Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

Article

Approach for the development of a Framework for the Identification of Activities of Daily Living using Mobile Devices' Sensors

Ivan Miguel Pires[1,2,3,*], Nuno M. Garcia[1,3,4], Nuno Pombo[1,3,4], Francisco Flórez-Revuelta[5] and Susanna Spinsante[6]

[1] Instituto de Telecomunicações, Universidade da Beira Interior, Covilhã, Portugal; impires@it.ubi.pt, ngarcia@di.ubi.pt, ngpombo@ubi.pt
[2] Altranportugal, Lisbon, Portugal; ivan.pires@it.ubi.pt
[3] ALLab - Assisted Living Computing and Telecommunications Laboratory, Computing Science Department, Universidade da Beira Interior, Covilhã, Portugal; impires@it.ubi.pt, ngarcia@di.ubi.pt, ngpombo@ubi.pt
[4] Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal
[5] Department of Computer Technology, Universidad de Alicante, Spain; francisco.florez@ua.es
[6] Department of Information Engineering, Marche Polytechnic University, Ancona, Italy; s.spinsante@univpm.it
* Correspondence: impires@it.ubi.pt; Tel.: +351-966-637-9785

Abstract: Sensors available on mobile devices allow the automatic identification of Activities of Daily Living (ADL). This paper describes an approach for the creation of a framework for the identification of ADL, taking in account several concepts, including data acquisition, data processing, data fusion, pattern recognition, and machine learning. These concepts can be mapped in a module of the framework, including the use and creation of several algorithms. For the development of a framework that works in several conditions, *e.g.* without Internet connection, these algorithms should take in account the hardware and software limitations of the mobile devices to run all main tasks locally. The main purpose of this paper is related to the presentation the sensors, algorithms, and architecture of the proposed approach.

Keywords: Activities of Daily Living (ADL); environment; sensors; mobile devices, framework; data acquisition; data processing; data fusion; pattern recognition; machine learning

1. Introduction

Sensors embedded in off-the-shelf mobile devices, *e.g.*, accelerometer, gyroscope, magnetometer, microphone, and Global Positioning System (GPS) receiver [1], are able to use in the creation of a method for the recognition of Activities of Daily Living (ADL) [2] and their environments. The recognition of ADL and their environments is part of research for the development of a personal digital life coach [3]. This research is related to the development of ambient assisted living (AAL) systems, and it is widely important for the monitoring of a plethora of situations in people with some disabilities, and elderly people.

Multi-sensor data fusion technologies may be implemented with mobile devices, because they have several sensors available in different categories, such as motion sensors, magnetic/mechanical sensors, acoustic sensors, and location sensors [4], improving the accuracy of the recognition of several types of activities, *e.g.*, walking, running, going downstairs, going upstairs, and standing, and environments, *e.g.*, bar, classroom, gym, library, kitchen, street, hall, watching TV, and bedroom. The recognition of physical activities may be performed with motion, and magnetic/mechanical sensors, and the environments may be recognized with acoustic sensors. The fuse of the data acquired from motion, magnetic/mechanical, and acoustic sensors increases the number of the ADL recognized,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

allowing the distinction between sleeping, and standing activities. In order to recognize the driving activity, the fuse of the data acquired from motion, magnetic/mechanical, acoustic, and location sensors allows the distinction between sleeping, standing, and driving activities.

The research about the identification of ADL and their environments using sensors has been studied during the last years, and several methods and frameworks [5-10] have been implemented using smartphones. However, this problem should be separated in a set of modules, such as data acquisition, data processing, data fusion, and artificial intelligence systems, and the frameworks developed in the previous studies commonly are only focused one some part of the problem. The *Acquisition Cost-Aware QUery Adaptation (ACQUA) framework* [11] has been adapted for data acquisition and data processing, but it does not include all steps needed for data processing. Related to data fusion methods previously developed, there are no structured frameworks implemented, but some methods have been developed related to the recognition of the ADL [12-14]. The frameworks and methods previously developed are no structured for the holistic approach of the identification of the ADL, and they mainly focused on the use of motion sensors for the identification of physical activities, but this paper presents a structure of a structured framework, using other types of sensors for the recognition of the environment.

In accordance with the previous work [4, 15, 16], the aim of this paper consists on the presentation of an approach for the creation of a framework for the identification of ADL and their environment mainly focused on the use of the data acquired from several sensors available in the off-the-shelf mobile devices. Around the concept of the fusion of the sensors' data, the selection of the sensors that are able to use for this purpose is the first step for the creation of the framework, defining a method for the acquisition of the data, and, consequently, processing of the data acquired. The processing of the data includes the data cleaning, data imputation, and extraction of the features. The final step for the development of the framework proposed consists on the use of the features extracted for the application of an artificial intelligence method, *i.e.*, the implementation of some types of Artificial Neural Networks (ANN) in order to choose the best method for the accurate recognition of the ADL and their environments.

The remaining sections of this paper are organized as follows: Section 2 presents the previous research studies performed in this topic, presenting a set of methods for each module. Section 3 presents a new approach for the development of a framework for the identification of ADL using the sensors available in off-the-shelf mobile devices, and the sensors and methods that will be used. Section 4 presents a discussion and conclusions about the new approach.

2. Related Work

Following the previous researches related to the new approach for the development of a framework for the identification of ADL and their environments, the research is organized the research about the sensors available on mobile devices separated by categories (subsection 2.1), the presentation of data acquisition methods (subsection 2.2), the definition of the methods for data processing (subsection 2.3), the presentation of data fusion methods (subsection 2.4), the presentation of artificial intelligence methods (subsection 2.5), and, finally, in the section 2.6, the methods to combine the possible the sensors' data with the users' agenda.

2.1. Sensors

Sensors are small components present in mobile device that allows the acquisition of data related to ADL and their environments [16]. These small components are able to capture different types of signals, such as electrical, mechanical, acoustic and others, in order to process and identify the ADL [1, 17].

The number and types of sensors available on mobile devices are different for each mobile platform, and, for the recognition of ADL, should be used the sensors available in a major part of mobile devices. In general, the sensors available in mobile devices are Magnetic/Mechanical sensors, Environmental sensors, Location sensors, Motion sensors, Imaging/Video sensors, Proximity sensors, Acoustic sensors, Optical sensors, and Force sensors [4].

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

Based on the classification presented in [4], the sensors available on Android devices are microphone, accelerometer, gravity, linear acceleration, gyroscope, rotation, magnetometer, pedometer, altimeter, humidity, ambient light, ambient, temperature, GPS, touch screen, microphone, and camera [18, 19]. In addiction of the platform dependent restrictions of the use of sensors, the hardware differences between devices can influence the existence of the sensors. Thus, the sensors available in a major part of mobile devices, presented in table 1, are the accelerometer, the gyroscope, the magnetometer, the GPS, the microphone, the touch screen, and the camera.

Table 1. List of sensors available in mobile devices.

Categories:	Sensors:
Motion sensors	Accelerometer Gyroscope
Magnetic/mechanical sensors	Magnetometer
Location sensors	GPS
Acoustic sensors	Microphone
Force sensors	Touch screen
Imaging/video sensors	Camera

2.2. Data Acquisition

Data acquisition consists in a process to receive the different types of data from the several sensors available in the off-the-shelf mobile devices, allowing the data collection in mobility. However, the data acquisition with sensors available in mobile devices has several challenges, because the environment is not controlled and the user is able to enable the data acquisition processing in several conditions, including the incorrect positioning of the mobile device, the uncontrolled data sampling rate, the possibility of the unavailability of all sensors used in the framework developed, and the environmental conditions [20].

In order to improve the data acquisition process, several frameworks have been developed, including *Acquisition Cost-Aware QUery Adaptation (ACQUA) framework* [11], *Orchestrator framework* [21], *ErdOS framework* [22], *LittleRock prototype* [23], *Jigsaw continuous sensing engine* [24], *SociableSense framework* [25], *Continuous Hand Gestures (CHG) technique* [26], and *Barbie-Q (BBQ) approach* [27].

The *ACQUA framework* allows the control of the order of the data acquisition, the correct segments of the data requested, the calibration of the data acquisition rates, the packet sizes and radio characteristics, the adaptation of the dynamic changes in query selective properties, and the support of multiple queries and heterogeneous time window semantics from all sensors available in off-the-shelf mobile devices, reducing the energy consumption of the real-time data acquisition [11].

