Non invasive Brain-Machine Interfaces

Final Report

Authors: Federico Carpi, Danilo De Rossi

Affiliation: University of Pisa, Interdepartmental Research Center “E. Piaggio”

ESA Research Fellow/Technical Officer: Carlo Menon

Contacts:

Federico Carpi
Tel: +39 050 2217064
Fax: +39 050 2217051
e-mail: f.carpi@ing.unipi.it

Carlo Menon
Tel: +31 (0) 71 56 58675
Fax: +31(0)715658018
e-mail: act@esa.int

Available on the ACT website
http://www.esa.int/act

Ariadna ID: 05/6402
Study Duration: 2 months
Contract Number: 19706/06/NL/HE

Table of Contents

Executive Summary

This document represents the final report of the study "*Non invasive brain-machine interfaces*" performed by the Interdepartmental Research Center "E. Piaggio" of the University of Pisa within the ARIADNA framework of activities promoted by the European Space Agency (ESA). Contents of the report are organized as follows. The first part presents a literature survey on the state of the art of brain-machine interfaces (BMI), with a particular emphasis on the non-invasive types. In order to discuss potential benefits deriving from the use even of additional interfaces, conceived as complementary and auxiliary for BMI, the second part reviews different types of non-invasive man-machine interfaces. Their working principles, implementations, possible applications and typical features are discussed. Such additional interfaces are considered as a useful help, especially for multi-task activities. The report then presents a selection of the most promising and feasible non-invasive BMI concept for space applications, as well as the most interesting man-machine interface concepts capable of working as auxiliary and complementary tools. In particular, selected concepts consist of EEG-based BMI, to be eventually used in combination with interfaces based on speech recognition, EMG activation and motion capture and gesture recognition. The final part reports potential fields of space applications for such types of interfaces.

Keywords: brain-machine interface, brain-computer interface, non-invasive interface, man-machine interface, EEG, fMRI, MEG, EMG, tele-operation.

1. Introduction

Brain-Machine Interfaces (BMI) or *Brain-Computer Interfaces* (BCI), also referred to as *Neuro-Prostheses,* are conceived as technological interfaces between a machine (usually a computer) and the brain of a user. They should permit the use to perform a certain task, usually without implementing any motor action. This implies that neural impulses generated by the user's brain are detected, elaborated and utilised by the machine, approximately in real-time, to perform definite tasks. As an example, information can be processed and employed to control mechanical systems (e.g. actuators) or electrical devices (e.g. electronic equipment).

Brains are characterised by every property that engineers and computer scientists detest and avoid. They are chaotic, unstable, non-linear, non-stationary, non-Gaussian, asynchronous, noisy, and unpredictable in fine grain, yet undeniably they are among the most successful 'devices' that evolution produced.

A great demand for brain-machine interfaces is arising nowadays, pushed by several promising scientific and technological recent results, which are encouraging the concentration of efforts in such a direction. The possibility of measuring, processing and decoding brain activity, so that to interpret neuronal signals, is regarded as the challenging possibility of bypassing damaged neural and/or motor structures in patients affected by motor disorders and paralyses.

Accordingly, BMI might represent advantageous systems of assistance for such patients, permitting them either to perform rehabilitation, or to communicate, or to receive a continuous assistance during their daily activities. The brain control of devices and systems of different type, such as computer virtual keyboards, home electronic equipment or aid system (e.g. wheel chairs), certainly represents one of the main goals of this discipline. Moreover, neuronal signals captured from the brain activity can be used for a different purpose: instead for controlling devices or systems deputed to perform external actions, they can be employed to trigger functional electrical stimulations (FES) of muscles or nerves. This is aimed at bypassing degenerated or interrupted biological electrical routes in patients being affected by neural disorders or degenerations, or having suffered from fatal accidents (e.g. spinal injuries). In such a case, the intention is to allow the brain to use a healthy portion of its own body as the effector of desired motor tasks.

In addition to biomedical applications, the availability of reliable, efficient and non-invasive brain-machine interfaces may provide advantages to different disciplines. The space field is one of them. In fact, as identified by ESA, as an example extra-vehicular activities may be performed by robotic systems teleoperated by astronauts by means of non-invasive brain-machine interfaces. Such interfaces may be used also to perform multi-task operations. However, in this case, it can be expected that an even higher level of reliability and efficiency can be obtained, by using not only brain-machine interfaces, but also by completing them with auxiliary systems, such as muscle-brain interfaces, as it is discussed further on in this study.

The concept of ‘reading’ the brain to detect intended actions and to use extrapolated signals to perform tasks has been developed so far in several ways by adopting different technical and methodological approaches and achieving different results. The main purposes of this work consist in providing a literature review on BMI, in the identification of the most promising state-of-the-art concept for non-invasive interfaces, and in the proposal of a suitable future time scale for the development of such an interface. Moreover, this study includes an additional part, which is only indirectly related to BMI. In fact, it is opinion of these investigators that a considerable benefit for a non-invasive BMI employed in certain types of tasks may derive from the concomitant use of an auxiliary interface, such as an EMG-based interface. This is particularly evident for certain types of multi-task activities, as it is going to be discussed in the following. Accordingly, this study reports additional information and discussions related to the development even of other types of non-invasive man-machine interfaces as auxiliary systems for BMI.

2. Brain-Machine Interfaces (BMI): a brief overview on the state of the art

A great deal of efforts in neuroscience, robotics, and computer science are today spent by many research groups to develop BMI. In order to provide a glance at this vast field, it can be useful to mention here at least some relevant examples [1-24], briefly reported below.

2.1 Non-Invasive Brain-Machine Interfaces

Experiments in humans utilizing modern invasive and non-invasive brain imaging technologies as interfaces have been conduced. The most commonly studied potential interface for humans has been electroencepalography (EEG), mainly due to its fine temporal resolution, ease of use, portability, and cost of set-up. However practical use of EEG as a BCI requires a great deal of user training and is highly susceptible to noise. In 2004 scientists of the Fraunhofer Society utilized neural networks to shift the learning phase from the user to the computer and thus recorded noticeable results within 30 minutes of training. Magnetoencephalography (MEG) and even functional magnetic resonance imaging (fMRI) have both been used successfully as rudimentary BCIs, in the latter case allowing two users being scanned in real-time to play Pong against one another by altering their haemodynamic response through various biofeedback techniques.
Recent studies have shown that imagining the execution of a particular sensori-motor task gives rise to almost the same pattern of neuronal activity in central nervous system as actual performance of the sensori-motor task. The state of the art is that correct decoding of EEG signal is possible to a very large extent. It is still not good enough for applications since the erroneous responses in a remaining 10% can lead to completely wrong actions. The main challenge in BCIs EEG-based is to identify the particular EEG signal components (features) that can be successfully used as control commands (Fig. 2.1).

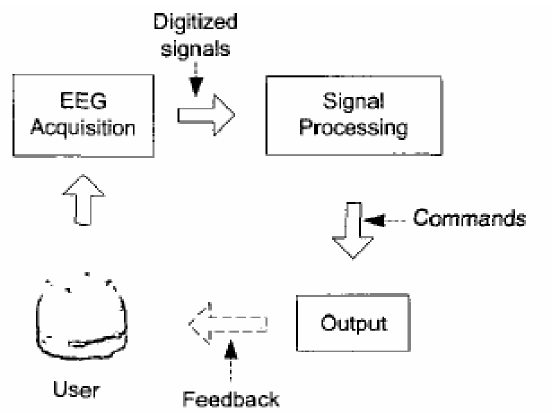


FIG 2.1: General BCI architecture (from [12]).

The main, but not unique, problem in these approaches is that single trial EEG data is very noisy, with data stemming from many sources. The characteristic responses to specific events are usually obtained by averaging signals from many trials, like in evoked signals. To successfully match single trial data, the relevant source of the signal needs to be separated out before it can be matched to average templates.

In matching, a similarity measure is applied to compare the signal trial with each template. This measure is still easily distorted by signals from non-relevant parts of frequency spectrum. After matching, a vector of similarity measurements for the different channels and the different templates needs to be classified into a category judgment. Optimal methods need to be found that do not over-fit data.

A rating of confidence of the judgment is much needed as outputting a wrong symbol may have a high cost in some situations. This aspect is lacking in all procedures proposed in literature so far. The system should work only in regions of high confidence.

EEG electrodes may bring many practical problems, like sensitivity to electromagnetic radiation, difficulty to place and position, varying conductance, usually a limited number of channels, and discomfort when used for a longer time.

There are few new concepts in the design of EEG measurement systems like miniaturized, battery-powered front-end close to patient, with fiber optic data transfer to the signal processing PC (see [7],[14]), or use of active electrodes, which have the property that the first amplifier stage is integrated within the electrode.

A group of the most important authors in the field of non-invasive BCIs gave the list of goals important for future progresses of these systems [15]. Future progress will depend on:

1) identification of those signals, whether evoked potentials, spontaneous rhythms, or neuronal firing rates, that users are best able to control;
2) development of training methods for helping users to gain and maintain that control;
3) delineation of the best algorithms for translating these signals into device commands;
4) attention to elimination of artefacts as electro-myographic and electro-oculographic activity;
5) adoption of precise and objective procedures for evaluating BCI performance.

At Graz University, the group of Pfurtscheller is one of the leading groups in Europe. They have extensive experience in recording and analyzing EEG signals with the aim to use them to restore functionality in patients who lost the ability to move their limbs (tetraplegia). They are also doing basic research, focusing on beta and gamma synchronization of cortical EEG activity. These oscillations in the frequency range between 15 and 70 Hz provide information about attention for perception and action. But the signal-to-noise ration in this frequency range is not high enough to use these signals for reliable on-line control. Their ultimate work [1] is aimed at assessing the feasibility of walking through a virtual city by using motor imagery. Therefore they combined EEG based BCI with Virtual Reality (VR) technology. A BCI transforms bioelectrical brain signals, modulated by mental activity (e.g. imagination of hand movement) into a control signal. This signal is used to walk forward/backward or to remain stationary inside a virtual city (Fig. 2.2). They demonstrated that the combination of a BCI with VR as a feedback means forms a feasible system for navigation in very simple virtual environments.

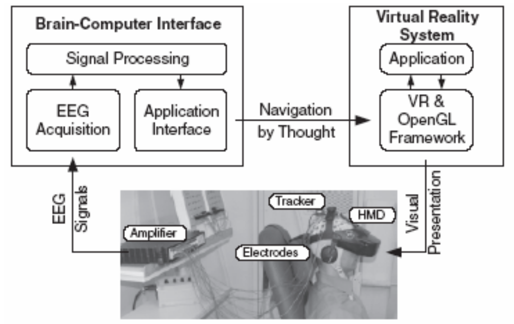


FIG 2.2: Schematic model of a combined framework, where the BCI system consists of the EEG as input, extracts and classifies EEG-parameters and calculates a control signal, which is sent to the VR system and influences there the visual feedback (from [1]).

Ultimate work in the field of EEG BCIs is made by the group from Taipei, Taiwan [2]. They used motor imagery electroencepalography, which embodies cortical potentials during mental simulation of left or right finger lifting tasks, to provide neural input signals to activate a BCI. The effectiveness of such an EEG based BCI system relies on two indispensable components: distinguishable patterns of brain signals and accurate classifiers (Fig. 2.3).

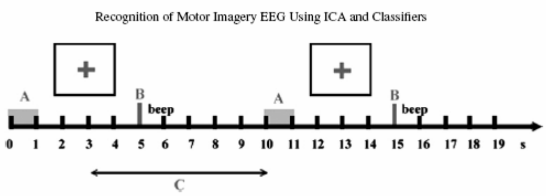


FIG 2.3:Timing of two consecutive 10-s trials of the motor imagery task. Each trial began with 1 s presentation of random noise during which subjects were allowed to blimk eyes(A).The subject was then instructed to stare at the fixation cross in the center of the monitor at 2 s and started to image right of left index finger lifting right after he\she heard an acoustic cue "beep"(with frequency 1kHz and 10 ms duration) at 5 s (B).Signals from 3 s to 10 s (C) in each trial (excluding bad epochs) were extracted to construct paired contralateral and ipsilateral rebound maps (from [2]).

They extracted two neural features termed contralateral and ipsilateral rebound maps, by removing artefacts from motor imagery EEG based on Independent Component Analysis (ICA). Results showed that, with the use of ICA, recognition rates for four classifiers (fisher linear discriminant (FLD), back propagation neural network (BP-NN), radial-basis functional neural network (RBF-NN) and support vector machine (SVM)) significantly improved (all classifiers gave rates higher than 70%).

Fundamental to effective BMI and neuroprostheses design is an understanding of how sensory and motor information are encoded, integrated and adapted by the nervous system. There are two current theories of sensorimotor integration which posit that neural information may be encoded centrally as an “internal model” of the environment or as a stochastic state-space model that modulates the activity of spiking neurons. Underlying both theories is a possible role for Bayes’ rule, as suggested by recent findings. In fact, it has been reported [3] that the brain may employ Bayesian internal models during certain types of sensorimotor learning, in order to optimize task-specific performance. Moreover, it has been recognised that the emergent activity of certain neural ensembles may be modelled as joint Bayesian point processes. These emerging concepts of neural signal processing have far reaching implications, from rehabilitation engineering to artificial intelligence. Because of the random variability of neuronal firing, sensory and motor signals are intrinsically stochastic in nature even when the environment is static. Bayesian learning and estimation may represent a kernel of human intelligence.

Steady-state visual evoked potentials (SSVEPs) are recorded from scalp over visual cortex reflecting visual information processing in the brain when the stimulation repetition frequency is higher than 6 Hz. A SSVEP-based BCI [4], applied by the group from Bejing University (China), has, as a potential advantage, a high information transfer rate. However, individual difference greatly affects its practical applications. This group presented the method of lead selection to improve the applicability of a SSVEP-based BCI system. ICA is employed to decompose EEGs over visual cortex into SSVEP signals and background noise. Optimal bipolar lead is selected by comparing signal correlation and noise correlation between different channels. The system with one optimal bipolar lead reached an average transfer rate about 42 bits/min. It has also been successfully applied to an environmental controller for motion-disabled.

A very important event in the field of EEG-based BCIs occurred in 2003, when it has been organized the BCI Competiton 2003 [5], with the aim to evaluate the current state of art. Signal processing and classification methods are essential tools in development of improved BCI technology. Four laboratories well versed in EEG-based BCI research provided six data sets in a documented format. Those data sets and their descriptions where available on the Internet. Researchers worldwide tested their algorithms and competed for the best classification result (Fig. 2.4). The data-sets covered slow cortical potentials, mu-rhythm, P300, motor imagery and finger tapping.

data set	research lab	contributor(s)
Ia	Massachussets Institute of Technology, Boston	Brett Mensh, Justin Werfel, Sebastian Seung
Ib	University of Tübingen	Vladimir Bostanov
IIa	Fraunhofer FIRST (IDA), Berlin	Gilles Blanchard, Benjamin Blankertz
IIb	University of Bielefeld	Matthias Kaper, Peter Meinicke, Ulf Großekathöfer, Thomas Lingner, Helge Ritter
IIb	Tsinghua University, Beijing	Xiaorong Gao, Neng Xu, Xiaobo Miao, Bo Hong, Shangkai Gao, Fusheng Yang
IIb	University of Tübingen	Vladimir Bostanov
IIb	Fraunhofer FIRST (IDA), Berlin	Benjamin Blankertz, Gabriel Curio (Charité Berlin, CBF)
IIb	Fraunhofer FIRST (IDA), Berlin	David Tax, Benjamin Blankertz
III	Fraunhofer FIRST (IDA), Berlin	Christin Schäfer, Steven Lemm (Charité Berlin, CBF)
IV	Tsinghua University, Beijing	Zhiguang Zhang, Yijun Wang, Yong Li, Xiaorong Gao

FIG 2.4: Table of winning teams of the BCI Competition 2003 for all competition data set (from [5])

The neuroinformatics group from Bielefeld University (Germany) proposed an algorithm [6] based on Support Vector Machines (SVM) to analyze EEG data from the P300 speller BCI paradigm. The oddball paradigm teaches that rare attended stimuli are accompanied by a positive deflection in the EEG after about 300 ms. This so-called P300 component is exposed by nearly every human being and is therefore independent of training of participants. They evaluated the performance of this technique and achieved high transfer rates up to 97.57 bits/min (47.26 bits/min). It is the highest bit rate for EEG-based BCIs that we found in literature. They also investigated generalization ability of classifier, obtaining transfer rates up to 61.04 bits/min.
SVM are also the kernel of work made by the German group [10]. For classifying EEG signals, they propose the use of the state-of-the-art feature selection algorithms Recursive Feature Elimination and Zero Norm Optimization, which are based on training of support vector machines (SVM). They adapt the methods for the purpose of selecting EEG channels, showing that for a motor imagery paradigm number of used channels can be reduced significantly without increasing classification error.
Another SVM-based approach is made by a Swiss group [12]. They presented a method for classifying EEG signals based on the information content of their correlative time-frequency-space representation (CTFSR). A SVM kernel is proposed that can be calculated in the time domain while it computes a similarity measure in the CTFSR space (Fig. 2.5).

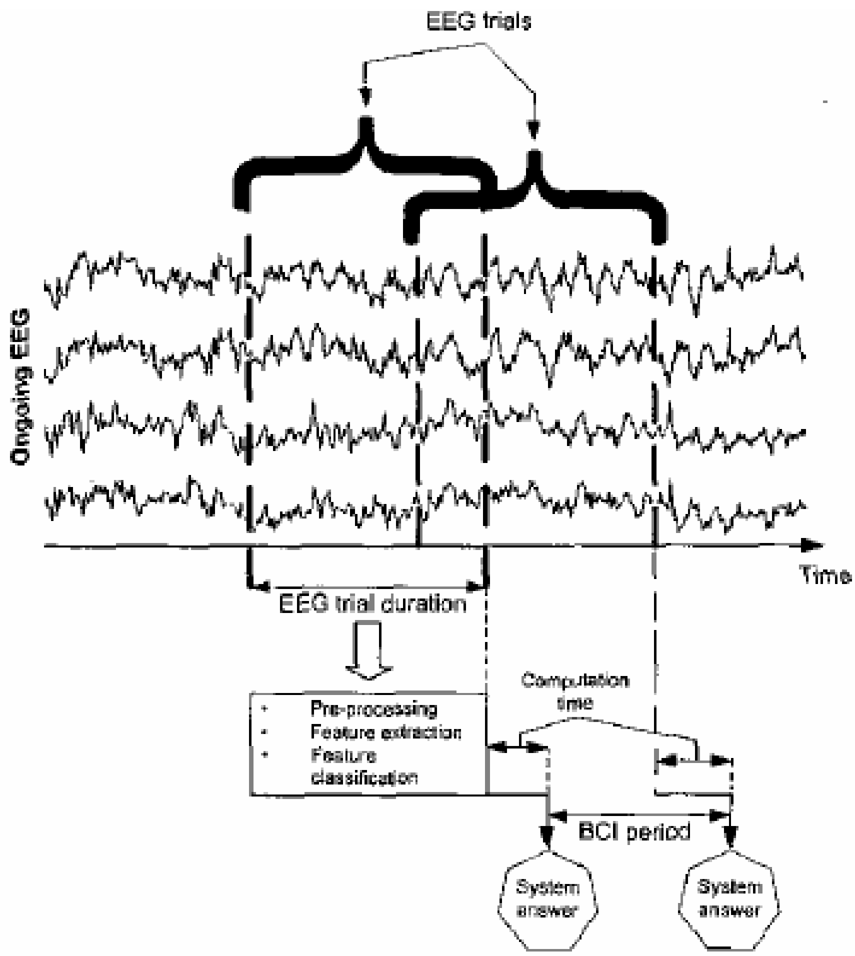


FIG 2.5: BCI scheduling (from [12]).

