International Journal of Automation, Mechatronics & Robotics– IJAMR
Volume 1 : Issue 2 [ISSN 2374 – 1546]
Publication Date : 25 June 2014

Electromyogram as a Viable Direct Human Input for Human Machine Interfaces

Interface Design with Case Studies

G Sajaysurya, J Suresh Kumaar, M Madhu, K Prakasan

***Abstract*— Direct Human Input (DHI) is an input methodology for Human Machine Interfaces (HMI). Use of DHI eliminates physical input devices as input data is collected directly from the signals generated by the human body. Various varieties of DHI exists of which Electromyogram (EMG) is found to have properties that makes it readily applicable. In this work, the capability of EMG to detect facial expressions of the user has been utilized to create a generalized interface that processes EMG signals into machine commands in accordance with the user's facial expressions. The interface is essentially a combination of dimension reduction and classification algorithms that classifies EMG signals into facial expressions to activate machine commands correlated with the estimated facial expression. Linear Discriminant Analysis and Support Vector Machines have been used for dimension reduction and classification respectively. The interface has been found to perform satisfactorily with a classification accuracy over 99%. Case studies were conducted to verify the performance of the proposed interface in real-time applications and the performance was found to be satisfactory.**

Keywords— Human Computer Interface, Direct Human Input, Electromyogram, Facial Expressions

I. Introduction

In the realm of Human Machine Interfaces (HMI), Direct Human Input (DHI) is relatively a newcomer and hence is sparsely applied. However, some Virtual Reality (VR) systems have adopted DHI [1]. DHI largely differs from all other kinds of inputs to an interface in that its format is native to the human body. Here, the data collected directly from the signals generated by the human body serve as the input [1]. Hence the need of a physical input device is eliminated as the user himself serves as the source of input. The common forms of DHI for HMI are motion, speech, electromyogram (EMG) and electroencephalogram (EEG). Motion inputs are captured by a video camera and the video is processed to identify specific commands which are pre-correlated with specific movements or gestures. Systems that utilize this input are widely commercialized. Microsoft's Xbox Kinect™ [2] is one such system. Speech has been successfully utilized as an input for HMI [3, 4] but it suffers from its inherent disadvantages particularly when used by people of different ethnic groups with language barrier. Further the interface must be capable of discerning commands from casual conversation.

G Sajaysurya, J Suresh Kumaar, M Madhu, K Prakasan
PSG College of Technology
India

Numerous works have been conducted to aid handicapped patients in locomotion using electroencephalogram (EEG) as input [5-7]. HMI that utilized EEG are termed as Brain Computer Interfaces (BCI). Here, the potential distribution on the scalp of the user is monitored for unique patterns pertaining to discrete input commands. These systems need costly amplifiers and the cheap ones will naturally degrade classification accuracy. BCI that use mu rhythm (motor imagery – imagining a physical limb movement) to get input commands are indeed successful [8] up to some extent but hinder when the user actually uses his limbs for physical locomotion. BCI needs highly controlled environments and the number of classes (discernable input commands) is very low. Most BCI are synchronous or cue based (analogous to synchronous circuits waiting for a clock pulse to act). This forces the user to express a trained thought only during a small window of time just after the trigger. This inhibits the capability of the user to have a free will. Further, identified potential distribution patterns on scalp may drift with external factors like temperature and time.

In spite of these issues, Göhring et al. [9] report successful usage of cheap commercial EEG gaming headsets to steer a real car using proprietary classification algorithms developed by emotiv systems [10]. But they do admit that the accuracy is not high enough to steer the car in open traffic. Even though BCI could be perfected in the near future BCI illiteracy is a much bigger problem to solve [11] as about 30% of human population could not be trained to use BCI.

Electromyogram (EMG) that captures muscle actuation is a viable DHI for HMI. Though high level motion inputs can easily be captured by a camera, subtle muscular activity like eye movements and facial expressions can be identified from EMG with high reliability. EMG signals are much stronger than that of EEG and hence with minimal training, EMG signals can be used for high speed real-time classification of muscular actuations into discrete inputs. Though EMG patterns are different for each person they exhibit minimal response to external factors. These features of EMG intrigued the authors to probe the applicability of EEG as a DHI for general HMI.

II. Interface Design

It has been intended to design an interface that takes facial expression of the user as input to give corresponding output commands to the machine of interest. The facial expression of the user would be identified by processing the EMG signals acquired from the user.

