

A neural algorithm for the non-uniform and adaptive sampling of biomedical data

Original

A neural algorithm for the non-uniform and adaptive sampling of biomedical data / Mesin, Luca. - In: COMPUTERS IN BIOLOGY AND MEDICINE. - ISSN 0010-4825. - STAMPA. - 71:(2016), pp. 223-230.
[10.1016/j.combiomed.2016.02.004]

Availability:

This version is available at: 11583/2638430 since: 2016-09-13T09:32:28Z

Publisher:

Elsevier

Published

DOI:10.1016/j.combiomed.2016.02.004

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

Elsevier postprint/Author's Accepted Manuscript

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>. The final authenticated version is available online at:
<http://dx.doi.org/10.1016/j.combiomed.2016.02.004>

(Article begins on next page)

**A NEURAL ALGORITHM FOR THE NON-UNIFORM AND ADAPTIVE
SAMPLING OF BIOMEDICAL DATA**

Luca Mesin

*Mathematical Biology and Physiology, Dipartimento di Elettronica e Telecomunicazioni,
Politecnico di Torino, Turin, Italy.*

Keywords: electromyography (EMG); electrocardiography (ECG); electroencephalography (EEG);
accelerometer; non-uniform sampling; compressive sensing; Nyquist limit

Running title: Non-uniform sampling of biomedical data

Address for correspondence:

Luca Mesin, Ph.D.

Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino; Corso Duca degli Abruzzi 24, Torino,
10129 ITALY

Phone: 0039-011-0904085; Fax: 0039-011 5644099; e-mail : [**luca.mesin@polito.it**](mailto:luca.mesin@polito.it)

Abstract

Background and Objective. Body sensors are finding increasing applications in the self-monitoring for health-care and in the remote surveillance of sensitive people. The physiological data to be sampled can be non-stationary, with bursts of high amplitude and frequency content providing most information. Such data could be sampled efficiently with a non-uniform schedule that increases the sampling rate only during activity bursts.

Methods. A real time and adaptive algorithm is proposed to select the sampling rate, in order to reduce the number of measured samples, but still recording the main information. The algorithm is based on a neural network which predicts the subsequent samples and their uncertainties, requiring a measurement only when the risk of the prediction is larger than a selectable threshold.

Results. Four examples of application to biomedical data are discussed: electromyogram, electrocardiogram, electroencephalogram and body acceleration. Sampling rates are reduced under the Nyquist limit, still preserving an accurate representation of the data and of their power spectral densities (PSD). For example, sampling at 60% of the Nyquist frequency, the percentage average rectified errors in estimating the signals are on the order of 10% and the PSD is fairly represented, until the highest frequencies. The method outperforms both uniform sampling and compressive sensing applied to the same data.

Conclusion. The discussed method allows to go beyond Nyquist limit, still preserving the information content of non-stationary biomedical signals. It could find applications in body sensor networks to lower the number of wireless communications (saving sensor power) and to reduce the occupation of memory.

1. Introduction

There are applications in which sampling at the Nyquist frequency is not efficient. For example, sparse signals are decomposed using a few significant components in compressive sensing (CS) [1]. Reducing the sampling rate is useful when many body sensors are used to monitor continuously the lifestyle of a healthy person or the condition of sensitive people [2]. Sensors are lightweight, non-invasive, wearable or embedded in cloth and they include a wireless communication with a storage and decision making system. Many physiological data can be sensed, e.g., acceleration, bioelectric activity, blood pressure, galvanic skin response, breathing [3]. Monitoring these data supports the individual self-assessment which allows to develop a personalized health care that helps healthy people to maintain their well-being [4]. Moreover, many diseases can benefit from a continuous monitoring, like cardiovascular problems, diabetes, Alzheimer's and Parkinson's diseases, renal failure, chronic obstructive pulmonary disease, post-operative conditions, stress or sudden infant death syndrome [5]. The remote patient monitoring could allow a rapid intervention when needed, developing an individualized care [6], with positive effects on the management of clinical services and on the quality of life of patients [7]. Many additional applications of body sensors are found, e.g., in military monitoring, interactive gaming [8], recognition of dietary activity [9], rehabilitation [10], personal information sharing and secure authentication [11].

Some recorded signals can show burst activity, reflecting the alternation of periods in which the investigated physiological system is either silent or active: for example, surface electromyogram (EMG) during cyclic tasks [12] or in pathological conditions (e.g., motor tremor induced by epileptic seizures [13]); electrocardiogram (ECG), with the QRS complex including most of the high frequency content [14] (unless pathological behaviours arise [15]); electroencephalogram (EEG), when the brain is performing different tasks or in pathological conditions (e.g., seizure in epileptic patients [16]); body acceleration during different activities [18] or in pathological conditions [19].

When data contain bursts, a high sampling rate is required if a uniform sampling is adopted, even if there are portions of the signal which could be down-sampled without loss of information. A non-uniform sampling, using a high sample rate only during the bursts, would allow both memory and energy saving (very important for wireless sensors, which are supplied with scarce resources [20]).

A predetermined non-uniform sampling schedule cannot be established for biomedical applications, but the sampling should be adaptively defined on the basis of the data. Adaptive techniques, like AZTEC, CORTES, SLOPE, or Fan [21], have been developed specifically for ECG data

compression. They can reduce considerably the number of acquired samples by a non-uniform sampling schedule, but they do not allow a real time modulation of the sampling frequency, as they require to measure each sample in order to decide if to keep it. A real time adaptive solution was proposed in [14], increasing the sampling rate when the signal curvature was high. The method showed higher performances than uniform sampling, but a-priori knowledge on the signal was required, limiting versatility [22]. In this respect, CS is considered more generally applicable [22], showing low compression ratio (CR, ratio between number of acquired versus original samples) and good accuracy (usually measured in terms of the percentage root mean square difference, PRD) [23][24]. However, it recovers the signal by an offline procedure applied on time epochs, introducing a delay. On the other hand, a real time adaptive sampling schedule could allow to save energy (to sample and transmit data) and to implement simple decision making control even on the sensor (e.g., a fall detector based on a threshold on the body acceleration). Thus, a versatile adaptive sampling algorithm could provide an important contribution.

The method proposed in [20] increased automatically the sampling rate when the investigated signal became unpredictable or provided high frequency contributions. A conceptual framework was discussed in [20], showing different general applications, but without optimizing the algorithm on specific data and the under-sampling was not measured relative to Nyquist frequency (problem affecting also the literature on CS applied to biomedical data [1], where approximation errors at specific CR are often provided without caring about the possible over-sampling of the original signal; see Discussion for details). The present work investigates these open issues. The algorithm proposed in [20] is improved, by an automatic tuning on the data, based on an offline analysis of a training set. The adaptive algorithm is compared with uniform sampling and CS with the same CR, when applied to different biomedical signals sampled at the Nyquist frequency: EMG, ECG, EEG and body acceleration.

2. Materials and Methods

2.1 Adaptive Sampling Algorithm

The adaptive sampling is based on a prediction algorithm and on its application to estimate the uncertainty of a predicted sample. It can be split into three parts:

1. selection of optimal predictors (based on the theory of time series embedding, Section 2.1.1);

2. training of an adaptive algorithm to predict the next sample (a multi-layer perceptron, MLP, was used, Section 2.1.2);
3. real time schedule of the sampling rate (based on the uncertainty of the prediction, Section 2.1.3).

2.1.1 Time series embedding

Embedding theory was applied to the data [25][26][27]. Specifically, the time series were supposed to be extracted from a deterministic physiological system described by a set of unknown deterministic rules

$$\frac{d\vec{x}}{dt} = \vec{F}(\vec{x}) \quad (1)$$

where \vec{x} is the vector of state variables of the system and \vec{F} is a set of functions called the vector field (defining the evolution of the state variables), which was assumed not to be an explicit function of time (i.e., the system was assumed to be autonomous). The time series $y(t)$ (where t from now on is a discrete time variable) was assumed to be extracted from the system through a measurement process described by an unknown function $g(\cdot)$ of the state variables:

$$y(t) = g(\vec{x}(t)) \quad (2)$$

Given a single measurement, a vector of time delayed versions (delayed coordinates) was built [25]

$$\vec{Y}(t) = \begin{bmatrix} y(t) \\ y(t-\tau) \\ \vdots \\ y(t-(m-1)\tau) \end{bmatrix} \quad (3)$$

where the time delay τ was chosen so that two delayed coordinates provided different information and the number m of elements of the vector is called the embedding dimension (as it is the dimension of the so called phase space in which the trajectory $\vec{Y}(t)$ is embedded) [25][26][27].

