3-2023

Wearables for in-situ monitoring of cognitive states: Challenges and opportunities

Meera RADHAKRISHNAN

Thivya KANDAPPU

Manoj GULATI

Archan MISRA
Singapore Management University, archanm@smu.edu.sg

Wearables for In-Situ Monitoring of Cognitive States: Challenges and Opportunities

Meera Radhakrishnan[1*], Thivya Kandappu[2], Manoj Gulati[2], Archan Misra[2]
[1]Institute of High Performance Computing, Agency for Science, Technology and Research,
[2]School of Computing and Information Systems, Singapore Management University
radham@ihpc.a-star.edu.sg,{thivyak, manojg, archanm}@smu.edu.sg

***Abstract*—We propose using wrist and ear-based sensing, via multiple novel and *complementary* modalities, to unobtrusively infer activity-aware, complex cognitive and affective states (such as confusion, boredom, and recall failure) of individuals. While state-of-the-art wearable devices are predominantly used (a) independently, with limited coordination among multiple devices, and (b) to capture macro-level physical activity and physiological state, we seek to expand the ambit of unobtrusive wearable sensing to capture the cognitive states while performing commonplace physical activities. Such states typically manifest via fine-grained, almost unobservable, *microscopic* head, face, and eye movements. We identify some of these fine-grained physical markers that serve as proxies for cognitive/affective states and show that earable-mounted pressure, EMG, and ultrasonic sensing hold promise for capturing such markers.**

***Index Terms*—Earables, Multi-device sensing, Cognitive states, Acoustic sensing, Electromyography, Pressure**

I. Introduction

The recent proliferation of wearable devices, with convenient form factors, presents unique opportunities for personal health monitoring. Smart wearables, such as wristbands and earable devices, can help capture a variety of kinetic human activity (e.g., [13], [20]) and physiological context (e.g., [21]). State-of-the-art sensing techniques typically utilize a single device at a time to distinguish an array of daily physical activities. In this paper, we explore the possibility of expanding the ambit of consumer wearable devices to unobtrusively capture various facets of *human cognitive functioning*, such as attention, boredom, and confusion, while engaged in regular daily lifestyle activities. If possible, the ability to utilize multiple wearables to collectively sense different modalities of human signals and biomarkers can spawn a wide range of novel and pervasive applications, such as (a) pervasive monitoring of an elderly individual's rate of cognitive decline while performing tasks such as cooking or cleaning, and (b) in-situ, real-time sensing of loss of learning efficacy during an online class. For example, as illustrated in Figure 1, an exemplar online application called *MetaTutor* can use *smart watches* to detect an array of hand gesture activities (such as scrolling, browsing and typing), which in turn opportunistically trigger *earables* to detect negative affect states (such as boredom and cognitive confusion) of students in real-time during an online learning session. Such detection can then be used to subsequently provide suitable corrective intervention (e.g., reminders, modified content).

*This work was done while the author was affiliated with Singapore Management University.

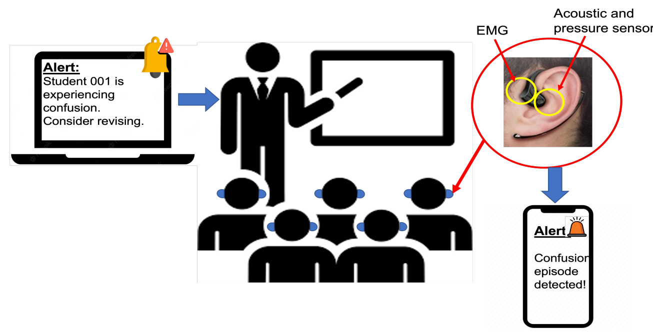


Figure 1: MetaTutor: An exemplar application

Several prior works in the literature (e.g., [17], [22], [10]) have established the use of different sensors on wrist-worn devices, such as smartwatches, for accurate recognition of various hand gestures. As such in this work, we principally explore the use of multiple often-novel, earable-embedded sensing modalities, including *electromyography (EMG), accelerometer, ultrasound, and barometric pressure* to capture subliminal head, facial, and eye movement cues that serve as *micro-markers* of cognitive state. We report on early experimental studies, that reveal successes and failures in the use of these sensing modalities, as well as identify additional practical challenges in the use of such sensors for real-world, *continual* sensing. Overall, we make the following **key contributions**:

- We propose a *multi-device, "wristable+ earable" sensing paradigm*, that combines sensing of gestural actions with sensing of subliminal head, face, and eye movements to infer cognitive states of individuals during daily activities. Besides offering complementary capabilities, their orchestration helps reduce the sensing energy overhead on the more battery-constrained earable device.
- We investigate various combinations of {sensing modality, in/near-ear placement}, such as EMG, accelerometer, barometric pressure, and ultrasound, placed in/around the ear, to determine their ability to capture the minute physical markers that are correlated with human cognitive and affective states. Through experimental studies, we confirm that common sensing techniques (such as inertial sensing) fail to recognize various core biomarkers e.g., pupil movements).

