

# Processing of sEMG Signals for Online Motion of a Single Robot Joint through GMM Modelization

Riccardo Valentini<sup>1</sup> and Stefano Michieletto<sup>1</sup> and Fabiola Spolaor<sup>2</sup> and Zimi Sawacha<sup>2</sup> and Enrico Pagello<sup>1</sup>

**Abstract**— This paper evaluates the use of Gaussian Mixture Model (GMM) trained through Electromyography (EMG) signals to online estimate the bending angle of a single human joint. The parameters involved in the evaluation are the number of Gaussian components, the channel used for model, the feature extraction method, and the size of the training set. The feature extraction is performed through Wavelet Transform by investigating several kind of configuration. Two set of experimental data are collected to validate the proposed framework from 6 different healthy subjects. Trained GMMs are validated by comparing the joint angle estimated through Gaussian Mixture Regression (GMR) with the one measured on new unseen data. The goodness of the estimated date are evaluated by means of Normalized Mean Square Error (NMSE), while the time performances of the retrieval system are measured at each phase in order to analyze possible critical situations. Achieved results show that our framework is able to obtain high performances in both accuracy and computation time. The whole procedure is tested on a real humanoid robot by remapping the human motion to the robotic platform in order to verify the proper execution of the original movement.

## I. INTRODUCTION

The behavior of human limbs and its correlation with muscle pulses are subject of many researches since they stand at the basis of bio-mechanical solutions for locomotor diseases, prostheses and exoskeletons. Knowledge in the field can also improve humanoids research to obtain more natural looking motions.

A widely used approach is to analyze signals resulting from skeletal muscle activity called Electromyography (EMG) due to their strong relation to strength, location, time and effort of the movement. Surface Electromyography (sEMG) signals are usually preferred, as they are able to extract a similar information in a less invasive way. Some widespread techniques adopted for pattern recognition are Fourier Transform [1], Integral Absolute Value (IAV), variance and zero crossing [2], Mean Absolute Value (MAV) [3], Root Mean Square (RMS), Mean Power Frequency (MPF) [4], or as proposed in [5] full wave rectification, filtering and normalization. The major drawback of these transformation methods, especially fast and short-term

Fourier Transform, is that they assume signal to be stationary [6]. Alternative approaches based on Wavelet Transform have more effective results since EMG signals are non-stationary. Daubechies adopted Wavelet Transform to analyze time series that contain non-stationary power at many different frequencies [7]. Laterza [8] showed that Wavelet Transform is a valuable alternative to represent time frequency signals. His work highlighted several advantages. In fact, Wavelet Transform are a linear multi-resolution representation of the original signal without cross-terms affections. While Guglielminotti [9] found out the good matching properties between an EMG signal and its Wavelet shapes. Recent works reinforced advantages in using Wavelet Transform for EMG analysis [10][11]. In particular, Chowdhury [12] emphasized the good results obtained when adopting Daubechies functions by investigating and analyzing various research studies on Wavelet Transform. Moreover, Wavelet information can be synthesized to obtain a more compact representation by using statistical features as stated by Subasi [13][14].

The considered number of channels also plays an important role in EMG analysis of human motion. Increase the number of channels obviously increases the amount of available information, on the other hand the classification efficiency could be affected by the presence of noise. Tsenov [15] suggested that a number of channels between three and four connected to the considered movement may be sufficient to obtain good results for a specific joint.

Several models have been used to classify or predict human motion starting from EMG analysis. Deterministic techniques rely on accurate kinematic and dynamic of the human body, scaled to match the physical characteristics of the considered subject. Probabilistic approaches automatically train a specific model able to identify or replicate movements coming from a person. They used EMG features to feed different machine learning algorithms (i.e., Neural Networks (NN) [12], Gaussian Mixture Model (GMM) [16], Hidden Markov Model (HMM) [17]) to obtain a probabilistic approximation of the human movement. Scheme [18] studied a wide number of probabilistic techniques by comparing them from different point of view. As highlighted by Huang [16], a wide range of parameters must be considered when choosing a classifier: accuracy, efficiency, adaptation capability to novel data, optimization properties with respect unique patters, amount of training data. She proved that GMM satisfies each of the listed criteria for classification purposes.

