Article

HMM Adaptation for Improving a Human Activity Recognition System

Rubén San-Segundo *, Juan M. Montero, José Moreno-Pimentel and José M. Pardo

Speech Technology Group, E.T.S.I. Telecomunicación, Universidad Politecnica de Madrid, 28040 Madrid, Spain; juancho@die.upm.es (J.M.M.); josemoreno@die.upm.es (J.M.-P.); pardo@die.upm.es (J.M.P.)
* Correspondence: ruben.sansegundo@upm.es; Tel.: +34-915-495700 (ext. 4228)

Abstract: When developing a fully automatic system for evaluating motor activities performed by a person, it is necessary to segment and recognize the different activities in order to focus the analysis. This process must be carried out by a Human Activity Recognition (HAR) system. This paper proposes a user adaptation technique for improving a HAR system based on Hidden Markov Models (HMMs). This system segments and recognizes six different physical activities (walking, walking upstairs, walking downstairs, sitting, standing and lying down) using inertial signals from a smartphone. The system is composed of a feature extractor for obtaining the most relevant characteristics from the inertial signals, a module for training the six HMMs (one per activity), and the last module for segmenting new activity sequences using these models. The user adaptation technique consists of a Maximum A Posteriori (MAP) approach that adapts the activity HMMs to the user, using some activity examples from this specific user. The main results on a public dataset have reported a significant relative error rate reduction of more than 30%. In conclusion, adapting a HAR system to the user who is performing the physical activities provides significant improvement in the system's performance.

Keywords: user adaptation; human activity segmentation; HMMs; smartphone inertial sensors

1. Introduction

The research on multisensor networks has increased significantly in the last 10 years, defining the Internet of Things (IoT) concept. These networks typically include cameras, indoor location systems (ILS), microphones, wearable sensors, etc. Using the information from sensors, computer-based systems can adapt their behaviors to the context conditions (increasing their intelligence) or they can report important information to the user (difficult to obtain through other means). Thanks to the increment of sensor neural networks, the number of possible research areas has also increased rapidly. One of these areas is psycho-motor training where an automatic system senses a psychical activity carried out by a person and provides feedback about the performance. When developing a fully automatic system for evaluating motor activities, one important aspect is to segment and recognize the different activities in order to focus the system analysis on some specific ones. This process must be carried out by a Human Activity Recognition (HAR) system. The recognition of human activities has received a lot of attention in the last five years due to the high number of promising applications and the increasing interest shown by government and commercial organizations.

This paper proposes a user adaptation technique for improving a HAR system based on Hidden Markov Models (HMMs). This system segments and recognizes six different physical activities (walking, walking upstairs, walking downstairs, sitting, standing and lying down) using inertial signals from a smartphone. This paper is organized as follows. Section 2 presents the state of the art. Section 3 shows an overview of the HMMs-based HAR system, describing the main modules. Section 4

describes the user Maximum A Posteriori (MAP) adaptation. Section 5 presents the experiments carried out in this work, including a detailed description of the dataset used in the experiments. The main discussions and conclusions are summarized in Section 6.

2. Background

HAR systems can be categorized according to the sensor type or the machine learning algorithm. According to the sensor type, it is possible to consider various types of sensors: on-body, object-placed or ambient sensors. On example of environment sensors is video cameras in monitoring areas [1,2]. Human activity can be also analyzed based on a rich variety of acoustic events. The determination of both the identity of sounds and their position in time may help to detect and describe that human activity [3]. Most environment sensors require an important infrastructure support: for example, the installation of video cameras in the monitoring areas. Additionally, people not always spend all their time in the same environment. In this respect, environmental sensors are limited by their infrastructure and cannot provide monitoring outside the specific environment. This limitation can be overtaken using on-body sensors [4,5]. Body-worn sensors add new possibilities to the human monitoring system [6]: not only by being able to measure body signals (e.g., physiological, motion, location) but also by providing portable and off-site user supervision at any location without the need of fixed infrastructure. In the literature there are different approaches for locating motion sensors in different body parts such as the waist, wrist, chest and thighs achieving good classification performance [7–9]. In [10], a chest-mounted accelerometer was used for classifying five Activities of Daily Living (ADL). However, the use of body sensors has important limitations such as the user's discomfort while wearing them (these sensors are usually uncomfortable for the common user) and energy-limited mobile devices (they do not provide a long-term solution for activity monitoring).

In recent years, smartphones and smartwatches have become widespread, increasing the number of possibilities for human-centered applications. These devices include embedded built-in sensors such as microphones, dual cameras, accelerometers, gyroscopes, etc. The inertial sensors are a very interesting for monitoring ADL. These devices have important advantages [11,12]: easy device portability, unobtrusive sensing provided by the embedded sensors and the processing power of new smartphones that allow online computation. Because of this, some works focused on HAR using smartphones have been developed [13–17].

