2013 IEEE International Conference on Robotics and Automation (ICRA)
Karlsruhe, Germany, May 6-10, 2013

Analysis of Human Behavior Recognition Algorithms based on Acceleration Data

Barbara Bruno, Fulvio Mastrogiovanni, Antonio Sgorbissa, Tullio Vernazza and Renato Zaccaria

***Abstract*— The automatic assessment of the level of independence of a person, based on the recognition of a set of Activities of Daily Living, is among the most challenging research fields in Ambient Intelligence. The article proposes a framework for the recognition of motion primitives, relying on Gaussian Mixture Modeling and Gaussian Mixture Regression for the creation of activity models. A recognition procedure based on Dynamic Time Warping and Mahalanobis distance is found to: (i) ensure good classification results; (ii) exploit the properties of GMM and GMR modeling to allow for an easy run-time recognition; (iii) enhance the consistency of the recognition via the use of a classifier allowing *unknown* as an answer.**

I. INTRODUCTION

The automatic monitoring of specific Activities of Daily Living (ADLs), adopted in gerontology for the assessment of the independence level of a person, is among the most challenging research fields of Ambient Intelligence. By providing information about a person actions, an automated system for the recognition of human activities can be a useful tool for both the human/robot cooperation in smart environments and the purposes of wearable robotics [1].

Tri-axial accelerometers are acknowledged in literature to provide the most useful information for activity recognition, whereas probability-based methods are the most commonly adopted technique for the modeling of the activities of interest. Nonetheless, there actually is no established solution for the comparison procedure, which aims at quantifying the similarity between run-time sensory data and each model. We propose a comparison and classification procedure for an automatic recognition system for simple ADLs, assuming the information coming from a single wrist-placed tri-axial accelerometer and Gaussian Mixture Modeling (GMM) and Gaussian Mixture Regression (GMR) as modeling method. The main contributions of the article are: (i) a requirement analysis for any comparison procedure dealing with acceleration data and probabilistic models; (ii) a novel comparison procedure based on Dynamic Time Warping and Mahalanobis distance, which exploits the properties of GMM and GMR to allow for an easy run-time classification.

The article is organized as follows: Section II reports a detailed literature analysis; Section III introduces the requirements of the application and describes the system architecture as well as the adopted modeling procedure, while the details of the recognition procedure are analyzed in Section IV; experimental results are provided in Section V; conclusions follows.

B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza and R. Zaccaria are with the University of Genova, Via Opera Pia 13, 16145 Genova, Italy. Corresponding author's email: `barbara.bruno@unige.it`.

II. RELATED WORK

Gerontology defines the *Activities of Daily Living* as daily activities requiring the use of specific basic motor and cognitive capabilities, such as *bathing*, *dressing*, *toileting*, *transferring*, *continence* and *feeding*. From the analysis of a person ability in carrying out ADLs it is possible to estimate their functional status and their level of autonomy [2]. The most commonly adopted sets of ADLs have been envisaged assuming that qualified medical staff is able to examine the person performance on a qualitative (yet informed by experience) basis. On these premises, the design and development of an automated system for the recognition of ADLs is particularly challenging: the design process which is required to translate the qualitative recognition of human behavior in appropriate quantitative and well-defined models is far from being well understood. The initial steps of the design process are the selection of a proper *sensing strategy* (e.g., wearable sensors versus sensors that are distributed in the environment) and the most suitable *modeling approach* (e.g., probability versus logic-based). On the matter of the *sensing strategy*, literature suggests three families of approaches.

