5-1-2010

NeuroPhone: Brain-Mobile Phone Interface using a Wireless EEG Headset

Matthew K. Mukerjee
Dartmouth College

NeuroPhone: Brain-Mobile Phone Interface using a Wireless EEG Headset

Matthew K. Mukerjee
Dartmouth College, Hanover, NH, USA

Dartmouth Computer Science Technical Report TR2010-666

ABSTRACT

Neural signals are everywhere just like mobile phones. We propose to use neural signals to control mobile phones for hands-free, silent and effortless human-mobile interaction. Until recently, devices for detecting neural signals have been costly, bulky and fragile. We present the design, implementation and evaluation of the *NeuroPhone* system, which allows neural signals to drive mobile phone applications on the iPhone using cheap off-the-shelf wireless electroencephalography (EEG) headsets. We demonstrate a mind-controlled address book dialing app, which works on similar principles to P300-speller brain-computer interfaces: the phone flashes a sequence of photos of contacts from the address book and a P300 brain potential is elicited when the flashed photo matches the person whom the user wishes to dial. EEG signals from the headset are transmitted wirelessly to an iPhone, which natively runs a lightweight classifier to discriminate P300 signals from noise. When a person's contact-photo triggers a P300, his/her phone number is automatically dialed. NeuroPhone breaks new ground as a brain-mobile phone interface for ubiquitous pervasive computing. We discuss the challenges in making our initial prototype more practical, robust, and reliable as part of our on-going research.

1. INTRODUCTION

Like mobile phones, neural signals are ever present in our everyday lives. Given the recent availability of low-cost wireless electroencephalography (EEG) headsets [3,12,13] and programmable mobile phones capable of running sophisticated machine learning algorithms, we can now interface neural signals to phones to deliver new mobile computing paradigms—users on-the-go can simply "think" their way through all of their mobile applications.

In this paper, we present the design, implementation and evaluation of the *NeuroPhone* system (see video demo [2]), a brain-mobile phone interface based on the wireless Emotiv EPOC EEG headset [3] and the iPhone. We demonstrate a mind-controlled address-book dialing app, which works on similar principles to a P300-speller [8] brain-computer interface: the phone flashes a sequence of photos of contacts from the address book and a P300 brain potential is elicited when the flashed photo matches the person whom the user wishes to dial. We also demonstrate a version of the same app which detects the much larger and more easily detectable EEG signals triggered by the user winking their eyes when the target photo appears. This "wink"-triggered dialing works robustly in noisy conditions. The P300, or "think"-triggered, dialer is very promising but at present less reliable. One could argue that other "hands off" types of actuation such as voice recognition is more suitable an interface to mobile applications. However, our goal is to best understand how firing neurons can drive mobile applications and what the current limitations in the state of the art are when using off-the-shelf wireless EEG headsets and phones.

In this paper, we discuss our broader vision of a brain-mobile phone interface and then present the initial design, implementation, and evaluation of the NeuroPhone system. Our initial results look promising showing that the iPhone is capable of processing raw neurosignals and classifying the P300 using a cheap, noisy commercial EEG headset. However, a number of challenges remain in developing a practical and robust brain-mobile phone interface not only capable of working in controlled laboratory settings but also out in the wild. Addressing these challenges is part of our on-going research.

2. BRAIN-MOBILE PHONE INTERFACE

We envision that many mobile applications can be reinvented; for example, instead of hand dialing your friend Tim while driving you can simply wink or think of him while your phone displays your contacts. We also imagine new many-to-one mobile applications; for example, a teacher of a foreign language is interested in seeing exactly how many students actually understood the last question she asked. The students are all wearing EEG headsets and their data is being streamed in real-time to the teacher's mobile phone. She simply takes out her mobile phone and it gives her up to the second statistics on each of her students. She quickly glances at the aggregate class statistics and realizing that the students really did understand her difficult question, proceeds with her lecture. Other scenarios may soon be possible; for example, a person enters a room (e.g., bar, club, meeting, classroom) and instantly has a sense of the overall emotional state of the space (i.e., happy, tension, frustration, sad, bored, hostile). There is prior work classifying EEG signals into different bands of frequencies corresponding to different emotions such as meditation and activity [10]. In addition, the Emotiv headset [3], which is designed primarily for gaming purposes, is also capable of detecting certain facial expressions (e.g., smile, laugh, shock – eyebrows raised, anger – eyebrows furrowed) and non-conscious emotions. If one could read the emotional state of people moving through a building then the notion of mood music would take on a literal sense.