The time delay τ and the embedding dimension m were computed as follows.

- Time delay. The mutual information of the original and delayed data was computed:

$$MI(\tau) = \iint_{A B} P_{AB}(a,b) \ln \left(\frac{P_{AB}(a,b)}{P_A(a)P_B(b)} \right) dadb \quad (4)$$

where the time series $y(t)$ and $y(t-\tau)$ are considered as random variables A and B , respectively, with joint probability density $P_{AB}(a,b)$ and marginal probabilities $P_A(a)$ and $P_B(b)$, respectively. The minimum between the delays corresponding to the first local

minimum or to a 90% decrease of $MI(\tau)$ was selected as the time delay τ of the delayed coordinates.

- Embedding dimension. Cao's method was used [27][28]. It is based on the number of points of the trajectory, described by the vector in (3), which are neighbours of other points of the trajectory itself. When increasing the embedding dimension by adding one element to the vector (3), neighbouring points which were close only due to the projection of the trajectory in a low dimensional space (false near neighbours) may turn away. Thus, the number of neighbouring points decreases by increasing the embedding dimension, till false neighbours are present. The minimum phase space dimension allowing to remove the false near neighbours was selected as the embedding dimension m : it allows to identify uniquely the dynamics of the trajectory and possibly to predict it. Specifically, Cao's method investigates the following function of the embedding dimension [28]

$$El(m) = \frac{E(m+1)}{E(m)}, \quad \text{where} \quad E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} \frac{\|Y_{m+1}(i) - Y_{m+1}(n(i, m))\|}{\|Y_m(i) - Y_m(n(i, m))\|} \quad (5)$$

where N is the number of considered samples of the time series, $\|\cdot\|$ is the absolute distance norm, $Y_m(i)$ is the i^{th} sample of the reconstructed vector with embedding dimension m and $n(i, m)$ indicates its nearest neighbour in the m -dimensional reconstructed phase space. The function $El(m)$ saturates when all false near neighbours are removed. Thus, such a function has a point of maximum curvature (corresponding to the correct embedding dimension m), separating a region of increase from a plateau. Such a point was estimated automatically, considering the best approximation of $El(m)$ by 2 lines [27].

The delay τ and the embedding dimension m were selected considering only a portion of data (the training set defined in Section 2.1.2).

2.1.2 Prediction algorithm

An MLP forecasted the subsequent sample of the time series with a constant interval using the delayed coordinates as inputs. Thus, given the time delay τ and the embedding dimension m (estimated as described in Section 2.1.1), in order to predict the value of the sample number $(i+1)$ of the time series, the MLP used as inputs the m values $y(i), y(i-\tau), \dots, y(i-(m-1)\tau)$.

Different MLPs were investigated, from which the optimal one was chosen as that with best generalization performances. The following procedure was adopted to select such an optimal MLP. The data were split into training, validation and test sets (50%, 25% and 25% of data, respectively).

MLPs with a single hidden layer were used (notice that a single hidden layer is sufficient to approximate any nonlinear function [29]). Sigmoidal activation functions were used for the hidden neurons. Their number was chosen in the range 10-50. The output neuron had a linear activation function. The MLPs were trained on the training data, applying the quasi-Newton algorithm [30] for a number of iterations in the range of 25-350. The optimal MLP was selected choosing the topology (i.e., the number of hidden neurons) and the parameters (i.e., the synaptic weights and bias, after a specific number of iterations of the optimization algorithm) with best performances on the validation set (measured in terms of the mean square error).

2.1.3 Sampling schedule

The sampling schedule was defined by selecting which sample to measure next, on the basis of the uncertainty of its prediction obtained by the optimal MLP described in Section 2.1.2. The algorithm was applied on the test set. The prediction and an estimate of its uncertainty were performed for each time sample, using the available (sampled or predicted) data [20]. All delayed coordinates (Section 2.1.1) used as inputs of the neural predictor (i.e., the optimal MLP, Section 2.1.2) were characterized by their (predicted or measured) values and uncertainties.

The prediction of a new value was obtained as follows: 1000 random input data were simulated, using for each data a uniform distribution centred around its value and with a range given by its uncertainty; the neural predictor was run for each of these 1000 inputs; the prediction was the median of the distribution of the obtained estimates.

The uncertainty of a data was chosen distinguishing between two cases, i.e., when the data was measured or predicted. The uncertainty of a measurement was defined by the user, considering the accuracy of the sensor and the expected noise level. The uncertainty of a predicted sample was defined as the mean of two terms (see [20] for details):

1. the range of predicted values obtained from the 1000 random tests indicated above (excluding 10 possible outliers);
2. the estimated prediction error obtained integrating in time the rate of prediction error, defined as the ratio between the estimation error (available when a new sample was measured and defined as the absolute difference between the measured and the predicted values) divided by the time delay from the last acquired sample; a memory term was also included, by computing the weighted average of the last two rates of prediction error (with weights 65% and 35%, respectively); finally, a saturation was considered, by imposing a minimum CR.

An additional measurement was required from a sensor when the associated uncertainty overcame a threshold, chosen by the user. This free parameter and the minimum CR mentioned above allow the user to tune the level of under-sampling to the application at hand.

2.2 Testing signals

Different experimental signals were used to test the algorithm. The first 3 signals were provided by different institutions, acknowledged at the end of the paper, whereas the last one was acquired by the author (in accordance with the Declaration of Helsinki).

1. Surface EMG was recorded from the tibialis anterior muscle of a healthy subject walking on a treadmill at a self-selected speed (bipolar electrodes placed on the muscle belly; ground electrode on the right patella; sampling rate 2048 Hz). The experiment lasted about 20 s. The signal had a bandwidth of about 400 Hz. It was low-pass filtered at 400 Hz offline (4th order anti-causal, anti-aliasing Butterworth filter) and resampled at 800 Hz, in order to be at the Nyquist limit.
2. ECG was recorded from the wrists of a healthy subject [31]. The experiment lasted about 75 s. The signal was sampled at 2 kHz, but its bandwidth was about 30 Hz. Thus, it was resampled at 60 Hz (as the EMG described above).
3. A scalp EEG was recorded from an epileptic patient for about 50 minutes. The channel O2 was selected (with reference on the earlobe). The EEG was sampled at 1024 Hz (Brain Explorer, EB-Neuro[®] amplifier), but the signal (after removing the frequency components under 0.5 Hz) had more than 95% of energy under 14 Hz: it was down-sampled to 32 Hz (notice that with this preliminary down-sampling, only theta, delta and alpha rhythms are available; however, for this representative application, the main objective is identifying the epileptic seizures).
4. Acceleration data were recorded for 5 minutes from a subject keeping fixed in a bag a Tablet embedding the 3 axial accelerometer LSM330DLC. Quite standing, walking and jumping were the investigated activities. The signal was sampled at 50 Hz. Its bandwidth was about 15 Hz. It was resampled at 32 Hz.

2.3 Test of the algorithm and comparison with other down-sampling methods

The algorithm was applied to the test data (i.e., a portion of the signals described in Section 2.2, as indicated in Section 2.1.2). Different thresholds were used, obtaining different CRs. Then, data were uniformly down-sampled with the same CRs, by cubic interpolation. From the two time series, the

samples of the original signal were estimated by cubic interpolation. The fit was assessed in terms of the averaged rectified error (ARE) and the percentage root mean squared difference (PRD)

$$ARE = 100 \frac{\sum_i |x_i - \hat{x}_i|}{\sum_i |x_i|} (\%) \quad PRD = 100 \frac{\sqrt{\sum_i |x_i - \hat{x}_i|^2}}{\sqrt{\sum_i |x_i|^2}} (\%) \quad (6)$$

where x_i and \hat{x}_i are the original and approximated data, respectively. The power spectral densities (PSD) of the three time series (original and estimated from either of the two under-sampled signals) were also compared.

Some tests were also performed using CS, with the principal components of the data as basis functions [32]. Specifically, principal components were computed from all adjacent epochs extracted from the training and validation sets. The duration of the epochs was chosen as that providing maximal accuracy on the test set (1 s for all considered signals except for EMG, for which epochs of 50 ms were used).