The *Orchestrator framework* promotes the distributed execution of the data acquisition using several mobile devices, and all devices executes a part of the data processing, avoiding to reduce the requirements related to the processing power and energy consumption [21].

The same purpose of *Orchestrator framework* is achieved from *ErdOS framework* and *LittleRock prototype*, distributing the data acquisition and processing processes by all resources available in the devices used, and reducing the energy needed to process the data collected from all sensors [22, 23].

The *Jigsaw continuous sensing engine* implements a method to control the different sample rates, adapting the data acquisition and processing for the different capabilities of the sensors [24].

The *SociableSense framework* has a mechanism to adapt the different sample rated of all sensors used and it is a cloud-based framework, reducing the locally data processing, but restricting the use of the framework to the availability of the Internet connection [25].

The authors of [26] implemented a *CHG technique* for the data acquisition with Windows Phone-based smartphones and low processing capabilities, capturing accelerometer and gyroscope data, storing the sensory data in the smartphone memory.

The *BBQ framework* applies a multi-dimensional Gaussian probability density function from all sensors, inferring the order of the data acquisition with conditional probabilities [27].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

Data acquisition process implemented in mobile devices may be performed without the use of frameworks, improving only the data processing to the different resources capabilities. The authors of [28-31] implement the data acquisition process from accelerometer data based in Apple iPhone and Android-based smartphones for the identification of several activities, including driving, walking, sitting, standing, running, and jumping activities. The authors of [32] implemented a Cursor Movement Algorithm to detect several activities, capturing the real-time data from the accelerometer and storing them into a local database in the mobile device.

Table 2 presents a summary of the data acquisition methods and their main characteristics for further implementation the in new approach for the development of a framework for the identification of ADL and their environments.

Table 2. Summary of the data acquisition methods.

Methods:	Advantages:
ACQUA framework	• Controls of the order of the data acquisition; • Controls the correct segments of the data requested; • Controls the calibration of the data acquisition rates; • Controls the packet sizes and radio characteristics; • Controls the adaptation of the dynamic changes in query selective properties; • Controls the support of multiple queries and heterogeneous time window semantics; • Adapted for low processing, memory, and energy capabilities.
Orchestrator framework	• Distributed execution of the data acquisition using several mobile devices; • Adapted for low processing, memory, and energy capabilities.
ErdOS framework	• Distributed execution of the data acquisition using several mobile devices; • Adapted for low processing, memory, and energy capabilities.
LittleRock prototype	• Adapted for low processing, memory, and energy capabilities.
Jigsaw continuous sensing engine	• Controls the different sample rates; • Adapted for low processing, memory, and energy capabilities.
SociableSense framework	• Cloud-based framework; • Needs a constant Internet connection; • Adapted for low processing, memory, and energy capabilities.
CHG technique	• Stores the sensory data in the smartphone memory; • Adapted for low processing, and energy capabilities.
BBQ framework	• Uses a multi-dimensional Gaussian probability density function from all sensors; • Adapted for low processing, memory, and energy capabilities.
Cursor movement algorithm	• Stores the sensory data in the smartphone memory; • Adapted for low processing, and energy capabilities.
No framework	• Adapted for low processing, memory, and energy capabilities.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

2.3. Data Processing

After the data acquisition, the sensors' data should be processed in order to prepare the data for the fusion of all sensors data, and, consequently, application of the methods for the recognition of ADL. At the start, the data processing should validate the integrity and quality of the data, and, then, applying data cleaning and/or data imputation techniques [33]. However, data processing depends on the environmental conditions, the types of sensors and data, the sensor failures, and the capabilities of the mobile devices [34]. Several techniques have been developed to reduce the memory and energy consumption of the data processing techniques.

The *ACQUA framework* is also used to optimize the data processing, using automated storage and retrieval system (ASRS) algorithms [11]. Other studies have presented methods to adapt the data processing methods to the low capabilities of the mobile devices, processing the data after splitting or using methods with low resources needed [20, 35-37].

The use of data cleaning methods, presented in the subsection 2.3.1, is important to decrease the influence of the environmental conditions or systems failures. In order to improve the results, when the data acquisition fails, subsection 2.3.2 presents the possible data imputation methods to correct the data acquired. However, these methods are not subject for the development of the new approach for the framework for the identification of ADL and their environments, assuming that the data acquired is sufficient for the extraction of the several features from the sensors' signal, presenting the feature extraction methods and possible features to extract in the subsection 2.3.3.

2.3.1. Data Cleaning

Data cleaning consists in the identification of the incorrect values, removing the outliers values and smoothing and filtering the invalid values obtained during the data acquisition process, commonly considered as noised values [38-40]. Using data cleaning methods, the dependency of the environmental conditions, position of the mobile device, and systems failures occurred during the data acquisition process is reduces. The selection of these methods depends on the type of data acquired and spatiotemporal characteristics of the data acquired.

The authors of [41] proposed a weighted moving average (WMA) algorithm that collects the sensors' data and computes the weighted moving average and applies the weighted moving average filter for the normalization and cleaning of the sensors' data.

The filters used for the motion and magnetic/mechanical sensors may have two types, these are the low-pass filter, the high pass filter, and the KALMAN filter [42, 43]. On the other hand, for acoustic data, the moving average filter can also be used for filtering this type of data, but the different types of Fourier transforms, such as Discrete Fourier Transform (DFT), Inverse Discrete Fourier Transform (IDFT), and Fast Fourier Transform (FFT) are also used to filter the acoustic data [44, 45].

Table 3 presents a summary of the data cleaning methods related to the different types of sensors, presented in the section 2.1, for further implementation the in new approach for the development of a framework for the identification of ADL and their environments.

Table 3. Relation between the types of sensors and the data cleaning techniques allowed.

Types of Sensors:	Data Cleaning Techniques:
• Motion sensors; • Magnetic/mechanical sensors.	• Low-Pass Filter; • High-Pass Filter; • KALMAN Filter; • Weighted moving average (WMA) algorithm; • Moving average filter.
• Location sensors	• The data cleaning technique is not important for this type of data acquired.
• Acoustic sensors	• Moving average filter; • Discrete Fourier Transform (DFT); • Inverse Discrete Fourier Transform (IDFT);

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

Types of Sensors:	Data Cleaning Techniques:
	• Fast Fourier Transform (FFT).
• Force sensors • Imaging/video sensors	• The data cleaning technique is not important for this type of data acquired.

2.3.2. Data Imputation

During the data processing, the verification of the existence of the faulty data is performed to validate the presence of inexistent values in some instants of the time series of data acquisition. The data imputation methods are mainly used for motion sensors, and magnetic/mechanical sensors. However, for the development of the new approach of the framework for the identification of ADL and their environments, the data imputation techniques were not considered, assuming that the sensors' data acquired is complete. Thus, in this section, the best methods for data imputation will be presented based on a literature review.

Faulty data may have different types that can be classified as Missing Completely At Random (MCAR), Missing At Random (MAR) and Missing Not At Random (MNAR) [46]. When the faulty data is randomly distributed during the time interval for the data acquisition, the classification of this data is MCAR, that is possible faulty data found in the new approach of the framework for the identification of the ADL and their environments. The other types of faulty data are MAR, verified when the faulty data is randomly distributed by different subsets of the data acquired, and MNAR, verified when the faulty data is distributed by defined instants of the data acquisition.

The K-Nearest Neighbor (k-NN) method are one of the most used methods for data imputation of data acquired from motion, and magnetic/mechanical sensors [47-50]. The k-NN method has several variants that can be used for data imputation, such as MKNNimpute (K-nearest neighbor imputation method based on Mahalanobis distance), SKNNimpute (sequential K-nearest neighbor method-based imputation), and KNNimpute (K-nearest neighbor imputation) [47, 48].

The clustering techniques are also used for the data imputation, including K-means clustering, K-means-based imputation, and fuzzy C-means clustering imputation [46, 51, 52], which are implement in the Imputation Tree (ITree) method presented in [46].

There are other methods related to data imputation, including multiple imputation [53], hot/cold imputation [54], maximum likelihood [55], Bayesian estimation [55], expectation maximization [49, 56, 57], discarding instances [12], pairwise deletion [12], unconditional mean imputation [12], conditional mean imputation [12], hot deck imputation [12], cold deck imputation [12], substitution method [12], linear regression [12], logistic regression [12], expectation-maximization (EM) algorithm [12], probabilistic neural networks [12], fuzzy min–max neural networks [12], general regression auto associative neural network [12], tree-based methods [12], multi-matrices factorization model (MMF) [58], mean imputation (MEI) [49, 57], Multivariate Imputation by Chained Equations (MICE) [49, 57], Fourier method [57], and Fourier and lagged k-NN combined system (FLk-NN) [49, 57, 59].