Many practical challenges remain to make this vision a reality. For example, research-grade EEG headsets [5] are expensive (e.g., tens of thousands of dollars) but offer a much more robust signal than the cheaper (e.g., $200-$500) headsets. As a result there is a significant amount of noise in the data of the cheaper headsets, requiring more sophisticated signal processing and machine learning techniques to classify neural events (e.g., P300). However, the cheaper headsets provide an encrypted wireless interface between the headset and computer allowing for mobility but complicating the design of a clean brain-mobile phone interface. Mobile phones are not designed to support continuous neural sensing applications. The energy cost of continuously streaming raw neural signals over the air interface and running classifiers on the phone is challenging. We imagine that brain-mobile phone interfaces will be used when and where the user is: walking in a busy street, in a car, on a bicycle, while shopping, sitting quietly in a library, etc. We show that many of these use cases present significant noise artifacts in the data complicating the design of a practical brain-mobile interface today. Filtering out components of the signal associated with artifacts (e.g., neural signals associated with walking or unintentional facial expressions) is needed to advance this vision.

We envision that wireless EEG headsets will become cheaper and more robust and that machine learning techniques developed for high end research-grade wired EEG headsets [5] can be effectively exploited by resource limited phones. As this vision gathers speed and noise issues are solved, EEG will be integrated into wearable fabric (e.g., baseball caps, woolen hats, bicycle helmets) or become the new wireless "earphones plus" (i.e., earphones plus a limited set of electrodes). This raises a number of interesting issues. For example, the NeuroPhone system relay (discussed later) transmits raw unencrypted neural signals over-the-air to the iPhone in IP packets. This leads to the notion of insecure "neural packets everywhere," opening up important privacy challenges that need to be addressed.

Figure 1: NeuroPhone in use

3. NEUROPHONE DESIGN

We create the NeuroPhone as a means of taking a first step towards this vision. The NeuroPhone system uses the iPhone to display pictures of contacts in the user's address book. The pictures are displayed and individually flashed in a random order. The user concentrates on the picture of a person s/he wishes to call in the case of the think mode of our application, called "Dial Tim". Utilizing the P300 neural signal, NeuroPhone recognizes which person the user is focused on and calls them. The wink mode is similar to the think mode where the user simply winks with the left or right eye to make the intended call. The wink mode relies on the much more clearly defined muscle movement signals in the raw EEG data, than the much more subtle neural signals [14]. Figure 1 shows a user with the headset and phone, and Figure 2 shows the application running. In what follows, we present an overview of the P300 signal and the wireless Emotiv EPOC EEG headset used by our Dial Tim application. We also discuss a number of design considerations that directed our initial implementation discussed in Section 4.

3.1 P300

When somebody concentrates on a task-specific stimulus (e.g., a highlighted image in Dial Tim) among a pool of stimuli (e.g., non highlighted images), the task-related stimulus will elicit a positive peak with a latency of about 300ms from the stimulus onset in subject's EEG signal. This positive peak is known as the P300 signal in neuroscience literature [8]. P300 is emanated in the central-parietal region of the brain and can be found more or less throughout the EEG on a number of channels. Figure 3 illustrates such a P300 signal captured using our headset, where the signal is bandpass filtered and averaged over multiple trials. A classic example experiment driven by P300 signals is the

