A VIRTUAL REALITY TESTBED FOR BRAIN–COMPUTER INTERFACE RESEARCH*

J. D. Bayliss and D. H. Ballard
Department of Computer Science
University of Rochester

1. The System

Recent BCI work has shown the feasibility of on–line averaging and biofeedback methods in order to choose characters or move a cursor on a computer screen with up to 95% accuracy [McFarland et al., 1993; Pfurtscheller et al., 1996; Vaughn et al., 1996; Farwell and Donchin, 1988]. Virtual reality (VR) promises to extend the realm of possible prototypes through allowing subjects to interact directly with the environment rather than a computer monitor while still maintaining environmental control. Furthermore, the safety of VR makes it an excellent candidate for BCI research on tasks such as driving.

The VR environment is rendered on a SGI Onyx with 4 R10,000 processors and an Infinite Reality graphics engine. A flexible program for graphics rendering enables researchers to easily switch environments. For immersion, subjects wear a head–mounted display (HMD) containing an eye tracker.

The heart of this system is a NeuroScan commercial package for EEG signal acquisition (called Acquire). After the EEG signal and trigger codes enter the Acquire program, they are grabbed from the acquisition buffer via a dynamic linked library (DLL) provided by NeuroScan. This library enables the locally written software to have access to the unprocessed data and trigger codes. The DLL is called from within a recognition and feedback program. This program chooses which data need to be sent for further processing via the Matlab program. This program may give audio feedback to the user after recognition occurs, send return information to the SGI through a serial port interface, save recognition data, calculate whether recognition has actually occurred (using trigger codes), and can read previously processed data from a Matlab file for a demonstration of the speed of recognition.

In order to enable easy use of different recognition routines, all routines are Matlab m–files. While compiled programs are faster than m–files, we have not had a problem with speed and find the general interface encourages the use of new computer algorithms for processing.

2. Assessment of Results

Several ways of assessing results are available. The most obvious is to analyze the EEG signals after a session. We use the NeuroScan analysis package as well as several locally written Matlab routines. On–line single trial EP recognition via different algorithms enables a direct assessment of BCI recognition abilities. During a session we record all visual data to videotape.

Our lab is also equipped with an eye tracker in the VR HMD. While the eye tracker is not necessary for BCI research, it may enable better analysis of results since subjects tend to look at what they're thinking about. It also allows comparisons between BCIs and eye tracking for particular subjects.

3. The Task

VR allows subjects to make on–line decisions in a dynamic environment. Thus, the best tasks for this environment involve interaction with physical objects. The flexibility of the VR environment allows a concentration on interface issues before building a BCI in the real–world.

To this end, we have used two environments; a driving environment to look at on–line driving issues and a two

*This research was supported by NIH/PHS grant1–P41–RR09283. It was also facilitated in part by a National Physical Science Consortium Fellowship and by stipend support from NASA Goddard Space Flight Center

bedroom apartment in order to look at issues related to controlling appliances automatically as well as simple speech (see Figure 1).

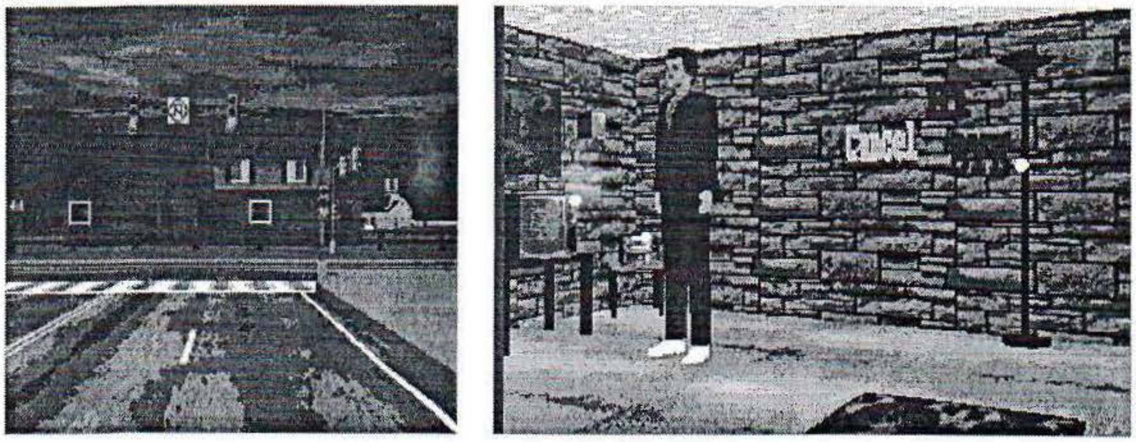

Figure 1. (Left) A typical stoplight scene in the virtual driving environment. (Right) The living room of a virtual apartment.