In general, these methods can be applied to data collection from motion, and magnetic/mechanical sensors, but the data imputation methods can also be applied to the acoustic data with k-NN methods and singular value decomposition (SVD) algorithms [60].

As the data imputation methods should be able to impute the empty instances of the data acquired by motion, and magnetic/mechanical sensors, the methods that are able to use with this purpose are MEI, EM, MICE, and FLk-NN [49]. However, k-NN can be applied with the comparison between the history of the data acquisition that is similar to the data acquired in the stream with faulty values [49]. The data imputation is not important for the acoustic, and location sensors.

2.3.3. Feature Extraction

The correct definition of the features increases the accuracy of the identification of ADL and their environments. This definition depends on the types of sensors and the data acquired, but the purpose of use is another variable that should be taken in account.

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

For the correct extraction of the features for the motion and magnetic/mechanical sensors' data, the Euclidean norm for each instant of outputs from the sensors defined as magnitude of vector (MV). Thus, the features that should be extracted from the motion and magnetic/mechanical sensors are the mean for each axis [61-64], variance of MV [65, 66], mean of MV [62, 65-70], median of MV [65, 69], maximum of MV [61, 65, 66, 68], minimum of MV [61, 65, 66, 68], standard deviation of MV [61, 62, 65-70], Root Mean Square (RMS) of MV [61, 65], average of peak frequency (APF) of each axis [61], maximum of each axis [61, 64, 69], minimum of each axis [61, 64, 69], standard deviation of each axis [61, 63, 64], RMS of each axis [61], cross-axis signals correlation [61, 62, 64, 68, 71], Fast Fourier Transform (FFT) spectral energy [65, 71], frequency domain entropy [71], FFT coefficients [65, 68], logarithm of FFT [71], skewness of each axis [62], kurtosis of each axis [62], average absolute deviation of each axis [62], time between peaks [67], Interquartile range of MV [66, 68], skewness of MV [66], kurtosis of MV [66], wavelet energy of MV [68], average of peak values [72], average of peak rising time [72], average of peak fall time [72], average time per sample [72], average time between peaks [72], slope for each axis [69], binned distribution for each axis [63], percentiles of MV [70], and zero crossing rate for each axis [64].

Related to the motion and magnetic/mechanical sensors' data, the most used features are mean, standard deviation, maximum, minimum, median, correlation, variance, and FFT spectral energy of MV.

For the correct extraction of the features for the acoustic sensors' data, the features that should be extracted are average [73], thresholding [73], minimum [73], maximum [73], distance [73], and MFCC (Mel-frequency cepstrum coefficients) [74, 75].

For the location sensors, the feature that should be extracted is the distance travelled between a time interval, in order to identify ADL with high distance travelled. The distance between two points captured by GPS receiver is the ellipsoidal distance, because the two points are acquired in the geodetic coordinate system, where the calculation of this distance is measured with the Vincenty formula [76-78].

Table 4 presents a summary of the features extracted for each type of sensors presented in the section 2.1, for further implementation the in new approach for the development of a framework for the identification of ADL and their environments.

Table 4. Relation between sensors and features extracted.

Types of Sensors:	Features:
• Motion sensors; • Magnetic/mechanical sensors.	• Mean, average of peak frequency (APF), maximum, minimum, standard deviation, Root Mean Square (RMS), cross-axis signals correlation, skewness, kurtosis, average absolute deviation, slope, binned distribution, and zero crossing rate for each axis; • Mean, median, variance, maximum, minimum, standard deviation, Root Mean Square (RMS), Fast Fourier Transform (FFT) spectral energy, frequency domain entropy, FFT coefficients, logarithm of FFT, Interquartile range, skewness, kurtosis, wavelet energy, and percentiles of MV; • Time between peaks, average of peak values, average of peak rising time, average of peak fall time, average time between peaks.
• Location sensors	• Distance between two points.
• Acoustic sensors	• Average; • Thresholding; • Minimum; • Maximum;

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

Types of Sensors:	Features:
	• Distance; • MFCC (Mel-frequency cepstrum coefficients).
• Force sensors; • Imaging/video sensors.	• These sensors are not useful for the development of the framework for the Identification of ADL and their environments.

2.4. Data Fusion

After the extraction of the features, the data acquired from all sensors should be fused to improve the accuracy of the identification of the ADL and their environments in the new approach for the framework proposed in this study [5]. The data fusion methods implemented should be related with the final purpose of the framework presented in the subsection 2.6.

Based on the literature studies presented by several authors [6, 14, 79, 80], the data fusion methods are grouped in four categories [6, 79, 80], these are the probabilistic methods, the statistical methods, the knowledge base theory methods and the evidence reasoning methods.

The probabilistic methods [6, 14, 79, 80] include Bayesian analysis methods, maximum likelihood methods, state-space models, evidential reasoning, possibility theory, Kalman Filter [81, 82], Particle filtering, k-Nearest Neighbor (k-NN), k-Means, optimal theory, uncertainty ellipsoids, Gaussian mixture model (GMM), weighted averages, and regularization.

The statistical methods [6, 79, 80] for data fusion include covariance intersection, cross-covariance, and other robust statistics. However, other statistical methods used for data fusion are dynamic time warping (DTW) [83], which measures the similarity between two temporal sequences, based on the raw data or the features extracted.

The knowledge base theory methods [6, 14, 79, 80] for data fusion include Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, Deep Learning, Fuzzy Logic, Topic models, and Genetics Algorithms.

The evidence reasoning methods [6, 79, 80] for data fusion include evidence theory, Bayesian network, Dempster-Shafer, and recursive operators.

Based on these categories of data fusion methods, several implementations have been performed and presented in several studies for the identification of a plethora of a real-life activities and environments. The Rao-Blackwellization unscented Kalman filter (RBUKF) [84] was implemented to fuse the data acquired from a compass, a gyroscope, and a GPS receiver. The Kalman filter was used to fuse the data acquired from the GPS receiver and the gyroscope in order to support a navigation system [85]. The Naïve Bayes classifier is used to fuse the data acquired from acoustic, accelerometer and GPS sensors to recognize different situations during daily life [86]. The Autoregressive-Correlated Gaussian Model was implemented in the KNOWME system [87]. Bayesian analysis and Kalman filter where used to data acquired from the several sensors available in mobile devices for the identification of the ADL [88]. The CHRONIOUS system implements several methods to recognize several ADL, such as Support Vector Machine (SVM), random forests, Artificial Neural Networks (ANN), decision trees, decision tables, and Naïve Bayes classifier, in order to fuse the data collection from several sensors available in mobile devices [89]. In [90], the authors used the empirical mode decomposition (EMD) applied to the inertial sensors available in a mobile device, including accelerometer, gyroscope, and magnetometer, for the identification of several ADL. The authors of [91] implements several methods for data fusion, including SVM, random forest, hidden Markov models (HMMs), conditional random fields (CRFs), Fisher kernel learning (FKL), and ANN for several sensors, such as Accelerometer, RFID, and Vital monitoring sensors for the correct identification of ADL.

Table 5 presents a summary of the data fusion methods that can be applied for each type of sensors presented in the section 2.1, for further implementation the in new approach for the development of a framework for the identification of ADL and their environments.

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

Table 5. Relation between the different types of sensors and some data fusion methods.

Types of sensors:	Data fusion methods:
• Motion sensors; • Magnetic/mechanical sensors; • Location sensors; • Acoustic sensors.	• Autoregressive-Correlated Gaussian Model; • Fuzzy Logic; • Dempster-Shafer; • Evidence Theory; • Recursive Operators; • Support Vector Machine (SVM); • Random Forests; • Artificial Neural Networks (ANN); • Decision Trees; • Naïve Bayes classifier; • Bayesian analysis; • Kalman Filter; • K-Nearest Neighbor (k-NN); • Least squares-based estimation methods; • Optimal Theory; • Uncertainty Ellipsoids.
• Force sensors; • Imaging/video sensors.	• These sensors are not useful for the development of the framework for the Identification of ADL and their environments.

