

Short Papers

An Unsupervised Approach for Automatic Activity Recognition Based on Hidden Markov Model Regression

Dorra Trabelsi, Samer Mohammed, *Member, IEEE*,
Faïcel Chamroukhi, *Member, IEEE*, Latifa Oukhellou, and
Yacine Amirat, *Member, IEEE*

Abstract—Using supervised machine learning approaches to recognize human activities from on-body wearable accelerometers generally requires a large amount of labeled data. When ground truth information is not available, too expensive, time consuming or difficult to collect, one has to rely on unsupervised approaches. This paper presents a new unsupervised approach for human activity recognition from raw acceleration data measured using inertial wearable sensors. The proposed method is based upon joint segmentation of multidimensional time series using a Hidden Markov Model (HMM) in a multiple regression context. The model is learned in an unsupervised framework using the Expectation-Maximization (EM) algorithm where no activity labels are needed. The proposed method takes into account the sequential appearance of the data. It is therefore adapted for the temporal acceleration data to accurately detect the activities. It allows both segmentation and classification of the human activities. Experimental results are provided to demonstrate the efficiency of the proposed approach with respect to standard supervised and unsupervised classification approaches.

Note to Practitioners—This paper was motivated by the problem of automatic recognition of physical human activities using on-body wearable sensors in a health-monitoring context. Three wearable sensors (accelerometers) are placed at the chest, the right thigh and the left ankle of the subject. The studied activities concern the static ones, the dynamic ones as well as transitions between those activities. Since the goal is to recognize human activities from only the raw acceleration data, the acquired acceleration signals are seen as multidimensional time series with regime changes due to the changes of activities over time. The activity recognition problem is formulated as a problem of multidimensional time series segmentation. Segmenting the time series according to different unknown regimes over time is equivalent to classifying the acceleration data into one set of activities; each activity being associated with a regime. The proposed approach does not require any annotation of the raw accelerations by experts to learn the model parameters. However, it assumes that the number of activities is known. This assumption can be a constraint in a context of exploratory data mining where the aim is to automatically cluster a large amount of data into different group of activities. To tackle this limitation, a selection criterion can be used to determine the number of the groups of activities.

Index Terms—Activity recognition, hidden Markov model (HMM), multivariate regression, unsupervised learning, wearable computing.

Manuscript received April 10, 2012; revised September 23, 2012; accepted March 11, 2013. Date of publication May 15, 2013; date of current version June 27, 2013. This paper was recommended for publication by Associate Editor A. Colombo and Editor A. Bicchi upon evaluation of the reviewers' comments. This work was presented in part at the IEEE IJCNN 2012, Brisbane, Australia, 2012. (*Corresponding author: Y. Amirat.*)

D. Trabelsi, S. Mohammed, and Y. Amirat are with University Paris-Est Créteil (UPEC), LISSI, 94400 Vitry-Sur-Seine, France (e-mail: dorra.trabelsi@etu.u-pec.fr; samer.mohammed@u-pec.fr; amirat@u-pec.fr).

F. Chamroukhi is with University Sud Toulon-Var, LISIS, UMR 7296, Bâtiment R, 83957 La Garde Cedex, France (e-mail: Faïcel.Chamroukhi@univ-tln.fr).

L. Oukhellou is with University Paris-Est, IFSTTAR, GRETTIA, F-93166 Noisy-le-Grand, France (e-mail: latifa.oukhellou@ifsttar.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASE.2013.2256349

I. INTRODUCTION

The aging population has recently gained an increasing attention due to its socio-economic impact. By 2050, the number of people in the European Union aged 65 and above is expected to grow by 70% and the number of people aged over 80 by 170%.¹ This demographic change poses increasing challenges for healthcare services and their adaptation to the needs of this aging population. Facing this problem or reducing its effect would have a great societal impact by improving the quality of life and regaining people independence to make them active in society. The aim is therefore to facilitate the daily activity lives of elderly or dependent people at home, to increase their autonomy and to improve their safety. In fact, most elderly prefer to stay at home in the so-called “aging in place” [1]. The emergence of novel adapted technologies such as wearable and ubiquitous technologies is becoming a privileged solution to provide assistive services to humans, such as health monitoring, well being, security, etc. Among which, activity recognition has a wide range of promising applications in security monitoring as well as human machine interaction [2]. A large amount of work has been done in this active topic over the past decades; nevertheless it is still an open and challenging problem [3].

Several techniques have been used to quantify these activities such as video-based sensors [4], wearable-based sensors, environmental sensors and object sensors (smart phones, RFID, etc.). Recently, the Kinect sensor that contains both RGB and Infra red (IR) cameras has been released by Microsoft [5]. This sensor has been used for human activity detection and recognition [6]. Even though this sensor has several advantages such as low cost, depth information and ability to operate in the day as well as night, it has however some disadvantages [7]. In fact, the Kinect sensor has limited field of view, poor performances in natural lighting, dependence on surface texturing and occlusion problem in cluttered environment. The wearable inertial-based system used in this study for activity recognition overcomes the above mentioned limitations. Recently, the use of wearable sensors-based systems for activity recognition has gained more attention on a large number of technological fields such as navigation, monitoring and control of aircrafts [8], [9], medical application [10], [11], localization and robots [12], [13]. Among the inertial sensors used for activity recognition, the accelerometers are the most commonly used [14]. They have shown satisfactory results to measure the human activities in both laboratory/clinical and free-living environment settings [15]. In addition, the latest advances in Microelectromechanical Systems (MEMS) technology have greatly promoted the use of accelerometers thanks to the considerable reduction in size, cost and energy consumption. Early studies in activity recognition used uniaxial accelerometers, while recent studies use mainly triaxial accelerometers [16], [17].

