On the use of wavelet denoising and spike sorting techniques to process ENG signals recorded using intra-neural electrodes

Luca Citi [a,b], Jacopo Carpaneto [a], Ken Yoshida [c,d],
Klaus-Peter Hoffmann [e], Klaus Peter Koch [e,1], Paolo Dario [a],
Silvestro Micera [a,*]

[a] *ARTS and CRIM Labs, Scuola Superiore Sant'Anna, Pisa, I*

[b] *IMT Institute for Advanced Studies Lucca, I*

[c] *Center for Sensory-Motor Interaction, Aalborg University, DK*

[d] *Biomedical Engineering Dept, Indiana University-Purdue University, Indianapolis, USA*

[e] *Fraunhofer Institute for Biomedical Engineering, St.Ingbert, D*

Abstract

Among the possible interfaces with the peripheral nervous system (PNS), intra-neural electrodes represent an interesting solution for their potential advantages such as the possibility of extracting spikes from electroneurographic (ENG) signals. Their use could increase the precision and the amount of information which can be detected with respect to other processing methods.

In this study, in order to verify this assumption, thin-film longitudinal intra-fascicular electrodes (tfLIFE) were implanted in the sciatic nerve of rabbits. Various sensory stimuli were applied to the hind limb of the animal and the elicited ENG signals were recorded using the tfLIFEs. These signals were processed to determine whether the different types of information can be decoded. Signals were wavelet denoised and spike sorted. Support vector machines were trained to use the spike waveforms found to infer the stimulus applied to the rabbit. This approach was also compared with previously used ENG processing methods.

The results indicate that the combination of wavelet denoising and spike sorting techniques can increase the amount of information extractable from ENG signals recorded with intraneural electrodes. This strategy could allow the development of more effective closed-loop neuroprostheses and hybrid bionic systems connecting the human nervous system with artificial devices.

Key words: ENG signals, spike sorting, neuroprostheses, intra-neural interfaces, bionics, bio-robotics, neuro-robotics, cybernetic prostheses.

1 Introduction

Interfaces with the peripheral nervous system (PNS) can be used as a means to create a bi-directional link between the user's nervous system and an artificial device. The neural prosthetic interface can be used to induce activity in nerve fibers through electrical stimulation to place information into the nervous system. Conversely, information from the nervous system could be retrieved by recording the electrical activity of the nerve. Given a chronically stable device acting on an appropriate set of nerve fibers, such an interface could be used as part of a FES (functional electrical stimulation) system to restore function to paralyzed limbs (Popović et al., 1993) or in a brain-controlled robotic limb application (Micera et al., 2006).

In all these applications, the processing of raw electroneurographic (ENG) recordings from the neural interface can be used to reduce noise and to estimate the neural information source. Currently, there are various electrodes under investigation as neural interfaces (see Navarro et al., 2005, for a review on the available electrodes). Their characteristics determine the choice of the signal processing method. Electrodes with limited selectivity (e.g., cuffs) can, generally, only record the compound activity formed by the superposition of action potentials belonging to many axons. Therefore, in most cases, the neural activity recorded in this way has been used for onset detection, for example for the control of a 1-DoF hand prosthesis (Stein et al., 1980) or of FES-systems in hemiplegics Hoffer et al. (2005). Even if these limits can be partly overcome by using multi-site cuff electrodes (Yoo and Durand, 2005) and advanced processing techniques (Micera et al., 2001; Cavallaro et al., 2003; Lin et al., 2007; Tesfayesus and Durand, 2007), more selective PNS interfaces may be necessary to access more specific information. In fact, higher selectivity interfaces make possible the identification of single spikes from single axons (or a small group of axons) and to access the natural frequency coded information (Micera et al., 2006) in ENG signals.

Among the possible choices, intraneural electrodes represent an interesting solution because of their trade-off between invasiveness and selectivity (Yoshida and Stein, 1999; Warwick et al., 2003; McDonnall et al., 2004). In particular, longitudinal intrafascicular electrodes (LIFEs), which are wire-based electrodes inserted longitudinally into the nerve (Li et al., 2005; Lawrence et al., 2004), have been used in the past (Goodall and Horch, 1992; McNaughton and Horch, 1994; Mirfakhraei and Horch, 1997) to identify single units in multi-unit peripheral nerve recordings using different features and classification schemes.

———
* Corresponding author.
  Email address: `micera@sssup.it` (Silvestro Micera).
[1] K P Koch is now with the Electrical Engineering and Information Technology Dept, University of Applied Sciences Darmstadt, D

For example, artificial neural networks allowed differentiation of 4 to 5 units with a 70 to 90% reliability with single channel or differential recordings and a 90 to 98% reliability with dual channel recordings (McNaughton and Horch, 1994).

These very promising results could be improved by using different processing algorithms. For example, wavelet denoising techniques have been shown in the past to be a valuable tool for the analysis of signals recorded from the central nervous system (CNS) (Oweiss and Anderson, 2001; Nenadic and Burdick, 2005) and from the PNS during microneurography (Diedrich et al., 2003).

At the same time, the selectivity of intraneural electrodes (e.g., extraction of single units) enables the development of approaches based on spike sorting techniques borrowed from cortical array signal processing (Wheeler and Heetderks, 1982; Bankman et al., 1993; Lewicki, 1998; Welsh and Schwarz, 1999; Schwartz, 2004; Zhang et al., 2004; Bar-Hillel et al., 2004; Nenadic and Burdick, 2005).

The combined use of wavelet denoising and spike sorting algorithms could increase the amount of information that is decoded from intraneural recordings in the PNS. This could allow the development of more effective neuroprosthetic systems.

The aim of the present study was to assess the performance of these methods while decoding neurally derived information recorded by using intraneural electrodes. The approach proposed in this manuscript was further compared with other previously used algorithms.