2.5. Identification of Activities of Daily Living

The definition of the methods identification of the ADL represents the final module of the new approach for the development of a framework for the identification of ADL and their environments, presented in the figure 1. The identification of the ADL and their environments depends on the sensors' data used. Thus, if the implemented method used the data acquired from motion and/or magnetic/mechanical sensors, commonly, it will identify the ADL. On the other hand, if the implemented method uses the data acquired from acoustic sensors, it will identify the environments. Finally, if the implemented method uses the location sensors, it probably is identifying activities will high movement, *e.g.*, driving, or it will try to identify the place where the ADL is performed.

In general, the identification of ADL is performed at the same time of the data fusion, because the methods used has the same techniques. The machine learning is a set of several techniques for artificial intelligence, including the techniques for the identification of ADL and their environments. The concept of machine learning will be presented in the subsection 2.5.1. In subsection 2.5.2, the pattern recognition methods are presented, which consists in a subset of the machine learning techniques.

2.5.1. Machine Learning

The Artificial Intelligence (AI) is one of the main areas for the development of computer science systems, and machine learning is composed by a subset of methods for AI, where the computers have the ability to learn and perform some tasks, taking in account the external conditions of the system in order to change the execution of some methods for obtaining of better results [92].

Machine learning methods are based on the creation and implementation of algorithms for the recognition and prediction of several situations based on the data acquired, and these methods are commonly classified in four categories [93, 94], such as Supervised learning, Unsupervised learning, Reinforcement learning, and Semi-supervised Learning and Active Learning.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

Supervised learning methods are based on the automatic adjustment of the network parameters, comparing the actual network output with the desired output previously defined in the data acquired, where the error obtained is the mean squared error (MSE) [94]. The input data involved in the supervised leaning should be labeled, in order to perform the comparisons.

Unsupervised learning methods consists on the correction of the results obtained based on the input data, attempting to obtain the signification patterns or features in the unlabeled input data, automatically learning with intuitive primitives like neural competition and cooperation [94].

Reinforcement learning methods are similar to supervised learning methods, but the exact desired output is *a priori* unknown [94]. Thus, these methods are learning based on the feedback provided during the execution of the algorithm by an artificial agent in order to maximize the total expected reward [94].

Semi-supervised Learning and Active Learning methods are methods that should be applied to dataset with a large collection of unlabeled input data and a few labeled examples to generalize the results and performance of the method, based on assumptions related to the probability of occurrence of some output.

For the development of a new approach for the development of a framework for the identification of ADL and their environments, the machine learning may be used, because it can be adapted to bioinformatics and human-related systems [95-98]. Pattern recognition methods, described in the subsection 2.5.2, consists on a subset of machine learning methods for the recognition of patterns [99], which are very useful in the development of the framework for the identification of ADL and their environments.

2.5.2. Pattern Recognition

The use of pattern recognition methods is the final part of research for the creation of a new approach for a framework for the identification of ADL and their environments. Several sensors, presented in the section 2.1, may be used with pattern recognition methods, which should be applied to the features extracted from the input data.

The methods implemented during the pattern recognition step are similar to the methods implemented for the data fusion, presented in the section 2.4. As reported early in this paper, the data fusion and pattern recognition may be confused, and the pattern recognition is performed at the same time of the data fusion. The categorization of the methods is similar to the methods applied for data fusion, and they are separated in four categories [6, 79, 80], these are the probabilistic methods, the statistical methods, the knowledge base theory methods and the evidence reasoning methods.

Several ADL may be recognized with pattern recognition methods, as example for the recognition of standing, and walking activities may be used ANN [100]. Several authors [7-10, 61-64, 66-71, 101-114] proposed the use of the ANN, the probabilistic neural networks (PNN), the deep neural networks (DNN), the SVM, the Random Forest, the Bayesian Network, the Sequential Minimal Optimization (SMO), the Logistic Regression, the Naïve Bayes, the C4.5 Decision Tree, the Logistic Model Trees (LMT), the J48 Decision tree, the K-Nearest Neighbor (KNN), and the Simple Logistic Logit Boost methods for the recognition of walking, running, jogging, jumping, dancing, driving, cycling, sitting, standing, lying, walking on stairs, going up on an escalator, laying down, walking on a ramp activities, cleaning, cooking, medication, sweeping, washing hands, and watering plants.

The Hidden Markov Model (HMM) and their variants are also a pattern recognition implemented in several studies related with the identification of ADL and their environments, such as the Hidden Markov Model (HMM) [66], the Hidden Markov Model Ensemble (HMME) [115], the Sliding-Window-based Hidden Markov Model (SW-HMM) [107]. The ADLs commonly identified by the HMM method are walking, walking on stairs, standing, running, sitting, and laying.

Table 6 presents a summary of the pattern recognition methods that can be applied for each type of sensors presented in the section 2.1, for further implementation the in new approach for the development of a framework for the identification of ADL and their environments.

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

Table 6. Relation between the different types of sensors and some pattern recognition methods.

Types of sensors:	Pattern recognition methods:
• Motion sensors; • Magnetic/mechanical sensors; • Location sensors; • Acoustic sensors.	• Support Vector Machines (SVM); • Decision trees (J48, C4.5); • Artificial Neural Networks (ANN); • Probabilistic Neural Networks (PNN); • Deep Neural Networks (DNN); • K-Nearest Neighbour (KNN); • Naïve Bayes; • Random Forest; • Logistic Regression; • Bayesian network; • Sequential minimal optimization (SMO); • Hidden Markov model (HMM); • Logistic Model Trees (LMT); • Simple Logistic Logit Boost.
• Force sensors; • Imaging/video sensors.	• These sensors are not useful for the development of the framework for the Identification of ADL and their environments.

2.6. Relation between the Identification of Activities of Daily Living and User Agenda

After the correct identification of the ADL and their environments, the results obtained should be compared with the users' agenda for the validation of the scheduled activities performed during the daily life. With this comparison and inputs from agenda, the monitoring of the lifestyle [116] for the development of a personal digital life coach [3] is one of the possibilities for the use of the new approach developed. However, the inputs from agenda can also be used to validate the accuracy of the framework developed. The correct identification of several patterns to validate the lifestyles and combine the results with the users' agenda is most important in the for monitoring elderly people [117].

3. Methods and Expected Results

The new approach proposed for the creation of the framework for the identification of ADL (figure 1) is based on the studies in [4, 15, 16], and it is composed with several stages. The stages for the framework are the selection of the sensors, the data acquisition, the data processing, including data cleaning, data imputation, and feature extraction, the data fusion, the identification of ADL with artificial intelligence, including pattern recognition, and other machine learning techniques, and, at the end, the combination of the results obtained with the data available in the users' agenda.

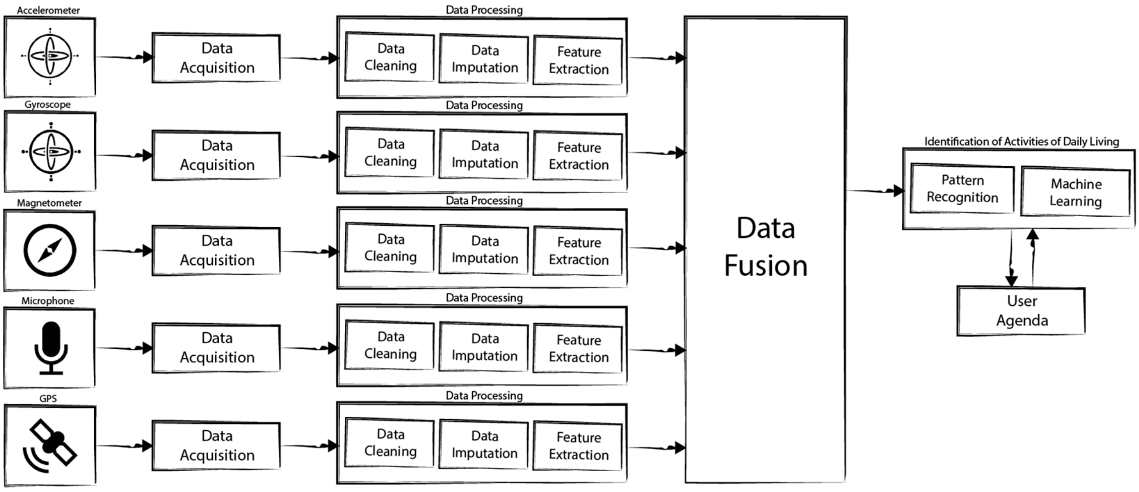

Figure 1. Schema for the framework for the recognition of Activities of Daily Living (ADL).