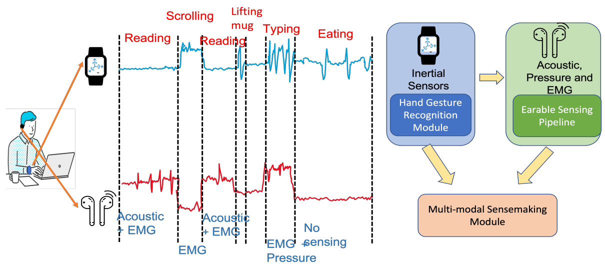


Figure 2: Multi-Device, Multi-Modal Sensing Paradigm

- We demonstrate that even minute contractions of the facial muscles, generated by subtle facial actions such as moving eye gaze, cause deformations in the ear canal *shape*. Such deformations generate discernible patterns in both (a) the inside-ear air pressure sensed by an in-ear differential barometric sensor and (b) the frequency response generated by an active in-ear ultrasonic transmitter.
- We demonstrate, using an initial 3-subject study, that even an unobtrusively and remotely placed EMG sensor (electrodes affixed in front and behind the ear) can *indirectly* capture 6 different typical facial movements with 80+% classification accuracy. Our EMG results differ from past work [9], where the electrodes are placed obtrusively (attached to different muscle sites on the face) to directly measure muscle activity. In addition, we identify the challenge of "muscle hyperpolarization" [14], which emerges when activities are performed continuously and naturally and which has been neglected in the design of past activity detectors.

II. Overall Multi-Device Paradigm

We first outline our vision of combining wrist-based and earable sensing to capture an individual's cognitive states during various daily activities. Figure 2 illustrates such a combined sensing paradigm. Inertial sensors (e.g., accelerometer and gyroscope) on the wrist-worn device first help detect micro-activities and gestures, such as performing housekeeping activities [12] or typing in a password/PIN on a mobile App [23]. Triggered by the detection of such activities, different earable-mounted sensors can then help capture microscopic facial or eye movements (such as eye gaze while searching for a household item or frowning while trying to recall a PIN), which can help reveal cognitive challenges associated with such activities.

We believe that this form of multi-device, longitudinal multi-modal monitoring provides a couple of benefits for a new paradigm of cognition and emotion monitoring:

- Using gestural activities as a trigger to dynamically activate earable sensing will help conserve the scarce battery resources of such earable devices. It is worth noting that the battery capacity on a representative *eSense*[1] earable is 40 mAh–i.e, $\sim \frac{1}{10}^{th}$ that of a typical smartwatch.
- Using a combination of earable and wrist-worn sensing helps provide a greater understanding of a user's activity and environmental context, and thereby improves inference accuracy. For example, the act of frowning may occur due to the challenge of recalling a password, as well as while watching an unfavorable political event on TV. To reliably identify the likely failures of short-term memory and recall, it is important to distinguish between these two contexts.

As wrist-worn activity sensing is relatively well understood, in the rest of the paper we primarily focus on evaluating the feasibility of using earable-based cognitive sensing.

III. Preliminary Exploration: Accelerometer, Pressure and Ultrasound Sensing

Our hypothesis is that the wearable device's placement, close to the brain, head, eyes and face, will enable its embedded sensors (such as an accelerometer) to capture various low-level physical movements associated with different cognitive and affective states. In addition, we hypothesize that eardrum movements, generated when the eyes move [8] and arising from the anatomical connection between the extraocular muscles and the inner ear, generate in-ear deformations that can possibly be captured via pressure changes or ultrasound scans.

