# STEAR: Robust Step Counting from Earables

**Jay Prakash**
University of Illinois at Urbana Champaign (UIUC)
Singapore University of Technology and Design (SUTD)

**Zhijian Yang**
University of Illinois at Urbana Champaign (UIUC)

**Yu-Lin Wei**
University of Illinois at Urbana Champaign (UIUC)

**Romit Roy Choudhury**
University of Illinois at Urbana Champaign (UIUC)

## ABSTRACT

This paper shows that inertial measurement units (IMUs) inside earphones offer a clear advantage in counting the number of steps a user has walked. While step-count has been extensively studied in the mobile computing community, there is wide consensus that false positives are common. The main reason for false positives is due to limb and device motions producing the same periodic bounce as the human walk. However, when IMUs are at the ear, we find that many of the lower-body motions are naturally "filtered out", i.e., these noisy motions do not propagate all the way up to the ear. Hence, the earphone IMU detects a bounce produced only from walking. While head movements can still pollute this bouncing signal, we develop methods to alleviate the problem. Results show 95% step-count accuracy even in the most difficult test case – very slow walk – where smartphone and fitbit-like systems falter. Importantly, our system *STEAR* is robust to changes in walking patterns and scales well across different users. Additionally, we demonstrate how *STEAR* also bring opportunities for effective jump analysis, often important for exercises and injury-related rehabilitation.

## 1 INTRODUCTION

Step-counting has been an important primitive for the mobile/wearable industry, including smartphones, smartwatches, fitbits, arm-bands, etc. Applications have used step-count to derive statistics like calorie burnt, exercise tracking, activity loggers, and even gait analysis for post-injury rehabilitation. Yet, there is wide agreement that step-count is still not accurate; random actions of the leg and hand lead to over/under estimates. For instance, shaking one's leg while seated can increment the step counter, as could playing drums, ping-pong, or video games.

The problem of accurate step-counting with IMUs is nontrivial. Briefly, the human body bounces as a user walks and this bounce manifests into a periodic sinusoidal signal in the IMU's accelerometer. Step count is essentially the frequency (or the number of peaks) of this sinusoidal signal. In reality, however, various other motions of the human body (and the device) get added to the IMU measurements, polluting the sinusoid, or injecting spurious periodicity even when the user is not walking. Hence, the technical challenge lies in continuously identifying and filtering out these pollutions, while robustly adapting to the user's variations in walking patterns. For instance, a user may walk differently during a stroll on a beach, during a walk to the office, while walking down the stairs, or when tipsy from an evening party. Quickly recognizing these unseen patterns, and yet, filtering out false positives, is a difficult problem. Finally, the IMU itself is cheap, hence noisy, adding an extra layer of complexity to signal processing.

With IMUs becoming popular in modern earphones, a natural question is: *does motion tracking in general, and step counting in particular, benefit from IMUs placed at the ear?* This paper finds that while tradeoffs exist, the net outcome is favorable. In particular, the bounce of the human body during a walk gets reflected at the earphone's IMU, but the random motions (of the leg and hand) get filtered out to a large extent. Said differently, the sinusoid from human walking emerges as a cleaner signal when measured from the ear. Of course, the head motions can still pollute the sinusoid, however, the head's movements are generally more constrained, and mostly rotations (as opposed to linear motions). All in all, the net IMU signal lends itself well to signal processing, resulting in robust step-count estimation.

Of course, translating the core opportunity into a robust system entails 2 key challenges. (1) The orientation of the earphone can vary as the user moves her head. If this orientation is not tracked continuously, the IMU data will not project correctly to the global reference frame (explained later), ultimately derailing the step counting system. (2) The bouncing motion varies over different sessions, and even within a session, the signal shape can change. Standard peak detector techniques falter because spurious spikes or fluctuations causes peak counts to get incremented. Filtering the signal around walking frequencies is difficult since the walking frequency is not known *a priori*. Any attempts to predict this frequency makes the system less robust to variations.

