Rakeness-based Compressed Sensing of Surface ElectroMyoGraphy for Improved Hand Movement Recognition in the Compressed Domain

Rakeness-based Compressed Sensing of Surface ElectroMyoGraphy for Improved Hand Movement Recognition in the Compressed Domain

Alex Marchioni*, Mauro Mangia†, Fabio Pareschi†‡, Riccardo Rovatti*† and Gianluca Setti†§
*DEI, †ARCES, University of Bologna, Italy - Email: {alex.marchioni, mauro.mangia2, riccardo.rovatti}@unibo.it
‡ENDIF, University of Ferrara, Italy - Email: fabio.pareschi@unife.it
§DET, Politecnico di Torino, Italy - Email: gianluca.setti@polito.it

***Abstract*—Surface electromyography (sEMG) waveforms are widely used to generate control signals in several application areas, ranging from prosthetic to consumer electronics. Classically, such waveforms are acquired at Nyquist rate and digitally transmitted trough a wireless channel to a decision/actuation node. This causes large energy consumption and is incompatible with the implementation of ultra-low power acquisition nodes. We already proposed Compressed Sensing (CS) as a low-complexity method to achieve substantial energy saving by reducing the size of data to be transmitted while preserving the information content. We here make a significant leap forward by showing that hand movements recognition task can be performed directly in the compressed domain with a success rate greater than 98% and with a reduction of the number of transmitted bits by two order of magnitude with respect to row data.**

I. Introduction

Surface ElectroMyoGraphy (sEMG) signals are fundamental in the analysis of muscle activity. Since they are composed of spike trains that neurons send to muscles, they are commonly used in rehabilitation medicine for controlling prosthetic devices [1], [2]. More recently, sEMG signals has been used in consumer applications such as sign language recognition [3] and remote control [4], where the device recognizes performed hand gestures by detecting patterns in muscles activation.

A classical hand movement recognition system based on sEMG signals is quite complex, since it is composed by three main stages: 1) signal acquisition (including electrodes transduction, signal conditioning and A/D conversion), 2) feature extraction and 3) classification. Notwithstanding such a complexity, wearability is a key feature for the system – especially in commercial applications– which set stringent constraints both in size and power consumption of the device. To solve the impasse, a typical approach relies on decoupling signal acquisition from gesture recognition task, the latter being performed on a more powerful hub. Yet, since sEMG signals need to be acquired from multiple electrodes and at 2-4 kHz frequency [1], [2], even the sole data transmission operation requires a great amount of energy that makes the acquisition node incompatible with a long battery life.

A possible solution to reduce the data size is to transmit only the features needed for classification which are extracted locally. In [5], the authors proposed a solution based on the transmission via impulse-radio ultra-wide-band (IR-UWB) of the locally computed Willison amplitude to extract the muscle force. This allows an overall power consumption reduction by more than an order of magnitude.

Yet, several applications of sEMG signals requires the knowledge of more sophisticated features. To cope with this, in [6], Compressed Sensing (CS) –as well as its improved version based on acquisition sequence maximizing *rakeness* [7], i.e. the signal energy acquired by every CS sample– was applied to sEMG signal acquisition in hand movement recognition context with the aim of reducing the size of data and still have access to the whole information content. CS is, in fact, very interesting for wearable applications since it allows considerable data compression with a very low computational at transmitter side. Complexity is moved at the receiver/hub side, where power consumption and size constraints are less stringent.

In this work we extend the CS approach in [6] by adopting two strategies that, when combined, are able to reduce both latency and data size without significant accuracy loss. First, we show that, in a hand movement recognition system, latency can be reduced by avoiding CS decoding (which is often the most time-consuming stage in CS processing chain) and extracting the features needed for classification directly in the compressed domain, similarly to what was done [8] for EEG signals. Then we demonstrate that a further substantial data compression can be achieved by reducing the number of bit used (and transmitted) to represent the compressed information without appreciably altering the performances of the classifier.

The paper is organized as follows: Section II briefly introduces (rakeness-based) CS. In Section III, a hand movement recognition system is first presented; then CS blocks are inserted in the processing chain to drastically reduce the size of the transmitted data and to perform the desired classification in the compressed domain. System performances are shown in Section IV and finally conclusions are drawn in Section V.

II. Rakeness-based Compressed Sensing

Compressed Sensing (CS) is an acquisition technique in which chunks of an input signal are represented with fewer scalars than the intrinsic limit indicated by the Nyquist-Shannon theorem. This is possible assuming that the signal to process is *sparse*, i.e. a proper basis exists such that the projection of any input waveform over that basis has only few terms significantly different from zero.