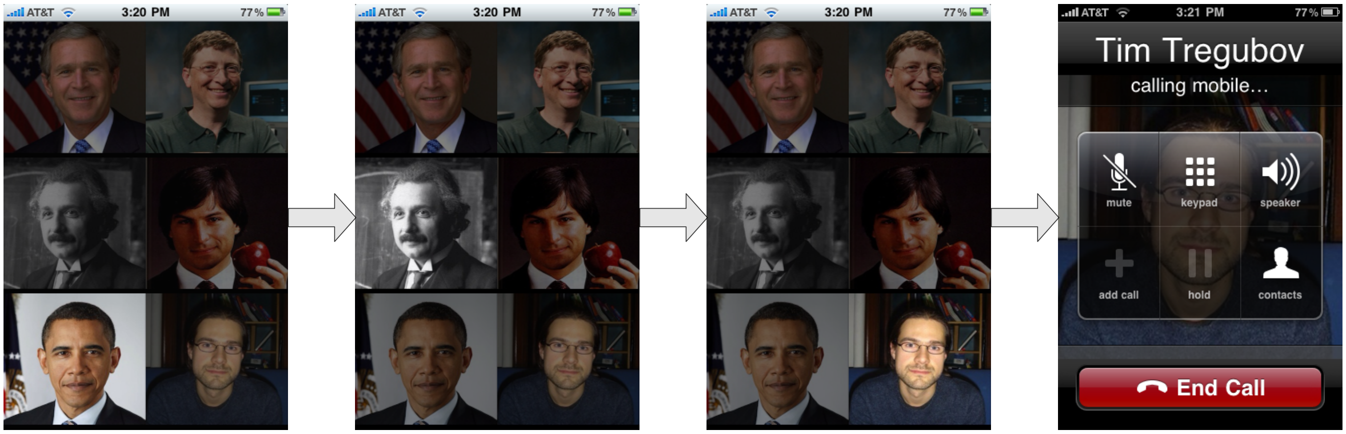

Figure 2: The Dial Tim application works on similar principles to P300-speller brain-computer interfaces: the phone flashes a sequence of photos of contacts from the address book and a P300 neural signal is elicited when the flashed photo matches the person whom the user wishes to dial. EEG signals from the headset are transmitted wirelessly to an iPhone, which natively runs a simple classifier to discriminate P300 signals from noise. When a person's contact-photo triggers a P300, their phone number is automatically dialled. In this case, the user wants to dial Tim, thus when his picture is flashed, Tim is automatically dialled.

P300 speller [4]. A grid of 6×6 alphanumeric characters is presented to a subject. The subject focuses on a specific character, while the rows and columns are randomly flashed. Whenever a row or column containing that specific character flashes, a P300 signal is elicited in the subject's EEG. The speller then predicts the specific character that the subject intends to select by determining the row and column that correspond to P300 signals in subject's EEG and takes the letter at the intersection point. While we focus on the P300 neural signal as a driver of the Dial Tim application we plan to study the suitability of other neural signals as part of on-going work.

3.2 Wireless EEG Headset

We use the Emotiv EPOC headset [3] which has 14 data-collecting electrodes and 2 reference electrodes (see Figures 5, 6, and 1). The electrodes are placed in roughly the international 10-20 system and are labeled as such [10]. The headset transmits encrypted data wirelessly to a Windows-based machine; the wireless chip is proprietary and operates in the same frequency range as 802.11 (2.4Ghz). The software that comes with the Emotiv headset provides the following detection functionalities: various facial expressions (referred to as "Expressiv" by Emotiv); levels of engagement, frustration, meditation, and excitement ("Affectiv"); subject-specific training and detection of certain cognitive neuro-activities such as "push", "pull", "rotate", and "lift" ("Cognitiv") [3]. Also built in the headset is a gyroscope that detects the change of orientation of subject's head. However, the headset is not meant to be an extremely reliable device, thus it is challenging to extract finer P300 signals from the EEGs this headset produces. But, as we stated in our vision, this headset can be easily deployed at large scale because of its low price, and can be extremely handy if we can extract useful signals (e.g., P300) from it through smart signal processing and classification algorithms running on the phone.