A. Preliminary Analysis: Accelerometer

Recent work has demonstrated that earable-based accelerometers can sense and detect more-visible human motion, such as walking/running [13], head tilt and turning [7], [19]) and jaw movements during eating [5]). However, the ability of one or more inertial sensors to detect subliminal physical cues such as eye blinks and gaze movements is largely unknown. Accordingly, we equip a human subject with a commercially available and widely adopted eSense device (embedded with a 6-axis IMU). Figure 3 depicts the accelerometer readings (from left ear) generated while the subject performs three actions: (a) nodding head, (b) left to right pupil movements when watching a video of a ball moving horizontally, and (c) reading aloud. We can observe that while the accelerometer accurately detects macro movements like nodding head, it fails to detect micro-facial movements, such as gaze and jaw movements (reading). Accordingly, alternative, uncommon sensing modalities will be needed to detect the weaker physical signals associated with cognitive and affective states.

B. Preliminary Analysis: Pressure Sensor

An innovative aspect of our work is the exploration of the physical phenomena called *ear canal deformation*: our conjecture is that relaxing/contracting facial muscles cause the ear canal to expand/shrink, changing its shape and volume, and indirectly modifying the internal air pressure when the acoustic *meatus* is sealed.

We start our investigation by measuring the changes in the barometric pressure of the ear canal while exerting various facial muscles. In particular, we attach a differential pressure sensor (Honeywell micro-pressure sensor[2] with a measurement

[1] eSense–http://www.esense.io/

[2] Honeywell sensor– https://tinyurl.com/honeywellmicropressuresensor

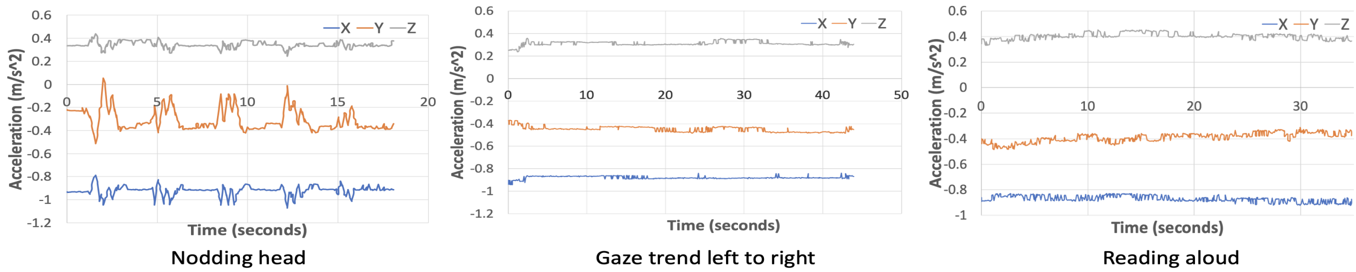


Figure 3: Accelerometer readings: different facial actions

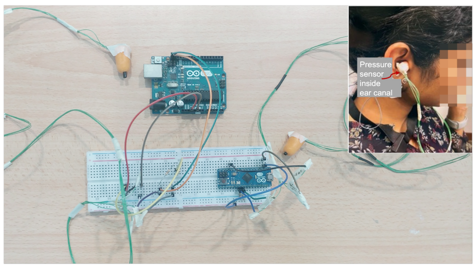
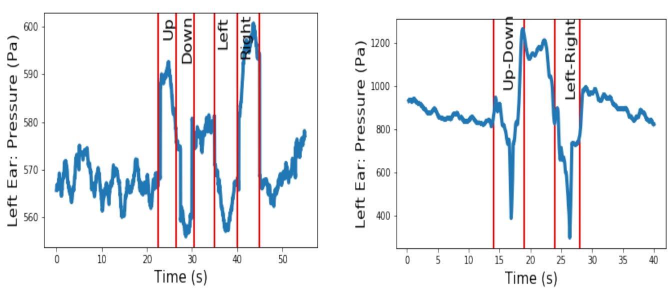


Figure 4: Pressure sensor setup



(a) Gaze Movement (b) Head Movement

Figure 5: Recorded pressure signals while (a) moving gaze, and (b) head up-down-left-right

range of 60 mbar-2.5 bar, or 6-250 kPa) on the tip of a standard foam-based earplug (preferred due to its convenient form factor and lightweight nature). As depicted in Figure 4, the outer end of the earplug was sealed with hot glue to secure the pressure sensor. Figure 5 plots the pressure variations recorded by our setup (after discounting the inhale-exhale process and pulsing effect) while the human subject performed the following activities: (a) gaze movement, and (b) turn head (up-down-left-right). We can observe that even subtle facial cues such as moving gaze up/down and left/right cause observable pressure variations inside the ear canal, arising from geometric changes of the ear canal (expansion or compression of the ear canal wall). However, identifying the muscle groups that cause such physical deformation using a pressure sensor is not trivial. For example, the contracting/relaxing of two distinct facial muscle groups (such as Zygomaticus Major and Frontalis) may generate non-distinguishable pressure signatures, presenting a challenge to reliable activity classification.

