Robotica (2016) volume 34, pp. 1209–1226. © Cambridge University Press 2014
doi:10.1017/S0263574714002197

Emotion detection for wheelchair navigation enhancement

Hachem A. Lamti†,*, Mohamed Moncef Ben Khelifa‡, Adel M. Alimi† and Philippe Gorce‡

† *Research Group on Intelligent Machine (REGIM) Laboratory, National School of Engineers in Sfax, Sfax, Tunisia*
‡*Ingénierie des Handicaps et de la bio-modélisation (HANDIBIO) Laboratory, South University, Toulon-Var, France*

(Accepted July 22, 2014. First published online: August 15, 2014)

SUMMARY
The goal of this study is to investigate the use of emotion as a braking system for wheelchair navigation. In the first part emotion is detected based on ElectroEncephalography (EEG) technology and emotion induction experiments. Using different techniques for features extraction (Welch and Wavelets), selection (Principal Component Analysis (PCA) and Genetic Algorithm (GA)) and classification (Support Vector Machine (SVM), Multi Layer Perceptron (MLP) and Linear Discriminate Analysis (LDA)), the best combination was assigned to Wavelets-GA-MLP. In the second part, in order to validate the impact of emotion as velocity modulator, a comparison between emotion-based and non emotion-based wheelchair navigation scenarios in a simulated environment was conducted. The assessment was based on four parameters: obstacles hit, navigation path, execution time and outbound points of gaze (POG). While the first two emotion introductions showed better results, this was not the case for the third. These findings can be utilized in order to prescribe a suitable wheelchair according to the subject pathology.

KEYWORDS: Emotion; Wheelchair navigation; EEG; User performance.

1. Introduction

Disabled or elderly persons with heavily reduced physical and/or mental abilities find it hardly possible to control a powered wheelchair using a conventional joystick. For these persons, shared paradigms were introduced to enhance navigation security; it consists of giving the user more or less control on a need basis.[1] Many studies incorporated this paradigm in order to conceive a suitable wheelchair according to the subject pathology. Vander Poorten *et al*.[2] established a bilateral communication channel between the wheelchair controller and the user based on haptic feedback; the local map of the wheelchair environment was built through the on-board sensors. This leads to the introduction of haptic collision avoidance and haptic obstacle avoidance algorithms that help the user to successfully maneuver the wheelchair backwards inside an elevator. Urdiales *et al*.[3] proposed the construction of a wheelchair based on the skills of the navigation profile.

They proposed a new method to extract a prototype user profile, using real traces based on more than 70 volunteers with different physical and cognitive skills, to determine the average behavior expected from the wheelchair user in order to cope with real situations. It was successfully tested on 18 volunteers affected by left and right brain strokes. Some other studies, such as that of Ren *et al*.,[4] tried to introduce a map matching based on Global Positioning Systems (GPS) in order to enhance wheelchair navigation. They also proposed an alternative to deal with difficulties faced by GPS during navigation, such as poor satellite availability, by introducing a fuzzy logic-based algorithm to perform matching wheelchair movements on sidewalks. In our pilot study,[5] a new approach based

* Corresponding author. E-mail: lamtihachem@gmail.com

on combining the user's gaze and his mental activities was introduced in order to assess wheelchair navigation performance in comparison with a standard gaze-based navigation. The results showed that the system performance was better using the combined modalities.

These projects centered the navigation enhancement on modifying the wheelchair system by either adding new on-board sensors such as GPS, cameras, etc., or by assessing user's performance by means of motor activities, such as haptic feedback, which are still centered on the wheelchair system and holds a posteriori reaction; the wheelchair corrective behavior is generated after the user commits an error during navigation.

The Brain Eyes WHEELchair Interface (BEWHEELI) project aims at proposing a new alternative for severely disabled people (example palsy and Locked-In patients) to command their wheelchairs. It's built on two major blocks: the command block, which ensures the migration from the use of joystick to the gaze/brain automated command (including the intermediate phase of gaze/joystick command) and the security block which deals with the wheelchair control by assessing human factors.

In this paper, we will focus on a security block that has an anticipatory aspect; through human factor, the wheelchair anticipates the next action to do. To the best of our knowledge, the proposal of anticipatory human factor-based wheelchair navigation is uncommon. In fact, during the exchange with experts, doctors, occupational therapist and psychologists, many human factors could influence wheelchair navigation, such as mental fatigue and emotions. In the current study, emotions are investigated and measured through mental activity.

From a clinical perspective,[25] severely disabled people are exposed to several forms of negative emotions that can influence their navigation performance. There are three common forms. The first is stress, resulting from a feeling of fear that follows the dysfunction of assisting systems. The next is excitement which must be treated with care as the subject can easily overflow the limits of normal excitement. And finally nervousness, an obsession with the idea of fast recovery of motor functionalities.

Being aware of the importance of these emotions, the main goal of this study is to setup an emotion detection module that enables the wheelchair system to account for the user's state. This information is then implemented to decrease the wheelchair velocity and assist the user during his navigation until he reaches his normal state (either a neutral or positive emotion such as a feeling of relaxation).

This manuscript is divided into three major parts. In the first part, scientific research projects related to brain computer interfaces (BCI) and emotions are exposed. In the second part, emotion induction experiment, techniques used for extraction, selection and detection of features are explained with a comparison of different techniques used. In the third part, the integration of emotions as a braking system, as well as the experimentation employed to assess navigation performance is described with results and discussion of different measures found. Finally, a short conclusion will introduce the next step for the project: the investigation of mental fatigue.

2. Overview

2.1. Brain computer interface

A BCI is a system that assists persons in communication with the external world by thought without relying on muscular or nervous activity.[6] A functional model of a BCI starts by monitoring the user's brain activity which is conveyed into brain signals processed to obtain features grouped into vector called "feature vector." The mapping of the latter results in commands to be executed by the system that displays feedback to the user in order to fine tune or modulate his brain activities.

In an EEG, the subjects mental activities appear as distinctive patterns. Recognized by BCI, these are mapped into commands, which are associated with certain actions. In scientific literature, five standard EEG bands are defined, mainly: delta (up to 4 Hz), theta (4 Hz–8 Hz), alpha (8 Hz–13 Hz) beta (13 Hz–30 Hz) and gamma (30 Hz–100 Hz). Each of these are prominent in different situations (e.g., delta band is more prominent when a person is sleepy, alpha when he is conscious, etc.). The current field of applications of EEG-based BCI systems vary widely; from wheelchair navigation,[7] evaluation of Brain-Computer Interface to categorize human emotional response,[8] assessment of cognitive loads[9] to neurofeedback training for children with attention deficit disorders.[10]

2.2. Emotion

Emotions consist of multiple components that may include intentions, action tendencies, appraisals, other cognitions, central and peripheral changes in physiology, and subjective feelings. Emotions are not directly observable, but are inferred from expressive behavior, self-report, physiological indicators, and context.

Emotions are classified into discrete and dimensional models. Discrete emotions could be divided into basic set (core set) and secondary set depending on the action tendencies.[11,12] Some other theories argue that emotions could be better measured as differing in degree on dimensions. The well-known dimensional models are the pleasure-arousal (pleasantness within a given emotional state and arousal for physiological activation)[13] and approach-avoidance (tendencies to approach or avoid the stimuli).[14] Note that the pleasure-arousal dimension is used in this article in line with several different research projects.

