uOttawa

L'Université canadienne
Canada's university

**FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES**

uOttawa

L'Université canadienne
Canada's university

**FACULTY OF GRADUATE AND POSDOCTORAL STUDIES**

Vladimir Hinic
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Electrical Engineering)
GRADE / DEGREE

School of Information Technology and Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Brain Computer Interface System for Communication and Robot Control Based on Auditory Evoked Event-Related Potentials

TITRE DE LA THÈSE / TITLE OF THESIS

Prof. E. Petriu
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Prof. V. Groza

Prof. M. Nakla

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

# Brain Computer Interface System for Communication and Robot Control Based on Auditory Evoked Event-Related Potentials

by

Vladimir Hinić

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the MASc degree in Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Information Technology and Engineering
Faculty of Engineering
University of Ottawa

*Your file Votre référence*
*ISBN:* 978-0-494-59860-3
*Our file Notre référence*
*ISBN:* 978-0-494-59860-3

# Abstract

*This study evaluates the suitability of the auditory, specifically spoken words, evoked P300 based Brain-Computer Interface (BCI) for the human-computer symbiotic control of Robotic Sensor Agents (RSA). This study is unique in using auditory evoked P300 for BCI communication and control, contrary to currently established P300 BCI research. The general, elaborate, and multi platform software framework BCIILab was developed as a part of BCI laboratory workbench used specifically in the research. Result of this study shows that it is possible to use auditory P300 based BCI for the complementary control of RSA on the strategic level. The problems encountered in classification during experiments were highlighted and suggestion for future research given.*

# Acknowledgements

I am greatly indebted to my graduate supervisor, Dr. Emil M. Petriu. He has provided invaluable advice, guidance, and support throughout my time at University of Ottawa. My graduate study experience has been profoundly enriched by his constant encouragement, enthusiasm and selfless interest in my success. Most importantly, he has also taught me how to become a better engineer and critic of my own work. I will always be grateful for his wisdom, understanding and patience on both an academic and personal level. I will cherish our passionate and lengthy discussion and exchange of ideas related to computer technology and robotics.

I would also like to extend a special thanks to Dr. Nicolas D. Georganas for expressing interest in my initial research ideas and introducing me to Dr. Emil M. Petriu and his research. EEG equipment that he selflessly landed me for this research was invaluable for experimental part of this project.

I am grateful to the faculty, staff and students in the Department of Electrical and Computer Engineering, and especially everyone in SMR and DISCOVER laboratory for making my graduate study stimulating and fun.

For giving me a second chance, personally, professionally and academically, I am ever so grateful to Canada and all of immigrants that made it this great country that it is today.

Finally, out of all my hart, I owe my partner and friend in life, Irena and our two sons Goran and Luka a very special thanks. Without their support, motivation and understanding I would not be able to break-free from daily routine and gather enough strength to go back to school and complete graduate studies. Thank you for coping with my student "blues", being a strange "dad" and odd working hours. Thank you for your unconditional support and love. This work belongs to you and I only hope that it is at the level you deserve.

# Table of Contents

1 Introduction ........ 1

1.1 Overview ........ 1

1.2 Human-Computer Symbiotic Partnership ........ 2

1.3 General Overview of BCI Research ........ 5

1.4 Neurophysiological basis for BCI ........ 7

1.5 Typical BCI Architecture ........ 11

1.6 Contribution ........ 13

1.7 Thesis Organization ........ 14

2 Background Information ........ 16

2.1 A Brief Introduction to EEG, EP, and P300 ERP ........ 16

2.2 Classification Using Support Vector Machine ........ 23

3 Literature Review ........ 33

3.1 State of the Art in P300 BCI ........ 33

3.2 BCI for Robotic Control ........ 37

4 Symbiotic Robot Control Using Spoken Words

Auditory Evoked ERP Based BCI System ........ 40

4.1 Problem Statement ........ 40

4.2 Research Outline ........ 41

4.3 BCI System Framework ........ 42

4.4 Experimental Setup ........ 55

5 Results ........ 61

6 Conclusions .................................................................................................... 67

References .................................................................................................... 69

# List of Tables

2.1: Spontaneous rhythmic brain potentials .......... 19

2.2: Typical SVM kernel functions .......... 30

4.1: EEG equipment used in BCI system .......... 43

4.2: Robot kits used in BCI system .......... 45

4.3: Open Source projects used in BCILab .......... 46

4.4: Navigator components and their role .......... 53

# List of Figures

1.1: Network of intelligent robotic sensors monitoring the environment. ............................ 3

1.2: Human-robot interaction for symbiotic RSA operation. ................................................ 4

1.3: A letter written using Thought Translation Device (TTD)

BCI by the ALS patient. .......................................................................................... 6

1.4: Restoration of hand grasp in a tetraplegic patient with high spinal cord injury. ............ 6

1.5: Different neurophysiological basis for present day EEG based BCI systems:

(A) Slow Cortical Potentials; (B) P300 ERP;

(C) Sensorimotor rhythm (ERS/ERD). ......................................................................... 9

1.6: The architecture of a typical BCI system. ................................................................... 11

1.7: An example of BCI application: BCI based pong game ............................................. 12

2.1: An early EEG recording, obtained by Hans Berger in 1924.

The upper tracing is EEG, and the lower is a 10 Hz timing signal. ............................ 16

2.2: The International Federation of EEG Societies 10-20 system seen from (A)

left and (B) above the head. A- Ear lobe, C - central, Pg – nasopharyngeal,

P - parietal, F - frontal, F - frontal polar, O - occipital. ............................................... 17

2.3: A taxonomy of brain potentials of interest for BCI research. ...................................... 18

2.4: ERPs evoked by visual (A) and auditory (B) stimuli. Schematic potential traces

on a logarithmic time-scale. Exogenous components are different for visual

(electroretinogram – ERG) and auditory (brainsteam auditory potentials – BEAPs

and midlatency auditory evoked potentials MEAPs) modalities. Endogenous

components, including P300, are very similar in both modalities. ............................. 20

2.5: Interface display used in P300 speller. The rows and columns of the matrix flashed alternately. The letters selected by the subject (BRAIN) are displayed at the top. ...... 21

2.6: Relationship between different errors and their dependency on VC dimension (i.e., complexity or capacity) of the model, labelled using varying terminology used in literature. ...... 24

2.7: Two simple linearly separable classes in R2 with possible separating hyperplanes. Hyperplane H3 is not properly separating classes. Both H1 hand H2 do separate two classes, but H2 does it with the maximum margin. ...... 25

2.8: Maximum-margin hyperplane and margins for a SVM trained with samples from two classes. Samples on the margin are called the support vectors. ...... 26

2.9: Example of nonlinear SVM for 3 input vectors. ...... 31

3.1: P300 speller paradigm interface used in BCI Competitions. ...... 33

3.2: Typical visual paradigms used in current P300 based BCI systems. ...... 36

3.3: A robot controlled by mental-task BCI and sample maze used by Millán et al. ...... 38

3.4: A humanoid control interface used by Bell et al. The robot arena is presented in A. The robot's eye view and choice of objects for selection by operator can be seen in B. After P300 was detected robot is instructed in picking the desired robot as presented in C. ...... 39

4.1: Components of BCI System for control of RSA. ...... 42

4.2: EEG Cap by Electro-Cap International, Inc. used in BCI setup ...... 44

4.3: EEG8 biological amplifier from Contact Precision Instruments Inc. ...... 44

4.4: One of the robots based on Bioloid Kit and used in the experimental setup. ...... 45

4.5: BCILab modules and their dependencies. ...... 47

4.6: Core classes of BCILab Signal module. .......... 48

4.7: Different BCI application paradigms implemented in BCILab. .......... 49

4.8: Signal components. .......... 50

4.9: BCILab user interface for: A) View of raw EEG data and P300 event trigger (top channel); P300 Speller paradigm; and Navigator paradigm. .......... 51

4.10: Signal components configuration for Navigator paradigm. .......... 52

4.11: XML configuration file for Navigator application paradigm. .......... 54

4.12: Experimental setup for the P300 BCI control of a RSA. .......... 55

4.13: Location of 8 electrodes used in experiment. .......... 56

4.14: Structure of a stimuli trial run. .......... 57

4.15: A simple strategic plan execution state machine. .......... 58

5.1: Target stimuli - EEG activity over the Cz area. The dashed line represents stimuli offset. The timeline starts from stimuli onset. .......... 61

5.2: Non-target stimuli - EEG activity over the Cz area. The dashed line represents stimuli offset. The timeline starts from stimuli onset. .......... 62

5.3: Possible P300 and its average latency for the Cz recording. The timeline starts from stimuli onset. .......... 63

5.4: A sample data window in relationship to the spoken word .......... 64

5.5: Classification accuracy as a result of trial run repetitions. .......... 65

## List Of Abbreviations

| | |
|---|---|
| AEP | Auditory Evoked Potential |
| ALS | Amyotrophic Lateral Sclerosis |
| BCI | Brain-Computer interface |
| BMI | Brain-Machine interface |
| CNV | Contingent Negative Variation |
| ECoG | Electrocorticogram |
| EEG | Electroencephalography |
| EP | Event Potentials |
| ERD | Event Related De-synchronization |
| ERN | Error Related Negativity |
| ERP | Event-Related Potentials |
| ERS | Event Related Synchronization |
| fMRI | functional Magnetic Resonance Imaging |
| HCI | Human-Computer Interaction |
| KKT | Karush-Kuhn-Tucker conditions |
| MCU | Micro Controller Unit |
| MEG | Magnetoencephalography |
| PET | Positron Emission Tomography |
| QP | Quadratic Optimization |
| RBF | Radial Basis Function |
| RSA | Robotic Sensor Agents |

| | |
|---|---|
| SCP | Slow Cortical Potentials |
| SEP | Somatosensory Evoked Potential |
| SLT | Statistical Learning Theory |
| SMO | Sequential Optimization Method |
| SNR | Signal to Noise Ratio |
| SRM | Structural Risk Minimization |
| SVM | Support Vector Machines |
| UML | Unified Modelling Language |
| VC | Theory of Vapnik-Chervonenkis |
| VEP | Visually Evoked Potential |

# 1 Introduction

## *1.1 Overview*

With the development of computer technology and related disciplines the traditional view of human-computer interaction (HCI) is gradually expanding into human-computer coupling and human-computer symbiotic partnership [32, 34, 36].

Humans posses the ability to quickly asses the overall situation based on information provided. Traditionally, they have been considered as performing decision making tasks in the man-machine systems [32]. A human can also be viewed as high bandwidth, dexterous and intelligent component when studied as an integral part of a human-computer symbiosis systems [20]. In this new relationship between man and machine the traditional role of the human as a controller or decision maker is shifting toward more intricate role where human desire and perception of the system and environment are used as indirect "contextual" sensory information [34]. Human senses and contextual state are fused in a multimodal sensory system and the system takes the role of the controller of the overall behaviour.

Motivated by predictions in development of human-computer symbiotic partnership [32, 36], the focus of this research is exploring a brain-computer interface (BCI) as an enabling technology in communication and robotic control in a human-robotic partnership [20]. This thesis presents a possible solution for this symbiotic partnership using auditory evoked P300 event related potential (ERP) based BCI system.

## *1.2 Human-Computer Symbiotic Partnership*

Based on the evolution of computing technologies, it is reasonable to expect that current human-computer interfaces (HCI) will evolve into more efficient multimedia symbiotic partnerships in which humans will contribute human-specific capabilities complementing those of the computer. Some of these capabilities are as follows [32]:

i) humans are very high-bandwidth creatures capable of perceiving more than a hundred megabits of information per second and able to recognize visual, auditory, olfactory, gustatory, and haptic stimuli;

ii) humans are dexterous, which allows them to precisely manipulate a wide variety of objects;

iii) the most important capability is that humans are still far more intelligent than any computer, are able to act on incomplete or ambiguous instructions, able to adapt to a variety of computer interfaces, and able to interact directly with other humans and asses individual and society-level human behaviour.

Figure 1.1 shows the architecture of a wireless network of stationary and mobile Robotic Sensor Agents (RSA), under development at the University of Ottawa [33]. Individual sensor-views are fused into a coherent Virtualized Reality Environment multimodal world model.

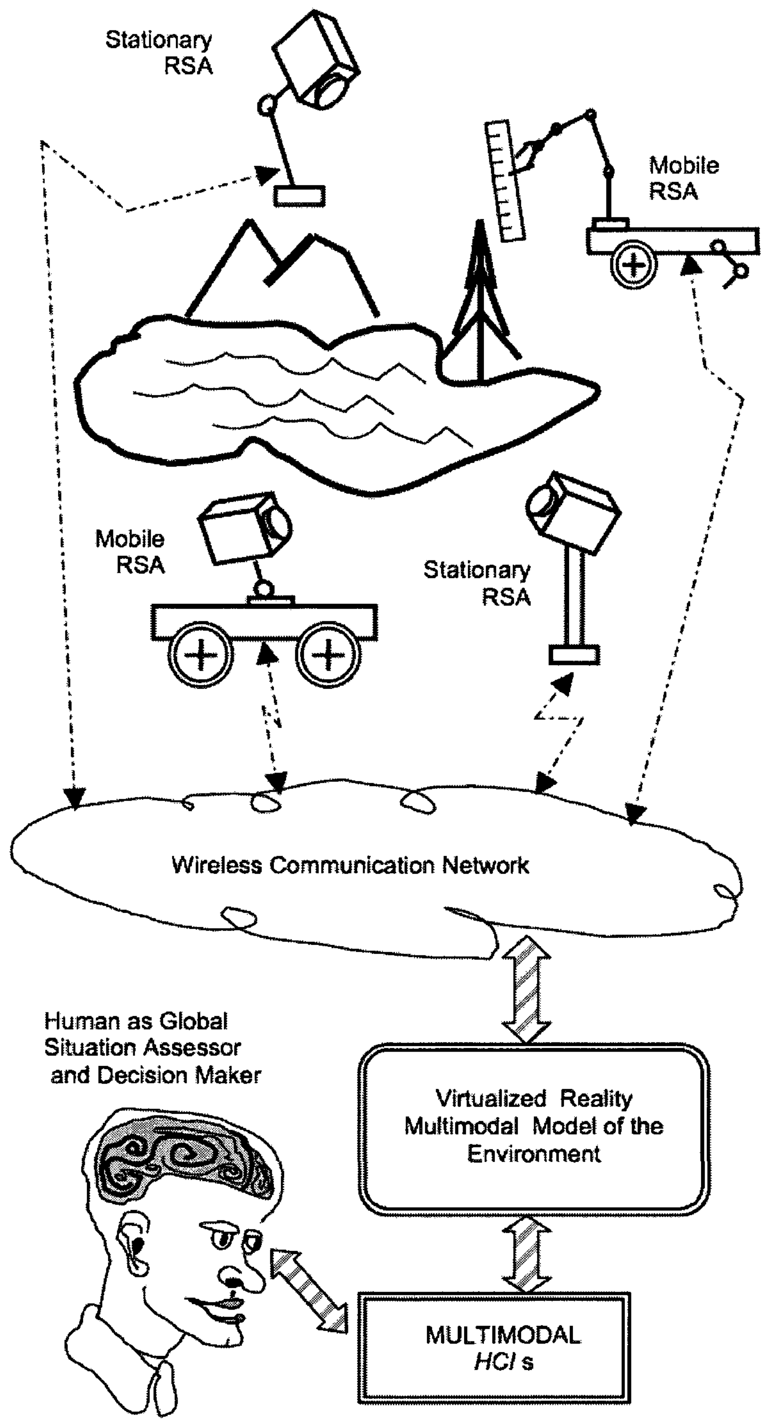


*Figure 1.1: Network of intelligent robotic sensors monitoring the environment (from [20]).*

This thesis presents the human-computer symbiotic cooperation as a complementary resource for environment perception (Figure 1.2) thus expanding the concept of human beings as peripheral devices for computer systems proposed by the authors in [32]. Here the Brain-Computer Interface (BCI) is used as enabling technology for the symbiotic human-computer control of a RSA.

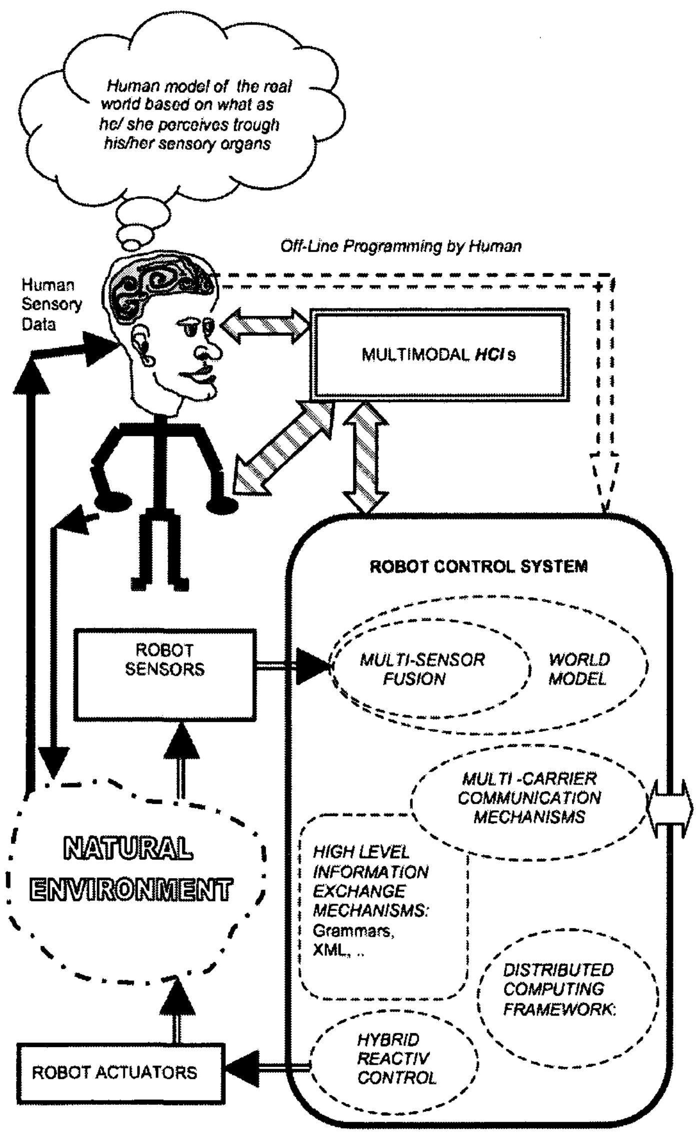


*Figure 1.2: Human-robot interaction for symbiotic RSA operation (from [20]).*

## *1.3 General Overview of BCI Research*

The brain-computer interface (BCI) or brain-machine interface (BMI) is based on a relatively old idea of recording and interpreting brain signals and using them directly for communication and control. The idea that individual intentions and thoughts could be interpreted developed in concordance with the invention of the first electroencephalography (EEG) device [1]. Unfortunately, the use of BCI techniques in research was limited in the past due to questionable practices and development outside of the established medical and scientific communities.

