Radiooculogram (ROG) for eye movement sensing with eyes closed

Zijing Zhang* and Edwin C. Kan
School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA.
*E-mail: zz587@cornell.edu

***Abstract*—This paper presents radiooculogram (ROG), a novel sensor for non-invasive eye movement (EM) monitoring with eyes closed. We have experimentally demonstrated accurate measurements of EM frequency and directions for 5 participants and benchmarked ROG with electrooculogram (EOG). Compared with biopotential-based sensors, ROG has higher user comfort due to touchless operation and can capture direct muscle activity even in deep tissues. This work on voluntary EM sensing can serve as the baseline implementation for eventual sleep rapid EM monitoring.**

Keywords—eye movement; muscle monitoring; non-invasive sensing.

I. INTRODUCTION

Eye movement (EM) measurement can derive profuse information in emotion perception [1], neurodegenerative diseases [2], and monitoring of sleep and dream stages [3]. It can also facilitate human-computer interface (HCI) and virtual reality (VR) applications [4][5]. Current eye tracking systems with eyes open by camera-based methods can achieve high accuracy, though still have concerns of privacy, complexity, and occlusions [6][7]. Sensing EM with eyes shut under low ambient light can be even more difficult for cameras. EM sensing with eyes closed during sleep is important for the detection of rapid eye movement (REM), a sleep phase characterized by random rapid EM with an inclination of vivid dreaming. REM as an important sleep stage can be an indicator of health and cognitive performance, such as brain maturation [8], memory consolidation [9], and learning facilitation [10]. Existing methods for REM recording during sleep mainly used biopotential signals from electrooculogram (EOG) [11] and electroencephalography (EEG) [12], as parts of the clinical polysomnography (PSG) [13]. However, the electrode-based sensor can be limited by low user comfort and skin irritation as well as ambiguity and interferences due to skin potentials and leaky neural signals.

Here we propose radiooculogram (ROG), a novel EM sensor based on radio-frequency (RF) signals that can accurately and non-invasively monitor internal eye muscle activities with eyes open or shut. We validated accurate measurement of EM frequencies and directions by a human study of 5 participants with selected longitudinal experiments. We further benchmarked ROG with synchronous EOG as the baseline comparison and physiological correlation. The main advantage of the proposed ROG system can be summarized as:

- Improved user comfort. EOG and EEG measurements demand numerous electrodes around the eye region with stable electrical contact, which are inconvenient, uncomfortable, and prone to face motion interference. ROG can operate without direct skin contact. For example, ROG can be integrated into the eyeglasses frame or implemented as an insert or at an outer layer of a sleep mask or a masquerade. As no skin contact electrode is involved and ROG sensors just need to be at a constant place relative to eyes, both comfort and design freedom are enhanced.
- Unmediated sensing of directional EM. While the biopotential-based sensors such as EOG, EEG, and electromyography (EMG) measure neural stimulus for muscle activity, ROG directly measures the muscle motion by coupling RF energy to deep internal muscles, i.e., muscle condition is the direct output, instead of an estimate derived from the measured stimulus. EOG and ROG can also be used together to retrieve the closed loop of stimulation and actuation.
- Baseline for sleep REM detection. While camera-based methods are difficult to use for sleep REM, ROG has the flexibility to operate when eyes are open or closed without privacy concern. This work on conscious participants performing voluntary EM can formulate a validation baseline for future sleep REM monitoring.

II. SENSOR SETUP AND EXPERIEMENT PROTOCOL

A. Sensor Setup

ROG is based on the near-field coherent sensing (NCS) [14][15] of ultra-high frequency (UHF) RF signals to monitor the dielectric boundary change of internal muscles during EM. Electromagnetic energy was coupled deep into the body by the near-field effect, and tissue motion was modulated on the channel characteristics. As shown in Fig. 1(a), one ROG sensing unit consisted of a notched miniature coaxial RF cable, where the metal shield of the 1-inch middle part was removed to allow a small amount of RF energy leaking into the user's upper face region [16]. The ROG system integrated four sensing units attached to an eye mask around the eyes, as indicated in Fig. 1(b). The ROG RF transceiver was implemented by software-defined radios (SDR) to drive the notched sensors and to interface with the host computer through USB. Two National Instrument Ettus B210 were used, each of which had two transmitter/receiver (Tx/Rx) ports as shown in Fig. 1(c). The two SDRs were synchronized by an external local oscillator (LO, BG7TBL-GPSDO) with 10 MHz reference and 1 PPS (pulse per second) baseband synchronization. The experimental setup on a user's face was shown in Fig. 1(d) for ROG and in Fig. 1(e) for EOG. The ROG system can be alternatively implemented in

wireless active [17] and passive [14] units, although the present prototype is a wired system for convenient benchmarking.

