EEG SIGNAL PROCESSING FOR EYE TRACKING

Mohammad Reza Haji Samadi, Neil Cooke

Interactive Systems Engineering Research Group, University of Birmingham, U.K.

ABSTRACT

Head-mounted Video-Oculography (VOG) eye tracking is visually intrusive due to a camera in the peripheral view. Electrooculography (EOG) eye tracking is socially intrusive because of face-mounted electrodes. In this work we explore Electroencephalography (EEG) eye tracking from less intrusive wireless cap scalp-based electrodes. Classification algorithms to detect eye movement and the focus of foveal attention are proposed and evaluated on data from a matched dataset of VOG and 16-channel EEG. The algorithms utilise EOG artefacts and the brain's steady state visually evoked potential (SSVEP) response while viewing flickering stimulus. We demonstrate improved performance by extracting features from source signals estimated by Independent Component Analysis (ICA) rather than the traditional band-pass preprocessed EEG channels. The work envisages eye tracking technologies that utilise non-facially intrusive EEG brain sensing via wireless dry contact scalp based electrodes.

Index Terms— ICA, SSVEP, VOG, eye tracking, visual attention

1. INTRODUCTION

The prevalent method for eye tracking is a camera pointed at the eye - Video-Oculuography (VOG). Eye image capture is inherently noisy due to the sensor technology and the unstable physical space between the camera and eye from occlusion, movement and varying light conditions; state-of-the-art head mounted cameras have $0.5°$ visual error and remote systems $2°$ visual error [1].

Brain signal processing via Electroencephalography (EEG) provides two additional eye movement information sources- Oculomuscular movements via Electrooculography (EOG), and the brain's response to flickering stimuli - the Steady-State Visual Evoked Potential (SSVEP) [2]. In this work, these two information sources are combined to develop an EEG-based eye tracking solution.

Section 2 reviews eye tracking methods and EEG. Section 3 describes the proposed EEG signal processing method. Section 4 reports the evaluation including details of the EEG/VOG dataset collected specifically for this study. Section 5 discusses future work.

2. BACKGROUND

2.1. Eye Movement and Tracking

The eye tracking problem can be defined as estimating a person's focus of foveal attention [3, 4, 5] and eye movement types such as saccades (large and rapid movements), smooth pursuits (slow movements), fixation (little movement) and blinking.

Eye tracking techniques can be classed as either Video-oculography (VOG) or non-optical. Electrooculography (EOG) [6] is a non-optical method for eye movement sensing based on changes measured in the steady corneoretinal potential difference. The potential changes are measured with a set of electrodes mounted around the eyes. If the eye moves from its central position, the positive pole of the eye (cornea) moves to one of the electrodes while the negative pole (retina) approaches the opposite electrode. The dipole orientation change effects the EOG signal amplitude. Although EOG can estimate $2°$ visual angle [5], it can be cumbersome and uncomfortable, normally limited to laboratory experiments and they be employed in daily life and mobile circumstances.

VOG records an eye image sequence with high resolution cameras. They rely on the software processing system for obtaining the eye position [5, 7]. Cameras are either head mounted or remote. VOG is more comfortable and less intrusive than EOG. However, remote eye trackers limit the visual field and mobility. Head-mounted trackers require users to wear devices which potentially limit physical activity.

2.2. Electroencephalography

The variation of the surface potential distribution over the scalp reflects the brain's functional activities which are recorded by placing a set of electrodes (sensors) on the scalp and measuring the voltage differences between electrode pairs. The resulting data is called the Electroencephalography (EEG) [8, 9].

EEG signals are contaminated by artefacts rendering them less usable. There are two categories of artefact- environmental and biological [10]. Environmental artefacts originate from outside of the body such as $50Hz$ or $60Hz$ power line noise. Biological artefacts are electrical activities originate from non-cerebral origins such as muscle artefacts (facial and corporeal) and EOG.

In Neuroscience and Brain-Computer Interface (BCI) studies, the EOG signal is an unwanted source which is normally removed to reveal 'clean' EEG containing neural activity only. Averaging, filtering and linear regression methods are commonly used to reject EOG but they typically either under compensate leaving residual artefacts, or over compensate removing neural information from the signal.

Improvement in EEG artefact rejection is demonstrated with the Blind Source Separation (BSS) method Independent Component Analysis (ICA) [11] which assumes that the observed EEG signals from electrode channels are a linear mixture of independent source signals (components). The EEG signal channels are decomposed into multiple components with a time-invariant invertible mixing matrix which is estimated by criterion to identify different sources from signal characteristics such as entropy or mutual information.

