Development of a neural interface for the control of a robotic hand

PhD Program in Biorobotics Science and Engineering
XX Cycle

By
Luca Citi
2009
The dissertation of Luca Citi is approved.

Program Coordinator: Prof. Paolo Dario

Supervisor: Dr. Silvestro Micera,
Scuola Superiore Sant’Anna, Pisa, Italy and
Swiss Federal Institute of Technology, Zurich, Switzerland

Supervisor: Prof. Paolo Dario,
Scuola Superiore Sant’Anna, Pisa, Italy

Tutor: Dr. Oliver Tonet

The dissertation of Luca Citi has been reviewed by:

Dr. Oliver Tonet

Dr. Silvestro Micera

IMT Institute for Advanced Studies Lucca
Scuola Superiore Sant’Anna
2009
To Pierpaolo
Contents

List of Figures xi
List of Tables xiii
Acknowledgements xv
Vita and Publications xvii
Abstract xxi

1 Introduction 1

2 Taxonomy of interfaces 5
  2.1 Hybrid Bionic Systems ................................. 5
  2.1.1 Overview ........................................... 5
  2.1.2 Types of users ...................................... 6
  2.1.3 Performance measures of HBS ...................... 7
  2.2 Devices .................................................. 8
  2.3 Interfaces .............................................. 10
  2.3.1 Cortical interfaces .................................. 10
  2.3.2 Non-cortical interfaces ............................ 14
  2.3.3 Summary of interface performance ................ 15
  2.4 Results .................................................. 16
  2.4.1 Throughput and latency ............................ 16
  2.4.2 Other measures ...................................... 20
  2.5 HMIs for hand prostheses ............................. 20
  2.6 Conclusions ............................................. 24
3 Interfacing the peripheral nervous system
  3.1 Neurophysiological bases for the use PNS neural interfaces
    3.1.1 Cortical Physiology
    3.1.2 Peripheral Physiology
    3.1.3 General considerations for PNS interfaces
  3.2 Overview of PNS interfaces
    3.2.1 Cuff electrodes
    3.2.2 Sieve electrodes
    3.2.3 Multielectrodes arrays
    3.2.4 Longitudinal intra-fascicular electrodes
    3.2.5 Considerations
  3.3 Recording and processing of neural signals recorded using longitudinal intra-fascicular electrodes
    3.3.1 Nerve recording capabilities
    3.3.2 Processing of neural signals from LIFEs
  3.4 Experiments using Intra-fascicular electrodes in amputees
    3.4.1 Works by Dhillon and Horch
    3.4.2 Works by Jia and colleagues

4 Offline processing of afferent neural activity in animal model
  4.1 Experimental setup
    4.1.1 thin-film LIFEs
    4.1.2 Animal preparation
    4.1.3 Sensory stimuli
    4.1.4 ENG signal acquisition
    4.1.5 Channel combination
  4.2 Signal preprocessing
    4.2.1 Downsampling
    4.2.2 Wavelet denoising
  4.3 Spike Sorting
    4.3.1 Templates creation
    4.3.2 Templates matching
  4.4 Classification
    4.4.1 Features
    4.4.2 Classifier
  4.5 Assessment of the results
    4.5.1 Validation scheme
    4.5.2 Inspection of results
    4.5.3 Comparison with other techniques
  4.6 Results
List of Figures

1 Classification of human-machine interfaces ......................... 10
2 Graphical representation of the requirements of devices and of the performance of separate subclasses of human-machine interfaces ................................ 17
3 Graphical representation of the requirements of space applications and of the performance of the four main classes of human-machine interfaces ................................. 18
4 Lateral surface of left cerebral hemisphere, viewed from the side ........................................................................ 28
5 Transverse section of human peripheral nerve showing the main structures of the nerves ................................................. 29
6 Cross section showing different types of sensory mechanoreceptors ................................................................. 30
7 Picture and unfolded overview of tfLIFE ............................. 35
8 Comparison between different neural interfaces in terms of selectivity and invasiveness .................................................. 36
9 One minute of recordings while some stimuli were applied to the paw of the rabbit ................................................... 46
10 One of the spike templates typically found by the spike sorting algorithm .............................................................. 47
11 DWT-based denoising of the same signal performed with time shifts from 0 to 7 samples ............................................ 49
12 Analysis of how different “spike waveforms” could aid in the detection of specific stimuli ........................................... 58
<table>
<thead>
<tr>
<th>Page</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Probability density functions and cumulative distribution functions of the inter-spike interval (ISI) for the three groups of templates during the different stimuli</td>
</tr>
<tr>
<td>14</td>
<td>Data flow within the data acquisition software</td>
</tr>
<tr>
<td>15</td>
<td>Screenshot of the acquisition software</td>
</tr>
<tr>
<td>16</td>
<td>Screenshot of QtProtocol</td>
</tr>
<tr>
<td>17</td>
<td>Screenshot of QtProtocol showing its scripting capabilities</td>
</tr>
<tr>
<td>18</td>
<td>Tiled view of the frames of one of the movies composed in order to instruct the subject to perform the “close hand” movement</td>
</tr>
<tr>
<td>19</td>
<td>Screenshot of Blender showing the <em>Emo</em> hand in the <em>octahedron</em> view</td>
</tr>
<tr>
<td>20</td>
<td>The figure shows how the <em>pinch grip</em> is obtained using 4 key-frames</td>
</tr>
</tbody>
</table>
### List of Tables

1. Stimuli applied to the hind paw of the rabbits .......................... 43
2. Performance in terms of percentage of correct classifications and channel capacity ........................................ 56
3. Results of fitting the table of percentage of correct classifications with a logistic regression .......................... 57
Acknowledgements

Part of the content of this dissertation is based on work authored by me and co-authors and published in journal papers or conference proceedings. Specifically:

1. Chapter 2 is partly based on the papers [TMC+08] and [CTM09].
2. Chapter 3 is partly based on the papers [MNC+08] and [CCY+08], and an unpublished work by Citi et al.
3. Chapter 4 is partly based on the papers [CCY+08] and [CCY+06].
4. Chapter 5 is partly based on unpublished work by Citi et al.
Vita

May 13 1979
Born, Pistoia, Italy

July 2004
Degree in Electronic Engineering
Final mark: 110/110 cum laude
Università degli Studi di Firenze, Firenze, Italy

Dec 2004 – Mar 2005
Research Officer
Dipartimento di Sistemi e Informatica
Università degli Studi di Firenze, Firenze, Italy

Mar 2005 – Apr 2009
PhD Candidate in Biortotic Science and Engineering
Jointly offered by IMT Institute for Advanced Studies Lucca
and Scuola Superiore Sant’Anna, Pisa, Italy

Jul 2007 – Aug 2007
Visiting PhD Student
Department of Anesthesia
Massachusetts Institute of Technology, Cambridge, USA

Jul 2008 – Aug 2008
Visiting Scientist
Department of Anesthesia
Massachusetts General Hospital
and Massachusetts Institute of Technology, Cambridge, USA

Jun 2008 –
Research Officer
School of Computer Science and Electronic Engineering
University of Essex, Colchester, UK
Related Publications

Referred journal papers


Referred conference papers


Others
Abstract

The restoration of sensorimotor functions for the control of artificial hands is a fundamental point in order to improve the quality of life of amputees. Current hand prostheses use electromyographic (EMG) signals, but are limited to a small number of channels and rely on visual feedback.

As the interface is the bottleneck for a wide class of hybrid bionic systems, we have developed a general framework to match task requirements with interface performance. We have carefully and thoroughly examined the case of the amputee user and its requirements: throughput and latency, but also user-friendliness, invasiveness, bi-directionality, and possibility of natural control of the prosthesis. From this analysis, we have concluded that, albeit suboptimal from a mere throughput and latency point of view, peripheral invasive interface can represent a promising medium-term solution.

We have compared the different types of interfaces with the peripheral nervous system, finding in longitudinal intrafascicular interfaces (LIFEs) a tradeoff between invasiveness and selectivity.

In order to assess the possibility of extracting complex information from LIFEs, we have run preliminary experiments with small animal models recording induced afferent information. Using the sophisticated signal processing techniques developed (wavelet denoising and spike sorting) and a robust classifier, we were able to discriminate four (or five) different classes of stimuli with performance in a range between 90% and 99%. These results confirmed and outperformed prior work carried out with different approaches.

There are plans to validate the approach with a human amputee. Hence, several steps have been taken in order to make possible the recording of neural signals from a human subject and allow online processing and control of the “Cyberhand” smart robotic hand prosthesis.
Chapter 1

Introduction

The user acceptance of last generation robotic limbs by amputees could be significantly increased by a system providing accurate, graded, distally referred touch and proprioceptive sensations. Furthermore, the functionality of such prostheses would be improved by a more natural control mechanism allowing the robotic limb to be felt as if it were a natural part of the body.

Considerable technologic progress has recently been made for developing advanced cybernetic hands. Underactuated robotic hands allow fine complex movements without the need for independently control each single joint. Highly sensored prototypes can provide the user with proprioceptive and sensory feedback, and can also implement on-board control loops that are able to perform stable holding of objects and deal with slip avoidance automatically.

The Cyberhand [CCM+06] developed at the ARTS Lab of Scuola Superiore Sant’Anna is an example of an underactuated sensored smart robotic hand. It represents the first integrated attempt to address simultaneously the following main objectives: to obtain a cosmetic and dexterous artificial hand; to develop a bio-inspired artificial sensory system; to allow a control which mimics as much as possible the natural control of biologic hands; and to delivery sensory feedback (touch, proprioception) to the subject.

The weak link in the hybrid bionic system, composed by the human subject and the robotic artifact, is the interface. In fact, a key component in these systems is a fast, intuitive, bidirectional interface between the controller of the device and the neural system of the user. Current
hand prostheses use electromyographic (EMG) signals, but are limited to a small number of channels and rely on visual feedback. For this reason, to achieve sensory feedback and a higher number of control channels, a different type of interface is required.

As the interface is the bottleneck for a wide class of hybrid bionic systems, we developed a general framework to match task requirements with interface performance. First, devices for domotics, rehabilitation and assistive robotics, and their requirements, in terms of throughput and latency, were identified. Second, human-machine interfaces were classified and their performance described, still in terms of throughput and latency. Then, device requirements were matched with performance of available interfaces. For the specific case of the hand prosthesis, EMG-based interfaces and central invasive interfaces partially meet the requirements in terms of throughput and latency. Nevertheless, significant issues arise if we consider other important factors, such as the unidirectionality and the use on non-homologous control strategies for the former, and a still limited biocompatibility for the latter. On the contrary, interfaces with the peripheral nervous system (PNS) which are suboptimal from a mere throughput and latency point of view, present many characteristics which enable them to represent a promising medium-term solution.

The neurophysiological bases for the use PNS neural interfaces and an overview of the different types of electrodes are presented in this thesis. The ideal properties of an electrode would be low invasiveness and high selectivity, but these are not attainable at the same time. For smart sensored hand prostheses, Longitudinal Intra-Fascicular Electrodes (LIFEs) could be a nice trade-off as a result of a reduced (although not absent) invasiveness and a good selectivity. These considerations were supported by previous experiments in literature, especially works by Dhillon and Horch, and by Jia and colleagues, which are briefly reviewed.

To assess the possibility of extracting complex information from LIFEs electrodes, preliminary experiments with animal models were conducted. These first experiments were performed on small animals and, therefore, induced afferent information was used instead of volitional efferent commands which are obtainable during experiments with primates. The assumption is that if an advanced signal processing technique gives access to several different types of sensory information, there are good possibilities that it can also decode volitional motor commands which can be used to control a robotic hand.

Thin-film longitudinal intra-fascicular electrodes (tfLIFE) were implanted in the sciatic nerve of rabbits. Various sensory stimuli were ap-
plied to the hind limb of the animal and the elicited electroneurographic (ENG) signals were recorded. We processed the signals to determine whether the different types of information could 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. We compared this novel approach 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.