C. Preliminary Analysis: Ultrasound Sensing

Literature suggests a multi-modal interaction between visual and auditory systems occurs at the eardrum, causing it to move during eye movements [8]. Such deformations may thus result in changes in the frequency response (phase and amplitude) of the reflected signal of a sound that is injected into the ear canal. While continual injection of audible sound will simply be unacceptable to a human, ultrasound signals are largely imperceptible and may offer a means of such *active* audio sensing. As a preliminary exploration, we utilized the eSense earbud to transmit an effectively inaudible ultrasound signal (with a frequency range of 18-20 KHz and chirp duration of 1 microsecond) into the ear canal and recorded the microphone response while the subjects performed varying actions such as 'close eyes', 'blink eyes', 'shake head', 'reading'. Figure 6 plots the histogram of the frequency responses for these distinct eye movements. We see that each movement is associated with a distinct frequency response pattern, thereby suggesting that inaudible acoustic sensing may help capture such subtle eye and facial movements.

IV. EMG Sensing

In this section, we shall focus on the biosignals captured by EMG electrodes that are placed near the ear (i.e., not on the target muscle that fires the electric potential) and perform a more detailed evaluation of their capabilities of sensing various minuscule facial muscle movements and actions.

A. Experiment Setup and Data Collection

To study the feasibility of our proposed vision, we conducted preliminary studies exploring the use of ear-based EMG (Shimmer EMG sensor[3]), with a sampling rate of 512 Hz, to capture muscle activities for different facial movements. More specifically, we attach one EMG channel per ear – the positive and negative electrodes of the EMG channel are attached respectively to the front and back of the same ear (depicted in Figure 7). The reference electrode is attached to a site farther away from the two ears (i.e., at the back of the neck).

Through multiple iterations, we find that our current placement of EMG electrodes provide the right balance between the ability to accurately capture subtle facial movements and enable continuous and unobtrusive monitoring. In contrast, a

[3]Shimmer sensor–https://shimmersensing.com/

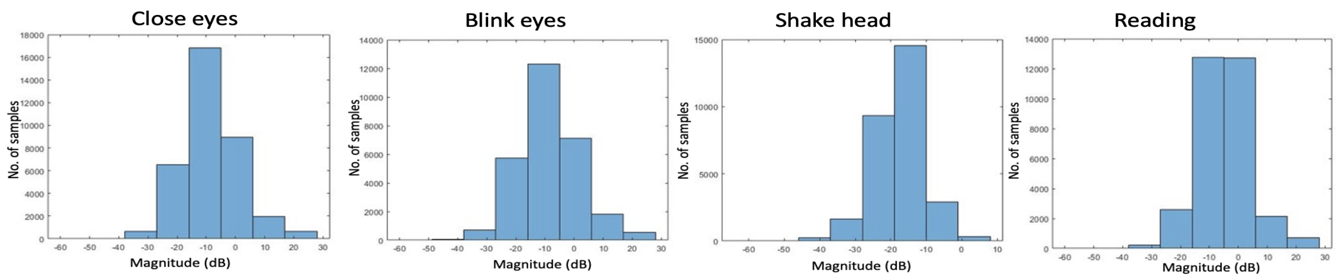


Figure 6: Histogram of frequency responses for different actions (Ultrasound sensing)

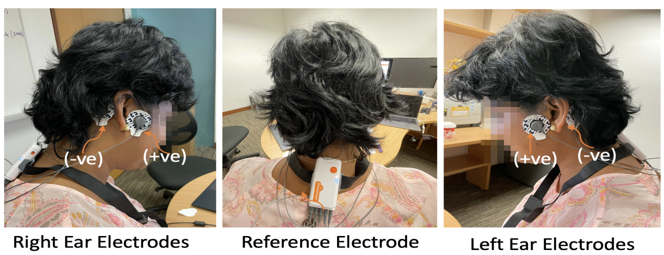


Figure 7: EMG electrode setup

commonly adopted behind-the-ear setup (similar to the one in [5]) fails to capture certain subtle actions such as pupil movements. Our setup is less obtrusive and more suitable for continual EMG monitoring.

