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 and active research fields. Purpose of the article is to propose a framework for the recognition of motion primitives, relying on Gaussian Mixture Modeling and Regression for the creation of the models of the activities. 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 an open classifier.**

I. INTRODUCTION

The automatic monitoring of specific Activities of Daily Living, defined and adopted in gerontology for the assessment of the level of independence of a person, is among the most challenging and active research fields of Ambient Intelligence. Although various approaches to the task have been tested, making use of a wide variety of sensors and recognition algorithms, tri-axial accelerometers are acknowledged to provide the most useful information and probability-based methods are the most commonly adopted technique for the creation of the models of the activities of interest. Nonetheless, there actually is no established solution for the comparison procedure, which aims at quantifying the similarities between the run-time sensory data and each of the known models.

We propose a comparison and classification procedure for an automatic system for the recognition of simple ADLs, assuming the information coming from a single wrist-placed tri-axial accelerometer and Gaussian Mixture Modeling and Gaussian Mixture Regression 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 and classification procedure based on Dynamic Time Warping and Mahalanobis distance, which ensures good classification results and exploits the properties of GMM and GMR modeling to allow for an easy run-time recognition.

The article is organized as follows: Section II reports a detailed analysis of the Literature of interest; Section III introduces the main requirements of the application and describes the architecture of the system and the adopted modeling procedure, while the details of the recognition procedure are analyzed in Section IV; experimental results are provided and discussed 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* (ADLs) 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 the ADLs it is possible to estimate his functional status and determine his level of autonomy [1].

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 and classification of ADLs is particularly challenging: the design process necessary 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 possible families of approaches:

- *Smart environments*. Heterogeneous sensors are distributed throughout an environment that must be purposely modified in advance. On the one hand, smart environments have the advantage of posing no limitation on the person movements and appearance [2]; on the other hand, environmental sensors 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 [3], [4], is often not sufficient for simpler actions [5], [6].
- *Wearable sensing*. Sensors are located on the person body using either wearable devices or purposely engineered articles of clothing. Literature proves that 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 [7], most systems adopt a single sensing device in order to enforce the user motion freedom [8], [9]. Single-device solutions are based either on the integration of different sensors [8], [9] or the use of a single sensing mode. On this matter, accelerometers are found to provide the most useful information [9] and single-device solutions exclusively relying on the acceleration information are successfully adopted in [10], [11], [12], [13].

- *Hybrid approach*. The integration between smart environments and wearable devices is an active and promising research trend in the field of complex actions recognition [14], but the complexity of integrating heterogeneous data coming from distributed sources seems not balanced by the associated increase in accuracy for the task of simple motions recognition.

The *modeling approach* defines the system internal representation of the activities of interest in terms of the available sensory data. Literature presents two classes of approaches:

- *Logic-based*. Each activity to monitor and recognize is encoded through sound and well-defined *rules*, i.e., ranges of admissible values for a set of relevant parameters. Most systems adopting the logic-based approach use decision trees to classify the run-time data by progressively narrowing the set of activities they could represent [10], [13], [15].
- *Probability-based*. Each activity is represented through a *model* and classification is performed by comparing run-time sensory data with the stored models through probabilistic distance measures. Most solutions adopt either Hidden Markov Modeling [7], [9] or Gaussian Mixture Modeling [12], [16] for the creation of the models. Probabilistic solutions are found to outperform logic-based solutions [13].

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 tri-axial accelerometer [17]. A modeling procedure based on the combination of Gaussian Mixture Modeling (GMM) and Gaussian Mixture Regression (GMR) has been preferred to other probability-based modeling methods for three reasons: (i) it allows for the creation of models of different resolution; (ii) the models can be projected in the space of the run-time acceleration data, thus allowing for an easy run-time comparison; (iii) the models can be either person-specific or general-purpose, according to the chosen modeling dataset.

The optimal placement of the sensing device is usually 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. Since most of the ADLs considered in the presented work are purely arm gestures, we decided to place the accelerometer on the person right wrist.

III. System Requirements and Architecture

A. Motion Primitives

We define *motion primitives* human movements that can uniquely identify an ADL and are associated with stereotyped and simple motions (see Table I) and adopt them to allow for a quantitative description and analysis of the considered activities. A number of motion primitives can correspond to the same ADL, so that when any one 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.