II. RELATED WORK ON HUMAN ACTIVITY RECOGNITION

Regarding the human activity classification, one can make the distinction between supervised and unsupervised classification approaches. Supervised classification techniques consist of inferring a decision rule from labeled training data. The use of the supervised activity classification approaches has shown promising results [18]. Some supervised approaches have enhanced the activity recognition

¹http://ec.europa.eu/health-eu/my_health/elderly/

process performances by using spatio-temporal information [19]. Regarding the algorithms used in the supervised context, one can cite k -Nearest Neighbor (k -NN) algorithm [20], multiclass Support Vector Machines (SVM) [21], and Artificial Neural Networks (ANN) including both MultiLayer Perceptron (MLP) [22], [23] and Radial Basis Function (RBF) networks [24].

Nevertheless, the collection of sufficient amounts of labeled data for a various and rich set of free-living activities may be sometimes difficult to achieve and computationally expensive [25]. On the other hand, unsupervised classification techniques try to directly construct models from unlabeled data either by estimating the properties of their underlying probability density (called density estimation) or by discovering groups of similar examples (called clustering). The unsupervised learning techniques are of particular interest for an exploratory analysis of large amounts of unlabeled data. They can also consist of a preliminary task to further run a supervised classifier based on the obtained partition of the data. The use of an unsupervised approach may be needed in such a context of activity recognition when it is difficult to have labels for the data.

Regarding the approaches used in the unsupervised context, one can cite the well-known k -Means algorithm [26], the Gaussian Mixture Models (GMM) approach [27] and the one based on Hidden Markov Model (HMM) [28], [29] or HMM with GMM emission probabilities [30]. Both the GMM and the HMM approaches use the EM algorithm [31].

The HMM has shown good results in earlier exploratory studies thanks to their main advantage of suitability to model sequential data which is the case of monitoring human activities. Indeed, the acceleration data are measured over time during physical human activities of a person and are therefore sequential over time. The EM algorithm [31] (also called Baum-Welch [32]) in the context of HMM is particularly adapted for unsupervised learning.

In this study, an unsupervised approach for human activity recognition is proposed. It combines an HMM-based model with the use of acceleration data acquired during sequences of different human activities. More specifically, the proposed approach is based on a Hidden Markov Model in a multiple regression context and will be denoted by MHMMR.

As the sequences of acceleration data consist in multidimensional time series where each dimension is an acceleration, the activity recognition problem is therefore formulated through the proposed MHMMR model as the one of joint segmentation of multidimensional time series; each segment is associated with an activity. In the proposed model, each activity is represented by a regression model and the switching from one activity to another is governed by a hidden Markov chain. The MHMMR parameters are learned in an unsupervised way from unlabeled raw acceleration data acquired during human activities.

The most likely sequence of activities is then estimated using the Viterbi algorithm [33]. The proposed technique is then evaluated on real-world acceleration data collected from three sensors placed at the chest, the right thigh, and the left ankle of the subject.

This study is an extension of the paper [34], where additional technical implementations are shown: Twelve activities and transitions are studied and performances of the proposed approach are evaluated and compared to those of some well known unsupervised and supervised techniques for activity recognition.

This paper is organized as follows. Section III presents the experimental protocol and the data acquisition platform. Section IV presents the proposed model and its unsupervised parameter estimation technique from unlabeled acceleration data. In Section V, the performances of the proposed approach are evaluated and compared to those of some well known unsupervised and supervised techniques for activity recognition.

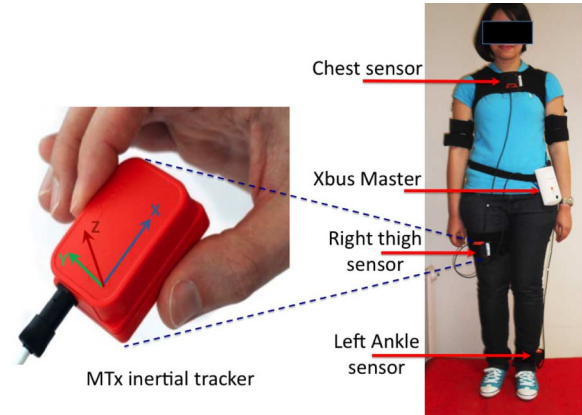


Fig. 1. MTx-Xbus inertial tracker and sensors placement.

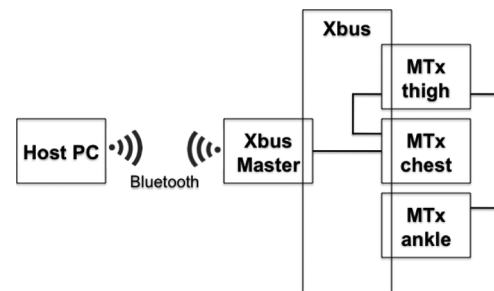


Fig. 2. Data gathering from the MTx-Xbus acquisition system.

III. DATA COLLECTION

In this study, human activities are classified using three sensors placed at the chest, the right thigh, and the left ankle, respectively, as shown in Fig. 1. Sensors placement is chosen to represent predominantly upper-body activities such as standing up, sitting down, etc., and predominantly lower body activities such as walking, stair ascent, stair descent, etc. The sensor's placement guarantees at the same time less constraint and better comfort for the wearer. The attachment of the sensors to the human body should be well fitted and secured (Fig. 1). These sensors consist of three MTx 3-DOF inertial trackers developed by Xsens Technologies [35]. Each MTx unit consists of a triaxial accelerometer measuring the acceleration in the 3-D space (with a dynamic range of $\pm 5g$, where g represents the gravitational constant). Our experiences show also that the measured ankle-sensor accelerations during the different activities do not exceed the limit of $\pm 5g$. The sampling frequency is set to 25 Hz, which is sufficient and larger than 20 Hz the required frequency to assess daily physical activity [36]. The sensors were fixed on the subject with the help of an assistant before the beginning of the measurement operation. Raw acceleration data are therefore collected over time when performing the activities. The MTx units are connected to a central unit called Xbus Master that is attached to the subject's belt. Fig. 2 shows the data gathering process from the Xbus-MTx acquisition system to the host PC. The Xbus Master is directly connected to the chest MTx unit while the remaining MTx units (thigh and ankle) are connected in series. Data transmission between the Xbus Master and the PC is carried out through a Bluetooth wireless link.

