# Considerations on Strategies to Improve EOG Signal Analysis

Tobias Wissel[1,2] and Ramaswamy Palaniappan[2]

[1]Institute for Electronics, Signal Processing and Communications (IESK),
Otto von Guericke University, Magdeburg, Germany

[2]School of Computer Science and Electronic Engineering
University of Essex
Colchester, UK

tobias_wissel@daad-alumni.de/rpalan@essex.ac.uk

Corresponding author:

*Dr Ramaswamy Palaniappan
Lecturer
School of Computer Science and Electronic Engineering
University of Essex
Wivenhoe Park, Colchester
CO4 3SQ, United Kingdom
T/F: +44(0)1206-872773/872788
Email: rpalan@essex.ac.uk; palani@iee.org

# Considerations on Strategies to Improve EOG Signal Analysis


## ABSTRACT

*Electrooculogram (EOG) signals have been used in designing Human-Computer Interfaces, though not as popularly as electroencephalogram (EEG) or electromyogram (EMG) signals. This paper explores several strategies for improving the analysis of EOG signals. The multi-resolution representation obtained from wavelet decompositions is applied in various fields of research. This article explores its utilization for the extraction of features from EOG signals compared with parametric, frequency-based approach using an autoregressive (AR) model as well as template matching as a time based method. Our results indicate that parametric AR modeling using the Burg method, which does not retain the phase information, gives poor class separation. Conversely, the projection on the approximation space of the fourth level of Haar wavelet decomposition yields feature sets that enhance the class separation. Furthermore, for this method the number of dimensions in the feature space is much reduced as compared to template matching, which makes it much more efficient in terms of computation. We also report on an example application utilizing wavelet decomposition and the Linear Discriminant Analysis (LDA) for classification, which was implemented and evaluated successfully. In this application, a virtual keyboard acts as the front-end for user interactions.*



*Keywords:* *Autoregressive, Electrooculogram, Wavelet decomposition, Virtual keyboard*


## INTRODUCTION

As part of the Human-Computer Interface (HCI) research field, the electrooculography (EOG) has been established as a promising modality among other sources like electroencephalography (EEG) or electromyography (EMG). In this context various eye-based applications have been proposed, acting as hands-free interfaces to a general computer system (Estrany, Fuster, Garcia, & Luo, 2009) or even wearable embedded system, in order to aid the context awareness of machines (Bulling, Roggen, & Tröster, 2008).

Applying a combined approach of Continuous Wavelet Transform (CWT) and threshold comparisons for feature extraction, Bulling et al. has shown that it is possible to detect saccade movements revealing clues about the context and in particular the reading activity of a subject (Bulling, Roggen, & Tröster, 2008).

Apart from those applications, parts of the research have been motivated by the idea to enhance the quality of life especially for disabled people who suffer from severe motor impairments (Barea, Boquete, Mazo, & López, 2002; Teccea, Gips, Olivieri, Pok, & Consiglio, 1998).

People suffering from neurodegenerative diseases like amytropic lateral sclerosis (ALS) are often locked into their own body, whereas the oculomotor system, being one of the last capabilities remaining, has a certain resistance to the derogating process (Dhillon, Singla, Rekhi, & Jha, 2009).

With respect to that motivation, several eye-based spelling devices using a virtual keyboard on a screen have been designed to provide a means of communication for that particular target group (Dhillon, Singla, Rekhi, & Jha, 2009; Hori, Sakano, Miyakawa, & Saitoh, 2006; Usakli & Gurkan, 2010).

Those interfaces were proposed as an alternative to already existing, commercial applications as described by Krolak et al. (Krolak & Strumillo, 2009), which are vision-based and entail the necessity of a bulky headgear.

The work presented here embraces the idea of a virtual keyboard to form the context for the investigation of different methods of feature extraction, which is the main focus of this study.

Further, most keyboard interfaces use the eye blink, for the selection command. However, eye blinks can also occur involuntarily, wherefore a long blink is proposed here as an alternative as it is more distinguishable from involuntary ones which tend to be shorter.

The methods underlying most of the approaches mentioned earlier are time-based and some are even designed without significant feature extraction stage in the reasoning that the interface should be kept simple and without complicated algorithms (Usakli & Gurkan, 2010).

Thus, the investigation being done here attempts to answer the question, whether simplicity does always correlate with efficiency. It aims at a comparative study of time, frequency-based methods as well as hybrid approaches, which combine both domains.

With regards to spectral analysis, only parametric methods are applicable to EOG signals, the amplitudes of which range between 50 to 3500 µV, with its major components below 35 Hz (Barea, Boquete, Mazo, & López, 2002). A non-parametric FFT would hence result in a very poor frequency resolution due to a small amount of samples available for a particular time interval.

Parametric approaches, on the contrary, attempt to model the system based on a set of coefficients. Those models, which are mostly autoregressive (AR) systems due to their good prediction results, have successfully been utilized for the analysis of EEG characteristics (Faust, Acharyaa, Allen, & Lin, 2008; Tseng, Chen, Chong, & Kuo, 1995), but have not sufficiently been taken into account for eye-based interfaces.

Further on, investigations have revealed that wavelet-based procedures are superior to frequency-based methods for certain applications (Fargues & Bennett, 1995) and are even found to be promising for biomedical systems (Unser, 1996).

The multi-resolution approach, considering the given signal at several distinct scales, was employed for signal compression (Bhandari, Khare, Santhosh, & Anand, 2007), which is noise reduction in more general terms, as well as for the extraction of specific information.

The latter aspect led Magosso et al. to propose solutions capable of detecting slow eye movements (SEM) during certain sleep phases in the EOG signals (Magosso, Ursino, Zaniboni, & Gardella, 2006) and moreover of identifying patterns in the brain activity during an epileptic seizure (Magosso, Ursino, Zaniboni, & Gardella, 2009).

