

Comparative Performance Analysis of a Commercial Wearable EOG Glasses for an Asynchronous Virtual Keyboard

Nathaniel Barbara
Department of Systems and Control
Engineering, Faculty of Engineering,
University of Malta, Msida,
MSD2080, Malta
nathaniel.barbara@um.edu.mt

Tracey A. Camilleri
Department of Systems and Control
Engineering, Faculty of Engineering,
University of Malta, Msida,
MSD2080, Malta
tracey.camilleri@um.edu.mt

Kenneth P. Camilleri
Department of Systems and Control
Engineering, Faculty of Engineering,
University of Malta, Msida,
MSD2080, Malta
kenneth.camilleri@um.edu.mt

Conventional electrooculography (EOG) signal acquisition systems, apart from requiring expensive equipment, necessitate electrode gel for conductivity and electrode wires, which may obstruct the field of view and restrict user movements, thus making the setup itself impractical for regular use. Recently however, a cheaper, wireless, gel-free, sleek and wearable alternative EOG recording device, inconspicuously incorporating dry electrodes mounted on an ordinary-looking pair of glasses, has been made commercially available. This work compares this device, known as the JINS MEME EOG glasses, against a gold-standard conventional gel-based EOG setup. Specifically, while gaze displacement estimation errors of $1.32 \pm 0.26^\circ$ (and $1.67 \pm 0.26^\circ$) in the horizontal (and vertical) directions were obtained using the conventional setup, the corresponding errors obtained using the JINS MEME were $1.97 \pm 0.34^\circ$ (and $1.85 \pm 0.30^\circ$). Saccades and blinks were also found to be reliably detected and labelled using these two EOG recording modalities, with average labelling accuracies exceeding 99%. The two modalities were also compared when used in real-time to interface with an asynchronous EOG-based virtual keyboard having a QWERTY layout. Specifically, average writing speeds across subjects of 11.9 ± 4.4 and 9.9 ± 3.6 characters per minute were obtained using the conventional EOG setup and the JINS MEME respectively, which have been shown to improve substantially with user experience with the system. These results have demonstrated that the JINS MEME offers a feasible wearable alternative to the conventional EOG setup, which is more practical for eye movement-based applications.

Electrooculography, wearable technology, mobile computing, human-computer interaction, virtual keyboard.

1. INTRODUCTION

Technology has literally seeped through the lives of human beings living in the 21st century, with applications being controlled by simple taps on touch screens or button presses on keyboards and remote controls. These applications however, could not be easily controlled by people suffering from mobility impairments, such as those diagnosed with Amyotrophic Lateral Sclerosis (ALS) or paralysed stroke patients, who may find it difficult to interact with such applications in their surroundings. In spite of the limitations imposed by such conditions, the eyes are generally regarded as being the last organs to be affected. Hence, an eye movement-based human-computer interface (HCI) system provides a good alternative communication channel to these intelligent devices through eye movements alone, thereby giving these individuals more independence and an enhanced quality of life (Usakli & Gurkan, 2010).

Off-the-shelf eye movement-based HCIs typically use videooculography (VOG)-based eye movement recording techniques. These refer to the use of video-cameras and image processing algorithms to track the user's ocular pose. VOG-based techniques are however known to be very computationally intensive, susceptible to the lighting conditions, generally require external illumination sources and sensitive to the user's movement, amongst other limitations. Although most VOG-based systems are desktop setups, wearable VOG-based trackers are also commercially available. These however, still require bulky and heavy head-mounted equipment, are not unobtrusive and still very expensive (Heide, et al., 1999; Bulling, et al., 2009).

Electrooculography (EOG), on the other hand, is another eye movement recording technique which records the electrical activity that is generated by the human eyes. Specifically, the eyes could be regarded to behave like an electric dipole, with the

negative and positive poles at the retina and cornea respectively. This creates a potential difference, known as the corneo-retinal potential, with a magnitude varying in the range of 0.4-1.0 mV, which creates an electric field. The electrical signal generated by this field is recorded through EOG using electrodes attached to the face, in close proximity to the eyes. Specifically, the conventional EOG setup requires two pairs of such electrodes, which are typically aligned horizontally and vertically with the eyes. These electrodes are usually attached to the face using adhesive plaster, where electrode gel is also used to enhance the conductivity between the skin and the respective electrodes (Heide, et al., 1999; Bulling, et al., 2009). Using such an EOG signal acquisition setup however, has some disadvantages as assistance is typically required to set up and the electrode gel may also dry out with long-term use and hence, might require re-application. Additionally, the setup has some ergonomic issues as the electrodes themselves are not aesthetically pleasing and comfortable to wear. This is particularly due to the multiple lead wires feeding bio-signal amplifiers which restrict the subject's movement and which may interfere with the subject's field of view, thus making the setup impractical for regular use (Bulling, et al., 2009).

Several attempts have been made in the literature to address the practicality of using EOG for eye movement-based HClS, specifically by mounting the electrodes on spectacle frames (Bulling, et al., 2009; Vehkaoja, et al., 2005). However, the aesthetic and wearable concerns of such systems are still an issue particularly because these setups still require head-mounted equipment (Vehkaoja, et al., 2005) and wired leads from the electrodes to a processor (Bulling, et al., 2009). Recently however, JINS Company Limited has developed a truly wireless and wearable EOG signal acquisition system: a device which inconspicuously incorporates three dry electrodes mounted as the bridge and nose pads of an ordinary-looking pair of glasses to record EOG data, as shown in Figure 1b (Kano, et al., 2015). This device, known as the JINS MEME EOG glasses, apart from it addressing the aesthetic concerns of the conventional setup, allows for more independent use. This is because its sleek design, the wireless communication links to computers and the use of dry electrode technology, which eliminates electrode gel application, enhance its practicality even further (Kano, et al., 2015).

However, since the JINS MEME has been very recently made commercially available, there is limited literature (Kano, et al., 2015; Barbara & Camilleri, 2016) on whether this modality offers a feasible alternative to the conventional setup. Therefore, this work aims to compare the MEME glasses against the conventional EOG setup, specifically in terms of: (i) the gaze displacement estimation performance; (ii) the saccade and blink

labelling performance; and (iii) when used in real-time to interface with an asynchronous EOG-based virtual keyboard. The methods used for gaze displacement estimation and eye movement detection and labelling, as well as the virtual keyboard, have been presented in our previous work (Barbara, et al., 2018).

