IST-2003-511598 (NoE)

COGAIN

Communication by Gaze Interaction

Network of Excellence

Information Society Technologies

D4.9 Online information on how to use control facilities to supplement gaze for control

Due date of deliverable: 28.02.2009
Actual submission date: 14.04.2009

Start date of project: 1.9.2004

Duration: 60 months

Public University of Navarra

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	x (report)
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	x (software)

Authors:

Mikel Ariz (UPNA)
Javier Navallas (UPNA)
Arantxa Villanueva (UPNA)
Javier San Agustin (ITU)
Rafael Cabeza (UPNA)
Martin Tall (ITU)

Contributors:

Gintautas Daunys (SU)

Table of Contents

Executive Summary

The objective of this deliverable is to present a repository based on online materials that can be used as supplementary technologies for gaze interaction. More specifically, materials and information for an EMG (Electromyographic) system are provided. Combining the eye tracking signal with muscle activity detection (Video-oculography, VOG + EMG) in multimodal interfaces can be a solution for the *Midas Touch* problem for many users with difficulties when blinking or fixating. Moreover, the EMG activation is faster than the techniques available using just the eye. Brain computer interfaces can also be a potential solution for this issue, but normally a lot of training is needed from the user and although the results are quite promising, we have prioritized the EMG-based system in this work.

The supplied material includes useful tools for constructing a multimodal interface based on a VOG and an EMG system. The target audience would be researchers with technical profile and programming skills which could easily understand the algorithms provided for muscle signal detection. Information about commercial systems (for non technical users) is also provided.

The supplied materials are:

- Files (private, for COGAIN members only):
  - DLL implementation of selected detection algorithms in temporal domain. Original code in Matlab and C (optimized) are provided (implementation).
  - Database of activations signals, recorded in our lab (DDBB).
  - Random click generator to simulate noise (noisy clicks).
- Deliverable D4.9 (this report; public, available online)
  - Description of the uploaded algorithms.
  - Short description about how to do your own muscle activation detector.
  - NIA system description.

This report (D4.9) is public, however, the material (DLL implementation and database files) are private and meant for COGAIN members only. The files have been uploaded to

http://www.cogain.org/internal/wp_areas/wp4/D4.9Files/ (password protected, members only).

1 Introduction

An EMG switch constitutes an efficient substitute for the mouse button. Initial research by Nelson et al. (1996) found that EMG clicking could be up to 20% faster than finger-based mouse clicking. In their study they used the Cyberlink Brainfingers, the system on which the Neural Impulse Actuator (NIA) is based. People who suffer from severe motor disabilities might still retain some degree of facial muscle activity. Even if they are not able to control a mouse or any other input device, they might be able to control an EMG forehead switch by performing slight muscle movements. It must be noted that due to natural facial muscle activity in a normal situation, a high number of undesired activations might occur when making use of an EMG switch. For instance, coughing or laughing will cause the EMG signal to rise, issuing one or more activations.

A combination of gaze pointing and EMG clicking seems to be an efficient, hands-free interface technique that can help people who suffer from severe disabilities to interface with a computer. A few studied have evaluated the potential of gaze pointing and EMG clicking in point-and-select tasks (i.e. pointing at a target and selecting it). Partala et al. (2001) found completion times to be shorter with gaze-EMG combination than when using a mouse. However, a high error rate (34%) was observed. Surakka et al. (2004) extended that study with a more in-depth Fitts' Law study, but did not find any differences in speed between gaze-EMG and mouse. Recently, Mateo et al. (2008) evaluated the four possible input combinations of gaze/mouse pointing and EMG/mouse clicking following the ISO 9241-9 Standard that measures input performance. They found gaze pointing to be faster than mouse pointing, but no significant differences were observed between EMG and mouse clicking in terms speed.

2 EMG Detection Algorithms Description

Various methods to detect muscle activity have been explored during the last years. Electromyography is an important topic not only from the point of view of this document, but also as the field that studies muscle behavior and activity modeling. Among the existing techniques for SEMG (Surface EMG), precise onset detection is one of the most sought-after objectives. Determining the contraction instant of the muscle accurately is of high importance for many purposes. When analyzing the EMG signal for certain physiology related applications it is necessary to determine with high accuracy (in the order of milliseconds) the time of the activity onset. In our case, this requirement can be relaxed to some extent in one or two orders. Thus, simpler methods can be used for our purposes. Some techniques try to analyze the signal in frequency domain while others limit their analysis to temporal domain. We have focused our attention on temporal analysis since the computational complexity is reduced.