We obtained data from 3 subjects who performed six unique actions such as turning head to left and right, tilting head to left and right, opening and closing mouth, nodding head, pupil movements to the left and right, and blinking eyes. For each experiment, the subjects were seated and instructed to minimise their motions – the experiment consists of two phases: during the (a) activity phase, the subjects were asked to perform facial actions while their eyes are closed (to minimise crosstalk) except for the pupil movement and blinking eyes tasks, and the (b) resting phase (for 10-15 seconds; eyes were closed for all the activities) was introduced both before and after the activity phase to minimise motion-related artifacts. For the pupil movement task, the subject was asked to watch a video of a ball repeatedly and horizontally moving from left to right. We collected multiple samples of each action and each of the experiments lasted for about 30-80 seconds. To collect the ground truth, we recorded videos.

B. EMG Data Analyses and Insights

Figure 8 plots the EMG signal variations – we can see that each facial action generates an observable trend in the EMG readings. Interestingly, for certain actions like tilting head, we observe an inverted trend in the signals across the right and left ear, indicating the opposing muscle contraction and extension of the ear muscles.

Our studies reveal additional challenges to overcome. In our analysis, we find evidence of the phenomena called "hyper polarization" [14] (depicted in Figure 9), whereby a muscle does not instantaneously revert to its resting state but exhibits hysteresis. The impact of such hysteresis is further illustrated in Figure 10(a) when the facial actions (i.e., gaze movement) were repeated continuously – we observe, for example, that the peak EMG signal amplitude varies considerably across successive identical left-to-right pupil movements (each movement separated by green dashed lines in the image). This implies that unlike other state-of-the-art activity classifiers (e.g., inertial sensor-based), real-world EMG-based activity classifiers cannot be stateless but must incorporate the temporal separation between consecutive micro-activities.

Key Takeaway: The biosignals generated by eye or facial muscle movements are significantly weaker, and often have significant temporal dependencies. The impact of hyperpolarization poses a serious limitation on the usage of EMG to monitor facial muscle movements as it requires a significantly longer "resting period" between any two subsequent activities, while, in reality, various muscle movements can co-occur (such as smiling while blinking eyes).

To correct the effect of hyperpolarization, we propose a simple linear detrending of the signal. We fit a linear regression model on the signal, obtain the trend (as shown by the orange line in the example Figure 10(a)) and subtract it from the signal. Figure 10(b) shows the detrended signal.

C. EMG-based Facial Action Classification

The pre-processed EMG signal from Shimmer sensor (i.e., analog-to-digital converted) is first mean corrected and then rectified. It is then passed through a low pass filter with a cut-off frequency of 2 Hz and an order of 4. We then extract time and frequency domain features (as listed in Table I) on the processed EMG signal. In addition to the standard features for both temporal and frequency domains, we compute 'Willison amplitude' (a widely adopted feature for EMG-based pattern recognition) – the number of times the EMG amplitude exceeds a predefined threshold (here the threshold $= 3 * mean(EMG) + SD(EMG)$, where SD is the standard deviation).

We use a supervised J48 Decision Tree classifier trained on the features extracted to classify the various facial movements. In addition to the data corresponding to the 6 actions, we also include data from 'no action' as a NULL class in the classifier. As our dataset is imbalanced, we use a cost-sensitive classifier such that there is a higher penalization if all classes except 'no action' are misclassified. Using 10-fold cross-validation,