2.3. Steady-State Visual Evoked Potential

The Steady-State Visual Evoked Potential (SSVEP) is an electrical response in the brain when the retina is stimulated by a flickering visual stimulus at a frequency higher than $6Hz$ [12]. SSVEP responses are an oscillation in the EEG signals at the same frequency as the visual stimulation. SSVEP can be detected reliably with minimal between-person variation by analysing the EEG signal's spectral content. SSVEP has become a popular technique for BCI [13, 14] and enables a person's focus of attention (FoA) to be estimated, providing the focus/stimulus is flickering. Thus, augmenting the presentation of a focus with a flickering luminosity enables detection of the focus of foveal visual attention without the use of VOG.

2.4. Novelty

Most studies in EEG are motivated by neuroscience and BCI rather than eye tracking; eye movement is seen as an unwanted EOG artefact. They normally use laboratory-grade EEG devices with a high number of channels (e.g. 64), each channel sensed by a wet-contact electrode. Scalp-based electrodes are normally complemented by others on the face and chest to assist with artefact rejection. Consequently, such systems are costly and not suited for real-world wearable applications. This study captures data specifically for eye tracking with a wearable consumer device with a lower number (16) of wireless dry contact scalp-only electrodes.

3. EEG EYE TRACKING SIGNAL PROCESSING

There are two types of eye tracking information that can be extracted from EEG channels - eye movement type from EOG (e.g. saccades, fixations and pursuit) and eye position in relation to the visual field from the SSVEP response detection and apriori knowledge of the positions of stimuli in the visual field. The following signal processing stages are proposed:

Pre-processing: EEG channels are band-pass filtered (1 –$40Hz$) to remove the slow drifts and high-frequency noise (e.g. mains hum)

Feature Extraction: Features are extracted from the pre-processed EEG channels. For eye movement type classification (EOG-based) each channel is epoched into $400ms$ temporal windows with 50% overlap and two features extracted per channel - the minimum and maximum amplitude. For SSVEP response estimation, each channel is epoched into $2000ms$ windows with 60% overlap and 2 features per stimulus are extracted per channel - the spectral power at the stimuli frequency (typically either $7Hz$, $10Hz$ or $12Hz$) and its second harmonic. The Power Spectral Density is estimated with Welch's method [15]. The extracted features from all channels are concatenated at the end.

Independent Component Analysis: Source component signals are extracted from channels with the Blind Source Separation (BSS) algorithm Extended-Infomax [16] Independent Component Analysis (ICA). The same features are extracted from the component signals as from the pre-processed EEG channels.

Classification: Classifiers (reported in this paper k-Nearest Neighbour (kNN) with Euclidean distance metric) are trained for eye movement type (4 classes - blink, fixation, saccade and pursuit) and to detect SSVEP (4 classes - $0Hz$, $7Hz$, $10Hz$ and $12Hz$). The value k is set to 1 (i.e. 1-NN). Separate classifiers for features extracted from EEG channels and ICA components are trained and compared in the evaluation.

4. EVALUATION

To evaluate the EEG-based Eye Tracker, 5 users' EEG and VOG is recorded under controlled cued tasks tracking stimuli which requires fixation, saccadic and pursuit eye movements.

Participants: The 5 healthy users have normal or corrected-to-normal vision and no neurological disorders. They sit in a comfortable chair $1m$ from a $19''$ computer monitor. They are instructed to relax and retain as placid as possible, to avoid the EEG channel contamination with electrical interference from muscles.

Apparatus: Both EEG and VOG apparatus are non-intrusive and designed for mobile (wireless) situations outside of laboratory settings. EEG signals are recorded from the 16 electrode wireless Emotiv EPOC EEG headset which has a $128Hz$ sample rate. This device has credible albeit lower performance than the higher-channel laboratory-based devices [17] although performance degrades over time [18]. VOG is captured from Tobii Eye Glasses, a head-mounted monocular eye tracker which samples at $30Hz$ and records visual angles up to $56°$ horizontally and $40°$ vertically. EEG and VOG data are synchronised by inserting timestamps in the VOG data.

