BCI-TOUCH BASED SYSTEM, A MULTIMODAL CAD INTERFACE FOR OBJECT MANIPULATION

Rohit Bhat*
Department of Mechanical and Aerospace Engineering
University at Buffalo, SUNY Buffalo, NY, USA
rohitbha@buffalo.edu

Rahul Rai
Department of Mechanical and Aerospace Engineering
University at Buffalo, SUNY Buffalo, NY, USA
rahulrai@buffalo.edu

Akshay Deshpande*
Department of Mechanical and Aerospace Engineering
University at Buffalo, SUNY Buffalo, NY, USA
adeshpan@buffalo.edu

Ehsan Tarkesh Esfahani
Department of Mechanical and Aerospace Engineering
University at Buffalo, SUNY Buffalo, NY, USA
ehsanesf@buffalo.edu

ABSTRACT

The aim of this paper is to explore a new multimodal Computer Aided Design (CAD) platform based on braincomputer interfaces and touch based systems. The paper describes experiments and algorithms for manipulating geometrical objects in CAD systems using touch-based gestures and movement imagery detected though brain waves. Gestures associated with touch based systems are subjected to ambiguity since they are two dimensional in nature. Brain signals are considered here as the main source to resolve these ambiguities. The brainwaves are recorded in terms of electroencephalogram (EEG) signals. Users wear a neuroheadset and try to move and rotate a target object on a touch screen. As they perform these actions, the EEG headset collects brain activity from 14 locations on the scalp. The data is analyzed in the time-frequency domain to detect the desynchronizations of certain frequency bands (3-7Hz, 8-13 Hz, 14-20Hz 21-29Hz and 30-50Hz) in the temporal cortex as an indication of motor imagery.

INTRODUCTION

During the initial stages of the design process, more specifically at the conceptual design stage, designers and engineers most often make use of the paper and pencil approach to illustrate their ideas and thoughts. Paper-based sketching gives designers the freedom to creatively express their ideas while at the same time allows them to rapidly generate a wide variety of concepts. Due to this factor, a vast amount of research involving computer-aided sketching systems has been done to provide designers with reliable sketching platforms that offer the same amount of comfort and freedom as paper-based sketching. Most often, after creating conceptual design elements, designers might find the need to convert their raw 2D sketches to refined line drawings which can be used in Computer Aided Design (CAD) programs. The use of CAD presents a medium for the analysis of generated concepts. Thus, combining the strengths of paper based sketching with CAD could help to provide a platform for flexible sketching and modeling in the early stages of design [1].

Creating 3D objects at the conceptual design stage can be done in two different approaches: 1- Creating or modifying 3D objects 2- Reconstructing 3D shapes from 2D drawings. “SKETCH” [2], “Teddy” [3] and “Hypermesh” are examples of 3D sketch based systems that are centered around the first approach, while a good example of 3D modeling based on the second approach is a freehand sketching interface for progressive construction of 3D objects developed by Masry et al [4]. This interface was based on a reconstruction algorithm that assigned depth values to each vertex, and subsequently reconstructed each stroke in a two stage reconstruction process. “The first stage tests a straight line sketch extracted from the original for the presence of prevailing angular trends and uses this information to determine a 3D axis system that maps 2D lines onto 3D lines. The second stage reconstructs curved strokes, under the assumption that they are planar.” [4] After the reconstruction process, the 3D object’s faces are identified and triangulated, after which users can then add additional strokes, and sketch directly onto the object’s faces, or even perform transformations such as rotation and translation of the object. While this interface is suitable for creating simple 3D surface objects, one of the main limitations of the algorithm is the fact that the user cannot perform

* These authors have equal contributions to this manuscript.
transformations such as 3D translations along the primary axes of the modeling environment.