The down-sampling was written as the multiplication of the data with a sensing matrix. Such a matrix was chosen using one of the following criteria: 1) a random selection of a percentage of the rows of the identity matrix (tests were performed also choosing a sensing matrix with entries equal to either 1 or 0, drawn from a binomial distribution with probability of generating 1 equal to 0.6, as suggested in [23]; equivalent results were obtained, so that this possibility is not further considered in the following); 2) selecting the rows of the identity matrix corresponding to the samples chosen by the adaptive algorithm. The coefficients were estimated minimizing their L_1 norm, imposing their sparseness [33]. The original signal was finally recovered as a sum of the basis functions weighted by the estimated coefficients.

3. Results

The signals were embedded obtaining the dimensions and delays listed in Table 1. The optimal MLP for the prediction of the subsequent samples was investigated as explained in Section 2.1.2, resulting in the selection of a number of hidden neurons and of iterations of the optimization algorithm indicated in Table 1. The algorithm was implemented in Matlab on a PC with the following characteristics: Intel(R) Core i7-2630QM, Quad-Core, clock frequency of 2 GHz, 6 GB of RAM and 64 bits operating system. A sequential, interpreted implementation was considered. The average processing time for epochs of duration 1 s is indicated in Table 1: it is always about an order of magnitude shorter than the processed epoch, indicating that the sampling schedule can be adjusted in real time.

Figure 1 shows an application of the method to surface EMG (similar figures are given in the Supplementary Material for the other test data). A high sampling rate is selected for the portions of largest activity (panels A and C) and low rate when the EMG has small amplitude (reflecting cross-talk from nearby muscles or noise). This is evident in the two magnifications in Fig. 1A, which show that the uniform sampling has a better reconstruction of the noisy portions of the signal (where the adaptive method reduces the sampling to a minimum), whereas the adaptive algorithm provides a better estimate of the bursts of activity. The comparison of the PSDs (Fig. 1D) indicates that the adaptive under-sampling allows to estimate properly the high frequency components, even under a drastic reduction of the number of samples. On the contrary, a uniform under-sampling cannot estimate the high frequency content, as the sampling frequency is under the Nyquist limit.

Figure 2 shows the comparison between the adaptive algorithm and uniform sampling, when applied to the four test data. Different uncertainty thresholds (ranging from about 1 and 2 times the measurement uncertainty) reflect into different CRs and reconstruction errors. Keeping the same CR, the ARE and the PRD are lower when using the adaptive algorithm instead of a uniform schedule.

Figure 3 shows the accuracy of CS in estimating the original signal when using either random samples or those selected by the adaptive algorithm, considering the same CRs. The performances of CS are superior in the latter case.

4. Discussion

Body sensor networks found many applications in self-monitoring and surveillance of patients [2][4][5][8][9][11][18][34]. They support the individual care of healthy people and the long-term outpatient monitoring, which is important in Holter acquisitions or when rare events are monitored (like epileptic seizures [16] or possible falls [17]). Remote monitoring requires wireless sensors to sample and transmit data to a base station. Decreasing the amount of measured samples is important to reduce power consumption and memory occupation [1][22][24]. This paper discusses a method to select automatically the information to be acquired, scheduling a non-uniform sampling that adapts to the investigated signal.

4.1 Application to physiological data

Some applications are provided.

1. The application on EMG (Figure 1) shows that the algorithm is able to automatically identify the muscle activity intervals to be sampled at a high rate.
2. When applied to ECG (Figure I of the Supplementary Material), the sampling rate was reduced between two heartbeats and increased at the beginning of the P wave, with QRS complex and T wave being sampled at about the maximal frequency.
3. The sampling rate was increased to improve the approximation of EEG sharp waves preceding or following a crisis and during the seizures (Figure II of the Supplementary Material). These events are quite rare; as a consequence, the average error was only a bit lower than that obtained by uniform under-sampling and the power spectrum was estimated only slightly better. However, the resolution was improved exactly during crucial events for the study of epileptic seizures [38].
4. Applied to accelerometer data, the algorithm used high sampling rate only during the main activities of the subject (Figure III of the Supplementary Material). Only a short delay (related to the selected minimum CR) was needed to identify the beginning of the activity and to boost the sampling frequency.

The PSDs of the signals were fairly well estimated, even when the average sampling rate was far below Nyquist limit. Moreover, the algorithm provided good accuracy in the estimate of test data, outperforming both uniform down-sampling schedule and CS with the same CRs (Figures 2 and 3).

4.2 Comparison with the literature

In literature, high performances are indicated for CS [1]. For example, low CRs were obtained with ECG, electro-oculogram (EOG) and needle EMG [22][23], which are quite sparse in time. Lower

performances were obtained when processing signals with more interference, like surface EMG [23]¹ and scalp EEG [24]. The proposed algorithm, as it is adaptive, provides satisfactory results in all considered tests. On the other hand, CS did not provide impressive performances on the data here investigated (Figure 3). Possibly, this is due to the dictionary, selected by principal component analysis, instead of further optimizing the basis functions for each data. However, the shown performances are not so far from those reported in literature, if we consider that our data were sampled at the Nyquist limit, whereas this is sometime questionable for the literature: for example, EEG was sampled at frequencies not lower than 200 Hz in [23][24] and the PRD was about 20% when the CR was 0.4, i.e., sampling at an average rate of 80 Hz (which is higher than the Nyquist limit for the EEG considered here); ECG was sampled at 360 Hz in [23][39], which is above the Nyquist limit (indeed, the 99% of the energy of the records of the MIT-BIH Arrhythmia Database considered in [23][39] is below 50 Hz); the needle EMG considered in [22] was sampled at 4 kHz, but 95% of its energy was below 700 Hz. Considering again our data (resampled at the Nyquist limit), the performances of CS were improved when it was applied to the samples selected by the proposed adaptive algorithm. This is somehow in line with [39], where improved performances of CS were obtained after pre-processing the ECG². Dynamic thresholding, selecting only the largest signal amplitudes to increase the sparsity, is another pre-processing technique that improved CS performances [22].

4.3 Power reduction in wireless sensor networks

The transmission module of a wireless sensor can be switched off for most of the time (saving energy) and activated only when needed [20]³. Thus, energy could be saved computing in advance which sample to acquire next and switching on the transmission module only when new data are acquired. Different possibilities could also be explored to further reduce data transfers: the transmission could be enabled only when a certain amount of data is available, or in specific conditions (e.g., when there is a burst of activity); data could be compressed [35]; transmission could be suppressed when the acquired data match a prediction model [36].

Notice that the proposed algorithm reduces the power needed to sample and transmit data, but it requires energy to perform computations. The base station could perform them [20][37], but then the

¹ Notice that the sampling frequency of the original EMG studied in [23] was only 250 Hz.

² The QRS complex was identified in [39], in order to fix the period of each heartbeat and to select a specific number of beats each time. However, this procedure is fit to the specific data (i.e., ECG from a healthy subject) and it is not clear how the proposed method would cope with a highly varying heart rate or occurrence of an arrhythmia [1].

³ Using a radiofrequency transmitter for wireless communication (e.g., nRF24L01 by Nordic Semiconductor, <http://www.nordicsemi.com/eng/Products/2.4GHz-RF/nRF24L01>), the times needed to switch on (from stand-by) or off the transmission module are about 150 μ s. Thus, the time to switch on and off the module is lower than the minimum interval between measurements (larger than a ms for EMG and larger than tenth of ms for the other data).

communication with the sensors should be properly scheduled or the transmission module should be in stand-by and not switched off, with lower energy saving. Different alternatives could be tested in future.

4.4 Limitations

The proposed algorithm requires to train the neural predictor on a dataset. This preliminary processing can be considered as a limitation of the method, which cannot be applied directly online. However, this pre-processing step results also in the high flexibility of the method, which can adapt to the data (learning how to predict them) and to the application (selecting the CR still guaranteeing the needed resolution).

The online application could be applied without degradation of the performances as long as the training set fairly represents the dynamics of the signal to be acquired. On the other hand, if the data undergo a drift (e.g., for acquisition problems) or if the physiological system generating them undergoes some variation (e.g., the person goes to sleep), a decrease of the performance of the algorithm is expected. This means that the prediction errors will increase and the algorithm will adopt a high sampling rate to correct it. In such a case, the recorded samples could be used to train again the predictor algorithm, in order to adapt to the variations of the investigated system.