State-of-the-art EOG-based gaze angle estimation solutions typically use two distinct models, one for each EOG signal component, to estimate the horizontal and vertical ocular displacements (Acuña, et al., 2014). This approach however, assumes that the horizontal EOG signal component is solely a function of the horizontal ocular displacement, and similarly the vertical EOG signal component is solely a function of the vertical ocular displacement. This may not be typically the case in practice, such as due to possible misalignment between the horizontal and vertical EOG electrode pairs and the horizontal and vertical ocular dipole axes, for which the two EOG components would have to be used jointly for general ocular pose estimation (Barbara, et al., 2018; Bulling, et al., 2008). Our previous work (Barbara, et al., 2018) has systematically investigated this and proposed a two-channel input regression model, using features extracted from both EOG signal components jointly, to estimate the gaze displacements. In this work, a similar model is adopted to cater for possible such misalignment between the electrodes and the ocular dipole axes in the case of the conventional setup, and also to cater for the unconventional non-orthogonal electrode configuration of the MEME glasses.

In contrast to state-of-the-art eye movement detection methods, typically restricted to the detection of blinks and saccades in four (Tangsuksant, et al., 2012) or eight (Yamagishi, et al., 2006) directions only, this work detects and discriminates between saccades of any displacement and in any direction, and blinks, by processing EOG signals in real-time. Specifically, this work uses a threshold-based method to detect eye movements, and a cascade of a parametric and a signal-morphological classifier based on the EOG peak and gradient features to distinguish between general saccades and blinks (Barbara, et al., 2018).

These techniques are then used to interface with an asynchronous EOG-based virtual keyboard (Barbara, et al., 2018), to compare the JINS MEME against the conventional EOG setup when used in real-time. The use of a regression model to estimate the gaze displacements allows users to reach any icon from any position on the screen directly and hence, without being restricted to hover over icons in discrete fixed-size steps as is typically the case in the literature (Tangsuksant, et al., 2012; Yamagishi, et al., 2006). Furthermore, the detection of eye movements in real-time allows the application to be controlled asynchronously and

hence, without having the user restricted to perform eye movements within specific intervals to interact, as in some state-of-the-art applications (Barbara & Camilleri, 2016; Nathan, et al., 2012).

The rest of the paper is divided as follows; Section 2 focuses on the acquisition and processing of EOG signals, where the methods adopted for gaze displacement estimation, eye movement detection and labelling as well as the virtual keyboard, which were used to compare the two EOG recording modalities, are also summarised. This is followed by Section 3, which presents the performance results obtained using the conventional gel-based EOG setup and the JINS MEME EOG glasses. These results are then discussed in Section 4, while Section 5 concludes this paper.

2. METHOD

This section presents the acquisition and pre-processing done on EOG signals recorded using both the conventional EOG setup and the JINS MEME. The methods used to estimate the ocular pose, detect and distinguish between different eye movements as well as the asynchronous EOG-based virtual keyboard are also presented.

2.1 Acquisition and processing of EOG signals

The EOG data acquisition protocol was approved by the University Research and Ethics Committee (UREC) at the University of Malta, and each subject provided their informed consent before each signal acquisition session. Subjects were placed approximately 60 cm away from a 24 inch LCD screen, with their head held stationary using ophthalmic chin and forehead rests. During each session, subjects were asked to follow onscreen instructions to perform different eye movements. These were recorded using two different EOG recording modalities, specifically (i) using a conventional EOG setup comprising g.tec g.LADYbirdPassive gel-based ring electrodes attached to the face using adhesive tape; and (ii) using the wearable JINS MEME EOG glasses.

For the conventional EOG setup, the electrode configuration of Figure 1a was used, with one pair of electrodes placed next to the outer canthi and another pair placed above and below the right ocular socket. A reference ('R') and a ground ('G') electrode were placed on the mastoid behind the left ear and on the forehead respectively, as shown. EOG signals were recorded using the g.tec g.USBamp bio-signal amplifier (g.tec medical engineering GmbH Austria) at a sampling frequency of 256 Hz. The potential differences between the horizontally- and vertically-aligned electrode pairs were computed, to yield what are generally referred to as the horizontal and vertical EOG signal

components, denoted by $EOG_h^{(G)}(t)$ and $EOG_v^{(G)}(t)$ respectively:

$$EOG_h^{(G)}(t) = V_1(t) - V_2(t) \quad (1)$$

$$EOG_v^{(G)}(t) = V_3(t) - V_4(t) \quad (2)$$

where $V_1(t)$, $V_2(t)$, $V_3(t)$ and $V_4(t)$ represent the EOG potential recorded from electrodes 1-4 shown in Figure 1a respectively.

Conversely, the electrodes of the JINS MEME are mounted as the bridge and nose pads of an ordinary-looking pair of glasses, as shown in Figure 1b. In this case, the horizontal and vertical EOG signal components, denoted by $EOG_h^{(M)}(t)$ and $EOG_v^{(M)}(t)$ respectively, are obtained using (Kanoh, et al., 2015):

$$EOG_h^{(M)}(t) = V_L(t) - V_R(t) \quad (3)$$

$$EOG_v^{(M)}(t) = -(V_L(t) + V_R(t))/2 \quad (4)$$

where $V_L(t)$ and $V_R(t)$ represent the bipolar voltage recorded by the electrodes marked as L and R in Figure 1b respectively, with electrode C serving as reference. The sampling frequency in this case was 100 Hz. EOG signals arising due to saccades and blinks in sample EOG data recorded from both the conventional EOG setup and the MEME glasses are shown in Figure 2.

EOG signals are typically contaminated by external noise sources, such as electrical grid interference and high-frequency electromyographic (EMG) activity due to underlying facial muscle contractions, such as during squinting or smiling. EOG signals generally also suffer from baseline drift, which refers to a low-frequency signal that interferes with the acquired EOG data, typically arising due to