- *Smart environments*. Heterogeneous sensors are distributed throughout an environment purposely modified in advance. On the one hand, smart environments have the advantage of posing no limitation on the person movements and appearance [3]; on the other hand, environmental sensors (such as cameras) are highly affected by any interference and provide information about a person activity at a level of detail which, while acceptable for the recognition of complex actions [4], is often not sufficient for simpler actions [5].
- *Wearable sensing*. Sensors are located on the person body using either wearable devices or purposely engineered clothes. The number, size and placement of the required sensors are crucial parameters for both the feasibility and effectiveness of the approach: with only a few exceptions [6], most systems adopt a single sensing device in order to enforce the user motion freedom [7]. Single-device solutions are based either on the integration of different sensors [7] or the use of a single sensing mode. On this matter, accelerometers are found to provide the most useful information [7] and single-device solutions exclusively relying on the acceleration information are successfully adopted in [8]. The optimal placement of the sensing device is determined by a number of factors, among which are the kind of activities to be monitored and the robustness of

the data-acquisition system with respect to the sensor location. As it is detailed in [9], a wrist placement is to be preferred when the activities to be monitored are hand gestures, whereas a chest or hip placement is usually chosen in case of full body motions.
- *Hybrid approaches.* The integration between smart environments and wearable devices is an active research trend in the field of complex action recognition [10], but the complexity of integrating heterogeneous data coming from distributed sources seems not balanced by the increased accuracy for simple motion recognition.

The *modeling approach* defines the system internal representation for the activities in terms of available sensory data. A literature analysis presents two classes of approaches.

- *Logic-based.* Each activity to be monitored is encoded through well-defined *rules*, i.e., ranges of admissible values for relevant parameters. Most systems adopting the logic-based approach use decision trees to classify run-time data by progressively narrowing the set of activities they can represent [8].
- *Probability-based.* Each activity is represented through a *model* and classification is performed by comparing run-time sensory data with stored models through probabilistic distance measures. Most solutions adopt either Hidden Markov modeling [6], [7] or Gaussian Mixture modeling [11]. Probabilistic solutions are found to outperform logic-based solutions [8].

The reported literature analysis lead to the design of a system following the wearable sensing approach and relying on the information provided by a single wrist-mounted tri-axial accelerometer [12]. A modeling procedure based on Gaussian Mixture Modeling and Gaussian Mixture Regression has been chosen for three reasons: (i) it allows for the creation of models of different resolution; (ii) the models can be projected in the space of run-time acceleration data, allowing for an easy run-time comparison; (iii) the models can be either person-specific or general-purpose, according to the chosen modeling dataset.

III. System Requirements and Architecture

A. Motion Primitives

We define *motion primitives* as movements that can uniquely identify an ADL and are associated with stereotyped motions (see Table I). Motion primitives allow for a quantitative description of the ADLs. A number of motion primitives can refer to the same ADL, so that when anyone of them is recognized it is possible to infer that the corresponding ADL has been executed. The recognition of sequences of motion primitives is out of the scope of this article.

B. System Architecture

The human activity recognition system presented in this work is composed of two distinct modules (Figure 1): the *model builder* works off-line for the creation of probabilistic models of relevant motion primitives starting from a provided set of recordings; the *classifier* works on-line for the classification of the run-time acceleration data via the comparison with the available models. A mandatory preliminary step to the execution of the *model builder* is the gathering of a representative training set: for each motion primitive, a relevant number of examples (henceforth referred to as *modeling trials* or *trials*) is obtained by providing volunteers with a wrist-mounted tri-axial accelerometer and asking them to perform each motion primitive multiple times. The corresponding acceleration data are recorded and each occurrence of any motion primitive is tagged. The result is a dataset which, for each motion primitive, contains a huge number of examples in the form of sequences of acceleration data along the three axes x, y and z. The selection of the examples to be included in the training set allows to choose between specific (i.e., biased towards the behavior of a specific individual) or general (i.e., not biased) models.

TABLE I

ADLs AND RELATED MOTION PRIMITIVES

Personal hygiene	teeth brushing hair combing
Mobility	stairs climbing stairs descending walking
Feeding	drinking from a glass pouring water in a glass eating with fork and knife eating with spoon
Communication	using the telephone
Functional transfers	getting up from the bed lying down on the bed standing up from a chair sitting down on a chair

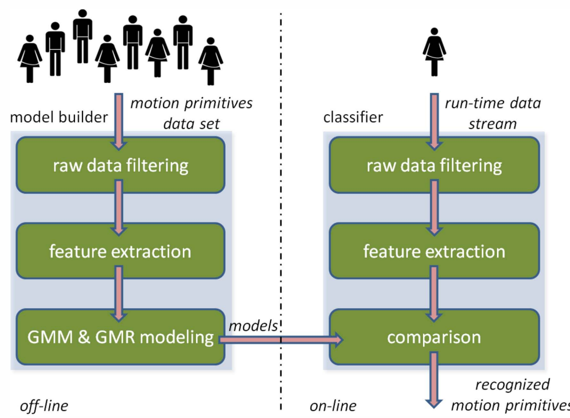

Fig. 1. System architecture.