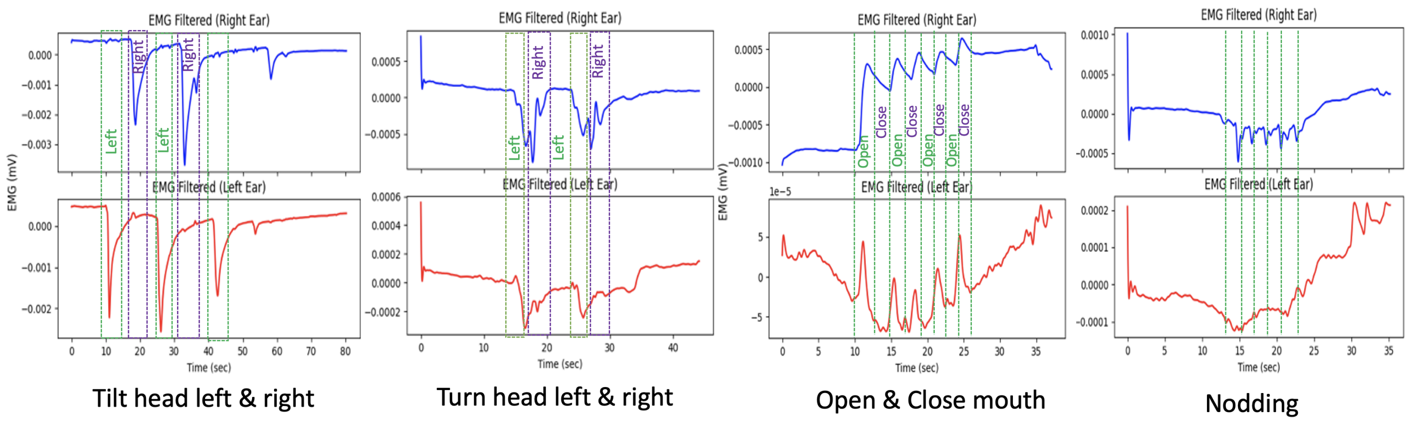


Figure 8: EMG variations for different facial actions. The dotted lines indicate the temporal labels of individual events.

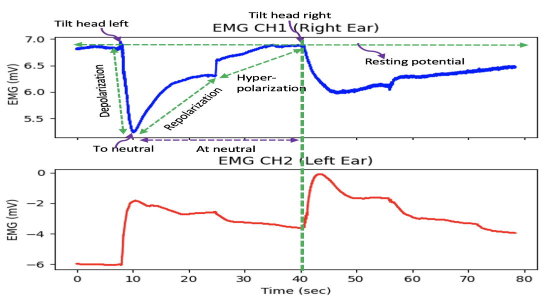


Figure 9: Effect of Hyperpolarization while tilting head to the left and right

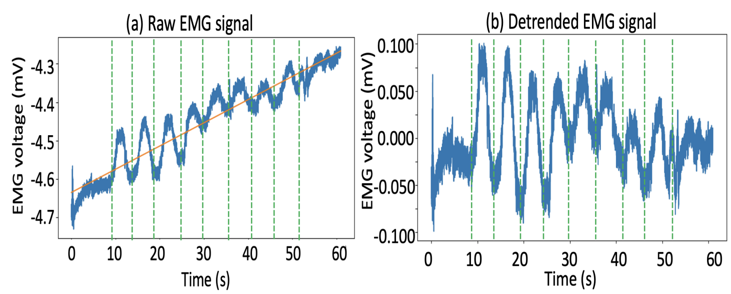


Figure 10: EMG signal variations captured from behind the right ear during left to right pupil movements

Table I: Features extracted on EMG Signal

Time Domain	Frequency Domain
Variance	Frequency Ratio
Root Mean Square	Mean Power
Integral	Total Power
Mean Absolute Value	Mean Frequency
Zero Crossing	Median Frequency
Willison Amplitude	Peak Frequency

we achieved promising results with an overall classification accuracy of **80.02%** (average precision and recall of 0.80 and 0.816, respectively).

V. Challenges and Future Directions

Our initial investigations and experimentation have helped identify the following open challenges that must be addressed.
Automatic EMG segmentation: While our initial results demonstrate how EMG hyperpolarization may be addressed, they are based on manual segmentation of each activity instance, based on ground truth video. Prior approaches for simple threshold-based segmentation of EMG signals are, however, inadequate for our scenarios and will need to be refined to tackle the significant amplitude variations that result from both (a) hyperpolarization during consecutive actions, and (b) simultaneous execution of multiple actions (e.g., eye roll and yawning).

