

Head Motion Tracking Through in-Ear Wearables

Andrea Ferlini

University of Cambridge
af679@cam.ac.uk

Cecilia Mascolo

University of Cambridge
cm542@cam.ac.uk

Alessandro Montanari

Nokia Bell Labs Cambridge
alessandro.montanari@nokia-bell-labs.com

Robert Harle

University of Cambridge
rkh23@cam.ac.uk

ABSTRACT

Head tracking is a fundamental component in visual attention detection, which, in turn, can improve the state of the art of hearing aid devices. A multitude of wearable devices for the ear (so called earables) exist. Current devices lack a magnetometer which, as we will show, represents a big challenge when one tries to use them for accurate head tracking.

In this work we evaluate the performance of eSense, a representative earable device, to track head rotations. By leveraging two different streams (one per earbud) of inertial data (from the accelerometer and the gyroscope), we achieve an accuracy up to a few degrees. We further investigate the interference generated by a magnetometer in an earable to understand the barriers to its use in these types of devices.

CCS CONCEPTS

• **Human-centered computing** → *Ubiquitous and mobile computing systems and tools.*

KEYWORDS

Earables, Head Motion Tracking, Visual Attention

ACM Reference Format:

Andrea Ferlini, Alessandro Montanari, Cecilia Mascolo, and Robert Harle. 2019. Head Motion Tracking Through in-Ear Wearables. In *Proceedings of 1st International Workshop on Earable Computing (EarComp'19)*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3345615.3361131>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EarComp'19, September 9, 2019, London, United Kingdom

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6902-2/19/09...\$15.00

<https://doi.org/10.1145/3345615.3361131>

1 INTRODUCTION

As of 2019, more than 466 million people are suffering from hearing impairments¹ [1]. By 2050, the World Health Organization forecasts that over 900 million people will be affected by hearing disabilities [1]. Subjects with hearing impairments (and non-impaired, too) struggle to identify and isolate the source of a sound. This is particularly true in social situations. Crowded places, where multiple conversations, involving several speakers happen at the same time, result in the phenomenon known as the *Cocktail Party Problem* [11].

Earables, like the eSense platform we evaluate in this paper, have an enormous and mostly unexploited potential. For instance, if adopted as hearing-aids, earables could be used both as sensors and actuators. Ears represent an extremely good vantage point to sense those behaviours that could be exploited to improve the performance of a hearing-aid (e.g. gaze tracking, head movements, etc.). Previous studies, especially from the medical community, highlight the importance of using directional hearing-aids leveraging the visual attention of the user [5, 14]. For example, Favre-Félix et al. [5] use electrooculography data (EOG) to characterize the visual attention of the patients. Their study shows how steering an hearing-aid by using EOG leads to better performance on the sentence correctness score. However, because of the challenging signal processing required when dealing with it, the authors suggest that EOG may not be the best enabler to steer hearing-aids with a sufficiently high degree of precision. Head movements are closely linked to eye-movements [5], and therefore they are considered a good proxy to sense visual attention, too.

Inertial motion tracking is a well known and studied problem. Yet, due to the lack of a reference point to re-calibrate the sensors, and to estimate the 3D orientation of the tracked object, tracking head movements with a device without a magnetometer represents a challenging task. To the best of our knowledge, because of the interference generated by the magnets of the speakers and in their cases, none of the earables in the market is equipped with a magnetometer. Indeed, like the eSense we are using in this work, the Apple

¹Hearing disabilities are considered so when referring to hearing loss greater than 40dB in adult subjects and 30dB in children.

AirPods², the Google Pixel Buds³, and the Samsung Galaxy Buds⁴ do not have a built-in magnetometer. *In this work, we focus on the evaluation of the eSense platform [7, 12] in tracking the head movements of a user concentrating on a specific spatial point.* To do so, we ran experiments with ten volunteers. We probed and stressed the robustness of the system by asking our volunteers to perform different activities, such as chewing and talking, while focusing on a series of targets placed at different spatial locations.