In order to validate the approach with afferent signals recorded from a human subject and try to let him actually control a robotic hand, the software described above was re-developed in order to work online. We developed a new approach to wavelet denoising based on the “overlap and save” method, leading to a significant speedup with respect to the standard procedure. The pieces of software performing wavelet denoising, spike sorting and classification were wrapped in LabView libraries in order to be callable from the acquisition software. The acquisition software, developed in the LabView language, is able to perform acquisition from the A/D card, visualization, preprocessing, spike sorting, classification, and transmission of the command to the robotic hand. We also developed a software for the timing of the training sessions in order to show the different grasps to the user and, at the same time, mark the recordings accordingly. This is essential for the training of the algorithm. The grasps shown to the user were in form of movies created using rendering software by shaping a hand model in order to perform basic movements and more complex grasps.

Finally, the achievements of this doctoral work are discussed and future works envisaged.

This thesis is organized as follows. Chapter 2 describes the general framework used to assess which interface is most suited to a given use scenario and then describes why we believe that interfaces with the PNS can represent a promising medium-term choice. Chapter 3 highlights some neurophysiological bases for the use of PNS interfaces, then gives a review of the main types of electrodes and describes why LIFEs were chosen in this work, also with the support of previous experiments by other groups. Chapter 4 describes the main scientific contribution of this doctoral work, which is a novel approach for the extraction of informa-
tion from neural signals; the results of the application of this method to recordings collected from anaesthetized animal models are also presented. Chapter 5 presents some additional work aimed at porting and adapting the approach above, in order to allow online use during a real experiment with a human amputee subject. Chapter 6 summarizes the work done and the scientific achievements of this doctoral work, discusses some open issues, and envisages future developments.
Chapter 2

Taxonomy of interfaces

The goal of this chapter is to introduce different types of Human-Machine Interfaces (HMIs), both traditional and neural interfaces. Then we, depict some representative use scenarios, and try to suggest which interfaces are suited for each scenario by comparing the requirements of a given application with the performance of a given interface. The spectrum of scenarios considered will be fairly broad in order to see whether and how neural interfaces can be integrated in our lives. Later the focus will be restricted to the case of an upper limb amputee.

2.1 Hybrid Bionic Systems

2.1.1 Overview

In everyday life, we increasingly interact with machines, such as computer, appliances, even robots. This interaction is mediated by a human-machine interface. The ensemble user-interface-device, comprising both artificial and biological components, is defined as Hybrid Bionic System (HBS).

From a control system viewpoint, we can follow the information flow that happens as we interact with a traditional HMI. Our intention to interact with the interface for a utilization task, e.g. grasp a knob, resides in dedicated neural networks within the brain and is translated into complex motor commands and then dispatched from the areas for motor planning and execution toward the target muscles through the cortico-
spinal and peripheral nervous fibres. The results of our action are then
gathered by our sensing system (eyes, touch and proprioceptive recep-
tors, etc.), translated into sensory signals and fed back to the Central Ner-
vous System (CNS) through the afferent pathways.

This scenario is over-simplified, but nonetheless it allows to clarify
the potentials of direct neural communication. A neural interface can
be defined as any system able to monitor neural activity and translate
a person’s intentions into commands to a device. In an ideal neural in-
terfaced HBS, the motor commands, instead of being sent to the physio-
logical musculo-skeletal effectors, will reach an artificial actuator (e.g., a
robotic limb); its action on the environment will be measured by a sensing
system composed of artificial sensors and the information gathered will
be fed back to the CNS as natural afferent signals.

The artificial system can be hooked in at various levels in the system
above, e.g. intercepting electromyographic signals, recording spike ac-
tivity at the level of the Peripheral Nervous System (PNS), or directly –
invasively or non-invasively – form the CNS. Similarly, the feedback can
be provided using the natural senses (e.g., through visual or vibrotactile
stimulation) or directly injected in the nervous system through stimula-
tion of afferent pathways.

2.1.2 Types of users

Several potential users can benefit from neural interfaces.

Central interfaces, often referred to as Brain-Computer Interfaces
(BCIs) or Brain-Machine Interfaces (BMIs), have so far been extensively
studied as a communication means for people that are affected by dis-
abilities – such as severely advanced stages of amyotrophic lateral scler-
osis (ALS), muscular distrophies, or brainstem lesions – which prevent
them to voluntarily control the muscles \textsuperscript{[WBM+02; Don02; MIM03]}. In
this case, despite having a normally working brain in terms of cognition
and self perception, they possess no communication means with the out-
side world and an interface with the CNS may represent their only way to
interact with other people and objects. For these cases, also interfaces of
modest efficiency will represent a significant improvement in daily living
abilities.

Different is the case of an amputee user. In this case, the peripheral
nerve is intact up to a few centimetres proximally to the amputation.
Therefore, in addition to central interfaces, they could also benefit from
the use of interfaces with the peripheral nervous system. These can be
more selective than non-invasive central interfaces, and their invasive-
ess can be more easily accepted than that of central invasive interfaces.

On the other end of the spectrum are able-bodied users. For these
users, an interface with the state-of-the-art performance conceived as an
alternative communication device is not useful. For them, a neural inter-
face would only be practical if conceived as augmenting, i.e. an interface
that allows users to perform actions in addition to what they already can
do with their normal abilities.

2.1.3 Performance measures of HBS

The performance of HBSs can be characterized by means of several pa-
rameters. Throughput and latency were chosen as initial measures for
determining whether a given interface and device are suitable to be inte-
grated in a HBS. Among the numerous factors that can be pinpointed,
they are probably the only ones easily quantifiable and comparable.
Therefore they seem a reasonable choice in order to perform a first se-
lection allowing to individuate which combinations of interfaces and de-
vices are in principle possible and which ones are surely not. Other im-
portant factors, need further to be considered for the final design of the
HBS, such as degree of invasiveness, user-friendliness, portability, set-up
time, need for training, cost/effectiveness balance, robustness to noise,
instantaneous and cumulative cognitive load required, temporal stabili-
ity, etc.

Throughput and latency will be described and then used to charac-
terize both the interface and the device components of a HBS, under the
hypothesis that performance of HMIs can be roughly compared indepen-
dently from task and method and across all types of users.

Throughput (also called bit rate, bandwidth, or information transfer
rate) is the rate at which a communicating entity sends or receives data,
i.e. the amount of data that is transferred over a period of time and is
measured in bit/s. It is therefore a measure of the channel capacity of
a communications link. Error probability has an influence on through-
put: error correction slows down the system and wastes communication
bandwidth. Latency is a time delay between the moment something is ini-
tiated, and the moment one of its effects begins (onset latency) or reaches
the azimuth/nadir (peak latency). The unit of latency is time (s).

In the following, classes of interfaces and devices are characterized.
For each class, a range for both throughput and latency is defined.
The throughput of devices \((TP_d)\) was calculated as the product of the number of bits per unit command \(b\) (in bit/command) and the number of commands per second \(\nu\) (command/s) that have to be sent to the device to be able to control it interactively.

\[
TP_d = b\nu .
\]

The throughput of interfaces \((TP_i)\) has been calculated as the Shannon information rate \([Sha48]\). This definition of throughput is also popular in the literature on brain-computer interfaces, having been first suggested by \([WRMP98]\). In most papers \(TP_i\) is not reported but the number of symbols, the error probability and the transfer rate (symbols/s) is stated or can be inferred. In these cases a symmetric \(N\)-symbol channel with symbol rate \(R\) and error probability \((1 - P)\) is hypothesized. Therefore the throughput \(TP_i\) (in bit/s) has been calculated as:

\[
TP_i = R \left( \lg_2 N + P \lg_2 P + (1 - P) \lg_2 \left( \frac{1 - P}{N - 1} \right) \right)
\]

There are other definitions of throughput, such as the Blahut-Arimoto (used in \([SKY+05]\)) and Nykopp (discussed in \([KVP05]\)). However the Shannon definition was chosen because it can be easily calculated also in studies where it is not reported and provides an acceptable measure for our needs, as will be clear in the following.

The value of latency for the device, depends on how interactive the system is intended to be, on how much feedback is needed to close the control loop. It is often difficult to say which is the biggest acceptable latency for communicating with the user, so that they still feel interacting with the device and not frustrated by the long waiting time or by byproducts of delayed control, such as the risk of overshoot and instability.

The value of latency for the interface, is usually reported or deducible from the description of the experimental protocol used to generate the physiological signal measured by the interface. The minimum value of latency is limited by physiological characteristics of the neural fibres. Latency is also bound by the time resolution of the technique used to measure the user’s intent or action.

2.2 Devices

An increasing number of electronic appliances have become more and more common in everyday environment, also thanks to portable equip-
ment. In the field of domotics, also called home automation, HMIs are used to control devices and appliances. Domotics can contribute to a better quality of life, and can be useful for disabled and elderly people to increase their independence and autonomy. Domotics can be applied to safety and remote surveillance, to the control of doors, windows, lights, indoor climate, multimedia and communication devices, and household appliances. The concept of environmental control is interesting also for other habitats, like an office, a car or outdoor environments.

In recent years, even robots have become more common in non-industrial environments. These robots are often called human-centred or human-friendly systems because the presence of the robot involves a close interaction between the robotic manipulation system and human beings. The most important applications of human-centred robots is rehabilitation and assistive robotics. Rehabilitation robots are in contact with the users and safety is a primary concern [Tej00]. Advances in rehabilitation robotics are required by the growth of elderly population and by injured people, to assist in rehabilitation procedures and to provide new functional prostheses and orthoses. Rehabilitation robots are important for the healing of neurological diseases [KVAH00]. Neural Prosthetics, i.e. movement restoration for people with motor disabilities, is indeed another key application for BMI technology. Cortical signals have been used to control a hand orthosis [PGM00], with the aim to restore the connection from the brain to a paralysed arm. A locked-in subject has also used neural signals to control a virtual hand [KBM00] in the hopes that simulation would provide clues to potentially incorporating functional electrical stimulation into a BMI system to restore movement. Rehabilitation devices have hereby been grouped into several categories, such as feeder robots, prosthetic hands (basic grasping function), wheelchairs. Two or more of these devices can also be combined in complex systems. Humanoid assistive robotics deals with robots for domestic assistance, patient care, and even human augmentation. They are more complex than rehabilitation robots, have more DOFs and are supposed to be more reactive. Combinations of robot systems (hands, arms, trunk, etc.), up to a complete humanoid robot, have been considered.

Detailed throughput-latency requirements of the above mentioned classes of devices can be found in [TMC08].
Figure 1: Classification of human-machine interfaces. Examples of signal acquisition techniques and of acquired signals are listed for each class.

2.3 Interfaces

At a first level (see Figure 1), interfaces can be divided into cortical and non-cortical. Second, they can be divided into invasive and non-invasive, by considering as invasive those interfaces that need skin incision.

2.3.1 Cortical interfaces

Cortical interfaces are all interfaces that exploit information collected from the human brain cortical relays, by various means, invasively or non-invasively.

Cortical non-invasive interfaces

Cortical non-invasive (C-NI) interfaces (often called BCIs in literature) can measure and correctly classify specific signals of brain activity intentionally and automatically produced by the subject and translate them into device control signals. Such signals are recorded from the scalp and suffer from the limitations of their transit through the extracerebral layers (severe amplitude reduction, filtering of frequencies particularly in
the high-frequency range, spreading of the generator source identification, increased contamination of the signal from the generator(s) by far-field volumetric potentials). Features commonly used in experimental studies derive from brain signals that include alterations of the electrical activity recorded through electroencephalography (EEG) such as mu or beta rhythms \cite{WMNF91}, event-related potentials (ERPs), including the P300 and N400 evoked potential and visual evoked potentials (VEPs), either transient (to individual, low-rate stimuli) or steady-state (to prolonged trains of high-rate, repetitive stimuli) \cite{FD88, Sut92, MMCJ00, KLRF05}, transient variations of the background rhythms, i.e. event-related (de)synchronization (ERS/ERD) \cite{PFK93}, slow cortical potentials (SCP) \cite{BGH99}, and activation patterns induced by mental task strategies \cite{CS03, KP00}. To avoid the need of skin preparation and electrolytic gels, dry recording electrodes are being studied \cite{Mas05}. Today’s wet electrodes are not suitable for daily use in normal living environment; dry electrodes would guarantee a good electrode/skin contact and allow acceptable signal-to-noise ratio for longer session times.