In contrast to the examination of energy distributions across all scales, as studied in the articles mentioned, this study utilizes the discrete wavelet transform (DWT) to obtain an efficient representation of the data within the chosen feature space.

The results are compared with AR-modeling as well as straightforward time-based template matching used by Usakli & Gurkan, (2010) for instance.

Efficiency is hence studied here as a combination of sufficient performance, which is necessarily accompanied by an optimal class separation, and small computation overhead.

After describing the experimental setup and data acquisition, the paper proceeds with a general elaboration of the methodology used. Finally the results are appraised and concluded with the online EOG-based virtual keyboard implementation.

# METHODOLOGY

The project scope involves the acquisition of EOG data, which is used for a subsequent off-line analysis to justify the methods for feature extraction and classification stage. Finally an on-line system is implemented and connected to the front-end speller.

## Data Acquisition

The EOG data was recorded from the five participants listed in Table 1 (all had normal or corrected to normal vision), whereas subjects 1 to 3 participated in two sessions, one of which was the basis for the off-line investigations. The other session, which was also recorded from the remaining two subjects, provided the training sets for the evaluation of the on-line implementation later on.

***Table 1**. Participant Listing*

| *Subject* | *Age* | *Gender* | *Vision* |
|---|---|---|---|
| *1* | *24* | *Male* | *Glasses* |
| *2* | *23* | *Male* | *Normal* |
| *3* | *27* | *Female* | *Normal* |
| *4* | *24* | *Male* | *LASIK* |
| *5* | *26* | *Male* | *Glasses* |

Active Ag/AgCl electrodes were placed above the right eyebrow and in a vertical plane on the malar prominence below the eye (Denney & Denney, 1984) to form the vertical pair and on the outer canthi of both eyes near the temples for the horizontal electrode pair.

Since these were used in terms of differential pairs, referenced to each other, a fifth reference electrode was not necessary. Only a less intrusive ground reference for the ADC hardware was attached to the left hand of the subject.

The participants were required to perform eye movements in four directions and long blinks as requested in separate sections on a computer monitor. The sessions, which were recorded for the investigation, also involved random sections, where the individual produced EOG objects (i.e. eye movements and blinks) following the order which the subject determined earlier. To prevent fatigue during the session, subjects were instructed occasionally to perform short blinks during the collection of ten EOG objects in each session.

## Pre-processing

Using the EOG signals that were sampled with 256 Hz, a vertical and a horizontal channel were computed by referencing, in order to eliminate common noise and drift effects.

Subsequently, they were passed to a cascade of two digital, low-order Bessel filters acting as a bandpass with the lower cut-off frequency at 1 Hz and the higher at 5 Hz. The filter type was chosen due to its linear phase characteristic in the passband without any ripples in the amplitude response. The lower cut-off frequency eliminates parts of the EOG signal for the benefit of a very stable baseline, the components of which overlap the information inseparable in the lower bands. Preliminary investigation found this to be optimal for the following stages of feature extraction and classification.

Finally, avoiding aliasing effects, down-sampling by a factor of four was carried out, which resulted in a Nyquist frequency of 32 Hz.

Further operations for on-line setup are described as follows. First, blocks of

112 data samples were obtained from a sliding window, which were moved in steps of 16 samples through the signal. Hence, the buffer will be large enough to cover the duration of a long blink and will be updated every 250 ms in the final on-line solution.

## Feature Extraction

The features are extracted with three different approaches.

First, a straightforward time-based approach takes the whole buffer content without any further operations, resulting in a 112 dimensional feature space. The classification in this case is exclusively done using a nearest neighbor classifier (with single neighbor, 1NN) similar to Usakli and Gurkan, 2010.

This classifier computes the Euclidean distance as a measure of similarity to all sets of stored class templates for each EOG object type and assigns the nearest class to the test feature vector. Thus, this approach is also referred to as template matching.

Second, due to the obvious limitations of FFT based methods, a parametric, frequency-based approach using AR-modeling was chosen. As a result, a set of coefficients was estimated approximating the physical system with the following equation, where $m$ is the chosen model order:

$$H(z) = \frac{1}{\sum_{j=1}^{m} a_j z^{-j}} \quad (1)$$

In this context, the method proposed by Burg (1975) is used in its order-recursive implementation to obtain the parameters. The algorithm calculates the coefficients $a_j$, being characteristic of the data, based on the optimization criterion:

$$\epsilon_r = \sum_{i=m}^{N-1} \left( \left| e_{fr} \right|^2 + \left| e_{br} \right|^2 \right) \rightarrow Min \quad (2)$$

where $N$ is the buffer size, $r$ the current order with regards to the recursion and $e_{fr}$ and $e_{br}$ the forward and backward prediction errors, respectively.

Using the Levinson-Durbin recursion, the method finally provides $m$ coefficients, which are used as features stretching an $m$ dimensional feature space. Due to inaccuracies and computations related to the autocorrelation function, this method was preferred to the Yule-Walker approach (Hoon, Hagen, Schoonewelle, & Dam, 1996).

By minimizing the Akaike Information Criterion (AIC), the necessary order was found and fixed to be nine corresponding to a sufficient estimation of the data block contents. This leads to ten distinct features separating the classes. However, the first coefficient is normally chosen to be 1 and hence can be discarded when used as features.

Finally, wavelet decomposition based on DWT is employed to obtain specific features in a particular time and frequency resolution.