Data were first down-sampled at the Nyquist limit and the performances were tested in terms of the recovery of high amplitude or energetic components (related to ARE and PRD, respectively). However, some signals contain important high frequency contributions with low energy. For example, EEG includes components in the beta and gamma bands (13-30 Hz and more than 30 Hz, respectively) which were excluded in this study. Different applications could require to maintain specific low energy components. In such cases, the performances of the algorithm should not be based on average reconstruction errors, but on the ability of extracting the information of interest. Applications could be proposed for specific biomedical signals (e.g., estimating the average heart rate or its variability from down-sampled ECG: a high reconstruction error could be tolerated, but no R peak should be lost).

4.5 Future works

The algorithm was implemented on a PC with average performances, using an interpreted routine written in Matlab working on a single processor. Using a compiled routine, a sensible reduction of the processing time could be obtained. Moreover, the hidden neurons of the MLP can process data in

parallel, reducing further the processing time depending on the number of available cores (e.g., using a GPU or an FPGA implementation, the prediction of the subsequent samples could be obtained largely reducing the number of clock cycles). Consider also that, as indicated in Section 2.1.3, 1000 predictions were performed for each sample (to predict its value and to assess the prediction risk), but preliminary tests show that equivalent results could be obtained reducing their number of an order of magnitude, as in [20] (which would reduce significantly the computational time).

A specific prediction algorithm (i.e., MLP with a single hidden layer) was considered. Preliminary tests were also performed with both a simpler and a more sophisticated regression algorithm: using a linear predictor, poor performances were obtained; on the other hand, considering an MLP with three hidden layers (with the same number of neurons as the optimal MLP with a single hidden layer, for each of the three layers), the performances were not statistically different from those of the simpler MLP, but with a larger computational cost. Thus, the MLP with a single hidden layer is preferable than the considered alternatives; however, further different possibilities could be tested in future.

Further work should also be devoted to the implementation of the proposed algorithm in a wireless body sensor system. Specific applications could then be tested, indicating the performances in terms of reduction of power and occupation of memory, still preserving the information of interest: for example, surface EMG in non-stationary conditions, extracting activity intervals or fatigue indexes [40]; sleep bruxism monitored from ECG and masseter EMG [41]; continuous EEG (e.g., during sleep [42], coma [43] or long detections waiting for an epileptic seizure [16]); body acceleration to detect possible dangerous conditions (e.g., possible falls of elderly people or of workers [17]).

The algorithm could also be used in a network of body sensors, either applied to each of them in turn, or integrating their joint information to improve the prediction [20].

5. Conclusions

This paper discusses a real time algorithm that schedules a non-uniform sampling adapted to the measured data, that reduces the number of acquired samples, but still preserving high frequency information, going beyond the Nyquist limit. The tests on experiments are promising (Figure 2 and

3), showing that the adaptive algorithm outperforms both uniform under-sampling and CS using the same CRs.

If applied in wireless sensor networks, the method lowers the required data transmissions, saving energy and allowing an efficient management of the resources. Moreover, it could be integrated with other techniques, for a further compression or a more accurate recovery of the original data. Many potential future investigations are expected, e.g., for the individual self-assessment of healthy people or the remote monitoring of patients.

Acknowledgements.

The author thanks Prof. Dario Farina (Department of Neurorehabilitation Engineering, University Medical Center, Goettingen), Prof. Eros Pasero (Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino) and Dr. Elisa Montalenti (Epilepsy Service, Department of Neuroscience, University of Torino) who provided the EMG, ECG and EEG, respectively.

Conflict of interest

The author states no conflict of interest.

Other declarations

There was no source of funding for this research. Ethical approval was given by the local committee for the acquisition of acceleration data.

References

- [1] D. Craven, B. McGinley, L. Kilmartin, M. Glavin, E. Jones, Compressed Sensing for Bioelectric Signals: A Review. *IEEE J Biomed Health Inform.* 19 (2015) 529-40.

- [2] P.S. Pandian, K. Mohanavelu, K.P. Safeer, T.M. Kotresh, D.T. Shakunthala, P. Gopal, V.C. Padaki, Smart Vest: wearable multi-parameter remote physiological monitoring system. *Med Eng Phys.* 30 (2008) 466-77.
- [3] N. Sharma, T. Gedeon, Objective measures, sensors and computational techniques for stress recognition and classification: a survey. *Comput Methods Programs Biomed.* 108 (2012) 1287-301.
- [4] F. Buttussi, L. Chittaro, MOPET: a context-aware and user-adaptive wearable system for fitness training. *Artif Intell Med.* 42 (2008) 153-163.
- [5] B. Latrè, B. Braem, I. Moerman, C. Blondia, P. Demeester, A survey on wireless body area networks, *Wireless Netw* 17 (2011) 1–18.
- [6] M. Patel, J. Wang, Applications, challenges, and prospective in emerging body area networking technologies, *IEEE Wirel Commun*, 17 (2010) 80–88.
- [7] S. Park, S. Jayaraman, Enhancing the quality of life through wearable technology. *IEEE Engineering in Medicine and Biology Magazine*, 22 (2003) 41–48.
- [8] E.M. Holz, J. Höhne, P. Staiger-Sälzer, M. Tangermann, A. Kübler, Brain-computer interface controlled gaming: evaluation of usability by severely motor restricted end-users, *Artif Intell Med*, 59 (2013) 111-120.
- [9] O. Amft, G. Troster, Recognition of dietary activity events using on-body sensors, *Artif Intell Med* 42 (2008) 121-136.
- [10] S. Patel, H. Park, P. Bonato, L. Chan, M. Rodgers, A review of wearable sensors and systems with application in rehabilitation, *J Neuroeng Rehabil*, 20 (2012) 9:21.
- [11] M. Chan, D. Estève, J.Y. Fourniols, C. Escriba, E. Campo, Smart wearable systems: current status and future challenges, *Artif Intell Med.* 56 (2012) 137-156.
- [12] A. Muro-de-la-Herran, B. Garcia-Zapirain, A. Mendez-Zorrilla, Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications, *Sensors*, 14 (2014) 3362-3394.

- [13] I. Conradsen, S. Beniczky, P. Wolf, T.W. Kjaer, T. Sams, H.B. Sorensen, Automatic multi-modal intelligent seizure acquisition (MISA) system for detection of motor seizures from electromyographic data and motion data. *Comput Methods Programs Biomed.* 107 (2012) 97-110.
- [14] R. Rieger, J.T. Taylor, An adaptive sampling system for sensor nodes in body area networks, *IEEE Trans Neural Syst Rehabil Eng.* 17(2009) 183-189.
- [15] R. Fensli, T. Gundersen, T. Snaprud, O. Hejlesen, Clinical evaluation of a wireless ECG sensor system for arrhythmia diagnostic purposes. *Med Eng Phys.* 35 (2013) 697-703.
- [16] A.J. Casson, E. Rodriguez-Villegas, Toward online data reduction for portable electroencephalography systems in epilepsy. *IEEE Trans Biomed Eng.* 56 (2009) 2816-2825.
- [17] A.K. Bourke, J.V. O'Brien, G.M. Lyons, Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm, *Gait & Posture* 26 (2007) 194–199.
- [18] D. Anguita, A. Ghio, L. Oneto, X. Parra, J.L. Reyes-Ortiz, Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine, *International Workshop of Ambient Assisted Living (IWAAL 2012)*. Victoria-Gasteiz, Spain. Dec 2012
- [19] L.Z. Popovic, T.B. Sekara, M.B. Popovic, Adaptive band-pass filter (ABPF) for tremor extraction from inertial sensor data. *Comput Methods Programs Biomed.* 99 (2010) 298-305.
- [20] L. Mesin, S. Aram, E. Pasero, A neural data-driven algorithm for smart sampling in wireless sensor networks, *EURASIP Journal on Wireless Communications and Networking*, 23 (2014) 1-8.
- [21] S.M. Blanchard, R.C. Barr, Comparison of methods for adaptive sampling of cardiac electrograms and electrocardiograms, *Med Biol Eng Comp.* 23 (1985) 377-86.
- [22] A.M. Dixon, E.G. Allstot, D. Gangopadhyay, D.J. Allstot, Compressed sensing system considerations for ECG and EMG wireless biosensors, *IEEE Trans Biomed Circuits Syst.* 6 (2012) 156-66.

