

Research on Human-Machine Interfaces
of Vigilance Estimation and Robot Control
based on Biomedical Signals

MA Jiaxin

Abstract

Biomedical signals, for example, blood pressure, heart rate, etc., are generated from our bodies. These signals show information about our health states, physical or mental activities. This thesis focuses on three kinds of biomedical signals: electrooculogram (EOG), electroencephalogram (EEG), and electromyogram (EMG), which are the electric signals generated from the eyes, the brain, and the muscles, respectively. Based on these biomedical signals, in this thesis it implements human-machine interfaces (HMI) of different functions, establishing multiple kinds of connections between human beings and machines.

The contents of this thesis are mainly about three proposed human-machine interfaces. These HMIs extract useful information from biomedical signals (EOG, EEG, and EMG), by which they can monitor the body states, or control the external devices.

The first proposed HMI is an EOG interface, applied for vigilance estimation. Vigilance is an index to measure the degree of concentration, it can be affected by monotonous task, fatigue, or sleepiness. By estimating vigilance from EOG signals, this HMI can be used on such cases like the prevention of fatigue driving. In order to explore the relationship between EOG features and vigilance, first, eye movements including slow eye movement (SEM), saccade, and blink were detected from EOG signals. Then a variety of EOG features were defined and extracted from those eye movements. These features were processed by a linear dynamic system (LDS) approach for noise reduction. Finally the denoised features were compared with a vigilance reference value. Experiments including 22 subjects and 36 data sets indicate that the combination of SEM, saccades, blink, and energy features of EOG has a high correlation coefficient with the vigilance reference, up to the average 0.75. In addition, in this study it also proposed an alternative placement of EOG electrodes for wearing the measuring device more conveniently. In the new placement, the EOG electrodes around eyes were moved to the forehead. Independent component analysis (ICA) was used to retrieve the same signals as if the electrodes were not moved.

The second proposed HMI is an EOG/EEG hybrid interface, applied for controlling robots. This hybrid interface works in two modes: an EOG mode where the robot receives commands generated by eye movements, and an EEG mode where the robot receives commands generated by event-related potentials (ERP), a kind of brain activity. While both eye movements and ERPs have been separately used for implementing assistive interfaces which help patients with motor disabilities in performing daily tasks, the proposed hybrid interface integrates them together. In this way, both the eye movements and ERPs complement each other. Therefore, it can provide a better efficiency and a wider scope of application. In this study, a threshold algorithm was designed for recognizing four kinds of eye movements which are blink, wink, gaze, and frown. In addition, an oddball paradigm

with stimuli of inverted faces is used to evoke multiple ERP components including P300, N170, and VPP. To verify the effectiveness of the proposed HMI, two different online experiments were carried out. One is to control a multi-functional humanoid robot, and the other is to control four mobile robots. In both experiments, the subjects could complete tasks effectively by using the proposed interface whereas the best completion time was relatively short and very close to the time operated by hand.

The third proposed HMI is an EMG interface, applied for controlling a prosthesis. By using surface EMG signals measured from the forearm, this HMI allows the user to control a prosthesis to perform four kinds of movements: open, close, pronate, and supinate. The nonnegative matrix factorization (NMF) algorithm is used to map muscle activities into these four movements through calculating a muscle synergy matrix. An EMG feature selection process along with a control scheme has been added, which smooths the output thereby stabilizing the movements. Using this method, the proportional control of multiple degrees of freedom (DOF) can be realized. This means the output does not only contain simple states of open, close, pronate, and supinate, but also contains all the transition states between open and close, pronate and supinate, as well as some combination states like “open and pronate”, “close and supinate”. Ten healthy subjects performed an online experiment comprised of two tests: (i) proportional control on single DOF, and (ii) simultaneous control on multiple DOFs. The results indicated that fluid hand/wrist movements could be successfully estimated from EMG.

Acknowledgments

This thesis was supervised by Professor Fumitoshi Matsuno. I would like to express my respect for him. He has always given me kind guidance and advice when I studied at the Kyoto University. Together with Professor Matsuno, Professor Tetsuo Sawaragi and Professor Naohide Tomita were the member of my thesis committee. I would like to thank all of them. I would also like to thank Professor Andrzej Cichocki, who supervised me during my visit to RIKEN, and Professor Nitish Thakor, who supervised me when I was at the Johns Hopkins University.

Moreover, I thank my colleagues of the Matsuno laboratory in the Kyoto University, the Cichocki laboratory in RIKEN, and the Thakor laboratory in the Johns Hopkins University. Particularly, I thank my contemporaries in the Matsuno laboratory, Ryo Ariizumi, Tae Hyon Kim, and Yuichi Ambe, who accompanied me for years and gave me helps. Also, I specially thank Yu Zhang of the Cichocki laboratory, and Michael Powell of the Thakor laboratory, both of whom gave me great help when I visited the laboratories they belong.

I would like to express my thanks to Chinese Scholarship Council, which supported me with full scholarship during my first three years in Kyoto University. Also, I would like to thank Japan Society for the Promotion of Science, which funded me during my six-month visit to the Johns Hopkins University.

Finally, I would like to express my heartfelt thanks to my parents, Bangwei Ma and Xingyan He, who have supported me not only through my college years, but also throughout my entire life.

Contents

1	Introduction	1
1.1	Human-machine interface	1
1.2	Biomedical signals	2
1.2.1	EEG	2
1.2.2	EOG	2
1.2.3	EMG	2
1.3	Biomedical signal based HMI	3
1.4	Paper organization	4
2	EOG based Vigilance Estimation	5
2.1	Introduction	5
2.2	Related Work	7
2.2.1	EOG Applications	7
2.2.2	Vigilance Estimation	8
2.3	Materials and Methods	9
2.3.1	Eye movements and EOG Features	9
2.3.2	Preprocessing	10
2.3.3	Eye Movement Detection	11
2.3.4	Feature Extraction	14
2.3.5	Smoothing and De-noising	16
2.3.6	Feature Combination	17
2.3.7	Placement methods	18
2.3.8	Experiment	19
2.3.9	Evaluation	21
2.4	Results	21
2.5	Discussion	23
2.5.1	Relation between EOG features and vigilance	23
2.5.2	Online Attempt	26
2.5.3	Electrode replacement	27

3	EOG and EEG based Robot Control	28
3.1	Introduction	28
3.2	Model of hybrid interface	31
3.3	EOG analysis	33
3.3.1	EOG acquisition	33
3.3.2	Eye movement detection algorithm	34
3.3.3	EOG calibration and thresholds calculation	39
3.4	EEG analysis	40
3.4.1	EEG acquisition	40
3.4.2	ERP paradigm	40
3.4.3	ERP classification	43
3.5	Experiments	43
3.5.1	Experiment introduction	43
3.5.2	Humanoid robot control	46
3.5.3	Mobile robots control	47
3.6	Results	48
3.6.1	Offline evaluation	48
3.6.2	NAO experiment	48
3.6.3	Kobuki experiment	51
3.7	Discussions	53
3.7.1	Performance	53
3.7.2	Practicality	55
3.7.3	Extensions	56
4	EMG based Prosthesis Control	57
4.1	Introduction	57
4.2	Methods	59
4.2.1	Synergy model and NMF	59
4.2.2	Feature Selection	62
4.2.3	Control Scheme	62
4.2.4	Experiment	64
4.3	Results	71
4.4	Discussion	75
4.4.1	Offline experiment	75
4.4.2	Online experiment	75
4.4.3	Practicality	76
5	Conclusion	78
5.1	Conclusion	78
5.2	Future work	79

List of Figures

2.1	The flow chart of the proposed EOG based vigilance estimation interface.	10
2.2	The waveform of a blink in vertical EOG and its differentiated shape. The x-axis is time and y-axis is EOG (μV).	13
2.3	The structure of state space of LDS. The hollow circles are hidden states, and the solid circles are observed states. Arrows mean inferable relations.	17
2.4	Comparison of EOG signals before and after using LDS approach. Here the blue line denotes the original EOG signal and the red one denotes the EOG signal smoothed by using the LDS approach.	17
2.5	Comparison of two placements of EOG electrodes. REF and GND in (a) correspond to the red circles in (b); A, B, C, and D in (a) correspond to the blue circles in (b).	19
2.6	The scene of the image identification experiment. The EOG data are collected by four electrodes on the EEG cap. The subject's task is to push the button which shares the same color with the image stimulus.	20
2.7	Relation between the local error rate (green curve) and fatigue level presented by SEM proportion (red curve). Here, the correlation coefficient is 0.92.	22
2.8	The correlation between EOG features and local error rate for all 36 sessions. The y-axis represents the value of correlation coefficient (γ). For the features which exhibit negative correlation with local error rate, their γ have been inverted for observation convenience. In the box plot, the five characteristic points from the top to the bottom down are: maximum, third quartile, median, first quartile, and minimum, which are indicated as P100, P75, P50, P25, and P0, respectively.	23
2.9	The correlation between EOG features and local error rate for 4 combination groups. Each group contains a number of EOG features which are reduced to one dimension by linear summation (Lin), principal component analysis (PCA), and manifold learning (ISO). The y-axis represents the value of correlation coefficient (γ).	24

2.10	The correlation between part of EOG features and local error rate, for all 36 sessions. The red thick line represents linear summation and other colored lines are individual features in Group 2. Central region corresponds to low γ while margin to high γ	24
2.11	The correlation coefficients for each movement. The 14 different kinds of movements are: 1) bending the head; 2) shaking the head; 3) clenching the teeth; 4) blinking; 5) rolling the eyes; 6) horizontal eye movement; 7) vertical eye movement; 8) swallowing; 9) squinting; 10) raising the eyebrows; 11) scowling; 12) smiling; 13) scratching a cheek; and 14) opening and closing the mouth.	25
3.1	The proposed model of the hybrid HMI	32
3.2	The placement of EOG electrodes. A: vertical EOG electrode; B, C: horizontal EOG electrode; REF: reference; GND: ground.	34
3.3	Typical waveforms of (a) blinks, (b) frown, (c) wink, and (d) gaze. For each, the upper plot is the original waveform whose y-axis is voltage (μV) and x-axis is time (s), and the lower plot is the differentiated result (first-order difference) of the upper one. Here each 3 second duration is considered as a single trial. Intervals between trials are omitted.	35
3.4	The waveform of a blink in vertical EOG and its differentiated shape [1]. The x-axis is time and y-axis is voltage (μV). The derivative is approximated by the first-order difference $\frac{x(t)-x(t-\Delta t)}{\Delta t}$ in our program.	36
3.5	Electrode positions for EEG recording. (Ground and reference electrodes are shown in Fig. 3.2.)	41
3.6	Grand average ERP waveforms derived from the target and non-target stimuli of the inverted face image. VPP, N170, and P300 can be clearly observed from channel Cz, P8, and Oz. The y-axes are voltage (μV) and x-axes are time (ms).	41
3.7	The timeline of a single run. In the training phase, a target cue was provided, and the number of trials K was 5. In the test phase, the number of trials K was 2 and feedback was provided. Each trial consisted of eight sub-trials in each of which one stimulus was randomly presented in one of the eight directions for 100 ms with an inter-stimulus interval of 100 ms.	42
3.8	Experimental robots.	43
3.9	Experimental layout. The solid arrows are routes of NAO, and the dashed arrows are routes of Kobukis.	46
3.10	The recall of six kinds of eye movements, detected by the proposed method and a pattern matching method. One kind of eye movement has 140 repetitions.	49

3.11	The ERP offline classification accuracy averaged on subjects using various trials. Vertical lines show maximal and minimal accuracies among all the subjects. The x-axis means how many trials were used.	49
4.1	From the training data, $W_i (i = 1, 2, \dots, K)$ is DOF-wise calculated by individual NMF algorithms, and thus $\mathbf{W}$ is obtained. For new testing data $f(\mathbf{Z})_{test}$, the estimated synergy activation is $\mathbf{H}_{est} = \mathbf{W}^{-1} f(\mathbf{Z})_{test}$	61
4.2	The block diagram of the control scheme.	64
4.3	The experimental setup (corresponding to #1 in Table 4.2).	66
4.4	The virtual prosthesis generated by the musculoskeletal modeling software (MSMS) [2]. In the offline experiment, the static pictures of these nine movements are displayed as cues. In the online experiment, coherent animation of the virtual prosthesis is used as the visual feedback.	66
4.5	Flowchart of the offline experiment. The sequence of movements is random.	67
4.6	An ideal demonstrations of online (a) test 1 and (b) test 2. Blue lines are the ideal output of control signals C_1 , C_2 , C_3 , and C_4 . Only the first halves of the two tests are shown.	68
4.7	One session of the offline experiment. (a) shows the average synergy activation levels for 9 different movements in the experiment and their standard variations. In (b), the blue lines are the time-varying synergy activations ($\widehat{H_k(t)}$ in Eq. 4.6) in trials (the first 2 seconds for preparation in each trial is omitted). The labels of x-axes represent different movements (R for rest, O for open, C for close, P for pronate, and S for supinate). The red dash lines are the target activation levels ($H_k(t)$ in Eq. 4.6).	70
4.8	The maximal (blue diamonds), minimal (red squares) and mean (green triangles) R^2 value of 14 offline experiment sessions for seven features.	71
4.9	An example of training session data from the online experiment. Red lines are activation level of four synergies, and blue lines are corresponding control signals C_1 , C_2 , C_3 , and C_4	72
4.10	Examples of online (a) test 1 and (b) test 2. Blue lines are the system output (control signals C_1 , C_2 , C_3 , and C_4) which the subject can see from the screen during the experiment.	73
4.11	The R^2 value calculated from all the ten subjects' online experimental data (test 1 and test 2), and also from the same data processed by low-pass filter but not the control scheme (test 1f and test 2f). The average R^2 values of test 1 and test 2 are 0.97 and 0.93, respectively.	74
4.12	One example of the experiment data of test 1 for comparison. Green lines are estimated synergy activations ($\mathbf{H}_{est}$). Red lines are $\mathbf{H}_{est}$ processed by the low-pass filter (with a output range $[0,1]$). Blue lines are $\mathbf{H}_{est}$ processed by the control scheme (control signals C_1 , C_2 , C_3 , and C_4).	74

List of Tables

2.1	List of EOG features analyzed in this work	15
3.1	Recent studies about hybrid brain or non-brain human-machine interfaces	30
3.2	A summary of different eye movements	36
3.3	Thresholds for eye movement detection	40
3.4	Commands for the two robot-control experiments	45
3.5	Results of the NAO experiment	50
3.6	Results of the Kobuki experiment	52
4.1	EMG Features	62
4.2	Experiment settings	65

Abbreviations

BCI	Brain-Computer Interface
BMI	Brain-Machine Interface
CNS	Central Nervous System
DWT	Discrete Wavelet Transform
DOF	Degree of Freedom
EEG	Electroencephalogram
EMG	Electromyogram
EOG	Electrooculogram
ERP	Event-Related Potential
FFT	Fast Fourier Transform
GUI	Graphical User Interface
HEO	Horizontal EOG
HMI	Human-Machine Interface
ICA	Independent Component Analysis
KDS	Karolinska Drowsiness Score
KSS	Karolinska Sleepiness Scale
LDS	Linear Dynamic System
LF/HF	low Frequency/High Frequency Ratio
NMF	Nonnegative Matrix Factorization
MAV	Mean Absolute Value
PCA	Principal Component Analysis
PERCLOS	Percentage of Eye Closure
PSD	Power Spectral Density
REM	Rapid Eye Movement
RMS	Root Mean Square
SEM	Slow Eye Movement
SNR	Signal to Noise Ratio
SSVEP	Steady State Visually Evoked Potential
SVM	Support Vector Machine
VEO	Vertical EOG
VPP	Vertex Positive Potential

Chapter 1

Introduction

1.1 Human-machine interface

Human-machine interface (HMI) is a general term to describe an intermediary system which allows interaction between a human being and a machine. Human machine interfaces vary widely, from control panels for nuclear power plants to the screen and input buttons on a cell phone. With an HMI, the human user can send commands to the machine to operate and control it, for example, control a robot to complete various tasks. Also, the use can receive information from the machine, which may include the machine's sensor information, the machine's analysis results, or the feedback of the user's commands. Therefore, the most common uses of HMIs include controlling and monitoring.

As the means of sending commands from the human user to the machine, usually, traditional HMIs have input devices like switches, buttons, handles, dials, pedals, remote-controllers, microphones, keyboards, mouses, and so on. For the human user to receive information from the machine, there are output devices like screens, meters, lamps, speakers, etc.. These input devices and output devices, together with the softwares and algorithms processing the information (data), are the key components of a human-machine interface.

This research focuses on a certain kind of human-machine interfaces whose inputs are biomedical signals. Biomedical signals are generated from our bodies, showing information about our health states, physical or mental activities. This information can be captured through physiological instruments that measure heart rate, blood pressure, oxygen saturation levels, blood glucose, nerve conduction, eye movements, muscle movements, brain activities and so forth.

1.2 Biomedical signals

In this research, three kinds of biomedical signals are involved. They are: electrooculogram (EOG) which relates to eye movements, electromyogram (EMG) which relates to muscle movements, and electroencephalogram (EEG) which relates to brain activities. All these biomedical signals are electrical signals which can be measured by electrodes. Traditionally, these signals are analyzed to provide useful information upon which clinicians can make decisions. For example, EMG are used in muscle rehabilitation. EEG are used in diagnosing Alzheimer's disease and epilepsy; EOG are used in identifying sleep stages and conditions.

1.2.1 EEG

Electroencephalography (EEG) is a technique measuring the electrical activity of the brain. Its resulting signal is called electroencephalogram (EEG). EEG is based on the propagation of electric impulses along a nerve fiber when the neuron fires. While a signal neuron cannot produce detectable potential, EEG reflects the summation of thousands or millions of neurons that have similar spatial orientation.

EEG is the most useful and important test in identifying epilepsy. It is also used to diagnose sleep disorders, coma, encephalopathies, and brain death. Compared with other brain imaging techniques such as MRI and CT, EEG has poor spatial resolution but extremely high temporal resolution, which makes it also possible to be used as a control signal.

1.2.2 EOG

Electrooculography (EOG) is a technique measuring the potential between cornea and retina, and the resulting signal is called electrooculogram (EOG). EOG is primarily used in ophthalmological diagnosis and in recording eye movements.

To measure eye movement, pairs of electrodes are typically placed above and below the eye (the resulting signal is called vertical EOG) and to the left and right of the eye (the resulting signal is called horizontal EOG). If the eye moves from center position toward an electrode, the electrode becomes near with the cornea which is positive, while the opposite electrode becomes near with the retina which is negative. Consequently, a potential difference occurs between the electrodes.

1.2.3 EMG

Electromyography (EMG) is a technique measuring the potential produced by muscle activities. The resulting signal is called electromyogram (EMG). Muscles are electrically

inactive at rest. Action potentials occur if muscles are stimulated. Increased strength of contraction leads to increased number of motor units being recruited, and thus produces an electrical signal with different rates and amplitudes.

EMG can be analyzed to detect muscle activation level. It is used as a diagnostics tool for detecting medical abnormalities, like neuromuscular diseases and disorder of motor control. EMG signals are also used as a control signal for prosthetic devices such as prosthetic hands, arms, and lower limbs.

1.3 Biomedical signal based HMI

Engineers are discovering new ways for the application of these biomedical signals, and thus it comes out the human-machine interfaces based on the biomedical signals. Until now, there have been various kinds of human-machine interfaces utilizing these biomedical signals.

For example, researchers have tried to build a connection between the human brain and the computer, and they created a special kind of HMI called brain-computer interface. A representative brain-computer interface is the EEG-based P300-speller. This interface allows the user to input words by only focusing on the flashing characters displayed on the screen as the unpredicted event (flash) can evoke a potential in EEG. This kind of interface can help disabled persons who lose the function of writing and speaking regain the abilities of communication. Although to extract useful information from brain signals like EEG is challenging, and its practicality is usually hampered by the low information transfer rate, the brain-computer interface has the potential of becoming a new interaction means without any muscle activity involved.

EOG is usually used for eye tracking. Compared with a video-based eye tracker, it can be more portable and inexpensive since it only needs electrodes and amplifiers. Through analyzing the eye movement, it can be used to recognize human's behaviors like reading a paper, watching a video, browsing the web, and so on. Also, it can be used in as a control interface, for example, an EOG-based remote-control TV. It is a very useful technology for providing support to those with disabilities which happen to restrict them from using other common HMIs.

EMG has been used in HMIs that control devices like exoskeletons, wheel-chairs, and prostheses. There has been some commercial products available using the techniques of EMG-based HMI. For example, HAL is a powered exoskeleton suit developed by Japan's Tsukuba University and the robotics company CYBERDYNE. It generates assist torque according to the user's surface EMG signals. MYO is an EMG armband developed by the Canada company ThalmicLabs. It can recognize about seven gestures. Those gestures can be used as control commands in presentations, video games, smartphone apps, and so on.

1.4 Paper organization

In Chapter 2, an EOG-based interface for vigilance estimation is proposed. Vigilance is an index to measure the degree of concentration; it can be affected by monotonous task, fatigue, or sleepiness. This study obtains a reference of vigilance from an experiment of long-time monotonous task. From recorded EOG signals, about 20 EOG features related to slow eye movement, saccade, blink, and energy are extracted, and the correlation between the EOG features and the vigilance reference is examined. The results indicate that some EOG features are highly related to vigilance and thus are possible to be used for vigilance estimation. In addition, the placement of EOG electrodes is discussed for realizing an easy wearing device.

In Chapter 3, an EOG/EEG hybrid interface for robot control is proposed. Two kinds of robots were used in this study. One is the humanoid robot NAO, the other is the mobile robot Kobuki. The HMI is designed to control those robots to perform different tasks. There are two ways of control: by EOG and by EEG. In EOG mode, the subject can use eye movements such as blink, wink, gaze, and frown to control the basic mobility of the robots. In EEG mode, the subject can send commands to the robot by focusing on the menu items, where the flashing items evoke potentials in EEG.

In Chapter 4, an EMG interface for prosthesis control is proposed. This HMI allows the user to control a prosthesis to perform four kinds of movements: open, close, pronate, and supinate. The input EMG signals are measured from the forearm, and processed by the nonnegative matrix factorization (NMF) algorithm. Through a training process, the NMF algorithm is able to map the multi-channel EMG signal into the four kinds of hand/wrist movements. By this approach, proportional control of multiple degrees of freedom (DOF) can be realized. This means the output does not only contain simple states of open, close, pronate, and supinate, but also contains all the transition states between open and close, pronate and supinate, as well as some combination states like “open and pronate”, “close and supinate”.

Finally, Chapter 5 is about the conclusion and the future works.

Chapter 2

Electrooculogram based Vigilance Estimation

2.1 Introduction

Driver fatigue causes considerable casualties in our daily life. Each year in the world, about 600,000 people die and 10,000,000 people injure due to the traffic accidents. In [3] it mentions: “A conservative estimation from US National Highway Traffic Safety Administration has shown that the annual traffic deaths due to driver fatigue count at least 100,000. The French National Police Department incident report indicates that traffic crashes due to driver fatigue, accounted for 14.9% of personal injuries and 20.6% of fatal accidents.” In addition, official statistics indicate that, among the total traffic accident deaths in 2007~2009 China, the deaths resulted directly from driver fatigue are 11.35%, 10.91% and 12.5%, respectively [4].

Compared to drunk driving, driver fatigue occurs spontaneously. It is a hidden danger that is difficult to be restrained by law or by self-control. Thus there is an urgent need of developing reliable technology for detecting driver fatigue. At present, a mainstream solution for monitoring driver fatigue is a kind of human-machine interface that uses infrared or normal camera to capture driver’s eye movements [5–7]. Since the eye movement patterns have been proved related to fatigue/sleepiness, the input of the HMI is the driver’s eye movements, and the output is the estimated fatigue (sleepiness) level. This video based HMI benefits from the advancement of image processing and computer vision technologies, and is convenient for users since the sensor devices do not contact human body. However, its reliability is relatively low and thus is inadequate for professional usage, e.g., the full-time car drivers, the locomotive engineers, or the aeroplane pilots. The main problem of this kind of HMI is that the infrared type only locates pupils and thus is limited in the variety of detectable features, while the normal type is vulnerable

to changeful environment conditions like illumination and face angles. Although the eye tracker is impeccable in these aspects, its expensiveness and ponderosity make it hardly a good choice.

To overcome the shortcomings of the existing video based methods, an alternative choice to solve this problem is to use a biomedical signal based HMI. For example, various approaches using electroencephalography (EEG) for vigilance estimation have been developed in the last several years [8–15]. Vigilance, sometimes also referred as alertness, is the ability to maintain attention over a prolonged period of time so it can be used to describe and measure driver fatigue. EEG directly reflects whether driver's awareness is active or inactive and cannot be faked purposely. Clinical researches usually take EEG as the gold standard due to its high accuracy. In spite of this, it is still difficult to bring an EEG-based HMI to practical driving situation mainly because of its extremely low signal to noise rate (SNR). In general but not experimental circumstances, the complicated background signals including electrooculogram (EOG), electromyogram (EMG) and so on would seriously hinder the identification of EEG. In addition, for the existing manufacturing techniques of dry-electrodes for acquiring EEG data, only rarely can they ignore the obstruction of hair. For these reasons, this study proposed another HMI based on EOG [16–18].

EOG is generated from the potential difference of retina and cornea. It reflects eye movements like the blinks and saccades. Normally two pairs of electrodes are used for EOG recording. The vertical EOG is recorded from two electrodes above and below one of eyes, and the horizontal EOG from two electrodes at the outer canthus of each eye. From the vertical and horizontal EOGs, it is possible to recognize and identify the eyelid and eyeball movements for vigilance estimation. Although EOG may not reflect the vigilance so directly as EEG does, it benefits from high SNR and easy recording. Compared with the video based methods, which utilize similar eye movement features, EOG is more likely to achieve high time resolution, high recognition rate, and large feature variety. Very recently, the characteristics of EOG have attracted wide attentions in various fields, such as activity recognition [19], sleeping staging [20], and visual fatigue estimation [21].

First and foremost, the researches on EOG and vigilance are still at the pilot stage and most of them only consider one kind of EOG features. In order to take full advantage of the rich information contained in EOG, it is significantly important to utilize and combine various EOG features together. For example, although blink parameters are usually effective and widely used, they have their limitations [22]. Introducing other eye movement features grants the system additional reliability and stability. Thus in this study, totally twenty different EOG features extracted from slow eye movements, saccades, blinks, and EOG energy are analyzed.

Second, almost all studies rarely mention the practicability of the existing methods. The practicability, like on-line processing ability, wearing convenience, etc., is very important for an HMI. In this aspect, a linear dynamical system (LDS) approach

is adopted, which is compatible with both off-line and on-line processing. This method could reduce the noise from EOG data without time delay. Moreover, a new EOG electrode placement is proposed, where the electrodes are placed on the forehead instead of around the eyes. This placement is considered more convenient for wearing the device. Independent component analysis (ICA) is used to retrieve the same signals as if the electrodes were at the original places.

2.2 Related Work

2.2.1 EOG Applications

EOG has been adopted in implementing human-machine interface related to device controlling and condition monitoring. Some examples of EOG-based HMI are introduced as follows.