By tracking instantaneous head movements as a proxy to track visual attention, our study shows how a system, that relies only on accelerometer and gyroscope, can still provide useful insights on where a person is facing. Despite the fact that head movements are user dependent, we obtained estimations with an average error that ranges from 5.4 degrees for short movements done by silent subjects, to 18.7 degrees for longer movements carried out by subjects who are chewing. This paper lays the foundations of a line of work aiming to sense and characterize human attention through earables, wearables that are neither socially-awkward, nor cumbersome, unusable or unrealistic (e.g. combining an hearing-aid with a pair of eye-tracking glasses). Lastly, it sheds light on how a magnetometer would behave if placed in an earable.

2 RELATED WORK

Earables. Earables are a relatively new concept. Despite their huge potential, few research works currently use them. Most focus on health monitoring and sensing. For instance, LeBoeuf et al. [10] prototyped an optomechanical sensor to sense blood flow and estimate oxygen consumption during daily activities. The studies carried out by Bedri et al. [3] and by Amft et al. [2] combine IMUs and microphones to detect and classify eating activities. The former aimed to detect in-the-wild chewing activities, whereas the latter focused on the classification of four different types of food through the analysis of eating activities.

Head Movements Tracking. Inertial motion tracking is a known challenge and a well explored area. One of the most recent works in the field, and the state-of-the-art, is the study carried out by Shen et al. [13]. In their paper, the authors widely discuss the 3D Orientation problem and present MUSE, a magnetometer-centric sensor fusion algorithm for orientation tracking. Their results found that MUSE outperformed all the previous state-of-the-art orientation tracking approaches. Prior to their work, the other state-of-the-art techniques, as A^3 [15], were heavily relying on the gravity to determine the object orientation in the space, using the

magnetometer data mainly to re-calibrate the system. However, as reported by Shen et al. [13], those previous works are mostly based on the following assumptions: (1) slow linear motion, with accelerometer data that have gravity as average; (2) slow rotational motion, with Gaussian errors that preserve the linearity of the system; (3) motion with frequent, fairly long pauses, needed to reset the gravity estimation.

Yet, because of the huge freedom and unpredictability that characterizes human motion, these assumption rarely hold when tracking head movements. Our work partially follows the approach used by LaValle et al. [9]. However, in their work, the authors use a significantly different hardware (an Oculus headset), which is equipped with a magnetometer.

This paper assesses and evaluates the performance of eSense, an earable equipped solely with accelerometer and gyroscope, to track head movements. In addition, it presents a preliminary study on the effects of a magnetometer if placed in an earable. Given that there are no available earables equipped with this kind of sensor, we believe these preliminary observations represent an interesting input to both the research community and industry.

3 PLATFORM OVERVIEW

In this section we introduce the eSense platform and the challenges of performing head movement tracking on a device that can not rely on the data from a magnetometer.

eSense Platform

The eSense platform [7, 12] consists in a pair of true wireless earbuds which have been augmented with kinetic, audio and proximity sensing options. The left earbud has a 6-axis Inertial Measurement Unit (IMU) with accelerometer and gyroscope and a Bluetooth Low Energy (BLE) interface which is used to stream data and to send periodic beacons that can be used to detect proximity to nearby devices. Both earbuds are also equipped with microphones to record external sounds. The benefit of eSense, contrary to other commercial earbuds, is the access to the raw data from the onboard sensors and the complete flexibility in the configuration of their parameters. In addition to serve as a well established and socially acceptable device, for example to listen to music and take phone calls, eSense allows to gather real-time sensor data, opening the door to novel sensing applications involving the head.

Challenges of Inertial Tracking

The primary goal of this work is to understand the accuracy achievable by a device that solely relies on accelerometer and gyroscope to track user's head movements. The biggest challenges we had to face while investigating that, were related to the device itself. The small form factor, together with the pure "wireless experience" and the relatively short battery life, represent non-trivial constraints to deal with.