Most detection algorithms consist of up to three basic processing stages (Staude et al., 2001):
• signal conditioning,
• detection unit,
• post-processor.

In the signal conditioning step a low pass filter is normally employed, although more sophisticated solutions, such as signal rectification followed by a low pass filter, can be found in the bibliography. The test unit uses some of the past samples to create an intermediate signal which is analyzed by the decision algorithm to look for activations. Normally, this is compared with a threshold value. Once an activation has been detected the post processor unit makes a more accurate analysis to find the exact onset instant. To follow, commonly used detection methods are described (Staude et al., 2001), describing their independent parameters, the function of comparison with the threshold that determines possible activations, and the delay introduced by the algorithm based on the independent parameters. This description is largely based in the work by Staude et al. (2001). For each one of the parameters the recommended values for a sampling frequency of 1kHz are provided in parenthesis (the parameters expressing number of samples should be changed according to sample frequency). It should be noted that, due to the relaxation of the onset temporal-estimation accuracy, we have not implemented the post-processing stages of some of the following algorithms, whenever it is devoted only to augment the precision of the onset temporal-estimation (Abbink, AGLR Step and AGLR Ramp). On the other hand, we did implement the post-processing of the other algorithms (Bonato and Lidierth), because they are part of the onset detection system. Below, we describe several of the methods, their parameters, test functions and formulas for delay estimation.

2.1.1 Hodges and Bui

Method based on thresholding (Hodges and Bui, 1996).

1. Independent parameters (see Figure 1):
    - **h** (2.5): threshold. When function g(k) exceeds this threshold, activation is detected.
    - **W** (50): size of window. The calculation of function g(k) for each sample is made taking a window of W samples. It is recommended to know, at least roughly, the duration of the activations, to define this parameter accordingly.
    - **M** (200): number of initial samples of the signal used for the estimation of the average and the standard deviation of our signal in a relaxed situation of the subject.
    - **fc** (50): cut-off frequency in Hertz of the low pass filter used.

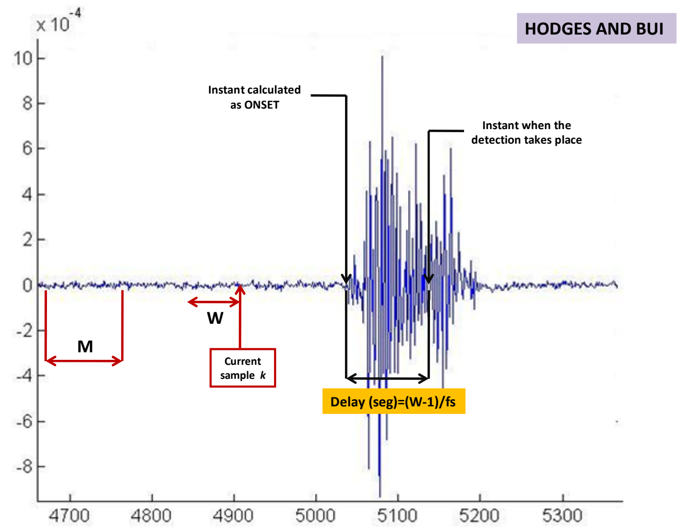

Figure 1. Hodges and Bui

2. Test function

   $g(k) = \frac{1}{\hat{\sigma}_0}(\bar{y}_k - \hat{\mu}_0)$ where $\bar{y}_k = \frac{1}{W}\sum_{i=k-W+1}^{k} y_i$ and $\hat{\sigma}_0$ and $\hat{\mu}_0$ are the standard deviation and the average estimated from the **M** first samples, respectively.

3. Delay estimation

   Ret (sec) = $\frac{W-1}{fs}$

   where f_s is the sampling frequency of the EMG signal.

2.1.2 Bonato et al.

Method based on thresholding (Bonato et al., 1998).