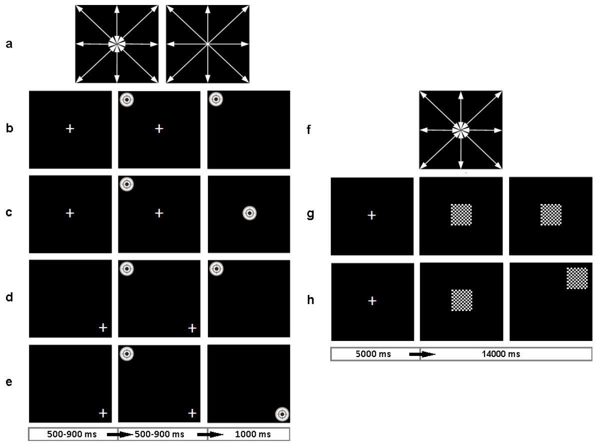

Fig. 1. Display to user over time (left to right). Row (a) Possible stimuli locations; (b) Elicit short saccade; (c) Elicit fixation in the display centre; (d) Elicit a long saccade; (e) Elicit fixation at the display edge; (f) Possible stimuli locations (g) Static SSVEP stimulation for fixations where stimulus is fixed at the screen centre; (h) SSVEP stimulation for pursuit where stimulus moves from the screen centre to the sides.

The $30Hz$ VOG data is up-sampled to match the $128Hz$ EEG data by linear interpolation; this approximates the saccades' trajectory in VOG.

4.1. Sessions recorded

Fixation, Saccade and Pursuit direction detection from EOG: Fixations and horizontal, vertical and diagonal saccades are elicited by visual cue. The EEG collected is used to detect and classify pursuit and saccade direction from the EOG features as outlined in section 3. Refer to Fig. 1, users are presented with a black square display and there are 9 possible stimulus locations around the screen edge. Sessions start with a white fixation cross appearing. To elicit short saccades, the cross appears in the centre. To elicit long saccades, the cross appears at the edge. After a random time between $500ms$ and $900ms$, the stimulus is presented at one of the nine edge locations. The cross disappears at a random interval $500ms$ to $900ms$ after stimulus presentation which provides the cue for users to make a saccade to the stimulus location. Stimuli are presented for a further $1s$ after the cue. Each user performs 112 cued fixations and 112 cued saccades which lasts approximately 11 minutes.

Focus of foveal Attention (FoA) during fixations and pursuit movements: Fixations and smooth pursuits are elicited by flickering stimuli. The EEG collected is used to detect and classify FoA fixations and pursuits direction from features outlined in section 3. Similar to the previous session, refer to Fig. 1, the screen is a black square. Again, sessions start with a white fixation cross appearing. After $5s$ the stimulus - a 10x10 checkerboard - appears flickering at one of three different frequencies suitable for SSVEP detection ($7Hz$, $10Hz$ or $12Hz$ [19]) for another $14s$. To elicit fixations the checkerboard remains still. To elicit pursuit it moves in one of 8 directions towards the screen edge. Each user performs 8 fixations and 8 pursuit movements for each flickering frequency (total 48 movements over approximately 11 minutes.

4.2. Classifiers

Eight kNN Classifier pairs (labelled $C1 - 8$ in the result Table 1) are trained on the EEG session data following the process outline in section 3. Each pair consists one classifier trained on EEG pre-processed channels and one classifier trained on source components estimated by ICA. $C1$ and $C2$ classify long and short saccade directions respectively from EOG features (8 class problem). $C3$ and $C4$ classify FoA fixations and pursuit movement respectively from SSVEP features (2 class problem). $C5$ and $C6$ repeat this but for SSVEP detection for a specific flickering frequency (4 class problem assuming 3 potential FoA flickering at different frequencies). $C7$ and $C8$ classifiers classify pursuit direction with and without SSVEP present (2 class problem).

The saccade and pursuit direction detection performance is assessed by the Root Mean Square Error (RMSE) of the difference between the detected angle and the real angle associated with the stimuli's position:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\theta_t - \hat{\theta}_t)^2} \tag{1}$$

,where N is the time length, and θ_t, $\hat{\theta}_t$ are the real and estimated angular stimuli position at time t, respectively.

4.3. Results

Table 1 summarises the classification performance results obtained through different classifiers ($C1$-$C2$). The classifiers' accuracy and RMSE before and after applying ICA are reported. As its evident from the results, the all classifiers' accuracy is improved and the between-person variation decreased significantly if features are extracted from ICA components (Fig. 2).

For $C1$, short saccade direction classification accuracy is almost 74% with a 14% improvement if features are extracted from ICA components. $C2$ benefits about 2% more from applying ICA than $C1$, result in 89% accuracy for long saccade direction classification. Considering RMSE, similar trend is obtained when comparing $C1$ and $C2$ (Fig. 2).

For $C3$ and $C4$, FoA estimation (i.e. SSVEP detection during fixations and pursuits respectively) accuracy again is shown to be improved (e.g. $C3$ by 6% to 97.90%) with ICA component features. Similarity between results suggests that the SSVEP response during pursuit movements is no harder

Table 1. Eye Movement Classification Results.

