Tobias Grosse-Puppendahl, Kiesstr. 24, 64283 Darmstadt, Germany
Telephone: +49 171 6506806, E-Mail: tobias@grosse-puppendahl.com

Master Thesis Summary - Jakob Karolus:
Detecting Falls with Consumer-Grade Smartphones using Active and Passive Ultrasound Sensing

Introduction

Sensor miniaturisation and streaming classification techniques can be used to recognize human behaviours and contexts. This is extremely valuable to realize smart environments, e.g. to support healthy and independent living. The most important parameters to sense include indoor location, gestures, or emergencies like falls. Up to now, activity recognition systems face a number of sensitive drawbacks. For example, camera-based systems induce privacy issues and are costly to deploy. Body-worn systems are inconvenient to wear over long periods of time. Highly visible systems may introduce social stigma and modify the well-known living environment. In this project, we lay the foundations for the use of a new, unobtrusive, physical principle to sense and recognize human activities.

Our novel approach leverages the properties of ultrasound to estimate the surrounding environment. In particular we aim to provide a method to detect emergency like falls by analyzing the frequency distribution of ultra-sonic sound waves emitted and recorded by the user's phone. In this scenario, a consumer-grade smartphone could potentially nullify the need for additional hardware equipment and provide the consumer with a simple to execute solution to detect emergencies in which the person might not be capable to contact any assistance. Out of necessity, we require a static setup for our technique to work. Because of this, the main application scenario corresponds to times where the phone is in such a static position, e.g. during night time.

Related Work

Acoustic Sensing

In the last decade, acoustic sensing has proven to be a very unobtrusive technology for sensing interactions within a human's environment. In contrast to ultrasound, acoustic sensing does not rely on actively transmitted signals. Although speech recognition is a well established research field, acoustic sensing has not yet pervaded other areas on a wide scale. The sensing method is based on the fact that almost every interaction with the environment leaves acoustic traces. This ranges from tapping onto a button, moving along a surface, or interacting with devices such as water taps.

Solid materials are very well-suited for transporting acoustic noise along their surface, making them an ideal candidate for passive acoustic sensing of humans and their environment. Choosing a suitable technology for sound-pickups is not straight-forward, developers face challenges in frequency responses and coupling to materials that strongly depend on the use case. Techniques range from using microphones, microphones combined with stethoscopes, piezo-electric vibration elements, to accelerometers (MEMS) [10].

Sensing interactions with a wide variety of materials has been applied for sensing hand movements and touches. Even when using walls as a material for conducting sounds, ranges up to 8 m from touching the surface to the microphone can be achieved [9]. In order to achieve these reasonable high resolutions, a common technique is to use stethoscopes with an attached microphone. When scratching along a surface with a finger nail, sound frequencies up to 3 KHz are generated and conducted within the material [9]. Using this approach, finger and hand movements along walls, tables, screens, and even fabric can be recognized.

In order to detect whole-body movements, one can distinguish between body-mounted and stationary acoustic sensors. In the domain of body-mounted sensors, multiple physical parameters can be captured using sound. For example, using the sound produced by muscle fibers enables Yamakawa et al. to detect different finger movements [24]. In contrast to muscles, conducting sounds through bone also allows for recognizing bending of a human elbow [21]. Chewing, eating and drinking activities can be captured with microphones mounted inside a human ear [2] or a stethoscope placed near the throat [26].

Popescu et al. employ acoustic sensing to detect fall situations [16]. Based on multiple microphones, the height of the sound-source is estimated. If the height is below a threshold of approximately 60 cm, an

Tobias Grosse-Puppendahl, Kiesstr. 24, 64283 Darmstadt, Germany
Telephone: +49 171 6506806, E-Mail: tobias@grosse-puppendahl.com

emergency situation can be derived. However, the authors state that the number of false positives is still too high for real-world deployments. Sound-source estimation can also be used for person localization by microphone arrays [5]. Such systems provide an accuracy of approximately 1 m, while being able to detect multiple sources of sound.

Ultrasound Sensing

Not all interactions leave such nicely analyzable traces as in acoustic sensing. When it comes to in-the-air interactions or detecting stationary states, passively picking up acoustic noise is not sufficient anymore. In contrast to acoustic sensing, actively emitting non-hearable sounds can overcome the problem of recognizing stationary objects. Here, sounds are emitted and the reflected signal is measured by one or multiple microphones. In ultrasound sensing, one has to distinguish between active messaging nodes and passive backscattering techniques. The latter ones analyze backscattered signals to infer the position of body parts or the location of a person [8]. Signals from other devices received and transmitted by nodes allows for implementing an active messaging system.