- [23] A.J. Casson, E. Rodríguez-Villegas, Signal agnostic compressive sensing for Body Area Networks: Comparison of signal reconstructions, Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2012) 4497-4500.
- [24] A.M. Abdulghani, A.J. Casson, Esther Rodríguez-Villegas, Compressive sensing scalp EEG signals: implementations and practical performance, J Medical and Biological Engineering and Computing, 50 (2012) 1137-1145.
- [25] H. Kantz, T. Schreiber, Nonlinear Time-series Analysis, Cambridge University Press, Cambridge, 1997.
- [26] M.G. Signorini, Nonlinear analysis of heart rate variability signal: physiological knowledge and diagnostic indications, Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE, 2 (2004) 5407-5410.
- [27] L. Mesin, A. Monaco, R. Cattaneo, Investigation of Nonlinear Pupil Dynamics by Recurrence Quantification Analysis, BioMed Research International, 2013 (2013) Article ID 420509.
- [28] L.Y. Cao, Practical method for determining the minimum embedding dimension of a scalar time series, Physical D, 110 (1997) 43-50.
- [29] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1998.
- [30] E. Barnard, Optimization for training neural nets, IEEE Trans Neural Networks, 3 (1992) 232-240.
- [31] L. Mesin, A. Múnera, E. Pasero, A low cost ECG biometry system based on an ensemble of Support Vector Machine classifiers, in Advances in Neural Networks: Computational and Theoretical Issues, Computational Intelligence for Information Communication Technologies, (Editors: S. Bassis, A. Esposito, C. F. Morabito, E. Pasero) Springer, 2016.
- [32] R. Masiero , G. Quer , D. Munaretto , M. Rossi , J. Widmer, M. Zorzi, Data acquisition through joint compressive sensing and principal component analysis, IEEE Global Telecommun. Conf. (IEEE GLOBECOM), (2009) 1-6.

- [33] E.J. Candès, J. Romberg, T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory*, 52 (2006) 489–509.
- [34] F. Pinciroli, C. Pagliari, Understanding the evolving role of the Personal Health Record, *Comput Biol Med.* 59 (2015) 160-163.
- [35] F. Dabiri, H. Noshadi, M. Sarrafzadeh, Behavioural reconfigurable and adaptive data reduction in body sensor networks, *International Journal of Autonomous and Adaptive Communications Systems*, 6 (2013) 207-224.
- [36] A. Banerjee, S. Nabar, S.K.S. Gupta, R. Poovendran, Energy-efficient long term physiological monitoring, *Proceedings of the 2nd Conference on Wireless Health*, Article No. 17, ACM New York, NY, USA, 2011.
- [37] M. Cordina, C.J. Debono, Maximizing the lifetime of wireless sensor networks through intelligent clustering and data reduction techniques, *Proceedings of the 2009 IEEE Wireless Communication and Networking Conference, Budapest (IEEE, Piscataway, 2009)*, pp. 1–6.
- [38] S.J. Smith, EEG in the diagnosis, classification, and management of patients with epilepsy, *J Neurol Neurosurg Psychiatry*. 76 (2005) ii2-7.
- [39] L.F. Polania, R.E. Carrillo, M. Blanco-Velasco, K.E. Barner, Compressed sensing based method for ECG compression, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (2011) 761-764.
- [40] C.J. De Luca Myoelectric manifestations of localized muscular fatigue in humans, *CRC Crit Rev Biomed Eng*, 11 (1984) 251–279.
- [41] T. Castroflorio, L. Mesin, G.M. Tartaglia, C. Sforza, D. Farina, Use of Electromyographic and Electrocardiographic Signals to Detect Sleep Bruxism Episodes in a Natural Environment , *IEEE Journal of Biomedical and Health Informatics*, 17 (2013) 994-1001.
- [42] X. Zhang, X. Dong, J.W. Kantelhardt, J. Li, L. Zhao, C. Garcia, M. Glos, T. Penzel, F. Han, Process and outcome for international reliability in sleep scoring. *Sleep Breath*. 2014
- [43] L. Mesin, P. Costa, Prognostic value of EEG indexes for the Glasgow outcome scale of comatose patients in the acute phase, *J Clin Monit Comput.*, 28 (2014) 377-385.

Table 1. Embedding dimension, time delay, number of hidden neurons of the optimal MLP, number of iterations of the optimization algorithm and processing time for 1 s of signal (mean \pm standard deviation) for the different considered signals.

Signal	Embedding dimension	Time delay (samples)	Hidden neurons	Iterations	Processing time for 1 s (ms)
EMG	8	7	10	50	79.5 \pm 20.1
ECG	9	8	28	50	158.6 \pm 6.3
EEG	7	2	28	250	121.6 \pm 1.9
Acceleration	5	8	26	125	80.8 \pm 9.8

FIGURES

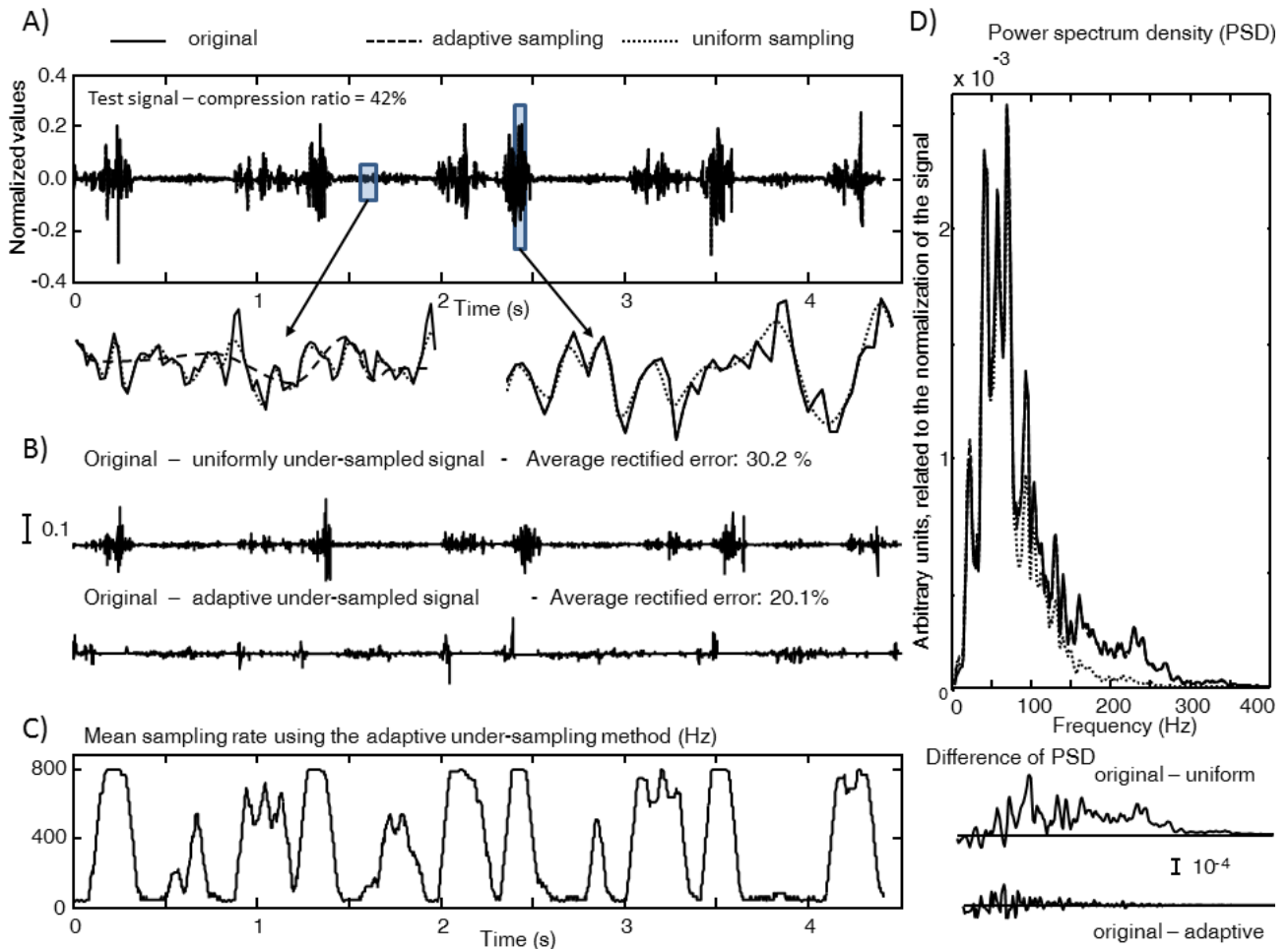


Figure 1. Example of application of the under-sampling algorithm to EMG (normalized data with respect to the range; uncertainty of a normalized measurement was assumed to be 10^{-3} ; maximal imposed reduction 90%). A) The signal is shown in the time domain superimposed to the under-sampled versions. The magnifications of two portions are also shown. B) The differences between the test signal and the approximations obtained by uniform or adaptive under-sampling are shown. C) Average sampling rate (computed in terms of the number of selected samples on sliding windows of 50 samples). D) PSD of the normalized test EMG shown in A and of the under-sampled versions, computed by the Welch algorithm using windows of 0.5 s with 25% of overlapping.