In each stage, the signal is decomposed into an approximation space $V_j$ stretched by scaling functions $\varphi_{j,n} = 2^{-j/2} \cdot \varphi(2^j x - n)$ and a detail space $W_j$ stretched by wavelet functions $\psi_{j,k} = 2^{-j/2} \cdot \psi(2^j x - k)$, with the following conditions fulfilled:

$$V_j = V_{j-1} \cup W_{j-1} \quad (3)$$

Hence

$$V_0 \subset V_1 \ldots \subset V_j \quad (4)$$

For feature extraction, only the projection on the approximation space is kept, which is needed to compute the next approximations of the subsequent, coarser decomposition stage. For this projection, the Haar basis, being piecewise constants on unit intervals, was used with its coefficients $h_0 = 1 / \sqrt{2}$ and $h_1 = 1 / \sqrt{2}$, where according to Mallat (1989):

$$\forall n \in \mathbb{Z},\ h_n = \langle \varphi_{\mathrm{j},0}, \varphi_{\mathrm{j-1,n}} \rangle \quad (5)$$

As proposed, based on the relationship between the coefficients of the refinement equations and FIR filters found by Mallat (1989), the implementation was realized with a bank of quadrature mirror filters, which are followed by a downsampling operator of factor two in each stage. Thus, after four stages of decomposition the input buffer results in seven approximation coefficients to be used as the representative feature set.

For an objective evaluation of the class separation capabilities for each proposed method, a measure $R_{ij}$ from the Davies Bouldin Index (Davies & Bouldin, 1979) is applied to the feature sets.

This measure computes the ratio between the compactness and distance of two classes $i$ and $j$ in the chosen feature space:

$$R_{ij} = \frac{d_i + d_j}{\|\mu_i + \mu_j\|} \quad \mathrm{i,j} \in \{1,2 \ldots, q\} \tag{6}$$

Assuming a number of $Q$ labeled reference vectors for each of the $q$ classes, the centroid $\mu_i$ and the class scatter $d_i$ are obtained as follows:

$$\mu_i = \frac{1}{Q} \sum_{\mathrm{r} \in \mathrm{C_i}} \mathrm{r} \quad \mathrm{i} \in \{1,2 \ldots, q\} \tag{7}$$

$$d_i = \frac{1}{Q} \sum_{\mathrm{r} \in \mathrm{C_i}} \|\mathrm{r} - \mu_i\| \quad \mathrm{i} \in \{1,2 \ldots, q\} \tag{8}$$

For all operations, the Euclidean distance was used to measure the norm for the separation of two objects in the feature space.

## Classification

Taking the extracted feature set into consideration, the classification stage was to decide to which of the defined classes the current signal sequence was to be assigned. In order to distinguish between all EOG objects and to discard involuntary short blinks, six classes have been set up:

| | | |
|---|---|---|
| $C_1$ | → | Eyes move down |
| $C_2$ | → | Eyes move up |
| $C_3$ | → | Long blink |
| $C_4$ | → | Eyes move left |
| $C_5$ | → | Eyes move right |
| $C_6$ | → | Short blink |

Apart from the 1NN classifier, which was the only classifier used with template matching, two other classification methods have been employed here: Artificial Neural Network (ANN) and Linear Discriminant Analysis (LDA).

The two stages related to the pre-processing were only triggered in case of a prominent activity in one of the channels. This reduced unnecessary computation and minimized wrong decisions due to a random sequence in the buffer, which would not represent any of the defined classes.

The detection was done by applying a threshold to the absolute value of the signal, which corresponds to the square root of the signal energy. If the threshold, which was adjusted for each subject, was exceeded, the activity was evaluated as containing significant eye movement or blink data.

Moreover, based on preliminary investigations, the number of classes to be separated is reduced to four for the vertical channel and to two for the horizontal channel, since left and right movements of the eye will only influence the potential induced to the electrodes on the outer canthi, whereas the other four EOG objects affect the vertical channel only.

For each channel, a neural network using the multilayer perceptron (*MLP*) structure was trained with reference feature vectors using the backpropagation algorithm along with a sigmoid transfer function for each neuron. The hidden unit size was fixed at 30.

The pooled covariance matrices and class means were computed based on the training data for both channels and passed to the linear statistical classifier (i.e. LDA).