1. Independent parameters (see Figure 2):
   - **h** (7.4): threshold. When function g(k) exceeds this threshold, the post processing is launched.
   - **n** (1): post processing parameter*
   - **m** (5): post processing parameter *
   - **T1** (50): post processing parameter*
   - **M** (200): number of initial samples of the signal used for the estimation of the average and the standard deviation of our signal in a relaxed situation of the subject.

   * Once the possible activation sample has been detected, the post processing consists of taking the **T1** following samples and divide them in blocks of **m** samples (**T1** must be multiple of **m**). In each one of the blocks it must happen that **n** of those **m** samples are larger than **h** in order to confirm that a muscle activation has happened. It is important to consider the duration of the activations in order to fix these parameter values.

2. Test function

$$g(k) = \frac{1}{\sigma_0^2}\left(y_{k-1}^2 + y_k^2\right)$$

where σ_0^2 is the signal variance calculated from the first **M** samples. Sample pairs are considered for the test function, thus ***g(k)*** is defined for ***k*** pairs..

3. Delay estimation

Ret (sec) = $\frac{2 \times T1}{fs}$

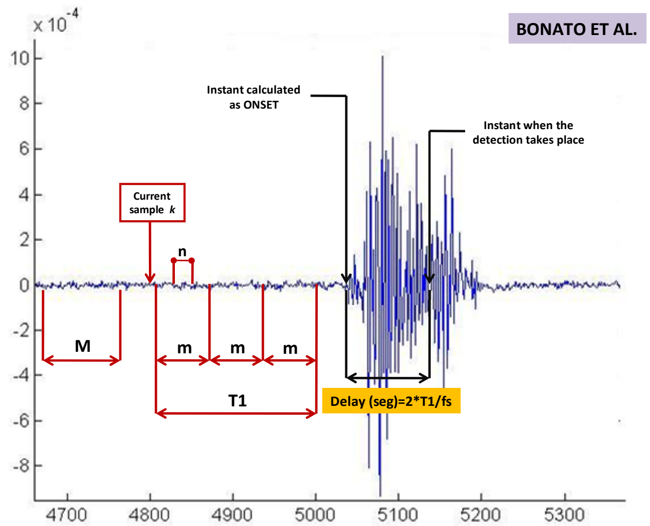

Figure 2. Bonato et al.

2.1.3 Lidierth

Method based on thresholding (Lidierth, 1986).

1. Independent parameters (see Figure 3):
   - **h** (3): threshold. When function g(k) exceeds this threshold, the post processing will be launched.
   - **W** (50): size of window. The calculation of function g(k) for each sample is made taking a window from W samples. It is recommended to know, at least roughly, the duration of the activations, to define this parameter accordingly.
   - **M** (200): number of initial samples of the signal used for the estimation of the average and the standard deviation of our signal in a relaxed situation of the subject.
   - **T1** (90): post processing parameter *
   - **T2** (15): post processing parameter *

* The post processing is identical to the one by Bonato et al., being **T2** parameter equivalent to **m** of the other algorithm and fixing **n** to the value of 1 sample. It is necessary to consider the duration of the activations to define **T1** and **T2.**

2. Test function

$$g(k) = \frac{1}{\hat{\sigma}_0}(\bar{y}_k - \hat{\mu}_0) \qquad \text{where} \qquad \bar{y}_k = \frac{1}{W}\sum_{i=k-W+1}^{k} y_i$$

and $\hat{\sigma}_0$ and $\hat{\mu}_0$ are the standard deviation and the average estimated from the M first samples, respectively.

