***Index Terms—*** Accelerometers, wearable sensors, movement classification, human motion recognition, daily activity.

## 1. INTRODUCTION

One of the most important tasks in pervasive computing is to provide accurate and opportune information on people's activities and behaviors. Applications in medicine, security, entertainment and tactical scenarios are examples of this effort. For instance, patients with obesity, diabetes or heart disease, are often required to fulfil a program of activity which follows a training schedule that is integrated within their daily activities [1]. Therefore, the detection of activities such as walking or running becomes quite useful to provide valuable information to the caregiver about the patient's behavior. Under conditions of daily living, human-activity recognition could be performed using objective and reliable techniques. In computer vision, complex sensors such as cameras have been used to recognize human activities. In general, tracking and activity recognition using computer vision-based techniques perform quite well in a laboratory or well-controlled environments. However, their accuracy falls under a real-home setting, due to the high-level activities that take place in the natural environments, as well as the variable lighting or clutter [2]. As a result, body-attached accelerometers are commonly used as an alternative in order to assess variable daily living activities.

The human motion detection problem using accelerometers is an emerging area of research. Sensors embedded in objects or attached on the body, are generally chosen to study movement patterns or human behavior. Accelerometers have been used widely, due to their low-power requirements, small size, non-intrusiveness and ability to provide data regarding human motion. In an ideal scenario, this data can be processed using signal processing and pattern recognition methods, in order to obtain a real-time recognition of human motion.

Several human-activity recognition systems have been proposed in the past, which include the use of accelerometers. Some of them analyze and classify different kinds of activity using acceleration signals [2], [3], while others apply them for recognizing a wide set of daily physical activities [4], or describe a human activity recognition framework based on feature selection techniques [5]. Bernecker et al [7], proposed a reclassification step that increases accuracy of motion recognition. Karantonis et al. [8] introduced an on-board processing technique for a real-time classification system, yielding results that demonstrate the feasibility of implementing an accelerometer-based, real-time movement classifier using embedded intelligence. Khan et al. [2] propose a system that uses a hierarchical-recognition scheme, i.e., the state recognition at the lower level using statistical features and the activity recognition at the upper level using the augmented feature vector followed by linear discriminant analysis. Several powerful machine learning algorithms have been proposed in the literature for the detection of human motion. The most widely used are the artificial neural networks [6, 8, 10], the naïve-Bayes [4] and the support vector machines [5].

In this paper, a human motion sensor-based detection framework is proposed. After evaluating different preprocessing methods [4, 8], it combines both statistical and

The reported research was partially supported by the ARMOR Project (FP7-ICT-2011-5.1 - 287720) "Advanced multi-paRametric Monitoring and analysis for diagnosis and Optimal management of epilepsy and Related brain disorders", co-funded by the European Commission under the Seventh' Framework Programme.