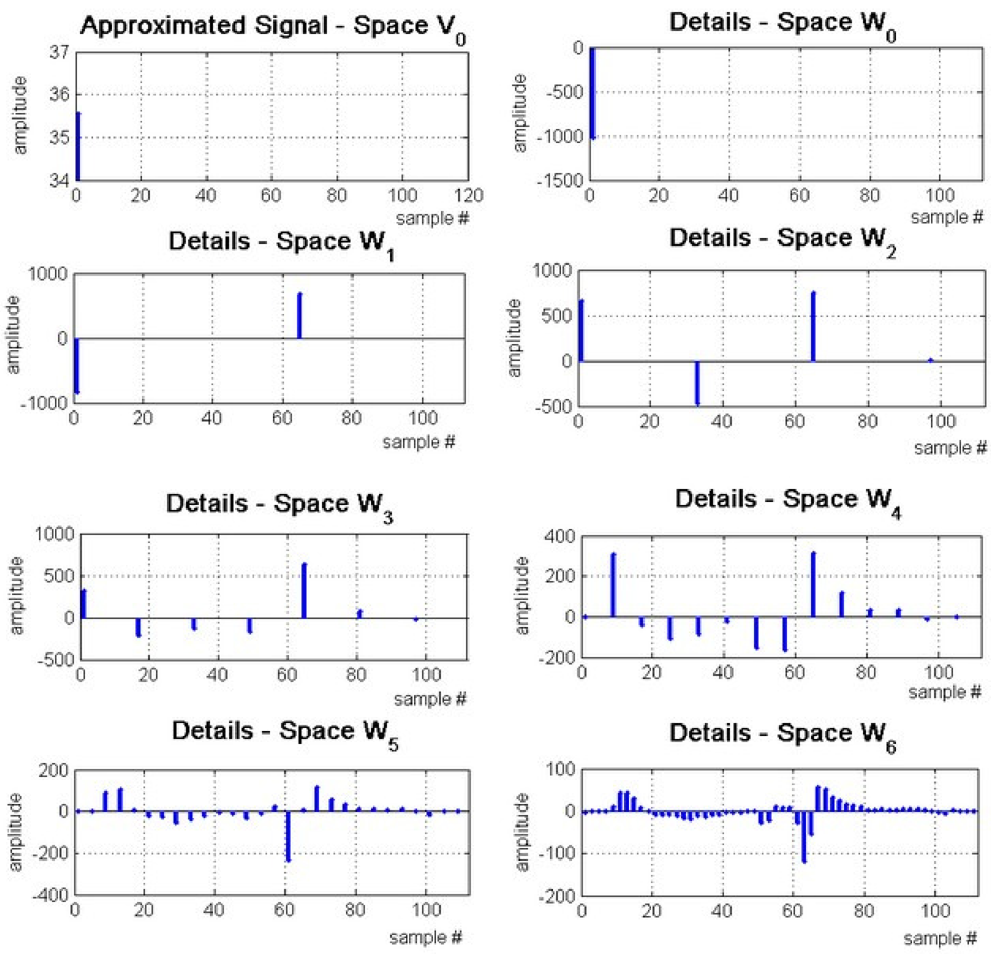


***Figure 1**. Haar wavelet decomposition for the class 'Eyes Down'*

The computation costs for both ANN and LDA were not significantly different simple because the training is done off-line and during on-line testing, both classifiers perform similarly in speed. In this context, adjustments that will change the amount of training data will not affect the execution time.

## RESULTS AND DISCUSSION

In order to explore the different vector spaces and the proportion of the overall information they contain, recorded EOG data have been decomposed using the 112 samples buffer content.

As an example, consider the class 'Eyes Down'. As it is apparent, the decomposition results in seven detail spaces $W_j$ and one approximation space $V_0$. In each level of decomposition, the algorithm takes every sample after convolving with the orthogonal filters accordingly.

By examining the different resolutions, it is obvious that the coefficients significantly exceed 500 μV from space $W_2$ onwards, which corresponds to the fact that its major components are located in the subsequent spaces.

Consequently, a decomposition using four stages has been chosen as appropriate, which stops with the projection indicated below:

$$V_4 = V_3 \cup W_3 \tag{9}$$

*Table 2. Class separation indices*

| | $C_1 \leftrightarrow C_2$ | $C_1 \leftrightarrow C_3$ | $C_2 \leftrightarrow C_3$ | $C_4 \leftrightarrow C_5$ | $C_1 \leftrightarrow C_6$ | $C_2 \leftrightarrow C_6$ | $C_3 \leftrightarrow C_6$ |
|---|---|---|---|---|---|---|---|
| Burg | 1.866 | 2.845 | 3.600 | 2.664 | 4.208 | 2.295 | 3.902 |
| Haar | 0.241 | 0.397 | 0.756 | 0.169 | 0.248 | 0.507 | 0.591 |
| Templates | 0.456 | 0.578 | 0.927 | 0.259 | 0.480 | 0.574 | 0.677 |

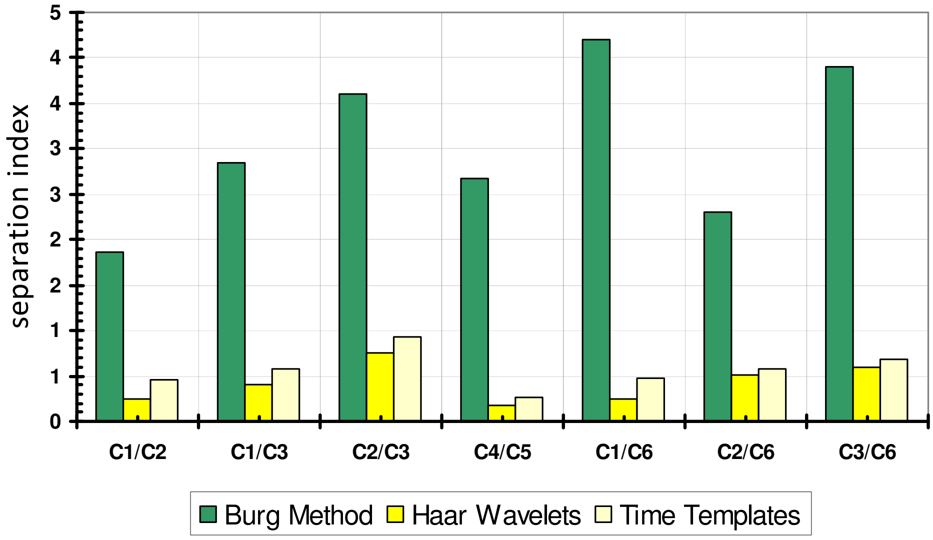


***Figure 2****. Illustrated separation indices*

After discarding the details in $W_3$, the algorithm results in seven order approximation coefficients, which constitute the new feature set to be sent to the classifier. Apart from this feature representation, nine coefficients from a ninth order AR-system have been computed using the Burg method in its recursive implementation.

Finally, the time sequences being used as feature sets by the template matching approach were also taken into consideration for evaluating the class separation capabilities by means of the Davies-Bouldin Index. The results comparing the scores for each of the three methods are listed in Table 2 and Figure 2.

First, it is apparent that the frequency based algorithm performed much worse compared to the others. For that method, the index values denote results that were always above 1.5 and significantly higher compared to the other feature sets. In this context a lower index indicates better separability. The results were confirmed by applying this feature extraction to the offline data in order to obtain the rate of correct detections. As expected, it was found that the classes were hardly distinguishable which led to unreliable detections with less than 50% success rate.

Looking at the theory behind this approach, one will notice that the criterion for deriving the AR-system is based on a minimal sum of squared prediction errors. Consequently the algorithm is not sensitive with regard to phase information, which is however an essential requirement for a good class separation for this data.

Considering the time characteristics of the EOG data, it could be seen that a considerable part of the information about the eye movement is carried by the phase component.