HAR is a machine learning problem, where a system extracts features from sensor signals, generates a model for each activity, and classifies the next activities based on these models. In the literature, different machine learning solutions have been applied to the recognition of activities including Naive Bayes [18], Decision Trees [19], and Support Vector Machines (SVMs) [20]. In many works, several approaches are compared: for example, Yang [21] uses the WEKA learning toolkit to compare the accuracy rates of several machine learning approaches: C4.5 Decision Trees, Naïve Bayes, k-Nearest Neighbor, and Support Vector Machines. In [13], three learning algorithms were evaluated: Logistic Regression, J48, and Multilayer Perceptron. Hidden Markov Models is a successful modeling strategy for classifying temporal sequences. HMMs offer dynamic time warping, have clear Bayesian semantics and well-understood training algorithms. HMMs are very robust against degradation, giving the possibility to be trained on one person and to be tested on another. In the last five years, there has been an increase in the number of HAR systems based on HMMs for modeling inertial signals (Table 1).

Table 1. Comparison of previous HAR works based on HMMs.

Ref.	Sensor	Target	Classes	Users	Time	User Adaptation	Performance
Mannini and Sabatini [11]	Five bi-axial accelerometers, located at the hip, wrist, arm, ankle and thigh	Activity recognition and segmentation	7 activities	13	29 min	NO	Error: 1.6%
Lee and Cho [22]	Accelerometer in a LG smartphone	Activity recognition	7 activities	3	339 min	NO	Error: 15.0%
Reddy et al. [23]	Accelerometer in a Nokia n95	Transportation mode recognition	5 modes	16	1200 min	NO	Precision and Recall > 93%
Wang et al. [24]	Tri-axial accelerometer (MMA7260)	Activity recognition	6 activities	13	~100 min	NO	Error: 2.8%
Witowski et al. [25]	Simulated data	Detecting physical activity	2 sedentary behavior vs. physical activity	-	1000 days	NO	Error: 18.8%
Trabelsi et al. [26]	Three MTx 3-DOF (Degree of Freedom) inertial trackers	Activity recognition and segmentation	12 activities	6	~100 min	NO	Error: 9.6%
San-Segundo et al. [17]	Accelerometer in a Samsung Galaxy S2	Activity recognition and segmentation	6 activities	30	220 min	NO	Error: 3.5% (without sequence model)
This paper	Accelerometer in a Samsung Galaxy S2	Activity recognition and segmentation	6 activities	30	220 min	YES	Error: 2.0%

According to Table 1, this work is the first one (known by the authors) that proposes a user adaptation technique for improving the performance of a HMMs-based HAR system. In the literature there are some references proposing other user adaptation techniques: a semi-supervised method [27] or a Multi-Classifier Adaptive-Training (MCAT) algorithm [28]. The MCAT algorithm consists of using a meta-classifier for combining several pattern recognition methods. This meta-classifier is trained with user-dependent adaptation data to improve the results for a specific user.

3. HAR System Overview

Figure 1 shows the general system architecture presented in [17], including an additional module for user MAP adaptation. The system is made up of four main modules: feature extraction, HMMs training, HMMs adaptation and activity recognition/segmentation. The main contributions of this paper are focused on HMMs training and adaptation modules. In order to consider only the influence of the HMMs in the results, the Activity Sequence Model (ASM) proposed in [17] has been deactivated in this work.
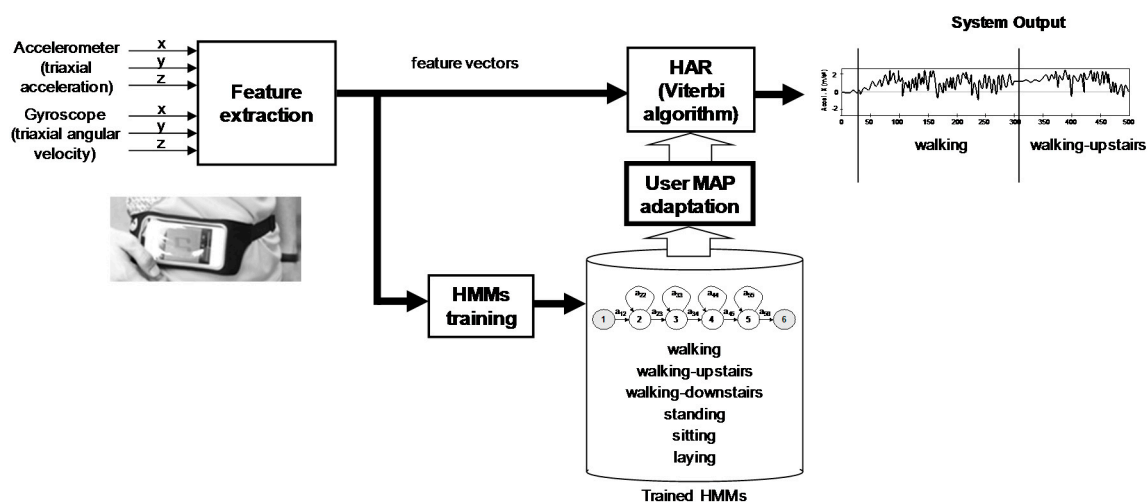

Figure 1. HAR system architecture.

The feature extraction module obtains the accelerometer and gyroscope signals, samples them with a sampling rate of 50 Hz and filters them for noise reduction (with a 20 Hz cut-off frequency). This sampling rate is sufficient for capturing human body motion: more than 95% of its energy is contained below 15 Hz [5]. Using a Butterworth low-pass filter (with a 0.3 Hz cut-off frequency), the sensor acceleration signals are divided into body acceleration and gravity. The Euclidean magnitude and time derivatives (jerk da/dt and angular acceleration dw/dt) are also obtained during the feature extraction process.