In the near-filed region, the dielectric boundary change of associated eye muscles during EM would couple into the leaked RF energy, and hence affected the signals between Tx and Rx. Four sensors at different positions around the eyes provided more observation diversity to improve the amplitude and direction resolution. We adopted the multiple-input multiple-output (MIMO) strategy to explore $N^2 = 16$ coupling channels from $N = 4$ sensing units to further enhance the spatial diversity [18].

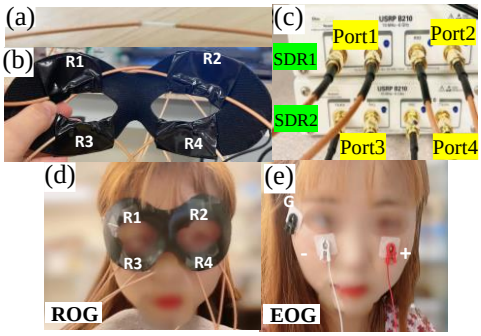

Fig. 1. The ROG system. (a) One ROG sensor unit by a notched transmission line; (b) Four ROG sensor units on a mask; (c) The SDR transceiver; (d) ROG on a participant's face; (e) EOG setup for baseline comparison.

The digital baseband of each Tx went through the digital-to-analog converter (DAC) and was then mixed with the carrier frequency f_{RF}, selected at 1 GHz. The RF power was less than −10 dBm or 0.1 mW, well under the safety limits set by Occupational Safety and Health Administration (OSHA). The RF signal leaked from the notched structure is coupled into internal muscle motion, received by Rx, and then demodulated and sampled by the analog-to-digital converter (ADC) to retrieve the baseband. The quadrature scheme was employed as the baseband tone f_{BB}. The multiple Tx channels utilized frequency-division multiple access (FDMA) by setting f_{BB} =10, 25, 40, and 125 kHz, respectively, for Tx1−Tx4. The system was configured as 4 self and 12 cross channels, which were all sampled at 10^6 samples per second (Sps), and further down-sampled to 500 Sps after demodulation.

B. Human Study Protocol

Two routines of human study on 5 volunteers were executed when eyes were closed. The ROG signals were similar with open eyes, but with interference from blinking. Blinking can be separately monitored by ROG as well, but not presented in this paper. Routine 1 was for EM frequency detection when the participant followed voice instructions and exercised EM with 10, 15, 20, 30, and 60 beats per minute (BPM). The eye exercise in each frequency had a duration of 30s with eyes moving left and right. Figs. 2(a)-(b) presented several examples of ROG (Tx3 – Rx3) and EOG waveforms. Participants were then instructed to move eyes in four directions in Routine 2. All directions had 2 versions of moving once and twice. Hence, we had 8 distinctive EMs, and each motion was performed in a time window of T_{win} =5s with around 24 repetitions. Unlike gaze localization with open eyes [19], the ground truth of EM direction and voluntary control of eyeball rotation were less precise when eyes were closed.

C. Benchmark with EOG

The reference EOG setup was by BIOPAC MP36R with the three EL513 electrodes around the eyes as + (under right eye), − (under left eye), and ground (left to left eye). ROG and EOG channels were synchronized in Labview and transferred to the host computer by USB. The same study protocol was performed on two participants with longitudinal iterations.

III. Signal Processing

A. EM Frequency Estimation

For the EM frequency testing in Routine 1, the signal was first bandpass-filtered from 0.05Hz to 2Hz to remove the DC drift and high-frequency noises. Then we utilized the moving average-crossing algorithm to first extract a moving-average curve in a given window length, and then label local maximum and minimal points [15]. The EM rate is estimated by counting the number of detected cycles over an epoch of 10s. One EM cycle includes moving eyeballs from left to right and then back to left. We have collected 16 quadrature channels from ROG, and each channel has amplitude and phase separately. We selected the channel with minimum covariance of the EM rate to output the final estimates. EOG was processed in a similar way, although there was only one channel in our setup. Notice that the EOG here was for baseline comparison, and not a state-of-the-art system in the EOG research field.