The first international meeting devoted to BCI was held in June 1999 at the Rensselaerville Institute near Albany, New York in order to bring worldwide research efforts together and legitimize BCI research area. BCI was defined as [46]:

> *"... a communication system that does not depend on the brain's normal output pathways of peripheral nerves and muscles."*

Since 1999, significant medical and scientific resources have been dedicated to the development of BCI. Today, BCI is the only assistive technology that allows communication with locked-in patients, such as patients with Amyotrophic Lateral Sclerosis (ALS) (see example on Figure 1.3), and as a promising technology for control of neuro-prosthetic devices for individuals with spinal cord injuries (see example on Figure 1.4) [46].

LIEBER-HERR-DAMBMANN-DAS-IST-MEIN-DRITTER-BRIEF-
DEN-ICH-NUR-MIT-MEINEN-GEDANKEN-GESCHRIEBEN-HABE.-
WELTWEIT-GIBT-ZWEI-MENSCHEN-DIE-DAS-KÖNNEN.-DAFÜR-
WERDE-ICH-VON-UNI-TÜBINGEN-DREIMAL-PRO-WOCHE-
TRAINIERT.-DAS-TRAINING-IST-HART-MACHT-ABER-
SPASS.-ICH-FREUE-MICH-ÜBER-IHREN-BESUCH.-
HERZLICHE-GRÜßE-AN-IHRE-FRAU.-IHR-JÜRGEN-BOROWICZ-

*Figure 1.3: A letter written using Thought Translation Device (TTD) BCI by the ALS patient (from [10]).*

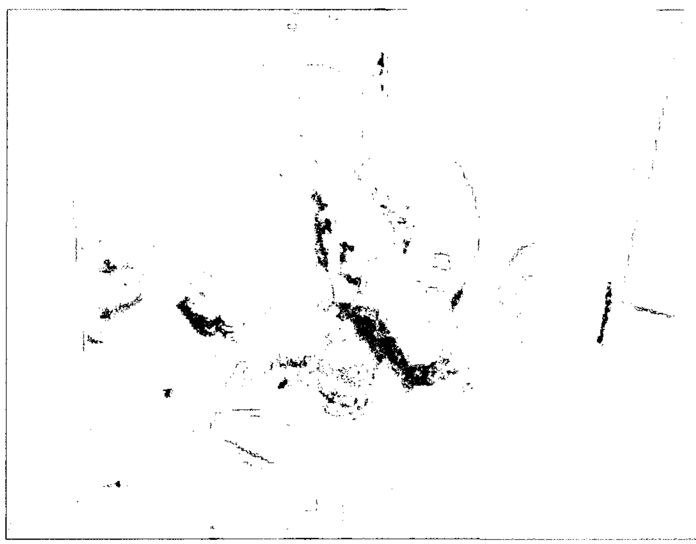

*Figure 1.4: Restoration of hand grasp in a tetraplegic patient with high spinal cord injury (from [36]).*

However, assistive technology is not the only application area for BCI. The BCI is also studied as an enabling technology for extending human abilities in operating complex equipment, art, gaming, and generally as a "prosthesis" for our creative minds [24].

## *1.4 Neurophysiological basis for BCI*

Due to the complexity of the brain signals and brain activities BCI systems have many different modalities [25, 47] based on:

1. Brain Signal

2. Experimental Strategy

3. Feedback

## Brain Signal

There are different types of measurements that can be used for recording brain activity. They include EEG, electrocorticogram (ECoG), direct cortical recordings, magneto-encephalography (MEG), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI). Generally, measurements of electrical or magnetic signals such as EEG, ECoG, and MEG have better temporal resolution while other methods based on recordings of chemical reactions have better spatial resolution [25, 31, 43]. Unlike other methods, MEG has significantly better spatial resolution in addition to temporal; however, the equipment is bulky and expensive and therefore not as practical. Cortical recordings also provide excellent temporal and spatial resolution and minimize noise due to artefacts. Direct cortical recordings are provided by implanted electrodes that measure electrical activity of a single neural cell or group of cells. However, surgical implants present considerable risk to the subject and have been tested primarily in animal experiments and in limited clinical studies on humans [24].

This research focuses on using EEG as a source of brain signals for communication and control [20]. EEG has high temporal resolution, readily available equipment, low acquisition cost and noninvasive nature.

## Experimental Strategy

The brain activity is very complex and recorded signals cannot be used to directly "guess" the thoughts or intentions of an individual. Instead, certain neurophysiological phenomena (Figure 1.5) are studied and used as triggering mechanisms for BCI [25, 46, 47].

Some of these phenomena are naturally occurring in the brain and therefore a subject does not have to be trained to generate them. BCI systems based on those naturally occurring phenomena are referred to as *pattern recognition approach.* An example of such "natural" phenomena is P300 ERP. The P300 is related to the focused attention of the subject to the task to be performed and represents a distinct increase in EEG amplitude that typically occurs 300 ms after the target stimuli is presented [31].

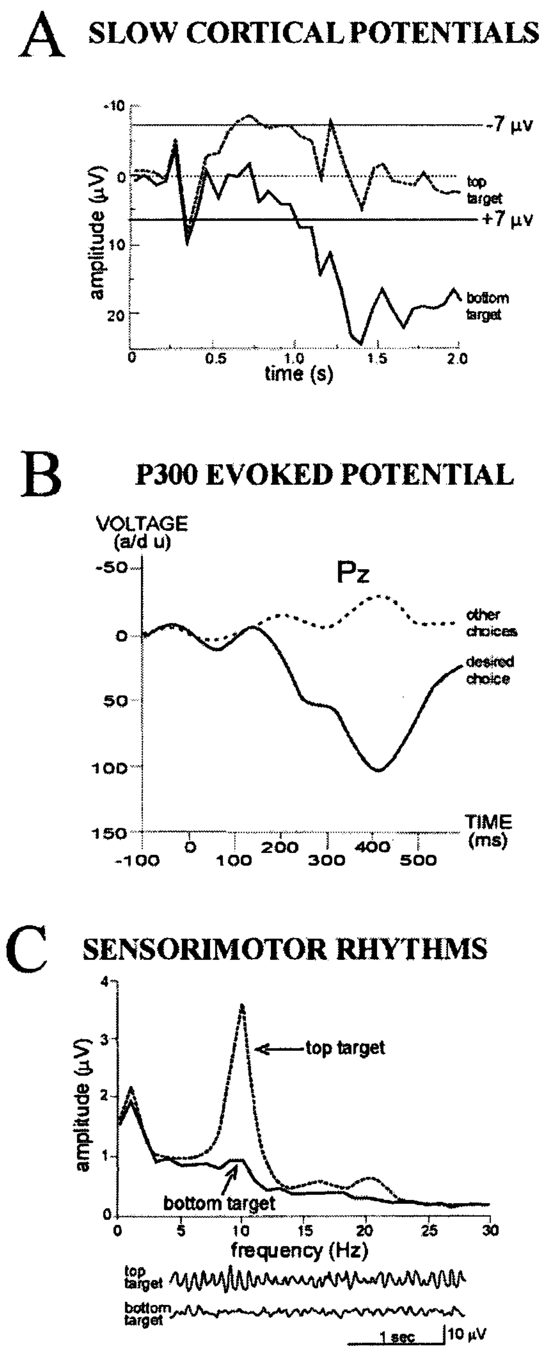


*Figure 1.5:Different neurophysiological basis for present day EEG based BCI systems (adopted from [47]): (A) Slow Cortical Potentials; (B) P300 ERP; (C) Sensorimotor rhythm (ERS/ERD).*

Other brain dynamics phenomena require training of the subject in order to generate dynamic features in the signal detectable by the BCI equipment. These BCI systems using the voluntarily control of brain activity are usually referred to as *operand conditioning approach.* Examples of this approach are the use of slow cortical potentials (SCP) [10, 25, 47] related to the emotional states, or *Mu* rhythm related to frequency synchronization (ESP) and de-synchronization (EDP) during motor imaginary movement [35, 46].

## Feedback

Feedback is crucial for every type of BCI [25, 27, 35]. For the operant conditioning approach feedback is essential since all adaptation in control is performed by the subject's brain. While not essential for pattern classification approach, feedback is valuable in improving learning and performance of the subject.

The feedback to the user of the BCI system is probably the least developed area in BCI research. The typical feedback is based on visual and auditory stimuli. However, the level of sophistication of the current BCI systems is very limited and not extensively studied compared to typical human-computer interface or ergonomic studies done when developing modern software and equipment.

## 1.5 Typical BCI Architecture

A general BCI system represents a closed loop system [25, 47]. The brain activity is monitored while collecting different biological signals. Those signals are then processed and feature extraction is performed. Based on the feature set extracted, a certain control action is performed. Finally, a feedback is provided to the user indicating a level of accuracy performed in desired action.

The typical BCI equipment consists [47] of a signal acquisition module, a signal processing module, an application interface, and an actual application feedback (Figure 1.6).

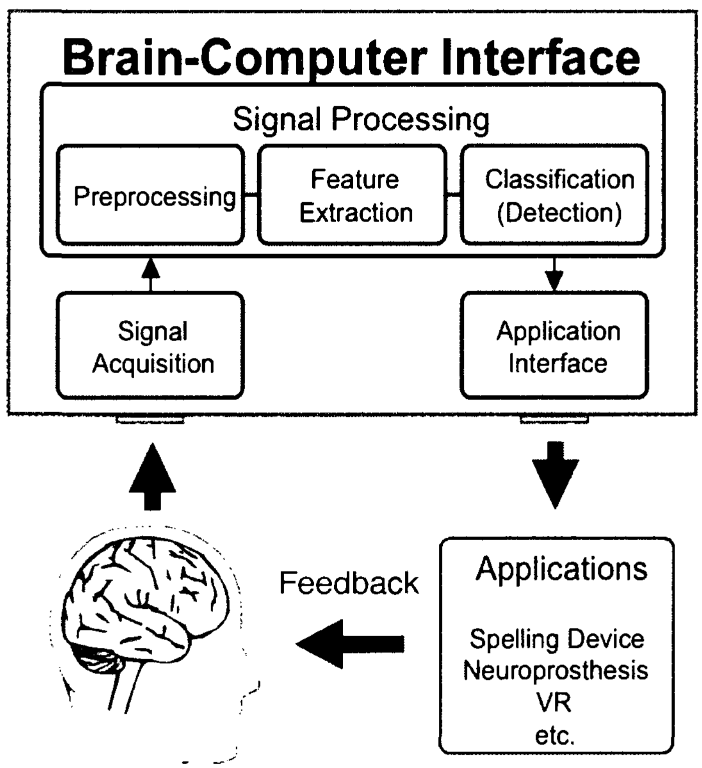


*Figure 1.6: The architecture of a typical BCI system.*

The implementation of the signal acquisition component depends upon the type of the recording used (e.g., EEG or cortical implants). Signal processing component includes preprocessing, feature extraction, and classification of the desired brain signal dynamic. Application interface links classification results with the actual BCI application and may include control of the cursor, virtual keyboard, control of the prosthetic device, or robot control. Finally, the application (see example Figure 1.7) provides the feedback for the subject regarding the outcome of the desired action as detected by the BCI system.


*Figure 1.7: An example of BCI application: BCI based pong game*

*(pictures taken at the II BCI workshop in Graz, 2004)*

## *1.6 Contribution*

The contributions of the thesis can be summarized as follow:

1. Investigation of the suitability of spoken words as a triggering mechanism for an auditory evoked P300 base BCI system.

2. Development of a BCILab – a multi-platform, open software framework for BCI research.

3. Implementation of spoken word auditory evoked P300 BCI system for symbiotic relationship between human operator and semiautonomous RSA.

Portions of the results of this research were published [20] and presented at the Instrumentation and Measurement Technology Conference, IMTC 2007, Warsaw, Poland, May 1-3, 2007.

## *1.7 Thesis Organization*

The thesis consists of four general sections. The first section provides motivation for the research and an introduction and background for the BCI research, EEG, evoked potentials (EP), ERP, and Support Vector Machines (SVM). The second section provides a literature overview of the state of the art in the field of BCI research related to P300 ERP and BCI robotic control. The third describes specific problems addressed in this research and details of the methods, experiments, and results. The fourth section is a discussion of research findings, conclusion of this study, and possible implications for future work.

These four general sections are subdivided in six chapters as follows:

- Chapter 1: background and motivation for this research as well as fundamental concepts and neurophysiological foundation in BCI.
- Chapter 2: background information related to brain's EEG, EP, and ERP and brief overview of SVM and it's discriminant function as a principal classification mechanism used in this research.
- Chapter 3: literature review of the present state-of-art in areas of P300 ERP based BCI and BCI in robotic control.
- Chapter 4: problem statement, methods, equipment used in experimentation, proposal for system architecture for the BCI test workbench, and details of performed experiments.
- Chapter 5: results of experiments performed using proposed BCI test workbench.

- Chapter 6: research findings, conclusions, potential implication of this research and possible future work in this area.

# 2 Background Information

## *2.1 A Brief Introduction to EEG, EP, and P300 ERP*

The first human EEG recording and the term electroencephalography can be credited to Hans Berger in 1924 (Figure 2.1). EEG represents a recording of electrical activity of the neurons in the human brain [26, 43]. The electrical activity is generated in the membrane of excitable cells, such as neurons in the human body. Neurons are linked together with their dendrites and axons. The electrical current through these is generated using ionic flow through membrane. EEG represents this electrical activity superimposed to the surface of the scalp. Therefore, EEG cannot record activity of a single neuron, but rather a combined activity of a large group of neurons [43].

*Figure 2.1: An early EEG recording, obtained by Hans Berger in 1924. The upper tracing is EEG, and the lower is a 10 Hz timing signal.*

EEG signal of a typical healthy human has an amplitude of approximately 100 □V and frequency range from 0.5 to 100 Hz [43]. The signal-to-noise ratio (SNR) of EEG signal is very poor. Therefore, high gain, high input impedance (at least 10 M□) biological amplifiers are normally used for recording.

The most commonly used EEG surface electrodes are silver/silver-chloride. They serve as transducers that convert ionic current to electron current, which is then amplified and recorded by the equipment. The electrodes are usually mounted on an EEG cap based on recommendation by International Federation of EEG Societies 10-20 standard (Figure 2.2). The EEG cap fulfills two needs: movement artifacts have to be reduced and the electrode placement has to be standardized.

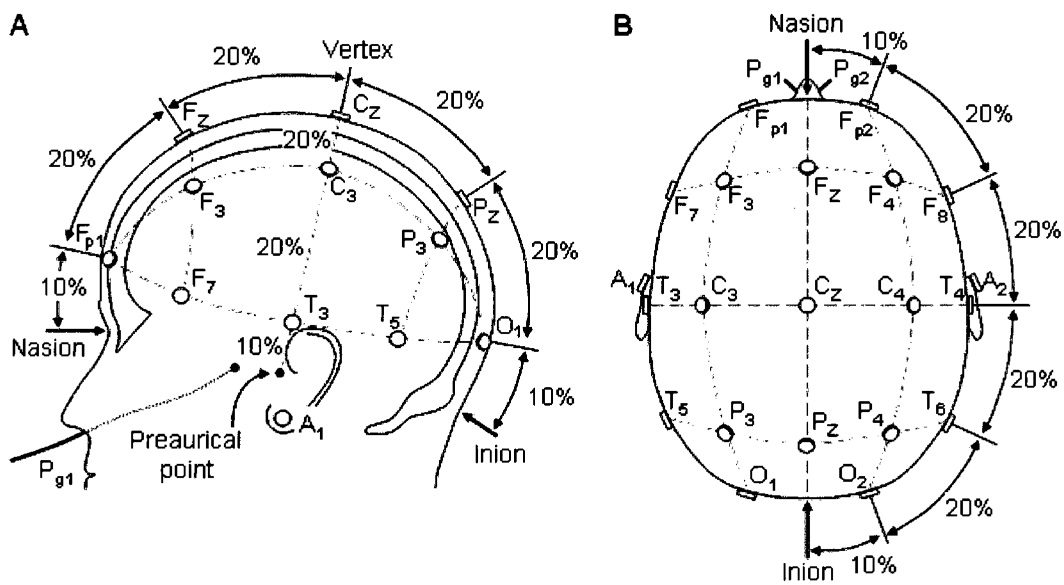


*Figure 2.2: The International Federation of EEG Societies 10-20 system seen from (A) left and (B) above the head (from [26]). A- Ear lobe, C - central, Pg - nasopharyngeal, P - parietal, F - frontal, F - frontal polar, O - occipital.*

There are several ways to configure electrodes in relation to the referenced signal. The most common methods are reference and bipolar recordings. In a reference method, signal from each electrode is usually referenced to the common, distant electrode (i.e., ear lobe A1/A2). The bipolar recording differential measurements are made between successive electrodes.

EEG has many features that are of interest for the BCI. All of those features are based on the varying potentials modulated on top of typically noisy brain activity. The potentials of interest for BCI can be divided in two large groups: spontaneous rhythmic potentials and potentials induced by particular event such as sensory stimulus or movement [47, 48]. Taxonomy of brain potentials used in BCI research is presented in Figure 2.3.