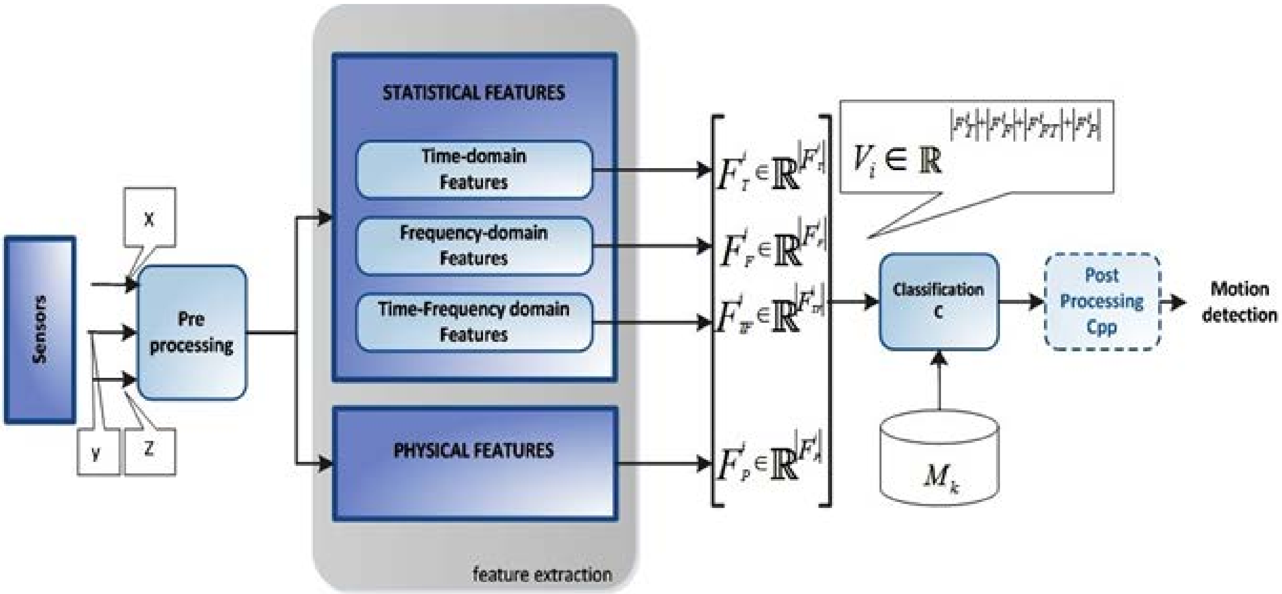


**Fig. 1.** Block diagram of the human motion detection framework.

physical features as indicated in[6]. The parameterized signals are processed by a classification model and decision is taken for input data. In relation to the other state-of-the-art methods, this method combines the best known preprocessing, feature extraction and classification techniques, in order to estimate which features are optimal for motion recognition and which combination between feature extraction and classification yields the highest motion detection results, since there is little known research that investigates this issue.The rest of this paper is organized as follows: In Section 2 we present the framework used for human motion detector. Section 3 offers details about the evaluation data and methodology followed and in Section 4 we present the achieved experimental results. Finally in Section 4 we conclude this work.

## 2. FRAMEWORK FOR MOTION DETECTION

In the present framework for motion detection, we assume as input tri-axial accelerometer data collected from sensors. The collection and transmission of the data to the detector is not part of this study. The block diagram of the motion detection framework is illustrated in Fig. 1. As shown in Fig. 1, the input to detection framework consists of 3-dimensional $(\mathrm{x}, \mathrm{y}, \mathrm{z})$ signal streams.

Each stream represents one movement direction in the sense of moving forward/backward, up/down and left/right. The correspondence between axes and directions depends on the sensor placement, which is manually configured at the beginning of each recording session. Preprocessing consists of applying a sliding window $W$ to the incoming streams, of constant length, resulting to $W_i$ frames, where $1 \le i \le I$ . The time shift between two successive frames is also constant and can result to overlapping or non-overlapping frame sequences.

After preprocessing the sensor data, each frame is processed by statistical and physical feature extraction algorithms. The statistical algorithms can briefly be divided totime, frequency and time-frequency domain methods. Physical algogrithms are derived based on physical interpretations ofhuman motion.The utilization of these methods ensures that we get as much possible information from the data retrieved. In detail, each incoming frame $W_i$ is processed in parallel by each of the feature extraction modules shown in Fig. 1. The estimated feature vectors, i.e. the time-domain features $F_T^i \in \mathbb{R}^{\left|F_T^i\right|}$, the frequency-domain features $F_F^i \in \mathbb{R}^{\left|F_F^i\right|}$, the time-frequency domain features $F_{TF}^i \in \mathbb{R}^{\left|F_{TF}^i\right|}$ and the physical features $F_P^i \in \mathbb{R}^{\left|F_P^i\right|}$ are concatenated to a single feature vector $V_i \in \mathbb{R}^{\left|F_T^i\right|+\left|F_F^i\right|+\left|F_{FT}^i\right|+\left|F_P^i\right|}$, with $1 \le i \le I$ .

After the decomposition of the sensor data to feature vectors $V_i$, one for each frame, the sequence of feature vectors is processed by a classification algorithm $C$ . Before the classification step, feature evaluation could be performed using feature ranking algorithms, in order to reduce the dimensionality of the feature space. To achieve the best classification performance, the dimensionality of the feature vector should be as small as possible, with regard to the most prominentand complementary features.

During the training phase a set of motion data (training data) with known labels, i.e. with a-priori annotated motion labels, is used to estimate one model, $M_k$, with $1 \le k \le K$, for each human motion $k$ of interest. At the test phase the unknown motion data (test data) will be pre-processed and decomposed to feature vectors as in the training phase. The classification algorithm $C$ will compare each test vector $V_j$, with $1 \le j \le J$, against each motion model $M_k$ in order to decide the corresponding motion class, i.e. $d_j = \arg\max_k \left\{ C\left(V_j, M_k\right) \right\}$, where $d_j$ is the motion classlabel assigned to the $j$-th test frame of the sensor data.

After classification a post-processing algorithm can be applied on the automatically labeled frames of the test recording in order to fine-tune the detected human motion classes.

