Mobile Robot Control by Neural Networks EOG Gesture Recognition

Choi Kyoung ho, Gifu University, 1-1 Yanagido, Gifu, g3812206@guedu.cc.gifu-u.ac.jp
Minoru SASAKI, Gifu University, 1-1 Yanagido Gifu, sasaki@cc.gifu-u.ac.jp

This paper describes the development of a neural networks gesture recognition system whereby one can control a mobile robot by using the components of his brain wave bio-potentials. Such a system may be used as a control device through human eye-movements, facial muscle, and brain wave bio-potentials. Neural networks are trained to classify EOG data into one of two classes corresponding to two cognitive tasks performed by eight training segments. The operator's forehead bio-potentials can be acquired and processed in Cyberlink™ as mobile robot control source signals. The computer analyzes an operator's the EEG(electroencephalographic) and EOG (electrooculographic) signals in real time. Neural networks analyze user's EOG signal in order to discern for the presense of a signal and then decide whether it corresponds to a valid command. In the course of EOG analysis, the neural network checks for example, turning the robot. A trained neural network can effectively recognize user intention, left or right based only on the EOG signal. The experimental results suggest that a mobile robot can be operated by human brain wave bio-potentials with neural networks.

Key words : Brain-Wave, Mind-Body Operated Devices, Human Computer Interfacing, Mobile Robot, Brain Computer Interface, EOG, EMG

1. Introduction

The purpose of this paper is to apply the neural networks to the EOG pattern recognition problem. Artificial neural networks have been successfully applied in several fields such as speech recognition, image processing, and pattern recognition. In this paper we take some steps towards the ambitious goal of constructing a Brain Computer Interface (BCI). To do this, we will use EOG patterns and make artificial neural networks "learn" to identify them. The "learning" is accomplished by applying the back propagation algorithm, which was first announced in 1986. This algorithm allows the solution of highly nonlinear problems. Also, in this work, numerous other techniques from the theory of signal analysis have been used to obtain representations and extract the features of interest from the signals for classification purposes[1]. The neural networks area has become one of the most promising fields of computer science with a prolific research and wide application from simple classification to forecasting. Applications also include data compression, control, and character recognition. Neural networks are not absent of problems. Perhaps the main one present in back propagation networks is the learning time; that is, the time it takes to the network to learn and associate a set of patterns with its corresponding outputs. Learning time is a mayor concern when the network is working in a constantly changing environment. If this were not the case, we could almost disregard the problem. Although the network can take a considerable amount of time learning, once this has been done, the network can be put into the working phase, where a response to a set of inputs can be considered instantaneous (compared with to learning time). When the environment is evolving however, learning time becomes a major factor. Neural networks may lead to the ultimate computer interface, a neural connection between human being and computers. The quality and reliability of neural interfacing, particularly if implants are being used, will have to be unquestionable. Fatal exceptions and errors might be impossible to recover from and might cause actual physical harm. In many real world situations, we are faced with incomplete or noisy data, and it is important to be able to make reasonable predictions about what is missing from the information available. This can be an especially difficult task when there isn't a good theory available to help differentiate the complex data. It is in such situations that neural networks may provide some answers. After years of hesitation neural networks are now widely accepted as important tools for pattern recognition. As their characteristics differ strongly from the traditional techniques, neural networks have a great influence on the recent developments of non-neural classifiers as well. In this work, a back propagation network is used to classify EOG signals as commands for a mobile robot to taken left or right. In the learning process, based on the network is trained to correctly classify the input EOG signals based on user feedback. EOG gesture recognition can be utilized in an interface situation to enable the user to make discrete choices by execution of different gestures. The EOG gesture recognition is accomplished by analysis of the EOG signal using simple recognition techniques. Finally, these gestures were used to control a mobile robot. A PC provided a GUI to streamline the entire process of training and control.

2. Neural Network Classifier

As stated previously, this paper applies neural networks to the problem of EOG pattern recognition and applies the results to the control of a mobile robot.