2 Methods

2.1 Experimental setup

2.1.1 thin-film LIFEs

A new version of the LIFEs, named the thin film LIFEs (tfLIFE) was used in the experiments (Yoshida et al., 2006; Hoffmann and Koch, 2005). These electrodes were developed on a micropatterned polyimide substrate which was chosen because of its biocompatibility, flexibility and structural properties (Stieglitz et al., 2005). After microfabrication, this substrate filament (shown in Fig. 1) is folded in half so that each side has 4 active recording sites. Therefore, tfLIFEs allow multi-unit peripheral nerve recordings at 8 recording sites per structure. A tungsten needle linked to the polyimide structure is used
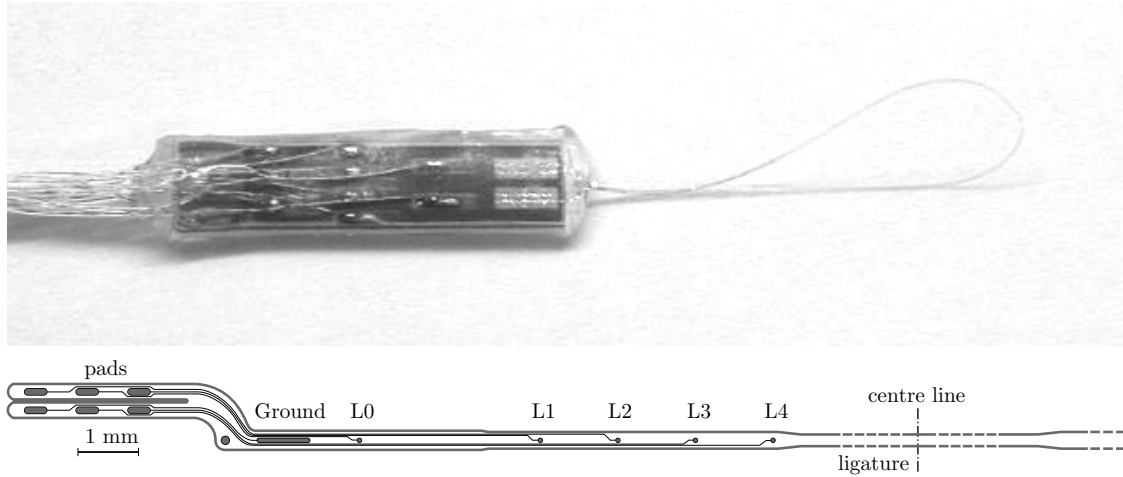
Fig. 1. Picture and unfolded overview of tfLIFE (Hoffmann and Koch, 2005). Total length: 60 mm. Length without pad areas: 50 mm. Each end of the tfLIFE carries a ground electrode (GND), an indifferent recording electrode (L0, R0) and the recording sites (L1-L4, R1-R4).

for implanting the electrode and is removed immediately after insertion.

2.1.2 Animal preparation

The experimental procedures were approved by the Danish Committee for the Ethical Use of Animals in Research. A set of protocols, including the one described in this paper, were carried out on a total of six adult (8-9 months old) female New Zealand White rabbits of approximately 4-4.5 kg. Anaesthesia was induced in the rabbits using an intramuscular injection of a Hypnorm/Dormacrom cocktail (0.15 mg/kg Midazolam (Dormicum, Alpharma A/S, Norway), 0.03 mg/kg Fetanyl and 1 mg/kg Fluranison combined in Hypnorm, Janssen Pharmaceutica, Belgium). TfLIFEs were implanted through a lateral access to the sciatic nerve between the biceps femoris and abductor cruris cranialis muscles. A second posterior access was created to expose the popliteal fat pad, which was removed to allow visualization of the branches of the sciatic nerve. The medial gastrocnemius nerve and lateral gastrocnemius/soleus nerves were identified by visual inspection, and by tracing the nerve to the muscle.

2.1.3 Sensory stimuli

The protocol described in this paper was carried out on five rabbits. During each session various sensory stimuli were applied to the hind limb of the rabbit and the elicited signals were recorded using the tfLIFEs. Stimuli were, for example, ankle flexion/extension, flexion/extension of one or more toes, and

stroking cutaneous receptive fields. The various stimuli were selected as a means to obtain adequate stimuli to activate mechanoreceptors on the paw, and second to localize the activity. Between 50-100 g force was exerted during the stroking stimulation to cause some stretching of the skin, but not sufficient to be considered painful. Both the type and the location of stimulation were retained in the experimental record. A similar technique was used by Goodall et al. (1993).

Although the implant location was kept constant between animals, the exact units and thus the sensors recorded by the electrode varied from animal to animal since the units recorded by the electrode are those that happened to be closest to the electrode site after implantation. The analysis conducted was on the activity from the set of units detected. The intent of the study was to obtain a representative set of neural activity given the current state of the art LIFE. The electrode and the implant site were not optimized to obtain the best set of units.

In one of the five sessions there was a clear increasing activity only during dorsiflexion and plantarflexion while no perceptible activity was correlated with the action of stroking superficial receptive fields. Since we wanted to test the algorithm in discriminating at least four classes of stimuli, this session was discarded.

Table 1 reports a more detailed description of the different stimuli applied during each of the four remaining sessions.

2.1.4   ENG signal acquisition

Signals were amplified by using an 8 channel custom built low-noise headstage amplifier. It had a gain of 1000x, a 1st order high-pass filter at 0.1 Hz, and an input impedance of approximately 10 MΩ. The recording chain following the amplifier consisted of an Axon Cyberamp 380 which high-pass filtered the data at 1 Hz with a 2nd order Bessel filter, to remove any residual DC offset, and post-amplified the signal obtaining an overall gain of 2500x. The signals were then acquired by a customized Alexis XT 16 bit digital tape recorder with a built-in anti-aliasing filter. Each channel was sampled simultaneously with the others with a sampling rate of 48 kHz.

2.1.5   Channel combination

The acquired data contained simultaneous monopolar recordings from different active sites of the tfLIFE electrodes (Fig. 1). For each animal, the best monopolar or differential recording pair for a given stimulus was chosen using an automated procedure.

Table 1
Stimuli applied to the hind paw of the rabbits.

Session	Stim. description	Stim. label coarse
A	Squeezing the foot	sqf
	Ankle flexion	af
	Toe extension	te
	Toe extension combined with ankle flexion	af_te
C	Stroking medial plantar area of paw	smpp
	Stroking distal plantar area of paw	sdpp
	Extension of toes (2nd, 3rd, 4th)	et
	Ankle flexion	af
D	Ankle flexion at 90º	a90
	Release from ankle flexion	a90r
	Ankle extension at 175º	a175
	Stroking paw with ankle at 90º	sp_a90
	Stroking paw with ankle at 175º	sp_a175
E	Ext. of 2nd and 3rd toes with ankle neutral	te_an
	Ext. of 2nd and 3rd toes with ankle flexed	te_af
	Flex. of 2nd and 3rd toes with ankle flexed	tf_af
	Ext. of 2nd and 3rd toes with ankle extended	te_ae

Starting from the assumption that the noise is Gaussian while the signal is not, signal amplitudes exceeding four standard deviations of the estimated noise amplitude were considered action potential peaks (Diedrich et al., 2003).

The following procedure was performed on every channel and differential combination of channels. The noise variance σ_n was evaluated on aquiescent epoch. For every epoch of stimulation a measure of spikiness s_{epoch} was evaluated as the ratio of the fraction of samples whose amplitude exceeded $4\,\sigma_n$ divided by what is expected for a Gaussian distribution ($6.3{\times}10^{-5}$). These measures were grouped by stimulus type and averaged, obtaining the spikiness of each stimulus type s_{stim}. The final figure of merit s_{sig} of the signal (i.e. of the combination of channels) was the minimum s_{stim} of the different stimulus types.