## 3. METHODOLOGY

The framework for motion detection described in the previous section was evaluated using the USC-HAD database [11], which is a freely available dataset provided by the University of Southern California. The dataset corresponds to well-defined low-level daily activities appropriate for evaluation of algorithms for human motion in healthcare scenarios. The database consists of data from 14 subjects, 7 male and 7 female, taken from a sensor placed at their front right hip, using sampling frequency equal to 100 Hz. MotionNode, the sensing platform used, integrates a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer. For the purpose of this paper, only the 3-axis accelerometer data were used. Subjects performed 14 different types of activities, including walking, running, sitting, standing, ascending and descending stairs. Ground truth was annotated by an observer while the experiments were being carried out.

### 3.1. Preprocessing

During pre-processing, each of the 3-dimensional $(\mathrm{x}, \mathrm{y}, \mathrm{z})$ signal streams was frame blocked with a sliding window of 1 second length, with time-shift 0.5 seconds [8]. Except the frame blocking, three preprocessing methods were used, namely the (i) GA, (ii) BA and (iii) Tilt [8], thus resulting to four preprocessing setups together with the baseline preprocessing (i.e. purely frame blocking of the data samples). In detail, method GA extracts the gravitational acceleration component from the data signal. Method BA extracts the body acceleration component from the data signal and Tilt method measures the change in angular velocity.

### 3.2. Feature Extraction

For each of the four preprocessed outputs per frame, statistical and physical features were extracted [6]. As considers, the statistical features they are briefly separated to time-domain, frequency domain and time-frequency domain [12]. Time-domain features are raw data values and include mean, median, variance, root mean square (rms), standard deviation (std), skewness, kurtosis and interquartile range (25%, 50%,75%). Frequency-domain features, which mainly represent the periodic structure of the signal are the Fourier transform, the spectral entropy, the spectral energy and the 3rd order autoregressive-filter (AR) coefficients. Time-frequency domain features are used to investigate both time and frequency characteristics of complex signals and in general employ wavelet techniques, such as wavelet coefficients or energy of wavelet coefficients.

In addition, physical features [6], include the movement intensity (MI), the eigenvalues of dominant directions (EVA), the averaged velocity along gravity direction (AVG), the correlation of averageacceleration along gravity and heading directions (CAGH), the dominant frequency (DF) and the averaged acceleration energy (AAE). For each one of the four pre-processing methods the estimated feature vector has dimensionality equal to 77. Thus, the final feature vector per frame is of dimensionality 4×77=308.

### 3.3 Feature Evaluation

After computing the motion features described in subsection 3.2, we estimate the importance of each motion feature, with respect to their discriminative ability on the motion type recognition task. For the evaluation of the importance of the motion features we relied on the ReliefF algorithm [17]. The ReliefF algorithm computes a vector $R$ of the estimations of the qualities of all the motion features. The ranking position of each feature is defined by its ranking score, i.e. the corresponding estimation of quality, $w \in \mathbb{R}$, which indicates the degree of importance of that feature.

### 3.4 Classification

For the classification of the estimated feature vectors $V$ we relied on four well known and widely used classification algorithms, namely the support vector machines (SVMs) implemented with the sequential minimal optimization method [13] using the polynomial kernel function (poly), a two-layered backpropagation multilayer perceptron (MLP) neural network [14], the k-nearest neighbor (IBk) algorithm [15]and the C4.5 (J48) decision tree [16]. All classifiers were implemented using the WEKA machine learning toolkit software [12].

## 4. EXPERIMENTAL RESULTS

The human motion detection framework presented in Section 2 was evaluated according to the methodology described in Section 3. To avoid overlap between the training and test datasets we followed a 10-fold cross validation protocol. The experimental results, in percentages, for the four evaluated classification algorithms and for each of the14 subjects of the dataset are presented in Table 1. The best performing algorithm for each subject is indicated in bold.