Other features recorded with different modalities include neuro-magnetic signals recorded through magnetoencephalography (MEG) \cite{TRP97, GLLM05}, blood oxygen level-dependent (BOLD) responses recorded through functional magnetic resonance imaging (fMRI) \cite{WMB04} and localized activity-related brain oxygenation measures recorded through near infrared spectroscopy (NIRS) \cite{CWMM04a}.

Current cortical non-invasive HMIIs are uni-directional interfaces. Brain signals can be used to drive a machine. Stimulating the CNS by means of non-invasive technologies, such as transcranial magnetic stimulation is not selective \cite{RBB94, RR07}. Therefore natural afferent pathways are used for communication feedback.

The extracted features must be translated into commands directed to the artificial device. Usually Artificial Intelligence methods are used. The training of the system – algorithm and human subject – should be performed at three subsequent stages: a) initial training of the algorithm with off-line recordings, b) subsequent tuning with periodic adjustments during on-line use, c) mutual adaptation and reinforcement of both the system and the subject (e.g., through bio-feedback).

Common BMI applications (mainly still at research stage) are, for example, moving a cursor on a computer screen \cite{WM04}, writing with a speller \cite{FD88}, controlling domotic appliances \cite{CAB06}, driving smart robots \cite{dRMRMG04a} or smart wheelchairs \cite{TMW05} and controlling the hand grasp through functional electrical stimulation \cite{MPSPR05}. The
control of a hand prosthesis has been seldom investigated \cite{GHHP99,ACC06,MPP08}.

Data collected and processed from the following studies have been considered to assess throughput and latency performance of cortical non-invasive interfaces:

- **ERP, ERD/ERS:** classification of mental states, related only to motor imagery \cite{KNL06}, or including also mental tasks (such as cube rotation or calculation) \cite{ONGP01,Nyk01,Leh02,dRMM03,dRMRMG04b,BLV08}, or imagination of sensory stimulation \cite{DBCM04}.

- **P300 evoked potentials:** selection of items in a sequence, such as four-choice paradigm, or arranged into square matrices, typically of size $6 \times 6$ \cite{DSW00,KR04,KMG04,SYT05, MKH03, TG05, SKM06, SKD06, BSCR08, HVED08, CPSC08, WGZ05}.

- **Slow Cortical Potentials:** 1-D cursor movement tasks \cite{BKG00, BMC04, KCBM07, FdNLD07, PHN05, KHR06}.

- **Sensorimotor cortex rhythms:** 1-D cursor movement tasks \cite{MSW03, FMPW04, BFdRM06, NFG08} and 2-D cursor movement tasks \cite{WM04, FMPW04, GGD06, VS06}.

- **Steady-State Visual Evoked Potentials:** 1-D cursor movement tasks \cite{MMC00}, 2-D control of a prosthetic hand \cite{MPP08}, and nominal selection of a variable number of targets \cite{KLF05, MPEWP08, CGGX02, WWG06}.

### Cortical invasive interfaces

Cortical invasive (C-I) interfaces are based on the voluntary control of the firing rate of individual neurons in the primary motor cortex. Neural signals recorded in cortical invasive interfaces range from small neuronal samples to large ensembles, including local field potentials (LFPs), spread over a single or multiple recording sites \cite{LN06}. Commonly used intra-cortical electrodes are microwires \cite{MA67}, multiple electrode arrays (MEAs) \cite{MNN97}, and neotrophic electrodes \cite{Ken89}.

In most of the cases, research groups have focused their efforts on the extraction of information from the primary motor cortex (M1) to drive a robotic arm. In fact, neurons in M1 arm area can provide
information about intended arm reaching trajectories \[\text{LN06}\]. Therefore, signals recorded from ensembles of M1 cortical neurons can be processed through different algorithms, such as neural networks or Kalman filters to predict arm end-effector trajectories usable to control a robotic system. This approach has been tested with very promising results in animal models \[\text{LN06}\] and recently in selected highly disabled subjects \[\text{HSF}^+\text{06}; \text{TFDH08}\]. Signals used in cortical invasive interfaces are usually generated by the subject through motor imagery tasks \[\text{LSW}^+\text{04}; \text{GHLP04}; \text{HSF}^+\text{06}\]. Also, interfaces exploiting LFPs generated by non-motor imagery (e.g. in the auditory cortex) have been investigated \[\text{WFG}^+\text{06}\].

An alternative, less invasive, recording modality is electrocorticography (ECoG) based on epidural or subdural implanted mesoelectrodes. In humans, many experiments exploit ECoG signals measured on epilepsy patients requiring invasive monitoring of cortical activity for localization and eventual resection of an epileptogenic focus \[\text{KB98}\].

Cortical invasive interfaces have the potential to deliver a sensory feedback \[\text{RHZ}^+\text{00}\]. However, most studies currently use them only for recording neuronal activity, relying on visual stimuli for feedback, because of side effects that must that be considered. In fact, when the cortex is stimulated in order to elicit sensory feedback, two main problems could occur: (a) undesired muscle contraction can be induced (see for example \[\text{WC97}\]) and (b) the re-organization of the somatosensory cortex \[\text{ESF}^+\text{97}\] can provoke misleading sensations.

Data collected and processed from the following studies with human subjects have been considered to assess throughput and latency performance of cortical invasive interfaces:

- 1-D cursor movement: \[\text{KB98}; \text{KBM}^+\text{00}; \text{LSW}^+\text{04}; \text{WFG}^+\text{06}\]
- 2-D cursor movement: \[\text{HSF}^+\text{06}; \text{TFDH08}\]
- Nominal selection of up to 4 mental states \[\text{GHS}^+\text{03}; \text{GHLP04}\]

Research on the use of cortical invasive interfaces as BMI has started on animals over three decades ago \[\text{Fet69}; \text{HST70}\]. Indeed, being more invasive than human studies, animal experiments have shown higher performance. Data from the following animal studies, on rats, cats and monkeys, have been included, in order to demonstrate the potential of invasive technology:

- Switches: \[\text{CMMN99}; \text{LWN00}\]
• 2-D cursor movement: \([\text{SHP}^{+02} \text{ SRY}^{+05}]\)
• 3-D movement of cursor and robot arm: \([\text{WSK}^{+00} \text{ TTS}^{02} \text{ TTS}^{03} \text{ CLC}^{+03} \text{ ATA}^{+08} \text{ AAS}^{+07}]\)

Besides MEA, also LFP have been exploited \([\text{BPAM}^{06}]\).

### 2.3.2 Non-cortical interfaces

Non-cortical interfaces are all interfaces that do not access the signals generated by the human cortex directly. The signals that drive the interface are measured in the peripheral nervous system, on the muscles, or are the result of muscular activity (change of body posture or physical interaction of the body with the interface).

#### Non-cortical non-invasive interfaces

Non-cortical non-invasive (NC-NI) interfaces, sometimes referred as Human-Computer Interfaces (HCIs), are operator interfaces terminals with which users interact in order to control other devices. The interaction can include touch, sight, sound or any other physical or cognitive function. HCIs have been divided into classes, according to the method used to detect the control command sent to the machine.

The Switch1 class includes contact switches, i.e. keyboards, touch screens, joysticks, buttons, etc. The Switch2 class includes non-contact switches, e.g. eye blinking systems, detecting user’s eye blink and using sequences of long and short blinks interpreted as semiotic messages \([\text{GBGB}^{01}]\), and a camera based finger counter \([\text{CB}^{02}]\). The Pointer class includes mice, laser pointers \([\text{OS}^{02}]\) and interfaces built on gaze trackers \([\text{Jac}^{90} \text{ SJ}^{00} \text{ CM}^{96} \text{ DLKT}^{03} \text{ XHP}^{+05}]\). The Speech class consists of dictation software and a small vocabulary automatic speech recognition system \([\text{UB}^{05} \text{ AGG}^{+05}]\). In addition to the references above, to compute the values of throughput and latency, comparative studies of common HCIs \([\text{Fit}^{54} \text{ CEB}^{78} \text{ PS}^{92} \text{ Hyr}^{97} \text{ MKS}^{01} \text{ OS}^{02} \text{ MS}^{03}]\) and the ISO Standard 9241-9 \([\text{ISO}^{00}]\) were used and some data was inferred from average speeds of touch-typists (for keyboards) \([\text{SC}^{92} \text{ KK}^{05}]\) and telegraphers (for a single switch). A few of these interfaces, such as those based on gaze movements or eye blinking, can be used by also by severely disabled people as an alternative to BMIs.

Moreover, electromyographic (EMG) interfaces were considered. Some currently available rehabilitation devices, such as hand prostheses,
exploit EMG signals recorded via surface electrodes [MKC07; SMCR08]. By activating specific muscles not necessarily related to the desired task, the user can select different predefined grasping patterns [ZCS04; CE05; EH03; AW05]. This approach offers advantages as robustness, simplicity to implement, and non-invasiveness but also presents several limitations, such as the need to use the non-homologous muscles to control the movements of the prosthetic device. Performance for EMG-based interfaces has been calculated from the above references.

**Non-cortical invasive interfaces**

The unidirectionality of EMG-based interfaces is the rationale of the recent attempts to directly connect the PNS with the artificial device by using non-cortical invasive (NC-I) interfaces, i.e. invasive intra-neural interfaces. Invasiveness is obviously considered a drawback and is acceptable only if it can lead to significant and long-lasting improvements in terms of reliability, selectivity, stability of the implant.

Low invasiveness and high selectivity are not attainable at the same time. Less invasive extraneural electrodes, such as cuff and epineural electrodes, have reduced selectivity. More invasive intra-fascicular electrodes, such as longitudinal intra-fascicular electrodes (LIFEs), MEAs and regenerative electrodes are more selective and they allow interaction with small groups of axons within a nerve fascicle. A review and comparison of different peripheral nerve interfaces can be found in Section 3.2 and in [NKL05].

Few experiments on non-cortical invasive interfaces have been published. Apart from the above mentioned papers, some additional numerical data have been calculated from [CCY06; CCY08; BSRD07].

### 2.3.3 Summary of interface performance

Figure 2 shows a summary of latency and throughput for all classes of HMIs. These values are not meant to be an exhaustive coverage of all available interfaces, nor to reproduce the performance of the interfaces in a quantitatively rigorous way. Each type of interfaces is represented by the convex hull containing all the points corresponding to the experimental data cited above. It is also worth noting that this is a simplified graphical representation. The main problem is that it depicts very well

---

1The convex hull is the smallest convex polygon containing a given set of points
the range of the values but does not give much information about how the data is distributed within this range. For most sessions and subjects, anyway, the typical performance lies toward the centre of the polygon.

MEG-based BCIs have recently shown performance comparable to EEG [KNL+06]. However MEG devices are expensive, immobile and extremely vulnerable to body-generated and urban magnetic noise, when operative outside magnetically shielded rooms. fMRI scanners are also expensive and immobile. fMRI-based BCIs, such as [YFC+04], suffer from poor temporal discrimination due to the haemoglobin relaxation time which produce BOLD effects. Conversely, NIRS-based BCIs, such as [CWMM04b] are inexpensive and portable. However they suffer from very low throughput (in the order of 0.01 bit/s). For all these reasons, BMIs based on MEG, fMRI and NIRS are not suitable to control a HBS (and especially a hand prosthesis) and are not included in this study.

2.4 Results

2.4.1 Throughput and latency

In this section the needs of the applications presented in Section 2.2 are matched with the performance of the interfaces described in Section 2.3. Identifying the areas of overlap allows to define realistically which applications can be driven by means of a given BMI and also which types of BMI are suitable for a given application. As said, this matching represents a necessary, but not sufficient, condition. Other key factors in the design of a HBS, will be considered later.