Finally, the channel combination with higher s_{sig} was chosen. This procedure reflects the considerations an expert of the field would make by visual inspection of the recordings, but has the advantage of being objective, reproducible and automated.
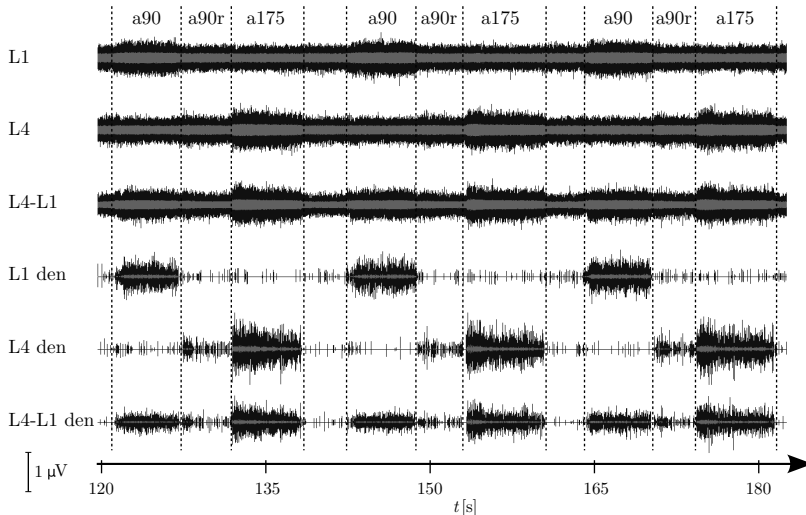
Fig. 2. One minute of recordings related to session D while some stimuli (a90, a90r, a175, see Table 1) were applied to the paw. The first two plots show the raw signal recorded by two different active sites of the tfLIFE while the third is the difference of the two signals. Below, the same signals after the wavelet denoising.

The combination with highest score was often from a differential configuration. There are three possible reasons for this: First, a differential configuration allows combining a site that responds better to one stimulus with another one that responds better to a different type as shown in Fig. 2. This leads to some loss of information but avoids the computational cost of processing more signals. Second, the recording configuration is a differential recording thus canceling noise from distant sources. Third, some characteristics of the typical spike waveform that can ease spike detection are enhanced by this configuration. In fact, most spike waveforms found were approximately composed of a positive wave followed by a negative one, or vice versa (solid line in Fig. 3). Therefore, the conduction speed of the action potential along the nerve fibre determines a delay between the sites that can superimpose upon the positive wave observed by one site and the negative one observed by the other. The difference of the two signals can enhance the performance (Fig. 3).

2.2 Signal preprocessing

2.2.1 Downsampling

Recordings acquired at 48 kHz were later downsampled to 12 kHz. This should not lead to significant loss of information since most of the physiological information is below 2 kHz as reported, among others, by Diedrich et al. (2003) who found 10 kHz to be the optimal sampling frequency for similar record-
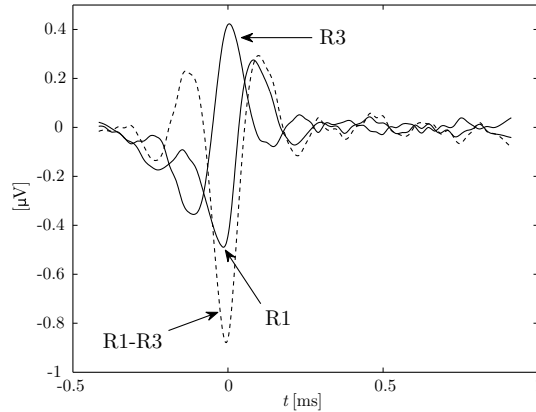
Fig. 3. One of the spike templates typically found by the algorithm. With solid stroke: the waveforms recorded by two active sites of the tfLIFE (R1 and R3). With dashed stroke: the difference of the two signals.

ings. Nevertheless, signals were sampled at 48 kHz to benefit of the pros of oversampling, mainly the lower phase distortion in the upper part of the band.

The data were downsampled using "sox", a multiplatform utility for audio files conversion. The algorithm used was a polyphase filter with Nuttal ($\sim$ 90 dB stopband) window and approximate filter length of 1024 samples.

2.2.2 *Wavelet denoising*

Wavelet denoising is a set of techniques for removing noise from signals and images. It has been used in biomedical signal processing to reduce background noise that can be approximated to a Gaussian distributed random source (e.g., Tikkanen, 1999; Oweiss and Anderson, 2001; Kim and Kim, 2003; Diedrich et al., 2003).

The main idea is to transform the noisy data into an orthogonal time-frequency domain. In that domain, thresholding was applied to the coefficients to remove the noise, and the coeffecients were finally transformed back into the original domain de-noised.

A decomposition scheme based on the *Translation-Invariant Wavelet Transform* (Coifman and Donoho, 1995) was used. It is equivalent to the Stationary Wavelet Transform and to the Undecimated Wavelet Transform. It was chosen since it is invariant to signal time shifts unlike the usual wavelet denoising performed via Dyadic Walelet Transform (DWT). This can be a key point when dealing with abruptly changing signals such as spikes in our case. Fig. 4 shows how the same spike can be captured, missed, or "incorrectly" extracted by a DWT-based denoising algorithm depending on the time the spike takes place (e.g., signals #0, #4, and #7).
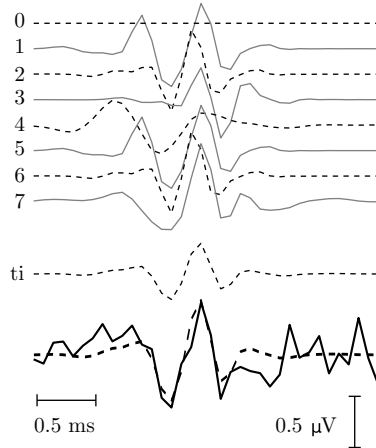
8

Fig. 4. DWT-based denoising of the same signal performed with time shifts from 0 to 7 samples. The usual denoising (signal #0) completely misses the spike while it would have detected it, in different ways, if the spike happened a few samples before or later (signals #1 to #7). The Translation-Invariant Wavelet Transform (the average of the 8 signals above, labelled with "ti") detects it and the outcome is always the same. At the bottom, the original signal (solid line) and a scaled version of "ti" (dashed line).

The *Symmlet 7* mother wavelet and *hard-thresholding* were used, as by Diedrich et al. (2003), because this choice outperforms other alternatives when working with similar neural signals.

Threshold selection is important because it determines how aggressive the denoising is. In this paper, the minimax threshold method (1) was used

$$\theta = \sigma \left(0.3936 + 0.1829 \log_2(N) \right) \tag{1}$$

σ being the standard deviation of the noise, and N the number of samples in the signal s. In statistics the minimax principle is often used to design estimators. The denoised signal can be thought of as an estimator of the unknown regression function. The minimax estimator minimizes the maximum mean square error over a given set of functions (Donoho and Johnstone, 1994). The use of the minimax leads to smaller thresholds as compared to the corresponding universal threshold by a factor 0.7 to 0.8 for the values of N used in this work. This threshold selection is more conservative than the universal one and is more adequate when significant details of the signal lie near the noise range. This is exactly the case with the low signal to noise ratio typical of neural signals.