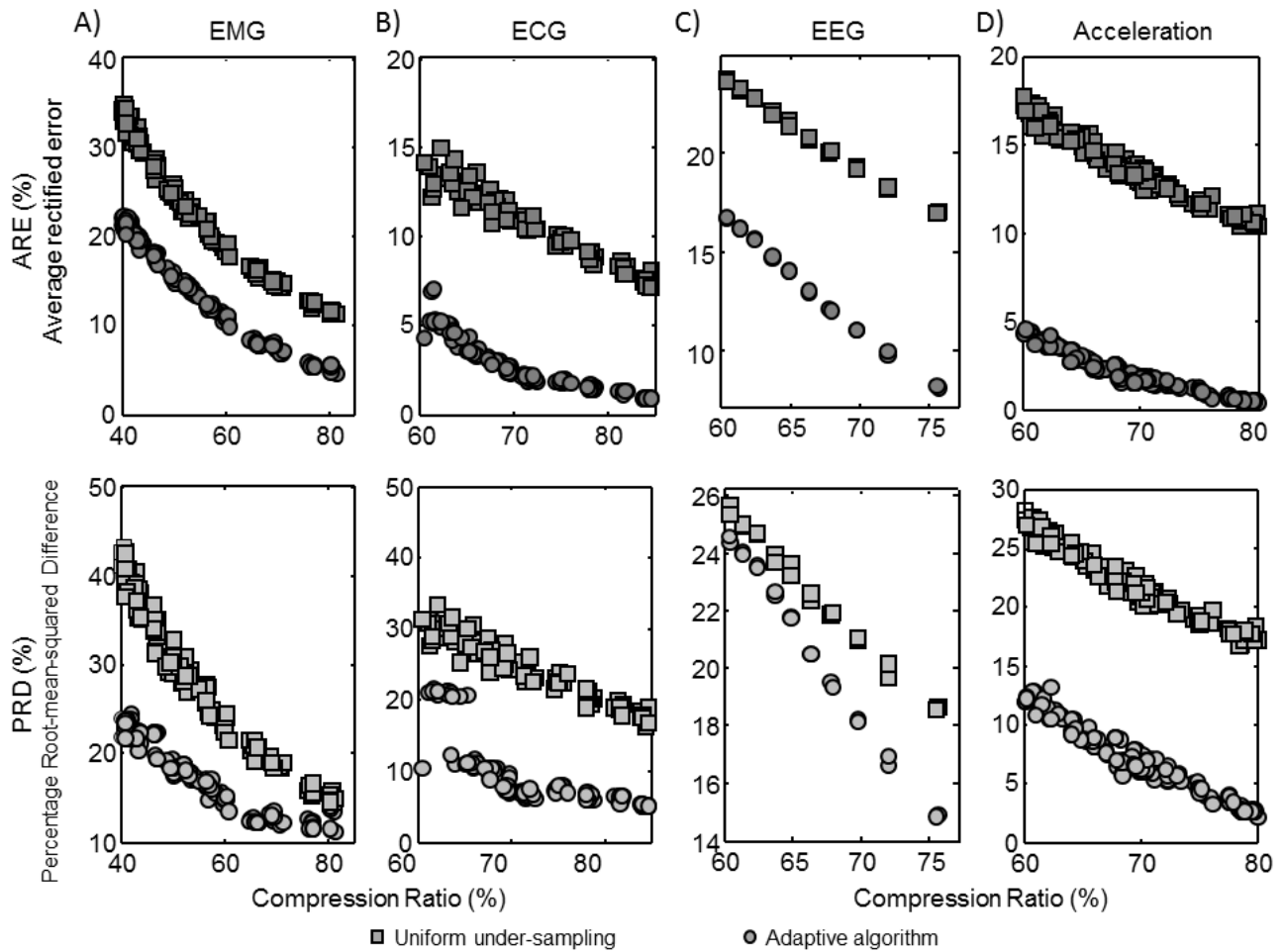


Figure 2. Average rectified error (ARE) and percentage root mean square difference (PRD) versus the compression ratio (i.e., the ratio between the number of acquired versus original samples) for the investigated signals: A) EMG, B) ECG, C) EEG and D) modulus of the acceleration vector. Both the uniform and the adaptive under-sampling methods are considered.

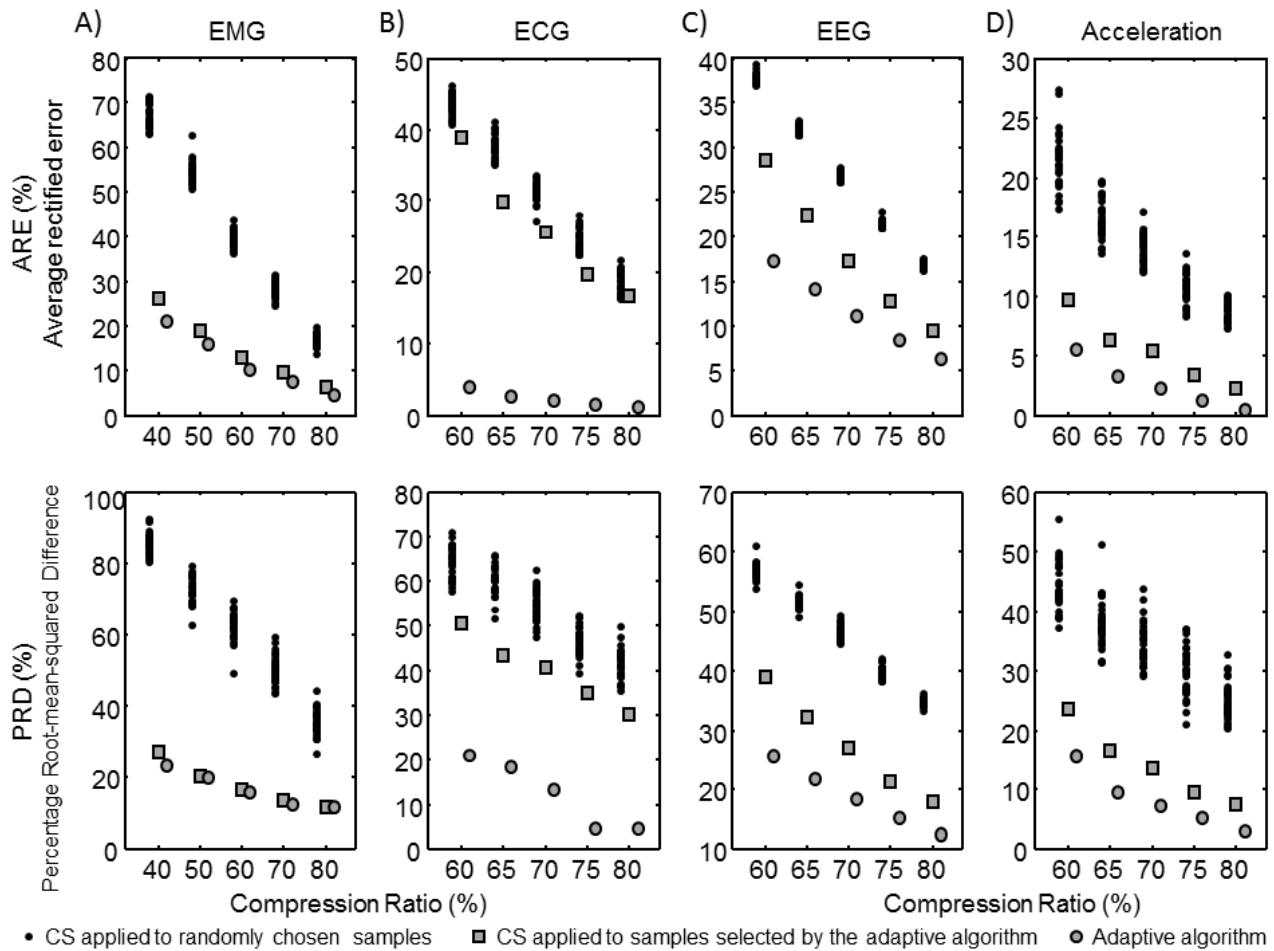


Figure 3. Comparison between compressive sensing (CS) and the proposed adaptive algorithm. Different compression ratios (i.e., ratio between the number of acquired versus original samples) are considered for each of the studied signals: A) EMG, B) ECG, C) EEG, D) modulus of the acceleration vector. CS was applied either to randomly chosen samples (30 random choices were considered) or to the samples selected by the adaptive algorithm.