<sup>1</sup>Riccardo Valentini, Stefano Michieletto, and Enrico Pagello are with Intelligent Autonomous Systems Lab (IAS-Lab), Department of Information Engineering (DEI), University of Padova, 35131 Padova, Italy

<sup>2</sup>Fabiola Spolaor, and Zimi Sawacha are with Movement Analysis Laboratory, Department of Information Engineering (DEI), University of Padova, 35131 Padova, Italy

riccardo.valentini@studenti.unipd.it,  
{michieletto, fspolaor, zimi.sawacha,  
epv}@dei.unipd.it

Recently, focus has been placed more on determining human continuous motions than on distinguishing high level actions. Han [19] developed a state space EMG model based on Hill Muscle Model for continuous estimation of elbow joint which however involves many physiological parameters and whose computational complexity of joint motion states makes it unsuitable for real-time applications. An EMG-driven virtual arm has been developed by Manal [20] which reproduces the movement according to Hill Muscle Model through recognition of muscles activation. An online estimation of elbow joint has been proposed in [21], where EMG patterns recognition is approached through a Hierarchical Projected Regression (HPR) algorithm that incrementally builds a tree-based knowledge library, whose components represent local regression models. In a previous work [22], we built a GMM-based probabilistic framework to verify its estimation capabilities of a joint bending angle by using as input EMG signals coming from several channels. We considered the whole signal representing the movement and at each instant we associated the corresponding bending angle and the EMG values for each channel. This model was able to obtain very good results for each subject involved in the study, but it was highly dependent by the time. Moreover, both training and testing phase of the model have been performed offline. In this work, we overcome the limitations of the previous framework by estimating a single-joint angle through a GMM trained with EMG features extracted through Wavelet Transform. Wavelet Transform allows us to consider only a small window of the whole signal. By using this technique, we are able to compute the synthesis value representing the signal just after a single raw sample has been collected. A Gaussian Mixture Regression (GMR) algorithm is used to retrieve the data from the trained model. This approach enables an autonomous extraction of the task-related information encoded in EMG signals, without loss of generality. Moreover, such a probabilistic framework based on Mixture of Gaussians (MoG) distributions only require a reduced number of parameters to be kept, resulting in lightweight models. Furthermore, a GMM/GMR probabilistic framework requires low training data to achieve good results. It also provides fast regression that perfectly matches with the use on an online application.

The remainder of the paper is structured as follows. Sect. II will describe the feature extraction procedure from EMG signals, the algorithm for estimating the joint angles, and the modelization technique. In Sect. III, a complete study composed by two experiments on different joint angles and subjects. The results of the study are also presented and discussed. Finally, Sect. IV will summarize the achieved results while proposing some future extensions of this work.

## II. METHODOLOGY

### A. Signal acquisition

Electromyography (EMG) signals were acquired with an active 8-channel wireless EMG system at 1000 Hz. The

eight EMG electrodes were placed on the left leg of each subject in order to cover the principal muscular groups active during two different tasks: a kick task and a stance phase of gait (step). The emg activity of the following muscles was recorded: *Rectus femoris*, *Vastus lateralis*, *Vastus medialis*, *Tibialis anterior*, *Gastrocnemius lateralis*, *Gastrocnemius medialis*, *Biceps femoris caput longus*, *Peroneus longus* for the kick, the *Tibialis anterior*, *Gastrocnemius lateralis*, *Peroneus longus*, *Extensor digitorum* for the step.

A 6 cameras stereophotogrammetric system (60 Hz, BTS srl) was used in order to acquire the 3D trajectories of 26 lower limbs anatomical landmarks according to Leardini *et al* [23]. The knee and the ankle joints kinematics were estimated synchronously to the Electromyography (EMG) signals and the flexion-extension angles were used to drive the analysis together with the emg signals. The knee joint flexion-extension was used to describe the kick task, whether the ankle joint was used for the step.

### B. Signal analysis

Wavelet Transform, in a similar way as Fourier Transform, has been used in numerous applications about processing and analysis of signals. Both these tools are used to look at the frequency components, anyway Wavelet Transform preserves the temporal aspect of the signal without resolution limits in frequency. Therefore, Wavelet analysis is able to extract signal information regarding both time and frequency. Wavelet Transform of a signal is a decomposition into several kernel functions called wavelets. These wavelets are generated using a base function, called *mother wavelet* or *wavelet function*. The mother wavelet is scaled and translated to provide multi-resolution analysis: wide wavelets are used for low frequencies, while narrow wavelets do the work at high frequencies. A complete and detailed description of Wavelet Transform and its supporting theory can be found in [24] and [25].

In this work, we are particularly interested in the selection of the mother wavelet. In fact, it plays a key role in the framework formulation because every mother wavelet yields to different results even when applied to the same signal. Chowdhury [12] successfully used Daubechies family (*db*) function to analyze sEMG signals. His research focused on the processing of sEMG and its use in different applications. The signal was process by means of some specific functions (*db2*, *db4*, *db6*, *db44* and *db45*) at decomposition level 4 in order to maintain the maximum amount of information. In a similar study, Phinyomark [11] was able to find good results by using *db7* as mother wavelet.