The sample sequences are grouped together in frames: fixed-width sliding windows of 2.56 s and 50% overlap (128 samples per frame with an overlap of 64 samples). From each frame, the system obtains a feature vector computing measures from the inertial signals. These features are traditional measures such as the mean, correlation, signal magnitude area (SMA) and auto regression coefficients [29]. This vector has been extended including more features from time and frequency domains generating a vector with a total of 561 features. The dataset used in this work (available at the UCI Machine Learning Repository) includes the already-computed feature vectors. For comparison, this work uses the same features proposed in [30]. The set of features computed from the time domain are mean, standard deviation, median absolute deviation, max, min, magnitude, energy, interquartile range, entropy, autoregression coefficients with the Burg order equal to 4, and the correlation coefficient between different axis. The set of features obtained from the frequency domain includes additional ones such as the index of the frequency component with the largest magnitude, the weighted average

of the frequency components, skewness, kurtosis and the energy in 64 bins covering the whole signal frequency range.

For HMMs development, we used the HTK toolkit [31]. This toolkit allows us to train the HMMs, adapt them to a new user, and recognize new activities using these HMMs. In the HAR used in this work, six HMMs are considered, one for every activity. Every model represents the sequence of observed feature vectors corresponding to each activity. Given a vector sequence, it is compared to all the models, computing their likelihood. The model with the highest likelihood is the activity recognized. A HMM can be seen as a finite state machine in which all the states can change every time unit t. Every state j generates a feature vector Ot considering a probability density $bj(Ot)$. There is also a probability to control transitions between states. For example, the transition between state i a state j can be governed by the discrete probability a_{ij}. Figure 2 represents a HMM with six states, associated with a sequence of six observations $O1$ to $O6$. In HTK, the entry and exit states of a HMM are non-emitting. This is to facilitate the construction of composite models (out of the scope of this paper). The association between the observed vector and states is $X = 1; 2; 3; 3; 4; 5; 5; 6$.
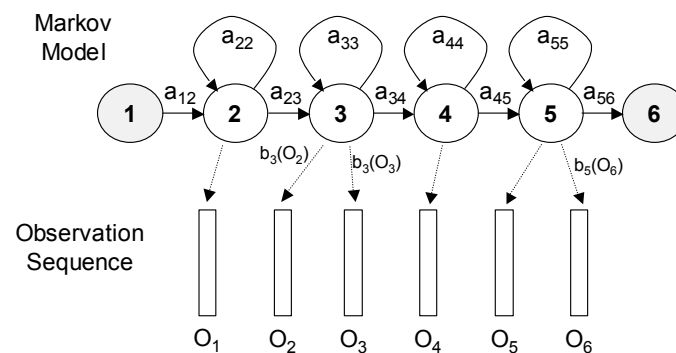

Figure 2. Example of a Hidden Markov Model.

The joint probability that O is generated by the model M moving through the state sequence X is calculated simply as the product of the transition probabilities and the output probabilities:

$$P(O, X/M) = a_{12}b_2(O_1)a_{23}b_3(O_2)a_{33}b_3(O_3)... \tag{1}$$

In a real application, only the observation sequence O is known and the underlying state sequence X is hidden. In this case, the required likelihood is computed by considering only the most likely state sequence. The main parameters of every model are a_{ij} and $b_j(O_t)$. The output distribution $b_j(O_t)$ can be modeled by using Gaussian mixtures, reducing the model parameters to the mean and variance of every Gaussian distribution. The model parameters can be determined automatically through a robust and efficient re-estimation procedure (estimation maximization, EM) considering a set of training examples corresponding to a particular activity:

(1) Initialize all Gaussian distributions with the mean and variance computed throughout the whole dataset.

(2) Calculate the forward and backward probabilities for all states j and times t.

(3) For each state j and time t, use the probability $Lj(t)$ and the current observation vector O_t to update the accumulators for that state.

(4) Use the final accumulator values to calculate new parameter values.

(5) If the value of $P = P(O/M)$ for this iteration is not higher than the value at the previous iteration, then stop; otherwise, repeat the aforementioned steps using the new re-estimated parameter values (from step 2).

In the estimation maximization (EM) algorithm, the main target is to maximize the likelihood of an activity respect to its HMM: Maximum Likelihood Estimation (MLE). In some applications, it

is possible to consider a discriminative training procedure. In this procedure, the main target is to maximize the differences between models. The HTK toolkit includes a tool for training HMMs in a discriminative way. In this work, the Maximum Mutual Information Estimation (MMIE) discriminative procedure has also been evaluated [32]. In this procedure, the main difference compared to MLE is the objective function to be optimized. In the case of the MMIE, the function to maximize is:

$$F_{MMIE} = \sum_{i=activities} \log \frac{P(O_i, X_i / M_i)}{\sum\limits_{n=activities\ sequences} P(O_i, X_n / M_i)} \tag{2}$$

The first term in the numerator is identical to the objective function for the MLE. In order to maximize Equation (2), the numerator must be increased while the denominator is decreased. Similar to the MLE, the MMIE has the target to maximize the likelihood of each observation given by the training sequences. In addition, the MMIE has a denominator term that can be reduced by decreasing the probabilities of other possible activity sequences. In conclusion, the MMIE attempts make the correct hypothesis more probable, and at the same time, it tries to make incorrect hypotheses less probable.