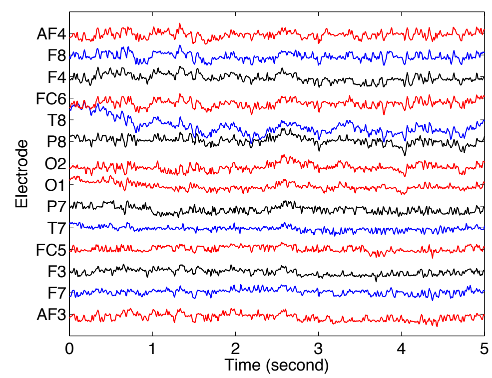

Figure 6: Raw data from the headset

3.3 Design Considerations

In what follows, we discuss a number of design considerations that relate to building a reliable and robust NeuroPhone system.

Signal to Noise Ratio (SNR): Since the Emotiv headset is not intended towards finer signal detection, there is more noise than usual on every electrode of the EEG. To compound this issue, EEG's are relatively noisy to begin with [9]. Assuming that this noise is relatively random, it has the potential to completely invalidate the data that we use to detect winks and P300 signals in the first place. We study various solutions to increase the SNR, such as bandpass filtering [10] and independent component analysis (ICA) [15]. A sensible approach to increase the SNR is to average the

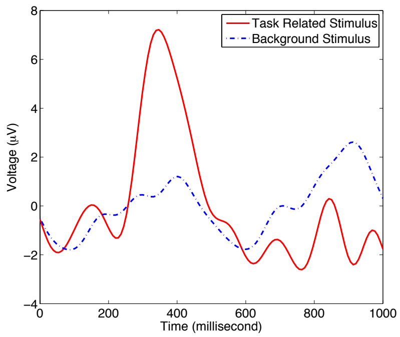

Figure 3: Multi-trial averaged band-pass filtered P300 Signal from one electrode. The difference from the peak of the P300 signal to the background noise is about 6μV

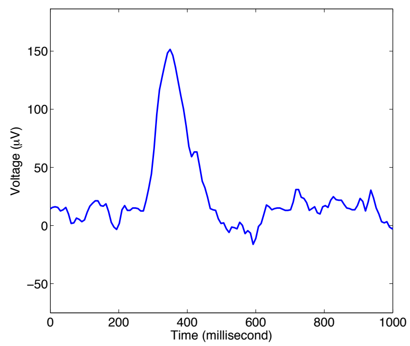

Figure 4: Unfiltered wink signal from one electrode. The difference from the peak of the wink signal to the the background noise is about 140μV

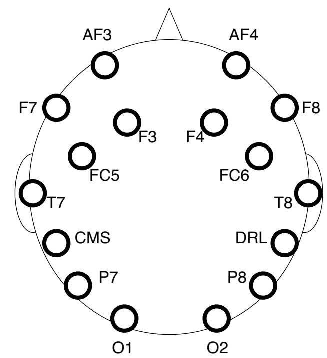

Figure 5: Electrode positions on the headset [3]

data over many trials, which is also a commonly used technique in neuroscience [11]. Naturally, this introduces delay in the acquisition of a reliable P300 signal, because we need to average several trials before actually start detecting the P300. However, in wink mode we can avoid averaging because wink signals (Figure 4) are much more easily detectable in raw EEG data than P300 signals (Figure 3).

Signal Processing: Although we are averaging data for a better SNR, we can still improve the EEG signals for better P300 detection. We use a bandpass filter to get rid of any noise that are not in the P300 frequency range [15]. Again this signal processing is unnecessary for wink mode because of the same reason why we do not need averaging.

Phone Classifiers: Typically, realtime EEG signal processing and classification algorithms are designed for powerful machines, not resource limited mobile phones. For example, Lotte et al. [9] use a weighted combination of various classifiers for EEG classification. These classification algorithms are not practical to run on the mobile phone because of power efficiency and resource issues. To address this challenge, we combine two approaches for efficient classification on the phone: i) we do not supply all channels from the headset to the phone for classification, rather, only the relevant subset of EEG channels; and ii) we implement lightweight classifiers, more specifically, a multivariate equal-prior Bayesian classifier is used for wink mode and a simple decision stump is used for the think mode.