²<https://www.ifixit.com/Teardown/AirPods+2+Teardown/121471>

³<https://medium.com/@justlv/google-pixel-buds-teardown-396183cbbc18>

⁴<https://root-nation.com/audio-en/headphones-en/en-samsung-galaxy-buds-review/>

The presence of multiple magnets in the case compelled the hardware manufacturer to put a 6 degree of freedom (6 DoF) IMU, instead of a 9 DoF, more complete, sensor. Practically, it means the platform is bounded to the 3 DoF of the accelerometer ($x_{acc}, y_{acc}, z_{acc}$) and the 3 of the gyroscope ($x_{gyro}, y_{gyro}, z_{gyro}$), lacking the presence of a magnetometer.

Accelerometer and gyroscope provide relative movement estimates that drift over time. Without the magnetic north as a reference, we could not rely on the state-of-the-art calibration and re-calibration techniques [13]. Besides, when tracking the motion of an object (or of the head of a person), it is crucial to initialize the system correctly, with the right 3D orientation of the object itself. Unfortunately, once again, the majority of the algorithms to solve the 3D Orientation Problem rely on the absolute direction reference provided by the magnetometer (combined with gravity and instantaneous gyroscope readings) [13, 15].

4 HEAD TRACKING METHODOLOGY

The data we collected from the eSense buds are accelerations (from the accelerometer) and rotational velocities (from the gyroscope). We had integrate the data to get an idea of where the person is paying its visual attention to.

Combining multiple IMUs. Tracking head motion from only one ear might not give precise enough results, especially considering the absence of external reference points, such as a magnetometer. As having more fine grained data points enhances the precision of the estimation, we combined the streams of inertial data coming from the IMU sensors in the two earbuds⁵. We leverage the assumption that the two different IMUs are recording, from different vantage points, the same rotation. Yet, combining the data correctly becomes crucial. We do so by concatenating and averaging the accelerometer and gyroscope data over a window of 200ms. Prior to that we resample and filter the readings from the IMU sensors. A further challenge comes from the orientation of the IMUs themselves. We obviate that by making our system independent from the human coordinate. To do that, we only consider the intensity of the rotation, rather the motion components along the 3 axis of the accelerometer and the 3 of the gyroscope. In fact, we only leverage the motion components to tell whether the rotation is positive (to the right) or negative (to the left).

Quaternions and Complementary Filter. Euler angles, better known by their components *yaw*, *pitch*, and *roll*, are the most common, and widely used, coordinate system to represents rotations. Despite their diffusion due to their ease of interpretability, they come with the problem know as Gimbal Lock [4]. To obviate the gimbal lock problem, it is common

⁵since only the left eSense buds are equipped with the IMU sensors, we had to use two left buds.

to switch to a better suited coordinate system: Quaternions [6, 9].

Integrated gyroscope measurements are subject to short-term drift, which may be more or less severe depending on the application. The situation worsen over time, as the error grows faster. To mitigate that, a common approach is to fuse the gyroscope readings with the accelerometer ones, as it is known to be more stable than the gyroscope in the short-time.

Our angular estimation is based on a complementary filter, which allows us to fuse gyroscope and accelerometer data, and follows the approach proposed by LaValle et al. [9]. We derive our estimation as follows:

$$q_{gyro} = \cos\left(\frac{\theta}{2}\right) + i\omega_x \sin\left(\frac{\theta}{2}\right) + j\omega_y \sin\left(\frac{\theta}{2}\right) + k\omega_z \sin\left(\frac{\theta}{2}\right) \quad (1)$$

where:

$$\omega = (\omega_x, \omega_y, \omega_z) = \left(\frac{gyro_x}{\|\omega\|}, \frac{gyro_y}{\|\omega\|}, \frac{gyro_z}{\|\omega\|}\right), \text{ and } \theta = \|\omega\|dt.$$

q_{gyro} is the quaternion that describes the instantaneous rotation of the head, based on the gyroscope data. We partially account for the gyroscope drift by adding the 3D orientation estimation and tilt correction of the head. To do that properly, we should rely on a combination of magnetometer and accelerometer data. Instead, since the eSense do not have a 9 DoF IMU, we could to only rely on the gravity as external reference to perform a rough orientation estimation and tilt correction.