C. Model Builder Module

The *model builder* module performs a number of steps (see Figure 1 on the left hand side).

Raw data filtering. The system employs a median filter to reduce the high frequency noise affecting the acceleration signal, having been demonstrated that median filters outperform linear filters when the signal-to-noise ratio is greater than 1 [13], which holds in our case. A window length of

$h = 3$ has been experimentally estimated as a good trade-off between noise reduction in high frequency and signal dynamics preservation in low frequency.

Feature extraction. Features are defined as quantities that can be extracted from the acceleration data to enhance the differences between motion primitives. Time-domain features are better suited to systems with real-time requirements [14] and *gravity* and *body acceleration* (i.e., acceleration generated by person movements) are the most commonly adopted features [8], [14]. The method for gravity and body acceleration extraction is implemented in two steps: (i) a low-pass filter is applied to the acceleration signal to isolate the gravity component; (ii) the gravity component is subtracted from the original signal to obtain the body acceleration component [14]. The explicit use of correlation among tri-axial acceleration data has been suggested in [8] and features taking correlation into account have been experimentally proved to lead to higher classification accuracy with respect to uncorrelated features [12]. Our system exploits two separate 4-dimensional features, namely gravity $g = (g_x, g_y, g_z, k)$ and body acceleration $b = (b_x, b_y, b_z, k)$.

GMM and GMR modeling. GMM and GMR generate, for each motion primitive, a probabilistic model of the associated acceleration pattern. The procedure builds expected curves of the features for each motion primitive from a set of modeling trials. A detailed description of the implemented GMM and GMR modeling procedure can be found in [12].

Let us assume that we have M different motion primitives. For each motion primitive m to learn, where $m = 1, \ldots, M$, let us assume that the training set includes S_m modeling trials and s is one of these trials. For each sample index k, where $k = 1, \ldots, K_m$, we define $g_k^s \in \mathbb{R}^4$ such that $g_k^s = (g_x(k), g_y(k), g_z(k), k)$ and $b_k^s \in \mathbb{R}^4$ such that $b_k^s = (b_x(k), b_y(k), b_z(k), k)$ as the gravity and body acceleration extracted from trial s, respectively. As an example, Figure 2 shows the gravity and body acceleration feature curves (i.e., g_k^s, b_k^s, $\forall k, s$) extracted from $S_m = 10$ trials modeling the *eating with knife and fork* motion primitive. We denote with ξ a generic feature, i.e., ξ can either correspond to gravity or body acceleration. The following definitions are in order.

- $\xi_k^s \in \mathbb{R}^4$ is the *data point* k of trial s, defined as:

$$\xi_k^s = (\xi_k^{s,a}, k), \tag{1}$$

  where $\xi_k^{s,a}$ stores the 3-axial acceleration information of the data point k, e.g., b_x, b_y and b_z for body acceleration, and k stores the sampling information.
- Ξ^ξ is the set of data points generating the feature curve ξ for all the S_m trials, defined as:

$$\Xi^\xi = \{\xi_1^1, \ldots, \xi_{K_m}^1, \xi_1^2, \ldots, \xi_{K_m}^{S_m}\}. \tag{2}$$

The S_m trials are synchronized (i.e., the starting index of the actual activity is the same in all the trials) and truncated (i.e., all the trials are composed of K_m data-points).