The experiments were performed at the LISSI Lab, University of Paris-Est Créteil (UPEC), by six different healthy subjects of different ages (who are not the researchers) in the office environment. In order to gather various and rich dataset, the recruited volunteer subjects have been chosen in a given margin of age (25–30) and weight (55–70) kg.

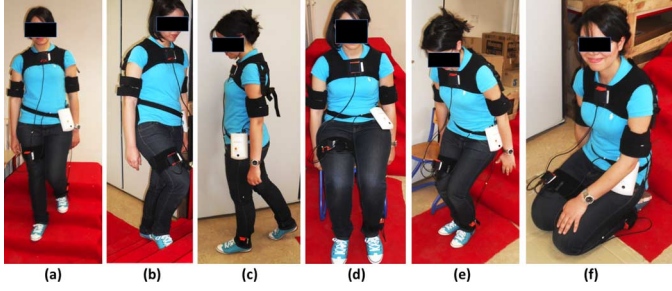


Fig. 3. Examples of some considered activities. (a) Climbing stairs down. (b) Climbing stairs Up. (c) Walking. (d) Sitting. (e) Standing up. (f) Sitting on the ground.

Twelve activities and transitions were studied and are listed as follows: Stairs down (A_1) – Standing (A_2) – Sitting down (A_3) – Sitting (A_4) – From sitting to sitting on the ground (A_5) – Sitting on the ground (A_6) – Lying down (A_7) – Lying (A_8) – From lying to sitting on the ground (A_9) – Standing up (A_{10}) – Walking (A_{11}) – Stairs up (A_{12}). The activities were chosen to have an appropriate representation of everyday activities involving different parts of the body (Fig. 3). The recognized activities and transition differ in duration and intensity level. Note that the activities A_3 , A_5 , A_7 , A_9 , and A_{10} represent dynamic transitions between static activities. Each subject was asked to perform the twelve activities in his own style and was not restricted on how the activities should be performed but only with the sequential activities order. In addition, the duration of each activity is not restricted to be the same as it may vary from one subject to another.

With three MTx sensor units, each one with a tri-axial accelerometer, a total of nine accelerations are therefore measured and recorded overtime for each activity. Since the goal is to recognize human activities from only the raw acceleration data, the acquired acceleration signals can be seen as multidimensional time series (of dimension 9) with regime changes due to the changes of activities over time. The activity recognition problem can therefore be formulated as a problem of multidimensional time series segmentation. Indeed, segmenting the time series according to different unknown regimes over time is equivalent to classifying the acceleration data into one set of activities; each activity being associated with a regime. This will be detailed in the next section that is dedicated to the proposed Hidden Markov Model Regression (HMMR) approach.

IV. SEGMENTATION WITH MULTIPLE HIDDEN MARKOV MODEL REGRESSION – MHMMR

In this section, the problem of activity recognition (classification) is formulated as the one of joint segmentation of multidimensional time series. Indeed, the acceleration data are presented as multidimensional time series presenting various regime changes. In such context, the goal is to provide an automatic partition of the data into different segments (regimes), each segment being considered afterwards as an activity.

Various modeling approaches of time series presenting regime changes have been proposed in literature. One can cite, in particular, the piecewise regression as one of the most adapted modeling approaches [37], [38]. The piecewise model has been applied in many domains including finance, engineering, economics, and bioinformatics [39]. In the piecewise regression model [38], data are partitioned into several segments, each segment being characterized by its mean polynomial curve and its variance. However, the parameter estimation in such method requires the use of dynamic programming algorithm [40], [41], which may be computationally expensive especially for time series with large number of observations. Moreover,

the standard piecewise regression model usually assumes that noise variance is uniform for all the segments (homoskedastic model). An alternative approach extended in this paper is based on HMMR [42]. This approach can be seen as an extension of the standard HMM [29] to regression analysis. Each regime is described by a regression model rather than a simple constant mean over time, while preserving the Markov process modeling for the sequence of unknown (hidden) activities. Indeed, standard HMM-based approaches use simple Gaussian densities as density of observation. However, in the HMM regression context, each observation is assumed to be a noisy polynomial function to better model very structured data as the acceleration data. The approach we propose further extends the HMM model to a multiple regression setting. This is due to the fact that the observed acceleration data are multidimensional. In the following, the Hidden Markov Regression Model for time series modeling is used by formulating its basic and multiple regression setting. In this framework, each observation, denoted by \mathbf{y}_i , represents the i th acceleration measurement while the associated state (class), denoted by z_i , represents its corresponding activity.

A. General Description of the Multiple Hidden Markov Model Regression (HMMR)

In Hidden Markov Model Regression (HMMR), each time series is represented as a sequence of observed univariate variables (y_1, y_2, \dots, y_n) , where the observation y_i at time t_i is assumed to be generated by the following regression model [42]:

$$y_i = \beta_{z_i}^T \mathbf{t}_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1), \quad (i = 1, \dots, n) \quad (1)$$