Detecting finger and hand movements in free air is often achieved by analyzing a backscattered ultrasound signal. Here, worn nodes that use ultrasound as a messaging system can rarely be found. [12] apply ultrasonic waves to unobtrusively recognize one-handed gestures. The authors employ the Doppler effect which introduces frequency shifts in the backscattered signal of a moving human body part. A single transmitter and three receiver microphones are sufficient to determine a 3D movement. Due to the availability of ultrasound capabilities in consumer hardware, ultrasound approaches have also been ported to consumer smartphones and laptops. For example, SoundWave makes it possible to recognize gestures in front of the screen in ordinary laptops [8].

In order to detect whole-body movements, worn nodes as well as passively backscattered signals can be used. Multiple microphones allow for ranging and localization based on the angle or time-difference of arrival. Moreover, exploiting physical effects like the Doppler effect helps to detect changes within the environment. This enables to recognize body parts moving away or towards a sound pickup [8]. Tarzia et al. [22] use a similar technique to determine user presence near a consumer laptop. Using a measurement window of only 10 s results in an accuracy of approximately 96 %, discriminating the two classes of absence and presence. Extending the windows to 25 s almost exceeds perfect accuracy.

Attaching ultrasound nodes to a human body represents a very convenient way to identify and track a person. Techniques which use simple backscattering are not able to identify objects and thus leave a certain amount of ambiguity in their results. Randell et al. present an active ultrasonic positioning system that combines RF and ultrasound for object and person localization [18]. Relying on a secondary modality with faster wave propagation enables to reliably calculate the time-of-flight of the ultrasound signal [11, 20]. This technique has found wide applications in people localization and the corresponding tags became very small or integrated into devices [3].

Transmitting dedicated information about the object and its manipulations requires an active messaging system. One of the first approaches combine radio frequencies with ultrasound to achieve a reference for time-of-flight measurements [17]. BeepBeep [15] solely relies on ultrasound generated by smartphones, that then supports ranging among objects. When more than three objects are involved, it is possible to determine a relative position based on the information about multiple devices. [14] follows a very similar approach: It enables users to wear a virtual uniform which emits sounds at specified frequencies. The system is intended to communicate the social state to other people. Sensing device location in a car was investigated by [25]: By classifying the mobile phone's position it is possible to differentiate between the driver using the phone and a passenger. Reynolds et al. [19] introduce an ultrasound position sensing system for tangible objects above LCD-screens. The authors present interactive 'pucks' that communicate by emitting and receiving ultrasound and reconstruct their position.

Fall recognition

Fall recognition is one of the hottest topics in Ambient Assistant Living, especially for elderly people. The fear of not being able to sustain oneself without the help of others is omnipresent [6]. Systems that are able to recognize emergency like falls and contact help can greatly boost one's living quality and keep the customers at ease.

Here at Fraunhofer IGD, we analyze the feasibility of recognizing falls via capacitive proximity sensing systems [7]. Compared to many current commercial systems, OpenCapSense is able to provide highspeed acquisition of proximity data necessary for fall detection as shown in case studies. This led to the development of CapFloor, a Fraunhofer industrial project, which allows for people localization and emergency detection using a capacitive floor. A grid of electrodes underneath the floor provides sensor values, which are analyzed by an embedded computer per apartment. The system is currently deployed in 23 apartments of an elderly home.

In a world full of sensors, capacitive sensing is not the only approach to solve the task of fall recognition. Other methods focus on evaluating sensor data from accelerometers, gyroscopes or use visual detection. However, some of the earlier systems including accelerometers or gyroscopes, require the user to wear a dedicated device, which is obstructive. Examples of these would include a waist-worn accelerometer by Zhang et al. [28], able to detect falls with 96.7% accuracy with an SVM algorithm, or a bi-axial gyroscope mounted on the torso utilized by Bourke and Lyons [4], which proved 100% successful in fall detection with a threshold-based algorithm.

Considering the application scenario and target audience, it is also highly probably that the user forgets to wear the device, rendering the system is useless. In this case, incorporating the device into everyday objects like phones seems more suitable, though this can make recognition more difficult, as the phone usually resides inside the owner's pocket. Zhang et al. [27] were able to achieve a mean ratio of correctness of 93.3% given six different fall categories using SVM pre-processing coupled with KFD (Kernel Fisher Discriminant) and kNN (Nearest Neighbor) algorithms for classification. In the work of Albert et al [1], they simulated four types of falls and evaluated five different machine learning classifiers on their dataset, which was recorded using the accelerometer of a mobile phone. SVM and logistic regression were able to identify a fall with 98% accuracy as well as classify its type with 99% accuracy.