Kara et al presented a study related to the surface skinning of 3D curve clouds drawn in a free form for conceptual shape design [5]. Their method takes into account scattered geometrical and topological data to produce an approximate surface by connecting the curve clouds which are in the form of 3D sketches. They use a guidance vector field calculated from the input curve clouds to bridge the gap between them and induce a surface without altering any of the geometrical or topological information [5]. Research has also been done to incorporate augmented reality environments with CAD packages for the creation and modification of free form surfaces through 3D sketch-based inputs [6]. Results have proved that product design systems using augmented reality environments can be efficiently used for the construction of free form surfaces by utilizing multiple shape representations such as point sets and analytical surfaces. Thus, with the help of multiple representations, additional modeling capabilities are provided. The inclusion of head mounted displays and glove-based interfaces has proved to aid in the fusion of virtual and CAD environments [7]. A study on 3D shape creation for industrial styling design using sketch-based inputs revealed that a rapid and intuitive design procedure can be practically carried out with free form curves and surfaces. The strategy includes processing initial wire-frame inputs so as to convert them into surfaces used for the styling design [7].

Thus, while a number of applications exist to model 3D objects, efficient applications need to also be developed to aid in intuitive manipulation of these objects. This task becomes especially difficult when the user-interface is a 2D interface since a large amount of ambiguity results by using 2D gestures to manipulate an object in 3D space. Due to this problem, research is being done to design multimodal interfaces to create 3D CAD models, in an attempt to reduce the ambiguity that results from using 2D gestures to create 3D models.

Sharma et al designed “MozArt”, a prototype interface that explores how multimodal inputs can be used to simplify conceptual 3D modeling for first-time CAD users [8]. “MozArt features a minimalist UI, and uses a touch table whose orientation can be changed depending on the needs of the user” [8]. This interface incorporates touch and speech interactions that are used for 3D modeling. While MozArt provides users with efficient results, the main problem with this interface was that an open microphone was used for speech input which resulted in false positives (spurious input) due to ambient sounds, especially in multi-user environments [8]. Due to these disadvantages with incorporating speech interactions in multimodal interfaces, we propose a system that is based on gesture control and information obtained from brain signals of users in order to reduce the ambiguity associated with 2D gestures to produce 3D object transformations.

The remainder of this paper is organized as follows. The next section describes how brain computer interfaces (BCI) operate, and provides details of related research that has been done to utilize these interfaces for CAD modeling. The section after that discusses the procedure that was adopted for data collection. This is followed by a description of the method that was used for data analysis, along with the results that were obtained. The final section concludes the paper.

**BRAIN COMPUTER INTERFACE**

Traditional input devices such as the keyboard and mouse have largely been used in CAD systems. The inception of pen and touch based systems along with haptic devices used in virtual environments and brain-computer interface systems however, has provided users with more intuitive and interactive CAD systems. Research and development in BCI have helped in incorporating the human thinking process in virtual environments. BCI produces a communication link between the thoughts of a user and the output device or system without taking into account motor output pathways such as physical movements or expressions. BCI systems identify and relate thinking patterns generated in brain signals of users to their intentions and thoughts. Most BCI systems involve recording electroencephalography (EEG) signals generated by placement of electrodes on the human scalp. Commercial advancements also aid in creating immersive virtual environments in which users can manipulate virtual objects by means of their brain activity.

With regard to motor imagery i.e. the imagery related to motion, signals are acquired during imagined sensorimotor rhythms (SMR). These SMRs are primarily detected based on the features of μ (8-13 Hz) and β (18-25) rhythms [9]. Changes in amplitudes of these frequency bands give rise to what is known as Event Related Desynchronization (ERD) and Event Related Synchronization (ERS). ERD is characterized by the decrease in the amplitudes of the frequency bands, whereas ERS results due to the increase in amplitudes of the frequency bands. The desynchronization of rhythms takes place due to perception of preparation of movement or the movement itself, and the synchronization of rhythms takes place after the movement and with relaxation [10]. It has been proved that applications such as the control of the 2D motion of a cursor on a screen can be accomplished by using SMR based BCI [11, 12].