where $z_i \in \{1, \dots, K\}$ is a hidden discrete-valued variable. In this application case, z_i represents the hidden class label (activity) of each acceleration data point and K corresponds to the number of considered activities. The variable z_i controls the switching from one polynomial regression model associated to one activity, to another of K models at time t_i . The vector $\beta_{z_i} = (\beta_{z_i 0}, \dots, \beta_{z_i p})^T$ is the one of regression coefficients of the p -order polynomial regression model z_i and σ_{z_i} is its corresponding standard deviation, $\mathbf{t}_i = (1, t_i, t_i^2, \dots, t_i^p)^T$ is the $p + 1$ dimensional covariate vector at time t_i and the ϵ_i 's are standard Gaussian variables representing an additive noise. The HMMR assumes that the hidden sequence $\mathbf{z} = (z_1, \dots, z_n)$ is a homogeneous Markov chain of first-order parameterized by the initial state distribution π and the transition matrix \mathbf{A} . It can be shown that, conditionally on a regression model k ($z_i = k$), y_i has a Gaussian distribution with mean $\beta_k^T \mathbf{t}_i$ and variance σ_k^2 . Regarding the multiple regression case, the model can be formulated as follows:

$$\begin{aligned} y_i^{(1)} &= \beta_{z_i}^{(1)T} \mathbf{t}_i + \sigma_{z_i}^{(1)} \epsilon_i \\ y_i^{(2)} &= \beta_{z_i}^{(2)T} \mathbf{t}_i + \sigma_{z_i}^{(2)} \epsilon_i \\ &\vdots \\ y_i^{(d)} &= \beta_{z_i}^{(d)T} \mathbf{t}_i + \sigma_{z_i}^{(d)} \epsilon_i \end{aligned} \quad (2)$$

where d represents the dimension of the time series (sequence) and the latent process \mathbf{z} simultaneously governs all the univariate time series components. The model (2) can be rewritten in a matrix form as follows:

$$\mathbf{y}_i = \mathbf{B}_{z_i}^T \mathbf{t}_i + \mathbf{e}_i \quad ; \quad \mathbf{e}_i \sim \mathcal{N}(\mathbf{0}, \Sigma_{z_i}), \quad (i = 1, \dots, n) \quad (3)$$

where $\mathbf{y}_i = (y_i^{(1)}, \dots, y_i^{(d)})^T$ is the i th observation of the time series in \mathbb{R}^d , $\mathbf{B}_k = [\beta_k^{(1)}, \dots, \beta_k^{(d)}]$ is a $(p + 1) \times d$ dimensional matrix of the multiple regression model parameters associated with the regime (class) $z_i = k$ and Σ_{z_i} its corresponding covariance matrix.

The Multiple HMMR model is therefore fully parameterized by the parameter vector $\theta = (\pi, \mathbf{A}, \mathbf{B}_1, \dots, \mathbf{B}_K, \Sigma_1, \dots, \Sigma_K)$. The next subsection gives the parameter estimation technique by maximizing the observed data likelihood through the Expectation-Maximization (EM) algorithm. The parameter vector θ is estimated using the well-known maximum likelihood method thanks to its very well-known attractive properties of consistency, asymptotic normality, and efficiency. Indeed, in our experiments, a considerable number of data points is acquired during time, which makes the sample size suitable to take advantage of the limiting properties of the maximum likelihood estimator. The log-likelihood to be maximized in this case is written as follows:

$$\begin{aligned} \mathcal{L}(\theta) &= \log p(\mathbf{y}_1, \dots, \mathbf{y}_n; \theta) \\ &= \log \sum_{\mathbf{z}} p(z_1; \pi) \prod_{i=2}^n p(z_i | z_{i-1}; \mathbf{A}) \prod_{i=1}^n \mathcal{N}(\mathbf{y}_i; \mathbf{B}_{z_i}^T \mathbf{t}_i, \Sigma_{z_i}). \end{aligned} \quad (4)$$

Since this log-likelihood cannot be maximized directly, this can be performed using the EM algorithm [31], [43], that is known as the Baum-Welch algorithm in the HMM context [32], [29]. This algorithm alternates between the two following steps.

1) *E-Step*: This step computes the conditional expectation of the complete-data log-likelihood given the observed data \mathbf{Y} , time \mathbf{t} and a current parameter estimation denoted by $\theta^{(q)}$

$$Q(\theta, \theta^{(q)}) = \mathbb{E} \left[\log p(\mathbf{Y}, \mathbf{z} | \mathbf{t}; \theta) | \mathbf{Y}, \mathbf{t}; \theta^{(q)} \right]. \quad (5)$$

It can be easily shown that this step only requires the calculation of:

- the posterior probability

$$\tau_{ik}^{(q)} = p(z_i = k | \mathbf{Y}, \mathbf{t}; \theta^{(q)}) \quad (6)$$

$\forall i = 1, \dots, n$ and $k = 1, \dots, K$ which is the posterior probability that \mathbf{y}_i originates from the k th polynomial regression model given the whole observation sequence and the current parameter estimation $\theta^{(q)}$;

- the joint posterior probability of the state k at time i and the state ℓ at time $i-1$ given the whole observation sequence and the current parameter estimation $\theta^{(q)}$, that is

$$\xi_{i\ell k}^{(q)} = p(z_i = k, z_{i-1} = \ell | \mathbf{Y}, \mathbf{t}; \theta^{(q)}) \quad (7)$$

$\forall i = 2, \dots, n$ and $k, \ell = 1, \dots, K$.

These posterior probabilities are computed by the forward-backward procedures in the same way as for a standard HMM [29]. More calculation details on this step can be found in [29].

