VibSense: Sensing Touches on Ubiquitous Surfaces through Vibration

Jian Liu[†], Yingying Chen[†], Marco Gruteser[§], Yan Wang[*]
[†]Stevens Institute of Technology, Hoboken, NJ 07030, USA
[§]Rutgers University, North Brunswick, NJ 08902, USA
[*]SUNY at Binghamton, Binghamton, NY 13902, USA
Email: [†]{jliu28, yingying.chen}@stevens.edu, [§]gruteser@winlab.rutgers.edu, [*]yanwang@binghamton.edu

***Abstract*—*VibSense* pushes the limits of vibration-based sensing to determine the location of a touch on extended surface areas as well as identify the object touching the surface leveraging a single sensor. Unlike capacitive sensing, it does not require conductive materials and compared to audio sensing it is more robust to acoustic noise. It supports a broad array of applications through either passive or active sensing using only a single sensor. In VibSense's passive sensing, the received vibration signals are determined by the location of the touch impact. This allows location discrimination of touches precise enough to enable emerging applications such as virtual keyboards on ubiquitous surfaces for mobile devices. Moreover, in the active mode, the received vibration signals carry richer information of the touching object's characteristics (e.g., weight, size, location and material). This further enables VibSense to match the signals to the trained profiles and allows it to differentiate personal objects in contact with any surface. VibSense is evaluated extensively in the use cases of localizing touches (i.e., virtual keyboards), object localization and identification. Our experimental results demonstrate that VibSense can achieve high accuracy, over 95%, in all these use cases.**

I. Introduction

As the form factor of our mobile and wearable devices shrinks, there exists an increasing need to support interaction beyond the confines of the device itself. Particularly on wearable devices, small touchscreens and interfaces can render complex input cumbersome. One approach to address this challenge is to support convenient interaction through sensing approaches that capture input from other surfaces, without directly touching the device. Such input usually comes in the form of touches, but we consider a broad interpretation that goes beyond a human touch and includes objects touching these surfaces.

Existing Solutions. Recently, several research teams [1–4] have developed gesture and activity recognition techniques that rely solely on measurable changes of the radio-frequency environment. These radio-based systems could be easily affected by surrounding changes that affect signal propagation, such as different furniture placement or people walking by. Another direction for extending interactions is using acoustic signals. This technique has been used to track phone movements [5], to tag and remember a phone's indoor locations [6], and recognize keystrokes on a nearby paper keyboard [7]. The accuracy of acoustic user interaction declines sharply in noisy environments. Additionally, several researches [8, 9] utilize visible light to locate a user's finger or reconstruct 3D human postures, respectively. However, visible light based interaction requires line-of-sight and is susceptible to interference from light sources. Capacitive and resistive touch sensing can also be implemented on external surface or devices [10, 11], but these approaches require electrically conductive surfaces and cannot be applied to all objects of daily life. More related are two recent studies: Toffee [12] uses acoustic time-of-arrival correlation to determine the direction of touches on a surface with respect to a device relying on multiple piezoelectric sensors. Touch & Activate[13] actively generates acoustic signals and records the sound patterns to identify how a user touches a small object with the vibration speaker and piezo-electric microphone directly attached to the object. These early studies are limited to devices with four well-separated sensors or support limited sensing distances.

Generalized Vibration-based Sensing over Extended Surfaces through a Single Sensor. In the quest for a touch sensing technique that is robust to environmental noise and can operate on surfaces constructed from a broad range of materials, we explore a different approach by pushing the limits of sensing physical vibrations. The impact of a touch on a surface such as a table or door causes a shockwave to be transmitted through the material that can be *passively* detected with accelerometers or more sensitive piezo vibration sensors. Moreover, when a vibrator (such as those built inside the mobile devices for unobtrusive notifications) *actively* excites a surface resulting in the alteration of the shockwave propagation, the presence of the object in contact with the surface can thus be sensed. VibSense supports generalized vibration sensing based on a low cost single sensor prototype that can receive vibration signals in both passive and active sensing scenarios. It can be attached to non-conductive surfaces such as a table or a door and sense touching objects or users. By relying on vibrating signals, the system is less susceptible to environmental interferences from acoustic or radio-frequency noise.