Continuous activity recognition and segmentation involves connecting several HMMs in sequence. In the HTK toolkit, the Viterbi algorithm is expanded to allow several models to be connected in the search space: the last state of every model is connected with the first state of all models (Figure 3). Each model in the sequence corresponds directly to its activity.
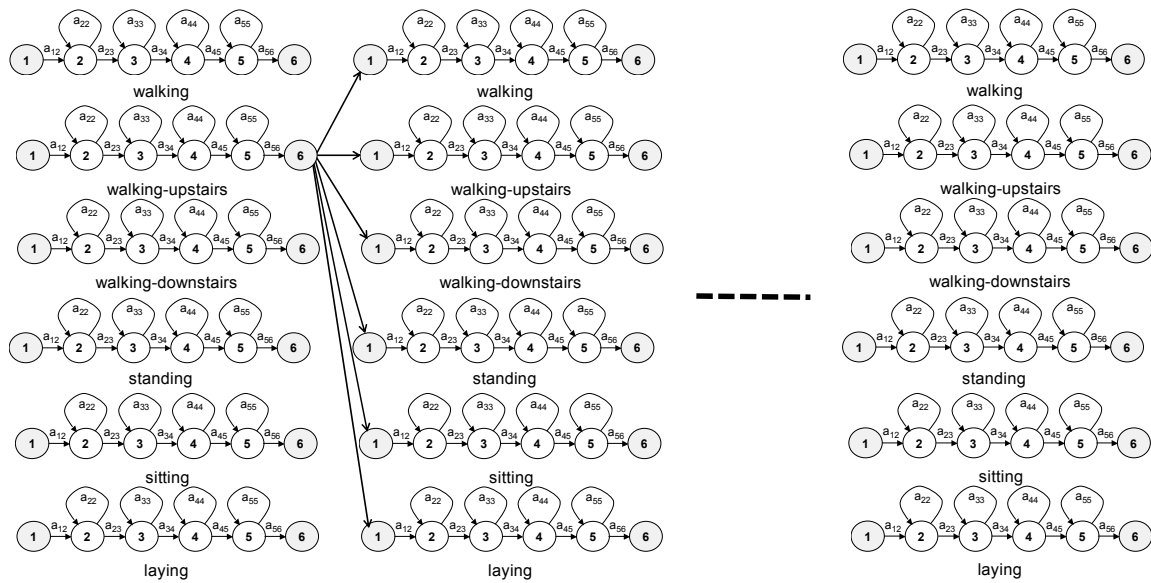

Figure 3. Search space for activity segmentation.

4. User Maximum A Posteriori Adaptation

In order to generate activity HMMs adapted to every user, the system trains general activity HMMs including information from all users. After training these general models, user-adapted activity HMMs are generated by adapting the general models to every user via a Maximum A Posteriori approach (sometimes referred to as Bayesian adaptation). MAP adaptation needs prior knowledge about the model parameter distribution (original HMMs). For MAP adaptation purposes, the informative priors that are generally used are the user-independent model parameters. The updated formula for the μ parameter of mixture component m is:

$$\mu_m = \frac{N_m}{N_m + MU} \mu_{general \cdot m} + \frac{MU}{N_m + MU} \mu_{USER \cdot m} \tag{3}$$

The estimated μ_m is a linear combination of the μ parameter of mixture component m in the original model ($\mu_{general \cdot m}$) and the μ parameter of mixture component m obtained considering only observation vectors from the user ($\mu_{USER \cdot m}$). *MU* is the adaptation coefficient that defines the weighting of user-dependent information compared to general HMMs. N_m is the occupation likelihood of the adaptation data (from the user) along the T frames, defined as:

$$N_m = \sum_{t=1}^{t=T} N(O_t; \mu_{general \cdot m}, \sum_{general \cdot m}) \tag{4}$$

After training the adapted activity HMMs for every user, the activity recognition and segmentation process is the same as the method used with user-independent HMMs: the process consists of computing the likelihood of the best model sequence when generating several activities.

5. Experiments Carried out in This Work

This section describes the dataset used in the experiments, the baseline results considering the training algorithms (MLE and MMIE), the user adaptation experiments, and at the end, the final results with a discussion.

5.1. Dataset Used in the Experiments

This work has been carried out using a public dataset available at the UCI Machine Learning Repository: the Human Activity Recognition Using Smartphones Data Set [30]. This dataset contains inertial information (from smartphone sensors: accelerometer and gyroscope) recorded from a group of 30 people (from 19 to 48 years old), performing six different physical activities several times. These activities are walking, walking upstairs, walking downstairs, sitting, standing and lying down. While performing these activities, every user carried a smartphone (Samsung Galaxy S II) for recording the inertial signals. These signals consisted of the three-axial linear acceleration and the three-axial angular velocity being sampled at a constant rate of 50 Hz. The dataset contains 13,182 s of recording including 400 activity instances from 30 users. An example of the recording process can be seen in a video [33].