Supplementary material

Figures I-III show examples of application of the adaptive under-sampling method to ECG, scalp EEG and acceleration data, respectively. The format is similar to that of Figure 1 of the paper, where surface EMG is considered. The signals were embedded (as described in Section 2.1 of the paper), obtaining embedding dimensions and time delays listed in Table 1 of the paper.

The under-sampling method automatically identifies the portions of larger activity as those to be sampled at higher rates (panels A and C of the figures). Considering the ECG (Fig. IA and C), the sampling rate is automatically increased when the heart is firing. The QRS complex is sampled at about the highest frequency, whereas, using a uniform sampling, the approximation error is maximal there. In the case of EEG, the energetic content of the spikes that precede the seizure is quite low. They are rare, but important epileptiform discharges indicating the incoming seizure (e.g., sharp waves of 70-200 ms or complexes; short spikes of 20-70 ms are instead removed by the preliminary down-sampling considered in the paper). The adaptive algorithm is able to sample them precisely (Fig. IIA), at the expense of reducing the fit of other portions of the signal, in which normal activity is present. Moreover, the two seizures are sampled at high frequency (compare Fig. IIA and C). In Figure III, the proposed algorithm adapts the sampling frequency to the acceleration data, increasing the number of measurements during the main activities of the subject (Fig. IIIC), obtaining a good approximation of the important portions of the signal (Fig. IIIA and 4B). Note the small delay needed by the adaptive algorithm to increase the sampling of a burst (magnification in Fig. IIIA). The only noticeable errors of the adaptive algorithm are at the beginning of the bursts (Fig. IIIB), when the method is still applying the intensive down-sampling of the data adopted during quiet standing.

In general, the approximation error is lower using the adaptive sampling than the uniform schedule. Moreover, the uniform sampling has a better reconstruction of the noisy portions of the signals (where the adaptive method reduces the sampling rate to a minimum), whereas the adaptive algorithm provides a better estimate of the bursts, reflecting the activity of the investigated system.

A further confirmation of the relevance of the algorithm is given by the analysis in the frequency domain. The comparison of the PSDs shown in panel D of Figures I-III indicates that the adaptive under-sampling allows to estimate properly the high frequency components present in the signals, even under a drastic reduction of the number of samples. On the contrary, a uniform under-sampling provides an approximation of the signals that cannot replicate the high frequency content, as the sampling frequency is under the Nyquist limit.

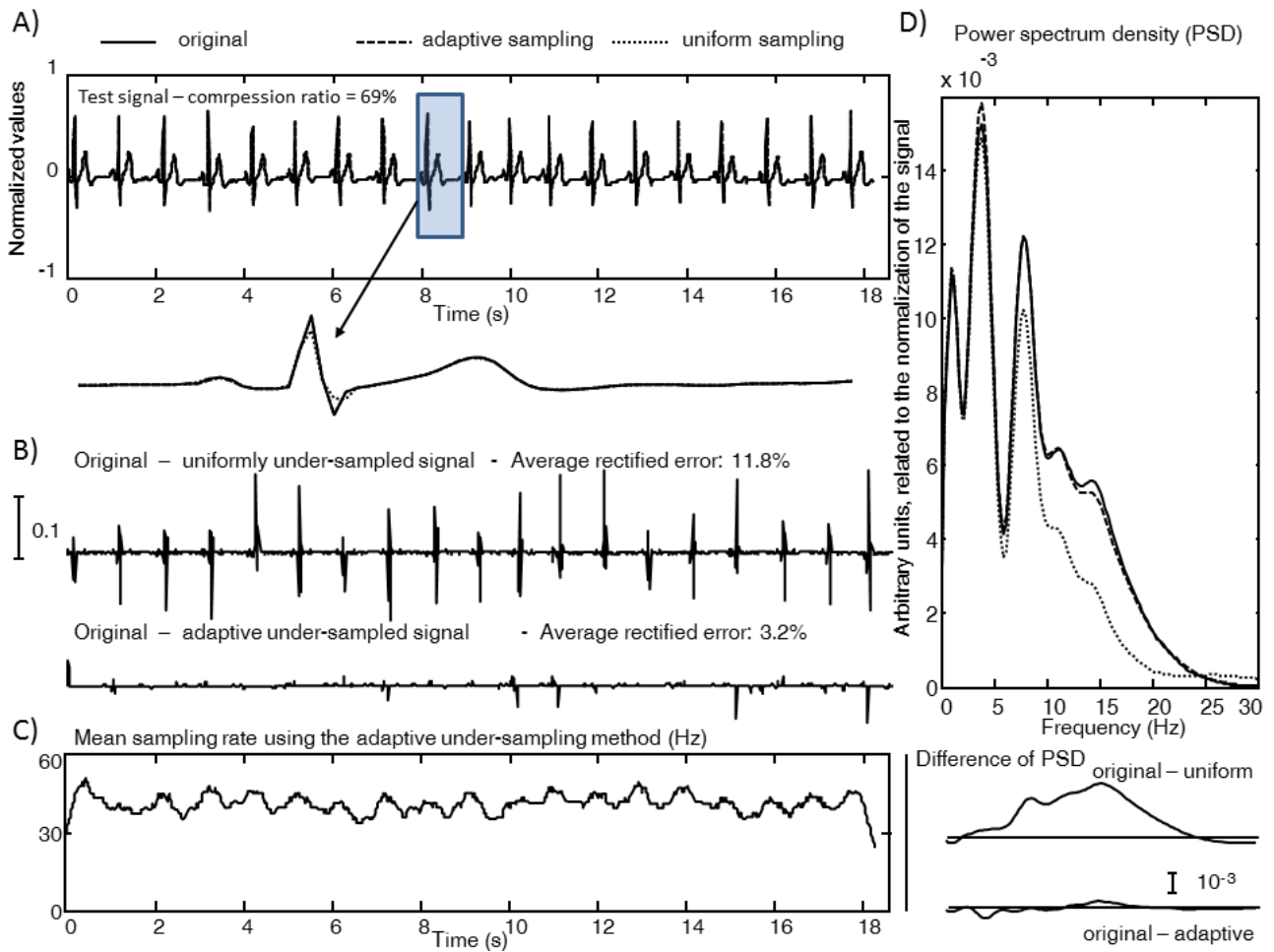


Figure I. Example of application of the under-sampling algorithm to ECG (normalized data with respect to the range; uncertainty of a normalized measurement assumed to be 10^{-2} ; maximal reduction 90%). A) The normalized signal is shown in the time domain superimposed to the under-sampled versions. A magnification of a portion is also shown. B) The differences between the test signal and the approximations obtained by uniform or adaptive under-sampling are shown. C) Average sampling rate (computed in terms of the number of selected samples on sliding windows of 50 samples). D) PSD of the normalized test ECG and of the under-sampled versions, computed by the Welch algorithm using windows of 0.5 s with 25% of overlapping.