The purpose of GMM and GMR modeling is to build a single, generalized version of each feature in the form of:

$$\hat{\Xi}^\xi = (\mu^\xi, \Sigma^\xi), \tag{3}$$

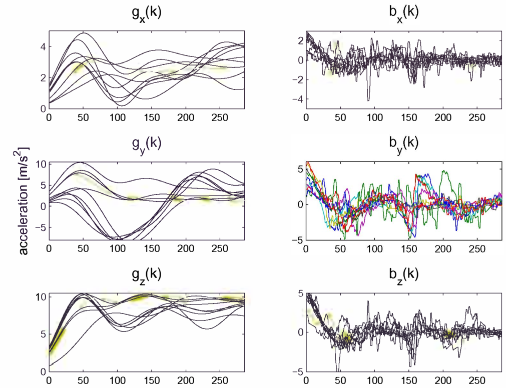

Fig. 2. Feature curves extracted from the trials of the *eating with knife and fork* motion primitive training set.

where μ^ξ is the *expected curve* modeling feature ξ for the motion primitive m and Σ^ξ is the covariance matrix associated with μ^ξ. The feature models $\hat{\Xi}^g$ and $\hat{\Xi}^b$ allow for a parametrized model of motion primitive m as:

$$\hat{\Xi}^m = (\hat{\Xi}^g, \hat{\Xi}^b). \tag{4}$$

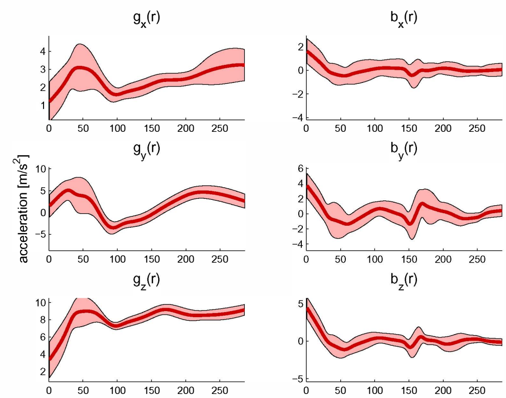

Fig. 3. 2D projections of the *eating with knife and fork* motion primitive model retrieved via GMR.

Figure 3 shows the feature models $\hat{\Xi}^g$ (left) and $\hat{\Xi}^b$ (right) of the *eating with knife and fork* motion primitive. The chosen modeling procedure allows for the creation of models existing in the run-time data domain and composed of a scalable number R_m of data points: it retrieves the expected acceleration value and the corresponding covariance matrix for every time instant that is present in a set $\Xi^{\xi,t}$, which does not necessarily have to coincide with the set of time instants of the modeling set trials, given that $R_m \leq K_m$.

IV. Classifier Module

A. Requirement Analysis for the Comparison Procedure

Acceleration data referring to the same motion primitive can significantly vary. Let us consider an example: Figure 4 shows gravity and body acceleration features extracted from two trials of the *lying down on the bed* activity (respectively, in blue and green) performed by two people. Although the trials have been synchronized, they still differ as far as two parameters are concerned.

- *Acceleration value*. The acceleration values recorded in the green trial and in the blue trial, albeit very close in the initial stages of the activity, grow progressively distant in later time instants.
- *Motion speed*. The acceleration curve recorded in the green trial corresponds to a locally shrunk version of the one recorded in the blue trial: the difference in speed between two trials can be time-dependent and it can vary non-linearly.

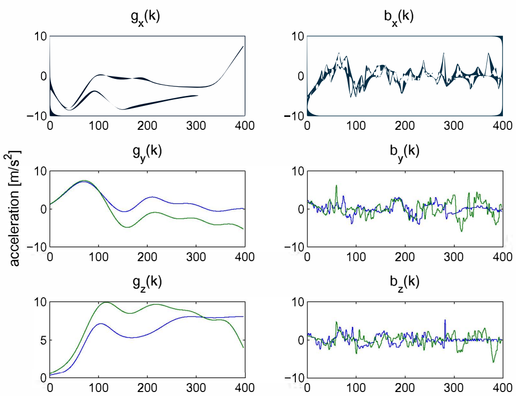

Fig. 4. Feature data recorded during two trials (respectively, in blue and green) of the *lying down on the bed* activity, performed by different users.