The group from LA University (USA) implemented a neural interface upon TinyOS-based sensing and communication platform [7]. The system amplifies, digitally encodes and transmits (by telemetry) two EEG channels of neural signals from an un-tethered subject to a remote gateway which routes signals to a PC (Fig. 2.6). This can be the foundation for chronic remote biological monitoring applications. The system is capable of amplifying, sampling, transmitting and reconstructing input signals at a rate of 480 8-bit samples per second.

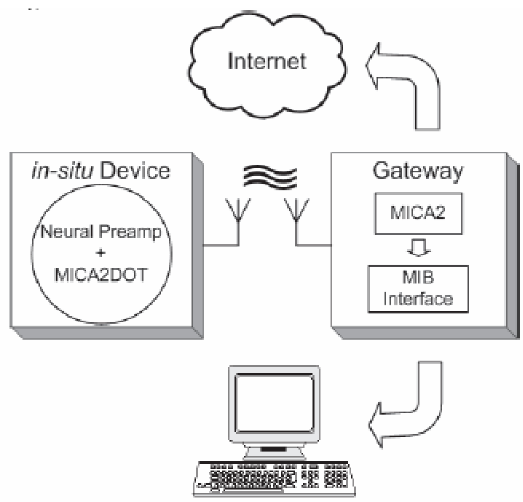


FIG. 2.6: Top level diagram of the neural interface system (from [7]).

Another remote BCI system is an ambulatory BCI (ABCI) [14]. It consists of a microcontroller-based circuit that acquires, digitizes and processes up to two amplified EEG signals and transmits the bandpass-limited power to a PDA for classification, translation and training ”game” control. The “game” allows user to train by providing a visual feedback signal that is to be cognitively controlled. The accuracy and response time are equivalent to a desktop BCI system (Fig. 2.7).

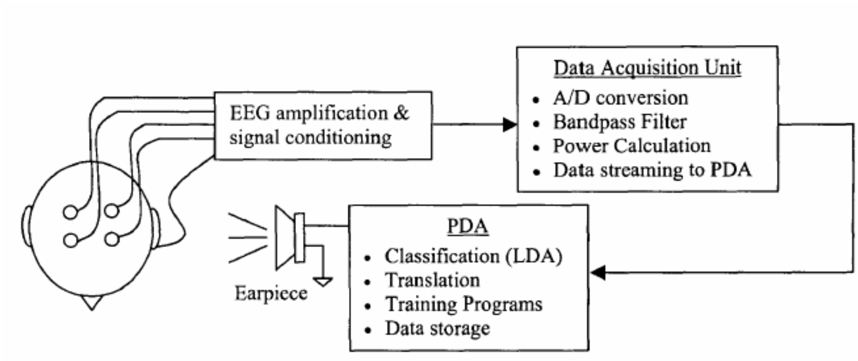


FIG. 2.7: Block diagram of a BCI system (from [14].

The Berlin Brain-Computer Interface (BBCI) [8] project is guided by the idea to train a computer by advanced machine learning techniques both to improve classification performance and to reduce the need for individual training. They presented two directions in which BCI can be enhanced by exploiting the lateralized readiness potential with the following aims: (1) for establishing a rapid response BCI system that can predict the laterality of upcoming finger movements before electromyographic onset even in time critical contexts and (2) to improve information transfer rates in the common BCI approach relying on imagined limb movements.
A very important result has been presented by a group from New York [9]. They showed that a non-invasive BCI that uses scalp-recorded EEG activity

and an adaptive algorithm can provide humans, including people with spinal cord injuries, with multidimensional point-to-point movement control that falls within the range of that reported with invasive methods in monkeys. In movement time, precision and accuracy, the results are comparable with those with invasive BCIs. The adaptive algorithm used in their non-invasive BCI identifies and focuses on EEG features that the person is best able to control and encourages further improvement in that control (Fig. 2.8). The results suggest that people with severe motor disabilities could use brain signals to operate a robotic arm without needing to have electrodes implanted in their brains.

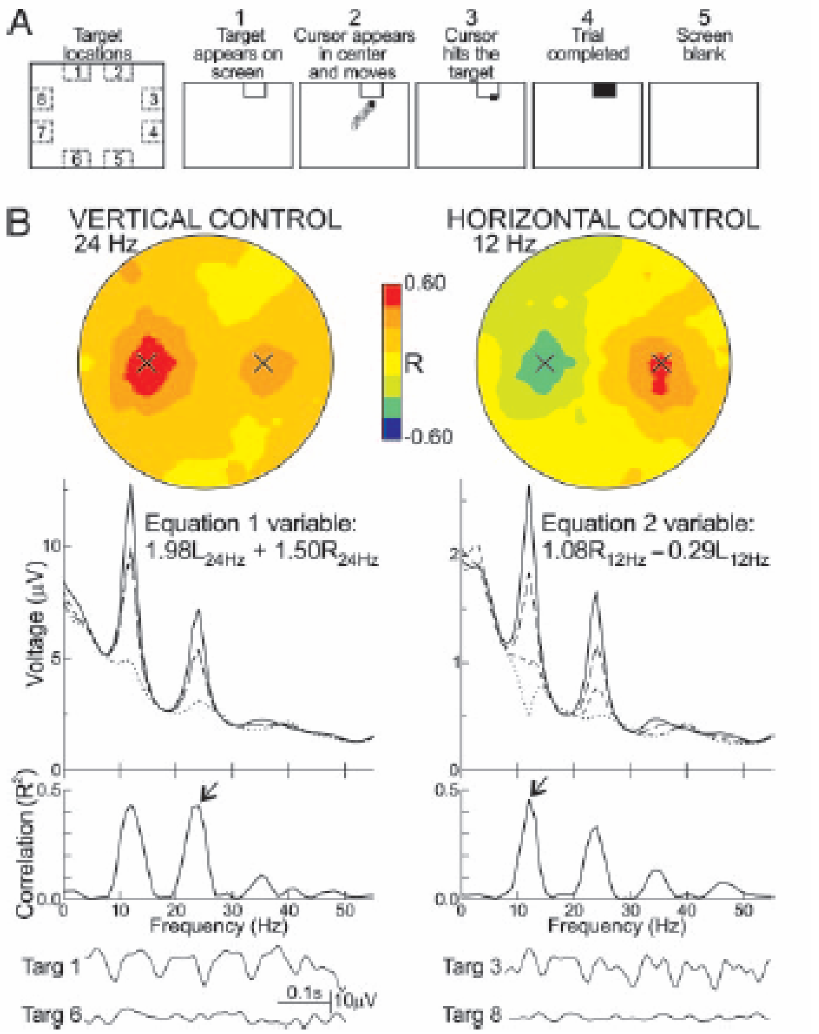


FIG. 2.8: (A) Protocol. The screen at left shows the eight possible target locations. The other screens show the sequence of events in one trial. 1, a target appears; 2, 1 s later the cursor appears and moves in two dimensions controlled by the user's EEG activity; 3, the cursor reaches the target; 4, the target flashes for 1 s; 5, the screen is blank for 1 s and than the next trial begins(Step 2 lasts up to 10 s. If the cursor does not reach the target in time, the trial jumps to step 5) (B) Topographical and spectral properties of user's EEG control (from [9]).

Another group showed [11] that two human subjects successfully moved a robot between several rooms by mental control only, using an EEG-based BMI that recognized three mental states. Mental control was comparable to manual control on the same task with a performance ratio of 0.74.
The novel idea introduced by this work is to control robots by mapping asynchronously high-level mental commands into a finite state automaton.

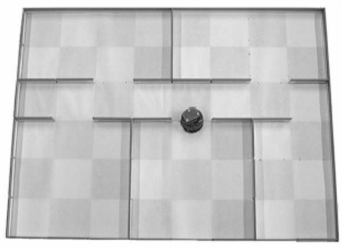


FIG. 2.9: The mobile robot in its environment, which consists of several rooms along a corridor. The robot is a two-wheeled vehicle. It has three lights on the top to provide feedback to the user and 8 infrared sensors around its diameter to detect obstacles (from [11]).

As an observation, it is here stressed that the fractal dimension of EEG has been proposed [13] as a significant component (new feature) for BCI classifiers. The fractal dimension is known as a good measure of the chaotic behaviour of EEG signals.
One of the classic experiments of cursor's control [16] was successfully performed by voluntary EEG modulation. Moving the cursor in one dimension, subjects were able to hit 100% randomly selected targets, while in two dimensions accuracies of 63% were achieved.
There are much less works in BCIs field where the voluntary signals are "read out" from magnetic resonance (fMRI and MEG). Reasons of that are several. The need for complicate, bulky, heavy and expensive driving and reading equipment is limiting attempts to use these types of BCIs. Another very limiting factor in applications of these BCI is the necessity that the subject does not move at all.

A group from Utah used fMRI to study brain activity [17] in subjects with spinal cord injuries while they are executing, or attempting to execute, movements of different limbs. They showed that their motor-cortical activation closely follows normal somatotopic organization in the primary and non-primary sensorimotor areas. They showed that it should be possible to access voluntary control signals by using a cortical neuroprosthesis.

2.2 Invasive Brain-Machine Interfaces

In this type of BMIs the signals originated by brain are recorded by implanted electrodes. Because of it's invasivity, the applicability of this approach is limited to animal experiments. Development of this systems have two basic directions: improvement of both implantable electrodes and signal processing methods.

The most important works in this field are those made by the Nicolelis's group. He works at Duke University, which started BCI experiments few years ago with rats, in which they recorded the simultaneous activity of large populations of neurons [18], distributed in the premotor, primary motor and posterior parietal cortical areas, while non-human primates performed two distinct motor tasks. Accurate real-time predictions of one- and three-dimensional arm movement trajectories were obtained by applying both linear and nonlinear algorithms to cortical neuronal ensemble activity recorded from each animal. In addition, cortically derived signals were successfully used for real-time control of robotic devices, both locally and through the Internet (Fig. 2.10). These results suggest that long-term control of complex prosthetic robot arm movements can be achieved by simple real-time transformations of neuronal population signals derived from multiple cortical areas in primates.

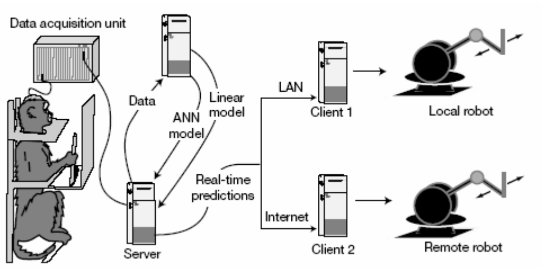


FIG. 2.10: Experimental set-up of one experiment of Nicolelis (from [18]).

In a famous experiment with a monkey [19], they demonstrated that primates can learn to reach and grasp virtual objects by controlling a robot arm through a closed-loop brain–machine interface (BMIc) that uses multiple mathematical models to extract several motor parameters (i.e., hand position, velocity, gripping force, and the electromyograms of multiple arm muscles) from the electrical activity of frontoparietal neuronal ensembles. As single neurons typically contribute to the encoding of several motor parameters, they observed that high BMIc accuracy required recording from large neuronal ensembles (Fig. 2.11).

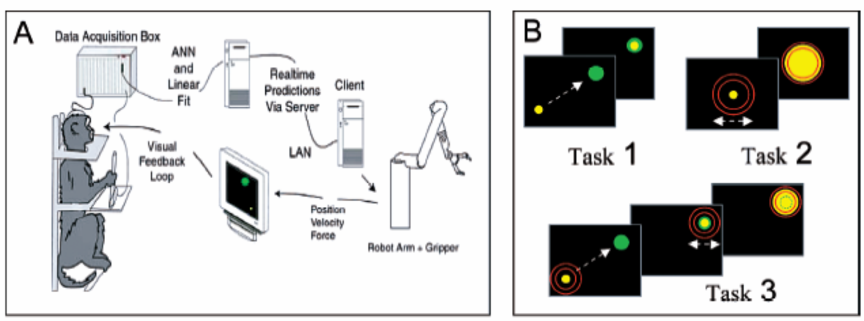


FIG 2.11: (A) Behavioural setup and control loops, consisting of the data acquisition system, the computer running multiple linear models in real time, the robot arm equipped with a gripper, and the visual display. Robot position was translated into cursor position on the screen, and feedback of the gripping force was provided by changing the cursor size. (B) Schematics of three behavioural tasks. In task 1, the monkey's goal was to move the cursor to a visual target (green) that

appeared at random locations on the screen. In task 2, the pole was stationary, and the monkey had to grasp a virtual object by developing a particular gripping force instructed by two red circles displayed on the screen. Task 3 was a combination of tasks 1 and 2. The monkey had to move the cursor to the target and then develop a gripping force necessary to grasp a virtual object (from [19]).

A German group (Mehring et al., 2003) recently reported that measuring local-field potentials in monkey motor cortex can be used to correctly predict the direction and velocity of arm movements to various targets in 90% of the cases.

One of the ultimate brain signal processing works [20] proposes the use of optimized brain–machine interface (BMI) models for interpreting the spatial and temporal neural activity generated in motor tasks. In this study, a nonlinear dynamical neural network is trained to predict the hand position of primates from neural recordings in a reaching task paradigm. They first developed a method to reveal the role attributed by the model to the sampled motor, premotor, and parietal cortices in generating hand movements. Next, using the trained model weights, they derived a temporal sensitivity measure to asses how the model utilized the sampled cortices and neurons in real-time during BMI testing.

A Statistical Encoding Model [21] for movement-related motor neurons using multi-electrode array recordings during a two-dimensional (2-D) continuous pursuit-tracking task avoids massive averaging of responses by utilizing 2-D normalized occupancy plots, cascaded linear-nonlinear (LN) system models and a method for describing variability in discrete random systems. It has been found that the expected firing rate of most movement-related motor neurons is related to the cinematic values by a linear transformation, with a significant nonlinear distortion in about one third of the neurons.

An American group demonstrated for the first time [22] that electrocorticographic (ECoG) activity recorded from the surface of the brain can enable users to control a one-dimensional computer cursor rapidly and accurately (Fig. 2.12). They first identified ECoG signals that were associated with different types of motor and speech imagery. Their results suggest that an ECoG-based BCI could provide people having severe motor disabilities with a non-muscular communication and control option that is more powerful than EEG-based BCIs and is potentially more stable and less traumatic than BCIs that use electrodes penetrating the brain.

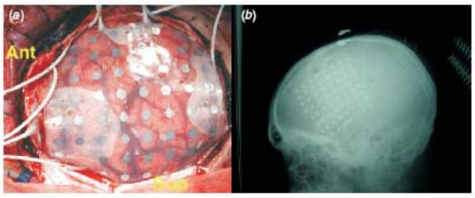


FIG. 2.12: Examples of electrode placement and ECoG signals. (a) Intra-operative placement of a 64-electrode subdural array. Inter-electrode spacing was 1 cm and electrode diameter was 2 mm; Ant: anterior. (b) Post-operative lateral skull radiograph showing grid placement (from [22]).

A rat car system and various flexible neural probes have been developed by a Japanese group [23]. The probe, which they have done is sufficiently flexible to enable minimally invasive recording. The rat car system uses signals recorded from a rat motor cortex to control the car carrying the rat. It is being developed as an application of BMI system.

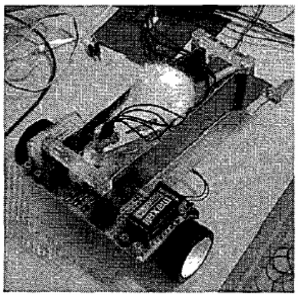


FIG. 2.13: Photo of the Rat Car System. The signals recorded from the rat motor cortex are utilized to control the car on which the rat is riding (from [23]).

A wireless data communication between brain implants and computer has been realized by the Pittsburgh University group [24]. This communication investigates this link and presents a new design using the mechanism of volume conduction of biological tissues. A theoretical model of volume conduction of the head is utilized to compute signal strength in data communication and the result is evaluated by a physical model (Fig. 2.14). The two-way data communication sensitivity of the volume conduction channel is found to be symmetric, as suggested by the reciprocity theorem. An x-shaped volume conduction antenna has been designed which not only enhances transmission/reception, but also minimizes the space for brain implantation. This investigation provides a new enabling technology to integrate brain function and the external computing environment with broad applications.

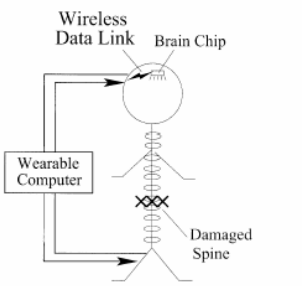


FIG. 2.14: Wireless data communication establishes a closed-loop information link bypassing the damaged spine, marked by "XXX" (from [24]).

2.3 References

[1] G.Pfurtscheller,R.Leeb "Walking through a virtual city by thought" .*Proceedings of the 26th Annual Conference of the IEEE EMBS,San Francisco,USA*. September 2004,pp.4503-4506

[2] Chin-I Hung and other "Recognition of motor imagery EEG using independent component analysis and machine classifiers". *Annals of Biomedical Engineering, Vol 33.,N.8* .August 2005,pp.1053-1070

[3] C.-S.Poon "Sensorimotor learning and information processing by Bayesian internal models". *Proceedings of the 26th Annual Conference of the IEEE EMBS, San Francisco,USA*. September 2004,pp.4481-4482

[4] Yijung Wang and other "Lead selection for SSVEP-based brain-computer interface". *Proceedings of the 26th Annual Conference of the IEEE EMBS,San Francisco,USA*. September 2004,pp.4507-4510

[5] "The BCI Competition 2003:Progress and Perspectives in detection and Discrimination of EEG single trials".*IEEE transactions on biomedical engineering,Vol.51,N.6*. June 2004,pp.1044-1051

[6] M.Kaper,H.Ritter "Generalizing to new subjects in brain-computer interfacing". *Proceedings of the 26th Annual Conference of the IEEE EMBS,San Francisco,USA*. September 2004,pp.4363-4366

[7] S.Farshchi and other "A TinyOS-bassed wireless neural interface". *Proceedings of the 26th Annual Conference of the IEEE EMBS,San Francisco,USA*. September 20044,pp.4334-4337

[8] M.Krauledat,R.Muller "Improving speed and accuracy of BCI using readiness potential features". *Proceedings of the 26th Annual Conference of the IEEE EMBS,San Francisco,USA*. September 2004,pp.4511-4515

[9] J.Wolpaw and J.McFarland "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans".*Pnas,Vol.101,N.51*.December 2004,pp.17849-17854

[10] T.N.Lal and other "Support vector channel selection in BCI".*IEEE trancactions on biomedical engineering,Vol.51,N.6*.June 2004,pp.1003-1010.

[11] J. del R. Millàn, F. Renkens, J. Mouriño, and W. Gerstner "Non invasive brain-actuated control of a mobile robot by human EEG". *IEEE trancactions on biomedical engineering, Vol.51, N.6*. June 2004,pp.1026-1033.