Let us consider the discrete-time representation of the input signal $x \in \mathbb{R}^n$, where n is the number of samples at Nyquist rate for given time window. Let also $D \in \mathbb{R}^{n\times n}$ be the sparsity basis and $\xi \in \mathbb{R}^n$ the vector containing the projection of x on D such that $x = D\xi$. If, for any possible instance x, the corresponding ξ has at most κ non-null elements ($\kappa \ll n$), then the class of signals is *κ-sparse* and compression can be achieved.

The idea grounding CS is to capture the information contained in x with its $m \leq n$ projections on a set of suitable *sensing vectors* such that:

$$y = Ax + \nu = AD\xi + \nu \tag{1}$$

where $y \in \mathbb{R}^m$ is the vector containing these m projections (also called *measurements*), $A \in \mathbb{R}^{m\times n}$ is the acquisition matrix whose rows are the m sensing vectors, and ν accounts for noise and non-idealities in acquisition process.

It is possible to recover x from y by looking for the sparsest vector over all possible ξ that satisfy (1), which can be shown to be equivalent to solve the following optimization problem:

$$\hat{\xi} = \underset{\xi\in\mathbb{R}^n}{\operatorname{argmin}} \|\xi\|_1 \qquad \text{s.t. } \|AD\xi - y\|_2 < \epsilon \tag{2}$$

where $\|\cdot\|_p$ is the l_p norm, ϵ accounts for the effect of ν and the reconstructed signal can be written as $\hat{x} = D\hat{\xi}$. From [9] it is known that, if the elements of A are instances of independent and identically distributed (i.i.d.) random Gaussian values or random antipodal values [10], reconstruction is guaranteed by adopting $m = \mathcal{O}(\kappa \log(n/\kappa))$.

Standard CS acquisition can be improved by exploiting an additional prior on the class of signals x, called *localization* $\mathfrak{L}_x$ [11], which measures the non-uniformity of the signal energy distribution. By indicating with $\mathbf{E}[\cdot]$, $\cdot^\top$ and $\mathrm{tr}(\cdot)$, the expectation, the transpose operator and the matrix trace, $\mathfrak{L}_x$ can be evaluated starting from an estimation of the input signal correlation matrix $C_x = \mathbf{E}[xx^\top]$, and by computing the deviation from the correlation matrix of a white process (i.e. the identity matrix I_n), that is

$$\mathfrak{L}_x = \frac{\mathrm{tr}({C_x}^2)}{\mathrm{tr}(C_x)^2} - \frac{1}{n} \tag{3}$$

Such an additional prior is used by rakeness-based CS [7] to maximize the average energy ρ collected by the measurements by adapting the correlation profile of the sensing vectors (rows of A) to C_x. This can be achieved by solving

$$\begin{aligned} \rho = \max_{C_A}\ & \mathrm{tr}(C_A C_x) && (4)\\ \text{s.t. } & {C_A}^\top = C_A && (5)\\ \text{s.t. } & C_A \succeq 0 && (6)\\ \text{s.t. } & \mathrm{tr}(C_A) = 1 && (7)\\ \text{s.t. } & \mathfrak{L}_A \leq \tau\mathfrak{L}_x && (8) \end{aligned}$$

where (5) and (6) ensure that C_A is a correlation matrix (i.e., symmetric and positive semidefinite), (7) normalizes the energy of the rows of A and (8), with $\tau \in [0, 1]$, sets the localization level of the sensing sequences with respect to $\mathfrak{L}_x$[1].

The analytic solution of the above optimization problem is given by

$$C_A = \frac{1}{2}\left(\frac{C_x}{\mathrm{tr}(C_x)} - \frac{I_n}{n}\right) \tag{9}$$

With C_A is possible to generate the sensing vectors by using jointly-Gaussian variables or by exploiting available techniques [12] capable of synthesize sequences formed by antipodal values $\{-1, +1\}$ with a given correlation.

III. System design and Hand Movement Recognition

The hand movements recognition task we consider requires to distinguish between three specific gestures, which are extremely relevant in upper limb prosthetic control: i) hand relaxed in natural position, ii) hand closed as in a power grasp, iii) hand open with fingers extension.

The dataset we use consists of a single record (65s duration) of an experiment where the subject, a healthy 32 years old male, was asked to perform 4 repetitions of the 3 movements for approximately 5 seconds each. The sEMG signal is acquired by mean of an armband equipped with 16 equally spaced electrodes and arranged in 2 rows. The 16 traces are amplified and filtered before being digitized with a 12 bit ADC working at 4 kHz sampling frequency.