Acceleration value and motion speed are the sole parameters where differences can occur between run-time acceleration data and probabilistic models. The two aspects can be elegantly dealt with by integrating two different distance metrics, namely the Mahalanobis distance [15] and the Dynamic Time Warping technique [16], to obtain a method to compare 4-dimensional acceleration signals. Furthermore, the two techniques can be used independently (e.g., Mahalanobis distance alone) when real-time requirements must be met.

B. Comparison Procedure

During the on-line phase (see Figure 1 on the right hand side), the monitored person is wearing a wrist-placed device providing acceleration data and a number of probabilistic models of motion primitives are available. The purpose of the comparison procedure is to *rank* the similarity between run-time acceleration data and each model m by computing the likelihood of its features (i.e., g and b) to the model features. The proposed comparison procedure is a hybrid distance metric based on the integration of Mahalanobis distance and Dynamic Time Warping (DTW): the former is a probabilistic distance measure used to compute the similarity between sets of random variables whose means and variances are known, whereas the latter is an algorithm used to compute the similarity between two numerical sequences characterized by different sampling frequencies (thereby taking into account the more general case of variations in time and/or speed).

Let us consider a moving horizon window Ξ_w of length N_w moving on the run-time acceleration stream:

$$\Xi_w = \{\xi_{1,w}, \ldots, \xi_{j,w}, \ldots, \xi_{N_w,w}\}, \tag{5}$$

where the window length is computed as:

$$N_w = \operatorname*{arg\,min}_{m=1,\ldots,M} K_m \tag{6}$$

corresponding to the length of the longest among the considered motion primitives. In all the experiments we set $N_w = 365$. On the basis of the M models available in the form of (4) and given a data stream in the form of (5), the goal is to find the model m^* with minimum distance to Ξ_w:

$$m^* = \operatorname*{arg\,min}_{m=1,\ldots,M} d(\hat{\Xi}^m, \Xi_w) \tag{7}$$

The *data filtering* and *feature extraction* procedures, which execute the very same algorithms of the off-line phase, allow for the creation of two sets $\Xi_w^{\xi,a}$ representing the run-time feature acceleration data. Adopting the notation of [17], we refer to $X = \hat{\Xi}^{\xi,a}$ as the set of acceleration values of the feature model $\hat{\Xi}^{\xi}$ and to $Y = \Xi_w^{\xi,a}$ as the run-time feature acceleration data. The Mahalanobis distance $d_M^{\xi}(r,j)$ between element $r = X_r = \mu_r^{\xi,a}$ and element $j = Y_j = \xi_{j,w}^a$ is computed as:

$$d_M^{\xi}(r,j) = \sqrt{(\mu_r^{\xi,a} - \xi_{j,w}^a)^T (\Sigma_r^{\xi,aa})^{-1} (\mu_r^{\xi,a} - \xi_{j,w}^a)}. \tag{8}$$

Given the two sets X and Y, we define the *warping path* $\phi(t)$, with $t = 1, \ldots, T$, as the mapping between the domains of X and Y and the domain swept by the t parameter $[1, T]$:

$$\phi(t) : X \times Y \rightarrow [1, T]. \tag{9}$$

The warping path defines the *optimal* distortion of X and Y to be properly compared. $\phi(t)$ is composed of two functions of t, namely $\phi_x(t) \in [1, R_m]$ and $\phi_y(t) \in [1, N_w]$, which remap in $[1, T]$ the indexes of, respectively, X and Y. Given that X and Y are represented in a probabilistic framework, the accumulated distortion induced by $\phi(t)$ between the warped series X and Y can be computed as:

$$d_\phi(X, Y) = \sum_{t=1}^{T} d_M^{\xi}(\phi_x(t), \phi_y(t)) \frac{m_\phi(t)}{M_\phi}, \tag{10}$$

where $m_\phi(t)$ is a per-step weighting coefficient and M_ϕ is the corresponding normalization constant, ensuring that the accumulated distortions are comparable. In particular, (10) yields the cumulative Mahalanobis distance between pairwise elements of X and Y weighted by the related distortion along t. In order for the mapping in (9) to be bijective (i.e., to preserve the time ordering between the two series), $\phi(t)$ is constrained to be monotone in each component, namely:

$$\begin{aligned} \phi_x(t+1) &\geq \phi_x(t), \\ \phi_y(t+1) &\geq \phi_y(t). \end{aligned} \tag{11}$$