[12] G. Garcia and others "Correlative exploration of EEG signals for direct brain-computer communication".2003

[13] A.Bashashati and other "Fractal dimension-based EEG biofeedback system". *Proceedings of the 25th Annual Conference of the IEEE EMBS,Cancun,Mexico*. September 2003,pp.2220-2223

[14] R.Weinmann, J.Schroeder "A microcontroller-based ambulatory brain-computer interface training system" . *Proceedings of the 25th Annual Conference of the IEEE EMBS,Cancun,Mexico*. September 2003,pp.2204-2207

[15] J.R.Wolpaw,N.Birbaumer,D.J.McFarland,G.Pfurtscheller,T.M.Vaughan"Brain-computer interfaces for communication and control".*Clinical Neurophysolog,N.113*.2002,pp.767-791

[16] A.Kostov, M.Polak "Parallel man-machine training in development of EEG-based cursor control". *IEEE transactions on rehabilitation engineerin,Vol8,N.2*.June 2000,pp.203-205

[17] S.Shoham and others "Motor-cortical activity in tetraplegics" *Nature*.October 2001

[18] Nicolelis and others "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates". *Nature,Vol 408*. November 2000,pp.361-365

[19] J.M. Carmena, M.A. Lebedev, R.E. Crist, J.E. O'Doherty, D.M. Santucci, D.F. Dimitrov, P.G. Patil, C.S. Henriquez, M.A. L. Nicolelis "Learning to control a brain–machine interface for reaching and grasping by primates". *PlosBiology, Vol 1,Issue 2* .September 2003,pp.193-208

[20] J.C.Sanchez, D. Erdogmus,M.A.L. Nicolelis, J.Wessberg "Principe Interpreting spatial and temporal neural activity through a recurrent neural network Brain–Machine Interface". *IEEE transactions on neural systems and rehabilitation engineering,Vol.13,N.2*. June 2005,pp.213-219

[21] S.Shoha, L. M. Paninski, M. R. Fellows, N. G. Hatsopoulos, J. P. Donoghue, Richard A. Normann "Statistical encoding model for a primary motor cortical brain-machine interface". *IEEE transactions on neural systems and rehabilitation engineering,Vol.52,N.7*. July 2005,pp.1312-1322

[22] E .C. Leuthardt, G.Schalk, J.R.Wolpaw,J.G. Ojemann and D.W.Moran "A brain–computer interface using electrocorticographic signals in humans". *JOURNAL OF NEURAL ENGINEERING*. June 2004,pp.63-71

[23] T.Suzuki and others "Development of flexible neural probes and their application to rat brain interfaces". *SICE Annual Conference in Sapporo,Hokkaido,Japan*. August 2004,pp.2377-2380

[24] Mingui Sun and others "Data communication between brain implants and computer". *IEEE Transactions On Neural Systems And Rehabilitation Engineering,Vol.11,N.2*. June 2003,pp.189-192

3. Non invasive man-machine interfaces besides BMI

Integrating human and robotic machines into one system offers multiple opportunities for creating assistive technologies that can be used in aerospace, biomedical and industrial applications. In this context, the development and use of non invasive man-machine interfaces is progressively gaining a considerable importance. The following sub-sections describe the most relevant types of such interfaces, different from a typical BMI.

3.1 EMG based interfaces

The electromyogram is an electrical signal generated by neuromuscular activity. It can be recorded non-invasively by using surface electrodes. Methods for effective recording and computer-aided analysis of EMG signals have been the object of study in the field of biomedical engineering for the last three decades. Typical applications based on EMG employ signals from the forearm [1]. EMG based interfaces generally involve signal acquisition from a number of several electrodes, signal processing (feature extraction) and real-time pattern classification [2,3]. Classification methods based on both statistical and neural network approaches have been made with satisfactory results. However, given the complexity of the task and the variability of the EMG signals these systems usually require calibration for each user or training of the pattern recognition algorithms. In a different fashion EMG signals have been used in conjunction with other physiological signals (skin conductivity, blood pressure and respiration) to detect the affective state of the user. Experimental results from a preliminary study show that even with simple processing techniques it is possible to detect brief muscle contractions in data acquired from moving subjects. The results encourage further development of this kind of interface [1]. The signal processing and pattern recognition strategies should be improved to achieve higher accuracy. At the same time, the efficiency of the interface can be increased introducing feedback. The use of dry electrodes is being considered to promote user acceptance. The EMG signals can be classified in realtime with an extremely high degree of accuracy for controlling a robotic arm-and-gripper. The human hand is a complex system, with a large number of degrees of freedom (DoF), sensors embedded in its structure, actuators and tendons, and a complex hierarchical control. Despite this complexity, the efforts required to the user to carry out the different movements are quite small [2]. On the contrary, prosthetic hands are just a pale replication of the natural hand, with significantly reduced grasping capabilities and no sensory information delivered back to the user. Several attempts have been carried out to develop multifunctional prosthetic devices controlled by electromyographic signals (myoelectric hands), harness (kinematic hands), dimensional changes in residual muscles, and so forth, but none of these methods permits the “natural” control of more than two DoF. Electromyographic signals, collected at the skin surface, have been used for the control of upper limb prosthetic devices since 1948 because they provide easy and non-invasive access to physiological processes that cause the contraction of the muscles. At present, the process of EMG signals is the most common approach used for the control of active prosthetic hands. In any case, the myoelectric signal permits the control of no more than one or two active DoF (generally, one DoF for the gripper and one for the wrist). Limitations in the mechanics of the prosthetic device and in the processing of EMG data make it difficult to control more. In past decades, and especially during the last years, many efforts have been carried out in order to implement effective control algorithms based on the processing of EMG signals. Starting from the first attempts in the late 1940s, several EMG-based algorithms have been developed and used to enhance the functionality and usability of prosthetic hands. As shown in figure 3.1.1, the formal scheme for the acquisition and analysis of the EMG signal for the control of prosthetic devices is composed of several modules:

- signal conditioning and pre-processing
- feature extraction
- dimensionality reduction
- pattern recognition
- offline and online learning

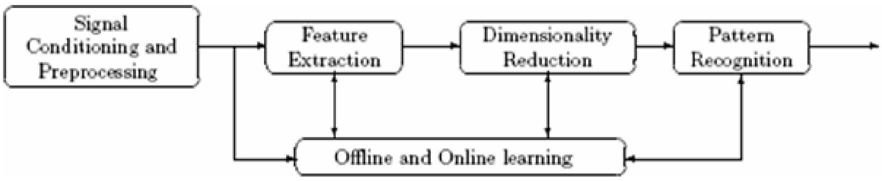


FIG. 3.1.1: Scheme for acquisition and analysis of an EMG signal (from [2])

The first module preprocesses the EMG signal in order to reduce noise artifacts and/or enhance spectral components that contain important information for data analysis. Moreover, it detects the onset of the movement and activates all the following modules. Many EMG-based control systems are able to control a single DoF in a prosthetic limb (hand open/close, wrist or elbow flexion/extension). These systems generally extract the EMG amplitude or rate of change by using two electrodes placed on two antagonist muscles (e.g., biceps and triceps brachii or flexor and extensor of the forearm, depending on the level of the amputation). This information is used to define the state of the hand and to control its speed or strength in a constant or even proportional way. Most commercial myoprocessors used in prosthetic control are now based only on one dimension of the EMG signal—the variance or mean absolute value. Several authors successfully contributed in refining variance estimation from the myoelectric signal, for example by applying a whitening filter or changing the smoothing window length in order to increase the number of states available from surface EMG signal. These techniques require a different muscle contraction for each controlled function, making the control of two or more joints very difficult. Other researchers have attempted to increase the information from one or two channels by using a time-series model. An example of processing of EMG signal is shown as figure 3.1.2.

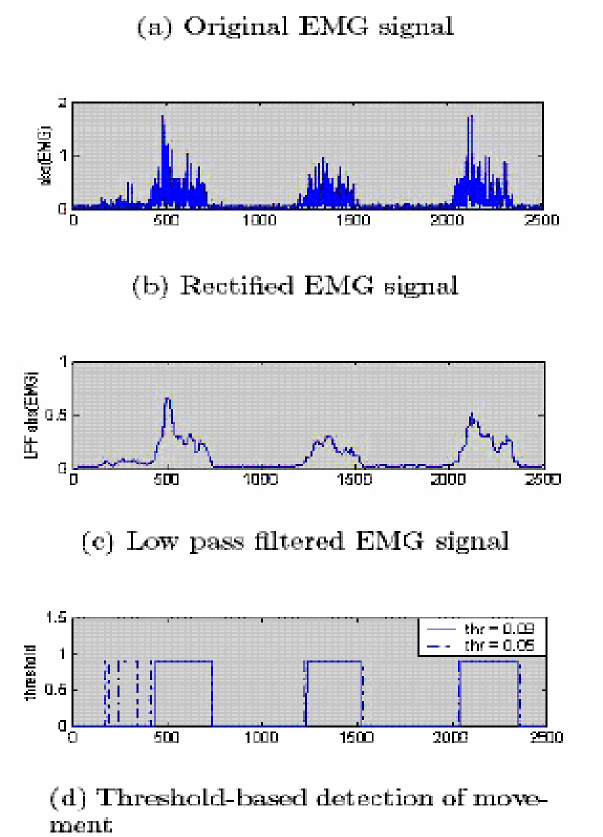


FIG. 3.1.2: Processing of an EMG signal recorded from a biceps brachii muscle (from[2]).

Despite some promising results, this method turned out to be sensitive to changes in signal amplitude. All these systems have been successfully implemented, but they cannot provide sufficient information to effectively control more than one DoF. Generally, all commercial myoelectric control systems are based on the common assumption that the instantaneous value of the myoelectric signal contains no information. Users are trained to produce a constant level of activation of muscles, and the prostheses are tuned according to these values. A steady-state EMG signal, however, has very little temporal structure because of the active modification of recruitment and firing patterns needed to sustain a contraction. The parameters that could be extracted to quantify its amplitude (e.g., variance, mean absolute value) or its frequency characteristics (e.g., Fourier spectrum, median frequency) are often not sufficient to distinguish between more than two classes of movement. Starting from the 1990s, researchers found that there is useful information in the transient burst of myoelectric signal. Hudgins and colleagues showed that there is a considerable structure in the myoelectric signal during the onset of a contraction. Furthermore, this structure seemed to be different for different kinds of contraction, and further works demonstrated that transient EMG signals have a greater classification capacity than do steady-state signals. For this reason several features are extracted. EMG signal patterns differ among individuals. Moreover, electrical impedance of the skin, electrode locations, time variations caused by fatigue, sweat, and so on differ from user to user and from time to time. It is clear that the EMG processing unit should adapt itself to these changes in order to minimize ill-discriminations, the device should "learn" how the user behaves and adjust its internal parameters relative to the operator's variation in real time. Most current prostheses, however, are tuned only in the offline phase. The user learns to reproduce one or two different signals, and the prosthesis is tuned to these signals. When the user cannot control the prosthesis properly, she/he should come back to the assistance center and retune the controller. With such a controller it is not possible to successfully control more than one active DoF, because the differences between tuned signals and actual ones tend to increase gradually with time. Nishikawa and his group proposed a real-time learning method that makes it possible to control up to ten different motions of the forearm from two channels of EMG with a success rate of up to 91.5%. The controller is composed of three modules: the analysis unit, which generates a feature vector containing useful information for discriminating motions from measured EMG signal; the adaptation unit, which makes a mapping function from the feature vector; and the trainer unit, which generates training data from the teacher signal sent by the operator and the feature vector at the moment.

APPLICATIONS

One of the most important aerospace fields of applications is devoted to interfacing a human arm with a powered exoskeleton (orthotic device). As an example, such a type of system was implemented in an elbow joint, naturally controlled by the wearer [4]. The Human–Machine interface was set at the neuromuscular level, by using a neuromuscular signal (EMG) as the primary command signal for the exoskeleton. The EMG signal along with the joint kinematics were fed into a myoprocessor, which in turn predicted the muscle moments on the elbow joint. An exoskeleton is an external structural mechanism whose joints correspond to those of the human body. It is worn by the human and the physical contact between the operator and the exoskeleton allows direct transfer of mechanical power and information signals. The exoskeleton structure under study was a two-link, two-joint mechanism, corresponding to the arm limbs and joints, which was mechanically linked (worn) by the human operator. The operator manipulated an external weight, located at the exoskeleton tip, while feeling a scaled-down version of the load. The remaining external load on the joint was carried by the exoskeleton actuator. The mechanical power of the machine integrated with the inherent human control system could perform tasks that need high forces in a very efficient manner. This is the underlying principle in the design of exoskeleton systems. Experimental tests have shown that synthesizing the processed EMG signals as command signals with the external-load/human-arm moment feedback, significantly improved the mechanical gain of the system, while maintaining natural human control of the system, relative to other control algorithms that used only position or contact forces. The results indicated the feasibility of an EMG-based power

exoskeleton system as an integrated human–machine system using high-level neurological signals. One of the human limits in performing physical tasks is the muscles’ strength, as opposed to strength limitation, humans possess naturally developed algorithms with complex and highly specialized control methods, using higher and lower neural centers, that enable them to perform very complicated tasks such as locomotion while avoiding object collision. In contrast, robotic manipulators can perform tasks requiring large forces or moments, depending on the nature of their structure and on the power of their actuators. However, their artificial control algorithms which govern their dynamics miss the flexibility to perform in a wide rage of fuzzy conditions preserving the same quality of performance as humans. It seems therefore that combining these two entities, the human and the robot into one integrated system under the control of the human, may lead to a solution which will benefit from the advantages offered by each subsystem. The exoskeleton system can be used for three conceptually different applications:
1) power amplifier;
2) master device of a master/slave teleoperator system;
3) haptic device.

In utilizing the exoskeleton as a human power amplifier, the human provides control signals for the exoskeleton, while the exoskeleton actuators provide most of the power necessary for performing the task. The human becomes a part of the system and applies a scaled-down force compared with the load carried by the exoskeleton. Using the exoskeleton as a master device in a master/slave teleoperation system enables the operator attached to the exoskeleton (master) to control a robotic arm (slave). In a bilateral mode, the forces applied on the robotic arm by the environment are reflected back to the master and applied by the exoskeleton structure and actuators on the operator’s arm. In this setup the operator feels the interaction between the robotic arm tool-tip and the environment. Employing the exoskeleton as a haptic device is a relatively new technology aimed to simulate human interaction with virtual object simulated in virtual reality. The operator is immersed in a virtual-realty environment wearing an exoskeleton. In that case a computer simulation is replacing the slave component and the realistic environment of the master/slave teleoperation with a virtual one. As a result, a virtual object in that virtual environment can be touched by the operator, whereas the exoskeleton structure and its actuators provide a force feedback, emulating the real object including its mechanical and texture properties. The exoskeleton, in that sense, simulates an external environment and adds the sense of touch (haptics) to the graphical virtual environment. Several mechanisms including arms, hands, and other haptic devices were developed for a wide range of applications. Throughout the last three decades, several designs of exoskeleton, as a human powered amplifier, have been developed and evaluated. In studying the evolution of these systems two basic types with a different human machine interface (HMI) seem to emerge, which may be defined as generations. During a time interval, the system will gather information regarding the muscle’s neural activation level based on a processed neuromuscular (EMG) signals and the joint position and angular velocity. This information will be fed into a myoprocessor (muscle model), which will in turn predict the moment that is going to be developed by the physiological muscle relative to the joint. The main advantage of establishing the interface at the neuromuscular level is the ability to estimate the forces that will be generated by the muscles before the mechanical contractions actually occur. This information will be fed in to the exoskeleton system such that by the time the physiological muscles contract, the exoskeleton amplifies the joint moment by a preselected gain factor. As a result, the reaction time of the human/machine system should decrease, resulting in a more natural control of the task. In line with this concept, a third generation of exoskeletons is proposed setting the HMI at the human neuromuscular junction. Studies have shown that the surface EMG recorded from bipolar electrodes during constant-force, constant-angle , nonfatiguing conditions can be modeled as a zero-mean correlation-ergodic, random process which is Gaussian distributed. Under those conditions the EMG signal can be correlated to the torque developed by the muscles with respect to a joint. Processed EMG signals along with system identification models which noninvasively estimated muscle forces and joint torques have been used as the control input to myoelectrically controlled prostheses. In operating a myoelectrically controlled powered prostheses, the human neural control system and the prostheses control system are separate entities. The human operator provides command signals in a feed-forward open-loop fashion utilizing only visual feedback as the primary source of information while maintaining a direct line of sight when attempting to grasp or place an object. Feedback information based on visual and auditory cues is slower, less automated, and less programmed than the normal feedback. As opposed to controlling a myoelectrically powered prostheses, in operating a myoelectrically powered exoskeleton (orthotic device) the human and the exoskeleton are mechanically linked, and therefore the human neural control system and exoskeleton control system coexist and have to cooperate by sharing the same kinematics and dynamics constraints. Moreover, when an exoskeleton is used, the nonisometric and nonisotonic conditions, which are valid assumptions for controlling a prosthetic device using EMG signals alone, do not hold, and therefore the muscle’s force cannot be estimated on the EMG signals. This is because the angle of the human limb joint, coupled with the exoskeleton joint, is constantly changing during the exoskeleton operation, and as a result the muscles attached to that joint are changing their length and end points velocities. Therefore, the muscle model (myoprocessor) has to take into account the muscle’s length and velocity in addition to the EMG signal, that defines the muscle activation level, for predicting the force that will be developed by the physiological muscle. In order to establish an interface at the neuromuscular junction, two basic conditions have to be fulfilled. The first condition is the capability to measure the biosignals. The myosignals of the muscles involved in the elbow flexion/extension movement are measured by surface electrodes, using noninvasive techniques. The second condition is the ability to simulate and to predict the functions of the human body subsystems and organs from the interface level (myosignals) down to the lower levels of the physiological hierarchy (skeletal muscle forces and moments) The term myoprocessor was used to define the component of the system that simulates the human skeletal muscle behaviour and provides an estimation of the muscle forces. The main purpose of the powered exoskeleton system is to amplify the load carrying capacity of a healthy operator; however, it can also be used as an upper limb orthosis for physically impaired humans. For a patient to employ any powered exoskeleton, he must have some minimal motor control abilities in order to generate neural signals. The powered exoskeleton improves the patient's limb performance while utilizing what remains from the natural motor control functions of the operator. Thus, instead of promoting muscle atrophy, this powered exoskeleton system could be therapeutic by enhancing further muscle development, due to resurgence of limb use. The human elbow joint complex can be considered as a 2 DOF joint including flexion–extension and pronation–supination joint movements. The exoskeleton, in its current mode, supported only the flexion–extension movement of the elbow joint.

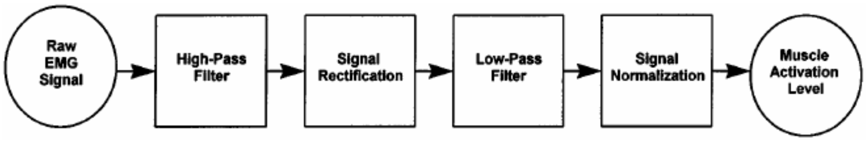


FIG. 3.1.3: Block diagram of an EMG signal processing algorithm for detecting muscle activation level (from [4]).

The algorithm for estimating the normalized muscle activation level (NAL), based on raw EMG signals, follows a signal processing procedure (Fig 3.1.3) which includes:
1) a high-pass filter;
2) full signal rectification (absolute value);
3) a lowpass;
4) a signal normalization with respect to the EMG mean signal during maximal voluntary isometric contraction.
The key element of the myosignal based exoskeleton as a powered assistive device, enabling the HMI at the neural level was the myoprocessor. This module predicted the elbow moments that would be developed by the physiological muscles. This prediction was then used as a primary command signal to the exoskeleton control system, which in turn operated the actuator, mounted on the exoskeleton elbow joint, to add its part of the moment developed at the elbow joint. From the system perspective (Fig 3.1.4), the control algorithm used three sets of feedback information:
1) dynamic feedback—the moments generated at the interfaces between the human arm, the external load, and the exoskeleton structure;
2) kinematic feedback—the elbow joint angle measured by an encoder (the angular velocity and the angular acceleration were calculated by finite differences and filtered by a Butterworth fourth-order digital filter with a cutoff frequency of 10 Hz at 3 dB). These signals were used by the myoprocessor;
3) physiological feedback—the operator used his inherent biosensors and receptors (high level feedback— visualization, low level feedback—muscle spindle, tendon organ, joint receptors).

This physiological feedback was not implemented directly in the exoskeleton control scheme. However, it was taken into consideration by matching the exoskeleton controller frequency bandwidth to the human operator frequency bandwidth.