In order to create a new approach for a framework for the identification of ADL and their environments, the architecture, presented in the figure 1, and set of methods presented in the section 2 is proposed for obtaining results with good accuracy. The architecture presented in the figure 1 has several modules, these are the activation of the selected sensors, the data acquisition, the data processing, including data cleaning, data imputation, and feature extraction, the data fusion, the identification of ADL with artificial intelligence, including pattern recognition, and other machine learning techniques. The selection of the methods for all modules is based on the verification of the most used methods.

Following the list of sensors available in off-the-shelf mobile devices, presented in the section 2.1, the sensors that will be used in the framework should be dynamically selected, according to the sensors available in the mobile device. Thus, the types of sensors selected to use in the framework will be motion sensors, magnetic/mechanical sensors, acoustic sensors, and location sensors. Related to the motion sensors, the accelerometer is available in all mobile devices, but gyroscope is only available on some devices, but to cover the execution of the framework in all devices, two different methods should be implemented in the framework with accelerometer and gyroscope, and only with accelerometer. Related to the magnetic/mechanical sensors, the magnetometer is only available on some devices, and this sensor should be included in the same method of motion sensors, combining this sensor with the method only with accelerometer, and the method with accelerometer and gyroscope. Related to the acoustic sensors, the microphone is available in all mobile devices. Related to the location sensors, the GPS is available in a major part of the mobile devices.

With the presentation of the data acquisition methods, in the section 2.2, and aligned with the limitations of the mobile devices, it was verified that the use of a framework for the data acquisition is not needed, using a mobile application for capture the of sensors' data. In addition, the data acquisition method, implemented in the mobile application, should acquire 5 seconds of data from all sensors every 5 minutes.

Following the creation of the new approach for a framework for the identification of ADL and their environments, the selection of data processing methods, presented in the section 2.3, should contains the data cleaning, data imputation, and feature extraction methods.

The data cleaning methods adapted for the framework depends on the types of sensors. For the accelerometer, gyroscope, and magnetometer sensors, the data cleaning method that should be applied is a low pass filter to remove the noise occurred during the data acquisition process. On the other hand, for the acoustic sensors, the data cleaning methods that should be applied is the FFT in order to extract the frequencies of the audio. Finally, for the location sensors, there are no data cleaning methods needed.

The data imputation methods is not important to implement in the development of a new approach for a framework for the identification of ADL and their environments, assuming that the data acquired from all sensors is always filled.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

Related to the feature extraction, the features needed to recognize the ADL and their environments should be selected based on each type of sensors. Firstly, the features selected for the accelerometer, gyroscope, and magnetometer sensors are the 5 greater distances between the maximum peaks, the average of the maximum peaks, the standard deviation of the maximum peaks, the variance of the maximum peaks, the median of the maximum peaks, the standard deviation of the raw signal, the average of the raw signal, the maximum value of the raw signal, the minimum value of the raw signal, the variance of the of the raw signal, and the median of the raw signal. Secondly, the features selected for the microphone are the standard deviation of the raw signal, the average of the raw signal, the maximum value of the raw signal, the minimum value of the raw signal, the variance of the of the raw signal, the median of the raw signal, and 26 MFCC coefficients. Finally, the features selected for the GPS receiver are the distance travelled during the acquisition time.

Before the presentation of the data fusion and pattern recognition methods that should be used for in the framework, the ADL and environments to recognize should be defined. This process should be executed in several phases, presented in the figure 2 and table 6, these are:

1. The identification of ADL with motion and magnetic/mechanical sensors;
2. The identification of the environments with acoustic sensors;
3. The identification of more standing activities with the fusion of the data acquired from motion, magnetic/mechanical and acoustic sensors;
4. The identification of more standing activities with the fusion of the data acquired from motion, magnetic/mechanical, acoustic and location sensors.

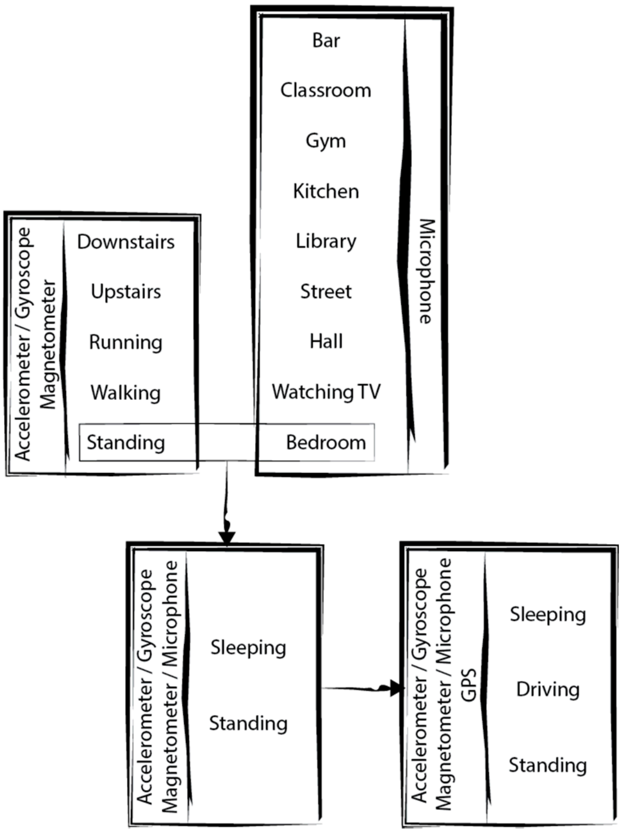

Figure 2. Sensors used for the recognition of Activities of Daily Living (ADL) and environments for each phase of development.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

Table 7. Sensors, Activities of Daily Living (ADL), and environments for recognition with the framework proposed.

		Accelerometer	Gyroscope	Magnetometer	Microphone	GPs
Activities	Downstairs	✓	✓	✓		
	Upstairs	✓	✓	✓		
	Running	✓	✓	✓		
	Walking	✓	✓	✓		
	Standing	✓	✓	✓	✓	✓
	Sleeping	✓	✓	✓	✓	✓
	Driving	✓	✓	✓	✓	✓
Environments	Bar				✓	
	Classroom				✓	
	Gym				✓	
	Library				✓	
	Kitchen				✓	
	Street				✓	
	Hall				✓	
	Watching tv				✓	
	Bedroom				✓	

Firstly, the proposed ADL to identify with the framework will be going downstairs, going upstairs, running, walking, and standing. Secondly, the proposed environments to identify with the framework will be bar, classroom, gym, kitchen, library, street, hall, watching TV, and bedroom. Thirdly, the proposed ADL to distinct with the framework will be sleeping, and standing. Finally, the proposed ADL to distinct with the framework are sleeping, standing, and driving.

Based on the list of data fusion methods and pattern recognition methods, defined in the sections 2.4 and 2.5, the method that should be applied in the framework are ANN, but, firstly, should implemented three types of ANN, in order to compare the different accuracies obtained with several parameters, and type of ANN with better accuracy should be implemented in the new approach for a framework for the identification of ADL and their environments. The selected method is ANN, because, based on the literature, it is the method that reports best accuracies and it is the most used. The different types of ANN that will be applied to the acquired data, in order to identify the best type of ANN are:

- Multilayer Perception (MLP) with Backpropagation;
- Feedforward neural network with Backpropagation;
- Deep Learning.

Regarding the data acquired from GPS receiver, it can be useful to increase the accuracy of the identification of the ADL and their environments, but it can also be used for the identification of the location where the ADL are executed, in order to improve the comparison with the users' agenda presented in the section 2.6. In the future, the use of users' agenda will be important to allow the creation of a personal digital life coach, or to validate the reliability of the implemented methods in the framework.

4. Discussion and conclusions

This paper presents the architecture of a new approach for a framework for the identification of ADL and their environments, using methods with a reported good accuracy. The development of the new approach for the development of a framework for the identification of ADL and their environments, based on the system presented in [4, 15, 16], is one of the steps for the creation of a personal digital life coach [3] using mobile devices.

The framework will be composed by several modules several, such as data acquisition, data processing, data fusion, and a module to implement artificial intelligence techniques for the identification of the ADL and their environments.