2) *M-Step*: In this step, the value of the parameter θ is updated by computing the parameter $\theta^{(q+1)}$ that maximizes the conditional expectation (5) with respect to θ . It can be shown that this maximization leads to the following updating rules. The updates of the parameters governing the hidden Markov chain \mathbf{z} are the ones of a standard HMM and are given by:

$$\pi_k^{(q+1)} = \tau_{1k}^{(q)} \quad (8)$$

$$\mathbf{A}_{\ell k}^{(q+1)} = \frac{\sum_{i=2}^n \xi_{i\ell k}^{(q)}}{\sum_{i=2}^n \tau_{ik}^{(q)}}. \quad (9)$$

Updating the regression parameter consists of performing K weighted multiple polynomial regressions. The regression parameter matrices updates are given by:

$$\begin{aligned} \mathbf{B}_k^{(q+1)} &= \left[\sum_{i=1}^n \tau_{ik}^{(q)} \mathbf{t}_i \mathbf{t}_i^T \right]^{-1} \sum_{i=1}^n \tau_{ik}^{(q)} \mathbf{t}_i \mathbf{y}_i^T \\ &= (\mathbf{X}^T \mathbf{W}_k^{(q)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_k^{(q)} \mathbf{Y} \end{aligned} \quad (10)$$

where $\mathbf{W}_k^{(q)}$ is a $n \times n$ diagonal matrix of weights whose diagonal elements are the posterior probabilities $(\tau_{1k}^{(q)}, \dots, \tau_{nk}^{(q)})$ and \mathbf{X} is the $n \times (p+1)$ regression matrix given by $(\mathbf{t}_1, \dots, \mathbf{t}_n)^T$. The updating rule for the covariance matrices is written as a weighted variant of the estimation of a multivariate Gaussian density with the polynomial mean $\mathbf{B}_k^{T(q+1)} \mathbf{t}_i$ such as:

$$\begin{aligned} \Sigma_k^{(q+1)} &= \frac{1}{\sum_{i=1}^n \tau_{ik}^{(q)}} \sum_{i=1}^n \tau_{ik}^{(q)} (\mathbf{y}_i - \mathbf{B}_k^{T(q+1)} \mathbf{t}_i)^T (\mathbf{y}_i - \mathbf{B}_k^{T(q+1)} \mathbf{t}_i) \\ &= \frac{1}{\sum_{i=1}^n \tau_{ik}^{(q)}} (\mathbf{Y} - \mathbf{X} \mathbf{B}_k^{(q+1)})^T \mathbf{W}_k^{(q)} (\mathbf{Y} - \mathbf{X} \mathbf{B}_k^{(q+1)}). \end{aligned} \quad (11)$$

V. RESULTS AND DISCUSSIONS

This section presents experiments carried out to validate the two main ideas explored throughout this paper, i.e., the segmentation and the classification of the human activity from raw acceleration data using a MHMMR approach within an unsupervised learning framework.² Series of experiments were conducted to evaluate the performance of the proposed approach and also to perform comparisons with well-known unsupervised and supervised classification approaches.

A. Performance Evaluation

Given a set of 9-acceleration data from three tri-axial accelerometer modules mounted on the chest, the right thigh, and the left ankle, the proposed approach allows both segmentation and classification of the 12 activities. Each obtained segment is indeed considered as an activity, achieving thus a classification task. We chose to take as a ground truth about the class of an activity, labeling obtained thanks to an expert. While the different subjects were performing the sequence of activities, an independent operator was asked to annotate the activities, thus providing a labeling of the dataset.³ The provided partition is indeed matched to the true labels (ground truth) by evaluating all the possible label switchings. The label switching leading to the minimum error rate is selected as the best class prediction. For the supervised classification approaches, data labels were used to both train and test the models. In this case, the performance was estimated through a 10-fold cross-validation procedure. Regarding the classification problem, confusion matrices between the annotated classes and the estimated classes for all the subjects in the database are computed. The criteria used to evaluate the performance of an approach are the correct classification rate and the prediction accuracy in terms of precision and recall.

In the following, the results of the MHMMR approach obtained on real acceleration data of human activities are first detailed, then they are compared to those of standard unsupervised and supervised classification approaches.

B. Classification Performance of the MHMMR

The following experiments were conducted to qualitatively assess the performances of the proposed approach in terms of automatic segmentation of human activity on the basis of raw acceleration signals. From the sequence of nine observed variables $\mathbf{y}_i = (y_i^{(1)}, \dots, y_i^{(9)})$ at each time step i for $i = 1, \dots, n$ corresponding to the three-axis accelerations measured by the three sensors, the MHMMR is used to

²Note that, in this study, the raw acceleration data are directly used without any feature extraction. Indeed, in many areas of application, a feature extraction step is needed before running the classifier and may itself lead to an additional computational cost, which can be penalizing in real-time applications.

³Note that the labels were not used to train unsupervised models; they were only used afterwards for the evaluation of classification errors.

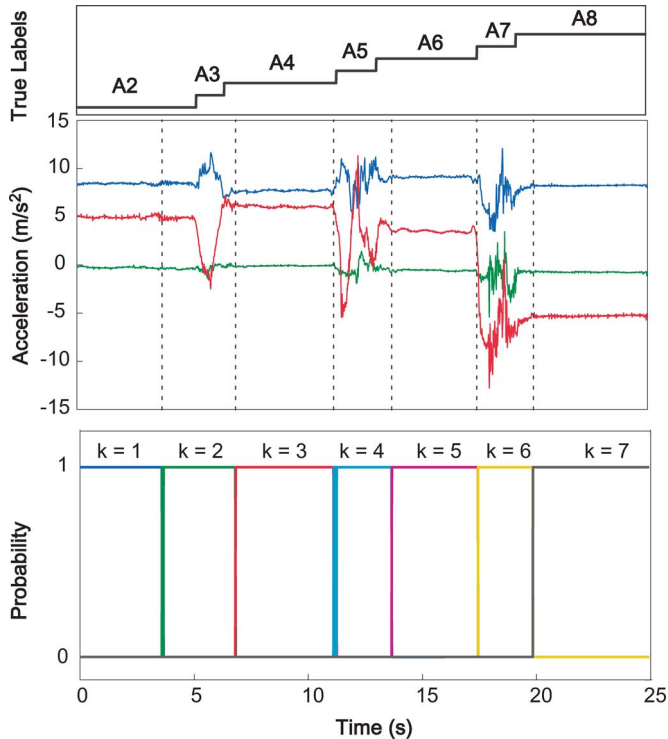


Fig. 4. MHMMR segmentation for the sequence (Standing A2 – Sitting down A3 – Sitting A4 – From sitting to sitting on the ground A5 – Sitting on the ground A6 – Lying down A7- Lying A8) for the seven classes $k = (1, \dots, 7)$.