To elicit emotions, many standard strategies such as films, pictures, sound and odors were utilized. Other methods such as autobiographic recall and social interactions were also used. These suggested that the behavior of others would affect our own emotional state. In order to measure emotions, the methods used are either objective or subjective. The subjective methods consist of questionnaire, adjective checklists and pictorial tools such as Activation-Deactivation Check List (AD-ACL), Positive and Negative Affect Schedule (PANAS) and Self Assessment Mankini (SAM). Objective methods use physiological manifestations derived from emotions, such as brainwave activity, heart rate, facial expression, vocal properties, and others. In this article, an emotion induction technique based on video sequence visualization was utilized. This will be explained in the next section.

2.3. System framework

The system framework (Fig. 1) involves two major parts: the inputs and the outputs. The inputs are formed by raw signals issued from an EEG, ElectroOcculoGraphy (EOG) and a joystick. These are filtered, sampled and processed depending on the modality used. In the current study, we focus most on EEG activity in order to extract emotions. The methodology used for emotion detection is illustrated in Fig. 2. The features extracted from bandwave signals are the Power mean (Pm) and the Root mean square (R_{ms}). Two different techniques have been adopted: the Welch method and wavelets. With the huge quantity of data in the extracted features due to the crossing of many parameters (bandwave per time unit per sensor per subject) a selection phase is needed. For this purpose, Principal Component Analysis (PCA) and Genetic Algorithm (GA) were applied. At the classification phase, three techniques were compared: Linear Discriminate Analysis (LDA), Multi Layer Perceptron (MLP) and Support Vector Machine (SVM). Before the application of different techniques on the raw data, a reference database was setup containing a temporal profile of each studied emotion class for different subject samples. We assume that the studied emotions are located in separated quadrants in the valence-arousal model. In this case, each combination of valence and arousal rates results in one emotion class. The second hypothesis assumes that the cerebral activity during visualization period matches the given rating in the valence-arousal scale. In this way, one can attribute a temporal profile to each chosen emotion.

Meanwhile, the Points of Gaze (POG) extracted from EOG camera were processed and classified to estimate the user's direction choice: left or right. A calibration algorithm based on nine-point visualization is used to estimate POG. These are used as inputs in the classification algorithm, which splits the screen into three quadrants (Right/Idle/Left) by calculating the maximum fluctuation of the eye gaze. Finally, right and left commands are sent to the Reference Velocity Calculation module. If the navigation direction is controlled by the user's gaze, forward and backward commands are calculated according to the joystick angle. In fact, an embedded cursor can convert the joystick angle to an analogical signal needed for the wheelchair wheels to move forward or backward.

The three described signals constitute the inputs of the Reference Velocity Bloc. The latter uses a special algorithm to calculate reference angular and linear velocity based on the rotation angle, raw linear velocity and the user's emotion. This algorithm will be detailed later, as it takes the emotion factor into consideration. If the latter is detected, the braking command is sent to the raw angular and linear velocities so that it decreases slightly. After obtaining the reference velocities, they are compared to the encoders recorded velocities mounted on the wheels, and processed by an analog

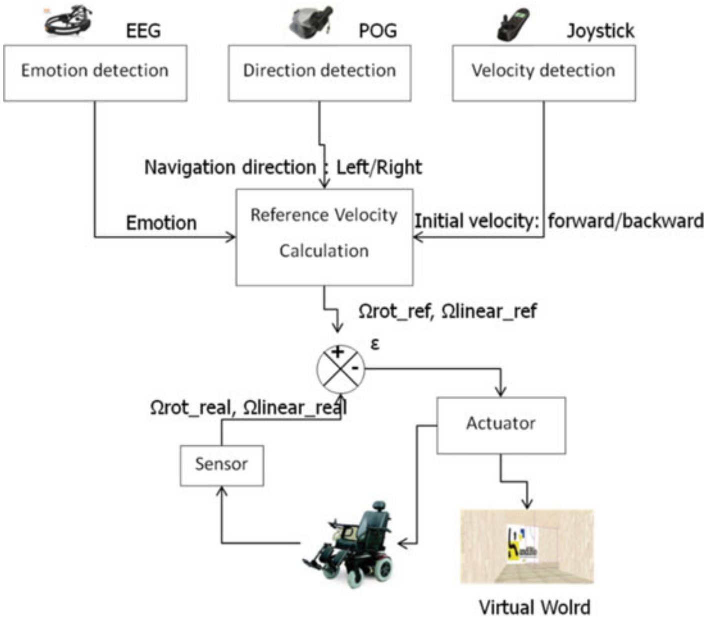


Fig. 1. System framework: the illustration of the use of emotions as velocity modulator.

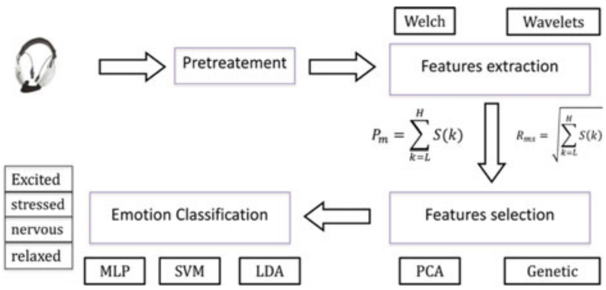


Fig. 2. Emotion detection methodology.

digital converter. After correction, they are transmitted to the wheelchair actuators in parallel to the virtual world projection.

2.4. *Experimental setup for emotion detection*

Our experiment consists of measuring the emotional states of the persons when viewing and listening to audio/video sequences. The recorded data will supply a reference database in order to assign to each emotion class a temporal EEG data profile, and thus applying different techniques offline before implementing an online emotion detection module.

Consequently, the authenticity of the collected emotions must be checked. This will be done by comparing emotions distribution with the valence/arousal common distribution, which is considered as our ground-truth. In this case, the association between signal features and emotion is deduced from the correlation between the EEG activity each second, with the rate given by the subject at the end of the experiment.

2.4.1. Participants. 40 healthy subjects (50% female, 50% male) participated in the experience. They were aged from 22 to 55 years old. Their vision was normal or corrected to be normal. Before the implementation of the experimental equipment, the subjects signed a consent form and answered a questionnaire to analyze their normal state[15] and verify if they consumed medicines or stimulating drinks such as coffee or tea.

2.4.2. Procedure. Once in the experimentation room, the experimenter explained to the participant that this study aims to measure his emotional state in regard to the audio/video sequences. Each

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

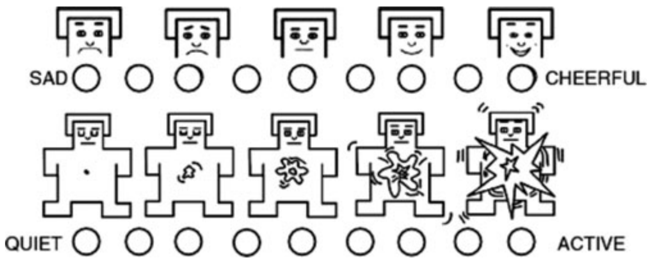


Fig. 3. SAM-Scales for valence and arousal.