In coping with these challenges, our system *STEAR* adopts simple but scalable methods. For orientation, we continuously integrate the gyroscope to compute the 3D projection matrix from the local to the global reference frame. Since human head motion is only rotations, such a gyroscope based scheme is sufficient. Once the 3D orientation is known, the accelerometer data is projected on to the global frame. Then, instead of peak detection, we apply a dynamic time warping (DTW) based scheme to cope with signal shape variations, and potential spurious fluctuations. Of course, some degree of pre and post processing is performed on the signal to ultimately extract the step count of the user.

We implement *STEAR* on Nokia's ESense earbuds [4], embedded with an IMU. The IMU samples at the highest supported sampling rate of 100Hz. We ask 7 users to walk in normal modes. We also ask users to perform activities such as jumps and skips. Our evaluation reveals the following:

• *STEAR* measures very slow steps with ≥ 95% accuracy, typically a hard problem on phones and watches.

• Earphones are as good, if not better than smartphones in calculating steps in other modes of walking.

• Earphones measure jumping characteristics better using unique properties like a trail of zero-acceleration while a person is in free fall. Smartphones are subjected to noise due to relative jitters and friction with pants and bags, ultimately affecting the jump analysis.

In sum, the contribution of this paper may be summarized as follows: *we show a natural opportunity that human walking motion is better reflected in earphones, and we design a system to harness this opportunity, ultimately resulting in a robust step-counter and jump detection method for earphones.*

The rest of the paper expands on this core contribution, starting with some ground measurements, followed by system design, evaluation, and conclusion.

## 2 GROUND MEASUREMENTS

This paper builds on the premise that IMUs on earphones are less impacted by noise and arbitrary limb motions compared to IMUs in smartphones, wrist-bands, and smartwatches. Figure 1 visually illustrates the motion trajectory of the head during a walk – this trajectory mimics a sinusoid. To verify whether a real IMU measures a sinusoid as well, we record IMU acceleration from a earbud and a phone during a normal walk. Their time and frequency representations are shown in Figure 2(a)-2(d). Evidently, earbuds exhibit much cleaner measurements in comparison to a relatively spread-out and noisy spectrum in phone's IMUs. The frequency spectrum also shows a clearer peak for ear IMUs compared to a mix of comparably strong frequencies for phones.

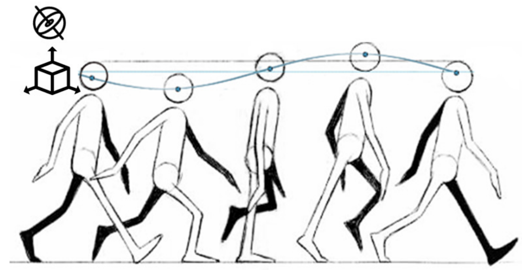

**Figure 1: Sinusoidal motion of a head during walk**

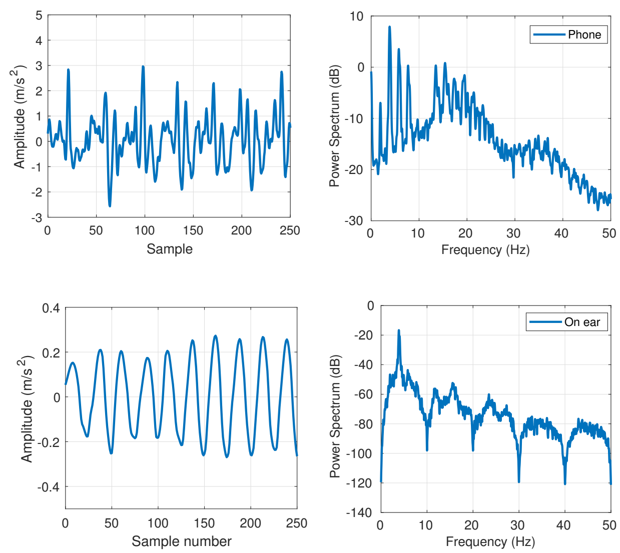


**Figure 2: Acceleration (a) as recorded from phone (b) spectrum at phone (c) as recorded from earable (d) spectrum at earable.**

Next, we study the IMU's signal-to-noise ratio (SNR) across different body locations and activities. SNR here is defined in terms of the ratio of signal power during a period of activity to the power of the noise during a time segment before the activity. Figure 4 shows the results. The participant walked slowly, normally, and ran for the third session; he wore a earbud on the ear, and carried a smartphone in hand and another in the pocket. Evidently, earbuds exhibit an advantage over the phone in the studied scenarios. These form the basis for a robust physio-analytics framework, developed in the rest of the paper.