identify the latent sequence $\mathbf{z} = (z_1, \dots, z_n)$ corresponding to the 12 activities. The number of classes K is fixed to 12 and the order of regression p is fixed empirically to three as it gives the best performance among several values of p . Model parameters are estimated from the data using the algorithm detailed in Section IV-A. Figs. 4 and 5 show the performance of the proposed method to segment the two following sequences:

- Sequence 1: Standing – Sitting down – Sitting – From sitting to sitting on the ground – Sitting on the ground – Lying down – Lying.
- Sequence 2: Standing – Walking – Climbing up stairs – Standing.

These figures represent the evolution of the acceleration data and the corresponding posterior probabilities for the two different sequences. Note that the posterior probability is the probability that a sample i will be generated by the regression model k given the whole sequence of observations $(\mathbf{y}_1, \dots, \mathbf{y}_n)$. It can be observed that the obtained sequences are interesting and promising despite some confusion between activities such as (A11, A12).

Table II shows that the MHMMR gives 91.4% as a mean correct classification rate averaged over all observations. It highlights the potential benefit of the proposed approach in terms of automatic segmentation and classification of human activity. Both the transitions and the stationary activities are well identified. Exhaustively, Table I gives the percentage of precision and recall for each activity. Indeed, one can observe that static activities (A2, A4, A6, and A7) are easier to recognize than dynamic activities (A1, A11, A12). In order to focus on the efficiency ratio of the three sensors used for activity recognition, the MHMMR algorithm has been evaluated using data from only two sensors. The classification results, given in Table II, show as expected, that the percentage of correctly classified instances decreases with the number of data sources. The worst result is obtained when the sensor placed at the thigh is not taken into account.

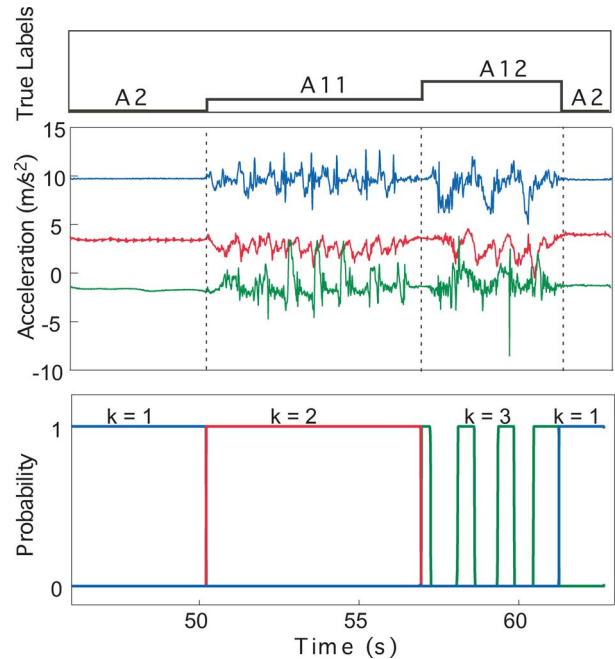


Fig. 5. MHMMR segmentation for the sequence (Standing A2 – Walking A11 – Climbing up stairs A12 – Standing A2) for the three classes $k = 1, \dots, 3$.

C. Comparison With Unsupervised and Supervised Classification Approaches

Correct classification rates and the standard deviations obtained with standard unsupervised and supervised classification approaches as well as the MHMMR approach are given in Tables III and IV. Compared to standard unsupervised classifiers, the proposed MHMMR outperforms them since it provides a classification rate of 91.4%, while only 60%, 72%, and 84% of instances are well classified with, respectively, the k -Means, the GMM, and the standard HMM approaches. Notice that, the GMM and k -means approaches are not well suitable for this kind of longitudinal data. In Table IV, it can be observed that the k -NN ($k = 1$) gives the highest classification rates with 95.8%, followed by the Random Forest with 93.5%. Then, the SVM gives 88.1% and the MLP gives 83.1%. However, the Naive Bayes gives the lowest classification rate with 80.6%. Table IV shows also that the k -NN ($k = 1$) has the best classification algorithm in terms of prediction accuracy since it achieves 95.9% of precision and recall. Compared to standard supervised classification techniques (using class labels), these results are very encouraging since the proposed approach performs in an unsupervised way. The main errors are due to the confusions located in transition segments. This is due to the fact that the transitions lengths are much shorter than the activities ones. Since the confusion matrix was computed using real labels supplied by a human expert, the obtained labels may not correspond perfectly to the expert labels, particularly, during transitions. Indeed, it is difficult to have the ground truth of the limit between an activity and a transition. Furthermore, the aforementioned supervised classification approaches require a labeled collection of data to be trained. Besides, they do not explicit the temporal dependence in their model formulation as they assume an independent hypothesis for the data; the data are treated as several realizations in the multidimensional space (\mathbb{R}^d) without considering possible dependencies between the activities. Moreover, it can be noticed that assigning a new sample to a class using the k -NN approach requires the computation of as many distances as there are examples in the dataset, which may lead to a significant computation time. Using the proposed approach, classification

TABLE I
RECALL VERSUS PRECISION OF THE MHMMR

Class	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
Precision (%)	71.9	96.4	78.4	95.7	92.3	98.9	97.6	92.5	82.6	82.6	83.2	95.6
Recall (%)	95	87.8	83.4	94	97.3	94.6	95.4	90.9	98.5	92.2	98.1	82.3