As can be seen in Table 1, the two discriminative classification algorithms, i.e. the SVM and MLP, achieved the best human motion recognition accuracy for all evaluated subjects. Specifically, the SVM algorithm using the polynomial kernel function significantly outperformed all other models. The superiority of the kernel-based algorithm is probably owed to the curse of dimensionality phenomenon, from which SVMs do not suffer [18]. The detection performance varies from 89.73% (for subject 05) to 94.73% (for subject 11). This range of 5% is owed to the different amount of available data for each subject (2732 instances for subject 05 vs. 5196 instances for subject 11). The evaluation results showed that the five most confused human motions across all subjects are elevator up, elevator down, walking left, walking right and walking downstairs. Finally, the less difficult to detect human motion was found to be sleeping. Slightly lower performance has been reported in [6], however no direct comparison can be done, due to the different evaluation dataset used in the two studies.

| Sub. | SVM | MLP | IBk | J48 |
|---|---|---|---|---|
| 01 | **94.25** | 93 .52 | 90 .87 | 88 .99 |
| 02 | **93.34** | 92 .01 | 90 .65 | 89 .61 |
| 03 | **93.29** | 91 .05 | 91 .42 | 91 .87 |
| 04 | **92.76** | 86 .71 | 85 .87 | 86 .16 |
| 05 | **89.73** | 87 .12 | 84 .66 | 86 .68 |
| 06 | **94.28** | 92 .74 | 92 .12 | 92 .06 |
| 07 | **93.87** | 93 .26 | 91 .93 | 91 .63 |
| 08 | **94.13** | 91 .87 | 88 .51 | 90 .42 |
| 09 | **93.57** | 92 .31 | 86 .78 | 89 .86 |
| 10 | **94.23** | 90 .39 | 89 .05 | 89 .33 |
| 11 | **94.73** | 93 .92 | 93 .96 | 92 .51 |
| 12 | **94.19** | 93 .40 | 91 .54 | 94 .04 |
| 13 | **94.12** | 91 .87 | 91 .51 | 90 .60 |
| 14 | **92.87** | 85 .42 | 82 .42 | 82 .21 |

**Table 1.** Motion detection accuracy (%) per subject and classification algorithm

The average achieved results for the task of human motion detection, in terms of percentages of accuracy, for the full motion feature vector are shown in Table 2.It is evident that the SVM classification algorithm outperformed all the other algorithms achieving 93.5% accuracy, which is competitive to the 89.3% accuracy reported in [6], when the single-layerclassifier is used.The best performance, approximately 95%, is achieved for subject 11, which is comparative to the performance of 93.1% for the same database in [6].The best performing SVM was followed by the Random Forest algorithm, which achieved approximately 1% lower performance. Both the decision tree (J48) and the neural network (MLP) achieved significantly lower performance. The IBk (k- Nearest Neighbors) algorithm yielded the worse accuracy, approximately 89%.

| Algorithm | Accuracy (%) |
|---|---|
| MLP | 90.37 |
| SVM | 93.52 |
| IBk | 88.72 |
| J48 | 89.33 |
| Random Forest | 92.25 |

**Table 2.**Average accuracy forhuman motion detection per classification algorithm

The advantage of the SVM algorithm is probably owed to the high dimensionality of the feature space ($V \in \mathbb{R}^{308}$) in combination with the amount of evaluated data, since SVMs do not suffer from the curse of dimensionality [18]. Moreover, in contrast to the rest evaluated algorithms SVM training will converge to the global optimal parameter values, and thus for a specific dataset will not provide suboptimal performance.

In a further step we present the top ranked features, as they were chosen by the ReliefF algorithm [17]. In Table 3 we present the 10 most discriminative features.

| Ranking | Features per Subject |
|---|---|
| 1 | Spectral entropy |
| 2 | CAGH of x axis (gravity) |
| 3 | Correlationxz |
| 4 | Median |
| 5 | Mean |
| 6 | AVG |
| 7 | Wavelet coefficients of x axis |
| 8 | DF |
| 9 | Rms |
| 10 | Std |

**Table 3.** Top-10 motion features according to the ReliefF criterion.

As can be seen in Table 3, within the most discriminative motion features for the task of human motion recognition are spectral entropy, the correlation between acceleration along gravity direction (CAGH), the correlation between gravity and heading direction, median, mean, average acceleration for gravity direction, wavelet coefficients for gravity direction (x axis), dominant frequency, root mean square and standard deviation.

Spectral entropy helps to differentiate between signals that have similar energy values but correspond to different activity patterns, such as walking in different directions and running. CAGH and AVG are selected, since they contribute to the discrimination among motions that have different velocities and intensities along the heading direction, like walking, running, and jumping.Correlation is especially useful in differentiatingbetween activities that involve translation in a single dimension, i.e. jumping and walking upstairs/downstairs. Wavelet coefficients can capture sudden changes in signals produced by motions like jumping or running.

Statistical features like mean or standard deviation achieve also a high score in the ranking, since they generally give an indication of the stability of the signal. These results are in agreement with [6], where CAGH, AVG, DF, mean, median and std were also found within the most discriminative features for motion classification tasks. At this point, the significance of the addition of physical features is evident, since three physical features are highly rated by the ReliefF algorithm. Frequency domain features also play an important role, as we meet spectral entropy and correlation at the top three ranked features.