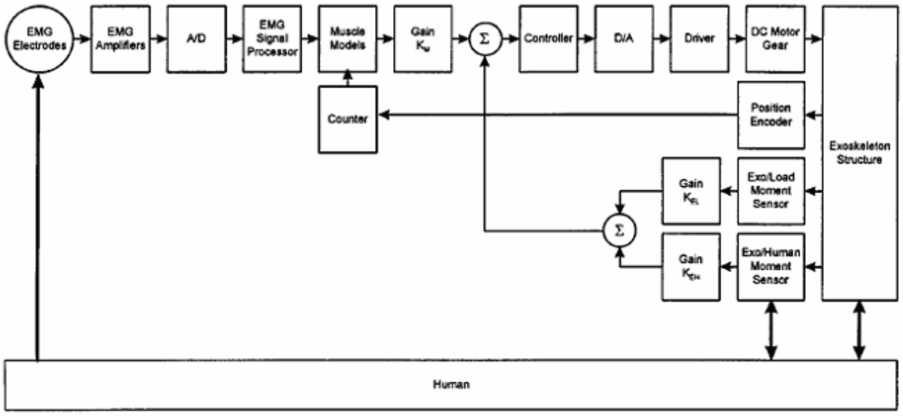


FIG. 3.1.4: Block diagram of a control system of an exoskeleton (from [4])

The EMG signals most probably increased the significance of the command signal leading to improved signal-to-noise ratio (SNR). Improving this ratio allowed to further increase the overall mechanical gain of the exoskeleton system. In addition, from the theoretical point of view, gain and time delays are linked together. Inherent time delays in a system reduce the phase margin and hence the stability will be reduced. By using the EMG signals in conjunction with the myoprocessor the system used parallel processing of the command signal as opposed to the second generation in which the physiological and the mechanical system were processing the command signal in a cascade fashion. Therefore, increasing the system gain is one of the leading advantages of the present concept. Recent studies [5] are developing electromyographic and electroencephalographic methods, which draw control signals for human-computer interfaces from the human nervous system. They have made progress in four areas: a) real-time pattern recognition algorithms for decoding sequences of forearm muscle activity associated with control gestures, b) signal processing strategies for computer interfaces using EEG signals, c) a flexible computation framework for neuroelectric interface research, d) non-contact sensors, which measure EMG or EEG signals without resistive contact to the body. It's defined a system that couples the human nervous system electrically to a computer as a neuroelectric interface: a sensing and processing system that can use signals from the brain or from other parts of the nervous system, such as peripheral nerves, to achieve device control. The focus is on using features from electroencephalograms (EEG) and electromyograms (EMG) as control signals for various tasks, such as aircraft or vehicle simulations and other graphic displays. In order to map EMG signal features to gestures, the model proposes mixtures of Gaussians within a Hidden Markov Model context. This system was tested with many trials over a two-year period in three subjects, who flew and landed high-fidelity simulations of a Boeing F-15 Eagle or a Boeing 757-200 freighter aircraft. Control of both aircrafts was adequate for normal maneuvers. For the 757, a real-time landing sequence under neuroelectric control was filmed at NASA Ames Research Center.

OTHER APPLICATIONS

Most of the studies in the field of EMG based interfaces are particularly devoted to the realization of cybernetic hands. Some of them aim at the development of systems for space applications and other at the terrestrial use of ultra lightweight arms and multi-finger hands on mobile platforms. For many operations, e. g. handing drawers, doors and bayonet closures in an internal lab environment, two finger grippers seem adequate and sufficient; the appropriate mechanical counterparts in the lab equipment are easily designed and realised even in a very late design stage. For more complex tasks however, future space robots need articulated multi-finger hands [6]. Replicating the performance of the human hand is beyond current technical capabilities. In fact, the human hand is extremely complex: it has 22 DoF, controlled by about 38 muscles in the hand (almost twice the number of DoF), and it incorporates about 17,000 tactile units of 4 different types with different receptive fields and different sensitivity to static and dynamic events. Commercial hand prostheses have a limited number of DoF (one or two for finger movements and thumb opposition), and thus they have low grasping functionality. In fact, they do not allow adequate encirclement of objects, compared to the adaptability of the human hand. Moreover, their low compliance leads to instability of the object in the presence of external perturbations, the main advantage of current prosthetic hand devices is that they can generate large grasping forces (>100 N)and are simple to implement and control, in particular by using EMG signal. During the last two decades several robotic and anthropomorphic hands have been developed. All these hands have a high number of DoF (up to 16), and a dexterity comparable to that of the human hand. Some examples of robotic hands are the Utah/MIT hand, the Stanford/JPL hand, the DLR hand, and the Robonaut Hand. Unfortunately, none of these hands can be used as prostheses, because their actuation and control systems are quite heavy and bulky, and thus they cannot be embedded within the hand. Fig 3.1.5 presents a comparison among the human hand, some hand prostheses, and some robotic hands. A design solution that could improve the dexterity of a prosthesis while maintaining intrinsic actuation (i.e., all the actuators embedded within the hand structure) is based on underactuated mechanisms (i.e., a mechanism that has fewer actuators than degrees of freedom). Under-actuated mechanisms allow grasping objects in a way that is closer to human grasping than independent actuation, but their main limitation is the control of this functionality.

	# of DoFs	Size of the hand (normalized)	# of Fingers	# of Sensors	# of Actuators	Opposable Thumb
Human Hand[3]	22	1	5	17'000	38	yes
MARCUS Hand[5]	2	1.1	3	5	2	no
Ottobock/SUVA Hand[7]	1 + 1	1	3	1	1	no
Utah/MIT Hand[78]	16	2 + control	4		32	yes
Stanford/JPL Hand[79]	9	1.2 + control	3		12	no
DLR Hand II[00]	13	1.5 + control	4	64	13	no
Robonaut Hand[81]	12 + 2	Astronaut gloved hand + control	5	43 + tactile	14	yes

FIG. 3.1.5: Comparison of robotic and prosthetic hands (from [2])

Electromyographic signal is a simple and easily obtained source of information on what the users of a prosthesis would like to do with their artificial hands. Surface electrodes are easy to use and manage, and they do not require surgery. Moreover, there are no harnesses that could limit the movement of the forearm. It is possible to control an active device with just one differential electrode placed on the residual limb, even in infants, the technology of EMG signal processing is making steady progress, evolution of the use of the EMG signal to actively control a prosthetic hand is showed in Fıgure 3.1.6.

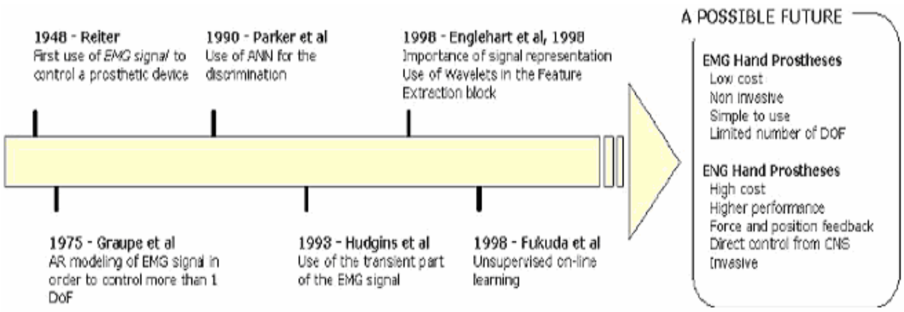


FIG. 3.1.6: Evolution of use of an EMG signal for prosthetic hands (from [2])

Reiter, in 1948, was the first to use the EMG signal to control a simple prosthetic device. Nowadays, all prosthetic devices used in clinical practice have one or two active DoF, directly controlled by a couple of electrodes placed on two antagonist muscles, either in proportional or on/off mode. The use of a larger number of electrodes to control more active DoF has several drawbacks because the coding of movements and the number of electrodes would greatly increment the problems in fabricating and using the socket. In recent years, EMG signals have been largely investigated both for the realization of multifunctional myoelectric prostheses and for the improvement of teleoperation of robotic devices; but as yet all these systems are not capable of successfully controlling a multifunctional hand. The major problem is the time-variant characteristics of the EMG signal, due to physiological changes in the muscles and to the changes in the coupling between skin and the electrodes. An equally important problem is the stochastic nature of the EMG, resulting in parameter estimation errors that, in turn, cause classification and/or control difficulties. Moreover, some control errors are generally introduced by the inability of the patient to reliably generate and reproduce the target contraction signals (operator errors). Some attempts to control multifunctional devices by using more than two electrodes have been made and the use of nerve–muscle graft has been proposed. However, increasing the number of electrodes is not useful in clinical practice, because it introduces additional discomfort in using the prosthesis. With these considerations in mind, two solutions for controlling hand prostheses could be envisaged. On the one hand, EMG-controlled prostheses could represent a “cheap” solution (i.e., low cost and noninvasive) for the restoration (even if partial) of some hand functions. On the other hand, a multifunctional “cybernetic” hand prosthesis with EMG-based control would be a more sophisticated solution. The user can control the grasping task adopting, the processing of the EMG signals, a consolidated technique as shown in new studies [8].

Myographic prostheses can be applied from the disarticulation of wrist up to the level of shoulder disarticulation. The advantages of this type of aid are: high grasping force, high degree of functionality. The factors that instead can advice against of the application are: the independent electromyografic signals are insufficient or not controllable, impossibility to control more functions at the same time, high weight. In some cases, the skeletal structure of the patient can support the weight of the prosthesis, especially in the cases where an electromechanical elbow is also used. Another important frontier in the field of prostheses limb regards the problem of the sensorization and the bio-feedback. In fact the prehensile function is not the only function carried out; with our natural hand we can not only seize an object but also detect the weight, the consistency, the roughness, the temperature. Different types of sensors and transducers are today available and several efforts are devoted to their integration inside prosthetic hands. Bio-feedback is a relevant issue. As an example, with the healthy limb it is possible to seize an object without continuously observing it, thanks to the tactile sensibility. Similarly, with the prosthetic hands of new generation it is possible to seize an object automatically, by simply supplying an opportune command. One of the major difficulties faced by users of prosthetic devices is the great mental effort needed during the first stages of training. In this respect, a mechanism to help patients during the learning stages, without actually having to wear the prosthesis all the time has been developed. The system is based on a real hardware and software for detecting and processing electromyografic (EMG) signal. The association of autoregressive (AR) models and a neural network is used for EMG pattern discrimination. The outputs of the neural network are used to control the movements of a virtual prosthesis, which mimics what the real prosthesis would be doing. This strategy resulted in rates of success of 100% when discriminating EMG signals collected from the upper arm muscle groups. The results show a very easy-to-use system which can greatly reduce the duration of the training stages. Since then a great effort has been applied on the control of artificial limbs for patients with congenital defects or who have lost their limbs in accidents or surgery. Control of such devices necessitates real-time classification of biosignals, e.g., electromyographic signals recorded from intact muscles. Results [9] have shown that a 4-

degrees-of-freedom robotic arm can be controlled in real-time using non-invasive surface EMG signals recorded from the forearm. The innovative features of this system include a physiologically-informed selection of forearm muscles for recording EMG signals, intelligent choice of hand gestures for easy classification, and fast, simple feature extraction from EMG signals. Most commercially available prosthetic devices have limited control (e.g., one degree-of-freedom in the case of a prosthetic gripper).
As a final remark, we stress that other studies in the domain of bioengineering have concentrated on the use of electromyographic signals for control of prosthesis, rehabilitation and computer interfaces for users with motor disabilities. Beyond medical applications, EMG has been proposed for control of computer interfaces. Examples include interfaces for musical expression, controls for consumer electronics [10] and videogames [7].

ADVANTAGES

The electromyogram can be recorded non-invasively using surface electrodes. Electromyographic (EMG) signals provide an extremely useful non-invasive measure of ongoing muscle activity. They could thus be potentially used for controlling robotic devices. Electromyographic signals have significant harmonics in the interval 25-3k Hz and can have amplitudes between approximately 100μV and 90mV, according to the type of signal and of electrodes used. Artefacts due to motion contain very low frequencies that can be effectively filtered without altering the useful EMG information.

DRAWBACKS

One of the major difficulties with EMG-based interfaces can be the great mental effort needed during the first stages of training. When working with myographic prosthesis, that effort increases dramatically [11]. For practical usage the number of EMG channels is limited to two, but the implementation of pattern recognition approaches can potentially lead to a much higher number of control commands [12]. One of the disadvantages of using surface electrodes is that such electrodes can be used only with superficial muscles and are sensitive to the electrical activity of a great area. In the last thirty years, many research efforts have been carried out in the myoelectric control field. Several techniques have been developed to control multifunctional prosthetic devices, and many of them showed promising results. Moreover, these techniques could be also applied in other fields, not only in the control of myoelectric prostheses. For example, algorithms for detecting the activation of muscles are quite useful in gait analysis However, despite all these efforts, EMG signal analysis seems to be quite limited in the number of possible functions that can be restored by using a few electrodes. Moreover, the EMG signal cannot provide any feedback to the user. A possible solution to overcome the limits of the EMG-based approach could be the realization of an interface between the peripheral nervous system (PNS) and the artificial device (i.e., a “natural” neural interface [NI]) to record and stimulate the PNS in a selective way. Recent developments in the technology of electronic implants and in the understanding of nerve functions have made it possible to fabricate selective neural interfaces that work by interchanging information between the nervous system and computerized artificial instruments. A biocompatible neural interface can restore some sensory feedback to the user by stimulating in an appropriate way the afferent nerves and can allow motor control of the prosthesis based on a “natural” EMG-based control. This will be possible by focusing appropriate research efforts on the technological development of the neural interface and on the characterization of the PNS afferent signals in response to mechanical and proprioceptive stimuli. When the user receives sensory feedback from the stimulation of the afferent nerves, and the prosthetic device is controlled directly through the efferent nerves, the user will be able to “feel” the device as a sort of part of the body [2].

3.2 Gaze-tracking based interfaces

Non-verbal communication is often used in social interactions. However, this additional information channel is rarely used in Human-Computer Interaction (HCI) as it requires complex gesture recognition and context relational models. Gaze tracking may propose an interesting compromise as it covers both aspects while being technologically affordable [13]. Gaze reflects our attention, intention and desire. Thus, detection of the gaze direction makes possible to extract such information that is valuable in Human-Computer Interaction. Computers integrated with gaze tracking function must potentially provide an intuitive and effective interactive system. Eye tracking consists in following the eyes movements and computing the gaze direction in order to integrate this information into a computerized system. Research in this field really started at the beginning of the 90s. Since then, the technology became more accurate, less cumbersome, and is today available as commercial products [13]. In parallel, the understanding and the modelling of the gaze behaviour improved, together with an extension of wide application domains: psychiatry, cognitive science, behavioural analysis, medicine, and Human-Computer Interaction. Among various possible applications of gaze tracking system, Human-Computer Interaction is one of the most promising fields [14]. Gaze information plays an important role in identifying a person's focus of attention. The information can provide useful communication cues to a multimodal interface. For example, it can be used to identify where a person is looking and what he is paying attention to. A person's gaze direction is determined by two factors: the orientation of the head, and the orientation of the eyes. While the orientation of the head determines the overall direction of the gaze, the orientation of the eyes determines the exact gaze direction and is limited by the head orientation. The clear vision of an object is possible only when its image falls on the center zone of an ocular portion, called fovea. Figure 3.2.1 shows the structure of the eye. In order to explore a scene it is necessary that the eyes complete the movements that concur to carry and to maintain stable on the fovea the image of interest objects. The ocular movements of a subject, therefore, can tell us exactly where he is watching, what is observing and for how much time. The saccades are the faster ocular movements that the oculomotor apparatus can complete and they have the task to move the visual axis during the exploration of the scene. The fixation is constituted from the pause between two successive saccades and represents the time interval during which the visual information is acquired. The visual exploration therefore is composed of a succession of saccades and fixations. The bidimensional tracing that the eyes complete during the exploration of a scene normally is defined whit the term of "scanpath" [15].

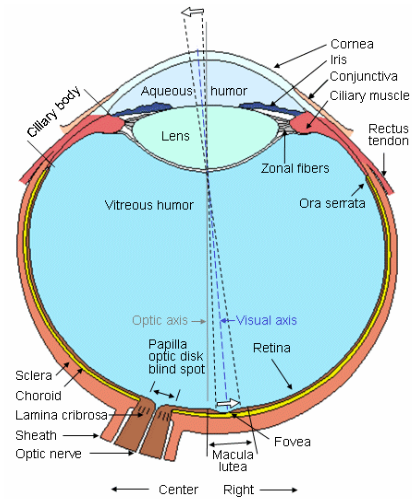


FIG. 3.2.1: Structure of the eye (from [18])

Gaze determines the user's current line of sight or point of fixation. The fixation point is defined as the intersection of the line of sight with the surface of the object (such as the screen) being viewed. Gaze may be used to interpret the user's intention for non-command interactions and to enable (fixation dependent) accommodation and dynamic depth of focus. Numerous techniques have been developed including some commercial eyes trackers. Video-based gaze estimation approaches can be partitioned into head-based approach, ocular-based approach, and the combined head and eye approach. The head based approach determines eye gaze based on the head orientation, in a feature vector to train a neural network to predict the two neck angles, pan and tilt, providing the desired information about head orientation. Gaze estimation by head orientation, however, only provides a global gaze since one's gaze can still vary considerably given the head orientation. Ocular-based approach estimates gaze by establishing the relationship between gaze and the geometric properties of the iris or pupil within the eyes. Specifically, the iris-based gaze estimation approach computes gaze by determining the iris location or shape distortions from its image while pupil-based approach determines gaze based on the relative spatial positions between pupil and the glint (corneal reflection). The most common approach for ocular-based gaze estimation is based on the relative position between pupil and the glint on the cornea of the eye. Assuming a static head, methods based on this idea use the glint as a reference point, thus the vector from the glint to the center of the pupil will describe the gaze direction. While contact-free and non-intrusive, these methods work well only for a static head, which is a rather restrictive constraint on the part of the user [16].

APPLICABILITY OF GAZE-TRACKING TECHNIQUES

Many interactive systems integrated with a gaze tracking system have been proposed. One features task selection by eye. Without using a mouse or keyboard, one can select an icon or a menu item by looking at it for a while. Selection does not necessarily require long dwell time. For example, separating the selection area from the menu enables very fast selection by eye. Integration of a gaze tracking system with other input devices also accelerates selection. Another way of using the eye is to extract the user's attention. Starker et al. proposed a method to detect the user's attention from the user's eye movement pattern while the user is looking at objects on the screen [14]. The potential benefits for incorporating eye movements into the interaction between humans and computers are numerous. For example, knowing the location of a user's gaze may help a computer to interpret a user's request and possibly enable a computer to ascertain some cognitive states of the user, such as confusion or fatigue. In addition, real time monitoring of gaze position permits the introduction of display changes that are contingent on the spatial or temporal characteristics of eye movements. Such methodology is referred to as gaze contingent display paradigm. For example, gaze may be used to determine one's fixation on the screen, which can then used to infer what information the user is interested in. Appropriate actions can then be taken such as increasing the resolution or increasing the size of the region where the user fixates. Another example is to economize on bandwidth by putting high-resolution information only where the user is currently looking. Gaze tracking is therefore important for HCI and intelligent graphics [16]. For several decades now, techniques have existed for tracking the gaze of a person. Until recently, most applications of eye tracking have been in psychological research for probing into subjects' perceptual or cognitive processes, for example when driving in traffic or reading, or for examining the process by which the eye-movements are determined. Especially, research into human visual search strategies has based itself on tracking subjects' gaze during "target object" search tasks, but also evaluation of computer displays has been based upon recordings of fixation patterns. Initially, the techniques were only usable for laboratory experiments: the eye tracking equipment was large and expensive, and required the user's head to be fixed, either in a frame or by using a bite-bar. Some eye tracking techniques using contact lenses are so obtrusive that they cannot be used for extensive periods, and are thus useless outside the laboratory. In these laboratory experiments, data would typically be collected and analysed off-line, i.e., after the subject had finished the experimental task. This state of affairs has changed, though; military research first developed head-up displays (displays integrated in the windscreens of aircraft, so that instrumental data is displayed "on top of" the surrounding flight scene) combined with eye tracking for guiding the missile system, thus freeing the pilot's hands for other tasks. This naturally required on-line processing of the tracking data, and this processing was not aimed at probing the pilot's perceptual or cognitive processes, but rather letting the pilot use the eyes as an extra manipulation channel. Later, the on-line processing of eye-gaze tracking data was extended to user-interfaces for non-military purposes. The techniques are still expensive, but less so now, and the equipment has been reduced in size. Some techniques have been made totally unobtrusive and now allow for (small) head movements, making them usable for enhancing the user-interfaces for disabled people-especially quadriplegics. The use of eye controlled word processors to aid disabled people in their daily life is an important development of eye-gaze tracking technology, our main interest lies in applications for the general community. If this is to have an effect, it must naturally also be based on on-line processing of the tracking data, to provide a real-time interface response to the user's eye movements [17]. The analysis of the ocular movements can be profitably used in study regarding the interaction whit the computer's screen can be one of a specific situation which is developed a visual exploration [15].