In this work, the main aim is to recognize the activity sequence carried by every user. For this evaluation, all activities carried out from the same user have been stored in the same file, defining a recording session. There are 30 sessions with an average number of 13.3 activities per session. The problem to solve in this paper consists of recognizing and segmenting all physical activities recorded in the same session (Figure 4). In the initial configuration of the dataset, it was divided in two sets, and 70% of the users were selected for training and 30% for testing the system. In this work, the 30 sessions have been randomly divided into six subsets. Every session includes all activities carried out by the same user, so all activities from the same user are included in the same subset. This characteristic avoids the person-dependent characteristic being influenced by the activity recognition or segmentation. In order to improve the significance of the results, a six-fold cross-validation procedure has been carried out. The cross-validation procedure uses four subsets for training the activity Hidden Markov Models (HMMs), one for validation (tuning the system parameters) and one for testing. This configuration has been repeated six times in a round-robin strategy. The results presented in this paper are average values obtained throughout the six-fold cross-validation procedure. In every experiment, the system is evaluated with all sessions (13,182 s) (defining a 95% confidence interval of $\pm 0.4\%$). In Sections 5.2 and 5.3, only the validation results were considered for tuning the adaptation coefficient and selecting the training procedure. The final results are presented in Section 5.4 with the testing subsets, using the best system configuration.
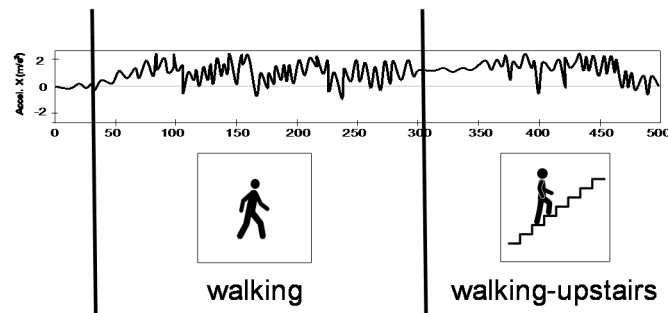
Figure 4. Example of activity recognition and segmentation.

Regarding the evaluation metrics, in this work we considered the Activity Recognition Error Rate (ARER): the percentage of time that has been wrongly assigned to an activity (Equation (5)).

$$ARER(\%) = 100 \frac{Time\ (\text{sec})\ wrongly\ classified}{Session\ duration\ (\text{sec})} \tag{5}$$

In addition to this measure, other possible metrics are precision and recall. The precision is the time correctly assigned to this activity (true positive) divided by the activity time detected by the system (including true and false positive times). Recall is defined as the time correctly assigned (true positive) divided by the actual duration of the activity. As it was shown in [17], there is an important correlation between these three measures. Because of this, ARER will be considered for system development, and in the final experiments, all metrics will be provided for comparison with further works.

5.2. Baseline Experiments Considering Different Training Algorithms

Figure 5 represents the ARER depending on the training procedure. As it is shown, the discriminative training strategy (MMIE) obtains slightly better results, although the differences are not statistically significant. For the next experiments, the MMIE strategy has been considered. In this task, the system is able to distinguish between static and dynamic activities, but the confusion between static activities (standing, sitting and lying down) is very high. A similar behavior occurs for the dynamic activities of walking, walking upstairs and walking downstairs. Figure 5 also includes the error points for all the subjects using the representation method proposed in [34]. Regarding the ARER distributions, both training procedures show very similar behaviors.
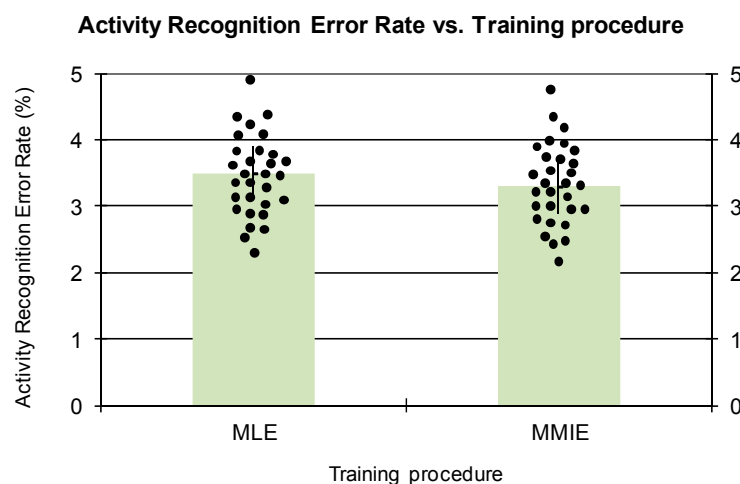

Figure 5. Activity recognition error rate depending on the training procedure including data points for all the subjects using the representation method proposed in [34].

5.3. User Adaptation Experiments

For the user adaptation experiments, all user sessions used for testing the system are randomly divided into two sub-sessions including 50% of the activities each. The average duration of an activity is 440 s. The first sub-session is used for adapting the HMMs to the user and the second sub-session is used for testing the system. There is not any overlap between these two sub-sessions. It is important to remark that the sub-session used for testing is the same along all the experiments to allow a fair comparison.