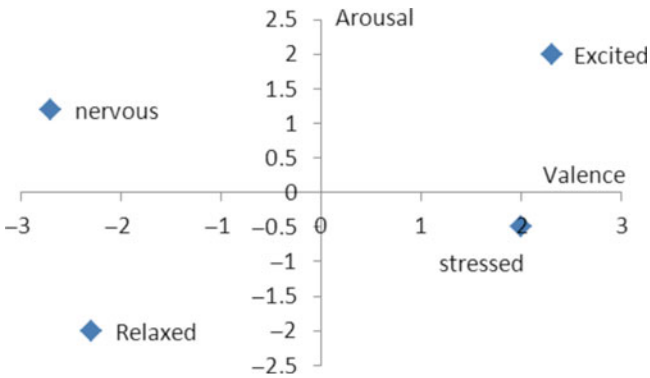


Fig. 4. mean ratings distribution according to emotion.

experience is based on the following steps. In the first stage, the experimenter asks the participant to complete a form containing his personal information (gender, age, left/right handed, alcohol/coffee/tea consumption, etc). Several studies[16] have shown the influence of sex and age on the emotional state of the subject. In the second stage, the experimenter asks the participant to choose four audio/video sequences to measure his emotional state in different situations such as excitement, nervousness, relaxation and stress. This is justified by the fact that the goal of emotion induction in this study is not to assess the stimulus on itself (as it was done to setup the IAPS system[26]) but rather to record the needed emotion independently of the used stimuli. To validate the results, emotions aroused by these tests are compared with reference emotions on the SAM scale. After the preparation of the chosen sequences, the experimenter sets the EEG headgear on the scalp of the participant. Before beginning the recording, he must be sure of the good quality of the emitted signals. In the third stage, the participant listens to the four sequences one by one.

The duration of each sequence is limited to 63 s with the first three seconds as baseline. There is a one-minute break before the beginning of the next session. After every test, the experimenter asks the subject to self-assess his emotional state from the dimensional model consisting of two manikin scales (Self Assessment Manikin, SAM) representing the two dimensions of arousal and valence.[17] The dimension of valence ranges from two poles; negative/bad and positive/good, whereas the arousal dimension spans between the two poles sleepy/calm for very low arousal and aroused/excited for very high arousal. In the third stage the participant should fill in each of the two scales, and indicate his level of valence and arousal by choosing a number between one and nine. (Fig. 3) depicts the two SAM scales.

2.4.3. Subjective rating analysis. Stimuli were selected to induce emotions in the four quadrants of the valence-arousal space (Low Arousal Low Valence, High Arousal Low Valence, Low Arousal High Valence, and High Arousal High Valence). (Fig. 4) summarizes the distribution of the mean ratings of different users according to the induced emotion. The choice of emotion was made intentionally in a way that each studied emotion belongs to only one quadrant (the combination of valence and arousal rates results in only one emotion class). The purpose is to avoid the overlap between emotions on the same quadrant. It can be seen that the results are well within those found in other scientific literature.

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

2.5. Feature extraction

According to the ratings, EEG signals were processed in order to extract the mean power (P_m) and the root mean square (R_{ms}) of five frequency bands (delta, theta, alpha, beta, gamma) in each second. In this case, each dataset is formed by two eventual configurations; either the P_m per band per sensor per user per second or the R_{ms} per band per sensor per user per second which results in a 5 bands × 14 sensors × 63 s × 40 subjects. The extraction step is applied using two methods: Welch method and discrete wavelets transform (DWT).

2.5.1. Welch method. Welch's method is a combination of the standard periodogram and Bartlett[18] methods. Let it be $x_i(n), i = 1, 2, \ldots, K$, K uncorrelated realizations of an aleatory process $x(n)$ over an interval $0 \le n \le L$. It can be directly expressed by:

$$\hat{S}_W(w) = \frac{1}{KLU} \sum_{i=0}^{K-1} \left| \sum_{n=0}^{L-1} w(n)\, x(n + iD) \exp(-jwn) \right|^2 \quad (1)$$

where $U = \frac{1}{N} \sum_{n=0}^{N-1} |w(n)|^2$, N is the length of the window $w(n)$, It has been proven that this method reduces the noise in power spectra whilst the resolution R depends on the chosen window and can be given by

$$R = \frac{1}{LT_s} \quad (2)$$

where T_s is the sampling period. Hence the lower L is the smoother Welch periodogram becomes. The initial signal was filtered between 1 and 64 Hz. In this case, delta (up to 4 Hz), theta (4 Hz–8 Hz), alpha (8 Hz–13 Hz) beta (13 Hz–30 Hz) and gamma (30 Hz–64 Hz) were kept for our study. The power spectral density (PSD) was computed every one second of the trial for each user. The Welch periodogram was computed using 512 points FFT and several Hamming windows of length 128, 64, 32, and 16 points with a 50% overlapping. For the bands found in each second, two parameters were computed: the mean Power (P_m) and the Root Mean Square (R_{ms}) calculated as:

$$P_m = \sum_{k=l}^{h} S(k) \quad (3)$$

$$R_{ms} = \sqrt{\sum_{k=l}^{h} S(k)} \quad (4)$$

where $S(k)$ are the sampled values of the periodogram $S(f)$ and l and h are the indexes of the higher and lower sampled frequencies for each band.

2.5.2. Discrete wavelets transform. The DWT is generated by the function:

$$\psi_{(a,b)}(t) = 2^{\frac{a}{2}} \psi\left(2^{\frac{a}{2}}(t - b)\right) \quad (5)$$

where a is called the scales and b the shifts, it is possible to approximate any function by dilating $\psi(t)$ with the coefficient 2^k and shifting the resulting function on interval proportional to 2^{-k}.

Using Debauches 8 wavelet, six levels of DWT are implemented. The first level (between 128 Hz and 64 Hz) contains noise as the needed waves only occur between 1 Hz and 64 Hz. In the second level (32 Hz and 64 Hz), Gamma waves can be detected as decomposition detail, then Beta waves (16 Hz and 32 Hz) and so on until reaching the sixth level (between 1 Hz and 4 Hz) where Delta waves constitute the approximation coefficients of the signal.

2.6. Feature selection

The feature vector contains two configurations with P_m and R_{ms} parameters of 5 band waves in the 14 sensors for the 63 s. This means that the feature vector suffers from 'high dimensionality curse'.[19]

Consequently, features must be selected according to its correlation proportionality with each other. In this section two selection techniques are used: PCA and GA.

2.6.1. Principal component analysis. PCA uses mathematical principals to transform a number of correlated variables into smaller numbers called principal components. For each second, the feature vector X is 200×80 length where rows $(x_i^1, \ldots, x_i^{80})$ represent the observations and columns $(x_1^j, \ldots, x_{200}^j)$ represent the variables.

From the correlation Matrix, the Eigen values and Eigen vectors are calculated to form principal components. The selection of the number of principal components axis to be kept is very important in PCA. The criteria used are generally empirical: the elbow method consists of detecting an elbow in the Eigen values plot or Kaiser Criterion. The latter consists on retaining only the Eigen values greater than the mean value.

2.6.2. Genetic algorithm. GA, a form of learning strategy, is an adaptive search technique which has proven its efficiency over a set of search methods.[20] Let it be X, the feature vector of length s. the inclusion or elimination of the corresponding feature is described as follow: if Xi = 0, then it represents elimination of the feature, otherwise 1 indicates its inclusion. Evaluation function choice is very important to obtain a successful application of GA. In the current problem, our purpose is to estimate the number of optimal features to be selected. The idea is to apply a fitness function on the correlation matrix M. The latter can be presented as follows:

$$M = \begin{pmatrix} 1 & \cdots & C_{i,j} \\ \vdots & \ddots & \vdots \\ C_{j,i} & \cdots & 1 \end{pmatrix} \tag{6}$$

where $C_{i,j}$ represents the correlation coefficient between feature i and feature j, the proposed fitness function F could be defined:

$$F = \min |M(i, j)| \tag{7}$$

where $M(i,j) = C_{i,j}$. For each iteration, the crossing chromosomes with the least correlation coefficient are kept for the next generations while the others are eliminated.

2.7. Classification

After selection of features, the classification was investigated using different classifiers; the LDA, the MLP and SVM. In a supervised learning environment, given the feature vector as input, the output could be one of the four studied emotions (stressed, excited, nervous, and relaxed).

