Smart Environments with Wide-Area Sensors

Yang Zhang

CMU-HCII-20-111
November 2020

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, USA 15213

Thesis Committee:
Chris Harrison (Chair)
Lining Yao
Scott E. Hudson
Gregory D. Abowd (Georgia Institute of Technology)

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Keywords: Ubiquitous Computing, Wide-Area Sensing, Internet of Things (IoT), Smart Environment.

Abstract

We live in an era where computer applications are no longer constrained by the proximity of big and expensive computers, and can extend to larger common spaces due to the ubiquity of sensors, electronics, and their ever-decreasing sizes and costs. This computing evolution has fueled the emerging concept of smart environments with applications such as context-aware computing, building and personal informatics, mental and physical health monitoring, and accessibility for the elderly and handicapped. However, smart environments are only as smart as what they can sense – the key to realizing this future is the development of accurate, reliable, and versatile activity sensing technology.

In this dissertation work, I identify the critical challenges faced in current activity sensing and propose a new direction for tackling these challenges. Specifically, instead of deploying many sensors throughout user environments, I propose a new sensing technique that involves the use of fewer but more powerful sensors than the existing methods. I call these sensors wide-area sensors. I built five systems based on capacitive sensing, radio frequency sensing, energy harvesting, and laser vibrometry. These systems achieve room-, building- and city-scale sensing by adapting everyday objects for sensing using low-cost instrumentation in concert with signals that can travel long distances. Additionally, I have conducted a series of background investigations and system performance evaluations to prove that such wide-area sensing systems can be low-cost, low-maintenance, and general-purpose while being able to sense rich signals. Finally, I summarize the contribution of this thesis and propose several future research efforts.

Acknowledgments

Foremost, I want to thank my advisor, Chris Harrison, for being a great advisor, who is always enthusiastic and caring for his students. I also want to thank my committee members, Lining Yao, Scott Hudson, and Gregory Abowd, for their feedback and inspirations. To my lab mates, Robert Xiao, Gierad Laput, Karan Ahuja, and Yasha Iravantchi – our hours working together in the lab driving projects to finish lines will be one of the most memorable and enjoyable experiences in my life. I am also indebted to all collaborators for without whom this thesis would not have been possible.

I also want to thank Queenie Kravitz and the entire HCII family for their continuous help and support. I am grateful to Alanson Sample and Ken Hinckley, who are my advisors during my internships. They have provided me a different aspect of my PhD study, and much great guidance, which will benefit my entire career. I am also grateful to my industry collaborators – Hrvoje Benko, Brian Tsang, and many others, from whom I have learned a lot about applying research in solving real-world problems.

Finally, I want to thank my parents for being my role models and raising me up to have a light heart. Their selfless love and support have fueled me to go through the ups and downs during life. Last but not least, my wife Zhuoshu (Dr. Li), who is my smarter half, has been a great source of motivation for me to become a better person and to live my life without regrets.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

In his well-known article "The Computer for the 21st Century" [155], Mark Weiser depicted the future of computing:

The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.

This pioneering vision inspired a wave of research that extends computer applications into common physical spaces, no longer constrained within the proximity of primary computing devices. Among these efforts, there is a significant body of research on smart environment applications (e.g., personal informatics and digital health), as well as context-aware computing for natural interactions where user attention is no longer required for every single computer action. These smart environments rely on accurate and robust sensing techniques for user activities and events in physical space.

Nonetheless, several key challenges remain in existing ubiquitous sensing techniques (e.g., high material and maintenance cost), preventing them from being widely adopted across society [50, 56, 99].

This thesis aims to tackle these challenges by proposing wide-area sensing, wherein one sensor is powerful enough to cover a wide area for practical sensing applications. Wide-area sensors can be cheaply and easily deployed at little monetary and maintenance costs while being general-purpose.

1.2 Activity Sensing in Smart Environments

In smart environments, activity sensing technologies digitize user activity and events in physical spaces to provide fundamental information to computing applications that are intended to improve user's productivity, security, and lifestyle. Examples of such activity sensing include environmental indicators (e.g., noise and traffic), personal informatics (e.g., exercise and food), industry surveillance (e.g., inventory management, infrastructure heath), and data-driven agriculture (e.g., precision farming).

Homeowners have been increasingly interested in upgrading their domestic environments with smart home technologies. According to GSMA Intelligence [2], smart home devices have become the biggest part of the consumer side of IoT in 2020 and will have more than 4 billion devices by 2025. Additionally, recent years have seen many successful commercial products such as Apple HomePod, Google Home, and Amazon Alexa. According to Canalys [1], in 2018 alone, 78 million smart speakers were shipped worldwide, which is more than double the amount shipped in the previous year. These increasingly popular smart home devices provide promising sensing platforms. In a perfect world, smart home devices could remind users when they drink too much coffee or watch too much digital content at night to promote sleep hygiene, spot anomalies in environments (e.g., water leakage, power failure) for home safety, log cooking, working, and exercising for users so that they can make better lifestyle decisions, which could improve their health and safety.

This thesis focuses on sensing user activities and events in common physical spaces to create sensory feeds as important building blocks for applications such as personal and environmental informatics, mental and physical health monitoring, context-aware computing, home security, and accessibility. There are also potential applications in industrial and agricultural settings where this thesis could be relevant.

1.3 Conventional Approaches and Key Challenges

All the aforementioned applications require activity sensing techniques as the fundamental building block. Let us use a common kitchen in Figure 1.1 as a concrete example. In this kitchen, a homeowner may have many questions: "did I forget to turn off the stove", "when will this round of washing finish", "is the drain clogged", and so on. To solve the problem of how to answer these questions, one of the most intuitive solutions, and perhaps the only solution before computers, is to use human operators. However, manual operation is costly (i.e., butlers cost an average of around $57k a year in the United States in 2020 according to ZipRecruiter [8]) and is difficult to scale (i.e., there are easily hundreds of things to monitor around-the-clock, which are beyond the capability of a single human operator). Therefore, a preferable solution is to use autonomous sensor-driven approaches, which have long been explored in the research domain [33, 37, 145, 154] and have seen some recent commercial success [16, 17, 19, 24].

However, even today's so-called smart home devices are far from being smart. Thanks to the internet, these devices (e.g., smart speakers) become all-knowing of facts that can be easily fetched online, yet they are are incapable of sensing even the closest things around them. As a result, these devices know more about medieval literature than things going on in their owner's kitchen. I believe the reason is that there are several fundamental challenges in existing activity sensing technologies.

Cameras are some of the most common sensors used in activity sensing due to their ability to remotely acquire high resolution data (i.e., pictures and videos). However, this data comes at the cost of user privacy. It is difficult to control what information is exposed in a picture. For example, a homeowner might want to use a camera to detect if the door is locked, but this also exposes who that homeowner allows into their home. Although privacy issues are universal across sensors (e.g., even electricity meters can reveal the number of people staying overnight), cameras are

one of the most notorious because pictures and photos can easily be interpreted by human eyes. As a result, users often have privacy concerns about cameras in their environments [3, 5, 172]. Similarly, microphone-based systems also provoke privacy concerns [93, 117].

To mitigate privacy concerns, another common activity sensing approach relies on object-borne sensors, which have limited sensing ranges. Compared to cameras and microphones, these close-range sensors give users more control of what exactly will be sensed. For example, an IMU-based sensor tag on a door can only sense the door, and is blind to the rest of the user's environments. However, due to this "shortsightedness", such sensor tags require a dense deployment to have practical sensing coverage (e.g., the Smart Dust project [154]). Several key challenges lie in such dense deployments of sensors:

1. Cost: Cost has been a main limiting factor for smart environments, which is why smart home technology has been more common for homes from $3 million up [7]. While the cost for an individual sensor might be low, the accumulated cost for a whole-room or -building installation is high. Figure 1.1 shows 17 sensor locations. If each sensor costs $20 (the average cost for a Notion sensor [24]), then the kitchen installation alone would cost $340, and the entire house coverage would likely cost more than $2,000, which is not financially feasible for average homeowners.
2. User Maintenance: Distributed sensors rely on distributed power, for which the most common solution is battery. Unfortunately, batteries requires user maintenance, which is one of the sparsest resources in the 21st century. Like monetary cost, maintenance cost also constrains the scalability of activity sensing with dense sensor deployments. Imagine a building with hundreds if not thousands of battery-powered sensors. Exchanging their batteries would become impossible, not to mention the added frustration of remembering their locations.
3. Sensing Versatility: Previous sensing approaches commonly rely on a dense deployment of

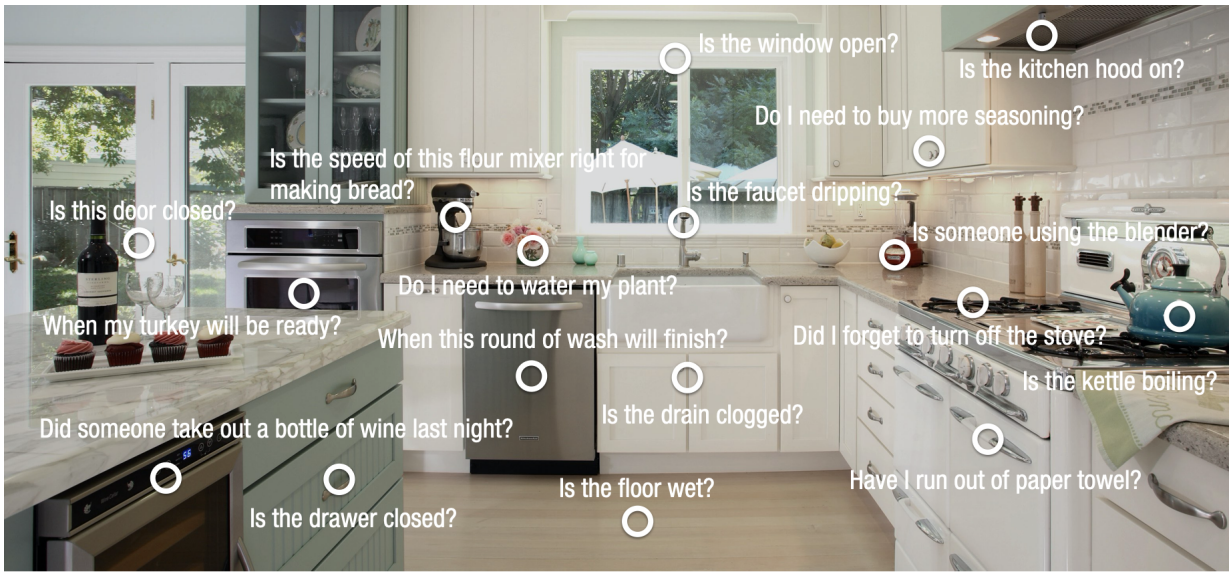


Figure 1.1: There might be many activities and events a user wants to know about in a kitchen.

special-purpose sensors. Each sensor is particularly designed for detecting one specific type of activity – e.g., water faucet sensors for sensing water flows, accelerometers for tracking moving objects, and electricity meters for monitoring electrical appliances. However, special-purpose sensors often demand user expertise, which makes it a challenge to retrofit them into existing environments. Additionally, being special-purpose means one sensor is not enough to detect multiple environmental facets. User environments might have a wide variety of activities and events, that no single sensor can detect in their entirety. Therefore, we must use a dense deployment of sensors to achieve effective coverage, which increases the challenges of reducing costs and maintenance.

1.4 Wide-Area Sensing

If we follow computing evolution (characterized by the ratio between users and computing devices they own [28]), we see a decreasing trend of user-to-computer ratio moving from many-to-one to one-to-many (Figure 1.2), thanks to smaller and cheaper computers. If we borrow this analogy and create a different metric – room-to-sensor ratio, which is the ratio between the number of rooms and sensors needed to make them smart, then it is natural to expect there to be a similar decreasing trend – from multiple rooms sharing one sensor to one room featuring multiple sensors. However, in order to overcome aforementioned key challenges in activity sensing, this thesis aims to shift this room-to-sensor ratio in the opposite direction, from multiple sensors per room, to one sensor per room, building floor, or even city block. This thesis argues that future smart environments will rely on fewer but more powerful sensors to sense activities and events throughout a wide area. The wide-area sensing approach I propose could be both much cheaper to scale, and considerably easier to maintain than conventional approaches.

There are several challenges involved in wide-area sensing. First, wide-area sensing requires sensors to have long sensing ranges, for example, several meters (room-scale), tens of meters (building-scale), and kilometers (city-scale). This is challenging because signals often attenuate over space and occlusions (e.g., furniture and walls), resulting in low signal-to-noise ratios. Even for remote sensors, signals from one activity could be interference for others, which makes it difficult to detect multiple concurrent activities. Finally, wide-area sensing must be versatile to be useful. There are existing sensors that can sense a wide area, such as occupancy sensors. However, they can only sense the presence of users, and are incapable of sensing other aspects that might be useful to smart environments. Therefore, high versatility is the key to the success of wide-area sensing.

In this thesis, I develop wide-area sensing systems that addressed these challenges, and demonstrate practical wide-area sensing for room-, building-, and city-scale smart environment applications. Figure 1.3 shows a taxonomy of this thesis work along a sensing-scale axis. On a high level, my systems adapt existing everyday objects for sensing with low-cost instrumentation (e.g., conductive paints, retroreflective materials, corner reflectors) to leverage signals that can travel long distances (e.g., laser and radio frequency waves). These efforts led to three publications at top conference venues, and two currently in submission.

1.5 Document Structure

In the next two chapters, I will review prior work and discuss my approaches to creating wide-area sensors. In Chapter 4, I describe Wall++, my first attempt at achieving room-scale activity sensing by converting walls into sensors. Chapter 5 describes Sozu, my first attempt at achieving building-scale activity sensing. Chapter 6 summarizes Vibrosight, a long-range laser vibrometry system that detects human activities through reconstructing surface-level vibrations on everyday objects. Chapter 7 describes an RF sensing system which achieves long-range vibrometry with frequency-modulated continuous wave (FMCW) radar. Chapter 8 describes another laser vibrometry system that is based on and extended from Vibrosight for city-scale activity sensing. Finally, in Chapter 9, I discuss the contributions of this thesis and propose several future research directions.

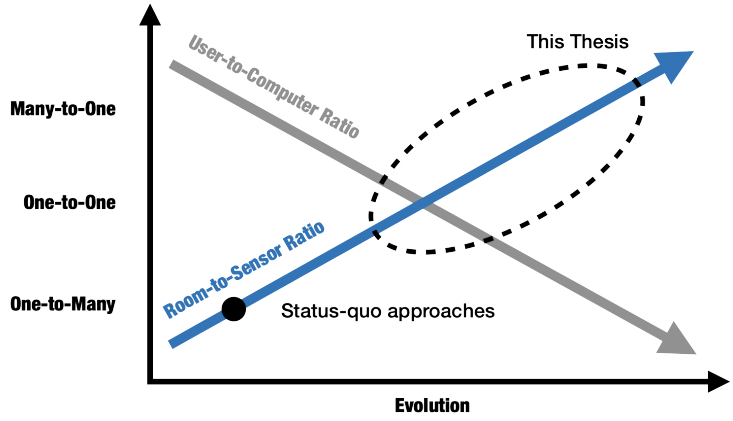


Figure 1.2: Smaller and cheaper computing devices result in an ever-decreasing user-to-computer ratio. People expect a similar decreasing trend of the ratio between the number of rooms and sensors. However, I propose wide-area sensing whereby each room will have fewer but more powerful sensors, resulting in an increasing room-to-sensor ratio.

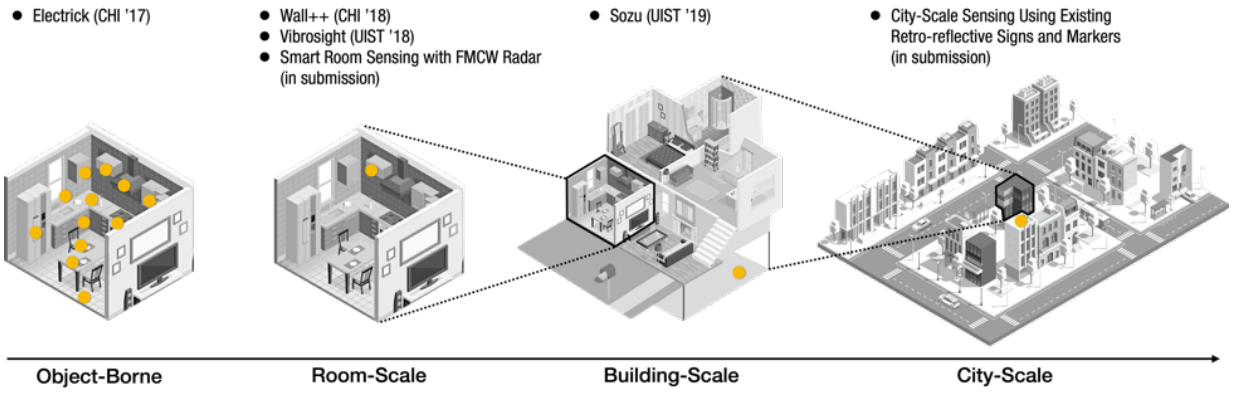


Figure 1.3: My thesis work in four sensing scales, from my early exploration of object-borne sensing, to room-, building-, and city-scale sensing. The orange circles denote sensor locations.

Chapter 2

Background

This thesis intersects with previous activity sensing systems for smart environments, the concept for which can be traced back to at least as early as the 1960s [71].

2.1 Special-Purpose Sensors

One straightforward approach to develop smart environments is to use a league of special-purpose sensors such as water flow sensors [36, 86], electricity monitor [23], and room indicators (e.g., humidity, temperature, occupancy sensors) [84, 111, 112, 113, 134] (Figure 2.1). These sensors are distributed around for e.g., a user's home, working collectively to detect user activities and environmental facets. Early pioneering research includes the Neural Network House [108], Georgia Tech's "Aware Home" [57, 81], and the SmartHome project at the University of Virginia, Medical Automation Research Center (MARC) [37].

Although this early research gave us a first glimpse into a smart-home experience, it has several drawbacks which have severely limited its use in practice. First, special-purpose sensors require sensor expertise from homeowners – a user has to know which sensors to use for the activity they want to sense, as well as the sensors' technical specifications such as field of view and sensitivity. In addition, reconfigurations of the room often mean users must recalibrate these sensors – a motion sensor pointed at a desk for occupancy detection must be calibrated

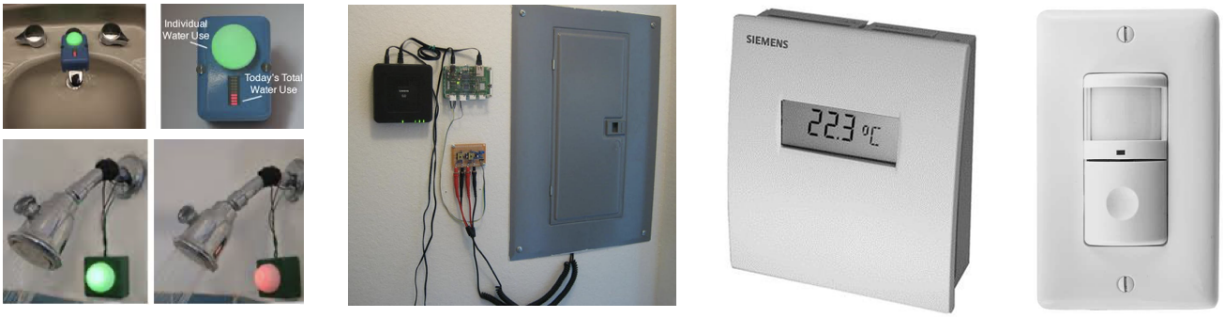


Figure 2.1: Examples of special-purpose sensors. From left to right: water flow sensors from UpStream [86], electricity monitor [23], a temperature sensor [12], and an occupancy sensor [27].

after the desk is moved. Finally, as previously discussed, special-purpose sensors require dense deployments, which scales with significant monetary and maintenance costs.

2.2 General-Purpose Sensor Tags

Due to the limitations of special-purpose sensors, people have sought more user-friendly approaches to provide a "stick-on and forget" user experience [145]. To achieve this, the sensor must be applicable to a wide range of activity types, or in other words, it must be versatile and general-purpose. Early research leveraged mechanical and magnetic switches [37, 145] which can be applied to many objects to detect the opening and closing of doors, cabinets, and appliance use, such as light and microwave use. However, these switches must still be strategically positioned to catch mechanical movements from target activities. Later, many sensors became more space- and power-efficient enough to fit into small tag form factors, which are more easily attached to objects of interest [33, 154]. Recently, sensor tag products (Figure 2.2) have been made available to consumers [16, 17, 24]. Although these sensor tags are more user-friendly for average homeowners than special-purpose sensors, they both require dense deployments, which undercuts their practicality for the high monetary and maintenance costs to scale.

2.3 Mobile Sensors

Instead of deploying static sensors into the environment, prior work also investigated sensors that can be carried around by users, most of which leverage motion sensors such as inertial measurement units (IMUs) [39, 51, 136, 146]. Other systems leverage signals such as sound [48, 78, 106], body-conducted acoustics [87, 90, 160], and computer vision [102, 114, 127]. It is also possible to use sensors that are not common to find on today's mobile devices. For example, researchers have leveraged electromagnetic interference [89, 103, 152] and RF backscatter [58, 123]. These sensors were often attached to a user's wrist, head, or limbs. Before the smartphone era, there had also been an effort to develop multi-sensor platforms, such as the Mobile Sensing Platform (MSP) [46]. Nowadays, smart devices like phones and watches often have multiple built-in sensors that greatly facilitate activity sensing research and its impacts. Millions of devices that are already around users are readily available for novel applications such as step count, hand washing timer, and sleep monitoring at the download of an app (Figure 2.3).

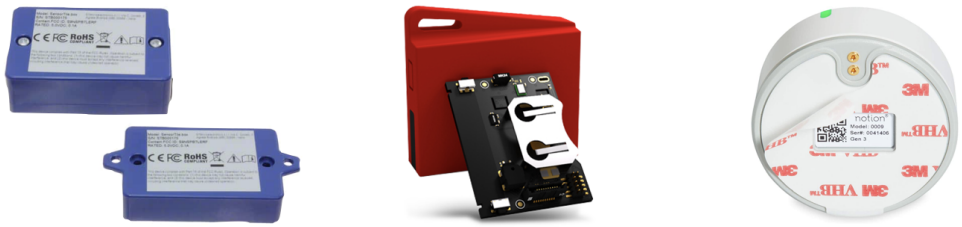


Figure 2.2: Examples sensor tag products, from left to right: STMicroelectronics SensorTile [16], TI SensorTag [17], and Notion multi-purpose sensor tag [24].

Figure 2.3: Three examples of activity sensing applications on mobile devices. From left to right: step counter [22], hand washing timer [6], and sleep monitoring [21].

2.4 Wide-Area Sensors

More pertinent to this thesis are remote sensing techniques, wherein a single sensor can cover a wide area for activity detection. This is commonly achieved either through cameras [88, 128, 173] and microphones [45, 92, 121], or through transforming common infrastructures into sensors [49, 59, 60, 119]. There are also active sensing approaches that involve the use of infrared signals (e.g., Microsoft Kinect, Intel RealSense), ultrasonic waves [135], and RF waves [29, 62, 97].

2.4.1 Camera

Many previous approaches have been used to detect activities at distance. Most popular are camera-based computer vision systems [88, 128, 173]. Recent research has also leveraged depth cameras which significantly ease the task of object segmentation [143]. In the product domain, there have been camera-based systems for activity sensing (e.g., Ring Security Systems [15] and Zensors [19]). Researchers have also used low-resolution thermal cameras, which capture black body radiation from the human body [79, 105]. For a more comprehensive review of this literature, please refer to [124]. Although camera-based systems are wide-area and versatile, their applications in environments like people's homes have received criticism for privacy concerns [3, 5].

2.4.2 Microphone

The microphone is another common sensor for detecting activities that generate sound. Prior work has explored detecting the presence of activities like eating and drinking [138], brushing teeth and washing hands [45], falling [55, 148], and gunshots [43, 96, 137]. Recently, the commercial success of smart speakers such as Google Home and Apple HomePod provides a viable microphone-based platform for activity sensing research (e.g., Listen Learner [157]). However, one of the most obvious limitations of microphones is that they cannot sense inaudible activities. Moreover, it is challenging for microphone-based systems to detect concurrent activities – signals from one activity would be interference for others when they have overlapped frequency components. Finally, microphones also pose privacy concerns due to their ability to disclose user

speech, which, like pictures and videos, might contain sensitive information that users do not want to reveal.

2.4.3 Wireless Sensing

Wireless sensing with radio frequency waves (e.g., Wi-Fi) has emerged as another promising activity sensing technique. First, we can use radio frequency waves to deliver power to a generic computing platform with sensors [130]. With backscatter techniques, smaller devices such as RFID can encode their information into signals that are reflected back to readers. Although commercial RFID systems are mostly created for presence detection, researchers have explored many novel uses such as activity sensing [41, 97, 123, 141]. However, passive RFIDs suffer from a limited sensing range (normally <15 meters between the reader and tags), which makes them able to sense activities at a room-scale, but not at a building-scale where activities of interest could be farther away from the sensor with multiple walls in between. Beyond RFIDs, previous research has also created novel backscatter devices to detect object grasp [62], 3D-printed object use [75], user location [109], and sound [34, 35]. Finally, researchers have leveraged radar techniques to enable instrumentation-free sensing applications. For example, prior work has demonstrated radar's capability in user identification and localization [149, 156], posture sensing [29, 147], object identification [161], and monitoring human vital signals [30, 153]. We recommend this recent survey [98] for a more comprehensive coverage of prior literature.

2.4.4 Infrastructure Mediated Sensing

Most related to this thesis is prior work that transforms common infrastructures into sensor – infrastructure-mediated sensing, from which this thesis took much inspiration. This sensing approach detects user activities through monitoring signals that propagate through infrastructures such as vibration on building floors [32, 118], pressure changes in water pipes [59, 60], sound from gas lines [49], air pressure in HVAC systems [120], and Electromagnetic Interference (EMI) from powerlines [67, 69, 72, 119]. Infrastructure-mediated sensing tackles the key conventional challenges in ubiquitous activity sensing on the deployability front. Specifically, such sensing systems allow users to quickly deploy a sensor at a single point while sensing activities throughout the entire space with minimum cost and maintenance effort. However, the sensing versatility of infrastructure-mediated sensing is constrained by the type of infrastructure used. For example, a powerline-mediated sensor could not detect the use of mobile tools or objects that do not run electricity. This thesis investigates opportunities to directly adapt everyday objects for remote sensing with low-cost and battery-free instrumentation. These instrumented objects then work in concert with centralized sensors through signals that propagate long distances (e.g., laser and radio frequency waves), so that they can be sensed afar and even through walls. With the constraints of infrastructures eliminated, the sensing systems in this thesis are highly versatile and general-purpose, being able to sense a wider range of activities.

Chapter 3

My Approaches

In this chapter, I summarize two key approaches I have taken to create wide-area sensors: 1) fabricating sensors onto large and ubiquitous everyday surfaces, and 2) transforming imperceivable signals to perceivable signals. My first approach embeds low-cost sensing onto everyday surfaces, with wide-area sensing achieved through large sizes and ubiquity of these surfaces. My second approach centers around a new sensing paradigm where a centralized sensor works in concert with peripheral mechanisms that adapt their host objects for remote sensing. These mechanisms transform their host objects' signals that are inherently challenging to be perceived by remote sensors, to perceivable signals. In this approach, I explored two specific techniques – one through self-powered mechanisms built on energy harvesting materials, and the other through lightweight or no augmentation on objects by leveraging retroreflectors.

3.1 Fabricate Sensors on Large Everyday Surfaces

To achieve wide-area sensing, one intuitive solution is to build large sensors (e.g., a room-size sensor for room-scale sensing). With sensors being large, it is easier for them to get closer to target activities and events. This is a positive thing for sensing because these large sensors' signal-to-noise ratios are high, which facilitates wide-area sensing. However, figuring out how to create such sensors in a low-cost manner is key to their practicality. I started with a preliminary exploration of an object-borne sensing approach, Electrick [165], wherein I used Electric Field Tomography and conductive paints, which can be easily coated on everyday surfaces to enable touch sensing on a wide range of objects for extremely low costs. Electrick costs less than $1 per square feet of the touch-sensitive area, which opens new exciting interaction opportunities for objects and surfaces that are large (e.g., furniture, walls) and inexpensive (e.g., tools, toys). With the same approach, I developed Wall++ [167], where I transformed walls, perhaps the most common structure in human environments, into sensors to detect user touch locations and poses for interactivity, as well as to sense active appliances for activity recognition (Figure 3.1). Chapter 4 will discuss Wall++ in detail.

3.2 Transform Imperceivable Signals to Perceivable Signals

Large sensors require more effort to deploy. This thesis has also explored ways to retrofit sensing into existing environments by small centralized sensors, so they can easily bring smartness to existing environments. These sensors could be integrated into future smart home devices, such as thermometers, lightbulbs, and speakers. However, small and centralized sensors must be able to sense things remotely. Human activities induce byproduct signals, many of which cannot travel long distances or through obstacles. For example, sunlight on a lawn, heat at a stove, airflow from a kitchen hood, as well as many other signals are inherently imperceivable to remote sensing, preventing a centralized sensor from providing whole-room or whole-house sensing coverage. To address this challenge, this thesis transforms imperceivable signals into signals perceivable to remote sensors with two techniques. Both techniques adapt objects of interest for remote sensing, which allow these objects to be more perceivable to centralized sensors.

3.2.1 Distributed Self-Powered Mechanisms

The first technique is to harvest energy from these signals and use this energy to generate different signals. This technique requires self-powered mechanisms that can be distributed on everyday objects. In Sozu [168], I built a set of small battery-free tags, each costing less than $10, to transform a wide range of activity signals into radio frequency broadcasts (Figure 3.2). For example, with the Sozu tag, vibration signals of a blender can be transformed into radio frequency broadcasts with piezoelectric materials. The presence of a radio frequency signal reveals the presence of a target activity, and we also leverage frequency modulation to distill rich information, such as direction, state, and intensity for a richer set of applications. Sozu self-powered mechanisms require significant amounts of effort to be instrumented on everyday objects. However, it is also possible to design sensors which can be easily retrofitted by users (i.e., set and forget) in their existing environment. I explored this concept in OptoSense [164], a system that allows users to attach self-powered photodetector arrays onto everyday objects for a wide range of activity sensing (e.g., medicine reminder, step count, liquid detection). Chapter 5 will discuss Sozu in detail.

3.2.2 Lightweight or No Augmentation of Objects with Active Sensing

Another technique is to enhance objects' response to active sensing with retroreflectors, which users can easily attach to objects, or find in existing infrastructures. These retroreflectors are ultra-low-cost and component-free, meaning they require extremely low monetary and time cost to deploy and maintain. Moreover, these retroreflectors return signals to their sources, resulting in increased signal-to-noise ratios, and thus critical performance improvements that make practical wide-area sensing possible. For example, with a vibrating blender, we can use laser vibrometry to transform its vibration signal into light intensity changes, which can be detected remotely. However, conventional laser vibrometry systems often have limited sensing ranges (i.e., < 5 m). In Vibrosight [166], I developed a compact laser vibrometry system to detect a wide range of activities, ranging from chopping vegetables to operating machines, by detecting vibrations induced by these activities (Figure 3.3). A similar sensing modality can be achieved with FMCW

radars, which I explored in another system [169]. Beyond indoor settings, I have also explored laser vibrometry for outdoor smart city applications by taking advantage of existing retroreflective markers and signs around the city [170]. These three systems will be discussed in detail in Chapter 6, 7, and 8.

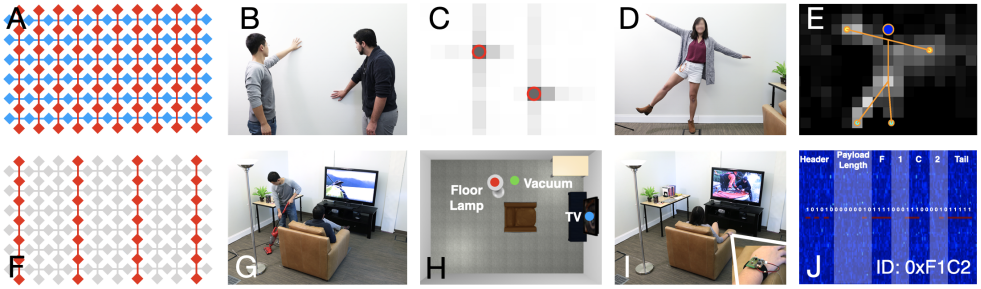


Figure 3.1: Wall++ in active mutual capacitive sensing mode (A) enables touch tracking (B,C) and pose estimation (D,E). Wall++ in electromagnetic sensing mode (F) enables appliance detection and tracking (G,H), as well as user ID (I,J).

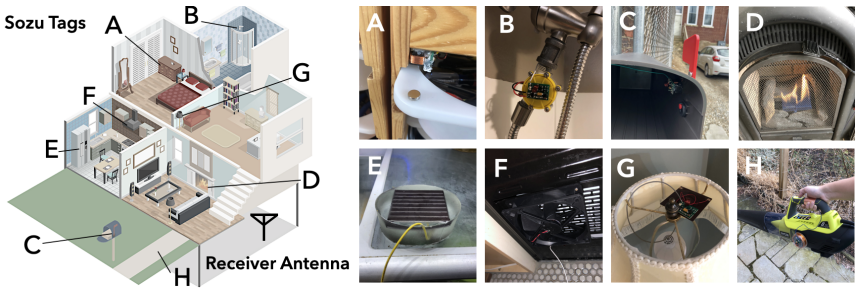


Figure 3.2: Sozu transforms energy from activities into RF broadcasts, monitored by a receiver antenna.

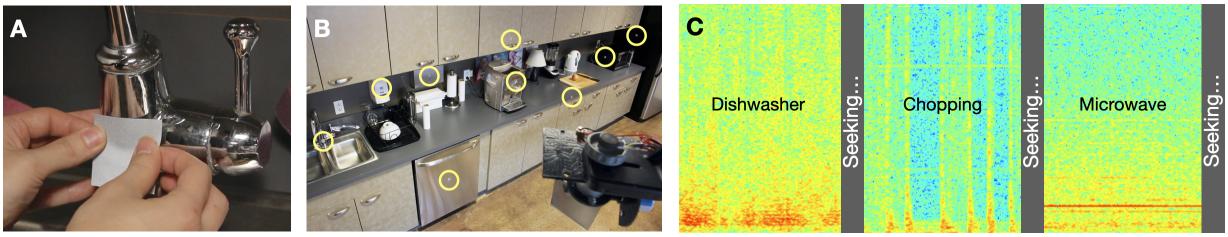


Figure 3.3: Users can tag objects (A) they wish to reveal to Vibrosight. The system (B, lower right) periodically scans the room to find retroreflective tags (highlighted with yellow circles), which can be used to sense vibrations at long distances. The resulting spectrograms (C) can be used to infer appliance use and human activities.

Chapter 4

Wall++: Room-Scale Interactive and Context-Aware Sensing

4.1 Introduction

Walls are everywhere, often making up more than half of indoor surface area, especially in residential and office buildings. In addition to delimiting spaces, both for functional and social purposes, they also hide infrastructure, such as wiring and HVAC. However, they are generally inactive structural elements, offering no inherent interactive or computational abilities (other than at small attached silos, e.g., thermostats and light switches), and thus present a tempting opportunity for wide-area sensing, especially considering their ubiquity.

We set out to identify methods that could recast walls as smart infrastructure, able to sense interactions and activities happening in a room. In addition to supporting these broad application domains, we also added process constraints. In particular, we sought a technical approach that was versatile and easy to apply, requiring no special tools (e.g., CNC machines) or skills (e.g., carpentry, electrical engineering). We also required our approach to be low-cost, so as to be economically feasible at room scale (even a small room, e.g., 2×2.5×2.5 m, has more than 20 m^2 of walls). Finally, we wanted our sensing approach to be minimally obtrusive, and ideally invisible.

We quickly identified paint as a particularly attractive approach. Walls are already painted, and the average homeowner has the requisite skills to paint a wall. While there are special tools for applying paint (e.g., brushes, rollers, painter's tape), these are all commodity supplies and readily available at home improvement stores. As we will discuss in greater depth later, we can apply a standard latex topcoat, which allows our technique to be wall-scale, and yet hidden in plain sight. Our ultimately selected method costs ~$20 per m^2 in materials at retail prices. These properties satisfied all our process criteria.

To enable user and environmental sensing, we drew upon two large bodies of work in the literature. First, we selected mutual capacitive sensing [66, 126, 140] for close-range interactions. Owing to its widespread use in smartphones and tablets, mutual capacitive sensing is well understood and robust, allowing us to readily adapt it to wall-scale applications. Second, we extended work in airborne electromagnetic (EM) sensing [83, 116, 125, 129, 158, 159]. This

required us to develop an electrode pattern that supports both of these sensing modalities. For user sensing, we investigated touch interaction, pose estimation, user identification and tracking. For environment sensing, we focused on context awareness though appliance recognition and localization.

Collectively, we call our process, materials, patterns, sensor hardware and processing pipeline, Wall++. As we detail in the following pages, optimizing for ease-of-application and reliability, as well as sensing range and resolution, required iterative experimentation, physical prototyping, simulation modeling and user studies. We believe our resulting system demonstrates new and interesting HCI capabilities and presents a viable path towards smarter indoor environments.

4.2 Electrode / Antenna Implementation

The basic principle of Wall++ sensing relies on patterning large electrodes onto a wall using conductive paint. Thus, as a first step, it was necessary to develop a reliable and economically-feasible way to add large electrodes to walls. To identify suitable materials and processes, we performed a series of tests with various conductive paints, backing materials, application methods, number of coats, and topcoats. We then explored different electrode patterns suitable for our applications, and optimized them for sensing range and resolution. In all tests, we used an LCR meter to measure electrical impedance at 100 kHz.

4.2.1 Phase 1: Paint and Backing Materials

Both capacitive sensing and airborne EM sensing require conductive electrodes in order to induce charges freely. Thus, our first task was to identify paints that were inexpensive, non-toxic and sufficiently conductive to support our application goals. We experimented with commercially-available carbon, water-based nickel, acrylic-based nickel, and silver paints. Simultaneously, we tested three common backing materials: wallpaper, drywall, and primed drywall. All paints were applied in a single coat with a roller.

Figure 4.1 shows the result of this 4×3 experiment. Despite its high conductivity, we removed silver paint from consideration due to its high cost (~$200 per m^2). Carbon paint was also eliminated due to its high resistance, which is incompatible with our technique. Among the remaining two nickel-based paints, we selected the water-based version, as it produced less odor and resulted in a smoother finish (Figure 4.1, bottom photos).

4.2.2 Phase 2: Application Method and Number of Coats

With our conductive paint selected, we next considered its application method. We varied both the number of coats and the tool used, both of which affect conductivity. We tested brush, spray, and roller applications with one, two, and three coats, resulting in a 3×3 test. Figure 4.2 shows these results, which consistently indicate that the surface conductivity increases with number of painted coats. Among the application methods, we selected roller painting, as it resulted in the highest conductivity and lower variance across the surface. As an added benefit, it was also the fastest application method.

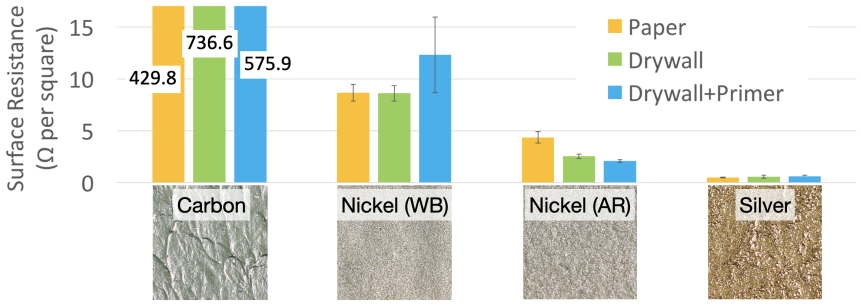


Figure 4.1: Conductivity test with different paints across three backing materials. Close-up of painted surface included.