The third area deals with the visual detection of a person, for instance in a video frame. Vishwakarma et al. [23] present a method to detect falls by extracting features from the bounding boxes of objects within a video. E.g. the angle between an object's bounding box and the ground can be used as an indicator whether we observe a fall or not. Together with a rule-based confirmation step, they managed to achieve 95% accuracy on single objects and 64% when there were multiple objects present. In a paper from Luštrek and Kaluža [13] radio tags are used to determine the location of body parts, thus enabling the authors to reconstruct the person's posture and movement. From the eight tested machine learning algorithms, the SVM showed the best results achieving over 95% classification accuracy.

As shown above, most of the presented systems use some kind of on-body device or require additional hardware (CapFloor). In this thesis, finding a way around additional sensors and carrying requirements is mandatory, since any wearables would be extremely obtrusive during night time. Therefore, ultrasound and/or acoustic sensing techniques are most suitable for this task.

Scientific Contribution

The scientific challenge is to understand and develop models of how human motion affects the frequency distribution of the recorded ultra-sonic sound waves. Based on these models, signal processing and machine learning techniques can be investigated to infer a user's activities and thereby recognizing emergency like falls during night time. The project will be conducted over a six-months period with the following scientific advancements:

1. Fall recognition utilizing the frequency distribution of recorded ultra-sonic waves.
2. Simply deployment and low-cost product for end-consumers via smartphone application.
3. Feasibility study of ultrasound frequency analysis on mobile devices.

Detailed Work Plan

The following points are expected to be completed in the scope of this work and thesis:

- Analysis of existing methods leveraging ultra-sonic sound waves and applications
- Implementation of an algorithm to analyze the frequency distribution

- Feature generation from a given analysis
- Utilizing Machine Learning techniques to create a prediction model which can recognize emergency like falls.
- Development of an Android app capable of emitting and recording ultra-sonic sound waves, as well as classifying these given the prediction model
- If needed: Webservice as a computational ressource for the system to fall back on, if the smartphone's hardware capabilities are not sufficient.
- Evaluation of different prediction models in the context of fall recognition.

References

[1] Mark V Albert, Konrad Kording, Megan Herrmann, and Arun Jayaraman. Fall classification by machine learning using mobile phones. *PloS one*, 7(5):e36556, 2012.

[2] Oliver Amft, Holger Junker, and Gerhard Tröster. Detection of eating and drinking arm gestures using inertial body-worn sensors. In *Proceedings of the 9th IEEE International Symposium on Wearable Computers (ISWC '05)*, pages 160–163. IEEE, 2005.

[3] Gaetano Borriello, Alan Liu, Tony Offer, Christopher Palistrant, and Richard Sharp. Walrus: Wireless acoustic location with room-level resolution using ultrasound. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services*, MobiSys '05, pages 191–203, New York, NY, USA, 2005. ACM.

[4] A.K. Bourke and G.M. Lyons. A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. *Medical Engineering and Physics*, 30(1):84–90, 2006.

[5] Ionut Constandache, Sharad Agarwal, Ivan Tashev, and Romit Roy Choudhury. Daredevil: Indoor location using sound. *SIGMOBILE Mob. Comput. Commun. Rev.*, 18(2):9–19, June 2014.

[6] Bundesarbeitsgemeinschaft der Senioren-Organisationen e.V. Ergebnisse einer befragung zum thema: Wohnen im alter. 2005.

[7] Tobias Grosse-Puppendahl, Yannick Berghoefer, Andreas Braun, Raphael Wimmer, and Arjan Kuijper. OpenCapSense: A Rapid Prototyping Toolkit for Pervasive Interaction using Capacitive Sensing. In *PerCom '13*, pages 152–159, 2013.

[8] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. Soundwave: Using the doppler effect to sense gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1911–1914, New York, NY, USA, 2012. ACM.

[9] Chris Harrison and Scott E. Hudson. Scratch input: Creating large, inexpensive, unpowered and mobile finger input surfaces. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, UIST '08, pages 205–208, New York, NY, USA, 2008. ACM.

[10] Chris Harrison, Desney Tan, and Dan Morris. Skinput: Appropriating the body as an input surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 453–462, New York, NY, USA, 2010. ACM.