Figure 2 shows the overlap of application needs and interface performance (convex hulls). Figure 3 is similar, but the different HMIs are grouped according to invasiveness (invasive/non-invasive) and to the location of the hybrid link (cortical/non-cortical).

At a first glance, it can be pointed out that applications that require little throughput and tolerate higher latency could be driven by any of the interfaces considered. These applications comprise devices for environmental control (domotics), feeders, and humanoid heads. This is not surprising: indeed, the control panel of domotic applications is usually a simple interface composed of switches and sliders, controls that are easily implemented by means of a BMI [GXCG03; CAB+06]. Feeder robots have even simpler interfaces: a trigger signal is needed to activate the robot, that then executes the feeding task autonomously without further feed-
Figure 2: Graphical representation, in terms of latency and throughput, of the requirements of devices (grey boxes) and of the performance of separate subclasses of human-machine interfaces (areas delimited by coloured convex hulls).
Figure 3: Graphical representation, in terms of latency and throughput, of the requirements of space applications (grey boxes) and of the performance of the four main classes of human-machine interfaces (areas delimited by coloured convex hulls).
back from the user. Even SCP-based low-throughput BMIs can be used to control feeder robots. However, feeder robots are more easily controlled by means of puff/sip switches, which only require breath control abilities. The humanoid head represents a two DOFs control, the steering of a camera (e.g. a robot-mounted camera). This application shares many aspects with interfaces allowing an impaired user to scroll the screen and reach icons and widgets on a computer desktop [CPSC08].

Concerning more complex rehabilitation applications, there is an overlap (or near overlap) for higher-performance interfaces with robotic devices that have few DOFs, such as a wheelchair. Sometimes shared control can be used to reduce the requested throughput and raise the maximum latency. In this case the human intervention is used to define the end point targets while the device is in charge of planning the trajectory and avoiding obstacles [dRMRMG04; TMW05].

Complex compound devices, namely the astronaut-equivalent arm and the whole astronaut-equivalent robot, require performance that is currently not attained by any of the interfaces.

EMG-based interfaces show some overlap with the robotic hand. In fact they are used in commercial hand prostheses, allowing basic grasp control of underactuated robotic hands.

Lower performance interfaces have been used to achieve sub-optimal control, e.g. basic binary (open/close) control grasping of a prosthetic hand [GHHP99; ACC+06] or a hand orthosis [PGM+00; KBM+00; MPSPR05].

Performance measured in monkeys suggest that cortical invasive interfaces could be used successfully for controlling prosthetic hands with greater interactivity. However, with cortical invasive interfaces, humans have not reached the same performance as monkeys. In [HSF+06], the quadriplegic human subject that received the 96-MEA, was able to control a computer cursor to interact with home appliances, operate the opening and closing of a prosthetic hand and perform rudimentary actions with a multi-jointed robot arm. An interesting note is that he could perform these actions even while conversing, which suggests that invasive interfaces have greater capabilities of discriminating shared output, i.e. simultaneous orders of different content.

Figure 3 shows no overlap between intra-neural PNS interfaces and robotic hands. The main reason is, probably, that the results reported here have been calculated from few preliminary works. Furthermore they are related to the use of only one intra-neural channel. The throughput should be significantly improved by the combined processing of several
contacts, e.g. involving more than one of the three nerves serving the hand sensorimotor control (median, ulnar and radial). It seems reasonable that the potential of PNS interfaces has not been exploited yet. The number of research groups focused on PNS interfaces, as well as the number of publications, is order of magnitudes smaller than, for example, the respective figures for EEG-based BMIs or invasive studies on monkeys.

2.4.2 Other measures

Performance measures reported above are mainly focused on information transfer capabilities (information throughput and ability to meet deadlines). A wider range of metrics includes general measures in terms of suitability of a device for a given task and class of target users.

A key parameter is the degree of invasiveness, for example because the risks related to surgical intervention inside the skull are not acceptable for augmentation devices or when a less invasive approach gives similar results.

A few additional measures are related to the user-friendliness of the interface, including comfort for the user, portability, easiness of use, set-up time and need for care holders intervention. Further parameters to be considered are the degree of bidirectional control (in terms of feedback), the training requirements, the cost/effectiveness balance, and robustness to noise.

Another key point is the instantaneous and cumulative cognitive load required. The instantaneous cognitive load of the interface can make it interfere with the task at hand while the interface should work as much transparently as possible. The cumulative cognitive load, instead, can reduce the temporal stability, i.e. for how long the user is able to drive the interface without degradation of performance, due to physical or cognitive tiring.

Also very important is the possibility of controlling the device in a natural way. This is especially important for robotic limbs which should be felt as a natural part of the body. It is desirable to use the homologous muscles to control the movements of the device.

2.5 HMIs for hand prostheses

From the previous section, it emerged that for the use of an advanced sensored robotic hand, none of the HMIs (excluding conventional inter-
faces) present a clear overlap in terms of throughput and latency. EMG interfaces seem to be the closest ones, followed by invasive interfaces and, finally, central non-invasive interfaces.

Nevertheless, from the last part of the previous section it was very clear that several other factors, besides throughput and latency, should be taken into account in order to choose the interface that is most suited for a given application. They have to be considered, discussed and weighted specifically for each interface and for each application and target user. Therefore in the following of this section, being the focus of this work on hand prostheses, the different neural interfaces will be reviewed discussing their usability for the control of hand prostheses in amputees.

**Cortical non-invasive HMIs**

Concerning the use of non-invasive cortical HMIs for the control of hand prostheses, they present the advantage of being non-invasive and low cost but also present some limitations. At the state of the art they are not very user-friendly (they are not very easy to operate, they have a non-neglectable per-session set-up time and there is need for careholders intervention) but this could be overcome once new electrode technologies are available. Other important factors to consider are that they are unidirectional, and definitely not robust to noise.

In order to let the user control the prosthesis in a natural way, the most suitable way would be to detect signals from the sensorimotor area of the brain. Ideally, the prosthesis should be moved by the brain activity alone generated by the hand controlling motor areas of the brain. This approach is currently not feasible because EEG do not allow recordings of the cortical activity distribution with such a high spatial resolution. It is more likely that non-invasive BMIs can be used to pass an easily-detectable low-bandwidth signal [GHHP99, KBM⁺00, MPSPR05, ACC⁺06, PGM⁺00] to an intelligent controller, capable of grasping objects with adaptable grip force through a sophisticated local control loop.

Finally, further investigation is required to rule out that the daily use of the these interfaces is made impossible by an excessive cognitive load or by interferences between the mental activity related to the BMI and the one related to the task at hand.
Cortical invasive HMIs

Cortical invasive interfaces present good performance but they also have some issues to take into account. First of all, they are invasive and they require the insertion of an artificial probe in the brain, the most important, delicate, and complex of our organs. Invasiveness is obviously considered a drawback and it is acceptable only if it can lead to significant improvements for the amputee.

Currently one of the main problems is the limited robustness of invasive CNS interfaces due to reasons such as the encapsulation with scar tissue around the recording area, the presence of proteins adsorbed onto electrode surface, and the micromovements between the brain and the interface damaging the nervous system and degrading the signal-to-noise ratio of the recorded signal [HA02; LBSL05; LUY+06].

Moreover, current strategies based on the extraction of joint trajectories from the primary motor cortex [Sch04] make the control of a multi-DoFs hand prostheses very difficult. In fact, it could be quite challenging to predict the independent or combined trajectories of individual fingers. This is a very important drawback of current cortical neuroprotheses even if some other approaches based on the extraction of grasping primitives by processing pre-motor signals could improve the situation [MCU+05].

Invasive cortical interfaces are in principle bi-directional. They can be used to deliver a sensory feedback [RHZ+00]. However, side effects induced by the direct stimulation of the primary sensory cortex must be considered. In fact, when the cortex area corresponding to the missing limb is stimulated in order to elicit sensory feedback, two main problems could occur: (a) undesired muscle contraction can be induced (see for example [WC97]) and (b) the re-organization of the somatosensory cortex [ESF+97] can provoke misleading sensations eventually sustaining a distressing “phantom limb” syndrome.

ECoG combines advantages over intra-cortical electrodes (no cortical invasiveness, reduced clinical risk, greater long-term stability) and EEG technology (larger amplitude of recordings, higher spatial resolution, reduced artifacts, less attenuation in the higher spectrum), while not incorporating many of their limitations [Mor03]. Nonetheless, ECoG is still an invasive technique requiring craniotomy and dural meningeal opening, which limits its use on specific clinical conditions.

In summary, while the central invasive approach can be promising for the future, open issues [MCB+06] and ethical aspects have to be investi-
gated before they can be considered suitable for amputees; such concerns cannot be overcome at the present.

Non-cortical non-invasive interfaces

EMG-based control offers advantages as robustness, simplicity to implement, and non-invasiveness but also presents several limitations. They allow control of a limited number of channels, and they are unidirectional. These limitations are becoming even more important with the development of more dexterous prosthetic arms with built-in sensors. Even more limiting, is the need for coding the different actions of the artificial hand using non-homologous muscles. For example, it is necessary to control the extension of the fingers of the prosthesis by using muscles of upper arm or forearm. EMG-based controlled prostheses suffer from the failure of making the arm feel a natural part of the body.

In the recent past some groups have tried to develop alternative methods useful to restore some similarity with the natural control scheme. For example, Kuiken and colleagues developed a new method based on the transferring of residual nerves of amputees to other muscles in or near the residual limb [KDL+05; KML+07b; ZLE+07]. The reinnervated muscles act as biological amplifiers of motor commands in the amputated nerves and the surface EMG can be used to enhance control of the robotic arm. Moreover, the delivery of sensory feedback by using external stimulation seems to be possible [KML+07a]. This approach has the interesting advantage that the nerve function correlates physiologically to the function it is controlling in the prosthesis and therefore operation is more natural and thus easier than current control paradigms. However, it requires a surgical intervention for nerve transplantation.

Non-cortical invasive interfaces

Peripheral nerve interfaces are invasive and require surgical intervention for the implant. Nevertheless, the risks during the implant and the possible future complications are less severe than with cortical interfaces. The degree of invasiveness depends on the type of probe used, ranging from low invasive extraneural (cuff and epineurial) electrodes to regenerative electrodes which require section of the nerve.

Once the electrodes are put in place and after a brief training by the user, their use should be transparent, requiring no set up time and no need for care holders intervention.
After a period of training, no cognitive load is required to operate the device in addition to what is required to operate a natural limb.

Among the different interfaces introduced here, peripheral nerve interfaces present the higher potential in terms of natural control of the prosthesis. In fact, it is in principle possible to record from the nerve fascicle that in the natural arm innervates a given muscle, to control the corresponding effector of the robotic hand. Fascicular groups destined to the same target remain localized within the nerve for some long distances, thus facilitating the selective interface of different fascicles within a given common nerve \[VGM93, BSN01, NVSS01\]. Therefore it seems advantageous to implant electrodes that can record from and stimulate individual, separate fascicles. Eventually, the use of interfaces with a large number of active electrode sites increases the chances of selective topographic stimulation or recording from more distinct nerve fascicles.

PNS interfaces are bi-directional and the stimulation of afferent fibres seem not to suffer from the same issues discussed above for cortical invasive interfaces. In fact, electrical stimulation of selective fascicles in the healthy portion of nerves in patients with chronic lesions or amputations has demonstrated that somatosensory localization remains accurate, despite the presumed central reorganization that takes place after nerve injury \[SBW+94, DLHH04\]. As large myelinated fibres are activated before small myelinated and unmyelinated ones, stimulation of afferent fibres can provide tactile or position sensations without concomitantly evoking pain sensations \[DKSH05\].

### 2.6 Conclusions

Considering all the factors discussed here, PNS interfaces can be considered a good medium term option for the control of robotic limbs.

In spite of the fact that their state-of-the-art performance in terms of throughput and latency is suboptimal for the control of a dexterous robotic hand, there is no theoretical reason for that and, probably, their potential has yet to be fully exploited.

Compared to central invasive interfaces, the invasiveness of PNS interfaces, and the related risks, can be acceptable for an amputee user.