As the noise is not assumed to be necessarily white, the noise standard deviation was estimated at each decomposition level l on a 45 seconds quiescent

9

epoch using

$$\sigma_l = \frac{\text{median}_k(|c_l(k)|)}{0.6745}. \tag{2}$$

In the standard wavelet denoising procedures, the thresholding function is applied to the details at each decomposition level while the approximation is left unchanged. In the present work, the approximation was completely discarded because it contained components below around $2^L f_n$, where L is the maximum level of decomposition and f_n is the Nyquist frequency, i.e., half the sampling frequency. $L = 3$ was chosen in order to filter out frequencies below 750 Hz.

2.3 Spike Sorting

If a sample of the denoised signal was greater than a detection threshold θ_d, then a time window around the spike was extracted. The window had a variable size in order to account for wide spikes - and consider the whole waveform - and narrow ones - and allow them to repeat at higher rates. The detection threshold was chosen to be 3 times the standard deviation of the samples in the quiescent epoch.

In order to identify similar spike waveforms in the neural signal, a spike sorting algorithm developed in C/C++ was used on the extracted windows. It consisted of a two phase process. During the first phase a set of spike templates was created. During the second phase the spikes in the signal were compared to each template and labelled as belonging to its best match.

2.3.1 Templates creation

If the extracted spike was similar to any template present in the set, the corresponding template was updated taking into account the new element, otherwise a new template was created. To assess similarity between the spike and the template, the spike was first aligned to the template using the lag that maximizes the cross-correlation, and then the following two criteria were checked:

- The correlation coefficient between the two was greater than 0.9;
- The ratio between the mean square difference of the two and the power of the template was less than 0.5.

Among the templates meeting these criteria, the one satisfying them best was chosen and the spike was labelled as belonging to that template. The template

was then updated as the weighted average of the new spike and the template, the latter having weight of the number of spikes that already joined the template. To reduce the influence of sporadic irrelevant spike templates, the ones formed by less than 0.5% of the overall number of spikes were discarded. The set of final templates was saved in an xml file for use in the next steps.

2.3.2   Templates matching

The second step was identical to the first one with the exception that the templates found in the first phase were not updated. Every spike was labelled with the number of the best matching class in terms of the above mentioned criteria.

2.4   Classification

2.4.1   Features

For each rabbit, the recordings were annotated with the types of stimulus applied in each phase of the experiment. For each stimulus type, 6 to 12 epochs, during 3 to 6 seconds, were labelled.

Each epoch was an example that was used to train the classifier or to test its generalization skills. The feature vector was made of the ratios between the number of spikes matching each template and the total number of spikes in the epoch, i.e. $\mathbf{F} = [f_1, \ldots, f_M]$ where M was the number of templates found during the templates creation phase and each f_i is given by

$$f_i = \frac{n_i}{\sum_j n_j}. \tag{3}$$

Therefore, the absolute spike rates were not used, but rather the relative spike rates of each waveform w.r.t. the others. This should prevent classification of the epochs based on the "quantity of activity" and favour the use of the "quality of activity" intended in terms of different waveforms for different stimuli.

2.4.2   Classifier

In order to infer the type of stimulus applied during a given epoch from the feature vector $\mathbf{F}$, a classifier based on *Support Vector Machines* (Cortes and Vapnik, 1995) was used making use of the open source library LIBSVM (Chang and Lin, 2001).

Support Vector Machines (SVMs) are a family of supervised learning methods for two-class classification problems. The type of SVM used is called ν-SVM (Chen et al., 2005). Among the various kernels available, the radial basis function (RBF) was chosen because it allows complex separation surfaces requiring a reduced number of hyperparameters to tune.

Even if this could lead to a more conservative result, it was decided to use fixed hyperparameters because tuning them by means of iterative methods would have required an additional cross-validation scheme. This would have further reduced the already small number of examples available to train and test the classifier. Using ν-SVM with RBF, hyperparameters ν and γ need to be chosen. The regularization parameter ν is is an upper bound on the fraction of margin errors and a lower bound on the fraction of support vectors. We chose $\nu = 0.4$ because we considered it a good trade-off between allowing training errors and favouring smooth separation surfaces. The parameter γ determines the radius of the RBF. We set $\gamma = 1/\rho^2$ where ρ is the radius of the smallest sphere in the input space that contains all feature vectors F of the training set. Keerthi (2002) reported that it is a good starting point for iterative methods.

To allow SVMs, and other binary classifiers, to handle multiclass problems, the latter must be decomposed into several binary problems. In this work we used a one-against-one approach (Huang et al., 2006) where, for a q-class classification problem, $q(q-1)/2$ machines were trained. Each SVM separates a pair of classes and, in the prediction stage, a voting strategy was used.

Each attribute of the feature vectors of the training set was scaled in the range $[-1, +1]$ and the scaling parameters were saved to later scale testing data in the same way. The main advantage was to avoid features in greater numeric ranges dominating those in smaller numeric ranges.

2.5   Assessment of the results

2.5.1   Validation scheme

A validation scheme was developed by using a single example per class (stimulus type) as the test set (retained "blind", i.e., never used to tune any parameter, to create spike templates, or to train the machine), and the remaining ones as the training set. For example, if a given session had 9 epochs for each of 5 different types of stimuli, the test set would consist of 5 epochs while the training set of 40 epochs. In order to obtain an average performance with small confidence intervals a random subsampling cross validation was used. The following procedure was repeated 5000 times:

- build the test set by randomly selecting one epoch for each stimulus type

(e.g., one set of 5 epochs among 9^5 combinations for the example above);

- build the training set by selecting the remaining epochs;
- run the template creation phase of the spike sort algorithm on the training set;
- run the template matching phase of the spike sort algorithm on the whole set;
- use the features from the training set to train a SVM machine;
- make the SVM machine predict the stimuli type for the test set and compare it with the ground truth.

2.5.2 Inspection of results

The results were evaluated in multiple ways in order to look at them from different points of view to clearly identify the potentials and limits of this approach.

The overall percentage of correct classifications (PC) was calculated as the ratio between the number of epochs whose stimulus type was correctly identified and the total number of epochs classified. This parameter can provide a rapid synthetic indication of the average performance of the system.

The confusion matrix was also inspected to identify subsets of classes the system recurrently confuses.

An interesting perspective is to analyse the results from an information theory point of view. The system can be interpreted as a discrete memoryless noisy communication channel where the actual stimulus type is the input and the predicted stimulus type is the output. For such channel the *mutual information* is a measure of how much information can be obtained from the input random variable (U) by observing the output one (Y). The mutual information depends on the probability distribution of the input variable U. The maximum of the mutual information over all possible distributions defines the *channel capacity*, i.e., the maximum amount of discrete information that the channel can carry. The channel capacity C has been evaluated by using the Arimoto algorithm (Arimoto, 1972).