Figure 6 represents the evolution of the ARER depending on the adaptation coefficient. This representation also includes the confidence intervals at 95% along the curve. When MU = 0 no adaptation is done and the system obtains the same results as when using general HMMs. When MU increases, the ARER decreases until reaching a minimum for MU = 15. After this value, the ARER increases because the available user-dependent data is limited and it is not productive to increase its weight for HMMs training.
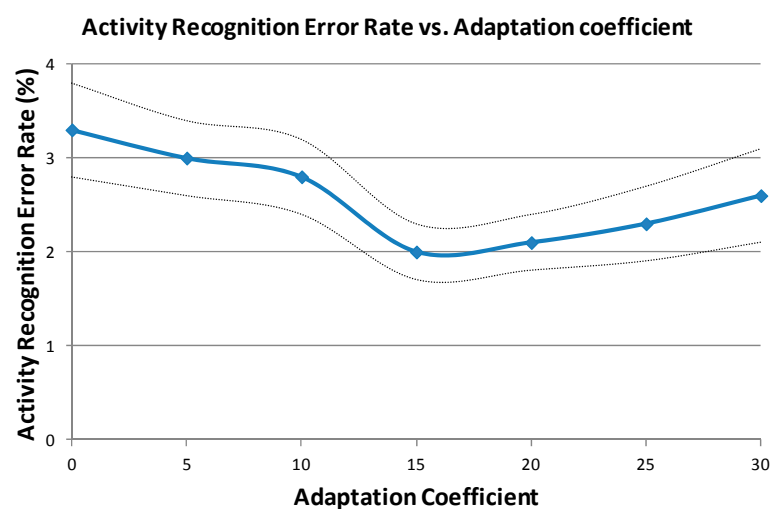

Figure 6. Activity recognition error depending on the adaptation coefficient.

The next figure (Figure 7) shows the ARER depending on the amount of data to adapt. Instead of using the whole sub-session (50% of the user session) for adapting the HMMs, different amounts have been considered: 10%, 20%, 30%, 40% and 50% of the user session.
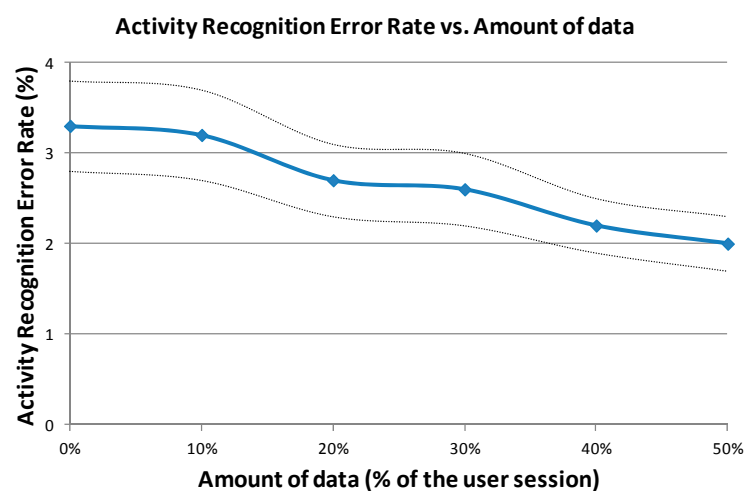

Figure 7. Activity recognition error rate depending on the amount of data for HMMs adaptation (% of the user session).

As it is shown in Figure 7, the ARER decreases when we increase the amount of data (this representation also includes the confidence intervals at 95% along the curve). Regarding the differences between users, the absolute ARER reduction varies from 0.5% (the lowest reduction) to 1.5% (the highest reduction). The curve presented in Figure 7 does not show any saturation tendency, so it means that if we increase the adaptation data, the ARER can continue decreasing. For further works, a bigger dataset will be considered in order to analyze this effect.

In order to complete the analysis, a new experiment has been carried by training the HAR system with the data of a single user. In this case, the first user sub-session has been used for training the system (instead of adapting the system) and the second sub-session for testing. In this case, the ARER increases to 9.4%. This significant degradation is due to the important reduction in the amount of data for training the system. This result supports the utility of the adaptation algorithm proposed in this paper as the best solution for developing a user-dependent HAR system when there is a small amount of data per user.

5.4. Final Experiments and Discussion

This subsection presents the final results on test datasets. These experiments have been carried out considering the best system configuration obtained from the analyses done on the validation sets (see previous subsection): MMIE HMMs training and user MAP adaptation using 50% of the user session for adaptation (first sub-session) and 50% for testing (second sub-session). The sub-session used for testing is the same in all the experiments. It is important to remark that there is not any overlap between testing and adaptation data. Table 2 shows the final results obtained on validation and test subsets. This table includes the activity segmentation error rates (%), recall (%), precision (%) and confidence intervals at 95%. The results on validation subsets have already been presented in previous subsections.

Table 2. Final segmentation results including activity recognition error rate (ARER), recall and precision metrics.

User Adaptation	Validation	Test		
	ARER (%)	ARER (%)	Recall (%)	Precision (%)
Baseline	$3.5\% \pm 0.4\%$	$3.5\% \pm 0.4\%$	$92.9\% \pm 0.4\%$	$92.6\% \pm 0.4\%$
MMIE training	$3.3\% \pm 0.4\%$	$3.3\% \pm 0.4\%$	$93.2\% \pm 0.4\%$	$93.1\% \pm 0.4\%$
MMIE training + User MAP adaptation	$2.0\% \pm 0.3\%$	$2.1\% \pm 0.3\%$	$95.2\% \pm 0.3\%$	$95.1\% \pm 0.3\%$

As shown, the final results on test subsets are slightly worse because the system was optimized on the validation subset. Similar to the conclusion obtained in previous subsections, the ARER error decreases significantly when adapting the HMMs to the user considering a MAP approach.