Given these constraints, for any X and Y many warping paths are possible, i.e., $\phi_1(t), \ldots, \phi_P(t)$. The DTW technique determines the optimal alignment $\phi^*(t)$ such that:

$$d^{\xi}_{\phi^*}(X,Y) = \underset{\phi_1(t),\ldots,\phi_P(t)}{\arg\min} \; d_{\phi}(X,Y). \tag{12}$$

The *overall distance* between the run-time data and a model is computed as the average distance on all the features:

$$d(\hat{\Xi}^m, \Xi_w) = \frac{1}{N_\xi} \sum_{\xi=1}^{N_\xi} d^{\xi}_{\phi^*}(\hat{\Xi}^{\xi,a}, \Xi^{\xi,a}_w). \tag{13}$$

C. Classification Procedure

The comparison procedure can yield a classification for *any* acceleration pattern. To avoid this drawback, a threshold mechanism is set-up to discriminate between unknown and potentially known motion primitives. We assume that the run time data stream Ξ_w is labeled as an occurrence of a motion primitive m whose model $\hat{\Xi}^m$ has the minimum distance $d(\hat{\Xi}^m, \Xi_w)$ among those below specific thresholds. Given a motion primitive m, we define the corresponding threshold τ_m as:

$$\tau_m = \frac{1}{N_\xi} \sum_{\xi=1}^{N_\xi} d^{\xi,f}(\Xi^{\xi,a,f}, \hat{\Xi}^{\xi,a}), \tag{14}$$

where $d^{\xi,f}$ is computed as the distance between the acceleration components of the model $\hat{\Xi}^{\xi,a}$ and the *farthest curve* $\Xi^{\xi,a,f}$, generated from the model itself, as:

$$\Xi^{\xi,f} = \{\xi^f_1, \ldots, \xi^f_r, \ldots, \xi^f_{R_m}\}, \tag{15}$$

where the single element ξ^f_r is represented as:

$$\xi^f_r = (\mu^{\xi,a}_r + \beta \Sigma^{\xi,aa}_r, \mu^{\xi,t}_r). \tag{16}$$

The parameter β is a fixed scaling factor experimentally set to take into account the *non generality* of the model, whereas $\Sigma^{\xi,aa}_r$ is the variance associated with the element r.

V. Experimental Validation

A. Results

System architecture. The system employs a *sensing bracelet* mounted on the right wrist of the user. The sensor is a tri-axial accelerometer with a sensing range of 3 *G* coding the information using 6 bit per axis, with a sampling frequency $F_s = 32$ *Hz*. *Data filtering* is performed by a median filter of size 3 and *feature extraction* relies on a low-pass Chebyshev I 5° order filter with $F_{cut} = 0.25$ *Hz*, $A_{pass} = 0.001$ *dB*, $A_{stop} = -100$ *dB*, $F_{stop} = 2$ *Hz*.

We recorded over 700 trials of 8 motion primitives from 16 volunteers (11 men and 5 women with age ranging from 19 to 83) (see Table II). Among the chosen motion primitives there are postural transitions (*getting up* from the bed, *sitting down* and *standing up* from a chair), reiterated activities (*climbing the stairs* and *walking*), complex activities associated to "standardized" motions (*drinking* and *pouring water*) and complex activities that different people execute in different ways (*eating*). *Climbing the stairs* & *walking* and *getting up* from the bed & *standing up* from a chair have very similar acceleration patterns, while the acceleration patterns of the simpler motions can be found in the initial stages of more complex motions, albeit of different nature (the initial stages of *climbing the stairs*, *walking* and *drinking* closely resemble the motion primitive *getting up* from the bed). Table III reports the details of the models we have built.