Today, several ways of tracking the direction of eye-gaze exist. None of these techniques are perfect in the sense that no single technique fully satisfies all the usability [17]. The ideal tracking device must:

a. Offer an unobstructed field of view with good access to the face and head
b. Make no contact with the subject
c. Meet the practical challenge of being capable of artificially stabilising the retinal image if necessary
d. Possess an *accuracy* of at least one percent or a few minutes of arc; e.g. not give a 10° reading when truly 9°. Accuracy is limited by the cumulative effects of nonlinearity, distortion, noise, lag and other sources of error
e. Offer a *resolution* of 1 minute of arc sec^{-1}, and thus be capable of detecting the smallest changes in eye position; resolution is limited only by instrumental noise
f. Offer a wide dynamic range of one minute to 45° (= 3000-fold) for eye position and one minute arc sec^{-1} to $800 \cdot sec^{-1}$ (= 50,000-fold) for eye velocity
g. Offer good temporal dynamics and speed of response (e.g. good gain and small phase shift to 100Hz, or a good step response).
h. Possess a real-time response (to allow physiological manoeuvers).
i. Measure all three degrees of angular rotation and be insensitive to ocular translation
j. Be easily extended to binocular recording
k. Be compatible with head and body recordings
l. Be easy to use on a variety of subjects

While most people would agree that all these requirements are desirable, we must note that they are not all prerequisites for acceptable eye-gaze tracking interfaces. The intention and attention of a person can be detected with the study of head motion (for the gesture) and gaze direction (for the attention) [17]. If we want to classify the current techniques of today by the way they make contact with the subject, there are basically three types of tracking techniques:

1) Corneal Reflection:

it consists of a measurement of the reflection of a light beam that is shone onto the eye. Typically, infrared light is used to distract the user as little as possible, and to avoid interference from other light sources like lamps. Head should not move, or head pose is measured by magnetic sensor, etc. Eye rotation is detected using IR reflection on the cornea. It's accurate in suitable conditions, but head mounted devices (figure 3.2.2) prohibits natural behaviour.

FIG. 3.2.2: Head mounted device for corneal reflection (from [17])

2) EOG (electro-oculography):

it consists of a measurement of electric potentials from the skin around the eyes. It measures eye movements relative to head position. It uses electrodes placed around the eye, as shown in figure 3.2.3.

FIG. 3.2.3: Example of EOG electrodes placed around the eyes (from [17])

3) Scleral Contact Lenses/Search Coil:

it exploits a search coil embedded into contact lenses (figure 3.2.4) that permits tracking of its position through electromagnetic fields. This method is the most intrusive; the use of a lens requires particular care and causes discomfort. However, it is highly accurate, but has a limited measurement range (almost 5°). It is used to measures eye movements relative to head position.

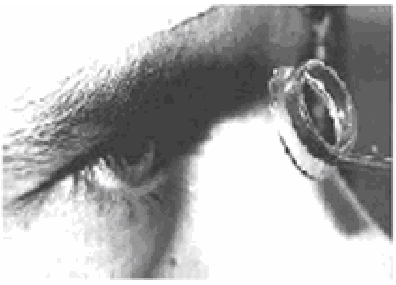


FIG. 3.2.4: Positioning of Scleral Contact Lens containing an induction coil (from [17])

In particular, the principle of the scleral search coil technique is based upon the magnetic induction of a small coil. The induction coil is embedded in a flexible ring of silicone rubber which adheres to the limbus of the human eye concentric with the cornea. Around the head of the subject an alternating horizontal and vertical magnetic field is generated and consequently an alternating voltage will be induced in the coil. After amplification and phase-locked detection two analog voltages are obtained which are proportional to the sine of the horizontal and vertical eye position. In addition to this coil, which is wound in the frontal plane, a second coil is wound in the sagittal plane. This combination coil simultaneously measures horizontal, vertical and torsional eye position. This technique is used for physiological research of the oculomotor system. Its high accuracy and bandwidth guarantees effortless recording of not only saccades, smooth pursuit, vergence, vestibular and optokinetic eye movements but also of miniature eye movements: tremor, drift and microsaccades.

This technique is not realistically applicable for space applications. On the contrary, corneal reflection and EOG are more feasible. These two are separately analyzed below.

3.2.1 Corneal reflection

With this technique a source of infrared light hits the cornea generating the corneal glare that renders the pupil luminous. A video camera resumes the position of the pupil and an opportune software reconstructs the movement completed from the look of the subject during the ocular movements. The systems for this videooculography can be of two types. In some systems the video camera is solidary with the head of the subject, being mounted on a helmet that the subject must wear. In other systems the video camera is solidary with the scene to explore - as an example the screen of the computer - and leaves the subject completely free. As an example, the system for EyeGaze (LC Technologies Inc. , Fairfax, Virginia) consists of a television CCD camera mounted under the screen of the computer and from a monitor that allows the visualization in real time of the eye of the subject and the corneal glare [18].

FIG. 3.2.5: Two examples of equipment for corneal reflection techniques (from [17]).

ADVANTAGES

Techniques based on corneal reflection are highly suitable gaze tracking [15]. They are accurate in best condition [17]. Although electro-oculography (EOG) can potentially provide the gaze direction, computer vision systems, especially with infrared illumination, have much better results. Recent improvements have made the setup more accessible, so that to be adaptable for desktops, large projection screens or head mounted displays. However, large head movements are still impossible, and many systems place the camera on the head [13].

DRAWBACKS

One of the problems with this gaze tracking system is that only local information, i.e. the images of the eyes, is used for estimating the user's gaze. Consequently the system relies on a relatively stable position of the users head with respect to the camera and the user should not rotate his head. Even minor head movement can fail these techniques [13]. This poses a significant hurdle to natural human computer interaction (HCI). Another serious problem with these gaze tracking systems is the need to perform a rather cumbersome calibration process for each individual. In fact, most of existing systems need some kind of personal calibration at the beginning of the measurement. The reasons for this are:

• Individual differences of eyeball size. Among adult persons, there is about 10 % individual difference in the radius of eyeball.

• Difficulty in measuring the position of the fovea. The fovea is the highest resolution area on retina, and the human visual axis can be defined as a vector from fovea to the center of the crystalline lens. On the other hand, existing gaze tracking techniques measure the estimated gaze direction from the center of cornea curvature and the center of the pupil, instead of the real visual axis. The estimated direction is then corrected by the personal calibration.

The latest research efforts are aimed at overcoming this limitation. Researchers from NTT in Japan proposed a new gaze tracking technique based on modelling the eyeball. Their technique significantly simplifies the gaze calibration procedure, requiring only 2 points to perform the necessary calibration. The method, however, still requires relatively stationary head, and there exists difficulty in acquiring accurate geometric eyeball model for each subject. IBM is also studying the feasibility of completely eliminating the need of gaze calibration procedure by using two cameras and by exploiting the geometry of eyes and their images. Other recent efforts also focus on improving the eye tracking robustness under various lighting conditions. To make the gaze tracking system more robust to user movement, it would be helpful to also use additional information such as the 3D position of the head relative to the camera to estimate the users gaze. Indeed, there are many drawbacks that have to be avoided when using gaze for interaction [13]:

- The "Midas touch": what the user is looking at is not necessarily what he wants to interact with. Although some studies try to integrate statistical analysis to better recognize the user's focus of attention [8], using exclusively eyes is not natural.
- Fatigue: voluntary and precise control of the gaze is tiring.
- Perpetual motion: while voluntary eye saccades (1 to 40° of the visual angle) corresponds to the visual search, micro-saccades (< 1°) still occur when the eye is focussing on a target. Although we know that the vision is suspended during saccades or eye blink, only few systems are able to really distinguish the fixation phases.

Gaze input is more appropriate for multimodal interaction. The combination with speech is very convenient as both can be well synchronized.

3.2.2 Electro-Oculogram (EOG)

The cornea of the eye is electrically positive relative to the back of the eye, as observed from Emil du Bois-Reymond (1848). Since this potential was not affected by the presence or absence of light, it was thought of as a resting potential. Actually it is not constant but slowly varying and is the basis for the *electro-oculogram* (EOG). This source behaves as if it were a single dipole oriented from the retina to the cornea. Such corneoretinal potentials are well established and are in the range of 0.4 - 1.0 mV. Eye movements thus produce a moving (rotating) dipole source and, accordingly, signals that are a measure of the movement may be obtained. The chief application of the EOG is in the measurement of eye movement [18]. Figure 3.2.6 illustrates the measurement of horizontal eye movements by the placement of a pair of electrodes at the outside of the left and right eye. With the eye at rest the electrodes are effectively at the same potential and no voltage is recorded. The rotation of the eye to the right results in a difference of potential, with the electrode in the direction of movement (i.e., the right canthus) becoming positive relative to the second electrode. Ideally, the difference in potential should be proportional to the sine of the angle. The opposite effect results from a rotation to the left, as illustrated. The calibration of the signal may be achieved by having the subject looking consecutively at two different fixation points located a known angle apart and recording the concomitant EOGs. Typical achievable accuracy is ±2° and maximum rotation is ±70° however, linearity becomes progressively worse for angles beyond 30°. Typical signal magnitudes range from 5-20 µV/°. Recording as of both horizontal and vertical ocular movements are possible but not those of torsion on the antero-posterior axis. In fact, these do not modify the direction of the dipole and they do not determine therefore potential differences on the derivation electrodes. The electrodes are located around to the eye: one superior and one inferior, in order to record the vertical movements; one external and one internal, in order to record the horizontal movements. Figure 3.2.6 shows the measurement of horizontal eye movement by means of a pair of electrodes.

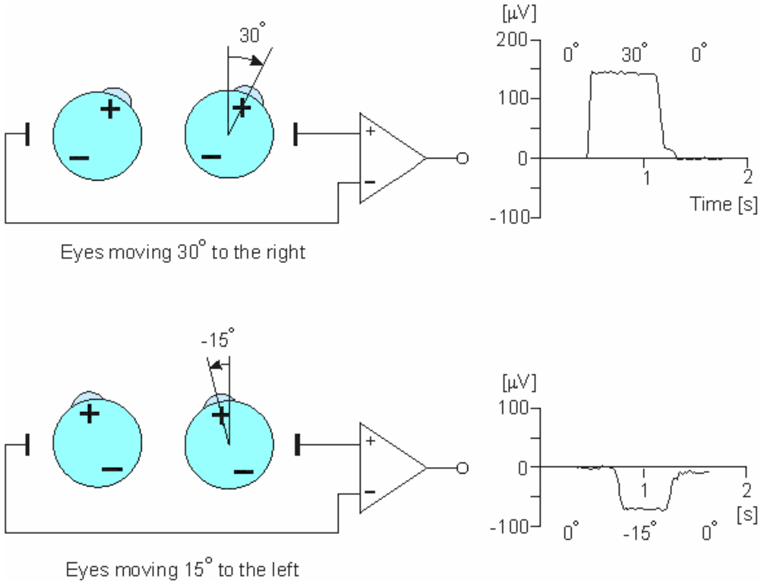


FIG. 3.2.6: Illustration of EOG signals generated by horizontal movements of the eyes (from [18]).

Is it possible to distinguish two subdivisions of the electrooculography: *saccadic response* and *nystagmography*. They are described below:

Saccadic Response: Saccadic movements describe quick jumps of the eye from one fixation point to another. The speed may be 20 – 700 °/s. *Smooth* movements are slow, broad rotations of the eye that enable it to maintain fixation on an object moving with respect to the head. The angular motion is in the range of 1 - 30°/s. The adjective *pursuit* is added if only the eye is moving, and *compensatory* if the eye motion is elicited by body and/or head movement. The aforementioned eye movements are normally *conjugate*, that is they involve parallel motion of the right and left eye. A normal saccadic response to a rapidly moving target is described in Figure 3.2.7. The stimulus movement is described here as a step, and eye movement speeds of 700°/s are not uncommon. The object of the oculomotor system in a saccade is to rapidly move the sight to a new visual object in a way that minimizes the transfer time. The parameters commonly employed in the analysis of saccadic performance are the maximum angular velocity, amplitude, duration, and latency. The trajectory and velocity of saccades cannot voluntarily be altered. Typical values of these parameters are 400°/s for the maximum velocity, 20° for the amplitude, 80 ms for the duration. When following a target moving in stepwise jumps, the eyes normally accelerate rapidly, reaching the maximum velocity about midway to the target. When making large saccades (≥25°), the eyes reach the maximum velocity earlier, and then have a prolonged deceleration. The movement of the eyes usually undershoots the target and requires another small saccade to reach it. Overshooting of the target is uncommon in normal subjects. Normally the duration and amplitude are approximately linearly correlated to each other. Several factors such as fatigue, diseases, drugs, and alcohol influence saccades as well as other eye movements' latency. Figure 3.2.7 illustrates eye movement in response to a step stimulus. Following a latency, the eye rapidly moves towards the new position, undershoots and moves a second time.

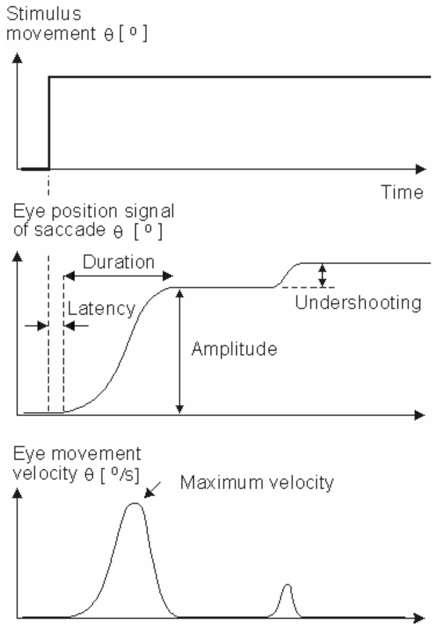


FIG. 3.2.7: Illustration of the eye movement response to a step light, whose horizontal position istantaneously shifts (from [18]).

Nystagmography: Nystagmography refers to the behavior of the visual control system when both vestibular (balance) and visual stimuli exist. Nystagmoid movement is applied to a general class of unstable eye movements, and includes both smooth and saccadic contributions. Based on the origin of the nystagmoid movement, it is possible to separate it into *vestibular* and optokinetic nystagmus. Despite their different physiological origin, these signals do not differ largely from each other.

Vestibular Nystagmus: Nystagmography is a useful tool in the clinical investigation of the vestibular system (Stockwell, 1988). The vestibular system senses head motion from the signals generated by receptors located in the labyrinths of the inner ear. Under normal conditions the oculomotor system uses vestibular input to move the eyes to compensate for head and body motion. This can occur with saccadic and/or pursuit motion. If the vestibular system is damaged then the signals sent to the oculomotor system will be in error and the confusion experienced by the patient results in dizziness. Conversely, for a patient who complains of dizziness, an examination of the eye movements arising from vestibular stimuli can help identify whether, in fact, the dizziness is due to vestibular damage. Inappropriate compensatory eye movements can easily be recognized by a trained clinician. Such an examination must be made in the absence of visual fixation (since the latter suppresses vestibular eye movements) and is usually carried out in darkness or with the patient's eye closed. Consequently, monitoring eye movement by EOG is the method of choice.

Optokinetic Nystagmus: Another example of nystagmoid movement is where the subject is stationary but the target is in rapid motion. The oculomotor system endeavors to keep the image of the target focused at the retinal fovea. When the target can no longer be tracked, a saccadic reflex returns the eye to a new target. The movements of the eye describe a sawtooth pattern. This is described as optokinetic nystagmus. This may also be provoked in laboratory by rotating a cylinder with dark stripes on a light background in front of a person's eyes.

APPLICATIONS

Eye tracking systems for uses with personal computers are under development [19]. The system is intended to provide a pointing device that could be useful for people with physical disabilities. The basis for this system is the use of Bio-Electrical signals from the user's body. In particular the use of the Electrooculogram and Visual Evoked Potentials has been investigated. Experiments have compared two algorithms for processing the signals and generate an effective output control. In theory, the EOG potential varies linearly with the rotation of the eye in its socket. In practice, the signal is corrupted by various sources, resulting in a baseline "drift" that obscures the eye movement signal. In the mentioned work [19] the authors have developed and tested two techniques for detecting and removing this drift from the EOG. The first uses an adaptive fuzzy-logic system to detect the difference between drift and eye movement. The second adds a detection scheme based upon detecting a visual evoked response (VEP) in the electroencephalogram (EEG) of the subject. Bioelectrical data collection and pattern recognition are performed by a real-time digital signal processor (DSP) system linked to a host computer (PC). The DSP system is configured with four channels used to measure EOG (vertical and horizontal for each eye) and one channel used to measure VEPs from over the occipital area, with the reference taken from the forehead for all channels.

ADVANTAGES

The advantages of the EOG include recording with minimal interference with subject activities and minimal discomfort. Furthermore, it is a method where recordings may be made in total darkness and/or with the eyes closed. Today the recording of the EOG is a routinely applied as a diagnostic method for investigating the human oculomotor system. The application of digital computers has considerably increased the diagnostic power of this method.

DRAWBACKS

Although both horizontal and vertical ocular movements can be recorded, no movements of torsion on the antero-posterior axis can be detected. In fact, these do not modify the dipole and they do not determine therefore potential differences them on the derivation electrodes [19]. The most important disadvantages relate to the fact that the corneoretinal potential is not fixed but has been found to vary diurnally, and to be affected by light, fatigue, and other qualities. Consequently, there is a need for frequent calibration and recalibration. Additional difficulties arise owing to muscle artifacts and the basic nonlinearity of the technique [18].

3.3 Motion-capture and gesture-recognition based interfaces

Man-machine interfaces based on motion capture and gesture recognition deserve a considerable attention. Four types of gesture and motion tracking techniques are separately described below:

3.3.1 Ultrasound trackers:

Ultrasound trackers have three components: a transmitter, a receiver and an electronic unit. The transmitter is a set of three ultrasonic speakers mounted on a fixed triangle frame. The receiver is a set of three microphones, mounted on a smaller triangular frame. The receiver is attached to the body part (head or hand) which needs to be tracked. For a given temperature the speed of sound is known and can be used to measure distances between the speakers and the microphones. A total of nine distances are measured in order to determine the position and orientation of the plane. The limitations to ultrasonic are low resolution, long lag times and interference from echoes and other noises in the environment. Two companies that provide ultrasonic tracking systems are Logitech and Transition State.

3.3.2 Magnetic trackers:

Magnetic trackers employ alternating low-frequency fields to determine the moving object's position and orientation. The low-frequency field is generated by a transmitter, which is an assembly of three stationary orthogonal antennas. A second set of orthogonal antennas is placed inside a receiver. The signal received is processed to determine the position and orientation of the receiver in relation to the transmitter. Limitations of these trackers are a high latency for the measurement and processing, range limitations, and interference from ferrous materials within the fields. The two primary companies selling magnetic trackers are Polhemus and Ascension.

3.3.3 Optical trackers

Several optical position tracking systems have been developed. One method uses a ceiling grid LEDs (light emitting diodes) and one or more cameras. The LEDs are pulsed in sequence and the cameras image is processed to detect the flashes. Two problems with this method are limited space (grid size) and lack of full motion (rotations). Another optical method uses a number of video cameras to capture simultaneous images that are correlated by high-speed computers to track objects. Processing time (and cost of fast computers) is a major limiting factor.