3. Delay estimation

$$\textbf{Ret (sec)} = \frac{W-1+2\times T1}{fs}$$

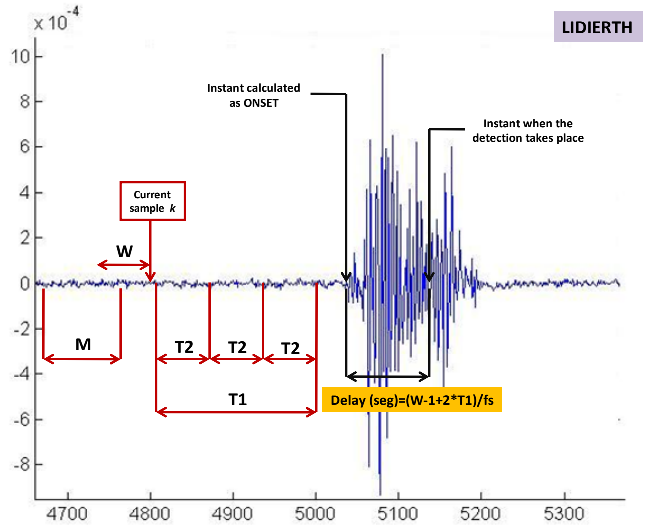

Figure 3. Lidierth

2.1.4 Abbink et al.

Method based on thresholding (Abbink et al., 1998).

1. Independent parameters (see Figure 4):
   - **h** (3): threshold. When function g (k) exceeds this threshold, an activation is detected and the post processing is launched.
   - **h2** (3): threshold value used for the post processing*
   - **N** (200): post processing parameter*
   - **M** (200): number of initial samples of the signal used for the estimation of the average and the standard deviation of our signal in a relaxed situation of the subject.

   *The post processing consists of taking, for each potential onset sample, the **N** previous and **N** posterior samples to test how many of the previous ones are larger than **h2** and how many of the posterior ones are larger than **h2**. The onset sample will be selected as the one that maximizes these both quantities. It is also important to consider the duration of the activations to fix **N** correctly.

2. Test function

$$g(k) = \frac{1}{\hat{\sigma}_0}(y_k - \hat{\mu}_0)$$

where $\hat{\sigma}_0$ y $\hat{\mu}_0$ are the standard deviation and the average estimated from the **M** first samples, respectively.

3. Delay estimation

Ret (sec) = $\frac{N}{fs}$

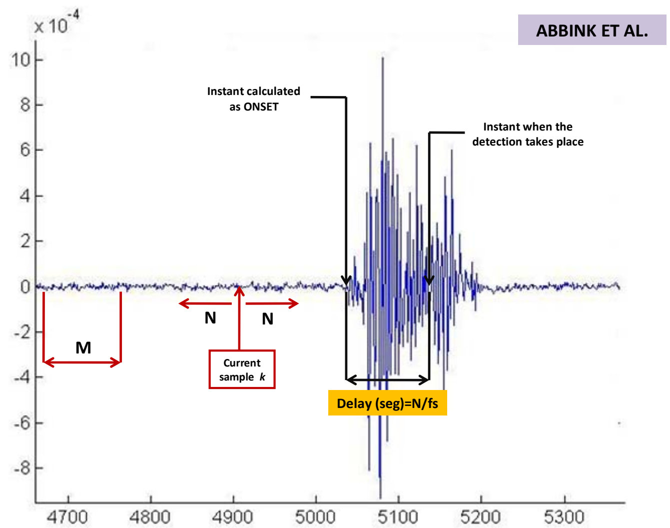

Figure 4. Abbink et al.

2.1.5 AGLR Step

Detector based on statistically optimal decision.

1. Independent parameters (see Figure 5):

- **h** (10): threshold. When function g(k) exceeds this threshold, the post processing will be launched*.
- **W** (25): window size. The estimation of the function g(k) for each sample is made using a window of W samples. It is advisable to know, at least approximately, the duration of the activations, to define the size of this window.
- **delta** (100): number of samples of the death zone for the post processing. This parameter is used to assure that the correct number of samples is taken into account for the estimation of the parameters.
- **M** (200): number of initial samples of the signal used for the estimation of the average and the standard deviation of our signal in a relaxed situation of the subject.

* The post processing consists on looking for the instant that maximizes the function S, making the searching around the sample for which the threshold has been crossed according to the test function.