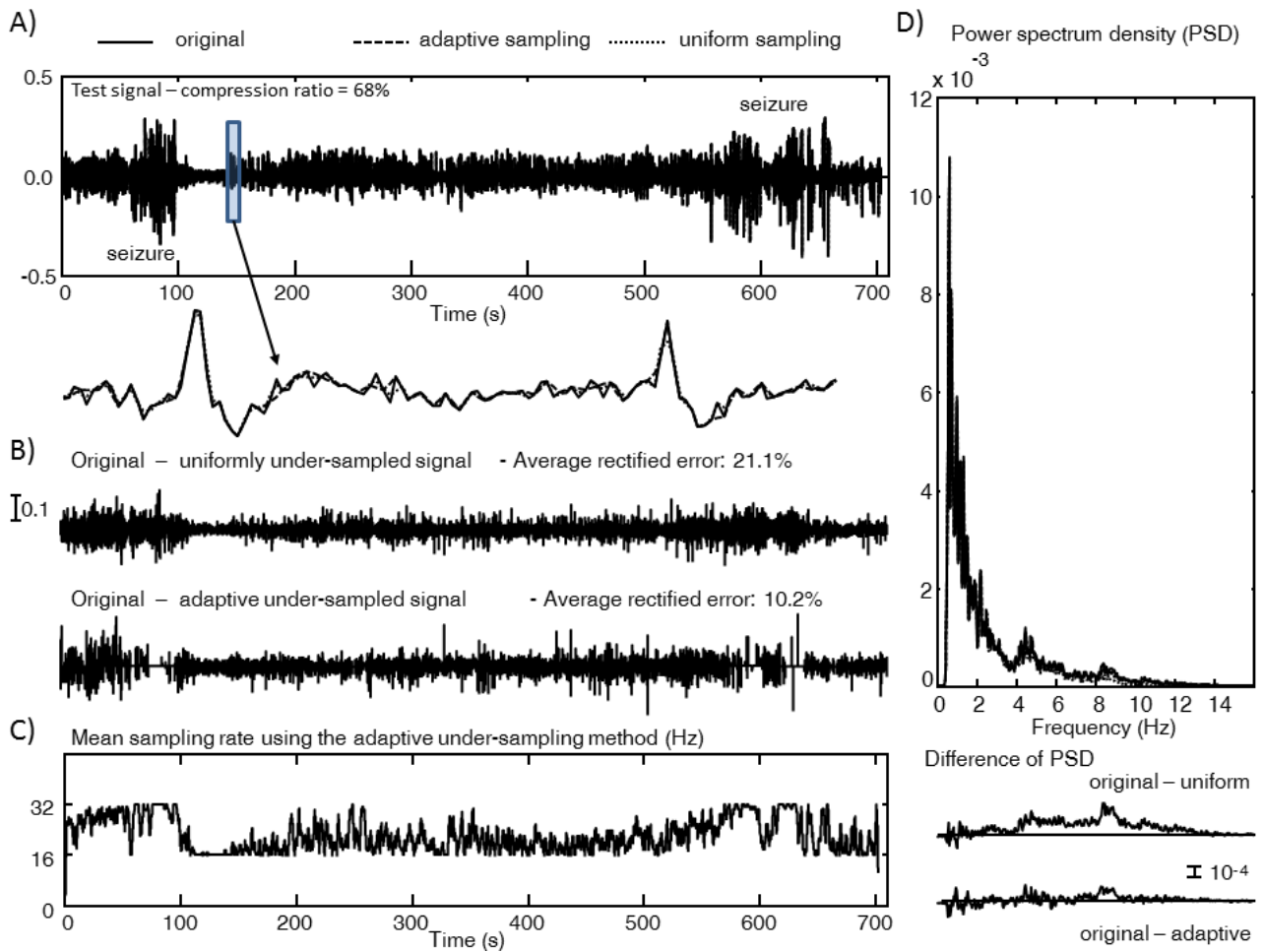


Figure II. Example of application of the under-sampling algorithm to EEG (normalized data with respect to the range; uncertainty of a normalized measurement assumed to be 10^{-2} ; maximal reduction 50%). A) The normalized signal is shown in the time domain superimposed to the under-sampled versions. The magnification of a portion is also shown. B) The differences between the test signal and the approximations obtained by uniform or adaptive under-sampling are shown. C) Average sampling rate (computed in terms of the number of selected samples on sliding windows of 50 samples). D) PSD of the normalized test EEG and of the under-sampled versions, computed by the Welch algorithm using windows of 25 s with 25% of overlapping.

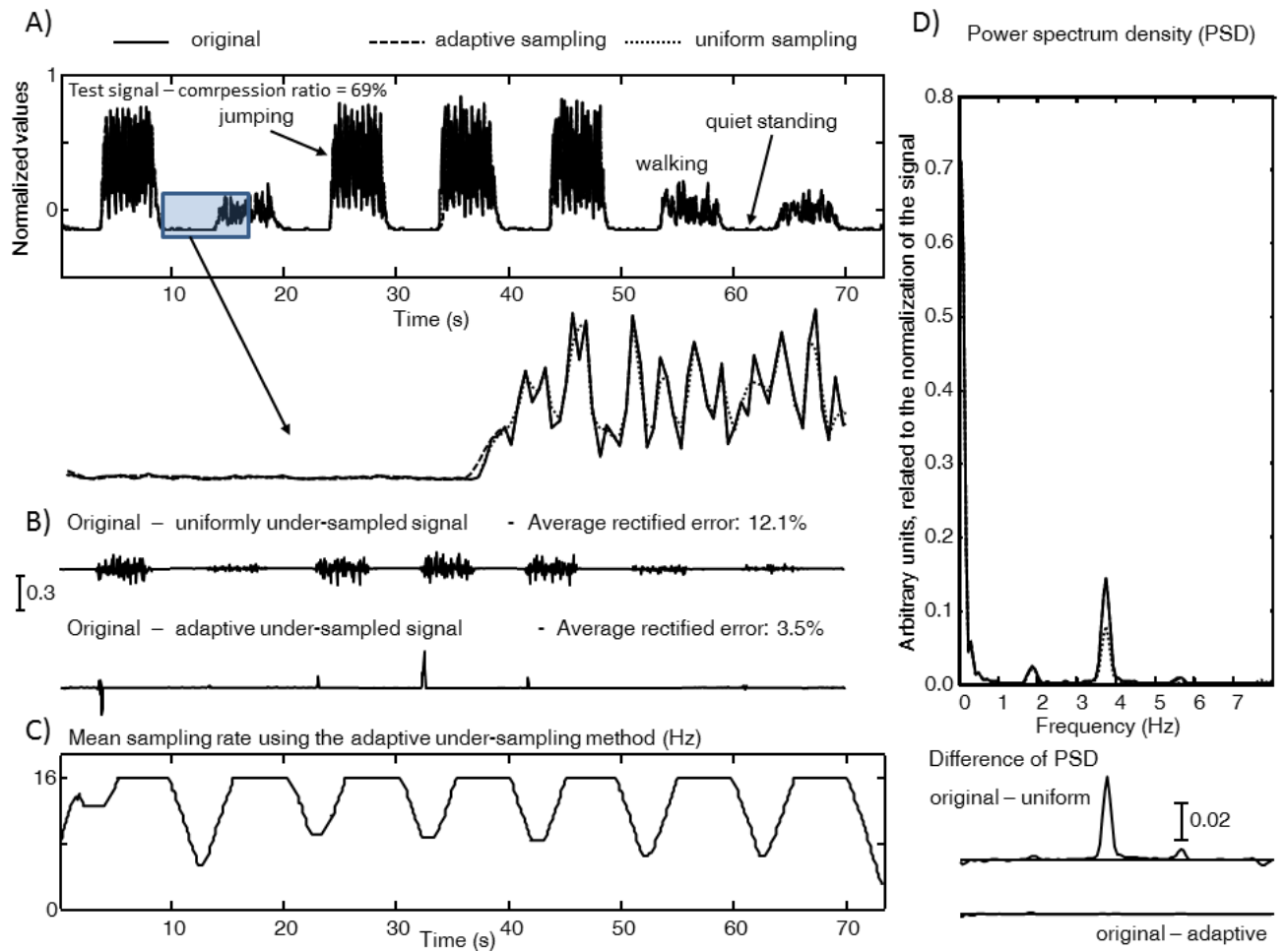


Figure III. Example of application of the under-sampling algorithm to acceleration data (normalized data with respect to the range; uncertainty of a normalized measurement assumed to be 10^{-2} ; maximal reduction 90%). A) The normalized signal is shown in the time domain superimposed to the under-sampled versions. A magnification of a portion is also shown. B) The differences between the test signal and the approximations obtained by uniform or adaptive under-sampling are shown. C) Average sampling rate (computed in terms of the number of selected samples on sliding windows of 50 samples). D) PSD of the normalized test acceleration and of the under-sampled versions, computed by the Welch algorithm using windows of 10 s with 25% of overlapping.