Wavelet Transform leads to a series of values representing the considered signal through a specific mother wavelet. In this study, we would like to synthesize the information provided by Wavelet Transform in a single value representing the wavelet decomposition as already tested in several works [26][11][13]. Mean Absolute Value (MAV) (Eq. 1), Root Mean Square (RMS) (Eq. 2), and Standard Deviation

(SD) (Eq. 3) methods have been selected due the good results reached in the literature.

$$\text{MAV} = \frac{1}{N} \sum_{k=1}^N |x_k| \quad (1)$$

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{k=1}^N x_k^2} \quad (2)$$

$$\text{SD} = \sqrt{\frac{\sum_{k=1}^N (x_k - E[x])^2}{N}} \quad (3)$$

where  $\{x_k\}$  are the wavelet coefficients.

Fig. 1 shows an example of the sEMG signals collected during this work, and the features resulting from wavelet analysis and computation of decomposition coefficients by applying MAV, RMS, and SD.

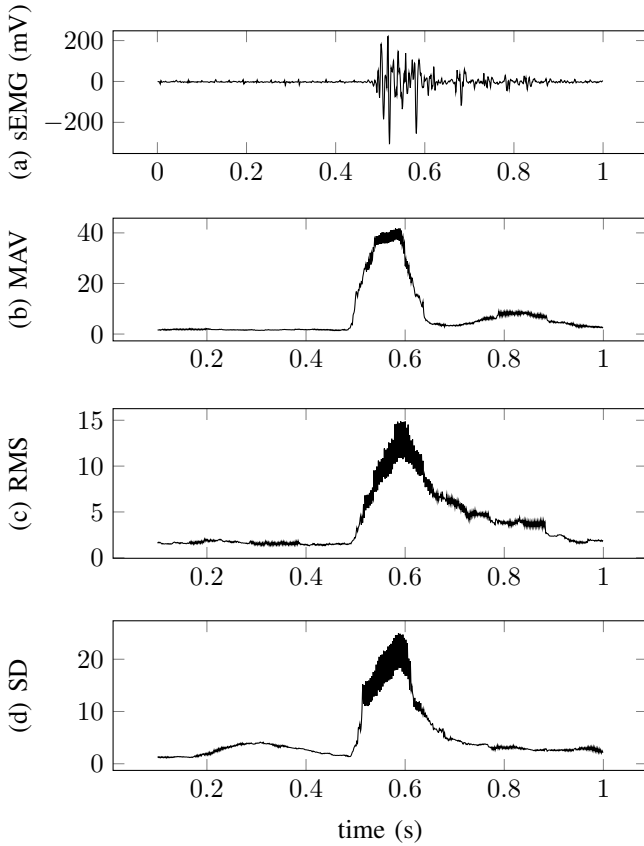


Fig. 1. Comparison between an original sEMG signal (a) and the corresponding computation through db2 Wavelet Transform and synthesis values, respectively MAV (b), RMS (c), and SD (d).

### C. Gaussian mixture model and regression

For modelization purposes, we exploited a stochastic approach in order to address the high variability of the input EMG signals. Information extracted from EMG was used as

input of a Gaussian Mixture Model (GMM) to estimate its correlation with the joint bending angle  $\alpha$ .

The aim of GMM is to obtain the weighted sum of  $K$  Gaussian components which best approximates the input dataset representing the set of activity trials used for the training. In this particular case, the total number of data samples was  $N = nT$ , where  $n$  is the number of trials used to train the system, and  $T$  is the number of observations acquired during each trial. A single data in input  $\zeta_j, 1 \leq j \leq N$  at the framework is described in Eq. 4.

$$\zeta_j = \{\xi(t), \alpha(t)\} \in \mathbb{R}^D \quad \xi(t) = \{\xi_c(t)\}_{c=1, \dots, C} \quad (4)$$

with:

- $\xi_c(t) \in \mathbb{R}$ , the value assumed from  $c^{th}$  EMG channel at the time instant  $t$ ;
- $\xi(t) \in \mathbb{R}^C$ , the set of value assumed from the considered channels,  $1 \leq C = |\xi| \leq 8$ , at the time instant  $t$ ;
- $\alpha(t) \in \mathbb{R}$ , the joint bending angle at the time instant  $t$ ;
- $2 \leq D \leq 9$ , the dimensionality of the problem.

With respect to our previous work [22], there is no direct use of the time instant  $t$  in the training data provided to build the model. The only constraint is a proper synchronization between different sensors collecting data. This solution allows us to process the information online, since there is no need of the whole signal before analyzing it.

The GMM was trained through the Expectation-Maximization (EM) algorithm [27]. The algorithm optimizes the parameters of the  $K$  Gaussian components by maintaining a monotone increasing likelihood during the local search of the maximum. This approach enables an autonomous extraction of the activity characteristic EMG signal while still maintaining an appropriate generalization. Finally, the resulting probability density function is computed:

$$p(\zeta_j) = \sum_{k=1}^K \pi_k \mathcal{N}(\zeta_j; \mu_k, \Sigma_k) \quad (5)$$

with  $\pi_k$  priors probabilities; and  $\mathcal{N}(\zeta_j; \mu_k, \Sigma_k)$  Gaussian distribution defined by  $\mu_k$  and  $\Sigma_k$ , respectively mean vector and covariance matrix of the  $k$ -th distribution.