Device Control

In the control field, EOG could be effectively implemented in human-machine interface while people can operate computers and robots through EOG signal inputs. Especially, EOG plays a great role in situations where operator's hands are busy or disabled. For example, the news has reported that British air army has invented a system carried on helmets allowing pilots to lock the target only by sightline, when their hands are freed to manage complex maneuvers in a battle. For disabled people, EOG based assistive interfaces enables them to perform the character input, wheelchair control, or robot-arm control [23,24].

Condition Monitoring

For condition monitoring, the EOG applications include sleep stage analysis, vigilance estimation, and activity recognition. According to R.K sleeping criteria (Rechtschaffen and Kales, 1968), slow eye movements and rapid eye movements are the important features representing the sleeping stages. Thus clinical researches usually use EOG to classify the sleeping stages. The estimation of vigilance, sometimes also referred as alertness, drowsiness or fatigue for some special cases like driving, is specifically studied in this work. It is developed from the sleeping research with the same principle, and it concerns the vigilance changes during the wake stage and the first sleeping stage. Finally, the activity recognition research uses EOG as input signal to judge the subject's activities as typing, reading, writing, watching a video, browsing the web, and so on [19].

2.2.2 Vigilance Estimation

Various methods for vigilance estimation have been developed in the last several decades. These methods can be classified into the following three main categories according to the vigilance indices used in the estimation systems.

Physiological Signals

Representative methods are by EEG, EOG, or ECG (electrocardiogram). There are also other methods which measure the breath, blood pressure, pulse, skin resistance, skin temperature [25], and so on. Among all of these vigilance indices, EEG has been widely accepted in medicine, but more focuses on clinical applications. EOG is being given attention and has development potential. Except the EEG which reflects the brain activity and EOG which reflects the eye movement, other physiological features still lack for enough theoretical and practical supports to prove whether they are strongly related to vigilance.

Behavioral Characteristics

This kind of methods locates the head or eyes, and observes the visible features like facial expressions, sightline directions, head movements, and eye movements [5]. Among them the best vigilance index is eye movements. PERCLOS, which stands for the Percentage of Eye Closure, has been confirmed as a reliable index of fatigue by Federal Highway Administration of the U.S., and has been successfully applied in on-vehicle products for fatigue driving detection [26]. This kind of methods is featured by computer vision techniques and contactless sensors, so EOG is not included.

Vehicle Parameters

This kind of methods is specialized for fatigue driving. It indirectly estimate the driver's fatigue by continuously measuring the driving parameters of vehicles, for example, the driving speed, the deviation from the traffic lanes [9], the abnormal movement of the steering wheel, and the grip strength to the wheel. Compared with the previous two kinds of approaches, this kind of methods is prone to influenced by vehicle types and driving environment.

2.3 Materials and Methods

2.3.1 Eye movements and EOG Features

EOG features can be extracted from different eye movements, for example, slow eye movement, saccade, and blink. While most existing studies only utilize a few EOG features from one kind of eye movement, in this study, a thorough investigation is made on a large number of EOG features for vigilance estimation.

Slow Eye Movements

Slow eye movement (SEM) refers to a kind of eye movement gradually increases in frequency when entering the first sleeping stage from wakefulness. SEM is a sign of first sleeping stage, and is closely related to the activity of alpha rhythm of EEG, which is a key feature of vigilance [27]. Moreover, the increase of SEMs also accompanies decrease of eyelid movements. All these evidences indicate the connection between SEMs and vigilance. At present, however, SEM is seldom utilized in vigilance studies. Hanke and colleagues proposed that LF/HF (the ratio of low frequency energy to high frequency energy) is a good vigilance index, where the low frequency is just the SEM frequency band [28].

Saccades and Rapid Eye Movements

Saccade and rapid eye movement (REM) are usually confused with each other in definitions, because they have similar amplitudes and waveforms. Some clarify that REM is a kind of eye movement during sleeping stages while saccade during wakefulness. The speed and frequency of saccades have been applied in the fatigue detection by a video based method (with a head tracker and a eye tracker). The frequency of REM extracted from EOG has also been analyzed as an index of alertness [29], and has been used in a prototype of PDA drowsiness detector [30].

Blinks

Blink features are most prominently used in video based systems. Almost all these systems choose blink duration or PERCLOS as one of the vigilance indices. EOG based methods are inclined to use blink features too. Blinks can be detected from EOG by using double thresholds method [1] or wavelet transform method [19]. By using EOG it is easier to access to some features like eyelid speed than by infrared devices. However, it does not make the best use of EOG if other feature types are not considered.

In this study, an EOG-based vigilance estimation interface is proposed. The overall architecture of this interface is shown in Fig. 2.1. All the included methods are coded

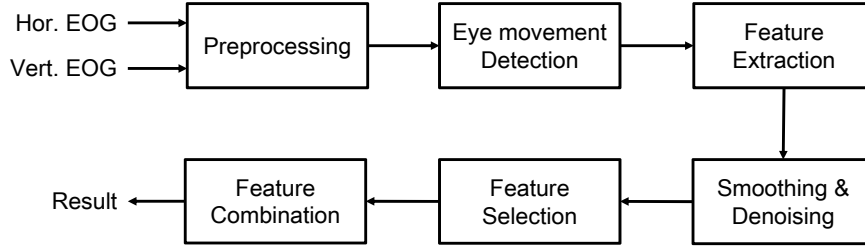


Figure 2.1: The flow chart of the proposed EOG based vigilance estimation interface.

by MATLAB. The EOG data together with the local error rate of subjects obtained from experiments are taken as the inputs of the interface.

The first step is eye movement detection and feature extraction. A time window of 8 seconds without overlap is applied to the vertical and horizontal EOGs. Three kinds of eye movements: SEM, saccade, and blink, are detected, and their features are extracted. The second step is de-noising and smoothing. A linear dynamical system (LDS) approach is used to remove the vigilance unrelated components while preserving the vigilance related parts. The final step is feature combination and analysis. Features with high relation to local error rate are selected and combined together. The final output is a time-varying curve which is the estimation to the subject's vigilance.

2.3.2 Preprocessing

Because the major components of EOG signals concentrate on low frequency band, filtering is usually necessary, which can safely remove the components above 30 Hz. Besides filtering, other preprocessing methods commonly used for physiological signals, like baseline drift removal and independent component analysis are not essential for EOG. These two issues are discussed as follows.

Baseline Drift Removal

Baseline drift is a steady movement of the baseline over time. In ECG signals it is the biggest hurdle in visualization of correct waveform and computerized detection of wave complexes based on threshold decision [31]. EOG and EEG signals are also vulnerable to baseline drift if their amplitudes are considered. Nowadays, wavelet transform is a prevalent method for removing the baseline drift from the original signals [32].

In this work, the eye movements detection methods are insensitive to baseline drift. During the SEM detection, discrete wavelet transform is performed and the approximation components of wavelet are removed, which is just similar to the baseline removal process.

For saccades and blinks, the first difference of signal is calculated before applying the thresholds. Thus the influence of baseline drift will be damped through the differential process.

Independent Component Analysis

For EEG signal processing, independent component analysis (ICA) is a significant procedure for artifact removal due to EEG's extremely low SNR [33]. Artifacts in EEG consist of EOG, ECG, and EMG. Even breath and galvanic skin response may have undesirable effect on EEG acquisition. But for EOG signal processing, only EMG artifact generated by facial movements may contaminate EOG to a certain extent and it is not of common occurrence. Therefore, ICA is relatively optional. However, if the EOG electrodes are moved to the forehead to improve the wearing experience, ICA is necessary to be used to decompose forehead signals into horizontal and vertical EOGs [17].

2.3.3 Eye Movement Detection

Let's first consider the case of standard electrodes placement for acquiring horizontal EOG and vertical EOG. The former represents horizontal eye movement and the latter represents more eyelid movement than vertical eye movement. SEMs and saccades are detected from horizontal EOG, and blinks from vertical EOG.

For video based systems except eye trackers, the available features are usually limited to eyelid opening and closing time, and so on. By using EOG, the access to SEMs and saccades becomes much easier. Also it is possible to exploit some features related to EOG signal energy. The detection methods for processing SEMs, saccades and blinks are introduced as follows.

Slow Eye Movements

The SEM detection method is referred from the work of Magosso and colleagues [34–37]. This method is based on discrete wavelet transform (DWT) and includes three steps: wavelet decomposition, energy computation, and discriminant function.

In the first step, DWT decomposes the input signal into subcomponents of different frequency bands. The 125 Hz input signal is processed by 30 Hz low pass filter followed by a 10-level DWT of Daubechies order 4 (db4) wavelet. This process divides the signal into components of 10 scales where the scale k contains $125/2^{k+1}$ Hz to $125/2^k$ Hz. Since SEMs are typically 0.2 Hz to 0.6 Hz, scale 7 to 10 are necessary to capture the SEM band. The reason of choosing db4 wavelet basis function is that its shape strongly resembles SEM waveforms [16].

The second step is using wavelet coefficients to compute the energy associated with each component. When SEMs become prominent, the energy of high scale components increases and that of low scale components decreases. Knowing the energy of each scale would help identify SEM periods. The energy of DWT is calculated by

$$D_i = \sum_{j=1}^{M/2^i} |d_{ij}|^2 \quad \text{for } i = 1, 2, \dots, 10, \quad (2.1)$$

where D_i is the total energy of the component of scale i , and d_{ij} is the j th instance of wavelet coefficients of D_i . Generally, for a discrete input signal with a length M , the number of wavelet coefficients in each scale is approximately equal to $M/2^i$. Eq. (2.1) means that the total energy of a scale is equal to the quadratic sum of all the wavelet coefficients in that scale.

Obviously, from D_i to D_{i+1} , the number of wavelet coefficients is approximately halved. When we compute the energy in a fixed time interval, one coefficient of D_{10} would correspond to 512 coefficients of D_1 , thus it is necessary to unify the coefficient resolution of each scale. The resolution of D_6 is taken as the standard. Then each two coefficients of D_5 would be grouped into one, and each coefficient of D_7 would be divided into two, with the total energy unchanged, and so forth.

For the third step, there is a discriminant function identifying SEM period by comparing the energy ratio with a threshold value as follows:

$$f = \frac{\sum_{i=7}^{10} W_i D_i}{\sum_{i=3}^5 W_i D_i + \sum_{i=7}^{10} W_i D_i}, \quad (2.2)$$

where W_i is the weight for D_i .

In Eq. (2.2), D_1 , D_2 and D_6 are excluded because D_1 and D_2 include high frequency noises, and D_6 is ambiguous in its behavior. This equation means the ratio of low frequency energy to total energy. In this function, the weights and the threshold were determined by a training process which contains visually inspection and validation. The trained weights and the threshold are adopted from the original work [35].

Saccades

Comparing with SEM or blink detection, the saccade detection algorithm used in this study is quite simple. By using wavelet method, it can precisely locate the saccades together with the fixations. However, for driver fatigue detection application, it rarely cares about the fixations. Thus a threshold algorithm which could detect saccades conveniently is used.

The algorithm is originally designed for detecting REMs with a velocity threshold of 15 degrees per second [29]. If the eye movement velocity exceeds this value, then that

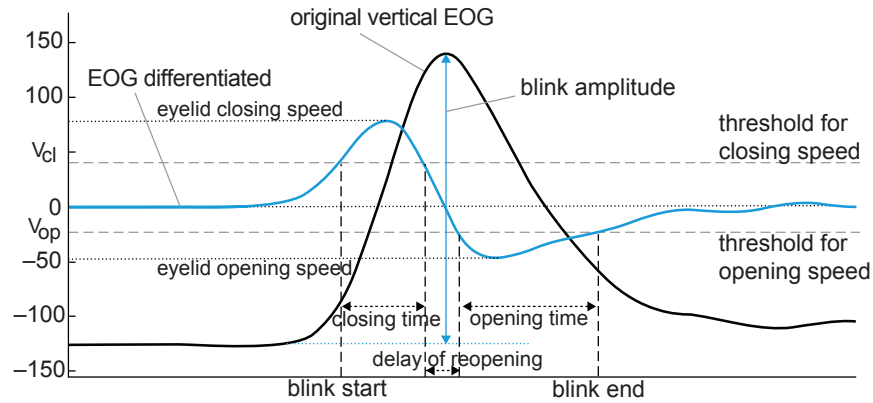


Figure 2.2: The waveform of a blink in vertical EOG and its differentiated shape. The x-axis is time and y-axis is EOG (μV).

time point is considered as the start time of REM. Saccades resemble REMs very much but only 10% to 30% higher in top velocity. Thus the threshold can be adjusted to detect the saccades. The algorithm consists of the following three steps:

Step 1: Process the horizontal EOG signal by 1 Hz to 8 Hz filter.

Step 2: Calculate the first difference of the signal for the speed of the eye movement. Then take absolute value for the velocity.

Step 3: Find all the points which are greater than the threshold. Those points abutting each other in time series belong to the same saccade. Here the threshold is set as 5.0 by observation.

Blinks

The blink detection algorithm is an improved version of the double thresholds method. Its key idea is to set two thresholds which represent the eyelid's closing speed and opening speed, respectively. The two thresholds can locate blink waveforms in EOG. This method is executed on the vertical EOG signal after difference, so it has resistance to baseline drift compared with single threshold method.

Figure 2.2 depicts a typical EOG waveform of blink, where the black curve is the vertical EOG, and the blue curve is the EOG after taking difference. When the black curve reaches top, the blue curve is just 0. This means when the eyelid is completely closed, its speed is equal to 0. V_{cl} and V_{op} are the two thresholds for judging eyelid closing and eyelid opening, respectively.

The algorithm for detecting blink consists of the following five steps:

Step 1: Process the vertical EOG signal by 10 Hz low pass filter. Because the frequency of blink is very low, the high frequency components unwanted are removed.

Step 2: Calculate the first difference of the signal. The result is the approximate speed of the eyelid.

Step 3: Apply the thresholds V_{cl} and V_{op} to the signal. A blink event should consist of four successive points where the first two are equal to V_{cl} and the last two are equal to V_{op} (not precisely equal because the data is discrete). All matched point series will be taken as blink candidates. For more accurate positioning, the beginning point and the ending point should be extended from $\{V_{cl}, V_{op}\}$ to $\{E_{cl}, E_{op}\}$.

Step 4: Combine the blink candidates which are very closed to each other. In some cases two adjacent blink events are so close that it is better to treat them as one blink. If the distance from the end point of one blink to the start point of another is less than a value δ_b , we will combine them into a single blink.

Step 5: Verify the duration and amplitude of blink candidates. This is to discriminant between the vertical eye movements and blinks. If the duration is less than 0.5 s, and the amplitude is larger than the given A_{min} , then the candidate will be regarded as a true blink. If the duration is larger than 0.5 s, then A_{min} needs to multiply a parameter λ which is larger than 1. This is because it is found that blinks longer than 0.5 s have larger amplitude than normal ones.

It should be noted that among all the parameters, the values of V_{cl} and V_{op} are the most important factors which greatly influence the detection result. Very slow blinks can also be safely detected if these thresholds are well chosen. In this paper, we use the same parameters with good average performance for all the subjects. The parameters are: $V_{cl} = 10$, $V_{op} = -10$, $A_{min} = 100$, $\lambda = 1.5$, $E_{cl} = 3$, $E_{op} = -3$, and $\delta_b = 1$.

2.3.4 Feature Extraction

Twenty different features belonging to four groups are extracted from the detected eye movements. Table 2.1 shows a short list of these features. All of the features are calculated within a 8-s time window. Because the data is very large, the step length is set to a maximum value 8-s in order to save calculation time. In on-line situation, it is reasonable to set the step length about 2-s, which means to output the vigilance estimation every two seconds.

- Three features for SEMs: proportion, mean and variance of amplitude.
- Four features for saccades: proportion, peak velocity, mean and variance of amplitude.
- Eleven features for blinks: duration, closing/opening time, delay of reopening, delay time ratio, interval, mean amplitude, closing/opening peak velocity, and closing/opening mean velocity.

Table 2.1: List of EOG features analyzed in this work

Variable	Feature	Description
F_1	SEM proportion	
F_2	SEM MAm	Mean amplitude of SEMs
F_3	SEM VAm	Amplitude variance of SEMs
F_4	Saccade proportion	
F_5	Saccade PVe	Peak velocity of saccades
F_6	Saccade MAm	Mean amplitude of saccades
F_7	Saccade VAm	Amplitude variance of saccades
F_8	Blink duration	Calculated from the start point of blink to the end point of blink
F_9	Closing time	
F_{10}	Opening time	
F_{11}	Delay of reopening	The duration of the eyelid keeping closed, similar to PERCLOS
F_{12}	Delay time ratio	Calculated by F_{11}/F_8
F_{13}	Blink interval	Calculated from the previous end point of blink to the start point of blink
F_{14}	Blink MAm	Mean amplitude of blinks
F_{15}	Closing PVe	Eyelid's peak velocity in closing phase
F_{16}	Opening PVe	Eyelid's peak velocity in opening phase
F_{17}	Closing MVe	Eyelid's mean velocity in closing phase
F_{18}	Opening MVe	Eyelid's mean velocity in opening phase
F_{19}	HEO LF energy	Calculated by $\frac{PSD(0\sim 1Hz)}{PSD(1\sim 20Hz)}$, on horizontal EOG
F_{20}	VEO LF energy	Calculated by $\frac{PSD(0\sim 1Hz)}{PSD(1\sim 20Hz)}$, on vertical EOG

- Two features for energy: the power spectral densities (PSDs) of low frequency components (0 Hz to 1 Hz) on horizontal EOG and on vertical EOG, respectively.

Among all the features, PSD is a special case which is not calculated from detected eye movements but directly from original EOG signal. The idea is using fast Fourier transform (FFT) to calculate the energy of a certain frequency band within the time window. The result has the form of LF/HF , where LF is the energy from 0 Hz to 1 Hz and HF is the energy from 1 Hz to 20 Hz. Those boundaries are chosen properly, and the DC component (energy at 0 Hz) is removed beforehand. PSD may also be counted as SEM and blink feature because 0 Hz to 1 Hz covers SEM band on horizontal EOG and blink band on vertical EOG.

2.3.5 Smoothing and De-noising

Moving average method has been widely used for signal smoothing. For EOG features, smoothing is also necessary since they are extracted by a short time window thus have large fluctuations. However, moving average is not suitable for on-line processing because it causes time delay. In moving average method, each time point is calculated by averaging its previous and following data, so does that mean it will produce an one-minute delay if the time window is chosen as two-minute length. To overcome this shortcoming, we introduce a linear dynamical system (LDS) approach [13], a semi-supervised dynamic model. The LDS approach is suitable for both off-line and on-line processing and has been successfully used in our previous work [16, 38].

The LDS approach is a kind of state space model. It is capable to use only the previous data to evaluate the state of the current time point, which means no time delay and thus it is suitable for on-line processing. The structure of the state space model is shown in Fig. 2.3, where z_i means hidden variable and x_i means observed variable. Internal relation (shown by arrows) exists in the hidden variables, but does not exist in observed variables due to observation noise and so on. The meaning of this model is that, through the relation among hidden variables and the relation between hidden variable and observed variable, we can deduce the probability distribution at the current moment. Specifically, we assume that the states we observed (x_i) are noise-free states (z_i) superimposed by gaussian noises, and the time transition among the noise-free states is also gaussian. According to this assumption, we can calculate the expectation and variation of these gaussian relations and then get z_i by knowing x_i .

EOG features include vigilance-related parts as well as vigilance-unrelated parts which are treated as noises. Usually, the vigilance-related parts are slow varying whereas the noises change irregularly. According to the discussion mentioned above, the LDS approach calculates the hidden states which correspond to vigilance-related parts, and the noises are removed in this process. The comparison of EOG signals before and after smoothing by

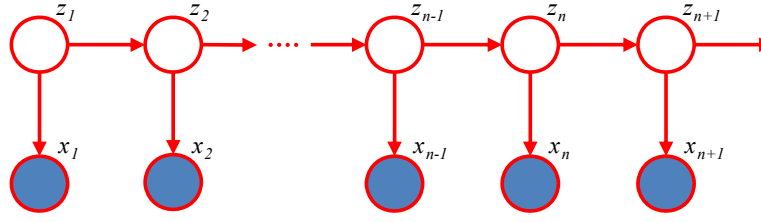


Figure 2.3: The structure of state space of LDS. The hollow circles are hidden states, and the solid circles are observed states. Arrows mean inferable relations.

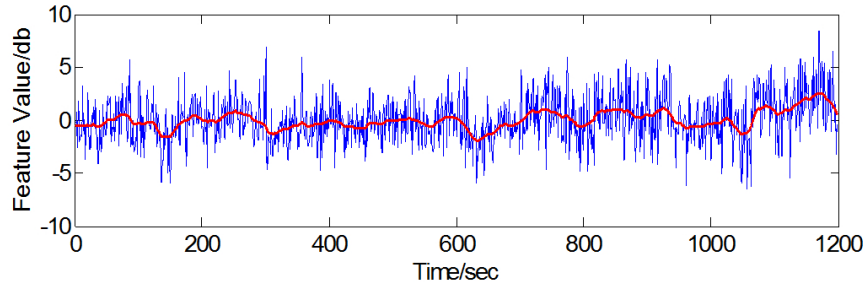


Figure 2.4: Comparison of EOG signals before and after using LDS approach. Here the blue line denotes the original EOG signal and the red one denotes the EOG signal smoothed by using the LDS approach.

using the LDS approach is illustrated in Fig. 2.4. Since the LDS approach is a semi-supervised method, no label information is needed.

2.3.6 Feature Combination

After using the LDS approach, a number of de-noised features are obtained. Then they should be verified, as well as their combinations since it is assumed that combination of features would yield better result than any single feature. EOG Features are combined by following three methods. The final result is revealed in next section.

Linear Summation

This is a simple combination method. All of the features are added together after normalization. Here, all the features have the same weight and thus make the same contribution to the final result.

Principal Component Analysis (PCA)

PCA is a commonly used dimension reduction method. It calculates the linear combinations of original features by multiplying a orthogonal projection matrix. The first projection direction is selected by assuring that the features have largest variance on that direction. Then all the following projection directions are also selected in this way while keeping orthogonal to each other. The projection result is a series of new features arranged in order of importance.

Manifold Learning

It is a non-linear unsupervised learning method for dimension reduction. It is based on multidimensional scaling method which remains the Euclid distance among data points unchanged after the dimension reduction. In this study, Isomap is adopted, which uses geodesic distance instead of Euclid distance for better adaptivity to data manifolding in the high dimension space.

2.3.7 Placement methods

Although placing electrodes near the eyes can obtain the most direct and noise-free EOG signals, this method has many limitations in practical applications. Thus, a new placement method is proposed based on the following criteria. First, electrodes should not affect normal operations or the sight of subjects. Second, electrodes should be easy to wear, like wearing a cap or band on the head. Third, given the electrode placement, the detection system should have a certain robustness. For example, it should avoid interference from artifacts or deliberate actions of subjects. Taking account all of these constraints, it is considered best to put electrodes on the forehead. Among the facial muscles, the forehead has less motion than most. It is normally bare so electrodes are not disturbed by hair on the scalp. It is also close to the eyes. Moreover, in this way it is really easy to implement all the electrodes on one band which makes the device quite practicable. The comparison between the traditional placement and the rearranged placement is shown as Fig. 2.5.

A prior study is made to verify whether it is possible to retrieve horizontal and vertical EOG signals (HEO and VEO) from forehead electrodes. Fast independent component analysis (FastICA) is used to decompose the forehead signals into horizontal and vertical EOG components. The input data is a matrix containing data from the four channels of observation signals from the rearranged placement, and the output is four channels of signals after separation. It is expected that among the output signal channels, one would represent HEO and another VEO. (SVM is used to identify which represents HEO and which represents VEO.)

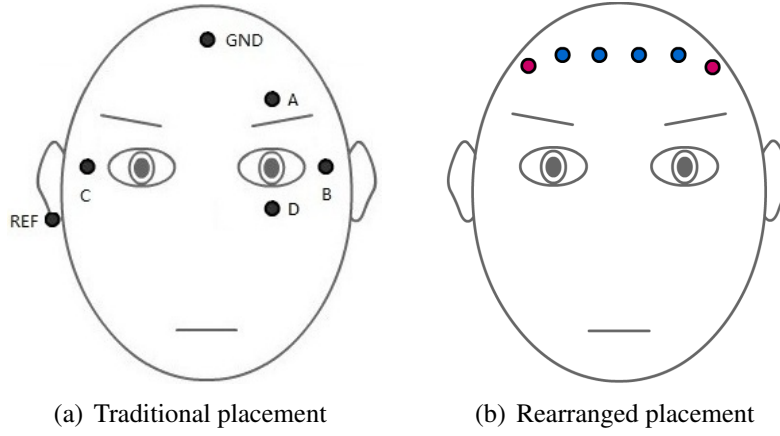


Figure 2.5: Comparison of two placements of EOG electrodes. REF and GND in (a) correspond to the red circles in (b); A, B, C, and D in (a) correspond to the blue circles in (b).

2.3.8 Experiment

Experiment of vigilance estimation

To collect EOG data for evaluating our proposed vigilance estimation method, an experiment of monotonous visual task [15] is carried out. The scene of the experiment is shown in Fig. 2.6.

The subject wears an electrode cap, and holds a button box. His/her task is to push the correct button according to the color of the image displayed on the screen. In this experiment, we use traffic signs in four colors which are red, green, yellow, and blue. Each color has 160 different signs, 640 in total. The sign is displayed for 0.5-s, then the screen turns black for a 5-s to 7-s interval before the next sign is presented. The whole procedure lasts for about 70-min.

When a sign appears on the screen, the subject is supposed to push the button of the same color with the sign. At the same time, the system records the correctness of each response (“no response” will count as incorrect) [16]. The more fatigued the subject is, the higher is his/her error rate. Then we can get a curve of error rate changing by time, which is taken as the reference of vigilance.

The physiological data in the experiment is recorded by NeuroScan system. The EOG signals were recorded in 32-bit resolution, sampled at 500 Hz, and filtered between 0.1 Hz and 100 Hz [16]. A camera monitors the subject’s condition, and the Stim software records his/her responses.

The subject is supposed to constrain his/her sleep time to less than six hours during the previous night in order to actually invoke the transition from wake to a sleepy state.

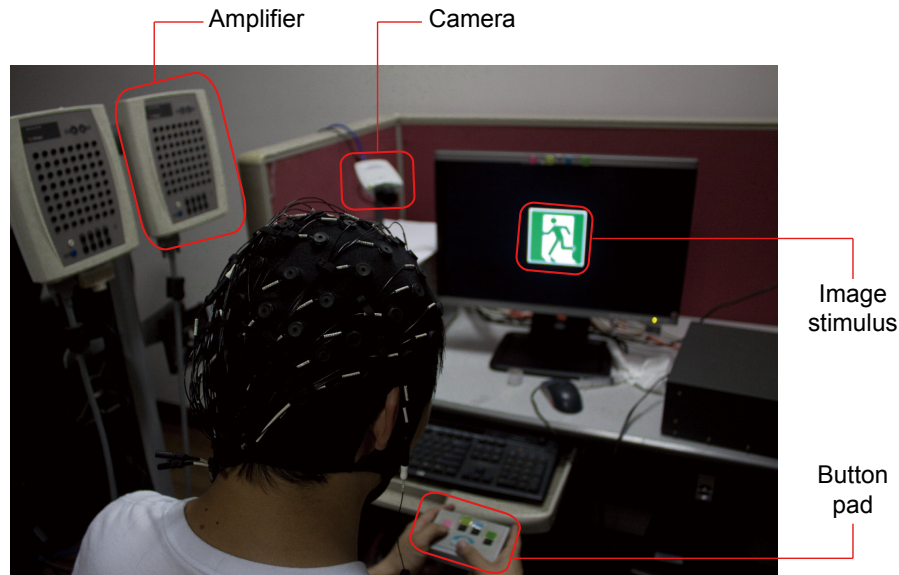


Figure 2.6: The scene of the image identification experiment. The EOG data are collected by four electrodes on the EEG cap. The subject's task is to push the button which shares the same color with the image stimulus.