Recent studies have been done on the use of BCI in CAD environments, the purpose of which was to utilize the user’s brain activity for the creation and modification of various CAD elements by engaging visual imagery in the design process. These environments allow the user to create primitive sketches and shapes, modify these shapes and move the shapes by translating and rotating them. In addition, users can also make use of viewing options such as zooming and panning. A study by Esfahani and Sundararajan [13] reported that the BCI can be efficiently used to distinguish between primitive shapes such as cubes, spheres, cylinders etc. that are imagined by users. In their study, data was collected in the form of brain activity from 14 locations on the human scalp and was analyzed with the help of Independent Component Analysis (ICA) and the Hilbert-Huang Transform (HHT). Research has also been done to evaluate the steady state visual evoked potential (SSVEP) as a feedback mechanism to examine the mental state of the user during motor imagery [14]. In this study carried out by Wang et al., subjects were asked to imagine the movement of flashing object in a particular direction. If the subject is mentally engaged in the task, the
SSVEP signal will be detectable in the visual cortex of the brain and therefore the motor imagery task can be confirmed. In addition, studies performed by Sree Shankar S. et al involved the recording of EEG and electromyogram (EMG) signals in order to activate and control different commands of a CAD package [15]. A few important considerations that should be taken into account for the use of BCI in CAD systems are:

1. **Geometry representation:** Before beginning the modeling process, the user must have a good mental judgment of the shape and size of the object, and must be able to gauge the dimensions and proportions of the object appropriately to generate the desired shape.

2. **Object manipulation:** By way of visual imagery, the BCI should be able to accurately recognize and carry out object manipulation as intended by the user so as to get the desired result. BCI could be precisely able to locate and orient the objects.

3. **Training data:** Training data should be made available to the BCI so as to enable the interface to recognize various kinds of operations in addition to user intentions.

**EXPERIMENT (MATERIAL AND METHODS)**

In order to understand the ambiguity that is associated with modeling a 3D object using a 2D user interface, consider the examples illustrated in Figure 1. The arrows marked on the figure represent gestures that are made on the user interface. Due to the 2D nature of the gestures, it is highly possible that the effect that they produce is one that is contradictory to the true intention of the user. By looking at the gesture indicated on Figure 1 (A), one possible outcome could be that the user desires to move the middle slab of the object in a downward direction, while keeping the remaining faces of the object stationary. Yet another possible outcome could be that the user desires to move the selected part in a downward and forward direction in 3D space. Similarly, consider the gesture indicated on Figure 1 (B). One possible interpretation of the gesture could be that the object should be translated along the positive x-axis as indicated in the figure. Another possible interpretation could be that the object should be rotated about the y-axis in a counter-clockwise direction as seen from above the model. From these various possible outcomes, it is clear that gesture modeling of a 3D object on a 2D surface has a great deal of ambiguity that is associated with it. The source of this ambiguity lies in the estimation of the depth of movement of the 2D gesture.

The experiments described in this section of the paper aim at addressing the problems associated with this ambiguity by investigating the possibility of incorporating a multi-modal user-interface that comprises sketch based modeling as the primary modality and BCI based modeling as a secondary modality. Our main hypothesis was based on the notion that if BCI could help identify the true intentions of the users of a 2D gesture-based interface, the ambiguity associated with these gestures could be significantly reduced, thus resulting in an efficient and intuitive multi-modal interface.

To verify our hypothesis, we conducted a series of experiments on two subjects. Both subjects had a background in engineering, and minimal experience with brain-computer interfaces. A total of 2 experiments were carried out, each comprising two tasks that were given to the subjects. These tasks involved simple 3D transformations of the object shown in Figure 1. Each task took an average of 15-20 seconds to perform, with a pause of 10-20 seconds in between tasks. Before commencing the experiments, the users were made to wear a neuroheadset that was used to detect specific brain patterns of the subjects while performing the experiments. The headset used was the Emotiv EPOC - a high resolution, multi-channel, wireless neuroheadset that uses a set of 14 sensors plus 2 references to tune into electric signals produced by the brain to detect the user’s thoughts, feelings and expressions in real time. The channel names based on the international 10-20 locations are: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. In the experiments conducted, the headset was used to obtain EEG signals from each subject for data analysis.