The main drawback in the learning process lies in the EM requirement of a prior specification for the model complexity (i.e., the number of components  $K$ ). On one hand, an over-estimation of this parameter might lead to over-fitting and, consequently, to a poor generalization; on the other hand, an underestimation will result to poor regression performances. To deal with this issue we introduced an entropy based selection of the best number of components,  $K$ , in the GMM.

Several entropy based model selection techniques has been proposed in literature (e.g., Bayesian Information Criterion (BIC) [28], Akaike Information Criterion (AIC) [29], Minimum Description Length (MDL) [30], and Minimum Message Length (MML) [31]). Although, in [32] the authors proposed a specific criteria to estimate the value of the  $K$

parameter in the case of EMG signals, in this work we preferred a more standard approach based on BIC. In our experiments the whole learning process has been repeated with different GMM complexities by using BIC (Eq. 6) as index of model quality with respect to the number of components  $K$ .

$$S_{BIC} = -2\mathcal{L} + n_p \log N \quad (6)$$

with:

- $\mathcal{L} = \sum_{j=1}^N \log(p(\zeta_j|\theta))$ , the log-likelihood for the considered model  $\theta$ ;
- $n_p = (K-1) + K(D + \frac{1}{2}D(D+1))$ , the number of free parameters required for a mixture of  $K$  components with full covariance matrix.

The log-likelihood measures how well the model fits the data, while the second term is introduced to avoid data overfitting and maintain the model general enough. In our experiments the best BIC value was obtained with  $K = 15$  components.

The Gaussian Mixture Regression (GMR) has been used to retrieve a smooth generalized version of the signal encoded in the associated GMM. So that, the conditional expectation of the considered joint angle  $\hat{\alpha}$  is calculated from the consecutive EMG signals  $\xi$  known a priori. As we already said, the  $k$ -th Gaussian component is defined by the parameters  $(\pi_k, \mu_k, \Sigma_k)$ , where:

$$\mu_k = \{\mu_{p,k} \mu_{\alpha,k}\} \quad \Sigma_k = \begin{bmatrix} \Sigma_{p,k} & \Sigma_{p\alpha,k} \\ \Sigma_{\alpha p,k} & \Sigma_{\alpha,k} \end{bmatrix} \quad (7)$$

with  $\mu_p$  and  $\Sigma_p$  respectively the mean and the covariance of the known a priori information. The conditional expectation and its covariance can be estimated respectively using Equation 8 and 9.

$$\hat{\alpha} = E[\alpha|\xi] = \sum_{k=1}^K \beta_k \hat{\alpha}_k \quad (8)$$

$$\hat{\Sigma}_s = Cov[\alpha|\xi] = \sum_{k=1}^K \beta_k^2 \hat{\Sigma}_{\alpha,k} \quad (9)$$

with:

- $\beta_k = \frac{\pi_k \mathcal{N}(\xi_c|\mu_{p,k}, \Sigma_{p,k})}{\sum_{j=1}^K \pi_j \mathcal{N}(\xi_c|\mu_{p,j}, \Sigma_{p,j})}$ , the weight of the  $k$ -th Gaussian component through the mixture;
- $\hat{\alpha}_k = E[\alpha_k|\xi] = \mu_{\alpha,k} + \Sigma_{\alpha p,k} (\Sigma_{p,k})^{-1} (\{\xi\} - \mu_{p,k})$ , the conditional expectation of  $\alpha_k$  given  $\{\xi\}$ ;
- $\hat{\Sigma}_{\alpha,k} = Cov[\alpha_k|\xi] = \Sigma_{\alpha,k} + \Sigma_{\alpha p,k} (\Sigma_{p,k})^{-1} \Sigma_{p\alpha,k}$ , the conditional covariance of  $\alpha_k$  given  $\{\xi\}$ .

Thus, the generalized form of the motions  $\hat{\zeta} = \{\xi, \hat{\alpha}\}$  required only the weight, mean and covariance of the Gaussian components calculated through the EM algorithm.

#### D. Procedure effectiveness

Finally, we exploited the Normalized Mean Square Error (NMSE) in order to evaluate the effectiveness of the GMM-based system. The selected function measures the Goodness

of Fit (GoF) by using as metric the NMSE between test and reference data, in our case  $\hat{\alpha}$  (the data estimated through the GMR) and  $\alpha$  (the angle calculated by means of the motion capture system):

$$GoF_{NMSE}(t) = 1 - \frac{MSE(\hat{\alpha}(t))}{MSE(E[\alpha(t)])} \quad (10)$$

with:

- $t$ , temporal instant from the beginning of the trial (ms);
- $\hat{\alpha}(t)$ , estimated angle at the instant  $t$ ;
- $MSE(x(t)) = \|\alpha(t) - x(t)\|^2$ , where  $\alpha(t)$  angle calculated through the motion capture at the instant  $t$ ;
- $E[\alpha(t)]$ , mean along the time of the angles given by the motion capture.