In rough terms, the class 'Eyes Left' could be seen as the class 'Eyes Right' multiplied with minus one. This corresponds to a phase shift of 180°

(though this is somewhat simplified explanation); the information about which is however eliminated by squaring the prediction error. As a conclusion the optimization criterion provides no clue about a possibly different phase. Both eye movements could have the same representation with respect to their AR coefficients.

This reasoning also applies for the other classes, which has been confirmed by the indices in Figure 2.

In the optimal case, the AR predictor generates white noise as its output, which means that it entirely compensates the physical system being modeled. The predictor filter hence models the amplitude of all components, whereas it suppresses the phase information. A particular phase shift in the random components of the white noise at the output has no influence on the noise energy represented by the prediction error.

Therefore this frequency-based method is not useful with regards to the separation of EOG objects. Furthermore, all the major frequency components for all objects are located in the same frequency bands, as confirmed by the wavelet decomposition earlier. Thus, a frequency-based algorithm, which aims at being successful, *must* retain the phase information, in order to keep the information necessary to distinguish between the classes.

The separation index also reveals an improvement in case of employing the wavelet based approach compared to template matching and the unprocessed time sequence.

Consequently, apart from reducing the number of dimensions significantly from 112 to 7, which will result in less computation requirements, the decomposition into the approximation space $V_3$ will also simplify the problem for the classifier, which could lead to higher performance.

Due to its localization in the time domain, the features obtained from the DWT retain the phase information in particular related to the frequency range extracted by the decomposition procedure. An example is shown for the movement 'Eyes Down' from the data of participant 3 in Figure 3.

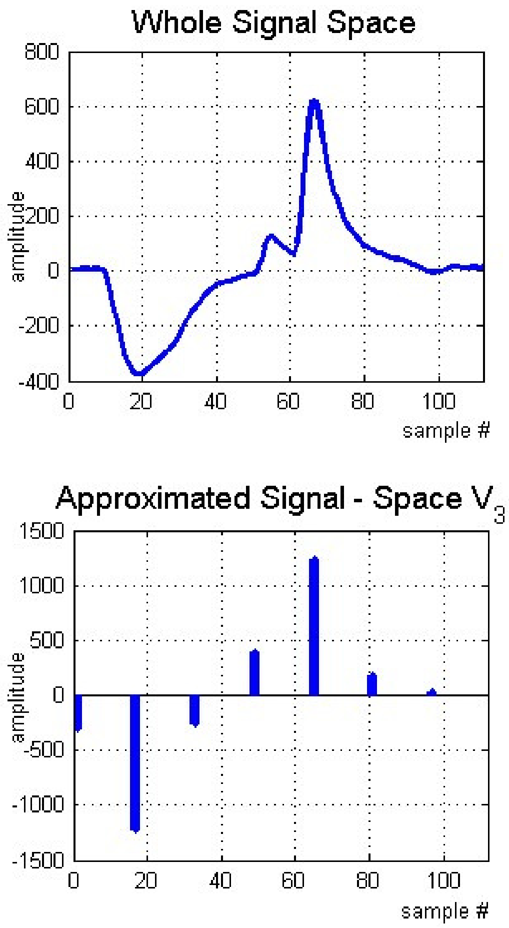


***Figure 3.*** *Features for 'Eyes Down'*

The features contain the information located in the vector space $V_3$, which corresponds to

$$V_3 = V_0 \cup \left(\bigcup_{j=0}^{2} W_j\right). \qquad (10)$$

For the classification step, template matching was combined with a 1NN classifier as suggested by Usakli and Gurkan (2010). Likewise, the proposed wavelet approach here was utilized along the LDA, ANN as well as 1NN.

*Table 3*. *Performance results*

| *Method* | *P1* | *P2* | *P3* | Average |
|---|---|---|---|---|
| DWT+1NN | 98.10% | 93.30% | 80.00% | 90.50% |
| DWT+ANN | 90.60% | 86.70% | 80.00% | 85.80% |
| templates | 96.20% | 93.30% | 80.00% | 89.80% |
| DWT+LDA | 98.10% | 93.30% | 93.30% | 94.90% |

Each of those four combinations (i.e. 1NN (template matching), DWT+LDA, DWT+ANN and DWT+1NN) was applied offline to the random sequences of the EOG data for each of the first three participants P1, P2 and P3. The performance in terms of correctly detected EOG objects is summarized in Table 3. The results in Table 3 shows that all methods worked well especially for participant 1 and slightly less well with participant 3; this could be caused by the 'Long Blink' and 'Eyes Up' signal characteristics which were very similar for participant 3. In this case, the combination of DWT and LDA performed better than the others, although it resulted in a similar outcome compared to DWT and 1NN for the other two subjects.

Moreover, the ANN classifier achieved hardly more than 90% of correct detections, which is in any case inferior to the performance of the other methods.

Considering the computation costs as a more practical aspect for both the feature extraction and classification stages, yields the following reasoning. Figures 4 and 5 illustrate the costs with respect to two different independent variables: the size of the input buffer and the number of labeled reference vectors (i.e. training data) per class. The diagrams indicate the number of multiplications and summations (being primitive operations), in which the algorithms can be decomposed.

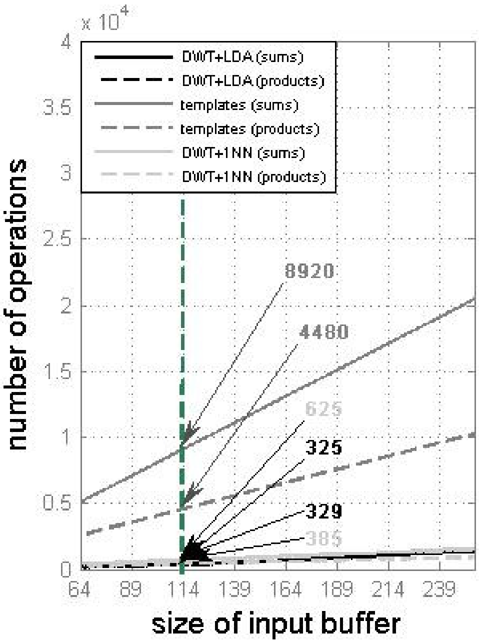


***Figure 4.*** *Number of operations versus buffer size for the vertical channel having four classes*

In this context, it has to be mentioned, that the square root function has not been counted for the kNN classifier, since there are different ways for implementing it depending on the aspired accuracy or exact realization. Referring to the latter point, a hardware implementation might use a LUT or the Taylor series, for instance.