$$\begin{aligned} q_{acc_body} &= 0 + i acc_x + j acc_y + k acc_z \\ q_{acc_world} &= pos[t] * q_{acc_body} * pos[t]^{-1} \\ q_{cf} &= \cos\left(\frac{\phi}{2}\right) + i\frac{n_x}{\|n\|} \sin\left(\frac{\phi}{2}\right) + j\frac{n_y}{\|n\|} \sin\left(\frac{\phi}{2}\right) + k\frac{n_z}{\|n\|} \sin\left(\frac{\phi}{2}\right) \end{aligned} \quad (2)$$

where q_{cf} is the implementation of the complementary filter in the space of quaternions, and where:

$$\begin{aligned} \phi &= (1 - \alpha) \arccos\left(\frac{q_{acc_worldy}}{\|q_{acc_world}\|}\right) \\ v &= \left(\frac{q_{acc_worldx}}{\|q_{acc_world}\|}, \frac{q_{acc_worldy}}{\|q_{acc_world}\|}, \frac{q_{acc_worldz}}{\|q_{acc_world}\|}\right) \\ n &= v * (0, 1, 0) \end{aligned}$$

We can now estimate the final rotation by doing:

$$final_position = q_{cf} * pos[t] \quad (3)$$

Notice that for each earbud, we account for the factory offset of both accelerometer and gyroscope by using the techniques described by Kok et al. in their work [8]. In addition, because of the absence of the magnetometer, we only focus on relative rotations (*delta motions*).

5 USER STUDY

In this section we describe the methodology we followed to investigate head motion tracking through earables. We detail how we collected the data and the results obtained tracking the head movements of our volunteers. Ethical approval was obtained to conduct the user study.

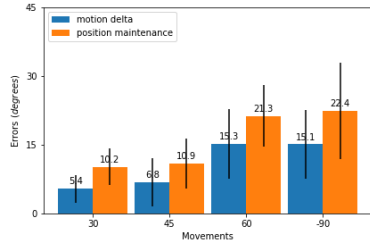


Figure 1: Mean error and standard deviation of the head movements estimation of 10 silent volunteers.

Data Collection Methodology

We recruited 10 volunteers to join our data collection campaign. Each individual was wearing two earbuds (both eSense left bud) connected via BLE (Bluetooth Low Energy) to an Android application running on a smartphone⁶ provided by us. The eSense buds collected, and streamed to the smartphone, at 100 Hz. For the sake of reproducibility we report the configuration of the two earbuds. Notice the two buds have the same configuration.

The experimental set up consisted of 4 targets (red cross) attached to the walls in an empty room. The angle between the position of the volunteer and the targets was known and represented ground truth. The targets were placed respectively at 30, 45, 60, and -90 degrees. We chose 30 degrees as the smallest angle we investigate assuming that for smaller angles people would mostly move their eyes, barely moving their heads. With -90 degrees, we wanted to show how our system could capture both clockwise and counterclockwise movements. We asked our volunteers to perform the following actions:

- Standing in silence and looking at different targets, and keep facing them, according to the instructions of the investigators;
- Standing, chewing a piece of chewing gum, and looking at different targets, and keep facing them, according to the instructions of the investigators;
- Standing, conversing with one of the researchers, and looking at different targets, and keep facing them, according to the instructions of the investigators.

Baseline: Silent Subject

We now present the results of what we consider our baseline. In this experiment, the volunteers were standing in silence, looking at the different targets according to a set of instructions provided by us. For each target, the volunteers started facing an initial reference, placed at 0 degrees. They then rotated their heads towards the given target (*delta motion*). Once there, we asked them to keep their head turned towards the target for about 5 seconds (*position maintenance*). We

⁶Google Pixel 2, https://en.wikipedia.org/wiki/Pixel_2

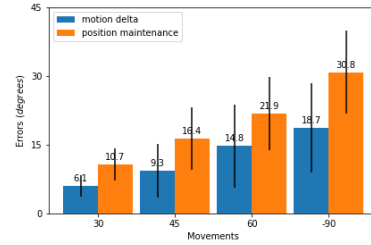


Figure 2: Impact of chewing activity on mean error and standard deviation of the head movements estimation of 10 volunteers.