With back propagation networks, learning occurs during a training phase in which each input pattern in a training set is applied to the input units and then propagated forward. The pattern of activation arriving at the output layer is then compared with the correct output pattern to calculate an error signal. The error signal is used to appropriately adjust the weights in each layer of the network. After a back propagation network has learned the correct classification for a set of inputs, it can be used on a properly configured robot control interface to control a robot's direction. The purpose of this project is to control mobile robot with a minimum of development in time. To stay with in this limitation, we chose small neural networks with appropriately sized hidden and output layers. This lead to satisfactory result.

The central idea behind this solution is that the errors for the units of the hidden layer are determined by back-propagating the errors of the units of the output layer. For this reason the method is often called the back propagation learning rule. Back-propagation can also be considered as a generalization of the delta rule for non-linear activation functions and multi-layer networks. The classification performance of neural networks depends on the initial weight values and on the data used to train and test. If the data contains noise or does not completely specify the target function, neural networks will over-fit the training data and it will not correctly interpolate and extrapolate the training data, i.e., it will not generalize well. Training the network is accomplished by initializing all weights to small, random values and then choosing the best result in the network's weight space to minimize a squared error function of the network's output. The error is between the network's output and the target value for each input vector. Eight trials were used for the training set. The error of the network on the validation data was calculated after every pass, or epoch, through the training data. After several epochs, the network state (its weight values) at the epoch for which the validation error is smallest was chosen as the network that will most likely perform well on new data. This best network is then used to control the real robot.

Two hidden layers were used. In our applications a feed-forward network with two hidden layers and one output layer unit were used with a sigmoid activation function for the units. In this application, the generalized delta rule involves two phases. During the first phase the input is presented and propagated forward through the network to compute the output values for each output unit. This output is compared with its desired value, resulting in an error signal for each output unit. The second phase involves a backward pass through the network during which the error signal is passed to each unit in the network and appropriate weight changes are calculated.

As we previously presented, gesture recognition is dependent upon the fact that different gestures occur as a result of different actions of bio-signal abstraction. There are two signals what used in this experiment. The detected EOG and EMG signals are dependent upon which muscle units in the vicinity of the electrodes have been recruited for the movements [2]. The recruitment order is stable for a given movement but varies when the type of movement is changed [3]. This results in EMG and EOG signal whose characteristics change for different gestures. This is the basis for any effort at gesture recognition. Neural network processing techniques are used to extract "features" from the EMG and the EOG time series. These features provide an abstraction of the underlying physical situation occurring during the execution of a gesture. The signals from Cyberlink are analyzed using the neural networks technique to obtain a proper motion[4]. In the context of this study, simple but effective back-propagation network approaches were employed for pattern recognition. This technique differentiates EMG and EOG at the same time. For the case of gesture recognition, the user executes the specific gestures indicating to the computer which of the gestures are being executed for the neural network's calculation. This procedure is necessary to calculate the proper weights for the neural network's use at the real control phase. Once trained, the neural networks can differentiate between the right and left gaze of the user as well as the lack of a proper EOG signal. The techniques described above served as the basis for a gesture recognition system designed to detect the direction of motion of an experimenter's eye movements. Surface electrodes were placed in the vicinity of the user's eyes on his forehead. Data from the electrodes as collected and used to train the users and to control the robot. Despite this, it is felt that a system utilizing both EOG and EMG data is warranted to accommodate users with disabilities who are unable to perform such clearly defined tasks as studied at the present time.

3. Application

The interface consists of two main windows. There is an easy to use data collection window. It is used to perform the neural network training with a GUI. The internal program gathers eight training segments from the user for the neural network. The user can move from this window to the other main window. The robot control window, by three eyebrow clicks. Internally, calculation of the weights for the neural network takes place the use of the training GUI and the robot control GUI. This is depicted in Fig.1.

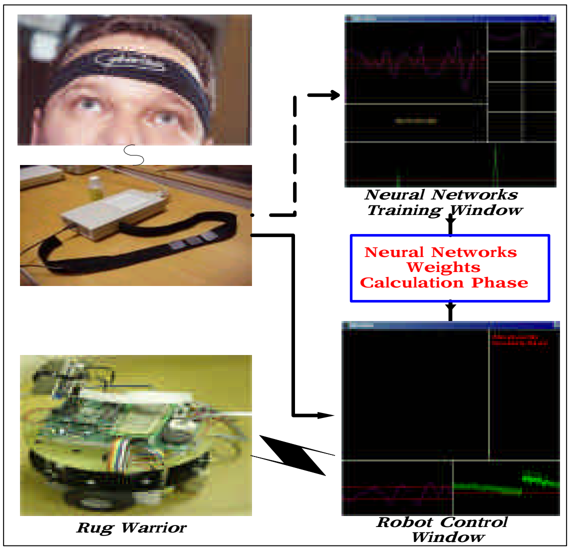

Fig. 1 System Flow.