4. IMPLEMENTATION

In this section, we discuss the implementation details of the wink mode and the think mode for the Dial Tim application. Due to the fact that headset only transmits encrypted data wirelessly and this data can be decrypted solely by Emotiv's closed-source SDK on a Windows machine, we use a laptop to relay the raw EEG data to the phone through WiFi. Upon receiving the EEG data, the phone carries out all the relevant signal processing and classification. The headset samples all channels at 128 samples/second, each of which is a 4-byte floating-point number corresponding to the voltage of a single electrode. The data rate of the EEG data streamed from the relay laptop to the mobile phone is 4kbps per channel. For each application mode, only relevant channels are streamed. Figure 7 shows the current system architecture. The phone uses simple machine learning techniques to determine user input (wink/non-wink or P300/non-P300). For the wink mode, we reverse mount the headset and only use the channels which are directly above the subject's eyes i.e., *O1* and *O2*. We develop a data collection program where the subject can easily label each wink. A multivariate Bayesian classifier is then trained and used for classification. We set equal prior such that it will not bias toward either wink or non wink classes. In the preprocessing step, we calculate variances over a 90% overlapping sliding window of the two channels. The variances are used as features and are fed to the classifier in the classification stage. During the offline training phase, 2D Gaussian distributions are estimated for both the wink and non-wink class, as illustrated in Figure 8. The two Gaussians are mostly separated, which results in good online classification performance.

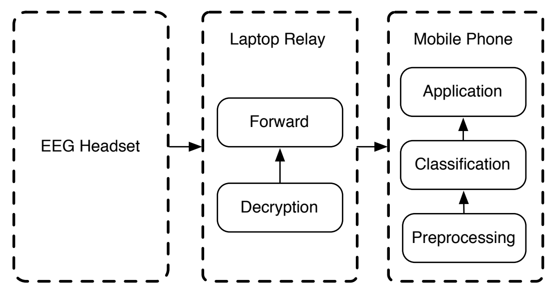

Figure 7: NeuroPhone system architecture

For the think mode of the application, which utilizes

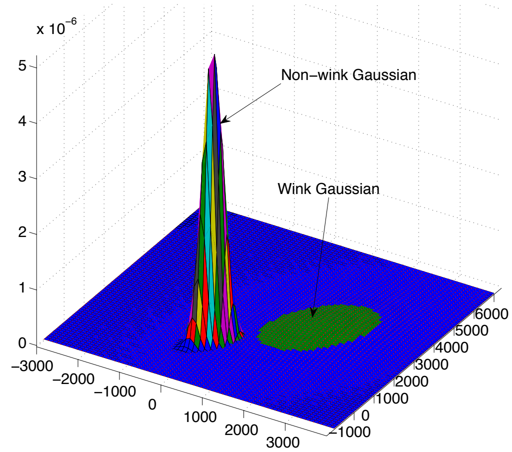

Figure 8: Gaussians for winks and non-winks

the P300 signal, we attempt to use similar 2D Gaussians. However, the distributions of the classes prove to be too overlapped for reasonable classification. As discussed in the design consideration section, to cancel out unnecessary noise we preprocess the data by filtering it with a 0-9Hz bandpass filter and averaging the signal over multiple trials. We do this preprocessing separately for all 6 stimuli corresponding to the six images of the Dial Tim application. Following this we only crop the signal segment that corresponds to the highest peak value at around 300ms after the stimulus onset. For classification, we use a decision stump whose threshold is set to the maximum value among the cropped signal segments for all 6 images.

5. EVALUATION

To evaluate our system, we test the wink and think modes in a variety of scenarios (e.g., sitting, walking) using two different Emotiv headsets and three different subjects. In what follows, we discuss our initial results.