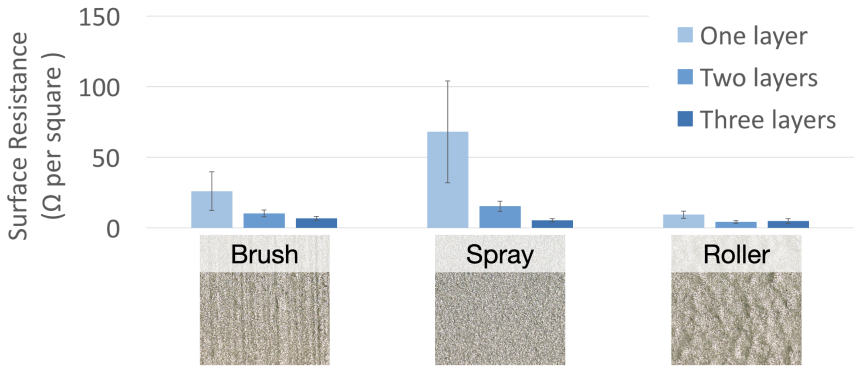


Figure 4.2: Conductivity test with different application methods and number of coats.

4.2.3 Phase 3: Topcoat

To improve appearance and durability, we studied the effect of topcoats on our electrodes' performance. We suspected that solvents from later paint coats could interact with the conductive paint layer, affecting its conductivity. We also wanted to see if varying surface permittivity of different topcoat materials affected performance. For this experiment, we tested no topcoat, acrylic, primer, latex paint and wallpaper. However, we did not find any significant differences across these conditions, and thus we recommend covering Wall++ in a standard architectural latex paint for improved durability, ease of cleaning and appearance.

4.2.4 Phase 4: Traces & Insulation

To connect our painted electrodes, we need to run thin traces between them. Crucially, transmitter electrodes must be insulated from receiver electrodes to project most of the electric field into the air, requiring insulation between trace intersections. Thus, as a precursor to exploring electrode pattern, we first needed to identify a trace option with high conductivity and good insulation. Our tests included three materials: copper tape (3.2 mm width), silver ink drawn by pen (1 mm width), and nickel paint applied with stencil and brush (1 mm width). Simultaneously, we tested three insulation materials: vinyl sticker, latex paint and primer.

Conductivity test shows that copper tape had the highest conductivity (0.13 Ω/cm SD=0.0),

followed by nickel paint (5.6 Ω/cm SD=4.9) and silver traces (63.5 Ω/cm SD=10.4). In the insulation test, we found that our nickel traces interacted with the latex paint and primer conditions, causing shorts, though it worked fine with vinyl stickers. Silver traces worked with all insulators, but had high variance in conductivity. Copper had the worst insulation due to a larger overlapping area, but the least variance, and for this reason, we chose it in combination with vinyl stickers (the most consistent of the insulators we tested).

4.2.5 Phase 5: Electrode Pattern

Having identified a reliable way to paint, connect and insulate conductive electrodes on walls, our next step was to select a pattern that enabled our desired applications. Fortunately, airborne EM sensing is not particularly sensitive to pattern geometry, and SNR is most of a function of antenna size. For example, previous work used copper patches or a simple wire antenna. Therefore, we chiefly optimized our design for mutual capacitive sensing, in which pattern plays a critical role. However, in Phase 7, we confirm the performance of our antenna designs in capturing airborne EM signals.

For our mutual capacitance sensing, we desired a pattern that 1) projected an electric field as far as possible, so as to provide the largest interactive volume, while also 2) offering sufficient resolution to enable fine-grained interactions, such as touch tracking. We studied five patterns common in the literature: lines, stripes, half circle, diamond and circle dot (Figure 4.3, top).

To best evaluate the electric field projection across these designs, we ran simulations using COMSOL. This provided a high-resolution view impossible to capture with hand measurements. We fixed the transmitter and receiver electrode size to 25 cm^2, except in our lines condition, which are purposely thin. We also fixed the distance between electrodes (i.e., pitch) to 5 cm, except in our lines and stripes conditions. We set the voltage difference between transmitters and receivers at 18 V.

Figure 4.3, bottom, shows our simulation results. Due to the short-range of its electric field projection, we eliminated lines as a candidate design. Projection range is improved with the increased electrode size in stripes, however there is too much inner capacitance between electrodes, which significantly reduces SNR and sensing range. The rest of the patterns do not suffer from this issue and have similar projection range. Ultimately, we selected the diamond pattern because it densely covers the surface, making it unlikely to miss user touches.

4.2.6 Phase 6: Pattern Optimization

After selecting the diamond pattern, there were two immediate parameters to tune – the size of the diamonds and the pitch. Intuitively, bigger diamonds and pitches should project larger electric fields. However, they also decrease the array's resolution. Therefore, we set out to find parameters that offered a balance between sensing range and resolution.

Figure 4.4, A-C, show electric field simulations at different electrode sizes (30, 50 and 70 mm) with a fixed 50 mm pitch. As expected, the bigger the electrode size, the farther the sensing range. Figure 4.4, D-F, show simulations at different pitches (30, 50 and 70 mm) with an electrode size fixed at 50 mm. Interestingly, bigger pitches do not improve sensing range. Combining what we discovered in this experiment, we settled on 70 mm electrodes with a 48 mm pitch – a common

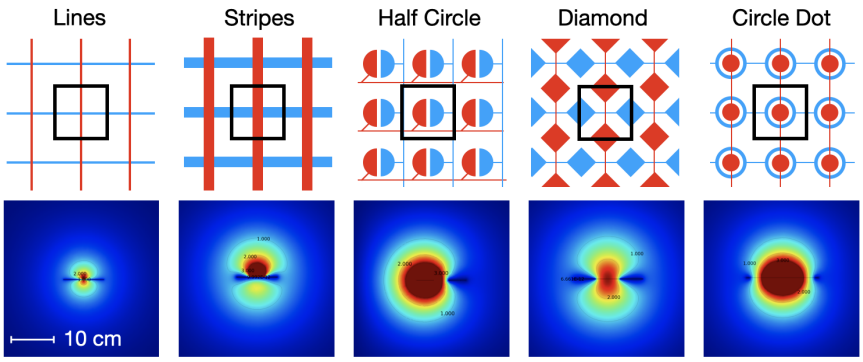


Figure 4.3: Top: electrode patterns we studied (transmitters in red, receivers in blue). Bottom: electric field simulations of electrodes in black region (higher voltages in red).

width of painter's tape, facilitating fabrication. As seen in Figure 4.3, a regular diamond pattern can be efficiently produced by laying down a crosshatch of painter's tape, and then using a paint roller.

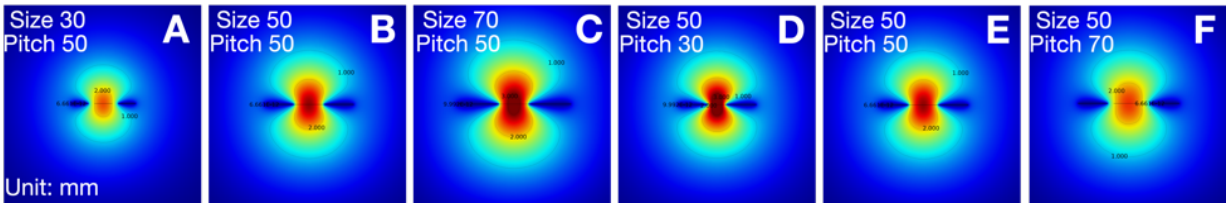


Figure 4.4: Simulations of diamond patterns with different sizes (A-C) and pitches (D-F).

4.2.7 Phase 7: Antenna Sensitivity

Phases 5 and 6 were primarily focused on mutual capacitance sensing. In this design phase, we wished to verify that our selected diamond pattern could robustly capture airborne EM signals. There are many ways to configure diamond patterns into an antenna array. For example, we could connect all columns and rows together to make one large antenna. However, this monolithic antenna eliminates the possibility of triangulating signal sources, discussed later.

Therefore, we investigated the idea of selecting a subset of diamond columns as antennas (as illustrated in Figure 3.1F). These need not be single columns, but could be several adjacent columns connected together. To see if this improved signal, we conducted a test in a shielded chamber with minimal EM noise. To be able to vary antenna size, we painted diamond electrode patterns on individual 1×8' foam boards, each of which acted as a single-column antenna, but which could be connected together to make a multi-column antenna. We varied the number of columns in the antenna unit from 1 to 3, with a known signal source placed 50 cm away.

Result indicated that larger antenna sizes offered improved signal strength. However, the improvement was minor – a three-columned-antenna only improved signal strength by 15% over a single column unit. Given the gain was modest, we decided to use single column antennas for circuit simplicity and improved spatial resolution.

Next, to better quantify the sensitivity of single-column antennas, we collected EM signals from 12 appliances at varying distances. As can be seen in Figure 4.5, all of our test appliances can be sensed within a 2-meter range, and some noisy devices up to 4 meters. We also found serendipitously that the human body broadcasts EM signals when holding and operating some appliances. For example, a hairdryer we tested had no visible signal unless a user was grasping it. We also found a class of appliances that only activate when touched (chiefly for power conservation), e.g., laptop trackpads and smartphone fingerprint readers. However, this has the interesting potential to allow for recognition of human activities at the moment of user engagement.

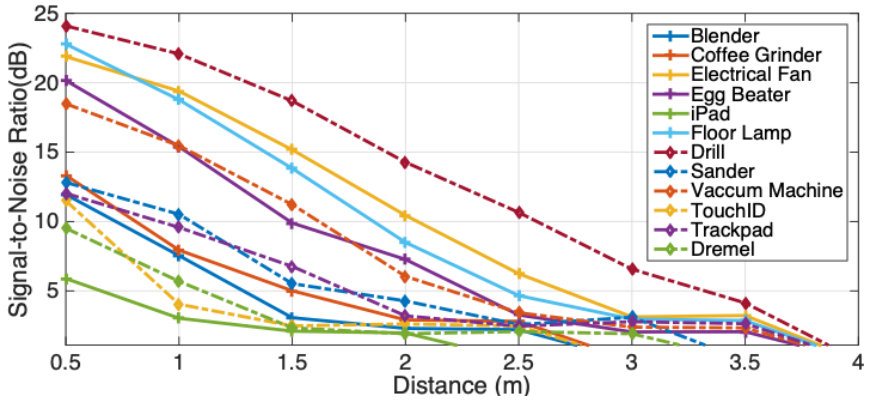


Figure 4.5: Received signal strength collected with appliances operating at different distances to the wall antenna.

4.2.8 Phase 8: Wall Construction

After we finalized our fabrication parameters, we painted a real wall at our institution, measuring 12×8' (3.7×2.4 m), seen in Figure 3.1 (B & D) and Figure 4.6. We used this wall to verify our previous focused experiments. This wall has 22 columns and 15 rows of electrodes, for a total of 37 coaxial cable connections to our custom sensing hardware. After we nickel painted the wall with a diamond pattern, we finished it with a standard latex paint. In total, the wall took roughly four hours to complete with a total material cost under $200. We anticipate that the time and material cost of a commercially-deployed solution would be significantly reduced with trained painters and bulk material purchase.

Figure 4.6: Painter's tape is laid down in a crosshatched pattern (A & B), and then painted en masse with e.g., a roller (C) to create a grid of regular diamonds (D).

4.3 Sensing Hardware Implementation

To enable user and environment sensing, Wall++ employs two distinct sensing modes: mutual capacitive sensing and airborne EM sensing. This required development of custom sensor boards (Figure 4.7), built around a Cortex M4 microcontroller running at 96 MHz (MK20DX256VLH7), powered by Teensy 3.2 firmware. Our main board (Figure 4.7A) plugs into two multiplexing boards, one designed for mutual capacitance sensing (Figure 4.7B) and another for EM sensing (Figure 4.7C). In the future, these could be integrated into a compact, single-board design.

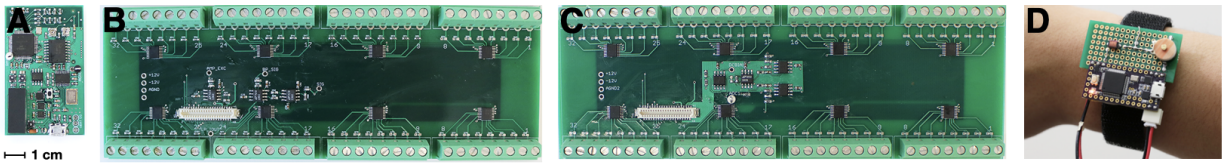


Figure 4.7: Different hardware components we developed. A) Main sensor board, B) capacitive sensing multiplexing board, C) EM multiplexing board, and D) signal-emitting wristband. Uniform scale.

4.3.1 Mutual Capacitance Sensing

To detect a user's hands and body pose, we use mutual capacitance sensing, which measures the capacitance between two electrodes. This sensing technique is the basis of modern touchscreens as seen in smartphones and tablets. When a body part is near a transmitter-receiver pair, it interferes with the projected electric field, reducing the received current, which can be measured. This is referred to as shunting mode sensing. On the other hand, if the user's body directly touches an electrode, it greatly increases the capacitance and received current. See [65] for a more thorough review of capacitive sensing in HCI.

In capacitive sensing mode, our main board uses an AD5930 DDS to generate a 100 kHz sine wave as the excitation signal. This signal is amplified to 18 V peak-to-peak by the multiplexing board (Figure 4.7B) and routed to a specified transmitter electrode column (Figure 3.1A, red). We use another set of multiplexers to connect a receiver electrode row (Figure 3.1A, blue) to our analog frontend, which is filtered and amplified. We use an AD637 RMS-DC converter to measure the amplitude of the received signal, which correlates to the capacitance between the current transmitter and receiver electrodes. We set the integration time for the AD637 to 100 microseconds (i.e., 10 periods of the excitation signal). The output of the converter is sampled by our microcontroller's built-in ADC.

Our 12×8' augmented wall has 22 columns and 15 rows. At any moment, only one transmitting column and one receiving row are selected for mutual capacitance sensing. The circuit measures the mutual capacitance between the two electrodes, which is most strongly affected by a user's body being proximate to (or touching) the intersection of the column and row. The circuit then moves on to the next row-column pair until all 22×15 = 330 measurements are collected. These measurements are then sent to a laptop over USB at 16.5 FPS.

4.3.2 Airborne EM Sensing

In EM sensing mode, no active signals are injected into the wall's electrodes. The multiplexing board (Figure 4.7C) features a differential amplification circuit with a 159 Hz high pass filter to remove DC components and powerline noise. One terminal of the input is connected to common ground, while the other terminal is cycled through columns, one at a time, each serving as a signal-column antenna. The signal is then amplified with a gain value of 100 and DC biased to AVDD/2 (1.65V) before sampling.

Our microcontroller's two built-in ADCs are configured into a high-speed, interleaved DMA mode, enabling a sampling rate of 4 MHz with 12-bit resolution. We collect 1024 ADC measurements and perform an on-board FFT computation. To better capture transient EM spikes, the board performs this measurement 20 times, and records the maximum value for each FFT bin as the result. This process takes ~20 milliseconds per column, resulting in an FPS of 6.2 for an 8-column-attenna setup.

4.4 Touch Tracking

Mutual capacitive sensing enables Wall++ to track a user's hand hovering above or touching a wall's surface. We first describe our software implementation, followed by our evaluation procedure and results.

4.4.1 Software

Due to fabrication inconsistencies, the raw capacitance measured at each row-column pair can vary. To compensate for this, we capture a background profile and convert all measurements into z-scores, which we calculate by subtracting the background mean from individual raw measurements and then dividing the differences by the background standard deviation. When a user touches the wall, a transmitter and receiver pair are capacitively shorted, which makes the touched region have a significantly higher capacitance than the captured background. We can visualize this as a pixel in a capacitive image (Figure 3.1C), which is thresholded to get touch coordinates. When a user's hand is hovering above a wall, it capacitively couples to many row-column pairs, appearing as a negative blob in the capacitive image. For hover tracking, we identify blobs of activated pixels, and interpolate the peak by calculating the center of mass in a 3×3 pixel area.

4.4.2 Evaluation

To investigate the hand tracking performance of Wall++, we recruited 14 participants (7 female, average age of 24). The heights of these participants ranged from 160 to 183 cm, with masses ranging from 50 to 90 kg. The study took roughly 40 minutes to complete and participants were compensated $20 for their time. We used our 12×8' wall as the test apparatus. A calibrated projector was used to render experimental prompts for participants. Additionally, a small plastic-runged ladder was provided if requested points were beyond a participant's reach (which also

provided a more challenging grounding condition to study).

We first asked participants to walk around for 10 minutes roughly one meter away from the wall. This provided 9900 no user present trials per participant. We then asked participants to "click" points digitally projected onto the wall's surface. When a point turned red, participants placed their hand to that point, allowing for 30 touch coordinates to be recorded over a 2 second period. The point then turned green, at which point participants held their hands roughly 10 cm from the point; 30 hover coordinates were recorded. No feedback about the tracking result was shown to participants. In total, 50 fully randomized points were requested from each of our 14 participants, resulting in 21,000 touch and 21,000 hover trials.

4.4.3 Results

Of the 138,600 no user present data points, representing 140 minutes of data, there were no touch or hover events reported by our system (i.e., 100% accuracy). Of our 21,000 touch trials, 97.7% (SD=2.4) were correctly labeled as touch events by our system (2.3% were incorrectly detected as hovers). Hover detection was 99.8% (SD=0.3) accurate.

Using requested coordinates and our system's reported coordinates, we calculated the Euclidian distance error for our touch and hover trials. We found a mean touch tracking distance error of 13.7 cm (SD=1.1) and a mean hover tracking distance error of 6.5 cm (SD=0.3). Figure 4.8 provides an interpolated error heat map across our wall's surface. There is one region of reduced accuracy, which we suspect is due to either a fabrication defect or possibly metal/electrical infrastructure behind the wall. Also, we did not see any reduction in accuracy for participants who used the ladder for reaching high points.

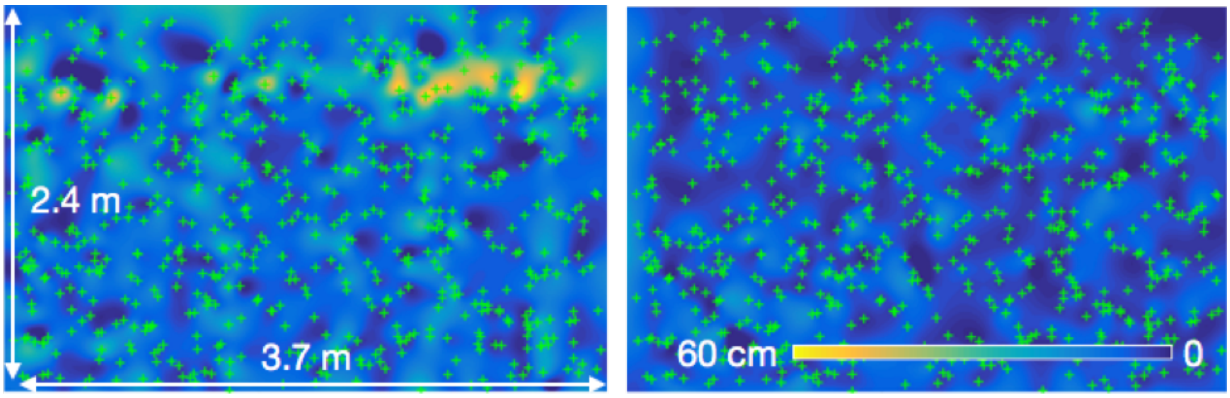


Figure 4.8: Touch (left) and hover (right) tracking distance error on our test wall (interpolated across surface). Green crosshairs show (50×14) 700 requested locations.

4.5 Pose Estimation

Wall++ can also estimate body pose of users if they are close to a wall. We now describe this software implementation, evaluation procedure and study results.

4.5.1 Software

As with touch and hover tracking, pose estimation uses a z-scored capacitive image as input. We first look for users by sliding a window of 3×15 antenna patches along the x-axis, searching for a blob of sufficient total activation. If a region is found to be above threshold, pose estimation is triggered for that bounding box. Along the center column, pixels above a second threshold are labeled as the torso. We then scan to the left and right of the blob, labeling bottom extents as feet and upper extents as hands. An example of these five key points is shown in Figure 3.1, D and E. We can use these key points to characterize different body poses; for evaluation, we included standing, left arm lifted, right arm lifted, both arms lifted, left leg lifted, and right leg lifted (Figure 4.9).

4.5.2 Evaluation

We used the same group of participants and apparatus as our touch tracking study. In total, there were five rounds of live testing. At the beginning of each round, we assigned participants to a random standing location 20 cm in front of the wall. They were then asked to perform the six test poses, sequentially, and in a random order. For each pose, we recorded 30 data points over a 2 second period, which resulted in 12,600 pose trials (5 rounds × 6 poses × 30 trials × 14 participants).

4.5.3 Results

Of the 12,600 trials we captured, 99.8% (SD=0.6) triggered our pose estimation pipeline. Overall, the system inferred the correct pose in 92.0% (SD=3.5) of trials; a confusion matrix is provided in Figure 4.9. The greatest source of error (63.5%) is from left leg and right leg being confused with stand. Figure 4.9, bottom-left, shows the averaged capacitive image for each pose (all trials and participants combined). We also found the torso key point accurately reflected a user's location along the wall, with a mean distance error of 8.6 cm (SD=2.2).

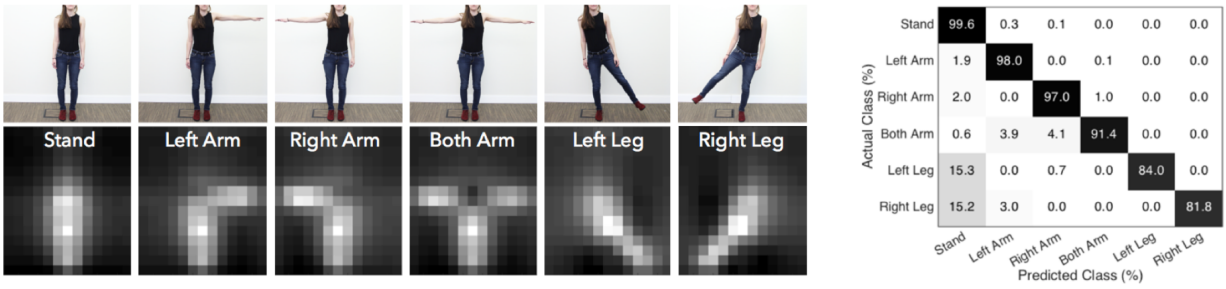


Figure 4.9: Left: six poses (top) and averaged capacitive images from the user study (bottom). Right: confusion matrix for 6 poses.

4.6 Appliance Detection

Wall++ captures airborne EM signals emitted by electrical appliances when running. In this section, we focus on detecting the on/off state of appliances (i.e., detection, but not localization). In being a room-scale sensing technique, Wall++ had to solve two important challenges, which differentiates us from prior work. First, unlike EM sensing with conductive media (e.g., powerlines [69, 72], human bodies [89, 150]), air substantially attenuates EM signals, which would generally preclude long-range airborne EM sensing. We overcome this by using large antennas. Second, unlike worn EM detectors, which can generally assume that only one appliance is grasped at any given time, Wall++ must handle simultaneous active appliances. For this, we use a special pipeline, described next.

4.6.1 Software

To help suppress persistently noisy EM bands (from e.g., fluorescent light ballasts), our system computes and uses z-scored FFTs. Before appliances can be recognized, they must be registered in our system. This is done by capturing data while an appliance is active, and recording its FFT signature. We then threshold this FFT to create a bitmask representing characteristic frequencies for that appliance.

When live data is being streamed from our sensor board, the incoming FFTs are bit-masked against each known appliance and passed to a corresponding, appliance-specific SMO classifier (Poly Kernel, E=1.0, output: running or not running). In essence, this bit-masking approach has the effect of making each appliance classifier blind to non-relevant EM bands, which enables multi-appliance detection, reduces training data collection, and improves overall robustness. We used a one-second classification hysteresis to reduce spurious appliance detections. The result of this process is a list of active appliances detected at each antenna. These sets are unioned to provide a list of active appliances in the room.

4.6.2 Evaluation

Room-scale appliance detection requires all walls to be augmented. However, our previous test apparatus was only a single wall (12×8'). To simulate a fully augmented room, we distributed 1×8' column antennas we had previously made for our Antenna Sensitivity study. As an added benefit, this apparatus allowed us to run our experiment in three different locations: office, kitchen, and workshop (Figure 4.10). At each location, we evenly distributed eight column antennas around the room periphery, working around windows and doorways as needed.

In each location, we tested six contextually-appropriate appliances: three fixed and three mobile. The locations of fixed appliances are color coded in Figure 4.10, while we randomized the position of mobile appliances according to a one-meter grid we laid out in each room. We omitted locations blocked by furniture, resulting in 12 test points in the Office, 14 in the Kitchen, and 26 in the Workshop.

To train our system for a room and its appliances, we collected EM signals using one antenna. In total, there were three rounds of training data collection. In each round, 90 data points were recorded over 15 seconds when no appliance was active. We then collected 90 data points for

each appliance while active (one at a time), during which we varied the distance between the appliance and the antenna up to 2 meters. We then trained a classifier for each appliance, using the background data (i.e., no appliance running) and the other five appliances as negative examples.

At each location, we performed three rounds of live testing at different times of day – morning (~8 am), noon (~12 pm), and late night (~11 pm) – when the building had different environmental conditions, occupancy load, etc. In each round of testing, we first recorded 10 minutes of data (3720 data points) when no appliances were active, to test for false positives resulting from background EM noise. We then activated all six appliances, one at a time, in a random order. As an added experiment, we also turned on all three fixed appliances simultaneously. In all trials, we turned on and off the appliances five times each. Real-time detection results were recorded.

4.6.3 Results

Across the 90 minutes (33,480 data points) of data collected when no appliance was turned on, 1.3% (SD=1.0) of trials were labelled as having an appliance running (i.e., false positives). Across all trials when appliances were running, 85.3% (SD=4.9) correctly classified the active appliances. We found that mobile appliances contributed 88.4% (SD=12.8) of the errors, mostly when they were at the center of rooms and far from any antenna.

We found no significant difference in accuracy across time of day. However, we did find that background noise changed over time, and thus we had to recapture the background profile for our z-score computation at the start of each session. This indicates that Wall++ will need a dynamic backgrounding scheme when deployed.

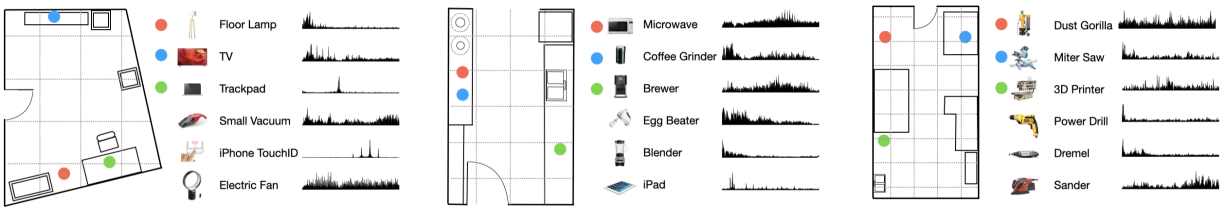


Figure 4.10: On the left: floor plans of an office (left), kitchen (center), and workshop (right). On the right: appliances and their EM profiles (0 to 2 MHz). One-meter-spaced grids are shown in dashed lines.

4.7 Appliance Localization

Airborne EM signals attenuate as they radiate in air, leading to different signal amplitudes across a distributed array of antennas, such as column antennas along the walls of a room. Wall++ leverages this affect to localize the source of EM signals, and even track the source if mobile. In this section, we describe our tracking pipeline, and its evaluation and results.

4.7.1 Software

Our tracking pipeline extends our Appliance Detection pipeline by additionally using the masked FFTs to calculate a received signal strength (RSS) as P_i in (4.1) for each known appliance at each antenna. According to the Friis transmission formula, the relation between the appliance's location and its RSS measured at the i-th antenna can be modelled as

$$f_i(x,y,A_T) : \frac{A_T G_i}{(x-x_i)^2+(y-y_i)^2} = P_i \tag{4.1}$$

Here, P_i is the received signal strength at the i-th antenna, G_ii is the sensitivity of antenna i, (x_i, y_i) are the coordinates of the i-th antenna, and A_T is the transmitter's radiated power. Therefore, an appliance's location can be obtained by solving an L2-norm minimization problem:

$$\min_{x,y,A_T} \sum_{i=1}^{L} \| f_i(x,y,A_T) - P_i \|_2 \tag{4.2}$$

We first calibrated our system using a known, single-tone transmitter to estimate G_i for the i-th antenna. Then, we used the Nelder-Mead optimization method from the Python scipy package to minimize 4.2.

Given a received signal P_i at i-th antenna with its respective location x_i,y_i and the sensitivity G_i, 4.2 can return both the unknown appliance location (x,y) and its radiated power A_T. Although different appliances have different radiated power, our algorithm does not depend on prior knowledge of an appliance's absolute radiated power A_T, as the computation is relative. Since we have three unknown parameters (x,y,A_T) in 4.2, at least $L \geq 3$ column antennas are needed to produce a location estimate. Intuitively, the more antennas, the more data the algorithm can use for convergence, improving localization accuracy.

4.7.2 Evaluation

For this evaluation, we used the same mobile appliances, rooms, and column antenna deployment as our Appliance Detection study (Figure 4.10). For each room, we collected 40 data points (over ~6 seconds) for three mobile appliances at all points on our one-meter grids (which acted as a spatial ground truth). As before, we omitted grid points blocked by furniture. In total, 6,240 data points (12+14+26 grid locations × 40 data points × 3 appliances) were collected for analysis.

4.7.3 Results

Our tracking algorithm localized appliances with a mean distance error of 1.4 m (SD=0.5). Figure 4.11 (top) illustrates this error across our one-meter room grids. As can be seen, accuracy varies considerably even in a single room (e.g., workshop – Figure 4.11, top-right – best tracking accuracy: 0.6 m; worst: 3.6 m), but overall suggests feasibility.

We also considered deployment in real-world locations where doors, windows and other infrastructure might block the placement of Wall++. To simulate this, and investigate how much it affects localization accuracy, we ran a post hoc study removing an increasing number of antennas from the room (Figure 4.12). More specifically, for each antenna count, we randomly selected a

Figure 4.11: Tracking distance error at our three test locations. Left to right: office, kitchen and workshop.

subset of antennas and reran our localization algorithm using only those data. We repeated this random selection three times for each antenna count and averaged the results.

As expected, tracking error increases as fewer antennas are available. However, even in the worst-case scenario, with only three antennas, we can still localize appliances to within 4 meters on average, which is coarse, but still potentially useful. It also appears likely that using more than 8 antennas would yield even better tracking accuracies. This would be the case in a fully realized installation of Wall++, as our recommend pattern has 8 column antenna every 1.35 m (Figure 4.6); a 4×4 m room would have roughly 100 antennas.

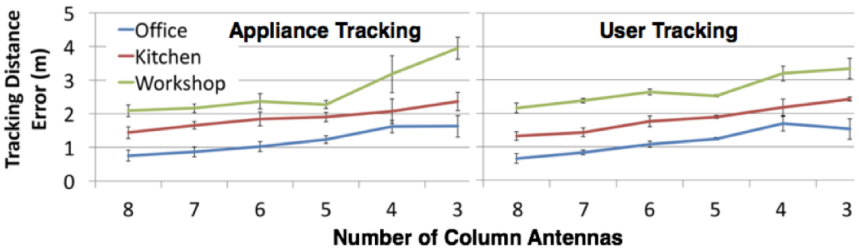


Figure 4.12: Tracking distance error using different numbers of column antennas (three locations tested).

4.8 User Tracking & Identification

We have already discussed how Wall++ can track appliances when they are radiating EM signals. This motivated us to build a small, signal-emitting wristband (Figure 4.7) to enable user localization and identification using the same physical setup and tracking pipeline as appliances. Our signal-emitting wristband uses a Teensy microcontroller attached to an LC tank, driven by a 3.3 V PWM signal set at a frequency between 800 kHz and 3 MHz. By enabling/disabling the drive pin, the emitted signal can be turned on or off, creating an on-off-keying (OOK) signal that we use to communicate with Wall++ (Figure 3.1 I & J) at a maximum speed of 300 baud. Though the throughput is low, it is more than sufficient to transmit a user ID, and even low-speed sensor data, such as heartrate. Figure 3.1 J shows a waterfall spectrogram from 1.4 to 1.6 MHz with a 1.5 MHz carrier frequency (seen as red line segments).

4.8.1 Evaluation

To evaluate Wall++ for user tracking, we conducted an evaluation using a similar procedure to our Appliance Localization study. We configured the wristband to emit a constant 1.5 MHz signal. In each of our three rooms, we asked 5 participants to wear our wristband, and stand on the one-meter grid points sequentially, during which we collected 40 data points (~6 seconds). As before, we omitted grid points that were blocked by furniture.

We also ran a basic study to investigate the data transmission performance over different distances between a participant and a wall antenna. For this test, we used one column antenna sampled at 120 FPS. As a proof-of-concept evaluation, we configured our wristband to output at 20 bits/sec, transmitting a 6-bit header, 8-bit payload length, 16-bit user ID, and 5-bit tail (35 bits in total; Figure 3.1 J). We recruited 5 participants to wear our wristband, and asked them to stand 1, 2, 3 and 4 meters away from the antenna. At each distance, we recorded 5 data transmissions from the wristband.

4.8.2 Results

Our user tracking results show an average distance error of 1.4 m (SD=0.6). This performance is almost identical to our Appliance Localization results. Figure 4.11, bottom, illustrates the tracking error across each room. We also ran a post hoc study to investigate how the number of antennas in a room would affect tracking accuracy (see counterpart study in Appliance Localization for procedure). As before, accuracy decreases with antenna count (Figure 4.12), but coarse tracking remains feasible with just three antennas.

With respect to data transmission performance, there were no bit errors for all trials collected within 3 meters of the antenna. However, at 4 meters, the bit error rate increased to 46.4% (SD=26.3). This was due to the carrier signal getting subsumed into background noise. Nonetheless, a 3-meter range would be sufficient for all three of our tested locations (i.e., no point is greater than 3 meters from wall). It is also likely that longer communication range can be achieved by using a higher amplitude emitted signal, or by applying standard error correction techniques.

4.9 Example Uses

Touch sensing, pose tracking, and activity detection are well trodden ground in HCI. Additionally, Wall++ can work in concert with many existing feedback mechanisms, including screens (e.g., TVs, smart appliances, wearables), voice interfaces (e.g., Google Home, Amazon Echo) and ambient displays (e.g., smart light bulbs). We offer some illustrative example uses:

Touch Tracking, for example, could enable flexible-placement of wall-borne buttons to e.g., turn on/off lights, or provide a number keypad to unlock a door. Wall++'s continuous touch tracking could allow slider-like input to adjust e.g., light level, room temperature, or music volume; discrete swipes could be used to change lighting mode, or move between songs. Pose Tracking could allow users to play video games with their backs near to a wall and control avatars on a TV across the room. Pose tracking could also be useful in inferring human activity and context when users are near to work surfaces, e.g., making dinner vs. coffee on a kitchen countertop. Desks are often pushed up against walls, where Wall++ can detect the presence of a user's legs for occupancy tracking. In narrow hallways, we can track users' locomotion (e.g., direction, speed, gait), perhaps even identifying occupants.

Activity Recognition is made possible by Wall++'s ability to detect appliance operation, and then track that appliance in a room (and potentially a whole building). This rich source of contextual information can directly inform smart environments and assistive virtual agents. For example, a room can automatically adjust its lighting and window blinds when Wall++ detects a TV waking from standby. Users could also subscribe to alerts when certain appliances turn off, such as a laundry machine or electric kettle.

4.10 Limitations

Cost. Since walls are pervasive and expansive, the cost of any wall treatment to enable Wall++ needs be low in order to be plausible. Our recommended materials and antenna pattern cost $21.30 per m^2 for the small number of walls that we augmented for this project. While significantly less expensive than conventional touchscreen technologies, it is still expensive for e.g., a home (which might have 100 m^2 of walls). We believe replacing copper tape with conductive paint traces, as well as purchasing materials in bulk, could significantly reduce cost.

Installation Complexity. Although Wall++ does not require any special materials or equipment, it still requires a fair amount of wiring effort, as each row and column needs to be connected to a sensor board, presumably hidden in or behind the wall. We also found that applying paint evenly is challenging – our "final" 12×8' wall still showed some fabrication variance, as seen Figure 4.8. While within the capabilities of a home DIY enthusiast, it is probably beyond the skill and comfort level of the average consumer. However, it is possible to have commercially produced antenna-backed wall paper as a future mitigation of installation complexity and cost.

Interference. Environmental EM noise from e.g., fluorescent lights, can affect Wall++. This is a minor issue in active mutual capacitance sensing, as our excitation signal dominates the received signal. However, in passive EM sensing mode, environmental noise can have a significant impact on SNR. Among our three tested locations, workshop had the noisiest EM environment, which no doubt contributed to it having the lowest accuracies.

Nearby Grounded Objects. We found that well-grounded objects near to a wall, such as a TV, attenuates the shunting effect of a user's body, which in turn interferes with our mutual capacitance sensing. We found a similar effect with airborne EM signals. This finding suggests that real-world installations should avoid using (i.e., skip or disable) antennas that are proximate to such objects. This issue might also be mitigated by using a superior background calibration process and an analog frontend with a programmable gain.

Sensing Range. Our implementation of body pose and airborne EM sensing have limited sensing range (roughly 0.5 and 3 meters respectively). Fortunately, for appliance detection and localization, we found that most appliances in real world settings are close to walls, chiefly because electrical power is provided along the walls (and not in the middle of rooms). While there are some inherent sensing limitations, we do believe that range can be increased with superior circuit topology and software improvements in the future.

4.11 Summary

In this chapter, we introduced Wall++, a low-cost sensing technique that can turn ordinary walls into smart infrastructure, able to sense interactions and activities happening in a room, and potentially throughout an entire building. Our multi-phase exploration of materials, application methods, and electrode patterns informed our proof-of-concept hardware and software implementation. Then, through a series of user studies, we demonstrated that Wall++ can robustly track user touches and poses, as well as detect and track appliances (or tagged users) in a room. These sensing modalities are important building blocks for a wide range of smart environment applications.

Chapter 5

Sozu: Self-Powered Radio Tags for Building-Scale Activity Sensing

5.1 Introduction

Human environments are incredibly diverse, with potentially hundreds of facets that would be valuable to feed into intelligent systems that could enhance everyday tasks. Such systems may wish to monitor fixed infrastructure (e.g., doors, cabinet drawers, faucets, toilets), movable objects (e.g., kitchen utensils, personal hygiene items, tools) and larger appliances (e.g., microwave, refrigerator, stove, laundry machine, coffee maker). Unfortunately, contemporary sensing approaches are generally ill-suited to the scale, diversity, and construction of homes and offices.

In this chapter, I will describe my effort to overcome previous limitations with Sozu, a new low-cost sensing system that can detect a wide range of events wirelessly, through walls and without line of sight, at whole-building scale. To achieve this in a battery-free manner, Sozu tags convert energy from activities that they sense into RF broadcasts, acting like miniature self-powered radio stations. A Sozu deployment consists of one antenna, which can be placed in an inconspicuous location, such as the basement. Users then attach Sozu "tags" to items and infrastructure of interest. To keep cost low, these tags are constructed from ultra-low-cost analog components, and thus cost only a few dollars each (i.e., no digital components, nor digital communication like Wi-Fi or Bluetooth). We describe the results from a series of iterative studies, culminating in a deployment study with 30 instrumented objects. Results show that Sozu is very accurate, with true positive event detection exceeding 99%, with almost no false positives. Beyond event detection, we show that Sozu can be extended to detect the state, intensity, count, and rate of events.