### Walking Pattern Variations

Figure 5(a) and (b) plot the time domain IMU data from multiple walking sessions – the former is from smartphones while the latter is from earbuds. The multiple lines in each graph are from different users. Clearly, smartphones show more variations, suggesting extraneous pollutions from different parts of the lower body. In contrast, earbuds are consistent,

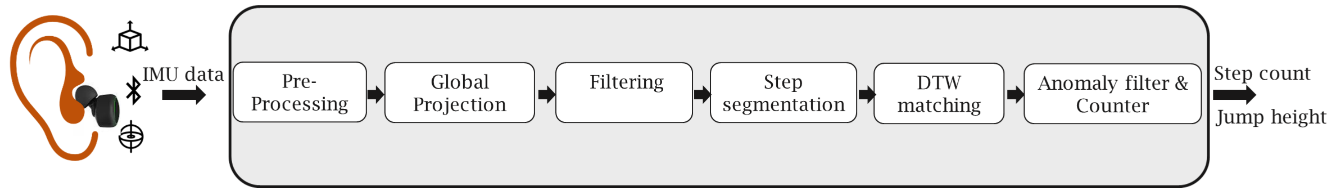


**Figure 3: System Architecture for Step Counting**

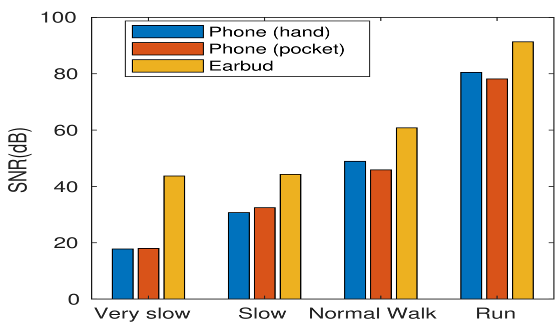


**Figure 4: SNR at Esense and phone for different activities**

again due to the effects of natural filtering. This further endorses the opportunity of robust step counting with earable IMUs.

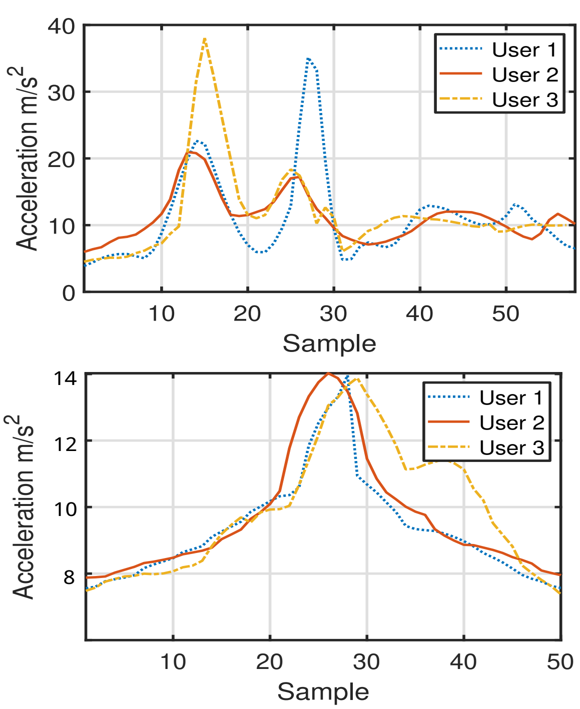


**Figure 5: Comparing templates of steps for different users (a) IMU in phone placed in pocket, (b) IMU in earbud.**

## 3 *STEAR*: SYSTEM DESIGN

*STEAR* has two modules: (1) a step counter that remains always ON, and (2) an exercise analytics module that is activated on-demand. We first describe the pipeline for counting steps, followed by methods for basic exercise, specifically jump analysis. Both these are simple and lightweight, lending themselves well to on-device, real-time processing, even on a small earable platform.