EXAMPLES:
Image-based tracking of the human hand is considered as a specific and very demanding problem. The relatively great amount of degrees of freedom and the complex kinematic structure make this task challenging. Approaches for human hand tracking are based on silhouette projections, information about movement or colored markers. Some applications aim at entire 3D reconstruction with complex hand models, but a real-time solution to this problem is hard to achieve. In most cases, no information about finger postures is used. Computation rather relies on optical flow, correlation of stereo color images or skin color segmentation. Frequently, camera images are used in order to recognize and classify static hand gestures. After preprocessing the images, tensors of gesture patterns are compared with the transformed actual images similar to an eigenspace method. A lot of interactive robot systems make use of both speech input and gesture and/or object recognition, e.g. as in to give instructions to the robot more naturally, or to clarify ambiguities in the input. Currently, facial expressions are being evaluated as a bidirectional information channel between humans and robots as well. However, those approaches focus on emotional human robot interaction and do not concentrate on commanding a robot assistant for manipulation tasks. Skin colour is used as the basic feature for a typical hand tracking. For recognizing gestures, it is necessary to focus the user's hand. Therefore, the camera follows hand movements. For the gesture classification, since hand contours are different in length and depend on rotational angle, translation and scale, a Fourier description of the silhouette is usually computed taking samples at equidistant points all over the outline (Fig 3.3.1).

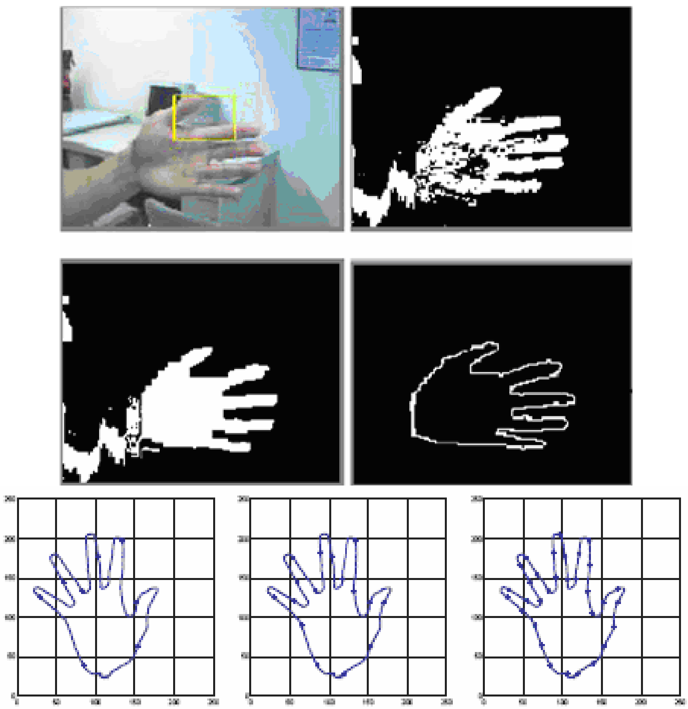


FIG. 3.3.1: gesture pre-processing and contour whit 8, 16 and 32 point (from [20])

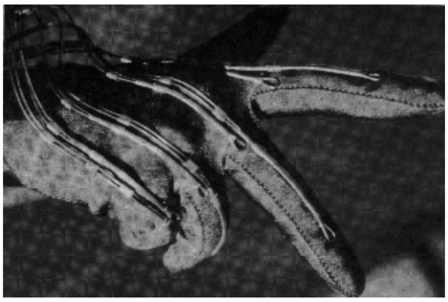
3.3.4 Wearable electromechanical trackers:

Wearable systems that integrate displacement or deformation sensors codify gesture into an appropriate electric response.

EXAMPLES:

a) DataGlove: this glove, developed by VPL Research, is made of lightweight lycra (Fig 3.3.2.a). The system that measures finger and hand movement consists of a position/orientation sensor and a set of coated fiber-optic cables that run along the back of each finger. The fibre optic cables running along the fingers are sectioned according to the joints of the hand. One end of each fibre optic cable is equipped with a light emitting diode (LED) and the opposite end is connected to a photosensor. The amount of light detected passing through the fibre is proportional to the degree to which the corresponding joint is bent. This information is sent to the system, which determines which fingers are being bent and by how much. The second measurement device incorporated to the system is a tracking mechanism that uses magnetic detection to determine the position and orientation of the hand position in relation to the whole scene [21].

b) 5TH GLOVE: this is a commercial product by iREALITY.COM, INC. It is equipped with fiber-optic flex sensors to generate finger-bend data (Fig 3.3.2.b) [22].



a)

b)

FIG. 3.3.2: a) DataGlove (from[2]) and b) 5HT GLOVE (from [22])

c) PowerGlove: this glove is partially derived from the DataGlove with the expensive fibre optic system replaced by flat plastic strain gauges. A small strip of plastic is coated with an electrically conductive ink and placed along the length of each finger. A small electrical current passing through the ink remains stable until a finger is bent. The computer can measure the change in the ink's electrical resistance and compute the finger position. The PowerGlove also uses two ultrasonic trackers to track the position and the orientation of the hand [23].

FIG. 3.3.3: PowerGlove (from [22])

d) CyberGlove: the CyberGlove from Virtual Technologies uses extremely thin strain gauges enclosed in the glove's material to measure how much the finger is bent. The CyberGlove uses up to 22 sensors to measure joint angles [23].

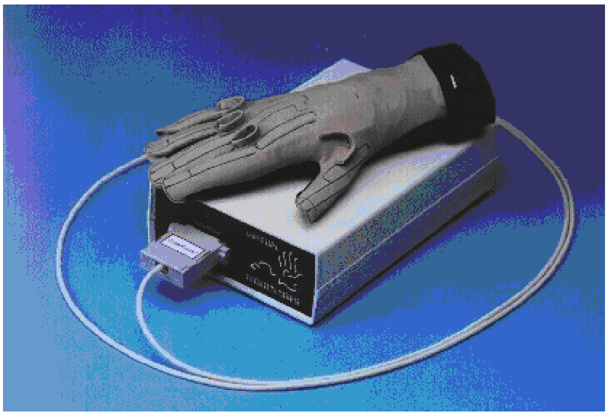


FIG. 3.3.4: CyberGlove (from[23])

e) Pinch Glove: the Pinch glove system does not measure finger joint angles. Instead, gloves are worn on both hands and contact between any two or more fingers completes a conductive path, allowing the definition of a variety of "pinch" gestures, which an application developer can map actions against. Over 1,000 gestures are theoretically possible. The gloves are constructed of a stretchable fabric and contain an electrical sensor in each fingertip. Each glove has a back-of-hand mount to accommodate a spatial tracker. The user's point of interaction in the virtual environment is represented by a 3-D cursor [23].

FIG. 3.3.5: Pinch Glove (from [23])

f) Sensorized garments based on piezoresistive polymer sensors

In order to endow garments with strain sensing capabilities for the monitoring of body-kinematics, such as position and movement of articulation segments, piezoresistive strain sensors (also known as strain gages) can be advantageously used.

In recent years, an innovative technology for wearable sensors based on electroactive polymers (EAP) has been developed by University of Pisa, Interdepartmental Research Centre 'E. Piaggio'. In order to confer strain sensing properties to garments, two types of piezoresistive polymer materials have been integrated into elastic fabrics: π-electron conjugated conducting polymers and carbon loaded elastomers.

Conducting polymers are a class of organic materials able to transport electricity. Since their discovery in the early seventies, several conducting polymers, such as polypyrrole, polyaniline and polythiophene, have been synthesised and their quality has been continuously improved [23, 24].

The research at University of Pisa has investigated their use for the development of organic piezoresistive sensors. To this aim, conducting polymer sensors have been fabricated by epitaxial deposition of a thin layer of polypyrrole (PPy) on a Lycra®/cotton fabric [25-28]. These sensors, showing both piezoresistive and thermoresistive properties [25,28], have been used to sensorise a glove and a leotard [27]. Similarly, CSIRO and University of Wollongong has integrated this kind of fabric based sensors into a knee sleeve [29]. The quasi-static characterisation of PPy-coated fabrics has indicated an average gage factor (GF) of about –12 [28,30,31]. Furthermore they have shown a temperature coefficient of resistance (TCR) of about 0.018 $°C^{-1}$ [28,30,31]. Despite the fact that the reported high absolute value of GF is suitable for strain gage implementations, two serious problems typically affect PPy-coated fabric sensors. The first one resides in the strong variation with time of their resistance (chemical instability) [28]. The second problem is represented by the high response time: following the application of a sudden mechanical stimulus, the sensor resistance reaches a steady state in a few minutes, strictly limiting the applicability of such devices [9]. Moreover, conducting polymer based sensors are not easily amenable to textile technology. In fact, they would require the problematic insertion in textile processes of tanks for the material synthesis and deposition/coating on fabrics. A new generation of high-performance strain sensors has been obtained by coating yarns and fabrics with carbon loaded elastomers (CLR), typically consisting of a silicone matrix filled with carbon black powder. Sensors are fabricated on a Lycra®/cotton textile by masked smearing of the conducting mixture. The same polymer/conductor composite is also used as material for the tracks of connection between sensors and an acquisition electronic unit, avoiding the stiffness of conventional metal wires. A GF of about 2.5, quite similar to those of metals, has been measured for CLR-coated fabrics, making this kind of device suitable for high-performance sensing [30, 31]. Moreover, CLR sensors show, as for PPy sensors, thermoresistive properties, with reported TCR values of about 0.08 $°C^{-1}$ [30, 31]. Such devices have been used to demonstrate prototype sensorised garments, including gloves, leotards, knee bands and sleeves (fig.3.3.6) [35].

FIG. 3.3.6: Sensorised glove, leotard and sleeve developed by University of Pisa, Interdepartmental Research Centre 'E. Piaggio'

All these systems are truly wearable fabrics, incorporating compliant polymer sensors, capable of recording body posture and gesture, with no discomfort for the subject wearing the garment and with negligible motion artifacts.

3.4 Speech recognition based interfaces

The goal of an automatic speech recognition system is to deduce meaningful linguistic units (i.e., words) from acoustic waveforms. Due to the random nature of the process and interferences, it is not possible to derive a deterministic formulation that provides a mapping between acoustic signal and conceptual meanings. Instead the problem is generally formulated in a probabilistic framework. In probabilistic setting, the speech recognition is stated as the estimation of 'most probable' linguistic representation of a 'given' acoustic waveform [38]. The mathematical formulation of this problem is:

$$\hat{W} = \arg\max_{W} P(W/O) \qquad (3.4.1)$$

,where O is a set of observations from the acoustic waveform and W is a random variable that takes its values from the possible linguistic representations in the language under consideration. P(W/O) is the conditional probability distribution of the linguistic representation given the observations. This conditional distribution constitutes the knowledge base of the recognizer. This knowledge base is constructed using statistical learning techniques and a priori expertise on speech production mechanisms. The role of a priori expertise on the domain is to provide a set of simplification assumptions that will guide the statistical machinery to extract relevant information for recognition. Because of the complexity of the speech production mechanisms there is no simple parametric representation of P(W/O) that involves both acoustic and linguistic information. The basic approach is to first divide the problem into acoustic and linguistic components that can be handled separately. This is achieved using a Bayesian reformulation:

$$\hat{W} = \arg\max_{W} P(O/W)P(W) \qquad (3.4.2)$$

In this formulation, the acoustic model, P(O/W), encodes the statistical distribution of speech acoustics given the linguistic labeling. P(W) is the probability assigned by the language model which encodes the a priori linguistic information. This approach is called the source-channel model. P(W) constitutes the language source with all the linguistic constraints of the underlying language and P(O/W) is the acoustic channel that outputs the speech signal based on the linguistic unit W. With this formulation the two components of the system can be constructed separately later to be combined in the recognition phase. The information source for the linguistic component is written text documents, supposed to be sufficient to represent the properties of language. The statistical approach consists of extraction of relevant information content from data and formulation of a parametric representation that is capable to encode this content. State-of-the-art automatic speech recognition (ASR) systems are based on probabilistic modelling of the speech signal using Hidden Markov Models (HMM). The goal of decoding process is to determine a sequence of states that the observed signal has gone through. There are three main problems: the evaluation problem; the decoding problem; the learning problem.

Recently, the acoustic modelling problem in speech recognition was reformulated within the probabilistic graphical models (PGM) formalism. PGM is a unifying framework for statistical learning which provides an abstraction of quantitative and qualitative components of a statistical model. Dynamic Bayesian networks (DBN) are a subset of PGM which are defined on directed acyclic graph structures. These models are defined with graph structures that encode the probabilistic relations between its variables through a set of associated conditional probabilities. One of the main advantages of PGM is the graphical abstraction that provides a visual understanding of the modeled process. Moreover they provide a powerful setting to specify efficient inference algorithms that can be specified automatically once the initial structure of the graph is determined. In feature extraction stage, the speech signal is considered as a quasi-stationary process consisting of consecutive frames that can be treated independently. The goal of front-end speech processing in ASR is to attain a projection of the speech signal to a compact parameter space where the information related to speech content can be extracted easily. Most parameterization schemes are developed based on the source-filter model of speech production mechanism. In this model, speech signal is considered as the output of a filter (vocal tract) whose input source is either glottal air pulses or random noise. For voiced sounds the glottal excitation is considered as a slowly varying periodic signal. This signal can be considered as the output of a glottal pulse filter feed with a periodic impulse train. For unvoiced sounds the excitation signal is considered as random noise. State of the art speech feature extraction schemes (Mel frequency cepstral coefficients (MFCC) and perceptual linear prediction (PLP)) are based on auditory processing on the spectrum of speech signal and cepstral representation of the resulting features. The spectral and cepstral analysis is generally performed using Fourier transform. The advantage of Fourier transform is that it possesses very good frequency localization properties. Feature extraction methods can be categorized into overlapping classes that share a number of common ideas. The most common ideas are related to filterbank processing, features inspired by the physiology of the auditory system, features utilizing perceptual knowledge, or inspired by phenomena that occur during speech production (e.g. modulations). A review of the proposed features for ASR systems indicates that cepstral analysis features have become one of the most common approaches. A popular alternative is the *PLP* or related features that are based on knowledge of the human auditory peripheral system. Finally, nonlinear speech processing techniques (e.g. modulations, fractals) have started to gain momentum. Many techniques share the concept of short-time processing. However, recently there have been introduced alternative methods - e.g. *RASTA*, *TRAP* - that filter out parts of the modulation spectrum or process frames that span longer time intervals. ASR is commonly described as converting speech to text. The reverse process, in which text is converted to speech (TTS), is known as speech synthesis. Speech synthesizers often produce results that are not very natural sounding. Speech synthesis is different from voice processing, which involves digitizing, compressing (not always), recording, and then playing back snippets of speech. Voice processing results are very natural sounding, but the technology is limited inflexibility and is disk storage-space-intensive compared to speech synthesis. Speech recognition developers are still searching for the perfect HMI, a recognition engine which understands any speaker, interprets natural speech patterns, remains impervious to background noise, and has an infinite vocabulary with contextual understanding. However, practical product designers can indeed use today’s speech recognition engines to make major improvements to today’s markets and applications. Selecting such an engine for any product requires understanding how the speech technologies impact performance and cost factors, and how these factors fit in with the intended application [41]. Speech recognition is more complicated than speech synthesis. However, it too can be thought of as having a front end and a back end. The front end processes the audio stream, isolating segments of sound that are probably speech and converting them into a series of numeric values that characterize the vocal sounds in the signal. The back end is a specialized search engine that takes the output produced by the front end and searches across three databases: an acoustic model, a lexicon, and a language model. The acoustic model represents the acoustic sounds of a language, and can be trained to recognize the characteristics of a particular user's speech patterns and acoustic environments. The lexicon lists a large number of the words in the language, along with information on how to pronounce each word. The language model represents the ways in which the words of a language are combined. Neither of these models is trivial. It's impossible to specify exactly what speech sounds like. Human speech rarely follows strict and formal grammar rules that can be easily defined. An indispensable factor in producing good models is the acquisition of very large volumes of representative data. An equally important factor is the sophistication of the techniques used to analyse that data to produce the actual models. Of course, no word has ever been said exactly the same way twice, so the recogniser is never going to find an exact match. And for any given segment of sound, there are very many things the speaker could potentially be saying. The quality of a recogniser is determined by how good it is at refining its search, eliminating the poor matches, and selecting the more likely matches. A recogniser's accuracy relies on it having good language and acoustic models, and good algorithms both for processing sound and for searching across the models. The better the models and algorithms, the fewer the errors that are made, and the quicker the results are found. Needless to say, this is a difficult technology to get right [43]. Automatic speech recognition can be viewed as a mapping from a continuous-time signal, the speech signal, to a sequence of discrete entities-for example, phonemes (or speech sounds), words, and sentences. The major obstacle to high accuracy recognition is the large variability in the speech signal characteristics. This variability has three main components: linguistic variability, speaker variability, and channel variability. Linguistic variability includes the effects of phonetics, phonology, syntax, semantics, and discourse on the speech signal. Speaker variability includes intra- and inter-speaker variability, including the effects of coarticulation - that is, the effects of neighboring sounds on the acoustic realization of a particular phoneme due to continuity and motion constraints on the human articulatory apparatus. Channel variability includes the effects of background noise and the transmission channel (e.g., microphone, telephone, reverberation). All these variabilities tend to shroud the intended message with layers of uncertainty, which must be unraveled by the recognition process (Figure 3.4.1).

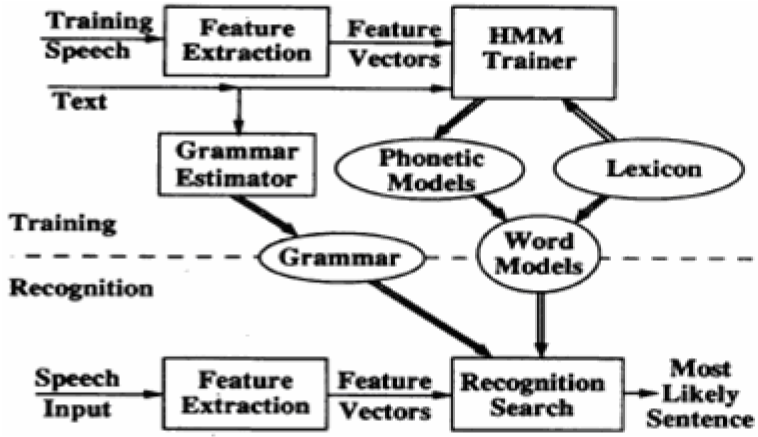


FIG. 3.4.1: Training and recognition process of a speech based interface (from [43])

The first stage in the continuous-to-discrete mapping that is required for recognition is performed by the analysis or feature extraction. Typically, the analysis consists of estimation of the short-term spectrum of the speech signal over a frame (window) of about 20 ms. The spectral computation is then updated about every 10 ms, which corresponds to a frame rate of 100 frames per second. This completes the initial discretisation in time. So, there is the need to discretise the spectrum into one of a finite set of spectra. Given a computed spectrum for a frame of speech, one can find the template in the codebook that is "closest" to that spectrum, using a process known as vector quantisation. In both training and recognition the first step in the process is to perform feature extraction on the speech signal.

As an observation, it worth noting here that even **sub-vocal recognition systems** are currently under development. They use wearable sensors to collect nerve signals transmitted from the brain to the vocal cords when the subject ‘reads silently to himself’. The sensors detect the nerve signals that generate this sub-vocal speech and relay those to a computer program. Applications of this technology include improved voice recognition systems, systems allowing the transmission of vocal commands in noisy environments.