5.2 Related Self-Powered Sensing Techniques

Most related to Sozu are approaches that are self-powered – sensors that capture energy from the environment (e.g., solar) or human activities (e.g., tool use). Self-powered sensing has seen some commercial success, for example, battery-free switches [13]. In the research literature, PowerBlade [53] sits on and siphons energy from electrical plugs to monitor energy consumption

of electrical appliances; WATTR [42] used variations in pipe water pressure for both sensing and energy harvesting, very much in the spirit of Sozu. Researchers have also harvested energy from powerlines [68], ambient temperature changes [171] and vibrations using the piezoelectric [132] and triboelectric effect [34, 35, 73].

Although these sensors do not require maintenance (meeting one of our ideal criteria), they all rely on digital components (microcontrollers and wireless modules), which increases cost. Additionally, these prior systems are specialized, focusing on particular categories of events, whereas Sozu aims to be a flexible and universal sensing system.

5.3 Deployment Overview

Sozu tags are essentially energy converters, turning energetic physical activities into wireless RF broadcasts. The system name – Sozu – comes from the traditional Japanese water feature, where a bamboo segment slowly fills with water, eventually causing its center of mass to shift, such that it pivots downwards, striking a rock (designed to scare away animals in gardens). This is similar in principle to our system, where one form of energy is converted into another for practical use.

The types of energy sources that Sozu can utilize dictates not only its RF emission strength, but also what facets it can sense in the environment. Therefore, our first step was to survey the environment for activities, classifying them into energy categories, and then investigating the effectiveness of harvesting implementations. Our next step was to determine how to best utilize this energy for radio broadcasts, taking into account our goals of long-range transmission without line of sight. We must also contend with, e.g., government regulations and background noise. This initial work informed our system implementation, which we describe in detail later.

5.4 Investigation One: Energy Harvesting

We identified seven energy categories from which Sozu tags could be powered. For each category, we describe example activities, along with proof-of-concept implementations and estimated cost. Later in this section, we discuss energy harvesting in practice using measurements we collected from 50 unique everyday objects/activities.

5.4.1 Motion

Motion is generated by a wide range of activities, including opening a drawer, popping a pill bottle, and pruning plants. To convert these motions into electrical energy, we used magnets to induce current in a wire coil (example integrations shown in Figure 5.1). We found two commercial harvesters that worked well for our purposes: a self-powered bicycle light ($2.70) and the ECO 200 energy bow made by EnOcean ($6.95).

Figure 5.1: Four Sozu-augmented objects that use motion energy: pill bottle, garden pruner, mailbox, and slipper

5.4.2 Vibration

Appliances such as food blenders and power tools generate vibrations that can be covered to electrical energy with piezoelectric materials. For this, we used a 50 mm piezo buzzer element ($1.36) weighted with a small nut. This harvester can be attached to movable objects with double-sided tape (Figure 2, left and middle). For stationary appliances, such as a grinding wheel, we can place the harvester underneath one of the feet (Figure 5.2, right).

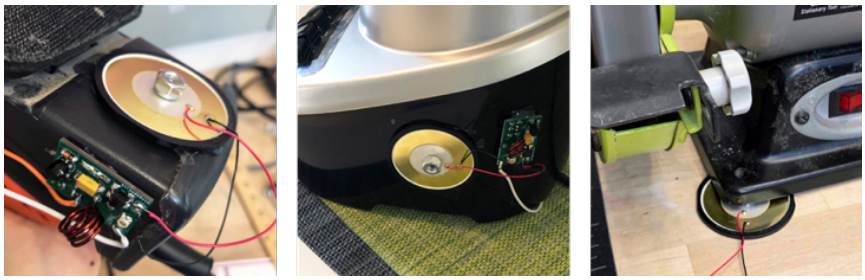


Figure 5.2: Sozu-augmented objects that convert energy from vibration: sander, blender, and grinding wheel.

5.4.3 Solar / Light

Light is another common form of energy, which can be used to infer events and activities. For example, illumination increases when room lights are turned on and refrigerator doors are opened (activating the internal light; Figure 5.3, left). Conversely, a car parking above a sensor or curtains being drawn decreases light level. To convert light energy, we use solar cells (58×55 mm Panasonic, $7.80).

5.4.4 Thermal

Many devices generate heat when operating, for example, gas-powered tools, hot glue guns, stoves and fireplaces. Sozu converts this thermal energy with Peltier junctions (~$2). We attach one side of a Peltier junction to an object of interest with thermal glue and tape, and on the other side, we

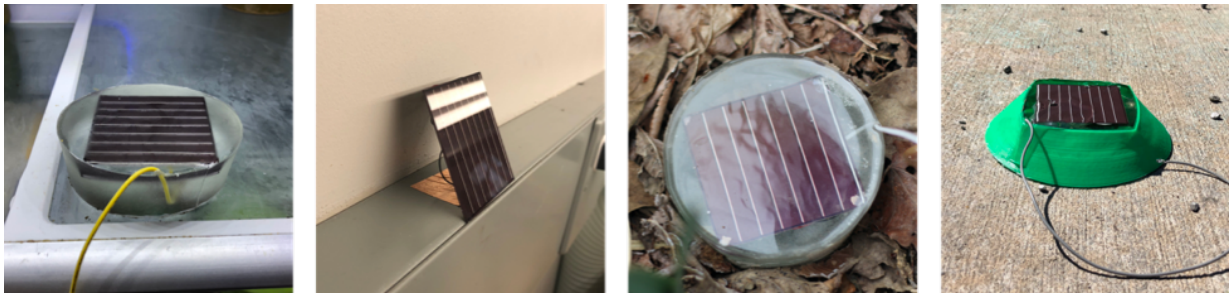


Figure 5.3: Sozu can also be powered by light energy: refrigerator door light, overhead work lights, and sunlight in a garden and parking lot.

use an aluminum heat sink to improve efficiency. Figure 5.4 shows four example instrumented objects.

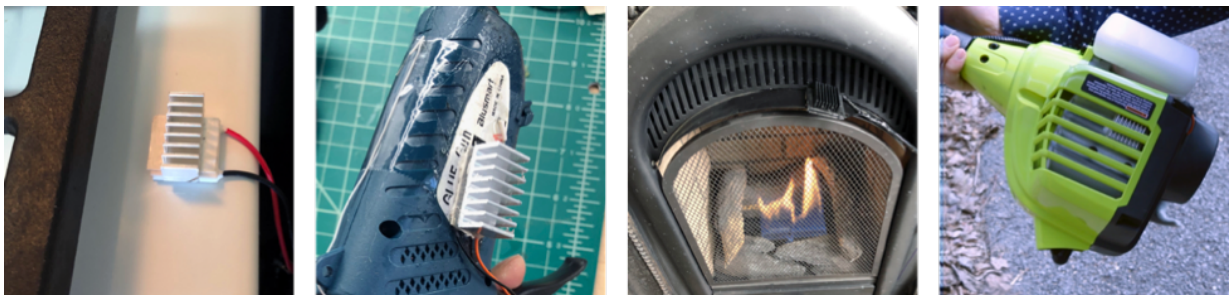


Figure 5.4: Sozu can be powered by hot surfaces, such as those found on stoves, hot glue guns, fireplaces, and gas tools.

5.4.5 Electromagnetic Radiation

Electrical appliances often emit electromagnetic (EM) radiation when in use, due to e.g., motors and switched-mode power supplies. We harvest this energy using a 60 mm diameter, 1500-turn coil with a ceramic core (~$5). These harvesters can be stuck (often magnetically) to devices, ideally close to motor coils and power regulation circuits (Figure 5.5).

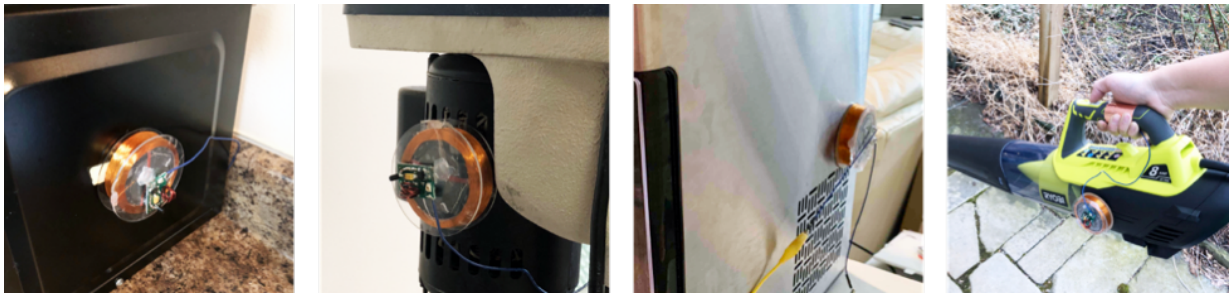


Figure 5.5: Sozu can be powered by EM radiation from objects such as a microwave, drill press, ice maker, and leaf blower.

5.4.6 Gas Flow

HVAC and machine exhaust generate flows of air and other gases, which can be easily harvested with DC brushless fans ($3.30), as seen in Figure 5.6.

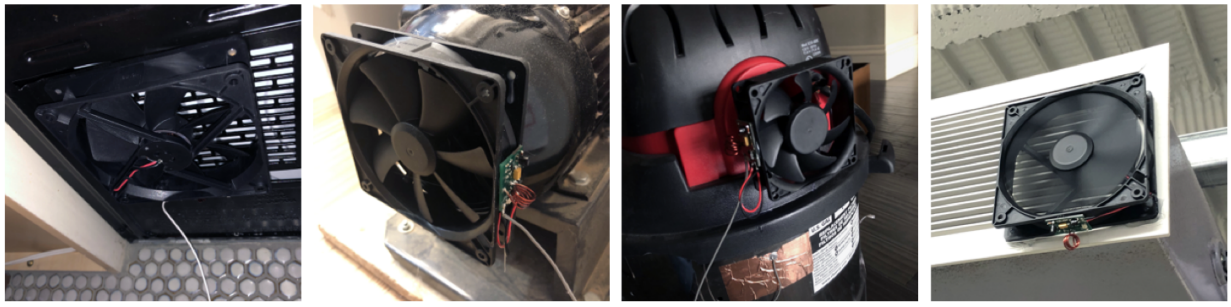


Figure 5.6: Sozu can be powered by gas flow from a kitchen hood, exhaust fan, ShopVac, and HVAC vent.

5.4.7 Water Flow

Finally, water flow is a high energy source, which Sozu can leverage to detect the use of faucets, showers, garden hoses, and even liquid containing vessels, such as watering cans (Figure 5.7). We selected two water-generators: one that fits US faucets ($3.5) and another that can be added inline to half-inch NPT water lines ($4).

Figure 5.7: Sozu can be powered by the flow of water from a sink faucet, garden hose, watering can, and showerhead.

5.4.8 Energy Harvesting in Practice

To better understand the energy budget for our Sozu implementation, we surveyed 50 objects (Table 5.1) to test energy harvesting in practice. For each test object, using harvesters described in the previous section, we recorded harvester output using an oscilloscope with a 10 kOhm load (close to the input impedance of our final Sozu tag design), which allowed us to estimate power output.

On average, our test objects yielded 2.7 mW of power (Table 5.1). We found our motion harvesters produced impulses, whereas other harvester types produced continuous output, either a constant signal (e.g., solar, thermal) or a periodic signal (e.g., vibration, EM radiation, gas and water flow). Overall, water flow produced the most power, followed by solar. Harvested power from the other energy categories was more variable, and more of a function of the object or activity. Finally, we found the Peltier junctions generated low voltages (<400 mV), and thus we augmented these harvesters with a LTC3108 boost converter ($6) to bring the output voltage up to a more usable 5V.

5.5 Investigation Two: RF Broadcast

With our power budget known, our next task was to identify an ideal frequency range for Sozu tag RF broadcasts. This was an iterative process, which we now describe.

5.5.1 Antenna Design & Multipath

According to antenna theory, the optimal size of an antenna is proportional to the wavelength of the radio wave. A 1 Mhz RF signal has a wavelength of 300 meters, much too long for practical use, even with e.g., a quarter-wave monopole antenna. For this reason, we use frequencies above 30 MHz (i.e., 2.5 m quarter-wave monopole antenna), approaching the physical size of commonplace objects (e.g., perimeters of appliances).

The wavelength of an RF signal also impacts its ability to penetrate obstacles, such as walls, floors and furniture. High-frequency signals (shorter wavelength) have more energy reflected by obstacles, which limits transmission range and also introduces multipath effects [31]. Therefore, we decided to only consider broadcast frequencies below 200 MHz.

5.5.2 Environmental Noise

To identify "quiet" parts of the RF spectrum in the range we identified previously, we used a USRP and ran a wide spectrum collection test from 1-200 MHz (62.5 kHz steps) using a wideband omnidirectional antenna. We collected data at six locations (two commercial buildings, two apartments, and two detached single-family homes, all in Pittsburgh) for 8 hours each. Figure 5.8 illustrates the power spectrum, which we computed by comparing the received power at each frequency with the noise floor (min within the current frame of powers). The range from ~90-100 MHz is densely occupied by FM radio stations, while ~30-35 MHz is used by mobile communication equipment. This left three relatively quiet bands: 15-30, 35-85, and 105-160 MHz. We dropped 15-30 MHz, as it did not provide much bandwidth for many tags, leaving us with two candidate ranges.

5.5.3 FCC Regulation

Sozu must also comply with FCC regulations, which permits low-power RF transmission in unlicensed frequencies [10]. However, there are many restricted frequencies that must be avoided.

Object Name	Activity or State	Energy Type	Waveform	Power (μW)
Faucet	Water running	Water flow	Periodic	26538.2
Shower Head	Spraying water	Water flow	Periodic	23215.2
Garden Hose	Spraying water	Water flow	Periodic	23105.2
Gas Trimmer	Engine running	Thermal	Constant	11700.0
Powered Drill	Drill spinning	EM radiation	Periodic	7227.1
Window Sunlight	Sunlight	Solar	Constant	4027.1
Refridgerator	Door open / light on	Solar	Constant	3804.3
Parking Spot (light)	Lit by sun	Solar	Constant	3609.8
Floor Lamp	On	Solar	Constant	3287.5
Garden Sunlight	Lit by sun	Solar	Constant	3164.0
Ice Maker	Making ice	EM radiation	Periodic	3071.6
Work Light	On	Solar	Constant	2759.0
Hedge Trimmer	Trimmer on	EM radiation	Periodic	2736.7
Hot Glue Gun	On	Thermal	Constant	2599.8
Toilet	Flushing	Motion	Impulse	1771.6
Hairdryer	Blowing hot air	Gas flow	Periodic	1634.5
Kitchen Hood	Fan exausting	Gas flow	Periodic	1521.1
ShopVac	Vacuuming	Gas flow	Periodic	1519.7
Hand-Held Vacuum	Vacuuming	Gas flow	Periodic	1190.4
Mitersaw	Blade spinning	EM radiation	Periodic	1176.1
Microwave	Running	EM radiation	Periodic	684.0
KitchenAid Mixer	Mixing	EM radiation	Periodic	466.0
HVAC Vent	Vent blowing	Gas flow	Periodic	390.2
Pruner	Cutting	Motion	Impulse	362.1
Watercan	Watering	Water flow	Periodic	348.8
Leaf Blower	Blowing air	EM radiation	Periodic	266.5
Air Purifier	Filtering air	Gas flow	Periodic	241.8
Drill Press	Drill on	EM radiation	Periodic	225.1
Projector	Projecter on	Solar	Constant	218.5
Sander	Sander on	Vibration	Periodic	217.5
Exhaust Fan	Exhausting	Gas flow	Periodic	205.3
Reciprocating Saw	Running	Vibration	Periodic	160.5
Fireplace	Fire lit	Thermal	Constant	119.4
Pill Bottle	Opening lid	Motion	Impulse	116.4
CNC	Router spinning	EM radiation	Periodic	115.7
Toaster	Toasting	Thermal	Constant	113.7
Battery-Powered Drill	Dill spinning	EM radiation	Periodic	106.6
TV	On	Solar	Periodic	81.0
Coffee Grinder	Grinding coffee	EM radiation	Periodic	43.9
Sliding Door	Opening / closing	Motion	Impulse	43.8
Blender	Blending	EM radiation	Periodic	41.9
Kettle	Boiling water	Thermal	Constant	41.0
Slipper	Walking	Motion	Impulse	29.2
Drawer	Opening / closing	Motion	Impulse	23.7
Mailbox	Indicator flag raised	Motion	Impulse	23.7
Grinding Wheel	Grinding	Vibration	Periodic	14.7
Solder Iron	On	Thermal	Constant	10.4
Blender	Blending	Vibration	Constant	6.8
Gas Stove	Stove lit	Thermal	Constant	4.8

Table 5.1: Mean energy generated by 50 activities we tested.

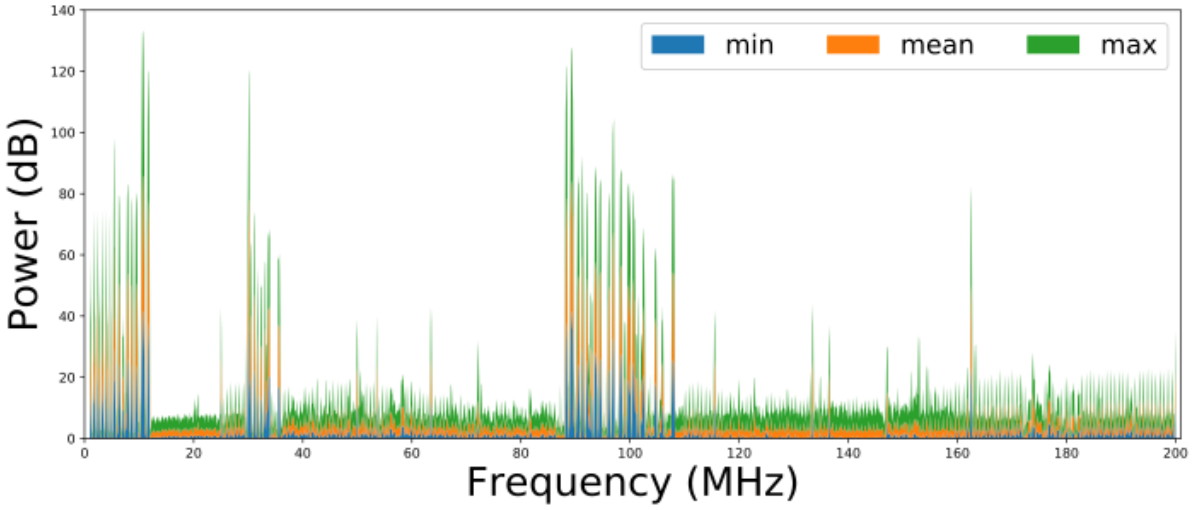


Figure 5.8: RF power density from 1 to 200MHz.

This eliminated the promising 105-160 MHz range from consideration, as it was mostly restricted. Thus, we ultimately selected the 35-85 MHz range for Sozu. In this frequency range, transmitters are limited to a field strength of 100 μV/m at 3m, which is equivalent to 3 mW transmitter power assuming isotropic radiation. This means, with an ideal passive antenna design (Power Gain = 1, impedance = 50 Ohm), the peak voltage (V_{peak}) of a sinusoidal RF signal has to be lower than 0.54 V. We strictly followed this in the Sozu RF circuit design to meet FCC guidelines.

5.5.4 Building Obstacle Penetration

From theory and literature, we knew our selected frequency range should have good building penetration. Nonetheless, we wished to measure this directly with our own equipment and across typical construction styles. We tested five walls: brick (30 cm thick), drywall + metal studs (15 cm), glass (2 cm), stone/masonry (65 cm), and precast concrete (20 cm). To capture data, we transmitted a swept-frequency signal from 35-85 MHz on one side of the wall, and measured the signal strength on the other side (symmetric 5 m separation between transmitter and receiver). We did not find significant differences in attenuation across frequencies, and so we average these results. Across all wall construction types, we found a mean attenuation of -12.04 dB (SD=7.83). The concrete wall had the most attenuation (-20 dB), whereas glass and drywall + metal studs had little effect (-4.10 and -3.30 dB respectively).

5.6 Sozu Implementation

A Sozu deployment consists of two main components: Sozu tags and an antenna receiver with attached computer. Sozu tags are distributed in the environment, transforming energy from activities into radio broadcasts. The antenna receiver is deployed at one central location (e.g., basement of a house), which monitors tag broadcasts. We now describe these components in detail.

5.6.1 Sozu Tags

Sozu uses a custom tag (Figure 5.9 and 5.10) which connects to an energy harvester (see Investigation One) and an antenna (discussed subsequently). The board itself is responsible for managing energy and generating an RF signal with a power consumption of 2.05 mW.

Energy Management: As discussed previously, most of our energy harvesters provided periodic signal, which we rectify with an ultra-low-forward-voltage rectifier. As the incoming power may not be enough energy to run our circuit, we designed an energy storage and latch mechanism to store a small amounts of energy, which can later be released. We use two capacitors (C_1 and C_2) in series to store rectified energy (Figure 5.9); the voltage at the junction between the two capacitors controls the gate of a silicon-controlled rectifier (SCR). A trigger voltage of 1.5 V activates the SCR, which then passes the stored energy to the RF oscillation circuit. The SCR turns off once the RF oscillation consumes all energy, and then the tag begins to store energy again for the next activation. Releasing the stored power results in an RF chirp lasting roughly half a second (e.g., Figure 5.11 right). We note that some activities provide sufficient continuous energy (e.g., water faucet, HVAC vent) that the SCR is always on, resulting in a continuous RF output (e.g., Figure 5.14 right).

Different transmission strengths and broadcast chirp durations are achievable by tuning the capacitance of C_1 and C_2. The ratio between C_1 and C_2 sets the initial voltage that powers the RF front end when the SCR is on. A smaller C_1:C_2 produces a higher initial voltage, which yields stronger signal strength and extended transmission range. However, we note that C_1:C_2 must be higher than 0.39 so that the FCC requirement is not exceeded.

To increase the duration of RF broadcasts, larger capacitors can be used. For our later evaluation, we used a C_2 of 100 μF and a shorted C_1. This was because all of our activities generated sufficient power, which did not need further energy storage. With this configuration, the minimum energy to generate one RF broadcast is 0.12 mJ, which is roughly the amount of energy needed to charge C_2 up to 1.5 V (i.e., the trigger voltage). This configuration also had the effect of limiting transmission power to below the FCC regulation.

RF Oscillation: To generate RF broadcasts, Sozu tags use an oscillation circuit based on a single-transistor LC oscillator. The output frequency can be calculated by:

$$f = \frac{1}{2\pi\sqrt{LC}} \tag{5.1}$$

where L is the inductance of a coated copper coil which measures 420 nH. C is the capacitance that is dominated by a user-adjustable capacitor (range of 8-50pF), which permits a tag to be set to a specified frequency. One version of our tag used a simple trimmer capacitor (adjustable with a screwdriver, Figure 5.10, left), while our second design used a 4-bit dip switch to connect/disconnect four capacitors (1, 2, 4, and 8 pF) in series with a 30 pF capacitor, treating frequency more like a binary ID (Figure 5.10, right).

We note that the transistor Miller capacitance (C_{Miller}) between base and common also contributes to C, through being in series of C4 and C3, and parallel to the trimmer capacitor. Considering the Miller capacitance, the C component of the oscillation circuit is calculated by:

$$C = \frac{1}{\frac{1}{C_{Miller}} + \frac{1}{C_3} + \frac{1}{C_4}} + C_{Trimmer} \tag{5.2}$$

This Miller capacitance is affected by the supply voltage, resulting in very basic frequency modulation (FM). We observed a variance of Miller capacitance (C_{Miller}) of around 1 pF with a supply voltage from 1.5 to 5 V, resulting in a frequency shift of ~0.5 MHz. This effect allows us to demodulate analog signals on the receiver end, enabling richer sensing opportunities, which we discuss later.

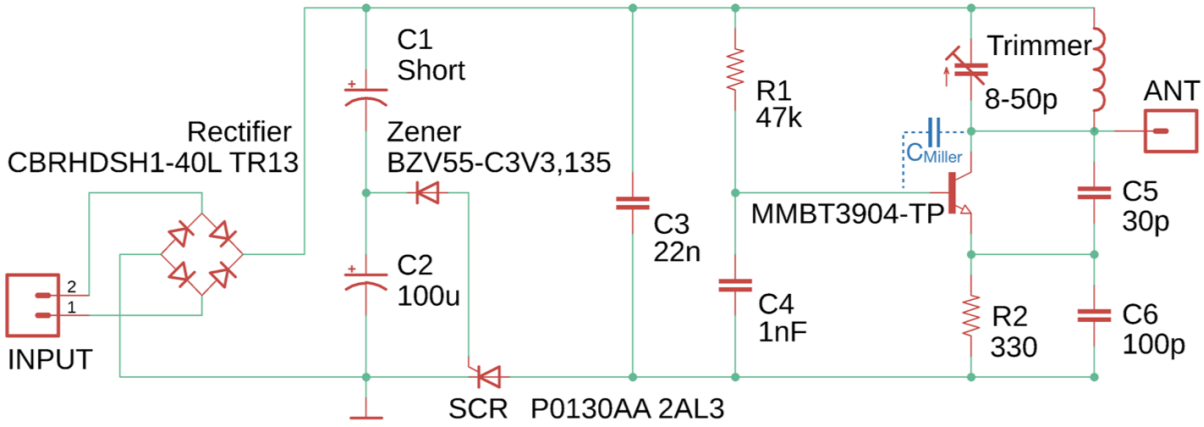


Figure 5.9: Sozu tag circuit schematic.

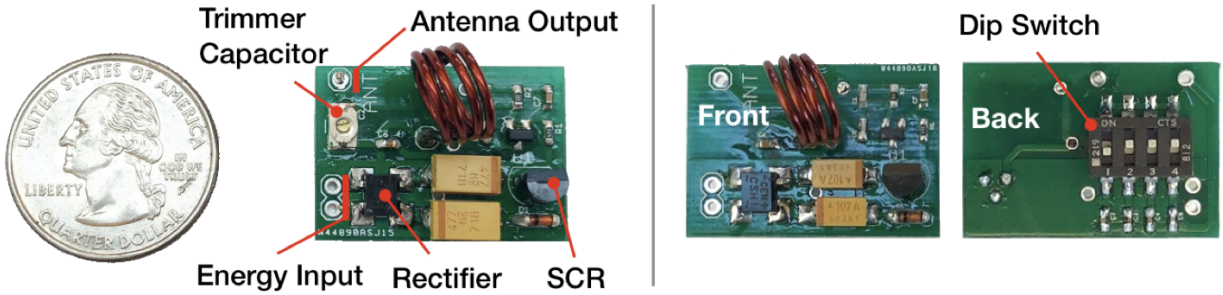


Figure 5.10: Two Sozu tag designs, one with broadcast frequency tuned with a trimmer capacitor (left) and one with a dip-switch (right).

5.6.2 Transmit Antenna

We investigated a wide variety of antennas for our Sozu tags, including chip antennas and PCB antennas. Ultimately, we selected simple monopole antennas, which can be easily integrated into many objects. We varied our antenna lengths (from 5 to 100 cm) for different objects, depending on their output power. In general, objects that produce less energy require longer antennas to help

maintain broadcast range. We used generic 18-gauge braided wire for our antennas, which was easily trimmed to a desired length.

Our wire antennas can often be hidden by making clever use of an object's geometry (e.g., running down a mailbox post), or simply be tucked behind larger objects and infrastructure. For smaller and mobile objects, where long antennas are more problematic, we found two alternatives. First, for objects with metallic enclosures (e.g., kitchen appliances), we can use the shell itself as an antenna. Second, the human body can be co-opted as a large antenna; for this, we connect our RF output to a copper patch where the user would grasp an object (example shown in Figure 5.14).

5.6.3 Bandwidth & Concurrent signals

In order to support detection of many concurrent activities, Sozu requires each tag to operate at a unique frequency. The frequency stability (i.e., bandwidth) of Sozu tags therefore decides the maximum number of concurrent activities that the system can sense simultaneously. To quantify this, we measured the frequency shift of a Sozu tag while varying the supply voltage from -5 to +5 V. We observe an average bandwidth of 0.52 MHz (SD=0.12). This suggests that Sozu could support up to ~96 tags (i.e., (35-85 MHz) / 0.52 MHz bandwidth). In practice, we found that moisture and temperature can affect the capacitance of our LC oscillator circuit, which further shifts the registered frequency. To account for this, we suggest using a 1 MHz bandwidth, especially for outdoor objects and activities (e.g., mailbox, parking spot).

5.6.4 Environmental Hardening

Outdoor placements, such as a garden or parking lot, require resilience to moisture and temperature change. To survive these environments, we fully encased some of our tags in clear epoxy resin (examples in Figure 5.3). The only external element was our wire antenna. This solved tag failures from moisture and extreme temperatures, but it did not solve small shifts in our LC oscillator circuit from temperature changes.

5.6.5 Sozu Receiver

To capture RF signals emitted by Sozu tags, we used a HackRF One software defined radio (SDR) connected to a wideband omnidirectional antenna. We recommend this antenna be placed in a discrete, but central location, such as an attic, closet or basement (Figure 5.11). The SDR streams data to a laptop over USB, after which it is processed by GNU Radio, which computes FFTs (non-overlapping windows, 8192 samples). This yields a stream of ~2400 FFT results per second (i.e., FPS), which are streamed over a local socket to a custom Java program for further computation and interactive control.

5.6.6 Activity Recognition

Detecting the presence of activities (i.e., on/off) can immediately power a wide variety of smart building applications, including automatic lighting with occupancy detection, alerts for unattended

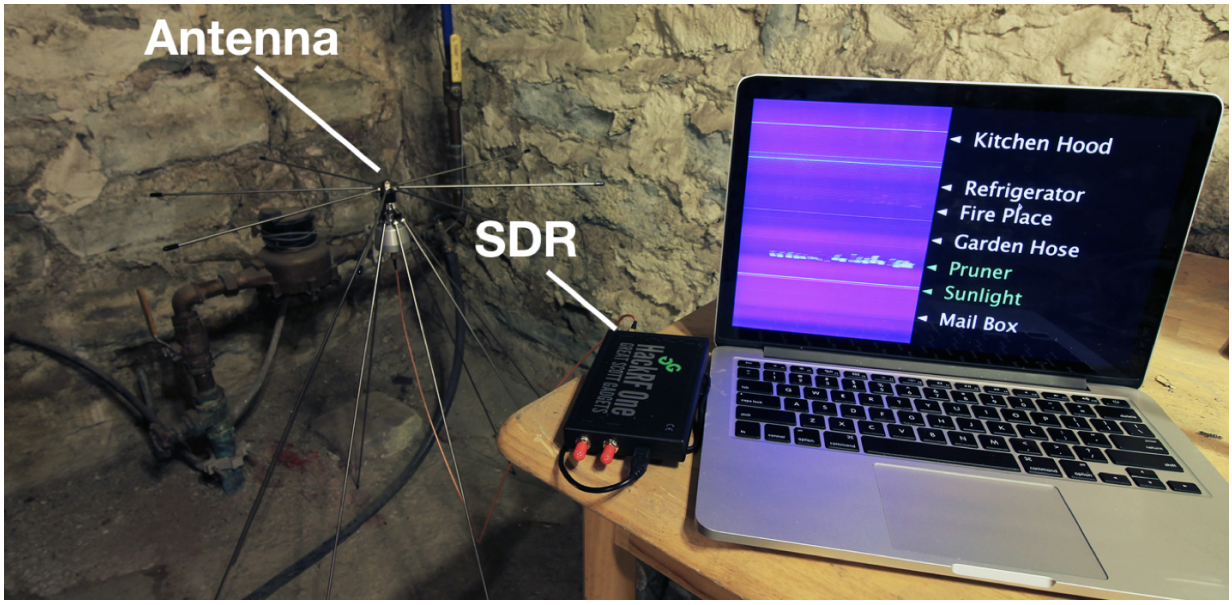


Figure 5.11: Sozu receiver setup deployed in a basement.

stoves, and medication reminders with smart pill bottles. As noted previously, detecting the presence of activities is equivalent to detecting the presence of RF signals at registered frequencies. Specifically, for each registered frequency, we find the min and max FFT bin within a ±0.5 MHz window and compute the difference. We found this range value compensates for varying antenna sensitivity across frequencies and interference from wide-band environmental noise (e.g., fluorescent light). To increase stability, we smooth this value with an exponential moving average, and then use a basic threshold to decide if an activity is on or off. This decision is further stabilized with a majority voter (300 ms history).

5.6.7 Open Source & Toolkit

To facilitate replication and deployment, we have open sourced our Sozu tag PCB design files and deployment software: `https://github.com/FIGLAB/Sozu`

5.7 Accuracy Evaluation

We deployed Sozu tags at three locations – an apartment, a detached house, and a commercial building. These environments offered different objects, room functions, and construction types. At each location, we deployed our Sozu receiver in an inconspicuous location, and augmented ten commonplace objects with Sozu tags (Figure 5.12). Of the 30 activities we chose to evaluate, 22 have been identified as important in prior work [60, 69, 91]. The eight "new" items are pruner, hedge trimmer, drawer, mailbox flag, pillbox, hot glue gun, garden sunlight, and window sunlight. The aforementioned papers were unable to sense unpowered handheld items (e.g., pruner) and distant unpowered objects (e.g., mailbox flag), and thus these items help to underscore Sozu's new capabilities.

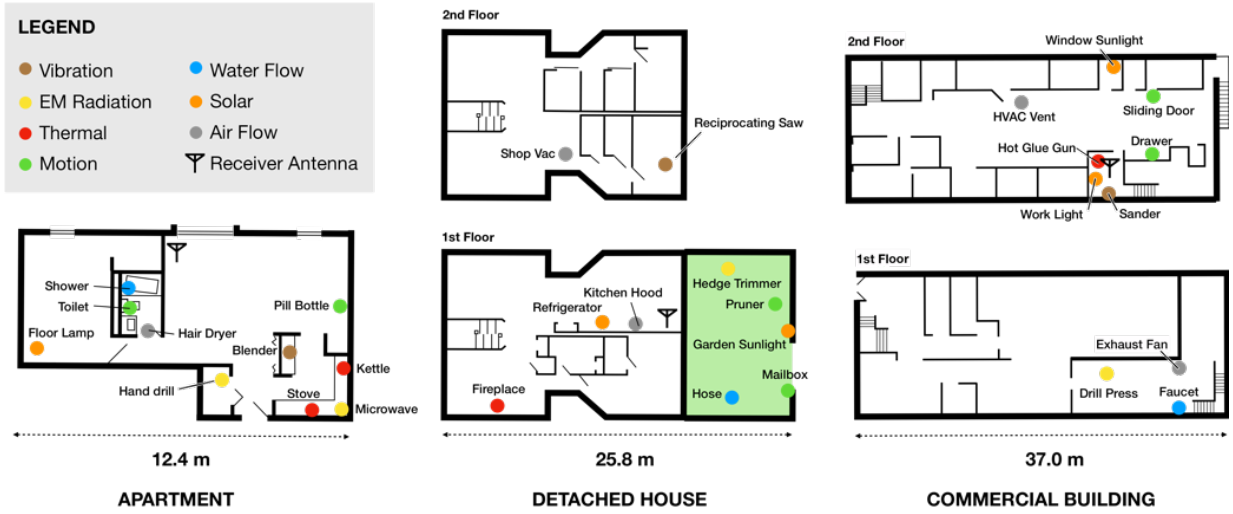


Figure 5.12: We deployed Sozu at three test locations. Objects in our accuracy evaluation are denoted by colored-coded circles.

5.7.1 Procedure

Our 30 tags were deployed for three weeks. At the end of each week of operation, the experimenter visited each location to check the durability of the Sozu tags, as well as test detection accuracy. To start, all activities were turned off (sunlight powered tags were covered), after which the experimenter collected ten instances of Sozu detections, spaced apart by one minute. Activities were then manually activated one at a time (random order), and a single instance was captured. The experimenter then turned the activity off (or waited for the activity to naturally end, e.g., toilet flush and refill) and repeated the collection process nine more times for that activity. Note that every "on" instance for a particular activity gave us nine "off" instances for all other activities. In sum, for each activity, we collected 30 "on" trials (10 instances × 3 weekly collections) and 300 "off" trials (10 instances × 3 weekly collection when all objects were "off", plus 9 other "on" objects per location × 10 instances × 3 weekly collections). In total, across all 30 activities we studied, we collected 9,900 on/off instances.

5.7.2 Result

All tags were functional at the end of the deployment (including those placed outside. The overall accuracy of on/off activity detection was 99.94%, with an even split of false positives and false negatives. Though this high accuracy is encouraging, we caution that further long-term deployment studies (multiple months, more locations) are required.

5.7.3 Latency

For all harvesting methods, other than thermal, the time between activity actuation and RF broadcast is under 500 ms. Our thermal harvesters required at least a 30 °C temperature differential

before generating sufficient voltage, and the large thermal mass of our test objects simply required time to heat up.

5.8 Beyond On/Off Detection

While on/off activity detection is a powerful building block, smart building applications can also benefit from fine-grained information, such as directionality and rate of activities. Sozu can offer such information with simple extensions of its detection process.

5.8.1 Directionality

Some objects have complex states with directional activities. Doors are a quintessential example, as a closed door has a very different function and meaning than an open one, and simply knowing "door moved" is not sufficient. As an example implementation of direction sensing, we strategically mounted magnets on the rail of a sliding door (Figure 5.13). This asymmetry results in two sets of chirps (a pair and a single chirp), the order of which can be used to infer direction, and thus door state.

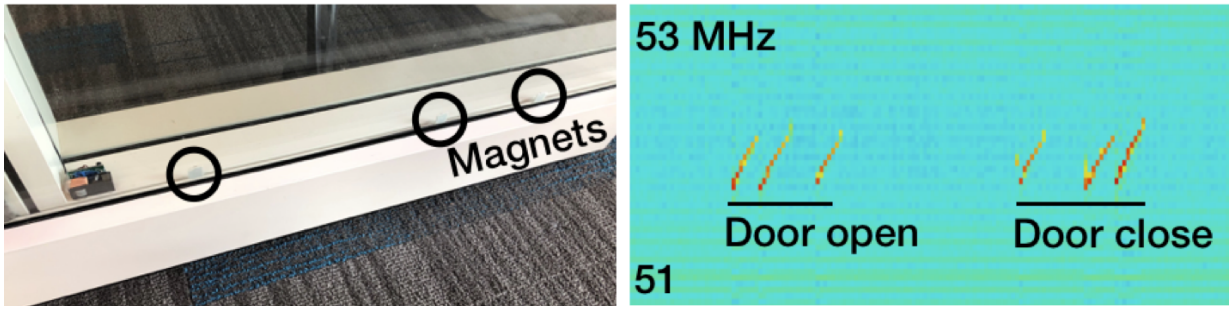


Figure 5.13: Strategic placement of magnets on a door frame allows the direction of a sliding door to be sensed.

5.8.2 Grasp Detection

Sozu can also detect when a user is grasping an object. We achieve this by connecting the user's body to our LC oscillation circuit with a conductive patch, which characteristically lowers the broadcast frequency. As an example, we augmented a hot glue gun (Figure 5.14), allowing Sozu to not only know if the glue gun is on, but also if the user is actively using it. With this data, we could e.g., alert a user if they have left a glue gun on without use for more than 15 minutes.