3.1. EEG Components Used

In order to test the feasibility of on–line recognition in the noisy VR environment, we recognized the P3 EP, discovered by [Chapman and Bragdon, 1964; Sutton et al., 1965]. It is a positive waveform occurring approximately 300–450 ms after an infrequent task–relevant stimulus.

Previous P3 research has concentrated primarily on static environments such as the continuous performance task [Rosvold et al., 1956]. In the visual continuous performance task, static images are flashed on a screen and the subject is told to press a button when a rare stimulus occurs or to count the number of occurrences of a rare stimulus. This makes the stimulus both rare and task relevant in order to evoke a P3. As an example, given red and yellow stoplight pictures, a P3 should occur if the red picture is less frequent than the yellow and subjects are told to press a mouse button only during the red light.

3.2. The Stoplight Experiment

We assumed a similar response would occur in a VR driving world if red stoplights were infrequent and subjects were told to stop their virtual cars at the red light. In order to make yellow lights more frequent, both green and red lights were preceeded by yellow lights. Red lights change to green after 3 seconds and the red light condition is triggered only when subjects are close to the stoplight so that subjects will have to begin stopping when the red light is triggered.

The subjects used a modified go cart in order to control the virtual car. We chose go cart driving because it is more like a "natural" driving task than driving and stopping with a mouse. While this choice may cause a more artifacts in the signal collection (due to turning the steering wheel and braking), most of the actual artifact in the data was discovered to be due to eye movement.

A trigger pulse containing information about the color of the light was sent to the EEG acquisition system whenever a light changed. While an epoch size from –100 ms to 1 sec was specified, the data was recorded continuously. Information about head position as well as gas, braking, and steering position were saved to an external file.

Eight electrodes sites (FZ, CZ, CPZ, PZ, P3, P4, as well as 2 vertical EOG channels) were arranged on the heads of five subjects with a linked mastoid reference. Electrode impedances were between 2 and 5 kohms for all subjects.

The EEG signal was amplified using Grass amplifiers with an analog bandwidth from 0.1 to 100 Hz. Signals were then digitized at a rate of 500 Hz and stored to a computer.

In order to determine that the P3 EP occurred only at red stoplights, we calculated the averages over red light and yellow light trials with trials where the subject ran a red light (approximately 2 per subject) removed. As expected, the data obtained while driving contained artifacts. In order to reduce these artifacts before averaging, we preprocessed the data and subtracted a combination of eye and head movement artifact using the linear regression technique described in [Semlitsch et al., 1986]. Results show that a P3 EP indeed occurs at red and not yellow lights [Bayliss and Ballard, 1999].

Table 1. Recognition Results ($p < 0.01$)

	Robust Kalman Filter %Correct		
Subjects	Red	Yel	Total
S1	55	86	77
S2	82	94	90
S3	74	85	81
S4	65	91	82
S5	78	92	87

Table 2. Return Subject Recognition Results

	Robust K- Filter %Correct		
Subjects	Red	Yel	Total
S4	73%	90%	85%
S5	67%	87%	80%

3.2.1. Results

While averages show the existence of the P3 EP at red lights and the absence of such at yellow lights, we needed to discover if the signal was clean enough for single trial recognition as the quick feedback needed by a BCI depends on quick recognition. We tried four methods for classification of the P3 EP: correlation, independent component analysis (ICA), a Kalman filter, and a robust Kalman filter. While all algorithms performed significantly better than correlation with the light averages ($p < 0.01$), we will only report the results of the best algorithm, the robust Kalman filter in Table 1. Approximately, 90 yellow light and 45 red light trials from each subject were classified. The reason we allowed a yellow light bias to enter recognition is because the yellow light currently represents an unimportant event in the environment. In a real BCI unimportant events are likely to occur more than user–directed actions, making this bias justifiable.

Data was preprocessed with the method described in the previous section. We used the robust Kalman filter framework formulated by Rao [Rao, 1997]. The robust Kalman filter is trained using red and yellow light averages from the maximal electrode site for obtaining the P3 for each subject. We used the whole trial epoch for recognition because it yielded better recognition than just the time area around the P3.