![Figure 1: A) AMBIGUITY IN DIRECTION OF TRANSLATION [+Z VS. -Y] B) AMBIGUITY BETWEEN TRANSLATION AND ROTATION](image)
tasks are very similar to each other. In both cases, the subjects used gestures similar to the arrow depicted in Figure 1 (B). Since two similar gestures were used to perform two separate transformations, these two tasks were ideal to carry out our analysis about how BCI could help reduce the ambiguity between these two outcomes.

Figure 2: Subject performing Experiment 1

In the second experiment, the subjects were provided with a left view of the same 3D object. The first task under this experiment was to translate the object by a variable distance along the positive y-axis. The second task involved rotating the object about the x-axis in a clockwise direction as seen from the right of the given view of the model. Once again, markers were used to indicate the start and end points of the tasks, and the EEG signals were recorded. These two tasks were chosen due to the same nature of ambiguity that results between similar gestures that are used to perform two different transformations.

DATA ANALYSIS

Data analysis primarily involved the study of EEG data from the 14 channels of the Emotiv EPOC neuroheadset that were placed on the scalp of the users. The analysis has been segmented into the following main sections:

1. Artifact Removal

In the experiments that were carried out, it is found that the data collected from the EEG brain signals of the user performing the task, is affected by various artifacts such as the eye blinks, closing of eyes and muscle movements. But, there is an effective way to remove the muscle movements, which mostly lie in the frequency bands of order higher than 30 Hz, by applying a low pass filter. Therefore, the main artifacts which we needed to deal were the eye blinks and eye movements.

A combination of Independent Component Analysis (ICA) and Empirical Mode Decomposition (EMD) was used for artifact removal. The objective of ICA is to represent the recorded EEG signals \([X = x_1(t), x_2(t) \ldots x_{14}(t)]\) as a linear combination of statistically independent sources \([S = S_1(t), S_2(t) \ldots S_{14}(t)]\) as shown in Equation (1):

\[
X = W*S
\]  

‘W’ is a weighting matrix and is calculated by minimizing the mutual information between the calculated sources. Usually, the EEG signals are cleaned by removing the artifact related independent component. But one of the major drawbacks of this approach is that independent components taken for removal may also contain some informative EEG activities. So, complete removal of the selected ICs leads to the loss of EEG data. Hence, to avoid this, we use the EMD to clean only the artifacts and keep the EEG data untouched.

The EMD decomposes the contaminated EEG source into a set of Intrinsic Mode Functions (IMFs) and uses a user defined threshold to obtain and remove artifact related components in that independent components\([17]\). The remaining IMFs are then summed together to form artifact-free sources from which the artifact-free EEG signals are obtained.

2. Feature Extraction

The artifact-free data was then analyzed by forming 7 symmetric channel pairs from the 14 primary channels. In addition, the following five frequency bands were considered: Theta (3-7 Hz), Alpha (8-13 Hz), Lower-beta (14-20 Hz), Upper-beta (21-29 Hz) and Gamma (30-50 Hz).

The power spectral density (PSD) of the brain signals in each channel was then calculated for both users. The PSD of a signal describes how the power of the signal is distributed over different frequencies. The PSD was calculated by computing the Fast Fourier Transforms (FFT) of the EEG signals. FFT was used to convert the collected data that was originally in the time domain to the frequency domain.

The PSD values of the EEG signals were used as an indication of ERS and ERD, which were ultimately used to classify the tasks that the subjects were performing. Feature selection was then done using the following criteria which was effectively put forward by Palaniappan [18].

\[
Power_{\text{diff}} = (P_i - P_j)/(P_i + P_j)
\]  

where, \(P_i\) and \(P_j\) are the power spectral densities of signals in symmetric channels \(i\) and \(j\) in the \(b^{th}\) spectral band.