3.1 Obtaining Signals

The last a few years have witnessed a rapidly growing body of research and technique development involving detecting human brain responses and putting these techniques to appropriate uses to help people with debilitating diseases or who are disabled in some way – the so-called "brain computer interface".

The most advanced work in designing a brain-computer interface has stemmed from the evolution of traditional electrodes[4]. There are essentially two main problems, stimulating the brain(input) and recording from the brain(output). Traditionally, both input and output were handled by electrodes consisting of metal wires and glass tubing.

The Cyberlink is best optimized to get the user's brain waves in real time with electrodes. The Cyberlink accommodates three signals from electrical input. It uses a headband with three embedded electrodes, thus eliminating the cost and health risks associated with invasive electrodes. The signals can be either EMG (Eye brow clicks), EEG (brain waves), or EOG (electric activity from the eyes) signals. The lowest frequency control signal is called the ElectroOcularGraphic(EOG) signal. This is the frequency region of the forehead bio potentials that is responsive primarily to eye movements and is typically used to detect left and right eye motion. The second type of control signal is called the ElectroEncephaloGraphic(EEG) signal. It is influenced by thoughts and broadband signals created from facial movements. This signal is typically used for vertical control of the cursor [1]. The third type of control signal is called the ElectroMyoGraphic(EMG) signal. The EMG signal primarily reflects facial muscle activity. In this experiment EMG and EOG signals were used to get discretized control commands. The Cyberlink communicates to the outside world using either a standard RS-232 serial link with 9600bps. In this system the Cyberlink is used to estimate the muscular exertion in EMG data and to filter the activity from the eyes. The obtained data of these signals are calculated and sent over a standard RS-232 serial link to a pc running the use interface software package.

3.2 Neural Networks Training GUI interface

Basic to neural networks is the need to be trained so that their weights can be adjusted to optimum values. A large collection of sample, inputs and outputs are needed for the training session. In this application, the training was by having the experimenter generate the sample data himself through a GUI with the electrodes on, the experimenter directs his gaze to the left or right, introducing to the program the direction of his gaze. The EOG signals captured from the electrodes are shown on the screen, By repeating this process, the neural network becomes trained and the experimenters gains practice. This direct training method is promising but further developments must be done to fully exploit its use for handicapped users

For the work here, we used simple back propagation neural networks with pure sigmoid activation functions. The neural networks were initialized with random weights(0 – 1). The training process adjusted these weights such that the output matched the experimenter's indication of the correct value.

Here are some equations that describe our neural

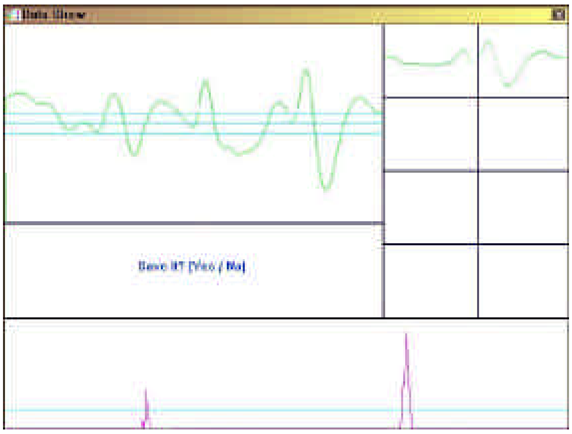

Fig. 2 Neural Networks Training GUI Window.