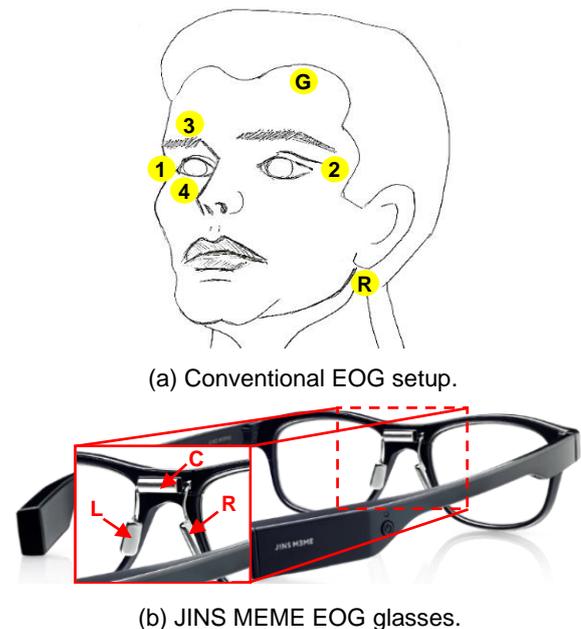
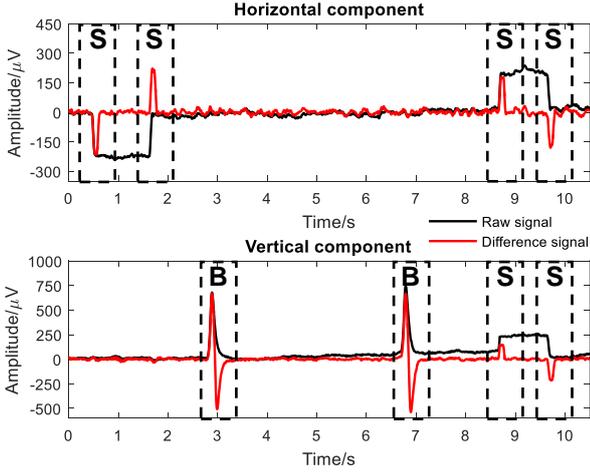
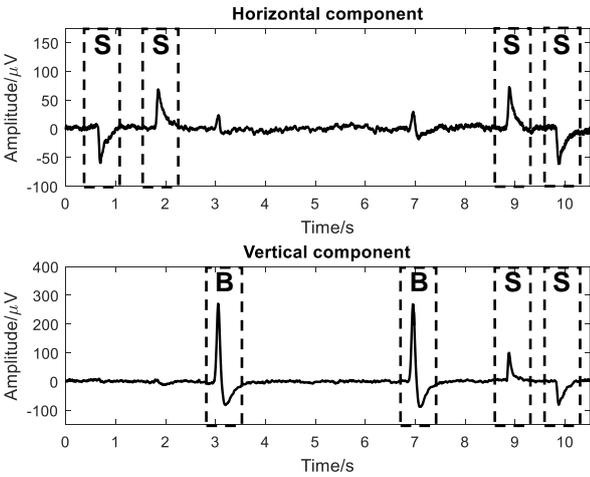


Figure 1: EOG electrode configuration for the two considered modalities.



(a) Signals recorded using the conventional EOG setup.



(b) Signals recorded using the JINS MEME.

Figure 2: Saccades (S) and blinks (B) as manifested in the EOG signals of the two considered modalities.

background signal interference or electrode polarisation (Bulling, et al., 2009). In this regard, considering first EOG data recorded using the conventional setup, a bandpass filter between 0 Hz and 30 Hz, as well as a 50 Hz notch filter were applied to mitigate high-frequency EMG noise and grid interference respectively. The baseline drift effect was mitigated by computing a difference signal as follows (Barbara, et al., 2018):

$$\Delta_{\tau}EOG_h^{(G)}(t) = EOG_h^{(G)}(t) - EOG_h^{(G)}(t - \tau) \quad (5)$$

$$\Delta_{\tau}EOG_v^{(G)}(t) = EOG_v^{(G)}(t) - EOG_v^{(G)}(t - \tau) \quad (6)$$

where $\Delta_{\tau}EOG_h^{(G)}(t)$ and $\Delta_{\tau}EOG_v^{(G)}(t)$ denote the horizontal and vertical differenced EOG components respectively. The value for τ considered in this work was chosen to be equal to the duration taken by the eye to traverse saccadically from one angular pose to an opposite extreme pose. This value was determined empirically to be 25 samples at a sampling frequency of 256 Hz, equivalent to $\tau = 0.0977$ s. The effect of such pre-processing on EOG signals recorded using the conventional setup is demonstrated in Figure 2a.

Contrastingly, due to the internal filtering involved in the case of the MEME glasses, the effect of the different noise sources discussed previously was found to be negligible. Hence, no further pre-processing was applied in this case. The pre-processed EOG signals recorded using the conventional setup, $\Delta_{\tau}EOG^{(G)}(t)$, (shown in Figure 2a) and the raw internally-filtered EOG signals from the MEME glasses, $EOG^{(M)}(t)$, (shown in Figure 2b) have similar characteristics and thus, similar gaze displacement estimation and eye movement detection algorithms are applied.

2.2 Gaze displacement estimation

Based on the signal characteristics of $\Delta_{\tau}EOG^{(G)}(t)$ and $EOG^{(M)}(t)$, the EOG peak amplitudes extracted from the horizontal and vertical EOG components arising due to saccadic movements, P_h and P_v respectively, are used to estimate the horizontal and vertical ocular angular gaze displacements, $\widehat{\Delta\theta}_h$ and $\widehat{\Delta\theta}_v$ respectively. Based on the observations of our previous work (Barbara, et al., 2018) and as discussed in Section 1, $\widehat{\Delta\theta}_h$ and $\widehat{\Delta\theta}_v$ are estimated using a two-channel input linear regression model, \mathcal{M} , which considers P_h and P_v jointly.

In general, \mathcal{M} involves linear combinations of fixed non-linear basis functions (BFs) of the input variables, which could be represented as (Bishop, 2006):

$$\mathbf{y}(x, \mathbf{W}) = \mathbf{W}^T \boldsymbol{\phi}(x) \quad (7)$$

where $\mathbf{y} = (\widehat{\Delta\theta}_h, \widehat{\Delta\theta}_v)^T$ is a $K = 2$ dimensional column vector output, $\mathbf{x} = (P_h, P_v)^T$ is a $D = 2$ dimensional column vector input, \mathbf{W} is an $L \times K$ matrix of parameters and $\boldsymbol{\phi}(x)$ is an L dimensional column vector whose elements are the BFs $\phi_j(x)$ for $j = 0, \dots, L - 1$, with L representing the total number of BFs considered.

The BFs $\phi_j(x_n)$ considered in \mathcal{M} are grouped according to the maximum order M . Specifically, different sets of polynomial BFs of the two dimensional input vector x_n up to the M^{th} degree were considered for any integer $M > 0$. Table 1 presents the BFs considered for $M \in \{1, 2, 3\}$; a similar scheme applies for higher orders (Barbara, et al., 2018).