2. Test function:

$$g(k) = S_{k-W+1}^{k} \qquad \text{where}$$

$$S_j^k = \frac{k-j+1}{2}\left(\frac{(1/(k-j+1))\sum_{i=j}^{k} y_i^2}{(1/M)\sum_{i=1}^{M} y_i^2} - \ln\left(\frac{(1/(k-j+1))\sum_{i=j}^{k} y_i^2}{(1/M)\sum_{i=1}^{M} y_i^2}\right) - 1\right)$$

3. Delay Estimation:

$$\textbf{Ret (sec)} = \frac{delta}{fs}$$

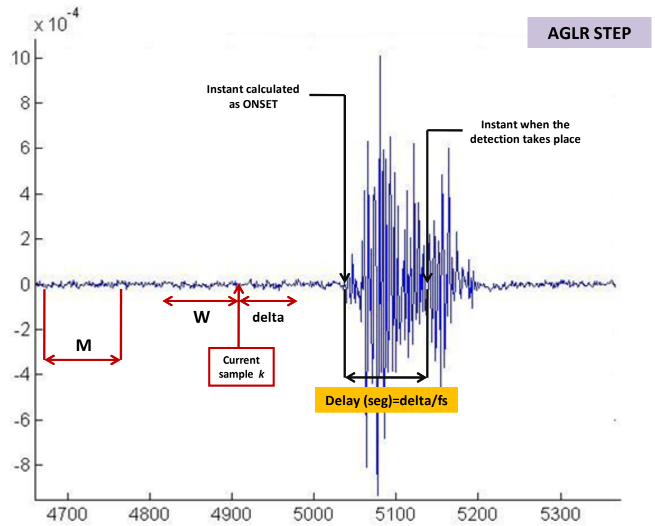

Figure 5. AGLR Step

2.1.6 AGLR Ramp

Detector based on statistically optimal decision.

1. Independent parameters (see Figure 6):
   - **h** (10): threshold. When function g(k) exceeds this threshold, the post processing will be launched*.
   - **W** (25): window size. The estimation of the function g(k) for each sample is made using a window of W samples. It is advisable to know, at least approximately, the duration of the activations, to define the size of this window.
   - **delta** (100): number of samples of the death zone for the post processing. This parameter is used to assure that the correct number of samples is taken into account for the estimation of the parameters.

- **tau** (5 - 40): ramp duration in msec. In order to use this method it is assumed that the shape of the muscle activity signal when changing from the relax state to the activation state resembles a ramp. With this parameter we control the slope of the ramp. It would be advisable to know beforehand the type of activations produced by the subject, to fix an appropriate value for tau. We can take as reference values changing in the range from 5 to 40 msec.
- **M** (200): number of initial samples of the signal used for the estimation of the average and the standard deviation of our signal in a relaxed situation of the subject.

* The post processing is identical to the one used in AGLR Step.

2. Test function:

$$g(k) = \hat{S}_{k-W+1}^{k} \qquad \text{where}$$

$$\hat{S}_j^k = \frac{1}{2}\sum_{i=j}^{k}\left[\left(\frac{1}{\hat{\vartheta}_0} - \frac{1}{\hat{\vartheta}_1 u(i,j)+\hat{\vartheta}_0}\right)y_i^2 + \ln\frac{\hat{\vartheta}_0}{\hat{\vartheta}_1 u(i,j)+\hat{\vartheta}_0}\right] \qquad \text{with}$$

$$\hat{\vartheta}_0 = \frac{1}{M}\sum_{i=1}^{M} y_i^2 \qquad y \qquad \hat{\vartheta}_1(j,k) = \frac{\sum_{i=j}^{k}(y_i^2-\hat{\vartheta}_0)}{\sum_{i=j}^{k} u(i,j)}$$

The function **u(i,j)** represents the shape of the activation signal. It is a ramp with a duration represented by tau.