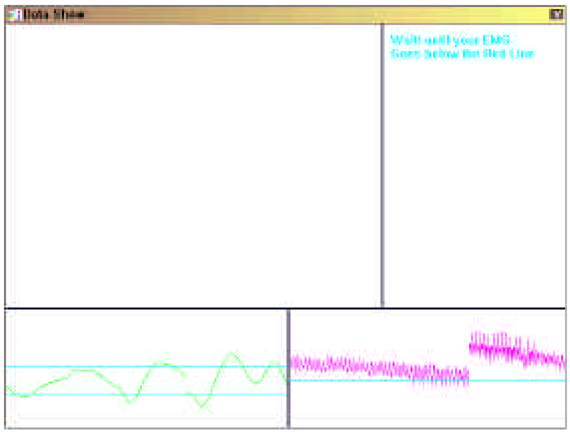

Fig. 3 The main interface window initial settings.

networks. Let the weight between neuron j and neuron i be w_{ij}. Let the net input to neuron j be net j, then

$$net_j = \sum_{i=1,n} w_{ij} o_i ,$$

where n is the number of units feeding into unit j and the activation value for neuron j is o_j, a function of the net input,

$$o_j = f(net_j) ,$$

The actual activation value of an output unit k, will be o_k and the target for unit k, will be t_k. The error δ_k is defined as

$$\delta_k = (t_k - o_k) f'(net_k)$$

We use the usual activation function

$$f(net_k) = \frac{1}{1 + e^{-net_k}} .$$

The derivative term is

$$o_k(1 - o_k) .$$

The formula to change the weight w_{ik} between the output unit k and unit j is

$$w_{jk} \leftarrow w_{jk} + \eta \delta_k o_j ,$$

where η is a small positive constant called the learning rate. We used a small value of 0.8. The error δ_j for a hidden net j is given by

$$\delta_j = f'(net_j) \sum_k \delta_k w_{kj} .$$

The weight change formula for a weight w_{ij} that goes between the hidden unit j and the input unit i is essentially the same as the equation given previously,

$$w_{ij} \leftarrow w_{ij} + \eta \delta_j o_i .$$

Employing these equations, the neural network was able to obtain weights which allowed the experimenter to control the mobile robot's direction after only a few epochs of training.

3.3 The Robot Control GUI Window

Three eyebrow clicks serve to change the active window from the neural network training window to the robot control window.

A combination of the proportional estimation and neural networks gesture recognition techniques discussed previously were used to implement several graphical user interface (GUI) objects (Fig.3). Specifically, a continuous eye movement display and a continuous facial muscle control display were displayed using the C++ programming language under Windows98. As mentioned previously, a gesture recognition system allows the user to pick from among a discrete set of choices. The proportion of facial muscular exertion and eye movements allows the user to control the mobile robot continuously.

These two modalities of communication complement each other in a human-computer interface situation. An example where both are used is the case of the graphic window mentioned previously. In this experiment, a specific gesture could be executed to select from one of controls displayed on a computer screen in the help window that is located in the upper right of the main window. Once selected, the muscular exertion estimate could then be used to get the specific code presented herein. The direction of motion of robot was determined by the direction of motion of the user's eye gaze. An internal neural network program uses the input EOG data and the weights in the training process to find the direction of the user's gaze. Once the direction is determined by the neural network, specific code is sent to the robot to obtain movement in the appropriate direction. The speed of motion was determined by the continuous exertion of the user. Once the neural network was trained, the user was able to move the robot any time. Even if the user EOG signal varied, the neural network was able to detect user eye direction effectively.

4. Robot Control

In this experiment, we used a mobile robot called Rug-Warrior Pro. As shipped, several control commands are preconfigured in the robot. The MindMouse is shipped with software for training the operator to use brain waves and small facial movements. The experimenter can use this software,

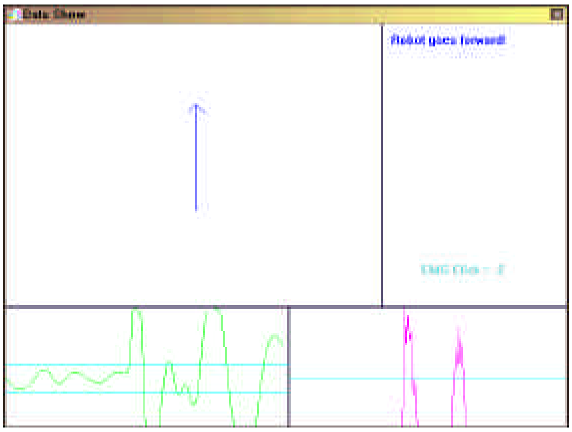

Fig. 4 Detection of a Move Forward command by EMG (2 eyebrow clicks).