APPLICATIONS

Results have been reported about a research project aimed at exploring ways to command an industrial robot using the human voice [41]. This feature can be interesting with several industrial, laboratory and clean-room applications, where a close cooperation between robots and humans is desirable. A demonstration was presented using two industrial robots and a personal computer (PC) equipped with a sound board and a headset microphone. The demonstration was coded using the Microsoft Visual Basic and C#NET 2003 and associated with two simple robot applications: one capable of picking-and-placing objects and going to predefined positions, and the other capable of performing a simple linear weld on a work-piece. The speech recognition grammar was specified using the grammar builder from the Microsoft Speech SDK 5.1. Moreover Speech interfaces have the potential to address the data entry bottleneck of many applications in the field of medical informatics [42]. Data entry has been identified as a key bottleneck in many biomedical applications. Large volumes of information must be gathered by clinicians and researchers to support patient care and clinical trials. This data must be collected and managed according to specific protocols. Often the situation exists where a clinician is occupied with patient care and cannot document his or her findings until later. This interval of time between the generation of information and its recording can compromise the data collection process. Despite considerable advances in computer architectures over the last 20 years, the keyboard and video display remain the principal means of entering and retrieving data. New human-computer interface modalities are needed which can automate the data collection process at the *source*, where the information is actually generated. Speech-driven computer interfaces can address two key concerns in biomedical computer interfaces: the demand for ease of use and constraints on the user's ability to work with the keyboard or mouse. Speech technology is still limited, however, with most successful systems using medium-sized vocabularies with well-defined grammar rules. As described in the literature, the main applications of speech include template-based reporting, natural language processing, multimodal integration of speech with other methods of input, and hands-busy data entry. The first two reflect the need for more intuitive interfaces. The latter two deal with limitations of traditional input using the keyboard or mouse. Template-based reporting has been applied to radiology, pathology, endoscopy, and emergency medicine. The potential advantage is that turnaround time is decreased and accuracy is increased by eliminating the need for dictation and transcription by clerical personnel. An alternative to template-based reporting explored methods that circumvent shortcomings in the current technology while maintaining the flexibility and naturalness of speech. The reduction of speech errors is typically viewed as a technical problem. Another study reported a reduction in spoken disfluencies by using more structured interfaces. A significant positive correlation was also observed between the increased acceptance and decreased diagnosis errors including domain knowledge into the user interface would be advantageous. Moreover speech interfaces play a very important role also in the manufacturing systems efficiency will increase if the interface is more natural or similar to the human way of commanding things [41]. Industrial-manufacturing systems would benefit very much from speech recognition for human machine interface (HMI) even if the technology is not so advanced. Gains in terms of autonomy, efficiency and agility seem evident. The modern world requires better products at lower prices, requiring even more efficient manufacturing plants because the focus is in achieving better quality products, using faster and cheaper procedures. This means autonomy, having systems that require less operator intervention to operate normally, better HMIs and cooperation between humans and machines sharing the same workspace as real co-workers. The final objective is to achieve, in some cases, semiautonomous systems, i.e. highly automated systems that require only minor operator intervention. In many industries, production is closed tracked in any part of the manufacturing cycle, which is composed by several in-line manufacturing systems that perform the necessary operations, transforming the raw materials in a final product. In many cases, if properly designed, those individual manufacturing systems require simple parameterization to execute the tasks they are designed to execute. If that parameterization can be commanded remotely by automatic means from where it is available, then the system becomes almost autonomous in the sense that operator intervention is reduced to the minimum and essentially related with small adjustments, error and maintenance situations. In other cases, a close cooperation between humans and machines is desirable although very difficult to achieve, due to limitations of the actual robotic and automation systems [43]. In the past decade, tremendous advances in the state of the art of automatic speech recognition by machine have taken place. A reduction in the word error rate by more than a factor of five and an increase in recognition speeds by several orders of magnitude (brought about by a combination of faster recognition search algorithms and more powerful computers), have combined to make high-accuracy, speaker independent, continuous speech recognition for large vocabularies possible in real time, on off-the-shelf workstations, without the aid of special hardware. These advances promise to make speech recognition technology readily available to the general public. As is often the case in technology, a paradigm shift occurs when several developments converge to make a new capability possible. In the case of continuous speech recognition, the following advances have converged to make the new technology possible: higher-accuracy continuous speech recognition, based on better speech modelling techniques; better recognition search strategies that reduce the time needed for high-accuracy recognition; and increased power of audio-capable, off-the-shelf workstations. The paradigm shift is taking place in the way we view and use speech recognition. Rather than being mostly a laboratory endeavour, speech recognition is fast becoming a technology that is pervasive and will have a profound influence on the way humans communicate with machines and with each other.

ADVANTAGES

Speech is a natural form of communication that is pervasive, efficient, and can be used at a distance. Nevertheless, with respect to speech interfaces, user acceptance is complicated by limitations in current technology. Often expectations of how a speech interface should work are biased by our experience with human-to-human interaction. User acceptance was influenced more by accuracy than speed. In addition, factors unrelated to the software itself affected acceptance, such as the level of domain expertise.

DRAWBACKS

Although research in this area has been active for many decades, robustness is still a key issue that should be considered. Thus more effort should be placed in order to accomplish satisfactory performance in adverse acoustic environments [38,41]. Speech recognition is not a common feature among industrial applications, namely because: the technologies of speech recognition and text-to-speech are relatively new although they are already robust enough to be used with industrial applications; the industrial environment is very noisy which puts enormous difficulties to automatic speech recognition (ASR) systems; and the industrial systems were not designed to incorporate these types of features, and usually do not have powerful computers especially dedicated to HMI. Noise is still a problem, but using a short command structure with a specific word as pre-command string it is possible to reduce enormously the noise effects. Recent studies [39] have shown that ASR performance is far from the human performance in a variety of tasks and conditions. Indeed, ASR to date is very sensitive to variations in the channel (desktop microphone, telephone handset, speakerphone, cellular, etc.), environment (non-stationary noise sources such as speech babble, reverberation in closed spaces such as a car, multi-speaker environments, etc.), and style of speech. A typical approach for achieving robustness of environment focuses on obtaining a clean signal through a head-mounted or hand-held directional microphone. However, this is neither tether-free nor hands-free, and it makes speech-based interfaces very unnatural. Moving the speech source away from the microphone can degrade the speech recognition performance due to the contamination of the speech signal by other extraneous sound sources. The research work in robust ASR in noise may be classified into three broad areas: 1) Filtering of the noisy speech prior to classification. In this class of techniques, represented by spectral subtraction, an estimate of the clean speech spectrum is obtained by subtracting an average noise spectrum from the noisy speech. A disadvantage of such techniques is that crucial speech information may be removed during the process. 2) Adaptation of the speech models to include the effects of noise. In this class of techniques, speech models are adapted to include the effects of noise in an attempt to obtain models that would have been obtained in matched conditions. 3) Use of features that are robust to noise. In this class of techniques, an attempt has been made to incorporate temporal and cross-spectral correlation between speech features modeled after the mammalian auditory processing. These signal-based and model-based techniques to make speech recognition independent of channel and environment have been attempted with limited success. Most of these methods make strict assumptions on the environment characteristics and require a sizable sample of the environment to get small improvements in speech recognition performance. Furthermore, modelling reverberation is a hard problem. In summary, current techniques are not designed to work well in severely degraded conditions [39].

3.5 References

[1] Enrico Costanza, Alberto Perdomo, Samuel A. Inverso, Rebecca Allen (2004). EMG as a Subtle Input Interface for Mobile Computing. In MobileHCI 2004: 6th International Symposium (Stephen Brewster, Mark Dunlop) Glasgow,UK, pp 426-430.

[2] M. Zecca, S. Micera, M. C. Carrozza, & P. Dario (2002). On the control of multifunctional prosthetic hands by processing the electromyographic signal, Crit Rev Biomed Eng, vol. 30(4-6), pp. 459-485.

[3] Beau Crawford, Kai Miller, Pradeep Shenoy, Rajesh Rao (2005) Real-Time Classification of Electromyographic Signals for Robotic Control. Proceedings of AAAI (Systems, Man, and Cybernetics Society) pp. 523-528.

[4] Jacob Rosen, Moshe Brand, Moshe B. Fuchs, and Mircea Arcan (2001).A Myosignal-Based Powered Exoskeleton System. IEEE Transaction on Systems, Man, and Cybernetics, Part A 31(3): 210-222.

[5] Trejo L. J., Wheeler K., Jorgensen C., Rosipal R., Hibbs A. (2002). Multimodal Neuroelectric Interface Development. poster presented at The Second International Brain-Computer Interface Workshop, The Rensselaerville Institute, Albany, NY.

[6] P.J. Kyberd, J.L. Pons (2003) A Comparison of the Oxford and Manus Intelligent Hand prostheses. Proceedings of the 2003 IEEE International Conference On Robotics & Automation. Taipei, Taiwan.

[7] M.C. Carrozza, E Vecchi, E Sebastiani, G. Cappiello, S. Roccella, M. Zecca, R. Lazzarini and P. Dario (2003) Experimental analysis of an innovative prosthetic hand with proprioceptive sensors. Proceedings of the 2003 IEEE International Conference On Robotics & Automation, vol.2, 2230-2235.

[8] J. Butterfass, G. Hirzinger, S. Knoch, H. Liu (1998) DLR's Multisensory articulated Hand Part I: Hard- and Software Architecture, in Proc. 1998 IEEE International Conference on Robotics and Automation, pp. 2081-2086.

[9] "Protesi mioelettriche", Available online at:
http://www-arts.sssup.it/moodle/mod/resource/view.php?id=627.

[10] Toshio Tsuji, Osamu Fukuda, Hiroki Shigeyoshi and Makoto Kaneko (2000) Bio-mimetic Impedance Control of an EMG-controlled Prosthetic Hand. Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems.

[11] Edgard Lamounier, Alcimar Soares, Adriano Andrade, Renato Carrijo (2002) A Virtual Prosthesis Control Based on Neural Networks for EMG Pattern Classification. In The 6th IASTED International Conference - Artificial Intelligence & Soft Computing (ASC 2002), Canada, Anaheim - CA - USA: ACTA Press, v. 1. p. 425-430.

[12] Osamu Fukuda, Toshio Tsuji, Makoto Kaneko and Akira Otsuka (2003) A Human-Assisting Manipulator Teleoperated by EMG Signals and Arm Motions. IEEE Transactions On Robotics And Automation, VOL. 19, NO. 2.

[13] Jan Ciger, Bruno Herbelin, Daniel Thalmann (2004) Evaluation of Gaze Tracking Technology for Social Interaction in Virtual Environments. Second International Workshop on Modelling and Motion Capture Techniques for Virtual Environments (CAPTECH).

[14] Takehiko Ohno, Naoki Mukawa, Atsushi Yoshikawa (2002). FreeGaze: A Gazetracking System for Everyday Gaze Interaction. Proceedings of the symposium on ETRA 2002: eye tracking research & applications symposium, pp.125-132.

[15] C. Robino, S. Ramat, D. Zambarbieri (2001) Analisi dei Movimenti Oculari nella Valutazione Quantitativa di Interfacce Grafiche. Simposio su Human-Computer Interaction, HCITALY 2001, SIGCHI-Italy.

[16] Qiang Ji , Zhiwei Zhu (2004). Eye and Gaze Tracking for Interactive Graphic Display. Proc of the 2nd International Symposium on Smart Graphics, Hawthorne, NY, USA, Springer-Verlag.

[17] Yoshio Matsumoto (2003) Real-Time Computer Vision for Human Interfaces. EURON Summer School. Hamburg, Germany.

[18] Jaakko Malmivuo, Robert Plonsey (1995) Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields (1995), Oxford University Press.

[19] David W. Patmore, R. Benjamin Knapp (1998). Towards an EOG-Based Eye Tracker for Computer Control. Proceedings of the third international ACM conference on Assistive technologies. Marina del Rey, California, 197 – 203, ACM Press New York, NY, USA

[20] O. Rogalla, M. Ehrenmann, R. Zollner, R. Becher and R. Dillmann (2002). Using Gesture and Speech Control for Commanding a Robot Assistant. Int. Workshop on Robot and Human Interactive Communication, Berlin, Germany.

[21] "State of the Art in Virtual Reality Technology", available online at: http://www.nicve.salford.ac.uk/CIBTG24/vr.html.

[22] Input Controllers, available online at:
http://www.genreality.com/inputcontrols.html#CyberTrack2

[23] A. G. MacDiarmid (2001) "'Synthetic metals': a novel role for organic polymers", Current Applied Physics, vol. 1, pp. 269-279.
[24] A. J. Heeger (2001) "Semiconducting and metallic polymers: the fourth generation of polymeric materials", J. Phys. Chem. B, vol. 105, no. 36, 8475 – 8491.
[25] D. De Rossi, A. Della Santa, A. Mazzoldi (1999) "Dressware: wearable hardware", Mat Sci Eng C, vol. 7, pp. 31-35.
[26] D. De Rossi, F. Lorussi, A. Mazzoldi, P. Orsini, E. P. Scilingo (2000) "Monitoring body kinematics and gesture through sensing fabrics", Proc. of the 1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine & Biology, Lyon, France, pp. 587-592.
[27] D. De Rossi, F. Lorussi, A. Mazzoldi, E. P. Scilingo, P. Orsini (2001) "Active dressware: wearable proprioceptive systems based on electroactive polymers", Proc. of the Fifth International IEEE Symposium on Wearable Computers, pp. 161-162.
[28] E. P. Scilingo, F. Lorussi, A. Mazzoldi, D. De Rossi, "Strain-sensing fabrics for wearable kinaesthetic-like systems", IEEE Sensors J., vol. 3, no. 4, pp. 460–467, 2003.
[29] B. J. Muro, J. R. Steele, T. E. Campbell and G. G. Wallace (2003) "Wearable textile biofeedback systems: are they too intelligent for the wearer?", Proc. of International Workshop-New Generation of Wearable Systems for e-Health: Towards a Revolution of Citizens' Health and Life Style Management?, Lucca, Italy, pp. 187-193.
[30] D. De Rossi, F. Carpi, F. Lorussi, R. Paradiso, E. P. Scilingo, A. Tognetti (2005) "Electroactive fabrics and wearable man-machine interfaces", in Wearable electronics and photonics, X.-M. Tao Ed. Cambridge: Woodhead Publishing Limited.
[31] D. De Rossi, F. Lorussi, E. P. Scilingo (2003) "Artificial kinesthetic systems for telerehabilitation", Proc. of International Workshop-New Generation of Wearable Systems for e-Health: Towards a Revolution of Citizens' Health and Life Style Management?, Lucca, Italy, pp. 129-133.
[32] D. De Rossi, F. Lorussi, A. Mazzoldi, P. Orsini, E. P. Scilingo (2003) "Active dressware: wearable kinesthetic systems', in Sensors & Sensing in biology & Engineering, J. Secomb Ed. New York: Springer Verlag, pp. 381-394.
[33] D. De Rossi, F. Carpi, F. Lorussi, A. Mazzoldi, R. Paradiso, E. P. Scilingo, A. Tognetti (2003) "Electroactive fabrics and wearable biomonitoring devices", AUTEX Research J., vol. 3, no. 4, pp. 180-185.
[34] A. Tognetti, F. Carpi, F. Lorussi, A. Mazzoldi, P. Orsini, E.P. Scilingo, M. Tesconi and D. De Rossi (2003) "Wearable Sensory-Motor Orthoses for Tele-Rehabilitation", Proc. of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Cancun, Mexico, September.
[35] D. De Rossi, F. Carpi, F. Lorussi, A. Mazzoldi, E. P. Scilingo, A. Tognetti (2002) "Electroactive fabrics for distributed, conformable and interactive systems", Proc. IEEE Sensors 2002, Hyatt Orlando, Florida.
[36] F. Lorussi, W. Rocchia, E. P. Scilingo, A. Tognetti, D. De Rossi (2004) "Wearable, redundant fabric-based sensor arrays for reconstruction of body segment posture", IEEE Sensors J., vol. 4, no. 6, pp. 807-818.
[37] F. Lorussi, A. Tognetti, M. Tesconi, P. Pastacaldi, D. De Rossi (2003) "Strain sensing fabric for hand posture and gesture monitoring", Proc. of International Workshop: New Generation of Wearable Systems for e-Health: Towards a Revolution of Citizens' Health and Life Style Management?, Lucca, Italy, pp. 175-180.
[38] Khalid Daoudi (2004) State of the art in speech and audio processing. Available online at: http://www-rocq.inria.fr/imedia/Muscle/WP5/Docs/WP5_SoA_Task3_sept04.pdf
[39] C. Neti, G. Potamianos, J. Luettin, I. Matthews, H. Glotin, D. Vergyri, J. Sison, A.Mashari, and J. Zhou (2000) Audio-Visual Speech Recognition, Final Workshop

2000 Report, Center for Language and Speech Processing, The Johns Hopkins University, Baltimore, MD.

[40] R. Brown (2006) Talking Windows - Exploring New Speech Recognition And Synthesis APIs In Windows Vista. MSDN Magazine.

[41] J. Norberto Pires (2005) Robot-by-voice: experiments on commanding an industrial robot using the human voice. Industrial Robot, An International Journal, Emerald Group Publishing Limited, Volume 32, Number 6.

[42] Michael A. Grasso, David Ebert, Tim Finin (1997) Acceptance of a Speech Interface for Biomedical Data Collection. Proceedings of the AMIA 1997 Annual Fall Symposium pp. 793-743.

[43] Makhoul, J. & Schwartz, R. (1995) State of the art in continuous speech recognition. Proc. Natl. Acad. Sci. USA 92, 9956-9963.

4. Space applications of BMI and other non invasive man-machine interfaces

4.1 BMI: preferred candidate concepts

Table 4.1.1 briefly summarises fundamental issues resulting from an analysis of ESA technical requirements (according to the ESA ARIADNA call AO4919 activity 05/6402) for space applications of BMI:

Table 4.1.1: Summary of technical requirements for space applications of BMI

ESA technical requirement	**Comments**	**Candidate concept**
Non-invasive brain-machine interfaces (BMI) as advantageous systems in space applications (e.g. tele-control of robotic systems operating in extra-vehicular activity)	Need for an interface characterised at least by the following fundamental properties: - non-invasivity - high reliability - high efficiency - high sensitivity - ease of use by the astronaut - sufficient comfort for the astronaut - electromagnetic compatibility with electronic equipment of the spacecraft cabin - low weight and volume of the driving and reading equipment	The need of respecting at least all these characteristics limits the types of BMI potentially useful. In particular, it tends to exclude, in principle, fMRI-based interfaces and MEG-based interfaces, for the following reasons: - both of these types require complicate, bulky, heavy and expensive driving and reading equipment; - both of these types are definitely not practical for astronauts, especially in consideration of the activities that they have to perform continuously. On the contrary, EEG-based interfaces represent a more suitable candidate concept, as summarised in the following Table 4.2.1.
Multi-teleoperations to be performed using BMI	A BMI may consent to a user to perform, in principle, multiple tasks following an adequate training. Nevertheless, considerable benefits for a BMI employed in multi-task activities may derive from the concomitant use of additional and auxiliary man-machine interfaces. They would permit to compensate, speed-up, make easier and more efficient some actions to be performed in parallel to others.	Different types of non-invasive man-machine interfaces to be used as auxiliary systems for a BMI can be considered: see summary in section 4.2. As an example, electromyographic signals detected from the arm muscles of an astronaut may permit him to control a robotic arm. Electrical activity recorded in proximity of his muscle could be elaborated and used as the control input for the robotic mechanism. Such an action may be performed in parallel to those controlled by the brain interface. Therefore, the two interfaces could work at the same time to enable easier and efficient implementations of multiple tasks.

An evaluation of such technical requirements suggests a reasonable exclusion, in principle, of fMRI-based and MEG-based interfaces, from the candidate concepts to develop non-invasive BMI for space applications, due to their lack of practicality.

On the contrary, EEG-based interfaces are expected to represent the most suitable candidate concept for a BMI to be employed for space applications.