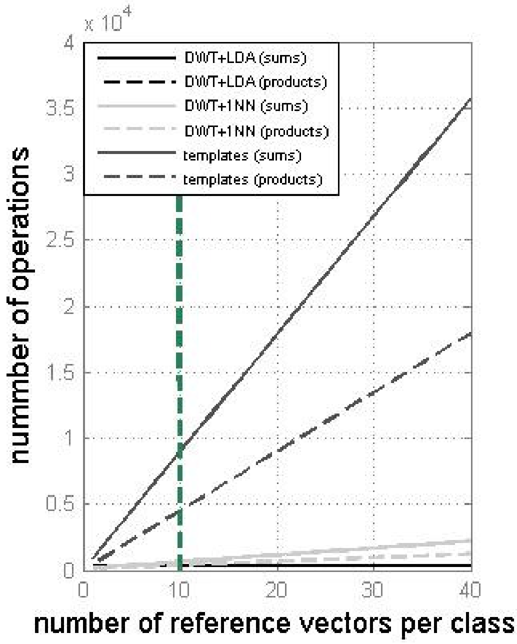


***Figure 5.*** *Number of operations versus amount of training data for four classes*

In any case, it will add a certain number of operations to the combinations utilizing kNN. Depending on the application, a tradeoff will always arise between hardware resources and computation time in order to cover those costs.

For the results, it can be deduced that template matching without any feature extraction turns out to be very inefficient. Applying this method entails more computation, whereas the computation cost of the other combinations after extracting the wavelet approximation coefficients is very close to each other.

Increasing the input buffer size seems not to be very useful from the practical point of view, since it will increase the latency and the hardware requirements of the system. Thus an implementation should aim at a reduction rather than at increased buffer size.

On the other hand, it is more likely that the training data per class might vary, since it is directly related to the performance of the system. In this case, the LDA and ANN classifiers are superior compared 1NN due to the independence on the amount of training data.

Conversely, the kNN classifier would have to compute more distance measures, if the number of reference vectors, i.e. training data, per class is increased.

Thus, its performance and the fact that the combination of DWT and LDA exhibits least cost for the chosen parameter set, leads to it being the chosen approach with regards to the implementation stage.

## IMPLEMENTATION

The algorithm explained earlier was embedded in a LabVIEW application running on a common desktop computer. A BioSemi recording device along with four active electrodes was used to acquire the EOG data from a subject interacting with the computer.

The implementation, the block diagram of which is shown in Figure 6, was placed into a general state machine which was working synchronously with one iteration requiring a time of 250 ms.

Taking the initial sampling rate of 256 Hz into account, 64 samples per iteration were passed as an input to the application illustrated in Figure 6.

The pre-processing step included the cascade of Bessel filters for each of the two channels, which were computed from the electrodes by referencing.

Since aliasing was avoided due to the elimination of higher frequency components by the cascaded filter, the signal was down-sampled again by a factor of four resulting in 16 samples per iteration for each channel.

This data was shifted into a register, which contained 2x112 samples by discarding the 16 oldest samples from the register. Based on the buffer size of 112 samples, the features were extracted by calculating the seven approximation coefficients obtained from the fourth decomposition level of the Haar wavelet transform.

This decomposition was efficiently implemented by a cascade of downsampling FIR filters as shown in Figure 7.

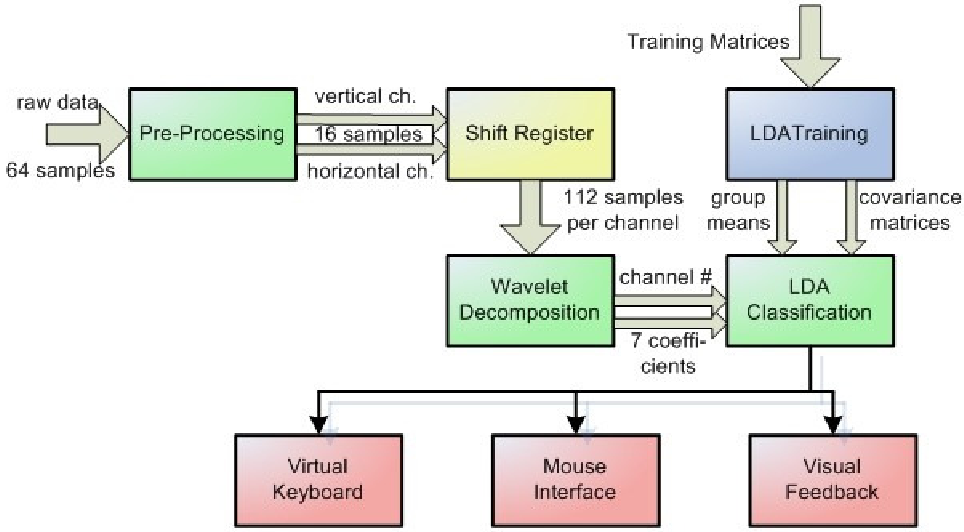


***Figure 6.*** *Flow diagram of the EOG interface*

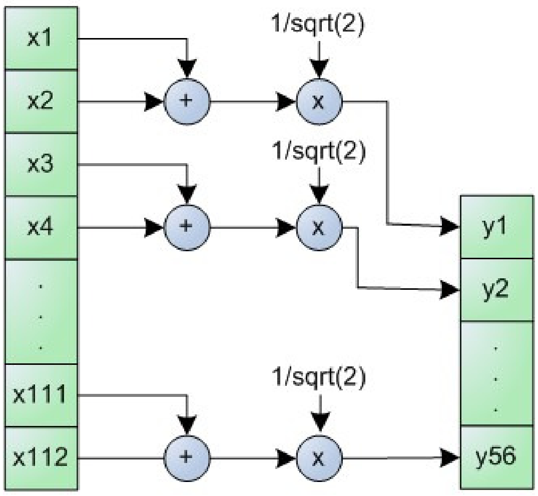


***Figure 7.*** *First level of wavelet decomposition*

Next, the feature extraction involved a threshold based onset detection for both channels, which triggered the extraction only if a prominent activity was detected in one of the channels. Thus, the data block in which the activity was detected and the number of the channel namely horizontal or vertical is passed to the LDA classifier.

