Demo Abstract: Deskbuddy: an Office Activity Detection System

Amelie Bonde
Carnegie Mellon University
Moffett Field, California
amelie@cmu.edu

Shijia Pan
Carnegie Mellon University
Pittsburgh, Pennsylvania
shijiapan@cmu.edu

Hae Young Noh
Carnegie Mellon University
Pittsburgh, Pennsylvania
noh@cmu.edu

Pei Zhang
Carnegie Mellon University
Moffett Field, California
peizhang@cmu.edu

ABSTRACT

We present Deskbuddy, a vibration-based system that can track a user's activities through their desk. Tracking sitting and other office related activities let us remind the user to have healthier working habits, as well as giving information about how office spaces are used. Many solutions have been proposed for office activity tracking, but they either require the user to wear a device, or they use cameras or microphones, which can make subjects uncomfortable. Our demo includes a small vibration sensor that sits on a table that can detect four office related activities. We capture the signal from the vibration sensor, extract features, and perform classification on the resulting features. The full functionality of the system will be shown in a video. In order for our demo to be more effective in a crowded environment, we have re-trained it to detect only typing versus not typing.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**; *Ubiquitous and mobile computing design and evaluation methods.*

KEYWORDS

structural vibration, office activity detection, cyberphysical systems

ACM Reference Format:
Amelie Bonde, Shijia Pan, Hae Young Noh, and Pei Zhang. 2019. Demo Abstract: Deskbuddy: an Office Activity Detection System. In *The 18th International Conference on Information Processing in Sensor Networks (co-located with CPS-IoT Week 2019) (IPSN '19), April 16–18, 2019, Montreal, QC, Canada.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3302506.3312490

1 INTRODUCTION

Too much sitting can increase risk of diabetes and cardiovascular risk, however, this risk can be mitigated if people periodically stand up or walk around[2]. Monitoring or detecting office activities, such as typing, talking, walking and sitting, allows for personalized reminders to take breaks. For example, if we can tell when they are typing or talking, we can make sure not to interrupt them in such periods. And when they are not in critical events, the system can remind them to rest regularly.

Existing research includes several systems to do activity detection in office environments, using various sensors. Mekruksavanich et. al. use smartwatches to detect sitting[5], and Rasheed et. al. use the accelerometers in a mobile phone to detect sitting, standing, and falling[9]. However, they require the user to wear or carry a device all the time, which may not be practical. Sung et. al use RGBD sensors to detect office environment activities, while Wojek et. al. use cameras and microphones[10][11]. Park et. al. use illumination sensors, microphones, and passive infrared sensors to detect various office activities, although they do not specifically detect sitting[8]. However, being monitored by cameras and microphones can be intrusive, and may make some users uncomfortable.

We present Deskbuddy, an activity recognition system utilizing desk vibration induced by human activity, including sitting, typing, speaking, and walking around. Deskbuddy uses a vibration sensor to capture the structural vibration caused by a person's activity. It then extracts features from the vibration signal and applies a classifier with heuristics, which will be described in the next section. As a demonstration of Deskbuddy's capabilities, we will present a video of Deskbuddy's full functionality, and include a desk-based typing detector that can work in a crowded, noisy environment.

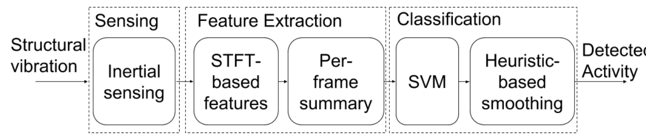

Figure 1: System Diagram

2 SYSTEM DESIGN

Deskbuddy consists of three modules, as shown in Figure 1. The Sensing Module acquires the desk vibration signal with a geophone placed on a desk (Section 2.1), as shown in Figure 2. Then, the Feature Extraction Module applies a short term Fourier transform (STFT) to get the power spectral density and extract features (Section 2.2). Finally, the Activity Classification Module classifies the activity through Support Vector Machine (SVM) and smooths noise in the result through heuristic rules (Section 2.3)).

IPSN '19, April 16–18, 2019, Montreal, QC, Canada

ACM ISBN 978-1-4503-6284-9/19/04.
https://doi.org/10.1145/3302506.3312490

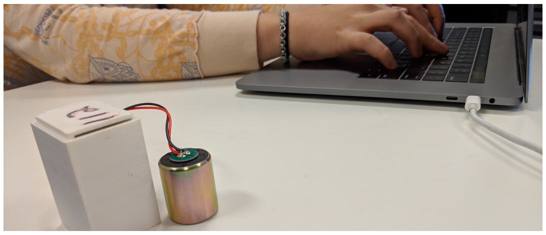

Figure 2: The sensor module

2.1 Vibration Sensing Module

Our sensing module relies on a geophone sensor, depicted in Figure 2. Geophones are widely used for seismic sensing and structural health monitoring because of their sensitivity to surface movement [7]. Its response curve shows a stable frequency response up to about 1000Hz, and its sensitivity to tiny motions fits our target sensing range (10^{-4} to 10^{-3}m/s) [1, 3].

We divide the detected signal into one second windows because one second is long enough to detect vibrations from human movement and speech, while being short enough for fine-grained detection of activity such as footsteps or typing strikes.