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

B. System Architecture

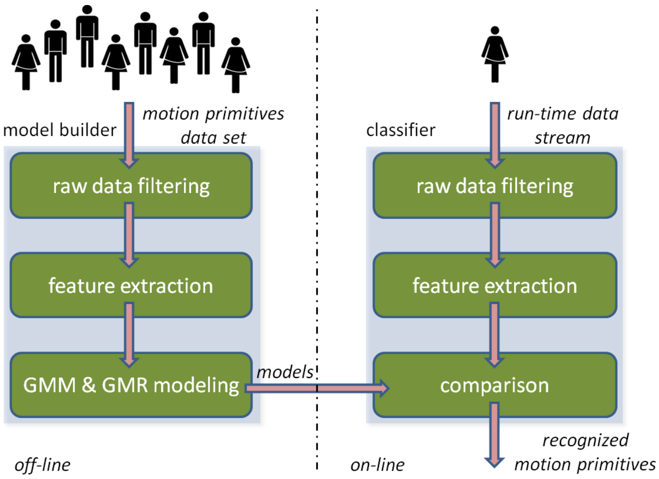

Fig. 1. System architecture.

The human activities recognition system presented in this work (see Figure 1) is composed of two distinct modules: the *model builder* works off-line for the creation of probabilistic models of relevant motion primitives starting from a provided dataset of human examples; 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 module* is the gathering of a representative training set: for each considered motion primitive, a relevant number of human examples is obtained by first providing volunteers with a wrist-mounted tri-axial accelerometer and then asking them to perform each motion primitive multiple times. The corresponding acceleration data are recorded and each occurrence of any motion primitive is tagged by a human observer. 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 of the modeling procedure allows to chose between general (i.e., not biased toward the behavior of a specific human individual) or specific (i.e., biased toward the behavior of a specific individual) models.

Once the training set is available, the *model builder* module performs a number of steps (see Figure 1 on the left hand side): *raw data filtering* is aimed at reducing high frequency noise, *feature extraction* discriminates between acceleration components induced by gravity and the body movements, respectively, *GMM and GMR modeling* generates the probabilistic models of the motion primitives. During the on-line phase (see Figure 1 on the right hand side), it is assumed that a monitored person is wearing a wrist-placed device providing acceleration data and that a number of probabilistic models of motion primitives to monitor are available. The on-line *raw data filtering* and *feature extraction* steps execute the very same algorithms of the off-line phase on the run-time acceleration data. The *comparison* step performs the classification by considering a temporal window moving over the run-time data stream and comparing it with the stored models. Given the computational requirements of this phase, the ability of GMM and GMR to generate models with variable resolution and complexity is fundamental for an effective classification procedure.

C. Model Builder Module

Data filtering. The proposed system makes use of 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 [18], which holds valid 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 in low frequency.

Feature extraction. Features are defined as quantities that can be extracted from the acceleration data to enhance the differences between the considered motion primitives. Time-domain features are better suited to systems with real-time requirements [15] and *gravity* and *body acceleration* (i.e., acceleration generated by the person movements) are the most commonly adopted features [12], [13], [15], [19]. The chosen 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 [12], [15].

The explicit use of correlation among tri-axial acceleration data has been suggested in [13], [20], [21], [22] and 4-dimensional features have been experimentally proved to lead to higher classification accuracy with respect to 2-dimensional features [17]. Our system makes use of 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 modelling. As a final step of the off-line phase, *GMM and GMR modeling* generates, for each motion primitive that must be detected at run-time, a probabilistic model of the learned acceleration patterns. The procedure builds "expected curves" of each motion primitive from a training set of human examples (henceforth referred to as "modeling trials" or "trials"). A detailed description of the GMM and GMR modeling procedure implemented in the system can be found in [17].

Let us assume to focus on a specific motion primitive m to learn, where $m = 1, \ldots, M$. For each motion primitive m, let us define s, where $s = 1, \ldots, S_m$, as one of the modeling trials part of the training set for m. All the S_m trials are assumed to be synchronized and truncated as to be composed of the same number K_m of data points. Figure 2 shows the gravity feature curves extracted from $S_m = 10$ trials modeling the *standing up from a chair* motion primitive.