By using this formula, the GoF costs vary between  $-\infty$  (bad fit) to 1 (perfect fit). The idea is to obtain a value representing the goodness of fit and not the error from the real value. Moreover, in this case, zero represents the value reached from a straight line in fitting the reference.

### III. EXPERIMENTS

Six healthy volunteers (S1-S6; age  $45.6 \pm 17.6$ ; one female) participated in the experiment. No motor related problems have been reported. The study was approved by the local ethic committee of Padova Hospital.<sup>1</sup> Two different experiments were performed.

#### A. Experiment 1: Kick

Subjects (S1-S3) were asked to naturally kick a ball (diameter 17 cm) from a sitting position. The ball was positioned on the floor at a fixed distance (2 cm) from the foot. After each kick, an operator was in charged of the ball reposition. As additional behavioral and motivational task, we asked to the subjects to try to shot the ball in a goal in front of them (distance 350 cm, width 122 cm, height 76.2 cm). It is worth to notice that the task was fully self-paced. Each participant performed several repetitions of the aforementioned task over a single recording session (day), 60 (train) + 5 (test) kicks were extracted for each subject.

Original signals have been elaborated to isolate every kick selecting a number of samples sufficient to cover the whole movement, in our case 2000 samples. In order to extract the features corresponding to the instant  $t$ , we considered the portion of sEMG signal between  $t - w$  and  $t$ , where  $w = 200ms$  is the window of significant past values used to compute the Wavelet transform. The window width has been selected looking at [33]. In his work, Smith has exhibited that values belonging to the interval 150-250 ms are significantly better with respect to other choices when dealing with EMG signals. The described procedure has been subsequently reiterated to cover the whole movement, obtaining 1800 features for each kick in the dataset. Therefore, a single feature is the Wavelet Transform obtained from a certain

<sup>1</sup>WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects

*mother wavelet* of the last  $w$  samples read at the instant  $t$ . Only detail coefficients of first level decomposition have been taken as suggested in [34] and [26]. From these coefficients a synthesis value has been computed using specific functions to represent the feature, as exposed in Sect. II-B.

A preparatory study has been conducted by looking at the most informative EMG channels. The objective of this study was to compare the muscles connected to the system results with the biological information about the muscles controlling the knee. The comparison aimed to verify the reliability of our framework and to give us a quantitative idea of the influence in the movement of each considered channel. The db2 *mother wavelet* and MAV synthesis feature have been applied to the raw signal provided from every single channel. The resulting values have been associated to the corresponding knee bending angle along time. A model for each channel has been trained by using the 60 training examples for every subject obtaining a total of 24 models. Estimated knee angles have been retrieved through GMR by using only the synthesis EMG values from the 5 testing kicks as input. GoF with respect to measured knee angles gave us the fitting score for each channel. We used the mean between the 3 subjects to sort them. Results are showed in Tab. I.

#	Channel	Subject 1	Subject 2	Subject 3	Mean
1	Rectus Femoris	0.5917	0.4523	0.1727	0.4056
2	Vastus Lateralis	0.7401	0.5403	0.5426	0.6077
3	Vastus Medialis	0.5917	0.8093	0.6266	0.6759
4	Tibialis anterior	0.0674	-0.2407	0.4706	0.0991
5	Gastrocnemius lateralis	-0.5058	0.3614	-0.3206	-0.155
6	Gastrocnemius medialis	0.5107	0.0233	0.191	0.2417
7	Biceps femoris caput longus	-0.0022	0.5074	-0.0281	0.1590
	Peroneus longus	0.1567	-0.0734	0.0380	0.0682

TABLE I

GoF VALUES RESULTING FROM TESTS USING ONLY ONE CHANNEL FOR TRAINING GMM WITH FEATURES EXTRACTED THROUGH DB2 AND MAV OF COEFFICIENTS.

The 3 most significant channels have been selected, according to the results coming from the preliminary study. Moreover, *Vastus lateralis*, *Vastus medialis* and *Tibialis anterior* are the most informative channels also from the physiological point of view. Using these channels, we investigated a series of mother wavelets and a set of synthesis functions in order to compare their performances when matched with our GMM/GMR framework. The mother wavelets come from Daubechies family, namely db2, db4, db6, db7, db44, and db45. The synthesis functions tested are MAV (Tab. II), RMS (Tab. III), and SD (Tab. IV). As before, we used the GoF mean between the 3 subjects as informative criteria to sort the results from the different models.