The experiment is performed after lunch when the subject is most likely to feel drowsy. During the previous day, the subject is restricted from drinks like liquor and coffee which would stimulate the brain.

Experiment of electrode replacement

The experiment for obtaining forehead EOG signals is a monotonous imitation task. The subject sits about two feet away from an LCD screen in a noise-free room. In each session, the screen displays 14 different face movement examples that would evoke various kinds of EOG and EMG artifacts. The subject is asked to perform the same action synchronously. All the 14 movements are chosen to represent possible human behaviors during vehicle driving; they include bending or shaking the head, clenching the teeth, blinking, eye rolling, horizontal or vertical eye movement, swallowing, squinting, eyebrow raising, scowling, smiling, cheek scratching, and opening and closing the mouth. The stimuli prompting the subjects to perform different movements are presented in random order. The structure of one trial is as follows: three seconds of counting down on a ready screen for preparation, then the instruction generating the specific movement is presented for 20 seconds. Finally, each trial ends with a brief black screen for relaxation. One experiment contains two sessions, each about 12 minutes, separated by a 5 minute

rest.

Ten sessions of EOG signals and response sequences from five subjects (four male and one female, all healthy and between 21 and 29) were recorded using the NeuroScan System. In order to evaluate the performance of the two electrode placement methods, both were used at the same time. That is, in addition to reference and ground electrodes, four electrodes were used to record HEO and VEO in the traditional way, and four more were used on the forehead. The recordings are sampled at 500Hz, and filtered between 0.1 and 100Hz.

2.3.9 Evaluation

In this study, the local error rate but not KDS and KSS [39] is used as the index of vigilance due to the following reasons: a) The criteria of KDS include the SEMs in EOG and the alpha and theta rhythms in EEG. These indices have already been taken as features in our EOG and EEG based research. b) For KSS, the subjects need to evaluate their drowsiness by themselves. This is biased by individual circumstances, and would harm the fluency of the experiment.

The evaluation method is mentioned in another work [15]: “The local error rate $e(t)$ is derived by computing the target false recognition rate within a 2-min time window at 8-s step as

$$e(t) = \frac{NumF_{(ST+2t-L/2, ST+2t-1+L/2)}}{NumT_{(ST+2t-L/2, ST+2t-1+L/2)}}, \quad (2.3)$$

where, ST is the start time for vigilance measurement, L is the 120-s (2-min) window length, $NumF_{(i,j)}$ is the number of false responses within the time window (i, j) , and $NumT_{(i,j)}$ is the number of total stimuli within the time window (i, j) .”

The result is represented by correlation coefficient (γ) of local error rate and EOG feature. γ ranges from -1 to 1. Higher absolute value indicates higher relevance. γ is calculated as follows:

$$\gamma = \frac{\sum_t (f(t) - \overline{f(t)})(e(t) - \overline{e(t)})}{\sqrt{\sum_t (f(t) - \overline{f(t)})^2 \cdot \sum_t (e(t) - \overline{e(t)})^2}}, \quad (2.4)$$

where $f(t)$ and $e(t)$ represent EOG feature value and local error rate, respectively. $\overline{f(t)}$ and $\overline{e(t)}$ are their average over time.

2.4 Results

The proposed method is evaluated on a data set including 36 sessions from 22 different subjects (some persons did experiments more than once). Each session is of the whole

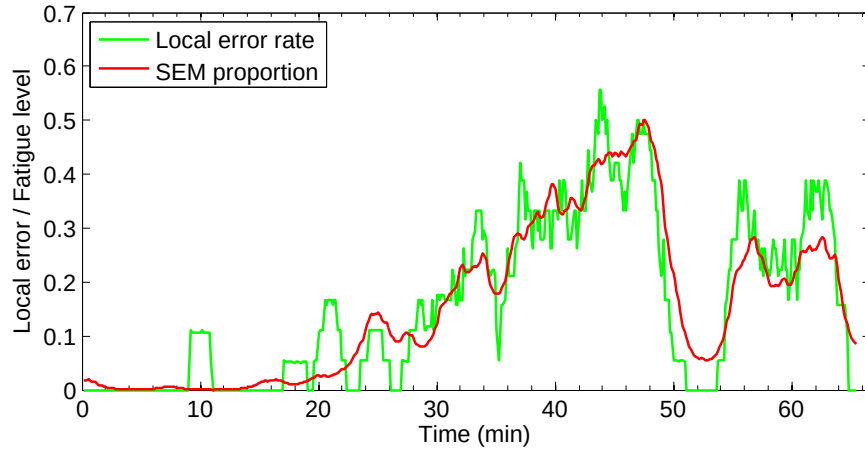


Figure 2.7: Relation between the local error rate (green curve) and fatigue level presented by SEM proportion (red curve). Here, the correlation coefficient is 0.92.

experiment data with the first 100 seconds and the last 160 seconds removed. There are more than 22 subjects involved in the experiments. However, for those who show no sign of fatigue during the experiment, their data are considered invalid. If a subject's maximum error rate does not exceed 0.35, his/her data will be discarded.

Figure 2.7 is an example of result which shows the relation between the local error rate and a single EOG feature (here is SEM proportion). From this figure it can be seen that the shape and trend of two curves are very similar. It indicates that for this subject, SEM proportion is a valid feature representing fatigue level.

Considering all 36 sessions of 22 subjects, the results are shown in Fig. 2.8. From these results depicted in Fig. 2.8, the features can be categorized into the following three classes according to their correlation with local error rate from high to low as well as their stability.

- Class $\mathcal{A}$ ($P_{25} > 0.5$ and $P_0 > 0$): SEM proportion, blink velocity (closing peak/mean, opening peak/mean), and the low frequency energy on horizontal and vertical EOGs.
- Class $\mathcal{B}$ ($0.2 < P_{25} < 0.5$): SEM mean amplitude, saccade proportion, saccade peak velocity, saccade mean amplitude, blink duration, delay of reopening, and blink mean amplitude. (Note that blink duration and delay of reopening would qualify for class $\mathcal{A}$ if one outlier session is removed from the data set)
- Class $\mathcal{C}$ ($P_{25} < 0.2$): SEM amplitude variance, saccade amplitude variance, closing time, opening time, delay time ratio, and blink interval.

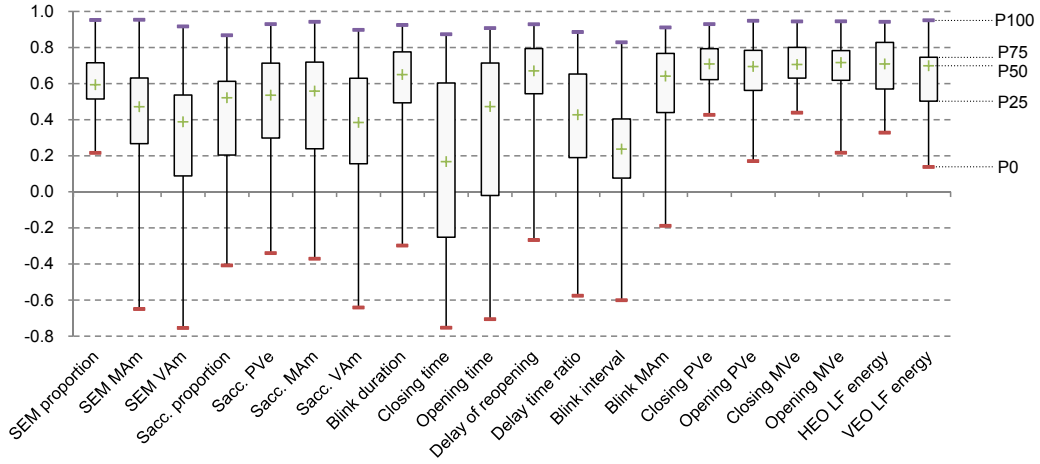


Figure 2.8: The correlation between EOG features and local error rate for all 36 sessions. The y-axis represents the value of correlation coefficient (γ). For the features which exhibit negative correlation with local error rate, their γ have been inverted for observation convenience. In the box plot, the five characteristic points from the top to the bottom down are: maximum, third quartile, median, first quartile, and minimum, which are indicated as P100, P75, P50, P25, and P0, respectively.

In order to confirm whether feature combinations bring better results, the following four different feature combinations has been tried: a) only closing and opening mean velocity; b) all class $\mathcal{A}$ features except closing and opening peak velocity; c) all class $\mathcal{A}$ features plus delay of reopening and blink mean amplitude; d) all class $\mathcal{A}$ and class $\mathcal{B}$ features. Linear summation, principal component analysis, and manifold learning are applied to examine these combinations. Figure 2.9 shows the results of these four groups of combinations. The first group consists of only two features thus yields relatively poor results compared with other three groups. Moreover, the performances of three combination methods are similar. Figure 2.10 illustrates the particular case of Group 2. Five individual features and their linear summation are shown in the figure. The mean value of γ of the linear summation is 0.75, while for individual features the highest is 0.70.

Regarding to the electrode replacement, the correlation coefficients between signals separated by FastICA and original HEO/VEO for each movement are illustrated in Fig. 2.11.

2.5 Discussion

2.5.1 Relation between EOG features and vigilance

From the experiment results mentioned above, the following observations can be drawn.

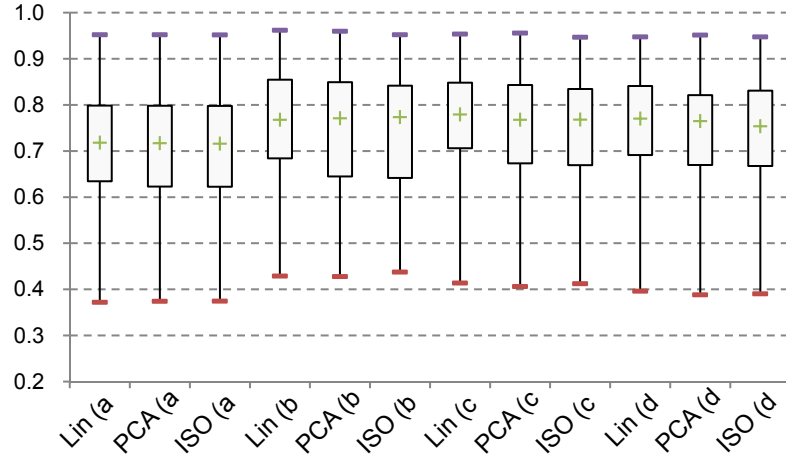


Figure 2.9: The correlation between EOG features and local error rate for 4 combination groups. Each group contains a number of EOG features which are reduced to one dimension by linear summation (Lin), principal component analysis (PCA), and manifold learning (ISO). The y-axis represents the value of correlation coefficient (γ).

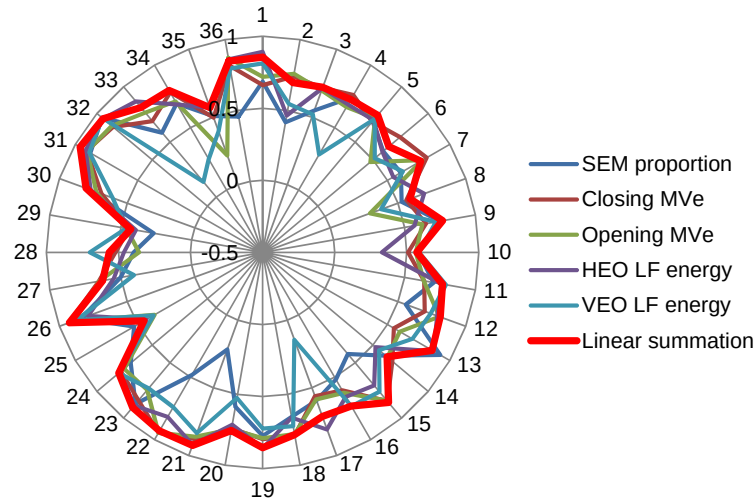


Figure 2.10: The correlation between part of EOG features and local error rate, for all 36 sessions. The red thick line represents linear summation and other colored lines are individual features in Group 2. Central region corresponds to low γ while margin to high γ .

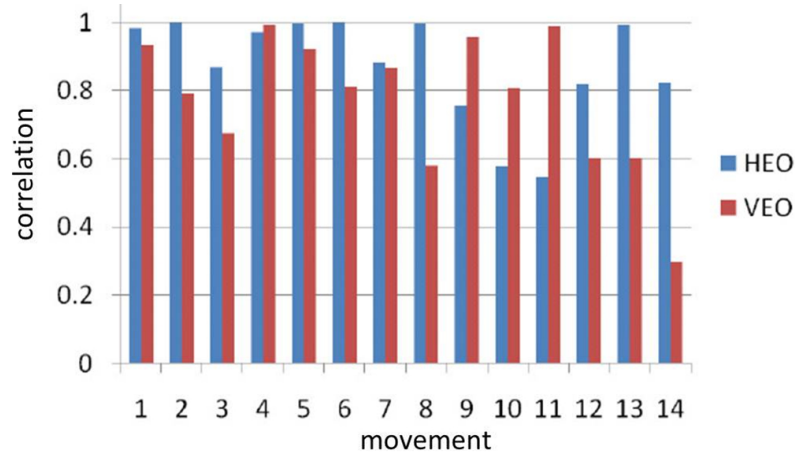


Figure 2.11: The correlation coefficients for each movement. The 14 different kinds of movements are: 1) bending the head; 2) shaking the head; 3) clenching the teeth; 4) blinking; 5) rolling the eyes; 6) horizontal eye movement; 7) vertical eye movement; 8) swallowing; 9) squinting; 10) raising the eyebrows; 11) scowling; 12) smiling; 13) scratching a cheek; and 14) opening and closing the mouth.

Blink Velocity

Among all the tested features, the blink velocity (closing peak and mean, opening peak and mean) best reflects vigilance transition (better than delay-of-reopening, or PERCLOS). When the subject fatigues, blink velocity drops obviously. This rule applies to both closing phase and opening phase, but the eyelid velocity of closing phase is a more reliable feature than that of opening phase. This fact indicates that when people are tired, the eyelid closing changes more clearly than the eyelid opening does.

SEM Proportion

SEM proportion also well reflects vigilance transition. However, for a few subjects, γ of SEM proportion is lower than expectation. This usually happens when the error rate is very high (near 100%), SEM proportion will decrease rather than increase. Research has shown that when people enter Stage 1 sleep, SEM proportion grows but when they continue entering Stage 2 sleep, SEMs will suddenly disappear. It is in accordance with the experiment results.

Low Frequency Energy

The low frequency energy (lower than 1 Hz) of EOG well reflects vigilance transition. On the horizontal EOG, the PSD of LF/HF is similar to SEM proportion in a sense. However,

the characteristic of downtrend in Stage 2 sleep has not been observed. In this aspect, low frequency energy on horizontal EOG is better than SEM proportion. On the vertical EOG, the low frequency energy presents a weaker but still satisfying connection to vigilance.

Blink Duration

Among all the blink features linked to time parameters, closing time, opening time, and blink interval are not found having any evident correlation with vigilance. Blink duration is a useful feature but its contribution seems entirely attributed to delay-of-reopening. This explains why delay time ratio does not work – when delay-of-reopening and blink duration have approximate meaning, their ratio makes no sense.

Saccades

The effect of saccades is not so prominent as SEMs and blinks. Since some works still claim the validity of saccades, this result may need to be verified again by using a detection method which can precisely locate the beginning and end point of saccades.

Amplitude

In general, the amplitude of SEMs and saccades shows no strong connection to vigilance. The amplitude of blinks is better but not good enough.

2.5.2 Online Attempt

In order to reduce the complexity of calculation and programming, off-line analysis is used to investigate the correlation between EOG features and vigilance in this work. However, the final goal is to realize a practical EOG based system, so it is important to assure that the method is compatible to on-line processing. The possibility of on-line processing for each step is discussed as follows.

Eye Movement Detection and Feature Extraction

For both eye movement detection and feature extraction process, the on-line environment is simulated by using 8-second sliding window. Long time window will cause delay to the result, while short time window may be unable to involve a complete long blink, so 8 seconds is a proper choice. For SEM detection a longer time window of 30 seconds is used (only the last 8 seconds is used for feature extraction). This is because the wavelet transform need enough signal length or the high scale wavelet coefficients will become meaningless.

LDS

LDS is completely compatible to on-line processing. In that case it uses only the past data but not future data to estimate the current state. The on-line result will slightly differ from off-line result, mainly expressed by less smoothing in the output. In the aspect of correlation, it will cause about 5% lost. However, this does not outweigh the merit of no time delay.

Result Decision

The off-line normalization is easy so there is no trouble on calculating the final correlation. But for on-line processing, it is difficult to implement normalization on a short time window without any prior knowledge to feature ranges. Although this study has shown that some features are rather reliable and highly related to vigilance, the individual difference forbids applying the same standard to different subjects. Even for the same subject, due to his/her various physiological status the result may vary. Individual calibration can partially solve this problem but it is trivial to carry out in the application. The next target is to find robust features which have relatively less individual difference in future research.

2.5.3 Electrode replacement

Figure 2.11 verifies that independent HEO and VEO can be separated from the signals detected at the forehead. When the actual HEO and VEO are already known, the correlation coefficients between ICA separated components and actual HEO/VEO during all the 14 kinds of movements were calculated. If for each trial the channel with the highest correlation coefficient was selected as the representation of HEO or VEO, their correlation value with HEO and VEO is 0.9036 and 0.7862 on average. On the other hand, if the representation channels are selected by SVM. The average correlation values of all the 5 subjects and all the data sessions are 0.8457 for HEO, and 0.7697 for VEO.

The most significant factor causing the decline of correlation values is the EMG interference in some weakly eye related movements like teeth clenching and scowling. In such movements, EMG artifacts still exist in the ICA separated signals. This interferes with the SVM classifier and makes it less likely to distinguish EOG from EMG. However, in such a case, even EOG detected in the traditional way may be contaminated by EMG artifacts, which means the correlation value between original EOG and ICA separated EOG may not be a reasonable measure of performance. In normal vigilance experiments, subjects do not usually make intense facial muscle movements which continuously generate strong EMG, so it can be expected that the actual performance would be higher.

Chapter 3

Electrooculogram and Electroencephalogram based Robot Control

3.1 Introduction

Brain-machine interface (BMI), also called brain-computer interface (BCI), is a communication system that allows direct connection between a human brain and a computer or other external device [40]. It is mainly designed for assisting people with severe motor disabilities, helping them re-establish communicative and environmental control abilities [41]. It may also apply to able-bodied people in some special situations where the other means of communication become unavailable or occupied. There are a variety of noninvasive techniques measuring brain activities: functional magnetic resonance imaging (fMRI) [42], near-infrared spectroscopy (NIRS) [43, 44], magnetoencephalography (MEG) [45], electroencephalography (EEG) [46], and so on. Among them, EEG has high time resolution, less environmental limits, and requires relatively inexpensive equipments [46]. It has been largely used in both clinical and research applications.

The EEG-based BCI modalities can be categorized into four different types: event-related desynchronization/synchronization (ERD/ERS) [47,48], steady state visual evoked potentials (SSVEP) [49], event-related potentials (ERP) [50–52], and slow cortical potentials (SCP) [53]. Among these, ERP and SSVEP-based BCIs are more practical than others because they support large numbers of output commands, and need little training time. ERPs are brain responses to specific cognitive tasks. P300 is one of the most often used ERP components. It is a positive deflection in EEG over parietal and occipital cortex, occurring approximately 300 ms after a rare but task-relevant stimulus [54]. P300-based BCI has relatively robust performance for target detection. Although, its information

transfer rate (ITR) is at a medium level, but unlike SSVEP-based BCI, it does not cause some subjects feeling annoyed or fatigued by the flickering stimuli [55, 56]. One of its representative applications is the P300 speller which is used for inputting characters [57].

Electrooculography (EOG) measures voltage fluctuations resulted from eye movements. EOG signals are generated by eye saccades or pursuit movements as well as blinks. EOG can be used to track the eye-gaze direction, doing similar work as an optical (video-based) eye tracker. It also contains highly recognizable information of eyelid movements such as blinks and winks. Although in EEG-based BCI, EOG signals are usually considered as a major noise source to be removed [58, 59], EOG alone can make up another kind of human machine interface. For an EOG-based system, the response speed can be considerably high, which is desirable especially for control applications. There have already been a lot of studies about designing EOG-based human-machine interfaces as described by [23, 60, 61].

However, most current techniques of EEG or EOG-based interfaces still face challenges that prevent these from being accepted by the majority in the clinical applications. For example, the main obstacle of the P300-based BCI (and almost all BCIs) is its relatively low ITR: the response time of the system is unsatisfactory for most daily tasks, and the accuracy is also not perfect. Moreover, P300 is a synchronous BCI system which receives inputs and generates outputs at specified time intervals. In other words, when users do not want to send commands, the system should be off to prevent unwanted outputs. Considering this point, additional means to make the system active/inactive are often needed. For example, SSVEP has been used in P300 BCI as a switch [62], but this approach still has some problems like causing vision tiredness.

For EOG interfaces, the main problem is that they do not adequately support large numbers of outputs. The number of reliable EOG commands is usually limited to several eyelid movement patterns. Comparing with video-based eye trackers, an EOG-based eye tracking system is light-weight but poor in accuracy, thus making it unsuitable for fine control of a graphical user interface (e.g. as a replacement of the computer mouse). Moreover, unlike BCI, EOG interface is muscle-dependant which means that it inevitably causes fatigue over a longer use.

Another problem is arises due to the needs of users. The expected users of such interfaces may have various kinds of needs (e.g. controlling a wheelchair to move around, sending messages to contact other people, and so on). Usually, a single EEG or EOG-based system can only manage one certain kind of task. It is rather difficult to have a universally robust system applicable to different situations.

To overcome the limitations and disadvantages, an extensive amount of work in hybrid BCI has been invested in recent years [63]. A hybrid BCI is typically a combination of two different types of BCI systems, or BCI and non-BCI systems (EOG, EMG, etc.). These hybrid systems allow for different subsystems to be chosen for different tasks, making them more flexible. Table 3.1 (partially from [63]) shows the recent studies about hybrid

Table 3.1: Recent studies about hybrid brain or non-brain human-machine interfaces

Literatures	Brain				Non-Brain	
	EEG		Non-EEG		EOG	EMG
	P300	SSVEP	ERD	NIRS		
[56] [64] [65] [66] [67]		✓	✓			
[68]			✓		✓	
[69] [70]			✓	✓		
[71]			✓			✓
[62] [72] [73]	✓	✓				
[74] [75] [76]	✓		✓			

brain or non-brain human-machine interfaces.

In this study, a novel EOG/EEG hybrid human-machine interface adopting eye movements and ERPs is proposed, and applied to robot control tasks. The proposed method has following novelties and features.

- The proposed hybrid HMI combines EOG and ERP interfaces. According to Table 3.1, there are no existing studies carried on EOG/ERP(P300) hybrid system.
- The EOG method used in this study can detect four kinds of eye movements: blink, wink, gaze, and frown, while other similar studies only focus on blink and gaze.
- Conventional ERP-based interfaces only utilize the P300 component. in this study the ERP paradigm uses inverted face images as stimuli which evoke not only the P300 component, but also the VPP and N170 components, and thus, is expected to give a better accuracy.
- Two different robot control experiments have been carried out to verify the proposed system. The experiments include single robot and multiple robots control.

The proposed hybrid interface works in two modes: an EOG mode and an EEG mode. The two subsystems are equally important and have separate functions: EOG for fast-response tasks and EEG (ERP) for menu-selection tasks. Therefore, the overall system becomes versatile and flexible. In many other studies on hybrid BCIs, one subsystem only works in assistance of the other. For example, among the SSVEP/P300 hybrid interfaces mentioned in Table 3.1, SSVEP is used as an on/off switch of P300 [62], an idle/non-idle

state detector of P300 [72], or is used to divide the character matrix of a P300 speller into several subareas to increase ITR [73]. All these implementations enhanced the original function of P300 interface but did not improve the versatility and flexibility.

The multi-ERP (P300, VPP and N170) paradigm based on stimuli of inverted faces belongs to one of the previous studies [77]. The basic idea of this hybrid HMI, as well as a detailed design for a humanoid robot control scheme, has been introduced in another pilot study [78]. So these parts are predigested, and the main contents in this paper are about the hybrid HMI details, the eye movement detection algorithm, and discussions.

3.2 Model of hybrid interface

In practice, a lot of control scenarios require multitasking. As mentioned in the introduction section, the significance of the proposed hybrid HMI is its versatility and flexibility. Combining eye movements and ERPs can potentially make full use of the advantages of both systems and help to overcome the disadvantages. By using eye movements, the system can achieve very high ITR, which compensates for the largest weakness of ERP interfaces. By using ERPs, a graphical user interface can be realized and large numbers of commands can be supported more easily. In addition, user experience like convenience is also an important feature. For using the EOG interface, repeatedly performing eye movements will easily accumulate fatigue on muscles. Although the ERP interface casts no physical burden on users, continuously watching the flashing cues on the screen can also lead to impatience and weariness, and thus decrease the system performance. With the hybrid interface, users do not need to constantly concentrate on the same operation mode, which potentially relieves users' burden, both on body and mind.

Figure 3.1 illustrates the whole model of the proposed hybrid interface. The upper part is EOG processing. This part works asynchronously, which means the system is always actively detecting the eye movements and the user can send EOG commands at any time. Input EOG signals are divided into vertical EOG and horizontal EOG to detect four kinds of eye movements: blink, frown, wink, and gaze. More specifically, the system detects double blinks, triple blinks, and frowns from the vertical EOG, and detects winks (left/right) and gazes (towards left/right) from the horizontal EOG. These are all common eye movements and can be performed easily (without great efforts) and non-voluntarily. The single blink is not involved in this model because people always spontaneously make single blinks. Generally, it is difficult to distinguish intentional blinks from spontaneous blinks. In this model, this problem is avoided by only concerning double and triple blinks but ignoring single blinks. The vertical gaze is also not involved because when compared with horizontal gazes, vertical ones are more unreliable and easy to be performed spontaneously.