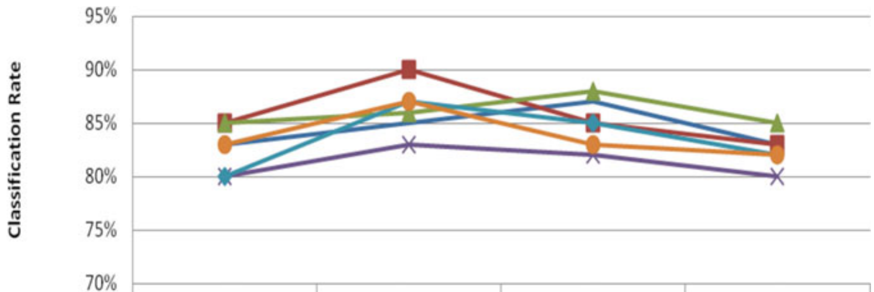
2.7.1. Linear discriminate analysis (LDA). The LDA is a linear combination of variables. They are presented in the form of:

$$y_{\mathrm{km}} = u_0 + u_1 X_{1\,\mathrm{km}} + u_2 X_{2\,\mathrm{km}} + \ldots + u_p X_{p\,\mathrm{km}} \tag{8}$$

where y_{km} is the value of the discriminate function for the case m on the group k as well as for $X_{i\,\mathrm{km}}$ which is the discriminate variable X_i for the case m on the group k, and u_i are the required coefficients. This implies that the number of discriminate functions is determined by the number of considered groups.

2.7.2. Multi layer perceptron (MLP). The used MLP is composed of an input layer with a variable size, the selected features of the input vector, a hidden layer with 20 neurons and an output layer with 4 neurons, which correspond to the number of emotions. The transfer function used is sigmoid and the cross-validation technique adopted is the test set validation technique; the database is divided into 3 sets: 70% for training, 15% for testing, and 15% for validation (and thus avoiding over-fitting).

2.7.3. Support vector machine (SVM). SVM maps input vectors into higher dimensional space to ease classification, then it finds a linear separation with the maximal margin in the new space. It



	128	64	32	16
Pm_MLP	83%	85%	87%	83%
Rms_MLP	85%	90%	85%	83%
Pm_LDA	85%	86%	88%	85%
Rms_LDA	80%	83%	82%	80%
Pm_SVM	80%	87%	85%	82%
Rms_SVM	83%	87%	83%	82%

Fig. 5. Classification performance using PCA and Welch periodogram, for different window lengths.

requires the solution of the following problem:

$$\begin{gathered} \min_{w,b,\varepsilon} \tfrac{1}{2} w^T w + C \sum_{i=1}^{l} \varepsilon_i \\ \text{subject to } y_i \left(w^T \vartheta\left(x_i\right) + b \right) \geq 1 - \varepsilon_i \\ \varepsilon_i \geq 0 \end{gathered} \tag{9}$$

where C is the penalty parameter of the error ε_i. The kernel used in this article is the Gaussian radial basis function. This could be expressed as follows:

$$K\left(x_i, x_j\right) = e^{(-\gamma x_i - x_j^2)} \text{ for } \gamma > 0 \tag{10}$$

A cross-validation technique is used to determine C and γ.

2.7.4. Results. The results found in this work are associated with many different combinations made from the extraction process (Welch and wavelets), feature selection (GA and Principal Component Analysis), and classification techniques (MLP, LDA, and Support Vector Machine). In order to reliably report results, we report the combined F-score (in percentage) based on the precision and recall measures. This measure takes into consideration the class balance and is commonly employed in information retrieval.[21]

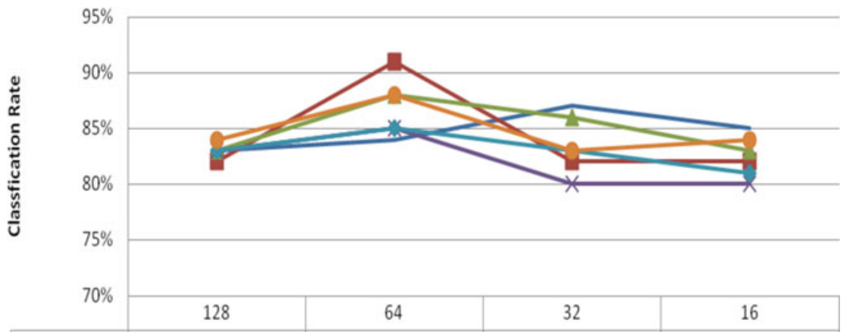
For the Welch technique, R_{ms} and P_m were compared in different window lengths (128 points, 64 points, 32 points and 16 points) using PCA and GA as features selectors. The results can be recapitulated in Figs. 5 and 6.

The figures show that classification accuracy, based on the Welch periodogram, depends on the window length chosen for extraction. The plots slopes increase from 128 to 64 and start dropping to reach the worst classification rate for 16 points. There is no noticeable difference between P_m and R_{ms} but a slight superiority is recorded for R_{ms} with the best classification for Rms_MLP64 with 90%.

MLP showed the best classification rates over LDA and SVM, while the latter also showed good performance. Also, the difference between PCA and GA is not significant although GA is the best overall classification as it is recorded for Rms_MLP64 with GA as feature selector.

This can be justified by the subjective criteria of PCA; in fact, despite the fact that there are a lot of techniques to pick the number of eigen values to keep, these can remain insufficient to select the right number of significant features, while for GA, the iterations take a lot of calculation time to achieve

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press



	128	64	32	16
Pm_MLP	83%	84%	87%	85%
Rms_MLP	82%	91%	82%	82%
Pm_LDA	83%	88%	86%	83%
Rms_LDA	83%	85%	80%	80%
Pm_SVM	83%	85%	83%	81%
Rms_SVM	84%	88%	83%	84%

Fig. 6. Classification performance using GA and Welch periodogram, for different window lengths.

the best selection even if the selected features seem to be more suitable than PCA. The results found using wavelets extraction, are recapitulated in Fig. 7.

The same already-mentioned notes are also noticed for the wavelets case. R_{ms} achieved the best classification rate over Pm. GA is more efficient than PCA. The best classification rate recorded is 93% assigned to MLP_Rms which is considered to be the best performance for all extraction, selection and classification techniques. The eventual explanation for the 7% error rate could be the fact that the extraction process was applied for each second of the 63 s of the whole experiment, which is not necessarily correct. While visualizing a video sequence, we cannot ensure that the same feeling is expressed during the whole experiment, it may happen for only a few seconds, not the 63 s. In our case, we assumed that each second of the experiment is associated with the expressed rating at the end of the experiment.

In motor imagery, by asking the user to imagine making a muscular activity, we are sure for the next two seconds that sensors are recording EEG patterns related to that action, while in emotion detection it is not guaranteed as it depends on many parameters such as the influence of the video sequence on the user, the time needed to switch between emotions and the fact that EEG fluctuations do not necessarily mean an emotion trigger. However, the classification rate obtained is sufficient to validate the presented approach of this project.

3. Implementation

In the previous section, emotion was detected using Wavelets as the extractor, GA as the feature selector and MLP as the classification technique. The goal of this article is to assess the influence of the introduction of emotions on navigation safety and the braking system. An experiment was set up to compare two navigation scenarios (Fig. 8).

The first scenario consists of controlling a wheelchair in a virtual environment going from point A to point G without taking user emotions into consideration. In the second scenario, if a negative emotion is detected (stress, excitement and/or nervousness), a braking command is sent to the wheelchair and to the virtual environment so it decreases its velocity as a preventive measure. If the detected emotion has not changed for a certain period of time, the wheelchair stops and a message instructing the user to concentrate and relax is displayed. The final goal of the project is to control the wheelchair using gaze and brain waves rather than a joystick. In this article, the first step taken was to generate left/right commands using gaze and forward/backward with a joystick. The studied parameters for assessment are: obstacles hit, navigation path, execution time and outbound POG, which will be explained in detail later.