The best result has been achieved by using db2 as mother wavelet and MAV features, corresponding to a  $gof_{NMSE} = 0.8093$ . Good results came also from db7 coupled with RMS, the two exploited a slightly lower  $gof_{NMSE} = 0.7982$ . Moreover, very different results came from each subject when using a specific mother wavelet. For example, the

Mother Wavelet	Subject 1	Subject 2	Subject 3	Mean
db2	0.8256	0.8484	0.7539	0.8093
db4	0.8021	0.7779	0.4415	0.6738
db6	0.8051	0.8397	0.3710	0.6719
db7	0.7732	0.8544	0.6568	0.7615
db44	0.7574	0.5983	0.6234	0.6597
db45	0.5426	0.7828	0.3928	0.5727

TABLE II

GoF ERROR OF THE THREE SUBJECTS OF DATASET FOR EACH DAUBECHIES WAVELET TESTED WITH MAV FEATURE.

Mother Wavelet	Subject 1	Subject 2	Subject 3	Mean
db2	0.5701	0.7409	0.8583	0.7231
db4	0.5806	0.8008	0.5825	0.6546
db6	0.5455	0.6945	0.8531	0.6977
db7	0.7177	0.8092	0.8678	0.7982
db44	0.6306	0.7931	0.6455	0.6897
db45	0.7258	0.8345	0.7572	0.7725

TABLE III

GoF ERROR OF THE THREE SUBJECTS OF DATASET FOR EACH DAUBECHIES WAVELET TESTED WITH RMS FEATURE.

bad performance of S3 with db6 and MAV or the good value obtained from S2 with db4 and RMS. We guessed that some mother wavelet could be more suitable for some subjects and less for others. Anyway, the number of examples populating our dataset are actually too small to lead to conclusive assumptions about this point. Instead, we could be more accurate about the use of SD as feature. It provided  $gof_{NMSE}$  values lower with respect to the other ones of at least 0.1, therefore SD is probably not suitable for this kind of analysis.

Finally, we set both channels (*Vastus lateralis*, *Vastus medialis*, *Tibialis anterior*) and feature extraction method (db2 + MAV) accordingly to the results of the previous tests and we focused on the cardinality of the training set. We increased the number of examples given to the system as input in order to track the trend of the GoF for each subject (Fig. 2).

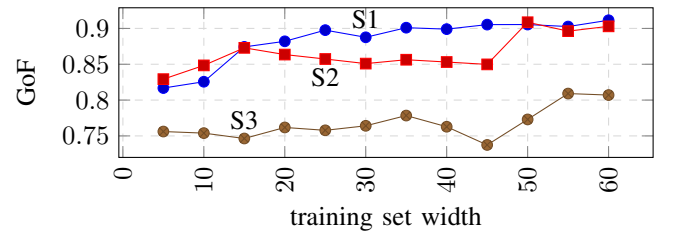


Fig. 2. Estimation accuracy with different size of the training set.

As expected, results showed an increasing tendency of GoF values in correspondence of a bigger size on the training dataset. Anyway, the trend is quite slow and the advantages of using a small amount of data probably overcomes improvements in term of performances.

Mother Wavelet	Subject 1	Subject 2	Subject 3	Mean
db2	0.7252	0.7171	0.6139	0.6854
db4	0.5967	0.8694	0.6563	0.7075
db6	0.4321	0.7946	0.5203	0.5823
db7	0.6760	0.8494	0.5909	0.7054
db44	0.6150	0.7642	0.6279	0.6690
db45	0.3603	0.7446	0.5189	0.5413

TABLE IV  
GoF ERROR OF THE THREE SUBJECTS OF DATASET FOR EACH  
DAUBECHIES WAVELET TESTED WITH SD FEATURE.

The framework has been subsequently paired with a real robotic platform, namely an Aldebaran NAO (Fig. 3), in order to acquire some measures by using an actual environment. The human knee angle retrieved through the GMR has been remapped to the NAO knee joint in order to respect the robot limitations. Therefore, we used the testing EMG signals

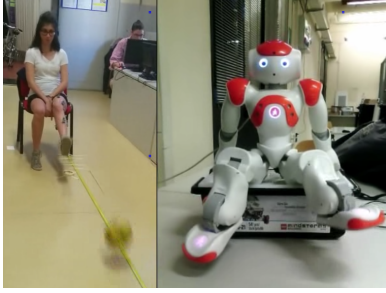


Fig. 3. Motion of a NAO robot controlled by EMG signals.

collected from the three subjects to directly actuate a single robotic joint by means of the GMM already generated offline from the training data. An accurate analysis of the computing time at each step (Tab. V) has been conducted by using an Intel® 64-bit computer with *i3* quad core CPU of 2.13 GHz and 4 GB of RAM. Computation of Wavelet Transform, feature extraction and regression have been tested for each subject. As highlighted by the collected data, our software is able to compute the pose messages for the robot in 2.4 ms, although the sending rate has been reduced to satisfy NAO bound of 50 Hz. The resulting robot movement is smooth and quite consistent with respect to the human motion. Nevertheless, we are planning some tests on different robots with a more demanding input framerate.