The priority selector is used to determine the priority of different types of eye move-

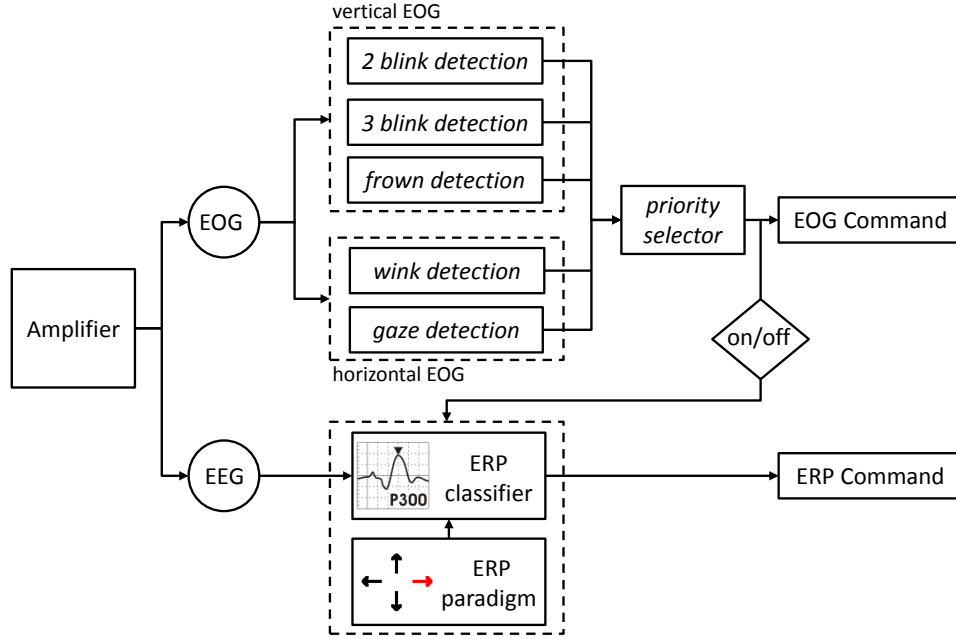


Figure 3.1: The proposed model of the hybrid HMI

ments. In this model, each kind of eye movement is detected by an individual sliding window on the corresponding EOG channel. Because all kinds of eye movements are detected in parallel, there can be a time in which two or more eye movements exist. A most common example could be: a triple blink always contains double blinks, so at a time both “triple blink” and “double blink” will be detected. Since there can only be one output, eye movements are assigned priorities. In this study it is like this: wink > frown > gaze > triple blink > double blink.

The lower part is EEG processing. Similar to a conventional P300 system, it includes an ERP paradigm and an ERP classifier. The ERP paradigm contains a dynamic graphical interface. It displays icons (e.g. arrows) on the screen, where each icon represents one ERP command. Stimuli of inverted face images are continuously flashed upon these icons in a random sequence. When the user focuses on one target icon, the stimulus flashed on that icon will evoke ERPs including P300, VPP and N170. At the same time, the ERP classifier analyzes the EEG signals of each time interval following a stimulus. The length of the time interval should be proper to contain the evoked ERPs (in this study it is 700 ms from the beginning of each stimulus). The classifier identifies which time interval is most likely to contain the evoked ERPs, and thus determines the output ERP command.

Since the ERP interface works synchronously, as mentioned in the previous section, there must be an external command to enable/disable it. An eye movement is a suitable choice for this task. The system has two modes: EOG mode and EEG mode. An eye

movement (in this study, frown) is used to switch between the two modes. In EOG mode, the ERP interface (including ERP paradigm and ERP classifier) is inactive, which means no image stimulus will be shown, and no EEG signal will be analyzed. In EEG mode, the ERP interface is active. By this design, the user can easily start and end an ERP trial any time through switching between the two modes. Although ERP interface can also have a “sleep” command to shut down itself (turning on is impossible), but it is definitely more convenient to use an eye movement to do so.

On the other hand, even in EEG mode, the system still continuously detects eye movements from EOG signals. One reason is just mentioned above that the eye movement is responsible for mode switching. For other kinds of eye movements unused in EEG mode, the system can be set as unresponsive. However, eye movements can have more roles than this in EEG mode, one example is to report the error of ERP classification. The accuracy of ERP classification varies across individuals and is hardly perfect for most people. Incorrect results lead to unwanted command being sent and executed. To avoid this, it is better to have an error report mechanism as an assurance. In the current setting, a 1 sec delay is added before sending out an ERP command. During this time if the user finds the result shown on the screen is incorrect, he or she can perform a wink immediately to cancel it out and then no command will be sent.

The specific contents of EOG and ERP commands can be determined depending on various situations and tasks. But generally, EOG commands suit tasks that require fast response, while ERP commands suit a form of selection menu. In our online robot control experiments, EOG commands are mainly used to control the robot’s move, and ERP commands are to let the robot perform some preprogrammed behaviors, or to select the control target from multiple robots.

3.3 EOG analysis

3.3.1 EOG acquisition

In this study, various eye movements from EOG signals are recognized and then applied as the commands of robot control. The EOG electrodes are placed as shown in Fig. 3.2 where electrode A is for recording vertical EOG; electrodes B and C are for recording horizontal EOG. All of the electrodes are monopolar. Other two electrodes: ground and reference, are shared with EEG recording. Unlike the commonly used dipolar placement of EOG electrodes (i.e., B-C dipolar and A-D dipolar), this placement shares the reference electrode with EEG, so that electrode D is saved.

The signals were recorded by g.USBamp with g.GAMMAbox (g.tec medical engineering, Austria). The original sampling rate of EOG was 256 Hz. It was down-sampled to 32 Hz because the proposed eye movement detection algorithm prefers smooth data.

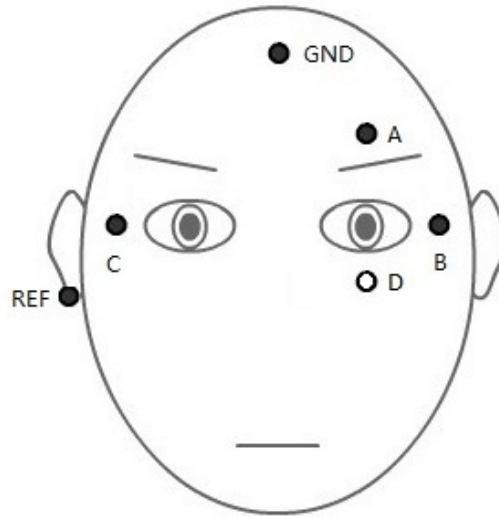


Figure 3.2: The placement of EOG electrodes. A: vertical EOG electrode; B, C: horizontal EOG electrode; REF: reference; GND: ground.

Higher sampling rate may cause undesirable fluctuations. The band-filter was chosen as 0.1-30 Hz (built-in). The 0.1 Hz lower-cutoff frequency is to eliminate the effect of baseline drift, and the 30 Hz upper-cutoff frequency is to remove high-frequency noise. In addition, the choice of lower-cutoff frequency influences the signal shape, which may also affect the optimal values of algorithm parameters. The algorithm used in this study was based on 0.1 Hz lower-cutoff frequency.

3.3.2 Eye movement detection algorithm

In this study, four kinds of common eye movements: multi-blink (blink twice or three times quickly), frown, wink (one-sided blink), and gaze (horizontal, towards left or right) are detected. All these eye movements have recognizable waveform shapes. While threshold method is commonly used for blink detection [1, 79], in this study, a simple and effective multi-threshold algorithm applicable to all these four kinds of eye movements is designed. This algorithm is similar to the blink detection algorithm which has been introduced in the previous chapter. But here the algorithm is to detect four kinds of eye movements simultaneously so some modifications have been applied.

Figure 3.3 illustrates the waveforms of the four kinds of eye movements, from which it can be observed that the waveforms of different eye movements have common characteristic. All of them consist of a positive peak in original EOGs, and a positive peak

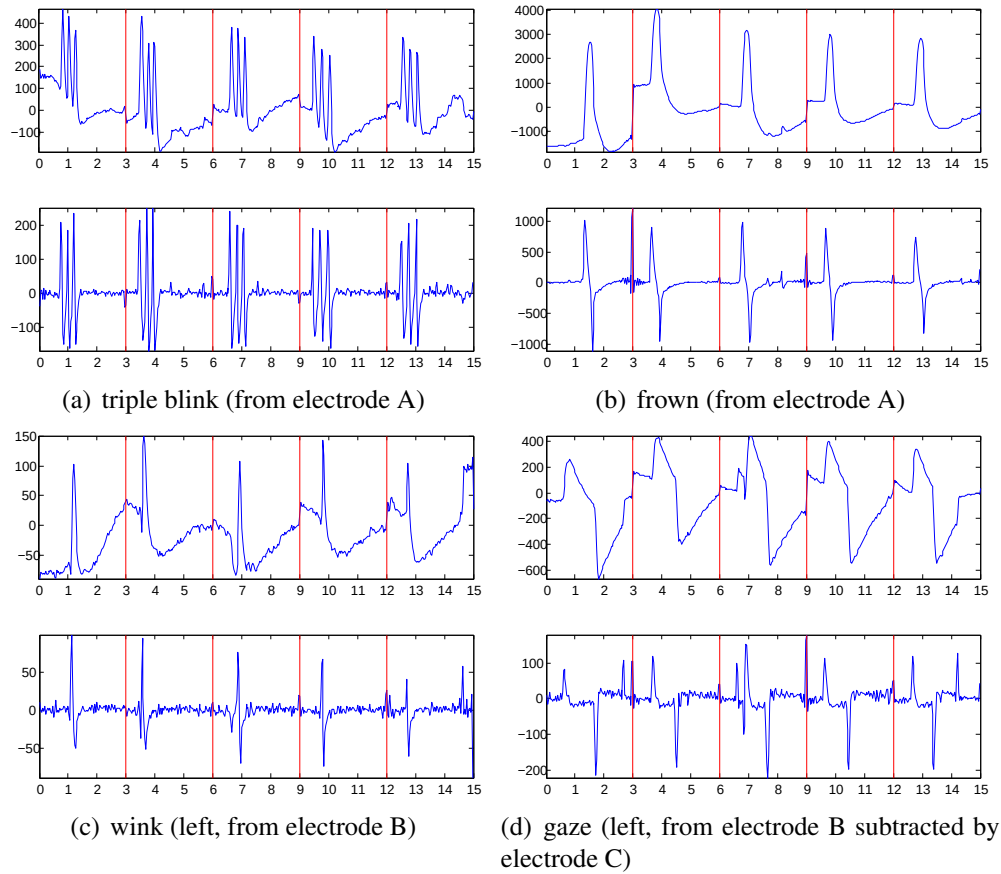


Figure 3.3: Typical waveforms of (a) blinks, (b) frown, (c) wink, and (d) gaze. For each, the upper plot is the original waveform whose y-axis is voltage (μV) and x-axis is time (s), and the lower plot is the differentiated result (first-order difference) of the upper one. Here each 3 second duration is considered as a single trial. Intervals between trials are omitted.

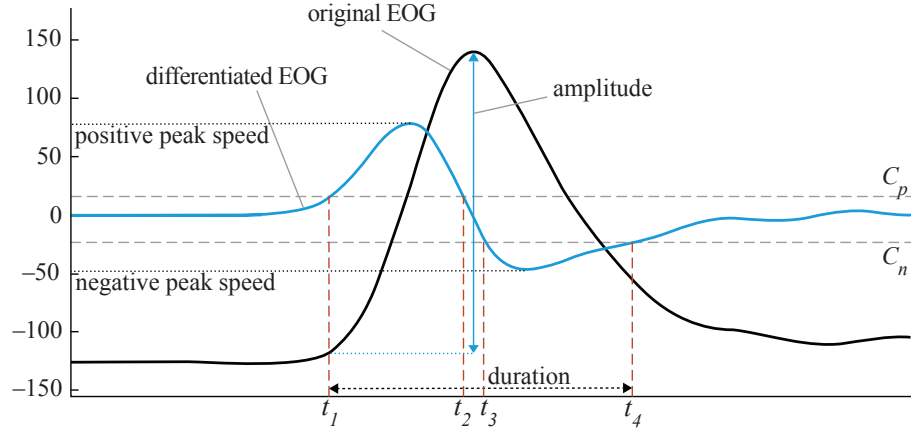


Figure 3.4: The waveform of a blink in vertical EOG and its differentiated shape [1]. The x-axis is time and y-axis is voltage (μV). The derivative is approximated by the first-order difference $\frac{x(t)-x(t-\Delta t)}{\Delta t}$ in our program.

followed by a negative peak in differentiated signals. This feature can be utilized to locate eye movements. Next, it can also be observed from Fig. 3.3 that different types of eye movements have different speed, amplitude, and time duration features. A summary can be made as shown in Table 3.2. According to Table 3.2, it is possible to distinguish different types of eye movements by the same algorithm. The proposed multi-threshold algorithm takes the speed, amplitude, and duration as three kinds of features to determine an eye movement event.

To explain the rough idea of the algorithm, see Fig. 3.4 as an example. The figure

Table 3.2: A summary of different eye movements

	blink	frown	wink	gaze
speed	M	L	S	S
amplitude	M	L	S	M
duration	S	L	M	M*
EOG channel	V	V	H	H**

L: large, M: medium, S: small, V: vertical, H: horizontal

* only including the eyeballs moving forth

** only considering horizontal gazes

shows a blink waveform and its differentiated shape. To locate the blink, it needs to find a positive peak (the interval between t_1 and t_2) and a negative peak (the interval between t_3 and t_4) on the differentiated signal. Then the blink event is just the interval between t_1 and t_4 . To further verify if it is a blink, three kinds of thresholds (speed, amplitude, and duration) were applied to identify its eligibility. These threshold values are determined by a calibration process before each experiment.

Assume that $f(t)$ is data of original EOG corresponding to a channel, and $f'(t)$ is the data of time differential of the original EOG signal. For detecting different kinds of eye movements $f(t)$ is also from different channels. As Fig. 3.2 shows the electrode positions, for blink and frown detection, $f(t)$ is from channel A; for wink detection, $f(t)$ is channel B (left) or C (right); for gaze detection, $f(x)$ is channel B–C (left) or C–B (right). The detailed steps of the algorithm are as follows.

Locate all the peaks

Assume there are following thresholds: S_{min} , S_{max} , A_{min} , A_{max} , D_{min} , and D_{max} , which represent minimal and maximal thresholds of speed (S), amplitude (A), and duration (D), respectively.

The first step is to find all the points t_1 , t_2 , t_3 , and t_4 which satisfy the following inequalities:

$$\begin{aligned} f'(t_1 - \Delta t) &< C_p, \quad f'(t_1) \geq C_p, \quad f'(t_1 + \Delta t) \geq C_p, \\ f'(t_2 - \Delta t) &\geq C_p, \quad f'(t_2) \geq C_p, \quad f'(t_2 + \Delta t) < C_p, \\ f'(t_3 - \Delta t) &> C_n, \quad f'(t_3) \leq C_n, \quad f'(t_3 + \Delta t) \leq C_n, \\ f'(t_4 - \Delta t) &\leq C_n, \quad f'(t_4) \leq C_n, \quad f'(t_4 + \Delta t) > C_n, \end{aligned} \tag{3.1}$$

where in the algorithm, by default it has $C_p = 10$, $C_n = S_{min}$. This step locates all the eligible peaks, where $[t_1, t_2]$ has a positive peak whose initial and final value is equal to C_p , and $[t_3, t_4]$ has a negative peak whose initial and final value is equal to C_n . Here C_p is given as a smaller number i.e. 10, because the more C_p is close to zero, the more accurate is t_1 , which is also the start point of eye movements. In the same way, if C_n is close to zero, the end point of eye movements t_4 will be accurate. However, in an online running environment, early detection is much more preferred than accurately locating the end point. That is why C_n is set equal to S_{min} , but not some value close to zero.

For clarity, the results are sorted to multiple pairs of $\{t_1, t_2\}$ (for positive peaks) and $\{t_3, t_4\}$ (for negative peaks).

Apply speed thresholds (S_{min} , S_{max})

The second step is to pick out all the eligible peaks that satisfy the following inequalities:

$$\begin{aligned} \max_{t \in [t_1, t_2]} f'(t) &\geq S_{max}, \\ \min_{t \in [t_3, t_4]} f'(t) &\leq S_{min}. \end{aligned} \quad (3.2)$$

This step is to make sure that the maximal f' value of each positive peak exceeds the threshold S_{max} , and the minimal f' value of each negative peak is lower than the threshold S_{min} . Note that if in the previous step $C_n = S_{min}$, here the second equation will be automatically satisfied.

The result comprises of some pairs of $\{t_1, t_2\}$ (positive peaks) and $\{t_3, t_4\}$ (negative peaks). Then, adjacent $\{t_1, t_2\}$ and $\{t_3, t_4\}$ are grouped together to form a complete eye movement candidate $\{t_1, t_2, t_3, t_4\}$.

Apply amplitude threshold (A_{min} , A_{max})

The third step is to pick out all the eligible eye movement events that satisfy the following inequality:

$$A_{min} \leq \max_{t \in [t_1, t_4]} f'(t) - f(t_1) \leq A_{max}, \quad (3.3)$$

where $\max_{t \in [t_1, t_4]} \{f'(t)\} - f(t_1)$ means the maximal voltage in an eye movement subtracted by its initial voltage, which leads to the amplitude.

Apply duration threshold (D_{min} , D_{max})

The last step is to pick out all the eligible eye movement events that satisfy the following inequality:

$$D_{min} \leq t_4 - t_1 \leq D_{max}. \quad (3.4)$$

Special cases for wink and gaze

As mentioned above, this algorithm is suitable for all four kinds of eye movements as long as the threshold values are properly set. But still, there are a few differences in processing wink and gaze.

First, wink and gaze have some similarities. Only using thresholds cannot separate them perfectly, so there is an additional step to distinguish them from each other. For gaze (left/right), the signals of channel B and C are almost inverse, while in the case of wink (left/right), the signals from these two channels are more likely to have the same trend. For example, a left wink leads to a large positive peak in channel B, and a small positive peak

or no peak in channel C. Therefore, it is able to compare the linear correlation between channel B and C, to determine if the eye movement event is a wink or a gaze. The criterion is:

$$\begin{aligned}\rho(f_B(T_1), f_C(T_1)) &> -0.8 \rightarrow \text{wink}, T_1 = [t_1, t_4] \\ \rho(f_B(T_2), f_C(T_2)) &< -0.8 \rightarrow \text{gaze}, T_2 = [t_1, t_2],\end{aligned}\tag{3.5}$$

where $\rho(X, Y)$ is Pearson's linear correlation coefficient ranging from -1 to +1. Here +1 means total positive correlation, 0 means no correlation, and -1 means total negative correlation.

The above criterion uses $\{t_1, t_2\}$ to determine a gaze instead of $\{t_1, t_4\}$. The reason is explained as follows. For a gaze, times of the positive peak and the negative peak in a differentiated EOG are highly separated (see Fig. 3.3), as the process of the eyes moving to one side and then moving back is much slower compared to other eye movements. As waiting for the occurrence of the negative peak will cause too much delay, for gaze detection the negative peak is completely ignored. So to detect gaze, for step (1) and (2), remove all the procedures related to t_3 and t_4 , and for other steps, replace t_4 by t_2 .

3.3.3 EOG calibration and thresholds calculation

A calibration process is needed to determine the specific threshold values. The time duration of eye movements is stable and can be fixed, but speed and amplitude values change in every experiment because the subject conditions and the electrode positions are inconsistent. Before a subject begins the real-time experiment, he or she is asked to go through a calibration process first. During the calibration, on the screen it shows the texts of “ready (1 s)→movement name (3 s)→relax (1 s)” repeatedly, where “movement name” includes “frown”, “triple blink”, “left wink”, “right wink”, “gaze left”, and “gaze right”, and is also displayed in this sequence. For each movement the cue is repeated for ten times.

In the calibration process, all the eye movements are time-locked so they can be easily detected even without thresholds. For frown, wink, and gaze, just simply find the peak amplitude in the middle 3 seconds of each trial, and define a short interval T (2 s for frown and wink, 1 s for gaze) centered on the peak value to represent the eye movements. For triple blink, since there are three individual peaks, predefined threshold values are used ($S_{max} = 10$, $S_{min} = -10$, $A_{min} = 150$) to locate all the three blink events. With these eye movement samples, positive peak speed S_p , negative peak speed S_n , and amplitude A for each eye movement can be calculated as

$$\begin{aligned}S_p &= \max_{t \in T} f'(t), \\ S_n &= \min_{t \in T} f'(t), \\ A &= \max_{t \in T} f(t) - f(t_0),\end{aligned}\tag{3.6}$$

where t_0 is the first time point of T . Then the thresholds are calculated as in Table 3.3. Thresholds, which are not mentioned in this table, are not necessarily needed for experiments.

3.4 EEG analysis

3.4.1 EEG acquisition

In this study, 8 electrodes were used to record EEG: Fz, Cz, P7, P3, Pz, P4, P8, and Oz (Fig. 3.5). The device is the same as EOG acquisition. The electrodes of ground and reference are shared with EOG electrodes (placed in forehead and ear lobe). These positions cover the areas where N170, VPP, and P300 occur.

The original sampling rate of EEG was 256 Hz and was down-sampled to 64 Hz. This was not specifically designed, but just for reducing redundant information that might result in over-fitting. The filter was the same which was used in EOG.

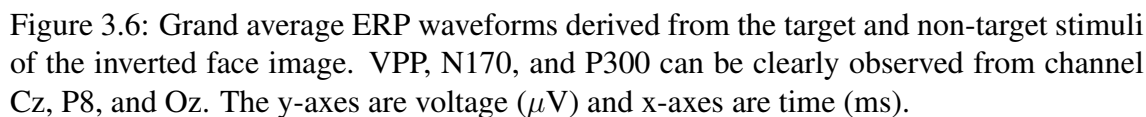
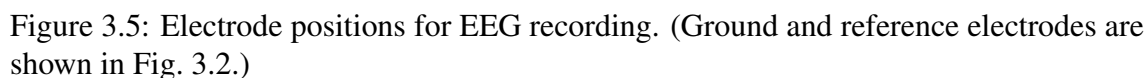
3.4.2 ERP paradigm

This study adopts a more advanced ERP paradigm which combines oddball presentation and inverted face perception. This paradigm mainly exploits three ERP components, namely VPP, N170 and P300, instead of only P300. According to a previous work [77],

Table 3.3: Thresholds for eye movement detection

Th.	Speed	Amplitude	Duration (s)
Blink	$S_{max} = 0.5 \min S_p^*$	$A_{min} = 0.8 \min A$	$D_{max} = 0.5$
	$S_{min} = 0.5 \max S_n$		
Frown	$S_{max} = 0.4 \min S_p$	$A_{min} = 0.8 \min A$	$D_{min} = 0.4$
	$S_{min} = 0.5 \max S_n$		$D_{max} = 2.0$
Wink	$S_{max} = 0.5 \min S_p$	$A_{min} = 0.8 \min A$	$D_{min} = 0.1$
	$S_{min} = 0.5 \max S_n$		$D_{max} = 0.5$
Gaze	$S_{max} = 0.8 \min S_p$	$A_{min} = 0.8 \min A$	$D_{max} = 0.5$
		$A_{max} = 1.4 \min A$	

* Here $\min S_p$ means the minimal positive peak speed of all the blinks in the calibration process.



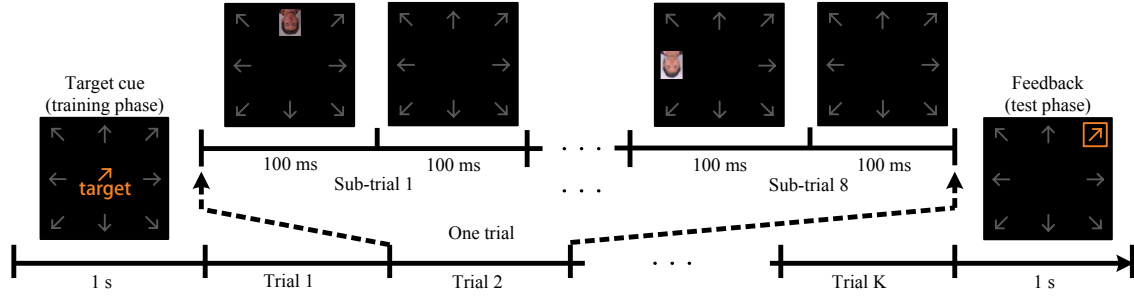


Figure 3.7: The timeline of a single run. In the training phase, a target cue was provided, and the number of trials K was 5. In the test phase, the number of trials K was 2 and feedback was provided. Each trial consisted of eight sub-trials in each of which one stimulus was randomly presented in one of the eight directions for 100 ms with an inter-stimulus interval of 100 ms.

it can significantly improve the target detection performance in contrast to the stimulus intensification pattern used in the conventional P300-based system.

Among the ERP components, N170 and VPP are evoked by the configural processing of facial image, and P300 is evoked by oddball event. Typical ERP waveforms are illustrated in Fig. 3.6.

The ERP interface used in this study has 8 arrow icons placed in 8 directions of screen (N, W, S, E, NW, NE, SW, SE). For one trial, stimuli (invert facial images) were displayed upon each arrow icon once, in random order. One stimulus was presented for 100 ms. After another 100 ms interval, the next stimulus was displayed. So one trial has $200 \text{ ms} \times 8 = 1.6 \text{ s}$.

It is possible to classify which icon the subject is focusing on by only analyzing a single trial, but to improve the classification accuracy, usually more than one trial are used to determine one output. Such group of trails is called a run. The training phase contains eight runs. Each run corresponds to one of the eight arrow icons as the target, and is of 5 successive trials. At the beginning of each run, there will be a 1 s cue instructing the subject which target he or she should focus on.

After training, subjects will go through a test phase which also contains eight runs. In the test phase each run only contains 2 trials so as to speed up the process and increase the difficulty. The test phase does not have cues at the beginning of each run; instead, the classification result (one of the 8 arrows) will be highlighted as a feedback to subjects after each run. With the feedbacks, subjects can be aware of their performance.

Figure 3.7 illustrates a detailed timeline of a single run.

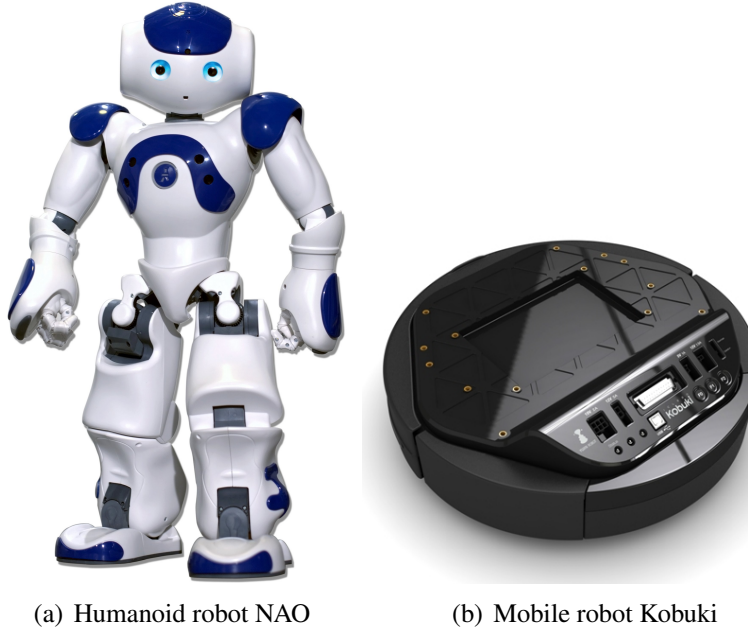


Figure 3.8: Experimental robots.

3.4.3 ERP classification

Linear discriminant analysis (LDA) was used to classify which target the subject is focusing on. Before the online experiment, EEG data collected from the training phase were used to train the LDA classifier first.