International Journal of Automation, Mechatronics & Robotics– IJAMR
Volume 1 : Issue 2 [ISSN 2374 – 1546]
Publication Date : 25 June 2014

TABLE I. FACIAL EXPRESSIONS BEING CONSIDERED

Facial Expression	Descriptive Action
	Left Eye being Winked
	Right Eye being Winked
	Eyebrows being Raised
	No Expression

The interface is designed to identify the facial expressions reported in Table 1.

A. Equipment

EMG signals were recorded using a commercially available emotiv EPOC neuroheadset [10]. Emotiv EPOC has been designed to record both EEG and EMG signals and is an easily available off the shelf solution for recording EMG signals. Sensors at 8 locations were monitored to predict the facial expression of the user. The eight locations (based of the international 10-20 nomenclature for EEG) are AF3, AF4, F7, F8, FC5, FC6, T7 and T8. The potential difference between each sensor at the 8 locations and the reference electrode are transmitted wirelessly from the neuroheadset to the computer. The eight channel data is then processed by the signal processing algorithm.

B. Analysis of EMG Signals

The way in which data is encoded in a signal largely determines the signal processing methodology. EMG signals are electric signals associated with muscle contraction. Central nervous system (CNS) is known to co-ordinate all the muscle contractions. An action potential (AP) generated by CNS traverses along nerves to reach muscles which stimulates contraction. The strength of muscle contraction has a direct correlation with the frequency of this stimulation [12]. Thus the major frequency component of the EMG signal (composed of APs) is a reliable measure of muscle contraction.

C. Signal Processing Algorithm

The signal processing algorithm is essentially a classification algorithm that classifies independent variables calculated from EEG signal into facial expressions. It then feeds the machine with the command correlated with the estimated facial expression. As the frequency domain of EMG signals is sensitive to changes in facial expressions, the power values of different frequency bands in the frequency domain of EMG signals were considered as independent variables for classification. EMG signals were transformed from the time domain into frequency domain using Fast Fourier Transform (FFT) [13] after applying a Hanning window. Power value quantifies the amount of signal with frequency within the specified frequency band. The power values of different frequency bands were estimated using (1).

$$power = 10 * \log_{10}\left(\frac{1}{h-l}\sum_{i=l*N/S_f}^{h*N/S_f}\left|c(i)\right|^2\right) \qquad (1)$$

Where,

h = higher frequency limit
l = lower frequency limit
N = number of discrete values over which FFT is performed
c = frequency domain values
i = index
S_f = sampling frequency

Generally the number of independent variables calculated is much higher than those actually required for classification so that a much larger group of users can be supported by the algorithm. Hence, dimension reduction is applied to extract user specific variables that contain almost all the data required for classification of signals for the particular user. This subset of independent variables is termed as features. These features serve as inputs for classifier algorithms that perform classification.

The proposed algorithm has two phases namely the training phase and the working phase as depicted in Fig 1. The training phase develops a suitable dimension reduction methodology and classifier design for the current user. The working phase, by using the dimension reduction methodology and the classifier design developed in the training phase, converts the EMG signals of the current user into machine commands. Thus the training phase of the algorithm must be executed once for every new user before executing the working phase. This signal processing algorithm performs dimension reduction by Fisher's Linear Discriminant Analysis (LDA) and the classification is performed by a Support Vector Machine (SVM) with radial basis kernel. The training phase requires recordings of EMG data associated with each facial expression of the user, for whom the algorithm is trained. For each facial expression, the algorithm records EMG data from all the eight locations (8 channel EMG) for 16 seconds.

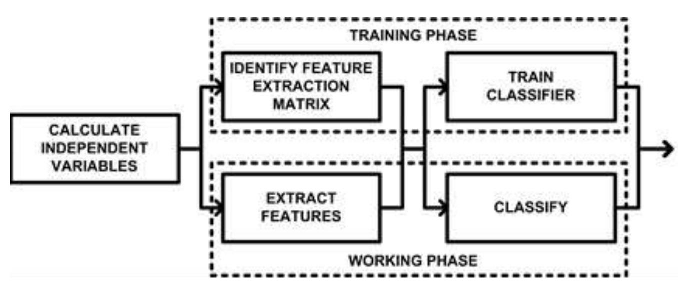

Figure 1. Signal Processing Methodology

It has been observed that the frequency band sensitive to facial expressions is 25Hz to 40Hz. To monitor variations within this band it has been divided equally into three bands. Hence the powers in these three bands were considered as independent variables. From each 16 second data, the algorithm determines independent variables in eight windows (one for each channel) of width 0.5 second, thus contributing 24 (3 variables for each of eight channels) independent variables that characterize the facial expression of the user at that particular second. The same process is repeated several times by rolling the window forward by 0.125 second till the entire 16 second data is exhausted. This procedure results in four classes of 24 dimensional data with each class associated with a particular facial expression.