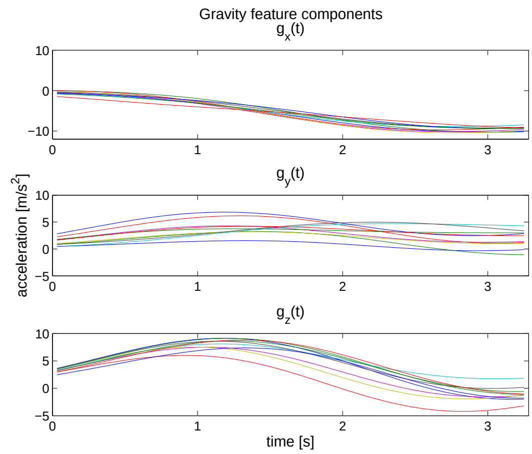

Fig. 2. Gravity feature curves extracted from 10 trials modeling the *standing up from a chair* motion primitive.

For each sample index k, where $k = 1, \ldots, K_m$, we denote with ξ the generic feature of interest, i.e., $\xi \in \{g, b\}$. The following definitions are in order.

- $\xi_k^s \in \mathbb{R}^4$ is the data point of the feature ξ that is extracted from trial s at point k, defined as:
$$\xi_k^s = (\xi_k^{s,a}, \xi_k^{s,t}), \tag{1}$$
where $\xi_k^{s,a}$ stores the 3-axial acceleration information of the data point k, e.g., g_x, g_y and g_z for gravity, and $\xi_k^{s,t}$ stores the timing information associated with data point k, e.g., k itself.
- Ξ^ξ is the set of all the data points ξ_k^s extracted for feature ξ from the S_m modeling trials, defined as:
$$\Xi^\xi = \{\xi_1^1, \ldots, \xi_{K_m}^1, \xi_1^2, \ldots, \Xi_{K_m}^{S_m}\}. \tag{2}$$

GMM partitions Ξ^ξ in a set of clusters:
$$\Xi^\xi = \{\Xi_1, \ldots, \Xi_c, \ldots, \Xi_{C_m^{\xi,*}}\}, \tag{3}$$
where each cluster Ξ_c is associated to a Gaussian function such that:
$$\Xi_c = (\mu, \Sigma, \pi, E)_c, \tag{4}$$
where:

- μ_c is the mean of the data points ξ_k^s belonging to Ξ_c;
- Σ_c is the covariance matrix associated with Ξ_c;
- π_c is the prior probability of Ξ_c;
- E_c is the cumulated posterior probability of Ξ_c.

The number $C_m^{\xi,*}$ of clusters must be optimally determined according to the considered feature ξ and motion primitive m so that the resulting groups enhance the distinctive traits of the motion primitive to enforce good classification chances: a novel algorithm based on the concept of *silhouettes* [23] is adopted. A detailed description of the algorithm can be found in [17].

The purpose of GMR is to build, starting from the $C_m^{\xi,*}$ GMM clusters Ξ_c, a single, generalized version of each feature of interest in the form of:

$$\hat{\Xi}^{\xi} = (\mu^{\xi}, \Sigma^{\xi}), \tag{5}$$

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 to obtain a parametrized model of motion primitive m as:

$$\hat{\Xi}^{m} = (\hat{\Xi}^{g}, \hat{\Xi}^{b}). \tag{6}$$

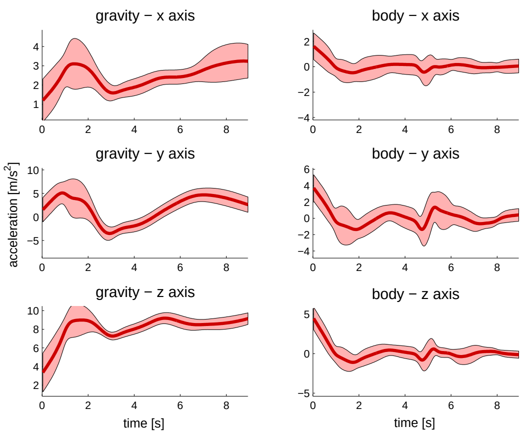

Fig. 3. 2D projections of the *eating with knife and fork* motion primitive model retrieved via GMR.

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 application of regression to the models computed as Gaussian mixtures in (4) allows for the creation of models existing in the run-time data domain and composed of a scalable R_m number of data points: GMR retrieves the expected acceleration value and 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 given by the trials of the modeling set, given that $R_m \leq K_m$. The resulting expected curves can have lower, and eventually variable, resolution with respect to the sampling frequency of the modeling trials.