Compared to EMG or EEG-based interfaces, peripheral nerve interfaces present the higher potential in terms of natural control of the prosthesis.
Also in terms of possibility to give feedback to the user, peripheral interfaces seem to be the best option available.
Chapter 3

Interfacing the peripheral nervous system

In this chapter, an overview of the neurophysiological bases for the use of PNS neural interfaces will be provided. Then, general considerations on interfacing the PNS will be discussed, before going through the different types of peripheral interfaces with particular attention to the pros and cons of each approach. Finally, limits and capabilities of longitudinal intra-fascicular electrodes (LIFEs) will be analyzed also by reviewing some past experiments recording and stimulating through LIFEs.

3.1 Neurophysiological bases for the use of PNS neural interfaces

3.1.1 Cortical Physiology

The primary motor cortex (MI), together with pre-motor areas, is responsible for the planning and execution of movements. MI contains neurons (Betz cells) whose long axons run down the spinal cord to form a synapsis with alpha motor neurons, which in turn connect to the muscles.

At a first approximation, the different body parts are represented somatotropically within MI. The arm and hand motor area is the largest, and occupies part of the anterior central gyrus.

Within the primary motor cortex, multiple yet discrete micro- and
macrozones exist. Several studies have shown that: (a) large, overlapping cortical areas converge onto single muscles; (b) outputs from any given cortical site diverge to multiple muscles; and (c) extensive horizontal interconnections between subregions of MI exist [HJ91; DLS92]. The overlap of contiguous joint representations within the motor cortex allows local interaction of output modules controlling different target muscles [KMMW78]. This, in turn, plays a key role in the coordination of actions needed to perform complex multi-joint movements.

The primary somatosensory cortex (SI) is located behind the primary motor cortex, configured to generally correspond with the arrangement of nearby motor cells related to specific body parts. A functional organization within the different areas of SI has been suggested. Area 3b would extract quantitative data relative to a stimulus in a given skin locus. Area 2 is thought to be more “associative”. Areas 1 and 2 would analyze more complex features (such as velocity and the presence of edges) of a stimulus involving different districts of the receptor surface [ITSH85]. Post-central somatosensory cortex is also a major source of sensory input to the motor cortex [JCH78].
Figure 5: Transverse section of human peripheral nerve showing the main structures of the nerves. A fascicle is a bundle of nerve fibres surrounded by a particularly specialized connective tissue, the perineurium. Within each fascicle, nerve fibres and blood capillaries are held together by a connective tissue made of a network of collagen fibrils, the endoneurium. In large nerves, the epineurium surrounds and holds together several nerve fascicles and larger blood vessels. From [Gra18] (public domain).

3.1.2 Peripheral Physiology

The somatic nervous system is the part of the peripheral nervous system responsible for the voluntary control of body movements, and the reception of external stimuli.

The majority of somatic peripheral nerves (Figure 5) are mixed nerves, conveying motor, sensory and autonomic information to and from the corresponding projection area.

Afferent sensory fibres carry different sensory impulses (e.g., touch, proprioception, thermal, painful stimuli). They can be myelinated or unmyelinated. They can originate at the periphery from specialized receptors in the skin, muscles and deep tissues (Figure 6). The receptor membrane transduces a tactile, proprioceptive, thermal or noxious stimulus into action potentials running through the axon. The intensity of the stimulus is mainly coded as impulse frequency. Each sensory neuron su-
Figure 6: Cross section showing different types of sensory mechanoreceptors. The intensity of the stimulus is mainly coded as impulse frequency. Some fibres primarily encode the intensity of the stimulus while others are more sensitive to the rate of change in stimulus intensity. Adapted from [Pha] (Creative Commons license).

A small bundle of fibres is called a funiculus; the funiculi are collected together into nerve fascicles that eventually give origin to branches innervating distinct targets [HFC98]. At the fascicular level, peripheral nerves are organized somatotopically and functionally: fascicular groups destined to the same target remain localized within the nerve for some long distances.
3.1.3 General considerations for PNS interfaces

Most of the peripheral nerve interfaces use an electric (as opposed to magnetic or optic) coupling to record or stimulate the nerve fibres. Therefore, in order to reduce the impedance of the coupling, electrodes are implanted as close as possible to the site: mostly around, adjacent or even within the peripheral nerve. To reduce the degree of invasiveness into the biological system, considerable attention of the research community is devoted to the development of more bio-compatible materials and structures. The goal is to allow chronic selective recording and stimulation without damaging the biological tissue.

Selective electrical interfaces are designed to contact with nerve fibres as selectively as possible. When interfacing a mixed electrode two types of selectivity must be taken into consideration: the topographic selectivity, which relates to the anatomic region (skin or muscle) served by the fibres in the fascicle; the functional selectivity, which refers to the types of functions carried by the different nerve fibres within the fascicle [PMH03].

Control signals for a certain movement of the prosthetic limb could be recorded from motor fibres that, in the natural limb, innervate the muscle(s) activated to actually perform the movement. The overall activity recorded from the corresponding motor nerve fascicle can give information about the desired movement and its intensity. Regarding functional selectivity during recording, the amplitude of action potentials recorded extracellularly, decreases with decreasing axon diameter (given the same distance between the fibre and the electrode). Therefore, detection above background noise is easier for impulses propagated in large alpha motor and mechanoreceptive sensory fibres than in small sensory and sympathetic fibres.

For the sensory feedback, stimulation of a small subset of sensory units should be able to evoke sensations that are both topographically circumscribed and of the desired modality. The somatotopical organization of the PNS favours the former but the latter could be more difficult to attain since sensory fibres of different modalities subserving the same anatomical area run together in the same fascicles [MCO90]. Fortunately, during stimulation, large myelinated fibres are activated before small myelinated and unmyelinated fibres (given the same distance between the fibre and the electrode) and this is advantageous since tactile or position sensations can be elicited without simultaneously evoking discomforting pain sensations [DKSH05]. However, smaller fibres near the
3.2 Overview of PNS interfaces

These neurophysiological considerations seem to confirm the interesting opportunities connected to the use of PNS neural interfaces. However, in order to select the most adequate interface for the control of advanced sensored prostheses, many different aspects of the various solutions need to be considered. A brief overview of PNS interfaces is given here, for a detailed review, see [NKL+05].

3.2.1 Cuff electrodes

Cuff electrodes have been used for many years to test FES systems in animal models and even in human trials. Therefore, they are currently considered as the most short/medium-term viable technology for chronic peripheral nerve interfaces. They wrap the nerve and record the combined activity of all the axons enclosed within it. Cuff electrodes allow the recording of single nerve action potentials from small fascicles, but they are unable to differentiate the different fibres. Therefore, they represent an implanted neural interface with relatively low invasiveness, good sensitivity but also low selectivity.

Multisite cuff electrodes [TM04; TD07], interfascicular cuff electrodes [TD97], innovative cuff structures [LD03; LGD06], and advanced processing algorithms [MJS+01; CMD+03; TD06] have increased the selectivity of cuffs. Despite these progresses, they do not provide a significant improvement over surface EMG approaches. In fact, a smart sensored robotic hand prostheses may require an interface with higher selectivity.

Albeit less invasive than other electrodes, compression exerted by the cuff in some instances may induce loss of nerve fibres [LTHS98]. Spiral cuff electrodes, composed of a thin and flexible polyimide insulating carrier and different configurations of platinum electrodes, proved to be a stable stimulation/recording device, whose physical properties avoid nerve compression or activity-induced axonal damage [RCS+00].
3.2.2 Sieve electrodes

Regeneration-type (or sieve) electrodes consist of a thin insulating plane with a grid of metallized via holes that is implanted between the two ends of a sectioned peripheral nerve. Regenerating axons eventually grow through the via holes, making it possible the recording of action potentials and the stimulation of individual axons or small groups. Regenerative electrodes can only be applied after transecting the peripheral nerve and encouraging the fibres in the transected nerve stump to regenerate axons through the interface.

Even if promising results have been achieved in experimental models [NCR98, CVCV02], some challenges remain, limiting their clinical usability [LCR05]. With silicon-based chips, only a low proportion of the proximal axons cross the via holes and regenerated fibres often show morphological abnormalities due to compressive axonopathy [NCB96].

Micromachined polyimide sieve electrodes [SNC96] have better biocompatibility [NCR98] than silicon dice, by providing a larger total open area and higher flexibility of the electrode. These electrodes allowed recording of compound action potentials and of bursts of single action potentials in response to sensory stimuli, and also selective stimulation of small motor nerve fascicles [CVCV02].

Nevertheless, also polyimide sieve electrodes present some open issues [LCR05]. The speed of regeneration is higher for small than for large nerve fibres, so that large myelinated sensory and motor axons are underrepresented within the holes of the electrode. Furthermore, regenerating axons grow at random and the regenerated nerves loss their normal topographical architecture, making selective interfacing of distinct nerve fascicles harder. Finally, in chronic animal experiments, there still appeared signs of compressive axonopathy at the sieve electrode level.

3.2.3 Multielectrodes arrays

Multielectrodes arrays (MEAs) [NRR94] are arrays of hundreds of penetrating silicon, glass or polyimide microelectrodes designed to selectively record or electrically stimulate neurons. Developed as an interface for the CNS, they have also been used in peripheral nerves and have been shown to be a selective recording device with single unit resolution and a low-current high-selectivity stimulation interface. They have been used in experimental works with animal models [BN00, MCN04], and also in a human volunteer [WGH03]. Unfortunately, they are highly invasive.
In fact they present high risk of nerve damage due to their rigid structure, the high electrode density and the invasiveness of the implantation procedure.

### 3.2.4 Longitudinal intra-fascicular electrodes

Longitudinal Intra-Fascicular Electrodes (LIFEs) are another kind of intraneural electrodes constructed from thin insulated conducting wires (such as Pt-Ir or metallized Kevlar fibres). They are inserted longitudinally into the nerve tissue and are designed to lay in-between and parallel to the nerve fibres.

Gathering signals from only a small number of axons, they allow more selective recordings than, e.g., cuff electrodes. Also, LIFE electrodes are less invasive than MEAs that are inserted transversely into the nerve leading to a higher risk of nerve damage. LIFEs have been under development as a neuroprosthetic device since the late 1980’s and has recently been implanted semi-chronically in amputees. Recently, a new version of LIFEs based on the use of a polyimide substrate (named thin film LIFEs (tfLIFEs), Figure 7) have been developed exploiting microfabrication techniques and making them amenable to mass production. These electrodes were developed on a micropatterned polyimide substrate which was chosen because of its biocompatibility, flexibility and structural properties. After microfabrication, this substrate filament 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 for implanting the electrode and is removed immediately after insertion.

During chronic implants in rat peripheral nerves, tfLIFEs induced only minimal damage and did not cause functional impairment. However, slight encapsulation by fibrous tissue around the LIFE was produced, a fact that can reduce the amplitude of recordable signals over time.

### 3.2.5 Considerations

In general, low invasive extraneural electrodes (such as cuff and epineurial) lack selectivity because they give simultaneous interface to
Figure 7: Picture and unfolded overview of tfLIFE [HK05]. 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).

As low invasiveness and high selectivity are not attainable at the same time, the two factors must be balanced taking the application and the situation into account. For smart sensored hand prostheses, LIFEs could be a nice trade-off thanks to a reduced (although not absent) invasiveness and a good selectivity.

3.3 Recording and processing of neural signals recorded using longitudinal intra-fascicular electrodes

3.3.1 Nerve recording capabilities

LIFEs can reliably record sensory signals providing a good interface to analyze the sensory information from peripheral nerves fascicles. Neu-
Figure 8: Comparison between different neural interfaces in terms of selectivity and invasiveness.

Nerve signals were recorded with chronically implanted LIFEs during periods of several months \[\text{[GLH91][LGH}^{+91}\]. Mechanical stimuli were used to selectively activate individual nerve fibres and determine the type of mechanoreceptors and their receptive field. Over a period of six months, there was a change in the recorded population, but the electrodes have continued to provide a representative sample of the implanted nerve. The total number of accessible units remained almost constant over time, and individual units persisted in recordings for up to six months.