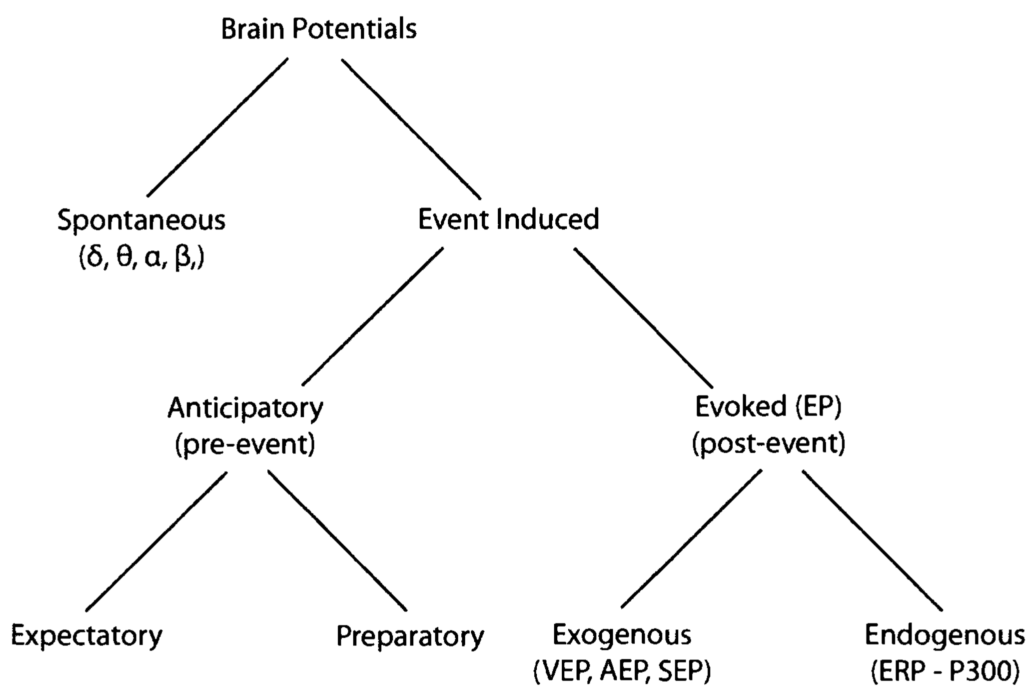


*Figure 2.3: A taxonomy of brain potentials of interest for BCI research (adapted from [13]).*

Spontaneous rhythmic activity of the brain is perhaps the most generally known feature of EEG and it was the first characteristic of human EEG that was studied [5, 11]. Typical rhythmical activity and associated frequency ranges are given in Table 2.1.

*Table 2.1: Spontaneous rhythmic brain potentials*

| Band | Frequency Range |
|---|---|
| delta (□) | < 3.5 Hz |
| theta (□) | 4 - 7.5 Hz |
| alpha (□) | 8 - 13 Hz |
| beta (□) | > 13 Hz |

Event potentials are of the special interest for BCI. The pioneering BCI research done by Vidal in 1973 was using visual evoked potential (VEP) as a communication mechanism [41]. Others are using contingent negative variation (CNV) potential [13]. However, the most interesting, and also the most popular potential for BCI research, is P300 ERP.

P300 ERP is a class of endogenous event potentials that are usually associated with higher thinking and they happen as a part of the normal recognition process of the brain. They can be evoked by any stimuli (visual, auditory, or somatosensory) or even without any external cause. The label P300 means that this potential represents a positive deflection in the EEG signal approximately 300 ms after stimuli onset (Figure 2.4).

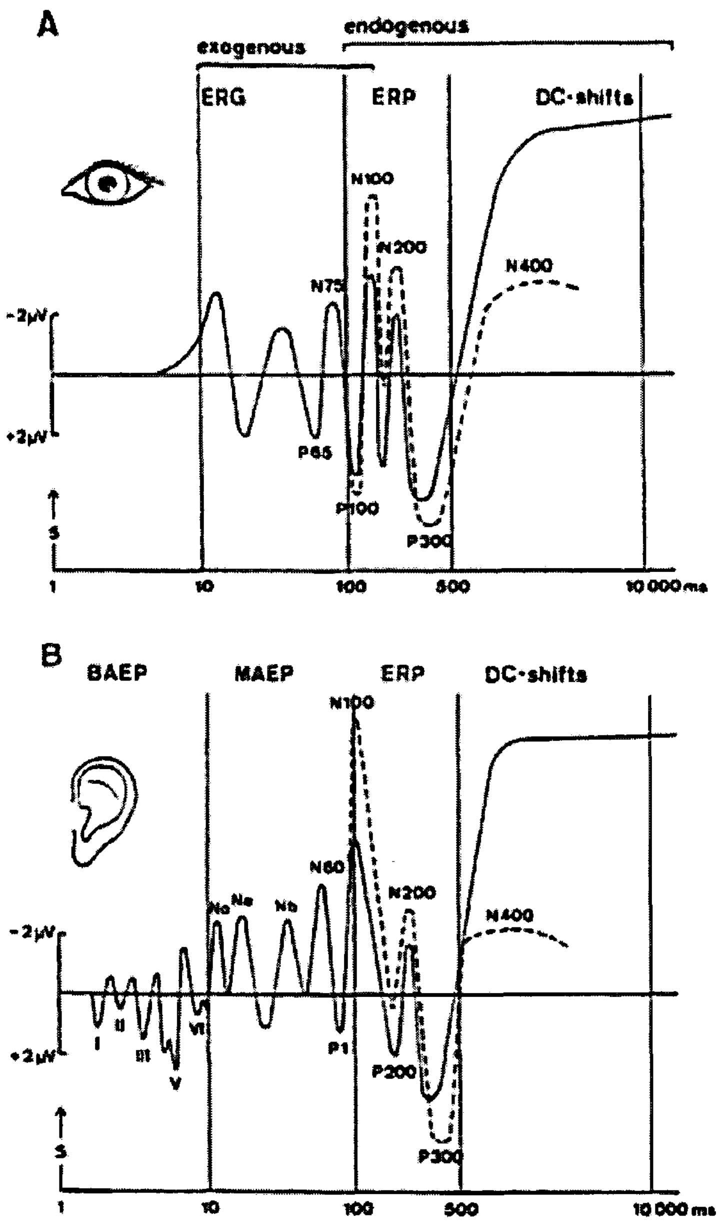


*Figure 2.4: ERPs evoked by visual (A) and auditory (B) stimuli (from [31]). Schematic potential traces on a logarithmic time-scale. Exogenous components are different for visual (electroretinogram – ERG) and auditory (brainsteam auditory potentials – BEAPs and midlatency auditory evoked potentials MEAPs) modalities. Endogenous components, including P300, are very similar in both modalities.*

Occurrence of P300 is mostly associated with the presentation of attended rare stimuli, which is also known as an "oddball" paradigm. If infrequent but expected target stimulus is presented, P300 corresponding to the target stimulus will be present in the subject's EEG. This feature of P300 was first used in BCI research by Farwell and Donchin in their now seminal 1998 publication [19]. They presented a spelling system using a matrix of letters organized in rows and columns (Figure 2.5). The rows and columns were then randomly flashing. The flashing of the column or row containing target (desired) letter then generated P300 that was detectable and used in the spelling process.

**MESSAGE**

**BRAIN**

**Choose one letter or command**

| | | | | | |
|---|---|---|---|---|---|
| A | G | M | S | Y | * |
| B | H | N | T | Z | * |
| C | I | O | U | * | TALK |
| D | J | P | V | FLN | SPAC |
| E | K | Q | W | * | BKSP |
| F | L | R | X | SPL | QUIT |

*Figure 2.5: Interface display used in P300 speller. The rows and columns of the matrix flashed alternately. The letters selected by the subject (BRAIN) are displayed at the top (from [19]).*

There are several characteristic of P300 that make it an attractive choice for present BCI research [1]:

- it is "naturally" present in the EEG signal without any training
- it is has a long history of clinical research and solid knowledge base
- it is present in EEG regardless of the differences in subjects age, gender, and intellect
- it is independent from the type of external stimuli and interface used to evoke it

The negative characteristic of P300 for BCI are [1]:

- it is driven by external sensory event and therefore BCI system usage must be synchronous; this is characteristic of all EP based BCI systems.
- it is very difficult to detect a P300 from a single trial due to very poor SNR.

## 2.2 *Classification Using Support Vector Machine*

Support vector machine (SVM) is a relatively novel approach in learning discriminant functions in classification problems from sparse training data while minimizing the expected probability of error. While SVM can be used in solving regression problems, this thesis will focus on classification only.

Theoretical background for SVM has been developed in statistical learning theory (SLT), structural risk minimization (SRM), and theory of Vapnik-Chervonenkis (VC) bound developed by Vapnik and his colleagues between 1960 and 1990 [15, 23]. Traditional classification adaptation algorithms, such as neural networks minimize the absolute empirical error or error squared. This approach could result in an overfitted model that has weak generalization ability (i.e., performance on unseen data). Unlike traditional approaches, SVM minimizes complexity of the model (i.e. VC dimension) while keeping the empirical error constant. The SRM and SLT prove that when VC dimension is low, expected probability of the error is low as well. The issue of underfitting and overfitting, as well as the idea of VC bound is illustrated in Figure 2.6.

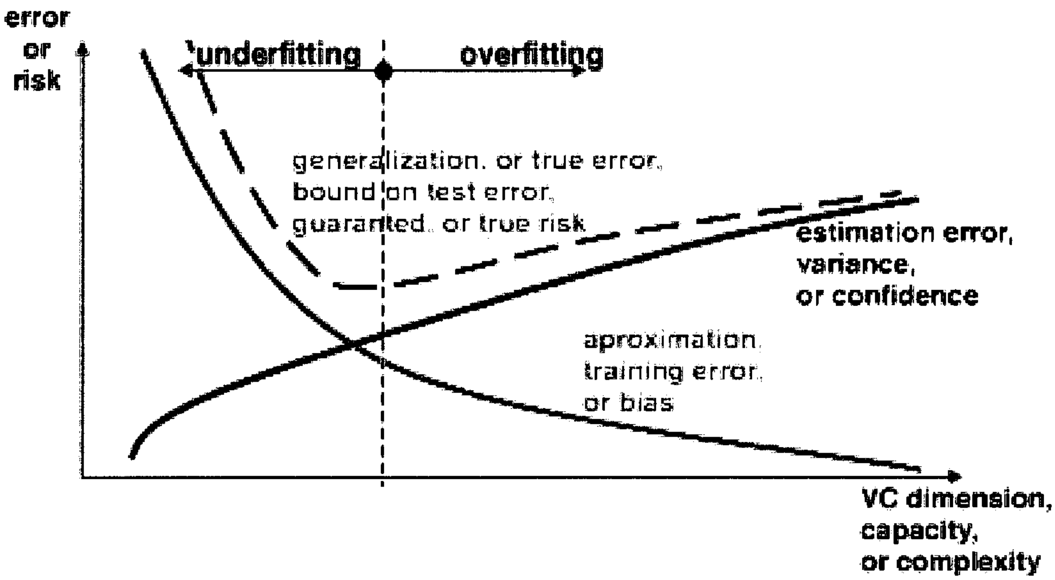


*Figure 2.6: Relationship between different errors and their dependency on VC dimension (i.e., complexity or capacity) of the model, labelled using varying terminology used in literature (adapted from [23]).*

The main idea behind SVM approach can be illustrated by examining a simple classification problem in two dimensional linearly separable input space [23]. While the two classes can be divided by the number of hyperplanes, it is intuitive that the one with a maximum margin between input vectors nearest to separating hyperplane (Figure 2.7) has the best generalization. There is a theoretical proof for the above intuitive observation in SRM. This type of linear classifier based on determining the maximum margin hyperplane is also known as the maximum-margin classifier.

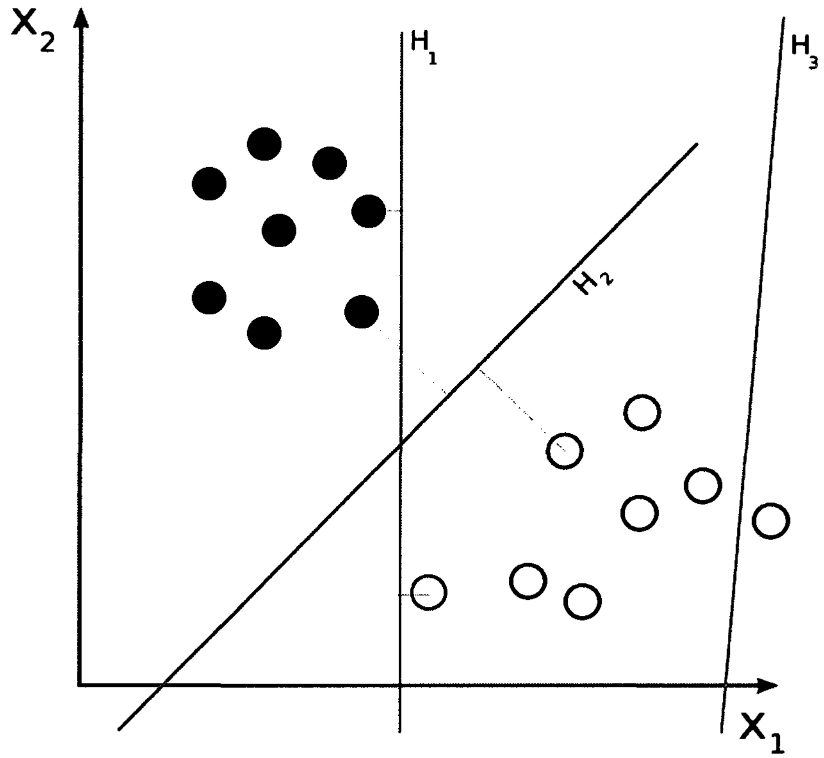


*Figure 2.7: Two simple linearly separable classes in $R^2$ with possible separating hyperplanes. Hyperplane $H_3$ is not properly separating classes. Both $H_1$ hand $H_2$ do separate two classes, but $H_2$ does it with the maximum margin (from Wikimedia Commons).*

Using more formal notation as in [15] and [23], the set of training data can be presented as

$$D = \{(x_i, c_i) | x_i \in \mathbb{R}^P, c_i \in \{-1, 1\}\}_{i=1}^{n} \tag{2.1}$$

where $c_i$ represents a label (i.e., -1 for one class and 1 for another) indicating the class to which point $x_i$ belongs. Each $x_i$ is $p$-dimensional real vector in input space. Assumption is made that training data set $D$ is linearly separable in space $\mathbb{R}^P$. Therefore, there is at least one hyperplane separating training data into two classes. The separating hyperplanes can be described as

$$w \cdot x - b = 0 \tag{2.2}$$

where $w$ represents normal vector perpendicular to the hyperplane and parameter $b$ determines the offset from the origin. Because of the assumption that training data are linearly separable, $w$ and $b$ can be chosen so that two parallel hyperplanes are separating the data while maximizing distance (margin) between them $\|w\|$ (Figure 2.8).

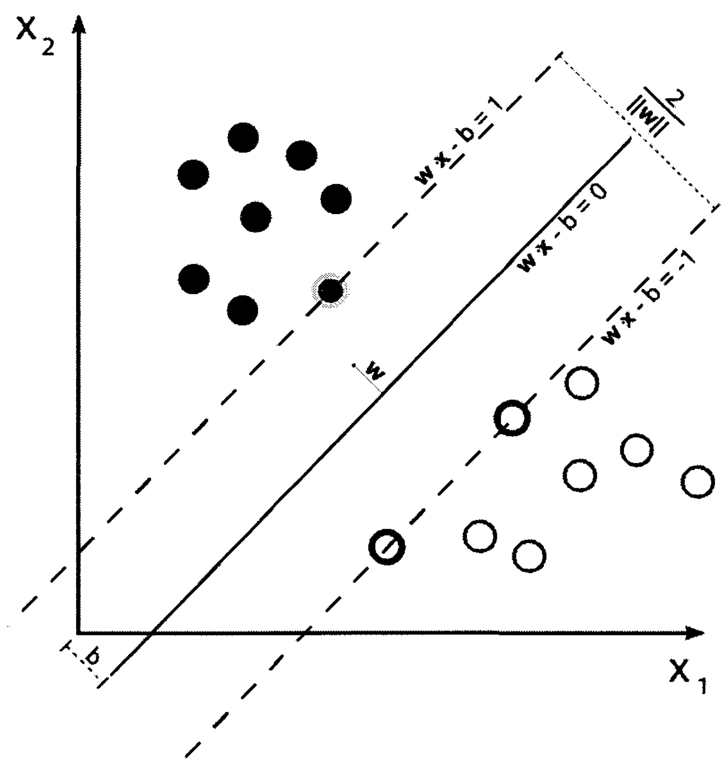


*Figure 2.8: Maximum-margin hyperplane and margins for a SVM trained with samples from two classes. Samples on the margin are called the support vectors (from Wikimedia Commons).*

The condition for these hyperplanes can be expressed as

$$\begin{aligned} w \cdot x - b &\geq 1 \\ w \cdot x - b &\geq -1 \end{aligned} \tag{2.3}$$

This can be written as an optimization problem

$$\begin{aligned} &\text{chose } w, b \text{ to minimize } \|w\| \\ &\text{subject to } c_i(w \cdot x_i - b) \geq 1, 1 \leq i \leq n \end{aligned} \tag{2.4}$$

In order to eliminate the norm of $w$ (which requires square root) we can express above optimization problem in its prime form as

$$\begin{aligned} &\text{minimize } \frac{1}{2}\|w\|^2 \\ &\text{subject to } c_i(w \cdot x_i - b) \geq 1, 1 \leq i \leq n \end{aligned} \tag{2.5}$$

This represents a classic nonlinear optimization problem with inequality constraints. This kind of optimization problems are normally solved by the saddle point of the Lagrange function [23] expressed as

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \{c_i(w \cdot x_i + b) - 1\} \tag{2.6}$$

where $\alpha_i$ represents Lagrange multipliers. The optimal saddle point $(w_0, b_0, \alpha_0)$ must be found in order to minimize $L$ with respect to $w$ and $b$ and maximize it with respect to positive $\alpha$. This problem can be solved in dual space (i.e., space of Lagrange multipliers).

The Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for the maximum of (2.6) because both objective function and constraints (2.5) are convex.

At the saddle point $(w_0, b_0, \alpha_0)$ derivatives of $L$ with respect to primal variables (i.e., $w$ and $b$) are equal to zero; thus, the dual variables Lagrangian can be expressed as

$$L_d(\alpha)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{n}c_i c_j \alpha_i \alpha_j x_i^T x_j \tag{2.7}$$

To find the optimal hyperplane dual Lagrangian must be maximized for positive $\alpha_i$.

This is now a standard quadratic optimization (QP) problem. The solution of this QP leads to following discriminant (i.e., optimal separating) hyperplane

$$d(x)=\sum_{i=1}^{n} w_{0i} x_{0i}+b_0=\sum_{i=1}^{n} c_i \alpha_i x^T x_i+b_0 \tag{2.8}$$

and final indicator function as

$$i_F=0=sign(d(x)) \tag{2.9}$$

It is interesting to notice that decision hyperplane depends only on the training data points on the optimal hyper plane margins (Figure 2.8). Those input points are called support vectors.