This stage decided on the class assignment based on the covariance matrix and the class means obtained from the offline training data. Strictly speaking, the application worked with two classifiers, one for each channel, which were selected according to the output of the feature extraction.

If there was no detected activity in any channel, both the wavelet decomposition and the classification steps will not be executed, which reduces unnecessary computation.

The threshold applied to the data stream corresponds to a simple decision between two cases. Since its purpose was only to detect whether there is a significant increase in the signal energy or not, it does not overly rely on prior knowledge about the EOG characteristics.

The front-end focuses on a virtual keyboard application, which is illustrated in Figure 8.

It can also provide an interface controlling the Windows mouse cursor as well as a simple visual and textual feedback using the processed data from the buffer’s content.

After receiving the classification results from the statistical classifier, they could be connected to any front-end by mapping the decisions to actual commands, which are interpreted by the machine.

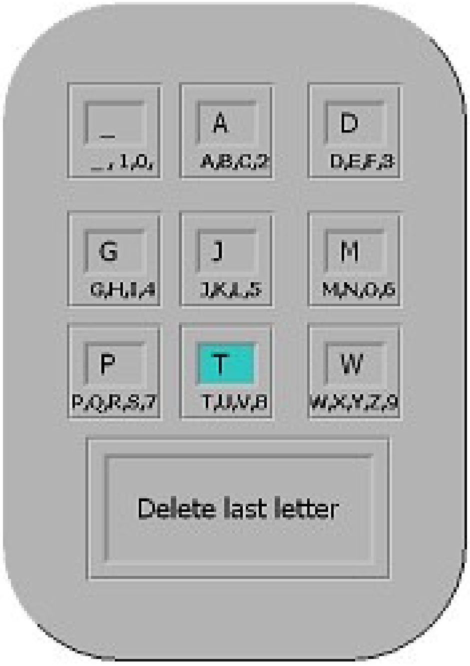


***Figure 8***. *Layout for the virtual keyboard*

In this particular keyboard, short blinks were not used as commands and hence were ignored. Consequently, the keyboard used only the four directions of the eye movements to move the cursor in discrete steps from one key to next one and the long blink was used as the selection command.

Differing from other virtual keyboard applications, this front-end intentionally utilized a mobile phone layout, which provided three to four letters per key (Dhillon, Singla, Rekhi, & Jha, 2009; Hori, Sakano, Miyakawa, & Saitoh, 2006; Usakli & Gurkan, 2010). This was seen as a promising alternative, since eye movements are expected to be much more tiring compared to blinks. Consequently, the layout would decrease the number of movements for the benefit of blinks, since the subject is required to select the correct letter by browsing through the options provided by each key.

As a structured way of designing the keyboard, a Mealy state machine was used, the output of which is generated depending on the current state of the machine and its input.

The basic layout consists of a 3x4 key array that covers the whole alphabet and a few special characters including ‘space’ as well as a return key to make corrections if necessary.

Additionally, it should be mentioned that after a completed classification, the classifier output is blocked until the buffer is emptied. This is to avoid a double classification of the same EOG data.

## EVALUATION

The interface was evaluated with five subjects as listed in Table 1. Before interacting with the system, training data was collected, i.e. ten reference vectors for each of the six classes.

Those were pre-processed and arranged in a matrix representation to be fed into as inputs to the LabVIEW online system. The actual testing and the training were done in two different sessions, since it was found that the algorithm works robustly across sessions without significant decrease in performance.

However, the system does not work well across subjects, i.e. it is subject dependent, since appearance of the EOG objects might differ due to the way a subject performs them. For example, the speed and angle in which a subject moves his or her eyes would be different for each subject. Furthermore, the signals might be weaker or stronger depending on the ability of the subject.

Each participant was asked to carry out the defined EOG objects, whereas the amount of correct classifications was used as a measure of the performance for the online system. Comparable to the offline measurements, the number of right detections was between 96% and 100%. The performance is slightly better compared to the offline, possibly as the subject could unconsciously adapt to the system.

The second task (as also used for evaluation purposes by Usakli & Gurkan, 2010)) was to write the word ‘WATER’ on the virtual keyboard. For this task, the number of corrections and the elapsed time were recorded as a measure of the performance for the virtual keyboard.

All the subjects were able to produce the word as the text output and none of them needed more than one correction. In total, each participant had two trials, in which he was required to write the specified word. The time needed to complete the five letters was used to calculate the average rate indicating how many letters could be written in one minute. For all subjects, the rate was between 3 and 5 letters per minute even though all subjects except subject 1 interacted with the system for the first time. A trained individual as participant 1 was able to finish the task within one minute. Movies on the operation of the virtual keyboard could be found at http://csee.essex.ac.uk/staff/palaniappan.

It should be mentioned that the fact of using long blinks as a class prevented the system from being faster, since the buffer size had to be chosen accordingly to cover the duration of a long blink, which could be more than one second.

Similar applications described in other articles were exclusively employing short blinks for selection, which would increase the writing performance in terms of time, but it will not be robust against involuntarily performed blinks.