Peer-reviewed version available at Sensors 2018, 18, 640; doi:10.3390/s18020640

The sensors used in the framework will be accelerometer, gyroscope, magnetometer, microphone, and GPS, in order to recognize several ADL, including going downstairs, going upstairs, running, walking, standing, sleeping, and driving, and their environments, including bar, classroom, gym, kitchen, library, street, hall, watching TV, and bedroom.

The sensors' data should be acquired and, before the extraction of the features of the sensors' data, some filters (*i.e.*, low pass filter, and FFT) should be applied. Afterwards, the data fusion and pattern recognition methods should be applied. The most used methods for data fusion and pattern recognition, that should be implemented in the framework, is the ANN to achieve the final purpose of the framework developed.

Acknowledgments: This work was supported by FCT project **UID/EEA/50008/2013** (*Este trabalho foi suportado pelo projecto FCT UID/EEA/50008/2013*).

The authors would also like to acknowledge the contribution of the COST Action IC1303 – AAPELE – Architectures, Algorithms and Protocols for Enhanced Living Environments.

Author Contributions: All the authors have contributed with the structure, content, and writing of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Salazar, L.H.A., et al., *A Systematic Literature Review on Usability Heuristics for Mobile Phones.* International Journal of Mobile Human Computer Interaction, 2013. **5**(2): p. 50-61.
2. Foti, D. and J.S. Koketsu, *Activities of daily living.* Pedretti's Occupational Therapy: Practical Skills for Physical Dysfunction, 2013. **7**: p. 157-232.
3. Garcia, N.M., *A Roadmap to the Design of a Personal Digital Life Coach*, in *ICT Innovations 2015*. 2016, Springer.
4. Pires, I., et al., *From Data Acquisition to Data Fusion: A Comprehensive Review and a Roadmap for the Identification of Activities of Daily Living Using Mobile Devices.* Sensors, 2016. **16**(2): p. 184.
5. Banos, O., et al., *On the use of sensor fusion to reduce the impact of rotational and additive noise in human activity recognition.* Sensors (Basel), 2012. **12**(6): p. 8039-54.
6. Akhoundi, M.A.A. and E. Valavi, *Multi-Sensor Fuzzy Data Fusion Using Sensors with Different Characteristics.* arXiv preprint arXiv:1010.6096, 2010.
7. Paul, P. and T. George, *An Effective Approach for Human Activity Recognition on Smartphone.* 2015 Ieee International Conference on Engineering and Technology (Icetech), 2015: p. 45-47.
8. Hsu, Y.-W., et al. *Smartphone-based fall detection algorithm using feature extraction*. in *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. 2016. Datong, China: IEEE.
9. Dernbach, S., et al. *Simple and Complex Activity Recognition through Smart Phones*. in *2012 8th International Conference on Intelligent Environments (IE)*. 2012. Guanajuato, Mexico: IEEE.
10. Shen, C., Y.F. Chen, and G.S. Yang. *On Motion-Sensor Behavior Analysis for Human-Activity Recognition via Smartphones*. in *2016 Ieee International Conference on Identity, Security and Behavior Analysis (Isba)*. 2016. Sendai, Japan: IEEE.
11. Misra, A. and L. Lim, *Optimizing Sensor Data Acquisition for Energy-Efficient Smartphone-Based Continuous Event Processing*, in *Mobile Data Management (MDM), 2011 12th IEEE International Conference on*. 2011, IEEE: Lulea. p. 88-97.
12. D'Ambrosio, A., M. Aria, and R. Siciliano, *Accurate Tree-based Missing Data Imputation and Data Fusion within the Statistical Learning Paradigm.* Journal of Classification, 2012. **29**(2): p. 227-258.
13. Dong, J., et al., *Advances in multi-sensor data fusion: algorithms and applications.* Sensors (Basel), 2009. **9**(10): p. 7771-84.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

14. King, R.C., et al., *Application of data fusion techniques and technologies for wearable health monitoring.* Med Eng Phys, 2017. **42**: p. 1-12.
15. Pires, I.M., N.M. Garcia, and F. Flórez-Revuelta. *Multi-sensor data fusion techniques for the identification of activities of daily living using mobile devices*. in *Proceedings of the ECMLPKDD 2015 Doctoral Consortium, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. 2015. Porto, Portugal.
16. Pires, I.M., et al. *Identification of Activities of Daily Living Using Sensors Available in off-the-shelf Mobile Devices: Research and Hypothesis*. in *Ambient Intelligence-Software and Applications–7th International Symposium on Ambient Intelligence (ISAmI 2016)*. 2016. Springer, Cham.
17. White, R.M., *A Sensor Classification Scheme.* Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on, 1987. **34**(2): p. 124-126.
18. Bojinov, H., et al., *Mobile device identification via sensor fingerprinting.* arXiv preprint arXiv:1408.1416, 2014.
19. Katevas, K., H. Haddadi, and L. Tokarchuk. *Sensingkit: Evaluating the sensor power consumption in ios devices*. in *Intelligent Environments (IE), 2016 12th International Conference on*. 2016. IEEE.
20. Bersch, S.D., et al., *Sensor data acquisition and processing parameters for human activity classification.* Sensors (Basel), 2014. **14**(3): p. 4239-70.
21. Seungwoo, K., et al. *Orchestrator: An active resource orchestration framework for mobile context monitoring in sensor-rich mobile environments*. in *Pervasive Computing and Communications (PerCom), 2010 IEEE International Conference on*. 2010.
22. Vallina-Rodriguez, N. and J. Crowcroft, *ErdOS: achieving energy savings in mobile OS*, in *Proceedings of the sixth international workshop on MobiArch*. 2011, ACM: Bethesda, Maryland, USA. p. 37-42.
23. Priyantha, B., D. Lymberopoulos, and L. Jie, *LittleRock: Enabling Energy-Efficient Continuous Sensing on Mobile Phones.* Pervasive Computing, IEEE, 2011. **10**(2): p. 12-15.
24. Lu, H., et al., *The Jigsaw continuous sensing engine for mobile phone applications*, in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*. 2010, ACM: Zürich, Switzerland. p. 71-84.
25. Rachuri, K.K., et al., *SociableSense: exploring the trade-offs of adaptive sampling and computation offloading for social sensing*, in *Proceedings of the 17th annual international conference on Mobile computing and networking*. 2011, ACM: Las Vegas, Nevada, USA. p. 73-84.
26. Gupta, H.P., et al., *A continuous hand gestures recognition technique for human-machine interaction using accelerometer and gyroscope sensors.* IEEE Sensors Journal, 2016. **16**(16): p. 6425-6432.
27. Deshpande, A., et al., *Model-driven data acquisition in sensor networks*, in *Proceedings of the Thirtieth international conference on Very large data bases - Volume 30*. 2004, VLDB Endowment: Toronto, Canada. p. 588-599.
28. Kubota, H., et al., *A Study of Data Acquisition and Analysis for Driver's Behavior and Characteristics through Application of Smart Devices and Data Mining.* Published by: The Society of Digital Information, 2016: p. 63.
29. Ayu, M.A., et al. *Recognizing user activity based on accelerometer data from a mobile phone*. in *Computers & Informatics (ISCI), 2011 IEEE Symposium on*. 2011. IEEE.
30. Banos, O., et al. *mHealthDroid: a novel framework for agile development of mobile health applications*. in *International Workshop on Ambient Assisted Living*. 2014. Springer.
31. Chavan, V.B. and N. Mhala. *Development of Hand Gesture Recognition Framework Using Surface EMG and Accelerometer Sensor for Mobile Devices*. 2015; Available from: https://www.irjet.net/archives/V2/i5/IRJET-V2I542.pdf.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