The presented solution is limited to linearly separable classification problems which are rare in practice. Soft margin method was introduced by Cortes and Vapnik in 1995 in order to address this issue [23]. In this method a separating hyperplane is constructed so that data are divided into classes as accurately as possible. The margin is then maximized on data that is correctly split while ignoring the points that are misclassified. This is achieved by adding a slack variable $\xi_i$ that represents a measure of misclassification of a training vector $x_i$.

Additionally a linear penalty function is introduced into original optimization problem (2.5) such that

$$\text{minimize } \frac{1}{2}\|w\|^2 + C\sum_i \xi_i$$
$$\text{subject to } c_i(w \cdot x_i - b) \geq 1 - \xi_i, 1 \leq i \leq n \tag{2.10}$$

where $C$ represents a penalty constant. The linear penalty function has a convenient property that the slack variable disappears from the dual problem and KKT conditions are still holding. Therefore, this new optimization problem can also be solved using QP techniques.

In order to adapt SVM for non-linear classification problem Boser, Gayon and Vapnik suggested using a kernel trick [23]. The kernel trick is computationally efficient method of mapping a non-linear classification problem space into higher dimension hyper-space where it might become linearly separable. It replaces a dot product between vectors with a kernel function. Specifically, when applying kernel trick to the decision function (2.8) we derive

$$d(x) = \sum_{i=1}^{n} c_i \alpha_i K(x, x_i) + b \tag{2.11}$$

where $K$ nonlinear kernel function replaces original dot product. Some of the typically used kernel functions [23] are listed in Table 2.2.

*Table 2.2: Typical SVM kernel functions.*

| Function | Type of Classifier |
|---|---|
| $x^T x_i$ | linear |
| $(\gamma x^T x_i + 1)^d, \gamma > 0$ | polynomial of degree $d$ |
| $\exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right)$ | gaussian |
| $\exp(-\gamma \|x - x_i\|^2), \gamma > 0$ | radial basis function (RBF) |
| $\tanh(\gamma x^T x_i + r)$ | sigmoid |

Once trained, SVM looks and behaves same as either linear classifier, polynomial models, neural networks, or fuzzy models.

The main advantages of SVM classifier over different traditional and heuristic classification methods are:

- a solid theoretical background in SLT and SRM that guaranties a global minima (i.e., unique solution).
- practical, as only a small number of independent parameters are needed for tuning of the model. For example, only one parameter (i.e., $C$ – error penalty) for linear SVM and two for RBF (i.e., $C$ and $\gamma$ ) are required.
- a generalization of several other classification methods depending on the kernel used. For example, perceptron (linear kernel), polynomial classifiers (polynomial kernel), multi-layer perceptron (sigmoid kernel), RBF and fuzzy classifier (RBF kernel).

Example of trained nonlinear SVM model for three input vectors is presented in Figure 2.9.

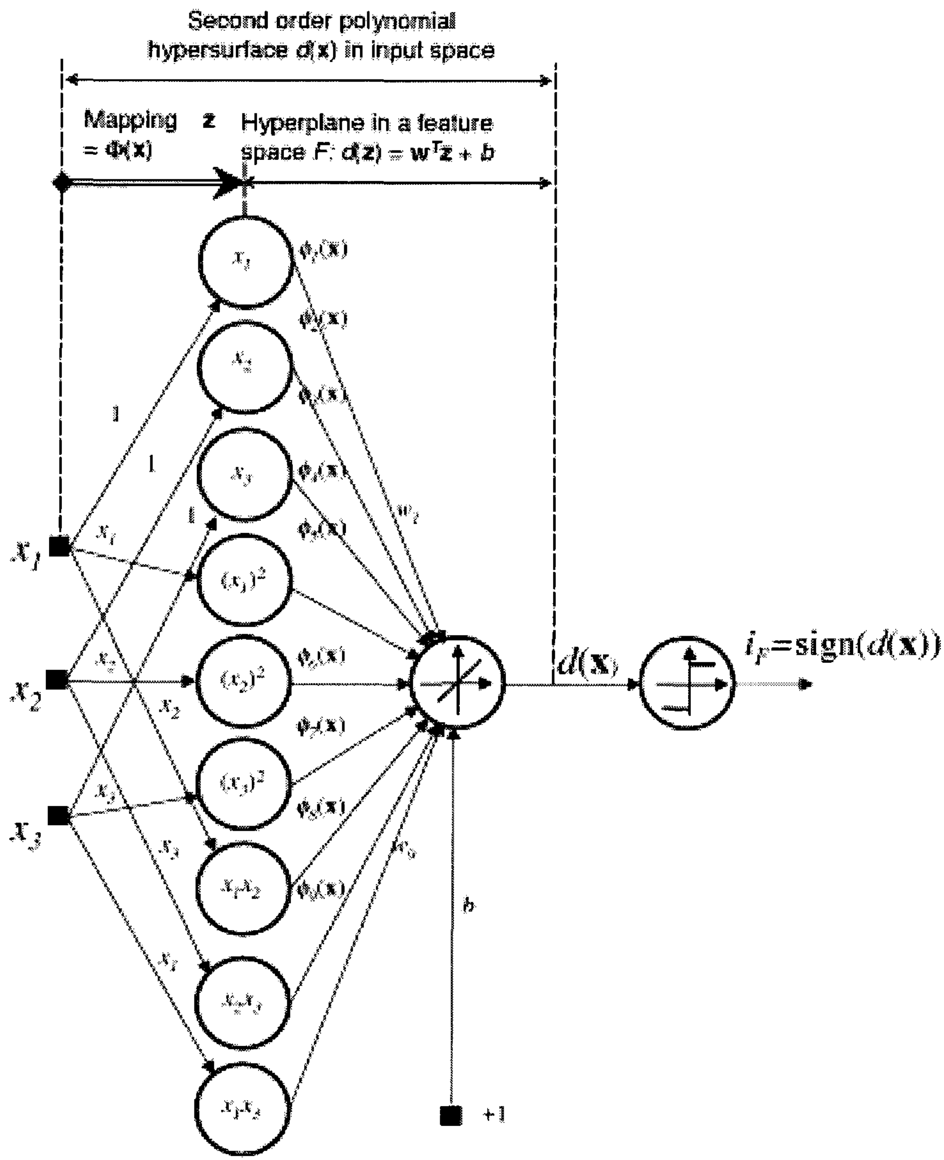


*Figure 2.9: Example of nonlinear SVM for 3 input vectors (from [23]).*

The main practical issue with using SVM is the memory requirement. The dot product or kernel function matrix can be very large for the real word problems as its dimension ( $N \times N$ ) directly depends on the number of training samples. Fortunately, this issue can be solved using decomposition techniques where large QP problems get decomposed to several smaller ones, effectively reducing the size of the required matrix [15]. Additionally, the variation of decomposition technique called sequential optimization method (SMO) [18] has been recently used to successfully both reduce the memory requirements and dramatically increase performance of the whole optimization algorithm. The SMO essentially decomposes optimization task to a problem of only two $\alpha_i$ parameters at each step; thus, it can be solved analytically while avoiding numerical QP optimization altogether.

# 3 Literature Review

## 3.1 State of the Art in P300 BCI

In order to validate signal processing and classification methods for BCI, several leading research groups organized the BCI competition in 2003 [11] and in 2005 [12]. Both competitions included a data set related to P300 BCI. The data sets were collected on several trails of a single subject in the first case, and multiple in the second case.

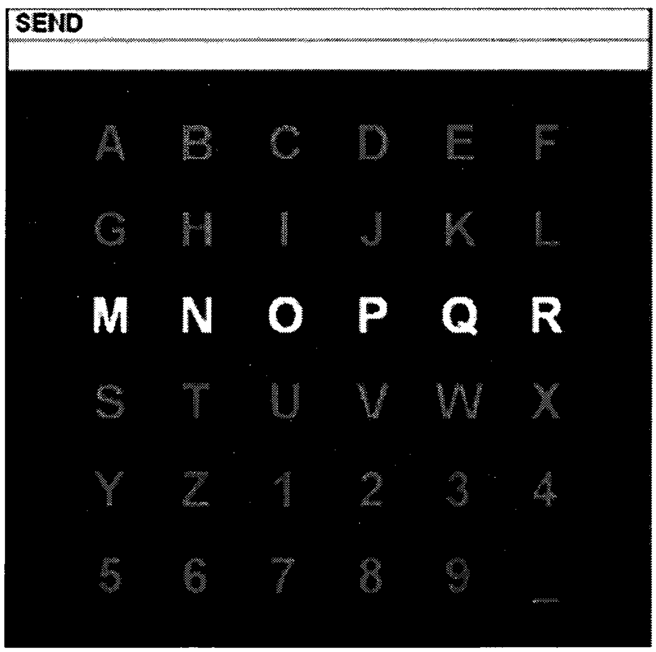


*Figure 3.1: P300 speller paradigm interface used in BCI Competitions.*

In the P300 speller paradigm used in the competition, the subject was presented with a 6 by 6 matrix of characters (Figure 3.1) [11, 12] similar to the one used by Donchin et al. [17], and originally by Farwell and Donchin [19]. In the experiment the subject was focusing attention on characters in a word (i.e., one letter at the time). All rows and columns of this matrix were successively and randomly flashing at a rate of 5.7Hz. Two out of 12 of those flashes contained the desired character (i.e., one for the row and one for the column with the desired letter) and consequently evoked P300. The goal of P300 based competition task was to use the provided data with associated spelled-out letters for the training of the classifiers and determine the letters spelled in the testing data set. The actual letters for the testing were not given prior to the competition. The competitor success was measured based on the percentage of success in "guessing" of the spelled-out letters in the testing data sets. Each letter spelled had 15 epochs of data collected.

In both 2003 and 2005 competition the winners in P300 classification category were classifiers using SVM. For 2003 competition, Kapper et al. [6] suggested using the nonlinear discriminant function (2.11) instead of the regular binary output of the SVM. This value represents a confidence or score that a single stimulus can determine the presence of P300. The maximum sum of confidence values over several epochs will determine the final desired symbol.

Therefore the sum of confidence for the row in P300 speller paradigm can be expressed as

$$d_r^{row} = \sum_{k=1}^{M} d\left(x_{rk}^{row}\right) \tag{3.1}$$

where $d\left(x_{rk}^{row}\right)$ represents the confidence value of the epoch from the *r*th row of the *k*th trial. The $M$ is number of epochs recorded for each letter.

The row with the maximum sum of confidence values is now

$$r_{sel} = \underset{r}{arg\,max}(d_r^{row}),\, r = 1, \ldots, 6 \quad (3.2)$$

The same principal is applied to the columns, which can be expressed as

$$c_{sel} = \underset{c}{arg\,max}(d_c^{col}),\, c = 1, \ldots, 6 \quad (3.3)$$

The intersection of the row and the column with the maximum sum of the confidence values is then chosen as a desired letter for spelling.

In their work Kapper et al [22] and Meinicke at al [28] suggested using a Gaussian kernel in order to capture nonlinear aspects of the problem. However, in 2005 competition, Rakotomamonjy et al [37] recognized the high variability of EEG signals between different subjects and trials and recommended using the linear kernel in order to prevent over-fitting of the classifier during the training.

In addition to P300 speller paradigm [17, 19] a significant amount of research was done in application of P300 based BCI in navigation and manipulation in a virtual environment [5, 6] as well as embedding of P300-like interaction in a regular user interface [8]. Recently, some researchers used P300 BCI for manipulation and control of wheelchair [39] and a humanoid robot [9].

While application and classification methods might be different, the common approach in all of this research is that P300 is evoked using a visual stimulus (Figure 3.2). Typically flashes of words symbols on the screen [8], or in the case of the virtual world or robot navigation flashes of different objects [5, 7, 9], are used to evoke P300. This is quite contrary to the clinical setting for the research of P300 auditory evoked P300, which is frequently used as well [31].

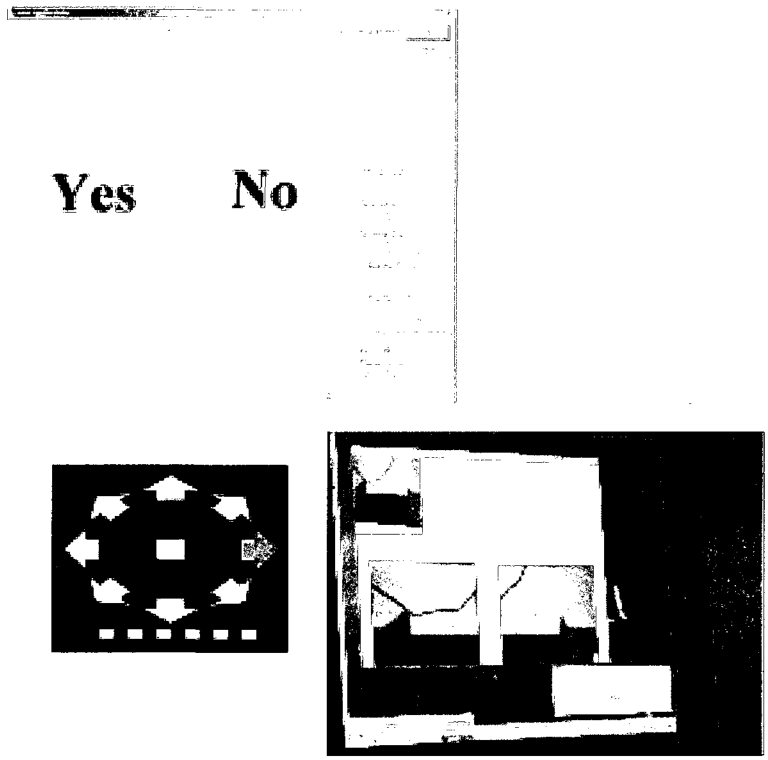


*Figure 3.2: Typical visual paradigms used in current P300 based BCI systems.*

Some limited research has been done in assessing the usability of the auditory evoked P300 in comparison to the visually evoked by [40]. However, this research was focused on the application of P300 BCI as an assistive technology.

## *3.2 BCI for Robotic Control*

The first publication about noninvasive BCI control of the robot resulted from research by Millán et al in 2003. This research focused on using a BCI paradigm based on classification of different mental tasks. The BCI was non-synchronous as it did not require an external stimulus. Instead, an operator was required to focus on one of the several mental tasks, such as: "relax", imagination of "left" and "right" hand (or arm) movements, "cube rotation", "subtraction", and "word association".

Millán and his colleagues realized that with slow bit rate in BCI communication, a semiautonomous robot was required instead of direct control. In their experiment the robot was traversing in a maze (Figure 3.3) using either "follow-left-wall" or "follow-right-wall" strategies and human operator was directing the robot as per which particular strategy should be used next. This simple method generated results in which BCI control was comparable to the manual control of the robot using presented strategies.

A study of P300 based BCI system for control of humanoid robot was published by Bell et al in April 2008 [9]. The autonomous humanoid robot demonstrated a complex behaviour by walking to a specific target and picking up a desired object as instructed by the human operator using BCI. The robot was equipped with a camera capable of recognizing objects of different colors. The view of the camera was presented to the operator (Figure 3.4). The operator was also presented with two choices of objects with different flashing rectangles around them. Elicited P300 was translated into instructions for the robot for which object to approach and pick.

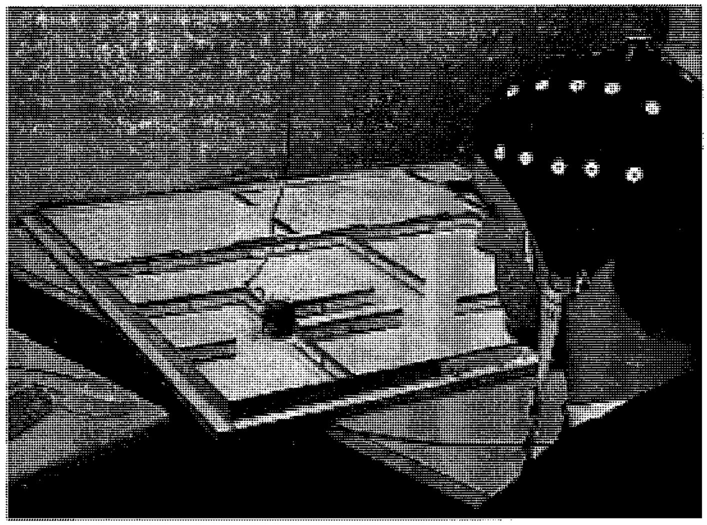

*Figure 3.3: A robot controlled by mental-task BCI and sample maze used by Millán et al [30].*

In both [29] and [9] BCI controlled robotic systems the slow communication channel from BCI is compensated by using semiautonomous robotic systems. This eases the burden from BCI direct control of the system to more adequate BCI as the guide for the overall system behaviour as suggested in [20].

It is also clear that the interface used in P300 BCI humanoid robot control [9] has higher degree of complexity. This is mainly due to the fact that visually evoked P300 is used. In this solution the operator is limited to the view of the screen rather than having an option to freely observe environment as in [30].