TABLE II
EFFECTS OF REDUCING THE NUMBER OF SENSORS WHEN USING MHMMR

Sensors	Percentage of correctly classified instances
Chest, thigh, ankle	91.4% \pm 1.65
Chest, ankle	83.9% \pm 1.98
Chest, thigh	86.2% \pm 2.03
Thigh, ankle	84% \pm 2.21

TABLE III
COMPARISON OF THE PERFORMANCE IN TERMS OF CORRECT CLASSIFICATION, RECALL AND PRECISION OF THE FOUR UNSUPERVISED CLASSIFIERS

	Correct Classification (%)	Precision (%)	Recall (%)
<i>k</i> -Means	60.2 \pm 2.48	60.4	59.8
GMM	72.3 \pm 2.05	71.8	73.5
HMM	84.1 \pm 1.84	83.8	84
MHMMR	91.4 \pm 1.65	89	95.6

TABLE IV
COMPARISON OF THE PERFORMANCE IN TERMS OF CORRECT CLASSIFICATION, RECALL AND PRECISION OF THE FIVE SUPERVISED CLASSIFIERS

	Correct Classification (%)	Precision (%)	Recall (%)
Naive Bayes	80.6 \pm 0.91	80.9	80.6
MLP	83.1 \pm 0.45	82.8	83.2
SVM	88.1 \pm 1.32	87.6	88.3
<i>k</i> -NN	95.8 \pm 0.32	95.9	95.9
Random Forest	93.5 \pm 0.78	93.5	93.5

needs the computation of the posterior probabilities, as many as there are activities. On the other hand, comparison with the unsupervised classification approaches (*k*-Means and the GMM) and the standard HMM shows that the proposed method gives relatively a high rate and better performances.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a statistical approach based on hidden Markov models in a regression context for the joint segmentation of multivariate time series of human activities. It is based upon the use of raw accelerometer data acquired from body mounted inertial sensors in a health-monitoring context. The main advantage of the proposed approach comes from the fact that the statistical model takes into account the regime changes over time through the hidden Markov chain; each regime being interpreted as an activity (a class). Furthermore, learning with this statistical model is performed in an unsupervised way using unlabeled examples only; parameter estimates are computed by maximizing a likelihood criterion, using a dedicated EM algorithm. Considering human activity recognition within an unsupervised learning framework can be particularly interesting within an exploratory data-mining context in order to automatically cluster a large amount of unlabeled acceleration data into different groups of activity. The comparison with well-known supervised classification approaches shows that the proposed method is competitive even when performed in an unsupervised way. This work can be extended in several directions, namely, integrating the model into a Bayesian context to better control the model complexity via choosing suitable prior distributions on the models parameters. Then, and perhaps more interestingly, another step to explore is to build a fully non Bayesian nonparametric

model which will be useful for any kind of complex activities and in which the number of activities will not have to be fixed. In terms of application, a promising perspective in a rehabilitation context would be to use the proposed approach for recognizing in an unsupervised framework the undesirable compensatory physical behaviors observed with stroke and injury patients.

REFERENCES

- [1] B. Kaluza, V. Mirchevska, E. Dovgan, M. Lustrek, and M. Gams, "An agent-based approach to care in independent living," in *Int. Joint Conf. Ambient Intell. (Aml-2010)*, Malaga, Spain, 2010, vol. 6439, Lecture Notes in Comput. Sci., pp. 177–186.
- [2] C.-H. Lu and L.-C. Fu, "Robust location-aware activity recognition using wireless sensor network in an attentive home," *IEEE Trans. Autom. Sci. Eng.*, vol. 6, no. 4, pp. 598–609, Oct. 2009.
- [3] H. Jinhui and V. B. Nikolaos, "Fast human activity recognition based on structure and motion," *Pattern Recognit. Lett.*, vol. 32, no. 14, pp. 1814–1821, 2011.
- [4] O. Brdiczka, M. Langet, J. Maisonnasse, and J. L. Crowley, "Detecting human behavior models from multimodal observation in a smart home," *IEEE Trans. Autom. Sci. Eng.*, vol. 6, no. 4, pp. 588–597, Oct. 2009.
- [5] *Kinect for Xbox 360*. Redmond, WA, USA: Microsoft Corp., [Online]. Available: <http://en.wikipedia.org/wiki/Kinect>
- [6] J. Sung, C. Ponce, B. Selman, and A. Saxena, "Unstructured human activity detection from RGBD images," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2012, pp. 842–849.
- [7] T. Gill, D. T. Anderson, R. H. Luke, and J. M. Keller, "A system for change detection and human recognition in voxel space using the Microsoft Kinect sensor," in *Proc. IEEE Appl. Imagery Pattern Recognit. (AIPR)*, 2011, pp. 1–8.
- [8] D. MacKenzie, *Inventing Accuracy A Historical Sociology of Nuclear Missile Guidance*. Cambridge, MA, USA: MIT Press, 1991.
- [9] M. Carminati, G. Ferrari, M. Sampietro, and R. Grasseti, "Fault detection and isolation enhancement of an aircraft attitude and heading reference system based on MEMS inertial sensor," *Procedia Chem.*, vol. 1, no. 1, pp. 509–512, 2009.
- [10] E. Jovanov, A. Milenkovic, C. Otto, and P. C. DeGroot, "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," *J. Neuro. Eng.Rehab.*, vol. 2, no. 6, 2005.
- [11] W. H. Wu, A. A. T. Bui, M. A. Batalin, D. Liu, and W. J. Kaiser, "Incremental diagnosis method for intelligent wearable sensor system," *IEEE Trans. Inf. Technol.B*, vol. 11, no. 5, pp. 553–562, Sep. 2007.
- [12] B. Barshan and H. F. Durrant-Whyte, "Evaluation of a solid-state gyroscope for robotics applications," *IEEE Trans. Instrum. Meas.*, vol. 44, no. 1, pp. 61–67, Feb. 1995.
- [13] C. W. Tan and S. Park, "Design of accelerometer-based inertial navigation systems," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 6, pp. 2520–2530, Dec. 2005.
- [14] L. Jiayang, Z. Lin, J. Wickramasuriya, and V. Vasudevan, "Uwave: Accelerometer-based personalized gesture recognition and its applications," *Pervasive and Mobile Computing*, pp. 657–675, 2009.
- [15] M. J. Mathie, B. G. Celler, N. H. Lovell, and A. C. F. Coster, "Classification of basic daily movements using a triaxial accelerometer," *Med. Biol. Eng. Comput.*, vol. 42, pp. 679–687, 2004.
- [16] N. Noury, P. Barralon, G. Virone, P. Boissy, M. Hamel, and P. Rumeau, "A smart sensor based on rules and its evaluation in daily routines," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Bio. Soc., (EMBS)*, Cancun, Mexico, 2003, pp. 3286–3289.
- [17] J. Y. Yang, J. S. Wang, and Y. P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," *Pattern Recognit. Lett.*, vol. 29, no. 16, pp. 2213–2220, 2008.
- [18] K. Altun, B. Barshan, and O. Tuncel, "Comparative study on classifying human activities with miniature inertial and magnetic sensors," *Pattern Recognit.*, vol. 43, no. 10, pp. 3605–3620, 2010.