Classification is first performed by using the root mean square (RMS) value computed on non-overlapping windows. As in real time applications latency should not exceed 150 ms [13], we choose windows of length 256 samples (that is 64 ms) to account for further delay along the processing chain. Our classifier is an artificial neural network (NN) with a hidden layer of 20 neurons, whose inputs are the RMS values computed on the 16 traces and the output classes correspond to the 3 hand movements. The NN is trained with the Leveberg-Marquardt backpropagation algorithm. The test set is constructed by randomly choosing three different movements among all possible instances and collecting all corresponding samples. The remaining data constitutes the training set. Performances are evaluated by averaging over 1000 different configurations. The figure of merit used to assess the classification performances is the Average Recognition Rate:

$$\mathrm{ARR} = \mathbf{E}\left[\frac{\#\text{ correct recognitions}}{\#\text{ signal windows}}\right]$$

By extracting RMS values directly from the acquired signal an $\mathrm{ARR}_x = 99.8\%$ is obtained[2].

To improve performance in terms of reducing energy for wireless transmission we add the CS encoder and decoder between acquisition and feature extraction stages, as shown in Figure 1, with the aim of reducing the amount of transmitted

[1] τ is needed to guarantee the sufficient randomness of the sensing waveforms to reconstruct uncommon signal instances and a usal choice is $\tau = 0.25$.

[2] Results reported here have been obtained using an extended set of data with respect to our previous work [6]. This had the consequence of increasing classification accuracy from 95% to 99.0%

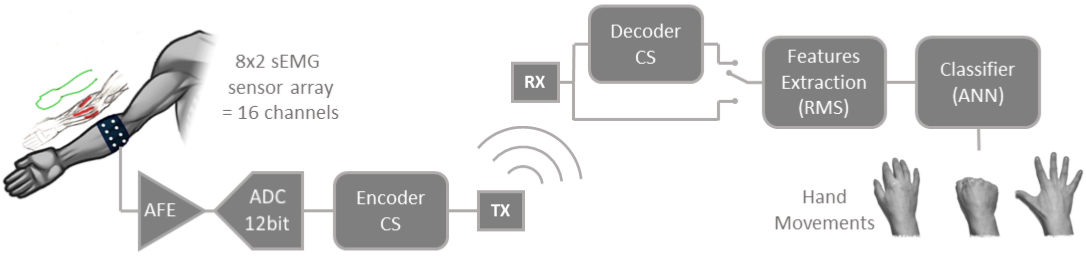

Fig. 1. Block diagram of the hand movement recognition system.

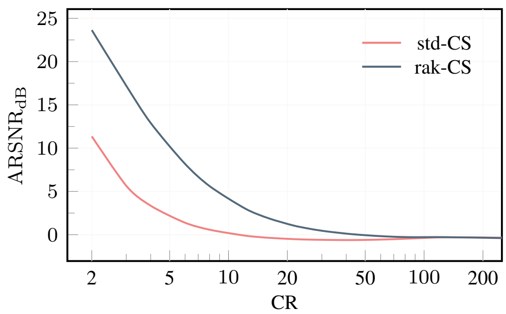

Fig. 2. Reconstruction performance in term of ARSNR as function of the compress ratio CR for both standard CS and Rakeness-based CS.

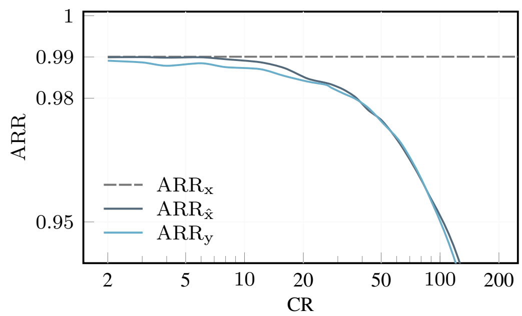

Fig. 3. Classification performance in terms of ARR as function of the compress ratio CR. The plot shows results obtained with RMS computed on reconstructed signal $\mathrm{ARR}_{\hat{x}}$ and directly on the measures ARR_y, with the upper bound $\mathrm{ARR}_x = 99.9\%$ obtained transmitting signal acquired at Nyquist rate without compression.

data. The CS coding block compresses the discrete sEMG signals into measures vectors that are transmitted to the receiver side. CS is applied independently to the 16 traces following (1) with the desired compress ratio $\mathrm{CR} = {}^{n}/_{m}$, where $n = 256$ is the dimension of the signal window (same length used to compute RMS values) and m is the number of the rows of A. Furthermore, to reduce complexity, sensing vectors are composed by antipodal values that, in case of standard CS (std-CS) are generated as i.i.d. random values and for rakeness-based CS (rak-CS) are the output of a sequences generator that uses C_A evaluated as in (9).