Physical Earable Design with Multi-Modal Sensors: Our results suggest that multiple sensing modes, positioned both inside and around the ear, will be needed to discern the multitude of head, facial and eye based markers of different cognitive activity. For example, both pressure sensors and ultrasonic transmitter/receivers may be needed inside the ear canal, while EMG sensors need to be placed behind and on top of the earlobe. The earable device will thus need a form factor that permits such diverse sensor placement and is lightweight enough to ensure human comfort.
Translating Biomarker Sequences into Cognitive States: Our investigations have hitherto focused purely on the ability to identify each individual biomarker instance. Most cognitive states, however, manifest themselves over longer periods of time (e.g., 10s of secs–mins), and involve the expression of a reasonably long, dynamic *sequence* of such biomarkers. While additional studies are needed to reliably map such biomarker sequences to the underlying cognitive state, the longer period may, however, permit the state estimation to be less sensitive to errors in individual biomarkers.
Ensuring Wrist-Earable Coordination: The wrist-based activity detection triggers need to be timed to ensure adequate observability of relevant physical micro-markers of cognition. In general, the longer the observation period, the higher the activity recognition accuracy. Because cognitive micro-markers can be fleeting and not necessarily exactly synced

with physical gestures, the system design must balance the desire for high precision (to avoid triggering the earable sensors needlessly) with the need for low latency (to avoid missing relevant head/eye/face micro-movement signals).

VI. Related Work

Earable Sensing: Earables have recently received heightened interest for supporting unobtrusive, longitudinal monitoring applications. Due to their close proximity to key body parts, such as the brain, eyes, and facial muscles, earables can consequently not only capture human physiological signals, such as body temperature [16], blood pressure [6] and brain activity [18], but also support monitoring of human activities [13], eating and drinking activities [4], [5], recognize emotions [3] and facial expressions (such as speech utterances and head/tongue movements [1], [2]).

Human Cognitive and Affective State Monitoring: In tandem with earables, head-worn wearables such as "smart glasses" (representing a wide variety of eye-mounted wearable devices, equipped with sensors such as LIDAR, RGB and IR cameras) have begun to re-emerge, driven by the increased interest in augmented and virtual reality applications. State-of-the-art research has recently shown how IR [15] or ultrasonic [11] sensing by such smart glasses can detect eye movements (including blinks) and pupillometry, with such physical context being then used to infer emotional state, albeit in controlled, laboratory settings where such movements are performed discretely rather than continuously. Capturing cognitive and affective states (such as attentiveness and confusion), however, remains a largely unexplored and challenging problem, as such states are typically associated with multiple, almost-subliminal physical actions and physiological changes (such as furrowing one's eyebrows or exhibiting a "blank" stare together with perhaps a shallower breathing rate).

VII. Concluding Remarks

In this work, we have proposed a vision of combining wrist-worn sensing with sensors embedded in ear-worn devices as a novel mechanism to capture various {head, face, eye} movements that serve as proxies of complex cognitive and affective states. To be specific, we have focused on three sensing modalities, namely, ear pressure, ultrasound, and EMG – our early experimental results reveal: (a) in-ear pressure and ultrasound sensing can detect activity-specific shape variations of the ear canal, and (b) an EMG sensor placed near the ear can classify 6 different facial movements with 80+% accuracy.

VIII. Acknowledgment

This research is supported partially by the Ministry of Education, Singapore, under its AcRF Tier 2 Funding (Grant No:MOE2019-T2-2-116) and by the National Research Foundation, Singapore under its NRF Investigatorship grant (NRF-NRFI05-2019-0007). Any opinions, findings and conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of the Ministry of Education, Singapore or the National Research Foundation, Singapore.

References

[1] Amesaka, T., Watanabe, H., and Sugimoto, M. Facial expression recognition using ear canal transfer function. *Proc. of the ISWC'19*, 2019.

[2] Ando, T., Kubo, Y., Shizuki, B., and Takahashi, S. Canalsense: Face-related movement recognition system based on sensing air pressure in ear canals. *Proc. of UIST'17*, 2017.

[3] Athavipach, C., Pan-Ngum, S., and Israsena, P. A wearable in-ear eeg device for emotion monitoring. *Sensors*, 19(18):4014, 2019.

[4] Bedri, A., Li, R., Haynes, M., Kosaraju, R. P., Grover, I., Prioleau, T., Beh, M. Y., Goel, M., Starner, T., and Abowd, G. Earbit: using wearable sensors to detect eating episodes in unconstrained environments. *In Proc. of the ACM IMWUT'17*, 1(3):1–20, 2017.

[5] Bi, S., Wang, T., Tobias, N., Nordrum, J., Wang, S., Halvorsen, G., Sen, S., Peterson, R., Odame, K., Caine, K., et al. Auracle: Detecting eating episodes with an ear-mounted sensor. *In Proc. of the ACM IMWUT'18*, 2(3):1–27, 2018.