5.8.3 Intensity

We can also leverage frequency modulation to encode analog signals. Specifically, we calculate the frequency difference between the current frequency (f_{curr}) and the registered frequency (f_{base}) of a tag. This frequency shift correlates to the output voltage of the harvester, which can be

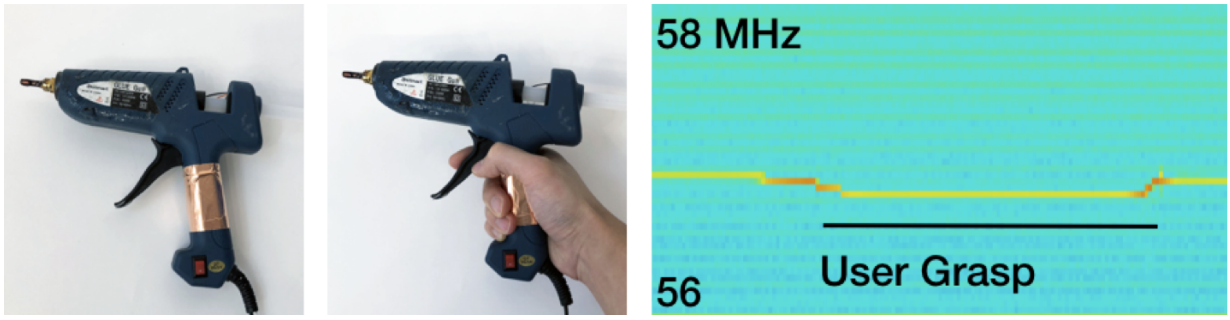


Figure 5.14: By including the user in the LC oscillator circuit via a conductive grip (copper tape), Sozu can detect user grasp, which manifests as a characteristic frequency shift.

associated with an analog dimension of the activity. As a demonstration, we placed a light meter (for ground truth) side-by-side with a Sozu tag powered by a solar cell (fbase = 48.6 MHz). We used a dimmable light to vary the environmental illumination from 312 to 1279 Lux. We recorded both the measured light intensity and the received tag frequency. Figure 15 (top) plots some example signals, and further shows that it is fairly straightforward to correlate frequency shift with light intensity (Figure 5.15, bottom).

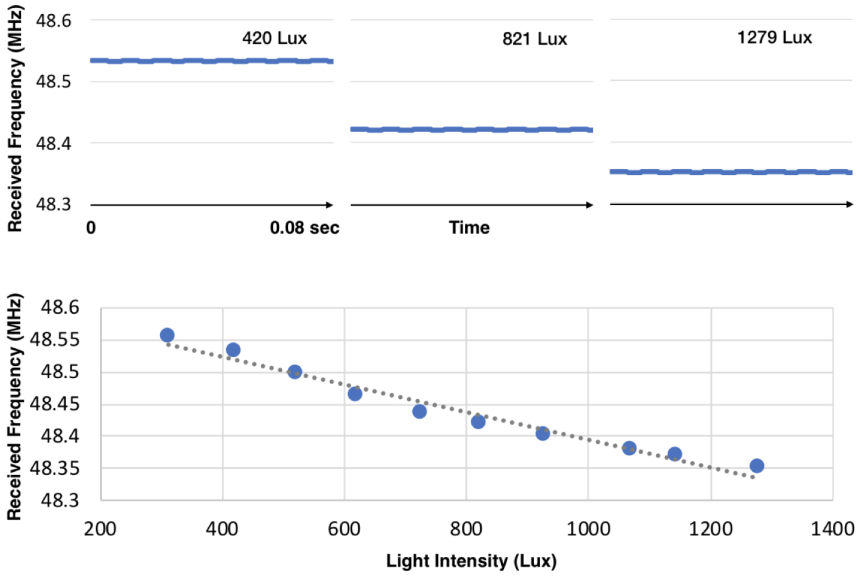


Figure 5.15: Top: received Sozu tag frequency at different light intensities. Bottom: mean received frequency vs. light intensity with linear regression plotted.

5.8.4 Rate

Sozu can also detect the rate of activities, such as water flow volume and motor RPM. In these cases, the energy harvesters provide periodic signals that correlate with the rate of the activity. This manifests as an oscillating frequency shift (see Figure 5.16, top), the rate of which can be measured by counting zero crossings.

As an example, we captured data from a gas-flow-powered Sozu tag at different wind speeds using a DC fan (placed next to an airflow meter for ground truth). The raw FM signal can be seen in Figure 5.16 (top), as well as the linear relationship between the detected frequency oscillation rate and actual wind speed (bottom). We performed a similar experiment for water flow; Figure 5.17 shows the linear correlation between the frequency oscillation rate and true water flow rate.

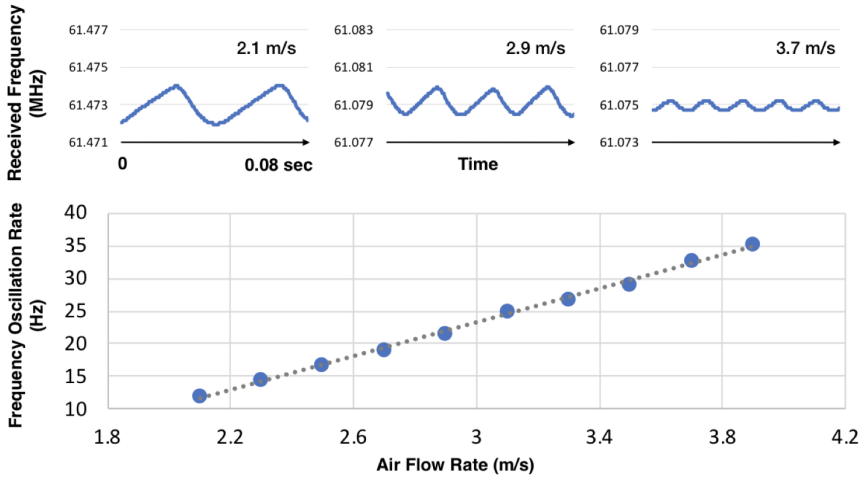


Figure 5.16: Top: received Sozu tag frequency at different wind speeds. Bottom: the linear correlation between Sozu-detected frequency oscillation rate and actual wind speed.

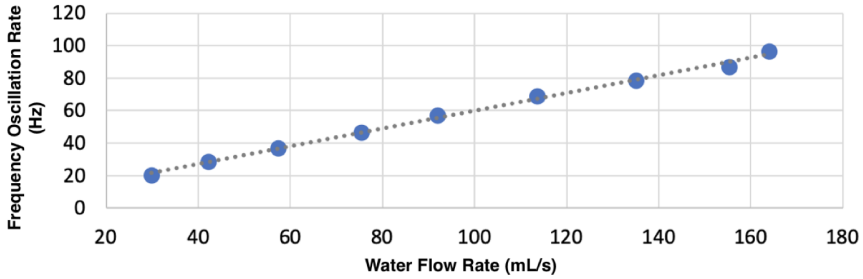


Figure 5.17: The linear correlation between Sozu-detected frequency oscillation rate and water flow rate.

5.9 Tag Localization

Knowing the location of activities would also be useful for smart buildings systems. For example, a shop manager may wish to track the location of power tools for safety reasons, while facilities staff may want to see what rooms have been vacuumed recently. Additionally, in multi-occupant environments (e.g., apartment buildings), a personal Sozu system should only decode a user's activities and not neighbors'. This could be also achieved by using different tags (with different registered frequencies) in different rooms (e.g., each faucet is unique), but this approach does not scale or work for movable devices, such as power tools and vacuum cleaners. Instead, one

could use tags registered to the same frequency (e.g., all faucets use the same frequency), and their location serves as extra bits of information for disambiguation. This approach would also expand the number of tags we could support in our 35 – 85 MHz frequency range.

Since RF broadcasts attenuate while propagating through space, we used multiple antennas to triangulate broadcasts and thus localize tags. Specifically, we used four receiver antennas and the Wentzel-Kramers-Brillouin (WKB) approximation [80] to model RF propagation:

$$P_R(q,b_i) = P_T(q) + L(q,b_i) + \gamma \sum_j d_j \eta_j \tag{5.3}$$

where $P_R(q,b_i)$ denotes the received signal power at the i-th antenna for the tag q, $P_T(q)$ denotes the transmission power, $L(q,b_i)$ is the path loss power, and $\gamma \sum_j d_j \eta_j$ is the shadowing term, which captures the impact of the attenuation because of the wall type η_j and thickness d_j. These factors were calculated using data from our previous Building Penetration study. We then corrected the signal power with a shadowing term, inferred the distances between the tag and the base stations using the path loss power, and used a least square error approach to compute the location.

5.9.1 Procedure

For testing, we deployed antennas at the four corners of a 13×5 m office space (illustrated in Figure 5.18). We first calibrated antenna sensitivity by collecting data from a Sozu tag of a known power on a 1 m grid on the first and second floors, as well as outside of the building.

After calibration, we used a gas-flow-powered Sozu tag powered by a small DC fan (simulating a vacuum cleaner) and collected data at 18 locations (8 points on the second floor, 6 points on the first floor and 4 points outside the building – see Figure 5.18). At each location, we collected 20 seconds of data, yielding 14,400 data points in total. We then calculated the Euclidian distance error between the true location and Sozu's estimated location to quantify the performance of our localization approach.

5.9.2 Results

Sozu was able to localize tags with an average error of 4.1 m (SD=1.9). Figure 5.18 breaks this result out across the 18 locations we tested. Although coarse, it is sufficiently accurate for room-level localization. This means if there are several e.g., faucets in a house, they could use a common frequency and be disambiguated by location data. We also used our study data to simulate zone-level localization (i.e., first floor, second floor, and outside), and in this case, all points were correctly localized.

5.10 Sozu Toolkit

To lower the barrier of entry in applying Sozu to activity sensing applications, we created a toolkit that consists of a Sozu tag connected to a solar cell, an RTL-SDR, a set of antennas, and demo code. We also wrote a quick start guide on how to set up the software, deploy

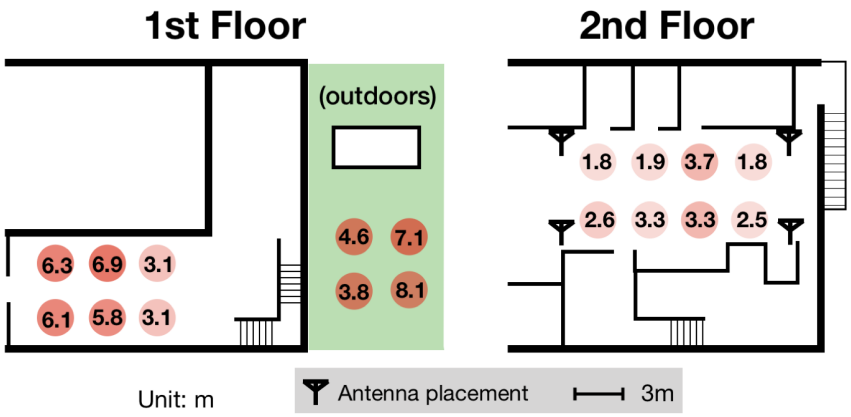


Figure 5.18: Locations (circles) tested in our localization study, with mean Euclidian error (meters) provided inside circles.

Sozu tags, and receive signals. All materials can be found in the Sozu repository: `https://github.com/FIGLAB/Sozu/tree/master/toolkit`.

5.10.1 User Study

To evaluate our Sozu toolkit, we conducted a user study with 8 students taking classes in Human-Computer Interaction at our institution (6 undergraduate and 2 graduate students). Among these students, four rated themselves as having little hardware development experience. Each student was given a Sozu toolkit and provided with any harvester they requested for their project. Students were responsible for generating an application concept, implementing their idea, and documenting their project with a video (see Auxiliary Materials). The students were given one week and were compensated $50 for their time. We recorded how long it took participants to set up the Sozu sensing pipeline, from receipt of the toolkit package to displaying signal on their laptop. We also kept track of bottlenecks, failures and successes with daily interviews.

5.10.2 Results

All participants successfully deployed Sozu tags and finished their projects within the one-week deadline. On average, it took 36 minutes (max 2 hours) to set up Sozu on their personal computers. We note that most of the time was spent on Python environment configuration. Participants found the solar cell we included in the toolkit package particularly helpful in getting started (i.e., "signal out of the box", no soldering needed). Participants also found our guide on how to harvest energy from different sources useful, though it was still the most challenging part of the project according to their feedback.

5.10.3 Example Student Projects

We now briefly describe the student projects, which both illustrate potential uses of Sozu and also the ease at which our approach can be deployed. See also Auxiliary Materials for student-made

videos of their projects.

Parking Occupancy: One student made a parking sensor with Sozu, which is powered by a solar cell. When the output frequency dips (or turns off entirely), it is inferred that a car has parked at the spot above the sensor (Figure 5.19).

Figure 5.19: Locations (circles) tested in our localization study, with mean Euclidian error (meters) provided inside circles.

Medication Reminder: Another student used Sozu to build a smart pill bottle that can detect if its lid is opened. If the lid is not opened when medication is scheduled to be taken, an alert can be sent to the user (Figure 5.20).

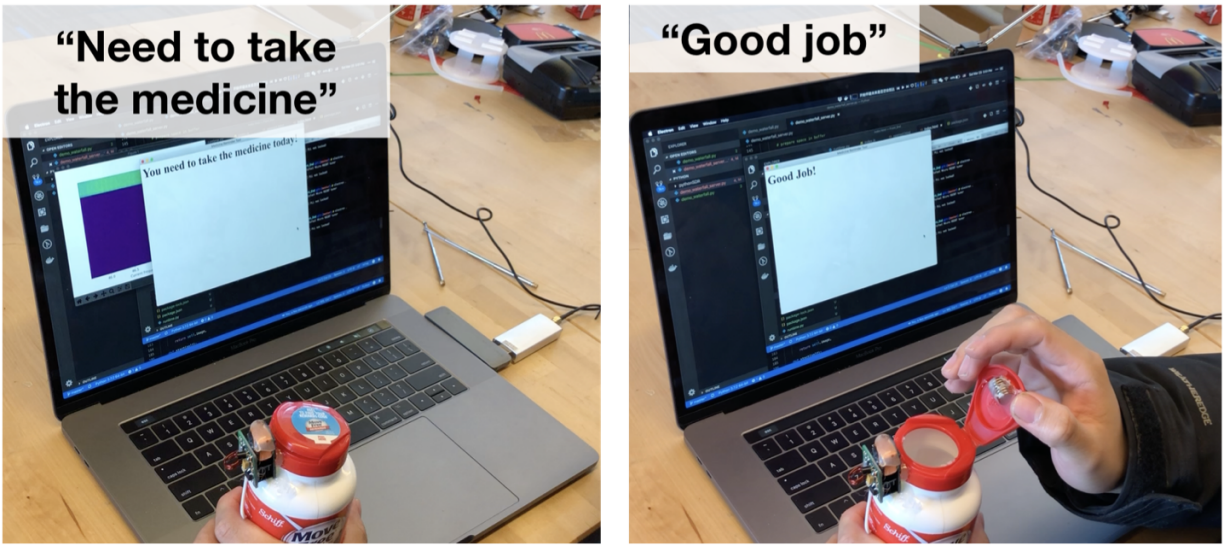


Figure 5.20: A Sozu-augmented pill bottle can detect when its lid is opened, or not opened, in which case an alert is sent.

Smart Recycling Bin: In this project, a student used Sozu to detect if a recycling bin was full, based on whether a solar-powered tag near the top was blocked by refuse (Figure 5.21).

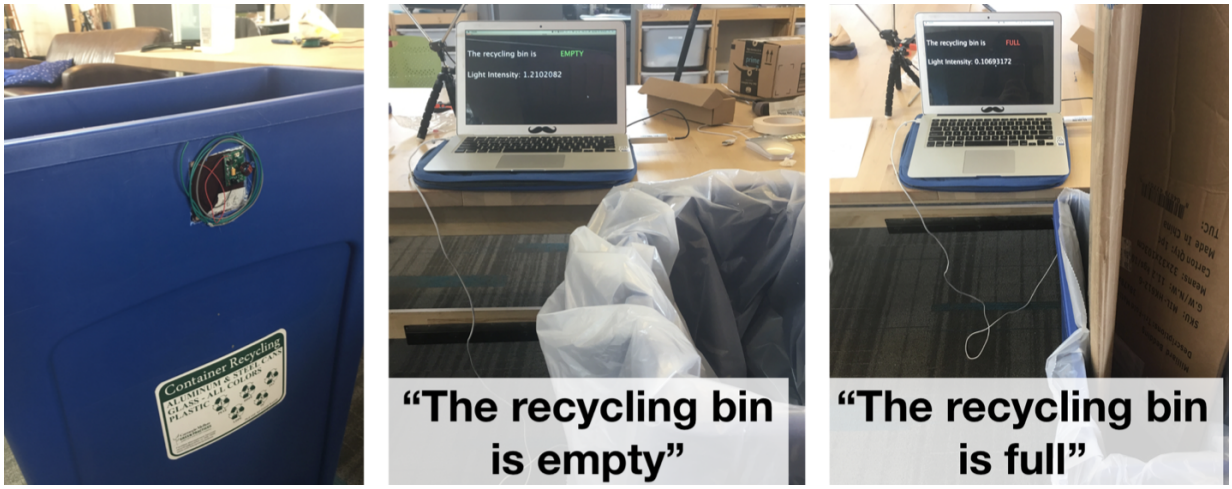


Figure 5.21: A recycling bin that alerts facilities staff when full.

Automatic Door for Wheelchairs: One student attached a motion-powered Sozu tag to a wheelchair and placed a receiver antenna by a doorway. The idea was to automatically open the door once a wheelchair's presence was detected. The sensitivity and orientation of the receiver antenna was tuned such that wheelchairs were only detected when ~1 m away (Figure 5.22).

Figure 5.22: A door could automatically open when a wheelchair approaches.

Consumables Monitoring: One of our student participants instrumented a bin in the lab containing tapes with a motion-powered Sozu tag. This allowed for automatic tracking of how often the bin was accessed. After a certain threshold, a reminder is sent to the shop manager to check on supplies (Figure 5.23).

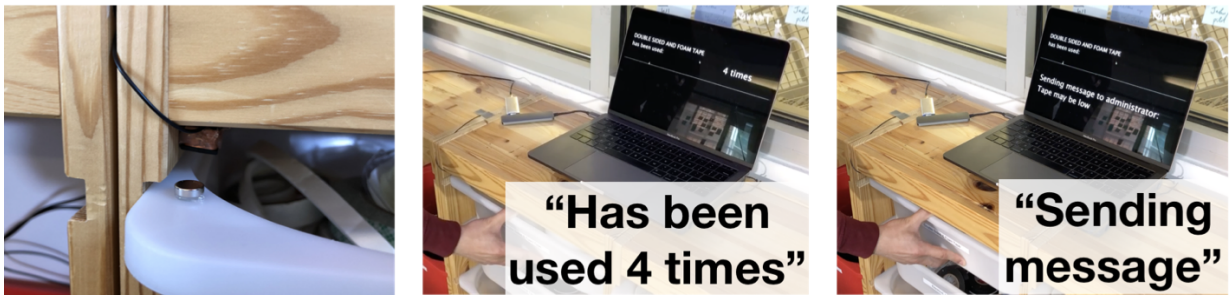


Figure 5.23: A Sozu-augmented storage bin can alert a shop manager to check on supplies after repeated access.

Foot Traffic: In this student project, a solar-powered Sozu tag was used in concert with a wall-powered laser pointer to count how many people passed through a doorway.

Meeting Room Occupancy: This project had a student place a solar-powered Sozu tag at the corner of a projection screen to detect when the digital projector was turned on, inferring meeting room occupancy.

Work Hour Monitoring: Lastly, a student made a chair occupancy sensor using a solar-powered Sozu tag cut into the back of a chair. Light is blocked or attenuated when someone is seated, and the student used it to track how long they are seated for work.

5.11 Summary

In this chapter, we discuss Sozu, a low-cost, self-powered, building-scale activity sensing approach. Instead of batteries, Sozu utilizes energy produced as a byproduct from many everyday activities. We optimized Sozu's implementation by first investigating energy harvesting opportunities. We then conducted a series of RF investigations to explore how to best utilize harvested energy. This effort led to the final design of Sozu tags that transforms signals which are innately challenging to be sensed remotely, into radio frequency waves that can travel long distances and through walls. Results from a multi-week evaluation across three multistory buildings and 30 activities suggest our approach offers very high detection accuracy, with few false positives. We also discuss how Sozu can be used to sense rich signals, such as direction, user grasp and rate of events. Finally, we put together a toolkit and had eight students participate in a week-long study. All participants were able to get Sozu running on their personal computers and build simple, yet illustrative example applications of their choosing, underscoring the ease-of-use and flexibility of Sozu.

Chapter 6

Vibrosight: Long-Range Vibrometry for Smart Environment Sensing

6.1 Introduction

In this chapter, we leverage vibrations, which have been shown to be a rich signal source for detecting a wide array of events (see e.g., [90]). Indeed, almost all physical activities generate vibration as a byproduct, whether it be chopping vegetables, writing on a whiteboard, typing on a laptop, running on a treadmill, or even sitting and reading a book. Likewise, many devices and appliances produce characteristic vibrations (e.g., faucets, kitchen appliances, HVAC, power tools, and even electronics if they contain fans). In response, there has been significant prior work leveraging vibrations for activity detection, including sensors coupled to floors and walls [32, 118, 133], as well as plumbing, gas and HVAC infrastructure [42, 49, 59, 60].

To utilize on this unique signal for smart environment sensing, we propose Vibrosight – a low-cost, vibration sensing approach that works at long distances. Specifically, Vibrosight transforms surface vibrations into laser intensity changes, which can be easily detected remotely. To use Vibrosight, users affix small, inexpensive, passive stickers (Figure 3.3A) to objects they wish to reveal to our system – surfaces and objects without tags are invisible to our system. By using a steerable mirror, we can direct our sensing to any point with line of sight. We use this ability to intensively scan a scene for tags (Figure 3.3B), and then once found, rapidly cycle between tags to sense the vibrations of their host surfaces. We use this data to produce vibrational spectrograms for each object (Figure 3.3C), which we feed to a machine learning pipeline for recognition. In our evaluation, we investigated sensing accuracy across 24 objects in four locations. Our system can detect activation at 98.4% accuracy, with a false positive rate of 0.7%. To underscore Vibrosight's robustness to interference, most of our study data were collected with multiple active appliances. Overall, we believe this work illuminates a new sensing approach with unique strengths.

6.2 Inspiration

We were initially inspired by light-based eavesdropping devices invented by Léon Theremin in the 1950s [64]. These setups bounced light off of distant, large, reflective surfaces – most often

windows – and measured the intensity of the reflected light. Vibrations, induced by e.g., voice, cause surfaces to oscillate, which in turn, slightly alters the light path. This manifests as an amplitude modulated signal that can be captured by a photosensor placed in the reflected light. Today, these devices use lasers as their light source and are called "laser microphones".

Unfortunately, traditional setups suffer from two significant drawbacks that preclude their immediate use for activity sensing. First, the method is compatible with very few materials and surfaces. Windows are traditionally used because they are reflective, large and somewhat elastic, and thus function in a similar manner to a microphone diaphragm; air-borne sounds, such as speech, cause the glass to oscillate. However, by using a window as a proxy to capture sounds, individual vibrational signals cannot be separated. A second, even greater drawback is that laser microphones are at the mercy of an environment's geometry. Unless the setup can be perfectly perpendicular to e.g., a window, the emitter and receiver must be placed apart in accordance with the reflective angle of incidence. This means sensing multiple surfaces almost certainly means having to deploy multiple sensing setups.

We found that these limitations could be removed by using inexpensive, retroreflective stickers (Figure 3.3A). These reflect light back to its source, allowing us to combine a laser and photosensor into single, compact sensing unit (Figure 3.3 and 6.7). Unlike co-opting a window as a diaphragm, our tags allow us to exclusively read the vibrations of a host surface. This spatial precision makes our technique robust to noisy environments and vibrational interference, and means we can support recognition of multiple simultaneous activities, which most prior systems do not support.

6.3 Related Laser Sensing Techniques

Lasers have long been used in sensing systems, including laser tachometers measuring rotary speed and laser velocimeters that measure the speed of surfaces and particles. It is also possible to measure distances with lasers (i.e., LIDAR), either by measuring parallax or phase shift of returned laser light. Some depth cameras, such as the original Microsoft Kinect and Apple iPhone X TrueDepth camera, include a structured-light laser emitter.

In the research domain, Iyer et al. [76] use retroreflective tags and a scanning laser to find phones in a scene and wirelessly charge them (and could be easily extended to work with our approach). Researchers have also used line lasers to track touch input, estimate finger angle on a touchscreen [144], and even construct a 3D model of a hand [82]. Reflected laser light produces a unique speckle pattern (one example shown in Figure 3), which researchers have also leveraged. For example, SpecTrans [131] uses the signal to identify objects, SpeckleSense [174] and SpeckleEye [115] track the motion of distant objects, and Smith et al. [139] demonstrate hand gesture recognition.

Most relevant to our work are non-contact laser vibrometry techniques, which fall into two main categories. The first is laser Doppler vibrometry, which uses the phase difference between reflected, Doppler-shifted light (due to movement of a surface) and an internal reference [77, 110]. There are also intensity-based methods that use high-powered zoom lenses to interrogate laser light falling onto a surface, and sense changes in the speckle pattern resulting from surface vibration [104, 151, 162]. Vibrosight is also intensity-based, but requires no complex optics due to our use of retroreflective tags. Commercial laser vibrometers – used for, e.g., mechanical analysis and

quality control – exclusively rely on laser Doppler vibrometry [14]. Expensive optics mean that even entry-level systems cost tens of thousands of dollars (vs. ~$80 for our Vibrosight prototype). These commercial systems sometimes ship with retroreflective stickers to improve SNR.

6.4 Background Experiments

Although laser vibrometry has been well studied, no previous work has attempted to use it for activity detection in everyday settings. Therefore, prior to the development of our system, we first ran a series of background experiments to verify the principles of operation and investigate parameters that affect sensing quality.

6.4.1 Experiment Apparatus

To generate vibrations of known frequency in our tests, we used a wide-band transducer (driven by an audio amplifier) connected to a signal generator. To measure the photocurrent, we used a PIN photodiode amplified by a standard transimpedance amplification circuit with a 16-2800 Hz bandpass filter. The photodiode was placed next to the laser diode, facing in the same direction. The amplified signal is sampled by an ADC at 5 kHz, which we render into a spectrogram. Of note, all spectrograms in this chapter are rendered using a logarithmic color scale.

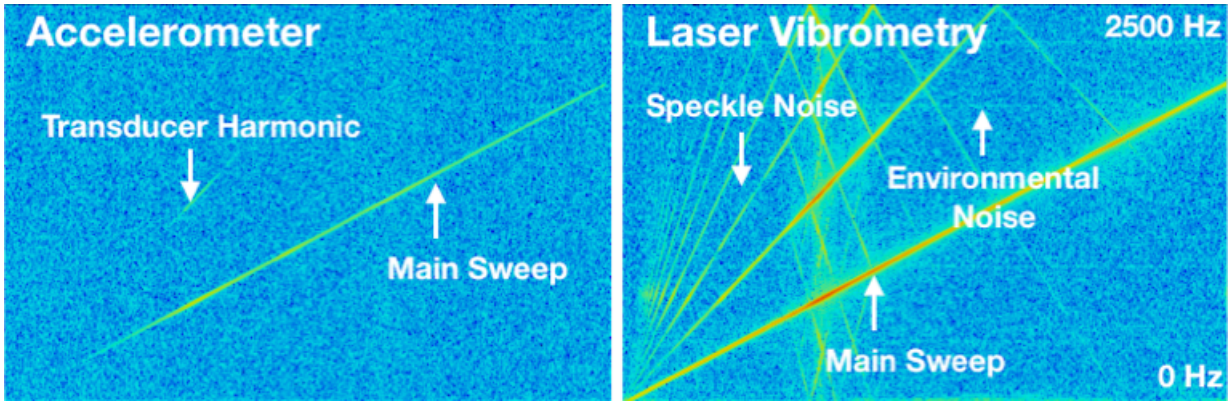


Figure 6.1: Vibration spectrograms of a swept-frequency signal (0 to 2 kHz) over 20 seconds, measured by an accelerometer (left) and our laser vibrometry sensor (right).

We configured the transducer to output a linear 0 to 2 kHz swept-frequency signal with a peak amplitude of 0.12 G. For reference, a generic microwave we tested had peak vibrational amplitude of 0.45 G. Figure 6.1, left, shows the vibration signal collected by an accelerometer (STMicroelectronics LIS3DH) affixed to the transducer. Figure 6.1, right, is data captured using a 3 mW red laser from 5 meters away using a mirror affixed to the transducer. Besides the main sweep, we also see higher-order harmonics (and their aliases) due to the mechanical properties of the transducer and the nature of the reflected laser speckle pattern, specifically the discrete speckles entering and leaving the photosensitive area results in speckle noise (Figure 6.2). The spectrograms also reveal unwanted environmental noise caused by artificial lighting and electrical noise (Figure 6.6, right).

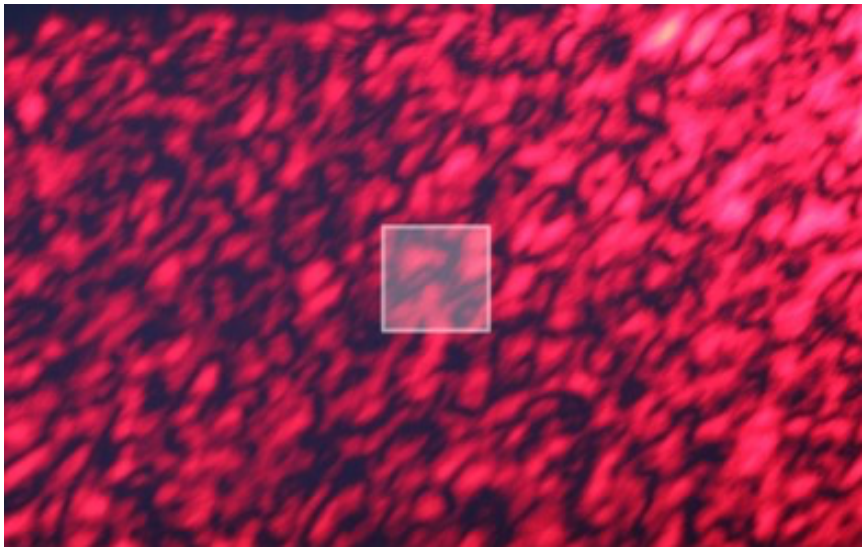


Figure 6.2: Example retroreflected speckle pattern as captured by a 22×14 mm CMOS sensor at 2 m. A small part of this pattern would land onto the 2.7×2.7 mm photosensor used in Vibrosight (illustrated at scale with white square).

6.4.2 Experiment 1 – Laser Wavelength

Laser vibrometry requires a laser beam to illuminate a surface of interest, which naturally leads to the first parameter to investigate: the wavelength of laser light. We tested three common wavelengths: 532 nm (green), 635 nm (red), and 940 nm (infrared) to see if performance varied. As before, we used a 0-2 kHz swept-frequency signal with a mirror attached to our transducer. We placed our laser-photodiode unit 1 m away; see Figure 6.3 for spectrograms.

The three laser wavelengths had nearly identical performance, though we did notice that the driver circuit of our green laser seemed to be more prone to harmonics of powerline noise. In real-world use, an infrared laser might be preferred, as it is invisible to humans (and eye-safe at the wattage we use). However, for ease of debugging, we selected the red laser for all future experiments.

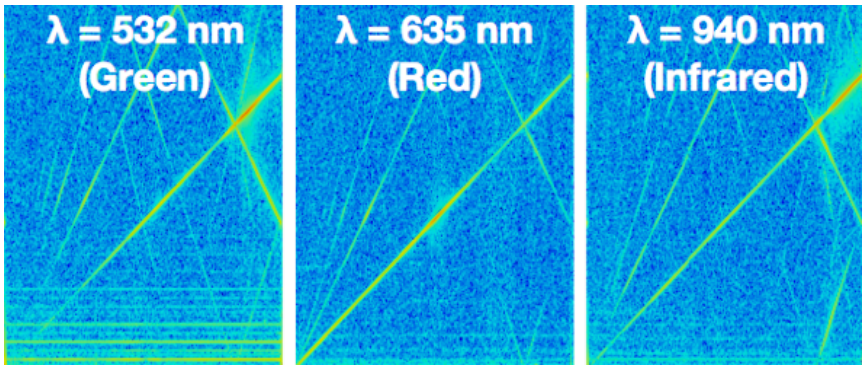


Figure 6.3: Vibration spectrograms of a swept-frequency signal (0-2 kHz) using lasers of different wavelengths.

6.4.3 Experiment 2 – Reflective Materials

Our next test was to investigate the performance of different reflective materials. In addition to the previously used mirror, we included (soda-lime) glass and three common retroreflective materials: retroreflective tape, retroreflective spray and a (bicycle) corner reflector matrix. We used the same setup and procedure as the previous test. Of note, glass uniquely permits the passage of some laser energy, reflecting approximately 30% (when normal to emitter), which indicates that sensing through glasses could be achieved if desired.

The resulting spectrograms can be seen in Figure 6.4. The retroreflective spray returned little signal, while the corner reflector generated noise due to physical resonance on the transducer. The mirror, glass, and retroreflective tape all performed well. As noted previously, retroreflective materials have the added benefit of enabling Vibrosight to sense objects at a range of angles, and so we adopted the retroreflective tape as our tag material of choice.

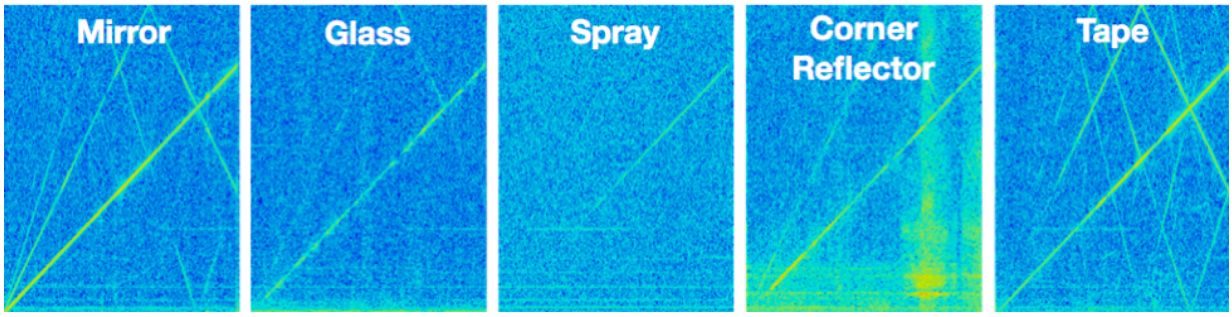


Figure 6.4: Vibration spectrograms of a swept-frequency signal (0-2 kHz) using different reflective materials.

6.4.4 Experiment 3 – Tag Distance and Angle

Users can tag objects in the environment at any angle and distance, and so we wished to quantify the performance of our selected retroreflective tape with these factors. For this test, we configured the transducer to output a constant 1 kHz vibration and measured the amplitude of the detected signal (at 1 kHz) with the noise floor subtracted. We tested distances from 1 to 8 meters (1 m interval) and angles from 0° to 60° (15° increments). As seen in Figure 6.5, the returned signal decreases as distance and angle increases. However, even at 8 meters and 60°, the returned signal is still detectable with our setup (0.04 Vpp). This range of sensing should be sufficient to cover typical indoor spaces. Longer distances and more acute angles could be enabled with more powerful lasers and superior analog frontends.

6.4.5 Experiment 4 – Artificial and Natural Light

In the previous tests, we found that the most significant source of noise comes from lighting in the environment. In this experiment, we investigated how light – both natural and artificial – affects our approach. We used the same setup as above, at 1 m, using a 0-2 kHz swept-frequency signal. We varied the lighting conditions: dark, daylight, incandescent and fluorescent light.

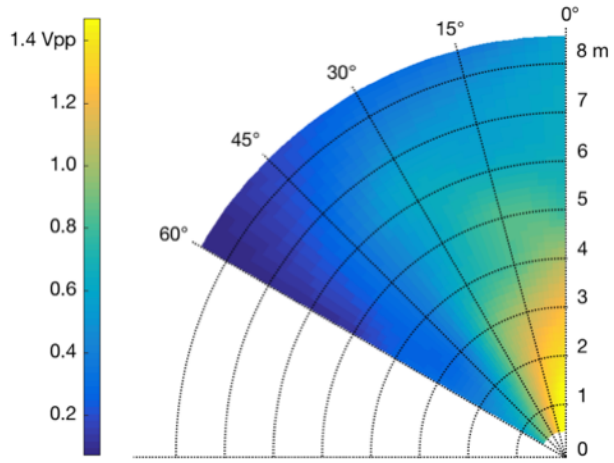


Figure 6.5: Signal amplitude at different retroreflective tag distances (1 - 8m with 1m interval) and angles (0, 15, 30, 45, 60°).

We found very little difference between dark and daylight conditions, suggesting that natural light level has little effect (Figure 6.6). This is because our circuit's bandpass filter removes the DC bias of a constant light source. However, artificial light oscillates, which is visible in the captured signal. Additionally, fluorescent lights tend to emit wide-band electrical noise.

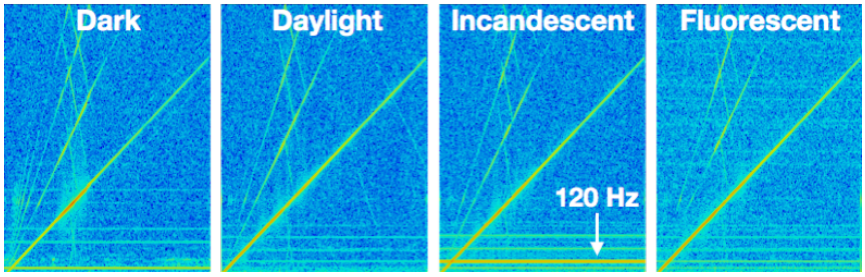


Figure 6.6: Vibration spectrograms of a swept-frequency signal (0-2 kHz) under different lighting conditions.

6.4.6 Experiment 5 – Speech

During previous tests, we serendipitously noticed our setup was insensitive to voice. As a test, we affixed a tag to a large window, roughly 1×2 m in size, which would be an excellent surface for a conventional laser microphone to capture speech. However, we found that our setup could not detect speech at any volume.

As a more formal test, we recruited four participants (two female) to read aloud a 51-word paragraph, 1 meter away from the window. Using the same setup as the previous experiments, we recorded vibrations while participants spoke. We then played the recordings back to participants, who all reported that they could not hear their voice. Our own analysis suggested there is no discernable signal. As a further test, we affixed an accelerometer to the window, similar to our

earlier transducer tests, and recorded vibration during speech. As with our laser vibrometer, we saw no signal, suggesting the magnitude of vibrations induced by speech is very small (under the 0.001 G noise floor of the accelerometer), and many orders of magnitude lower than vibrations from e.g., motor-driven appliances. That said, it might be possible to recover voice with an analog frontend optimized for human speech frequencies. However, by limiting the signal gain in hardware, we obtain a beneficial privacy-preserving effect.

6.5 Implementation

To make our sensor compact and easy-to-deploy, we built a custom sensor board and pan-tilt mirror platform (Figure 6.7). In total, our prototype hardware cost roughly $80.

6.5.1 Sensor

Our sensor (Figure 6.7) consists of two main elements. First is a 3 mW, class III A, red laser diode (Quarton VLM-635), which is rated as safe for brief, direct exposure to the naked eye. In our scanning mirror design, the duty cycle at any one position is low, and as we discuss later, we can detect when the beam is interrupted and turn off the laser as an extra safety measure (similar to [76]). The second component is a 2.7×2.7 mm square PIN photodiode fitted with a 635±18 nm bandpass optical filter. To further reduce the effect of ambient light and improve directionality, we house our photodiode in a 2 cm long, black plastic tube.