Fig. 5 The Rug-Warrior moves.

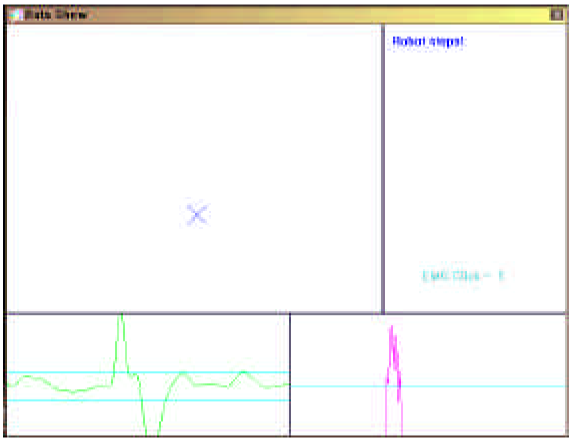

Fig. 6 Detection of “Stop” command by EMG (1 eyebrow click).

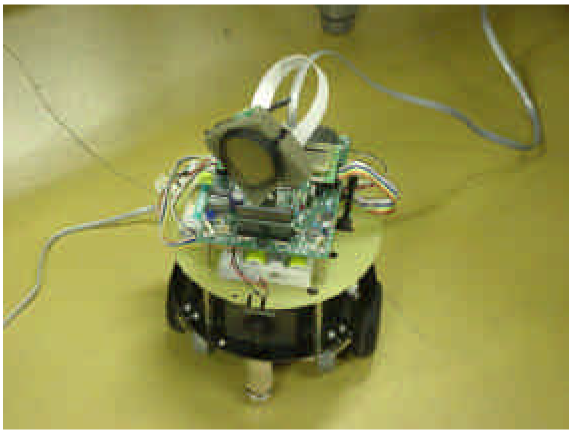

Fig. 7 The Rug-Warrior comes to a stop.

which included challenging programs such a labyrinth and MindTetris, and Cyber Trainer to teach himself to use EOG, EMG and EEG to control a variety of activations. After going through training the user is able to control the mobile robot effectively The raising of an eyebrow is termed ‘clicking’ and one or two clicks is associated with the raising of the eyebrow once or twice. These clicks can give two commands such as “stop” and “go forward” to the robot. The threshold borderline was displayed in the EMG signal display window at the lower right of the main window. By observing this display, the user can become familiar with EMG control. The experimenter can decide which muscle excitations are used make the robot go forward (Fig. 4, Fig. 5) or to stop (Fig. 6, Fig. 7).

Initially, the experimenter must wait his or her EMG signal to reach the bottom line of EMG signal window. Two motion commands are also derived from the direction of the experimenter’s gaze. The mapping here is obvious, with the robot turning right or left according to the direction of the experimenter’s gaze. Figures 8-9 show these turning commands in action. The threshold borderlines were shown in the EOG signal display window on the lower left of the main window. Table 1 shows the mapping between the physiological input and the desired robot command. Figures 4 through 11 show

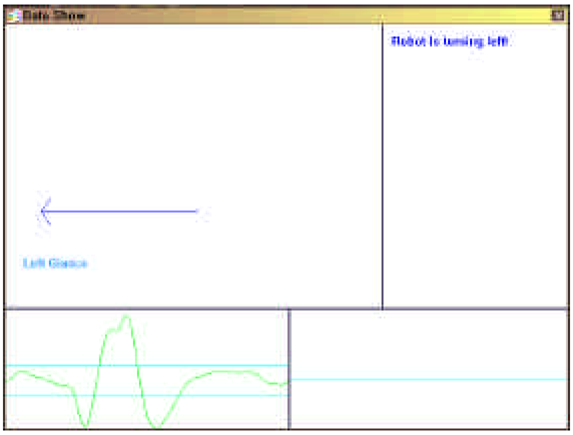

Fig. 8 Detection of “Turn Left” command (left gaze).

Fig. 9 The Rug-Warrior turns left.

all the commands in action. The experimenter looks from the center to the left/right and then back to the center to generate a left/right gaze.