There are two possible approaches for classification. On the one hand, one can use the CS decoder to reconstruct the signal, give it as input to the feature extraction stage which computes the RMS value for each signal trace, and passes them to the NN for classification. In this work, we use the SPGL1 toolbox[3] in the CS decoder to achieve reconstruction by solving (2). On the other hand, we show here that one can by-pass the CS decoder and extract the feature directly from the CS measures.

IV. Numerical evidence

We first confirm the benefits of adopting rak-CS with respect to std-CS. To do so, performance is evaluated (in terms of reconstructed signal quality) in all windows composing the sEMG record by computing the Average Reconstruction Signal to Noise Ratio

$$\mathrm{ARSNR} = \mathbf{E}\left[\frac{\|x\|_2}{\|x-\hat{x}\|_2}\right]$$

[3]http://www.cs.ubc.ca/∼mpf/spgl1/

Figure 2 shows ARSNR as function of the compression ratio for both std-CS and rak-CS, obtained by averaging over 16000 sEMG signal windows (1000 per signal trace). Clearly, rak-CS outperforms std-CS with a gap that increase by decreasing CR and reaches 12 dB for $\mathrm{CR} = 2$. Since the rak-CS obtains better results than std-CS also in following analysis, from now on, we only consider the former approach.

With the aim of demonstrating that classifying hand movements in the compressed domain is possible, we compared results when classification is performed after signal reconstruction and when the CS decoder is bypassed. Figure 3 shows ARR as a function of CR in both cases. Although the NN classifier performs better with RMS computed on reconstructed signal $\mathrm{ARR}_{\hat{x}}$ than on measures ARR_y the gap is almost negligible. Both curves remain at approximately the same level of the reference case (that is $\mathrm{ARR}_x = 99.0\%$, corresponding to sEMG signals acquired at Nyquist rate and transmitted without compression) until $\mathrm{CR} \leq 5$ and still enjoy $\mathrm{ARR} > 0.98$ for $\mathrm{CR} \leq 40$.

The above analysis does not however offer a complete picture in terms of the required energy for the wireless transmission of information. To this aim, in fact, instead of considering $\mathrm{CR} = {}^{n}/_{m}$, we need to adopt $\mathrm{CR}_{\mathrm{bit}} = {}^{nb_x}/_{mb_y}$, where b_x and b_y are the number of bit used to represent the each sample of the signal and of the measure vectors, respectively. Since each element of x is coded with b_x = 12 bit and each measure is obtained via (1) with n signed additions, a generic element of y needs to be represented with

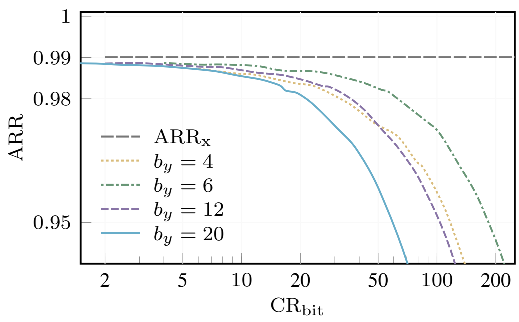

Fig. 4. Classification performance in term of ARR as function of the bit compression ratio $\mathrm{CR_{bit}}$, for different values of b_y.

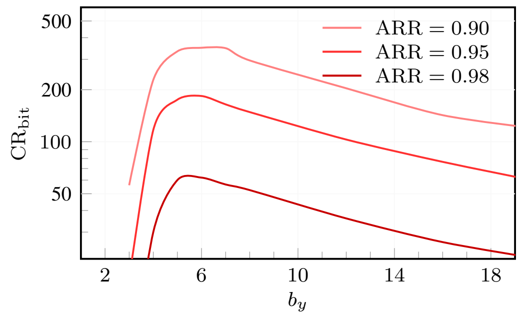

Fig. 5. Maximum obtainable bit compression ratio $\mathrm{CR_{bit}}$ that guarantees a certain level of hand movement recognition rate ARR as function of the number of bit used to represent the measures b_y.