LDA linearly combines these 24 independent variables into a new variable such that the new variable is a better class discriminator. As LDA is supplied with the details regarding the class to which each 24 dimensional data belongs, it optimizes the linear combination in such a way that the new variable has minimum within-class variance and maximum between-class variance. The maximum number of these new variables is one less than the number of classes [14]. As there are four classes, three new variables can be extracted thus reducing a 24 dimensional data into a 3 dimensional data. These three variables serve as features for classification. The co-efficient of the three linear combinations that define the three new variables can be consolidated into a single matrix termed as transformation matrix. The 24 dimensional data, when post-multiplied by this transformation matrix yields the three dimensional data. This transformation matrix is user-specific and is used by the algorithm for dimension reduction in the working phase. The algorithm uses the same transformation matrix to transform the four classes of 24 dimensional data associated with the training phase into four classes of 3 dimensional data which would serve as the training data set to train the classifier.

As the class details associated with the three dimensional training data set is known, the SVM identifies the optimal hyperplane that linearly separates two classes of multi-dimensional data [14]. If the classes are not linearly separable, the radial basis kernel transforms the data into a higher dimension till linear separability is achieved. This identification of hyperplane is termed as classifier training. As a hyperplane cannot linearly separate more than two classes, all possible hyperplanes between any two of the four classes are identified thus identifying six ($4C_2$) hyperplanes or essentially training six binary classifiers.

Once the training phase is executed, the algorithm switches to the working phase. The working phase of the algorithm records packets of EMG data in real-time. Each packet of data contains EMG data of length 0.5 second. The algorithm calculates the 24 independent variables associated with the EMG data packet and dimensionally reduces it to 3 features using the transformation matrix developed during the training phase. These 3 feature values are supplied as a three dimensional data point to the classifier. Based on the side of the hyperplane to which the point lies, the six classifiers provide six different predictions of the class to which the 3 dimensional point belongs. The six predictions would be considered as votes and the facial expression associated with the winning class is considered to be the current facial expression of the user. Based on this predicted facial expression the algorithm sends the associated command to the machine being interfaced.

III. Results and Discussion

The signal processing algorithm of proposed HMI has been checked in real-time and the results are reported in Table 2. In this context, accuracy refers to the ratio of the number of attempts during which the facial expression is correctly interpreted into the corresponding machine command to the total number of attempts made.

TABLE II. Facial Expressions being Considered

Facial Expression	Classification Accuracy %	Smoothened Accuracy %
Left Eye being Winked	97.24	99.92
Right Eye being Winked	95.87	99.82
Eyebrows being Raised	93.94	99.63
No Expression	93.39	99.56

The classification accuracies signify the performance of the components of signal processing algorithm and in turn the performance of HMI. Though the accuracies are greater than 90%, it is impossible to use them in real-time application as they still lead to numerous misclassifications. To increase the accuracies smoothening techniques can be adopted. It could be programmed in such a way that only two same consecutive facial expression predictions can evoke the associated navigational command. The improbability of two consecutive misclassifications has also been reported in Table 2 as “Smoothened Accuracy” which is high enough for real-time applications. This result supports the fact that the proposed interface is a viable HMI that processes EMG input. Smoothening is done at the cost of time lag between user action and machine’s response for command.

IV. Case Studies

As the performance of the proposed interface has been satisfactory, the authors performed case studies by using the interface in variable applications to verify its applicability. Three different case studies were made by employing the interface for navigation in virtual environment, orienting a 3D object in a Computer Aided Design (CAD) environment and to teleoperate an electric car.

A. Navigation in Virtual Environment

The physical input device that generally provides the computer with the navigational commands has been replaced with the proposed interface. Vizard Software Toolkit [15] has been used to create a virtual environment of an Italian Palazzo for the study. Vizard SDK [16] has been used to create the software interface to process EMG signals into navigational commands. The correlation between the facial expression and the navigational commands are reported in Table 3.