Classifer	Description	RMSE (°)		Accuracy (%)	
		Before ICA	After ICA	Before ICA	After ICA
$C1$	Short saccade in one of 8 directions	19.02	9.02	73.84	87.58
$C2$	Long saccade in one of 8 directions	19.30	7.81	73.45	89.25
$C3$	Fixation from SSVEP detection (all frequencies) (2 class)	N/A	N/A	91.38	97.90
$C4$	Pursuit from SSVEP detection (all frequencies)(2 class)	N/A	N/A	92.27	97.79
$C5$	Fixation-FoA from SSVEP detection ($7Hz$, $10Hz$,$12Hz$)(4 class)	N/A	N/A	83.87	96.38
$C6$	Pursuit-FoA from SSVEP detection ($7Hz$, $10Hz$,$12Hz$) (4 class)	N/A	N/A	86.59	96.05
$C7$	Pursuit direction (no SSVEP) (8 class)	11.51	1.79	81.64	97.34
$C8$	Pursuit direction (SSVEP present) (8 class)	12.44	**1.63**	82.88	97.57

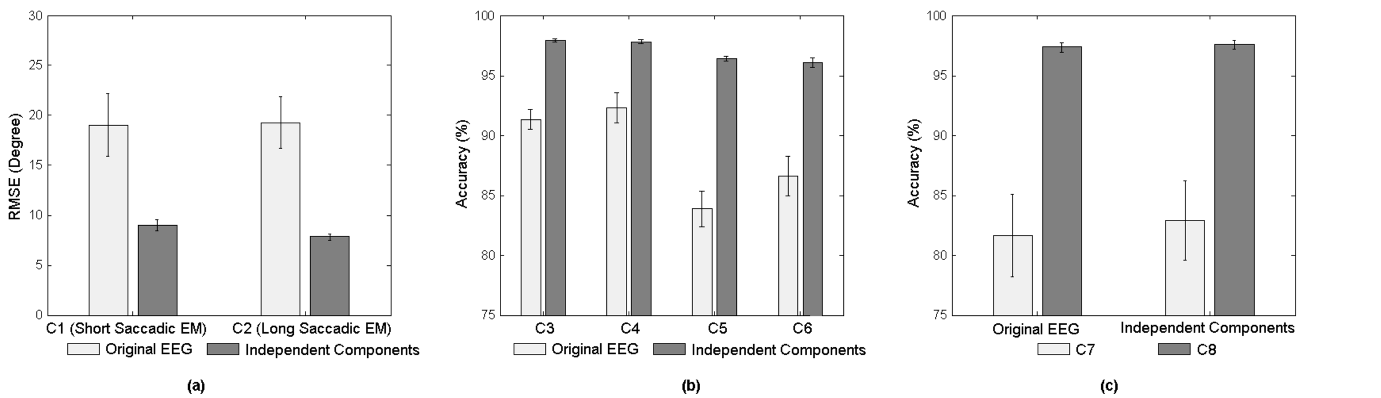

Fig. 2. , (a) RMSE for saccadic eye movement direction detection; $C1$: Short saccade; $C2$: Long saccades, grey bars represent RMSE when features are extracted from original EEG channels, black bars represents RMSE when features are extracted from ICA estimated sources; (b) SSVEP frequency classification accuracy;(c) Pursuit direction detection; $C7$: without SSVEP, $C8$: with SSVEP. The error bars represent standard errors.

to detect than fixations despite the potential for more EEG signal contamination due to EOG.

For $C5$ and $C6$, FoA estimation for SSVEP given 3 potential stimuli flickering has 96% accuracy for pursuit movements and fixations with features extracted from ICA components.

Finally, results from $C7$ and $C8$ demonstrate that the presence of SSVEP in the visual field does not hinder the classification of pursuit direction with both classifiers show similar results.

Overall, eye movement direction classification accuracy is best in $C8$ (i.e. 97.57%) with the least RMSE (i.e. $1.63°$) where there is pursuit eye movement rather than saccadic eye movement.

5. DISCUSSION

A signal processing scheme for extracting eye tracking data from wireless portable scalp-based EEG has been proposed and evaluated. Feature extraction from EEG source components estimated by ICA results in better performance compared to feature extraction from pre-processed EEG channels; notably between-person variation reduces. This work is the first we are aware that uses source components for SSVEP classification rather than EEG channels. Future work will refine the eye movement type classification in the presence of other EEG artefacts such as body movement. The VOG data captured in this study affords the opportunity for hybrid VOG/EEG multimodal eye tracking. The modulation of real-scenes with flickering stimuli via augmented reality technologies to estimate focus of foveal visual attention will potentially lead to VOG-free eye tracking, without the social and physical intrusion of face mounted electrodes.