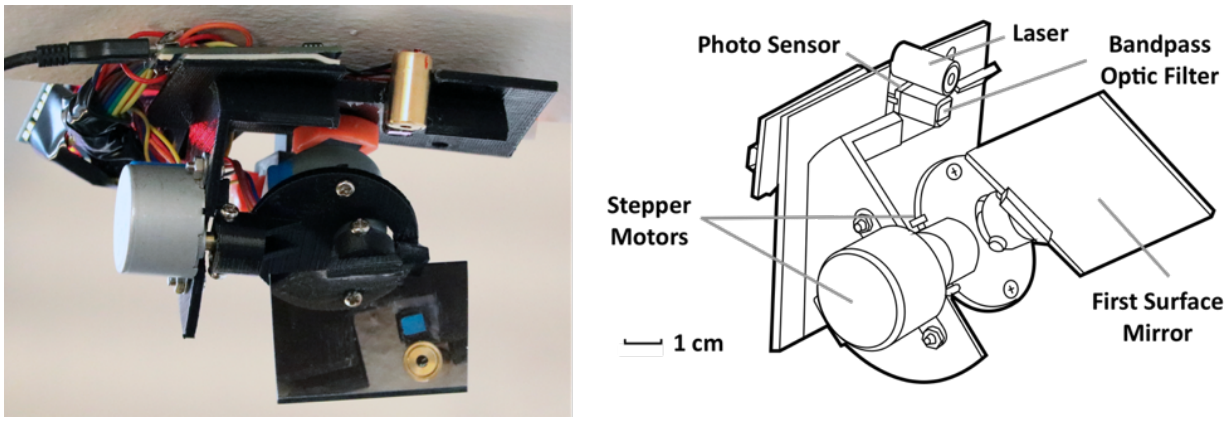


Figure 6.7: Left: Vibrosight deployed on a ceiling. Right: Illustration of key components.

6.5.2 Pan-Tilt Mirror Platform

To sense multiple objects in an environment, we designed a mirror-control platform, built from two BYJ-48 stepper motors and a 5×5 cm first surface mirror (Figure 6.7). This platform has a 120° (pan) by 60° (tilt) field of view with an angular resolution of 0.18° (i.e., full step mode). We run our stepper motors at 30°/sec. To remove gear backlash, our driver software seeks to a point 6° above and to the left of a point, before translating to the desired target.

6.5.3 Driver/Sensor Board

Our driver/sensor board (Figure 6.8) is built around an MK20DX256VLH7 microcontroller powered by Teensy 3.2 firmware. Eight digital pins are used to control the two stepper motors. The board also features a 16-2800 Hz bandpass analog frontend with a programmable gain. The amplified signal is sampled by the built-in ADC on the microcontroller at 5 kHz. A second analog frontend (no bandpass) measures the absolute reflected light intensity. Data is sent to a laptop over USB for visualization and further computation.

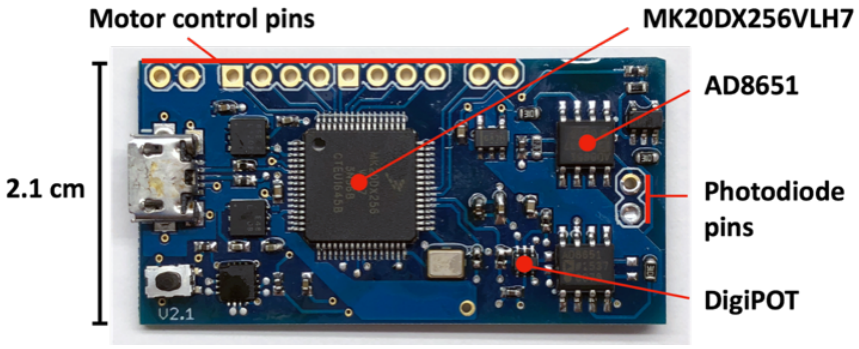


Figure 6.8: Vibrosight driver/sensor board.

6.5.4 Tags

We used 3M 79961 Scotchlite Reflective Striping Tape for our tags. To ensure Vibrosight can find tags, even at long distances, they cannot be smaller than the step size of our mirror-control platform. The formula to calculate the minimum size of a tag given stepper motor angular resolution and sensor distance is:

$$Tag\ Size \geq 2 \times \sin(Angular\ Resolution/2) \times Distance \tag{6.1}$$

For example, when sensing at 4 m with our prototype's 0.18° angular step size, the minimum tag size would be 1.3×1.3 cm. At longer sensing distances, larger tags must be used. We selected 5×5 cm as our standard tag size, which was sufficient for all of our tests and locations. At small-volume retail prices, these tags cost 30 cents each.

6.5.5 Tag Search

To find tags, Vibrosight raster scans its field of view from top to bottom while measuring the absolute reflected light intensity at each step. If the laser hits a tag, the intensity of reflected light significantly increases, manifesting as a peak in our scanned signal. When the raster scan is complete, we spatially cluster all peaks to extract the theta and phi of probable tags. On our prototype hardware, this scanning process takes 20 minutes and is chiefly limited by the speed of our stepper motors. We envision this process being triggered by users when they tag a new object they wish to register with the systems. However, tag search could also occur opportunistically throughout a day, such that new objects are found within a few hours automatically.

6.5.6 Tag Labeling

Our search process finds tags in a scene, but cannot identify what objects they are attached to. One option is for users to provide labels, for example, our system could flash a visible laser onto a tag and ask "what is this?" (e.g., by voice interface or smartphone app). Alternatively, our system could automatically identify objects based on their vibrational signal (e.g., pre-loaded library) when they first run, which we show is possible in our subsequent evaluation.

6.5.7 Signal Processing & Featurization

After our system finds all tags in its field of view, it cycles between them in a round robin fashion. We used a traveling salesman solver to find the (approximate) shortest path between known tags. While seeking, the laser is turned off. Once a target is reached, our system waits half a second for mechanical oscillations to dampen. The laser then turns on and one second of data is recorded. We compute a spectrogram using a 50% overlapping, 512 window-sized FFT, which we use to compute the max, mean, and standard deviation for each frequency band (256×3 values), resulting in 768 features for machine learning. Detection

For each object (i.e., tag), we train a two-class classifier: active and inactive. We used a Random Forest implementation (default parameters) provided by Weka. When Vibrosight seeks to an object, the corresponding classifier is fed live data to detect activity. As each object classifier is independent, we innately support the detection of simultaneous events, and can easily add new object classifiers.

6.5.8 Open Source

We make all source files (PCB design, 3D models, embedded firmware, and software) available online for replication: `https://github.com/FIGLAB/Vibrosight`. We also include all data collected in our evaluation, described next.

6.6 Evaluation

To investigate the accuracy of Vibrosight, we ran a multi-stage evaluation in four indoor contexts: kitchen, gym, office and workshop. For portability, we deployed Vibrosight on a tall tripod, which we placed in the corner of each room. Within each context, we selected six common objects to test (Figure 6.9), which we tagged with line of sight to our sensor. Active signals from all tested objects can be seen in Figure 6.10; Table 6.1 provides the distances and angles of the tags.

Before collecting vibration data, we first ran three rounds of tag search as a test. We recorded the number of missed tags and any false positive tags found, manually correcting the data before proceeding. In each location, Vibrosight cycled between tags, collecting three seconds of data from each object for four conditions:

No Objects Active – All objects were turned off, with each object tag being sensed ten times in total (for each tag: 10 cycles $\times$ 3 seconds of data).

Figure 6.9: 4 test locations and 24 objects (outlined in yellow). From left to right: Kitchen, Gym, Office and Workshop.

One Object Active – One object at a time was turned on, followed by ten rounds of data collection for all objects (for each tag: 10 cycles × 6 active objects × 3 seconds of data).

Two Objects Active – Two objects at a time were turned on, followed by ten rounds of data collection for each pairwise combination (for each tag: 10 cycles × 15 active object pairings × 3 seconds of data).

All Objects Active – All six objects in the room were turned on simultaneously, with each object being sensed ten times in total (for each tag: 10 cycles × 3 seconds of data) This procedure yielded 5,040 active instances and 11,520 inactive instances from our 24 objects.

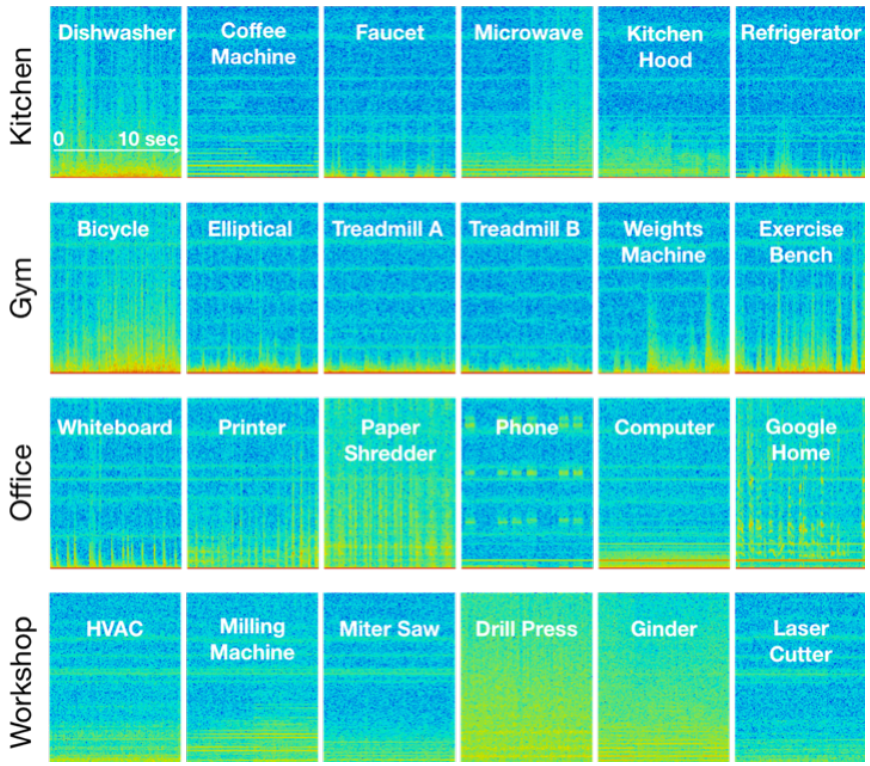


Figure 6.10: Vibration spectrograms (0 to 2.5kHz) of our 24 test objects when active.

6.7 Results

6.7.1 Tag Search Accuracy

Across the 12 rounds of tag search we performed (4 locations × 3 rounds, with each location containing 6 tagged objects), Vibrosight failed on a single tag twice, resulting in a tag search accuracy of 97.2%. We note that no tags were missed when results from all three passes were combined. There were no false positive tags.

6.7.2 Usage Detection Accuracy

To investigate Vibrosight's ability to detect if an object is active, we ran a leave-one-round-out cross-validation. Specifically, we trained detection classifiers on nine rounds (i.e., seek cycles) of collected data, and tested on the remaining round (all combinations, results averaged). This procedure prevents data with time adjacency from being included in both the train and test dataset.

Across all objects, locations and conditions, Vibrosight achieved an average detection accuracy of 98.4% (SD=1.6). False positive activations are low: 0.7% (SD=0.5). We break out True Positive and False Positive rates across objects in Table 6.2. We found no correlation in accuracy between tag distance and angle.

Location	Object	Distance to Sensor	Tag Angle	True Positive Rate	False Positive Rate
Kitchen	Dishwasher	1.18 m	34 °	98.1 %	0.8 %
	Coffee Machine	1.21 m	29 °	88.1 %	0.4 %
	Faucet	1.84 m	40 °	96.2 %	4.4 %
	Microwave	2.14 m	45 °	98.6 %	1.5 %
	Kitchen Hood	1.92 m	10 °	98.6 %	0.9 %
	Refrigerator	1.56 m	33 °	93.8 %	1.1 %
Gym	Bicycle	1.61 m	12 °	97.6 %	0.2 %
	Elliptical	2.50 m	8 °	100.0 %	0.2 %
	Treadmill A	3.48 m	13 °	99.0 %	0.0 %
	Treadmill B	4.78 m	35 °	99.0 %	0.0 %
	Weights Machine	5.89 m	42 °	99.0 %	0.0 %
	Exercise Bench	4.23 m	28 °	98.6 %	0.4 %
Office	Whiteboard	2.28 m	28 °	97.1 %	0.2 %
	Printer	2.68 m	38 °	97.6 %	0.8 %
	Paper Shredder	2.79 m	10 °	98.1 %	0.2 %
	Phone	2.18 m	18 °	82.4 %	0.2 %
	Computer	1.26 m	39 °	93.3 %	1.5 %
	Google Home	1.28 m	42 °	93.8 %	1.0 %
Workshop	HVAC	7.68 m	8 °	99.5 %	0.0 %
	Milling Machine	4.09 m	14 °	97.6 %	0.4 %
	Miter Saw	8.73 m	28 °	97.1 %	0.8 %
	Drill Press	8.25 m	28 °	98.6 %	0.0 %
	Grinder	5.25 m	32 °	100.0 %	0.0 %
	Laser Cutter	3.44 m	20 °	87.1 %	1.3 %

Table 6.1: Deployment details and accuracies for the 24 objects.

6.7.3 Vibrational Interference

Vibrational interference can occur when objects vibrate so strongly that signals propagate through a structure and shake other objects. This can cause false positives, as inactive objects can begin to vibrate. Another failure mode is when an active appliance's vibration signal changes due to the superimposition of an external vibration, resulting in a new (i.e., unlearned) signal that can decrease true positive accuracy.

To understand the effects of vibrational interference on our system, we break out results into the four data collection conditions in Table 6.2. There was a small, but significant decrease in accuracy (2.3%) from one object active (no interference possible) to two objects active (potential for interference) conditions (paired t-test; $p<.05$). The confusion matrices suggest that objects that sat on a common surface (e.g., table) experienced some interference. For example, in our kitchen location, the microwave and coffee machine sat on a common granite countertop, and interference between these two devices was responsible for 33.4% of all detection errors in the kitchen.

	True Positive Rate	False Positive Rate
No object active	NA	1.3%
One object active	97.9%	0.5%
Two objects active	95.6%	0.7%
All objects active	97.5%	NA

Table 6.2: Accuracies across data collection conditions.

6.7.4 Sampling Duration

Shorter sampling durations per object are beneficial, as it allows Vibrosight to cycle between tags in a scene more rapidly, allowing for detection of shorter activations and with less latency. However, less data could also lead to a decrease in detection accuracy.

To investigate this effect in our study, our procedure recorded three seconds of data from each object per seek. In a post hoc experiment, we followed the same validation procedure as above, but varied our train and test data in duration from 0.1 to 3.0 seconds. We used the same 768 machine learning features, which are duration independent. As expected, accuracy drops as sampling duration decreases – Figure 6.11 plots this result. However, the drop in accuracy is not consistent across devices. Those with intermittent vibrations, such as the phone ringing (which contains pauses of ~1 second) are most impacted, while devices with consistent output (e.g., coffee grinder) are unaffected down to around 0.2 seconds. One second sampling duration appears to be a good tradeoff between speed and accuracy.

6.7.5 Robustness Across Time

To test Vibrosight's performance across time, we ran a second session of data collection separated by one week. We placed our setup in a similar (but not exactly the same) location. The data

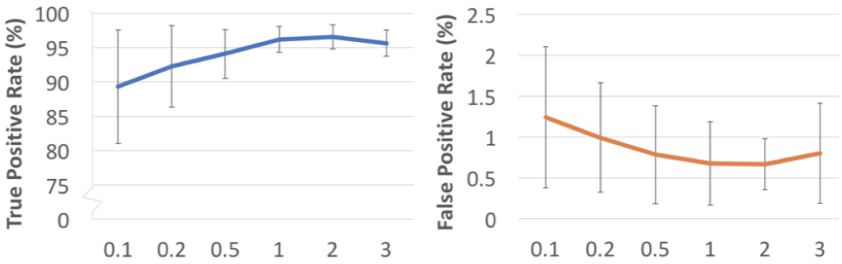


Figure 6.11: Average accuracy vs. sampling duration (seconds).

collection procedure was the same as before, except that we collected five rounds (i.e., seek cycles) of data instead of ten (resulting in 2,520 active and 5,760 inactive instances). We ran this newly collected data through our pre-trained classifiers from week one. Average accuracy was 96.2% (SD=2.1), with a false positive rate of 2.3% (SD=2.3), which was not significantly different than week one accuracy, suggesting robustness across time.

6.7.6 Object Identification

So far, we have only discussed Vibrosight's ability to detect if objects are active or inactive. However, we found that vibrations of different appliances are quite unique (see Figure 6.10) and might be used to infer object type without user labeling. For this test, we evaluated the accuracy of a 24-class object classifier (SMO, Poly Kernel with E=1.0) using a leave-one-round-out cross-validation (week one data, all conditions, active instances only). Across our 24 objects, average identification accuracy was 92.1% (SD=5.9); the confusion matrix is offered in Figure 6.12. We can see that most of the confusion occurs within a location. We suspect this is due to the classifier overfitting to environmental noise (both lighting and EMI) and also vibrational interference between objects, where one object's vibration is being picked up by another. Additionally, in the case of the gym, we included two objects of the same type (treadmill). If we discount confusion between the two treadmills as an error, average object identification accuracy is 92.8% (SD=5.5).

6.8 Supplemental Studies & Applications

We now describe a series of focused, supplemental studies, as well as additional applications enabled by Vibrosight.

6.8.1 Sensing Modes of Operation

We found that many appliances vibrate uniquely when in different modes of operation, especially motor-driven appliances. Figure 6.13 offers three examples: a dishwasher at different cleaning phases, a treadmill when a user is walking/jogging/running, and a milling machine at different RPMs. Vibrosight can easily leverage these characteristic signals to enable richer, object-specific classifiers.

Actual \ Predicted	Dishwasher	Coffee Machine	Faucet	Microwave	Kitchen Hood	Refrigerator	Bicycle	Elliptical	Treadmill A	Treadmill B	Weights Machine	Exercise Bench	Whiteboard	Paper Shredder	Printer	Phone	Computer	Google Home	HVAC	Milling Machine	Miter Saw	Drill Press	Grinder	Laser Cutter
Dishwasher	91.9	0.0	0.5	0.0	5.2	0.0	0.0	0.0	0.0	0.0	0.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Coffee Machine	2.9	89.0	2.4	1.4	1.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
Faucet	1.9	1.4	83.3	0.5	1.0	8.6	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.5	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
Microwave	0.0	7.1	0.5	87.1	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
Kitchen Hood	2.4	0.0	0.5	4.8	91.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0	0.0
Refrigerator	1.0	1.0	7.6	0.0	0.0	90.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bicycle	0.5	0.0	0.0	0.0	0.0	0.0	88.6	2.4	1.0	1.0	1.9	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Elliptical	0.0	0.0	0.0	0.0	0.0	0.0	1.4	88.1	5.2	1.9	1.4	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Treadmill A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.6	87.1	1.4	2.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Treadmill B	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.3	7.6	87.6	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Weights Machine	0.5	0.0	0.0	0.0	0.0	0.0	6.2	0.5	1.0	4.8	81.9	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Exercise Bench	0.5	0.0	0.0	0.0	0.0	0.0	1.9	3.8	1.0	1.4	6.2	84.3	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0
Whiteboard	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	95.7	0.5	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Printer	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	97.6	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Paper Shredder	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	2.9	91.9	1.9	0.0	0.5	0.0	0.0	0.0	1.0	0.0	0.0
Phone	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.5	97.6	0.5	1.0	0.0	0.0	0.0	0.0	0.0	0.0
Computer	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	99.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Google Home	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0
HVAC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.1	0.0	1.0	0.0	1.0	0.0
Milling Machine	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.6	0.0	0.0	0.0	1.4
Miter Saw	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.0	0.0	0.0	1.0
Drill Press	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.0	1.0	0.0
Grinder	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	99.0	0.0
Laser Cutter	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	1.4	5.7	0.0	0.0	91.4

Figure 6.12: Confusion matrix of object identification (%).

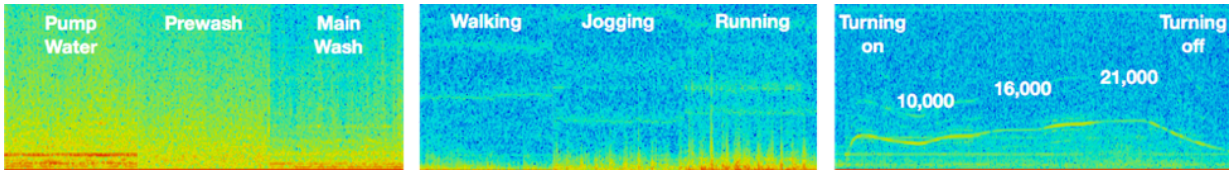


Figure 6.13: Vibration spectrograms (0 - 500 Hz) of a dishwasher (left), treadmill (middle), and milling machine (right) in different modes of operation.

6.8.2 Detecting Objects and Activities on Work Surfaces

One potential limitation of Vibrosight is that it is ill-suited for tracking small appliances that move around, such as blenders and staplers. We found, however, that we can tag host surfaces to capture vibrations. For example, Figure 6.14 shows data and detection results from a tagged work surface, on top of which a user hammers, hand files and sweeps. This property allows Vibrosight to detect a much wider range of activities beyond large fixed appliances.

To test this ability more formally, we revisited our four test locations and ran a supplemental study. We affixed a tag to one exemplary surface in each location and developed a new set of test classes: kitchen countertop (classes: chopping, coffee grinder running, kettle boiling, blender running), office table (classes: typing, writing), workshop table (classes: hammering, hand filing, sawing), and gym mat (classes: jumping jacks, push-ups). Similar to our main study, we collected 3 seconds of data for each class (including no activity) × 10 rounds. We then ran a leave-one-round-out cross-validation study (per location) using the same featurization and machine learning pipeline as before (i.e., SMO, Poly Kernel, E=1.0). Vibrosight was able to detect activities using surface vibration data with a mean accuracy of 89.7% (SD=2.3).

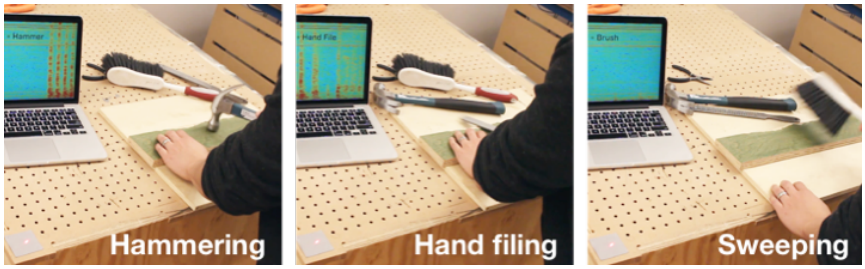


Figure 6.14: Vibrosight detects activities on a work surface. Signal and classification shown on laptop for illustration.

6.8.3 Human Movement

Human movement inevitably induces vibrations in the environment. This is especially true for objects such as seating and beds. Figure 6.15 shows the signal from a sofa as a user sits down, which can be used to detect use and occupancy. Similarly, Vibrosight could monitor sleep restfulness if a tag is placed on a headboard or bedframe.

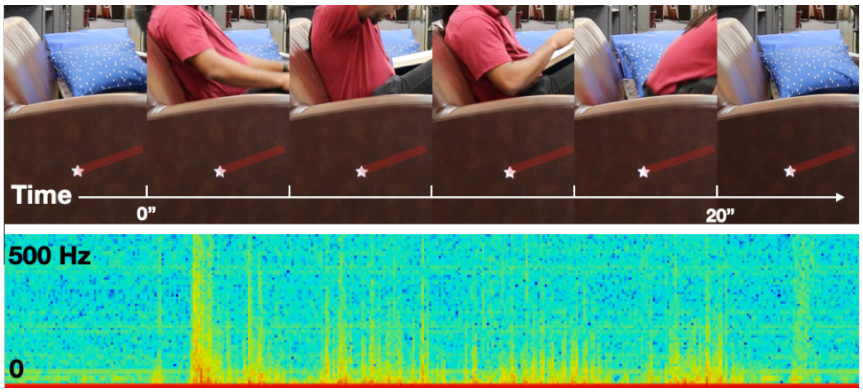


Figure 6.15: Vibrosight can detect when this couch is in use.

6.8.4 User Input

Vibrations resulting from knocking or tapping on a tagged surface are very apparent to Vibrosight. Thus, it is possible to support basic user input. Figure 6.16 offers one example where double knocking on a laundry machine instantiates a notification request (as Vibrosight also knows when the laundry machine cycle is complete).

Figure 6.16: A user double knocks on a laundry machine to request a completion notification.

6.8.5 Data Communication

For devices with speakers, motors or other actuators, it is possible to encode information in the form of vibrations, which can be detected and decoded by Vibrosight. This opens an interesting communication channel for richer applications. For example, devices could emit their brand and model number (allowing a specialized classifier to be loaded), or an IP address to expose additional functionality.

As a proof of concept, we attached a retroreflective tag to the bezel of an LCD TV and used on-off-keying with a 1 kHz carrier emitted through the build-in speakers to broadcast a UID. Figure 6.17 offers an example transmission, which includes a 6-bit header, 4-bit payload length, 12-bit payload, and 5-bit tail.

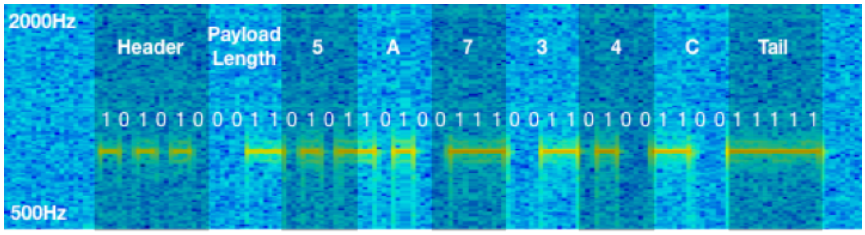


Figure 6.17: Vibrosight can detect encoded vibrations from a TV.

6.8.6 Leveraging Occlusion

Occlusion of tags can also be used to infer activity information. For example, it is possible to strategically attach a tag so that it is occluded when an object is in use, for example, an eyewash station (Figure 6.18). Occlusion is easily detected by measuring the reflected light intensity.

Figure 6.18: Tag occlusion is easily detected by measuring the reflected light intensity, and can be used to infer use.

6.9 Limitations

Laser safety was a primary concern from the beginning of the project. To reduce the potential of harm, we limited ourselves to lasers rated Class III A or below. It may be possible to use even lower-powered lasers with superior band-pass optics, photodiodes and analog frontend. We also endeavored to reduce the duty cycle as much as possible, not only by scanning, but also with occlusion detection. Specifically, once we seek to a tag and receive no reflected light, we infer the tag is occluded, and immediately turn off the laser (within 1 ms), limiting accidental exposure.

Similar to other optical approaches, our system requires line of sight and is affected by occlusion. To partially mitigate this, we most often deployed Vibrosight near the ceiling, which affords a good view of rooms. This is also a less conspicuous location and where sensors are typically deployed in commercial settings. If we detect that a tag is occluded, we can simply suspend classification, as opposed to producing an errorful result. Another limitation is that Vibrosight cannot observe multiple objects simultaneously, as our current implementation reads data from tags in a round-robin manner (pausing at each for e.g., one second). This limits our current system to events with sufficient duration. Fortunately, many appliances and activities have durations on the order of tens of seconds, minutes or even hours. There are also immediate ways to improve our prototype with e.g., faster stepper motors and multiple lasers. Nonetheless, there are transient events (e.g., stapling papers) that can occur and dissipate within ~100 ms, which would require a non-time-multiplexed approach.

6.10 Summary

In this chapter, we have presented Vibrosight, a sensing approach for smart environments that uses laser vibrometry, which transforms surface vibrations into laser intensity signals to facilitate wide-area sensing. Users tag objects they wish to reveal to our system, which can then sense vibrations on these objects at long distances. Through a series of experiments and evaluations, we demonstrate that our approach robustly detects activities at high accuracies. We also explored a range of supplemental uses for our system. We believe Vibrosight is a new and practical way to enable a wide range of interactive and context-aware applications for smart environment. We hope to encourage future work by open sourcing all project materials.

Chapter 7

Smart Room Sensing with FMCW Radar

7.1 Introduction

In Vibrosight, we demonstrated that vibration signals could reveal much information about user activities in their environments. In this chapter, we present a new approach to sense vibrations based on frequency-modulated continuous wave (FMCW) radar. These small, solid state sensors, typically used in automobile and security applications, are generally configured to track large objects, such as people [29] and vehicles [4]. Although prior research has used these sensors to infer activity by tracking people, they have yet to be used for tracking room and object states (i.e., physical facets). In particular, their low-level signals hold the potential for fine-grained sensing of environments in a cost-efficient manner, with the sensor centrally located in e.g., a smart speaker. While some objects and appliances can be tracked without any instrumentation through vibration sensing, we found our system could recognize a much wider array of devices when instrumented with an RF corner reflector. These can be made from foil-backed paper, costing mere cents to produce. We also show how simple mechanisms can be made incorporating corner reflectors to further extend sensing capabilities.

7.2 Principle of Operation

FMCW radar is commonly used to track objects. In this section, we quickly review our system's principle of operation based on this technology.

7.2.1 Frequency-Modulated Continuous Wave (FMCW) Radar

The simplest FMCW setup consists of only one pair of transmitter and receiver antennas, with which the distance between target objects and the radar, as well as their radial velocities, can be measured. This data manifests as a range-doppler map (Figure 7.3). It is also possible to use multiple transmitter and receiver antennas with beamforming to estimate the angle of the target objects, which yields a range-azimuth map (Figure 7.3). Our signal processing and detection are based on these two fundamental data structures which we will discuss in greater detail.

7.2.2 Types of Signals

There are three fundamental types of signals in the range-doppler and the range-azimuth maps – position, amplitude, and doppler – each revealing a unique property of the target object.

The 2D positions of the target objects can be measured by tracking the peaks on the range-azimuth map. We found this signal useful for sensing activities that involve positional changes, such as use of drawers and sliding doors. Besides positions of these peaks, their amplitudes (i.e., heights) also contain useful information. In general, amplitudes indicate target objects' ability to reflect RF energy, which correlates with activities. For example, a fridge door changes its angle to the radar and thus reflects different amount of RF energy. This manifests as a peak on the range-azimuth map with its magnitude encoded with the angle information of the fridge door. Finally, there are doppler signals which reveal their target objects' radial velocities to the radar. Such signals can be referred to on range-doppler maps, which we leverage to sense activities that involve periodic movements (e.g., the movements of drill presses or ceiling fans) and vibrations (e.g., those that occur when a microwave is running).

7.2.3 Corner Reflectors

Our principle of operation also intersects with the corner reflector mechanism. Specifically, as we will later discuss in greater detail, we attached small ultra low-cost corner reflectors to appliances and objects to achieve better signal-to-noise ratios (SNR). Corner reflectors are often used to enhance radar cross-section (RCS) for their ability to reflect RF signals back to their sources (i.e., retroreflectivity). Therefore, corner reflectors have been widely used in RF sensing applications such as terrain measurements and remote vehicle tracking.

The retroreflectivity of corner reflectors was achieved through by having three mutually perpendicular metallic surfaces, each best reflecting RF waves in its normal direction. With the angle of incidences often less than 40°, the incoming RF waves are bounced by these surfaces sequentially and are redirected back to their incoming direction. Prior literature has investigated corner reflectors with various geometries and materials [54, 85].

7.3 Initial Explorations

To investigate how well a FMCW radar can capture signals from everyday appliances, we went to three locations where we collected signals from a wide range of applications with the AWR1642 FMCW radar from TI. These three locations include a kitchen, an office, and a workshop.

First, we found that many appliances yield distinguishable signals when they are in operation (Figure 7.1). For example, in the kitchen, the microwave and the dishwasher surfaces vibrate, causing radial displacements that can be sensed by the RF sensor. The spinning fan of the kitchen hood also manifests as strong doppler shifts on the RF sensor. Interestingly, the faucet also produces significant signals for its dripping water's radial velocity changes concerning the radar.

Another important finding is that the distance and angle between objects and the radar are significant factors on SNR. Specifically, due to the attenuation of RF energy over distance, objects that are far away from the sensor yield lower SNRs and are challenging to sense especially when

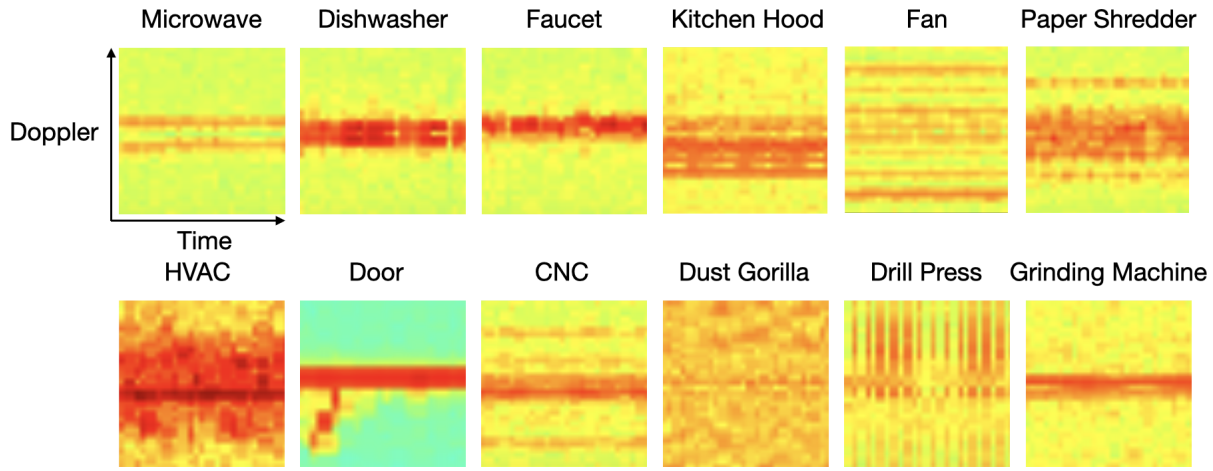


Figure 7.1: Example RF doppler signals from 12 appliances and objects.

the amplitude of their motion/vibration is small. Similarly, objects with larger angles of incidence to the radar also reflect relatively less RF energy back to its source, which results in lower SNRs. For example, we found that a running dishwasher yielded strong signals if the RF sensing beam is orthogonal to its vibrating door, but the SNR rapidly decreases as the angle of incidence increases, and is almost zero at 35°.

For these appliances and objects, we leverage corner reflectors, which have been commonly used to improve the radar reflectance of objects. To evaluate how much the corner reflector improves SNR, we ran several explorations and experiments with a microwave. Specifically, we investigated the effectiveness of a corner instrumentation with two other factors: distance and angle of incidence.

7.3.1 Sensing Distance with Corner Reflector

In this test, we placed a microwave with the radar pointing perpendicularly to its door. We recorded SNRs in decibels when we varied the distance between the microwave and the radar from 1 to 8 m with a 1-m interval. We then placed a commercial trihedral corner reflector with the edge length of 9.5 cm on the microwave top, and aligned its front with the microwave front. With this setup, we collected the measurements with the same procedure. Figure 7.2 left shows the result, which indicates little change in SNR with the instrumentation of corner reflector. This result is hardly surprising for the significantly larger RCS of the microwave door than that of the corner reflector, which makes the microwave a dominant signal source. This result suggests that corner reflectors are of little help to improve sensing distance when target objects already have surfaces at close-to-perpendicular angles to the radar.

7.3.2 Incidence Angle with Corner Reflector

In this test, we varied the angle of incidence of the radar on the microwave door from 0 to 60° with a step of 10°, at a fixed 5-m distance. Same as the previous test, we collected measurements

before and after we instrumented the microwave with the corner reflector. Figure 7.2 right shows the result, which indicates an average improvement of 1.16 dB (SD=0.48) up to a 40°angle of incidence. In practice, we found this improvement to be very important for sensing objects that can be at various angles to a radar device in user environments.

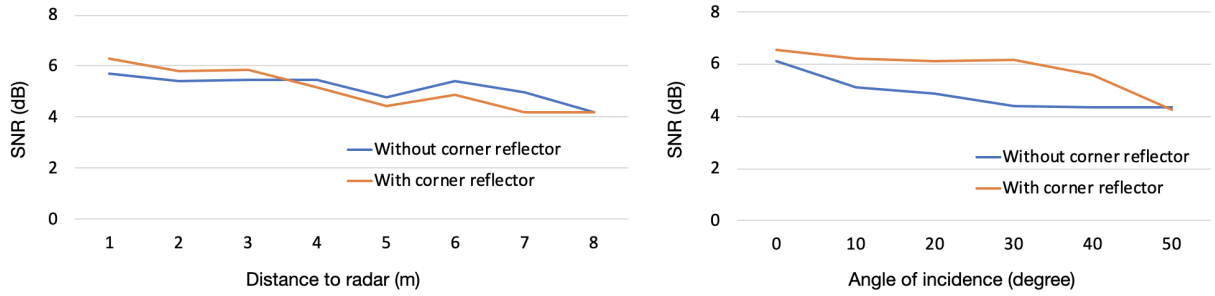
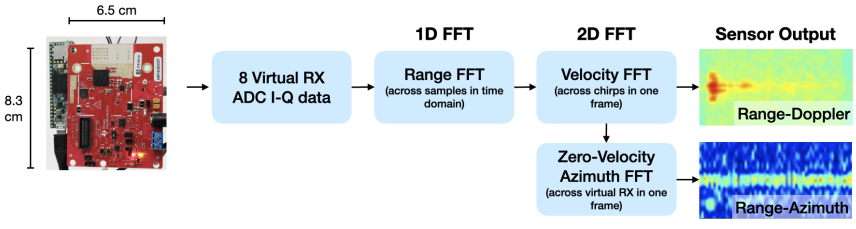


Figure 7.2: Measured SNRs of a microwave at different distances (left) and angles (right).

7.4 Implementation

Our system is based on TI AWR1642 FMCW radar which features two transmitter and four receiver antennas with an operating frequency band of 76 - 81 GHz. This configuration of multiple transmitter-receiver pairs allows us to acquire a range-azimuth map with a 160-degree azimuth field of view and an angular resolution of 15°. Figure 7.3 left shows the underlying processing pipeline that outputs the range-doppler and range-azimuth maps. Figure 7.3 right shows the configuration for the AWR1642 used in our system, which enables a maximum unambiguous velocity of around 3 m/s and a maximum sensing distance of around 10m.

The radar streams data through SPI to a Teensy 4.1, which streams the data to a laptop via USB for further signal processing and detection at 10 FPS. The total sensor bundle costs $386 with the core components (i.e., the SoC and the radar chip with antenna) costing less than $30. We group our system's sensing modalities into three categories, each featuring a unique corner reflector configuration.



Parameter	Configuration
Start Frequency	77 GHz
Sweep Slope	25 MHz/us
Sweep Bandwidth	3225 MHz
Sampling Frequency	2000 ksps
Number of chirps in one frame	32
Number of samples per chirp	256
Idle time between chirps	154 us
Number of transmitters	2
Number of receivers	4
Frame rate	15 fps

Figure 7.3: Our standard FMCW signal processing pipeline (left) and the configuration of the TI AWR1642 FMCW radar in our system (right).

7.4.1 Sensing Vibration and Motion without Reflectors

We identified a wide range of activities that can be sensed without enhancement of reflectors. Such activities often have a relatively large magnitude of vibration or motion such as running faucet and operating CNC machines.

Our processing pipeline for sensing vibration and motion is solely based on the range-doppler map. For each target activity, we first define a 1-m range of interest centering around its location on the range-doppler map. Within this range of interest, we engineered several features for our machine learning detection pipeline. Specifically, we measure three time-varying signals which include 1) the sum of pixels, 2) the sum of x-components of optical flow, and 3) the sum of y-components of optical flow. We then calculate FFTs on these three signals with a window size of 32, the results of which are concatenated. We also concatenate the maximums of pixels along each velocity bin on the range-doppler map to make the final feature vector (i.e., $16 \times 3 + 32 = 80$ data points). We build a two-label classifier with Scikit-Learn (v 0.22.2) SVM (linear kernel) to detect presence of activities (i.e., the on/off state of appliances and objects).