Table 3 includes the same experiments considering the original dataset partition (70% for HMMs training and 30% for testing) for comparing to previous works. These results show that using a user MAP adaptation, it is possible to significantly improve the segmentation results obtained in previous works on this dataset. When adapting the HMMs to the user, the HAR has better modeling for recognizing the physical activities carried out by this specific user. This adaptation is the main contribution of this paper.

Table 3. Final results considering the original dataset partition including activity recognition error rate (ARER), recall and precision metrics.

System	Test		
	ARER (%)	Recall (%)	Precision (%)
Anguita et al. 2013 [30]	4.0% ± 0.4%	-	-
San-Segundo et al. 2016 [17]	3.2% ± 0.3%	93.3% ± 0.4%	93.1% ± 0.4%
MMIE training	3.1% ± 0.3%	93.8% ± 0.3%	93.9% ± 0.3%
MMIE training + user MAP adaptation	2.0% ± 0.3%	95.3% ± 0.3%	95.2% ± 0.3%

6. Conclusions

This work has proposed a user adaptation technique for improving a HAR system based on HMMs. This system segments and recognizes six different physical activities (walking, walking upstairs, walking downstairs, sitting, standing and lying down) using inertial signals from a smartphone. The system is composed of a feature extractor for obtaining the most relevant characteristics from the inertial signals, a module for training the six HMMs (one per activity), and the last module for segmenting new activity sequences using these models. This paper has evaluated two different HMMs training strategies: the first one, a generative approach (Maximum Likelihood Estimation), and the second, a discriminative one (Maximum Mutual Information Estimation). The discriminative training strategy (MMIE) obtains slightly better results, although the differences are not statistically significant.

The main contribution of this paper is the analysis of a user adaptation technique for adapting the HMMs to the user who performs the different activities. The user adaptation technique consists of a Maximum A Posteriori (MAP) approach. The final results on a public dataset [30] have reported significant error rate reduction: from 3.2% to 2.0% ARER (more than 30% relative error rate reduction). In conclusion, adapting a HAR system to the user who is performing the physical activities reports significant improvement in the activity segmentation process.

Author Contributions: Rubén San-Segundo implemented the HAR system using HMMs, executed the experimental work for the adaptation process, analyzed the results, drafted the initial manuscript and revised the manuscript. Juan Manuel Montero conceptualized the HMM adaptation process, supervised the experiments and revised the manuscript. José Moreno-Pimentel integrated the discriminative training procedure, executed the experiments related to discriminative training, helped to draft the initial manuscript and revised the final version. José Manuel Pardo analyzed the results, provided feedback and revised the final manuscript.

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

References

1. Poppe, R. Vision-based human motion analysis: An overview. *Comput. Vis. Image Underst.* **2007**, *108*, 4–18. [CrossRef]
2. Poppe, R. A survey on vision-based human action recognition. *Image Vis. Comput.* **2010**, *28*, 976–990. [CrossRef]
3. Temko, A. Acoustic Event Detection and Classification. Ph.D. Thesis, Polytechnic University of Catalonia, Barcelona, Spain, 2009.
4. Lukowicz, P.; Ward, J.A.; Junker, H.; Stäger, M.; Tröster, G.; Atrash, A.; Starner, T. Recognizing workshop activity using body worn microphones and accelerometers. In Proceedings of the 2nd International Conference Pervasive Computing, Vienna, Austria, 21–23 April 2004; pp. 18–22.
5. Karantonis, D.M.; Narayanan, M.R.; Mathie, M.; Lovell, N.H.; Celler, B.G. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Trans. Inf. Technol. Biomed.* **2006**, *10*, 156–167. [CrossRef] [PubMed]

6.　Bao, L.; Intille, S.S. *Activity Recognition from User-Annotated Acceleration Data*; Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Nierstrasz, O., Rangan, C.P., Steffen, B., Terzopoulos, D., Tygar, D., et al., Eds.; Pervasive Computing: Linz/Vienna, Austria, 2004; pp. 1–17.

7.　Casale, P.; Pujol, O.; Radeva, P. Human activity recognition from accelerometer data using a wearable device. In *Pattern Recognition and Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2011; p. 289.

8.　Krishnan, N.; Narayanan, C.; Colbry, D.; Juillard, C.; Panchanathan, S. Real time human activity recognition using tri-axial accelerometers. In Proceedings of the Sensors, Signals and Information Processing Workshop, Sedona, AZ, USA, 11–14 May 2008.

9.　Nishkam, R.; Nikhil, D.; Preetham, M.; Littman, M.L. Activity recognition from accelerometer data. In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence, Pittsburgh, PA, USA, 9–13 July 2005; pp. 1541–1546.

10.　Hanai, Y.; Nishimura, J.; Kuroda, T. Haar-like filtering for human activity recognition using 3d accelerometer. In Proceedings of the IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop (DSP/SPE), Marco Island, FL, USA, 4–7 January 2009; pp. 675–678.

11.　Mannini, A.; Sabatini, A.M. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* **2010**, *10*, 1154–1175. [CrossRef] [PubMed]

12.　Vinh, L.T.; Lee, S.; Le, H.X.; Ngo, H.Q.; Kim, H.I.; Han, M.; Lee, Y.-K. Semi-markov conditional random fields for accelerometer-based activity recognition. *Appl. Intell.* **2011**, *35*, 226–241. [CrossRef]

13.　Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. *SIGKDD Explor. Newslett.* **2011**, *12*, 74–82. [CrossRef]

14.　Brezmes, T.; Gorricho, J.L.; Cotrina, J. Activity recognition from accelerometer data on a mobile phone. In *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 796–799.