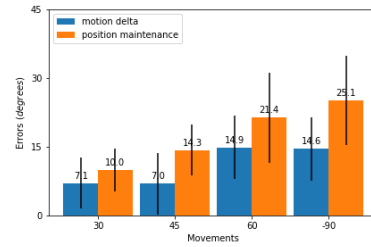


Figure 3: Impact of speech on mean error and standard deviation of the head movements estimation of 10 volunteers.

estimated the movements done by the volunteers processing the readings of accelerometer and gyroscope according to what described in Section 4. Figure 1 respectively reports the mean errors of the motion delta estimation and the position maintenance estimation when the users were rotating their heads clockwise 30, 40, and 60 degrees, and counterclockwise to -90 degrees. From the chart, we can immediately appreciate how the position maintenance errors are greater than the motion delta ones. This is due to the inertial sensors' drift that heavily affects the integration. Because of our long term application, we are interested in instantaneous movements (deltas) we therefore mostly care about motion delta errors. Another interesting observation is how the errors grow for longer movements (i.e. greater angles), indicating a higher precision of the system for short movements (i.e. small angles). Moreover, the high standard deviation that characterizes the mean errors denotes a strong user dependency of the motion estimation. In fact, for some volunteers, we even registered sub-degree motion delta accuracy in some movements.

Impact of Chewing Activity

Once we evaluated our system on the simplest case, we started testing the robustness of our motion estimation. We gave chewing-gum to our volunteers, and asked them to repeat the same sequence of movements. The chewing activity of the volunteers generated spurious vibrations that were inevitably picked up by the inertial sensors in the earbuds. To

make our system more robust to this kind of noise, we tuned the parameter α of our complementary filter, aiming for the best performance in all the three types of experiments. Figure 2 depicts the mean errors of the estimation. If compared with our baseline, as expected, the errors are slightly higher. As we observed in the previous case, because of the drift, the motion delta mean errors are smaller than the position maintenance ones.

Impact of Speech

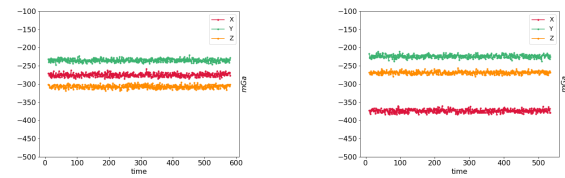
Lastly, we assessed how speech affects our estimation. To do so, we asked our volunteers to talk. As for the chewing experiment, speaking generates unwanted vibrations and micro-movements that are captured by the sensors, and we are interested to study the robustness of our system. Figure 3 shows how our system behaves with talking subjects. As before, the mean errors of the estimation in this third case are slightly higher than the ones for silent users. As in both the previous cases, the motion delta accuracy is higher and decreases for longer movements. Comparing Figure 2 and Figure 3 we can observe that chewing activity seems to have a comparable impact on our motion estimation as speech. The high standard deviation of the mean errors further remarks the findings we got from the previous experiments.

Summary. Evaluating the performance of the eSense platform, our work investigates the potential of ears as a vantage point to sense visual attention through head movements. We achieve estimations with an average error that ranges from 5.4 degrees for short movements in the least challenging situation, to 18.7 degrees for longer movements, in noisier circumstances. In the reminder of this paper, we try to further improve the accuracy of our system, investigating whether it would be feasible add a magnetometer, to gain an external absolute reference point. This would let us to track absolute and incremental movements and not only relative motion.

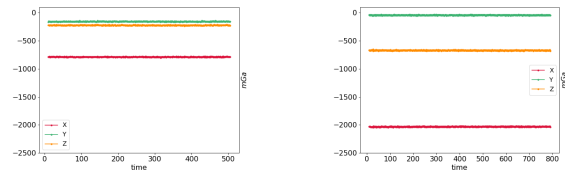
6 WHY IS THE MAGNETOMETER MISSING?

In order to asses if we could use state-of-the-art motion tracking approaches to further improve the precision of our motion estimation, we studied how the magnets used to hold the earbuds into the case and the magnet in the speaker affect the readings of a magnetometer. We did that by placing a STEVAL-STLCS01V1 sensor tile⁷ at different distances from one earbud, and we plotted the data captured by the magnetometer in Figure 4.