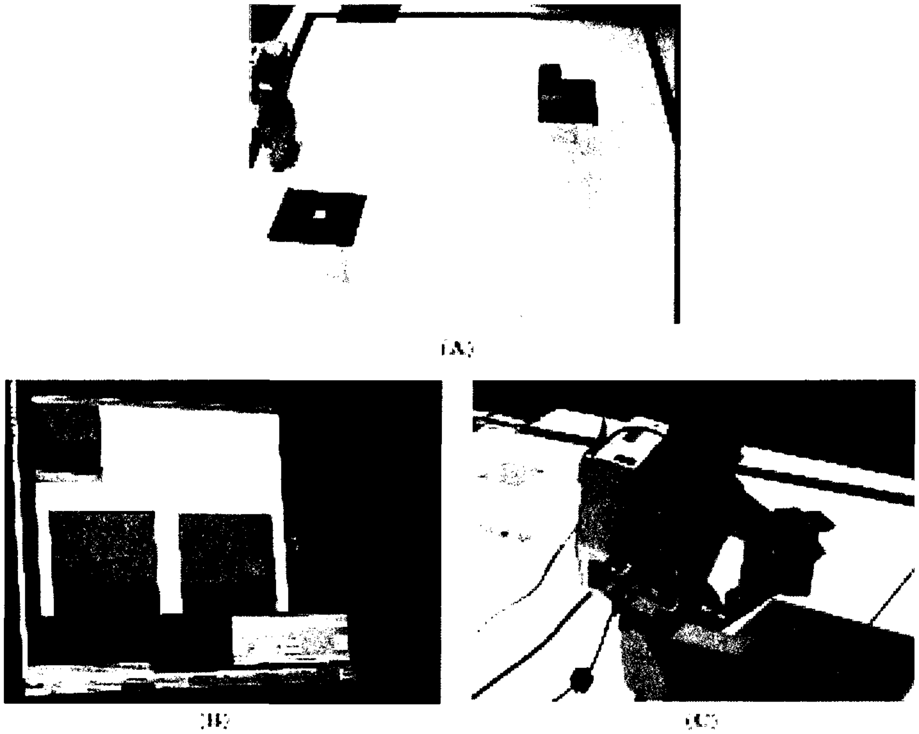

*Figure 3.4: A humanoid control interface used by Bell et al [9]. The robot arena is presented in A. The robot's eye view and choice of objects for selection by operator can be seen in B. After P300 was detected robot is instructed in picking the desired robot as presented in C.*

The issue with visual stimuli in the P300 generator in real environments or virtualized environments as presented above is that it is typically noisy and saturated with other relevant information for the task at hand. Additionally, visual stimuli implementation is usually complex and requires synergy of all other visual information for presentation on a single device (i.e., computer screen).

# 4 Symbiotic Robot Control Using Spoken Words Auditory Evoked ERP Based BCI System

## *4.1 Problem Statement*

This research is motivated by recent developments in BCI field, especially in BCI for robotic communication and control. It has a novel perspective in considering BCI as an enabling technology in further human-computer symbiosis.

The main goal of this research is to investigate the suitability of spoken words auditory evoked P300 base BCI system for communication and control of semiautonomous robotic sensor agents in a controlled environment operated by healthy human subjects. The secondary goal is to establish a standard software workbench or framework for future BCI research.

Currently, P300 based BCI systems, including systems for robotic control, are all based on visually evoked ERP. This is contrary to the established clinical procedures where visual and auditory evoked P300 are both used. Visually evoked P300 represents additional challenges and increases complexity in application interface used in BCI. This research is proposing a novel approach in replacing visually evoked ERP with auditory stimulus as spoken words related to the task at hand. This approach would effectively free operator's visual perception and allow for a focus on a task at hand rather than BCI communication mechanics.

The spoken word auditory evoked P300 BCI would provide a relatively simple and inexpensive interface of the system as no special equipment other than speakers or headphones are required. It also allows for selective delivery of stimuli to selected operators without having to rely on computer monitor or heads-up-displays.

## *4.2 Research Outline*

The work for this research is divided into three distinct phases in order to perform a systematic approach in archiving the set goals:

1. Construction and configuration of the BCI system framework.
2. Collection of P300 EEG data for preliminary off-line analysis.
3. Performance of online experiments in human-RSA collaboration using P300 BCI.

The first phase provides a foundation and address the secondary goal of the research. The second phase investigates suitability of spoken word auditory stimulus as a generator of P300 feature in subject's EEG. The third phase uses the results from the first two and includes the description of live experiments with a human subject guiding a semiautonomous RSA. The last two phases ultimately answer the primary question of this research.

## *4.3 BCI System Framework*

One of the prerequisites for this research is to construct BCI framework so that reliable and reproducible experiments can be performed. Conceptual model of BCI framework and it's relationship with other components in the system can be presented using UML deployment diagram as in Figure 4.1.

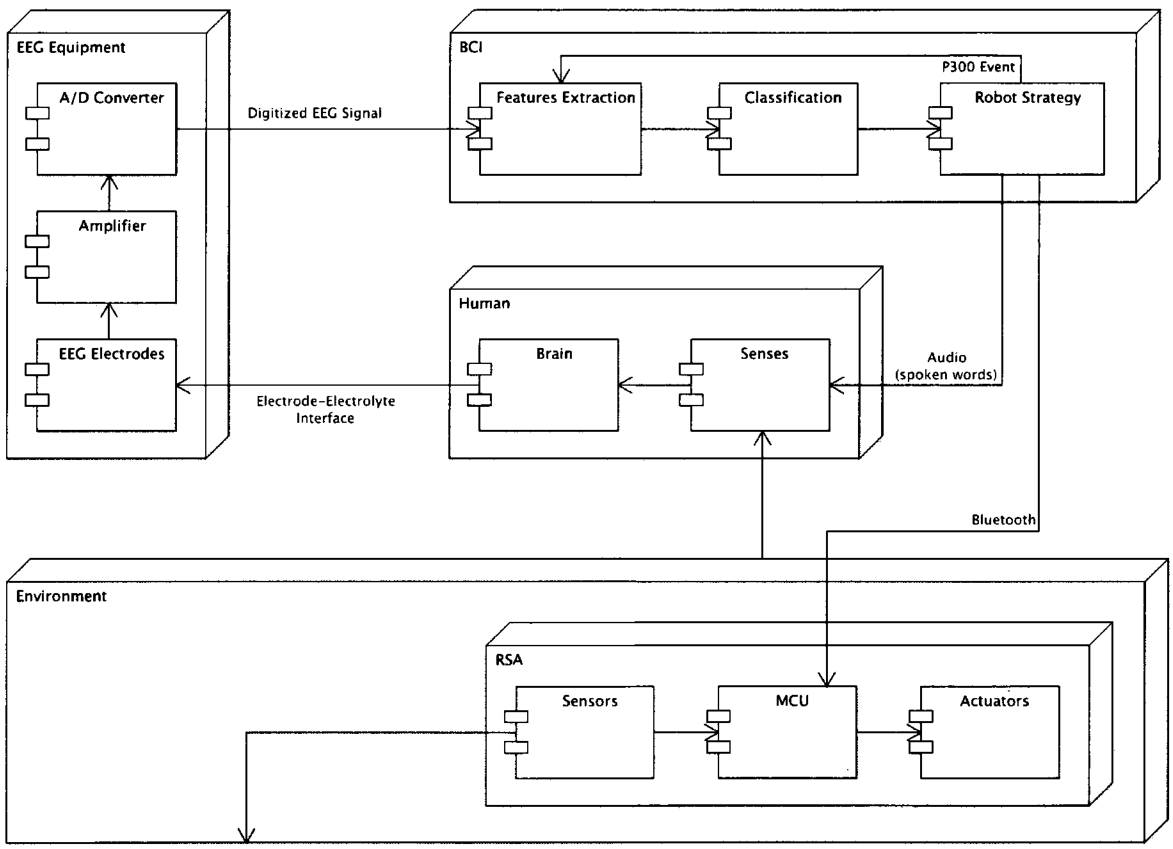


*Figure 4.1: Components of BCI System for control of RSA.*

The EEG equipment performs acquisition of raw EEG signals of the human operator's brain activity evoked by auditory stimulus (spoken words) and generated by the BCI system.

Those signals, amplified and digitized, are then passed to the BCI System where feature extraction and final classification are performed. Based on classification of the selected EEG feature, the appropriate robot strategy is selected. The choice of the strategy is then sent to the RSA and confirmed to the human operator using auditory signal (spoken words). The RSA stores the new, selected strategy and uses it when the next decision point (i.e., obstacle) is reached. The RSA performs obstacle detection using onboard sensors. The current, chosen strategy is stored in onboard memory and executed by RSA's MCU unit that collect sensory data and controls robot's actuators (i.e., wheels). The full feedback loop is achieved by human operator's observation of RSA performance in the environment.

The EEG equipment used include standard and commercially available EEG cap (Figure 4.2), biological signal amplifier (Figure 4.3), and A/D converter (see Table 4.1).

*Table 4.1: EEG equipment used in BCI system.*

| Component | Equipment | Vendor |
|---|---|---|
| EEG Cap | E1 Electro Cap | Electro-Cap International, Inc. |
| EEG Amplifier | PSYLAB EEG8 | Contact Precision Instruments Inc. |
| A/D Converter | NI6009 | National Instruments |

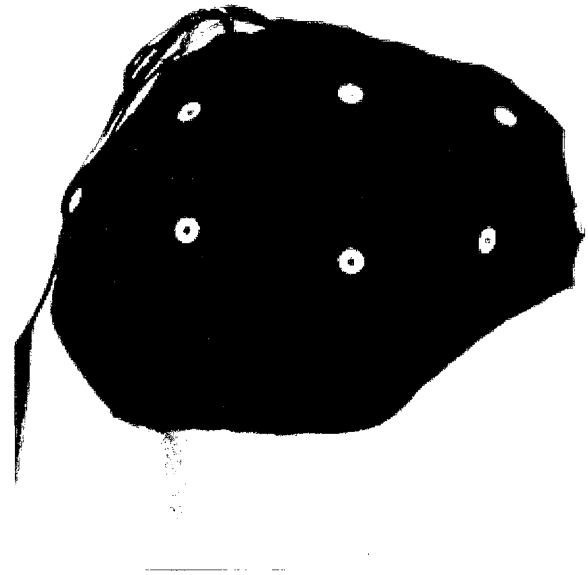

*Figure 4.2: EEG Cap by Electro-Cap International, Inc. used in BCI setup*

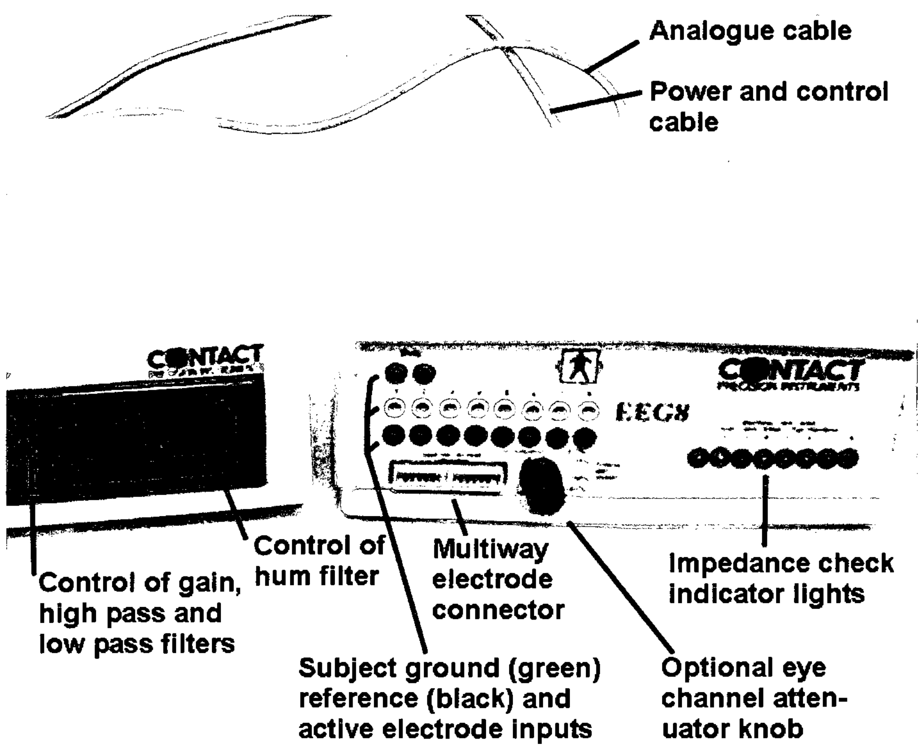


*Figure 4.3: EEG8 biological amplifier from Contact Precision Instruments Inc.*

RSA included wheeled robots equipped with infrared or ultrasound sensors for obstacle avoidance and bluetooth for communication with the system. Wheeled robots were built using commercially available kits (see Table 4.2). Figure 4.4 illustrates one of the robots used in the experimental setup.

*Table 4.2: Robot kits used in BCI system.*

| Robot Kits | Vendor |
|---|---|
| Mindstorms NXT [2] | The LEGO Group |
| Bioloid Kit [4] | ROBOTIS, Inc. |

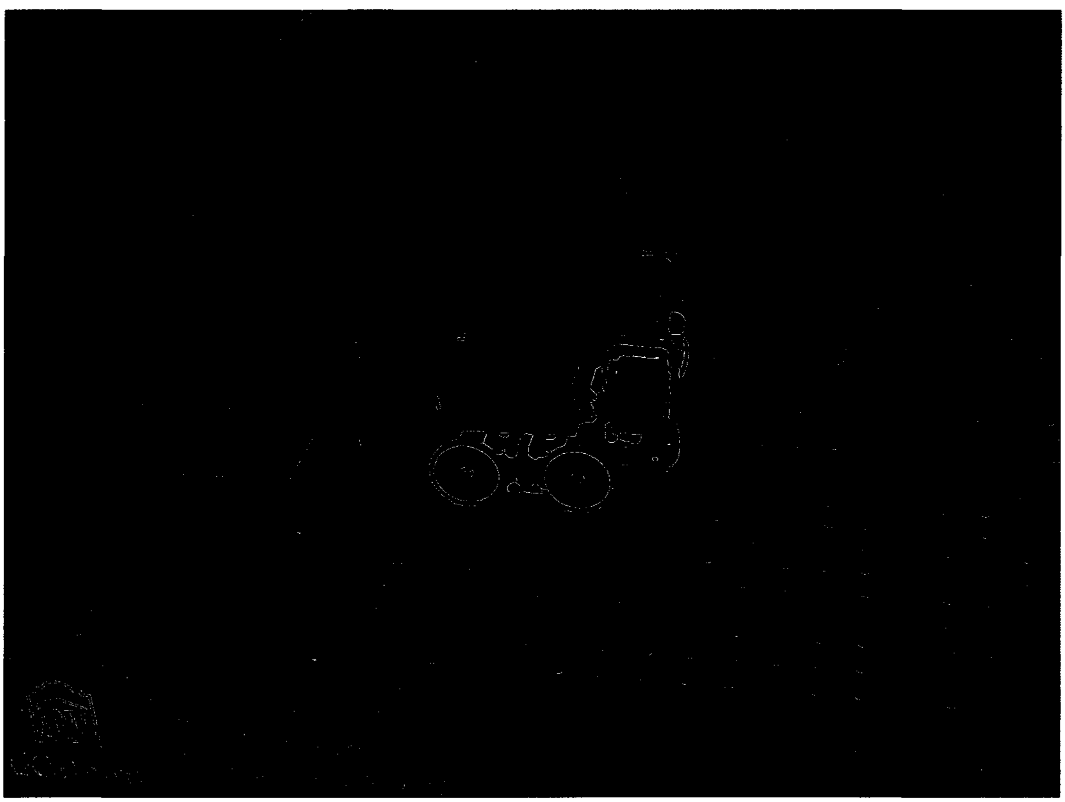

*Figure 4.4: One of the robots based on Bioloid Kit and used in the experimental setup.*

The BCI software framework *BCILab* was developed specifically for the secondary goal of the research. *BCILab* is implemented using *Java* language and on top of *NetBeans Platform*. The *NetBeans Platform* is a generic software application framework for *Java*. It provides the services common to almost all mature desktop applications [45], such as window management, menus, settings and storage, update management, and file access.

*BCILab* is developed as a multi platform application. For this research it was used on MS Windows and OS X operating systems. The software developed consist of different components (or modules in *NetBeans* terminology), each with a distinct purpose. Several open source projects were used for the development of *BCILab* (Table 4.3).

*Table 4.3: Open Source projects used in BCILab.*

| Project | Short Description |
|---|---|
| NetBeans Platform | A framework for the development of desktop applications [3] |
| FreeTTS | A speech synthesizer written entirely in Java [42] |
| JChart2D | A precise, runtime visualization of data in Java [44] |
| libSVM | A library for Support Vector Machines [16] |
| rxtx | A library for serial and parallel communication in Java [21] |

List of the key modules in *BCILab* and their dependency is presented in UML component diagram (Figure 4.5).

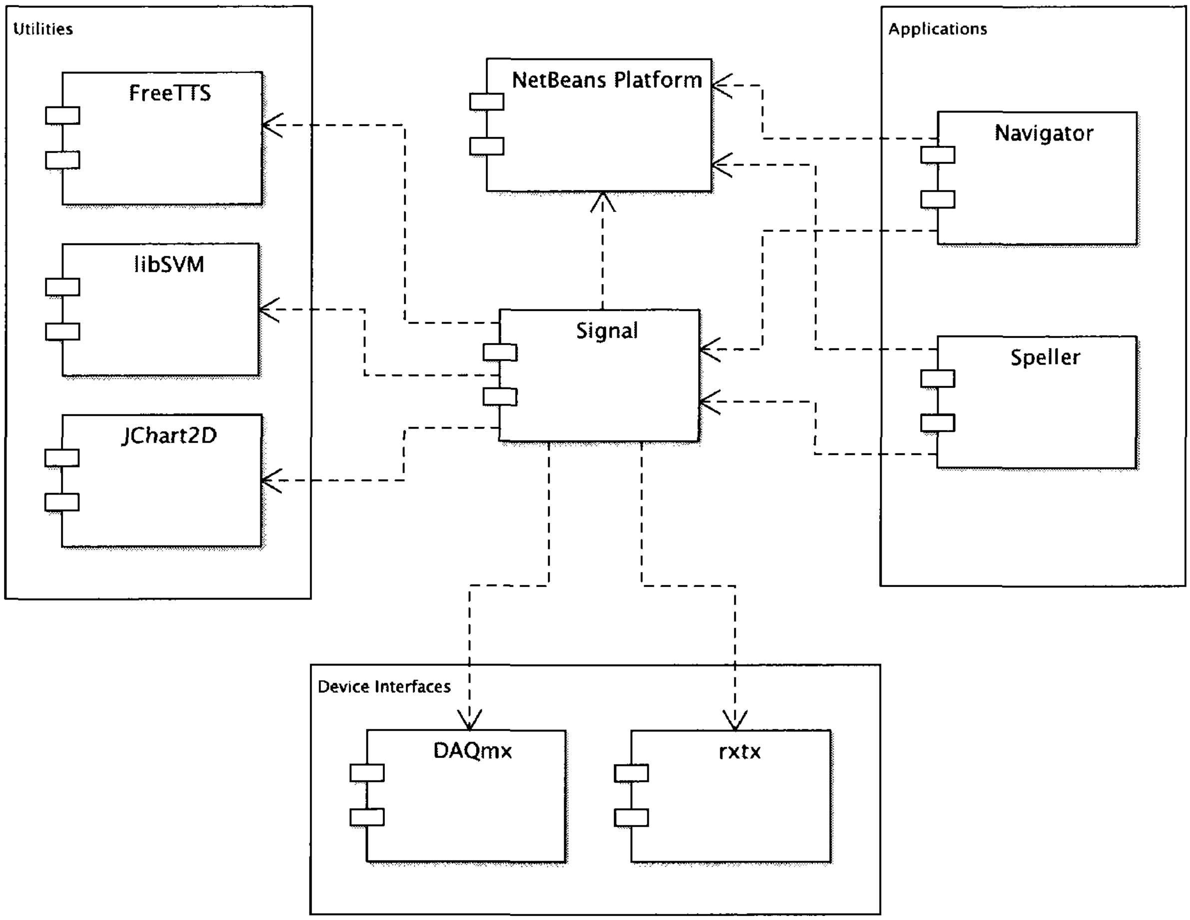


*Figure 4.5: BCILab modules and their dependencies.*

The core module of the system is *Signal*. This component contains all classes and logic for EEG signal filtering, scaling, feature extraction, classification and communication, and control of RSA. This component is loosely modelled based on physics of signal propagation which is reflected in names and purpose of all implemented *Java* classes.