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

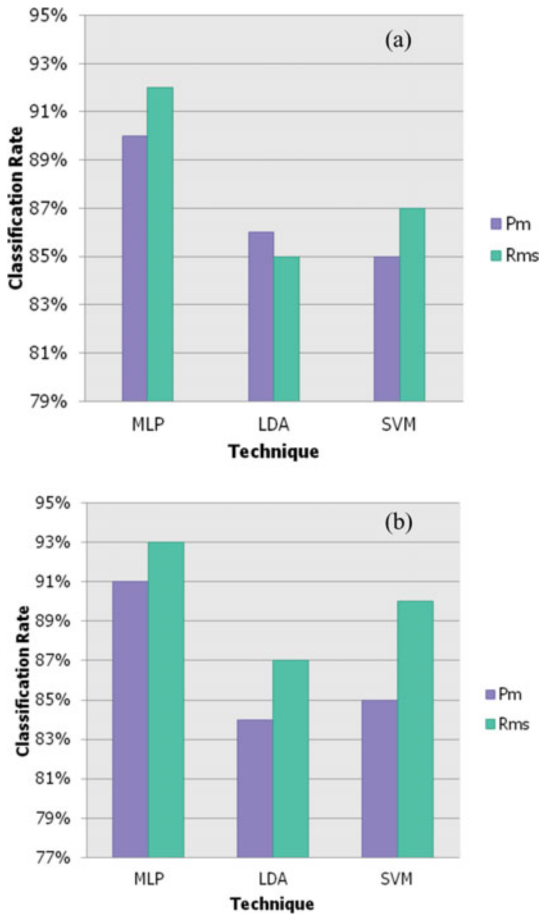


Fig. 7. Classification performance using Wavelets-PCA (a), classification performance using Wavelets-GA (b).

Fig. 8. Virtual world wheelchair navigation.

3.1. *Experimental setup*

The experimental environment consists of many components; these can be divided into hardware and software.

3.1.1. Hardware framework. As the experiment targets wheelchair navigation, the use of a wheelchair is crucial. For this purpose an Invacare Storm 3G Ranger *X* branded wheelchair is used. Equipped with a joystick, encoders were added to its wheels so the wheelchair velocity can be digitized and

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

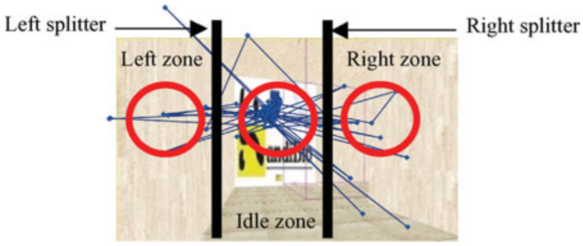


Fig. 9. Example of implementation of the min-right/max-left algorithm.

treated. This can be useful to control a virtual world projected on a 180 degrees panoramic screen to help the immersion of the user in the world. To calculate POG, an ASL EyeTrac 6 branded eye tracker was placed in front of the user. A specific algorithm was used for system calibration and for dividing the screen into command zones. Alternatively, the Emotiv Epoc headgear is added to estimate emotions.

3.1.2. Virtual world. The virtual world was developed using the Reality Factory Engine.[22] It consists of a virtual maze formed by many rooms labeled from A to G and a hallway. Each room has the special features needed to induce frustration within the user.

Room A is the starting point for the navigation. Room B, contains some obstacles, but not many. With more obstacles, room C leads to rooms D and E. The first is filled with coronas that generate a high illumination effect in order to disturb the user to the point that they find the door to navigate to the hallway. The second is a very dark room filled with obstacles and only illuminated with light coming from a torch. Room F is the most difficult because it contains nine moving obstacles spread all over the room. Reaching room G is the final goal.

The gaze calibration was measured using a nine-point calibration algorithm. For each iteration, a special point is projected on the screen and the user has to look at it. Meanwhile, the system estimates the corresponding POG based on the feature vector defined as the distance between the central point of the corneal reflection and the central point of the pupil reflection. After the success of the operation, the screen is divided into command zones. These are respectively left, right and idle. A min-right/max-left algorithm was implemented for this purpose. The idea is to ask the user to look straight ahead at the screen at the gaze concentration point. The maximum point position and the minimum are recorded. Then, while looking to the right, the minimum point is recorded. The same operation is then processed, using the maximum point position for looking left. The zone separations are calculated from points as follows:

$$X_{\text{zone\_left}t=\frac{X_{\text{point\_center}}-X_{\text{point\_max\_left}}}{2}} \tag{11}$$

$$X_{\text{zone\_right}=\frac{X_{\text{point\_min\_right}}-X_{\text{point\_center}}}{2}} \tag{12}$$

where $X_{\text{point\_center}}$ is the X coordinate of the center point of the POG when the user looks straight. $X_{\text{point\_max\_left}}$ is the X coordinate of the maximum point of the gaze concentration when the user looks left and $X_{\text{point\_min\_right}}$ is the X coordinate of the minimum point of the gaze concentration when the user looks right. An example is presented in the Fig. 9. While it is undeniable that a wheelchair is considered to be a wheeled robot, it does have some important differences. In fact, because a robot is provided with a communication protocol, the latter manages the whole process in order to convert commands into velocities. This is not necessarily the case for a wheelchair, which is considered to be an analogical device. In a real wheelchair, the user has to move the joystick forward to accelerate, and backward to decelerate. In the virtual world, as it's basically a video game, to move forward or backward or to turn left or right the user has to push keyboard keys.

The idea is to link both systems by simulating keyboard keys in the virtual world each time the user moves the joystick forward or backward. The same idea is applied for turning left and right based on the user's POG and command zone detection.

Table I. Emotion detection frequency during experiments (H = Healthy subject, P = Palsy subject).

	Frequency									
Emotion	H1	H2	H3	H4	H5	H6	H7	H8	P1	P2
Relaxation	10	8	9	7	7	6	9	6	3	2
Stress	0	0	0	2	0	4	0	3	7	8
Nervousness	0	2	1	4	3	0	0	3	6	5
Excitement	2	1	0	1	1	1	2	1	2	0

3.2. Procedure

After sitting comfortably in the wheelchair, the calibration process starts by asking the user to look at the nine points of the screen to estimate his POG. After detecting the command zones, the scene is projected on the panoramic screen. The way to command the wheelchair in the environment is explained and the plan of the maze is displayed. The user is asked to navigate from the room A to room G following three different paths or scenarios: the first path, which is the simplest one, consists of going directly from A to G through the hallway. In the second path, the user has to go through B, C, D, hallway and G. The most difficult path consists of going through B, C, E, F, hallway and G. This concludes the first trial. In the second trial, which occurs after a week in order to inhibit the learning effect, the user has to perform the same tasks but with the introduction of emotion detection through the EEG headgear. In this case, the wheelchair velocity depends not only on the joystick and gaze as inputs but becomes moderated through emotions.

3.3. Results

10 subjects (8 healthy and 2 palsy) took part in the experiment.

The frequency of each detected emotion for all subjects were as follows in Table I.