A 700 ms data segment was extracted from the beginning of each flash stimulus and baseline corrected by a 100 ms pre-stimulus interval. A total of 320 such data segments consisting of 40 targets and 280 non-targets were derived from each subject. To reduce the feature dimensionality, each data segment was further downsampled to 15 temporal points (approximately 21 Hz). The training samples (i.e., feature vectors) were then formed by the concatenation of 15 temporal points at 8 channels from the data segments. That is, the dimensionality of feature vector is $8 \times 15 = 120$. The extracted feature vectors were then fed to train the LDA classifier for the subsequent online application.

3.5 Experiments

3.5.1 Experiment introduction

In this study, two kinds of online experiments were carried out to verify the proposed system. One was humanoid robot control, in which the experimental robot was NAO

(Aldebaran robotics, Inc., Fig. 3.8(a)). The other was multiple mobile robots control, where the experimental robots were Kobuki (Yujin Robot Co. Ltd., Fig. 3.8(b)). For simplicity, hereinafter the two experiments are called the “NAO experiment” and the “Kobuki experiment”. The experimental platform was MATLAB Simulink (The MathWorks, Inc.). The details of the two experiments are introduced in this section.

Totally 13 subjects (numbered S1~S13) were involved in the two experiments. Among them, two (S7 and S10) are female and others are male. All the subjects are able-bodied, aged from 22 to 30. Four subjects: S1, S2, S3, and S4 participated in the NAO experiment, and ten subjects: S1, S5~S13 participated in the Kobuki experiments. In addition, S1~S4 have some knowledge about BCI while others do not.

In both experiments, EOG and EEG were acquired as the descriptions of Section III and IV. For each subject, EOG was calibrated and EEG was trained before doing the online experiments. The target of the online experiments was to control robots to complete a series of tasks. Subjects needed to send different commands to robots in EOG mode and EEG mode. In EOG mode, the system was asynchronous, which means whenever the user performs an eye movement, the robot responds. The timing of sending commands was held by the user. If the user did not send any command, the robot would make no action. The eye movement detection algorithm must be robust enough to prevent any false detection which leads to undesired output. On the other hand, EEG mode was synchronous, which means the user had to follow the system’s predefined pace. In this mode, even the user did not look at the screen, the system would still analyze the EEG and output most probable results. So, when the user was idle, he needs to switch to EOG mode in order to inactivate the ERP paradigm.

In EEG mode, one output was generated after 16 stimuli were given (8 icons by 2 trials) where each stimulus consisted of 100 ms highlight and 100 ms dark, which was 3.2 s. Including the time of showing the result and any other interval of delay, the total time to generate an ERP command was about 6 s, whereas using eye movement to generate a command only took around 1 s. Considering this fact, commands requiring fast response such as moving and turning are assigned to EOG mode; other functions which do not require fast response are assigned to EEG mode. In addition, the left/right eye movements respectively correspond to leftward/rightward movements of robots. The detailed commands for both experiments are listed in Table 3.4.

Note that in the NAO experiment, when a wrong ERP command was generated, the subject needed to “frown twice” to cancel that command (frown once to stop the current output and switch to EOG mode, then frown again to switch back). Since this operation was too ineffective, a new command “right wink” was implemented in EEG mode to report classification error and cancel the corresponding execution. However, this function was unused in the Kobuki experiment where the ERP commands were only related to robot selection, not behavior execution.

Table 3.4: Commands for the two robot-control experiments

Mode	Command	NAO Exp.	Kobuki Exp.
EOG mode	double blink	stop	stop
	triple blink	go head	go head
	left wink	turn 90° left	keep turning left
	right wink	turn 90° right	keep turning right
	left gaze	head 90° left	—
	right gaze	head 90° right	—
	frown	mode switch	mode switch
EEG mode	icon 1	receive an object	select robot No.1
	icon 2	hand over an object	select robot No.2
	icon 3	dance	select robot No.3
	icon 4	sit down	select robot No.4
	icon 5~8	—	—
	frown	mode switch	mode switch
	right wink	—*	error report**

* unimplemented

** implemented, but unused

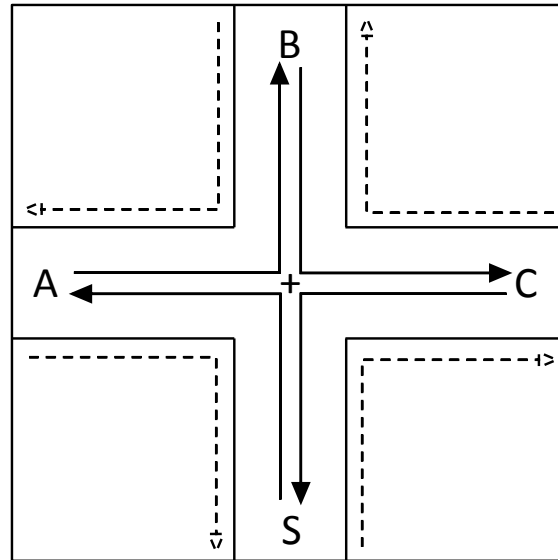


Figure 3.9: Experimental layout. The solid arrows are routes of NAO, and the dashed arrows are routes of Kobukis.

3.5.2 Humanoid robot control

NAO is a multi-functional humanoid robot, programable to perform various complex behaviors. The NAO experiment stimulates a scenario that a person controls a humanoid robot to make simple communication with other people. The detailed experiment scenario is described as follows.

- The robot starts from the point S (see Fig. 3.9), and moves to the point A by a provided route, then receives an object from a person at the point A.
- Holding the object, the robot moves to the point B, and gives the object to a person at the point B.
- Then the robot moves to the point C and performs a dance.
- Finally the robot moves back to the point S and sits down.

The whole experiment was divided into 4 sessions. These sessions consist of similar tasks. Take the first session as an example, the detailed experiment steps are:

- The robot gets ready at the point S (standing, in EOG mode).
- The robot moves ahead, until gets to the center point (+), then stops.

- The robot looks left, then looks back to center.
- The robot turns 90° left, then goes ahead.
- The robot stops at the point A, then switches to EEG mode.
- The robot performs a provided behavior by ERP command.

For other sessions, only the robot positions and the provided behaviors are different. The four sessions ($S \rightarrow A$, $A \rightarrow B$, $B \rightarrow C$, $C \rightarrow S$) were scheduled separately.

3.5.3 Mobile robots control

The purpose of the Kobuki experiment is to simulate a scenario that a person controls multiple robots to move around and gather information. In this experiment, four Kobuki robots were used and moved to destinations one by one. These robots only have basic function of mobility. For simplicity, in this experiment they were placed near each other under the direct observation of the subject. But if necessary, they can also be mounted with a laptop and sensors like a camera and a GPS, which can enable them to remotely perform complicated tasks.

The experimental layout (Fig. 3.9) was designed similar to the previous experiment. Four Kobuki robots numbered 1~4 were placed at the right side of the marks (so they were unlikely to run into each other) S, C, B, and A, respectively. The detailed experiment steps are:

- The subject switches to EEG mode, and selects the No.1 near the point S (see Fig. 3.9).
- The No.1 moves ahead, and stops when near the center point.
- The No.1 turns right, and stops when facing its destination.
- The No.1 moves ahead, and stops when reaching its destination (near C).
- Repeated until all four robots are selected and moved to their destinations.

Each subject completed the experiment without any break, and repeated the experiment again after the first completion. The time for each completion was recorded.

3.6 Results

3.6.1 Offline evaluation

For all the 13 subjects involved in online experiments, their offline data (EOG calibration and ERP training) were recorded and evaluated. Since one subject participated in two different experiments, there are totally 14 groups of data, and the results are listed as follows.

EOG

14 groups of calibration data are used for evaluation. Each group includes 6 kinds of eye movements: triple blink, frown, left wink, right wink, left gaze, and right gaze, with 10 repetitions of each. A comparison has been made between the proposed threshold method and a standard pattern matching method. The proposed method calculated the thresholds according to Table 3.3. For the pattern matching method, the implementation is as follows: 10 repetitions of eye movements are averaged, and the result is considered as a standard eye movement pattern. Then a sliding window on original signals is compared with the standard pattern, by calculating the Pearson's linear correlation coefficient ρ . If ρ is larger than a predefined threshold, it means an eye movement is detected. A proper length of the time window has been carefully chosen, as well as a proper threshold of ρ for each kind of eye movement.

Figure 3.10 shows the recall of each kind of eye movement detected by the two methods. Lower recall means higher miss rate.

ERP

Figure 3.11 depicts the classification accuracy averaged on all subjects using one to five trials. The classification accuracy is calculated by cross-validation. Since each subject took 8 ERP runs (one run includes five trials) in training process, the cross-validation took 7 runs as training data, and the rest one as test data. This process was repeated 8 times, and each time the test data was a different run.

3.6.2 NAO experiment

The performance of subjects and experiment details are shown in Table 3.5. The time listed in the table does not include any behavior performing (e.g. dance) so that each session has the same expected time cost. For comparison and evaluation, the experiment was also completed by using Choregraphe, a GUI software platform for NAO, with a

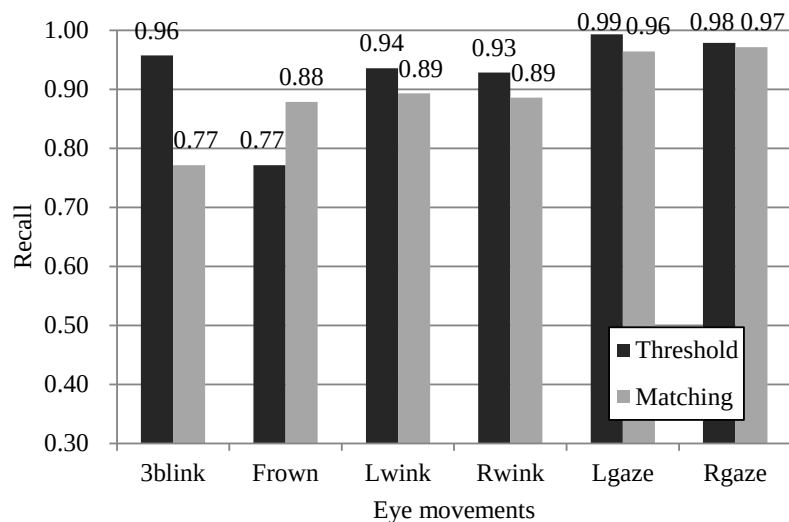


Figure 3.10: The recall of six kinds of eye movements, detected by the proposed method and a pattern matching method. One kind of eye movement has 140 repetitions.

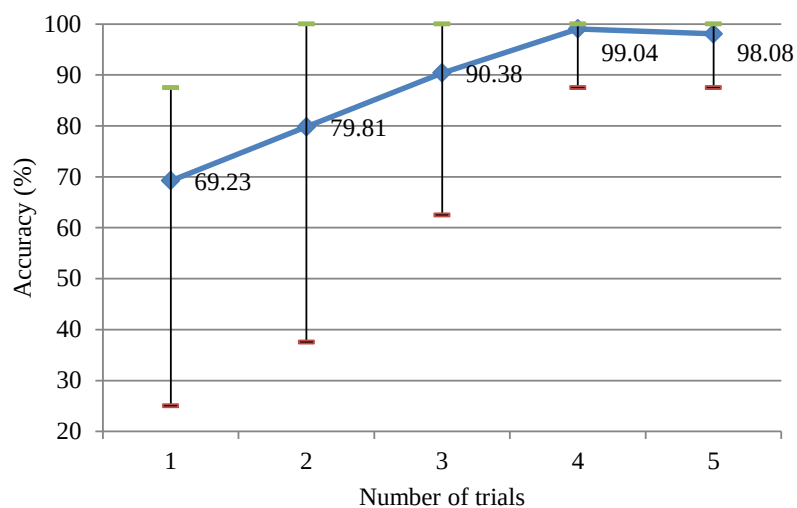


Figure 3.11: The ERP offline classification accuracy averaged on subjects using various trials. Vertical lines show maximal and minimal accuracies among all the subjects. The x-axis means how many trials were used.

Table 3.5: Results of the NAO experiment

Subject	Session	No. of commands	Time(s)	Avg. time (s)	ERP accuracy	EOG accuracy
S1	1	12	62.9	70.2	4/4	35/39
	2	9	57.0			
	3	11	58.3			
	4	11	102.6			
S2	1	10	57.2	72.6	4/9	41/45
	2	9	52.0			
	3	11	59.1			
	4	24	122.2			
S3	1	14	78.3	67.5	4/6	35/40
	2	9	57.3			
	3	10	56.2			
	4	13	78.2			
S4	1	15	62.7	54.1	4/4	33/38
	2	9	52.2			
	3	9	52.5			
	4	9	49.0			

Note: The minimum number of commands for one session is 9. If operated by hand, the average time cost is 49.1 s.

mouse. Operated by hand, the completion time of one session was about 49.1 s, which was averaged by four runs.

One session without any mistake should include nine commands (eight EOG, one ERP) which are sequentially: go ahead, stop, look left, look center, turn left, go ahead, stop, EEG mode, and behavior selection. So the minimal number of commands for one experiment is 32 EOG commands and 4 ERP commands. In Table 3.5, the results are evaluated by: i) the total number of detected commands, ii) the time cost, iii) the ERP accuracy, and iv) the EOG accuracy. The ERP and EOG accuracy is expressed as “number of effective commands/number of detected commands”. For ERP, the number of effective commands is always equal to 4, because a complete NAO experiment contains 4 ERP commands. For EOG, the effective commands include the minimum necessary commands to complete the task (which is 32), plus the commands to correct earlier mistakes. For example, if the robot received a wrong command and did “look right” instead of “look left”, there should be an additional command ordering the robot to “look center” first. Although this command does not belong to the minimum necessary commands, it is still counted as an effective command. Additionally, in Table 3.5 for each subject, the denominator of “ERP accuracy” plus that of “EOG accuracy” means the total number of commands, so it should always be equal to the sum of “No. of commands”.

The EOG accuracy of online experiment shows the number of false detections, while the miss rate (false negative) can only be referred from the offline evaluation. This is because in the online experiment the EOG detection process was asynchronous: the subject could perform any eye movement at any time. So it is difficult to tell which eye movements were missed. However, the miss rate can also be estimated from the time cost. If the time cost was very long in spite of high EOG accuracy, it is very likely that there is relatively a larger number of misses.

3.6.3 Kobuki experiment

Table 3.6 shows the performance and details about the Kobuki experiment. Four robots were used in the experiment, and each was given seven commands: EEG mode, robot selection, EOG mode, go ahead, turn right, go ahead, and stop. Therefore the minimal number of commands for one experiment should be 28 (4 ERP and 24 EOG). The evaluation is similar to the NAO experiment, except that this experiment was not timed in separated sessions, and each subject completed this experiment twice.

For this experiment, the hand-operated time was also calculated for comparison. Because Kobuki has no corresponding GUI software platform, an operation menu in Simulink was created as its GUI which enables speed adjustment, direction adjustment, and robot selection. By using this operation menu, the whole experiment could be completed in about 122.7 s

Table 3.6: Results of the Kobuki experiment

Subject	Exp.	No. of commands	Time(s)	Avg. time (s)	ERP accuracy	EOG accuracy
S1	1st	29	123.7	124.9	8/9	52/56
	2nd	36	126.1			
S5	1st	45	190.7	172.5	8/10	52/76
	2nd	41	154.2			
S6	1st	35	160.9	157.5	8/9	52/62
	2nd	36	154.1			
S7	1st	51	212.0	221.5	8/9	56/103
	2nd	62	230.9			
S8	1st	37	147.5	136.5	8/9	52/59
	2nd	31	125.5			
S9	1st	34	137.3	140.8	8/10	50/61
	2nd	37	144.3			
S10	1st	37	147.0	141.8	8/9	51/65
	2nd	37	136.6			
S11	1st	35	136.4	140.3	8/8	51/67
	2nd	40	144.1			
S12	1st	39	141.1	134.2	8/10	55/63
	2nd	34	127.2			
S13	1st	34	138.0	141.6	8/9	50/60
	2nd	35	145.1			

Note: The minimum number of commands for one session is 28. If operated by hand, the average time cost is 122.7 s.

3.7 Discussions

3.7.1 Performance

EOG

According to the results shown in Fig. 3.10, when compared to the pattern matching method, the proposed threshold method has a distinct advantage in detecting triple blinks, but remains below expectations in frown detection. Regarding wink and gaze detection, both methods have very high recall, but the proposed method performs better.

The main reason of the performance gaps between the proposed method and the pattern matching method is the duration of blink and frown. First, blinks are usually very quick, so a long blink and a short blink will have relatively large deviation. Because the pattern matching method checks the similarity between a blink candidate and the standard pattern, large deviation of blink candidates will lead to a high miss rate. For a triple blink, if the method misses at least one of the three blinks, the detection will fail. On the contrary, the threshold method does not have this problem as long as the blink candidates do not exceed the threshold.

The case of frown detection is just opposite. Generally a frown is more difficult to detect relative to other eye movements because it is usually close to a facial expression and is more affected by individual habits. But as its duration is usually very long, the deviation of frown candidates becomes small. In that case the pattern matching method will have some advantages. For the threshold method, the duration threshold was set to 2 s, so frowns longer than 2 s will have no chance to be detected. Subjects have been told not to make very long frowns because that is not effective in doing a control experiment. However, some frowns were made longer than 2 s in the calibration process.

Although the thresholds have been summarized as in Table 3.3, they need to be determined for individual subjects. A safe strategy is to manually adjust the thresholds for every subject, and in that way, the gap in frown detection can also be filled. However, only using pattern matching method to detect frowns, and using the proposed method to detect other eye movements could be a simpler solution.

ERP

The feature extraction based on P300 activity with high classification accuracies have already been discussed in many studies [80–84]. However, a P300-based BCI usually still requires several (more than five) trials average to obtain good performance. The inverted face paradigm adopted in this study used stimuli of facial images with loss of configural face information. The configural face perception has been demonstrated to evoke not only significantly stronger P300 but also VPP and N170 [52, 77, 85]. The outstanding effects of the inverted face paradigm can help the BCI to accurately decide the desired command

with only two trials or even single trial [77]. This is why the inverted face paradigm was chose instead of the conventional P300 paradigm for the system development.

In fact, many studies have suggested other components, in addition to P300, may also contribute significantly to the target detection in an ERP-based BCI [52, 77, 86, 87]. The adopted inverted face paradigm exploits a joint effect of multiple ERP components (including N170, VPP and P300) for target detection, which has been confirmed by giving very good classification accuracy while using the ordinary LDA even in a small sample size scenario [77]. Whereas other algorithms such as SVM, regularization FDA, etc. could be a better choice for classification when only a low number of samples are available, they require a regularization parameter that is typically determined by the time-consuming cross-validation. Considering a trade-off between efficiency and accuracy, the simple LDA was adopted for classification in the developed BCI system.

Online experiments

In the NAO experiment, all the four subjects got a satisfactory result (see Table 3.5), except for the 4th session of S1 and S2. Although S1 has a high EOG accuracy, but as mentioned above, the online accuracy does not show the miss rate (which is only shown in offline evaluation). According to the data, when S1 tried to make frowns in his 4th session, they were not detected. The data showed that his frowns in the 4th session became weaker than those he did in the calibration process, where the most possible reason might be tiredness. S2 failed the ERP command for 5 times successively in his 4th session, but before that his ERP was 100% accurate. It is possible that the subject was distracted or tired at that time, nevertheless, such kind of successive failures never happened on other subjects. Apart from these two sessions, the experiment was successful. Especially, S4 gave an outstanding result, whose average completion time was 54.1 s, which means only 5 s longer than operating by hand. While he was unaccustomed to the experiment in his first session, all of his session 2, 3, and 4 were completed by minimal necessary commands and without any error.

Regarding the results of the Kobuki experiment (Table 3.6), most subjects had a stable performance except S5 and S7. Among all the 10 subjects, only S1 had participated in the NAO experiment and was familiar with BCI. This may be the reason that he had a clear advantage in terms of time. His average completion time was only 2.2 s longer than the hand-operated time. S5 had poor performance in his first attempt, he failed some winks, and confused double blinks and triple blinks many times. However, in his 2nd attempt he made a clear improvement where his completion time was reduced from 190 s to 154 s. In the case of S7, she spent obviously longer time than the other subjects. The reason for this was that she had some troubles in making winks. More specifically, her intentional winks were as weak as normal blinks, so the algorithm confused these two eye movements. As a result, she could not effectively make use of the hybrid HMI.

The proposed hybrid HMI received positive feedback from most subjects. The results show that the proposed hybrid HMI can be potentially applied to able-bodied people due to its good effectiveness. However, considering the fact that a few people might not be able to perform certain kinds of eye movements, more attention is needed for irregular eye movement detection.

3.7.2 Practicality

In this study two online experiments were carried out: one using a multi-functional humanoid robot, the other using multiple mobile robots. On the basis of these two experiments, the proposed hybrid interface can also be expanded to applications requiring multiple controllable agents with multifunction. In that case, the ERP paradigm should include two ERP menus: one for target selection, and the other for behavior selection. The two ERP menus can be switched by ERP commands or eye movements.

The proposed control strategy is suitable for both direct control and remote control. Mounted with cameras, the robots are able to perform various tasks remotely, which can be applied for the case when the user has difficulty with mobility. On the other hand, the proposed method can also be used to assist the user's self-move. For example, nowadays there is a new concept called BCI smart home which is designed for assisting people with disabilities in their daily house life [88]. A smart BCI house usually includes a BCI wheelchair for moving between rooms, and some other BCI operations for using household electrical appliances like TV, lights, and so on. Compared to traditional EEG-based wheelchair which is relatively poor in mobility, using the proposed hybrid HMI, the user can move dexterously and freely, unrestricted by predefined routes. Moreover, the EOG mode and ERP mode are supposed to work well together in this application. For example, it is easy to imagine a scene such as a user moved into living room by EOG controlled wheelchair, then he chose to operate the TV from an ERP menu which listed all the operable targets in the living room. The TV turned on, and all the available channels appeared on the ERP menu. He chose a channel, made a wink to gradually turn up the volume, and winked again when he felt the volume was just good.

Compared with pure BCI or brain-based hybrid BCI such as SSVEP/P300, the proposed hybrid HMI has advantages in versatility and flexibility. Of course, since it takes eye movements as commands, those disabled persons who cannot control eye movements normally (e.g. serious facial paralysis) are unable to use this hybrid interface. Because most disabled people have preserved eye movement capabilities, the proposed hybrid HMI shows promise.

3.7.3 Extensions

As mentioned above, the hybrid EOG/EEG interface requires the user to, at least, preserve normal function of making eye movements. However, even able-bodied people may have troubles in making certain kinds of eye movements. The eye movement detection is implemented by using predefined thresholds which are only optimized for regular eye movements. Unfortunately, the proposed method still has some flaws when the eye movements are not regular, which can be seen from the result of the S7 in the Kobuki experiment. For this reason, powerful algorithms are needed to correctly recognize irregular eye movements. We have tried extracting standard eye-movement patterns from the subject's EOGs and doing online matching, but it was only advantageous in frown detection. Using classifiers can be a potential solution, except for two problems. First, EOG mode works asynchronously, so the training set of classifier will be largely biased where most data belongs to null class. Second, the training time is another important issue. Short training time affects the accuracy, while longer training time harms user experience. These difficulties need to be overcome to find a proper solution.

In aspect of ERP paradigm, the most important issue has always been to increase the information transfer rate. When the classification accuracy has been improved through using N170 and VPP together with P300, error report is another way to achieve the goal. If the classification error is reported in time, a second trial can be presented at once, or even the second probable result can be directly selected. As a result, the overall time consuming will be reduced. In this study, it has been proposed that eye movements can be used to report errors. However, there is another interesting method using error-related potential, which enables the system to automatically identify whether the subject considers the result correct or not [89, 90]. If this automatic error detection can achieve a high accuracy, the ERP paradigm will be more convenient to use.

Chapter 4

Electromyogram based Prosthesis Control

4.1 Introduction

Prosthetic limbs generally available to amputees include simple grasping mechanisms. A standard one-to-one mapping control paradigm typically includes two myoelectric sites to detect EMG signals from flexor and extensor muscles on the forearm. Users can only control a single DOF of the prosthesis with the activities of these muscles. Multiple-DOF control (switching between different DOFs) needs to be realized through incorporation of some additional commands (e.g. co-contraction). Modern prosthetic hands have multiple fingers which offer higher degrees of freedom [91]. However, electromyogram (EMG) based control algorithms do not have comparable control capability [92]. Higher DOFs increase the functional complexity of the prosthesis and, hence, decoding subjects' intent becomes difficult.

Pattern recognition (machine learning) methods can enhance prosthesis control. In myoelectric prosthesis related studies, various pattern recognition algorithms have been used. Artificial neural network [93], support vector machine [94], Bayesian network [95], the Gaussian mixture model classifier [96], and hidden Markov models [97] are some widely used algorithms. Decoding has also been performed for multiple DOF control cases, such as individual finger movement [98]. Through training, a pattern recognition method can “learn” various states such as “open hand” or “close hand” and recognize them perfectly. However, a prosthesis does not have limited discrete outputs. Between “open hand” and “close hand” there is a series of transition states. Proportional control of prostheses cares about how to recognize these transition states; however, training all of them is time consuming and unrealistic. Moreover, for a multiple-DOF prosthesis, even when each DOF is well trained, it does not mean that the user can simultaneously control

two DOFs. For example, usually a classifier recognizes single states “open” and “pronate” but cannot respond to the combination state “open and pronate” correctly. While regression methods are more appropriate for solving these problems due to their continuous nature [99] [100], they still require supervised training. The training data must involve all of the transition states, and all of them must be correctly labeled. Such training processes are also complicated and time consuming. As a result, a method which can recognize untrained transition states or combination states, with less training time, is preferred.

Neuroscience can support building models for natural movements, where a typical representation is the muscle synergy. Unlike normal pattern recognition methods which are based on statistical rules, this theory offers an understanding of the underlying processes of limb movements, so it can potentially recognize untrained states. According to the theory of muscle synergies, when a subject’s central nervous system (CNS) controls the body movements, it does not send a single command to multiple muscles individually. Instead it activates a small set of coordinated movements or a cascade of movement primitives called synergies. Each synergy is responsible for activating a larger set of muscles. The theory has been tested on frogs, mammals, primates and humans [101] [102] [103] [104]. An example of a marionette can also illustrate the basic idea: only several strings can control the complex movements of the puppet because the manipulator’s hand (equivalent to the central nervous system CNS) sends out higher order commands to activate the strings (synergies) and achieves control of multiple joints of the puppet (human limbs). Synergies and muscles possess many-to-many mappings, and the number of synergies needed is typically far less than the number of muscles. This approach of modeling limb motion can also be regarded as a dimension reduction mechanism from muscles to synergies. If the synergies can be calculated for an amputee, then it would be possible to represent the extent of a single movement, as well as the simultaneous activation of multiple movements.