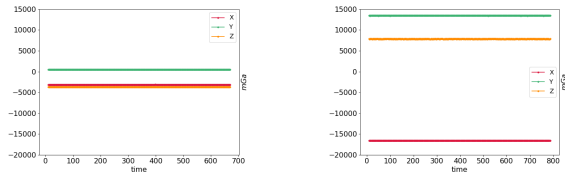
To start off, we put down the magnetometer at a distance of 5 cm from the earbud (Figure 4a). We proceeded keeping the earbud still in the initial position, while moving the sensor tile closer to the bud. Figure 4b shows the magnetometer readings when 3 cm apart. We can immediately appreciate



(a) 5cm from the earbud. (b) 3cm from the earbud.



(c) 1.5cm from the earbud. (d) 1.25cm from the earbud.



(e) 1cm from the earbud. (f) Inside the earbud case.

Figure 4: Magnetometer readings from a STEVAL-STLCS01V1 device at different distances from one eSense earbud. Notice how the scale of the plots changes from top to bottom.

how the magnets start affecting the data, introducing an offset. It is worth noticing how the offset does not fluctuate when the magnetometer is fixed. As expected, the readings change at different distances from the earbud. The closer we get to the earbud, the higher the influence of the magnets is. We further moved the STEVAL-STLCS01V1 at a distance of 1.50 cm, 1.25 cm, and 1 cm, as respectively depicted in Figure 4c, Figure 4d, and Figure 4e. Here, we can clearly notice how the magnetic field generated by the magnets in the earbud overtakes the Earth’s one, flattening all the readings. Eventually, we put the sensor tile inside the case (Figure 4f). The readings skyrockets, as the earbud’s magnetic field adds a significant offset to the Earth’s.

Because of the constant offset at different locations, we decided to delve deeper into the behaviour of the magnetometer, collecting more data samples while moving the device. We repeated the same set of movements twice. In the beginning, we moved the STEVAL-STLCS01V1 alone, without any direct external interference caused by either the earbud or by the vicinity of a metallic source (Figure 5a). In the following experiment, we recorded the data, performing the very same movements, but placing the STEVAL-STLCS01V1 inside the case of the earbud, swapping it with the eSense’s existing PCB (printed circuit board) (Figure 5b). While in the first case, where there was not interference, the magnetometer

⁷<https://www.st.com/en/evaluation-tools/steval-stlcs01v1.html>

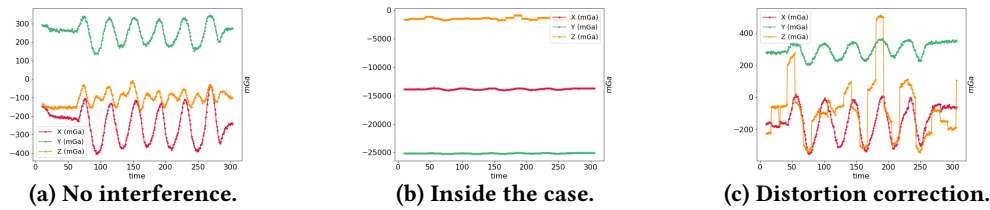


Figure 5: We run two experiments where a volunteer was asked to shake his head while having the a STEVAL-STLCS01V1 close to the ear. (a) We first collected data without magnetic interference. (b) We then asked our volunteer to repeated the same movements with the STEVAL-STLCS01V1 placed inside the eSense case. (c) This way we could observe how correcting the offset introduced by the magnets, the magnetometer placed in the earbud is still able to record motion data.

was able to record the motion events, from Figure 5b we are unable to observe any motion-related data. However, the constant trend of the offset let us apply the calibration technique know as Hard Iron Distortion. As a result, we managed to recover most of the motion related information, especially along x and y (Figure 5c). These preliminary results provide an initial indication about the possibility of integrating a magnetometer even with the presence of strong magnets in the sensor's vicinity. However, experiments with different conditions (e.g. when the buds are playing music) and a detailed analysis of the resulting data are needed to confirm our insights. We leave this for future work.