IV. Classifier Module

A. Requirement Analysis for the Comparison Procedure

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

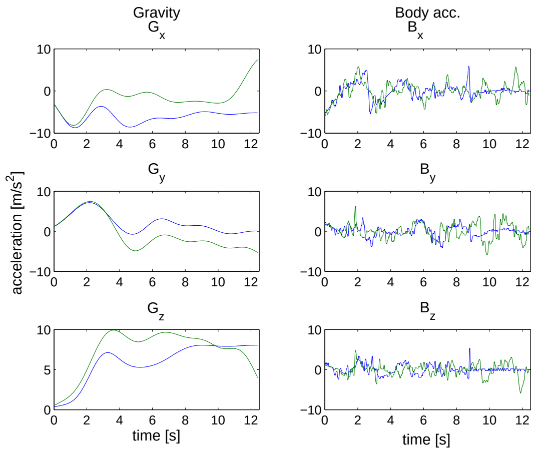

Fig. 4. Feature data recorded during two trials (respectively, in blue and green) of the *lying down on the bed* activity, performed by different users.

- *Acceleration value*. The acceleration value recorded in the green trial is different from the acceleration value recorded in the blue trial for any time instant k.
- *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.

Acceleration value and motion speed are the sole parameters where differences can occur between run-time acceleration data and the probabilistic models. The two aspects can be elegantly dealt with by integrating two different distance metrics, namely the Mahalanobis distance [24] and the Dynamic Time Warping technique [25], to obtain a simple method to compare 4-dimensional acceleration signals both in the acceleration value and time domains. Furthermore, the two techniques can be used independently (e.g., Mahalanobis distance alone) when computational resources are limited and real-time requirements must be met.

B. Comparison Procedure

The purpose of the comparison procedure is to *rank* the similarity between run-time acceleration data and each previously learned model m by computing the likelihood of its features (i.e., g and b) with respect 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 also 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 as follows:

$$\Xi_w = \{\xi_{1,w}, \ldots, \xi_{j,w}, \ldots, \xi_{N_w,w}\}. \quad (7)$$

The window length can be tuned according to the available computational resources; we set $N_w = 365$.

On the basis of the M models available in the form of (6) and given a data stream in the form of (7), the goal is to find the model m^* with minimum distance to Ξ_w:

$$m^* = \underset{m=1,\ldots,M}{\arg\min}\, d(\hat{\Xi}^m, \Xi_w) \quad (8)$$

The *data filtering* and *feature extraction* procedures allow for the creation of two sets $\Xi_w^{\xi,a}$ representing the run-time feature acceleration data. Adopting the notation introduced in [26], 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:

$$\begin{aligned} d_M^{\xi}(r,j) &= d(\mu_r^{\xi,a}, \xi_{j,w}^a) = \\ &= \sqrt{(\mu_r^{\xi,a} - \xi_{j,w}^a)^T (\Sigma_r^{\xi,aa})^{-1} (\mu_r^{\xi,a} - \xi_{j,w}^a)}. \end{aligned} \quad (9)$$

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

$$\phi(t) : X \times Y \rightarrow [1, T]. \quad (10)$$

The warping path defines the *optimal* distortion of X and Y to be properly compared as time series. $\phi(t)$ can be thought as made up 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 time series X and Y can be computed starting from (9) as:

$$d_\phi(X, Y) = \sum_{t=1}^{T} d_M^{\xi}(\phi_x(t), \phi_y(t)) \frac{m_\phi(t)}{M_\phi}, \quad (11)$$

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, (11) 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 (10) 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:

$$\begin{aligned} \phi_x(t+1) &\geq \phi_x(t), \\ \phi_y(t+1) &\geq \phi_y(t). \end{aligned} \quad (12)$$

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:

$$d_{\phi^*}^{\xi}(X, Y) = \underset{\phi_1(t),\ldots,\phi_P(t)}{\arg\min}\, d_\phi(X, Y). \quad (13)$$

The optimal alignment ϕ^* is the deformation of the time axes of X and Y minimizing the distance between the two sequences.