A detection frequency is defined as the number of times the detection algorithm is triggered by the specific emotion. Notice the difference between healthy and palsied users, especially the average relaxation (7 for healthy and 2.5 for palsy) and in accordance, the increase of stress (7.5 palsy and 1 for healthy) and nervousness (5.5 for palsy and 1.5 for healthy) levels. In their ratings, palsied subjects reported being scared from the virtual maze especially when facing moving obstacles and driving into dark rooms. This can justify the high average obtained for stress and nervousness. Also, P9 reported his familiarity with video games. Consequently, better performance was noticed in comparison with P10; especially in relaxation frequency and excitement level.

The performances of the subjects were studied depending on four essential parameters: number of obstacles hit, navigation path deviation, execution time and outbound POG. The first one was calculated based on the obstacles that the user hit when he was navigating from A to G. For the navigation path deviation, an optimal path for the three scenarios was calculated using a path finder algorithm. Then, the distance between points of the followed path in each scenario and the optimal path are calculated. Finally, the standard deviation from the recorded distance points were calculated and reported. In addition, execution times for both trials were exposed in order to better assess the tradeoff of the applicability of emotion integration in navigation per scenario/trial.

At the end, the unbound POG were calculated. During the experiment the user's POG was recorded. Due to distracting situations such as searching within the environment, lack of concentration and emotional disturbance, the user was able to look outside the bounds delimited by the command zones already calculated. This parameter can help us to study the user's psychological state while navigating in the controlled environment.[23] If the number of outbound points is important, this means that the user was, in the major part of the navigation, uncomfortable. The combination of those parameters can help us to decide if the introduction of emotion is viable or not. To be noted, the discussed results assess healthy users and palsied ones separately. It is assumed that the duration of one-week separation between the emotion-based and non-emotion-based navigations is enough to exclude the learning effect.

3.3.1. Obstacles hit. The mean number of obstacles hit by all subjects in the two trials (with and without emotion) for the three scenarios (easy, medium, and hard) are reported in the Figs. 10 and 11.

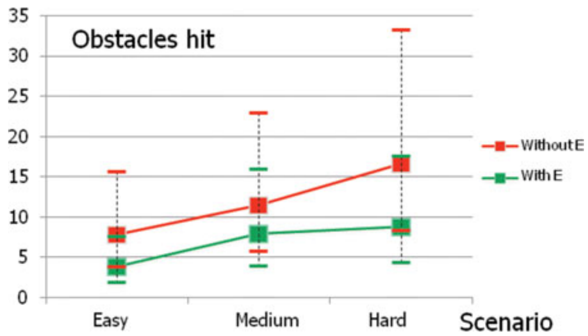


Fig. 10. Number of obstacles hit in during the two trials for healthy subjects.

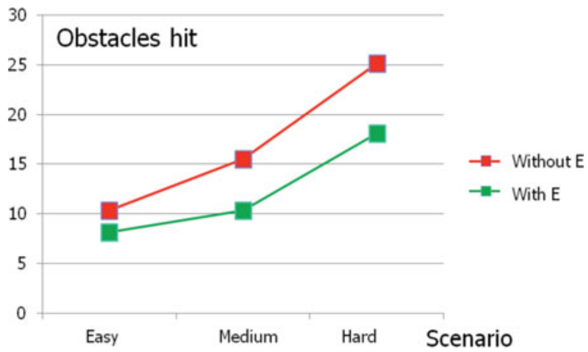


Fig. 11. Number of obstacles hit in during the two trials for palsy subjects.

For healthy subjects, the number of obstacles hit increases in accordance to the difficulty of the scenario. It reaches its maximum in the third scenario, where the moving obstacles established in the room F are very difficult to avoid. This can be deduced from the variance that increases considerably in the third scenario. While in the first scenario, even if the path to follow is very easy, the user finds it difficult to master the wheelchair navigation; especially with the introduction of the gaze to turn left and right. The difference resides in the way the trials slopes increase. In the first trial, the slope increases considerably, especially between the second and the third scenario. While in the second trial the slope is logarithmic and the number of obstacles hit in the second and the third scenarios are very close. The overall obstacles hit dropped from the first trial to the second by half.

These results are not remarkable for the palsied users; in fact, the two curves have the same shape, especially in the third scenario, where the number of obstacles hit is very high in the two cases. Even though, in case of emotion-based navigation, the number of obstacles hit is lower than in non-emotion-based navigation.

3.3.2. Navigation path. During the experimental sessions, the wheelchair navigation path is recorded and then compared to an optimal path calculated using a pathfinder algorithm. It is to be noted that the comparison is based only on path shape and no navigation time is taken into consideration. An example of navigation trajectory for the two trials in the three scenarios is shown in Figs. 12(a)–(c).

Notice that the harder the scenario is, the greater the deviation from the optimal path. This could be explained by the fact that the subject may be confused by the obstacles he faces during navigation and loses the control of his wheelchair or the path he had intended to follow. The distance between each point of the followed path and the optimal one is calculated for the two trials. After this, the standard deviation is calculated for each user, and finally the mean of all standard deviations are reported in Figs. 13 and 14.

The deviation from the optimal path increases considerably going from the easiest scenario to the hardest one. It reaches its maximum in the third scenario. Again, the moving obstacles cause a lot of problems in maintaining a straight trajectory. It may deviate the wheelchair from its initial path, but the same conclusion could be established when comparing both trials: the deviation rate dropped by half and the trials slopes are very different. It should also be noted that in the second trial the

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

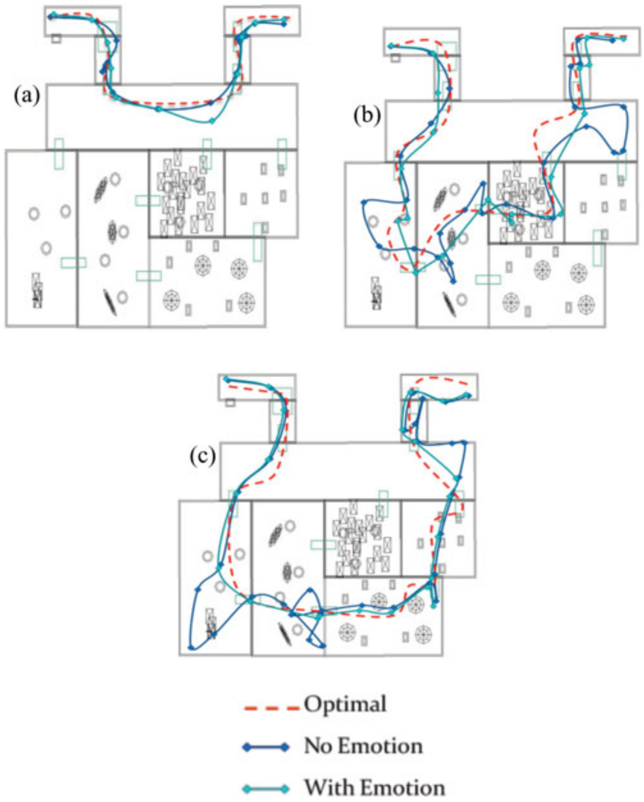


Fig. 12. Navigation path for the three scenarios: easy (a), medium (b) and hard (c).

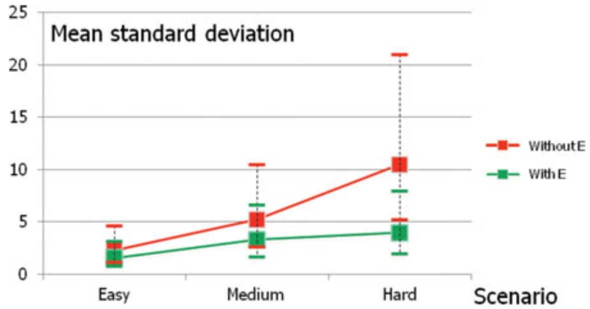


Fig. 13. Mean standard deviations in the two trials for the three scenarios for healthy users.