6. REFERENCES

[1] Dan Witzner Hansen and Qiang Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *Pattern*

Analysis and Machine Intelligence, IEEE Transactions on, vol. 32, no. 3, pp. 478–500, 2010.

[2] Matthew Middendorf, Grant McMillan, Gloria Calhoun, and Keith S Jones, “Brain-computer interfaces based on the steady-state visual-evoked response,” *Rehabilitation Engineering, IEEE Transactions on*, vol. 8, no. 2, pp. 211–214, 2000.

[3] Kyung-Nam Kim and RS Ramakrishna, “Vision-based eye-gaze tracking for human computer interface,” in *Systems, Man, and Cybernetics, 1999. IEEE SMC’99 Conference Proceedings. 1999 IEEE International Conference on*. IEEE, 1999, vol. 2, pp. 324–329.

[4] Kevin Smith, Sileye O Ba, Daniel Gatica-Perez, and Jean-Marc Odobez, “Tracking the multi person wandering visual focus of attention,” in *Proceedings of the 8th international conference on Multimodal interfaces*. ACM, 2006, pp. 265–272.

[5] Carlos H. Morimoto and Marcio R. M. Mimica, “Eye gaze tracking techniques for interactive applications,” *Comput. Vis. Image Underst.*, vol. 98, no. 1, pp. 4–24, Apr. 2005.

[6] Ali Bülent Usakli, Serkan Gurkan, Fabio Aloise, Giovanni Vecchiato, and Fabio Babiloni, “On the use of electrooculogram for efficient human computer interfaces,” *Computational Intelligence and Neuroscience*, vol. 2010, pp. 1, 2010.

[7] Arantxa Villanueva and Rafael Cabeza, “Models for gaze tracking systems,” *Journal on Image and Video Processing*, vol. 2007, no. 3, pp. 4, 2007.

[8] Hans Berger, “Über das elektrenkephalogramm des menschen,” *European Archives of Psychiatry and Clinical Neuroscience*, vol. 87, no. 1, pp. 527–570, 1929.

[9] Saeid Sanei and Jonathon A Chambers, *EEG signal processing*, Wiley. com, 2008.

[10] John N Demos, *Getting started with neurofeedback*, WW Norton & Company, 2005.

[11] Ricardo Nuno Vigário, “Extraction of ocular artefacts from eeg using independent component analysis,” *Electroencephalography and clinical neurophysiology*, vol. 103, no. 3, pp. 395–404, 1997.

[12] Zhonglin Lin, Changshui Zhang, Wei Wu, and Xiaorong Gao, “Frequency recognition based on canonical correlation analysis for ssvep-based bcis,” *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 12, pp. 2610–2614, 2006.

[13] Leonard J Trejo, Roman Rosipal, and Bryan Matthews, “Brain-computer interfaces for 1-d and 2-d cursor control: designs using volitional control of the eeg spectrum or steady-state visual evoked potentials,” *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 14, no. 2, pp. 225–229, 2006.

[14] Po-Lei Lee, Jyun-Jie Sie, Yu-Ju Liu, Chi-Hsun Wu, Ming-Huan Lee, Chih-Hung Shu, Po-Hung Li, Chia-Wei Sun, and Kuo-Kai Shyu, “An ssvep-actuated brain computer interface using phase-tagged flickering sequences: a cursor system,” *Annals of Biomedical Engineering*, vol. 38, no. 7, pp. 2383–2397, 2010.

[15] Peter Welch, “The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms,” *Audio and Electroacoustics, IEEE Transactions on*, vol. 15, no. 2, pp. 70–73, 1967.

[16] Anthony J Bell and Terrence J Sejnowski, “An information-maximization approach to blind separation and blind deconvolution,” *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[17] Matthieu Duvinage, Thierry Castermans, Mathieu Petieau, Thomas Hoellinger, Guy Cheron, and Thierry Dutoit, “Performance of the emotiv epoc headset for p300-based applications,” *Biomedical engineering online*, vol. 12, no. 1, pp. 56, 2013.

[18] L Mayaud, M Congedo, A Van Laghenhove, M Figère, E Azabou, and F Cheliout-Heraut, “A comparison of recording modalities of p300 event related potentials (erp) for brain-computer interface (bci) paradigm,” *Neurophysiologie Clinique/Clinical Neurophysiology*, 2013.

[19] Danhua Zhu, Jordi Bieger, Gary Garcia Molina, and Ronald M Aarts, “A survey of stimulation methods used in ssvep-based bcis,” *Computational intelligence and neuroscience*, vol. 2010, pp. 1, 2010.