15.　San-Segundo, R.; Montero, J.M.; Barra-Chicote, R.; Fernández, F.; Pardo, J.M. Feature Extraction from Smartphone Inertial Signals for Human Activity Segmentation. *Signal Proc.* **2016**, *120*, 359–372. [CrossRef]

16.　Wu, W.; Dasgupta, S.; Ramirez, E.E.; Peterson, C.; Norman, G.J. Classification accuracies of physical activities using smartphone motion sensors. *J. Med. Intern. Res.* **2012**, *14*, e130. [CrossRef] [PubMed]

17.　San-Segundo, R.; Lorenzo-Trueba, J.; Martínez-González, B.; Pardo, J.M. Segmenting human activities based on HMMs using smartphone inertial sensors. *Pervasive Mob. Comput.* **2016**, *30*, 84–96. [CrossRef]

18.　Jatoba, L.C.; Grossmann, U.; Kunze, C.; Ottenbacher, J.; Stork, W. Context-aware mobile health monitoring: Evaluation of different pattern recognition methods for classification of physical activity. In Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, USA, 20–25 August 2008.

19.　Maurer, U.; Smailagic, A.; Siewiorek, D.; Deisher, M. Activity recognition and monitoring using multiple sensors on different body positions. In Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06), Cambridge, MA, USA, 3–5 April 2006.

20.　Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. Energy Efficient Smartphone-Based Activity Recognition using Fixed-Point Arithmetic. *J. Univ. Comput. Sci.* **2013**, *19*, 1395–1314.

21.　Yang, J. Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In Proceedings of the 1st ACM International Workshop on Interactive Multimedia for Consumer Electronics (IMCE '09), Beijing, China, 23 October 2009.

22.　Lee, Y.S.; Cho, S.B. Activity Recognition Using Hierarchical Hidden Markov Models on a Smartphone with 3D Accelerometer. In Proceedings of the 6th International Conference, Wroclaw, Poland, 23–25 May 2011; Volume 6678, pp. 460–467.

23.　Reddy, S.; Mun, M.; Burke, J.; Estrin, D.; Hansen, M.; Srivastava, M. Using mobile phones to determine transportation modes. *ACM Trans. Sens. Netw.* **2010**, *6*, 13. [CrossRef]

24.　Wang, J.; Chen, R.; Sun, X.; She, M.F.H.; Wub, Y. Recognizing Human Daily Activities from Accelerometer Signal. *Procedia Eng.* **2011**, *15*, 1780–1786. [CrossRef]

25.　Witowski, V.; Foraita, R.; Pitsiladis, Y.; Pigeot, I.; Wirsik, N. Using Hidden Markov Models to Improve Quantifying Physical Activity in Accelerometer Data—A Simulation Study. *PLoS ONE* **2014**, *9*, e114089. [CrossRef] [PubMed]

26. Trabelsi, D.; Mohammed, S.; Chamroukhi, F.; Oukhellou, L.; Amirat, Y. An unsupervised approach for automatic activity recognition based on hidden Markova model regression. *IEEE Trans. Autom. Sci. Eng.* **2013**, *3*, 829–335. [CrossRef]

27. Cvetković, B.; Luštrek, M.; Kaluža, B.; Gams, M. Semi-supervised Learning for Adaptation of Human Activity Recognition Classifier to the User. In Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI11), Barcelona, Spain, 16–22 July 2011; pp. 24–29.

28. Cvetković, B.; Kaluža, B.; Gams, M.; Luštrek, M. Adapting activity recognition to a person with Multi-Classifier Adaptive Training. *J. Ambient Intell. Smart Environm.* **2015**, *7*, 171–185.

29. Khan, A.M.; Lee, Y.-K.; Lee, S.Y.; Kim, T.-S. Human activity recognition via an accelerometer enabled-smartphone using kernel discriminant analysis. In Proceedings of the 5th International Conference on Future Information Technology, Busan, Korea, 21–23 May 2010; pp. 1–6.

30. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A Public Domain Dataset for Human Activity Recognition Using Smartphones. In Proceedings of the 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), Bruges, Belgium, 24–26 April 2013.

31. Young, S.; Evermann, G.; Gales, M.J.F.; Hain, T.; Kershaw, D.; Liu, X.; Moore, G.; Odell, J.; Ollason, D.; Povey, D.; et al. *The HTK Book*; Cambridge University Engineering Department: Cambridge, UK, 2006.

32. Chow, Y.L. Maximum mutual information estimation of HMM parameters for continuous speech recognition using the N-best algorithm. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP-90), Albuquerque, NM, USA, 3–6 April 1990; Volume 2, pp. 701–704.

33. Activity Recognition Experiment Using Smartphone Sensors. Available online: https://www.youtube.com/watch?v=XOEN9W05_4A (accessed on 30 August 2016).

34. Weissgerber, T.L.; Milic, N.M.; Winham, S.J.; Garovic, V.D. Beyond bar and line graphs: Time for a new data presentation paradigm. *PLoS Biol.* **2015**, *13*. [CrossRef] [PubMed]