4. Delay Estimation

Ret (sec) = $\frac{delta}{fs}$

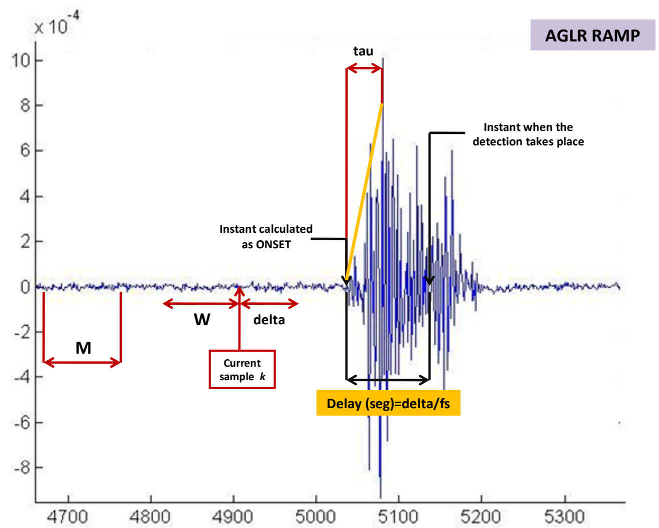

Figure 6. AGLR Ramp

3 How to Build Your Own EMG System

Apart from commercial solutions, such as the NIA (see section 4), you can build your own EMG system if the following hardware pieces are available:

- Surface Electrodes.
- Signal Amplifier.

3.1 Surface Electrodes

Surface electrodes are attached to the skin around the muscle area, the forehead is many times employed since the muscle signal is faster compared to other muscles, such as the hand. The forehead is many times used as activation muscle since it is an easy-to-control activation and can be an alternative for users with hand movement limitations. There are different types of electrodes, some of them use some kind of conducting gel (pre-gelled) and normally they are single-use electrodes. Dry electrodes can be used more than one time.

In the designed system surface electrodes have been employed. They are disposable, self-adhesive pre-gelled electrodes. These are common electrodes used for EEG or EOG purposes. Three electrodes should be used, two of them attached to the skin around the muscle zone and a third one working as the ground. In the system developed in our lab, the forehead is selected for placing the activity detection electrodes. The required action can be raising the eyebrows or tightening the jaw. The reason for having two electrodes is that we employ the difference signal between them providing a high common-mode rejection ratio. The third electrode working as ground is placed in the wrist. In the following Figures 7 and 8 a user wearing the electrodes is shown.

Figure 7. Electrode working as ground

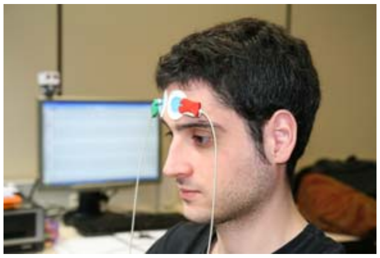

Figure 8. Two electrodes are placed on the forehand to detect muscle activations when raising the eyebrows or tighten the jaw.

3.2 Signal Amplifier

The objective of the signal amplifier is to adapt the signal (amplification, digitalization) to be transmitted to the computer. There are specific devices for biosignals. We have employed an amplifier that is connected directly to the USB interface of the computer and its software (DLL) allows interfacing with different programming languages. The amplifier has specific cables to be connected to the electrodes (as shown in Figure 9). The amplifier provides the possibility to connect more than one signal. In our case one input is used.

Figure 9. The electrodes are connected to amplifier inputs. The color code corresponds to Figures 7 and 8 to differentiate the ground from the detection activity electrodes.

The amplifier offers the possibility of adjusting the sampling frequency (number of samples per second) and buffer size (how many samples are delivered at the same time to the computer). It is important to know the nature of the signal we want to record, since having a low sampling frequency can produce a loss of information and our signal might not be properly reconstructed later (aliasing).

It is also important to consider the objective of our application. If the purpose is to analyze signals offline we do not need the samples to be delivered immediately to the computer to be processed. We can use longer buffers and save the data in the hard disk to be analyzed later. However, if we need a real time EMG activation detector such as the one proposed here, it is important to process the data in real time. Having long buffers can produce significant delays in the system if the buffer is not delivered immediately. We can find limitations for the sampling frequency and buffer size in the amplifier.

In our case a sampling frequency of 2400 Hz has been selected. The algorithm parameters dependent on the sampling frequency (the ones whose units are the number of samples) were modified consequently, multiplying them by a factor of 2.4. The buffer sizes selected were 144 and 20 samples for offline and online analysis, respectively. Regarding the sampling frequency it is important that it fulfills Nyquist criteria, i.e. sampling frequency should be at least two times larger than the bandwidth of the signal to be sampled.