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

$$d(\hat{\Xi}^m, \Xi_w) = \frac{1}{N_\xi} \sum_{\xi=1}^{N_\xi} d_{\phi^*}^{\xi}(\hat{\Xi}^{\xi,a}, \Xi_w^{\xi,a}). \quad (14)$$

C. Classification Procedure

The comparison procedure can yield a classification for *any* acceleration pattern, including those patterns that are not modeled in advance. To avoid this drawback, a threshold mechanism is set-up to discriminate between unknown and potentially known motion primitives. Since the distance between the run time data and each model is numerically model-dependent 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:

$$\tau_m = \frac{1}{N_\xi} \sum_{\xi=1}^{N_\xi} d^{\xi,f}(\Xi^{\xi,a,f}, \hat{\Xi}^{\xi,a}), \quad (15)$$

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:

$$\Xi^{\xi,f} = \{\xi_1^f, \ldots, \xi_r^f, \ldots, \xi_{R_m}^f\}, \quad (16)$$

where the single element ξ_r^f is represented as:

$$\xi_r^f = (\mu_r^{\xi,a} + \beta \Sigma_r^{\xi,aa}, \mu_r^{\xi,t}). \quad (17)$$

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

V. Experimental Validation

A. Results

System architecture. The system makes use of a *sensing bracelet* mounted on the right wrist. The adopted sensor is a tri-axial accelerometer able to sense accelerations up to $3G$ and to code the information using 6 bit per axis, with a sampling frequency $F_s = 32Hz$. *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 $F_{cut} = 0.25Hz$, $A_{pass} = 0.001dB$, $A_{stop} = -100dB$, $F_{stop} = 2Hz$.

We recorded over 700 trials of 8 motion primitives from 16 volunteers (11 men and 5 women with age ranging from 19 to 83) to be used both for the creation and the validation of the models (see Table II). Among the chosen motion primitives there are simple 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

TABLE II

MOTION PRIMITIVES DATASET

Motion primitive	# recorded trials	# volunteers
Climb the stairs	102	10 (7M, 3F)
Drink from a glass	100	10 (6M, 4F)
Eat with fork and knife	14	4 (2M, 2F)
Get up from the bed	101	10 (5M, 5F)
Pour water in a glass	100	10 (6M, 4F)
Sit down on a chair	100	10 (6M, 4F)
Stand up from a chair	102	10 (6M, 4F)
Walk	100	10 (7M, 3F)

TABLE III

MODELLING DATASET

Model	# modelling trials	# volunteers	K_m
Climb the stairs	20	5	250
Drink from a glass	20	5	270
Eat with fork and knife	10	3	350
Get up from the bed	20	5	260
Pour water in a glass	20	5	365
Sit down on a chair	20	5	155
Stand up from a chair	20	5	168
Walk	20	5	170

execute in very different ways (*eating*). Moreover, *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 and final 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.

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

$$d(\hat{\Xi}^m, \Xi_w) = \frac{1}{N_\xi} \sum_{\xi=1}^{N_\xi} d_M^\xi(\hat{\Xi}^{\xi,a}, \Xi_w^{\xi,a}). \qquad (18)$$

Purpose of the experiment is to compare the accuracy rates of the two techniques, in order to evaluate the feasibility of the approach in an application with soft real-time constraints.

Table VI focuses on the *getting up* from the bed motion primitive and reports the true positive and true negative rates obtained using models composed of different numbers R_m of data-points. 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 proper 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 the selection of the data-points to be considered using non-constant and non-linear sampling.

TABLE IV

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

Model	τ_m	TP	TN
Climb the stairs	$2.0386e^{+3}$	20%	93.34%
Drink from a glass	$8.9103e^{+3}$	100%	83.34%
Get up from the bed	$1.0244e^{+4}$	60%	66.67%
Pour water in a glass	$8.8823e^{+3}$	100%	80%
Sit down on a chair	$2.8519e^{+3}$	0%	93.34%
Stand up from a chair	$3.7048e^{+3}$	60%	83.34%
Walk	$2.5884e^{+3}$	40%	70%

TABLE V

RECOGNITION ACCURACY WITH COMPARISON BASED ON MAHALANOBIS DISTANCE

Model	τ_m	TP	TN
Climb the stairs	16.8507	38.3%	85.78%
Drink from a glass	133.6197	90.91%	86.59%
Eat with fork and knife	85.5632	88.89%	96.67%
Get up from the bed	97.6461	93.33%	61.64%
Pour water in a glass	108.1443	95%	71.55%
Sit down on a chair	37.5798	0%	100%
Stand up from a chair	63.8687	35.71%	93.96%
Walk	46.1097	88.33%	79.91%

B. Discussion

At the best of our knowledge, there is no system for the automatic recognition of ADLs relying on two features only. A comparison with the accuracy rates achieved by similar systems relying on broader feature sets [10], [11], [12], [13], [15] for commonly considered activities such as *drinking*, *getting up*, *sitting down*, *standing up* and *walking* 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. More precisely, the results reported in 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 Gaussian Mixture Modeling and Regression and Mahalanobis distance subject to soft real-time constraints and ensuring good classification accuracy.