[6] Bui, N., Pham, N., Barnitz, J. J., Zou, Z., Nguyen, P., Truong, H., Kim, T., Farrow, N., Nguyen, A., Xiao, J., et al. ebp: A wearable system for frequent and comfortable blood pressure monitoring from user's ear. *Proc. of MobiCom'19*, 2019.

[7] Ferlini, A., Montanari, A., Mascolo, C., and Harle, R. Head motion tracking through in-ear wearables. *Proc. of EarComp'19*, 2019.

[8] Gruters, K. G., Murphy, D. L., Jenson, C. D., Smith, D. W., Shera, C. A., and Groh, J. M. The eardrums move when the eyes move: A multisensory effect on the mechanics of hearing. *Proceedings of the National Academy of Sciences*, 115(6):E1309–E1318, 2018.

[9] Inzelberg, L., Rand, D., Steinberg, S., David-Pur, M., and Hanein, Y. A wearable high-resolution facial electromyography for long term recordings in freely behaving humans. *Scientific reports*, 8(1):1–9, 2018.

[10] Jiang, S., Kang, P., Song, X., Lo, B. P., and Shull, P. B. Emerging wearable interfaces and algorithms for hand gesture recognition: A survey. *IEEE Reviews in Biomedical Engineering*, 15:85–102, 2021.

[11] Liu, J., Li, D., Wang, L., and Xiong, J. Blinklistener: " listen" to your eye blink using your smartphone. *In Proc. of IMWUT'2021*, 5(2):1–27, 2021.

[12] Liu, K.-C., Hsieh, C.-Y., and Chan, C.-T. Transition-aware housekeeping task monitoring using single wrist-worn sensor. *IEEE Sensors Journal*, 18(21), 2018.

[13] Ma, D., Ferlini, A., and Mascolo, C. Oesense: employing occlusion effect for in-ear human sensing. *Proc. of MobiSys'2021*, 2021.

[14] Maier-Hein, L. Speech recognition using surface electromyography. *Master's thesis, Institut fü r Theoretische Informatik Universität Karlsruhe (TH), Karlsruhe, Germany*, 2005.

[15] Nie, J., Hu, Y., Wang, Y., Xia, S., and Jiang, X. Spiders: Low-cost wireless glasses for continuous in-situ bio-signal acquisition and emotion recognition. *Proc. of IoTDI'2020*. IEEE, 2020.

[16] Ota, H., Chao, M., Gao, Y., Wu, E., Tai, L.-C., Chen, K., Matsuoka, Y., Iwai, K., Fahad, H. M., Gao, W., et al. 3d printed "earable" smart devices for real-time detection of core body temperature. *ACS sensors*, 2(7):990–997, 2017.

[17] Park, T., Lee, J., Hwang, I., Yoo, C., Nachman, L., and Song, J. E-gesture: a collaborative architecture for energy-efficient gesture recognition with hand-worn sensor and mobile devices. *Proc. of SenSys'2011*, 2011.

[18] Pham, N., Dinh, T., Raghebi, Z., Kim, T., Bui, N., Nguyen, P., Truong, H., Banaei-Kashani, F., Halbower, A., Dinh, T., et al. Wake: a behind-the-ear wearable system for microsleep detection. *Proc. of MobiSys'2020*, 2020.

[19] Radhakrishnan, M., Misra, K., and Ravichandran, V. Applying "earable" inertial sensing for real-time head posture detection. *Proc. of PerCom'2021 Adjunct (PerHealth Workshop)*. IEEE, 2021.

[20] Ren, Y., Wang, C., Yang, J., and Chen, Y. Fine-grained sleep monitoring: Hearing your breathing with smartphones. *Proc. of INFOCOM'15)*, pages 1194–1202. IEEE, 2015.

[21] Röddiger, T., Wolffram, D., Laubenstein, D., Budde, M., and Beigl, M. Towards respiration rate monitoring using an in-ear headphone inertial measurement unit. *Proc. of EarComp'19*, 2019.

[22] Shen, S., Wang, H., and Roy Choudhury, R. I am a smartwatch and i can track my user's arm. *Proc. of MobiSys'2016*, 2016.

[23] Wang, C., Liu, J., Guo, X., Wang, Y., and Chen, Y. Wristspy: Snooping passcodes in mobile payment using wrist-worn wearables. *IEEE INFOCOM*, 2019.