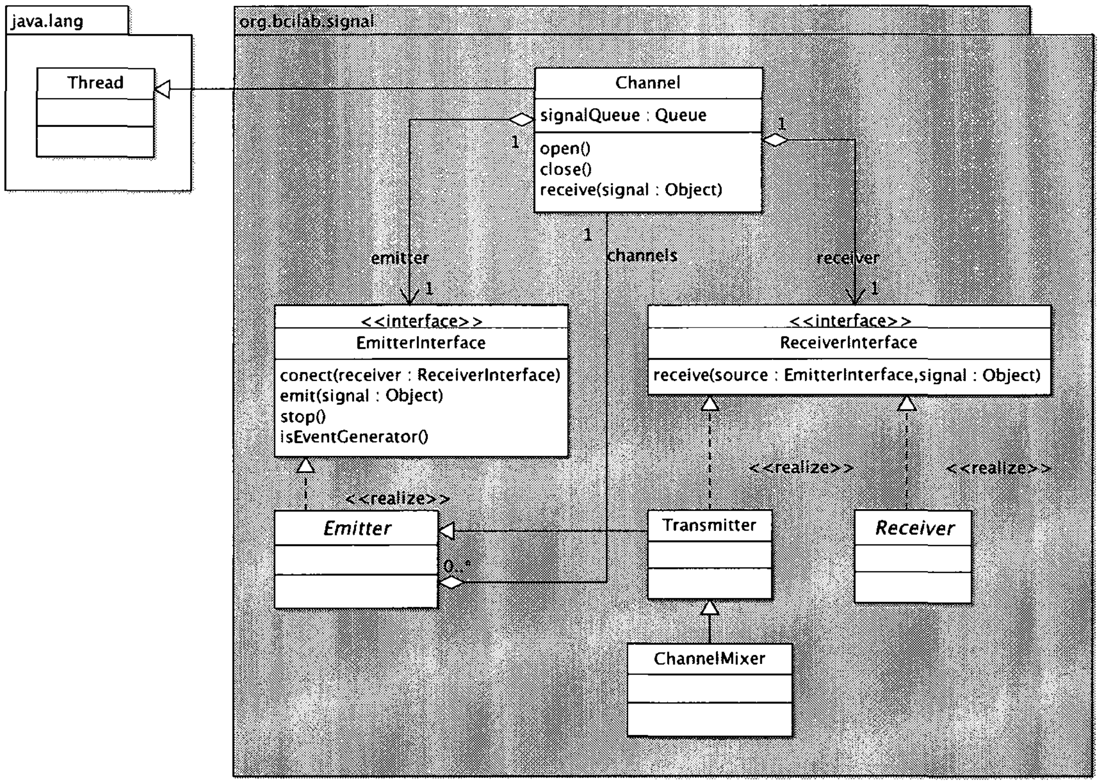


*Figure 4.6: Core classes of BCILab Signal module.*

The most important classes of the *Signal* module are *Emitter*, *Transmitter*, *Receiver* and *Channel* (see Figure 4.6). *Emitter* usually emits the signal by broadcasting it to the appropriate transmitters or receivers. *Transmitter* both receives the signal and emits it further. *Receiver* is the final destination of the emitted signals. The *Channel* resembles a *medium* through which signal travels and it is used by the *Emitter* when broadcasting the signal. The *Channel* has a capacity (i.e. size of the associated *Queue*) and is implemented as a *Thread* in order to execute each instance of the class (broadcast) concurrently.

All three basic components, *Emitter, Transmitter,* and *Receiver,* have specific implementation in different components used in the varying system configurations.

Different BCI application paradigms are implemented as separate modules and are based on an *Application* class that provides common functionality for loading and saving component configuration and parameters (Figure 4.7). Two paradigms were implemented. The original P300 speller application was used to validate the overall framework by recreating original Farwell and Donchin experiment [19]. The RSA communication and control was implemented as the *Navigator* paradigm.

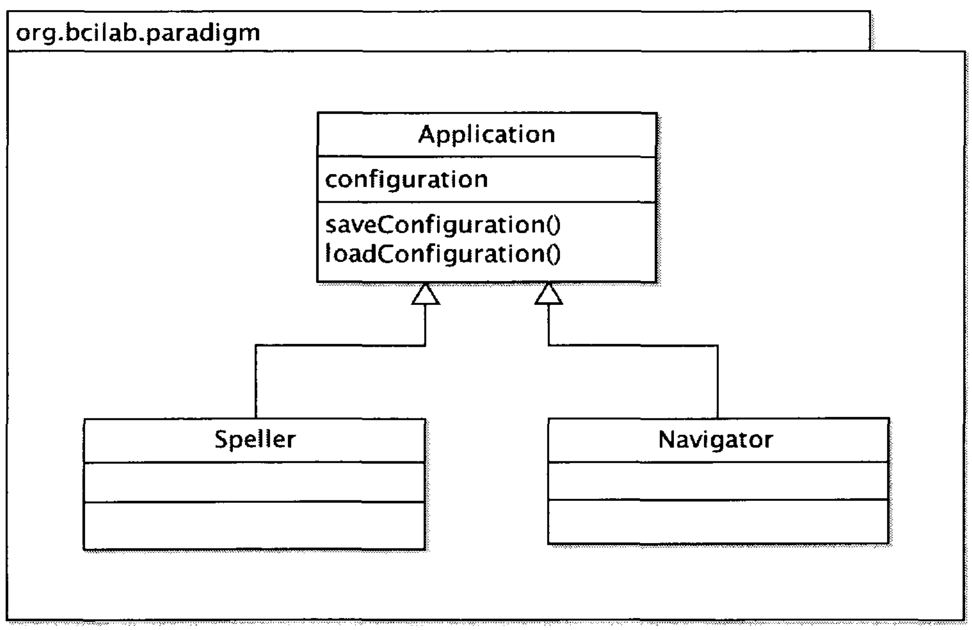


*Figure 4.7: Different BCI application paradigms implemented in BCILab.*

The class diagram of all *Signal* components and their relationship to application paradigms used in experiments is illustrated in Figure 4.8

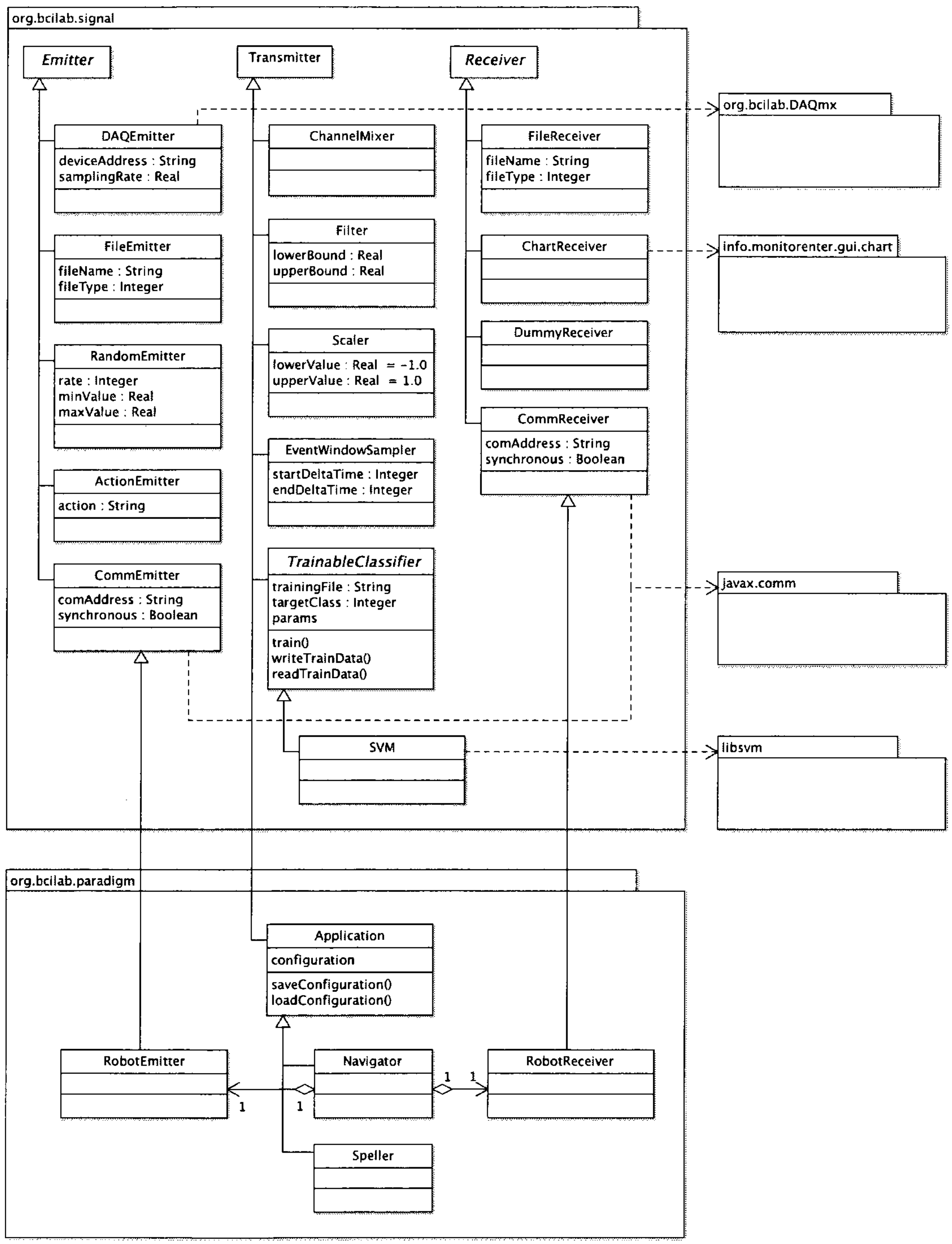


*Figure 4.8: Signal components.*

Each application might have a unique user interface appropriate for the implemented paradigm. The example of different paradigms and view of the raw EEG data in *BCILab* is presented in Figure 4.9.

The user interface for *Navigator* is only used during classifier training and monitoring during experiments. It is not used by the subject during online experiments since the interface was based on audio (i.e., spoken words) stimulus.

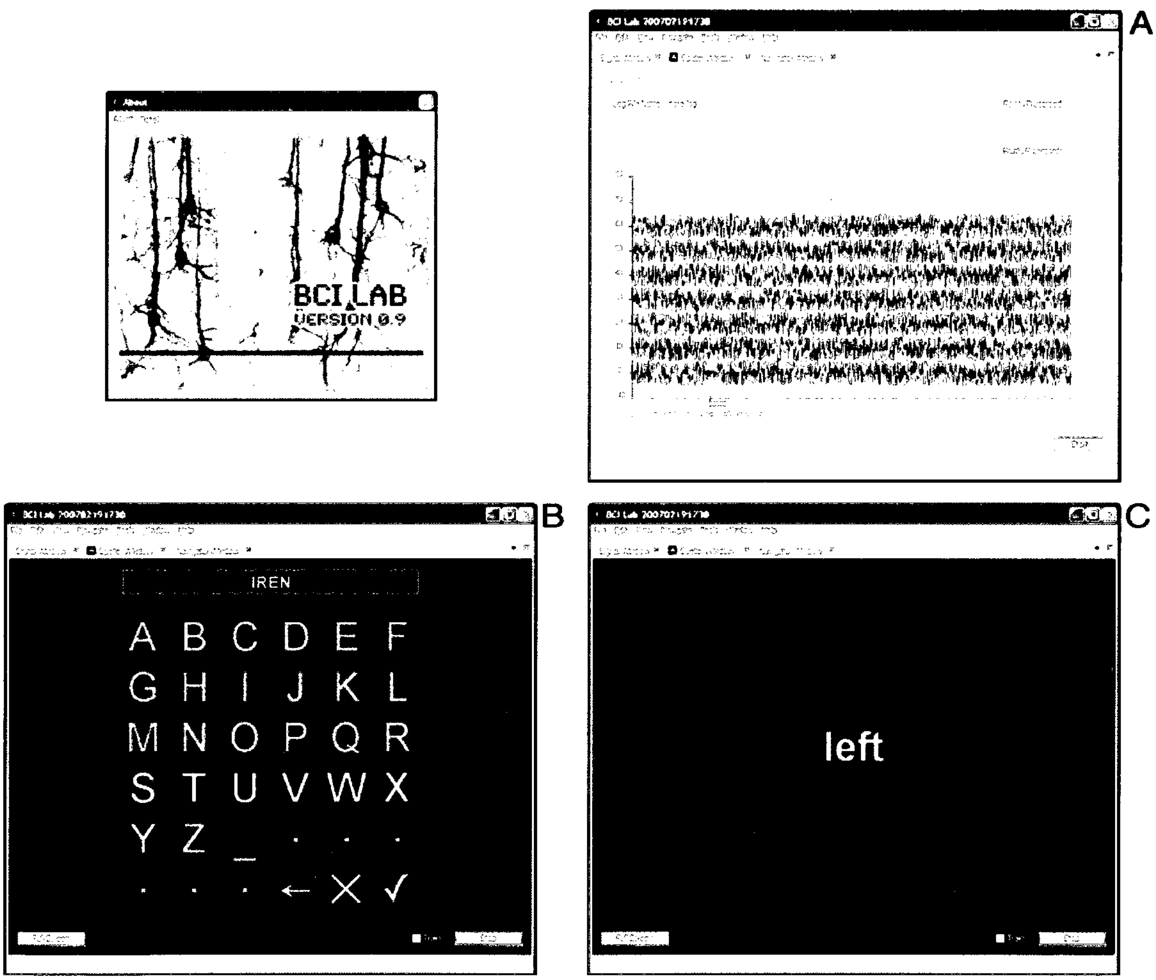


*Figure 4.9: BCILab user interface for: A) View of raw EEG data and P300 event trigger (top channel); P300 Speller paradigm; and Navigator paradigm.*

The configuration of *BCILab Signal* components and signals that they are emitting/receiving in *Navigator* paradigm is presented in Figure 4.10.

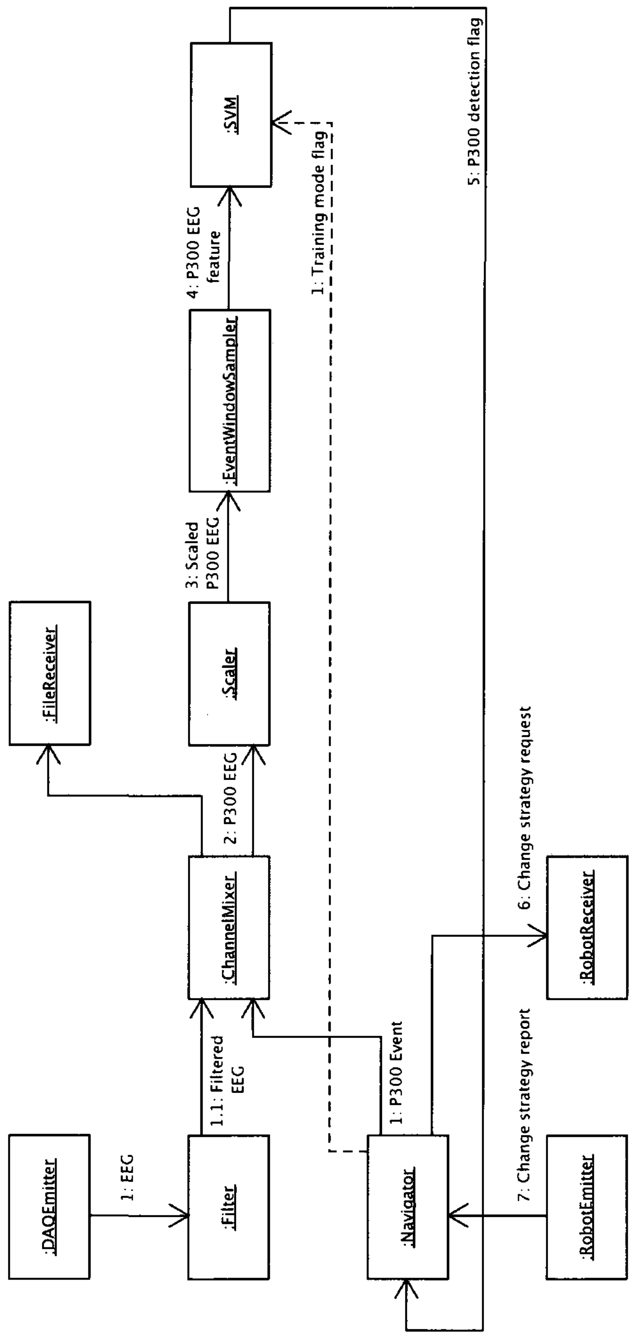


*Figure 4.10: Signal components configuration for Navigator paradigm.*

Description and role of each component is listed in Table 4.4.