For the wink mode, we collect multiple sessions of data from all subjects while they sit relaxed or walk, then train an equal-prior Bayesian classifier using a set of five sessions of data from a single subject sitting relaxed. This classifier is then applied to the rest of the data to test whether it can generalize to unseen data by calculating the classification precision (i.e., percentage of classified "winks" that are actually real winks), recall (i.e., percentage of real winks that are actually classified as winks) and accuracy (i.e., percentage of all events that are correctly classified). The experiment results are shown in Table 1. As can be seen from the table, the classifier performs well on data collected for sitting-relaxed scenarios but walking results in a decline in performance. The decline of recall suggests that while the subjects are walking, a small amount of blinks are contaminated such that the classifier fails to pick them up; thus, representing false negatives. There is a larger decline in precision, which suggests that in additional to the increase in false negatives reflected by the recall, there is also a increase in false positives; noisy peaks in EEG data caused by walking are erroneously picked up by the classifier as blinks. Despite the performance decline of the wink classifier when applied to more noisy data, we can, however, still observe that it is robust in reasonably noisy scenarios.

	Sitting Relaxed	Walking
Precision	92.35%	86.15%
Recall	99.39%	96.70%
Accuracy	95.58%	92.58%

Table 1: Wink classification results

For think mode, we test on the same set of subjects. We carry out the P300 experiments with the subjects using the Dial Tim application while sitting still, sitting with loud background music and standing up. We average the data over a set time interval. The experiment results are shown in Table 2. First, the accuracy increases as the data accumulation time increases, which coincides with the intuition that averaging over more data improves the SNR for the expected P300 signals, leading to higher accuracy. Second, P300 signals are quite susceptible to external noises, illustrated by the fact that when subjects are sitting still, we have the best accuracies, whereas accuracy decrease when considerable auditory noise is introduced. Accuracy further declines when the subjects stand up, which potentially adds more noises due to subjects' muscle controls and physical movements. Third, even though different experiment settings result in different P300 detection accuracies, more data accumulation and averaging generally yields better detection accuracies.

Time	Sitting	Music (Sitting)	Standing
20s	77.78%	44.44%	33.33%
50s	77.82%	66.67%	66.67%
100s	88.89%	88.89%	66.67%

Table 2: Think classification accuracies. Times in the first column indicate the different time durations of data accumulation for averaging. Contact pictures are flashed once every half a second in random order; each of the 6 pictures has a 1/6 chance for each flash. Accuracy measures the proportion of correctly classified sessions. Note that chance level classification accuracy would be $1/6 \approx 16.67\%$.

While our initial results are promising for a limited set of scenarios many challenges remain. Currently, to get usable P300 signals from the user, we need to average their data over a large number of trials. This is typically how neural signals are handled. However, this general "unresponsiveness" of the system proves to be rather frustrating for the end user. There has been recent works on single-trial classification of EEG data [6, 11]. We are currently investigating how to reliably carry out classification using such single-trial data approaches. We also carry out P300 experiments while subjects are walking and driving which yields low accuracies due to noise. We plan to study the application of different processing and classification algorithms capable of dealing with large induced noise from such activities. The CPU usage for our application on the iPhone is 3.3%, and the total memory usage is 9.40MB, of which 9.14MB are for GUI elements, meaning that the actual preprocessing and classification components of our application are quite lightweight, using minimal amounts of memory. However, continuous use of NeuroPhone streaming raw EEG channels to the phone using WiFi and running processing and classification pipelines would lead to battery drain. We plan to study duty cycling the phone to solve this problem.

6. RELATED WORK

There is a limited amount of related work in this area. A number of groups [4, 6, 15] use research/professional quality EEG devices that offer higher quality signals but are expensive and based on wired headsets not wireless. In contrast, consumer-oriented EEG headsets [3, 12, 13] are considerably cheaper and noisier but at the same time are more geared toward gaming applications rather than the types of classification we have used them for. Typically, these headsets are wireless, enabling mobile uses. [7, 12] are more closely related to NeuroPhone. [7] develops a wireless EEG headband prototype with 4 electrodes targeting forehead non-hairy skin area, which is not suitable for P300. [12] is a commercially available headset with a single electrode not powerful enough for the types of applications we have in mind such as Dial Tim. These projects discuss connecting neural signals to mobile phones just to display visualization and simple frequency-domain analysis of the signal, not to drive mobile applications themselves. In essence, the phone is used as a mobile display and not as a phone.