32. Sarkar, M., et al. *An Android based human computer interactive system with motion recognition and voice command activation*. in *Informatics, Electronics and Vision (ICIEV), 2016 5th International Conference on*. 2016. IEEE.
33. Pires, I.M., et al., *Validation Techniques for Sensor Data in Mobile Health Applications.* Journal of Sensors, 2016. **2016 %@ 1687-725X**.
34. Lane, N.D., et al., *A survey of mobile phone sensing.* IEEE Communications magazine, 2010. **48**(9).
35. Pejovic, V. and M. Musolesi, *Anticipatory Mobile Computing.* ACM Computing Surveys, 2015. **47**(3): p. 1-29.
36. Lin, F.X., A. Rahmati, and L. Zhong. *Dandelion: a framework for transparently programming phone-centered wireless body sensor applications for health*. in *Wireless Health 2010*. 2010. ACM.
37. Postolache, O., et al. *Enabling telecare assessment with pervasive sensing and Android OS smartphone*. in *2011 IEEE International Workshop on Medical Measurements and Applications Proceedings (MeMeA)*. 2011.
38. Jeffery, S.R., et al. *Declarative Support for Sensor Data Cleaning*. in *Proceedings of PERVASIVE'06 Proceedings of the 4th international conference on Pervasive Computing*. 2006. Dublin, Ireland: ACM.
39. Tomar, D. and S. Agarwal, *A survey on pre-processing and post-processing techniques in data mining.* International Journal of Database Theory and Application, 2014. **7**(4): p. 99-128.
40. Park, K., et al. *Human behavioral detection and data cleaning in assisted living environment using wireless sensor networks*. in *Proceedings of the 2nd International Conference on PErvasive Technologies Related to Assistive Environments*. 2009. ACM.
41. Zhuang, Y., et al. *A weighted moving average-based approach for cleaning sensor data*. in *Distributed Computing Systems, 2007. ICDCS'07. 27th International Conference on*. 2007. IEEE.
42. Li, Z., et al., *A vondrak low pass filter for IMU sensor initial alignment on a disturbed base.* Sensors, 2014. **14**(12): p. 23803-23821.
43. Graizer, V. *Effect of low-pass filtering and re-sampling on spectral and peak ground acceleration in strong-motion records*. in *Proc. 15th World Conference of Earthquake Engineering, Lisbon, Portugal*. 2012.
44. UiO. *Fourier analysis and applications to sound processing - UiO*. 2017 27 Aug. 2017]; Available from: http://www.uio.no/studier/emner/matnat/math/MAT.../v12/part1.pdf.
45. Ninness, B., *Spectral Analysis using the FFT.* Department of Electrical and Computer Engineering, The University of Newcastle, Australia.
46. Vateekul, P. and K. Sarinnapakorn, *Tree-Based Approach to Missing Data Imputation*, in *Data Mining Workshops, 2009. ICDMW '09. IEEE International Conference on* 2009, IEEE: Miami, FL. p. 70-75.
47. Ling, W. and F. Dong-Mei, *Estimation of Missing Values Using a Weighted K-Nearest Neighbors Algorithm*, in *Environmental Science and Information Application Technology, 2009. ESIAT 2009. International Conference on*. 2009, IEEE: Wuhan. p. 660-663.
48. García-Laencina, P.J., et al., *K nearest neighbours with mutual information for simultaneous classification and missing data imputation.* Neurocomputing, 2009. **72**(7-9): p. 1483-1493.
49. Rahman, S.A., et al. *Imputation of Missing Values in Time Series with Lagged Correlations*. in *Data Mining Workshop (ICDMW), 2014 IEEE International Conference on*. 2014. IEEE.
50. Batista, G.E. and M.C. Monard, *A Study of K-Nearest Neighbour as an Imputation Method.* HIS, 2002. **87**(251-260): p. 48.
51. Hruschka, E.R., E.R. Hruschka, and N.F.F. Ebecken, *Towards Efficient Imputation by Nearest-Neighbors: A Clustering-Based Approach*, in *AI 2004: Advances in Artificial Intelligence*. 2004, Springer Berlin Heidelberg. p. 513-525.

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

52. JiaWei, L., TaoYang, and YanWang, *Missing value estimation for microarray data based on fuzzy C-means clustering*, in *High-Performance Computing in Asia-Pacific Region, 2005. Proceedings. Eighth International Conference on*. 2005, IEEE: Beijing. p. 6 pp.-616.
53. Ni, D., et al., *Multiple Imputation Scheme for Overcoming the Missing Values and Variability Issues in ITS Data.* Journal of Transportation Engineering, 2005. **131**(12): p. 931-938.
54. Smith, B., W. Scherer, and J. Conklin, *Exploring Imputation Techniques for Missing Data in Transportation Management Systems.* Transportation Research Record, 2003. **1836**(1): p. 132-142.
55. Qu, L., et al., *A BPCA based missing value imputing method for traffic flow volume data*, in *Intelligent Vehicles Symposium, 2008 IEEE*. 2008, IEEE: Eindhoven. p. 985-990.
56. Jiang, N. and L. Gruenwald, *Estimating Missing Data in Data Streams*, in *DASFAA'07 Proceedings of the 12th international conference on Database systems for advanced applications*. 2007, Springer-Verlag Berlin: Heidelberg. p. 981-987.
57. Rahman, S.A., et al., *Combining Fourier and lagged k-nearest neighbor imputation for biomedical time series data.* J Biomed Inform, 2015. **58**: p. 198-207.
58. Huang, X.Y., et al., *Multi-matrices factorization with application to missing sensor data imputation.* Sensors (Basel), 2013. **13**(11): p. 15172-86.
59. Rahman, S.A., et al., *Combining Fourier and lagged k -nearest neighbor imputation for biomedical time series data.* Journal of Biomedical Informatics, 2015. **58**: p. 198-207.
60. Smaragdis, P., B. Raj, and M. Shashanka, *Missing Data Imputation for Time-Frequency Representations of Audio Signals.* Journal of Signal Processing Systems, 2010. **65**(3): p. 361-370.
61. Bayat, A., M. Pomplun, and D.A. Tran, *A Study on Human Activity Recognition Using Accelerometer Data from Smartphones.* 9th International Conference on Future Networks and Communications (Fnc'14) / the 11th International Conference on Mobile Systems and Pervasive Computing (Mobispc'14) / Affiliated Workshops, 2014. **34**: p. 450-457.
62. Khalifa, S., M. Hassan, and A. Seneviratne. *Feature selection for floor-changing activity recognition in multi-floor pedestrian navigation*. in *Mobile Computing and Ubiquitous Networking (ICMU), 2014 Seventh International Conference on*. 2014. Singapore, Singapore: IEEE.
63. Zhao, K.L., et al., *Healthy: A Diary System Based on Activity Recognition Using Smartphone.* 2013 Ieee 10th International Conference on Mobile Ad-Hoc and Sensor Systems (Mass 2013), 2013: p. 290-294.
64. Zainudin, M.N.S., et al., *Activity Recognition based on Accelerometer Sensor using Combinational Classifiers.* 2015 Ieee Conference on Open Systems (Icos), 2015: p. 68-73.
65. Fan, L., Z.M. Wang, and H. Wang, *Human activity recognition model based on Decision tree.* 2013 International Conference on Advanced Cloud and Big Data (Cbd), 2013: p. 64-68.
66. Liu, Y.Y., et al., *An Hidden Markov Model based Complex Walking Pattern Recognition Algorithm.* Proceedings of 2016 Fourth International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Services (Ieee Upinlbs 2016), 2016: p. 223-229.
67. Piyare, R. and S.R. Lee, *Mobile Sensing Platform for Personal Health Management.* 18th Ieee International Symposium on Consumer Electronics (Isce 2014), 2014: p. 1-2.
68. Chen, Y.F. and C. Shen, *Performance Analysis of Smartphone-Sensor Behavior for Human Activity Recognition.* Ieee Access, 2017. **5**: p. 3095-3110.
69. Vavoulas, G., et al., *The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones.* Proceedings of the International Conference on Information and Communication Technologies for Ageing Well and E-Health (Ict4awe), 2016: p. 143-151.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