Depending on the amplifier one can modify configuration options such as gain, low pass or band pass filtering among others. A bandpass filtering between 5 and 500 Hz can be appropriate to work with the EMG signal. It is also advisable to remove the 50Hz interference by using a narrow Notch filtering around that frequency.

3.3 Programming Issues

Once the acquisition and the interface with the amplifier have been adjusted, we can play with the EMG signal and use different detection algorithms to detect muscle activations. If we work in an offline analysis we can consider a separate application that opens the file signal and process it looking for activations. If a real time detector is desired we need to integrate the algorithm into the acquisition part and process the samples buffer as soon as it is delivered. Much attention should be paid in the synchronization of both loops (acquisition and processing). The acquisition should keep working continuously to prevent any loss of samples. On the other hand, the processing part should analyze each buffer and be finished before the next buffer is delivered by the acquisition loop. Otherwise that buffer will be lost and replaced with the next one, meaning that the signal would be interrupted.

When using the EMG input as an activation signal for the eye tracking system, an action should take place after each of the activations. If a “mouse click” is desired a message has to be sent to the OS to simulate that action. Each platform presents different forms to simulate this mouse action. In fact, this should be done by calling the Windows API in user32.dll. According to the platform different interfacing forms can be found for this function. Normally, once a muscle activation occurs, the algorithms can produce more than one activation (close in time from each other). We can control the behavior of our system by establishing a rejection period after a click, meaning that only one click per activation is possible, or consider different mouse actions depending on the duration of each one of the activations.

The VOG-EMG interface might be more difficult if the eye tracking system uses another cursor different from the one in Windows to show the user’s the estimated gaze coordinates on the screen. If that is the case, the Windows’ mouse click would not be appropriate since the clicks would not happen where the actual cursor is pointing. In that case, a more careful and dedicated (for each system) interface would be required between the EMG and VOG system.

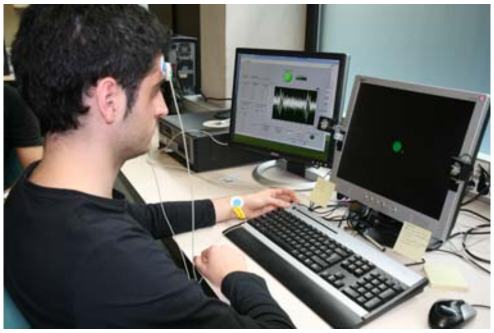

Figure 10. A user working with the eye tracking and EMG system.

4 The Neural Impulse Actuator (NIA) System

The Neural Impulse Actuator (NIA) is a commercial device that allows the user to interface with a computer by means of three different biosignals, namely electromyogram (EMG), electrooculogram (EOG) and electroencephalogram (EEG). It is sold by the company OCZ, which is specialized on gaming hardware. The device costs around 125€.

To detect the three different biosignals, the device uses three surface electrodes mounted on a headband that the user wears on her forehead. The headband is connected to an interface box that filters and amplifies the signal coming from the electrodes. Figure 11 shows the headband and the interface box.

Figure 11. NIA's headband and interface box.

The interface box is connected to a computer through a USB port. The software provided with the system analyzes the biosignal and decodes it in different frequencies. Figure 12 shows a screenshot of the different biosignals detected by the NIA.