7.4.2 Sensing with Static Reflectors

As we discussed earlier, some objects require reflector enhancements. The most straightforward approach is to directly attach a corner reflector to a target object. This static configuration means the reflectors are fixed to their host objects with no structural changes, as opposed to dynamic configurations which we will discuss later.

Fabricating Low-Cost Corner Reflectors

Commercial corner reflectors are heavy because they are made of metal sheets. For example, the one used in our pilot test weighs 276 g. Additionally, they often cost more than $50. Obviously, they are too heavy and expensive to be instrumented onto everyday objects. Therefore, we set out to investigate approaches to make low-cost corner reflectors, ideally of low weight and compact sizes. Our final corner reflector prototype is made out of tin-coated cardboard papers (Figure 7.4 left). The total cost of materials, which include tin-coated paper, Scotch tape, and glue to adhere corner reflectors onto objects, is less than 10 cents per unit.

In a comparison study with commercial-grade corner reflectors, we found a comparable RCS of our ultra low-cost corner reflector with two commercial grade ones (Figure 7.4 right). The two commercial grade corner reflectors yield a 0.63 and a -1.34 dBsm RCS. In comparison, the RCS of our corner reflector measures -2.42 dBsm, which we found sufficient. Because our corner reflector is lightweight (< 3 g), we can easily attach them to a wide range of objects and surfaces such as microwaves, kitchen hoods, and dishwashers.

Processing Pipeline

Corner reflectors can be used to enhance SNRs for vibration and motion sensing. The processing pipeline for objects with vibration and movements in the previous section applies here. In addition, reflector enhancement enables the unique sensing of translational movements, which we achieved by tracking positions of reflectors (i.e., peaks on the range-azimuth map). Figure 7.5 shows an

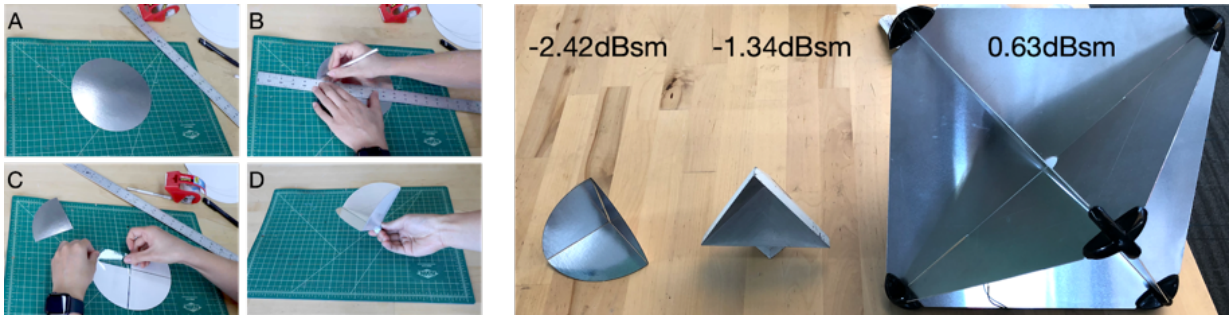


Figure 7.4: Our corner reflector is mostly made of ultra low-cost tin-coated cardboard paper (left). Comparison between our corner reflector and two commercial-grade ones (right).

example where we can detect the state of a sliding door by tracking a corner reflector attached to its top.

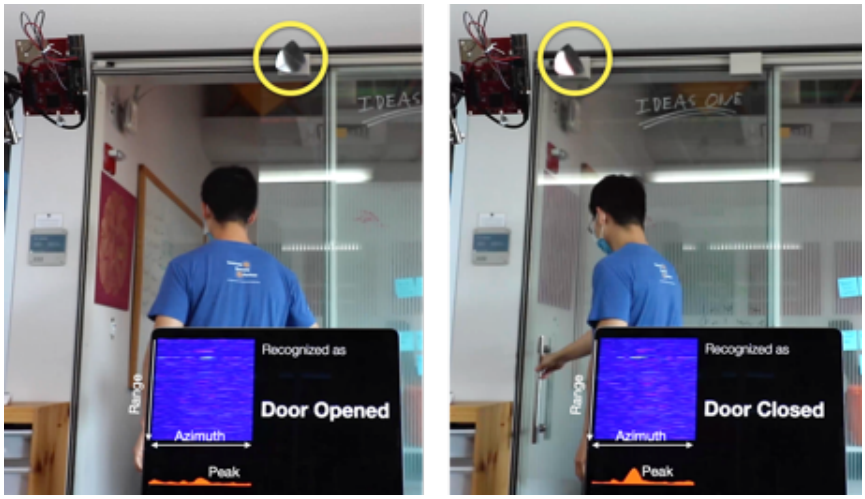


Figure 7.5: We instrument a corner reflector on a sliding door to monitor its state by tracking the position of the reflector.

In another application, we designed a water level sensor for water dispensers. To transform water level into RF characteristics for detection, we fabricated a corner reflector with hollow plastic structures covered in thin aluminum foil which floats at the water level (Figure 7.6). We then track the position of its peak on the range-azimuth map as an indicator of the water level. Note that in this application, the radar was rotated 90° so we can use azimuth data to infer elevation changes.

7.4.3 Sensing with Dynamic Reflectors

Beyond relying on static corner reflectors, we also explored dynamic reflector mechanisms which change their geometries when their host objects are at different states, resulting in characteristic RCS changes that can be used for a wider range of applications. Same as the static reflector

Figure 7.6: Tracking a sensor that floats at the water level for monitoring the amount of water left in the dispenser.

sensing, our signal processing for dynamic reflector mechanisms is also based on the range-azimuth map where we threshold and track peaks within regions of interest. We now describe the mechanisms we designed as well as their applications.

Door Sensor. To enhance the change of RCS between closed doors and opened doors, we designed a foldable door sensor (Figure 7.7 left) which can be easily instrumented on door hinges. When the door is fully opened, the door sensor forms a corner geometry that best reflects RF signals, resulting in a maximum RCS (-2.4dBsm). When the door is not fully opened, the door sensor loses its corner geometry, which results in a significantly smaller RCS (-5.4dBsm). The RCS of the door sensor is then thresholded for detection.

HVAC Register. We also used dynamic corner reflector mechanisms to translate events such as airflow from HVAC that are otherwise nearly invisible to RF sensors, to mechanical movements that radars can easily detect. Figure 7.7 right shows our register design which is built around a stainless-steel fan with 14 blades. This register is connected to the inner side of a damper, which hides it from users' sight. Once the HVAC is on, the airflow drives the register to spin, which results in strong doppler signals. Moreover, airflow at different speeds drives the register to spin at different speeds, which results in distinguishable patterns for airflow speed sensing.

Figure 7.7: Two dynamic reflector mechanisms designed for door state detection (left) and HVAC airflow sensing (right).

Toggle Switch. Besides activity recognition, our system also provides a feasible path to wireless room-scale interactions. As an example, we designed a toggle switch (Figure 7.8 left), which encodes the switch state by the angle of a corner reflector (i.e., up=40°, and down=20°). This change in angle results in an RCS change from -5.4 to -8.3dBsm.

Dimmer Switch. Similarly, we designed a dimmer switch (Figure 7.8 right), which changes its RCS by occluding a varying fraction of a corner reflector per its switch position. The resulted RCS changes can be used in regressions to infer the switch position.

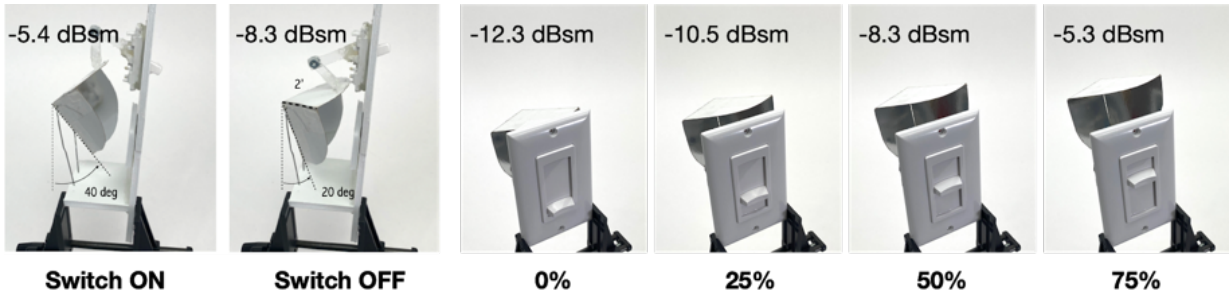


Figure 7.8: Two controllers based on corner reflector mechanisms: a toggle switch (left) and a dimmer switch (right).

7.5 Evaluation

To investigate the performance of our system, we conducted a study at three locations: a kitchen, an office, and a workshop where we collected data from four common appliances and objects at each location (Figure 7.9).

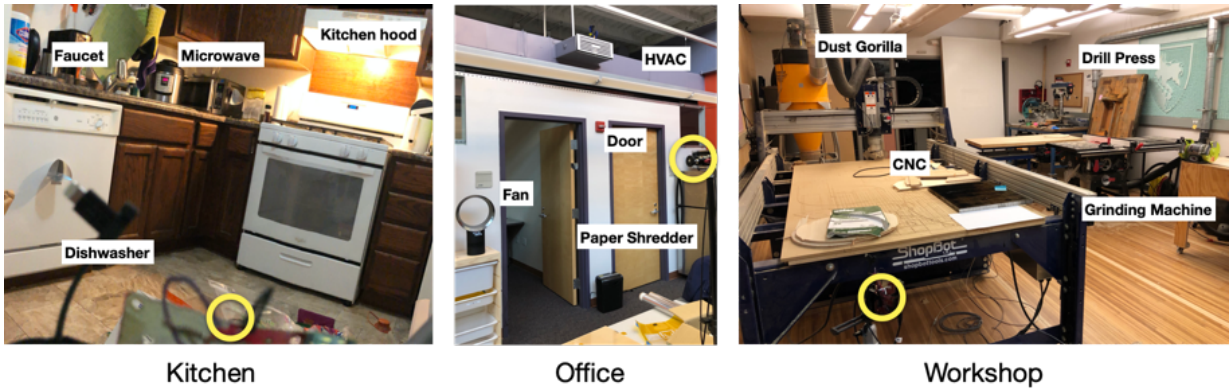


Figure 7.9: The three locations (kitchen, office, and workshop) where we conducted the study. The tested appliances and objects are labeled. The sensor locations are denoted by yellow circles.

7.5.1 Procedure

At each location, we first attached the radar where all test objects were within its field of view. We then manually measured the 2D locations of these test objects concerning the radar to calibrate the regions of interest, after which the data collection started. At each location, we first turned on one object, during which we collected 50 data points labeled as positives. We then turned off the object and collected 50 more data points labeled as negatives, before moving on to the next object. The order of the objects was randomized. After we collected all four objects within the location, we then collected 50 data points from a user walking around at that location also as negatives, which completes one round of data collection. In total, we collected three rounds of data with a 10 minutes gap between the second and the first, and a 1-week gap between the third and the first round. We used the first round of data collection to train the machine learning models (i.e., one

for each location), which were tested on data collected from the second and the third round. In total, we collected 4,050 data points (i.e., (4 objects x (50 positives + 50 negatives) + 50 walking) x 3 locations x 3 rounds).

7.5.2 Results

On average, we achieved an on/off detection accuracy of 96.2% (SD=5.9) across three locations and two test rounds (Figure 7.10). Specifically, round two (i.e., testing with a 10-minute gap after training) achieved an accuracy of 97.3% (SD=3.8) and 95.0% (SD=8.2) for round three (i.e., testing with a 1-week gap after training). We suspect that the slight decrease in accuracy was caused by environmental configuration changes and small sensor displacements. Within the errors, we found that 55.6% of false positives were caused by user walking, and 84.3% of false negatives were from dishwasher running. Although the total percentages of errors were small, the results on the sources of error suggest further investigation that might lead to superior machine learning classifiers that better distinguish user signals, and are more immune to environmental changes.

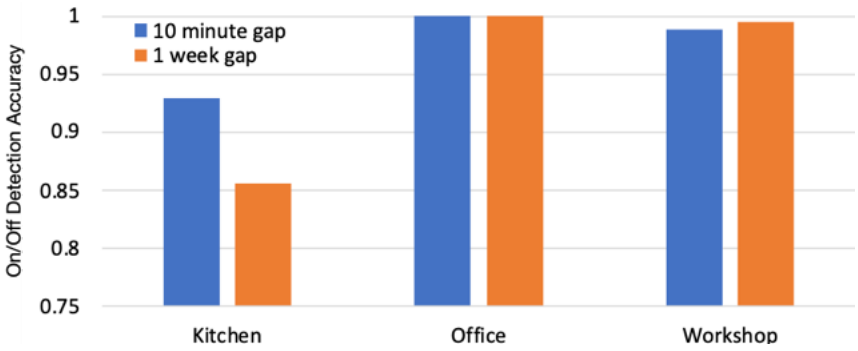


Figure 7.10: Accuracies of on/off detection across three locations and two data collection rounds.

7.6 Limitation and Future Work

The first limitation of our system is that we cannot sense through objects due to the high operation frequency band we used. Although the use of high-frequency bands improved our overall sensitivity toward small displacements (i.e., vibration), most RF energy at high frequencies will be reflected and scattered by everyday surfaces, which means our sensing approach favors line-of-sight detection. The occlusion layers must be thin and not conductive for our system to sense through them. The toggle switch and dimmer switch demos worked through a layer of foam board. Our current system could not sense objects behind brick walls, wood studs, or appliances with metallic shells.

Another limitation of our current system is related to its sensing dimension, resolution, and field of view. Although our current system can only track objects on the range-azimuth plane (i.e., 2D localization), it is not the limitation of FMCW radar. Highly complex antenna arrays that can achieve elevation tracking as the third dimension for 3D localization with higher spatial resolutions exist. Similarly, a wider field of view (even 360°) is possible with multiple multiplexed

antenna arrays each covering a different section of an environment. We plan to further investigate this in the future.

Finally, our current system relies on user-calibrated regions of interest. Although we did not demonstrate concurrent activity sensing, we anecdotally found that activities that are sufficiently spaced apart could be sensed simultaneously for their little interference on the range-doppler map and the range-azimuth map. However, future work is merited to further investigate concurrent activity detection with potentially more advanced beamforming techniques and Angle of Arrival approaches to better disambiguate signals from different sources in complex environments.

7.7 Summary

We developed an FMCW radar system for sensing activities and interactions in a room. We explored three sensing modalities with different configurations that leverage low-cost corner reflector mechanisms which we designed to work in concert with the FMCW radar. These three sensing modalities cover a wide variety of smart home applications that rely on appliance/object use, state of operation detection, and airflow speed and water level monitoring. With a two-week evaluation at three locations, we have demonstrated that our system can be accurate and robust. We envision our system to be integrated into future smart appliances such as speakers, light bulbs, and home assistants to enhance their ability to infer user context for more fluid and powerful interactions in a wider spectrum of smart room applications.

Chapter 8

City-Scale Sensing Using Existing Retroreflective Signs and Markers

8.1 Introduction

Wide-area sensing also applies to urban environments, which are vast and complex, with a wide range of facets that city planners, public transit agencies, business owners, citizens, and many other stakeholders might wish to sense [61, 70, 122, 163]. This sensor data can be used by cities to increase mobility [38], improve operational efficiency [107], boost quality of life [122], and become more sustainable [94]. For this reason, the concept of smart cities is gaining traction, but due to high cost, city-scale deployments are rare today except in mature and narrow use cases, such as traffic flow [18, 19] and gunshot detection [43, 137]. This leaves an exceptionally long tail of small, but interesting smart city applications with no viable source of data and without budgets to put them into practice.

The primary reason for such high costs is that today's smart city sensing paradigm is built around massive deployment of physical sensing infrastructure. Chicago's Array of Things project [20], launched in 2015, is a notable effort in North America. Backed by at least $4M in funding, the project has managed to deploy 130 sensing nodes on poles through the city by January of 2020 – a glacial and expensive rollout by any measure. In general, these smart city nodes are weather-hardened devices that require power and connectivity, and often cost thousands of dollars to install on a pole, not including hardware purchase and maintenance costs, which are almost certainly many times more expensive. Thus, the central thesis of our research is that to unlock the potential of smart cities, a truly tectonic shift in city-scale sensing is needed to 1) dramatically reduce deployment costs, 2) virtually eliminate maintenance, and 3) generalize across a wide array of sensing applications that can improve cities. Today, 4.1 billion people – a majority of the world's population, live in urban areas, and thus cities occupy a central role in modern human life. Understanding cities is to understand ourselves, and if cities can be made smarter through sensors and computing, we can improve quality of life for many people, making this very much an HCI research problem.

In this chapter, we present one candidate method for ultra-wide-area sensing that could be applied to the expansive urban landscape and meet our aforementioned three criteria. More

specifically, we suggest a novel application of laser vibrometry, used extensively in fields such as industrial quality control, aeronautics, and structural health monitoring [52, 95, 101]. These sensors (described in greater detail later) are generally used to sense at a few meters range. However, the sensing range can be significantly increased by attaching a retroreflective sticker (or equivalent marker) to the surface of interest, which greatly increases the returned laser light and permits accurate measurement. Importantly, the "smarts" and expense is limited to the vibrometer itself (e.g., the entry-level Polytec VibroGo costs around $30k [14]), with retroreflective tape costing less than 1 USD per yard. Of course, given the number of facets we might wish to sense across a city, this approach does not necessarily meet our first criteria (i.e., having to go out and specifically tag all of these objects would drive up deployment cost).

The second key insight that drove this research project was the fact that many cities, very fortuitously, have already deployed retroreflectors by the tens of thousands. Street signs, construction barriers, road studs, lane stripes, traffic control signage, license plates, bicycle reflectors, and many other objects in the urban domain are made with retroreflective materials to improve visibility at night. Figure 8.3 offers a small sampling of the items we found during our explorations. In a five-city survey we conducted, described later, we found an average density of about 7,000 retroreflectors per square kilometer. In essence, cities have already unknowingly invested and built out all of the sensing infrastructures our approach needs.

Our approach allows these existing reflectors (requiring no power or connectivity) to be functionally appropriated as low-resolution accelerometers at extreme ranges. As much prior work has shown, accelerometers can be rich signal sources used to power a wide range of domestic and commercial activity recognition tasks (for e.g., [90, 91]). We show how many of the same techniques can be applied to the civic domain, and we showcase eleven example uses to illustrate the potential and generalizability of our approach, meeting our third criteria. Before that, we describe the core concepts that underpin the physics of our approach and reviewing key related work. We then describe our proof-of-concept system, which cost us around $1000 to build (much less in volume production), before moving to an evaluation that validates our sensing approach. Finally, we conclude this work with a wide variety of smart city applications enabled by our sensing technique.

8.2 Principles of Operation

As a primer, we quickly review some relevant physics. Traditional reflectors, like a mirror one might have at home, bounce energy at the same angle at which a ray approaches the surface, mirrored across the surface normal. In contrast, retroreflectors return energy back along the exact same vector as the source. Like reflection, retroreflection is possible with all energy waves, including radio frequency, light, and sound. This ability to return energy towards the source of emission is particularly useful for urban signage and markers, where illumination from vehicle headlights makes them highly visible without the need for powered artificial lighting at the sign (i.e., the sign is completely passive). There are two general types of retroreflector geometry used for wavelengths in and around the visible light spectrum (Figure 8.1). The first are spherical beads, made of glass or plastic that have a carefully formulated index of refraction such that light escapes the sphere at the same angle of incidence. A second bead variety uses half-mirrored

spheres to achieve retroreflection. Finally, retroreflectors made from right-angle prisms (also called corner reflectors) are very common. For a more in-depth review of the physics of reflection, we recommend [40].

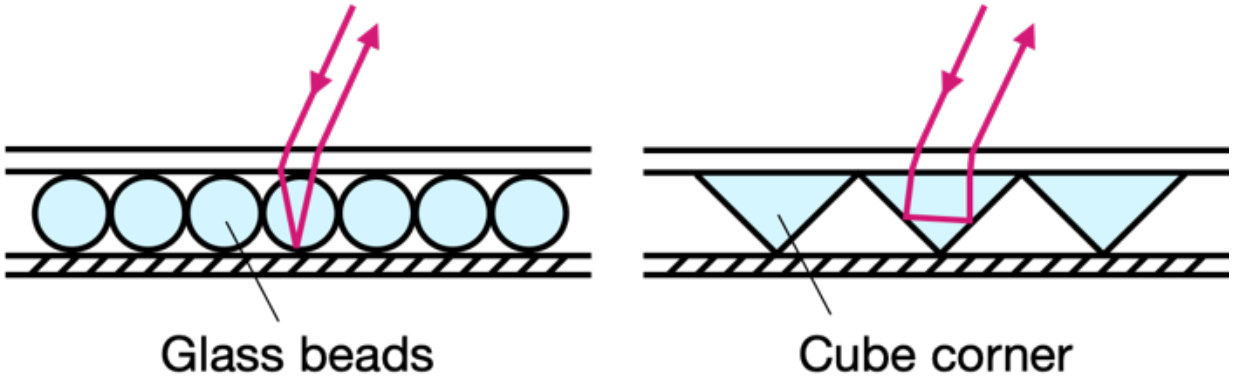


Figure 8.1: Two typical micro-structures of retroreflective materials: beads and cube corners.

There are two main types of laser vibrometry. The first type is built around interferometry [44, 63], where minute surface vibrations of an object cause laser light to return at a slightly different phase than the emitted signal, which is then used as a reference. Mixing these two signals together (using sophisticated internal optics) creates constructive and destructive interference, which can be read by a sensitive photodetector. The second type of laser vibrometry takes advantage of the reflected laser speckle pattern [166] (Figure 8.2). Object vibrations cause the reflected pattern to oscillate, and thus a photosensor placed in the light field will observe the speckle pattern translating through its field of view, causing the received intensity to rise and fall. This AC signal encodes an object's vibrations, and the DC component is also useful for detecting the presence of retroreflectors in the environment, and if they become occluded at a later point. For our proof-of-concept hardware, we use the intensity-based approach, but both types of laser vibrometry are applicable.

8.3 Retroreflectors In-the-wild Survey

While we anecdotally knew that there were many retroreflective signs and markers deployed in cities, we wished to more formally catalog their uses and placements, and also derive an estimate of their total volume. For this, we surveyed two block regions adjacent to the city centers of five mid-sized cities – two in North America, two in Europe, and one in Asia. We only cataloged along public streets, and we did not survey any interior spaces, nor exterior spaces that existed inside of the street grid (e.g., surface parking lots). We also treated vehicles as a special category, as they often feature seven or more discrete retroreflectors (depending on local law, e.g., [11]). Rather than counting these as separate retroreflectors, we counted them as one in our tally.

In total, we found 1,084 retroreflectors in a cumulative area of 172,813 / m^2 (which includes block interiors which we did not survey). This equates to an average retroreflector density of 7,066.7 / km^2. Table 8.1 shows the details of our results. In our hometown, a mid-sized city of ~150 km^2, it means there is on the order of a million deployed retroreflectors, each of which is

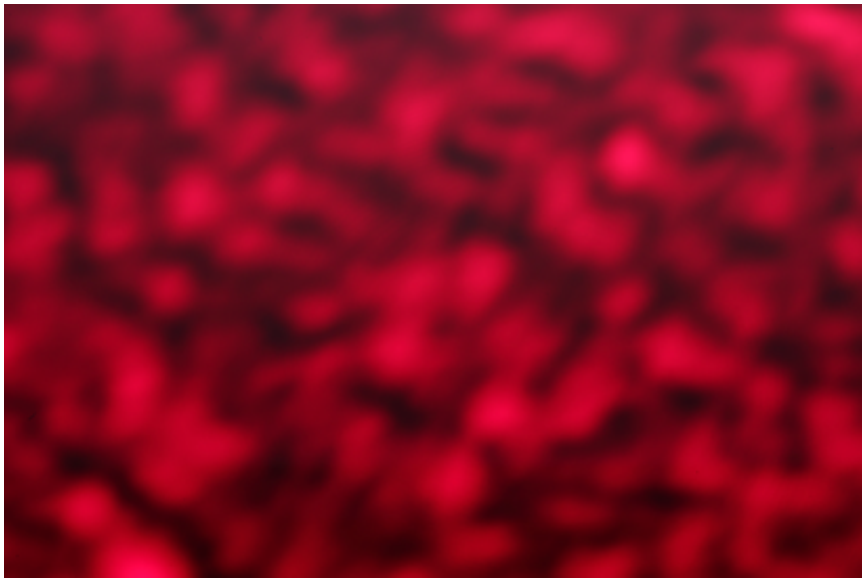


Figure 8.2: Speckle pattern reflected off of a remote retroreflective marker.

a potential smart city sensor. Figure 8.3 shows several representative retroreflectors we found during the survey. We categorized retroreflectors into six categories: 1) Retroreflectors placed on structures, such as buildings, covered bus stops and phone booths. 2) Retroreflectors on poles, such as traffic light and power line poles, as well as smaller poles used for signs (e.g., bus stop, yield, parking information, one-way street). 3) Retroreflectors on vehicles, which are a safety requirement that allow cars to be seen even when their lights are off. 4) Retroreflectors on trees, which are used to reduce collisions, but seem comparatively rare. 5) Retroreflectors attached to roads, such as road studs and fixed bollards, most often to help demark traffic flow (we note that painted road markings are often retroreflective, but we did not include them in this tally). Finally, 6) we found some miscellaneous reflectors that did not fit neatly into our previous categories, but what might be broadly described as "street furniture" [26] (e.g., reflectors attached to parking payment kiosks, trash dumpsters, fire hydrants, and mailboxes).

City	On Structure	On Pole	On Vehicle	On Tree	On Road	On Misc.	Total Count	Perimeter (m)	Area (m^2)	Density (count/km^2)
Ann Arbor	0	104	89	3	0	0	196	854.7	45491.0	4308.5
Pittsburgh	24	188	65	4	39	8	328	864.9	33038.3	9927.9
Harbin	17	11	99	0	11	0	138	906.1	47430.1	2909.5
Enfield	5	52	173	0	0	0	230	677.7	26990.5	8521.5
Stuttgart	11	103	71	0	6	1	192	573.9	19863.2	9666.1

Table 8.1: Results from our survey of retroreflective markers across five international cities.

8.4 System

In this section we describe our system, which includes hardware and software components working together to achieve our vision of long-range vibrometry sensing. Towards the end of our pipeline, we create several different processed signal streams that are used to power different sensing applications.

Figure 8.3: Various types of retroreflective signs and markers have existed throughout urban environments.

8.4.1 Hardware

Our system (Figure 8.4) consists of two major hardware components: 1) a pan-tilt motor platform, and 2) a sensor bundle. The pan-tilt motor platform is set on a tripod and controls the angle of the sensor with a resolution of 0.00625° — corresponding to a 10 mm step at a distance of 100 meters. This pan-tilt platform allows us to scan the environment for retroreflector search, and to cycle through different sensing directions when multiple retroreflectors are present.

The sensor bundle consists of a 200 mW 650 nm laser, a 1000-m range finder, and a photodetector (Thorlabs PDA100A2), calibrated by a secondary precision pan-tilt platform. We inserted an optical band-pass filter (Thorlabs FB650-40) inside a 13.5cm optical tube to make the photodetector more immune to the ambient light (e.g., sunlight, reflections from windows, street lights, etc.). The output of this photodetector is amplified and sampled by a digital audio preamp (U-PHORIA UMC404HD) before streamed to a laptop for detection.

8.4.2 Eye Safety

Our laser output power is 200 mW (Class III B), which is hazardous for eye exposure at short distances, but eye-safe at distances longer than the Nominal Ocular Hazard Distance (NOHD) [74]. At constant laser power, NOHD is inversely proportional to laser divergence. As a safety measure, we adjusted the divergence of our laser in accordance with ANSI Z136.1 [9]. At our tightest divergence (1.62 milliradians), the NOHD is 95 m. By defocusing our laser, we can make it eye-safe down to around 46 cm, where the energy density is under the recommended 2.5 mW/cm^2

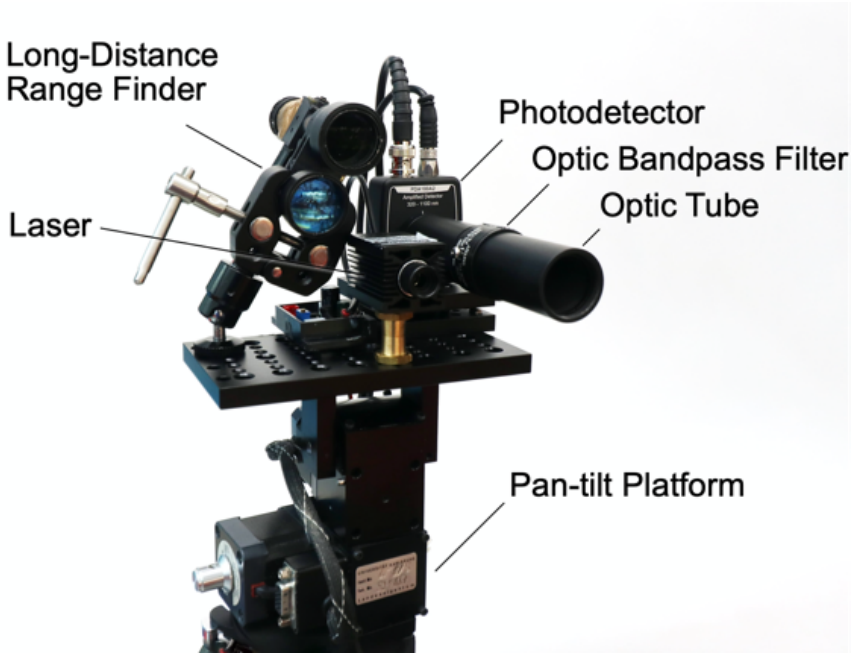


Figure 8.4: Our system consists of two major hardware components: a pan-tilt motor platform and a sensor bundle.

exposure. There are two additional ways to improve safety. First, we envision the strategic placement of our system at higher vantage points to remove the possibility of intersecting with traffic and pedestrians inside the NOHD. Finally, our system measures the reflected light intensity within 50 microseconds of laser power. This is well under the 250-millisecond exposure hazard threshold detailed in [9] if we are operating under 2.5 mW/cm^2. We use this 50-microsecond interval in our reflector search (described in the next section), meaning even accidental contacts with humans is non-damaging. Likewise, when we are capturing signal from known reflectors, we can easily detect if they are occluded (e.g., by a car or human), and we turn off the laser (again within 50 microseconds) and cycle to the next target.

8.4.3 Retroreflector Search

Before we can leverage existing signs and markers, the first step is to identify existing reflectors in their environment. For this, our system scans the environment to get a depth map and a reflectance map (Figure 8.5), based on which we find the reflectors with our search algorithm. For each scan, we first define the scope of search. Then our system performs a row-by-row scan with an angular resolution of 0.05°. At each angle, we measure the distance between the environment and our system with the rangefinder for generating the depth map. After the distance measurement, we toggle the laser with 50 microsecond intervals and measure the output of the photodetector, before moving our sensor to the next angle. By measuring the differences in the photodetector outputs between when the laser is on and off, we can build up a reflectance map that is independent from the environmental lighting. This enables us to detect small reflectors at a high level of accuracy.

Our retroreflector search algorithm uses both the depth map and the reflectance map. We

begin with a denoising process — by using the depth map to scale the raw reflectance values (measurements further away are scaled up more heavily), we mitigate the effect of stray light in the foreground. After this initial processing step, we eliminate all reflectance values under a predetermined threshold, and run a mean-shift clustering algorithm on the remaining points.

Figure 8.5 shows an example of our retroreflector search. Specifically, we deployed our sensor at the outside of a commercial building, pointing at a parking lot area of roughly 1,800 m^2. We defined a scan scope of 110 azimuth degrees by 16 elevation degrees, and started the retroreflector search. After the search was completed, we compared our results with our manually logged ground-truth counts and locations of the retroreflectors in the environment.

Our system found 22 out of 29 retroreflectors in the environment. Two of the missed items were oriented at extreme angles relative to our rig — and while they did register on the reflectance map, the sensed light was not bright enough to escape our thresholding process. The other five items were in tight groups which were detected by our system, but identified as only two items, as opposed to five distinct retroreflectors for being too close together.

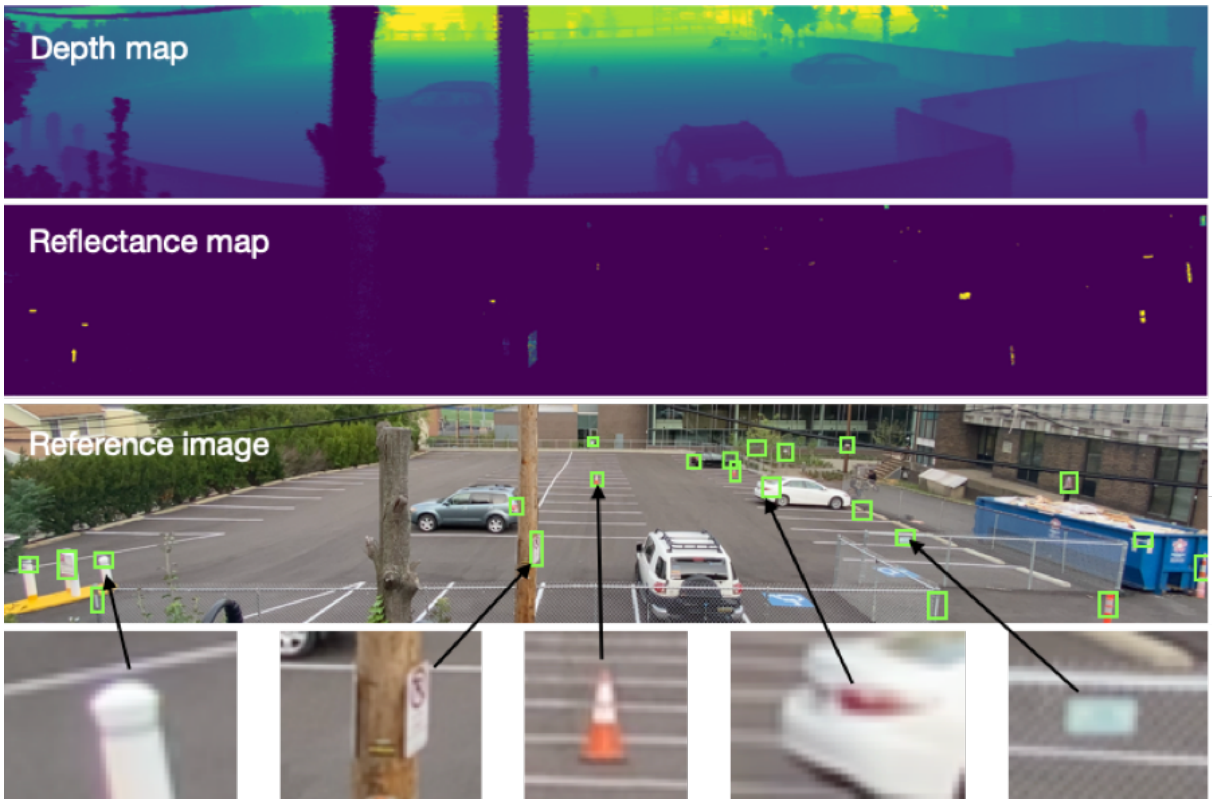


Figure 8.5: Depth map, reflectance map, and reference image of a parking lot where we conducted the retroreflector search study. Green rectangles highlight the identified retroflectors.

8.4.4 Presence/Occlusion Sensing

To sense the presence and the occlusion of retroreflectors, we used a Teensy 4.1 running at 600 MHz with a built-in ADC to measure the output voltage of the photodetector. A resistor attenuator

is used to downscale the photodetector output to the dynamic range of the Teensy 4.1 ADC. Each measurement takes 50 microseconds to complete. These measurements are then thresholded for the presence/occlusion detection, wherein values higher than the threshold value indicate retroreflectors' presence and not being occluded.

8.4.5 Vibration and Motion Sensing

Vibration and motion sensing requires a higher sampling rate, for which we used an audio preamp which samples the photodetector output at 44.1 kHz. We process the raw measurements with FFT (window size = 2048), the result of which is bandpass filtered. Note that we apply different band filters for different applications. We then sum the non-DC components as an indicator of the signal power, which is thresholded for several applications. There are also applications that require more complicated detection for which we use a standard machine learning pipeline (i.e., SVM with a linear kernel) for classification and regression.

8.4.6 Comparison to Accelerometer

To investigate our system's ability to sense vibration and motion, we conducted a comparison study with a gold standard sensor: an accelerometer. Specifically, we used a large diaphragm speaker (54 cm diameter) powered by a 600 W amplifier to output a swept frequency from 2 Hz to 2 kHz. We physically attached an analog accelerometer (ADXL335) and retroreflective tape (3M 79961 Scotchlite) to the center of the diaphragm. The analog accelerometer was read by a USB audio preamp at 44.1 kHz, while our laser vibrometry setup measured vibration using the retroreflective tape at a distance of 128 m. Figure 8.6 illustrates the result; the main sweep is clearly visible. Note the harmonics seen are due to the physical design of the speaker, which cannot output a pure sine wave without secondary oscillations.

To complement to our controlled experiment, we also selected one real-world example. We affixed an accelerometer to a mailbox and recorded the signal in the same manner as above. The USPS logo was already retroreflective, so we used that in conjunction with our system (at 30 m). The signal of someone depositing a letter into the mailbox is shown in Figure 8.7.

In both experiments, we found that although the laser vibrometer signal was considerably coarser and noisier than an actual accelerometer, the fidelity is more than sufficient to power most uses (e.g., it is very apparent when the mailbox door is opened). Additionally, we found that our vibrometer signal has constant low-frequency components due to ambient vibrational and optical noise, which we remove with bandpass filters in applications. Overall, we believe this to be a very strong result given that our system is sensing at a distance (vs. being affixed to the object and requiring local power).

8.5 Evaluation

There are many factors that affect the sensitivity of our system, including retroreflector type, laser color, retroreflector color, retroreflector angle, and distance from the sensor. It was not possible to test all combinations of these parameters, and so instead we investigate each factor

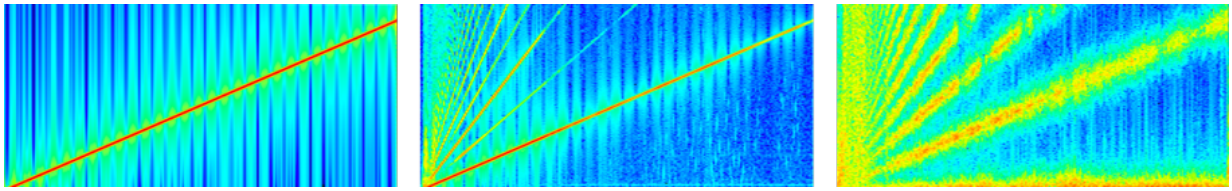


Figure 8.6: FFT spectrum of a 2 Hz - 2 kHz swept-frequency signal (left), the recorded signal from a surface accelerometer on the speaker diaphragm (center), and the recorded signal from our long-distance vibrometer at 128 m, pointing at a retroreflective marker on the speaker diaphragm (right). All spectrograms in this chapter are presented on a log scale.