7. CONCLUSION

We have presented the evaluation of an initial prototype that brings together neural signals and phones to drive mobile applications in new ways. One could argue that connecting the wireless Emotiv EPOC EEG headset and iPhone is just a simple engineering exercise. We believe the NeuroPhone system is an important development precisely because it is simple to engineer using cheap but noisy commercial components. NeuroPhone opens up new opportunities and challenges in ubiquitous sensing and pervasive computing. For example, sniffing packets could take on a very new meaning if brain-mobile phone interfaces become widely used. Anyone could simply sniff the packets out of the air and potentially reconstruct “thoughts” of the user. Spying on a user and detecting something as simple as them thinking yes or no could have profound effects. Thus, securing brain signals over the air is an important challenge.

8. ACKNOWLEDGEMENT

I would like to thank my family and friends for all that they have done to get me to this point. Next I would like to thank my advisors, Andrew T. Campbell, Tanzeem Choudhury, and Rajeev D. S. Raizada. The work detailed in this paper would not have come to fruition without the help of Shaohan Hu, Mashfiqui Rabbi, and Hong Lu.

This paper was accepted in the second ACM SIGCOMM workshop on networking, systems, and applications on mobile handhelds (MobiHeld ’10) and will be published as [1].

9. REFERENCES

[1] A. T. Campbell, T. Choudhury, S. Hu, H. Lu, M. K. Mukerjee, M. Rabbi, and R. D. S. Raizada. NeuroPhone: Brain-Mobile Phone Interface using a Wireless EEG Headset. In *Proceedings of The Second ACM SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds (MobiHeld’10)*. ACM New York, NY, USA, 2010.

[2] Demo-Video. Neurophone. `http://www.cs.dartmouth.edu/~shu/neurophone`.

[3] EmotivSystems. Emotiv - brain computer interface technology. `http://emotiv.com`.

[4] L. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6):510–523, 1988.

[5] GugerTechnologies. g.tec - guger technologies. `http://www.gtec.at/`.

[6] K. Li, R. Sankar, Y. Arbel, and E. Donchin. P300 Based Single Trial Independent Component Analysis on EEG Signal. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pages 404–410, 2009.

[7] C. Lin, L. Ko, C. Chang, Y. Wang, C. Chung, F. Yang, J. Duann, T. Jung, and J. Chiou. Wearable and Wireless Brain-Computer Interface and Its Applications. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pages 741–748, 2009.

[8] D. E. J. Linden. The P300: where in the brain is it produced and what does it tell us? *Neuroscientist*, 11(6):563–76, Dec 2005.

[9] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain-computer interfaces. *J Neural Eng*, 4(2):R1–R13, Jun 2007.

[10] J. Malmivuo and R. Plonsey. *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, New York, 1995.

[11] A. Mouraux and G. Iannetti. Across-trial averaging of event-related EEG responses and beyond. *Magnetic resonance imaging*, 26(7):1041–1054, 2008.

[12] NeuroSky. Neurosky - experience the mindset. http://www.neurosky.com/.

[13] OCZTechnology. nia game controller OCZ technology. http://www.ocztechnology.com/products/ocz_peripherals/nia-neural_impulse_actuator.

[14] A. B. Usakli, S. Gurkan, F. Aloise, G. Vecchiato, and F. Babiloni. On the use of electrooculogram for efficient human computer interfaces. *Comput Intell Neurosci*, page 135629, 2010.

[15] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang. BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications. *IEEE Transactions on Biomedical Engineering*, 51(6):1067–1072, 2004.