$b_y = b_x + \log_2(n)$ = 20 bit to avoid accuracy loss.

Perhaps surprisingly, such level of accuracy is not necessary to correctly perform the hand movement recognition task, so that the number of transmitted bits can be very significantly reduced, without any appreciable decrease of ARR. To show this, let us explore the effect of reducing b_y by quantizing the measures. The quantization range $[-r, +r]$ has been fixed to include 99% of the measure values (that is so that $\mathrm{Rr}(|y| < r) = 0.99$) so that the quantization step size is set to $q = r/2^{b_y-1}$. Hence we consider b_y ranging from 1 bit to 20 bit.

Figure 4 shows classification performance in terms of ARR as a function of $\mathrm{CR_{bit}}$ for few significant values of b_y.

The graph shows that $b_y = 20$ is not the best choice since any reduction ot its value causes an increase in performance. On the other hand, also when the number of bit per measure is too low, performance tends to deteriorate. Interesting enough, when $b_y = 6$ results are better than both cases $b_y = 12$ and $b_y = 4$. The non-monotonic trend in $b_y$indicates the existence of *an optimal working point*. To highlight this, we can fix a desired value for ARR and numerically determine the maximum value of $\mathrm{CR_{bit}}$ for different values of b_y. The result is shown in Figure 5 for $\mathrm{ARR} = 0.90, 0.95, 0.98$.

Each of the three curves follows a common trend: with full resolution $b_y = 20$ and until b_y reaches 8, a small reduction of b_y do not significantly affect the recognition ratio so $\mathrm{CR_{bit}}$ increases. Then the accuracy starts to deteriorate but it is compensated by an increment of the measures. Eventually, when b_y is too low to represent the signal information, classification performance drops. The optimal working point for the considered system correspond to the curves maximum that lies at $b_y = 6$.

Results show that $\mathrm{CR_{bit}} = 62$ is reachable by quantizing measures with 6 bit and maintaining $\mathrm{ARR} > 98\%$. With $\mathrm{ARR} > 95\%$ the system achieve a compression ratio close to 200.

V. Conclusion

This work considered a classical hand movements recognition system and introduced the rakeness-based CS techniques to reduce the bandwidth needed to transmit the sEMG signal information content. The analysis shows that feature extraction does no require signal reconstruction and can be performed in the CS domain without impairing the hand gesture recognition rate. Such a bypass of the CS decoder ensures lower latency that is fundamental in real-time applications. Moreover, we shown that data to be transmitted can be further reduced by quantizing the compressed signal and that an optimal working point exists for each value of ARR.

Acknowledgements

The authors would like to thank M. Paleari and P. Ariano from Italian Institute of Technology (IIT), who provided the dataset used in this work.

References

[1] B. Peerdeman *et al.*, "Myoelectric forearm prostheses: state of the art from a user-centered perspective." *Journal of rehabilitation research and development*, vol. 48 6, pp. 719–37, 2011.

[2] C. Castellini and P. van der Smagt, "Surface emg in advanced hand prosthetics," *Biological Cybernetics*, vol. 100, pp. 35–47, Jan 2009. [Online]. Available: https://doi.org/10.1007/s00422-008-0278-1

[3] J. Wu, L. Sun, and R. Jafari, "A wearable system for recognizing american sign language in real-time using imu and surface emg sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, pp. 1281–1290, Sept 2016.

[4] T. Labs. (2018, June) Myo: Gesture control armband. [Online]. Available: https://www.myo.com/

[5] M. Crepaldi *et al.*, "A quasi-digital radio system for muscle force transmission based on event-driven ir-uwb," in *2012 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Nov 2012, pp. 116–119.

[6] M. Mangia *et al.*, "Compressed sensing based on rakeness for surface electromyography," in *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*, Oct 2014, pp. 204–207.

[7] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, pp. 1001–1014, May 2012.

[8] M. Shoaib *et al.*, "A 0.6 - 107 µw energy-scalable processor for directly analyzing compressively-sensed eeg," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, pp. 1105–1118, April 2014.

[9] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, April 2006.

[10] J. Haboba *et al.*, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, pp. 443–459, Sept 2012.

[11] M. Mangia *et al.*, "Rakeness-based design of low-complexity compressed sensing," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, pp. 1201–1213, May 2017.

[12] R. Rovatti, G. Mazzini, and G. Setti, "Memory-m antipodal processes: Spectral analysis and synthesis," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, pp. 156–167, Jan 2009.

[13] T. R. Farrell and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, pp. 111–118, March 2007.