70. Torres-Huitzil, C. and M. Nuno-Maganda, *Robust smartphone-based human activity recognition using a tri-axial accelerometer.* 2015 Ieee 6th Latin American Symposium on Circuits & Systems (Lascas), 2015: p. 1-4.
71. Anjum, A. and M.U. Ilyas, *Activity Recognition Using Smartphone Sensors.* 2013 Ieee Consumer Communications and Networking Conference (Ccnc), 2013: p. 914-919.
72. Kumar, A. and S. Gupta, *Human Activity Recognition through Smartphone's Tri-Axial Accelerometer using Time Domain Wave Analysis and Machine Learning.* International Journal of Computer Applications, 2015. **127**(18): p. 22-26.
73. Hon, T.K., et al., *Audio Fingerprinting for Multi-Device Self-Localization.* IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015. **23**(10): p. 1623-1636.
74. Sert, M., B. Baykal, and A. Yazici. *A Robust and Time-Efficient Fingerprinting Model for Musical Audio*. in *2006 IEEE International Symposium on Consumer Electronics*. 2006.
75. Ramalingam, A. and S. Krishnan, *Gaussian Mixture Modeling of Short-Time Fourier Transform Features for Audio Fingerprinting.* IEEE Transactions on Information Forensics and Security, 2006. **1**(4): p. 457-463.
76. Vincenty, T., *Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations.* Survey Review, 1975. **22**(176): p. 88-93.
77. Karney, C.F.F., *Algorithms for geodesics.* Journal of Geodesy, 2013. **87**(1): p. 43-55.
78. Karney, C.F.F. and R.E. Deakin, *FW Bessel (1825): The calculation of longitude and latitude from geodesic measurements.* Astronomische Nachrichten, 2010. **331**(8): p. 852-861.
79. Khaleghi, B., et al., *Multisensor data fusion: A review of the state-of-the-art.* Information Fusion, 2013. **14**(1): p. 28-44.
80. Pombo, N., et al., *Medical decision-making inspired from aerospace multisensor data fusion concepts.* Inform Health Soc Care, 2015. **40**(3): p. 185-97.
81. Durrant-Whyte, H., M. Stevens, and E. Nettleton. *Data fusion in decentralised sensing networks*. in *4th International Conference on Information Fusion*. 2001.
82. Tanveer, F., O.T. Waheed, and Atiq-ur-Rehman, *Design and Development of a Sensor Fusion based Low Cost Attitude Estimator.* Journal of Space Technology,, 2011. **1**(1): p. 45-50.
83. Ko, M.H., et al., *Using dynamic time warping for online temporal fusion in multisensor systems.* Information Fusion, 2008. **9**(3): p. 370-388.
84. Zhao, L., P. Wu, and H. Cao, *RBUKF Sensor Data Fusion for Localization of Unmanned Mobile Platform.* Research Journal of Applied Sciences, Engineering and Technology, 2013. **6**(18): p. 3462-3468.
85. Walter, O., et al. *Smartphone-based sensor fusion for improved vehicular navigation*. in *Positioning Navigation and Communication (WPNC), 2013 10th Workshop on*. 2013.
86. Grunerbl, A., et al., *Smart-Phone Based Recognition of States and State Changes in Bipolar Disorder Patients.* IEEE J Biomed Health Inform, 2014.
87. Thatte, G., et al., *Optimal Time-Resource Allocation for Energy-Efficient Physical Activity Detection.* IEEE Trans Signal Process, 2011. **59**(4): p. 1843-1857.
88. Bhuiyan, M.Z.H., et al., *Performance Evaluation of Multi-Sensor Fusion Models in Indoor Navigation.* European Journal of Navigation, 2013. **11**(2): p. 21-28.
89. Bellos, C., et al., *Heterogeneous data fusion and intelligent techniques embedded in a mobile application for real-time chronic disease management.* Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 8303-6.
90. Ayachi, F.S., et al., *The use of empirical mode decomposition-based algorithm and inertial measurement units to auto-detect daily living activities of healthy adults.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2016. **24**(10): p. 1060-1070 %@ 1534-4320.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

91. Debes, C., et al., *Monitoring activities of daily living in smart homes: Understanding human behavior.* IEEE Signal Processing Magazine, 2016. **33**(2): p. 81-94 %@ 1053-5888.
92. Koza, J.R., et al., *Automated design of both the topology and sizing of analog electrical circuits using genetic programming*, in *Artificial Intelligence in Design'96*. 1996, Springer. p. 151-170.
93. Russell, S., P. Norvig, and A. Intelligence, *A modern approach.* Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs, 1995. **25**: p. 27.
94. Du, K.-L. and M.N.S. Swamy, *Fundamentals of Machine Learning*, in *Neural Networks and Statistical Learning*. 2014, Springer. p. 15-65.
95. Zhang, Y. and J.C. Rajapakse, *Machine learning in bioinformatics*. Vol. 4 %@ 0470397411. 2009: John Wiley & Sons.
96. Witten, I.H., et al., *Data Mining: Practical machine learning tools and techniques*. 2016: Morgan Kaufmann.
97. Schapire, R.E., *The boosting approach to machine learning: An overview*, in *Nonlinear estimation and classification*. 2003, Springer. p. 149-171.
98. Michalski, R.S., J.G. Carbonell, and T.M.X. Mitchell, *Machine learning: An artificial intelligence approach*. 2013: Springer Science & Business Media.
99. Bishop, C.M., *Pattern recognition and machine learning*. 2006: springer.
100. Lorenzi, P., et al., *Mobile Devices for the Real-Time Detection of Specific Human Motion Disorders.* Ieee Sensors Journal, 2016. **16**(23): p. 8220-8227.
101. Lau, S.L., et al. *Supporting patient monitoring using activity recognition with a smartphone*. in *Wireless Communication Systems (ISWCS), 2010 7th International Symposium on*. 2010. York, UK: IEEE.
102. Lau, S.L. *Comparison of orientation-independent-based-independent-based movement recognition system using classification algorithms*. in *Wireless Technology and Applications (ISWTA), 2013 IEEE Symposium on*. 2013. Kuching, Malaysia: IEEE.
103. Duarte, F., A. Lourenco, and A. Abrantes. *Activity classification using a smartphone*. in *e-Health Networking, Applications & Services (Healthcom), 2013 IEEE 15th International Conference on*. 2013. Lisbon, Portugal: IEEE.
104. Fahim, M., S. Lee, and Y. Yoon, *SUPAR: Smartphone as a ubiquitous physical activity recognizer for u-healthcare services.* Conf Proc IEEE Eng Med Biol Soc, 2014. **2014**: p. 3666-9.
105. Bajpai, A., et al., *Quantifiable fitness tracking using wearable devices.* Conf Proc IEEE Eng Med Biol Soc, 2015. **2015**: p. 1633-7.
106. Nguyen, P., et al., *User-friendly Activity Recognition Using SVM Classifier and Informative Features.* 2015 International Conference on Indoor Positioning and Indoor Navigation (Ipin), 2015: p. 1-8.
107. Wang, C., et al. *SW-HMM: A Method for Evaluating Confidence of Smartphone-Based Activity Recognition*. in *Trustcom/BigDataSE/ISPA, 2016 IEEE*. 2016. Tianjin, China: IEEE.
108. Lau, S.L. and K. David. *Movement recognition using the accelerometer in smartphones*. in *Future Network and Mobile Summit, 2010*. 2010. IEEE.
109. Zhang, L., X. Wu, and D. Luo. *Real-Time Activity Recognition on Smartphones Using Deep Neural Networks*. in *Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), 2015 IEEE 12th Intl Conf on*. 2015. Beijing, China: IEEE.
110. Cardoso, N., J. Madureira, and N. Pereira, *Smartphone-based Transport Mode Detection for Elderly Care.* 2016 Ieee 18th International Conference on E-Health Networking, Applications and Services (Healthcom), 2016: p. 261-266.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 January 2018 doi:10.20944/preprints201801.0068.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 640; doi:10.3390/s18020640

111. Vallabh, P., et al., *Fall Detection Using Machine Learning Algorithms.* 2016 24th International Conference on Software, Telecommunications and Computer Networks (Softcom), 2016: p. 51-59.
112. Filios, G., et al., *Hierarchical Algorithm for Daily Activity Recognition via Smartphone Sensors.* 2015 Ieee 2nd World Forum on Internet of Things (Wf-Iot), 2015: p. 381-386.
113. Tang, C.X. and V.V. Phoha, *An Empirical Evaluation of Activities and Classifiers for User Identification on Smartphones.* 2016 Ieee 8th International Conference on Biometrics Theory, Applications and Systems (Btas), 2016: p. 1-8.
114. Li, P., et al., *An Automatic User-Adapted Physical Activity Classification Method Using Smartphones.* IEEE Trans Biomed Eng, 2017. **64**(3): p. 706-714.
115. Kim, Y.J., B.N. Kang, and D. Kim, *Hidden Markov Model Ensemble for Activity Recognition using Tri-axis Accelerometer.* 2015 Ieee International Conference on Systems, Man, and Cybernetics (Smc 2015): Big Data Analytics for Human-Centric Systems, 2015: p. 3036-3041.
116. Brdiczka, O. and V. Bellotti. *Identifying routine and telltale activity patterns in knowledge work*. in *Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on*. 2011. Palo Alto, CA, USA: IEEE.
117. Costa, A., et al., *Sensor-driven agenda for intelligent home care of the elderly.* Expert Systems with Applications, 2012. **39**(15): p. 12192-12204.