In fact, interfaces based on electroencephalogram are characterized by relative simplicity and non-invasiveness. Although the quality and resolution of brain signals measured detected via EEG skin electrodes are not comparable to those recorded by means of implanted electrodes (due to a reduced spatial resolution and increased noise), such a non-invasive EEG method is of course preferable. Detected signals have been demonstrated (section 2) to permit to mentally operate devices and systems, e.g. robots in indoor environments. This is possible owing to the combination of advanced robotics, opportune protocols for the analysis of online EEG signals and machine learning techniques. This technology is in continuous development, thanks to recent scientific and technological results. In fact, first of all basic and clinical research has yielded detailed knowledge of the signals that comprise the EEG. For the major EEG rhythms and for a variety of evoked potentials, their sites and mechanisms of origin and their relationships with specific aspects of brain function are no longer wholly obscure. Second, the extremely rapid and continuing development of inexpensive computer hardware and software supports sophisticated online analyses of multichannel EEG. In addition, there have been significant advances in the development and use of electrophysiological recording methods. Nevertheless, beyond these recent advances, a number of challenging issues are still open as described in section 2. This must be carefully taken into consideration in perspective of potential space applications.

Interestingly, different types of non-invasive man-machine interfaces may offer potential advantages as auxiliary and complementary systems for such a type of BMI. This might result particularly advantageous whenever different tasks have to be performed at the same time (multi-task activities). For instance, this can be the case of an astronaut having to deal with multiple operations (e.g. multi-teleoperations) to control robotic arms operating in extra-vehicular activity. In such a situation, the astronaut may benefit from having the possibility of sending multiple control signals by using different types of interfaces at the same time. Each interface may be dedicated to a specific set of tasks. This may reduce the amount of information to be processed by the single brain interface and, therefore, may increase the accuracy, sensitivity and efficacy of the overall operation. Moreover, depending on the type of activity to be performed, a suitable interface (with specific properties and performance) could be selected. Nevertheless, depending on the type of auxiliary interface used, a further specific training of the operator, as well as specific signal elaboration strategies, should be adopted. The main features of the potential auxiliary interfaces presented in section 3 are summarized in section 4.2.

4.2 Non invasive man-machine interfaces besides BMI: summary evaluation

Among the techniques analysed in section 3, some of them are considered particularly promising for future space applications. Some of these are related to systems which have been already demonstrated for commercial civil applications. As an example, speech based interfaces are already available on the market. On the other hand, improvements are necessary for some types of such interfaces, in perspective of more demanding fields of applications, such as space. Table 4.2.1 summarizes fundamental properties of the considered different types of interfaces.

Interfaces based on speech recognition are today rather advanced. However, as discussed in section 3, they still present specific problems. For instance, better methodologies for improving recognition performance are necessary, as previously described. Moreover, the potential usability of this technique must be evaluated with respect to the particular application. As an example, the astronaut must speak also with colleagues and ground station. So, the feasibility is actually related to the application.

Concerning motion capture, the feasibility is high with wearable systems, while optical systems need too much cumbersome and complicate equipment. With opportune and personalized calibration strategies, a suitable reliability can be obtained.

A lower reliability is typically achieved with EMG based interfaces. Regardless the particular elaboration scheme adopted, this is frequently due to possible failures from the user in the control of his muscular contraction (e.g. due to fatigue).

The worst properties are shown by gaze tracking techniques. They typically are characterized by a low resolution. This is particularly problematic if it is necessary to recognize between more than two states. The reliability of these techniques is considered not sufficient for delicate tasks, as in the case of potential space applications. Moreover, they are not readily applicable: equipment for corneal reflection is not practical, while EOG electrodes are a source of discomfort. Accordingly, gaze tracking techniques are not considered as appropriate candidates for space applications.

On the contrary, both speech recognition, motion capture and EMG activation can be considered as valid strategies for non invasive interfaces. These technologies could advantageously be used as auxiliary tools for BMI, as previously described.

Table 4.2.1: Non invasive man-machine interfaces besides BMI: fundamental properties

Interface	**Number of controllable degrees of freedom**	**Reliability**	**Feasibility**	**Main advantages**	**Main disadvantages**
Speech recognition	Not defined	High	High	• No devices to be worn • Readily and easily applicable • Applicable in any light/darkness condition • The user can perform any other action at the same time	• Not applicable in noisy environments • Not compatible with usual verbal communications (inside cabin or with ground station)
Sub-vocal commands (vibration of vocal cords and tongue)	Not defined	High/ Medium	Medium	• Applicable in noisy environments • Applicable in any light/darkness condition • The user can perform any other action at the same time	• Need to wear a device • Not compatible with usual verbal communications (inside cabin or with ground station)
Motion capture and gesture recognition (wearable systems)	Equal to the number of monitored joints	High/ Medium	High/ Medium	• Applicable in any light/darkness and noisy condition • Exploitation of body commands	• Need to wear a device • The user can't perform any other action at the same time with that portion of the body
EMG activation	2 (recommended)	Medium	Medium	• Applicable in any light/darkness and noisy condition • 'Biomimetic' activation • Skin electrodes applicable beneath standard suits, without any additional device.	• Need to wear skin electrodes • Need of considerable and accurate training of the operator • High tendency to muscular fatigue
Gaze-Tracking: corneal reflection	1-2	Low	Low	• No devices to be worn • Applicable in noisy environments	• Need of absence of obstacles for light propagation • Tendency to muscular fatigue
Gaze-Tracking: EOG	1	Low	Low	• Applicable in any light/darkness condition and noisy environment	• Discomfort (skin electrodes around eyes) • Tendency to muscular fatigue

4.3 Technology readiness and perspectives

Currently, numerous systems based on the use of the speech technology are available on the market. On the contrary, for motion capture the market offers still few usable systems. Accordingly, a higher time scale for possible exploitations in the space field can be foreseen for interfaces based on motion capture with respect to speech recognition. An even higher time scale is envisaged for the EMG based technology. In this case, one of the major difficulties, which currently limit the wide spreading of this technique, is the considerable training required to properly use it. Table 4.3.1 summarizes the technology readiness and perspectives (tentative indication of possible time scale) of the different methodologies.

Table 4.3.1: Non invasive man-machine interfaces besides BMI: technology readiness and perspectives for space applications

Interface		**Technology readiness**	**Comments**	**Perspective time scale for possible space applications**
BMI	EEG activation	Low	No reliable commercial examples	> 10 years
Non BMI	Speech recognition	High	Already established commercial uses	1 year
	Sub-vocal commands	Medium	No reliable commercial examples	3 years
	Motion capture and gesture recognition (wearable systems)	Medium	The first few commercial products are appearing on the market	5-10 years
	EMG activation	Medium/ Low	No reliable commercial examples	> 10 years

4.4 Examples of potential space applications

Both BMIs and the different kinds of non-invasive man-machine interfaces described above may result particularly useful for several types of potential space applications. Some of the most relevant ambits of possible use are briefly mentioned below.

4.4.1 Spacecraft robotic systems

The possibility of controlling spacecraft robotic systems, such as robotic arms or tele-manipulators, is certainly one of the main fields of potential space application for both BMIs and different non-invasive interfaces.
As an example, Fig. 4.4.1 shows the robotic arm of the space shuttle. Astronauts may control operations of this system through elaborations of brain signals or alternative body inputs.

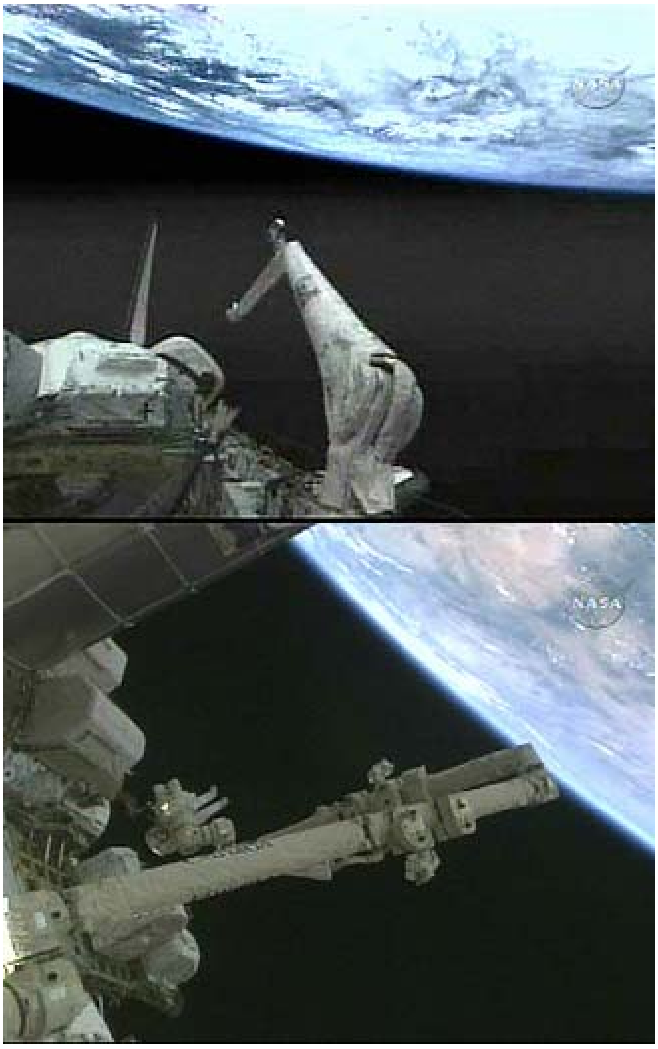


FIG. 4.4.1: Space shuttle robotic arm (NASA public pictures)

As reported in sections 2 and 3 for the different considered interfaces, ongoing research efforts spent by several groups all over the world are progressively consolidating the scientific and technological knowledge necessary to face these (and many other) kinds of delicate applications.
The human control of a robotic arm of the type shown in the example of Fig. 4.4.1 opens a series of new challenging issues. Nevertheless, some of them could result not so far from those related to other kinds of systems studied for different applications: for instance, artificial robotic hands to be used as prostheses. As an example, this could be the case when the robotic system should be managed by the astronaut as a sort of ‘appendix’ of his own body. Actually, in such a case several lessons learned from those different systems may result particularly useful. Accordingly, the reader is here referred to Section 3.1, where some basic issues related to EMG controls of prosthetic hands is reported.

4.4.2 Autonomous vehicles

Tele-controls of autonomous vehicles (to be used for instance for explorations, reparations or maintenance) are another example of potential field of application for the considered interfaces. The scientific bases for such tasks are currently being explored and successful results have been already reported (see ref. [11] of section 2). In that work, two human subjects were able to drive a robot between several rooms (Fig. 4.4.2) by mental control only, using an EEG-based BMI that recognized three mental states.

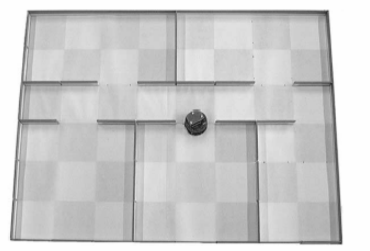


FIG. 4.4.2: A mobile robot (two-wheeled vehicle) in its environment, consisting of several rooms along a corridor. The robot had three lights on the top to provide feedback to the user and 8 infrared sensors around its diameter to detect obstacles (from ref. [11] of section 2).

The performance of mental controls was comparable to manual controls on the same task, with a performance ratio of 0.74. This work introduced as a novel idea to control robots by mapping asynchronously high-level mental commands into a finite state automaton (see ref. [11] of section 2).

4.4.3 Cabin instrumentation and equipment

Providing commands to cabin instrumentation or equipment may be accomplished not only manually, but also exploiting man-machine interfaces. The most promising types of useful interfaces for such purposes are certainly those based on speech recognition or sub-vocal commands. They have been reported in section 3.4. Both of them permit to the user to provide easily and quickly direct commands. Accordingly, astronauts may take advantage of these systems, to control cabin instrumentation or equipment without any manual command. For instance, they could advantageously provide instructions both remotely and preserving their hands free.
The current higher reliability of interfaces based on speech recognition is paid with a substantial uselessness in noisy environments. In such cases, sub-vocal interfaces are more suitable, although their performance is currently lower. More generally, the second type of systems could allow astronauts to silently control instrumentation or equipment aboard a spacecraft.

4.4.4 Exoskeletons

Human exoskeletons are conceived as wearable systems (Fig. 4.4.3) typically intended to serve for augmenting different types of human functions, such as muscular power or protection from dangerous environments.

FIG. 4.4.3: CAD rendering of an arm exoskeleton (adapted from [1]).

Exoskeleton concepts are currently being studied for both soldiers, labourers (lifting of heavy weights), nurses (lifting of patients) and patients requiring rehabilitation (e.g. after stroke or spinal cord ingiuries).
They may result particularly useful as auxiliary systems for astronauts too. The controllabilty of such a type of systems is a relevant potential space applicaton for both BMIs and different types of non-invasive man-machine interfaces. Section 3.1 reports some basic issues related to EMG controls of exoskeletons.
Several studies have been committed by ESA in this ambit of application. Therefore, according to a specific ESA request, this topic is not further considered in this report.

4.4.5 Extra-vehicular robotic activities

In addition to the considered types of applications, different examples are possible too, although they appear as quite challenging for the current maturity of the technology. For example, a robotics team at NASA's Johnson Space Center in Houston is developing a new breed of space robots called Robonaut [2]. Robonaut, designed to be as human-like as possible, will be remotely controlled to work in Extra-Vehicular Activity (EVA) environments, allowing astronauts to remain safely inside the spacecraft. This humanoid design criterion was key, because, over the last 50 years, space flight hardware has been made for human servicing and space walks (Fig. 4.4.4) are the main contingency for repairing on-orbit failures.

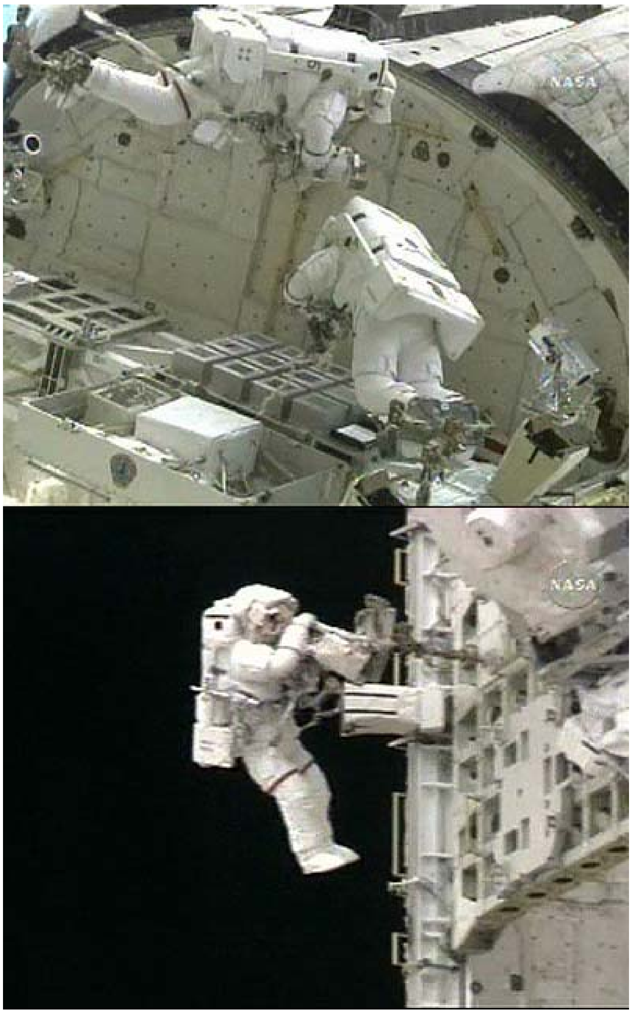


FIG. 4.4.4: Astronauts' space walks for extra-vehicular activity (NASA public pictures)

Robonaut's value is not only in assuming some of the mundane repair chores aboard the spacecraft, but it will respond much more quickly to emergencies than a human can. Since it can take several hours to suit up a human in preparation for a spacewalk, the Robonaut is better suited for unexpected emergency situations, since no life support equipment or supplies are required. A typical use will be from the International Space Station. Robonaut will be housed on-board, in its own locker, ready for an assignment. Although the robot is designed to be as human-like as possible, some adaptations have been made to avoid some of the disadvantages of the human design. For instance, in a weightless environment like space, human feet are a hindrance and are usually bolted down, in a portable foot, to avoid unnecessary drift. Therefore, Robonaut does not have legs rather it is connected to the space craft using a "tail" design. Once outside the spacecraft, Robonaut will be controlled by telepresence and will mimic the movements made by an astronaut. If the application is too complex to be done without human intervention, the Robonaut can still be successfully employed as an assistant, removing hardware from lockers, setting up tools, and having everything ready for the human worker. Robonaut is being designed and built to interface with external space station systems that only have human interfaces. To this end, the Robonaut hand provides a high degree of anthropomorphic dexterity ensuring a compatibility with many of these interfaces [3]. Another application concerns, as an example, the Hubble Space Telescope (HST) that was carried into orbit in 1990 aboard the Space Shuttle Discovery (STS-31). High focal length designs are too large to fit in the Orbiter payload bay as a single unit. Instead, components of these space telescopes could be launched separately for a rendezvous in Low Earth Orbit (LEO) for assembly and final repositioning. These and other larger, lighter, more extendable space structures will require greatly expanded EVA and Extra-Vehicular Robotic (EVR) capabilities as well as new and innovative structural systems. The recent emergence of highly dexterous space robots dramatically increases the opportunities for humans and robots working together in space. These machines can help conserve EVA hours by relieving human astronauts of many routine chores and assisting them in more complex tasks. Robots can take risks unacceptable to humans, perform contingency EVA operations in minutes, instead of hours, and setup worksites in preparation for the arrival of human astronauts. The Filled-Aperture Infra-Red (FAIR) telescope is one example of the new generation of space science platforms requiring expanded EVA/EVR capabilities. Boasting an extremely long focal length, the FAIR design is too large and flimsy to be carried into orbit as a single pre-integrated assembly. Instead, the spacecraft subassembly and components of the telescope are launched aboard the Space Shuttle, while the propulsion stage is launched separately on an expendable vehicle. For this task, humanoid robots, controlled by remote human tele-operators, may serve in various roles supporting the astronaut. Higher-fidelity tests involving more sophisticated gravity compensation and more realistic mobility, communication time delay, lighting conditions, etc. will complicate some operational aspects of EVR but simplify others. The multi-agent team featured represents only one particular instance of humans and robots working together. It represents a novel integration of existing technology prototypes and test beds intended to meet test objectives. However, it should only be interpreted as a part of the solution for increasing EVA capability and productivity [4].

4.4.6 Final remark

Some examples of potential space applications for BMIs and other non-invasive man-machine interfaces have been considered. It is worth noting that for all these examples it could be useful to combine the action of a BMI with another non invasive interface. In fact a BMI could require and may benefit from an auxiliary system to be used for specific tasks. This is the case of multi-task operations, which require several efforts from the astronauts. The challenge is to properly combine these technologies, by making the system robust and intelligent. Some tasks may be performed by using signals detected by a BMI while others, at the same time, may be accomplished by exploiting alternative means of communication enabled by different types of non-invasive interfaces. Fulfilment of such issues may open completely new approaches to manage space operations.

4.5 References

[1] Jacob Rosen, Blake Hannaford, Stephen Burns (2003) Neural Control of an Upper Limb Powered Exoskeleton System - Grant Report, First NSF Robotics and Computer Vision (RCV) Workshop – October 2003, Las Vegas, Nevada.

[2] Gay Engelberger (2001) NASA's Robonaut. Industrial Robot: An International Journal, Volume 28, Number 1, pp. 35-39.

[3] Lovchik C., Aldridge H., and Diftler (1999). Design of the NASA Robonaut Hand, Proceedings of the ASME Dynamics and Control Division, DSC-Vol. 67, Nashville, Tennessee.

[4] Fredrik Rehnmark, Nancy Currie, Robert O. Ambrose and Christopher Culbert (2004) Human-Robot Teaming In A Multi-Agent Space Assembly Task. World Automation Congress Tenth International Symposium on Robotics with Applications Seville, Spain.