- [19] A. H. Khalili and H. Aghajan, "Multiview activity recognition in smart homes with spatio-temporal features," in *Proc. 4th ACM/IEEE ICDSC*, USA, 2010, pp. 142–149.
- [20] C. L. Liu, C. H. Lee, and P. M. Lin, "A fall detection system using k-nearest neighbor classifier," *Expert Syst. Applicat.*, vol. 37, pp. 7174–7181, 2010.
- [21] H. Qian, Y. Mao, W. Xiang, and Z. Wang, "Recognition of human activities using SVM multi-class classifier," *Pattern Recognit. Lett.*, vol. 31, no. 2, pp. 100–111, 2010.
- [22] L. Dong, L. Che, L. Sun, and Y. Wang, "Effects of non-parallel combs on reliable operation conditions of capacitive inertial sensor for step and shock signals," *Sens. Actuators A: Phys.*, vol. 121, no. 2, pp. 395–404, 2005.
- [23] J. Y. Yang, J. S. Wang, and Y. P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," *Pattern Recognit. Lett.*, vol. 29, pp. 342–350, 2008.
- [24] G. S. Faber, I. Kingma, S. M. Bruijn, and J. H. Van Dieen, "Optimal inertial sensor location for ambulatory measurement of trunk inclination," *J. Biomechanics*, vol. 42, no. 14, pp. 2406–2409, 2009.
- [25] B. Cvetkovic, M. Lustrek, B. Kaluza, and M. Gams, "Semi-supervised learning for adaptation of human activity recognition classifier to the user," in *Int. Joint Conf. Artif. Intell. (STAMI)*, 2011, pp. 24–29.
- [26] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, second ed. New York, NY, USA: Wiley, 2000.
- [27] F. R. Allen, E. Ambikairajah, N. H. Lovell, and B. G. Celler, "Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models," *Physiol. Meas.*, vol. 27, no. 10, pp. 935–951, 2006.
- [28] J. F. S. Lin and D. Kulic, "Automatic human motion segmentation and identification using feature guided HMM for physical rehabilitation exercises," in *IEEE/RSJ Int. Workshop Conf. Intelligent Robots and Systems (IROS), Robot. Neurology Rehab.*, 2011.
- [29] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [30] A. Mannini and A. Sabatini, "Machine learning methods for classifying human physical activity from on-body accelerometers," *Sensors*, vol. 10, pp. 1154–1175, 2010.
- [31] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statist. Soc.*, vol. B, 39, no. 1, pp. 1–38, 1977.
- [32] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions," *Ann. Math. Statist.*, vol. 41, pp. 164–171, 1970.
- [33] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. 13, no. 2, pp. 260–269, Apr. 1967.
- [34] D. Trabelsi, S. Mohammed, L. Oukhellou, and Y. Amirat, "Activity recognition using body mounted sensors: An unsupervised learning based approach," in *IEEE World Congr. Computat. Intell. – Int. Joint Conf. Neural Networks (IJCNN)*, Brisbane, Australia, Jun. 10–15, 2012, pp. 1–7.
- [35] [Online]. Available: <http://www.xsens.com>
- [36] C. Bouten, K. Koekkoek, M. Verduin, R. Kodde, and J. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 3, pp. 136–147, Mar. 1997.
- [37] V. E. McGee and W. T. Carleton, "Piecewise regression," *J. Amer. Statist. Assoc.*, vol. 65, pp. 1109–1124, 1970.
- [38] V. L. Brailovsky and Y. Kempner, "Application of piecewise regression to detecting internal structure of signal," *Pattern Recognit.*, vol. 25, no. 11, pp. 1361–1370, 1992.
- [39] F. Picard, S. Robin, E. Lebarbier, and J. J. Daudin, "A segmentation/clustering model for the analysis of array CGH data," *Biometrics*, vol. 63, no. 3, pp. 758–766, 2007.
- [40] R. Bellman, "On the approximation of curves by line segments using dynamic programming," *Commun. Assoc. Comput. Mach. (CACM)*, vol. 4, no. 6, p. 284, 1961.
- [41] H. Stone, "Approximation of curves by line segments," *Math. Comput.*, vol. 15, no. 73, pp. 40–47, 1961.
- [42] M. Fridman, "Hidden Markov model regression," Univ. Minnesota, Inst. Mathematics, Minneapolis, MN, Tech. Rep., 1993.
- [43] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York, NY, USA: Wiley, 1997.