#### B. Experiment 2: Step

Subjects (S4-S6) were asked to walk at self selected speed on a 8m walkway. Several gait trials were acquired and, out of them, 8 (train) + 2 (test) stance phases of gait were extracted for each subject.

The aim of this experiment was to confirm the results obtained from the first one in a more challenging scenario. We considered a very complex but common movement in which a greater amount of muscles is used and therefore activated. Moreover, we focused on a different joint to be sure

Step	Method	Time( $\mu$ s)
Wavelet Transform	db2	581.3955
	db4	1827.9829
	db6	1596.3840
	db7	2322.7949
Feature extraction	MAV	0.8448
	RMS	2.2616
	SD	2.7085
Regression	-	1774.6145
Angle remapping	-	7.4357
Minimum sum	db2 + MAV	2408.9559
Maximum sum	db7 + SD	4100.1179

TABLE V  
ANALYSIS OF THE COMPUTATIONAL TIME ( $\mu$ S) NEEDED FROM THE  
FRAMEWORK AT EACH PHASE.

that the model does not rely on some peculiar characteristics of the knee. The preliminary step has been intentionally skipped in order to use the 3 most informative channels from a biological point of view, while the mother wavelet and the synthesis function came from the previous experiment. All the available training steps has been used to compute the GMM representing the motion for each subject. It is worth to notice that the speed of the gait influence directly the number of samples considered for the subjects. The more rapidly they walked, the less samples we obtained for analyzing the movement. Again, GoF has been used to evaluate the framework performances (Tab. VI).

ID subject	samples	NMSE
S4	651	0.9052
S5	550	0.8344
S6	701	0.8170

TABLE VI  
GoF OF ANKLE MOTION ESTIMATION DURING THE GAIT.

Despite the lower number of repetitions and the lower number of samples due to shorter movement, the results are consistent with respect to the previous dataset emphasizing the performances of the proposed framework when facing a small amount of input data.

## IV. CONCLUSION

This paper proposed a method to estimate a single-joint angle by means of Surface Electromyography signals for online purposes. Wavelet Transform has been chosen in order to extract features from the raw signal. Several mother wavelet function has been individuated and tested from some successfully used previously in literature to select the most informative one. Different synthesis functions have also been analyzed alongside the mother wavelet. Physiological

information has been encoded through Gaussian Mixture Model, while joint angles related to new sequences of unseen Electromyography data has been estimated by using Gaussian Mixture Regression on two different set of motions, namely kick and gait. Results exhibited an high accuracy in joint angle estimation ( $GoF > 0.8$ ) for the 6 considered subjects even using very few (5-10) examples for training purposes. The estimated joint angles have been remapped onto a robotic joint in order to online control a humanoid platform. The framework was able to estimate a new joint angle in 2.4 ms, providing sufficient data to obtain a smooth robot motion and even for being suitable for real-time applications.

As further work, we plan to apply the described method to a multiple joint motion and to test the resulting framework on a higher number of subjects and on a different robotic platform. Furthermore, we will investigate the time that exist between the EMG sample and the actual human motion, also called Electro Mechanical Delay (EMD). Modeling such delay will guarantee the performances of the whole system while introducing some space for safety controls in case of human interaction, for example when applied to a prosthesis or an exoskeleton.