Currently, typical algorithms to calculate muscle synergies are dimension reduction algorithms such as principal component analysis (PCA) [105], independent component analysis (ICA) [106], and non-negative matrix factorization (NMF) [107]. These three algorithms decompose a larger matrix into low dimension ones while applying certain restrictions, namely orthogonality, statistical independence, and non-negativity, respectively. Among them, NMF is considered close to the natural representation of synergies and outperforms PCA and rivals ICA [108]. In this work, four hand/wrist movements which are open, close, pronate, and supinate from surface EMG was decoded by using the NMF algorithm. Furthermore, this study designed an original control scheme to process the noisy output of NMF, and convert it to an eligible control signal.

Ajiboye and Weir [109] previously used the NMF algorithm to classify postures. They used 11 postures as training data to predict 33 postures. They have validated the usefulness of the NMF in prediction but their approach was discrete classification and did not show the strength of a synergy based method. Jiang *et al.* [110] presented an offline

decoding result of 4-direction wrist movements (flexion/extension and radial/ulnar deviation) plus pronation/supination, by using a semi-supervised NMF algorithm. They also examined the influence of muscle crosstalk and force transfer by comparing several methods. Choi and Kim [111] proposed a similar approach for 4-direction wrist movements with a complete supervised NMF algorithm and carried out an online experiment with visual feedback. Jiang and Choi validated that the NMF algorithm can effectively decode the 4-direction wrist movements, but a significant performance loss was reported when dealing with pronate/supinate movement [110]. Furthermore, they did not experiment with open and closed hands. Common prosthetic applications require opening and closing of the hand, as well as wrist pronation and supination. Therefore, this study focuses on opening and closing of the hand and the pronation/supination of the wrist. Compared with 4-direction wrist movements, these movements involve more muscles. If more muscles located closely are activated, the crosstalk that contaminates the surface EMG signals will be much more significant and the movements are expected to be more difficult to decode [110]. To improve the performance, an offline experiment for feature selection was first carried out, to test which features are best for calculating muscle synergies.

Another objective of this study is to design a control algorithm for the prosthetic hand. Although the synergy based paradigm offers continuous output, this output is too variable to be directly interpreted as a control signal of the prosthesis, and thus needs post-processing. Typically used filters are incapable of suppressing this fluctuation satisfactorily. Since the control of a prosthesis is a delicate process and users' experiences can be hampered by undesired output, it is important to generate smooth control signals. Unfortunately, few previous studies include this variety of suggested post-processing. Their results are generally continuous but fluctuating, or are discretized approximations of the continuous output [100]. In this work, a control scheme is designed to achieve direct control of the prosthesis. Online experiments are made to validate its reliability.

4.2 Methods

4.2.1 Synergy model and NMF

When a person tries to perform a movement, his CNS continuously sends commands to activate muscle synergies. Suppose the time-varying command sent by CNS is $\mathbf{H}(t) = [h_1(t), h_2(t), \dots, h_K(t)]^T$, where $h_i(t)$ refers to the expected activation level on the i th synergy at time t . Here K is the total number of synergies involved in movement-related tasks. The actual value of K is unknown, but for simplification it can be assumed that one simple movement is represented by one synergy [110] [111]. This assumption becomes inappropriate when a large variety of movements are involved. In that case, a reduced set of synergies should be found [109], but this study does not go that far

but focuses on a control paradigm. In this study, the task includes 4 basic movements (open/close/pronate/supinate), so the number of synergies is $K = 4$, where each synergy corresponds to one of these movements.

The muscles' activation levels $\mathbf{V}(t)$ can be expressed by $\mathbf{H}(t)$ multiplying a synergy matrix $\mathbf{W}$.

$$\begin{aligned}\mathbf{V}(t)_{M \times 1} &= \mathbf{W}_{M \times K} \times \mathbf{H}(t)_{K \times 1} \\ &= \begin{bmatrix} W_1 & W_2 & \cdots & W_K \end{bmatrix} \times \begin{bmatrix} h_1(t) \\ h_2(t) \\ \vdots \\ h_K(t) \end{bmatrix} \\ &= \sum_{i=1}^K W_i h_i(t),\end{aligned}\tag{4.1}$$

where $W_i = [w_{1i}, w_{2i}, \dots, w_{Mi}]^T$. $W_i h_i(t)$ is the effect of the i th synergy to all the M muscles. For a single muscle j , its activation level $V_j(t)$ is equal to the element of the j th row of $\sum_i^K W_i h_i(t)$.

From this synergy based model, the equation

$$\mathbf{V} = \mathbf{W} \times \mathbf{H},\tag{4.2}$$

is expected to be solved because the values of $\mathbf{H}$ contain the details of each movement. However, $\mathbf{V}$ (unmeasurable) is neither equal to nor linearly correlated to the surface EMG signal $\mathbf{Z}$ (which can be easily measured). In this case, a common solution is to assume that $\mathbf{V}$ has an approximately linear correlation with $f(\mathbf{Z})$, which is some feature extracted from $\mathbf{Z}$ [109] [110] [111]. Then the Eq. (4.2) can be treated as

$$f(\mathbf{Z}) \approx \mathbf{W} \times \mathbf{H}.\tag{4.3}$$

The NMF algorithm decomposes the non-negative matrix $f(\mathbf{Z})_{N \times O}$ into non-negative factors $\mathbf{W}_{N \times K}$ and $\mathbf{H}_{K \times O}$ (O is the number of observations, equal to the length of the time-series data of EMG). The result is not an exact factorization but a lower-rank approximation to the original matrix. NMF seeks to minimize the cost function expressed as

$$\delta = \frac{1}{2} \|f(\mathbf{Z}) - \mathbf{W} \times \mathbf{H}\|_F^2,\tag{4.4}$$

with the multiplicative updating process

$$\begin{aligned}\mathbf{H}_{n+1} &= \mathbf{H}_n * \frac{\mathbf{W}_n^T \times f(\mathbf{Z})}{\mathbf{W}_n^T \times \mathbf{W}_n \times \mathbf{H}_n + \epsilon} \\ \mathbf{W}_{n+1} &= \mathbf{W}_n * \frac{f(\mathbf{Z}) \times \mathbf{H}_{n+1}^T}{\mathbf{W}_n \times \mathbf{H}_{n+1} \times \mathbf{H}_{n+1}^T + \epsilon},\end{aligned}\tag{4.5}$$

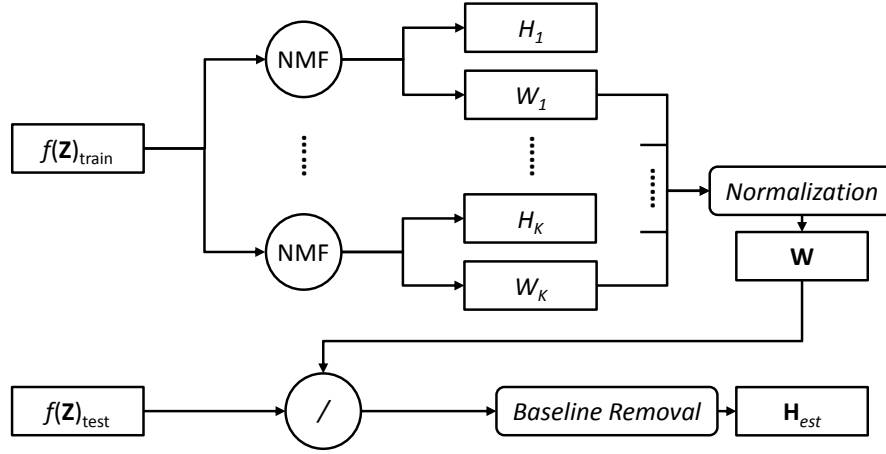


Figure 4.1: From the training data, $W_i (i = 1, 2, \dots, K)$ is DOF-wise calculated by individual NMF algorithms, and thus $\mathbf{W}$ is obtained. For new testing data $f(\mathbf{Z})_{test}$, the estimated synergy activation is $\mathbf{H}_{est} = \mathbf{W}^{-1} f(\mathbf{Z})_{test}$.

where in Eq. (4.4) Frobenius norm $\|\mathbf{A}\|_F = \sqrt{\sum_i \sum_j |a_{ij}|^2}$, and in Eq. (4.5) ϵ is a very small number [112].

The NMF algorithm is naturally unsupervised. However, a complete unsupervised algorithm is not powerful enough to solve the problem. Thus Jiang *et al.* [110] modified it to create a semi-supervised version by calculating W_i “DOF-wisely”, which means at any time only one W_i is calculated, but not the whole $\mathbf{W}$. Choi and Kim [111] chose to do this in a supervised way by fixing the initial $\mathbf{H}$ and not updating it during the NMF iterations. In this study, the DOF-wise method is chosen because it is unnecessary to label the training data.

The detailed usage of the NMF algorithm in our application is illustrated in Fig. 4.1. For the output $\mathbf{H}_{est}$, it is preferable that the corresponding H value should be equal to 1 if a synergy is fully activated and equal to 0 if a synergy is inactivated. This can be achieved by using normalization and baseline removal. Although for $\mathbf{H}_{est}$, the full activation values and the baseline values of synergies are unknown, they can be approximately estimated from $\mathbf{H}_{train}$ which is equal to $\mathbf{W}^{-1} f(\mathbf{Z})_{train}$. In the normalization step, $\mathbf{W}$ is scaled so that each W_i is multiplied by $\overline{H_i(t_{act})}$ (the average value over the “fully activated period” of the i th synergy of $\mathbf{H}_{train}$). This step ensures that $H_i(t)$ (from $\mathbf{H}_{train}$) is approximately equal to 1 if the i th synergy is fully activated at time t . As a result, $\mathbf{H}_{est}$ is also normalized. After normalization, the baseline removal step first records the baseline of each synergy which is defined as $\overline{H_i(t_{rest})}$ (the average value over the “rest period” of the i th synergy of $\mathbf{H}_{train}$). Because $\mathbf{H}_{est}$ is expected to have similar baseline values with $\mathbf{H}_{train}$, subtracting $\overline{H_i(t_{rest})}$ from $\mathbf{H}_{est}$ can remove the baselines.

Table 4.1: EMG Features

Feature	Description
MAV	Mean absolute value
MSV	Mean square value, equivalent to variance
RMS	Root Mean Square, equivalent to standard variance
WL	Waveform Length, calculated by $\sum_{k=1}^{N-1} x_{k+1} - x_k $
ZC	Zero Crossings, the number of times the waveform crosses zero
SS	Slope Sign Changes, the number of times the slope of the waveform changes its sign
Log	Log-Detector, calculated by $\exp(\frac{1}{N} \sum_{k=1}^N \log x_k)$

4.2.2 Feature Selection

Different EMG features have been taken as input to the NMF algorithm, such as root mean square (RMS) [109], mean square value (MSV) [110], and mean absolute value (MAV) [111]. Among these studies, only [110] states that MSV could be considered linearly correlated to muscle activation level $\mathbf{V}$. For Eq. (4.3) to hold, it is important to find proper features which have an approximately linear correlation with $\mathbf{V}$. On the other hand, if too many features are adopted, the decomposed matrices of NMF may also deteriorate because irrelevant input can convey large influence to the output. Jiang's research also mentioned that fewer EMG channels yield better performance when the muscle crosstalk is dominant [110]. This is why a feature selection process is needed to carefully choose the input of the NMF algorithm.

The NMF algorithm requires non-negative input. Based on this prerequisite, candidates of time-domain features are chosen as in Table 4.1. Most of them are frequently suggested features in pattern recognition studies [113] [114] [92]. Each feature has eight dimensions, corresponding to eight EMG channels. On each channel, all of the features are extracted by using a sliding window with an overlap. The lengths of the window and the overlap are mentioned in Table 4.2 afterward. In training phases, each feature is divided by its maximum value so that it is normalized to a range of $[0, 1]$. In test phases, the same values are used to approximately normalize the features.

4.2.3 Control Scheme

As illustrated in Fig. 4.1, after the NMF algorithm calculates the decomposed matrices $\mathbf{W}$ and $\mathbf{H}$, for any new coming data $f(\mathbf{Z})_{test}$, the estimated synergy activation is $\mathbf{H}_{est} =$

$\mathbf{W}^{-1}f(\mathbf{Z})_{test}$. This output $\mathbf{H}_{est}$ contains information on the time-varying activations of synergies, but with too many disturbing fluctuations. A control scheme is necessary to smooth the $\mathbf{H}_{est}$ and to qualify it as the direct control signal of the prosthesis. Until now, people have not paid much attention to this problem in laboratory studies, but it is inevitable for an integrated practical system. The traditional idea for smoothing is applying a low-pass filter but according to experiments, a low-pass filter is not enough to suppress the fluctuations and help the subject maintain stable control signals. To prove this, a control scheme for the prosthetic application is customized, and a comparison with a low-pass filter is made in this study.

Considering the characteristics of the prosthetic application, the following features of the control signal $\mathbf{C}$ are desired.

- The control signal $\mathbf{C} = [C_1, C_2, C_3, C_4]^T$ where C_1 , C_2 , C_3 , and C_4 are control signals for open, close, pronate, and supinate motions, respectively.
- The range of $\mathbf{C}$ should be $C_i \in [0, 1]$, where 0 stands for an inactivated synergy (movement) and 1 for a fully activated one.
- $\mathbf{C}$ should respond to an apparent increase or decrease in $\mathbf{H}_{est}$ with little delay.
- $\mathbf{C}$ should be stable, reluctant to change its value if the user is keeping his hand still.
- Open and close, pronate and supinate cannot be simultaneously activated, which means at the same time, only one of C_1 and C_2 , and one of C_3 and C_4 are valid.

According to the rules above, the control scheme algorithm is designed as shown in Fig. 4.2. The algorithm is composed of two steps. In step 1, first, the conditions for increasing and decreasing the output C_i are defined. Because in this application a stable output is highly preferred, the change of C_i should be carefully limited. According to the algorithm, only if the value of $H_i(t)$ is larger than $C_i(t-1)$ by a fixed number b_{up} , $C_i(t)$ is allowed to increase. Similarly, only if $H_i(t)$ is smaller than $C_i(t-1)$ by b_{down} (or $H_i(t)$ lower than a predefined baseline, 0.1 in this study), $C_i(t)$ is allowed to decrease. Both the increasing and decreasing of $C_i(t)$ use the same step variable ΔC to constrain the change rate. After that, the lower bound and the upper bound of C_i are applied. Through this process, the values of $C_i(t)$ ($i = 1, 2, 3, 4$) are temporarily determined.

Step 2 is to check the validity of the synergy pairs which are open-close ($\{C_1(t), C_2(t)\}$) and pronate-supinate ($\{C_3(t), C_4(t)\}$). First, some irregular cases where $C_i(t)$ and $C_j(t)$ ($\{i, j\} = \{1, 2\}$ or $\{3, 4\}$) are both very large have been observed in experiments. This usually happens when a subject maintains a movement with too much strength, which creates fluctuations on the other synergy. In that case, for example, the current output “100% open” can be suddenly turned to “100% close” (i.e. $C_2(t)$ becomes larger than

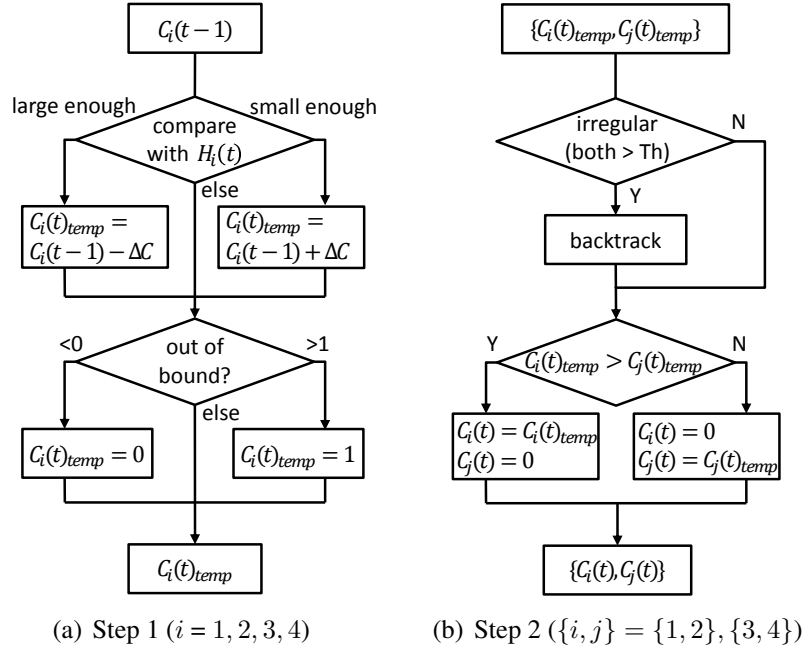


Figure 4.2: The block diagram of the control scheme.

$C_1(t)$) if the subject opens the hand too hard. To avoid this, in step 2, $C_i(t)$ and $C_j(t)$ are not allowed to go beyond a threshold simultaneously. If this happens, $C_i(t)$ and $C_j(t)$ will be backtracked to their previous states which are $C_i(t-1)$ and $C_j(t-1)$. Finally, because $C_i(t)$ and $C_j(t)$ are mutually exclusive, only the larger one is preserved, and the smaller one is assigned zero.

The parameter values mentioned above are: $b_{up} = 0.4$, $b_{down} = 0.4$, $\Delta C = 0.05$, and $Th = 0.7$. These values were determined by trial-and-error.

4.2.4 Experiment

Apparatus

The experiments of this study were conducted at two different institutions (The Johns Hopkins University, USA, and Kyoto University, Japan). The two institutions did not share the same experimental devices, so two different sets of devices were used to measure EMG signals. One was NI USB-6009, and the other was g.USBamp with g.GAMMAbox. The details are listed in Table 4.2. Both settings worked properly to measure EMG signals and no obvious difference in their performance is found.

Eight pairs of bipolar electrodes were placed in uniformly-spaced circles around the forearm. The electrode positions were not specifically fixed for easy attaching/detaching

Table 4.2: Experiment settings

Set	Device	Electrodes	Sample rate	Filter	Window length/ Overlap	Time resolution
#1	NI USB	8-channel bipolar	1000 Hz	30-300 Hz bandpass, 60 Hz notch	200/175	0.025s
#2	g.USBamp	8-channel bipolar	600 Hz	5-250 Hz bandpass, 60 Hz notch (built-in)	100/85	0.025s

consideration. Another electrode connected to ground was placed on the elbow near the bone to serve as the reference for the amplifier.

All experiments were performed on a personal computer platform. The source code was written in MATLAB 2011b leveraging the MATLAB Simulink environment. MSMS [2] supported the visualization of a virtual prosthesis on the computer screen. The virtual prosthesis has a large number of DOFs. It receives input commands through UDP ports and is animated in real time. In training mode it can provide visual cues to instruct the subject to make movements, and in testing mode it can work like a real prosthesis giving the subject visual feedback. The experimental setup is shown in Fig. 4.3, and a more detailed view of the virtual prosthesis is shown in Fig. 4.4.

Participants

Two kinds of experiments were conducted. One is offline, for data analysis, feature selection, and NMF parameter adjustment. The other is online, for evaluating the real-time performance and robustness with the NMF algorithm and the proposed control scheme.

Four healthy subjects participated in the offline experiment. One of them and another nine healthy subjects (total of 10) participated in the online experiment. In both experiments, subjects sat in a chair in front of a computer screen and rested their elbow (left or right, of their choice) on a flat surface.

Procedures (offline experiment)

In the offline experiment the subject performs nine movements: rest, open, close, pronate, supinate, open and pronate, close and pronate, open and supinate, and close and supinate in a random sequence.

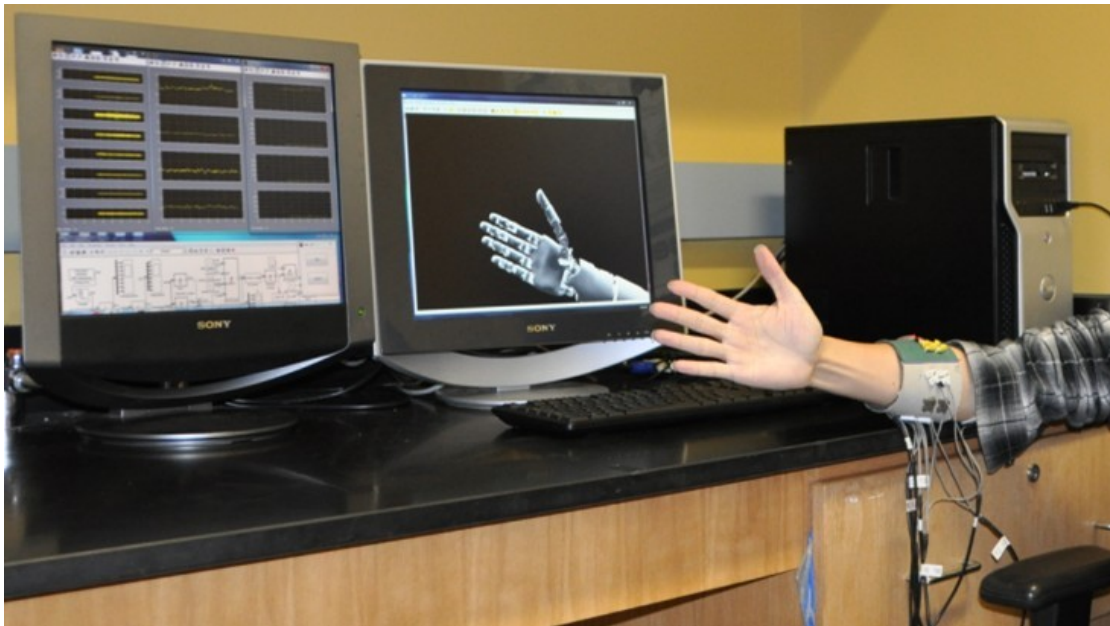


Figure 4.3: The experimental setup (corresponding to #1 in Table 4.2).

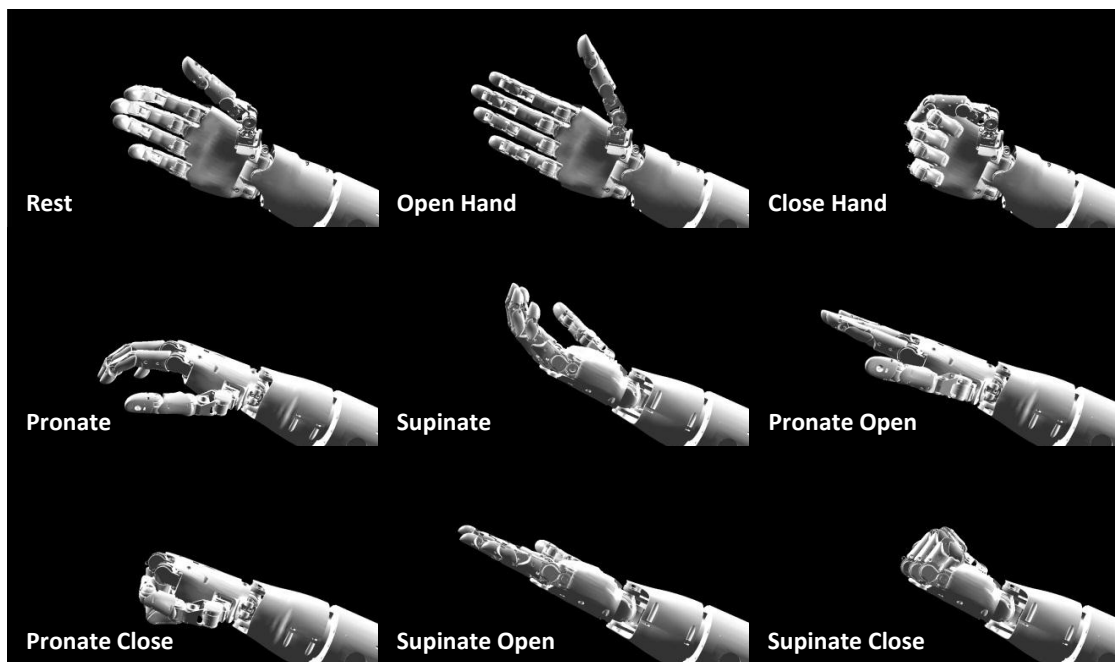


Figure 4.4: The virtual prosthesis generated by the musculoskeletal modeling software (MSMS) [2]. In the offline experiment, the static pictures of these nine movements are displayed as cues. In the online experiment, coherent animation of the virtual prosthesis is used as the visual feedback.

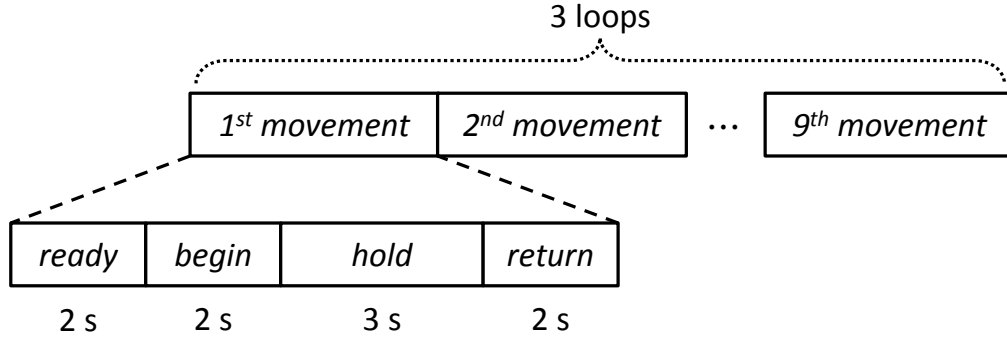


Figure 4.5: Flowchart of the offline experiment. The sequence of movements is random.

Performing one movement is treated as one trial. It lasts for 9 seconds and consists of four steps: preparation, movement, hold, and return. Visual cues, as shown in Fig. 4.4, are presented at the beginning of each trial, and the hint text “ready”, “begin”, “hold”, and “return” is also displayed together with cues. During the first 2 seconds, the subject is notified what movements he is supposed to make. During the next 2 seconds the subject gradually performs the movement until it reaches the target position with a certain strength. After that he holds the position for 3 seconds. In the final 2 seconds he returns his hand to the initial rest state. Nine trials, which consist of nine randomly ordered movements, compose a sub-session. The subject needs to complete three consecutive sub-sessions as a whole session. Therefore, the full length of one experiment session is $9 \times 9 \times 3 = 243$ seconds, as shown in Fig. 4.5.

Each subject completed at least three sessions, using medium strength, low strength, and high strength, respectively. This range was used to determine the best strength for training.

Procedures (online experiment)

The online experiment includes one training session and two test sessions.

The first step is a training session which is a simple version of the offline experiment. The subject only needs to perform five basic movements in a row: rest, open, close, pronate, and supinate. One trial containing one movement still lasts for 9 seconds, so it takes 45 seconds for the whole session. In these trials, the subject makes a movement at the very beginning and returns to a rest state at the last second without cues. The first trial “rest” is only used to record the baseline, and the other four trials are used to calculate the synergy matrix $\mathbf{W}$. $\mathbf{W}$ has four columns ($W_i, i = 1, \dots, 4$) which represent four synergies. Each W_i is calculated from one trial of training data by using the NMF algorithm.

After the subject completes the training session, he practices to become familiar with the real-time prosthesis control, and then does the online tests.