When stimulating the skin area supervised by the implanted nerve, changing the type of stimulation modulated the active area of the nerve, the frequency and the conduction velocity of action potential peaks \[LZZ^{+05}\]. This confirms that the identification of different types of functional neural activity might be possible with intra-fascicular electrodes. Activity recorded in response to sensory stimulation of low-threshold mechanoreceptors showed short spike bursts with sufficient signal-to-noise ratio.

The capabilities for selective recording of different populations of afferent nerve fibres was assessed by applying the same stimulus to the six plantar pads of the rat hind paw while recording from each one of the eight different electrode sites in one tf-LIFE inserted in the sciatic nerve \[\text{[NLV}^{+07}\]. Different contacts recorded neural signals coming from different small areas of the skin of the hind paw.
3.3.2 Processing of neural signals from LIFEs

The possibility of extracting action potentials from neural signal recordings opens up research opportunities in neuroscience but also for the development of HBSs. The relatively high LIFE electrode selectivity shown in the previous section enabled to achieve very interesting results in terms of decoding neural information. For example, in the past, wire LIFEs have been used [GH92; MH94; MH97] to differentiate single units in multi-unit peripheral nerve recordings using different features and different classification schemes. In particular, different techniques (single channel, differential, dual channel) were used to sensory ENG signals from cat radial nerves using LIFEs. Among several classification techniques tested, an artificial neural networks allowed differentiation of 4 to 5 units with 70 to 90% reliability with single channel or differential recordings and 90 to 98% reliability with dual channel recordings [MH94].

These performance can be improved by identifying spikes generated by different axons (or small group of axons) related to the natural frequency coding of information [MCB+06]. In particular, the shape of the spike is determined by the relative orientation of the nodes of Ranvier in the nerve fibre and the inhomogeneity of the intra-fascicular space. Thus, the signals related to different nerve fibres can be identified and extracted on the basis of the shape records from a multiunit recording. Moreover, the spike type activity is generally larger in amplitude than the background activity with a better signal-to-noise ratio.

In this case, it is possible to use spike detection and sorting algorithms already used to process cortical signals recorded by invasive electrodes [Lew98; NB05; WH82; BJS93; WS99; HHC+00; BHSS04; ZWZ+04]. In particular, the use of discrete and continue wavelet transforms can be used to process spikes recorded from the CNS [OA01a; NB05] and also individual action potentials recorded from the PNS by microneurography [DCB+03b].

In the recent past, this approach has been used with ENG signals recorded using LIFEs to verify whether it is possible to record and detect spikes using tLIFEs in a stable and robust way. In [AYCM07] data were collected implanting a 2-channel PtIr LIFE in the lateral gastrocnemius/soleus (LG/S) or the medial gastrocnemius (MG) nerve of rabbits over a 32 week period. LIFE recorded the ankle position during ramp and hold flexion extension displacements. Spikes were detected from the denoised data and then divided into classes corresponding to which nerve fibre had elicited the spike. The algorithm for spike classification was im-
plemented by using the method described in [ZWZ+04]. The creation of class templates and the comparison of the spikes with the templates was based on the sum squared error and correlation coefficient between the aligned spike window and the template window. In these experiments, LIFEs were able to record clear spikes (and to sort different classes of spikes in a robust way) for more than eight months. However, in most of the cases the recordings were not robust for the initial drifting of the electrodes while it was possible to extract the same kinds of spike templates after 4-8 weeks. Even if these findings are not able to show that a stable connection can be achieved during chronic implants, they indicate the possibility of using LIFEs for a long period of time to record and process ENG signals detecting spike activities.

3.4 Experiments using Intra-fascicular electrodes in amputees

3.4.1 Works by Dhillon and Horch

Dhillon and Horch [DLHH04; DKSH05; DH05] gave the first demonstration of direct neural sensory feedback and control of a prosthetic arm. They demonstrated that interfacing individual fascicles of peripheral nerve stumps in amputees using intra-fascicular electrodes is possible. Stimulation produced graded, discrete sensations of touch or movement which the amputees referred to their phantom limb. Attempted movements of the phantom limb produced detectable motor neuron activity which could be used as graded control signals.

Tactile sensations elicited by nerve stimulation were distally referred, mainly to digit tips, localized to small receptive fields. Increasing the intensity of stimulation led to a spread of the sensation. Proprioceptive sensations, while initially vague, with practice were perceived as either movements of the whole digit or of individual joints, or as a feeling of a change in joint position.

For recording motor neuron activity, the subjects were instructed to attempt phantom limb movements. Motor signals were recorded, preprocessed and sent to a loudspeaker. Once the subjects, using the acoustic feedback, had learnt to generate motor activity associated with a phantom motion, a simple computer game was used to evaluate their ability to modulate the spike rate and, correspondingly, to control a phantom limb. All of the subjects were able to generate motor activity associated
with missing limb movements. With training, they were able to generate modulated motor activity and control the cursor position in the computer game, but with varied success in scoring hits. After the first training phase, the time taken to reach a target and the success rate remained approximately constant, suggesting that the subjects were using an innate sense of motor control, rather than something newly learnt to deal with an unnatural experience.

3.4.2 Works by Jia and colleagues

Jia and colleagues [JKZ+07] sampled potentials from peripheral nerve stumps with intra-fascicular electrodes to study residual motor transmission and explore the feasibility of nerve signal-controlled artificial limbs. They used intra-fascicular electrodes with a spring-like structure which they previously successfully tested in animal experiments.

Six electrodes were inserted into the ulnar, radial, and median nerves in the stump of an amputee. Potentials were recorded while 32 groups of electrophysiologic tests were conducted under volitional control. Actions included finger extension and flexion, forearm pronation and supination, and wrist extension and flexion. Each action was carried out with light, intermediate, and full efforts.

Then, the electrodes were interfaced to a nerve signal-controlled artificial limb. Under volitional control of the subject, finger extension of the artificial limb was triggered successfully by the radial nerve signal, while flexion was not.

Electrical signals gathered from the ulnar nerve and median nerve did not trigger any action of the prosthesis (i.e., the median and ulnar nerves were not activated during movements like finger flexion that are controlled by median and ulnar innervated muscles). One possible reason is that at the implanted position, there were mixed motor and sensory fascicules while the radial nerve was implanted near almost pure motor fascicules.

Even if only one subject was enrolled in this test and only part of the attempted movement could be successfully performed, this experiment shows that a long-term amputee can generate motor neuron activity related to phantom limb movement and control an artificial limb.
Chapter 4

Offline processing of afferent neural activity in animal model

To assess the possibility of extracting complex information from longitudinal intra-fascicular (LIFEs) electrodes, preliminary experiments with animal models were conducted. These first experiments were conducted on small animals and, therefore, induced afferent information was used instead of volitional efferent commands which are obtainable during experiments with primates. The assumption is that if an advanced signal processing technique gives access to several different types of sensory information, it has good chances to being able to also decode volitional motor commands which can be used to control a robotic hand.

4.1 Experimental setup

4.1.1 thin-film LIFEs

A new version of the LIFEs, named the thin film LIFEs (tfLIFE) was used in the experiments [YHK06; HK05]. These electrodes were developed on a micropatterned polyimide substrate which was chosen because of its biocompatibility, flexibility and structural properties [SSK05]. After microfabrication, this substrate filament (shown in Figure 7 pag. 35) 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 for implanting the electrode and is removed immediately after insertion.

4.1.2 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 chapter, 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 Fentanyl and 1 mg/kg Flurazidone 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.

4.1.3 Sensory stimuli

The protocol described in this chapter 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 in [GHML93].

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

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

### 4.1.4 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 2\textsuperscript{nd} 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.

4.1.5 Channel combination

The acquired data contained simultaneous monopolar recordings from different active sites of the tfLIFE electrodes (Figure 7 pag. 35). For each animal, the best monopolar or differential recording pair for a given stimulus was chosen using an automated procedure.

The procedure was inspired by a set of rules a human expert could have followed in order to identify the best recording site or combination of recording sites:

1. Sites with better SNR were favoured. In detail, what was considered SNR was the spikiness during stimulation w.r.t. the quiescent epoch.

2. If the chosen site did not “respond” to every type of stimulus but some other site carried complementary information, the two were combined into a differential configuration.

The reason for the first rule is straightforward. The second rule could be better explained using Figure 9 as example. The four available sites for that session were L1, L4, L2 and L3 (the latter two are similar to, respectively, L1 and L4 but noisier). L4 had clearer spiking activity but L4 alone was unable to carry information about the a90 stimulus. Therefore combining L4 and L1 could be advantageous because together they carry information about all stimuli with decent SNR. The main reasons for taking the difference of two sites are:

1. 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 Figure 9. This leads to some loss of information but avoids the computational cost of processing more signals.

2. The recording configuration is a differential recording thus cancelling noise from distant sources.
3. 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 Figure 10). 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 (Figure 10).

Once the inspiring rules were identified, a metrics able to translate the rules above into a set of measurements that could be replicated and automated, was designed.

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 [DCB+03a].

The following procedure was performed on every channel and differential combination of channels. The noise variance $\sigma_n$ was evaluated on acquiescent epoch. For every epoch of stimulation a measure of spikiness $s_{\text{epoch}}$ 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 $s_{\text{stim}}$. The final figure of merit $s_{\text{sig}}$ of the signal (i.e. of the combination of channels) was the minimum $s_{\text{stim}}$ of the different stimulus types.

Finally, the channel combination with higher $s_{\text{sig}}$ 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.

4.2 Signal preprocessing

4.2.1 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, in [DCB+03a] who found 10 kHz to be the optimal sampling frequency for
Figure 9: 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.
similar recordings. 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 (∼90 dB stopband) window and approximate filter length of 1024 samples.

4.2.2 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., Tik99, OA01b, KK03, DCB+03a].

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 coefficients were finally transformed back into the original domain de-noised.

Within this framework, wavelet denoising can be performed with
different modalities depending on the chosen type of decomposition, mother wavelet, thresholding function and threshold selection.

A decomposition scheme based on the Translation-Invariant Wavelet Transform \[CD95\] 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 Wavelet Transform (DWT). This can be a key point when dealing with abruptly changing signals such as spikes in our case. Figure 11 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). The idea behind the Translation-Invariant De-Noising is, given a range of shifts \( H = \{ h = 0, 1, \ldots, h_{\text{max}} \} \), to perform a DWT-based denoising on every shifted version of the signal \( s \) and then to average the results:

\[
T(S) = \text{Ave}_{h \in H} \text{Sh}_{-h}(\text{IDWT}(\text{Th}(\text{DWT}(\text{Sh}_h(s))))))
\] (4.1)

where \( \text{Sh} \) is the circular shift operator, DWT and IDWT are the wavelet decomposition and reconstruction up to level \( L \), and \( \text{Th} \) is the thresholding operator. Although the algorithm described by (4.1) implies the evaluation of \( 1 + h_{\text{max}} \) times the DWT-IDWT pair, a more computationally efficient algorithm exists \[CD95\]. In the current work the brute force approach of (4.1) with \( 2^L \) time shifts was implemented in C/C++ and used.

The Symmlet 7 mother wavelet was used, as in \[DCB^{+03a}\], because this choice outperforms other alternatives when working with similar neural signals.

A hard-thresholding function was used, the mathematical expression is

\[
\text{Th}^H(c) = \begin{cases} 
  c & |c| \geq \theta \\
  0 & |c| < \theta
\end{cases}
\] (4.2)

The selection of the threshold is important because it determines how aggressive the denoising is. Among the different approaches, the most used are the ones proposed by Donoho and Johnstone\[DJ94\], such as the universal threshold and the minimax threshold.

The commonly used method is the "universal threshold", defined as

\[
\theta = \sigma \sqrt{2 \ln(N)}
\] (4.3)

being \( \sigma \) the standard deviation of the noise, and \( N \) the number of samples in the signal \( s \). In \[DCB^{+03a}\] a modified version of the universal threshold
Figure 11: 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).

has been used by scaling it by a factor $k$. They found $k = 0.8$ to minimize false and missed detections on simulated neurographic signal.