### Step Count

Figure 3 shows the processing pipeline for the proposed step counter, composed of the following modules:

■ **Pre-Processings:** Signal pre-processing mainly includes cancelling the bias in the gyroscope. Bias in gyroscope readings gets magnified after integration and affects the projection matrix to be discussed next. *STEAR* identifies when the device is static[1], and during this static window, computes the average of the gyroscope readings. If the readings are only due to (zero-mean) noise, the average should tend to zero. However, if the gyroscope exhibits a bias, it should be revealed here as a DC value. We subtract this average value over subsequent measurements, thereby compensating the bias. Clearly, larger the static window, better is the estimate, and we use many minutes for averaging.

■ **Global Projection:** IMU sensors report readings in its local reference frame. However, to understand motion trajectories like the one in Figure 1, the data needs to be projected to the global framework. This is because as the device moves, the $<X, Y, Z>$ coordinates are constantly changing/rotating, and the IMU measures the acceleration with respect to its *instantaneous* $<X, Y, Z>$ axes. To explain with an analogy, an IMU is like an airplane passenger who is able to tell that the plane is accelerating "forward" or taking a "left" turn (i.e., in her instantaneous $<X, Y, Z>$ axes), but it is hard for the passenger to track the plane's trajectory on the global world map. To track global movement, the IMU's orientation needs to be constantly rotated to keep the plane horizontal and moving North, and the acceleration needs to be measured in

[1]This is not difficult and is performed by checking if the magnitude of the acceleration is same as gravity.

this fixed global orientation. Said differently, the acceleration needs to be projected to this global 3D orientation.

*STEAR* performs this projection by integrating gyroscope data to estimate global orientation (a standard process in literature). Since gyroscopes drift, we reset it at static instants. As before, static instants are detected when the earphone's acceleration equals gravity and gyroscope measures zero readings (or just noise).

Let's denote the local accelerometer reading at static points as vector $a_{local} = [a_1, a_2, a_3]$. We find a rotation matrix $R$, which rotates $a_{local}$ to $a_{global} = [0, 0, g]$. Note that $R$ is not unique, because we did not specify the horizontal rotation. However, this will not cause a problem because we only care about vertical movements. After that, for each gyroscope reading at time $t_i$, we constantly apply delta rotation matrix $\Delta R_i$, representing the delta rotation from time $t_i$ to $t_{i+1}$, to the original rotation matrix $R$ to give the projection for each time stamp afterwards.

■ **Filtering:** Although earable IMUs offer better SNR, filtering can still be useful. Given that pollutions from some head motions are possible, we can still eliminate them since they are at low frequencies. We apply both low and high pass filter (using moving averages).

■ **Step Segmentation and DTW matching:** Traditional smartphone-based step counting uses peak detection on the accelerometer data. Given the diversity of gaits and smartphone placements, it is hard to separate peaks from walking steps, and those from (unrelated) limb or device movements. This causes errors in today's counters. With head motion, on the other hand, the peaks are cleaner and thereby lends itself to matching against a walking template. *Of course, there is no global template since humans walks with different step lengths, frequencies, and sways. Nonetheless, these variations can be modeled as a compression or expansion of a simple walking template.* This protect the step-counter from detecting spurious peaks and noisy false positives, while being robust to variations in walking patterns. To this end, the measured acceleration is first normalized based on the amplitude, and then matched with the template using a dynamic time warping (DTW) algorithm. DTW accommodates the variations of walks while detecting the peaks quite accurately. Importantly, no training is necessary.

■ **Anomaly (Step or Not) Filter and Counter:** We set a threshold on the DTW score to determine whether a potential step is an actual step or not. The threshold is set as two-thirds the standard deviation of all the DTW scores from a recent period of walking. This is of course a heuristic, however, results show that it adapts to changes in walking patterns. If one intends to reduce false positives (but can tolerate false negatives), the threshold can be increased.