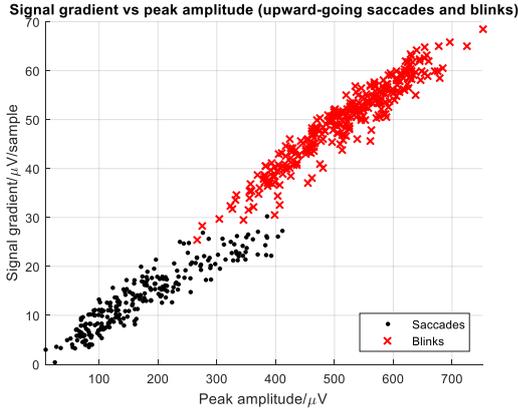
2.3 Real-time eye movement detection and labelling

The aim behind this algorithm is to detect and distinguish between different eye movements, specifically saccades and blinks, from EOG signals in real-time. Our previous work (Barbara, et al., 2018) has demonstrated that the distributions of the EOG peak, P_v , and gradient, G_v , features extracted from $\Delta_{\tau}EOG_v^{(G)}(t)$ corresponding to saccades and

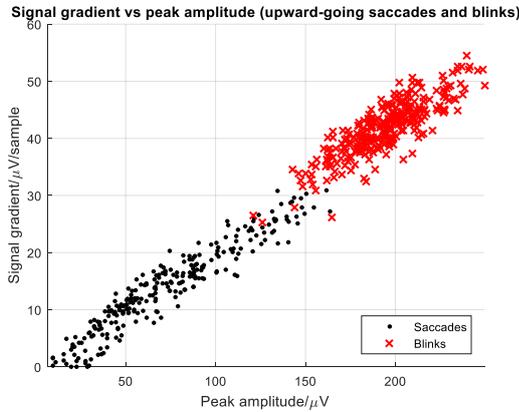
Table 1: The BFs considered in \mathcal{M} , grouped according to the maximum order, M , where $\mathbf{x}_n = (x_{1n}, x_{2n})^T$.

BFs		$\phi_0(\mathbf{x}_n) = 1$	$\phi_1(\mathbf{x}_n) = x_{1n}$	$\phi_2(\mathbf{x}_n) = x_{2n}$	$\phi_3(\mathbf{x}_n) = (x_{1n})^2$	$\phi_4(\mathbf{x}_n) = (x_{2n})^2$	$\phi_5(\mathbf{x}_n) = x_{1n}x_{2n}$	$\phi_6(\mathbf{x}_n) = (x_{1n})^3$	$\phi_7(\mathbf{x}_n) = (x_{2n})^3$	$\phi_8(\mathbf{x}_n) = (x_{1n})^2x_{2n}$	$\phi_9(\mathbf{x}_n) = x_{1n}(x_{2n})^2$...
M	1	✓	✓	✓								
	2	✓	✓	✓	✓	✓						
	3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

blinks are characterised by two clusters, implying that these two features could be used to discriminate between the two ocular movements. As shown in Figure 3, this characteristic was also observed when considering MEME acquired EOG data. Therefore, a similar eye movement detection and discrimination method to that in (Barbara, et al., 2018) was adopted for both modalities as shown in Figure 4. The algorithm is summarised below, where $\Delta_r EOG^{(G)}(t)$ and $EOG^{(M)}(t)$ are henceforth denoted in general as $EOG(t)$.



(a) Data recorded using the conventional EOG setup.



(b) Data recorded using the JINS MEME.

Figure 3: (P_v, G_v) feature distribution corresponding to upward-going saccades and blinks for one subject. Similar distributions were obtained across subjects.

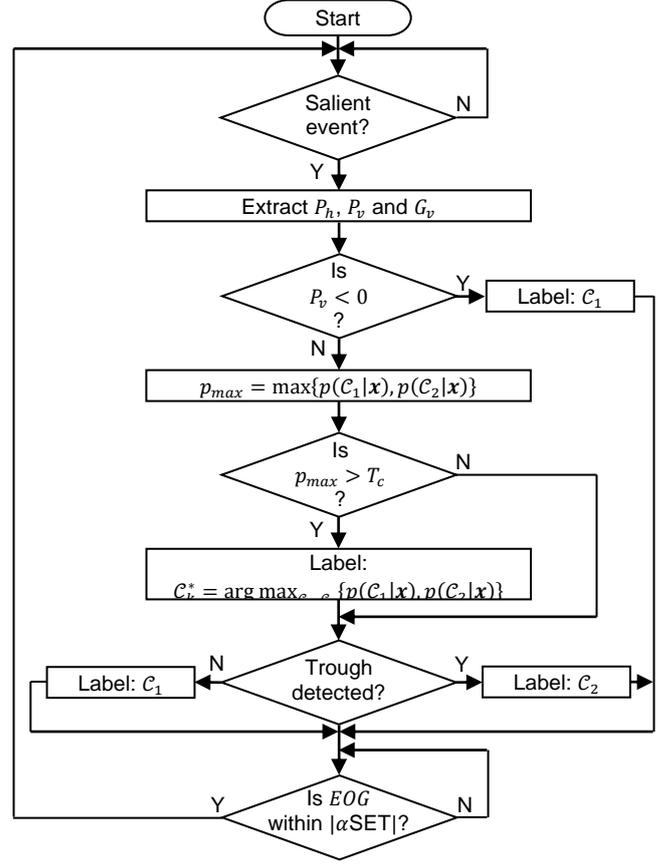


Figure 4: A flowchart showing the process of detecting and labelling saccades and blinks from EOG signals.

Specifically, a salient event threshold (SET) is applied to both EOG signal components to detect the onset of any eye movement, which is generally referred to as a salient event. Figure 5 shows a general oblique saccadic movement, where the instant at which the onset is detected is labelled by marker 'X'. The peak amplitude values P_h and P_v in the horizontal and vertical EOG signal components respectively, together with the vertical EOG signal component gradient, G_v , corresponding to the detected salient event are extracted. P_v and G_v are then used in a classifier to label the detected salient event as a saccade or blink. If the event is labelled as a saccade, P_h and P_v are passed to the regression model for gaze displacement estimation; conversely, if the event is labelled as a blink, no gaze displacement is estimated. Since, as shown in Figure 2, only upward-going saccadic movements are morphologically similar to blinks, the algorithm labels salient events with $P_v < 0$ immediately as saccades.

The P_v and G_v features are assigned to one of 2 classes C_k , with $k = 1, 2$ representing saccades and blinks respectively, using an optimal Bayes' classifier. Specifically, the saccade and blink (P_v, G_v) feature distributions are modelled as bivariate Gaussians and are used to obtain the class posterior probabilities $p(C_k|x)$.

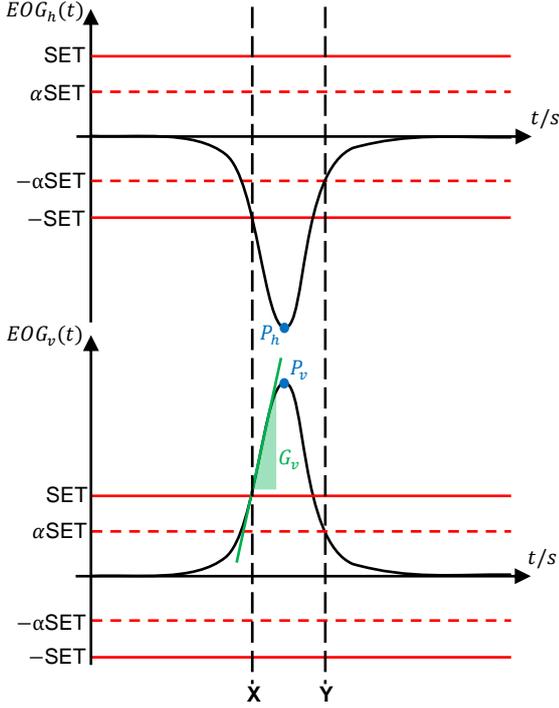


Figure 5: Diagrammatic representation of salient event detection from (pre-processed) EOG signals.