Consequently, we push the limits of vibration-based sensing on ubiquitous surfaces through VibSense along multiple dimensions. First, it provides an extended sensing area to demonstrate the power of both passive and active vibration sensing. Second, the system can passively localize the vibration source on a surface and enable localizing touches in emerging applications, such as paper keyboards. Third, it can also actively differentiate and localize objects when placing on a surface to support personal object (e.g., smartphones or keys) localization and identification. The main contributions of this work are summarized as follows:

- Pushing the limits of vibration-based sensing as a powerful touch/ object sensing alternative that does not require conductive materials and is robust to acoustic noise.
- Extending passive vibration sensing to allow distinguish touches on any surface using a single receiver that can precisely work for an imaginary/paper keyboard.
- Exploring the capabilities of active vibration-based sensing in applications such as differentiating objects placed on a surface, as well as locating these objects.
- Prototyping a passive/active vibration transceiver by using a low-cost piezo-electric sensor as well as touch identification and localization algorithms that rely on power spectral density profiles.
- Demonstrating experimentally that VibSense can achieve high accuracy of localizing touches, and personal object localization and identification in both passive as well as active vibration sensing scenarios under real environments.

II. Vibration Effect

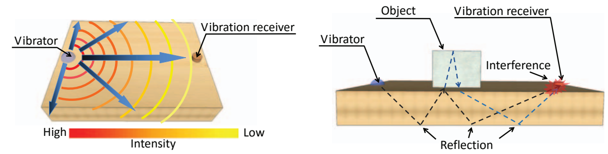
A. Propagation Characteristics

Vibration Signal Attenuation. When a vibration signal travels through a medium, the energy of the signal spreads out in an omnidirectional way and diminishes with its traveled distance due to the wave attenuation caused by the medium. Figure 1(a) illustrates the diminishing energy of the vibration signal during its propagation. The amplitude of the signal after attenuation can be modeled as the following [14]:

$$A(d) = A_0 e^{-\alpha \times d}, \tag{1}$$

where A_0 is the initial amplitude, d is the propagation distance from the vibration source, and α is the attenuation coefficient. The attenuation coefficient quantifies the intensity of the signal attenuation resulted from a particular medium. The value of attenuation coefficient varies in accordance with the type of materials and the frequency of vibration signals. For example, the attenuation coefficients of some typical materials, such as wood, would be around 0.11 and 0.07 under the different frequencies of 250 Hz and 4 kHz, respectively [15].

Equation 1 shows that the amplitude of the vibration signal is governed by the distance of propagation and attenuation coefficient. Under a fixed propagation distance, vibration signals with different frequencies will experience different attenuation. Whereas signals with the same frequency will also experience different attenuation when traversing through different materials and paths. The propagation media of vibration could consist of different materials. For instance, when putting a vibrator on a desk, the vibration signal will traverse through the desk and also the cup placed on it. Such an attenuation diversity reflects the details of the type of material and the path the vibration signal goes through, resulting in rich sensing information.



(a) Signal attenuation. (b) Signal interference.

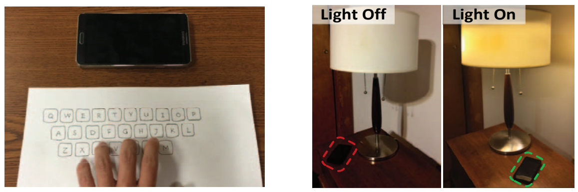
Fig. 1. Illustration of the propagation characteristics of vibration signals.

Vibration Multi-path Interference. Interference happens when two vibration signals superpose to form one resultant signal. The amplitude of the resultant signal may be either greater or smaller than those of the two participating ones depending on the phase difference between the two original signals. When a vibration signal hits the boundary of two propagation media (e.g., the contacting area when an object is placed on a desk) as illustrated in Figure 1(b), reflection and diffraction happen. Part of the wave is reflected back into the same medium and generates a new wave of the same frequency transmitting along a different propagation path, and part of the wave traverses into the different medium becoming a new wave with a different frequency. These new waves coming from multiple paths will meet and create various interference effects among each other. Such diverse multi-path interference effects could be captured by a vibration receiver and used to discriminate different propagation scenarios such as placing a cup on desk. The coexistence of both attenuation and multi-path interference when vibration signals traverse through different materials allows fine-grained discrimination to support various sensing-based applications.