2.5.3 Comparison with other techniques

The results achieved using the method proposed in this manuscript were also compared with the performance of previously used methods which can be implemented in analog circuitry (i.e., not requiring complex denoise techniques nor spike detection and sorting). This comparison was carried out in order to understand the relative importance of each processing step (i.e., the wavelet denoising and the spike sorting).

Two preprocessing methods were compared: the wavelet denoising stage previously introduced (abbreviated to WD) and a bandpass filter (FIR). In the latter case the input data were processed with a 90-tap equiripple FIR filter with cutoff at 700 and 2000 Hz (as in Diedrich et al., 2003).

Two different feature creation methods were compared. The first one (abbreviated to Srt) is the previously introduced spike sorting algorithm that only accounts for relative spike activity and gives the feature vector $\mathbf{F}$. The second one (abbreviated to RBI) is a traditional RBI (rectified and bin-integrated) algorithm evaluated as the mean over the epoch of the RBI computed on 50 ms windows.

The validation scheme is the same as before but the preprocessing stage is either WD or FIR, and the feature creation is Srt, or RBI.

To assess statistical significance of the possible improvement introduced by the use of WD over FIR, and Srt over RBI, the table of percentage of correct classifications was fitted by a logistic regression. The analysis was performed with the statistical software package R. The hypothesized relationship was $(\mathbf{Corr}, \mathbf{Wrong}) \sim \mathbf{Prep} * \mathbf{Proc} + \mathbf{Sess}$ that, in the notation of R, means that the dependent variable (number of correct and wrong classifications) depends ($\sim$) on a the explanatory variables $\mathbf{Prep}$ (accounting for the preprocessing: WD or FIR) and $\mathbf{Proc}$ (accounting for the processing: Srt or RBI) and their interaction ($*$). The independent variable $\mathbf{Sess}$ for the session, was also included to take into account the differences among the datasets.

3 Results

3.1 Quantitative results

The performance of the system for each of the four experimental sessions was assessed. Measures of performance in terms of percentage of correct classifications and channel capacity are summarized in the WD/Srt column of Table 2.

Table 2 also presents a comparison of the different processing configurations. The first column gives the results of the method introduced in this work (WD/Srt). The last column reports the results of a typical ENG approach (FIR/RBI) using a bandpass filter between 700 and 2000 Hz and RBI. The other columns present the performance of hybrid configurations.

Table 2
Performance of sessions with 4 (A, C, E) or 5 (D) stimuli in terms of percentage of correct classifications (PC) and channel capacity (C). For PC, the radius of a 95% confidence interval (rounded up) is reported in parentheses. WD/Srt is the approach introduced in this work; FIR/RBI is a typical ENG approach using a FIR bandpass filter and rectified-bin integration; the others are hybrid configurations.

Session	Measure	WD/Srt	FIR/Srt	WD/RBI	FIR/RBI
A	PC [%]	$92.6_{(0.4)}$	$66.7_{(0.7)}$	$68.6_{(0.7)}$	$52.8_{(0.7)}$
	C [bit/symb]	1.59	0.87	1.01	0.84
C	PC [%]	$99.1_{(0.2)}$	$95.2_{(0.3)}$	$88.8_{(0.5)}$	$69.0_{(0.7)}$
	C [bit/symb]	1.92	1.72	1.50	1.28
D	PC [%]	$94.5_{(0.3)}$	$80.4_{(0.5)}$	$76.0_{(0.6)}$	$66.5_{(0.6)}$
	C [bit/symb]	2.01	1.64	1.61	1.54
E	PC [%]	$90.1_{(0.5)}$	$60.7_{(0.7)}$	$92.6_{(0.4)}$	$89.1_{(0.5)}$
	C [bit/symb]	1.51	0.64	1.71	1.51

3.2 Logistic regression

As mentioned before, a logistic regression was performed in order to assess statistical significance of the possible improvement introduced by the use of WD over FIR, and Srt over RBI. The hypothesized relationship was $(\mathbf{Corr}, \mathbf{Wrong}) \sim \mathbf{Prep} * \mathbf{Proc} + \mathbf{Sess}$. The values of the estimated coefficients for the explanatory variables are listed in Table 3. To interpret the results one can consider that the presence of a given condition X_i over the baseline (FIR+RBI) scales the odds (i.e., $PC/(1 - PC)$) by e^{β_i}. Positive values of β_i indicate increasing the performance while negative values indicate decreasing performance. It is worth noting that the use of the wavelet denoise alone increases the odds by a factor 1.96 ($e^{0.674}$) while the use of the spike sorting alone increases the odds by a factor 1.43 ($e^{0.358}$). The combined use of wavelet denoise and spike sorting increases the odds by a factor 7.40 ($e^{0.674+0.358+0.970}$). This means that the proposed algorithm introduces a significant improvement over the current techniques, that both steps are important to achieve this goal, and that when used together the improvement is bigger than the sum of the improvements when either one is used.

3.3 How the system works

For completeness, an example that can help understanding how the system internally works is also reported. Two minutes of session D are considered.

Table 3
Results of fitting the table of percentage of correct classifications with a logistic regression. All the values are statistically significant ($p < 0.001$).

Explan. variable X_i	coeff. β_i
(Intercept)	+0.218
PrepWD	+0.674
ProcSrt	+0.358
SessC	+1.205
SessD	+0.524
SessE	+0.788
PrepWD:ProcSrt	+0.970

Fig. 5 shows 5 out of the 17 templates found. Templates #3 and #4 are considered together because they have similar properties and behaviour. The same was done for templates #9 and #10. Looking at the plots in Fig. 5 it is clear how their relative and absolute spike activities varied in time as the experimenter changed the stimulus applied to the paw. This also results from Fig. 6 that shows the probability density functions of the inter-spike interval (ISI) for the three groups of templates (#1, #3+4, and #9+10) during the different stimuli.

4 Discussion

4.1 General considerations

The aim of this work was to verify whether the combined use of wavelet denoising and spike sorting algorithms could allow to increase the amount of information which can be decoded from ENG signals recorded using tfLIFEs (and in general intraneural electrodes).