A probability threshold T_c is then applied to the maximum posterior class probability $p_{max} = \max\{p(C_1|\mathbf{x}), p(C_2|\mathbf{x})\}$ obtained. If $p_{max} > T_c$, the salient event is instantly labelled according to the maximum a posteriori (MAP) decision rule, specifically as $C_k^* = \arg \max_{C_1, C_2} \{p(C_1|\mathbf{x}), p(C_2|\mathbf{x})\}$. Conversely, if $p_{max} \leq T_c$, a reject option is applied and the event is labelled according to the morphology of the signal instead of using the classifier. Recall that, the blink signal morphology in both $\Delta_\tau EOG_v^{(G)}(t)$ and $EOG_v^{(M)}(t)$ is characterised by a negative trough which immediately follows a positive crest, as shown in Figure 2. In this regard, the presence or absence of a negative trough is used to label the event as a blink or a saccade respectively, specifically by determining whether a negative threshold B^- in the vertical component is exceeded within a time interval T_d after the occurrence of the positive peak. To address the unlikely occurrence of misclassifications using the MAP decision rule at $p_{max} > T_c$, this trough detection method is also run in the background to detect and correct any classifier labelling errors.

The algorithmic cycle of Figure 4 waits until the EOG potential in both components falls back to baseline level, specifically within $|\alpha SET|$, where $0 \leq \alpha \leq 1$; this is marked by label 'Y' in Figure 5. At that instant, the algorithm starts scanning afresh for new salient events. The α parameter has to be determined judiciously such that it is sufficiently small not to permit false detections arising due to noise in the decaying signal trail, and sufficiently

large not to delay the restart of the scanning process in order not to miss consecutive saccades (Barbara, et al., 2018).

Since the blink trough amplitude values generally vary from one subject to another, the B^- trough threshold value is determined for each subject individually. Specifically, B^- is determined from a set of recorded blinks using (Barbara, et al., 2018):

$$B^- = \lambda \text{med}\{\mathcal{E}_{train}\} \quad (8)$$

where \mathcal{E}_{train} is a dataset containing blink trough amplitude values extracted after a subject-specific training data recording session, $\text{med}\{\mathcal{E}_{train}\}$ represents the median value of \mathcal{E}_{train} and λ is a constant between 0 and 1. Choice of the median statistic was to ensure that the B^- threshold estimation is robust against outlier values, typically arising due to subject-related mistakes during data collection.

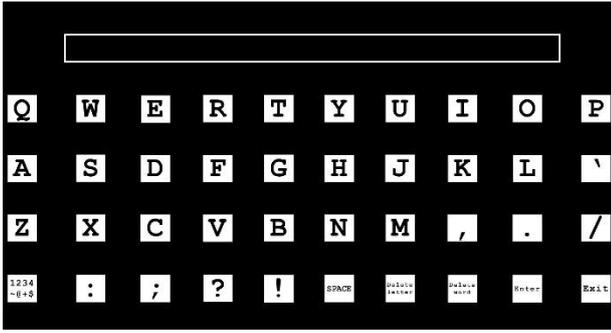
2.4 EOG-based virtual keyboard

The real-time eye movement detection algorithm to detect and label saccades and blinks in real-time as well as the two-channel input linear regression model for gaze displacement estimation discussed in the previous two sections were then used to interface with an asynchronous EOG-based virtual keyboard whose interface will be discussed next.

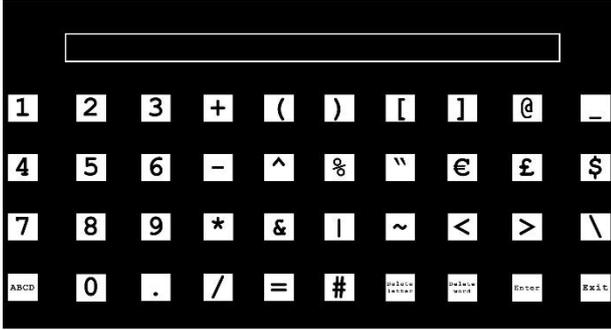
As shown in Figure 6, the virtual keyboard organises letters in a QWERTY arrangement together with numbers, punctuation symbols and other characters on two 40-icon menus. Both menus also include action icons to delete last transcribed character or word, emulate the pressing of the 'Enter' key, exit the application and toggle between the two menus.

Typing using this EOG-based virtual keyboard requires the user to perform saccadic movements to reach the desired icon and to select it using a dwell time-based validation technique, specifically by fixating for a period of 2 s. In this regard, when the eye movement detection technique discussed in Section 2.3 detects the i^{th} saccade, the corresponding horizontal and vertical angular gaze displacements, denoted by $\widehat{\Delta\theta}_{h_i}$ and $\widehat{\Delta\theta}_{v_i}$ respectively, are estimated using the two-channel input regression model outlined in Section 2.2. $\widehat{\Delta\theta}_{h_i}$ and $\widehat{\Delta\theta}_{v_i}$ are then used to determine the user's point of gaze (POG), which is mapped to the centroid of the nearest icon on the screen.

A green circular progress bar, centred over this nearest icon, is displayed to indicate the user where his/her POG has been mapped to and the time left before the icon is successfully selected. During this period the user is allowed to blink naturally, as any single blink gestures are ignored by the system and hence, do not reset the dwell-timer. Once the 2 s dwell period elapses the icon is highlighted in green



(a) Menu A.



(b) Menu B.

Figure 6: Icon layout of the virtual keyboard.

to provide confirmation that the icon was successfully selected.

If however, based on $\widehat{\Delta\theta}_{h_i}$ and $\widehat{\Delta\theta}_{v_i}$, the subject's POG is estimated to fall within the writing bar region, the system temporarily ignores the user's ocular activity, while an orange circular progress bar is displayed on the icon coinciding with the last estimated POG. This allows the user to verify that the correct character(s) have been successfully transcribed in the writing bar and refocus his/her POG on the indicated icon. A duration of 3 s was found to be a reasonable duration for this purpose.

This application also addresses the Midas touch problem (Barbara & Camilleri, 2016) by allowing users to pause and resume the system at their intent, while any ocular movement is ignored by the system during a no-control state. This was implemented through double-blink detection where specifically, if two consecutive blinks are detected within a 1 s interval, the system is toggled between its active and paused states. When the system is paused, a red-coloured circular cursor is displayed over the icon corresponding to the last estimated POG; this indicates where the user must focus his/her gaze at before performing another double-blink gesture to resume. This pause-resume feature is particularly useful to recover after a gaze displacement mis-estimation, in which case the user can perform a double-blink gesture to pause the system, focus his/her gaze on the indicated icon and resume by performing a second double-blink gesture.