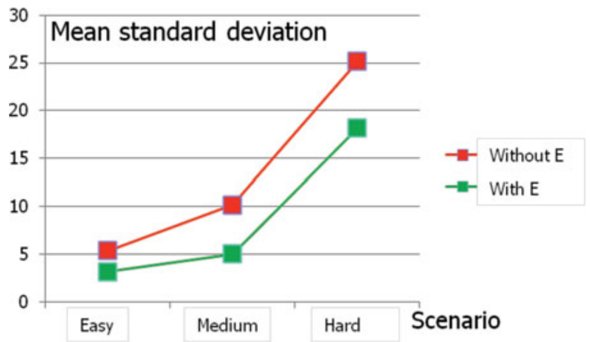


Fig. 14. Mean standard deviations in the two trials for the three scenarios for palsy users.

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

Table II. Execution time (in seconds) per trial/scenario for all subjects (H = Healthy subject, P = Palsy subject).

		Execution time (in seconds)									
Trial	Scenario	H1	H2	H3	H4	H5	H6	H7	H8	P1	P2
Unassisted	Easy	26	39	37	60	55	110	33	150	322	422
	Medium	35	67	63	90	97	140	45	200	502	615
	Hard	62	76	70	110	103	230	60	307	620	708
Assisted	Easy	26	38	45	75	66	106	35	208	344	430
	Medium	37	74	83	120	100	142	48	260	512	627
	Hard	65	80	105	160	154	340	70	430	636	712

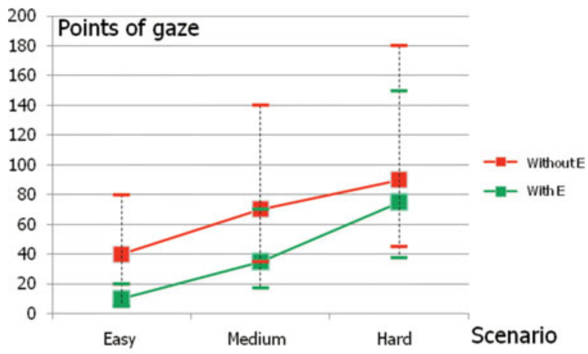


Fig. 15. Outbound points of gaze for the two trials and three scenarios for healthy users.

deviation rate between the second and the third scenarios is very close. This means that the subject can control his wheelchair easily, even if the scenario is harder. Again, the variance of the mean standard deviation increases considerably according to the scenario difficulty, for palsied subjects, the performance is worse than for healthy users. It can be reported that the deviation distance increases drastically, especially in the hardest scenario, but can also confirm that the standard deviation is better in the case of emotion-based navigation.

3.3.3. Execution time. Execution time of assisted/unassisted navigations is summarized in the Table II.

The execution time varies depending on the samples and users. Notice that for palsied subjects, execution times are very long. Notice also that the time needed to accomplish navigation tasks becomes longer when passing from the unassisted to the assisted trial. This could be explained by the fact that, when a negative emotion is detected, wheelchair velocity is decreased (and can be stopped in the case that the negative emotion is persistent). Consequently, the wheelchair takes more time to reach its goal. It also depends on the frequency of the appearance of the detected emotions during navigation, which can be more or less influential on the subject's performance. A correlation study between emotion frequency and navigation performance could reveal more interesting results about the relationship between the two. Depending on the context of the study, the tradeoff between navigation speed and security should be weighed. Although this parameter showed better performances in case of unassisted navigation, this criterion is not the most important as navigation safety is the highest priority.

3.3.4. Points of gaze. This parameter is very important as it can reflect the satisfaction of the subject during navigation. Even if the user controls the turn left/right by his gaze, it is revealed that he can perform some sporadic fixations that can reflect one of the following scenarios: either he's searching on the screen, confused by the obstacles or lacks concentration. It can be seen by the appearance of some POG outside the command zones. The task in this situation is to calculate the number of those points.

The number of outbound POG is calculated for each subject and the mean is presented in Figs. 15 and 16.

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

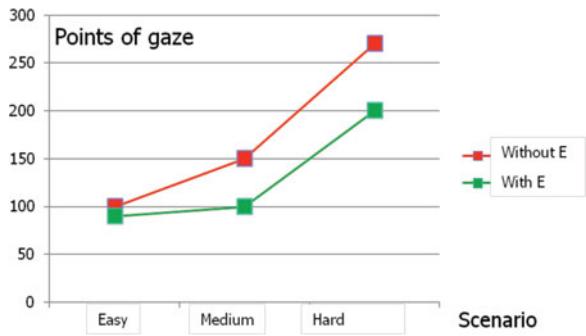


Fig. 16. Outbound points of gaze for the two trials and three scenarios for palsy users.

The results found in this study are different from the previous ones. The number of POG increased considerably in both trials. It can also be noted that in the second trial, the increase was even higher than in the first trial despite the fact that the overall number in the second trial is lower than the first one. A justification for this is again the moving obstacles that play an important role in the navigation performance; even if the user concentrates on commanding his wheelchair, he performs some unintentional fixations in the moving obstacles that can disturb his normal state. As a matter of fact, the moving obstacles had a lot of influence on the results found within the four parameters. This may imply that during wheelchair navigation, one has to give a lot of importance to the moving entities, such as humans, animals or objects that may affect the choice of suitable techniques and results.

For palsied subjects, the results are again very bad in both of the trials and the three scenarios. The POG detected increase proportionally to the scenario difficulty. It is well reported that for those subjects, the integration of emotions was embarrassing for them, especially because they are not familiar with this approach. This may explain the high number of outbound POG when comparing them with healthy subjects. Nonetheless, the latter showed dissatisfaction with semi-autonomous systems. This fact was also reported in various other studies.[24]

3.4. Discussion

The Emotion detection techniques were applied on a database of 40 healthy users but were tested afterward on both healthy and palsied subjects. This could explain the bad results found on the latter. In fact, due to the lack of sufficient disabled samples, the collected data was based only on normal EEG activity. The total number of severely disabled samples (only two) is not enough to apply intelligent techniques and bring comparative studies about navigation enhancement using emotion or not. It can be seen from the previous figures that relevant conclusions were observable on healthy subjects but not on palsied ones.

In this case, it was reported that within some of the parameters (obstacles hit and path deviation) integrating emotions can give better results, especially by assuming that a separation of one week between the two trials is sufficient to omit learning effect. Another hypothesis that can be made is that by adding training sessions, the subject performance can be improved. It was also reported that the moving obstacles have a consistent impact on the navigation performance, as the performance variances increase considerably when introducing this parameter. Moving obstacles can refer to real situations such as walking persons in streets or, driving cars on roads, and must be deeply studied to enhance more emotion-based navigation.

In the case of palsied subjects, and despite the fact that the number of subjects is very low, it is observable that emotion-based navigation gives encouraging results that can be proven later by enlarging the severely disabled subjects database. As for the chosen emotions that correspond to the real problems faced by severely disabled people, the use of only four emotions from each quadrant needs to be enlarged too. In fact, in the same valence/arousal quadrant, relaxation and sleepiness can be found. If relaxation is considered a positive emotion, sleepiness is revealed to be very dangerous in the navigation scenario and consequently, adding other emotions is crucial.