B. Potential Applications

Vibration effects caused by diverse signal attenuation and interference, either intentional or non-intentional, have high potential to discriminate the propagation conditions in a fine-grained manner. Such effects can be utilized to support a broad array of application domains.

1) Localizing Touches: Instead of interacting with mobile devices using the touch screens with restricted sizes, there have been active research on developing ubiquitous human-computer interaction techniques to mitigate this constraint. Existing approaches either use acoustic signals [7, 12] or laser projections to construct virtual keyboards. Compared to acoustic-based approaches, vibration-based sensing is more robust to various environmental sounds. And the approach only requires a single low-cost vibration receiver, making it scalable comparing to the laser projection based solution. When users touch a surface, each touch generates vibration signals with different frequencies. The vibration signals spread out in the medium and experience various attenuation and interference resulted from different surface materials and multiple paths dominated by the location of the touch on the surface. The diverse vibration signals could be exploited to discriminate the fine-grained location of each touch on the surface via fingerprint-based approaches. In this work,



(a) Recognizing keystrokes on a paper keyboard. (b) Localizing objects.

Fig. 2. Potential applications of VibSense: (a) is passive sensing, (b) is active sensing.

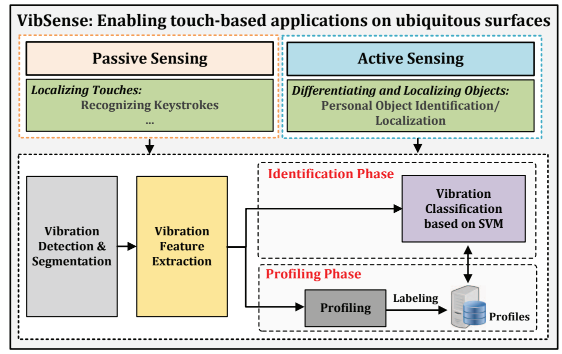


Fig. 3. Overview of VibSense.

we explore the limit of localizing touches through passive vibration sensing in human-computer interaction (HCI), for example, building a ubiquitous keyboard for mobile devices as shown in Figure 2(a).

2) Personal Object Localization/Identification: Fine-grained localization/ identification of objects plays an increasingly important role in smart home, smart healthcare, and smart cities. The existing approaches rely on pressure sensors [16] or tactile sensors [17], but their sensing capability is restricted in a small area, which limits the usage in practice. To address these issues, we explore the limit of vibration sensing on extended surfaces through VibSense, which enables localizing/ identifying personal objects (i.e., objects with fixed shapes and weights that are usually carried by a person, such as keys and phones) on any solid surface (Figure 2(b)). By exploiting the vibration signal propagation principles in physics, Vibsense pushes the limits of vibration sensing to provide fine-grained localization of a personal object placed on any solid surface through capturing the attenuation and interference effects to the vibration signals, resulting from the characteristics of the object (e.g., weight, shape, location and material).

III. VibSense Overview

The main objective of this work is to develop a general system that can realize the vibration-based sensing modality and explore its limit for various domains requiring fine-grained information. Toward this end, we design a low-cost vibration-based sensing system, *VibSense*, which aims to work on ubiquitous surfaces for localizing touches, object identification and differentiation.

The vibration-based sensing could be separated into *Passive Sensing* and *Active Sensing* depending on whether the vibration source is known to the system. VibSense can support both types of sensing and facilitate different touch-based applications, including recognizing keystrokes on a surface, personal object localization and identification. VibSense takes as input time-series amplitude measurements of vibration signals from a vibration receiver. After receiving the vibration signals, the system performs *Vibration Detection & Segmentation* (Section IV-B) to detect and obtain the useful segment of the received vibration signals. Next, the system utilizes the *Vibration Feature Extraction* (Section IV-D) to extract the unique vibration feature (e.g., power spectrum density) from the segmented signals in the frequency domain.