3. RESULTS

In this section, the gaze displacement estimation and eye movement detection performance results obtained using EOG signals recorded from both the conventional EOG setup as well as from the JINS MEME are presented. Furthermore, the results obtained when the two EOG recording modalities were used in real-time to interface with the EOG-based virtual keyboard are also presented.

3.1 Gaze displacement estimation performance

For this performance analysis, a total of 600 saccadic movements, of random angular displacements varying in the range of $\pm 40^\circ$ horizontally and $\pm 20^\circ$ vertically, and 300 blinks were recorded from six healthy subjects (4 males and 2 females, mean age = 23.3 ± 3.9). The EOG data recorded using the conventional setup and the JINS MEME glasses simultaneously was pre-processed as discussed in Section 2.1 and any subject-related mistakes, such as blinking when subjects were instructed to perform a saccade or vice versa, were manually identified and discarded from further analysis. The saccadic EOG peak amplitudes corresponding to the remaining saccadic events were divided into four folds, each with 120 trials.

The gaze displacement estimation performance obtained using the JINS MEME was compared to that obtained using the gold-standard conventional EOG setup using the root mean square error (RMSE), defined as:

$$E = \sqrt{\frac{1}{N} \sum_{j=1}^N \left\{ \left(\Delta\theta_{h_j} - \widehat{\Delta\theta}_{h_j} \right)^2 + \left(\Delta\theta_{v_j} - \widehat{\Delta\theta}_{v_j} \right)^2 \right\}} \quad (9)$$

where N denotes the total number of trials in each fold, $\widehat{\Delta\theta}_{h_j}$ and $\widehat{\Delta\theta}_{v_j}$ represent the estimated horizontal and vertical angular gaze displacements respectively, while $\Delta\theta_{h_j}$ and $\Delta\theta_{v_j}$ denote the target horizontal and vertical angular gaze displacements respectively. Additionally, the mean absolute angular errors in the horizontal and vertical directions were also computed.

The optimal model order M^* was determined systematically, by statistically comparing the two-fold cross-validated RMSE obtained for increasing $M \in \{1, 2, 3, 4\}$. Specifically, the subject-specific M^* was determined by noting that minimum order beyond which there is no statistically significant improvement in estimation performance. For EOG signals recorded using both the conventional setup and the JINS MEME, a subject-dependent optimal order ranging between $M^* = 1$ and $M^* = 3$ was required. Using the subject-specific M^* values, the two-fold cross-validated RMSE and mean absolute angular errors in the horizontal and vertical

directions were computed, as tabulated in Table 2. These results show that a better gaze displacement estimation performance was obtained using EOG signals recorded through the conventional gel-based setup, with a difference in the average error of 0.65° and 0.18° in the horizontal and vertical directions respectively. This is also reflected in the overall gaze displacement RMSE, where the RMSE obtained using the conventional EOG setup was found to be statistically significantly better than the RMSE obtained using the JINS MEME glasses ($p < 0.05$); this was found to be consistently the case across all six subjects.

3.2 Eye movement detection performance

In this case, the same EOG data recorded in Section 3.1 was used to evaluate this performance. The optimal values for the T_c , T_d , λ , SET and α parameters pertaining to the real-time eye movement detection and labelling algorithm discussed in Section 2.3 were determined using the method reported in (Barbara, et al., 2018). The optimal values for these parameters for the two EOG recording modalities are summarised in Table 3. As shown, while the values for the probability threshold T_c and λ (used to estimate the blink trough detection threshold) are the same for EOG data recorded using the conventional setup and using the JINS MEME, the other parameter values are different for the two EOG recording modalities.

The longer value for the duration threshold T_d obtained for MEME acquired EOG data reflects the difference in the crest-to-trough intervals of the blink signal morphology in $\Delta_\tau EOG_v^{(G)}$ and $EOG_v^{(M)}$, as

Table 2: Gaze displacement estimation performance; HE: horizontal error, VE: vertical error.

Subject	Conventional setup			JINS MEME		
	RMSE /°	HE /°	VE /°	RMSE /°	HE /°	VE /°
S1	2.47	0.96	1.57	2.83	1.54	1.51
S2	2.67	1.32	1.66	3.83	2.28	1.96
S3	3.09	1.25	2.05	3.72	1.74	2.37
S4	2.55	1.25	1.43	3.33	1.78	1.88
S5	2.93	1.75	1.39	3.80	2.38	1.64
S6	3.02	1.41	1.89	3.47	2.12	1.76
Average	2.79 ± 0.26	1.32 ± 0.26	1.67 ± 0.26	3.50 ± 0.38	1.97 ± 0.34	1.85 ± 0.30

Table 3: Optimal T_c , T_d , λ , SET and α parameter values.

Parameter	Conventional setup	JINS MEME
T_c	0.75	0.75
T_d	0.23 s	0.38 s
λ	0.4	0.4
SET	$60 \mu V - 70 \mu V$	$25 \mu V - 35 \mu V$
α	0.0 – 0.1	0.2 – 0.3

shown in Figure 2. Furthermore, the relatively smaller optimal SET values required in the case of signals recorded using the MEME glasses reflect the relatively smaller dynamic range of data recorded using this modality when compared to the dynamic range of EOG data recorded using the conventional setup. On the other hand, the relatively higher α values required in the case of MEME acquired EOG signals reflect the fact that the EOG signals after any eye movement decay at a slower rate, when compared to the pre-processed EOG signals acquired using the conventional setup. This is clearly observable in Figure 2.

Using these optimal parameter values, the 10-fold cross-validated saccade, blink and eye movement labelling accuracy was evaluated, as tabulated in Table 4, where an eye movement refers to a blink or a saccade. In order not to bias the results, saccades with $P_v < 0$ were not considered for this performance analysis, as such events are labelled with certainty as saccades, as discussed in Section 2.3. The difference in eye movement labelling performance obtained using the two EOG recording modalities considered in this work was not found to be statistically significant ($p > 0.05$).

3.3 The EOG-based virtual keyboard

The JINS MEME EOG glasses were also compared against the conventional EOG setup when used in real-time to interface with the asynchronous virtual keyboard discussed in Section 2.4. Specifically, this was tested by 9 healthy subjects (5 males and 4 females, mean age = 23.6 ± 3.0) who were instructed to transcribe two strings, S_1 : 'HELLO' and S_2 : 'GOOD DAY', each for three times using the two EOG recording modalities separately. The order in which the two modalities were used in the trials was randomised to ensure unbiased results. Subjects were instructed to transcribe the strings correctly and completely and to select the *exit* icon after each trial. Prior to these trials, subjects performed a training session during which 100 saccades and 50 blinks were recorded, which were used to estimate the system parameters as discussed previously.