Finally the participants were asked to fill in a questionnaire, in order to receive a subjective feedback about the system. All participants were of the opinion that blinks were less tiring compared to eye movements and hence supporting the phone layout rather than the QWERTY layout. On the whole, they agreed that the EOG interface was intuitive and easy to use.

## CONCLUSIONS

This paper has considered several strategies to improve the analysis and processing of EOG signals. It was shown that frequency based approaches, in particular the AR parametric approach using Burg's method, is not a good feature extraction approach in order to achieve a robust class separation.

The introduced method of Haar wavelet decomposition was found to be superior to other approaches with regard to class separation and especially in terms of computation overhead due to the reduction of the dimension size in the feature space.

The combination of this feature extraction stage and the statistical LDA classifier was implemented with a virtual keyboard acting as the user front-end and evaluated to give promising detection rates.

In particular, real-time or embedded systems would benefit from improved feature extraction since hardware resources and processing power are very limited for such systems. Moreover this would allow using more sophisticated classifiers instead of the computationally intense kNN approach as used by others. Systems, which have to meet strict timing and have to be predictable, would benefit from the independence on the amount of training data as shown for the case of using LDA.

## REFERENCES

Barea, R., Boquete, L., Mazo, M., & López, E. (2002). System for assisted mobility using eye movements based on Electrooculography. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on , 10* (4), 209 - 218.

Bhandari, A., Khare, V., Santhosh, J., & Anand, S. (2007). Wavelet based compression technique of Electro-oculogram signals. *3rd Kuala Lumpur International Conference on Biomedical Engineering 2006. 15*, pp. 440-443. Kuala Lumpur: Springer.

Bulling, A., Roggen, D., & Tröster, G. (2008). It's in your eyes - Towards context-awareness and Mobile HCI Using Wearable EOG Goggles. *ACM Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp 2008)*, (pp. 84-93). Seoul, Korea.

Bulling, A., Roggen, D., & Tröster, G. (2008). Robust recognition of reading activity in transit using wearable electrooculography. *Proceedings of the 6th International Conference on Pervasive Computing (Pervasive 2008)* (pp. 19-37). Sydney, Australia: Springer.

Burg, J. P. (1975). *Maximal Entropy Spectral Analysis, PhD Thesis.* Stanford University, Department of Geophysics.

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *Pattern Analysis and Machine Intelligence, IEEE Transactions on , PAMI-1* (2), 224 - 227.

Denney, D., & Denney, C. (1984). The eye blink electro-oculogram. *British Journal of Ophthalmology , 64*, 225-228.

Dhillon, H. S., Singla, R., Rekhi, N. S., & Jha, R. (2009). EOG and EMG based virtual keyboard: A brain-computer interface. *Computer Science and Information Technology, 2nd International Conference on* (pp. 259-262). Beijing, China: IEEE.

Estrany, B., Fuster, P., Garcia, A., & Luo, Y. (2009). EOG signal processing and analysis for controlling computer by eye movements. *Pervasive Technologies Related to Assistive Environments, Proceedings of the 2nd International Conference on.* Corfu, Greece: ACM.

Fargues, M. P., & Bennett, R. (1995). Comparing wavelet transforms and AR modeling as feature extractiontTools for underwater signal classification. *Proceedings of the Conference Record of the Twenty-Ninth Asilomar Conference on Signals, Systems and Computers (ASILOMAR '95)* (pp. 915 - 919). Pacific Grove, CA , USA: IEEE.

Faust, O., Acharyaa, R., Allen, A., & Lin, C. (2008). Analysis of EEG signals during epileptic and alcoholic states using AR modeling techniques. *IRBM , 29* (1), 44-52.

Hoon, M. J., Hagen, T. H., Schoonewelle, H., & Dam, H. v. (1996). Why Yule-Walker should not be used for autoregressive modelling. *Annals of Nuclear Energy , 23* (15), 1219-1228.

Hori, J., Sakano, K., Miyakawa, M., & Saitoh, Y. (2006). Eye movement communication control system based on EOG and voluntary eye blink. *Computers Helping People with Special Needs, Lecture Notes in Computer Science , 4061*, 950-953.

Kiamini, M., Alirezaee, S., Perseh, B., & Ahmadi, M. (2008). A wavelet based algorithm for ocular artifact detection in the EEG signals. *Multitopic Conference, 2008. INMIC 2008. IEEE International*, (pp. 165-168). Karachi.

Krolak, A., & Strumillo, P. (2009). Eye-blink controlled human-computer interface for the disabled. *Human-Computer Systems Interaction, Advances in Soft Computing , 60*, 123-133.

Magosso, E., Ursino, M., Zaniboni, A., & Gardella, E. (2006). A wavelet based method for automatic detection of slow eye movements: A pilot study. *Medical Engineering and Physics , 28* (9), 860–875.

Magosso, E., Ursino, M., Zaniboni, A., & Gardella, E. (2009). A wavelet-based energetic approach for the analysis of biomedical signals: Application to the electroencephalogram and electro-oculogram. *Applied Mathematics and Computation , 207* (1), 42–62.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence , 11* (7), 674-693.

Teccea, J. J., Gips, J., Olivieri, C. P., Pok, L. J., & Consiglio, M. R. (1998). Eye movement control of computer functions. *International Journal of Psychophysiology , 29* (3), 319-325.

Tseng, S.-Y., Chen, R.-C., Chong, F.-C., & Kuo, T.-S. (1995). Evaluation of parametric methods in EEG signal analysis. *Medical Engineering & Physics , 17* (1), 71-78.

Unser, M. (1996). Wavelets, statistics, and biomedical applications. *Proceedings of the 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing (SSAP '96)*, (pp. 244 - 249). Corfu , Greece.

Usakli, A. B., & Gurkan, S. (2010). Design of a novel efficient human–computer interface: An electrooculagram based virtual keyboard. *Instrumentation and Measurement, IEEE Transactions on , 59* (8), 2099 - 2108.