7 FUTURE WORK

Head movement tracking represents only a first step towards the characterization and sensing of human attention. The behavioural cues we want to sense are not only limited to head movement. After studying how the magnetometer would work if placed in the case, and after evaluating the eSense platform to track head motion, this paper motivates a second revision of the hardware. This should include a wide variety of sensors, such as a 9 DoF IMU, EOG support for gaze tracking, and electroencephalography (EEG) sensors. Such a device, could represent a step towards a platform to sense attention and act as an advanced hearing-aid. We acknowledge that the suggested improvements will pose significant technical challenges, such as battery-life and form factor, which we will gradually explore and tackle in future work.

8 CONCLUSION

In this work, we evaluated eSense as an earable device to perform in-ear head motion tracking. Our technique combines multiple streams of data, and, despite the absence of a magnetometer in the inertial sensor equipped in eSense, achieves results precise up to a few degrees. Although the accuracy of our estimation decreases for longer movements, it performs well also in more realistic situations (e.g. with the subjects speaking or chewing). Besides, our preliminary study on the magnetometer represents an interesting input to take into account in the development of future earables.

ACKNOWLEDGMENTS

This work is supported by Nokia Bell Labs through their donation for the Centre of Mobile, Wearable Systems and Augmented Intelligence.

REFERENCES

- [1] World Health Organization. Deafness and hearing loss. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>, 2019.
- [2] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster. Analysis of chewing sounds for dietary monitoring. In *Proceedings of the 7th International Conference on Ubiquitous Computing, UbiComp'05*, 2005.
- [3] A. Bedri, R. Li, M. Haynes, R. P. Kosaraju, I. Grover, T. Prioleau, M. Y. Beh, M. Goel, T. Starner, and G. Abowd. Earbit: Using wearable sensors to detect eating episodes in unconstrained environments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3), Sept. 2017.
- [4] J. Diebel. Representing attitude: Euler angles, unit quaternions, and rotation vectors. *Matrix*, 58(15-16):1-35, 2006.
- [5] A. Favre-Félix, C. Graversen, R. K. Hietkamp, T. Dau, and T. Lunner. Improving speech intelligibility by hearing aid eye-gaze steering: Conditions with head fixated in a multitalker environment. *Trends in Hearing*, 22:2331216518814388, 2018.
- [6] W. R. Hamilton. *Lectures on Quaternions: Containing a Systematic Statement of a New Mathematical Method*. 1853.
- [7] F. Kawsar, C. Min, A. Mathur, and A. Montanari. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing*, 2018.
- [8] M. Kok, J. D. Hol, and T. B. Schön. Using inertial sensors for position and orientation estimation. *arXiv preprint arXiv:1704.06053*, 2017.
- [9] S. M. LaValle, A. Yerushova, M. Katsev, and M. Antonov. Head tracking for the oculus rift. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 187-194, May 2014.
- [10] S. F. LeBoeuf, M. E. Aumer, W. E. Kraus, J. L. Johnson, and B. D. Duscha. Earbud-based sensor for the assessment of energy expenditure, hr, and vo2max. *Medicine and science in sports and exercise*, 2014.
- [11] J. H. McDermott. The cocktail party problem. *Curr. Biol.*, 19(22), 2009.
- [12] C. Min, A. Mathur, and F. Kawsar. Exploring audio and kinetic sensing on earable devices. In *Proceedings of the 4th ACM Workshop on Wearable Systems and Applications, WearSys '18*, New York, NY, USA, 2018.
- [13] S. Shen, M. Gowda, and R. Roy Choudhury. Closing the gaps in inertial motion tracking. *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking - MobiCom '18*, 2018.
- [14] B. G. Shinn-Cunningham and V. Best. Selective attention in normal and impaired hearing. *Trends in Amplification*, 12(4):283-299, 2008.
- [15] P. Zhou, M. Li, and G. Shen. Use it free: Instantly knowing your phone attitude. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, MobiCom '14*, 2014.