In this case, the two EOG recording modalities were compared against each other using two different

Table 4: Saccade (S), blink (B) and eye movement (EM) labelling accuracy.

Subject	Conventional setup			JINS MEME		
	S/%	B/%	EM/%	S/%	B/%	EM/%
S1	100.0	100.0	100.0	100.0	100.0	100.0
S2	100.0	100.0	100.0	100.0	98.0	99.0
S3	100.0	100.0	100.0	100.0	100.0	100.0
S4	100.0	100.0	100.0	100.0	100.0	100.0
S5	99.5	100.0	99.8	100.0	100.0	100.0
S6	100.0	100.0	100.0	100.0	100.0	100.0
Average	99.9	100.0	100.0	100.0	99.7	99.8

performance measures, specifically the writing speed (WS), in characters per minute (cpm), and the icon selection accuracy (ISA). The WS is defined as (Tangsuksant, et al., 2012; Yamagishi, et al., 2006; Nathan, et al., 2012):

$$WS = \frac{1}{N} \sum_{j=1}^N \left\{ 60 \frac{|S_j|+1}{D_j} \right\} \quad (10)$$

where N represents the total number of trials, $|S_j|$ denotes the length of the string transcribed in the j^{th} trial and D_j represents the time taken in seconds to transcribe S_j . Given that subjects were asked to transcribe the strings correctly and completely, then $|S_1| = 5$ and $|S_2| = 8$, where the +1 in the numerator represents the *exit* icon selection after each trial.

On the other hand, the ISA is defined as (Yamagishi, et al., 2006; Nathan, et al., 2012):

$$ISA = \frac{1}{N} \sum_{j=1}^N \left\{ 100 \frac{|S_{correct_j}|}{|S_{selected_j}|} \right\} \quad (11)$$

where N represents the total number of trials, $|S_{correct_j}|$ denotes the total number of intentionally-selected icons and thus also accounts for intentional *delete letter* icon selections, while $|S_{selected_j}|$ denotes the total number of icons selected during the j^{th} trial and thus also accounts for all unintentionally-selected icons and selections of the *exit* and *delete letter* icons.

The average WS and ISA results obtained using the two EOG recording modalities are tabulated in Table 5. The difference between the 11.9 ± 4.4 cpm and 9.9 ± 3.6 cpm WSs obtained when the conventional EOG setup and the JINS MEME EOG were used to interface with the virtual keyboard respectively were found to be statistically significant ($p < 0.05$), whereas no statistically significant difference has been found between the ISA results obtained using the two modalities ($p > 0.05$).

The theoretical maximum WS performance that could be achieved using the proposed EOG-based virtual keyboard is 30 cpm, which results due to the 2 s dwell time attributed to the successful selection of each icon. This is however, based on the assumption that no time is taken to traverse to the next icon after successfully selecting an icon, which is not the case in practice. Therefore, since the WS results documented in this work account for such delays, as well as other delays to correct mistakes arising due to system or user errors, the quoted performance results are realistic and practical values of the typing rate.

Typical user errors occurring during typing include instances where subjects fail to perform a double blink gesture immediately after a mis-estimation occurs to pause the system straightaway and gaze at the icon indicated by the red-coloured cue before resuming. Instead, users would typically follow their

Table 5: Average WS and ISA results obtained when the virtual keyboard was interfaced using the conventional EOG setup and the MEME EOG glasses.

Subject	Conventional setup		JINS MEME	
	WS /cpm	ISA /%	WS /cpm	ISA /%
S1	12.7±6.9	100.0	9.3±2.8	97.0
S2	12.7±4.6	98.5	10.3±1.9	100.0
S3	11.3±3.7	100.0	13.1±4.1	100.0
S4	12.1±5.0	100.0	6.5±1.4	100.0
S5	11.0±3.7	100.0	12.6±1.2	100.0
S6	7.8±2.0	100.0	7.2±4.4	100.0
S7	12.7±2.6	97.9	11.7±2.9	100.0
S8	16.4±2.5	100.0	11.8±3.2	96.4
S9	10.4±4.4	96.7	6.5±1.0	100.0
Average	11.9±4.4	99.2	9.9±3.6	99.3

natural instinct of looking at the icon to which their POG is mis-estimated to, resulting in the estimated POG to move further away from the desired icon. Others who performed the double-blink gesture immediately after a mis-estimation, also spent a considerable amount of time before performing a second double blink gesture to resume the system and continue typing.

3.3.1. Testing of the virtual keyboard by an experienced user

The above-mentioned user errors are however, expected to reduce drastically with user familiarity with the system. To demonstrate this, similar trials to those discussed in Section 3.3 were performed by an experienced user (S10). In this case, average WSs of 20.8 ± 2.4 cpm and 18.6 ± 3.3 cpm were obtained using the conventional EOG setup and the JINS MEME respectively, with an ISA of 100% using both systems. In view of these results it is clear that, with user familiarity, the WS performance is expected to improve substantially using both EOG recording modalities.

In addition, with further user familiarity, the dwell-time period required to successfully select an icon could also be reduced, thereby allowing for higher WSs to be achieved. In fact, this was tested by the same experienced user S10 using a revised dwell-time period of 1.5 s, who achieved average WSs of 29.8 ± 3.1 cpm and 26.8 ± 2.2 cpm using the conventional EOG setup and the MEME respectively. These results demonstrate that, by reducing the dwell period by just 25%, the average WS could easily improve by almost 50%.

4. DISCUSSION

The gaze displacement estimation and eye movement labelling performance results obtained using the conventional EOG setup and the JINS

MEME EOG glasses are discussed in this section together with the performance achieved when using the two modalities in real-time to interface with the proposed EOG-based virtual keyboard.

The angular gaze displacement estimation performance results obtained in Section 3.1 have demonstrated that a statistically significantly better estimation performance was obtained using EOG signals recorded through the conventional gel-based EOG setup when compared to that obtained using MEME acquired EOG data. This result could be justified by the fact that the EOG peak-gaze displacement data acquired using the MEME glasses is characterised by a higher variance when compared to that achieved using the conventional gel-based setup. For the virtual keyboard application demonstrated here however, this still permitted a full QWERTY keyboard, as shown in Figure 6, to be interfaced using the JINS MEME.

On the other hand, although the eye movement event labelling performance was found to be 0.2% better using the conventional EOG setup over the MEME glasses, this difference was not found to be statistically significant. Therefore, these results demonstrate that saccades and blinks could be reliably detected and distinguished in real-time from EOG signals recorded using both modalities.