As in EMG control, an experimenter definitely needs to practice using EOG signal window to obtain better performance with the robot.

These experimenter’s generated signals are analyzed by the GUI program running under Windows98. The robot command codes are sent in almost real time over a standard RS-232C.

5. Conclusion

Two classes of brain wave bio-potential signals were discussed. Proportional estimation of muscular

Table. 1 Mapping of physiological input to robot commands.

Physiological Input	Robot Command
Left gaze	Robot turns left direction
Right gaze	Robot turns Right direction
Eyebrow clicks one time	Robot goes forward
Eyebrow clicks two times	Robot stops movement

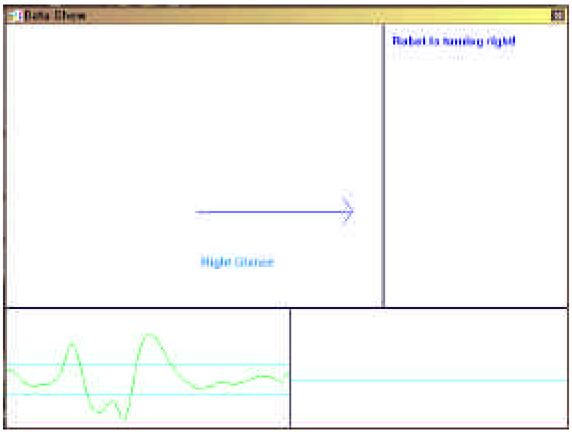

Fig. 10 Detection of "Turn Right" command (right gaze).

Fig. 11 The Rug-Warrior turns right.

exertion from eye movement and eyebrow clicks provide interface tools which can be used to control robot in a continuous nature. We developed a neural network to recognize a user' s gestures allowing him to make discrete choices among available commands on control a robot. An application was presented which integrated these two signals into a functional human interface. To obtain weights for the neural network, the user trained it directly through a GUI. Application of the techniques described here are not limited to just this case of the robot control. It is clear that other machines can be controlled in the same manner. The use of a GUI is especially helpful in the training process. Many other standard elements typical to the graphical computing environments common today could benefit by direct control using brain wave bio-potentials. Also, inclusion of a robust learning structure like the one employed here makes it possible for the system to adapt to the brain wave bio-potentials, different users, making particularly suited for disabled users.

This specific control method gives good opportunity to people who are physically dependent and can not control their environment and thus have few opportunities for spontaneous exploration [5][6]. It also proves that the Cyberlink system has great potential for controlling a wide variety of machines using a person' s brainwave bio-potentials.

The specific gestures used in such a system as this can be tailored to match an individual' s capabilities. It seems that with a little effort the system can be made reliable enough to be used in the field.

The field of neural interfacing is still extremely new. The potential of this technology is tremendous for use in many fields: medical, games, virtual reality and for normal everyday use. There are many different types of devices of varying usefulness, each with their own applications. Many companies already market this technology and many research groups are looking into finding new and innovative uses for it. This technology still has its problems and difficulties to overcome, but the future looks very bright for it indeed.

References

[1] A.S.Gevins and A.Remond. Methods of Brain Electrical and Magnetic Signals,volume 1 of Handbook of Electroencephalography and Clinical Neurophysiology (revised series). Elevier Science Publishers B.V., New York, NY, 1987.

[2] E. Doherty, C. Bloor, G. Cockton "Yes/No"- A Mind-Body operated device application for severely motor impared persons**.**

[3] Allanson, J. "A toolkit for Exploring Electro-Physical Human Computer Interaction" Human-Computer Interaction, Interact 99, IFIP TC, 13 International Conference on Human-Computer Interaction, 30th August, September, 1999.

[4] Junker A., United States Patent 5,692,517 December 2, 1997.

[5] Albert. M, Cook, Kathy howery, "Robot-Enhanced Discovery and Exploration for Very Young Children with Disabilities." Technology and Persons with Disabilities Conference 1999.

[6] W.T. Nelson, et al., "Brain-Body-Actuated Control: Assessment of an Alternative Control Technology for Virtual Environments," Proceedings of the 1996 IMAGE Conference, Scottsdale, Arizona, June 23-28, 1999