Hence, by considering 7 data channels and 5 frequency bands, we were able to arrive at 35 features to use for classification.
3. Classification

As described in the previous section, each experiment that was performed involved one translation and one rotation based task. Four classes were formed based on these tasks:

- Class 1 – Translation along positive x-axis
- Class 2 – Rotation about y-axis
- Class 3 – Translation along positive y-axis
- Class 4 – Rotation about x-axis

The primary motive to resolve the ambiguity was the correct classification between translation and rotation gestures made by the users. A binary classifier was used to classify between the extracted features. We experimentally found that the classification was dependent upon the subjects performing the tasks. Hence, for each of the subjects, certain best features were computed, and these features were used for the classification of tasks performed by the user.

RESULTS AND DISCUSSIONS

The graph shown in figure 4 depicts the PSD of the EEG signals of the seven symmetric channels for subject 1 while performing experiment 1. The green markers indicate the start of each task in the experiment, and the red markers indicate the completion of the tasks. Areas represented in white represent high values of PSD, whereas areas in black represent low PSD values.

![Figure 4 PSD OF ROTATION AND TRANSLATION](image)

By analyzing the graph it can be seen that the PSD values indicate brain activity that could most likely be attributed to motor imagery. Binary linear and quadratic classifiers were then used to classify between translation and rotation-based tasks based on features that were generated by calculating the PSD of 7 symmetric channel pairs for 5 different frequency bands. The initial results of the classification between classes 1 and 2 are shown in figure 5. This plot indicates the classification accuracy before artifact removal for 30 training samples and 30 testing samples for each class.

A t-test was used to arrive at the twelve best features to use for the classification. Based on the plot obtained in figure 5, the classification accuracies for both users were found to range between 60 – 65%.

In order to improve the accuracies of the classifiers, the same classification was carried out on artifact-free data with reference to only the Alpha band, and the upper and lower Beta bands. The resulting plot is shown in figure 5. It was found that after the removal of artifacts from the EEG signals obtained from the users, the classification accuracies for both users increased to 70%.

Based on the results obtained, a confidence interval was then calculated in order to arrive at an upper and lower bound for the accuracy that could be obtained by evaluating the data based on 30 testing samples.

$$e_T = (e_s \pm Z_n \sqrt{(e_s(1-e_s)/N})$$

(3)

$$\text{Confidence Interval} = 1 - e_T$$

(4)

where $Z_n$ is the z-score based on a normal distribution, $e_s$ is the sample error, $e_T$ is the true error and N is the total number of testing samples.

The sample error was calculated for a 90% confidence interval ($Z_n = 1.645$) based on which, an upper bound of 83.76% in accuracy was found to be obtainable along with a lower bound of 56.24%.
CONCLUSION

Our paper describes a novel method to resolve the ambiguities that result from using 2D gestures on a touch-based system. The experiments that were performed included translation and rotation-based transformations of a 3D CAD model using 2D gestures. In our approach, we used BCI to help identify the true intentions of the users while performing the aforementioned tasks. The BCI helped identify motor imagery-based brain activity that could be attributed to the object manipulations performed by the subjects.

In the analysis of the EEG data collected from the subjects performing the experiments, we obtained a classification accuracy of 70% for both subjects for differentiating between rotation and translation-based tasks after the removal of the artifacts from the brain signals.

The work presented in this paper is based on ongoing research that is aimed at incorporating BCI into multimodal interfaces for 3D CAD modeling. Future work will include conducting more experiments on a large number of subjects in order to obtain more data for analysis. Other improvements also include performing adaptive boosting to further improve the classification accuracy and conducting the same experiments with the help of a more advanced BCI headset. Motor imagery can be efficiently detected when placing sensors around the motor cortex of the user’s brain. Since the Emotiv EPOC is not fully equipped with these sensors, the results obtained can be further improved by utilizing a BCI headset that has the required sensors to collect this data.

REFERENCES