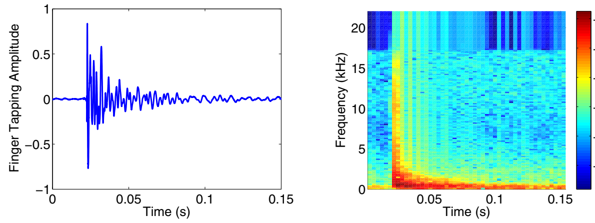
Next, the extracted vibration features are used by two phases in VibSense: *profiling* and *identification*. In the profiling phase, the extracted features are considered to be the unique signature corresponding to the characteristics of the object's touches on the medium, for example, keystrokes' locations, or the weight and size of a smartphone on a nightstand. These features are labeled with corresponding ground truth (i.e., location, object type, etc.), and saved to build an object profile. In the identification phase, the collected vibration samples are used to extract vibration features, which serve as inputs to a vibration classifier via *Vibration Classification based on SVM*. The classifier compares the extracted features with the signatures in the preconstructed profile to identify the target object and determine its location. The details of the classification are elaborated in Section IV-E.

IV. VibSense Design

In this section, we first describe the touch vibration signals (i.e., finger tapping) in passive sensing, and present how to detect and segment the vibrations, then describe the predefined vibration signals in active sensing and the unique vibration signatures being extracted. We finally show how to discriminate different touches or localize/ identify objects through classification.

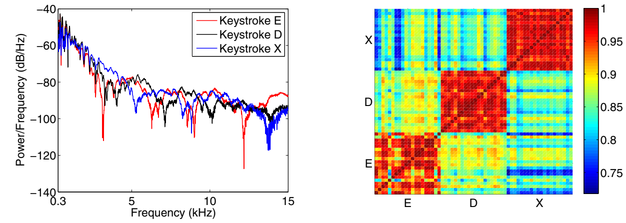
A. Unknown Vibration Source in Passive Sensing

The vibration signals collected in the passive sensing are generated by unknown vibration sources depending on specific application needs. The capability of passive sensing enables us to localize vibration sources in a fine-grained manner on ubiquitous surfaces. In particular, VibSense explores the limit of localizing close-by touches when tapping on any surface. When tapping on a medium (e.g., desk), we find that the received vibration from a finger click consists of a broad range of frequencies. The length of a finger click is usually around $0.1s$, and the highest frequency of the vibration could reach $15kHz$. Figure 4(a) and Figure 4(b) display an example of signal patterns from a finger click on the desk in the time domain and frequency domain respectively. The observed frequency band and tapping duration could guide the system to segment each keystroke/tapping and extract vibration features accurately when constructing the virtual keyboard/buttons. VibSense further utilizes the power distribution in the observed frequency band of the received



(a) Time domain patterns. (b) Frequency spectrum.

Fig. 4. Passive sensing: received vibration signals from finger tapping a desk surface.



(a) (b)

Fig. 5. Finger click vibrations of three nearby keys 'E' and 'D' and 'X' in a hand-written paper keyboard: (a) PSD pattern of keystroke vibrations; and (b) Pearson correlation between PSD of the three keys, each key is clicked 20 times.

vibration signals to discriminate close-by touches and support various applications (Section IV-D).

B. Vibration Signal Segmentation

After receiving vibration signals, VibSense utilizes an energy-based approach to detect and determine the segment of useful vibration signals. In particular, it calculates the short time energy levels of the received vibration signals by accumulating the square of their amplitudes in a sliding time window:

$$A(t) = \sum_{n=t}^{t+S} a^2(n), \tag{2}$$

where S is the length of the sliding time window and $a(n)$ is the amplitude of the received vibration signals.