To make this evaluation, different stimuli were applied to the paw of rabbits while recording the ENG signals using tfLIFEs. The results achieved in the discrimination by using a wavelet denoising algorithm, a spike sorting technique, and a SVM classifier together for only one channel (Table 2) show that, with four out of the five initial animals, it was possible to discriminate four (or five) different classes of stimuli with performance in a range between 90% and 99%. The technique outperformed prior work carried out with different approaches. Moreover, the channel capacity was in a range between 1.5 and 2 bits/symbol (approximately 20 to 50 bits/minute) which is comparable to
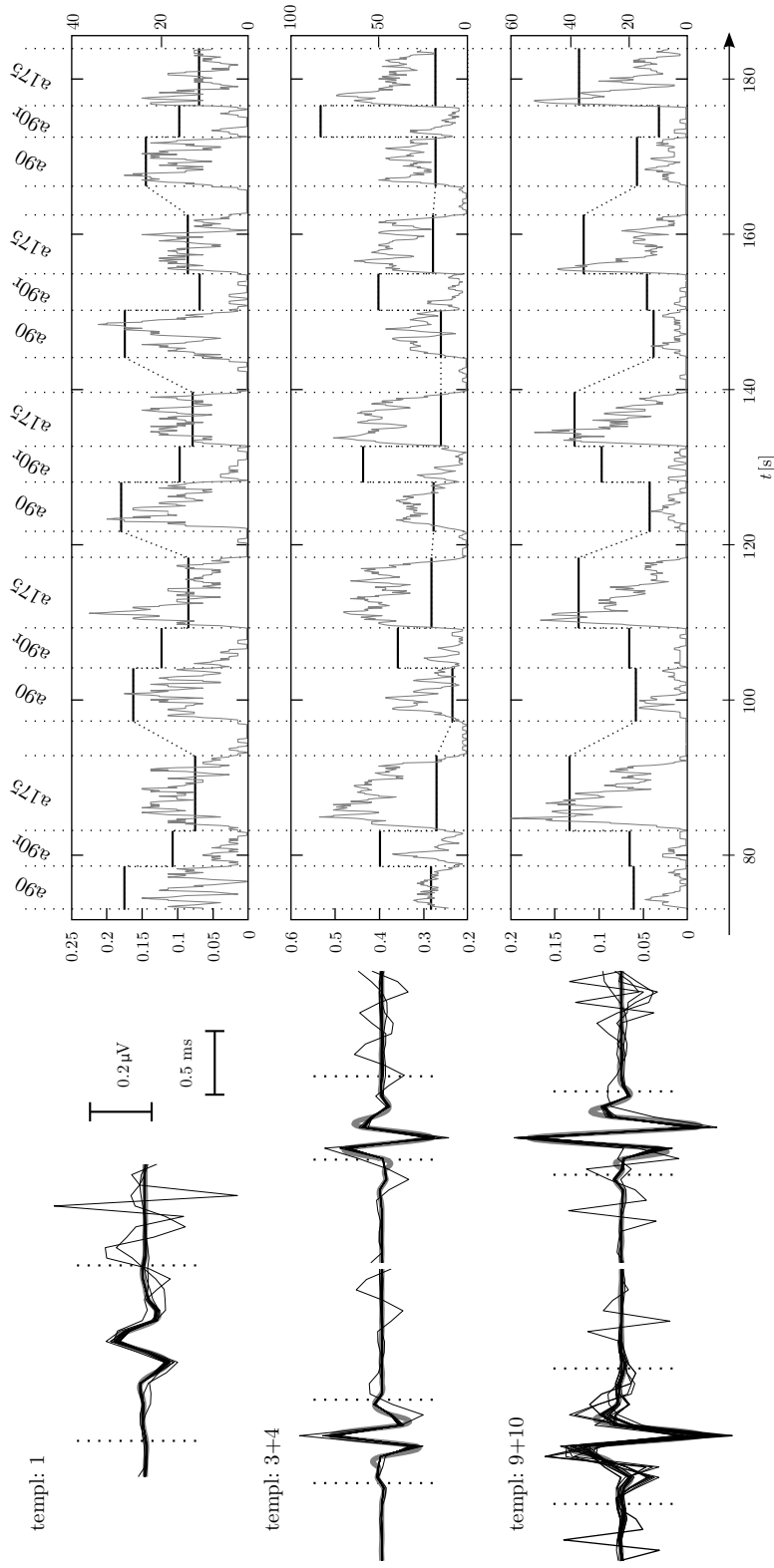
16

Fig. 5. Analysis of how different "spike waveforms" could aid in the detection of specific stimuli (a90, a90r, a175, see Table 1). On the left, some of the spike waveforms found are given. The thick gray lines represent the waveforms identified during the "template creation" phase and their support is delimited in the plots by the dotted vertical bars. The thin black lines represent some of the spikes identified as belonging to that specific template. On the right, the black lines (scales on the left) represent the ratio f_i between the occurrence of the correspondent "spike waveform" shown on the left and the total spikes detected for each epoch. The gray lines (scales on the right) are the absolute spike rate of the corresponding spike waveforms evaluated on 1-s sliding windows.
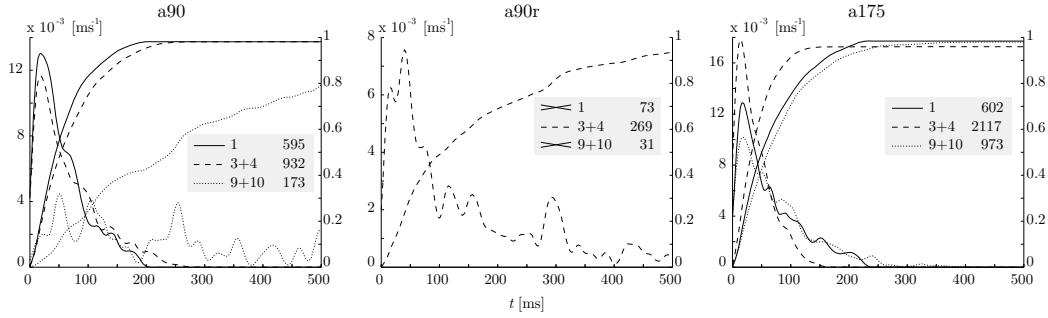
17

Fig. 6. Probability density functions (scales on the left) and cumulative distribution functions (scales on the right) of the inter-spike interval (ISI) for the three groups of templates (#1, #3+#4, and #9+#10) during the different stimuli of Fig. 5. The pdf has been estimated with kernel density estimation (or Parzen windows method) using a standard gaussian kernel and bandwidth $h = 7.5$ms. The number of ISIs used is reported in the legend near the template. When the number of ISIs was insufficient to obtain a meaningful pdf (i.e., the firing rate was very low), the corresponding plot has been omitted.

other invasive and non-invasive interfaces (Tonet et al., 2008).

It is important to point out that our attention was not focused on the use of the absolute spike rates but rather on the relative spike rates of each waveform w.r.t. the others. This approach was chosen to verify whether different "spike waveforms" extracted from tfLIFE signals could aid in the detection of specific stimuli. By using the absolute spike rate there is a further performance increase of 0 to 2% but for the sake of brevity and clarity this result was not considered in the paper.

The results have been achieved by using only the best channel available (as in Dhillon et al., 2004) and it is likely that the use of several channels simultaneously could further improve the situation as also reported by McNaughton and Horch (1994). Only one channel was used for the sake of simplicity; in fact, using additional electrodes (e.g., two channels) greatly increases the number of features for the classifier while maintaining the number of examples for the training. To avoid risks of overfitting, the options would be either to include an automated feature selection stage (e.g., PCA or ICA) to keep the dimension of the training vector reasonably small, or greatly increase the number of sessions. At the present stage the aim of the manuscript was not to state the maximum amount of information achievable, but just to give preliminary evidence that spike sorting techniques can be useful with ENG signals recorded through LIFEs and to stimulate further investigations in that direction. However, this is an open issue to be addressed in the future.