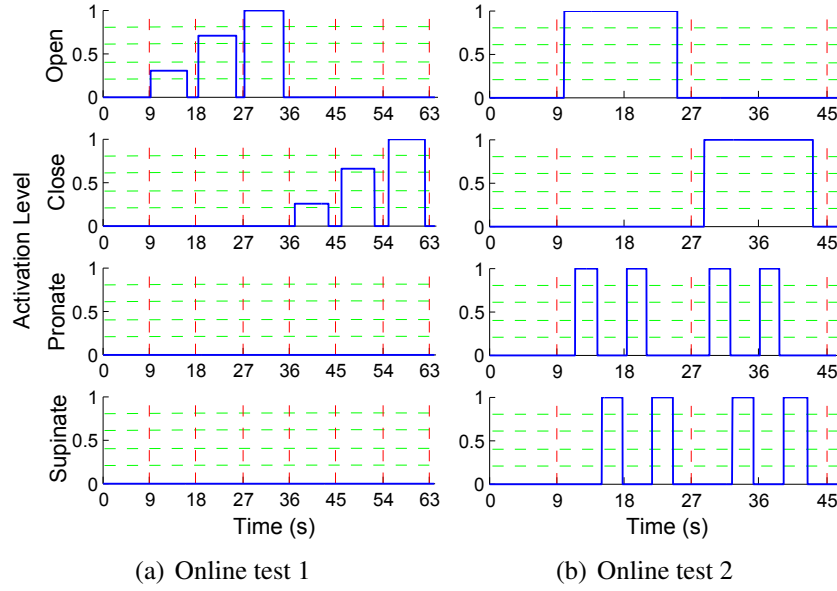


Figure 4.6: An ideal demonstrations of online (a) test 1 and (b) test 2. Blue lines are the ideal output of control signals C_1 , C_2 , C_3 , and C_4 . Only the first halves of the two tests are shown.

The tests are done to verify how precisely the subject can control the prosthesis with the proposed algorithm and the control scheme, and whether the subject can maintain the control signal in a stable manner. For this purpose the subject is instructed to look at a scope window which dynamically displays the value of the 4-channel control signal $\mathbf{C}$ on the screen. In this way, the subject can be aware of the “exact value” (in a range of $[0, 1]$) reflected by his movement.

In test 1, the subject is to perform four basic movements (open/close/pronate/supinate) at different strength levels. The sequence of movements is fixed as “open-close-pronate-supinate”. For each movement, the subject completes three trials (9 s each) from low strength to high strength. As the y-axes of the scope window are divided into five intervals from 0 to 1 by the division of 0.2, the subject is asked to try to make the control signal C_i fall into the 2nd interval ($[0.2, 0.4]$), the 4th interval ($[0.6, 0.8]$), and reach the maximum ($[1, 1]$), respectively. The signal should be maintained as stable as possible until the end of 7th second. The first half part of test 1 is demonstrated in Fig. 4.6(a). The full length of test 1 is $9 + 4 \times 3 \times 9 = 117$ seconds (an additional trial at the beginning is for preparation).

Test 2 verifies the subject’s ability to control the four varieties of hand/wrist combined movements: pronate open, pronate close, supinate open, and supinate close. Since there is strong mutual interference between the synergies of hand movement and the synergies of wrist movement, it is very difficult for subjects to precisely control the combined move-

ments. Therefore, the subject is not required to make movements by various strength levels, but only by the maximal strength.

Test 2 has four trials. In each trial, the subject first starts a primary movement, keeps it at maximum, then combines it with two secondary movements twice. In the 1st trial, the subject first activates “open” to maximal level, then activates “pronate-supinate”, and then again, without releasing the “open”. Each combined movement should be kept at maximum for 2 seconds at least. The timing of making each movement is controlled by the subject himself as long as he completes the series of movements in one trial of 18 seconds. The sequence of the trials is still “open (with combinations, the same below)”, “close”, “pronate”, and “supinate”. The first half of test 2 is demonstrated in Fig. 4.6(b). The full length of test 2 is $9 + 4 \times 18 = 81$ seconds (including the first trial for preparation).

Measures

The overall performance of both the offline and online experiments is evaluated by obtaining the R^2 value used in previous studies [102] [110] [111]. It expresses the relative distance between the estimated value and the target value:

$$R^2 = 1 - \frac{\sum_{k=1}^K \sum_t^T (\widehat{H_k(t)} - H_k(t))^2}{\sum_{k=1}^K \sum_t^T (H_k(t) - \overline{H_k(t)})^2}, \quad (4.6)$$

where $H_k(t)$ is the target value for k th synergy at time t , whose temporal average is $\overline{H_k(t)}$. $\widehat{H_k(t)}$ is the corresponding estimation, equal to the (k, t) th element of $\mathbf{H}_{est}$ (offline), or $C_k(t)$ (online). The numerator is the total mean squared error of the estimates, and the denominator is the total variance. R^2 value is a global measurement of estimation performance. Larger values mean the estimation is closer to the target with an upper bound of unity.

In the offline experiment, the data are simply labeled such that $H_k(t)$ has only two values: 0 for inactive and 1 for active. In the online experiment, except for the target value 1 representing full activation, there are two target intervals $[0.2, 0.4]$ and $[0.6, 0.8]$. The most common values appearing in those intervals are taken as the corresponding label value. For example, for one trial if the target was $[0.2, 0.4]$, and the subject held his movement mostly at 0.35, then the label value should also be 0.35; if the subject held his movement mostly at 0.5, the label should be equal to 0.4 because it cannot be out of range.

T in Eq. 4.6 represents the time intervals of interest. The value is determined by each experiment setting. For the offline experiment, it is the 5~7th seconds in each trial when the cue is “hold”. For online test 1, T is expanded a little longer to 4~7th seconds in each trial, where the 1~3rd seconds are for the subject to adjust his movement. For online test 2, T should be a 2-second interval when the combined movements are fully activated. In

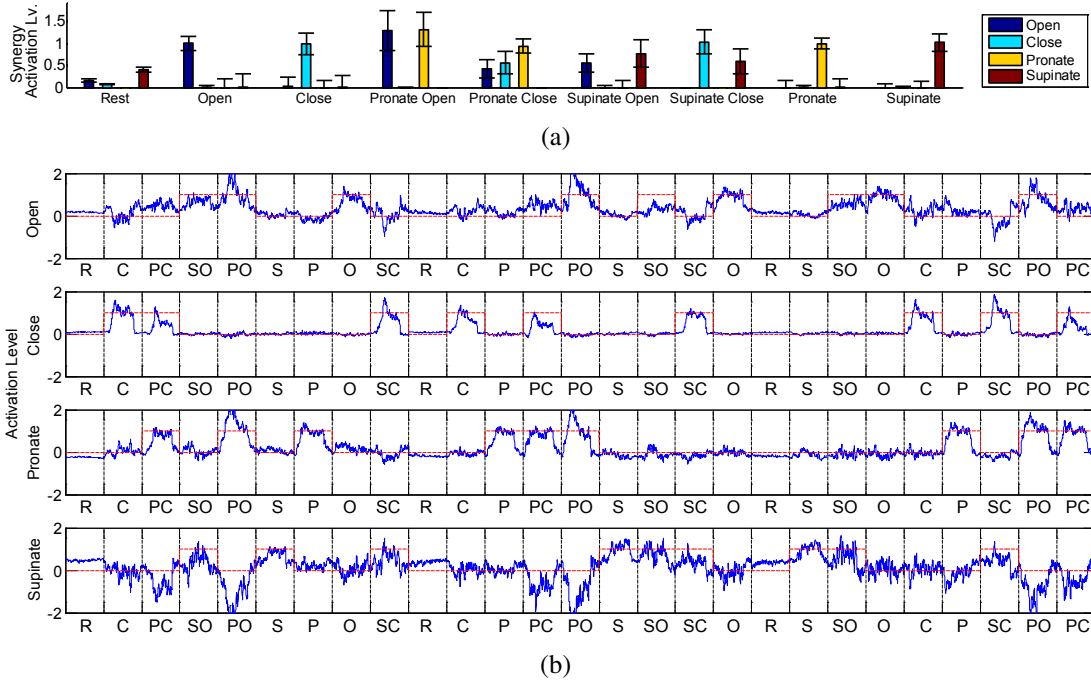


Figure 4.7: One session of the offline experiment. (a) shows the average synergy activation levels for 9 different movements in the experiment and their standard variations. In (b), the blue lines are the time-varying synergy activations $\widehat{H_k(t)}$ in Eq. 4.6) in trials (the first 2 seconds for preparation in each trial is omitted). The labels of x-axes represent different movements (R for rest, O for open, C for close, P for pronate, and S for supinate). The red dash lines are the target activation levels ($H_k(t)$ in Eq. 4.6).

this test, the starting point of each movement is uncertain and is manually labeled over the data.

Data analysis

Four subjects completed 14 offline sessions which include five sessions at medium strength, four sessions at low strength and five sessions at high strength. Figure 4.7 shows an example of one complete offline session. Note that in the offline experiment, the subject did not receive visual feedback; the control scheme was not used; the baseline removal was also not applied. In the figure there are a lot of fluctuations, as well as non-zero baselines. In that case the R^2 value can hardly achieve a high level. But as the main target is to compare the different features, these negative effects can help us find the robust ones. Despite these limitations, the synergies of “close” and “pronate” have ideal shapes in the figure, and Fig. 4.7(a) also shows that the nine movements are well represented by average

synergy activations.

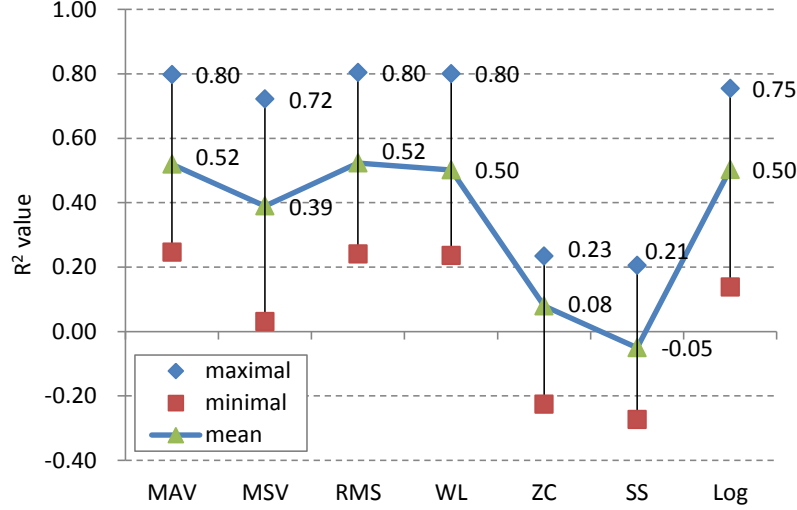


Figure 4.8: The maximal (blue diamonds), minimal (red squares) and mean (green triangles) R^2 value of 14 offline experiment sessions for seven features.

After analyzing the data of the offline experiment, the performance of different features is shown as in Fig. 4.8. Through comparing the maximal, minimal and mean R^2 values of different features, it can be concluded that the best three features are MAV, RMS, and WL. Log and MSV are nearly as good. ZC and SS are frequently used in classification methods but are obviously not oriented to the NMF algorithm.

As per the results of the offline experiment, the three best performing features: MAV, RMS, and WL, are adopted in the online experiment. The total dimension of the feature set is 24 (3×8 channel). Because features have different ranges, each one is normalized by a scaling value equal to the maximum of that feature in the training data.

4.3 Results

Figure 4.9 shows an example of the training result. In the figure, red lines are synergy activation levels which are calculated by applying the NMF algorithm to the input EMG data, and blue lines are the output of the proposed control scheme. Whenever a training session is completed, this figure will be displayed for evaluating the quality of training. Relatively smooth red lines (as Fig. 4.9) indicate high training quality.

Figure 4.10(a) shows an example of online test 1, where the subject completed the task well. There is little redundant control or noise during the whole test. With the exception of the first pronate trial which is slightly over-activated, other movements fall in the desired

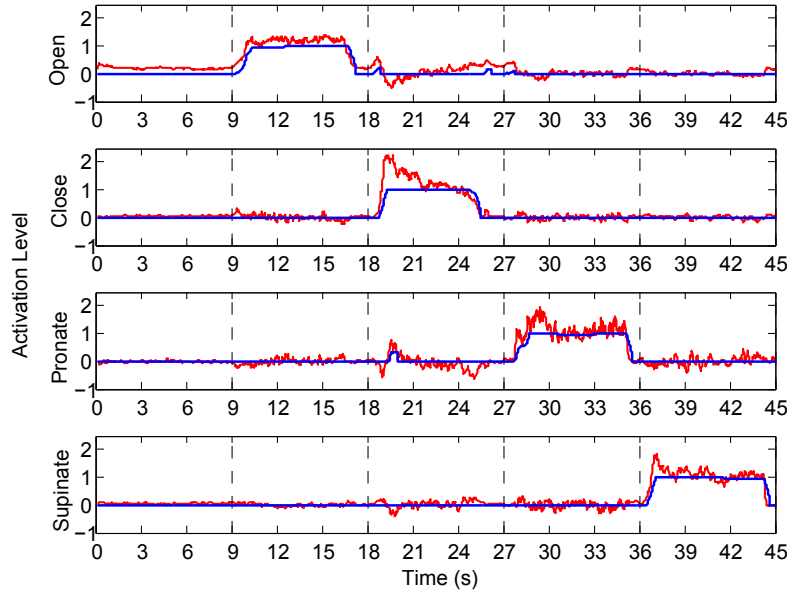


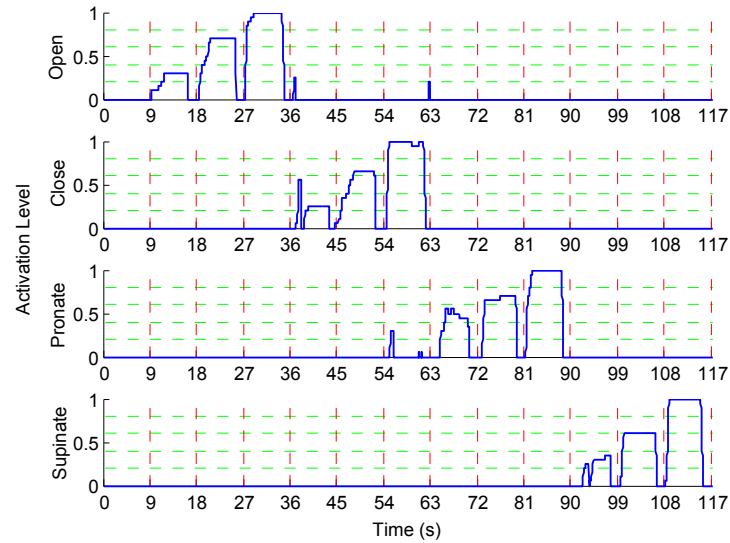
Figure 4.9: An example of training session data from the online experiment. Red lines are activation level of four synergies, and blue lines are corresponding control signals C_1 , C_2 , C_3 , and C_4 .

intervals, and they are also stably maintained in most trials. This example shows NMF is evidently capable of fluidly controlling individual DOF.

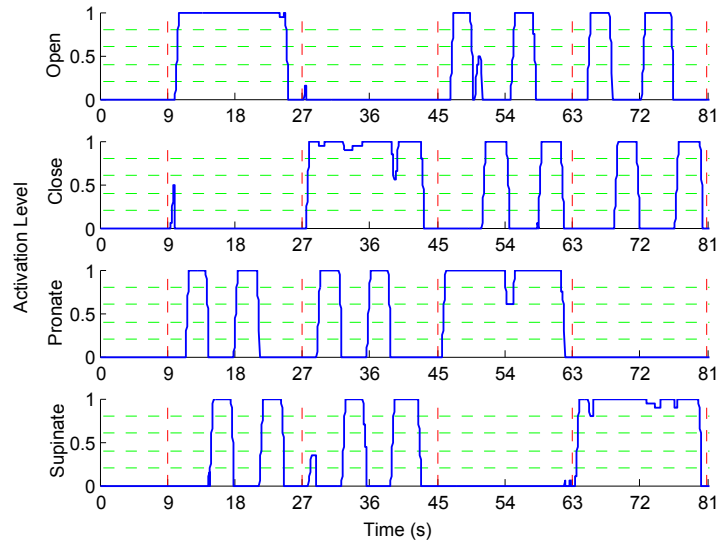
Figure 4.10(b) shows an example of online test 2. The subject was also able to finish the task adequately. Generally, multiple DOFs are more difficult to control than an individual DOF is, because wrist and hand movements are likely to interfere with each other. To simultaneously control two DOFs precisely is still too challenging to complete; thus the subject is only required to achieve maximal activations in this test.

Figure 4.11 shows the evaluation of the online experiment. The average R^2 values of test 1 and test 2 among 10 subjects achieve 0.97 and 0.93, respectively, while using a low-pass filter the R^2 values are 0.78 and 0.77. The result indicates that the proposed control scheme always has a satisfactory result and outperforms the low-pass filter. Note that online filters have significant time delays, so in this comparison an offline zero-phase FIR filter (0.5 Hz cutoff) is applied.

Figure 4.12 shows a detailed example of the comparison between the original data, the output of the filter, and the output of the control scheme for test 1. The figure shows that the filter works well for eliminating the high-frequency noise but is unsuitable for the task of “making movements and maintaining them”. From each trial in the figure, though the subject was told to maintain the movement, $\mathbf{H}_{est}$ has a trend to spontaneously decrease after it reaches a maximum, while the proposed control scheme can effectively



(a) Online test 1



(b) Online test 2

Figure 4.10: Examples of online (a) test 1 and (b) test 2. Blue lines are the system output (control signals C_1 , C_2 , C_3 , and C_4) which the subject can see from the screen during the experiment.

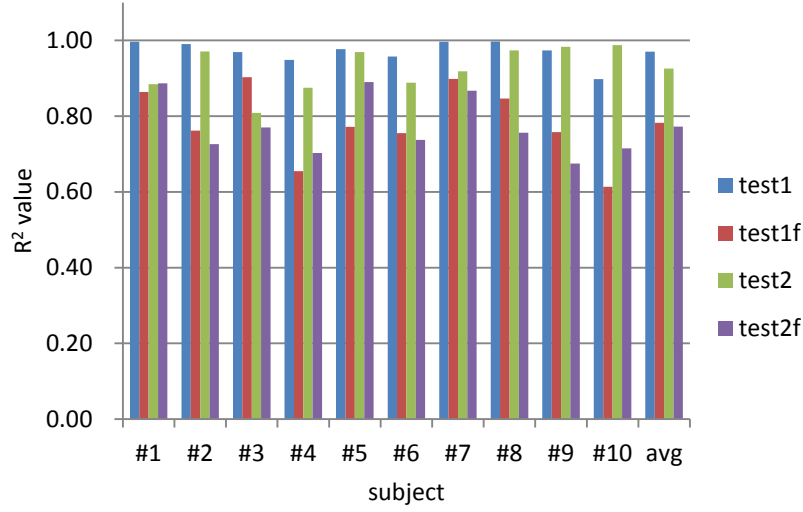


Figure 4.11: The R^2 value calculated from all the ten subjects' online experimental data (test 1 and test 2), and also from the same data processed by low-pass filter but not the control scheme (test 1f and test 2f). The average R^2 values of test 1 and test 2 are 0.97 and 0.93, respectively.

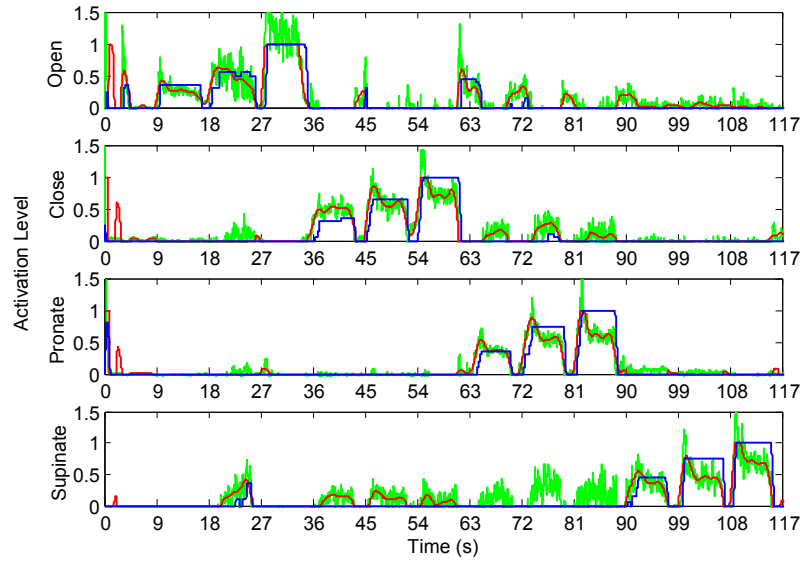


Figure 4.12: One example of the experiment data of test 1 for comparison. Green lines are estimated synergy activations ($\mathbf{H}_{est}$). Red lines are $\mathbf{H}_{est}$ processed by the low-pass filter (with a output range $[0,1]$). Blue lines are $\mathbf{H}_{est}$ processed by the control scheme (control signals C_1 , C_2 , C_3 , and C_4).

prevent the signal from an unwanted drop most of the time. Of course the drop does not occur every time (see 36~45 s of “close”), and using the online filter, the subject may give a better response to compensate for the drop. However, it is still difficult to suppress fluctuations without using the control scheme.

Besides maintaining the movement and reducing the fluctuations, the proposed algorithm can also suppress the wrongly activated synergy to some extent. See the 4th plot in Fig. 4.12 as an example. According to experiments, the “supinate” synergy should be only active in the interval of 90~117 s. However, green lines indicate that the “supinate” synergy has been wrongly activated in the intervals of 18~27 s and 36~90 s which are seven trials in total. Among them, the last three trials (63~90 s) were not truly affected because the “pronate” synergy was dominant during that period. For the other trials, the output of the low-pass filter (red lines) just reproduced the shape of the affected data, whereas the control scheme kept a nearly desirable zero output.

4.4 Discussion

4.4.1 Offline experiment

The results of offline experiments indicate that time-domain features (especially those with first order property) are suitable to NMF, but frequency-domain ones may not be. Through the experiments, it has also been found that MAV and RMS features only perform well when the signal has a zero mean, which means it is better to either remove the DC component beforehand, or to subtract the mean values when calculating the features. On the other hand, WL are robust for signals with either zero mean or non-zero mean, since its definition has no relation with mean values.

Aside from the performance of different features, the experiments also reveal some other information. First, it is difficult to judge which is better, high strength or normal strength, but it is certain that using low strength always leads to the worst result. Second, the initial condition of the NMF algorithm can vary, for example, whether $f(\mathbf{Z})$, the input of NMF, should be trimmed so that only fully activated intervals are preserved; whether the $\mathbf{H}$ should be given an initial value, or even be kept unchanged during the NMF updating process. Tests on these various initial conditions have been made and no significant differences were found. The initial condition is most likely to affect the scaling of the output but not its trend over time.

4.4.2 Online experiment

Judging from the performance of the subjects, nearly 100% completed test 1 with high accuracy, though some needed more practice than others. There are two problems iden-

tified by the online experiments. The first is when activating one movement, unrelated movement from different synergies may also be activated by chance. The second is that the activation level is difficult to keep stable at low strength levels. For the first problem, through practice the subject can gradually learn how to suppress unrelated movements. For the second one, the poor performance is mainly due to noise (fluctuations) in $\mathbf{H}_{est}$. With the low strength level, the input data fluctuations near the baseline are difficult to correct with the control scheme. Repeating the training process helps solve this problem.

Test 2 is more difficult for most subjects. Although the individual movements are likely to be completed without problems, activating and maintaining two movements simultaneously are different from performing them individually. In that case, the muscle crosstalk will have considerable influence on the result. When the subject is maintaining one movement and tries to activate another, the first movement becomes uncontrollable and its level drops. In the experiments, all the subjects successfully completed test 2, but the overall practice time was longer than that of test 1.

4.4.3 Practicality

For surface EMG measuring, every time the electrodes are detached/attached again, the electrodes' positions change, which requires training again. Moreover, after wearing the prosthesis for a period of time, the electrodes may be slightly displaced, and the states of the skin and muscle may also be changed (due to fatigue, temperature, etc.). As a result, the original training data are likely to become unsuitable over time. Until now, there has been no reliable solution to prevent the original training data from degrading. One alternative is to shorten the training process so that the system can recover. In the experiment, the training time is 54 s in total, 9 s for each movement (including rest). Shorter training time is also possible but has not been verified. Moreover, the direct cause of performance drop is baseline drift. Having a countermeasure against baseline drift may help avoid retraining and further raise the efficiency.

It is not clear how many DOFs can be supported by this kind of prosthesis control method. Currently, the 4-movement control paradigm (open, close, pronate, and supinate) tested in this study has acceptable performance. However, if the number of synergies increase with DOFs, subjects would be more likely to activate unrelated synergies when performing any movement. A better solution may be to find a set of synergies whose number is less than the number of DOFs, where such synergies form basis vectors for a wide variety of movements. In [109] Ajiboye and Weir used an NMF-based method to build a predictive framework for 33 hand postures with only 8 synergies. Their classification method, however, does not involve proportional control. When synergies and movements do not have one-to-one correspondence, the usage of the activation levels of multiple synergies to represent degrees of a single movement is an issue to be addressed.

In this study, considering the easy attaching/detaching property of the prosthesis, the

circularly arranged electrodes are fixed at the proper position on the forearm. In other studies, the electrodes were often placed on specified muscle sites. Both settings have their advantages and disadvantages. Setting the electrodes on the correct sites requires a great deal of preparation, especially for amputees, so the specified electrodes placement is more applicable to implantable myoelectric sensors than to surface EMG. However, it was found that the specified electrodes placement does improve the quality of signals, because it lowers the muscle crosstalk. If conditions permit or good performance is highly desired, the specified electrode placement can be taken into consideration.

Chapter 5

Conclusion

This thesis discussed about three kinds of human machine interface. They are for: 1) vigilance estimation, 2) robot control, 3) prosthesis control. All of them are based on biomedical signals.

5.1 Conclusion

Firstly, a vigilance estimation interface based on EOG is proposed. Compared with other studies, a quite large experimental data set containing 36 sessions from 22 subjects is gathered. Moreover, as many as 20 kinds of EOG features are extracted. Analysis to these data further indicates which kinds of EOG features have higher correlation with vigilance. Among all the EOG features, SEM proportion, eyelid velocity of blink, and low frequency energy ratio are prominently effective. Single feature has the highest average correlation of 0.70. After combination of multiple features, the correlation can reach 0.75. The feature combination results indicate the importance of using multiple features. This conclusion is not limited to EOG features but also other kinds of features, such as EEG and video. This work reveals the feasibility of realizing an EOG based vigilance estimation system, and also provides important theoretical basis and design ideas to deal with the problem of detecting driver fatigue.

Secondly, a novel EOG/EEG hybrid human-machine interface is proposed. The hybrid interface works in two modes, where EOG mode detects eye movements including blink, frown, wink and gaze, and EEG mode adopts multi-component ERPs to judge the user's visual focus. The usefulness of the hybrid interface is verified by two online experiments about controlling different robots. In each experiment, most subjects have satisfactory performance, and a few subjects were even able to complete the tasks with high effectiveness comparable with hand operation. The experiment results suggest that there is a complimentary relationship between eye movements and ERP paradigm, and also suggest

that the introduced hybrid interface is promising for BCI related applications.