*Table 4.4: Navigator components and their role.*

| Component | Role |
|---|---|
| *DAQEmitter* | Accepts raw EEG data from A/D converter |
| *Filter* | Bandpass filter |
| *ChannelMixer* | Joins raw EEG with P300 event |
| *FileReceiver* | Saves data to a file for offline analysis |
| *Scaler* | Scales EEG data (default between -1 and 1) |
| *EventWindowSampler* | Extracts desired time window for P300 event |
| *SVM* | Detects presence of P300 in EEG |
| *Navigator* | The main controller of the system (paradigm) |
| *RobotReceiver* | Sends strategy change command to RSA |
| *RobotEmitter* | Reports strategy change from RSA |

*Signal* components can be configured directly in code or in application configuration file. Configuration file is specified in XML format [14]. The example of application configuration file for *Navigator* paradigm corresponding to Figure 4.10 is listed in Figure 4.11.

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<!--
    Navigator Paradigm Experiment
    Created by Vladimir Hinic on Nov. 11, 2008
-->
<BCILab paradigm="Navigator">

    <component id="daq" type="DAQEmitter" connectedTo="filter">
        <param samplingRate="200" />
    </component>

    <component id="filter" type="Filter" connectedTo="mixer">
        <param lowerBound="0.1" />
        <param upperBound="40.0" />
    </component>

    <component id="mixer" type="ChannelMixer" connectedTo="scaler,file" />

    <component id="scaler" type="Scaler" connectedTo="eventSampler" />

    <component id="eventSampler" type="EventWindowSampler" connectedTo="svm">
        <param startDeltaTime="1100" />
        <param endDeltaTime="2800" />
    </component>

    <component id="svm" type="SVM" connectedTo="navigator">
        <param trainingFile="experiment1" />
        <param kernel="rbf" />
        <param alpha="0.01" />
        <param C="0.1" />
    </component>

    <component id="navigator" type="Navigator" connectedTo="robotReceiver" />

    <component id="robotReciver" type="RobotReceiver">
        <param comAddress="COM5" />
    </component>

    <component id="robotEmitter" type="Navigator" connectedTo="navigator">
        <param comAddress="COM6" />
    </component>