### Exercise Analytics: Jump Measurement

We focus on only one instance of exercise analysis: *jumping*. How high a user can jump is an important metric for various health checks [6]. Earphone IMUs offer an opportunity since it is fixed at the ear and does not jiggle/jitter like a phone in a pocket. Even smartwatches show some vibration since it is not tightly worn on the wrist, and the user would move her hands during the jump.

We propose to measure jump-height by observing that the accelerometer reads ZERO when the user is off the ground (see Figure 6). This is because the accelerometer actually measures the reactive force on the body, which is 0 after the feet leave the ground. This yields the exact start and end time of the jump, and applying basic physics and kinematics during this time window, we expect to calculate the height. This avoids the need to do double integrate the accelerometer (which results in a poor estimate due to noise integration).

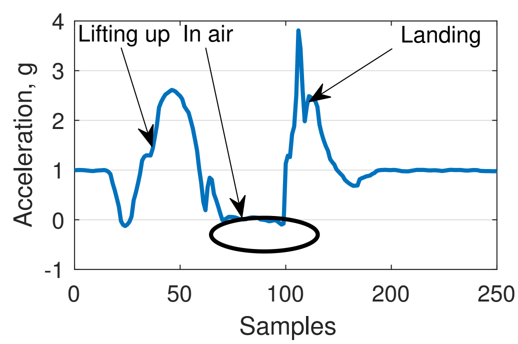


**Figure 6: Jumping analysis: 3 stages. Accelerometer reading is 0 when the user is in the air.**

■ **Jump Height Calculation:** We detect jumping by setting a threshold $\epsilon$, slightly above zero. When the accelerometer reading is $< \epsilon$, we believe the user is in air. Denoting $t$ as the total time in air, and the rise/fall movement is symmetric in time (i.e., $t/2$), we calculate the jump height simply as $h = 1/8gt^2$, where $g$ is known acceleration due to gravity.

## 4 IMPLEMENTATION AND EVALUATION

*STEAR* is built using onboard IMU of an Esense [4] earbud from Nokia. The IMU data stream, sampled at 100 Hz, is recorded through an Android application, nRF Connect by Nordic Semiconductor [7], and piped to MATLAB. The smartphone baseline result is obtained from OnePlus 3T smartphones running 3 most popular step counting apps, downloaded from the Android Play Store.

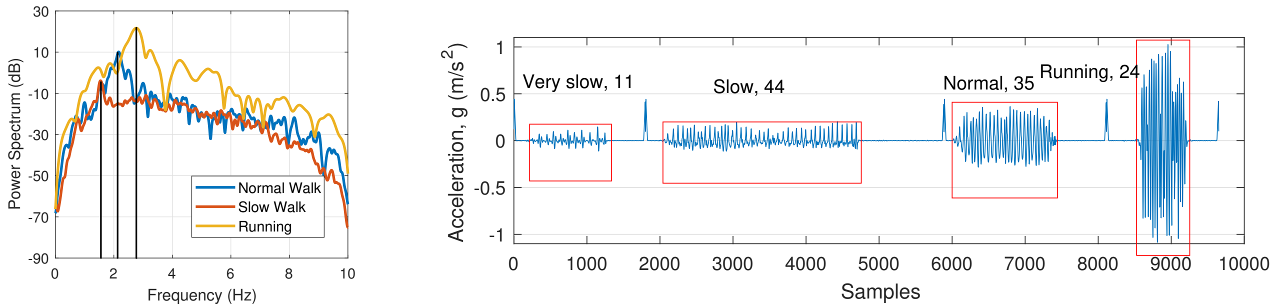


**Figure 7: (a) Acceleration spectrum across walking modes. (b) Time domain, filtered, signals for walking modes for same user.**

## Step Count

Four sets of experiments were performed with the participants wearing Esense on the ear and carrying smartphones in their pockets and hands. Participants were requested to walk in 4 modes: (a) very slow, (b) slow, (c) normal and (d) run. Figure 7(a) shows that the earbud captures the shift in the peak frequency with increasing walking speed. Figure 7(b) also visualizes the time-domain signals, showing the variations possible for the same person.