TABLE II
MOTION PRIMITIVES DATASET

Motion primitive	# recorded trials	# volunteers
Climbing the stairs	102	10 (7M, 3F)
Drinking from a glass	100	10 (6M, 4F)
Eating with fork and knife	14	4 (2M, 2F)
Getting up from the bed	101	10 (5M, 5F)
Pouring water in a glass	100	10 (6M, 4F)
Sitting down on a chair	100	10 (6M, 4F)
Standing up from a chair	102	10 (6M, 4F)
Walking	100	10 (7M, 3F)

TABLE III
MODELLING DATASET

Model	# trials	# volunteers	K_m
Climbing the stairs	20	5	250
Drinking from a glass	20	5	270
Eating with fork and knife	10	3	350
Getting up from the bed	20	5	260
Pouring water in a glass	20	5	365
Sitting down on a chair	20	5	155
Standing up from a chair	20	5	168
Walking	20	5	170

Recognition accuracy. Table IV reports the recognition accuracy rates achieved using the comparison procedure described in Section IV, while Table V reports the recognition accuracy rates achieved using a comparison procedure exclusively relying on Mahalanobis distance, i.e. adopting:

$$d(\hat{\Xi}^m, \Xi_w) = \frac{1}{N_\xi} \sum_{\xi=1}^{N_\xi} d^{\xi}_{M}(\hat{\Xi}^{\xi,a}, \Xi^{\xi,a}_w). \tag{17}$$

With TP we denote the *true positive* rate (i.e., the number of occurrences of the activity that were correctly labeled) and with TN we denote the *true negative* rate (i.e., the number of occurrences of different activities that were not mistakenly labeled as occurrences of the considered activity). Table VI focuses on the *getting up* from the bed motion primitive and reports the true positive and true negative rates obtained using models with different R_m. The purpose of the experiment is to estimate the relation between the recognition accuracy and the resolution of the models, so that it is possible to identify a value for R_m that (i) ensures an acceptable recognition accuracy and (ii) minimizes the memory space required to store the models. The experiments consider models with constant sampling, but the GMR procedure allows for non-constant and non-linear sampling.

B. Discussion

At the best of our knowledge, there is no system for the automatic recognition of ADLs relying on two features

TABLE IV

RECOGNITION ACCURACY WITH COMPARISON BASED ON DYNAMIC TIME WARPING AND MAHALANOBIS DISTANCE

Model	τ_m	TP	TN
Climbing the stairs	$2.0386e^{+3}$	20%	93.34%
Drinking from a glass	$8.9103e^{+3}$	100%	83.34%
Getting up from the bed	$1.0244e^{+4}$	60%	66.67%
Pouring water in a glass	$8.8823e^{+3}$	100%	80%
Sitting down on a chair	$2.8519e^{+3}$	0%	93.34%
Standing up from a chair	$3.7048e^{+3}$	60%	83.34%
Walking	$2.5884e^{+3}$	40%	70%

TABLE V

RECOGNITION ACCURACY WITH COMPARISON BASED ON MAHALANOBIS DISTANCE

Model	τ_m	TP	TN
Climbing the stairs	16.8507	38.3%	85.78%
Drinking from a glass	133.6197	90.91%	86.59%
Eating with fork and knife	85.5632	88.89%	96.67%
Getting up from the bed	97.6461	93.33%	61.64%
Pouring water in a glass	108.1443	95%	71.55%
Sitting down on a chair	37.5798	0%	100%
Standing up from a chair	63.8687	35.71%	93.96%
Walking	46.1097	88.33%	79.91%

only. A comparison with the accuracy rates achieved by similar systems relying on broader feature sets [8], [14] lead us to think that the loss in information deriving from the reduced feature set is adequately compensated by the adopted modeling and comparison techniques. Table V and Table VI suggest that it is possible to design a system for the automatic recognition of a few selected motion primitives relying on GMM and GMR and Mahalanobis distance subject to soft real-time constraints and ensuring good classification accuracy. Table IV and Table V suggest that when the *model builder* is given a large modeling dataset the models retrieved by GMR merge the differences in speed between the trials, so that the adoption of DTW does not significantly increase the recognition accuracy. Longer models tend to prevail over shorter models, suggesting the need to consider a balanced set of motion primitives. *Climbing*, *sitting down* and *standing up* are associated with low thresholds, suggesting that the corresponding modeling datasets are biased.