</BCILab>
```

*Figure 4.11: XML configuration file for Navigator application paradigm.*

## 4.4 Experimental Setup

The experimental setup used in the research is shown in Figure 4.12. A human operator equipped with visual and audio HCI, and BCI, controls the movements of an intelligent semiautonomous RSA at the strategic level. The RSA navigates through a maze using an onboard tactical-level obstacle-avoidance algorithm and implements strategic level plans for navigation similar to work presented in [30].

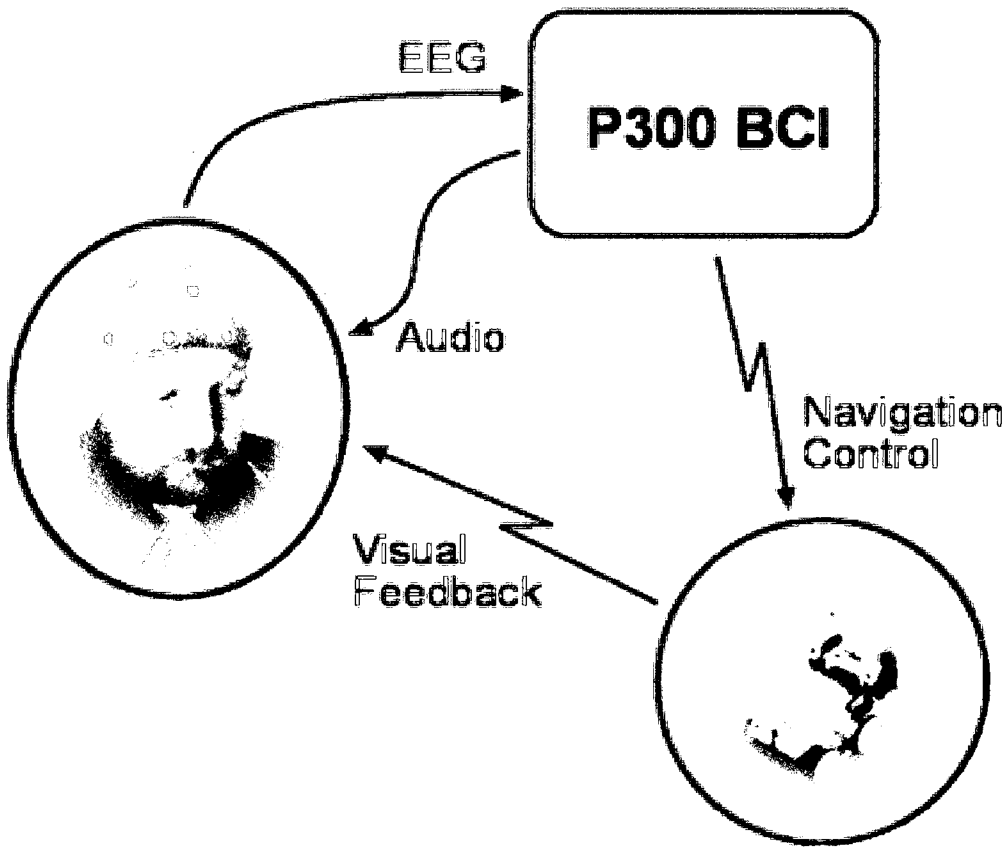


*Figure 4.12: Experimental setup for the P300 BCI control of a RSA.*

The human operator, who can see the RSA and the maze, wears an EEG cap as input to the BCI, and headphones as to receive auditory stimuli for P300 offering possible options for the RSA's change in strategy. The desired strategy as determined by BCI is then executed by the RSA.

## BCI Setup

A healthy 40 year old male subject wore a commercial EEG cap with integrated electrodes (white spots in Figure 4.12) mounted according to the international 10-20 system. EEG signals were recorded from 8 standard fronto-central-parietal locations (Figure 4.13): F3, F4, C3, Cz, C4, P3, Pz, and P4. The recordings were referenced to the left earlobe and grounded at the right earlobe. The electrode impedance did not exceed 5 kΩ. Amplified signal was bandpass filtered with cut-off frequencies of 0.1 Hz and 40 Hz and sampled at 200 Hz.

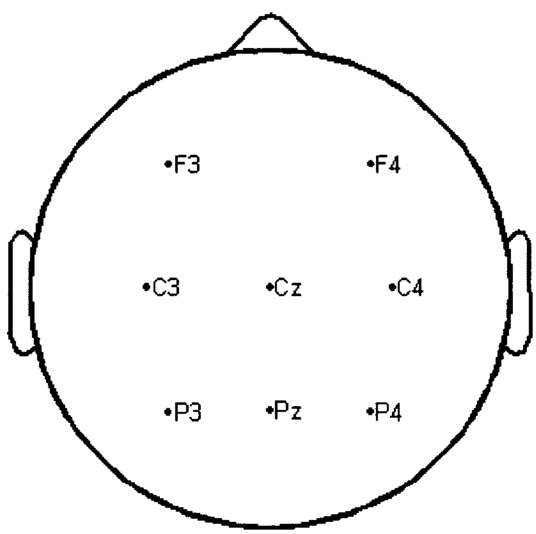


*Figure 4.13: Location of 8 electrodes used in experiment.*

The auditory stimuli were represented as spoken words. The words were: "left", "right", "ahead", "go", "stop", and "now". Only words "left" and "right" were used as target stimuli for the actual control of the RSA. Other words were simply used to decrease the probability of target stimuli appearing and therefore increasing P300 amplitude, which is inversely proportional to the probability of stimuli occurrence [6, 19].

The spoken words were of a medium quality, 16kHz diphone male voice generated using a speech synthesis system [42]. The average durations of the spoken words "left" and "right" were 2021 ms and 1787 ms, respectively. The average duration of all words was 1900 ms.

The auditory stimuli were presented to the operator in a number of repeating trial runs. Each trial run consists of 6 stimuli (6 spoken words) with 100 ms pause between them (Figure 4.14). The average duration of a trial was 12 s.

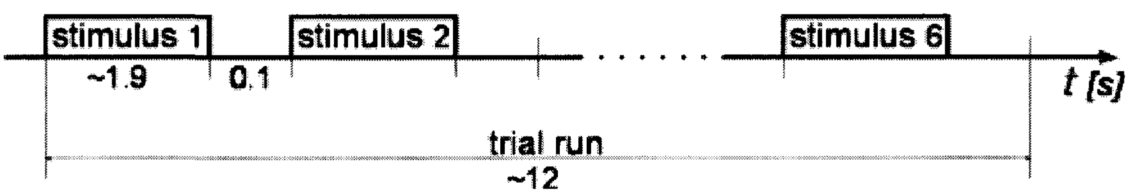


*Figure 4.14: Structure of a stimuli trial run.*

The objective for the operator was to attend target stimulus by making a mental note or silently counting target presentations and ignoring all non-target stimuli. The BCI system would then detect presence of P300 after target stimulus and the absence of P300 after non-target stimuli, and based on that, determine the desired command for the RSA.

The RSA would then send a verbal feedback informing the operator about the strategic plan that will be executed as a result of desired command.

## RSA Setup and Control

In our experimental setup of the RSA consists of a wheeled robot equipped with three frontally mounted infrared sensors. The sensors were used for obstacle avoidance and navigation in our experiments. The Bluetooth was used for communication with the BCI system. The robot implements two simple strategic level plans. The default plan is "follow the left wall" and the alternate is "follow the right wall". The robot changes execution of the plan based on input from the BCI.

The simple state machine representing the robot's strategic level plan execution is shown in Figure 4.15.

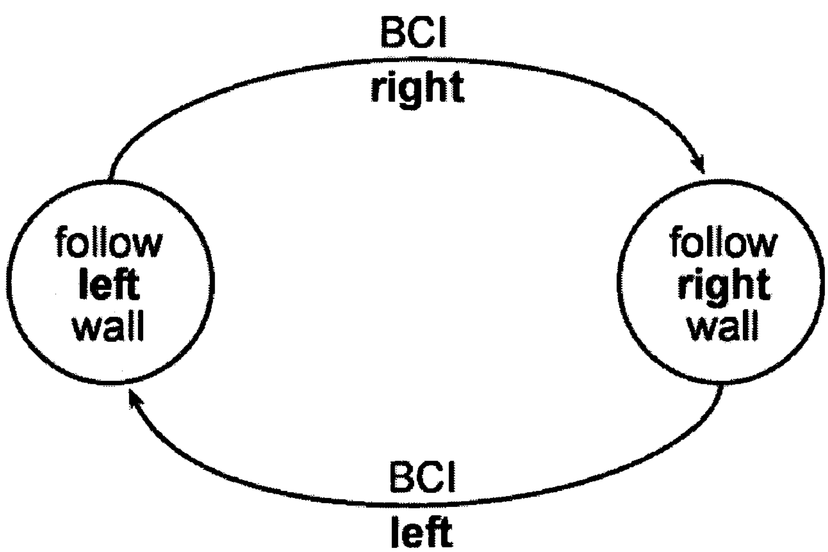


*Figure 4.15: A simple strategic plan execution state machine.*

## Feature Extraction and Classification

Since P300 is detectable in the amplitude domain, feature extraction is relatively simple and consists of selection of the data window and normalization of the samples.

The sample data window was selected experimentally in relationship to the stimulus onset time and appearance of P300 in the EEG signal. The selected data window of 1.7 s was then scaled to the range between -1 and 1 as suggested by [16]. Since this was performed online, scaling was done for each data window rather than for the whole recorded signal. This has both positive and negative side effects. The negative side effect is that normalization is not as effective as it would be if the whole signal would be normalized first. The positive side effect is that by normalizing each data window separately the whole signal is effectively shifted toward the center removing any low frequency or DC potentials.

Therefore, the size of the final input sample for classification is 340 points per EEG channel corresponding to 1.7 seconds for the data window, times recorded 8 channels, which is 2720.

SVM was used as a classifier with discriminant function interpreted as a score or confidence value as suggested by [22]. This score is then added for each repetition and the maximum one is considered to be desired word used for change in navigation of the strategy by the RSA. Based on formulas (3.1) and (3.2), this can be presented as

$$w_{sel} = arg\,\max_{w}(d_w^{word}),\ w = 1, \ldots, 6 \tag{4.1}$$

Where $w$ represents the word index and $w_{sel}$ is the index of the selected word.

The kernel function for SVM was RBF as recommended by [16]. It can be shown that RBF kernel function can represent both the linear and the polynomial as special cases. Since different lengths of spoken words resulted in shifting of P300 peek from sample to sample, certain non-linearity was introduced into classification problem and RBF was the better choice than the linear classifier.

# 5 Results

Two sets of experiments were performed: (i) the offline data collection and analysis in order to determine the best classification parameters and perform SVM training, and (ii) online execution based on parameters and training performed during the offline analysis.

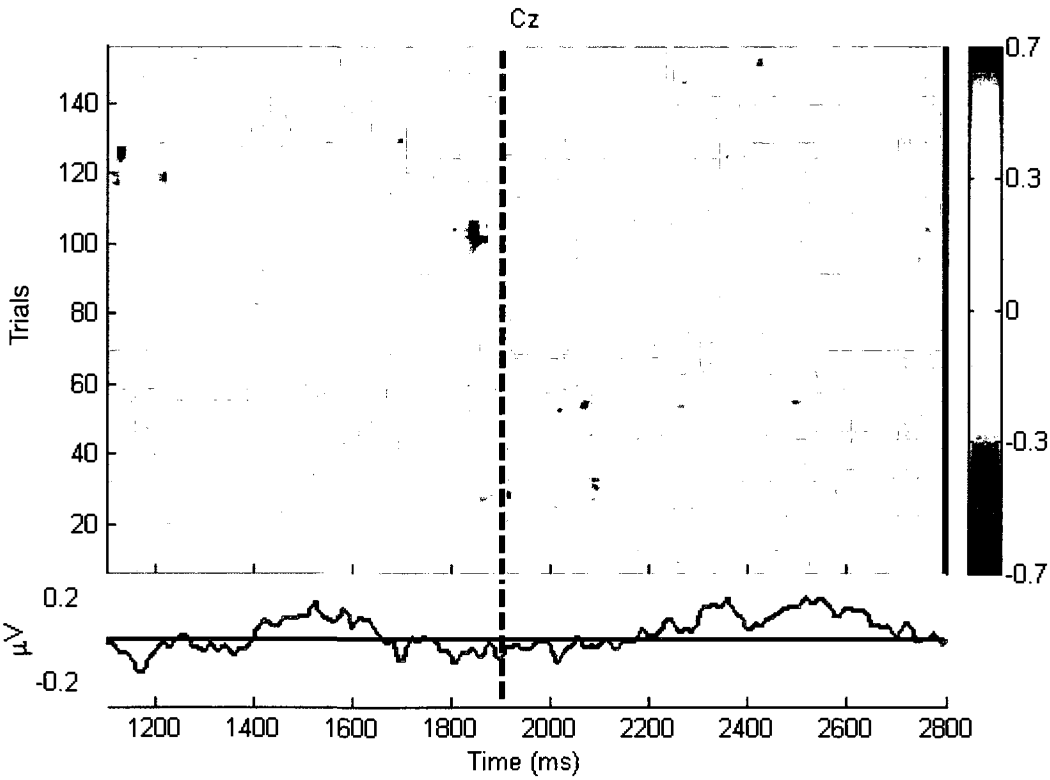


*Figure 5.1: Target stimuli - EEG activity over the Cz area. The dashed line represents stimuli offset. The timeline starts from stimuli onset.*

For offline data collection, 10 repetitions of trail runs over 16 sessions were used (8 for target left and 8 for target right). Based on [13], P300 can be expected somewhere between 300-900ms after stimulus offset. After analyzing data a positive peek that looked like P300 was located in the expected time interval for the target stimuli (Figure 5.1) and absence of P300 for the non-target stimuli (Figure 5.2).

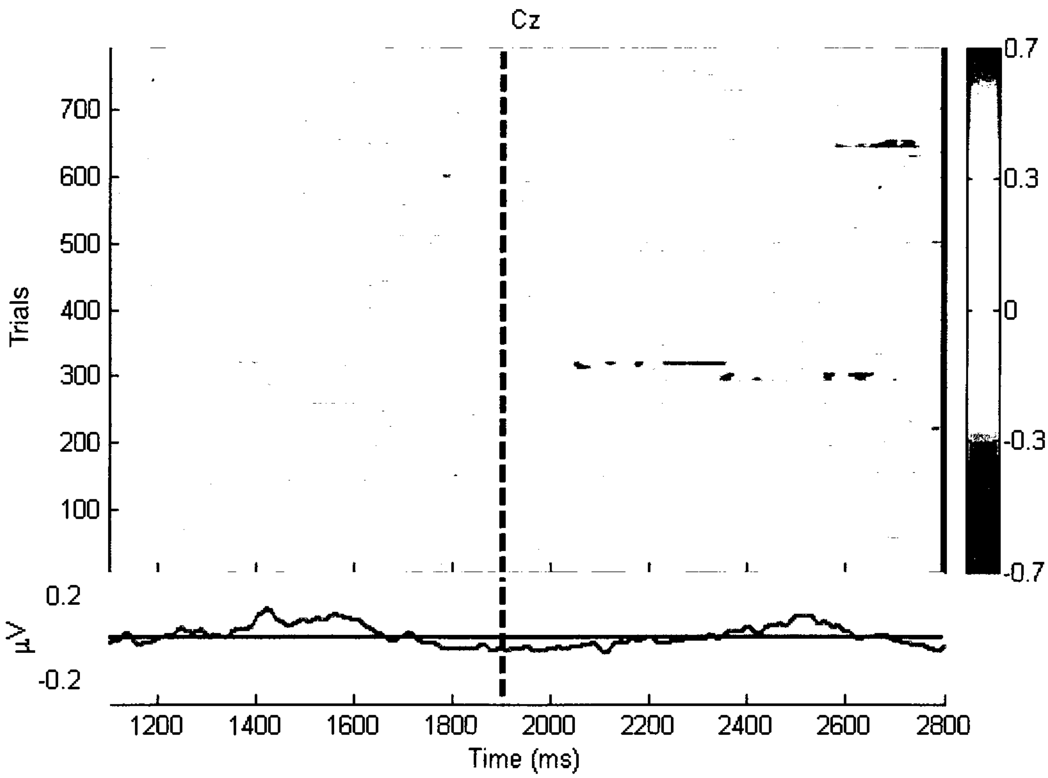


*Figure 5.2: Non-target stimuli - EEG activity over the Cz area. The dashed line represents stimuli offset. The timeline starts from stimuli onset.*

This possible existence of P300 and its average latency in target stimuli is also visible for all recorded channels in Figure 5.3.

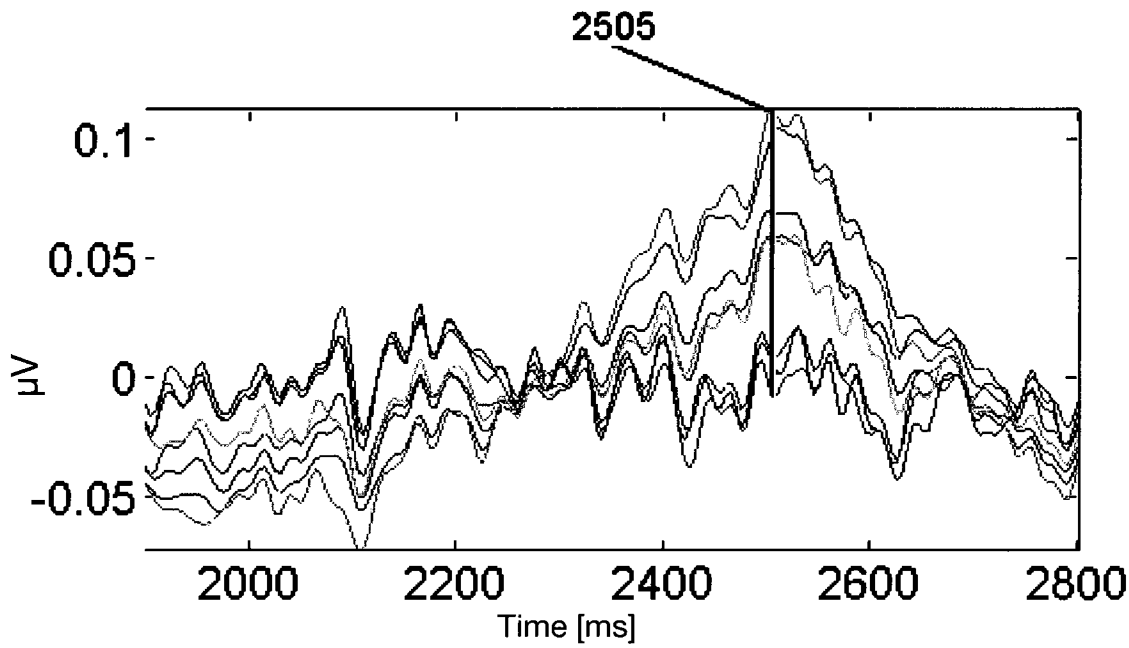


*Figure 5.3: Possible P300 and its average latency for the Cz recording. The timeline starts from stimuli onset.*

Additionally, experiments determined that the data window from 1100 ms to 2800 ms after stimulus onset provided the optimal classification results (Figure 5.4). The significance of the sample before stimulus offset is not clear in this research, but could be due either to the early recognition of spoken words (before the whole word is spoken) or some other non P300 physiological effect relevant for classification of the targeted stimuli.

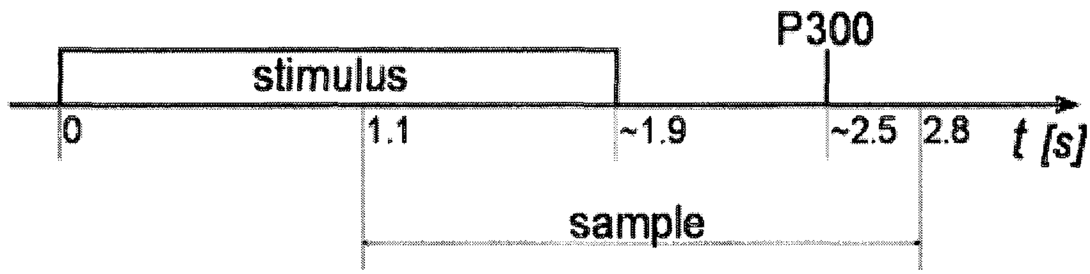


*Figure 5.4: A sample data window in relationship to the spoken word onset/offset and expected P300 latency.*

In order to determine classification accuracy the cross validation was performed on the 16 sessions collected, with 60% of trials randomly selected for training and the remaining 40% for testing. The same numbers of trials for left and right were used for both the training and the testing set. Chronological order for all repetitions for a single trial during our tests was preserved. In order to balance positive and negative samples in a training set, one available sample corresponding to the target stimulus and only one out of 5 samples corresponding to the non-target stimulus were used. The remaining 4 non-target stimuli were dismissed.

The classification tests were performed using different numbers of trial run repetitions. Results were collected for 1, 2, 3, 4 and all 10 repetitions (Figure 5.5).

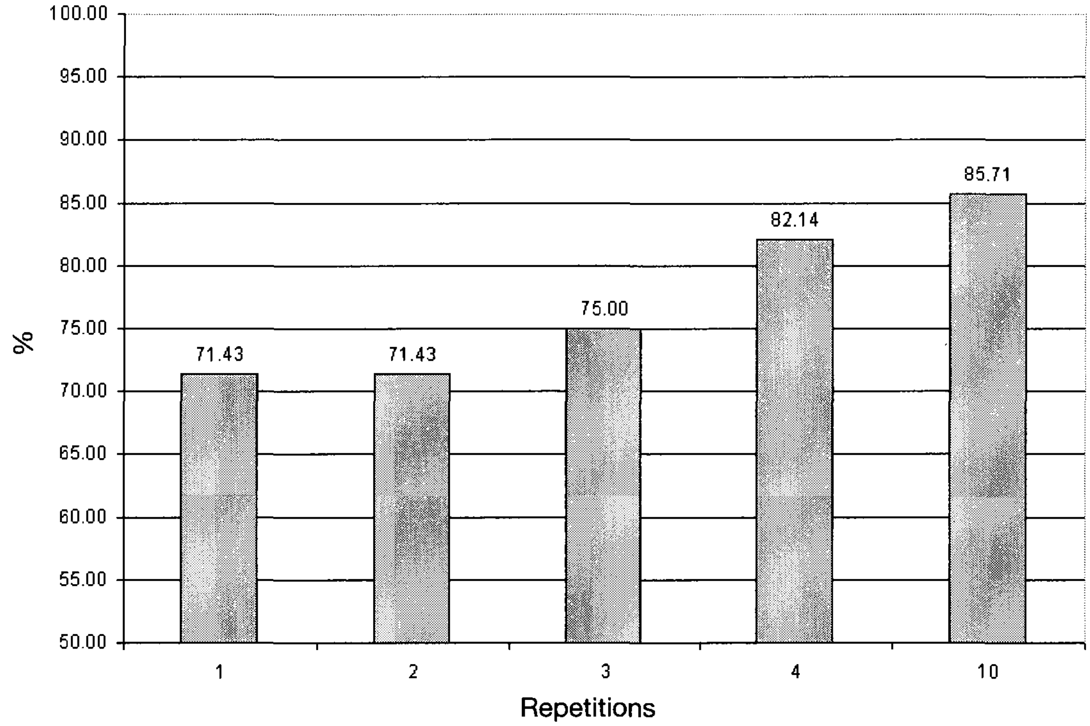


*Figure 5.5: Classification accuracy as a result of trial run repetitions.*

As expected, the best classification results of 85.71% were obtained with all 10 of the available repetitions. The classification results with a single trial run were 71.43%. Pure chance would give a result of approximately 16.67%. The classification results are better than in [40] where Step-Wise Discriminant Analysis (SWDA) was used as a method of classification.

Trained SVM is used for the online experiment where a fixed number of trial run repetitions were used. The tests were performed using 1 and 4 repetitions. The time for SVM classification is 1s per sample and it is concurrent with sample acquisition. Therefore, the total time for classification of the target command is 13s in the case of a single repetition and 49s in the case of 4 repetitions.

Preliminary subjective tests were performed in the indoor environment of the online system with the speed of the robot constant and at 1/3 of its length per second as described in [29]. The results were positive and show that the operator has control of the RSA in planning its navigation strategies.

The single repetition tests resulted in the subject "feeling" more in control than in 4 repetitions tests, even with lower classification accuracy. This could be explained by the fact that in the test setup, the operator was awaiting award by the RSA's change in strategy rather than focusing on the navigation task. This resulted in a desire for more frequent change in the RSA's strategy where the operator readily accepted negative classification or the wrong commands.

# 6 Conclusions

This research shows that it is possible to control RSA's on a strategic level using P300 based spoken word, auditory based BCI. The issue of the slow feedback loop when using BCI is avoided by directing the RSA only on a high, strategic level. Additionally, the flexible, reliable, multi platform, and *Java* based BCI framework was implemented for this and possible further research.

This research contributed the following: proof of the suitability of spoken words as a triggering mechanism for an auditory evoked P300 base BCI system; BCILab – a multi-platform, open software framework for BCI research; and implementation of spoken word auditory evoked P300 BCI system for symbiotic relationship between human operator and semiautonomous RSA.

The usage of spoken words in order to generate P300 for BCI is possible, but it introduces some challenges. The main issue is the duration of a stimulus and its latency jitter. This jitter reduces the effect of P300 and decreases classification accuracy. For example, by just introducing one more word in addition to left and right the classification accuracy drops significantly. However, when using auditory stimuli, the suitability of spoken word for the human operator, in comparison to other auditory stimulus, (different frequency or patterns) outweighs this issue.

There are several ways to increase the speed of this BCI system. One is to increase the accuracy of a single trial run classification and avoid repetitions. Other ways would be to reduce the number of different stimuli types (words). In [40] only 4 words were used and satisfactory classification results were achieved. Finally, rather than controlling the navigation strategy of the RSA, even higher level control could be implemented. For example, direct-to-target can be applied and tactical level navigation could be left to the RSA. We are currently working on increasing the accuracy of a single trial run and the number of recognized words by applying SVM assemblies as suggested in [38].

Possible enhancement could be made in areas of better feature extraction and potential usage of an error detection mechanism based on error-related negativity (ERN). The ERN is a negative deflection in the ongoing EEG that occurs approximately 100ms after an error has been committed [49, 50].

# References


[1] B. Allison, "P3 or not P3: Toward a Better P300 BCI," *PhD Thesis*, University of California, San Diego, 2003.

[2] "Lego Mindstorms NXT," *The LEGO Group*, available at http://mindstorms.lego.com

[3] "NetBeans Platform," *Sun Microsystems*, available at http://platform.netbeans.org

[4] "Educational Robot Kit: Bioloid," *ROBOTIS, Inc.*, available at http://www.robotis.com/zbxe/bioloid_en

[5] J. D. Bayliss and D. H. Ballard, "A virtual reality testbed for brain–computer interface research," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 188-190, June 2000.

[6] J.D. Bayliss, "A flexible brain-computer interface, " *PhD Thesis,* University of Rochester, 2001.

[7] J. D. Bayliss, "Use of the evoked potential P3 component for control in a virtual apartment," *Neural Systems and Rehab. Eng.*, 2003.

[8] J. Bayliss, S. Inverso, A. Tentler, "Changing the P300 brain computer interface," *Cyber Psychology & Behavior*, 2004.

[9] C. Bell, P. Shenoy, R. Chalodhorn, R. Rao, "Control of a humanoid robot by a noninvasive brain--computer interface in humans," *J. Neural Eng.*, 2008.

[10] N. Birbaumer, et al., "The thought translation device (TTD) for completely paralyzed patients," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 190-193, June 2000.

[11] B. Blankertz, K.-R. Muller, G. Curio, T. M. Vaughan, G. Schalk, K. R. Wolpaw, A. Schlogl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schroder, N. Birbaumer, "The BCI competition 2003: progress and perspectives in detection and discrimination of

EEG single trials," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, 2004, pp. 1044-1051.

[12] B Blankertz and K Muller and D Krusienski and G Schalk, "The BCI competition III: validating alternative approaches to actual BCI problems," *Neural Systems and Rehab. Eng.*, 2006.

[13] A. Bozinovski, "Human-computer interfaces and anticipatory brain potentials," *The 6th International Conference for Informatics and Information Technology*, 2008.

[14] T. Bray, J. Paoli, C.M. Sperberg-McQueen, E. Maler, F. Yergeau, "Extensible Markup Language (XML) 1.0 (Fifth Edition)," *W3C Recommendation*, 2008, available at http:// www.w3.org/TR/2008/REC-xml-20081126.

[15] C. J.C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, 1998, pp. 121–167.

[16] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," 2007, Software available at http://www.csie.ntu.edu.tw/_cjlin/libsvm.

[17] E. Donchin, K. Spencer, R. Wijesinghe, "The mental prosthesis: Assessing the speed of a P300-based brain–computer interface," *IEEE Trans. Rehab. Eng.*, vol. 8, June 2000, pp. 174-179.

[18] R. Fan, P. Chen, C. Lin, "Working Set Selection Using Second Order Information for Training Support Vector Machines," *The Journal of Machine Learning Research*, 2005.

[19] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroenceph. Clin. Neurophysiol.*, pp. 510-523, 1988.

[20] V. Hinic, E.M. Petriu, T.E. Whalen, "Human-computer symbiotic cooperation in trobot-sensor networks," *IEEE Instrum. Meas. Mag.,* Conference Proceedings, 2008.

[21] T. Jarvi, "RXTX: Native library for Java serial communication ," available at http://www.rxtx.org.

[22] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "BCI competition 2003 - data set IIb: Support vector machines for the p300 speller paradigm," *IEEE Trans. Biomed. Eng.*, 2004.

[23] V. Kecman, "Learning and Soft Computing: Support Vector Machines, Neural Networks and Fuzzy Logic models," *The MIT Press*, Cambridge, MA, 2001.

[24] M. Lebedev and M. Nicolelis, "Brain--machine interfaces: past, present and future," *Trends in Neurosciences*, 2006, pp. 536-546.

[25] J Lehtonen, "EEG-based Brain Computer Interfaces," *MSc Thesis*, Helsinki University of Technology, May 2003.

[26] J. Malmivuo and R. Plonsey, "Bioelectromagnetism - principles and applications of bioelectric and biomagnetic fields," *Oxford University Press*, 1995.

[27] D. McFarland, L. McCane, J. Wolpaw, "EEG-based communication and control: short-term role of feedback," *IEEE Trans. Rehabil. Eng.*, 1998.

[28] P. Meinicke, M. Kaper, F. Hoppe, M. Heumann, and H. Ritter,"Improving transfer rates in brain computer interfacing: a case study," *in Advances in Neural Information Processing Systems 15*, S. Becker, S. Thrun, and K. Obermayer, Eds., Cambridge, MA, 2003, MIT Press.

[29] J. R. Millán, F. Renkens, J. Mourino, W. Gerstner, "Non-invasive brain-actuated control of a mobile robot," *IJCAI 2003 Proceedings*, 2003, pp. 1121-1126.

[30] J. R. Millán, F. Renkens, J. Mourino, W. Gerstner, "Noninvasive brain-actuated control of a mobile robot by human EEG," *IEEE Trans. Biomed. Eng.*, 2004.

[31] E. Niedermeyer and F. Silva, "Electroencephalography: Basic Principles, Clinical Applications, and Related Fields,", 2004.

[32] E.M. Petriu, T.E. Whalen, "Computer-Controlled Human Operators," *IEEE Instrum. Meas. Mag.*, Vol. 5, No. 1, pp. 35-38, 2002.

[33] E.M. Petriu, T.E. Whalen, R. Abielmona, A. Stewart, "Robotic sensor agents: a new generation of intelligent agents for complex environment monitoring," *IEEE Instrum. Meas. Mag.*, Vol. 7, No. 3, pp. 46 -51, 2004.

[34] E.M. Petriu, T.E. Whalen, I. Rudas, D. Petriu, M. Cordea, "Human-instrument symbiotic partnership for multimodal environment perception," *IEEE Instrum. Meas. Mag.,* Conference Proceedings, 2008, pp. 1263-1268.

[35] G. Pfurtscheller, C. Neuper, G. Muller, B. Obermaier, "Graz-BCI: state of the art and clinical applications," *IEEE Trans. Neural Systems and Rehab. Eng.*, 2003.

[36] R Picard, "Human-computer coupling," *Proceedings of the IEEE*, 1998, pp.

[37] A. Rakotomamonjy, V. Guigue, G. Mallet, and V. Alvarado, "Ensemble of SVMs for improving brain-computer interface p300 speller performances," *15th International Conference on Artificial Neural Networks*, 2005.

[38] A. Rakotomamonjy and V. Guigue, "BCI competition III: dataset II- ensemble of SVMs for BCI P300 speller," *IEEE Trans. Biomed. Eng.*, 2008, pp. 1147-1154

[39] B. Rebsamen, E. Burdet, C. Guan, C. Teo, Q. Zeng, "Controlling a wheelchair using a BCI with low information transfer rate," *Rehab. Robotics*, 2007.

[40] E.W. Sellers, "A P300-Based Brain-Computer Interface: Testing an Alternative Method of Communication," *Ph.D. dissertation*, University of South Florida, 2004.

[41] J Vidal, "Toward Direct Brain-Computer Communication," *Annual Reviews in*

*Biophysics and Bioengineering*, 1973.

[42] W. Walker and P. Lamere and P Kwok, "FreeTTS -- A Java Based Text to Speech Synthesis System," , Speech Integration Group, Sun Microsystems Laboratories, available at http://freetts.sourceforge.net/

[43] J.G Webster, "Medical Instrumentation: Application and Design," *John Wiley & Sons*, New York, 1998.

[44] A. Westermann, "JChart2D: Java library for precise runtime visualization of numerical tuples in an x-y plot," available at http://jchart2d.sourceforge.net

[45] G. Wielenga, J. Tulach, T. Boudreau, "Rich client programming: plugging into the NetBeans platform," *Prentice Hall*, 2007.

[46] J.Wolpaw, N.Birbaumer, W.Heetderks, D. McFarland, "Brain-computer interface technology: a review of the first international meeting," Rehabilitation Engineering, 2000.

[47] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," Clin. Neurophysiol., vol. 113, pp. 767-791, 2002.

[48] J. Wolpaw, et al., "BCI Meeting 2005-Workshop on Signals and Recording Methods," *IEEE Trans. Neural Systems and Rehab. Eng.*, 2006.

[49] G. Schalk, J.R, Wolpaw, D.J. McFarlrald, G. Pfurtscheller, "EEG-based communication: presence of an error potential," *Clin. Neurophysiol.* no. 111, 2000, pp. 2138-2144

[50] L. Parra, C. Spence, A. Gerson, P. Sajda, "Response error correction-a demonstration of improved human-machine performance using real-time EEG," *Neural Systems and Rehab. Eng.*, 2003.