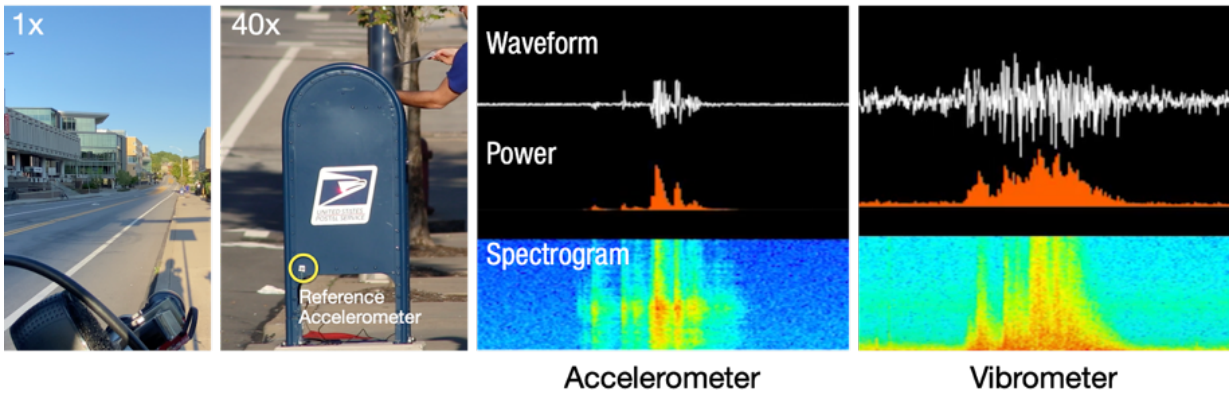


Figure 8.7: Sensing the vibration on a mailbox from someone depositing a letter, with an accelerometer attached to the mailbox surface, and with our laser vibrometer at 30 m.

sequentially. All data was collected in daylight, between midday and late afternoon. While great care and effort were taken to ensure all measurements were as comparable as possible, operating in an uncontrolled environment (i.e., the real world) meant that some variability in our measurements was unavoidable. Small inconsistencies arose from e.g., the sun changing angle over the duration of an experiment, variations in haze, and even light paths moving or oscillating due to air turbulence and temperature variations (i.e., mirages). The measured ambient illumination varied from 52 to 84 klux. More importantly, the main trends are apparent (e.g., signal decreases as reflector angle increases). Thus, we encourage readers not to place emphasis on any single reading.

8.5.1 Test Retroreflectors

We selected 16 markers and retroreflective materials (Figure 8.8) that were representative of the wide variety of retroreflective surfaces we found in the wild. Retroreflector A and B are commercial retroreflective tapes that can be adhered to any surface. C through F are commercial traffic cones and bollards that have a tape-like reflective material bonded to their surface. G is an aluminum sign with a retroreflective layer (also tape-like). H is a metal-bodied road stud with an inset plastic reflector of small prisms. I through N are also molded plastic prism arrays with various sizes, shapes, colors, and quality. Finally, O and P are large prism reflectors typically used in surveying. As described in the Principles of Operation section, there are also two common

ways to make retroreflective materials, using beads or prisms. For reference, A-E uses beads, while reflectors F-P use prisms. While prism-based reflectors are generally considered to be more efficient, in our later tests, we only found a small effect on retroreflector type. Instead, we found that the signal varied much more based on construction quality, including the types of plastics used and the quality of the mirror surface finish.

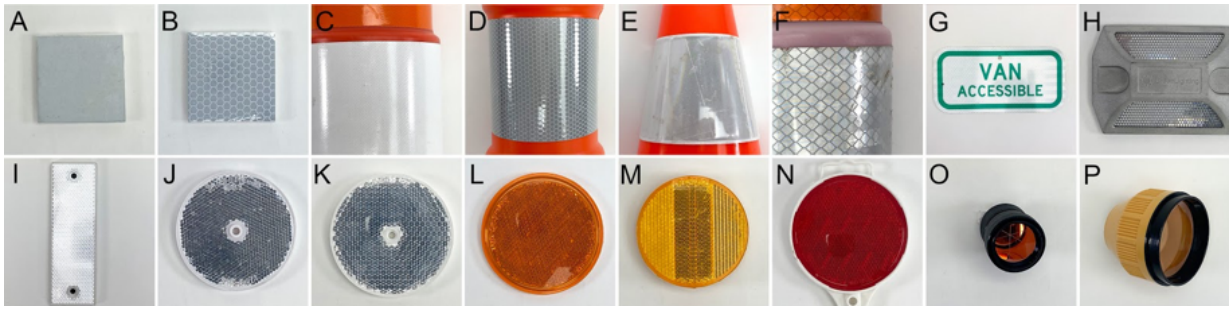


Figure 8.8: The 16 retroreflectors we tested in our evaluation - A to G are tape retroreflectors with no physical depth (micro-prisms), H to N are retroreflectors with a physical depth (molded plastic with reflective coating), and O and P feature large single prisms. We selected A, B, E, and M for our in-depth motion and vibration evaluation.

8.5.2 Test Apparatus

To provide a controlled motion and vibration signal, we use the same 54 cm diameter speaker, powered by a 600 W amplifier, as described in Section 5.6. As before, different retroreflectors were affixed to the center of the speaker, and actuated with a signal from a laptop. Given that our tests ranged up to 920 m, in outdoor conditions, we decided to load our signal generation apparatus into a van for mobility (Figure 8.9). The engine was turned off during tests so as not to introduce any extraneous vibrations.

8.5.3 Laser Color

To explore whether the wavelength of light had a significant effect on SNR, we tested three different 200 mW lasers: red (650 nm), green (532 nm), and infrared (940 nm). Unlike the other experiments in this section, all measurements were completed without an optical bandpass filter. Retroreflector J (clear plastic) was used at a distance of 32 m. We did not observe any significant difference in signal strength, measuring 0.15 V, 0.19 V and 0.17 V for the red, green, and infrared lasers, respectively. For all subsequent experiments, the red laser was used.

8.5.4 Reflector Color

Although markers can be purchased in many different colors, they are most often red, orange, or clear (white) in urban use cases. This color effect is achieved with a semi-translucent, tinted plastic cover material, which allows light with a desired color to pass (and reflect it back out), but attenuates other wavelengths. For this test, we purchased a matched set of four colored

Figure 8.9: We loaded our signal generation apparatus into a van for mobility, which allowed us to test various distances in outdoor conditions.

retroreflectors. Using our red laser, we measured their reflectance at 32 m. Unsurprisingly, the red retroreflector had the strongest signal at 0.34 V, followed closely behind by the clear (0.24 V) and orange (0.23 V) retroreflectors. Signal then drops off significantly as we move along the color spectrum, with the green retroreflector at 0.01V and no measured signal with our blue retroreflector.

8.5.5 Reflector Angle

In general, the design of retroreflectors means that they are most efficient at reflecting when perpendicular to a light source, with reflectively dropping off as the angle of incidence increases. To measure this across our set of 16 retroreflectors. For the effect of the lasers' angle of attack, we investigated all angles between 0° and 75° in 15° steps, where 0° presents the laser being perpendicular to the retroreflector. We performed all angle tests at 32 m distance and indoors with constant lighting conditions. As expected, we found the higher angle of incidence, the weaker the signal (0° yields the highest reflectance, and 75° and beyond result in almost no reflectance; Figure 8.10).

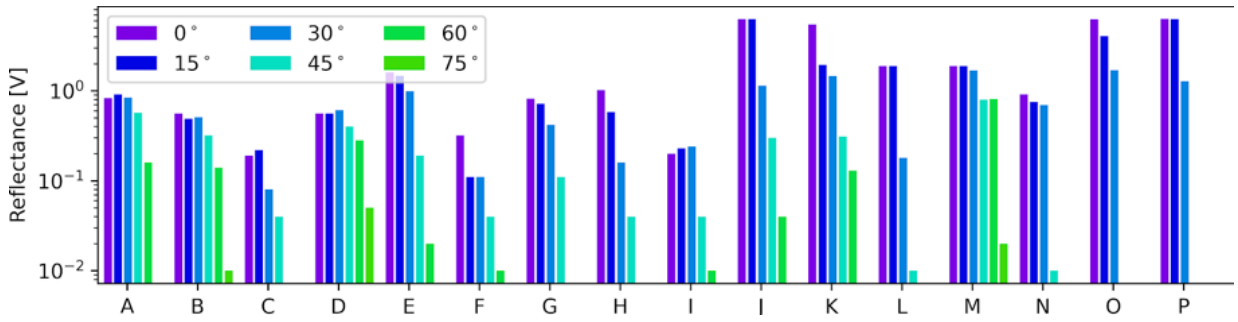


Figure 8.10: Reflection intensity for the 6 different angles grouped by the 16 types of retroreflectors.

8.5.6 Distance - Presence/Occlusion

With the presence and occlusion test, we want to investigate the maximal distance at which we can sense the reflective signal. This test was used to measure the quality of the reflectors and how well our sensor can pick up slight signal changes. We had to leave the building and find a long straight stage of land in which we can test the reflectance at different distances. This test was performed on a 920 m long straight street. We tested the reflectivity from 16 m on in steps of power of two up to 512 m and a maximum of 920 m, and measured deltas of the photodetector output when the laser was toggled on and off (see Figure 8.11). As expected, the reflection intensity decreases with distance. We also found variances across retroreflectors with the retroreflector O and P yielded best reflection at long distances due to their large prism sizes. However, all retroreflectors yielded detectable reflection intensity at the maximum distance (920 m).

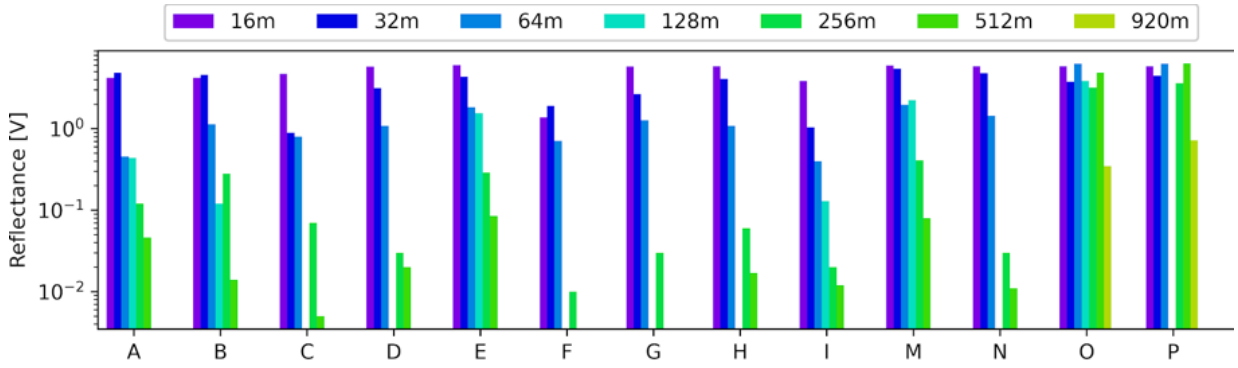


Figure 8.11: Reflection intensity for the seven different distances grouped by the 16 types of retroreflectors.

8.5.7 Distance - Motion

We also investigated our system's sensitivity to motion at different distances and on different types of retroreflectors. We selected 5 representative retroreflectors in this test. Specifically, we generated a period displacement signal with the speaker, from 0 to 2 Hz with an average amplitude of 4 cm. We then cross correlate this excitement signal with the measurements to calculate the received signal power, the result of which is shown in Figure 8.12. We found that

our system is capable of detecting the motion signal from all retroreflectors at all distances we tested. Overall, received signal power attenuates as distance increases, and there are differences across retroreflectors. For example, E and M yielded smaller attenuations, achieving a better sensitivity at long distances. This result well correlates well with what we learned from the Distance - Presence/Occlusion test where E and M also yielded more constant reflectance than other reflector types.

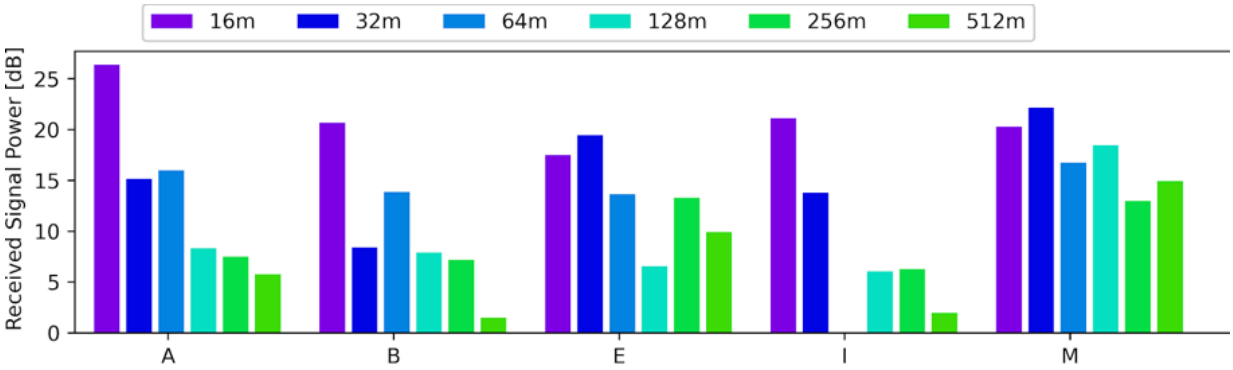


Figure 8.12: Received signal powers from the distance-motion test on five types of retroreflectors.

8.5.8 Distance - Vibration

To investigate our system's sensitivity to vibration, we conducted another test with the same procedure as the motion test, except that we generated a swept-frequency vibration with the speaker, from 2 Hz to 2 kHz. Figure 8.13 shows the result, which indicates that our system is capable of detecting vibrations from all reflectors at all distances. The sensitivity characteristics are very similar to the motion test, with distance being a significant attenuation factor on the received signal power and variances across retroreflectors. E and M again yielded more constant received signal powers across distances.

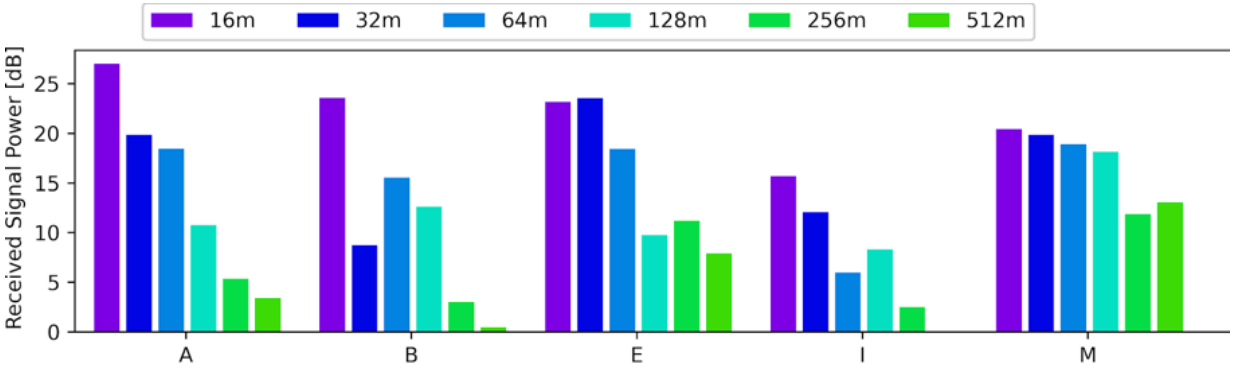


Figure 8.13: Received signal powers from the distance-vibration test on five types of retroreflectors.

8.5.9 Implications for Use

Our main results indicate that our system works in outdoor daylight conditions. It can detect the presence/occlusion of retroreflectors up to 512 m, some even up to 920 m. With the results showing little attenuation across certain types that feature large prisms, we suspect that our system could sense these retroreflectors even at tens of kilometers. It is also possible to detect motion and vibration from them up to 512 m. We found variances across angles of incidence, retroreflector colors and types. However, if our system has line of sight of a retroreflector with an angle of incidence of less than 45°, and the reflector color is compatible with the laser, then it is very likely that we can appropriate it as a sensor.

8.6 Example Uses

The practical applications for vibrometry in urban settings are immense. Given our signal is very accelerometer-like, we can generally subsume all prior applications across innumerable papers that have affixed powered accelerometers to objects and infrastructure (e.g., [91]). To illustrate this breadth of uses, we describe a dozen example applications spread across three signal categories.

8.6.1 Vibration Sensing Examples

The richest of our three signal categories is vibration. This signal allows for fine-grained detections, and more importantly, the rejection of false positive events that do not match the characteristic spectral signatures of what we are trying to sense. We now describe four illustrative example uses:

Mailbox Utilization

Using the retroreflective USPS logo on the mailbox seen in Figure 8.7, it is very apparent when someone opens the hinged door to deposit a letter. A simple incrementing count could be used to track the number of visitors, which could inform mailbox placement, but perhaps more interestingly, be used to dynamically schedule pickup based on the estimated volume of mail pieces. In this example, the mailbox is immediately adjacent to a road, so cars passing by also induce a vibration, but this signal is very different and thus easily ignored.

Bus Arrival Detection and On-Time Performance Tracking

Bus stops are often demarcated by a sign, which serves both riders and drivers alike. We found that buses pulling up to a stop creates a characteristic air front that causes the sign to flex and oscillate (Figure 8.14), much like a tuning fork. Smaller vehicles do not induce the same effect. Larger vehicles, such as trucks, can produce the same signal if they similarly decelerate to a stop next to the sign, though this is less typical. However, at least in some cases, this signal could be useful to power live bus monitoring apps and perhaps track on-time performance.

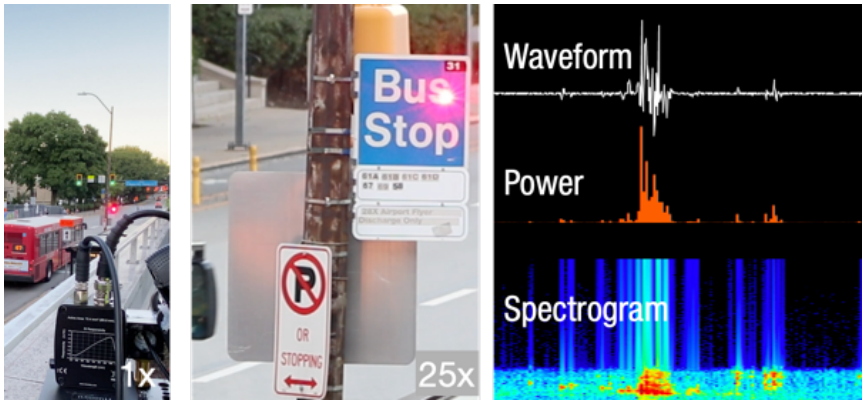


Figure 8.14: We threshold the power of unique frequency bands induced by a passing bus for detection.

Traffic Flow

Roads are painted with a wide variety of markings, including lanes, turns, and pedestrian crossings. Retroreflective glass beads are often incorporated into road paint to make the markings more visible at night. This means there are many millions of miles of "sensors" already installed on roads for us to use. As an example, we used a white shoulder line for detecting passing vehicles. The mass of a passing vehicle induces a low frequency oscillation of the road surface that can be easily detected and then counted over time to measure traffic volume (Figure 8.15). Anecdotally, we also observed the shape of the waveform changes by vehicle velocity and type (number of axles, vehicle weight) and velocity.

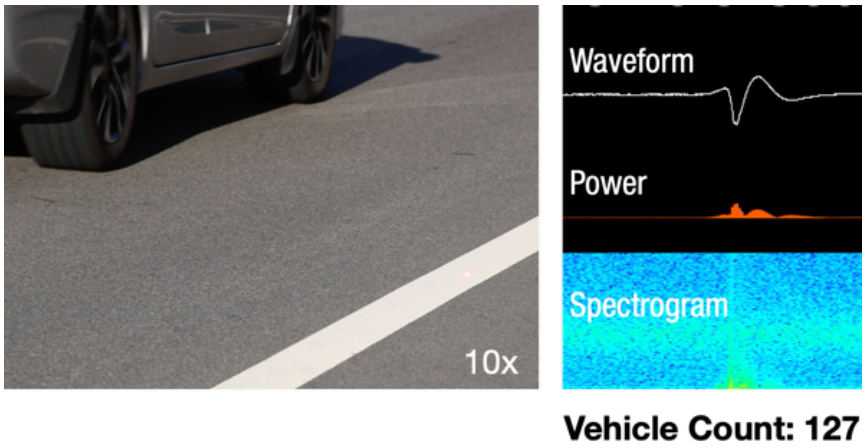


Figure 8.15: We detect peaks on the raw waveform to count passing vehicles.

Vehicle Type Detection

Although we observed that vehicle type might be inferred directly from road markings (previous example), a more direct measurement is enabled by sensing engine vibration using the license

plates of cars, which are retroreflective. As can be seen in Figure 8.16, the signals can be quite expressive, and we were able to train a vehicle classifier for cars, buses, and motorcycles. It may even be possible to infer engine type (Inline 4 vs. V6, diesel vs. gasoline) and RPM, though we leave this to future work. A downside of sensing vehicles in this manner is that it requires rapid scanning of an intersection to find license plates and take a measurement, and thus it is likely that only subsampling of traffic is possible.

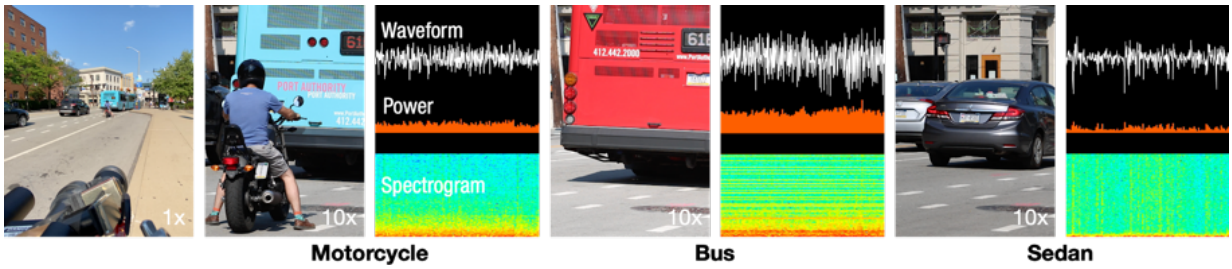


Figure 8.16: Type of vehicles can also be detected based on distinctive frequency components in their vibration.

8.6.2 Motion Sensing Examples

We categorized low frequency oscillations as movement, and built several example application classifiers not using spectral data, but rather mean power over time.

Block-scale Wind Monitoring

The large surface area of signs means they tend to catch the wind, causing deflection and oscillations, often at centimeter scales. Such gross changes can be easily captured by our sensor at very long-ranges. As an example, we recorded a pole-mounted street sign affixed to a wooden pole on a breezy day. Figure 8.17 shows three example signals at different wind intensities. Note that wind signals appear as paired bumps – this is due to the wind deflecting the sign one way, and then having the sign swing back to its neutral position.

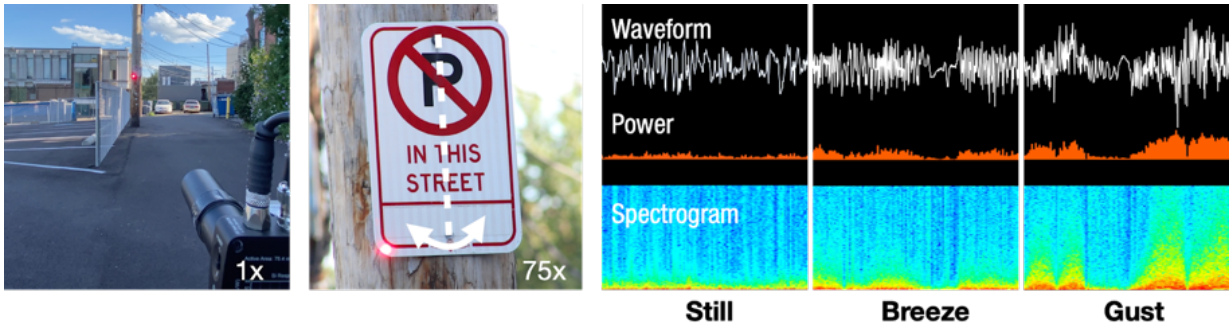


Figure 8.17: We monitor the wind speed through monitoring the mean power of non-DC frequency components over time.

Urban Tree Health

Like signage, trees also sway in response to wind (Figure 8.18). Interestingly, the sway frequency is correlated with leaf cover (i.e., biomass suspended on a trunk, forming a spring-mass system) and hydration (because water uptake affects the mass of the foliage) [47]. This signal could be used to detect anomalous trees in the cityscape to proactively dispatch care or schedule replacement (healthy urban trees reduce air pollution, provide shade, and beautify spaces, among many benefits). Although we did not build a fully working version of this example application (lacking ground truth hydration data), we know that our system can readily detect the sway of trees using reflectors found on trunks (placed to help prevent collisions), and thus a model could certainly be built (as has been shown with physical accelerometer deployments [47].

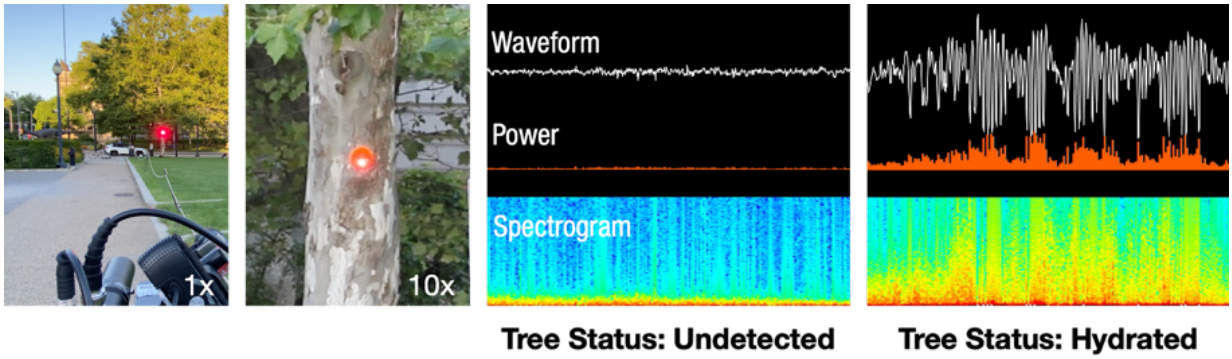


Figure 8.18: The water stress level of trees could be monitored through monitoring their sway frequency.

Structural Health

IMUs have long been used to measure the movement and vibrations of infrastructure, such as buildings and bridges, to monitor structural health [100]. A change in signal often indicates a change in physical condition, triggering an onsite inspection. With our system, instead of having to deploy expensive physical sensors, we can simply use retroreflectors already attached to the structure for other purposes. As an example, we recorded data from an overpass, and recorded the signals of passing cars (Figure 8.19). As with our tree health example, we did not have ground truth data for this bridge that permitted a structural health inference, but we believe our recorded signals demonstrate this could be possible.

8.6.3 Occlusion/Presence Sensing Examples

Our third and final signal category is occlusion and presence sensing. In short, when our laser beam is blocked or the retroreflector is moved, our system sees a dramatic drop in received light intensity. This binary signal has no spectral or intensity information, though in some cases, the rate at which the signal turns on and off can be used as a useful signal.

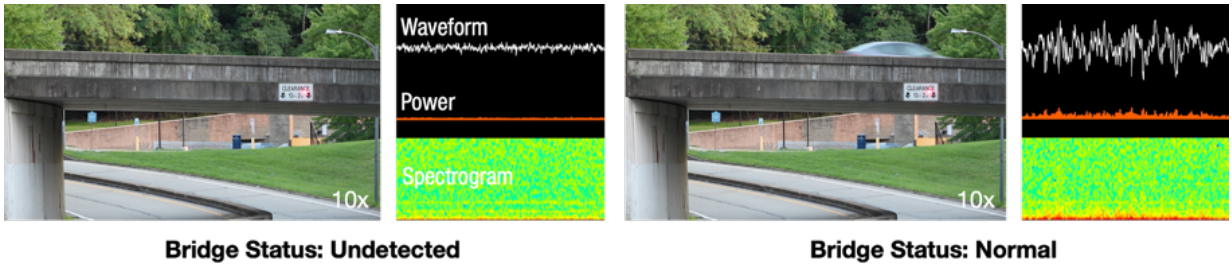


Figure 8.19: Our system could monitor bridge structural health by tracking the magnitude and frequency of its vibration as vehicles drive by.

Parking Spot Occupancy

Parking spots in cities are a limited resource, and many expensive sensing solutions are sold today to mitigate this issue. However, the high costs involved (often requiring installing a limited-life powered sensor in the pavement) has meant that publicly-owned, smart parking lots are comparatively rare today. With our system, road studs placed in the center of parking spots could be used as occlusion detectors (Figure 8.20). Given the sensing range of our unit, a whole parking lot could be sensed.

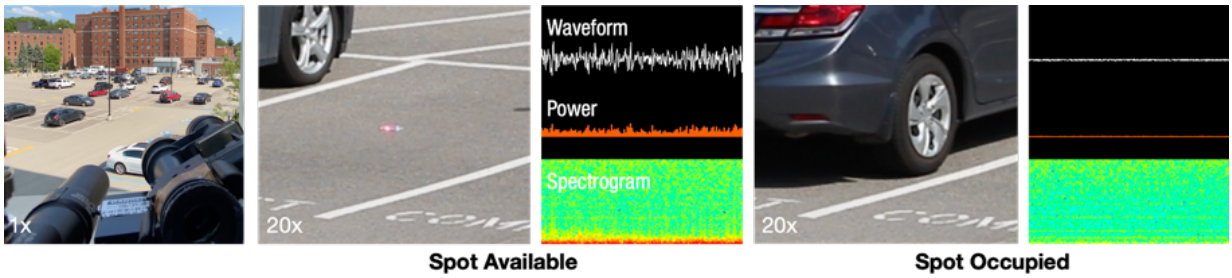


Figure 8.20: We detect parking spot occupancy based on occlusion detection of ground reflectors.

Hyper-Local Weather Monitoring

Technologies like doppler weather radar are excellent at estimating rainfall volume down to areas around 1 km^2 in size. For more localized measurements, physical weather stations with rainfall collectors must be deployed. With our system, we found that rainfall volume could be estimated using the brief signal interruptions caused by droplets passing through the laser path (Figure 8.21). As the distance to each retroreflector is known (see Depth map in Figure 8.5), the density of droplets can be estimated. This allows a single laser vibrometry base station to use its collection of identified retroreflectors to sample the local rainfall around it, potentially on a block-scale. While we did not collect any formal data, we also saw that fog and air pollution had an attenuating effect on signal strength, which our system could leverage to power a wide array of use cases.

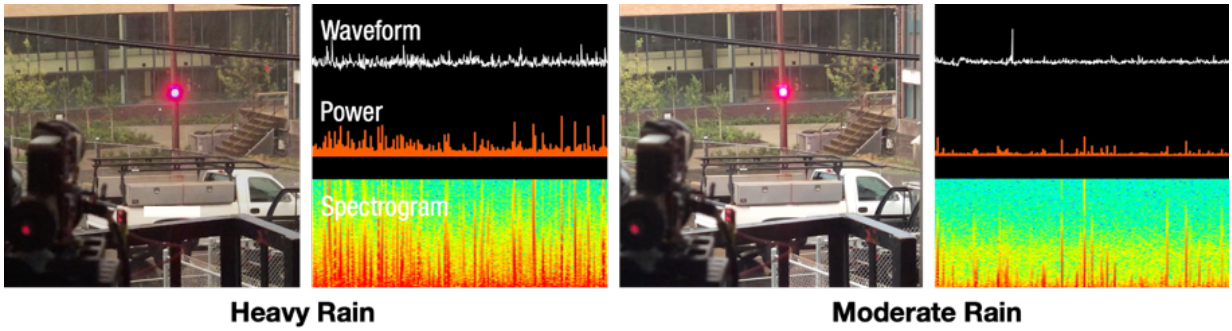


Figure 8.21: Rain intensity can be detected through tracking the power of non-DC frequency components.

Sidewalk / Road Access

Powered or manually-operated movable traffic bollards are used to regulate access to pedestrian and vehicular routes. As these are traffic control devices, they typically feature a retroreflective marker, which we can use to infer their states (Figure 8.22).

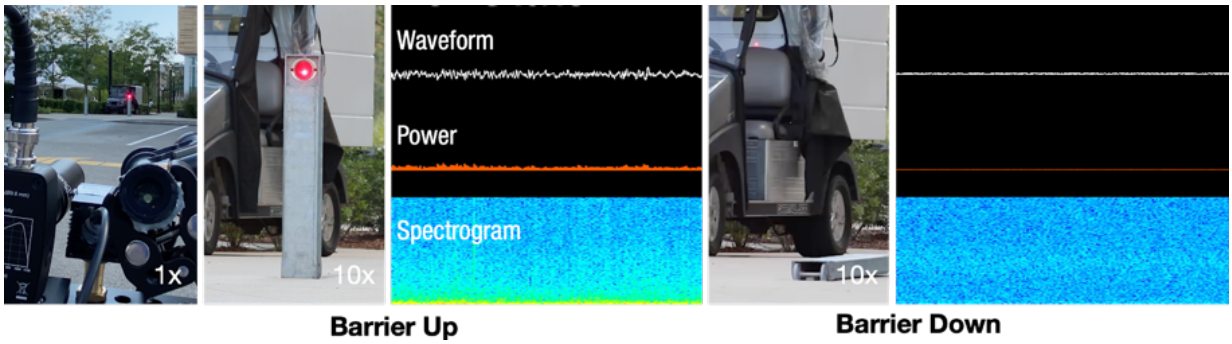


Figure 8.22: We detect the road barrier state through thresholding raw reflectance measurements.

Sanitation Collection

Finally, there are many objects that occupy the streetscape that are movable, such as refuse dumpsters (Figure 8.23). By knowing their rough parked location, our system can scan to find them, and then periodically check to see if they have been moved or taken away. Such data could be used to push alerts about refuse collection, for example, to building managers so they can optimize collection schedules.

8.7 Limitations

We summarize several limitations of our system which we summarize here. The most obvious one is that our system relies on the line of sight, which means if retroreflective signs and markers are not visible to our sensor, then we will not be able to get signals from it. However, we found

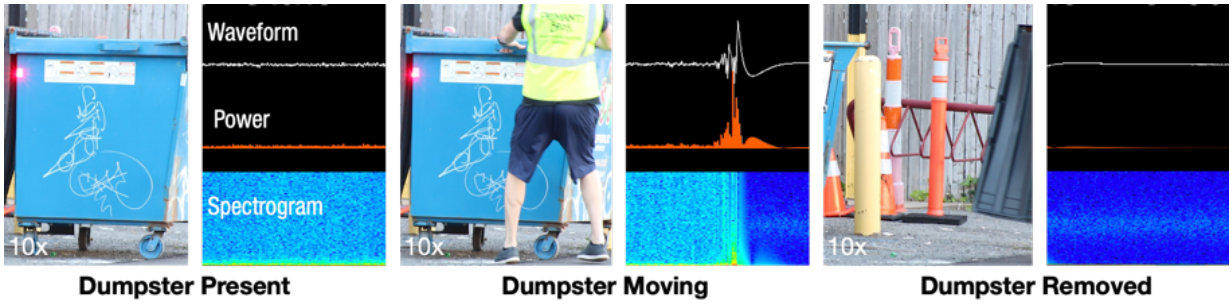


Figure 8.23: We detect the state of a dumpster through thresholding raw reflectance measurements.

that with the vast amount of densely distributed signs and markers throughout the city, as we have demonstrated in the survey, it would be more than sufficient with even just a fraction of them turned into sensors.

Second, the time-multiplex nature of our current sensing mechanism means that our system cannot sense transient events. Sensing transient events would require longer sensing duration or systems with multiple sensing beams. Our system is best suited for sensing continuous events (e.g., a tree swaying in the wind) or minute-scale changes to the environment (e.g., a car parking in a spot or a dumpster being emptied). We will investigate systems with multiple sensing beams to achieve transient activity detection in the future.

Finally, our system has privacy implications. We acknowledge that long-distance sensing can be inherently privacy-invasive when it is not carefully designed in real-world applications, even though this was not the focus of our system. Our take on this aspect is that the general vision of smart cities will inevitably raise privacy concerns – As smart cities knowing increasingly more about people who live in it, there will be more fronts where people's privacy could be compromised. Nonetheless, we hope that by demonstrating that reflectors can be turned into sensors, people and the city planners can be well informed when making decisions. We are motivated to explore and build systems that could encourage privacy-sensitive applications in the future.

8.8 Summary

We developed a long-range laser sensing system that turns retroreflective signs and markers into sensors for smart city applications. Our system is mostly based on laser vibrometry with which we can detect minute motion and vibration on reflectors, which often carry rich signals that can be used for block-scale weather and traffic sensing, infrastructure health and resource utilization monitoring, vehicle type detection, and beyond. Through a series of studies, we have characterized our system's performance across many factors, the results of which indicate that our system provides a feasible path to achieve practical city-scale activity sensing for future smart cities.

Chapter 9

Conclusion

9.1 Thesis Contributions

First, this thesis proposes a new direction for activity sensing. Instead of relying on densely distributed sensors, future ubiquitous applications could feature sparse sensor deployments with wide-area sensors. My observation is that existing activity sensing approaches are significantly bounded by three key challenges: cost, user maintenance, and sensing versatility. In response, this new direction allows users to achieve activity sensing at a much lower cost, and maintenance effort, that can thus be easily scaled for practical use. Additionally, through careful designs and engineering, wide-area sensors can be made to be general-purpose, being able to sense a wider range of activity than was previously possible. I believe all of these advantages are critical to making activity sensing applicable across society.

Second, this thesis summarizes two approaches that I have been practicing in order to create wide-area sensors. The first approach to achieve wide-area sensing is fabricating large sensors.The key to this approach's practicality is to create low-cost sensors that can cheaply scale with their sizes, which I achieved in Electrick [165] and Wall++ [167]. The second approach is transforming imperceivable signals from target activities into perceivable signals by low-cost and self-powered transformers [164, 168] or component-free retroreflectors [166, 169, 170]. I believe the methodology in this thesis is highly generalizable for future wide-area sensing research.

"If all you have is a hammer, everything looks like a nail" [25]. If one believes in this saying, then one will understand the importance of having a broad toolset. We want to have the right tool to solve the problem. A broad toolset is even more important to HCI researchers because the context of users and their environments could be highly dynamic, resulting in a very complex problem domain. In this sense, the third contribution of this thesis is that it broadens the toolset for activity sensing. Specifically, I developed fully-functional wide-area sensing systems, among which Vibrosight and Sozu are open-sourced to further facilitate future research that builds on top of these systems. Through real-world evaluations, these systems demonstrate high promises of sensing a wide range of user activities, which I believe will be critical in solving real-world HCI problems.

9.2 Future Directions

9.2.1 Encouraging Privacy-Sensitive Systems

First, wide-area sensors have privacy implications. Enabling sensors to gather information from increasingly longer distances inevitably raises questions about privacy. Moreover, seemingly harmless signals can be turned into invasive surveillance with sufficient computational power. What makes privacy more complicated is that people weigh privacy risks against perceived benefits [93, 142, 172]. Unlike mobile sensing where sensors can be personalized (e.g., deciding whether to turn on smartphone location services according to personal preferences), wide-area sensing involves many users (e.g., family members, locals) who might have very different privacy preferences. Nonetheless, I believe advancing sensing capabilities will help reveal threat models and raise users' awareness of these threats, which could be an important first step to improving privacy. For example, in the city-scale laser sensing project, we intended to use laser vibrometry to sense vibration from street signs in order to detect bus arrival. However, this sensing capability might accidentally reveal speech from people nearby. One solution is that people should avoid sensitive speech when they are near street signs; however, it is also possible to entirely eliminate this privacy risk by improving the engineering standard of street signs, so that speech cannot induce enough vibration on them for sensing, thereby shifting burdens for preserving privacy from individuals to systems. Solving this issue would require a broader discussion on ethics and security, which is beyond the scope of this thesis. However, I am committed to collaborating with stakeholders and researchers in this domain to fully investigate potential privacy implications in the future.

9.2.2 Taking Material Approaches to Enhance Sensing

I will further investigate the idea of taking material approaches to enhance sensing. Materials have played a pivotal role in transmitting, conducting, and receiving signals, which is why they are important to sensing. In fact, all systems in this thesis included materials with special properties to help with sensing – Wall++ uses conductive paints, Sozu and OptoSense use harvesters composed of materials that turns energy into electricity, and my proposed systems based on active sensing use retroreflectivce materials. Instead of instrumenting electronics everywhere, this thesis, in some sense, shifts the focus from electronics to materials that can be easily applied or found around users, resulting in systems that are easy to deploy and maintain. Beyond the materials that have been demonstrated in this thesis, there are many more to be explored. For example, materials that change their shapes, conductivity, transparency, or reflectivity in response to environmental factors such as humidity, sound, light, and temperature. Engineering materials optimized for wide-area sensing requires expertise in material science, which will be essential to future work in this research direction.