■ **Very slow walk:** It is crucial to be able to track very slow steps for assisting medical recovery in patients, older adults, and for slow walks inside houses or in beaches. Many prior work [2]-[5] report that traditional equipment and applications are inaccurate for adults walking at a speed of <0.9 m/s. Figure 8 verifies this result – smartphone miss most of the instances of slow walks since the peaks are buried under noise and random limb motions (since the peak amplitudes are weak for slow walks). With *STEAR*, due to of high SNR and nearly sinusoidal observations at earbuds, we are able to achieve higher accuracy in step-count. Figure 8 shows the comparison. We tracked 3 users moving very slowly and taking small steps, emulating walking old adult or patients. *STEAR* achieves accuracy >97% – 49 out of 50 steps – while smartphones under-counts as 5/50.

■ **Slow walk:** Slow walk corresponds to small and low-speed steps which we usually take while moving inside the home or while talking to a friend or while moving in groups. We tracked 3 users moving slowly and taking small steps, emulating such slow walks. We are able to achieve an accuracy of >98% on average. Tracking such steps is important for applications like indoor localization [9], where GPS like capabilities are unavailable, and pedestrian dead-reckoning is a candidate solution.

■ **Normal walk:** A typical walk produces promising SNR which reflects in near-perfect, >99%, accuracy. Of course, this is the reason phones and watches perform quite well since the peaks from each step rise above the noise floor. Several apps are tuned to identify this mode of walking.

■ **Running:** *STEAR* is able to count running steps with high accuracy as well. Figure 7(b) shows that there exists an opportunity of exploitation here, in the form of counting number of $-g + \epsilon$ acceleration points. In other words, when the runner's legs both are off the ground, the IMU accelerations shows an instantaneous ZERO measurement, which precisely counts the number of steps. Thus, large random hardware noise is the only reason to mis-count steps during running.

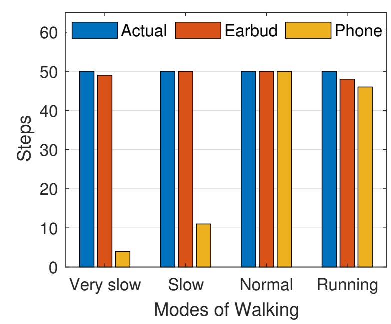


**Figure 8: Steps counts under different walking modes**

We observe that when a user has a phone in his hand and walks with a typical gait, both earbuds and phones perform equally well, as shown in Figure 9, and achieve near-perfect accuracy. But walking with the phone in pockets – both left-pocket (LP) and right-pocket (RP) – or playing with it, leads to over-estimation of step-counts.

■ **Jumping:** Medical practitioners suggest numerous styles of jumping for recoveries [6]. As shown previously in Figure 6, our evaluation suggests that we can exploit models of jump to find out instants of (a) rising up for the jump (smaller peak), (b) landing back on the ground (larger peak) and (c) time spent in the air (based on zero acceleration). Naturally, it is possible to count number of jumps with near-perfect accuracy. Also, the height of a jump, given by $gt^2/8$, can be calculated from the length of the horizontal line between

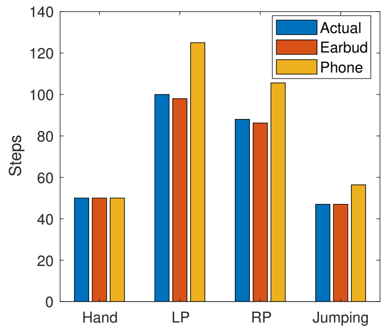


**Figure 9: Steps counts under different scenarios**

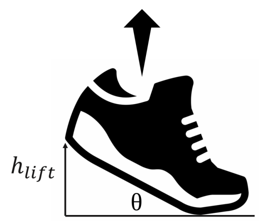


**Figure 10: Jump height compensation as per shoe lift**

**Table 1: Jump height estimation**

| Trials | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Actual height (cm) | 21 | 20 | 17 | 19 | 20 |
| Calculated height (cm) | 20 | 21 | 15 | 17 | 18 |

rising and landing. This unique opportunity, a long run of 0 acceleration, can be exploited to identify jump from a series of activities as well. But we needed to compensate for, $h_{lift}$, the height of the heel when toe lifts off from the ground because acceleration goes to 0 only when the whole body is completely in air (see Figure 10). *STEAR* measures jump height with an error of $+/-2$cm while the count of jumps is accurate to the ground truth, as presented in Table 1.