4. Conclusion and Perspectives

In this paper, an enhancement to the standard wheelchair navigation was added using emotion detection as a braking system. Different techniques for extraction (Welch and Wavelets), selection (PCA and GA) and classification (MLP, LDA and SVM) were tried, but Wavelets, GA and MLP were decided to be the best combination. The system was tested on a virtual environment and assessed on four parameters: obstacles hit, navigation path deviation, execution time and outbound POG. For the first and second parameters, it was clear that the introduction of emotions is efficient as it decreases the number of obstacles hit and the deviation rate from the optimal trajectory. However, this was not the case for the third parameter, which are the outbound POG. Although the overall number decreases between the first and second trials, it still shows dependency on the scenario. When the scenario becomes more difficult, the number of outbound points of gaze becomes higher. This may reflect dissatisfaction among the subjects. The concept of dissatisfaction towards autonomous or semi autonomous systems was already reported in many studies. Another point that should be discussed is the limitation of the palsied sample. While the results are persistent for healthy users and can be compared on the basis of navigation enhancement based on emotion integration, this is not the case for palsied subjects. Only two subjects are not enough to bring further discussion. The authors suggest enlarging the palsied sample database, because the results were encouraging for healthy ones. Last, the reliability of the use of EEG to detect emotions has not been proven yet; in this series of experiments, the correlation study was based only on EEG measures and subjects ratings, which tend to be very subjective. Also, in this study only four emotions were treated as they belong to one of the four quadrants of the valence-arousal scale. However, as it was mentioned in many studies,[27] EEG signs of emotional activation can vary in temporal scales (the increase/decrease of certain EEG amplitudes in accordance with emotional activations), frequency scales (linking low/high frequencies, thus brainwave signals, to specific emotions) or even in spatial scales (asymmetrical activation of the cerebral hemispheres, theta associated with activation being observed in frontal and central areas). Consequently, in order to obtain consistent EEG information, more electrodes should be placed in different areas of the brain, which may lead to other issues such as bulkiness, invasiveness, and applicability of EEG modality in the wheelchair navigation context.

Finally, emotion activation can be studied through its influence on different physiological sensors. For this purpose, the authors suggest the introduction of other sensors such as Electromyography (EMG) and Electrocardiography (ECG). The fusion between the mentioned sensors could reveal which combination holds the most viable information about levels of emotion.

References

1. D. Powell and M. K. O'Malley, "The task-dependent efficacy of shared-control haptic guidance paradigms," *IEEE Trans. Haptics* **5**(3), 208–219 (2012).
2. E. B. Vander Poorten, E. Demeester, A. Hüntemann, E. Reekmans, J. Philips and J. De Schutter, "Backwards Maneuvering Powered Wheelchairs with Haptic Guidance," *Proceedings of the 2012 International Conference on Haptics: Perception, Devices, Mobility, and Communication Volume Part I, EuroHaptics'12*, Berlin, Heidelberg, (2012) pp. 419–431.
3. C. Urdiales, E. J. Perez, G. Peinado, M. Fdez-Carmona, J. M. Peula, R. Annicchiarico, F. Sandoval and C. Caltagirone, "On the construction of a skill-based wheelchair navigation profile," *IEEE Trans. Neural Syst. Rehabil. Eng.* **21**(6), 917–927 (Nov. 2013).
4. M. Ren and H. A. Karimi, "A fuzzy logic map matching for wheelchair navigation," *GPS Solut.* **16**(3), 273–282 (2012).
5. H. A. Lamti, M. M. Ben Khelifa, P. Gorce and A. M. Alimi, "A brain and gaze-controlled wheelchair," *Comput. Methods Biomech. Biomed. Eng.* **16**(1), 128–129 (2013).
6. J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.* **113**(6), 767–791 (2002).
7. P. Steinier, HANDIMOBILITY, (Date: 1 Jul. 2009, Date of access: 15 May 2010), <http://www.handimobility.org/blog/?p=4953>
8. K. Crowley, A. Sliney, I. Pitt and D. Murphy, "Evaluating a Braincomputer Interface to Categorise Human Emotional Response," *Proceedings of the 2010 10th IEEE International Conference on Advanced Learning Technologies, ICALT '10*, Washington, DC, USA, (2010) pp. 276–278.
9. E. Haapalainen, S. J. Kim, J. F. Forlizzi and A. K. Dey, "Psychophysiological Measures for Assessing Cognitive Load," *Proceedings of the 12th ACM International Conference on Ubiquitous Computing, Ubicomp '10*, New York, NY, USA, (2010) pp. 301–310.

https://doi.org/10.1017/S0263574714002197 Published online by Cambridge University Press

10. C. Gani, N. Birbaumer and U. Strehl, “Long term effects after feedback of slow coritcal potentials and of theta-beta-amplitudes in children with attention-deficit/hyperactivity disorder,” *Int. J. Bioelectromagnetism* **10**(4), 209–232 (2008).
11. M. B. Arnold, *Emotion and Personality*, vol. 1 , (New York: Columbia University Press, 1960) pp. 11–13.
12. P. Ekman, W. V. Friesen and P. Ellsworth, *Emotion in the Human Face* (Oxford: Oxford University Press, 1972).
13. J. A. Russell, “A circumplex model of affect,” *J. Pers. Soc. Psychol.* **39**(6), 1161–1178 (Dec. 1980).
14. T. C. Schneirla, “An Evolutionary and Developmental Theory of Biphasic Processes Underlying Approach and Withdrawal,” *Nebraska Symposium on Motivation*, Lincoln, (1959) pp. 1–42.
15. H. Hécaen, ≪ *Les Gauchers : étude Neurophysiologique*≫ , vol. 39, (Paris: Presses Universitatires de France, 1986) pp. 128–129.
16. J. Condry and S. Condry, “Sex differences: a study in the eye of the beholder,” *Child Dev.* **47**, 812–819 (1976).
17. G. Garcia Molina, T. Tsoneva and A. Nijholt, “Emotional Brain-Computer Interfaces,” *Proceedings 3rd International Conference on Affective Computing and Intelligent Interaction—ACII 2009*, Amsterdam (10–12 Sep. 2009) pp. 138–146.
18. L. Feldman Barrett and J. A. Russell, “Independence and bipolarity in the structure of current affect,” *J. Personality Soc. Psychol.* **74**(4), 967–984 (1998).
19. K. De Jong, “Learning with genetic algorithms: an overview,” *Mach. Learn.* **3**(2–3), 121–138 (1988).
20. J. H. K. Friedman, “On bias, variance, 0/1-loss, and the curse-of-dimensionality,” *Data Min. Knowl. Discovery* **1**(1), 55–77 (1997).
21. S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt and I. Patras, “DEAP: a database for emotion analysis; using physiological signals,” *IEEE Trans. Affective Comput.* **3**(1), 18–31 (Mar. 2012).
22. D. Queteschiner, Reality Factory, (Date: 13 Oct. 2007, Date of access: 10 Mar. 2012), <http://www.realityfactory.info/cms/tutorials.html>.
23. K. Arai and R. Mardiyanto, “Method for psychological status estimation by gaze location monitoring using eye-based human-computer interaction,” *Int. J. Adv. Comput. Sci. Appl.* **4**(3), 199–206 (2013).
24. V. Sharma, R. Simpson, E. LoPresti and M. Schmeler, “Evaluation of semiautonomous navigation assistance system for power wheelchairs with blindfolded nondisabled individuals,” *J. Rehabil. Res. Dev.* **47**(9), 877–890 (2010).
25. J.-P. Chevance, ≪ *Existence et Infirmité Motrice Cérébrale*≫ , (AFP Ecoute Infos, Dec. 2007) pp. 1–11.
26. P. Lang, M. Bradley and B. Cuthbert, *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*, (University of Florida, USA, Tech. Rep. A-8, 2008).
27. G. G. Knyazev, “EEG correlates of personality types,” *Netherlands J. Psychol.* **62**(2), 78–87 (Dec. 2006).