The performance of the system presented in this study outperformed previously used (e.g., FIR+RBI) methods (see Table 2). In particular, it was

18

possible to show that the combined use of both the modules of our approach (wavelet denoising and spike sorting) was able to considerably increase the classification performance (as also confirmed by a logistic regression procedure).

This method can also be implemented to work unsupervised, continuously, in real-time by using compact, light-weight and battery-powered devices. In fact, several works reported the feasibility of similar algorithms running on small low power devices, such as stationary wavelet-based denoising algorithm on programmable DSP and FPGA (Montani et al., 2003), and algorithms for spike detection, alignment and spike sorting on VLSI architectures (Zumsteg et al., 2005; Zviagintsev et al., 2006). Signals recorded from chronic setups are dynamic in that the responses to repeated stimuli or motor commands change in time. To cope with this issue, the thresholds for the wavelet denoising and the spike detection can be updated adaptively as well as the spike templates of the spike sorting algorithm (Mirfakhraei and Horch, 1997).

4.2   Use of this methodology for the control of limb prostheses

Even if the ENG signals recorded and processed in this study were related to neural afferent information (instead of motor commands) the algorithms developed could be used in the near future to extract voluntary commands from efferent motor signals. This represents a very important step in developing an ENG-based control system for limb prostheses, extending the results shown by Dhillon and Horch (2005).

However, in this case, a slightly different strategy in terms of classification algorithm could be necessary. In fact, in order to select different grasping tasks it could be necessary to identify the presence of several specific spike classes related to a complex multi-muscle limb activity. Some of these spike classes could also be related to the recruitment of the same muscle but in this case it would simply require the development of a more flexible classifier embedding this kind of information.

Moreover, information such as spike rate could be correlated to the force the user would like to carry out. The combined use of spike sorting techniques and fuzzy logic algorithms could allow the development of this kind of classifier. All these considerations will be validated in animal models during the next months.

Even though the results are encouraging, it is important to point out that there is a big difference between inferring items of a reduced set of events and doing so with the device in practical use. The use of qualitative general stimuli such as brushing and stroking, instead of focusing on stimuli to quantitatively characterize single unit responses, was an attempt to generate a more natural data set of responses. The problem is well known in the field of cochlear implants where in early days a single-electrode cochlear implant could only provide some closed-set speech recognition with no open-set recognition at all while now modern devices can provide 70-80% open-set speech recognition (Zeng, 2004).

Another possible reason for performance degradation in real implants is that the anesthetized animal model used in this work provides a relatively quiet environment for recording neural activity. The noise level would be much higher in a real neuroprosthetic setting because of increased movement and background neural activity.

Moreover, the issue related to the possible overlap of different spike waveforms should be addressed by developing specific techniques.

Finally, it is important to point out that the implantation of tfLIFEs (and in general intra-neural electrodes) is blind. Thus, in some cases the position of the site was not optimal for recording useful neural information. This is the most important reason for the problems encountered with one of the five rabbits. Moreover, electrode drift during the implantation is possible. This limitation could be addressed by increasing the number of tfLIFEs implanted or by moving the different electrical contacts embedded with micro-actuators in the structure of the interface. Preliminary investigations by Bossi et al. (2007) with tfLIFEs actuated in order to follow the signal have been carried out.

5 Conclusions

In this paper, the possibility of decoding information from ENG signals recorded with tfLIFEs by using wavelet denoising and spike sorting techniques has been investigated. The results achieved seem to indicate that the approach proposed could provide better performance than different perviously used methods.

It is important to point out that, although tfLIFE signals were used in this work, the method is not limited to ENG signals recorded from tfLIFEs. It

can be applied to other intrafascicular techniques, such as signals from Utah probes, traditional LIFEs, polyLIFEs, sieve electrodes, and microneurography, where the electrode selectivity allows resolution of separable unit spike activity.

Further dedicated experiments will be carried out in order to assess whether this approach could be applied to efferent and afferent ENG signals to develop innovative FES systems and brain-controlled artificial limbs.

Acknowledgements

References

Arimoto S. An algorithm for computing the capacity of arbitrary discrete memoryless channels. IEEE Trans Inform Theor, 1972;18:14–20.

Bankman IN, Johnson KO, Schneider W. Optimal detection, classification, and superposition resolution in neural waveform recordings. IEEE Trans Biomed Eng, 1993;40(8):836–841.

Bar-Hillel A, Spiro A, Stark E. Spike sorting: Bayesian clustering of non-stationary data. In Advances in Neural Information Processing Systems, NIPS 2004. Vancouver, 2004.

Bossi S, Menciassi A, Koch K, Hoffmann KP, Yoshida K, Dario P, Micera S. Shape memory alloy microactuation of tf-LIFEs: preliminary results. IEEE Trans Biomed Eng, 2007;54.

Cavallaro E, Micera S, Dario P, Jensen W, Sinkjaer T. On the intersubject generalization ability in extracting kinematic information from afferent nervous signals. IEEE Trans Biomed Eng, 2003;50(9):1063–1073.

Chang CC, Lin CJ. LIBSVM: a library for support vector machines, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

Chen PH, Lin CJ, Scholkopf B. A tutorial on ν-support vector machines. Applied Stochastic Models in Business and Industry, 2005;21(2):111–136.

Coifman RR, Donoho DL. Translation-invariant de-noising. In Wavelets and Statistics, Lecture Notes in Statistics. SpringerVerlag, 1995, 125–150.

Cortes C, Vapnik V. Support-vector networks. Machine Learning, 1995; 20(3):273–297.

Dhillon G, Horch K. Direct neural sensory feedback and control of a prosthetic arm. IEEE Trans Neural Syst Rehabil Eng, 2005;13:468–472.

Dhillon GS, Lawrence SM, Hutchinson DT, Horch KW. Residual function in peripheral nerve stumps of amputees: implications for neural control of artificial limbs. J Hand Surg-AM, 2004;29(4):605–18.

Diedrich A, Charoensuk W, Brychta RJ, Ertl AC, Shiavi R. Analysis of raw microneurographic recordings based on wavelet de-noising technique and classification algorithm: wavelet analysis in microneurography. IEEE Trans Biomed Eng, 2003;50(1):41–50.

Donoho DL, Johnstone IM. Ideal spatial adaptation by wavelet shrinkage. Biometrika, 1994;81(3):425–455.

Goodall EV, Horch KW. Separation of action potentials in multiunit intrafascicular recordings. IEEE Trans Biomed Eng, 1992;39(3):289–295.