TABLE III. FACIAL EXPRESSIONS AND CORRESPONDING NAVIGATIONAL COMMANDS

Facial Expression	Navigational Command
Left Eye being Winked	Turn Left
Right Eye being Winked	Turn Right
Eyebrows being Raised	Go Straight
No Expression	Do not move

The performance of the interface has been found to be satisfactory.

B. Orienting 3D Object

Instead of using a mouse to orient a 3D model in a CAD environment, the proposed interface has been employed in this study. NX Unigraphics 7.0 [17] has been selected as the CAD package to perform testing as it provides NX Open Application Programming Interface (API). The software interface that maps facial expressions into orientation commands has been programmed using NX Open API through C++. The orientation commands are mono directional rotation commands about the three axes of the co-ordinate system. The mapping between the facial expression and orientation commands is reported in Table 4.

TABLE IV. FACIAL EXPRESSIONS AND CORRESPONDING ORIENTATION COMMANDS

Facial Expression	Navigational Command
Left Eye being Winked	Counter clockwise rotation about Z axis
Right Eye being Winked	Counter clockwise rotation about Y axis
Eyebrows being Raised	Counter clockwise rotation about X axis
No Expression	Do not move

The users were able to easily orient the 3D models but the limited number of discernable facial expressions proved to be a problem when bidirectional rotation is preferred. Bidirectional rotation needs seven different discernable facial expressions.

C. Teleoperating an Electric Car

An attempt has been made to teleoperate an electric car using the proposed interface. The user interacts with a server computer that contains the proposed interface programmed as a software. The navigational commands extracted by the server would be communicated to the on-board computer in the electric car through TCP/IP sockets over the Wide Area Network (WAN). A Raspberry Pi Model – B [18] computer serves as the on-board computer that receives the navigational commands and correspondingly steers the electric car. The facial expressions and the corresponding navigational commands are similar to the mapping used for navigation in virtual environment tabulated in Table 3. The electric car used to perform the case study is shown in Fig. 2. The general performance of the interface was found to be satisfactory in this case but the time lag induced by signal processing algorithms resulted in slow response which may at times be detrimental for this application.

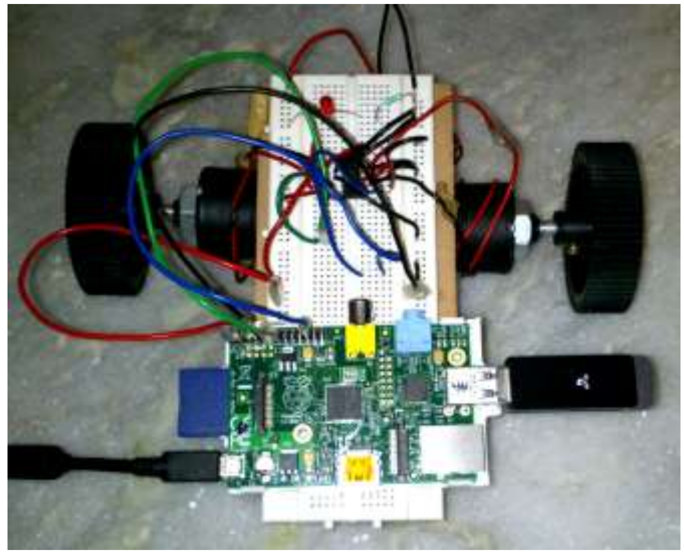

Figure 2. Electric Car with Onboard Computer

These case studies prove the applicability of the proposed interface and in turn the applicability of EMG as DHI.

V. CONCLUSION

Electromyogram has been identified as a viable candidate for Direct Human Input for Human Machine Interface. An interface that processes the EMG signals into commands for the machine has been developed. Three case studies that have been carried out to access the applicability of the proposed interface gave acceptable results.

In the future the authors intend to improve the signal processing algorithm that constitute the interface so that time lag is reduced and more number of facial expressions could be discerned thereby expanding the number of possible applications of EMG as DHI.

ACKNOWLEDGMENT

The authors would like to thank the Management and the Principal of PSG College of Technology, India for providing the facilities required for this work. The authors would also like to thank the members of the Virtual Reality Laboratory of PSG College of Technology for their assistance.