B. EM Direction Estimation

For the EM direction testing in Routine 2, after obtaining 1D time waveforms from 32 channels, we first applied bandpass filtering (0.05 Hz to 10 Hz) and normalization. The waveforms were then segmented into motion-based windows of T_{win} = 5 s, each containing one instructed EM. We transformed the 1D waveforms to 2D spectrograms using continuous wavelet transform (CWT) by Morelet and Gaussian mother wavelets. Finally, the 2D image-like data was fed into the deep learning network as the classifier to differentiate all EM directions. We adopted vision transformer (ViT) [20], a deep learning model in natural language processing (NLP) and computer vision, for classification. Other machine-learning models such as convoluted neural network (CNN) have been used on the time-domain waveforms as well, but with slightly inferior performance.

IV. Results and Analyses

A. EM frequency estimation

Fig. 2(c) shows the EM rate in BPM calculated from EOG (blue) and ROG (red) in comparison with the ground truth (green) from instruction in Routine 1. Fig. 2(d) shows the correlation of the EM rate from ROG (red markers) and EOG (blue markers) against the ground truth. In the left figure, both ROG and EOG achieved high correlation to the ground truth with Pearson coefficients denoted as $r_{ROG} = 0.99$ and $r_{EOG} = 0.98$. In the right figure, the Bland-Altman plot presents the agreement by the mean (m) and limits of agreement (LoA). The X axis is the average of the estimation and ground truth, and the Y axis is the difference. Both EOG and ROG achieved low m and narrow

The work was supported by NSF ECCS-CCSS 1945918.

LoA. Note that *m* is positive for both sensors, which implies that the ground-truth EM rate is higher than the estimated results. This is hypothesized that the participant may not perfectly follow the instruction especially for very fast EM at 60 BPM. Fig. 2(e) further presents the correlation between ROG and EOG in a similar format. In comparison with Fig. 2(d), correlation between ROG and EOG is higher than that to the ground truth, indicating higher consistency between the two sensors. Table I summarizes all correlation and B&A statistics of the EM rate using only ROG across 5 participants.

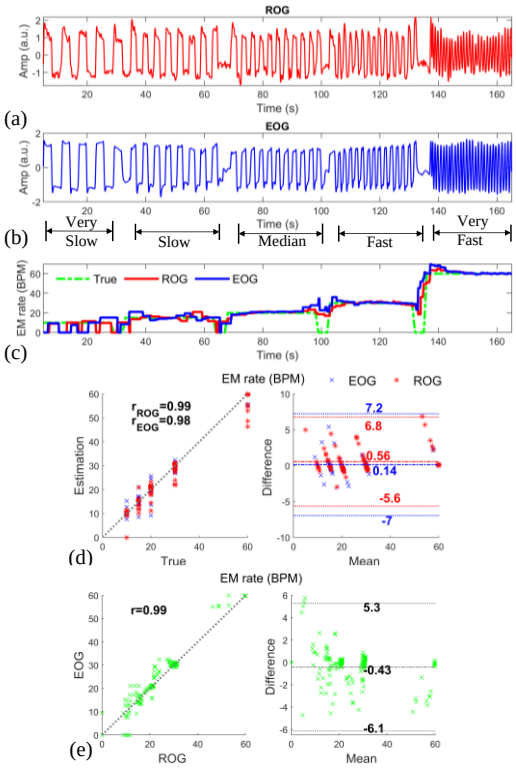

Fig. 2. EM frequency estimation. (a) ROG amplitude from Tx3- Rx3 and (b) EOG waveform samples. (c) EM rate in BPM. Correlation and agreement (d) between the ground truth and estimation from ROG and EOG, and (e) between ROG and EOG. Left: The scatter plot with denoted Pearson's correlation coefficient. Right: The Bland-Altman plot for biases and limits of agreement.