## 5 RELATED WORK

**Wearables based step counter:** Step counters are nothing new and have been implemented on a lot of mobile/wearable devices. The technique is not difficult, but it can suffer from errors. According to a recent measurement study [1], there is an 18.48% error in step counting over a 24-hour free-living period. Our work with an accuracy of > 95% provides a new opportunity to do better step counting in daily life.

**Jump analytics using inertial sensors:** There are an abundance of work that tracks various kinds of human motion, [6], MUSE [8] and SensorTape [3], to name a few. The closest to our work is [6], where they also calculate jump height using IMUs. However, the use of double integration to get vertical displacement, is subjected to noise. Our method, which explores the opportunity where acceleration equals to zero, is much more robust.

## 6 CONCLUSION

This paper shows the promise of robust step counting through ear mounted IMUs in modern earphones. The core opportunity emerges from the observation that the human body serves as a natural "filter", eliminating the noisy movements and only allowing certain walk-related vibrations to propagate up to the ear. We conjecture that this is due to the anatomical structure of the body – the joints in the skeletons and the muscles and tissues – which absorbs higher frequency movements, however, understanding the reasons for such filtering is a topic of study in another field (e.g., kinesiology). We benefit from this natural opportunity by demonstrating that physio-analytics can be improved with earables, not only in robustly counting steps but also in measuring the height of human jumps. Our ongoing investigation is focussed on further improvements to sensing these actions, and exploring other unique motion-related opportunities from earable IMUs.

## REFERENCES


[1] Hyun-Sung An, Gregory C Jones, Seoung-Ki Kang, Gregory J Welk, and Jung-Min Lee. 2017. How valid are wearable physical activity trackers for measuring steps? *European journal of sport science* 17, 3 (2017), 360–368.

[2] David R Bassett and Dinesh John. 2010. Use of pedometers and accelerometers in clinical populations: validity and reliability issues. *Physical therapy reviews* 15, 3 (2010), 135–142.

[3] Artem Dementyev, Hsin-Liu Cindy Kao, and Joseph A Paradiso. 2015. Sensortape: Modular and programmable 3d-aware dense sensor network on a tape. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, 649–658.

[4] Fahim Kawsar, Chulhong Min, Akhil Mathur, and Allesandro Montanari. 2018. Earables for Personal-Scale Behavior Analytics. *IEEE Pervasive Computing* 17, 3 (2018), 83–89.

[5] Jessica B Martin, Katarina M Krc, Emily A Mitchell, Janice J Eng, and Jeremy W Noble. 2012. Pedometer accuracy in slow-walking older adults. *International journal of therapy and rehabilitation* 19, 7 (2012), 387–393.

[6] Bojan Milosevic and Elisabetta Farella. 2015. Wearable Inertial Sensor for Jump Performance Analysis. In *Proceedings of the 2015 Workshop on Wearable Systems and Applications (WearSys '15)*. ACM, New York, NY, USA, 15–20. DOI:https://doi.org/10.1145/2753509.2753512

[7] Nordic Semiconductor. 2019. nRF Connect for Mobile. (Aug 2019). Retrieved Aug 12, 2019 from https://play.google.com/store/apps/details?id=no.nordicsemi.android.mcp&hl=en$_{U}$S

[8] Sheng Shen, Mahanth Gowda, and Romit Roy Choudhury. 2018. Closing the gaps in inertial motion tracking. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. ACM, 429–444.

[9] He Wang, Souvik Sen, Alexander Mariakakis, Romit Roy Choudhury, Ahmed Elgohary, Moustafa Farid, and Moustafa Youssef. 2012. Unsupervised indoor localization. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 499–500.