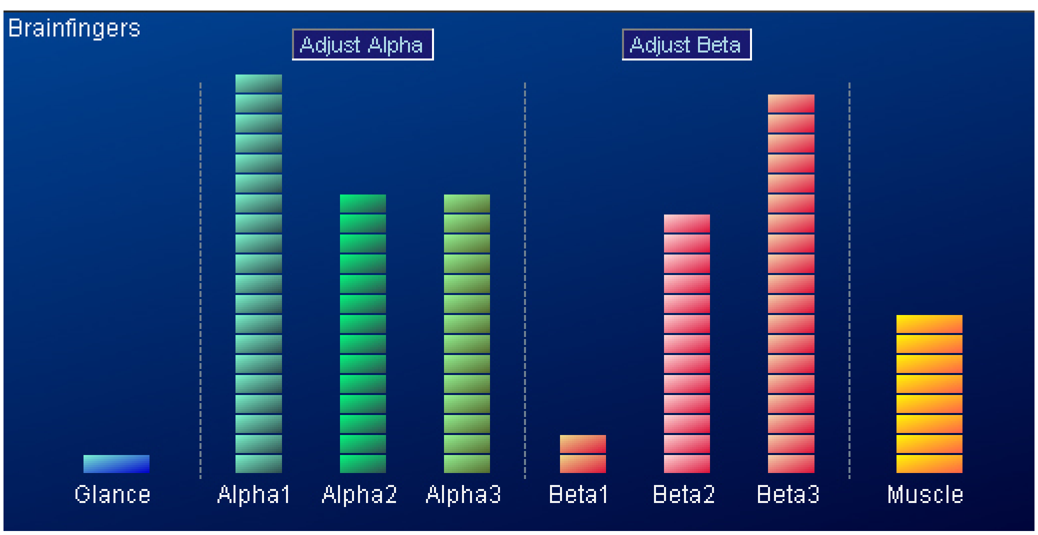

Figure 12. Screenshot of the biosignal detected by the NIA and decoded in different frequencies.

The lowest frequencies correspond to the electrooculographic (EOG) signals ("Glance"), which are detected when the user performs lateral eye movements. This signal can be mapped into a one-axis cursor movement, or an on/off switch activated by fast eye movements.

The mid frequencies correspond to the electroencephalographic (EEG) signal, which primarily reflects the mental activity of the user, and to some extent the facial muscle activity. This signal is divided into 2 bands, called alpha and beta, with 3 different frequencies in each band. It is possible for the user to learn to control each of these signals, which would then be mapped into a different input to the computer.

The highest frequency corresponds to the electromyographic (EMG) signal, which is related to facial muscle activity. The signal is detected when the user activates one or more of the muscles on her face. In particular, raising the eyebrows and tightening the jaw usually gives the best results. It is an analog signal, i.e. the amplitude of the signal will increase as the muscle activity is increased. However, the most common use of the EMG signal is digital in order to control an on/off switch.

Figure 13 shows a screenshot of the EMG signal detected by the NIA. The line in the middle represents the threshold that is used to determine whether a sample corresponds to an active muscle state. In this case we can see one muscle activation performed by the user, which would be translated into a click.

Figure 13. EMG signal detected by the NIA. One muscle activation has been performed.

References

Abbink, J., Van Der Bilt, A., & Van Der Glas, H. (1998). Detection of onset and termination of muscle activity in surface electromyograms. *J. Oral Rehab*., 25, 365–369.

Bonato, P., D'Alessio, T., & Knaflitz, M. (1998). A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait. *IEEE Transactions on biomedical engineering*, 45 (3), 287–299.

Hodges, P., & Bui, B. (1996). A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. *Electroencephalography and Clinical Neurophysiology/ Electromyography and Motor Control*, 101 (6), 511–519.

Lidierth, M. (1986). A computer based method for automated measurement of the periods of muscular activity from an EMG and its application to locomotor EMGs. *Electroencephalogr Clin Neurophysiol*., 64(4), 378–380.

Mateo, J., San Agustin, J., & Hansen, J. P. (2008). Gaze beats mouse: hands-free selection by combining gaze and emg. *In CHI '08 extended abstracts on human factors in computing systems* (pp. 3039–3044). New York, NY, USA: ACM.

Nelson, W. T., Hettinger, L. J., Cunningham, J. A., Roe, M. M., Haas, M. W., Dennis, L. B., Pick, H. L., Junker, A., and Berg, C. (1996). Brain-body-actuated control: assessment of an alternative control technology for virtual environments. *Proceedings of the 1996 IMAGE Conference* (pp. 225-232).

Partala, T., Aula, A., and Surakka, V. (2001). Combined voluntary gaze direction and facial muscle activity as a new pointing technique. *Proceedings of INTERACT 2001*, Amsterdam, 100–107.

Staude, G., Flachenecker, C., Daumer, M., & Wolf, W. (2001). Onset detection in surface electromyographic signals: a systematic comparison of methods. *EURASIP J. Appl. Signal Process.*, 2001 (1), 67–81.

Surakka, V., Illi, M., & Isokoski, P. (2004). Gazing and frowning as a new human-computer interaction technique. *TAP*, 1 (1), 40-56.