We further compared the temporal correlation between ROG and EOG waveforms. We extracted the optimal time lag that can maximize the cross-correlation between the two waveforms. When the time lag = 0.032 s, cross-correlation achieves the maximum value of 0.96, which indicates that the ROG waveform has a time delay following EOG events, as EOG detects the neural stimulation of EM and ROG detects the actual EM. Table II presents the statistics of time lag and correlation between EOG and ROG in two subjects during longitudinal tests, where $r_{R\&E}$ is the Pearson coefficient. The consistently high accuracy in the longitudinal studies, when the setup was refitted each time, suggests that the mask and sensor placement variation on the face may cause some waveform variations, but would not affect much on the extraction of EM rate and direction. Table III presents the correlation and B&A statistics of EM rate estimation using ROG and EOG in comparison with the ground truth in the longitudinal tests, where both ROG and EOG remain highly accurate.

TABLE I. CORRELATION AND B&A STATISTICS OF THE EM RATE ESTIMATION USING ROG FOR EACH SUBJECT

Subject No.	r_{ROG}	$m \pm \sigma$ (BPM)
1	0.987	0.70 ± 2.90
2	0.986	0.75 ± 3.02
3	0.985	1.04 ± 3.21
4	0.982	1.00 ± 3.11
5	0.984	1.46 ± 3.26
Mean	0.985	0.99 ± 3.10

TABLE II. Time lag and correlation between ROG and EOG

Subject No.	Iteration No.	Time Lag (s)	Max Corr.	$r_{R\&E}$	$m \pm \sigma$ (BPM)
1	2	0.090	0.88	0.99	0.28 ± 2.24
2	2	0.054	0.97	0.95	–1.75 ± 6.02
2	3	0.052	0.92	0.99	0.31 ± 3.00
2	4	0.032	0.96	0.99	–0.43 ± 2.91
Mean		0.057	0.93	0.98	–0.40 ± 3.54

TABLE III. CORRELATION AND B&A STATISTICS OF EM RATE ESTIMATION USING ROG AND EOG AGAINST THE GROUND TRUTH

Subject No.	Iter. No.	r_{ROG}	$m \pm \sigma$ (ROG)	r_{EOG}	$m \pm \sigma$ (EOG)
1	2	0.985	0.11 ± 3.13	0.992	0.39 ± 2.26
2	2	0.979	0.95 ± 4.10	0.950	2.75 ± 5.68
2	3	0.970	1.43 ± 4.50	0.979	1.74 ± 4.07
2	4	0.986	0.56 ± 3.16	0.983	0.14 ± 3.61
Mean		0.980	0.76 ± 3.72	0.976	1.26 ± 3.90

B. EM direction estimation

Routine 2 for EM direction estimation include 5 subjects with 947 samples of 8 classes of EM, namely 4 directions (up, down, right, and left) and 2 instances (once and twice). We built the training model within each participant and performed *k*-fold (k = 7) cross validation to estimate the mean accuracy for each participant. An overall accuracy was averaged on results from all participants. Fig. 3 shows the normalized confusion matrix by ViT. ROG can achieve high accuracy for distinguishing different directions. Note that the class 'D' (down) has relatively lower accuracy than other classes. It may be difficult for participants to follow the instruction to move eyeballs downwards in a consistent way with eyes closed. We also collected data from EOG in the benchmark experiment using the same protocol and signal processing procedures. Accuracy by one EOG channel only achieved 57.3%.

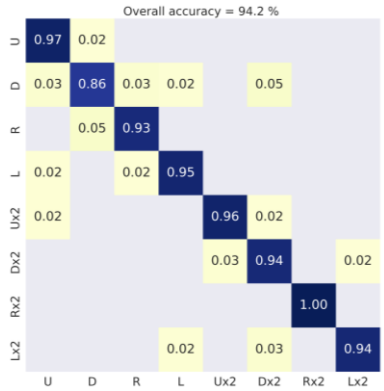

Fig. 3. The confusion matrix showing the overall accuracy of 94.2% for EM direction detection on all 5 subjects.

V. CONCLUSION

In this paper, we present a new non-invasive and touchless radiooculogram (ROG) for EM monitoring with eyes closed.

ROG can accurately detect the EM frequencies in a broad range and recognize different EM directions. In comparison with conventional EOG, ROG has high accuracy and improved user comfort without requiring direct skin contact. ROG can capture direct muscle actuation during EM with less ambiguity and interference. The consistent delay of ROG trailing EOG events indicates the lag of muscle actuation after the neural stimulation. In the future, ROG can be a promising alternative for sleep REM monitoring in clinical studies.