In this paper, the minimax threshold method (4.4) was used

$$\theta = \sigma (0.3936 + 0.1829 \log_2(N)) \quad (4.4)$$

$\sigma$ 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 [DJ94]. 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 is similar to the optimal scaling factor $k$ determined in [DCB⁺03a]. 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 epoch using

\[
\sigma_l = \frac{\text{median}_k(|c_l(k)|)}{0.6745}.
\] (4.5)

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 this work, instead, the approximation is completely discarded. The rationale is that the approximation contains components below approximately \( 2^L f_n \) where \( L \) is the maximum level of decomposition and \( f_n \) is the Nyquist frequency, i.e., half the sampling frequency. Therefore we decided not to precede the wavelet denoising by a high-pass filter and just discard the approximation. We chose \( L = 3 \) in order to filter out frequencies below 750 Hz.

### 4.3 Spike Sorting

If a sample of the denoised signal was greater than a detection threshold \( \theta_d \), 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.

#### 4.3.1 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.

### 4.3.2 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.

### 4.4 Classification

#### 4.4.1 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. \( F = [f_1, \ldots, f_M] \) where \( M \) was the number of templates found during the templates creation phase and each \( f_i \) is given by

\[
    f_i = \frac{n_i}{\sum_j n_j}.
\]
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.

### 4.4.2 Classifier

In order to infer the type of stimulus applied during a given epoch from the feature vector $F$, a classifier based on Support Vector Machines [CV95] was used making use of the open source library LIBSVM [CL01].

Support Vector Machines (SVMs) are a family of supervised learning methods for two-class classification problems. By means of a technique known as the "kernel trick" they tackle non-linear classification problems applying linear classification techniques. Feature vectors are non-linearly mapped to a very high-dimension space. In the transformed feature space a hyperplane is constructed to split the examples as cleanly as possible, while maximizing the distance to the nearest cleanly split examples. The type of SVM used are called $\nu$-SVM [CLS05]. Among the various kernels available we chose the radial basis function (RBF) because it allows fairly complex separation surfaces while still requiring a reduced number of hyperparameters to tune [HCL03].

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 $\nu$-SVM with RBF, hyperparameters $\nu$ and $\gamma$ need to be chosen. The regularization parameter $\nu$ 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 $\gamma$ determines the radius of the RBF. We set $\gamma = 1/\rho^2$ where $\rho$ is the radius of the smallest sphere in the input space that contains all feature vectors $F$ of the training set. In [Kee02] it is reported that this value 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. A number of approaches are possible [HWL06], and the most commonly used are one-against-one and one-against-the-rest. In this work we used
a one-against-one approach \cite{HWL06} 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.

### 4.5 Assessment of the results

#### 4.5.1 Validation scheme

In machine learning, cross-validation is a practice consisting on splitting the dataset of examples into two (or more) subsets such that supervised training is initially performed on one subset (the training set), while the another is retained “blind” (the test set) for later use in validating the trained machine and assessing its generalization abilities.

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.

4.5.2 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.

As the main goal of the system is to allow the user to route information through the interface and drive a robotic artifact, probably the best perspective to analyze the results is 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 channels one of the most important figures of merit is the mutual information, a measure of how much information can be obtained from the input random variable (U) by observing the output one (Y). The mutual information of U relative to Y is given by

\[
I(U; Y) = \sum_{u \in \Omega_U, y \in \Omega_Y} p(u, y) \log_2 \frac{p(u, y)}{p(u) p(y)}
\]  

(4.7)

where \( \Omega_U \) and \( \Omega_Y \) are the sets of possible values of, respectively, inputs and outputs. In our case they coincide and are the set of types of stimuli.

The mutual information depends on the probability distribution of U. The maximum, over all possible distributions, of the mutual information defines the channel capacity, i.e., the maximum amount of discrete information that the channel can carry

\[
C = \sup I(U; Y).
\]  

(4.8)

The channel capacity \( C \) was found with the Arimoto algorithm [Ari72].
4.5.3 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 denoising 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 \textsuperscript{[DCB+03a]}].

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 $F$. 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 $(\text{Corr, Wrong}) \sim \text{Prep} \ast \text{Proc} + \text{Sess}$ that, in the notation of R, means that the dependent variable (number of correct and wrong classifications) depends ($\sim$) on a the explanatory variables Prep (accounting for the preprocessing: WD or FIR) and Proc (accounting for the processing: Srt or RBI) and their interaction ($\ast$). The independent variable Sess for the session, was also included to take into account the differences among the datasets.

4.6 Results

4.6.1 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.

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

**4.6.2 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 $(\text{Corr, Wrong}) \sim \text{Prep} \ast \text{Proc} + \text{Sess}$. 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 $e^{\beta_i}$. Positive values of $\beta_i$ indicate increasing the performance while negative values indicate decreasing performance. It is worth noting that the use of the wavelet denoising 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 denoising 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.

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

<table>
<thead>
<tr>
<th>Explan. variable $X_i$</th>
<th>coeff. $\beta_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>+0.218</td>
</tr>
<tr>
<td>PrepWD</td>
<td>+0.674</td>
</tr>
<tr>
<td>ProcSrt</td>
<td>+0.358</td>
</tr>
<tr>
<td>SessC</td>
<td>+1.205</td>
</tr>
<tr>
<td>SessD</td>
<td>+0.524</td>
</tr>
<tr>
<td>SessE</td>
<td>+0.788</td>
</tr>
<tr>
<td>PrepWD:ProcSrt</td>
<td>+0.970</td>
</tr>
</tbody>
</table>

### 4.6.3 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. Figure 12 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 Figure 12, 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 Figure 13 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.

### 4.7 Discussion

#### 4.7.1 General considerations

The aim of the work presented in this chapter was to test, on afferent signals of animal models, whether the combined use of wavelet denoising
Figure 12: 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 grey 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 grey lines (scales on the right) are the absolute spike rate of the corresponding spike waveforms evaluated on 1-s sliding windows.
Figure 13: 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 Figure 12. The pdf has been estimated with kernel density estimation (or Parzen windows method) using a standard Gaussian kernel and bandwidth $h = 7.5\text{ms}$. 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.

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 intra-neural 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 other invasive and non-invasive interfaces (see Chapter 2).

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 here.

The results have been achieved by using only the best channel available (as in [DLHH04]) and it is likely that the use of several channels simultaneously could further improve the situation as also reported in [MH94]. 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 this work 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 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 [MDMS03], and algorithms for spike detection, alignment and spike sorting on VLSI architectures [ZKO+05;ZPG06]. 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 [MH97].
4.7.2 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 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 in [DH05].

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.

4.8 Conclusions

In this chapter, the results of investigations about the possibility of decoding information from ENG signals recorded with tfLIFEs by using wavelet denoising and spike sorting techniques were reported. The results achieved seem to indicate that the approach proposed could provide better performance than different previously 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 intra-fascicular 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.

Goal of the next chapter is to make some step further towards the application of this approach to efferent and afferent ENG signals in human subjects to develop innovative brain-controlled artificial limbs.
Chapter 5

Towards online experiments with human subjects

In this chapter we outline some of the steps made in order to allow the processing of efferent signals during an implant with a human amputee subject.

5.1 The acquisition chain

5.1.1 Hardware

The hardware side of the acquisition chain is constituted by a National Instruments NI PCIe-6251 acquisition board. It is a high-speed multi-function board incorporating several advanced features. The functions we need for our purposes are the digital inputs for the triggers marking the timing of the experiment, and the analog inputs for the neural recordings.

There are 24 digital inputs with TTL voltage levels.

The board allows sampling of up to 16 analog channels in single-ended configuration or 8 channels in differential configuration. The sampling is not simultaneous but sequential, meaning that the different channels are sampled at slightly different times. When sampling more than one channel, the overall maximum sampling rate is 1 MS/s. The signals are quantized with 16-bit and the input range can vary from $[-0.1V, 0.1V]$ to $[-10V, 10V]$. 
Figure 14: Data flow within the data acquisition software. The recordings are acquired in raw format from the A/D card and saved together with the triggers. Then they are converted to the corresponding voltage at the probe. The preprocessing can be performed using the wavelet denoising algorithm described in this work, or with FIR filtering. The features can be represented by the output of the spike sorting algorithm described in Section 4.2.2 or using simple cuff-like rectified bin integration (RBI). Finally the data are classified and the inferred grasp is translated to a command for the Cyberhand.

The signals gathered by the electrodes are pre-amplified before being fed to the acquisition board. The pre-amplifier is conform for use with human subjects and ensures galvanic isolation between the subject and the data acquisition system. While the inputs of the pre-amplifier are differential, the outputs are single-ended, therefore with this configuration 16 channels can be simultaneously recorded by the acquisition board.

5.1.2 Software

The software used to acquire the signals was programmed using the LabVIEW [Lab] visual programming language. LabVIEW is developed by National Instruments, the manufacturer of the acquisition board, therefore it provides a reasonably easy interface to the board.

In order to understand the role of the acquisition software, let us follow the flow of the signals. Through specialized blocks (in the visual LabVIEW language, the block is the equivalent of a function in textual languages) the acquisition board is initialized with the proper settings, such as the number of channels and the sampling frequency.

The raw data from the analog inputs of the board (the channels used for the recordings) and from the digital input port (the triggers) are available within the language through built-in blocks. These data are saved to files in their original raw format in order to preserve all the information
while limiting the amount of disk space used.

The recordings of the neural signals are then scaled from a/d units to microvolts. As preprocessing step, the user can decide whether to use the wavelet denoising algorithm presented in Section 4.2.2 or a simpler FIR bandpass filter between 800 Hz and 2500 Hz. The signal is also downsampled by a factor 4.

Then the neural signals are processed in order to extract features for the classification. The user can decide whether to use the spike sorting algorithm described in Section 4.3 or a simpler cuff-like rectified bin integration (RBI) approach. In both cases all channels are processed independently and the features of all channels are concatenated in a composite feature vector.

Finally the data are classified using a support vector machine previously trained (together with the spike sorting algorithm) offline on antecedently recorded datasets. The output of the classifier represent the
inferred grasp which is finally sent through the serial port as a command for the Cyberhand.

The algorithms for wavelet denoising, spike sorting and SVM classification were developed in C, compiled as dynamic link libraries, and accessed from within LabVIEW using the Call Library Function Node. This approach has the advantage of allowing partial reuse of code already developed for the offline analysis described in Chapter 4. Furthermore, despite visual programming languages are intended to be easier than textual ones, during development of complex non-standard algorithms, their usage can become even more cumbersome than that of a textual language. Finally, the computation overhead of visual languages with respect to C or Fortran, increases dramatically if the program is made of many simple operations repeated a great number of times. This is exactly the situation of implementing non-standard signal processing like in our case.

5.2 Timing the experiment

5.2.1 QtProtocol

During an experimental procedure, the subject is instructed to perform different movement attempts in different phases of the protocol. In order to be able to inspect the recordings offline, or to perform the training of any kind of machine learning algorithm, the recordings should be somehow marked or labelled. This enables the operator who is processing the data offline, to know the exact onset and the type of activity performed by the subject in every part of the recordings.

If instructions are given vocally, recording an audio track together with the neural signals could suffice. The clear drawback is that the operator needs to manually transcribe the recordings before being able to perform any kind of automated processing of the data. In addition, as the instructions are given by a person, they can slightly vary during the protocol.

We have chosen to use a computer program instead. This software performs the timing of the experiment, uses animations to give instructions to the subject on the movement he/she should try to replicate, and sends a trigger signal to the acquisition software in order to set the corresponding marker in the recordings.

We called this software QtProtocol. It has been developed in the C++
programming language using the Qt4 libraries [QtS]. The program consists of one window split into two vertically stacked panes. The upper pane shows the animations and the instructions which are given by the subject. The lower pane allows the operator the see and edit the entire sequence of instructions given to the user Figure 16 and is hidden during the real experimental session.