Finally, this study proposes an interface to control a prosthetic hand by EMG signals based on muscle synergies. The muscle synergy model suggests a framework to transform commands of the central nervous system to a set of complex muscular movements. This study focuses on controlling four kinds of hand/wrist movements of the prosthesis: open, close, pronate, and supinate. Nonnegative matrix factorization algorithm is used to decode the EMG signals, which has a short training time. The most important feature of this HMI is it supports proportional control on a single DOF, and simultaneous control on multi-DOFs. The proposed HMI has good online performance because a control scheme is applied to smooth the control signals. The results indicate that fluid hand/wrist movements could be estimated from EMG. The average R^2 values achieved by all subjects for the single-DOF test is 0.97, and for the multiple-DOF test is 0.93.

5.2 Future work

For the vigilance estimation interface, the future work includes two aspects. The first and principle is to consider the individual differences on vigilance estimation. Many EOG features have shown a good correlation with the trend of vigilance transition, but it needs extra work to make sure those features can correctly correspond to the absolute value of vigilance. The second is about the realization of an optimal rearrangement of EOG electrodes. Although the experiment about using ICA to separate the horizontal EOG and vertical EOG from the forehead EOG has been carried out, it is still necessary to verify the forehead EOGs on vigilance estimation, and consider the on-line availability of this forehead approach.

For the other two interfaces related to control, until now, experiments have been only carried on able-bodied people. However, the prosthesis control interface is targeted on amputees, and the robot control interface is also mainly for disabled persons. Thus it is necessary to perform further experiments on such subjects, to verify how well they can use the proposed interfaces, and how their performance differs from able-bodied people. Further optimization is necessary for the eye movement detection method, and its ability to detect irregular eye movements is expected to be improved. The prosthesis control experiment can be assigned more specific tasks, such as grasping objects and rotating handles. These tasks can better reflect whether the proposed method is suitable of daily usage. Another objective is to try to increase the number of supported DOFs while ensuring their stability, because it is desirable that the final product can cover as many amputees' daily needs as possible. It is also expected that these interfaces can be also applied to other traditional devices to bring fresh new user experience.

Bibliography

- [1] B. Jammes, H. Sharabty, and D. Esteve, “Automatic eog analysis: a first step toward automatic drowsiness scoring during wake-sleep transitions,” *Somnologie-Schlafforschung und Schlafmedizin*, vol. 12, no. 3, pp. 227–232, 2008.
- [2] R. Davoodi and G. E. Loeb, “A software tool for faster development of complex models of musculoskeletal systems and sensorimotor controllers in simulinkTM,” *Journal of Applied Biomechanics*, vol. 18, no. 4, pp. 357–365, 2002.
- [3] S. Seng, X. Xu, X. Yang, and D. Miao, “Research on the driver fatigue monitoring method based on the dempster-shafer theory,” in *Proceedings of Chinese Control and Decision Conference*, pp. 4176–4179, IEEE, 2010.
- [4] L. Gan and Q. Liu, “Eye detection based on rank order filter and projection function,” in *Proceedings of the International Conference on Computational Science and Its Applications*, vol. 1, pp. 642–645, IEEE, 2010.
- [5] Q. Ji and X. Yang, “Real-time eye, gaze, and face pose tracking for monitoring driver vigilance,” *Real-Time Imaging*, vol. 8, no. 5, pp. 357–377, 2002.
- [6] Q. Ji, Z. Zhu, and P. Lan, “Real-time nonintrusive monitoring and prediction of driver fatigue,” *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, pp. 1052–1068, 2004.
- [7] M. Bhuiyan, “Driver assistance systems to rate drowsiness: A preliminary study,” *New Advances in Intelligent Decision Technologies*, vol. 199, pp. 415–425, 2009.
- [8] T. Jung, S. Makeig, M. Stensmo, and T. Sejnowski, “Estimating alertness from the EEG power spectrum,” *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 1, pp. 60–69, 1997.
- [9] C. Lin, L. Ko, I. Chung, T. Huang, Y. Chen, T. Jung, and S. Liang, “Adaptive EEG-based alertness estimation system by using ICA-based fuzzy neural networks,” *IEEE Transactions on Circuits and Systems-I*, vol. 53, no. 11, pp. 2469–2476, 2006.

- [10] P. Davidson, R. Jones, and M. Peiris, "EEG-based lapse detection with high temporal resolution," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 5, pp. 832–839, 2007.
- [11] L. Shi, H. Yu, and B. Lu, "Semi-supervised clustering for vigilance analysis based on EEG," in *Proceedings of the International Joint Conference on Neural Networks*, pp. 1518–1523, IEEE, 2007.
- [12] L. Shi and B. Lu, "Dynamic clustering for vigilance analysis based on EEG," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 54–57, IEEE, 2008.
- [13] L. Shi and B. Lu, "Off-line and on-line vigilance estimation based on linear dynamical system and manifold learning," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6587–6590, IEEE, 2010.
- [14] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 121–130, 2011.
- [15] L. Shi and B. Lu, "EEG-based vigilance estimation using extreme learning machines," *Neurocomputing*, vol. 102, pp. 135–143, 2013.
- [16] J. Ma, L. Shi, and B. Lu, "Vigilance estimation by using electrooculographic features," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6591–6594, IEEE, 2010.
- [17] H. Cai, J. Ma, L. Shi, and B. Lu, "A novel method for EOG features extraction from the forehead," in *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3075–3078, IEEE, 2011.
- [18] Z. Wei and B. Lu, "Online vigilance analysis based on electrooculography," in *Proceedings of the International Joint Conference on Neural Networks*, pp. 1–7, IEEE, 2012.
- [19] A. Bulling, J. Ward, H. Gellersen, and G. Troster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 741–753, 2011.
- [20] G. Garcia-Molina, F. Abtahi, and M. Lagares-Lemos, "Automated nrem sleep staging using the electro-oculogram: A pilot study," in *Proceedings of the 34th Annual*

International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2255–2258, IEEE, 2012.

- [21] J. Yu, B. Lee, and D. Kim, “EOG based eye movement measure of visual fatigue caused by 2d and 3d displays,” in *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 305–308, IEEE, 2012.
- [22] N. Galley, R. Schleicher, and L. Galley, “Blink parameters as indicators of driver’s sleepiness—possibilities and limitations,” *Vision in Vehicles. Amsterdam: Elsevier*, 2004.
- [23] R. Barea, L. Boquete, and M. Mazo, “System for assisted mobility using eye movements based on electrooculography,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 4, pp. 209–218, 2002.
- [24] Y. Chen and W. Newman, “A human-robot interface based on electrooculography,” in *Proceedings of the International Conference on Robotics and Automation*, vol. 1, pp. 243–248, IEEE, 2004.
- [25] K. Wright, J. Hull, and C. Czeisler, “Relationship between alertness, performance, and body temperature in humans,” *The American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, vol. 283, no. 6, pp. R1370–R1377, 2002.
- [26] D. Dinges, M. Mallis, G. Maislin, and J. Powell IV, *Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management*. National Highway Traffic Safety Administration, 1998.
- [27] M. Atienza, J. Cantero, R. Stickgold, and J. Allan Hobson, “Eyelid movements measured by nightcap predict slow eye movements during quiet wakefulness in humans,” *Journal of Sleep Research*, vol. 13, no. 1, pp. 25–29, 2004.
- [28] S. Hanke, J. Zeitlhofer, G. Wiest, W. Mayr, and D. Moser, “Automated vigilance classification based on EOG signals: Preliminary results,” in *Proceedings of the World Congress on Medical Physics and Biomedical Engineering*, pp. 428–431, Springer, 2009.
- [29] K. Hyoki, M. Shigeta, N. Tsuno, Y. Kawamuro, and T. Kinoshita, “Quantitative electro-oculography and electroencephalography as indices of alertness,” *Electroencephalography and Clinical Neurophysiology*, vol. 106, no. 3, pp. 213–219, 1998.

- [30] T. Chieh, M. Mustafa, A. Hussain, S. Hendi, and B. Majlis, "Development of vehicle driver drowsiness detection system using electrooculogram (EOG)," in *Proceedings of the 1st International Conference on Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering*, pp. 165–168, IEEE, 2005.
- [31] V. Chouhan and S. Mehta, "Total removal of baseline drift from ECG signal," in *Proceedings of the International Conference on Computing: Theory and Applications*, pp. 512–515, IEEE, 2007.
- [32] M. Tinati and B. Mozaffary, "A wavelet packets approach to electrocardiograph baseline drift cancellation," *International Journal of Biomedical Imaging*, vol. 2006, pp. 1–9, 2006.
- [33] S. Makeig, A. Bell, T. Jung, and T. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in Neural Information Processing Systems*, vol. 8, pp. 145–151, 1996.
- [34] E. Magosso, F. Provini, P. Montagna, and M. Ursino, "A wavelet based method for automatic detection of slow eye movements: a pilot study," *Medical Engineering and Physics*, vol. 28, no. 9, pp. 860–875, 2006.
- [35] E. Magosso, M. Ursino, F. Provini, and P. Montagna, "Wavelet analysis of electroencephalographic and electro-oculographic changes during the sleep onset period," in *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4006–4010, IEEE, 2007.
- [36] E. Magosso, M. Ursino, A. Zaniboni, F. Provini, and P. Montagna, "Visual and computer-based detection of slow eye movements in overnight and 24-h eog recordings," *Clinical Neurophysiology*, vol. 118, no. 5, pp. 1122–1133, 2007.
- [37] E. Magosso, M. Ursino, A. Zaniboni, and E. Gardella, "A wavelet-based energetic approach for the analysis of biomedical signals: application to the electroencephalogram and electro-oculogram," *Applied Mathematics and Computation*, vol. 207, no. 1, pp. 42–62, 2009.
- [38] D. Nie, X. Wang, L. Shi, and B. Lu, "EEG-based emotion recognition during watching movies," in *Proceedings of the 5th International IEEE EMBS Conference on Neural Engineering*, pp. 667–670, IEEE, 2011.
- [39] K. Kaida, T. ÅKerstedt, G. Kecklund, J. Nilsson, and J. Axelsson, "Use of subjective and physiological indicators of sleepiness to predict performance during a vigilance task," *Industrial Health*, vol. 45, no. 4, pp. 520–526, 2007.

- [40] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [41] F. A. Mussa-Ivaldi and L. E. Miller, "Brain-machine interfaces: computational demands and clinical needs meet basic neuroscience," *TRENDS in Neurosciences*, vol. 26, no. 6, pp. 329–334, 2003.
- [42] J.-H. Lee, J. Ryu, F. A. Jolesz, Z.-H. Cho, and S.-S. Yoo, "Brain-machine interface via real-time fMRI: preliminary study on thought-controlled robotic arm," *Neuroscience Letters*, vol. 450, no. 1, pp. 1–6, 2009.
- [43] N. Naseer and K.-S. Hong, "Classification of functional near-infrared spectroscopy signals corresponding to the right-and left-wrist motor imagery for development of a brain-computer interface," *Neuroscience Letters*, vol. 553, pp. 84–89, 2013.
- [44] N. Naseer, M. J. Hong, and K.-S. Hong, "Online binary decision decoding using functional near-infrared spectroscopy for the development of brain-computer interface," *Experimental Brain Research*, vol. 232, no. 2, pp. 555–564, 2014.
- [45] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler, "An MEG-based brain-computer interface (BCI)," *Neuroimage*, vol. 36, no. 3, pp. 581–593, 2007.
- [46] J. R. Wolpaw, G. E. Loeb, B. Z. Allison, E. Donchin, O. F. do Nascimento, W. J. Heetderks, F. Nijboer, W. G. Shain, and J. N. Turner, "BCI meeting 2005-workshop on signals and recording methods," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 138–141, 2006.
- [47] G. Pfurtscheller and F. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [48] Y. Li, J. Long, T. Yu, Z. Yu, C. Wang, H. Zhang, and C. Guan, "An EEG-based BCI system for 2-D cursor control by combining Mu/Beta rhythm and P300 potential," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2495–2505, 2010.
- [49] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, p. 046002, 2009.

- [50] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, “Single-trial analysis and classification of ERP components—a tutorial,” *NeuroImage*, vol. 56, no. 2, pp. 814–825, 2011.
- [51] Y. Zhang, G. Zhou, Q. Zhao, J. Jin, X. Wang, and A. Cichocki, “Spatial-temporal discriminant analysis for ERP-based brain-computer interface,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 2, pp. 233–243, 2013.
- [52] T. Kaufmann, S. Schulz, C. Grünzinger, and A. Kübler, “Flashing characters with famous faces improves ERP-based brain-computer interface performance,” *Journal of Neural Engineering*, vol. 8, no. 5, p. 056016, 2011.
- [53] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, “A spelling device for the paralysed,” *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.
- [54] S. Sutton, M. Braren, J. Zubin, and E. John, “Evoked-potential correlates of stimulus uncertainty,” *Science*, vol. 150, no. 3700, pp. 1187–1188, 1965.
- [55] B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Graser, “BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 107–116, 2010.
- [56] C. Brunner, B. Allison, C. Altstätter, and C. Neuper, “A comparison of three brain-computer interfaces based on event-related desynchronization, steady state visual evoked potentials, or a hybrid approach using both signals,” *Journal of Neural Engineering*, vol. 8, no. 2, p. 025010, 2011.
- [57] L. Farwell and E. Donchin, “Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials,” *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.
- [58] H. V. Semlitsch, P. Anderer, P. Schuster, and O. Presslich, “A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP,” *Psychophysiology*, vol. 23, no. 6, pp. 695–703, 1986.
- [59] G. Gratton, M. G. Coles, and E. Donchin, “A new method for off-line removal of ocular artifact,” *Electroencephalography and Clinical Neurophysiology*, vol. 55, no. 4, pp. 468–484, 1983.

- [60] L. Y. Deng, C.-L. Hsu, T.-C. Lin, J.-S. Tuan, and S.-M. Chang, "EOG-based human-computer interface system development," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3337–3343, 2010.
- [61] R. Barea, L. Boquete, M. Mazo, and E. López, "Wheelchair guidance strategies using EOG," *Journal of Intelligent and Robotic Systems*, vol. 34, no. 3, pp. 279–299, 2002.
- [62] R. C. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous P300 BCI with SSVEP-based control state detection," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1781–1788, 2011.
- [63] S. Amiri, A. Rabbi, L. Azinfar, and R. Fazel-Rezai, "A review of P300, SSVEP, and hybrid P300/SSVEP brain-computer interface systems," in *Brain-Computer Interface Systems—Recent Progress and Future Prospects*, InTech, 2013.
- [64] B. Allison, C. Brunner, V. Kaiser, G. Müller-Putz, C. Neuper, and G. Pfurtscheller, "Toward a hybrid brain-computer interface based on imagined movement and visual attention," *Journal of Neural Engineering*, vol. 7, no. 2, p. 026007, 2010.
- [65] G. Pfurtscheller, T. Solis-Escalante, R. Ortner, P. Linortner, and G. R. Müller-Putz, "Self-paced operation of an SSVEP-based orthosis with and without an imagery-based brain switch: a feasibility study towards a hybrid BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 409–414, 2010.
- [66] A. Savić, U. Kisić, and M. Popović, "Toward a hybrid BCI for grasp rehabilitation," in *Proceedings of the 5th European Conference of the International Federation for Medical and Biological Engineering*, pp. 806–809, 2012.
- [67] J. Li, H. Ji, L. Cao, D. Zang, R. Gu, B. Xia, and Q. Wu, "Evaluation and application of a hybrid brain computer interface for real wheelchair parallel control with multi-degree of freedom," *International Journal of Neural Systems*, vol. 24, no. 04, 2014.
- [68] Y. Punsawad, Y. Wongsawat, and M. Parnichkun, "Hybrid EEG-EOG brain-computer interface system for practical machine control," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1360–1363, 2010.
- [69] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K.-R. Müller, and B. Blankertz, "Enhanced performance by a hybrid NIRS-EEG brain computer interface," *Neuroimage*, vol. 59, no. 1, pp. 519–529, 2012.

- [70] M. J. Khan, M. J. Hong, and K.-S. Hong, "Decoding of four movement directions using hybrid NIRS-EEG brain-computer interface," *Frontiers in Human Neuroscience*, vol. 8, 2014.
- [71] R. Leeb, H. Sagha, R. Chavarriaga, and J. del R Millán, "A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025011, 2011.
- [72] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 11, pp. 3156–3166, 2013.
- [73] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A speedy hybrid BCI spelling approach combining P300 and SSVEP," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 2, pp. 473–483, 2014.
- [74] B. Rebsamen, E. Burdet, Q. Zeng, H. Zhang, M. Ang, C. Teo, C. Guan, and C. Laugier, "Hybrid P300 and Mu-Beta brain computer interface to operate a brain controlled wheelchair," in *Proceedings of the 2nd International Convention on Rehabilitation Engineering & Assistive Technology*, pp. 51–55, 2008.
- [75] Y. Su, Y. Qi, J.-x. Luo, B. Wu, F. Yang, Y. Li, Y.-t. Zhuang, X.-x. Zheng, and W.-d. Chen, "A hybrid brain-computer interface control strategy in a virtual environment," *Journal of Zhejiang University SCIENCE C*, vol. 12, no. 5, pp. 351–361, 2011.
- [76] H. Riechmann, N. Hachmeister, H. Ritter, and A. Finke, "Asynchronous, parallel on-line classification of P300 and ERD for an efficient hybrid BCI," in *Proceedings of the 5th International IEEE/EMBS Conference on Neural Engineering*, pp. 412–415, 2011.
- [77] Y. Zhang, Q. Zhao, J. Jin, X. Wang, and A. Cichocki, "A novel BCI based on ERP components sensitive to configural processing of human faces," *Journal of Neural Engineering*, vol. 9, no. 2, p. 026018, 2012.
- [78] J. Ma, Y. Zhang, Y. Nam, A. Cichocki, and F. Matsuno, "EOG/ERP hybrid human-machine interface for robot control," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 859–864, 2013.
- [79] A. Bulling, J. A. Ward, H. Gellersen, and G. Troster, "Eye movement analysis for activity recognition using electrooculography," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 4, pp. 741–753, 2011.

- [80] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayouth, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the P300 speller," *Journal of Neural Engineering*, vol. 3, no. 4, p. 299, 2006.
- [81] V. Abootalebi, M. H. Moradi, and M. A. Khalilzadeh, "A new approach for EEG feature extraction in P300-based lie detection," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 1, pp. 48–57, 2009.
- [82] A. Turnip, K.-S. Hong, and M.-Y. Jeong, "Real-time feature extraction of P300 component using adaptive nonlinear principal component analysis," *Biomedical Engineering Online*, vol. 10, no. 1, p. 83, 2011.
- [83] J. N. Mak, D. J. McFarland, T. M. Vaughan, L. M. McCane, P. Z. Tsui, D. J. Zeitlin, E. W. Sellers, and J. R. Wolpaw, "EEG correlates of P300-based brain–computer interface (BCI) performance in people with amyotrophic lateral sclerosis," *Journal of Neural Engineering*, vol. 9, no. 2, p. 026014, 2012.
- [84] A. Turnip and K.-S. Hong, "Classifying mental activities from EEG-P300 signals using adaptive neural network," *International Journal of Innovative Computing Information and Control*, vol. 8, no. 9, pp. 6429–6443, 2012.
- [85] R. J. Itier, M. Latinus, and M. J. Taylor, "Face, eye and object early processing: what is the face specificity?," *NeuroImage*, vol. 29, no. 2, pp. 667–676, 2006.
- [86] B. Z. Allison and J. A. Pineda, "ERPs evoked by different matrix sizes: implications for a brain computer interface (bci) system," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 11, no. 2, pp. 110–113, 2003.
- [87] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient p300-based brain–computer interface for disabled subjects," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [88] G. Edlinger, C. Holzner, and C. Guger, "A hybrid brain-computer interface for smart home control," in *Human-Computer Interaction. Interaction Techniques and Environments*, pp. 417–426, Springer, 2011.
- [89] A. Combaz, N. Chumerin, N. V. Manyakov, A. Robben, J. A. Suykens, and M. M. Van Hulle, "Towards the detection of error-related potentials and its integration in the context of a P300 speller brain–computer interface," *Neurocomputing*, vol. 80, pp. 73–82, 2012.

- [90] M. Spüler, M. Bensch, S. Kleih, W. Rosenstiel, M. Bogdan, and A. Kübler, "Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI," *Clinical Neurophysiology*, vol. 123, no. 7, pp. 1328–1337, 2012.
- [91] C. Cipriani, M. Controzzi, and M. C. Carrozza, "The smarthand transradial prosthesis," *Journal of NeuroEngineering and Rehabilitation*, vol. 8, no. 1, p. 29, 2011.
- [92] M. Zecca, S. Micera, M. Carrozza, P. Dario, *et al.*, "Control of multifunctional prosthetic hands by processing the electromyographic signal," *Critical Reviews in Biomedical Engineering*, vol. 30, no. 4-6, pp. 459–485, 2002.
- [93] B. Karlik, M. Osman Tokhi, and M. Alci, "A fuzzy clustering neural network architecture for multifunction upper-limb prosthesis," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 11, pp. 1255–1261, 2003.
- [94] S. Bitzer and P. van der Smagt, "Learning EMG control of a robotic hand: Towards active prostheses," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2819–2823, 2006.
- [95] T. Sanger, "Bayesian filtering of myoelectric signals," *Journal of Neurophysiology*, vol. 97, no. 2, pp. 1839–1845, 2007.
- [96] Y. Huang, K. Englehart, B. Hudgins, and A. Chan, "A gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 11, pp. 1801–1811, 2005.
- [97] A. Chan and K. Englehart, "Continuous myoelectric control for powered prostheses using hidden Markov models," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 1, pp. 121–124, 2005.
- [98] F. Tenore, A. Ramos, A. Fahmy, S. Acharya, R. Etienne-Cummings, and N. Thakor, "Decoding of individuated finger movements using surface electromyography," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1427–1434, 2009.
- [99] R. Smith, F. Tenore, D. Huberdeau, R. Etienne-Cummings, and N. Thakor, "Continuous decoding of finger position from surface EMG signals for the control of powered prostheses," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 197–200, 2008.

- [100] R. Smith, D. Huberdeau, F. Tenore, and N. Thakor, "Real-time myoelectric decoding of individual finger movements for a virtual target task," in *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2376–2379, 2009.
- [101] A. d'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature Neuroscience*, vol. 6, no. 3, pp. 300–308, 2003.
- [102] A. d'Avella, A. Portone, L. Fernandez, and F. Lacquaniti, "Control of fast-reaching movements by muscle synergy combinations," *The Journal of Neuroscience*, vol. 26, no. 30, pp. 7791–7810, 2006.
- [103] S. Overduin, A. d'Avella, J. Roh, and E. Bizzi, "Modulation of muscle synergy recruitment in primate grasping," *The Journal of Neuroscience*, vol. 28, no. 4, pp. 880–892, 2008.
- [104] G. Torres-Oviedo and L. Ting, "Muscle synergies characterizing human postural responses," *Journal of Neurophysiology*, vol. 98, no. 4, pp. 2144–2156, 2007.
- [105] A. Daffertshofer, C. J. Lamoth, O. G. Meijer, and P. J. Beek, "PCA in studying coordination and variability: a tutorial," *Clinical Biomechanics*, vol. 19, no. 4, pp. 415–428, 2004.
- [106] C. James and C. Hesse, "Independent component analysis for biomedical signals," *Physiological Measurement*, vol. 26, no. 1, p. R15, 2004.
- [107] D. Seung and L. Lee, "Algorithms for non-negative matrix factorization," *Advances in Neural Information Processing Systems*, vol. 13, pp. 556–562, 2001.
- [108] M. Tresch, V. Cheung, and A. d'Avella, "Matrix factorization algorithms for the identification of muscle synergies: evaluation on simulated and experimental data sets," *Journal of Neurophysiology*, vol. 95, no. 4, pp. 2199–2212, 2006.
- [109] A. Ajiboye and R. Weir, "Muscle synergies as a predictive framework for the EMG patterns of new hand postures," *Journal of Neural Engineering*, vol. 6, no. 3, p. 036004, 2009.
- [110] N. Jiang, K. Englehart, and P. Parker, "Extracting simultaneous and proportional neural control information for multiple-DOF prostheses from the surface electromyographic signal," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1070–1080, 2009.

- [111] C. Choi and J. Kim, “Synergy matrices to estimate fluid wrist movements by surface electromyography,” *Medical Engineering and Physics*, vol. 33, no. 8, pp. 916–923, 2011.
- [112] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, “Algorithms and applications for approximate nonnegative matrix factorization,” *Computational Statistics and Data Analysis*, vol. 52, no. 1, pp. 155–173, 2007.
- [113] D. Tkach, H. Huang, and T. Kuiken, “Study of stability of time-domain features for electromyographic pattern recognition,” *Journal of NeuroEngineering and Rehabilitation*, vol. 7, no. 1, p. 21, 2010.
- [114] A. Phinyomark, C. Limsakul, and P. Phukpattaranont, “A novel feature extraction for robust EMG pattern recognition,” *Journal of Computing*, vol. 1, no. 1, pp. 71–80, 2009.

Publication List

Publications Related to The Dissertation

- [1] Jiaxin Ma, Lichen Shi, and Baoliang Lu, “An EOG-based Vigilance Estimation Method applied for Driver Fatigue Detection”, *Neuroscience and Biomedical Engineering*. In press.
- [2] Jiaxin Ma, Yu Zhang, Andrzej Cichocki, and Fumitoshi Matsuno, “A Novel EOG/EEG Hybrid Human-Machine Interface Adopting Eye Movements and ERPs: Application to Robot Control”, *IEEE Transactions on Biomedical Engineering*. In press.
- [3] Jiaxin Ma, Nitish Thakor, and Matsuno Fumitoshi, “Hand and Wrist Movement Control of Myoelectric Prosthesis Based on Synergy”, *IEEE Transactions on Human-Machine Systems*. In press.
- [4] Jiaxin Ma, Yu Zhang, Yunjun Nam, Andrzej Cichocki, and Fumitoshi Matsuno, “EOG/ERP Hybrid Human-Machine Interface for Robot Control”, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and System*, pp. 859-864, Tokyo, Japan, 2013.
- [5] Haoyu Cai, Jiaxin Ma, Lichen Shi, and Baoliang Lu, “A Novel Method for EOG Features Extraction from the Forehead”, in *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3075-3078, Boston, Massachusetts USA, 2011.
- [6] Jiaxin Ma, Lichen Shi, and Baoliang Lu, “Vigilance Estimation by Using Electrooculographic Features”, in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6591-6594, Buenos Aires, Argentina, 2010.

Publications Unrelated to the Dissertation

- [1] Hirotaka Nakamura, Kazuyuki Kon, Jiaxin Ma, Takahiro Katayama, Atsushi Kojina, Tomohiro Shirakawa, Hitoshi Fujii, Ryusuke Horibe, and Fumitoshi Matsuno, “Control of a walking assistance cart using classification of road condition and human intention”, in *Proceedings of the 58th Annual Conference of the Institute of Systems, Control and Information Engineers*, Kyoto, Japan, 2014. (in Japanese)
- [2] Jiaxin Ma, Jun Zhang, and Jinglu Hu, “Glomerulus Extraction by Using Genetic Algorithm for Edge Patching”, in *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 2474-2479, Trondheim, Norway, 2009.