REFERENCES

[1] Y. Wang, Z. Lv, and Y. Zheng, "Automatic emotion perception using eye movement information for E-healthcare systems," *Sensors*, vol. 18, no. 9, p. 2826, 2018.

[2] T. J. Crawford *et al.*, "Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's disease," *Biological Psychiatry*, vol. 57, no. 9, pp. 1052-1060, 2005.

[3] S.-F. Liang *et al.*, "Development of an EOG-based automatic sleep-monitoring eye mask," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 11, pp. 2977-2985, 2015.

[4] S. Z. Homayounfar *et al.*, "Multimodal smart eyewear for longitudinal eye movement tracking," *Matter*, vol. 3, no. 4, pp. 1275-1293, 2020.

[5] A. Poole and L. J. Ball, "Eye tracking in HCI and usability research," in *Encyclopedia of Human Computer Interaction*: IGI Global, 2006, pp. 211-219.

[6] R. A. Naqvi, M. Arsalan, G. Batchuluun, H. S. Yoon, and K. R. Park, "Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor," *Sensors*, vol. 18, no. 2, p. 456, 2018.

[7] N. Valliappan *et al.*, "Accelerating eye movement research via accurate and affordable smartphone eye tracking," *Nature Communications*, vol. 11, no. 1, pp. 1-12, 2020.

[8] M. S. Knoop, E. R. de Groot, and J. Dudink, "Current ideas about the roles of rapid eye movement and non–rapid eye movement sleep in brain development," *Acta Paediatrica*, vol. 110, no. 1, pp. 36-44, 2021.

[9] S. Diekelmann and J. Born, "The memory function of sleep," *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 114-126, 2010.

[10] M. Tamaki *et al.*, "Complementary contributions of non-REM and REM sleep to visual learning," *Nature Neuroscience*, vol. 23, no. 9, pp. 1150-1156, 2020.

[11] A. Boukadoum and P. Ktonas, "EOG-Based Recording and Automated Detection of Sleep Rapid Eye Movements: A Critical Review, and Some Recommendations," *Psychophysiology*, vol. 23, no. 5, pp. 598-611, 1986.

[12] M. A. Cruz-Aguilar, I. Ramírez-Salado, M. Hernández-González, M. A. Guevara, and J. M. Del Río, "Melatonin effects on EEG activity during non-rapid eye movement sleep in mild-to-moderate Alzheimer´ s disease: a pilot study," *International Journal of Neuroscience*, vol. 131, no. 6, pp. 580-590, 2021.

[13] N. Cooray, F. Andreotti, C. Lo, M. Symmonds, M. T. Hu, and M. De Vos, "Detection of REM sleep behaviour disorder by automated polysomnography analysis," *Clinical Neurophysiology*, vol. 130, no. 4, pp. 505-514, 2019.

[14] X. Hui and E. C. Kan, "Monitoring vital signs over multiplexed radio by near-field coherent sensing," *Nat. Electron.*, vol. 1, no. 1, pp. 74-78, 2018,

[15] P. Sharma, X. Hui, J. Zhou, T. B. Conroy, and E. C. Kan, "Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate," *NPJ Digital Medicine*, vol. 3, no. 1, pp. 1-10, 2020.

[16] Z. Zhang, P. Sharma, J. Zhou, X. Hui, and E. C. Kan, "Furniture-integrated respiration sensors by notched transmission lines," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 5303-5311, 2020.

[17] X. Hui, J. Zhou, P. Sharma, T. B. Conroy, Z. Zhang and E. C. Kan, "Wearable RF near-field cough monitoring by frequency-time deep learning", *IEEE Trans. Biomed. Circuits & Sys*, vol. 15, no. 4, pp. 756 – 764, 2021

[18] X. Hui, T. B. Conroy, and E. C. Kan, "Multi-point near-field RF sensing of blood pressures and heartbeat dynamics," *IEEE Access*, vol. 8, pp. 89935-89945, 2020.

[19] Y. Wang, X. Ding, G. Yuan, and X. Fu, "Dual-cameras-based driver's eye gaze tracking system with nonlinear gaze point refinement," *Sensors*, vol. 22, no. 6, p. 2326, 2022.

[20] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.