REFERENCES

[1] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev, *3D User Interfaces: Theory and Practice*: Pearson Education, 2004.
[2] *Xbox Kinect*. Available: http://www.xbox.com/en-IN/Kinect
[3] M. D. J. McNeill, H. Sayers, S. Wilson, and P. Mc Kevitt, "A Spoken Dialogue System for Navigation in Non-Immersive Virtual Environments," *Computer Graphics Forum*, vol. 21, pp. 713-722, 2002.
[4] C. Shoupu, Z. Kazi, M. Beitler, M. Salganicoff, D. Chester, and R. Foulds, "Gesture-speech based HMI for a rehabilitation robot," in *Southeastcon '96. Bringing Together Education, Science and Technology., Proceedings of the IEEE*, 1996, pp. 29-36.

[5] T. Felzer and B. Freisleben, "HaWCoS: the "hands-free" wheelchair control system," presented at the Proceedings of the fifth international ACM conference on Assistive technologies, Edinburgh, Scotland, 2002.
[6] B. Rebsamen, T. Chee-Leong, Q. Zeng, V. M. H. Ang, E. Burdet, C. Guan, H. Zhang, and C. Laugier, "Controlling a Wheelchair Indoors Using Thought," *Intelligent Systems, IEEE,* vol. 22, pp. 18-24, 2007.
[7] K. Tanaka, K. Matsunaga, and H. O. Wang, "Electroencephalogram-Based Control of an Electric Wheelchair," *Robotics, IEEE Transactions on,* vol. 21, pp. 762-766, 2005.
[8] Y. Jeon, C. S. Nam, Y.-J. Kim, and M. C. Whang, "Event-related (De)synchronization (ERD/ERS) during motor imagery tasks: Implications for brain–computer interfaces," *International Journal of Industrial Ergonomics,* vol. 41, pp. 428-436, 2011.
[9] D. Göhring, D. Latotzky, M. Wang, and R. Rojas, "Semi-autonomous Car Control Using Brain Computer Interfaces," in *Intelligent Autonomous Systems 12*. vol. 194, S. Lee, H. Cho, K.-J. Yoon, and J. Lee, Eds., ed: Springer Berlin Heidelberg, 2013, pp. 393-408.
[10] *emotiv*. Available: http://www.emotiv.com
[11] T. Dickhaus, C. Sannelli, K.-R. Müller, G. Curio, and B. Blankertz, "Predicting BCI performance to study BCI illiteracy," *BMC Neuroscience,* vol. 10, pp. 1-2, 2009/07/13 2009.
[12] A. C. Guyton and J. E. Hall, *Textbook of medical physiology*: Elsevier Saunders, 2006.
[13] E. O. Brigham, *The Fast Fourier Transform and Its Applications*: Prentice Hall, 1988.
[14] A. Fielding, *Cluster and Classification Techniques for the Biosciences*: Cambridge University Press, 2007.
[15] *Vizard Software Toolkit*. Available: http://www.worldviz.com/products/vizard
[16] *Vizard SDK*. Available: http://docs.worldviz.com/vizard/SDK_Introduction.htm
[17] *Siemens NX Unigraphics*. Available: http://www.plm.automation.siemens.com/en_us/products/nx/
[18] *Raspberry Pi*. Available: http://www.raspberrypi.org/

About Author (s):

G Sajaysurya received his B.E. degree in Production Engineering from Anna University, Chennai, India in 2013. His current research interests include Industrial Automation, Intelligent Human Machine Interfaces and Design Optimization.

J Suresh Kumaar received his B.E. degree in Production Engineering from Anna University, Chennai in 2013. His current research interests include Intelligent Human Machine Interfaces and Industrial Robotics.

M Madhu received his M.E degree in Manufacturing Engineering from Anna University, MIT Campus, Chennai in 2009. He is currently an Assistant Professor in the Department of Production Engineering in PSG College of Technology, Coimbatore, India. His main research areas are Virtual Reality, Hydraulic systems design and Robotics.

K Prakasan received his M.Tech degree in Aircraft Production Engineering from Indian Institute of Technology Madras, Chennai in 1991 and PhD degree in Mechanical Engineering from Indian Institute of Science, Bangalore, India in 1997. He is currently Professor and Head of the Department of Production Engineering in PSG College of Technology, Coimbatore. His main research areas are Metal castings, Metal Matrix Composites, Product Life-cycle Management and Product Data Management.