## REFERENCES

- [1] M. Costa, L. Pereira, R. Oliveira, R. Pedro, T. Camata, T. Abrao, M. Brunetto, and L. Altinari, "Fourier and wavelet spectral analysis of EMG signals in maximal constant load dynamic exercise," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pp. 4622–4625, Aug 2010.
- [2] S. Lee and G. Saridis, "The control of a prosthetic arm by EMG pattern recognition," *Automatic Control, IEEE Transactions on*, vol. 29, pp. 290–302, Apr 1984.
- [3] C. Loconsole, S. Dettori, A. Frisoli, C. A. Avizzano, and M. Bergamasco, "An EMG-based approach for on-line predicted torque control in robotic-assisted rehabilitation," in *Haptics Symposium (HAPTICS), 2014 IEEE*, pp. 181–186, IEEE, 2014.
- [4] T. Lalitharatne, Y. Hayashi, K. Teramoto, and K. Kiguchi, "A study on effects of muscle fatigue on EMG-based control for human upper-limb power-assist," in *Information and Automation for Sustainability (ICIAFS), 2012 IEEE 6th International Conference on*, pp. 124–128, Sept 2012.
- [5] M. Sartori, M. Reggiani, E. Pagello, and D. Lloyd, "Modeling the Human Knee for Assistive Technologies," *Biomedical Engineering, IEEE Transactions on*, vol. 59, pp. 2642–2649, Sept 2012.
- [6] A. Ismail and S. Asfour, "Continuous wavelet transform application to EMG signals during human gait," in *Signals, Systems and Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference on*, vol. 1, pp. 325–329 vol.1, Nov 1998.
- [7] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *Information Theory, IEEE Transactions on*, vol. 36, pp. 961–1005, Sep 1990.
- [8] F. Laterza and G. Olmo, "Analysis of EMG signals by means of the matched wavelet transform," *Electronics Letters*, vol. 33, pp. 357–359, Feb 1997.
- [9] P. Guglielminotti and R. Merletti, "Effect of electrode location on surface myoelectric signal variables: a simulation study," in *9th Int. Congress of ISEK*, vol. 188, 1992.
- [10] U. Sahin and F. Sahin, "Pattern recognition with surface EMG signal based wavelet transformation," in *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, pp. 295–300, 2012.
- [11] A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "Application of Wavelet Analysis in EMG Feature Extraction for Pattern Classification," *Measurement Science Review*, vol. 11, no. 2, pp. 45–52, 2011.
- [12] R. H. Chowdhury, M. B. I. Reaz, M. A. B. M. Ali, A. A. A. Bakar, K. Chellappan, and T. G. Chang, "Surface Electromyography Signal Processing and Classification Techniques," *Sensors*, vol. 13, no. 9, pp. 12431–12466, 2013.
- [13] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.
- [14] A. Subasi, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction," *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 227–244, 2007.
- [15] G. Tsenov, A. Zeghib, F. Palis, N. Shoylev, and V. Mladenov, "Neural networks for online classification of hand and finger movements using surface emg signals," in *Neural Network Applications in Electrical Engineering, 2006. NEUREL 2006. 8th Seminar on*, IEEE, 2006.
- [16] Y. Huang, K. B. Englehart, B. Hudgins, and A. D. Chan, "A gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses," *Biomedical Engineering, IEEE Transactions on*, vol. 52, no. 11, pp. 1801–1811, 2005.
- [17] A. D. Chan and K. B. Englehart, "Continuous myoelectric control for powered prostheses using Hidden Markov Models," *Biomedical Engineering, IEEE Transactions on*, vol. 52, no. 1, pp. 121–124, 2005.
- [18] E. J. Scheme, K. B. Englehart, and B. S. Hudgins, "Selective classification for improved robustness of myoelectric control under nonideal conditions," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 6, pp. 1698–1705, 2011.
- [19] J. Han, Q. Ding, A. Xiong, and X. Zhao, "A state space emg model for the estimation of continuous joint movements," 2014.
- [20] K. Manal, R. V. Gonzalez, D. G. Lloyd, and T. S. Buchanan, "A real-time EMG-driven virtual arm," *Computers in Biology and Medicine*, vol. 32, no. 1, pp. 25 – 36, 2002.
- [21] Y. Chen, X. Zhao, and J. Han, "Hierarchical projection regression for online estimation of elbow joint angle using EMG signals," *Neural Comput and Applic*, vol. 23, p. 1129–1138, 2013.
- [22] S. Michieletto, M. Antonello, L. Tonin, R. Bortoletto, F. Spolaor, E. Pagello, and E. Menegatti, "GMM-based single-joint angle estimation using EMG signals," in *Proceedings of the 13th Conference on Intelligent Autonomous Systems (IAS-13)*, 2014.
- [23] A. Leardini, Z. Sawacha, G. Paolini, S. Ingrassio, R. Nativio, and M. G. Benedetti, "A new anatomically based protocol for gait analysis in children," *Gait & posture*, vol. 26, no. 4, pp. 560–571, 2007.
- [24] L. Chun Lin, "A Tutorial of the Wavelet Transform,"
- [25] Y. Sheng, "Wavelet transform," *The transforms and applications handbook*, pp. 747–827, 1996.
- [26] I. Elamvazuthi, G. Ling, K. Nurhanim, P. Vasant, and S. Parasuraman, "Surface electromyography (sEMG) feature extraction based on Daubechies wavelets," in *Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on*, pp. 1492–1495, June 2013.
- [27] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, vol. 39, no. 1, pp. 1–38, 1977.
- [28] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.
- [29] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *2nd International Symposium on Information Theory*, pp. 267–281, 1973.
- [30] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2743–2760, 1998.
- [31] C. S. Wallace and D. L. Dowe, "Minimum message length and kolmogorov complexity," *The Computer Journal*, vol. 42, no. 4, pp. 270–283, 1999.
- [32] J. Chu and Y. J. Lee, "Conjugate prior penalized learning of Gaussian mixture models for EMG pattern recognition," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007.
- [33] L. H. Smith, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 19, no. 2, pp. 186–192, 2011.
- [34] F. A. Mahdavi, S. A. Ahmad, M. H. Marhaban, R. Mohammad, and A. T., "Surface Electromyography Feature Extraction Based on Wavelet Transform," *International Journal of Integrated Engineering*, vol. 4, no. 3, pp. 1–7, 2012.