The comparison between the recognition rates reported in Table IV and Table V suggests that when the *model builder* is given a large modeling dataset (which is composed of trials

TABLE VI

RELATION BETWEEN THE RECOGNITION ACCURACY AND THE MODEL RESOLUTION R_m

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

showing different acceleration values and motion speed) the expected curves and associated covariance matrices retrieved by GMR merge the differences in the two aspects, so that the adoption of Dynamic Time Warping does not significantly increase the recognition accuracy. The test of the two comparison procedures using smaller modeling datasets is subject of current work.

Table IV and Table V also prove that longer models tend to prevail over shorter models, thus suggesting the need to always consider a balanced set of motion primitives. The low recognition rates of the *climbing*, *sitting down* and *standing up* motion primitives are due to the low thresholds they are associated to, suggesting that the corresponding modeling dataset are biased toward a specific behavior.

VI. CONCLUSIONS

Purpose of the article is to propose a comparison and classification procedure for a system for the automatic recognition of simple ADLs ensuring acceptable recognition accuracy with limited information at disposal. The proposed system assumes the information coming from a single wrist-placed tri-axial accelerometer and requires the extraction of two time-domain features only (namely gravity and body acceleration) from the acceleration data. The combined adoption of a modeling procedure based on Gaussian Mixture Modeling and Regression and a comparison procedure based on Dynamic Time Warping and Mahalanobis distance allows to obtain a recognition accuracy comparable with that of more complex systems, relying on broader feature sets.

REFERENCES

[1] 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.

[2] 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.

[3] F. Mastrogiovanni, A. Scalmato, A. Sgorbissa, and R. Zaccaria, "An integrated approach to context specification and recognition in smart homes," in *Proceedings of the Sixth International Conference on Smart Homes and Health Telematics (ICOST 2008)*, Ames, Iowa, June 2008.

[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] P.Turaga, R. Chellappa, V. Subrahmanian, and O. Udrea, "Machine recognition of human activities: a survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 11, pp. 1473–1488, 2008.

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

[7] 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.

[8] U. Maurer, A. Rowe, A. Smailagic, and D. Siewiorek, "Location and activity recognition using ewatch: a wearable sensor platform," *Ambient Intelligence in Everyday Life*, vol. 3864, pp. 86–102, 2006.

[9] 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.

[10] M. Mathie, B. Celler, N. Lovell, and A. Coster, "Classification of basic daily movements using a triaxial accelerometer," *Medical & Biological Engineering & Computing*, vol. 42, pp. 679–687, 2004.

[11] M. Lee, A. Khan, and T. Kim, "A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation," *Personal and Ubiquitous Computing*, vol. 15, no. 8, pp. 887–898, 2011.

[12] F. Allen, E. Ambikairajah, N. Lovell, and B. Celler, "Classification of a known sequence of motions and postures from accelerometry data using adapted gaussian mixture models," *Physiological Measurement*, vol. 27, no. 10, pp. 935–953, 2006.

[13] 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.

[14] 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.

[15] 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.

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

[17] 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.

[18] 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.

[19] C. Bouten, K. Koekkoek, M. Verduin, R. Kodde, and J. Janssen, "A tri-axial accelerometer and portable data processing unit for the assessment of daily physical activity," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 3, pp. 136–147, 1997.

[20] R. Herren, A. Sparti, K. Aminian, and Y. Schutz, "The prediction of speed and incline in outdoor running in humans using accelerometry," *Medicine & Science in Sports & Exercise*, vol. 37, no. 7, pp. 1053–1059, 1999.

[21] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," in *Proceedings of the Second International Conference on Pervasive Computing (PERVASIVE 2004)*, Vienna, Austria, April 2004.

[22] Y. Cho, Y. Nam, Y.-J. Choi, and W.-D. Cho, "Smartbuckle: Human activity recognition using a 3-axis accelerometer and a wearable camera," in *Proceedings of the Second International Workshop on Systems and Networking Support for Healthcare and Assisted Living Environments (HealthNet 2008)*, Breckenridge, Colorado, USA, June 2008.

[23] P. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[24] 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.

[25] 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.

[26] 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.