9.2.3 Extending Wide-Area Sensing into Broader Application Domains

Additionally, this thesis can be extended into broader application domains. The sensors created in this thesis are important building blocks to achieve higher-order smart environment applications

which have not yet been explored. Such applications span across digital health, accessibility, Industry 4.0, precision agriculture, and beyond. The same sensing system that detects user activities and environmental facets at home can be readily applied to detect activities and events in industrial and agricultural settings. Expanding wide-area sensing from humans to vehicles, machines, plants, and a wide range of environmental aspects requires further effort. This thesis chooses to demonstrate a wide spectrum of sensing modalities, over specific applications that require continuous development of a certain sensing modality. In other words, it has targeted breadth over depth. However, future work will spend more effort on specific applications, which will require long term deployments, from which we will draw many insights that are not possible to have with preliminary feasibility tests. This effort will make wide-area sensing even more practical for smart environments.

9.2.4 Combining Wide-Area Sensing with Mobile Platforms

Emerging mobile platforms such as autonomous cars and drones create an inviting opportunity for wide-area sensing. These platforms provide a viable path to the mobility of wide-area sensors so that they will no longer be constrained to fixed deployment locations, but instead are capable of moving around for improved coverage and optimal sensing perspectives. Aside from outdoor settings, we have seen an increasing number of indoor moving platforms, such as autonomous vacuum cleaners, home security drones, and robots for personalized companionship. All of these platforms can be turned into sensors using the wide-area sensing techniques proposed in this thesis, perhaps more competently than the approaches that require sensors to be placed on or close to the things they aim to sense. Returning to the illustration on Figure 1.2, mobile wide-area sensors make it possible for a many-to-many ratio between rooms and sensors. This is an exciting new front for smart environment sensing.

9.2.5 Exploring Novel Sensing Technologies

Finally, I am eager to continue exploring novel sensing technologies. Applications create the need for technology. In turn, as we deepen our understanding of sensing and signals, we will unlock previously impossible applications. This positive feedback loop between application and technology could be the powerhouse of this thesis's future work. Additionally, thanks to advances in electronics, sensors have become much cheaper and faster than ever. I believe more previously bespoke sensors in industry and medical applications, many of which are too expensive and bulky for mobile uses in everyday settings, will become feasible and unlock unique HCI applications. For example, Vibrosight is roughly in the order of 10 times lighter and more compact than commercial laser vibrometry devices, making it possible for integration into smart devices, such as lightbulbs and speakers. Would it even be possible to achieve another 10 times miniaturization, so vibrometry could be used to power mobile applications for health monitoring (e.g., on smartwatch), context-aware computing (e.g., for AR/VR headsets) and beyond. As we deepen our understanding of the world, many signals and phenomena currently impossible to use will become accessible for practical uses, which could fundamentally change how we interact with computers. I am extremely excited about all of these possibilities, and am looking forward to my future explorations.

Bibliography

[1] Smart speaker market booms in 2018, driven by Google, Alibaba and Xiaomi. 2018. URL https://www.canalys.com/static/press_release/2019/pr20190226smart-speaker-market-boomsCanalys.pdf. Accessed Nov. 16, 2020. 2

[2] Ubiquitous computing – chips with everything. 2019. URL http://dl.magazinedl.com/magazinedl/The%20Economist%20USA/2019/The%20Economist%20USA%20-%20September%2014,%202019(magazinedl.com).pdf. Accessed Nov. 16, 2020. 2

[3] An Airbnb I recently stayed at had hidden cameras in the bedrooms, my friends and I were recorded having sex, 2019. URL https://old.reddit.com/r/legaladvice/comments/cx5uzr/an_airbnb_i_recently_stayed_at_had_hidden_cameras/. Accessed Nov. 16, 2020. 3, 9

[4] Texas Instruments, product application note, 2019. URL https://www.ti.com/lit/wp/spyy002b/spyy002b.pdf. Accessed Nov. 16, 2020. 73

[5] Airbnb refunds guest who found indoor cameras during his family's stay, 2019. URL https://www.washingtonpost.com/technology/2019/01/17/airbnb-refunds-guest-who-found-indoor-cameras-during-his-familys-stay/. Accessed Nov. 16, 2020. 3, 9

[6] How to use the handwashing feature on the apple watch, 2019. URL https://appleinsider.com/articles/20/09/19/how-to-use-the-handwashing-feature-on-the-apple-watch. Accessed Nov. 16, 2020. xiii, 9

[7] A House With a Brain. 2019. URL https://www.wsj.com/articles/a-house-with-a-brain-11561048479. Accessed Nov. 16, 2020. 3

[8] What is the average butler salary by state. 2020. URL https://www.ziprecruiter.com/Salaries/What-Is-the-Average-Butler-Salary-by-State. Accessed Nov. 16, 2020. 2

[9] ANSI Z136.1 Safe Use of Lasers, n.a. URL https://www.lia.org/resources/laser-safety-information/laser-safety-standards/ansi-z136-standards/z136-1. Accessed Nov. 16, 2020. 87, 88

[10] FCC regulation on unlicensed low-power radio frequency transmitter, n.a. URL https://www.fcc.gov/media/radio/low-power-radio-general-information#UNLICENSED. Accessed Nov. 16, 2020. 38

[11] 49 CFR 393.11 - Lamps and reflective devices, n.a. URL https://www.govinfo.gov/app/details/CFR-2011-title49-vol5/CFR-2011-title49-vol5-sec393-11. Accessed Nov. 16, 2020. 85

[12] 2K Shops, product information, n.d. URL https://www.2kshops.com/siemens-qaa2061d/. Accessed Nov. 16, 2020. xiii, 7

[13] Enocean, gmbh, n.d. URL https://www.enocean.com/products. Accessed Nov. 16, 2020. 33

[14] Polytec, product information, n.d. URL https://www.polytec.com/us/vibrometry/products/single-point-vibrometers/ivs-500-industrial-vibration-sensor/. Accessed Nov. 16, 2020. 57, 84

[15] Ring, product information, n.d. URL https://shop.ring.com/pages/security-system. Accessed Nov. 16, 2020. 9

[16] STMicroelectronics, product information, n.d. URL https://www.st.com/en/evaluation-tools/steval-stlkt01v1.html. Accessed Nov. 16, 2020. xiii, 2, 8

[17] Texas Instruments, product information, n.d. URL https://www.ti.com/tool/TIDC-CC2650STK-SENSORTAG. Accessed Nov. 16, 2020. xiii, 2, 8

[18] World Sensing, product information, n.d. URL https://www.worldsensing.com/application-areas/traffic-flow-monitoring/. Accessed Nov. 16, 2020. 83

[19] Zensors, product information, n.d. URL https://www.zensors.com. Accessed Nov. 16, 2020. 2, 9, 83

[20] Array of Things, project information, n.d. URL https://arrayofthings.github.io. Accessed Nov. 16, 2020. 83

[21] Fitbit, product information, n.d. URL https://www.fitbit.com/global/us/home. Accessed Nov. 16, 2020. xiii, 9

[22] Google, product information, n.d. URL https://www.google.com/fit/. Accessed Nov. 16, 2020. xiii, 9

[23] Real-time web based household power usage monitor, n.d. URL https://www.instructables.com/Real-time-Web-Based-Household-Power-Usage-Monitor/. Accessed Nov. 16, 2020. xiii, 7

[24] Notion, product information, n.d. URL https://getnotion.com. Accessed Nov. 16, 2020. xiii, 2, 3, 8

[25] If all you have is a hammer, everything looks like a nail, n.d.. URL https://en.wiktionary.org/wiki/if_all_you_have_is_a_hammer,_everything_looks_like_a_nail. Accessed Nov. 16, 2020. 103

[26] Street furniture, n.d.. URL https://en.wikipedia.org/wiki/Street_furniture. Accessed Nov. 16, 2020. 86

[27] Zoro, product information, n.d. URL https://www.zoro.com/hubbell-wiring-device-kellems-occupancy-sensor-pir-1200-sq-ft-white-ws2000nw/i/G8492976/. Accessed Nov. 16, 2020. xiii, 7

[28] Gregory D Abowd. Beyond weiser: From ubiquitous to collective computing. *Computer*, 49(1):17–23, 2016. 4

[29] Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand. Capturing the human figure through a wall. *ACM Transactions on Graphics (TOG)*, 34(6):1–13, 2015. 9, 10, 73

[30] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 837–846, 2015. 10

[31] Achmad Affandi, Ghais El Zein, and Jacques Citerne. Investigation on frequency dependence of indoor radio propagation parameters. In *Gateway to 21st Century Communications Village. VTC 1999-Fall. IEEE VTS 50th Vehicular Technology Conference (Cat. No. 99CH36324)*, volume 4, pages 1988–1992. IEEE, 1999. 38

[32] Majd Alwan, Prabhu Jude Rajendran, Steve Kell, David Mack, Siddharth Dalal, Matt Wolfe, and Robin Felder. A smart and passive floor-vibration based fall detector for elderly. In *2006 2nd International Conference on Information & Communication Technologies*, volume 1, pages 1003–1007. IEEE, 2006. 10, 55

[33] Giuseppe Anastasi, Alessio Falchi, Andrea Passarella, Marco Conti, and Enrico Gregori. Performance measurements of motes sensor networks. In *Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 174–181, 2004. 2, 8

[34] Nivedita Arora and Gregory D Abowd. Zeusss: Zero energy ubiquitous sound sensing surface leveraging triboelectric nanogenerator and analog backscatter communication. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*, pages 81–83, 2018. 10, 34

[35] Nivedita Arora, Steven L Zhang, Fereshteh Shahmiri, Diego Osorio, Yi-Cheng Wang, Mohit Gupta, Zhengjun Wang, Thad Starner, Zhong Lin Wang, and Gregory D Abowd. Saturn: A thin and flexible self-powered microphone leveraging triboelectric nanogenerator. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2 (2):1–28, 2018. 10, 34

[36] Ernesto Arroyo, Leonardo Bonanni, and Ted Selker. Waterbot: exploring feedback and persuasive techniques at the sink. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 631–639, 2005. 7

[37] Tracy S Barger, Donald E Brown, and Majd Alwan. Health-status monitoring through analysis of behavioral patterns. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 35(1):22–27, 2004. 2, 7, 8

[38] Johan Barthélemy, Nicolas Verstaevel, Hugh Forehead, and Pascal Perez. Edge-computing video analytics for real-time traffic monitoring in a smart city. *Sensors*, 19(9):2048, 2019. 83

[39] Abdelkareem Bedri, Richard Li, Malcolm Haynes, Raj Prateek Kosaraju, Ishaan Grover, Temiloluwa Prioleau, Min Yan Beh, Mayank Goel, Thad Starner, and Gregory Abowd.

Earbit: using wearable sensors to detect eating episodes in unconstrained environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1 (3):1–20, 2017. 8

[40] Austin Lyons Bischoff and Darcy M Bullock. Sign retroreflectivity study. 2002. 85

[41] Michael Buettner, Richa Prasad, Matthai Philipose, and David Wetherall. Recognizing daily activities with rfid-based sensors. In *Proceedings of the 11th International Conference on Ubiquitous Computing*, pages 51–60, 2009. 10

[42] Tim Campbell, Eric Larson, Gabe Cohn, Jon Froehlich, Ramses Alcaide, and Shwetak N Patel. Wattr: a method for self-powered wireless sensing of water activity in the home. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, pages 169–172, 2010. 34, 55

[43] Jillian Carr and Jennifer L Doleac. The geography, incidence, and underreporting of gun violence: new evidence using shotspotter data. *Incidence, and Underreporting of Gun Violence: New Evidence Using Shotspotter Data (April 26, 2016)*, 2016. 9, 83

[44] P Castellini, M Martarelli, and EP Tomasini. Laser doppler vibrometry: Development of advanced solutions answering to technology's needs. *Mechanical Systems and Signal Processing*, 20(6):1265–1285, 2006. 85

[45] Jianfeng Chen, Alvin Harvey Kam, Jianmin Zhang, Ning Liu, and Louis Shue. Bathroom activity monitoring based on sound. In *International Conference on Pervasive Computing*, pages 47–61. Springer, 2005. 9

[46] Tanzeem Choudhury, Gaetano Borriello, Sunny Consolvo, Dirk Haehnel, Beverly Harrison, Bruce Hemingway, Jeffrey Hightower, Predrag Pedja, Karl Koscher, Anthony LaMarca, et al. The mobile sensing platform: An embedded activity recognition system. *IEEE Pervasive Computing*, 7(2):32–41, 2008. 8

[47] Dominick M Ciruzzi and Steven P Loheide. Monitoring tree sway as an indicator of water stress. *Geophysical Research Letters*, 46(21):12021–12029, 2019. 99

[48] Brian Clarkson, Nitin Sawhney, and Alex Pentland. Auditory context awareness via wearable computing. *Energy*, 400(600):20, 1998. 8

[49] Gabe Cohn, Sidhant Gupta, Jon Froehlich, Eric Larson, and Shwetak N Patel. Gassense: Appliance-level, single-point sensing of gas activity in the home. In *International Conference on Pervasive Computing*, pages 265–282. Springer, 2010. 9, 10, 55

[50] Sunny Consolvo, Beverly Harrison, Ian Smith, Mike Y Chen, Katherine Everitt, Jon Froehlich, and James A Landay. Conducting in situ evaluations for and with ubiquitous computing technologies. *International Journal of Human-Computer Interaction*, 22(1-2): 103–118, 2007. 1

[51] Sunny Consolvo, David W McDonald, Tammy Toscos, Mike Y Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, et al. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1797–1806, 2008. 8

[52] C Cristalli, N Paone, and RM Rodríguez. Mechanical fault detection of electric motors

by laser vibrometer and accelerometer measurements. *Mechanical Systems and Signal Processing*, 20(6):1350–1361, 2006. 84

[53] Samuel DeBruin, Branden Ghena, Ye-Sheng Kuo, and Prabal Dutta. Powerblade: A low-profile, true-power, plug-through energy meter. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pages 17–29, 2015. 33

[54] Armin Walter Doerry. Reflectors for sar performance testing. Technical report, Sandia National Laboratories, 2008. 74

[55] Charalampos Doukas and Ilias Maglogiannis. Advanced patient or elder fall detection based on movement and sound data. In *2008 Second International Conference on Pervasive Computing Technologies for Healthcare*, pages 103–107. IEEE, 2008. 9

[56] W Keith Edwards and Rebecca E Grinter. At home with ubiquitous computing: Seven challenges. In *International Conference on Ubiquitous Computing*, pages 256–272. Springer, 2001. 1

[57] Irfan A Essa. Ubiquitous sensing for smart and aware environments: technologies towards the building of an aware home. In *Position paper for the DARPA/NSF/NIST Workshop on Smart Environments*, 1999. 7

[58] K. P. Fishkin, M. Philipose, and A. Rea. Hands-on rfid: wireless wearables for detecting use of objects. In *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, pages 38–41, 2005. 8

[59] James Fogarty, Carolyn Au, and Scott E Hudson. Sensing from the basement: a feasibility study of unobtrusive and low-cost home activity recognition. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, pages 91–100, 2006. 9, 10, 55

[60] Jon E Froehlich, Eric Larson, Tim Campbell, Conor Haggerty, James Fogarty, and Shwetak N Patel. Hydrosense: infrastructure-mediated single-point sensing of whole-home water activity. In *Proceedings of the 11th International Conference on Ubiquitous Computing*, pages 235–244, 2009. 9, 10, 44, 55

[61] Jennifer Gabrys. Programming environments: environmentality and citizen sensing in the smart city. *Environment and Planning D: Society and Space*, 32(1):30–48, 2014. 83

[62] Chuhan Gao, Yilong Li, and Xinyu Zhang. Livetag: Sensing human-object interaction through passive chipless wifi tags. In *15th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 18)*, pages 533–546, 2018. 9, 10

[63] Guido Giuliani, Simone Bozzi-Pietra, and Silvano Donati. Self-mixing laser diode vibrometer. *Measurement Science and Technology*, 14(1):24, 2002. 85

[64] Albert Glinsky. *Theremin: ether music and espionage*. University of Illinois Press, 2000. 55

[65] Tobias Grosse-Puppendahl, Christian Holz, Gabe Cohn, Raphael Wimmer, Oskar Bechtold, Steve Hodges, Matthew S Reynolds, and Joshua R Smith. Finding common ground: A survey of capacitive sensing in human-computer interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3293–3315, 2017. 21

[66] Tobias Alexander Große-Puppendahl, Alexander Marinc, and Andreas Braun. Classification of user postures with capacitive proximity sensors in aal-environments. In *International Joint Conference on Ambient Intelligence*, pages 314–323. Springer, 2011. 15

[67] Manoj Gulati, Shobha Sundar Ram, and Amarjeet Singh. An in depth study into using emi signatures for appliance identification. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-efficient Buildings*, pages 70–79, 2014. 10

[68] Manoj Gulati, Farshid Salemi Parizi, Eric Whitmire, Sidhant Gupta, Shobha Sundar Ram, Amarjeet Singh, and Shwetak N Patel. Capharvester: A stick-on capacitive energy harvester using stray electric field from ac power lines. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–20, 2018. 34

[69] Sidhant Gupta, Matthew S Reynolds, and Shwetak N Patel. Electrisense: single-point sensing using emi for electrical event detection and classification in the home. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, pages 139–148, 2010. 10, 25, 44

[70] Gerhard P Hancke, Gerhard P Hancke Jr, et al. The role of advanced sensing in smart cities. *Sensors*, 13(1):393–425, 2013. 83

[71] Richard Harper. *Inside the smart home*. Springer Science & Business Media, 2006. 7

[72] George William Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992. 10, 25

[73] Xu He, Yunlong Zi, Hua Yu, Steven L Zhang, Jie Wang, Wenbo Ding, Haiyang Zou, Wei Zhang, Canhui Lu, and Zhong Lin Wang. An ultrathin paper-based self-powered system for portable electronics and wireless human-machine interaction. *Nano Energy*, 39:328–336, 2017. 34

[74] Roy Henderson and Karl Schulmeister. *Laser safety*. CRC Press, 2003. 87

[75] Vikram Iyer, Justin Chan, and Shyamnath Gollakota. 3d printing wireless connected objects. *ACM Transactions on Graphics (TOG)*, 36(6):1–13, 2017. 10

[76] Vikram Iyer, Elyas Bayati, Rajalakshmi Nandakumar, Arka Majumdar, and Shyamnath Gollakota. Charging a smartphone across a room using lasers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4):1–21, 2018. 56, 61

[77] Martin Johansmann, Georg Siegmund, and Mario Pineda. Targeting the limits of laser doppler vibrometry. *Proc. IDEMA*, pages 1–12, 2005. 56

[78] Haik Kalantarian and Majid Sarrafzadeh. Audio-based detection and evaluation of eating behavior using the smartwatch platform. *Computers in Biology and Medicine*, 65:1–9, 2015. 8

[79] Dharmendra Chandrashekar Kallur. *Human localization and activity recognition using distributed motion sensors*. PhD thesis, Oklahoma State University, 2014. 9

[80] Chitra R Karanam and Yasamin Mostofi. 3d through-wall imaging with unmanned aerial vehicles using wifi. In *2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 131–142. IEEE, 2017. 49

[81] Julie A Kientz, Shwetak N Patel, Brian Jones, ED Price, Elizabeth D Mynatt, and Gregory D Abowd. The georgia tech aware home. In *CHI'08 Extended Abstracts on Human Factors in Computing Systems*, pages 3675–3680. 2008. 7

[82] David Kim, Otmar Hilliges, Shahram Izadi, Alex D Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. Digits: freehand 3d interactions anywhere using a wrist-worn gloveless sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, pages 167–176, 2012. 56

[83] Younghun Kim, Thomas Schmid, Zainul M Charbiwala, and Mani B Srivastava. Viridiscope: design and implementation of a fine grained power monitoring system for homes. In *Proceedings of the 11th International Conference on Ubiquitous Computing*, pages 245–254, 2009. 15

[84] Neil Klingensmith, Joseph Bomber, and Suman Banerjee. Hot, cold and in between: enabling fine-grained environmental control in homes for efficiency and comfort. In *Proceedings of the 5th International Conference on Future Energy Systems*, pages 123–132, 2014. 7

[85] Eugene F Knott, John F Schaeffer, and Michael T Tulley. *Radar cross section*. SciTech Publishing, 2004. 74

[86] Stacey Kuznetsov and Eric Paulos. Upstream: motivating water conservation with low-cost water flow sensing and persuasive displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1851–1860, 2010. xiii, 7

[87] Gierad Laput and Chris Harrison. Sensing fine-grained hand activity with smartwatches. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019. 8

[88] Gierad Laput, Walter S Lasecki, Jason Wiese, Robert Xiao, Jeffrey P Bigham, and Chris Harrison. Zensors: Adaptive, rapidly deployable, human-intelligent sensor feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1935–1944, 2015. 9

[89] Gierad Laput, Chouchang Yang, Robert Xiao, Alanson Sample, and Chris Harrison. Em-sense: Touch recognition of uninstrumented, electrical and electromechanical objects. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pages 157–166, 2015. 8, 25

[90] Gierad Laput, Robert Xiao, and Chris Harrison. Viband: High-fidelity bio-acoustic sensing using commodity smartwatch accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 321–333, 2016. 8, 55, 84

[91] Gierad Laput, Yang Zhang, and Chris Harrison. Synthetic sensors: Towards general-purpose sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3986–3999, 2017. 44, 84, 96

[92] Gierad Laput, Karan Ahuja, Mayank Goel, and Chris Harrison. Ubicoustics: Plug-and-play acoustic activity recognition. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, pages 213–224, 2018. 9

[93] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. Alexa, are you listening? privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–31, 2018. 3, 104

[94] Fabio Leccese, Marco Cagnetti, and Daniele Trinca. A smart city application: A fully controlled street lighting isle based on raspberry-pi card, a zigbee sensor network and wimax. *Sensors*, 14(12):24408–24424, 2014. 83

[95] WH Leong, WJ Staszewski, BC Lee, and F Scarpa. Structural health monitoring using scanning laser vibrometry: Iii. lamb waves for fatigue crack detection. *Smart Materials and Structures*, 14(6):1387, 2005. 84

[96] Glynn Lewis, Scott Shaw, Michael Crowe, Clay Cranford, Kevin Torvik, Peter Scharf, and Bob Stellingworth. Urban gunshot and sniper location: technologies and demonstration results. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Defense and Law Enforcement*, volume 4708, pages 315–323. International Society for Optics and Photonics, 2002. 9

[97] Hanchuan Li, Can Ye, and Alanson P Sample. Idsense: A human object interaction detection system based on passive uhf rfid. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2555–2564, 2015. 9, 10

[98] Xinyu Li, Yuan He, and Xiaojun Jing. A survey of deep learning-based human activity recognition in radar. *Remote Sensing*, 11(9):1068, 2019. 10

[99] Yang Li, Jason I Hong, and James A Landay. Design challenges and principles for wizard of oz testing of location-enhanced applications. *IEEE Pervasive Computing*, 6(2):70–75, 2007. 1

[100] Jerome P Lynch and Kenneth J Loh. A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock and Vibration Digest*, 38(2):91–130, 2006. 99

[101] Xianghong Ma, Lawrence Bergman, and Alexander Vakakis. Identification of bolted joints through laser vibrometry. *Journal of Sound and Vibration*, 246(3):441–460, 2001. 84

[102] Takuya Maekawa, Yutaka Yanagisawa, Yasue Kishino, Katsuhiko Ishiguro, Koji Kamei, Yasushi Sakurai, and Takeshi Okadome. Object-based activity recognition with heterogeneous sensors on wrist. In *International Conference on Pervasive Computing*, pages 246–264. Springer, 2010. 8

[103] Takuya Maekawa, Yasue Kishino, Yasushi Sakurai, and Takayuki Suyama. Recognizing the use of portable electrical devices with hand-worn magnetic sensors. In *International Conference on Pervasive Computing*, pages 276–293. Springer, 2011. 8

[104] Peter Martin and Steve Rothberg. Laser vibrometry and the secret life of speckle patterns. In *Eighth International Conference on Vibration Measurements by Laser Techniques: Advances and Applications*, volume 7098, page 709812. International Society for Optics and Photonics, 2008. 56

[105] Shota Mashiyama, Jihoon Hong, and Tomoaki Ohtsuki. Activity recognition using low resolution infrared array sensor. In *2015 IEEE International Conference on Communications*

(ICC), pages 495–500. IEEE, 2015. 9

[106] David Minnen, Thad Starner, Jamie A Ward, Paul Lukowicz, and Gerhard Troster. Recognizing and discovering human actions from on-body sensor data. In *2005 IEEE International Conference on Multimedia and Expo*, pages 1545–1548. IEEE, 2005. 8

[107] Saraju P Mohanty, Uma Choppali, and Elias Kougianos. Everything you wanted to know about smart cities: The internet of things is the backbone. *IEEE Consumer Electronics Magazine*, 5(3):60–70, 2016. 83

[108] Michael C Mozer. The neural network house: An environment hat adapts to its inhabitants. In *Proc. AAAI Spring Symp. Intelligent Environments*, volume 58, 1998. 7

[109] Rajalakshmi Nandakumar, Vikram Iyer, and Shyamnath Gollakota. 3d localization for sub-centimeter sized devices. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, pages 108–119, 2018. 10

[110] Hani H Nassif, Mayrai Gindy, and Joe Davis. Comparison of laser doppler vibrometer with contact sensors for monitoring bridge deflection and vibration. *Ndt & E International*, 38 (3):213–218, 2005. 56

[111] Mitsuhiro Ogawa and Tastuo Togawa. Monitoring daily activities and behaviors at home by using brief sensors. In *1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology. Proceedings (Cat. No. 00EX451)*, pages 611–614. IEEE, 2000. 7

[112] Mitsuhiro Ogawa, S Ochiai, Ken Shoji, Minori Nishihara, and Tatsuo Togawa. An attempt of monitoring daily activities at home. In *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Cat. No. 00CH37143)*, volume 1, pages 786–788. IEEE, 2000. 7

[113] Mitsushiro Ogawa, Ryoji Suzuki, Sakuko Otake, Takeshi Izutsu, Tsutomu Iwaya, and Tatsuo Togawa. Long term remote behavioral monitoring of elderly by using sensors installed in ordinary houses. In *2nd Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology. Proceedings (Cat. No. 02EX578)*, pages 322–325. IEEE, 2002. 7

[114] Katsunori Ohnishi, Atsushi Kanehira, Asako Kanezaki, and Tatsuya Harada. Recognizing activities of daily living with a wrist-mounted camera. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3103–3111, 2016. 8

[115] Alex Olwal, Andrew Bardagjy, Jan Zizka, and Ramesh Raskar. Speckleeye: gestural interaction for embedded electronics in ubiquitous computing. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 2237–2242. 2012. 56

[116] Sophocles J Orfanidis. Electromagnetic waves and antennas. 2002. 15

[117] Antti Oulasvirta, Aurora Pihlajamaa, Jukka Perkiö, Debarshi Ray, Taneli Vähäkangas, Tero Hasu, Niklas Vainio, and Petri Myllymäki. Long-term effects of ubiquitous surveillance in the home. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 41–50, 2012. 3

[118] Shijia Pan, Amelie Bonde, Jie Jing, Lin Zhang, Pei Zhang, and Hae Young Noh. Boes:

building occupancy estimation system using sparse ambient vibration monitoring. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2014*, volume 9061, page 90611O. International Society for Optics and Photonics, 2014. 10, 55

[119] Shwetak N Patel, Thomas Robertson, Julie A Kientz, Matthew S Reynolds, and Gregory D Abowd. At the flick of a switch: Detecting and classifying unique electrical events on the residential power line (nominated for the best paper award). In *International Conference on Ubiquitous Computing*, pages 271–288. Springer, 2007. 9, 10

[120] Shwetak N Patel, Matthew S Reynolds, and Gregory D Abowd. Detecting human movement by differential air pressure sensing in hvac system ductwork: An exploration in infrastructure mediated sensing. In *International Conference on Pervasive Computing*, pages 1–18. Springer, 2008. 10

[121] Vesa Peltonen, Juha Tuomi, Anssi Klapuri, Jyri Huopaniemi, and Timo Sorsa. Computational auditory scene recognition. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages II–1941. IEEE, 2002. 9

[122] Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos. Sensing as a service model for smart cities supported by internet of things. *Transactions on Emerging Telecommunications Technologies*, 25(1):81–93, 2014. 83

[123] Matthai Philipose, Joshua R Smith, Bing Jiang, Alexander Mamishev, Sumit Roy, and Kishore Sundara-Rajan. Battery-free wireless identification and sensing. *IEEE Pervasive Computing*, 4(1):37–45, 2005. 8, 10

[124] Ronald Poppe. A survey on vision-based human action recognition. *Image and Vision Computing*, 28(6):976–990, 2010. 9

[125] Niraniini Rajagopal, Suman Giri, Mario Berges, and Anthony Rowe. A magnetic field-based appliance metering system. In *2013 ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, pages 229–238. IEEE, 2013. 15

[126] Jun Rekimoto. Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 113–120, 2002. 15

[127] Helge Rhodin, Christian Richardt, Dan Casas, Eldar Insafutdinov, Mohammad Shafiei, Hans-Peter Seidel, Bernt Schiele, and Christian Theobalt. Egocap: egocentric marker-less motion capture with two fisheye cameras. *ACM Transactions on Graphics (TOG)*, 35(6): 1–11, 2016. 8

[128] Marcus Rohrbach, Sikandar Amin, Mykhaylo Andriluka, and Bernt Schiele. A database for fine grained activity detection of cooking activities. In *2012 IEEE Conference on Computer Vision and Pattern recognition*, pages 1194–1201. IEEE, 2012. 9

[129] Anthony Rowe, Mario Berges, and Raj Rajkumar. Contactless sensing of appliance state transitions through variations in electromagnetic fields. In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, pages 19–24, 2010. 15

[130] Alanson P Sample, Daniel J Yeager, Pauline S Powledge, Alexander V Mamishev, and Joshua R Smith. Design of an rfid-based battery-free programmable sensing platform. *IEEE Transactions on Instrumentation and Measurement*, 57(11):2608–2615, 2008. 10

[131] Munehiko Sato, Shigeo Yoshida, Alex Olwal, Boxin Shi, Atsushi Hiyama, Tomohiro Tanikawa, Michitaka Hirose, and Ramesh Raskar. Spectrans: Versatile material classification for interaction with textureless, specular and transparent surfaces. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2191–2200, 2015. 56

[132] Edward Sazonov, Haodong Li, Darrell Curry, and Pragasen Pillay. Self-powered sensors for monitoring of highway bridges. *IEEE Sensors Journal*, 9(11):1422–1429, 2009. 34

[133] Javier Schloemann, VVN Sriram Malladi, Americo G Woolard, Joseph M Hamilton, R Michael Buehrer, and Pablo A Tarazaga. Vibration event localization in an instrumented building. In *Experimental Techniques, Rotating Machinery, and Acoustics, Volume 8*, pages 265–271. Springer, 2015. 55

[134] James Scott, AJ Bernheim Brush, John Krumm, Brian Meyers, Michael Hazas, Stephen Hodges, and Nicolas Villar. Preheat: controlling home heating using occupancy prediction. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, pages 281–290, 2011. 7

[135] Oliver Shih, Patrick Lazik, and Anthony Rowe. Aures: A wide-band ultrasonic occupancy sensing platform. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*, pages 157–166, 2016. 9

[136] Muhammad Shoaib, Stephan Bosch, Ozlem Durmaz Incel, Hans Scholten, and Paul JM Havinga. Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors*, 16(4):426, 2016. 8

[137] Robert L Showen, Robert B Calhoun, and Jason W Dunham. Acoustic location of gunshots using combined angle of arrival and time of arrival measurements, January 6 2009. US Patent 7,474,589. 9, 83

[138] Jae Mun Sim, Yonnim Lee, and Ohbyung Kwon. Acoustic sensor based recognition of human activity in everyday life for smart home services. *International Journal of Distributed Sensor Networks*, 11(9):679123, 2015. 9

[139] Brandon M Smith, Pratham Desai, Vishal Agarwal, and Mohit Gupta. Colux: Multi-object 3d micro-motion analysis using speckle imaging. *ACM Transactions on Graphics (TOG)*, 36(4):1–12, 2017. 56

[140] Joshua Smith, Tom White, Christopher Dodge, Joseph Paradiso, Neil Gershenfeld, and David Allport. Electric field sensing for graphical interfaces. *IEEE Computer Graphics and Applications*, 18(3):54–60, 1998. 15

[141] Joshua R Smith, Kenneth P Fishkin, Bing Jiang, Alexander Mamishev, Matthai Philipose, Adam D Rea, Sumit Roy, and Kishore Sundara-Rajan. Rfid-based techniques for human-activity detection. *Communications of the ACM*, 48(9):39–44, 2005. 10

[142] Yunpeng Song, Yun Huang, Zhongmin Cai, and Jason I Hong. I'm all eyes and ears:

Exploring effective locators for privacy awareness in iot scenarios. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020. 104

[143] Jaeyong Sung, Colin Ponce, Bart Selman, and Ashutosh Saxena. Unstructured human activity detection from rgbd images. In *2012 IEEE International Conference on Robotics and Automation*, pages 842–849. IEEE, 2012. 9

[144] Yoshiki Takeoka, Takashi Miyaki, and Jun Rekimoto. Z-touch: an infrastructure for 3d gesture interaction in the proximity of tabletop surfaces. In *ACM International Conference on Interactive Tabletops and Surfaces*, pages 91–94, 2010. 56

[145] Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. Activity recognition in the home using simple and ubiquitous sensors. In *International Conference on Pervasive Computing*, pages 158–175. Springer, 2004. 2, 8

[146] Edison Thomaz, Irfan Essa, and Gregory D Abowd. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1029–1040, 2015. 8

[147] Yonglong Tian, Guang-He Lee, Hao He, Chen-Yu Hsu, and Dina Katabi. Rf-based fall monitoring using convolutional neural networks. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–24, 2018. 10

[148] B Uğur Töreyin, Yiğithan Dedeoğlu, and A Enis Çetin. Hmm based falling person detection using both audio and video. In *International Workshop on Human-Computer Interaction*, pages 211–220. Springer, 2005. 9

[149] Baptist Vandersmissen, Nicolas Knudde, Azarakhsh Jalalvand, Ivo Couckuyt, Andre Bourdoux, Wesley De Neve, and Tom Dhaene. Indoor person identification using a low-power fmcw radar. *IEEE Transactions on Geoscience and Remote Sensing*, 56(7): 3941–3952, 2018. 10

[150] Cati Vaucelle, Hiroshi Ishii, and Joseph A Paradiso. Electromagnetic field detector bracelet. *on Ubiquitous Computing*, page 109, 2008. 25

[151] AA Veber, A Lyashedko, E Sholokhov, A Trikshev, A Kurkov, Y Pyrkov, AE Veber, V Seregin, and V Tsvetkov. Laser vibrometry based on analysis of the speckle pattern from a remote object. *Applied Physics-Section B-Lasers and Optics*, 105(3):613, 2011. 56

[152] Edward J Wang, Tien-Jui Lee, Alex Mariakakis, Mayank Goel, Sidhant Gupta, and Shwetak N Patel. Magnifisense: Inferring device interaction using wrist-worn passive magneto-inductive sensors. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 15–26, 2015. 8

[153] Yong Wang, Wen Wang, Mu Zhou, Aihu Ren, and Zengshan Tian. Remote monitoring of human vital signs based on 77-ghz mm-wave fmcw radar. *Sensors*, 20(10):2999, 2020. 10

[154] Brett Warneke, Matt Last, Brian Liebowitz, and Kristofer SJ Pister. Smart dust: Communicating with a cubic-millimeter computer. *Computer*, 34(1):44–51, 2001. 2, 3, 8

[155] Mark Weiser. The computer for the 21st century. *ACM SIGMOBILE Mobile Computing and Communications Review*, 3(3):3–11, 1999. 1

[156] Christoph Will, Prachi Vaishnav, Abhiram Chakraborty, and Avik Santra. Human target detection, tracking, and classification using 24-ghz fmcw radar. *IEEE Sensors Journal*, 19 (17):7283–7299, 2019. 10

[157] Jason Wu, Chris Harrison, Jeffrey P Bigham, and Gierad Laput. Automated class discovery and one-shot interactions for acoustic activity recognition. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020. 9

[158] Robert Xiao, Gierad Laput, Yang Zhang, and Chris Harrison. Deus em machina: on-touch contextual functionality for smart iot appliances. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 4000–4008, 2017. 15

[159] Chouchang Yang and Alanson P Sample. Em-id: Tag-less identification of electrical devices via electromagnetic emissions. In *2016 IEEE International Conference on RFID (RFID)*, pages 1–8. IEEE, 2016. 15

[160] Koji Yatani and Khai N Truong. Bodyscope: a wearable acoustic sensor for activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 341–350, 2012. 8

[161] Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. Radarcat: Radar categorization for input & interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 833–841, 2016. 10

[162] Zeev Zalevsky, Yevgeny Beiderman, Israel Margalit, Shimshon Gingold, Mina Teicher, Vicente Mico, and Javier Garcia. Simultaneous remote extraction of multiple speech sources and heart beats from secondary speckles pattern. *Optics Express*, 17(24):21566–21580, 2009. 56

[163] Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, and Michele Zorzi. Internet of things for smart cities. *IEEE Internet of Things Journal*, 1(1):22–32, 2014. 83

[164] Dingtian Zhang, Jung Wook Park, Yang Zhang, Yuhui Zhao, Yiyang Wang, Yunzhi Li, Tanvi Bhagwat, Wen-Fang Chou, Xiaojia Jia, Bernard Kippelen, et al. Optosense: Towards ubiquitous self-powered ambient light sensing surfaces. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(3):1–27, 2020. 12, 103

[165] Yang Zhang, Gierad Laput, and Chris Harrison. Electrick: Low-cost touch sensing using electric field tomography. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2017. 11, 103

[166] Yang Zhang, Gierad Laput, and Chris Harrison. Vibrosight: Long-range vibrometry for smart environment sensing. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, pages 225–236, 2018. 12, 85, 103

[167] Yang Zhang, Chouchang Yang, Scott E Hudson, Chris Harrison, and Alanson Sample. Wall++ room-scale interactive and context-aware sensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2018. 11, 103

[168] Yang Zhang, Yasha Iravantchi, Haojian Jin, Swarun Kumar, and Chris Harrison. Sozu: Self-powered radio tags for building-scale activity sensing. In *Proceedings of the 32nd*

Annual ACM Symposium on User Interface Software and Technology, pages 973–985, 2019. 12, 103

[169] Yang Zhang, Cathy Fang, and Chris Harrison. Smart room sensing with fmcw radar. 2021. 13, 103

[170] Yang Zhang, Jesse T. Gonzalez, Sven Mayer, and Chris Harrison. City-scale sensing using existing retro-reflective signs and markers. 2021. 13, 103

[171] Chen Zhao, Sam Yisrael, Joshua R Smith, and Shwetak N Patel. Powering wireless sensor nodes with ambient temperature changes. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 383–387, 2014. 34

[172] Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. User perceptions of smart home iot privacy. *Proceedings of the ACM on Human-Computer Interaction*, 2 (CSCW):1–20, 2018. 3, 104

[173] Zhongna Zhou, Xi Chen, Yu-Chia Chung, Zhihai He, Tony X Han, and James M Keller. Activity analysis, summarization, and visualization for indoor human activity monitoring. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(11):1489–1498, 2008. 9

[174] Jan Zizka, Alex Olwal, and Ramesh Raskar. Specklesense: fast, precise, low-cost and compact motion sensing using laser speckle. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pages 489–498, 2011. 56