[11] S. Helal, B. Winkler, Choonhwa Lee, Y. Kaddoura, L. Ran, C. Giraldo, S. Kuchibhotla, and W. Mann. Enabling location-aware pervasive computing applications for the elderly. In *Pervasive Computing and Communications, 2003. (PerCom 2003). Proceedings of the First IEEE International Conference on*, pages 531–536, March 2003.

[12] K. Kalgaonkar and B. Raj. One-handed gesture recognition using ultrasonic doppler sonar. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 1889–1892, April 2009.

[13] Mitja Luštrek and Boštjan Kaluža. Fall detection and activity recognition with machine learning. *Informatica*, 33(2):197–204, 2009.

[14] Aleksandar Matic, Alban Maxhuni, Venet Osmani, and Oscar Mayora. Virtual uniforms: Using sound frequencies for grouping individuals. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 159–162, New York, NY, USA, 2013. ACM.

[15] Chunyi Peng, Guobin Shen, and Yongguang Zhang. Beepbeep: A high-accuracy acoustic-based system for ranging and localization using cots devices. *ACM Trans. Embed. Comput. Syst.*, 11(1):4:1–4:29, April 2012.

[16] M. Popescu, Yun Li, M. Skubic, and M. Rantz. An acoustic fall detector system that uses sound height information to reduce the false alarm rate. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4628–4631, Aug 2008.

[17] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, MobiCom '00, pages 32–43, New York, NY, USA, 2000. ACM.

[18] Cliff Randell and Henk L. Muller. Low cost indoor positioning system. In *Proceedings of the 3rd International Conference on Ubiquitous Computing*, UbiComp '01, pages 42–48, London, UK, UK, 2001. Springer-Verlag.

[19] Matthew Reynolds, Alexandra Mazalek, and Glorianna Davenport. An acoustic position sensing system for large scale interactive displays. In *Sensors, 2007 IEEE*, pages 1193–1196, Oct 2007.

[20] Andreas Savvides, Chih-Chieh Han, and Mani B. Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, MobiCom '01, pages 166–179, New York, NY, USA, 2001. ACM.

[21] Kentaro Takemura, Akihiro Ito, Jun Takamatsu, and Tsukasa Ogasawara. Active bone-conducted sound sensing for wearable interfaces. In *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology*, UIST '11 Adjunct, pages 53–54, New York, NY, USA, 2011. ACM.

[22] Stephen P. Tarzia, Robert P. Dick, Peter A. Dinda, and Gokhan Memik. Sonar-based measurement of user presence and attention. In *Proceedings of the 11th International Conference on Ubiquitous Computing*, Ubicomp '09, pages 89–92, New York, NY, USA, 2009. ACM.

[23] Vinay Vishwakarma, Chittaranjan Mandal, and Shamik Sural. Automatic detection of human fall in video. In Ashish Ghosh, RajatK. De, and SankarK. Pal, editors, *Pattern Recognition and Machine Intelligence*, volume 4815 of *Lecture Notes in Computer Science*, pages 616–623. Springer Berlin Heidelberg, 2007.

[24] Shumpei Yamakawa and Takuya Nojima. A proposal for a mmg-based hand gesture recognition method. In *Adjunct Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST Adjunct Proceedings '12, pages 89–90, New York, NY, USA, 2012. ACM.

[25] Jie Yang, Simon Sidhom, Gayathri Chandrasekaran, Tam Vu, Hongbo Liu, Nicolae Cecan, Yingying Chen, Marco Gruteser, and Richard P. Martin. Detecting driver phone use leveraging car speakers. In *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking*, MobiCom '11, pages 97–108, New York, NY, USA, 2011. ACM.

[26] Koji Yatani and Khai N. Truong. Bodyscope: A wearable acoustic sensor for activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 341–350, New York, NY, USA, 2012. ACM.

[27] Tong Zhang, Jue Wang, Ping Liu, and Jing Hou. Fall detection by embedding an accelerometer in cellphone and using kfd algorithm. *International Journal of Computer Science and Network Security*, 6(10):277–284, 2006.

[28] Tong Zhang, Jue Wang, Liang Xu, and Ping Liu. Fall detection by wearable sensor and one-class svm algorithm. In De-Shuang Huang, Kang Li, and GeorgeWilliam Irwin, editors, *Intelligent Computing in Signal Processing and Pattern Recognition*, volume 345 of *Lecture Notes in Control and Information Sciences*, pages 858–863. Springer Berlin Heidelberg, 2006.