VI. CONCLUSIONS

We propose a comparison and classification procedure for a system aimed at the automatic recognition of simple ADLs. The system assumes the information coming from a single wrist-placed tri-axial accelerometer and extracts two time-domain features from the acceleration data. The adoption of a modeling procedure based on GMM and GMR and a comparison procedure based on DTW and Mahalanobis distance allows to obtain recognition results comparable with those of systems relying on broader feature sets.

TABLE VI

RECOGNITION ACCURACY RELIANCE ON R_m

R_m/K_m	TP	TN
1	93.33%	60.34%
1/2	70%	51.72%
1/4	45%	53.45%

REFERENCES

[1] B. Argall and A. Billard, "A survey of tactile human-robot interactions," *Robotics and Autonomous Systems*, vol. 58, 2010.

[2] S. Katz, A. Chinn, and L. Cordrey, "Multidisciplinary studies of illness in aged persons: a new classification of functional status in activities of daily living," *Journal of Chronic Disease*, vol. 9, no. 1, pp. 55–62, 1959.

[3] T. Yamazaki, "Assistive technologies in smart homes," in *F. Mastrogiovanni and N.Y. Chong (Eds.) Handbook of Research on Ambient Intelligence and Smart Environments*. Hershey, Pennsylvania, USA: IGI Global, pp. 165–181.

[4] F. Mastrogiovanni, A. Sgorbissa, and R. Zaccaria, "Activity recognition in smart homes: from specification to recognition," *International Journal of Fuzzy and Intelligent Systems*, vol. 21, no. 1-2, pp. 33–48, 2010.

[5] R. Poppe, "A survey on vision-based human action recognition," *Image and Vision Computing*, vol. 28, no. 6, pp. 976–990, 2010.

[6] D. Minnen and T. Starner, "Recognizing and discovering human actions from on-body sensor data," in *Proceedings of the 2005 IEEE International Conference on Multimedia and Expo (ICME'05)*, Amsterdam, The Netherlands, July 2005.

[7] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. LaMarca, L. LeGrand, A. Rahimi, A. Rea, G. Borriello, B. Hemingway, P. Klasnja, K. Koscher, J. Landay, J. Lester, D. Wyatt, and D. Haehnel, "The mobile sensing platform: an embedded activity recognition system," *IEEE Pervasive Computing Magazine*, vol. 7, no. 2, pp. 32–41, 2008.

[8] G. Krassnig, D. Tantinger, C. Hofmann, T. Wittenberg, and M. Struck, "User-friendly system for recognition of activities with an accelerometer," in *Proceedings of the 2010 International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'10)*, Munich, Germany, May 2010.

[9] D. Olguin and A. Pentland, "Human activity recognition: accuracy across common locations for wearable sensors," in *Proceedings of the 2006 IEEE International Symposium on Wearable Computers (ISWC'06)*, Montreaux, Switzerland, October 2006.

[10] F. Mastrogiovanni, A. Bagnasco, G. Allasia, and A. Giuni, "The shell project: a secure habitat for an enhanced long living - design and implementation guidelines for a new concept of people assistance," in *Proceedings of the Third International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010)*, Rome, Italy, November 2010.

[11] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[12] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Human motion modeling and recognition: a computational approach," in *Proceedings of the 8th IEEE International Conference on Automation Science and Engineering (CASE 2012)*, Seoul, Korea, August 2012.

[13] A. Castro and D. Donoho, "Does median filtering truly preserve edges better than linear filtering?" *The Annals of Statistics*, vol. 37, no. 3, pp. 1172–1206, 2009.

[14] D. Karantonis, M. Narayanan, M. Mathie, N. Lovell, and B. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 156–167, 2006.

[15] P. Mahalanobis, "On the generalized distance in statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, 1936.

[16] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 1, pp. 43–49, 1978.

[17] T. Giorgino, "Computing and visualizing dynamic time warping alignments in r: the dtw package," *Journal of Statistical Software*, vol. 31, no. 7, pp. 1–24, 2009.