Although a better gaze displacement estimation and eye movement labelling performance was achieved using the conventional EOG setup when compared to the performance obtained using the JINS MEME, this resulted in an improvement of only 2 cpm (18.6%) in the WS performance when using the real-time application. Therefore, these results show that JINS MEME is a reliable alternative device that could be used to interact in real-time with a state-of-the-art EOG-based virtual keyboard (Barbara, et al., 2018) which allows users to reach any icon from any location on the screen directly and asynchronously, unlike EOG-based virtual keyboards developed in the literature (Barbara & Camilleri, 2016; Tangsuksant, et al., 2012; Yamagishi, et al., 2006; Nathan, et al., 2012). In addition, the marginally better performance obtained using the conventional EOG setup comes at the cost of a system which is substantially more expensive than the JINS MEME and which is meant for in-lab studies. Furthermore, since the JINS MEME has a sleek design and does not require application of electrode gel, due to the incorporated dry electrodes, no set up assistance is required, unlike using the conventional gel-based EOG setup, which makes the JINS MEME more practical for regular use. The JINS MEME is also a wireless and wearable device, thus making it more comfortable to use without having the user's movement restricted by any lead wires feeding bio-signal amplifiers, while the field of view remains unobstructed.

Furthermore, the results in Section 3.3.1 have also demonstrated that the WS performance is also expected to improve substantially with user experience with the system. Specifically, while the user errors are expected to reduce drastically with user familiarity with the system, the dwell-time required to select an icon could also be reduced. In fact, average WSs of 29.8 ± 3.1 cpm and 26.8 ± 2.2 cpm using the conventional setup and the JINS MEME respectively were achieved by an experienced user and a revised 1.5 s dwell-time setting. Hence, these results also demonstrate that with user experience, state-of-the-art WSs could be achieved using the wearable JINS MEME device, with its concomitant advantages over the conventional gel-based setup.

5. CONCLUSION

This work focused on the comparison of two EOG signal acquisition systems; specifically, a gold-standard, gel-based, wired conventional setup and the JINS MEME EOG glasses, which is a relatively cheaper, wireless, sleek, gel-free, wearable alternative and which has been recently made commercially available. When these two systems were used for gaze displacement estimation, an estimation error of $1.32 \pm 0.26^\circ$ and $1.67 \pm 0.26^\circ$ in the horizontal and vertical directions respectively was obtained using the conventional setup, whereas the corresponding errors using MEME acquired EOG data were $1.97 \pm 0.34^\circ$ and $1.85 \pm 0.30^\circ$ respectively. Furthermore, both saccades and blinks were found to be reliably detected and labelled using the two EOG recording modalities, with labelling accuracies higher than 99% using both systems. When the conventional setup and the wireless MEME glasses were used to interface with an asynchronous EOG-based QWERTY virtual keyboard, comparable WSs of 11.9 ± 4.4 cpm and 9.9 ± 3.6 cpm respectively were achieved, results which have been shown to improve substantially with user experience with the application. In view of these results, it has been shown that the relatively cheaper, wireless, gel-free, sleek and wearable JINS MEME device offers a practical alternative EOG recording modality when compared to the standard, gel-based, wired conventional system, which is feasible for use in eye movement-based HCI systems.

6. ACKNOWLEDGEMENTS

The research work disclosed in this publication is funded by the ENDEAVOUR Scholarship Scheme (Malta). The scholarship may be part-financed by the European Union - European Social Fund (ESF) under Operational Programme II - Cohesion Policy 2014-2020, "Investing in human capital to create more opportunities and promote the well-being of

society". The authors would also like to thank JINS Company Limited for their support.

7. REFERENCES

- Acuña, O. V., Aqueveque, P., & Pino, E. J. (2014) Eye-tracking capabilities of low-cost EOG system. 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, 26-30 Aug. 2014, 610-613. IEEE, Piscataway, NJ.
- Barbara, N., & Camilleri, T. A. (2016) Interfacing with a speller using EOG glasses. IEEE International Conference on Systems, Man, and Cybernetics, Budapest, 9-12 Oct. 2016, 1069-1074. IEEE, Piscataway, NJ.
- Barbara, N., Camilleri, T. A. & Camilleri, K. P. (2018) EOG-Based Eye Movement Detection and Gaze Estimation for an Asynchronous Virtual Keyboard. Biomedical Signal Processing and Control (Under Review).
- Bishop, C. M. (2006) Pattern Recognition and Machine Learning. Springer, New York, NY.
- Bulling, A., Roggen, D., & Tröster, G. (2008) It's in Your Eyes: Towards Context-Awareness and Mobile HCI Using Wearable EOG Goggles. Proceedings of the 10th international conference on Ubiquitous computing, Seoul, 21-24 Sept. 2008, 84-93. ACM, New York, NY.
- Bulling, A., Roggen, D., & Tröster, G. (2009) Wearable EOG Goggles: Seamless Sensing and Context-awareness in Everyday Environments. Journal of Ambient Intelligence and Smart Environments, 1 (2), 157-171.
- Heide, W., Koenig, E., Trillenber, P., Kömpf, D., & Zee, D. S. (1999) Electrooculography: technical standards and applications. Electroencephalography and clinical neurophysiology. Supplement, 52, 223-240.
- Kanoh, S., Ichi-nohe, S., Shioya, S., Inoue, K., & Kawashima, R. (2015) Development of an eyewear to measure eye and body movements. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, 25-29 Aug. 2015, 2267-2270. IEEE, Piscataway, NJ.
- Nathan, D. S., Vinod, A. P., & Thomas, K. P. (2012) An electrooculogram based assistive communication system with improved speed and accuracy using multi-directional eye movements. 35th International Conference on Telecommunications and Signal Processing, Prague, 3-4 July 2012, 554-558. IEEE, Piscataway, NJ.
- Tangsuksant, W., Aekmunkhongpaisal, C., Cambua, P., Charoenpong, T., & Chanwimalueang, T. (2012) Directional eye movement detection system for virtual keyboard controller. 5th Biomedical Engineering International Conference, Ubon Ratchathani, 5-7 Dec. 2012, 1-5. IEEE, Piscataway, NJ.
- Usakli, A. B., & Gurkan, S. (2010) Design of a novel efficient human-computer interface: An electrooculogram based virtual keyboard. IEEE Transactions on Instrumentation and Measurement, 59 (8), 2099-2108.
- Vehkaoja, A. T., Verho, J. A., Puurtinen, M. M., Nöjd, N. M., Leikkala, J. O., & Hyttinen, J. A. (2005) Wireless Head Cap for EOG and Facial EMG Measurements. Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society, Shanghai, 1-4 Sept. 2005, 5865-5868. IEEE, Piscataway, NJ.
- Yamagishi, K., Hori, J., & Miyakawa, M. (2006) Development of EOG-Based Communication System Controlled by Eight-Directional Eye Movements. International Conference of the IEEE Engineering in Medicine and Biology Society, New York, 30 Aug.-3 Sept. 2006, 2574-2577. IEEE, Piscataway, NJ.