## 5. CONCLUSION

In the present work we present a human motion detection framework based on 3-dimensional $(x, y, z)$ sensor-based data. The framework uses four different preprocessing methods and the motion signals are parameterized by statistical and physical features. The experimental results indicated an average performance of approximately 93.5% for 14 everyday human motion activities, with the best performance achieved being approximately 95%. The application of the feature selection algorithm ReliefF, validated the hypothesis that the physical features added in the experimental process allowed the improvement of the motion detection accuracy. A combination between the features used along with the SVM classification was proven to yield the best results for the current study. We deem the future addition of subject-independent scenario in the experiments, to examine which approach is best for the human motion detection problem.

## REFERENCES


[1] Y. Jia, “Diatetic and exercise therapy against diabetes mellitus”, in *Second International Conference on Intelligent Networks and Intelligent Systems*, 2009, pp. 693–696.

[2] A. Khan, Y. Lee, S.Y. Lee, and T. Kim, “Triaxial Accelerometer-Based Physical-Activity Recognition via Augmented-Signal Features and a Hierarchical Recognizer”, in *IEEE transactions on information technology in biomedicine*, vol. 14, no. 5 ,2010.

[3] J. Mantyjarvi, J. Himberg, and T. Seppanen, “Recognizing human motion with multiple acceleration sensors”, in *IEEE Int. Conf. Syst*., Man, Cybern., vol. 2, 2001, pp. 747–752.

[4] M. Sekine, T. Tamura, M. Akay, T. Fujimoto, T. Togawa, and Y. Fukui, “Discrimination of walking patterns using wavelet-based fractal analysis”, in *IEEE Trans. Neural Syst. Rehabil. Eng*., vol. 10, no. 3, Sep. 2002 ,pp. 188–196.

[5] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, “Frequent Bit Pattern Mining Over Tri-axial Accelerometer Data Streams for Recognizing Human Activities and Detecting Fall”, in *Procedia Computer Science*, Volume 19, 2013, pp. 56–63.

[6] M. Zhang, and A. Sawchuk, “A Feature Selection-Based Framework for Human Activity Recognition Using Wearable Multimodal Sensors”, in *BodyNets 2011*,University of Southern California, Los Angeles, 2011.

[7] T. Bernecker, F. Graf, H. Kriegel, and C. Moennig, “Activity Recognition on 3D Accelerometer Data”, Technical Report, 2012.

[8] D.M. Karantonis, M.R Narayanan, M. Mathie, N.H. Lovell, and B.G. Celler, “Implementation of a Real-Time Human Movement Classifier Using a Triaxial Accelerometer for Ambulatory Monitoring” in *IEEE on information technology in biomedicine*, vol. 10,2006.

[9] A.M. Khan, Y.K. Lee, and S.Y. Lee, “Accelerometer’s Position Free Human Activity Recognition Using A Hierarchical Recognition Model”, in *IEEE HealthCom*, 2010.

[10] N. Ravi, N. Dandekar, P. Mysore , and M.L. Littman, “Activity Recognition from Accelerometer Data”, Department of Computer Science, Rutgers University, 2005.

[11] M. Zhang, and A. Sawchuk, “USC-HAD: A Daily Activity Dataset for Ubiquitous Activity Recognition Using Wearable Sensors” in *UbiComp ‘12*, USA, 2012.

[12] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, “Activity Recognition Using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey”, University of Twente, The Netherlands, 2010.

[13] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, and K.R.K Murthy, “Improvements to Platt's SMO Algorithm for SVM Classifier Design. Neural Computation”, 2001, 13(3):637-649.

[14] T.M. Mitchell, *Machine Learning*, McGraw-Hill International Editions, 1997.

[15] D. Aha, and D. Kibler, “Instance-based learning algorithms”, *Machine Learning*, vol. 6, 1991, pp. 37-66.

[16] R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA, 1993.

[17] M. Robnik-Sikonja, and I. Kononenko, “An adaptation of Relief for attribute estimation in regression”, in *Proc. of the 14th international conference on machine learning*, ICML ’97, pp. 296–304.

[18] Y. Ohtaki, H. Inooka, K. Sagawa, A. Suzuki, Z. Xiumin, M. Okutsu, and R. Nagatomi, “Recognition of daily ambulatory movements utilizing accelerometer and barometer”, in*Proceedings of the second IASTED International Conference*, 2004.