In order to look at the reliability of the robust Kalman filter two of the Subjects (S4 and S5) returned for another VR driving session. The results of this session using the robust Kalman Filter trained on the first session are shown in Table 2. The recognition numbers for red and yellow lights between the two sessions were compared using correlation. Red light scores between the sessions correlated fairly highly – 0.82 for S4 and 0.69 for S5. The yellow light scores between sessions correlated poorly with both S4 and S5 at around –0.1. This indicates that the yellow light epochs tend to correlate poorly with each other due to the lack of a large component such as the P3 to tie them together.

3.3. The Apartment Environment

The stoplight experiment showed that EPs could be reliably detected in a VR environment. In order to take full advantage of the benefits of VR, a two bedroom virtual apartment was constructed. Various items in the apartment perform a function. For instance, the light, tv, and stereo may turn on/off. Simple verbal utterances such as "hi" and "bye" may be said to the graphical figures in the apartment. In order to enable users to back up if they choose a wrong option, a "cancel" option has been installed.

3.3.1. An Apartment Communication Protocol

Since single trial P3 epochs are used in picking an option, we have adopted an interface design similar to that proposed by Farwell and Donchin [Farwell and Donchin, 1988]. In order to evoke a P3 they have constructed a matrix of flashing options. The idea behind this is that if a user wants to pick an option, (s)he will look at the box and a P3 will be evoked when the box infrequently flashes.

Our design differs in that the flashing "buttons" are actually connected to the objects they control or are on the wall of the apartment if there is no physical object involved in the command. In this way, the commands available at any one time are dependent on the context of the environment and the user may ignore certain commands by attending to another part of the room.

One basic reason for constructing this environment is that the user interface of the BCI may greatly affect recognition ability. A pilot experiment in the apartment seems to confirm this. Six different flashing options (detailed above) were available and flashed in a round robin fashion at approximately 1 every 4 seconds.

In a pilot experiment 10 tasks were attempted and 9 were completed. The average time for the completion of one simple task was 2.8 minutes. Now, at first glance this number appears horrible until one looks at the completion time for the tasks individually. If the two stereo commands accomplished are removed from the average, then the average completion time for one task drops to 1.1 minutes. The subject could not seem to pick the stereo command and eventually gave up trying to pick it on the tenth task. The reason for this appears to be the location/size of the stereo object.

References

[Bayliss and Ballard, 1999] J.D. Bayliss and D.H. Ballard, "Single Trial P300 Recognition in a Virtual Environment," *Computational Intelligence: Methods & Applications Proceedings, Soft Computing in Biomedicine*, June 1999.

[Chapman and Bragdon, 1964] R.M. Chapman and H.R. Bragdon, "Evoked responses to numerical and non–numerical visual stimuli while problem solving.," *Nature*, 203:1155--1157, 1964.

[Farwell and Donchin, 1988] L.A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event–related brain potentials," *Electroenceph. Clin. Neurophysiol.*, pages 510-523, 1988.

[McFarland et al., 1993] D.J. McFarland, G.W. Neat, R.F. Read, and J.R. Wolpaw, "An EEG–based method for graded cursor control," *Psychobiology*, 21(1):77-81, 1993.

[Pfurtscheller et al., 1996] G. Pfurtscheller, D. Flotzinger, M. Pregenzer, J. Wolpaw, and D. McFarland, "EEG–based Brain Computer Interface (BCI)," *Medical Progress through Technology*, 21:111-121, 1996.

[Rao, 1997] R. P.N. Rao, "Kalman Filter Model of the Visual Cortex," *Neural Computation*, 9(4), 1997.

[Rosvold et al., 1956] H.E. Rosvold, A.F. Mirsky, I. Sarason, E.D. Bransome Jr., and L.H. Beck, "A Continuous Performance Test of Brain Damage," *J. Consult. Psychol.*, 20, 1956.

[Semlitsch et al., 1986] H.V. Semlitsch, P. Anderer, P. Schuster, and O. Presslich, "A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP," *Psychophys.*, 23:695-703, 1986.

[Sutton et al., 1965] S. Sutton, M. Braren, J. Zublin, and E. John, "Evoked potential correlates of stimulus uncertainty," *Science*, 150:1187-1188, 1965.

[Vaughan et al., 1996] T.M. Vaughan, J.R. Wolpaw, and E. Donchin, "EEG–Based Communication: Prospects and Problems," *IEEE Trans. on Rehabilitation Engineering*, 4(4):425-430, 1996.