In order to allow a higher degree of flexibility, the sequence of actions is given in form of a script. This allows the operator to tune the protocol quickly without the need to recompile the whole QtProtocol application. The base scripting language is javascript which is a fairly powerful scripting language. Figure 17 shows an example of a real script used to perform the timing of the sequence, give the subject written instructions and show him/her the movement he/she is supposed to try to replicate.

Additional ad-hoc functions are provided to the programmer of the script. A brief description of these functions are visible in the grey help box in Figure 17 together with a simpler usage example. The key functions are loadmovie and playmovie, used to load and play the movies with the desired hand movement; instr, to show an instruction message below
the movie; and \textit{trigger}, which sends a trigger through the parallel port to the acquisition card in order to be recorded together with the neural signals.

The implementation of these functions is coded in C++ at the design stage of QtProtocol and they are exposed to the future javascript programmer through the QtScript interface.

5.2.2 Emo

The instructions about the movement to attempt are given to the subject in form of written messages and by means of movies showing a virtual hand performing the target movement (Figure 18).

The movies were created using Blender, a 3D graphics application able to render animations of rigid and soft bodies. Animations of a high quality virtual hand were preferred to recording a real hand performing the same actions. The main reason was that a virtual hand might be easier to be imagined as a “representation” for the missing hand than a human
hand belonging to another person. The latter, in fact, could in principle produce a reaction similar to what experienced by Clint Hallam, the first recipient of a human hand transplant which had the hand re-amputated at his request [Whi00].

The virtual hand used, was originally designed for *Emo*, one of the main characters of the computer-generated short film “Elephants dream” [Ele06] whose content and sources were released under the Creative Commons Attribution license.

In Blender, a soft object is made up by an armature and a mesh attached to it. Changing the relative positions of the bones, makes the mesh smoothly deform in order to adapt to the new configuration of the armature. In the case of the hand, the armature is made up by bones which mostly replicate the actual bones of the hand, while the mesh represents the skin. Virtual bones connect the base to the tip of each finger. These bones can be rotated and scaled and, thanks to the hierarchy between the bones and to constraints, the actual bones representing the phalanges rotate and curl accordingly. This mechanism can be thought as a kind of underactuation which allows the control of the hand position with less variables. Figure 19 shows an octahedron view of the hand and its virtual bones, together with the names of the cinematic variables which determine the hand shape. In order to create the animation of a complex
Figure 19: Screenshot of Blender showing the Emo hand in the octahedron view. The octahedrons are a representation of the virtual bones used to determine the shape of the hand without the need of defining the position of every single joint. The position of the real bones is automatically determined by Blender in order to satisfy constraints and hierarchies between between the bones.

For each movement, the configuration of the bones must be specified for some key frames. Then, the position between key frames is automatically determined through interpolation. Figure 20 shows how the pinch grip is obtained using 4 key-frames.

For each movement, the frames (e.g. Figure 18) rendered by Blender were encoded in forward (rest to target grasp) and reverse order (target grasp to rest) in two .mng movie files which QtProtocol is able to play. QtProtocol shows the forward movement from the rest position to the target one, waits some seconds in this hold position, then plays the reverse movie to return to the rest position.
Movies were created for the following movements:

- **simple movements:**
  - adduction of the thumb
  - opposition of the thumb
  - flexion of the carpal-metacarpal joint of the fingers
  - flexion of the interphalangeal joints of the fingers
  - adduction of the fingers
  - abduction of the index finger
  - abduction of the little finger
  - flexion of the interphalangeal joints of the little finger
  - extension of wrist
  - flexion of wrist
  - hyperextension of the fingers
  - close the hand
  - close the hand slowly
  - close the hand ballistic

- **complete grasps:**
  - lateral grip (like holding a key)
  - power grip (like holding a suitcase)
  - spherical grip (like holding an orange)
  - tip grip (like holding a needle)
  - tripod grip (like holding a pen)

- **rest** (to characterize the quiescent noise level).
Figure 20: The figure shows how the *pinch grip* is obtained using 4 key-frames. The first key-frame is the starting neutral position. The movement starts with the flexion of the carpal-metacarpal joint of the thumb until the second key-frame is reached. Then the movement continues with the flexion of the inter-phalangeal joints of the thumb and of the index finger. As the range of movement for the thumb is shorter than for the index finger, its flexion stops with the third key-frame, while the flexion of the index continues until the fourth and last key-frame. Once the four key-frames are defined, the intermediate positions are automatically determined by Blender through interpolation.
Chapter 6

Discussion and Conclusions

6.1 Discussion of the results

The restoration of sensorimotor functions for the control of artificial hands is a fundamental point in order to improve the quality of life of amputees. Several research groups are currently striving to improve the performance and reduce the hurdles of human machine interfaces. In the case of the amputee, this could open the possibility for the user to control and feel the prosthetic hand as a natural part of the body.

In the first part of this thesis we have presented a procedure to assess which interfaces are best suited for a given application by using throughput and latency as first prerequisites. We have also shown that this first stage must be followed by a well-thought analysis of several additional factors. We have carefully and thoroughly examined the case of the amputee user and its requirements, such as user-friendliness, invasiveness, bi-directionality, and possibility of natural control of the prosthesis. From this analysis, we have concluded that, albeit suboptimal from a mere throughput and latency point of view, peripheral invasive interface can represent a promising medium-term solution as a result of their reduced invasiveness (compared to cortical invasive interfaces), their bi-directionality, and their potential in terms of natural control of the prosthesis. In fact, it is in principle possible to record from the nerve fascicle that in the natural arm innervates a given muscle, and use the neural information to control the corresponding effector of the robotic hand. Similarly, it is possible to stimulate the afferent fibres which convey the
somatosensory sensations coming from the hand.

We have compared the different types of interfaces with the peripheral nervous system, finding in longitudinal intra-fascicular interfaces (LIFEs) a tradeoff between invasiveness and selectivity. Even if more selective interfaces might be considered for the future, LIFEs are currently the most adequate solution for the development of a neurally-controlled cybernetic prosthesis. We have reviewed some works where LIFEs have been successfully used in experiments with human subjects to record efferent activity and provoke distally referred sensations.

In order to assess the possibility of extracting complex information from LIFEs electrodes, we ran preliminary experiments with small animal models recording induced afferent information. Using the sophisticated signal processing techniques developed (wavelet denoising and spike sorting) and a robust classifier, we were able to discriminate four (or five) different classes of stimuli with performance in a range between 90% and 99%. These results confirmed and outperformed prior work carried out with different approaches (see Section 4.7.1).

These first experiments were conducted on small animal models recording induced afferent activity. We thought that this was a necessary step before attempting experiments with higher order animals or human subjects. The assumption was that if an advanced signal processing technique is able to give access to several different types of sensory information, it has good chances to being able to also decode volitional motor commands which can be used to control a robotic hand. Now that this preliminary experiment has given a positive outcome, there are plans to validate the approach with a human amputee.

Hence, several steps have been taken in the last part of this doctoral work in order to make possible the recording of neural signals from a human subject and allow online processing and control of the Cyberhand [CCM+06].

6.2 Open issues

6.2.1 The subject

One of the problems to consider when attempting to interface peripheral nerves in amputee limbs is the possible presence of neuromas. In fact, at the stump level, axons try to regenerate and, in the absence of a corresponding distal portion, the regeneration is ineffective, unregulated and
characterized by numerous disorganized axonal sprouts of small calibre embedded in scar tissue [SP85].

The proximal nerve stump may also suffer from progressive atrophy and degeneration of myelinated nerve fibres [CLD79; MS81]. As the amplitude of recorded action potentials varies approximately proportionally to the radius of the axon [SNJ+77], the signal-to-noise ratio of nerve recordings might decrease because of the atrophy and degeneration of motor neurons after amputation.

Furthermore, after long-term limb amputation, dynamic changes in the CNS can lead to the expansion of adjacent, intact, limb regions into cortical areas formerly representing the missing parts of the limb [ESF+97; CCH02].

All these possible issues may reduce the number of amputees which can successfully receive an implant of a cybernetic hand.

6.2.2 The interface

As briefly reviewed in Section 3.4, LIFEs have been successfully implanted in human subjects with upper limb amputation, offering the proof of concept that these type of interfaces may be used as bidirectional interface for the control of advanced cybernetic hands.

However, the electrodes were left only for a few days within the subjects. The usability of these electrodes outside a controlled laboratory environment is dependent on their performance along time, their safety and reliability as chronically implanted devices. Therefore, experimental studies need to be performed in laboratory animals and human beings to assess biocompatibility and stability of neural electrodes implanted over at least some months.

Animal studies indicated that LIFEs made on polymeric materials are biocompatible, have a low mechanical mismatch with the nerve and do not induce significant functional damage [LYKNss]. Nevertheless, the formation of fibrous tissue [MKH04], even if may not cause damage by itself, can increase the impedance of the electrodes and consequently reduce the already limited signal-to-noise ratio of the neural interface.

Finally, it is important to point out that the implantation of tfLIFEs (and in general intra-neural electrodes) is relatively blind (limited repositioning procedures can be attempted during the surgical insertion). For this reason, in some cases their position could not be the best one in order to get useful neural contact. This is the most important reason for
the problems encountered with one of the five rabbit during the experiments described in Chapter 4. Moreover, electrodes can drift during the implant. 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 [BMK+07] with tfLIFEs actuated in order to follow the signal have been carried out.

6.2.3 The processing

Even though the results presented in this doctoral work (Chapter 4) are very 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 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 [Zen04].

Another possible reason for performance degradation in real implants is that the anaesthetized animal model used in the experiments presented in Chapter 4 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.

It is possible that some form of adaptation or learning, due to the plasticity of the central nervous system, can take place after the implant. In this case, the processing algorithm should adapt online in order to account and exploit these changes. At the cost of a more complex algorithm, the mutual learning of the biological and artificial side of the hybrid bionic system can consolidate and improve the performance in time.

6.3 Future works

As briefly highlighted in the previous section, several open issues still need to be addressed before neurally controlled robotic hands can be permanently implanted in amputee users. As this field is highly interdisciplinary, some of the open questions should be addressed by medical research groups, others by chemical and material engineers, others by experts of signal processing and artificial intelligence.
In this doctoral work we have confirmed the possibility of accessing complex neural information from LIFEs electrodes. But more importantly, we have developed advanced signal processing techniques able to increase the classification accuracy with respect to previous methods. We proved this with afferent signals recorded from anaesthetized small animal models. The natural next step is to verify whether these results can be confirmed with motor commands recorded from awake human amputee subjects.

To this end, during the last part of this doctoral work, the software setup for the online processing of neural signals has been developed. The hardware setup has been arranged by the biomedical microengineering group at Campus Biomedico in Rome. Everything is ready in order to start a series of experiments with human amputees. As previous literature [DKSH05] reported that some of the subjects were unable to provoke volitionally induced neural activity, several possible recipients should be individuated.

If the results are confirmed with human subjects, this can represent an important step towards neurally controlled robotic hands.
References


81


G. Dhillon and K. Horch. Direct neural sensory feedback and control of a prosthetic arm. *IEEE Transactions on Neural System and Rehabilitation Engineering*, 13:468–472, 2005. [34] [38] [61]


G. S. Dhillon, T. B. Krüger, J. S. Sandhu, and K. W. Horch. Effects of short-term training on sensory and motor function in severed nerves of long-term human amputees. *Journal of Neurophysiology*, 93(5):2625–2633, 2005. [24] [31] [34] [38] [77]

G. S. Dhillon, S. M. Lawrence, D. T. Hutchinson, and K. W. Horch. Residual function in peripheral nerve stumps of amputees: implications for neural control of artificial limbs. *Journal of Hand Surgery - AM*, 29(4):605–18, 2004. [24] [34] [38] [60]


[Ele06] Elephants dream. Online [www.elephantsdream.org](http://www.elephantsdream.org), 2006. 69


[Lab] LabVIEW. Online at www.ni.com/labview 64


[XHP+05] M. Xiao, J. Hyppolite, M. Pomplun, S. Sunkara, and E. Carbone. Compensating for the eye-hand span improves gaze control in