2.2 Feature Extraction Preserving Quasi-Stationary Data

The Feature Extraction Module extracts the spectrogram of the windowed signal for further analysis. In order to get a frequency response that applies equally to the whole window, we chose small windows of 50 ms, determined to preserve the quasi-stationarity of the data. [6]. Then we calculate the Discrete Fourier Transform (DFT) for each small window and take the squared magnitude of that to get the spectrogram. We observed that frequency information up to 800 Hz encompassed the human induced structural vibrations we empirically observed in different structures, and is within the frequency response range of our sensor. The frequency information is divided into bands for each small window and then summarized across time using the mean, maximum and minimum values.

This gives us a set of 30 features for our Activity Classification Module, which intuitively represents the frequency information of the structural vibration from human motion, aggregated across one second intervals.

2.3 Activity Classification Module

For the Activity Classification Module, we use a support vector machine (SVM) with a linear kernel. We trained the SVM with 40 minutes of data, about 10 minutes (600 1 second windows) from each class (sitting, typing, talking and walking).

We then use a heuristic-based smoothing algorithm on the data. The heuristic rule applied is to segment talking or typing activities with brief pauses (i.e., less than 10s) as a single episode [4]. This allows *Deskbuddy* to recognize the duration of each activity along with its type.

3 RESULTS AND DEMO

We tested our system on a half hour of data from the same subject we trained on, who was asked to switch between the four activities every minute or so. We acheived an overall classifcation accuracy of 94%, with typing having the lowest accuracy at 87% correctly classified (it was often misclassified as sitting or walking). This could be because the activity of typing includes sitting since the subject sat while they typed. Typing and walking are both impulsive activities, so they may have some similarities that at times caused misclassification. Our video demo shows the real-time detecting and classification of four office related activities (typing, talking, walking and sitting) with *Deskbuddy* placed on the desk. We will also provide a live demo at the conference demo venue. Due to the noisy vibration condition at the venue, we plan to focus on demonstrating the typing detection. The participants can test *Deskbuddy* by typing on the desk and our *Deskbuddy* will be able to present in real-time if the participant is typing, even at different locations of the desk, with different individuals.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation Graduate Research Fellowship Program (under Grant No. DGE 1745016), Carnegie Mellon University, and CMKL University. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of Carnegie Mellon University, the National Science Foundation, or the U.S. Government or any of its agencies.

REFERENCES

[1] Amelie Bonde, Shijia Pan, Zhenhua Jia, Yanyong Zhang, Hae Young Noh, and Pei Zhang. 2018. VVRRM: Vehicular Vibration-Based Heart RR-Interval Monitoring System. In *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications*. ACM, 37–42.

[2] David W Dunstan, Bethany Howard, Genevieve N Healy, and Neville Owen. 2012. Too much sitting–a health hazard. *Diabetes research and clinical practice* 97, 3 (2012), 368–376.

[3] Jonathon Fagert, MOSTAFA MIRSHEKARI, SHIJIA PAN, PEI ZHANG, and HAE YOUNG NOH. 2017. Monitoring Hand-Washing Practices using Structural Vibrations. *Structural Health Monitoring 2017* shm (2017).

[4] Kristina Lundholm Fors. 2015. Production and Perception of Pauses in Speech. (2015).

[5] Sakorn Mekruksavanich, Narit Hnoohom, and Anuchit Jitpattanakul. 2018. Smartwatch-based sitting detection with human activity recognition for office workers syndrome. In *Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON), 2018 International ECTI Northern Section Conference on*. IEEE, 160–164.

[6] Kuldip Paliwal and Kamil Wójcicki. 2008. Effect of analysis window duration on speech intelligibility. *IEEE signal processing letters* 15 (2008), 785–788.

[7] Shijia Pan, Tong Yu, Mostafa Mirshekari, Jonathon Fagert, Amelie Bonde, Ole J Mengshoel, Hae Young Noh, and Pei Zhang. 2017. Footprintid: Indoor pedestrian identification through ambient structural vibration sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 89.

[8] Homin Park, Jongjun Park, Hyunhak Kim, Jongarm Jun, Sang Hyuk Son, Taejoon Park, and JeongGil Ko. 2015. ReLiSCE: utilizing resource-limited sensors for office activity context extraction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45, 8 (2015), 1151–1164.

[9] Muhammad Babar Rasheed, Nadeem Javaid, Turki Ali Alghamdi, Sana Mukhtar, Umar Qasim, Zahoor Ali Khan, and M Haris Baidar Raja. 2015. Evaluation of human activity recognition and fall detection using android phone. In *Advanced Information Networking and Applications (AINA), 2015 IEEE 29th International Conference on*. IEEE, 163–170.

[10] Jaeyong Sung, Colin Ponce, Bart Selman, and Ashutosh Saxena. 2012. Unstructured human activity detection from rgbd images. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 842–849.

[11] Christian Wojek, Kai Nickel, and Rainer Stiefelhagen. 2006. Activity recognition and room-level tracking in an office environment. In *Multisensor Fusion and Integration for Intelligent Systems, 2006 IEEE International Conference on*. IEEE, 25–30.