Goodall EV, Horch KW, McNaughton TG, Lybbert CM. Analysis of single-unit firing patterns in multi-unit intrafascicular recordings. Med Biol Eng Comput, 1993;31(3):257–267.

Hoffer JA, Baru M, Bedard S, Calderon E, G Desmoulin P. Initial results with fully implanted Neurostep FES system for foot drop. In 10th Annual Conference of the International FES Society. 2005.

Hoffmann KP, Koch KP. Final report on design consideration of tLIFE2. Tech. rep., IBMT, 2005.

Huang TK, Weng RC, Lin CJ. Generalized Bradley-Terry models and multi-class probability estimates. Journal of Machine Learning Research, 2006; 7:85–115.

Keerthi S. Efficient tuning of SVM hyperparameters using radius/margin bound and iterative algorithms. IEEE Trans Neural Network, 2002; 13(5):1225–1229.

Kim KH, Kim SJ. A wavelet-based method for action potential detection from extracellular neural signal recording with low signal-to-noise ratio. IEEE Trans Biomed Eng, 2003;50(8):999–1011.

Lawrence SM, Dhillon GS, Jensen W, Yoshida K, Horch KW. Acute peripheral nerve recording characteristics of polymer-based longitudinal intrafascicular electrodes. IEEE Trans Neural Syst Rehabil Eng, 2004;12(3):345–348.

Lewicki MS. A review of methods for spike sorting: the detection and classification of neural action potentials. Netw Comput Neural Syst, 1998; 9(4):R53–R78.

Li LJ, Zhang J, Zhang F, Lineaweaver WC, Chen TY, Chen ZW. Longitudinal intrafascicular electrodes in collection and analysis of sensory signals of the peripheral nerve in a feline model. Microsurgery, 2005;25(7):561–565.

Lin CCK, Ju MS, Cheng HS. Model-based ankle joint angle tracing by cuff electrode recordings of peroneal and tibial nerves. Med Biol Eng Comput, 2007;45(4):375–385.

McDonnall D, Clark GA, Normann RA. Interleaved, multisite electrical stimulation of cat sciatic nerve produces fatigue-resistant, ripple-free motor responses. IEEE Trans Neural Syst Rehabil Eng, 2004;12(2):208–215.

McNaughton TG, Horch KW. Action potential classification with dual channel intrafascicular electrodes. IEEE Trans Biomed Eng, 1994;41(7):609–616.

Micera S, Carrozza M, Beccai L, Vecchi F, Dario P. Hybrid bionic systems for the replacement of hand function. In IEEE Proceedings, vol. 94(10). 2006.

Micera S, Jensen W, Sepulveda F, Riso R, Sinkjaer T. Neuro-fuzzy extraction of angular information from muscle afferents for ankle control during standing in paraplegic subjects: an animal model. IEEE Trans Biomed Eng, 2001;48(7):787–794.

Mirfakhraei K, Horch K. Recognition of temporally changing action potentials in multiunit neural recordings. IEEE Trans Biomed Eng, 1997;44(2):123–131.

Montani M, De Marchi L, Marcianesi A, Speciale N. Comparison of a programmable DSP and FPGA implementation for a wavelet-based denoising algorithm. In Proceedings of the 46th IEEE International Midwest Symposium on Circuits and Systems. 2003, 602–605.

Navarro X, Krueger TB, Lago N, Micera S, Stieglitz T, Dario P. A critical review of interfaces with the peripheral nervous system for the control of neuroprostheses and hybrid bionic systems. J Peripher Nerv Syst, 2005; 10(3):229–258.

Nenadic Z, Burdick JW. Spike detection using the continuous wavelet transform. IEEE Trans Biomed Eng, 2005;52(1):74–87.

Oweiss KG, Anderson DJ. Noise reduction in multichannel neural recordings using a new array wavelet denoising algorithm. Neurocomputing, 2001; 38–40:1687–1693.

Popović DB, Stein RB, Jovanović KL, Dai R, Kostov A, Armstrong WW. Sensory nerve recording for closed-loop control to restore motor functions. IEEE Trans Biomed Eng, 1993;40(10):1024–1031.

Schwartz AB. Cortical neural prosthetics. Annu Rev Neurosci, 2004;27:487–507.

Stein RB, Charles D, Hoffer JA, Arsenault J, Davis LA, Moorman S, Moss B. New approaches for the control of powered prostheses particularly by high-level amputees. Bull Prosthet Res, 1980;10-33:51–62.

Stieglitz T, Schuettler M, Koch KP. Implantable biomedical microsystems for neural prostheses. IEEE Eng Med Biol Mag, 2005;24(5):58–65.

Tesfayesus W, Durand DM. Blind source separation of peripheral nerve recordings. J Neural Eng, 2007;4:157–167.

Tikkanen PE. Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal. Biol Cybern, 1999;80(4):259–267.

Tonet O, Marinelli M, Citi L, Rossini PM, Rossini L, Megali G, Dario P. Defining brain-machine interface applications by merging interface performance with device requirements. Journal of Neuroscience Methods, 2008;:91–104.

Warwick K, Gasson M, Hutt B, Goodhew I, Kyberd P, Andrews B, Teddy P, Shad A. The application of implant technology for cybernetic systems. Arch Neurol, 2003;60(10):1369–1373.

Welsh JP, Schwarz C. Multielectrode recording from the cerebellum. In

Nicolelis M, editor, Methods for Neural Ensemble Recordings. CRC Press, Boca raton, 1999, 79–100.

Wheeler BC, Heetderks WJ. A comparison of techniques for classification of multiple neural signals. IEEE Trans Biomed Eng, 1982;29(12):752–759.

Yoo PB, Durand DM. Selective recording of the canine hypoglossal nerve using a multicontact flat interface nerve electrode. IEEE Trans Biomed Eng, 2005;52(8):1461–1469.

Yoshida K, Hennings K, Kammer S. Acute performance of the thin-film longitudinal intra-fascicular electrode. In First IEEE/RAS-EMBS International Conference BioRob. Pisa, 2006, 296–300.

Yoshida K, Stein RB. Characterization of signals and noise rejection with bipolar longitudinal intrafascicular electrodes. IEEE Trans Biomed Eng, 1999;46(2):226–234.

Zeng FG. Trends in cochlear implants. Trends Amplif, 2004;8(1):1–34.

Zhang P, Wu J, Zhou Y, Liang P, Yuan J. Spike sorting based on automatic template reconstruction with a partial solution to the overlapping problem. J Neurosci Meth, 2004;135(1-2):55–65.

Zumsteg ZS, Kemere C, O'Driscoll S, Santhanam G, Ahmed RE, Shenoy KV, Meng TH. Power feasibility of implantable digital spike sorting circuits for neural prosthetic systems. IEEE Trans Neural Syst Rehabil Eng, 2005; 13(3):272–279.

Zviagintsev A, Perelman Y, Ginosar R. Algorithms and architectures for low power spike detection and alignment. J Neural Eng, 2006;3(1):35–42.