We then use a threshold-based approach to detect the starting point p_s of the segment of useful vibration signals. The ending point p_e of the segment can be derived by $p_e = p_s + T_a$, where T_a is the estimated time length of original vibration signals determined by specific applications. In this work, T_a is set to 0.1s for passive sensing applications which covers the duration of most passive vibration signals (e.g.,finger tapping). The segmented vibration signal is then normalized with respect to the maximum amplitude of each to tackle different intensities of taps.

C. Vibration Signals in Active Sensing

As discussed in Section II-A, the attenuation and interference of vibration is strongly affected by the frequency of vibration signals. In active sensing, VibSense utilizes a vibrator to generate vibration. The vibration signals need to satisfy two aspects: i) contain a broad range of frequencies to increase the diversity of vibration features in the frequency domain; and ii) have sufficient vibration power (i.e., magnitude) to be transferred to the receiver end to support an expanded physical transmission medium (e.g. a large desk). Specifically, the frequencies of the vibration signals increase logarithmically with time, which can be represented as:

$$f(t) = f_0 \times (\frac{f_1}{f_0})^{\frac{t}{T}}, \tag{3}$$

where f_0 and f_1 are the initial and ending frequencies used from the frequency band, and T is the time duration of the generated vibration signal. In this work, we use $T = 1$s to maintain the balance between good performance and low annoyance caused by the vibration. The initial and final frequencies are determined by the hardware used in the prototype of VibSense. We empirically choose a relatively low frequency range (i.e., $f_0 = 300$Hz and $f_1 = 12$kHz) in VibSense to support a larger sensing area, since the magnitude of the vibration signals generated by a vibrator would be greatly decreased under the higher frequency range. We discuss the details of the vibrator and the generated vibration in Section V-C. Generally, VibSense could transmit the vibration signals repeatedly with a short time interval to keep its continuous sensing capability, and we use the similar method as discussed in Section IV-B to detect and segment each vibration signal.

D. Vibration Feature Extraction

Equation 1 shows that the effect of the channel is reflected through the amplitude of the received vibration signals, which is dominated by multi-dimensional factors, including the propagation distance, vibration frequencies, and material of the object touching the surface. Each transmission medium can be considered as a frequency selective channel for vibration signals resulting in different power and amplitude for the received vibration signal in the frequency domain. We thus choose vibration features based on the power of received vibration signals in the frequency domain in VibSense.

Specifically, VibSense utilizes the power spectral density (PSD) of the received vibration signals in both passive and active sensing as the basis for feature extraction to perform localizing touches and differentiating/ localizing objects. The PSD reflects the power distribution of the sensed vibration signals at each specific frequency, which can well capture the attenuation and interference effects influenced by vibration source, propagation medium, and objects contacting the medium surface. The PSD of a received vibration signal r_i can be estimated by:

$$PSD_i = 10\log_{10}\frac{(abs(FFT(r_i)))^2}{f_s \times n}, \tag{4}$$

where n is the number of samples of the received signal r_i, f_s is the sampling rate, and $FFT(\cdot)$ is the fast Fourier transform operation.

To demonstrate the capability of using PSD feature to support both passive and active sensing in VibSense, we show the results from two preliminary experiments: virtual keyboard construction on a desk surface (passive sensing) and object location differentiation on a table (active sensing). Figure 5(a) shows the distinguishable PSD features of the received vibration signals collected in a passive sensing scenario, i.e., when a user taps multiple times without

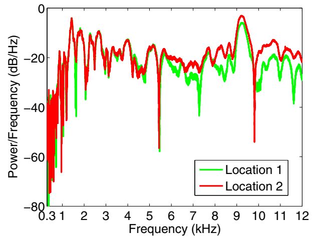


Fig. 6. PSD of the received vibration signals when an object is placed at two locations of a wooden table.

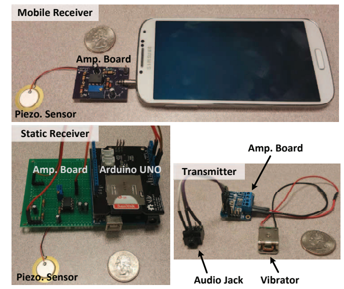


Fig. 7. The hardware design of VibSense.

controlling intensity at each of three close-by positions on a desk, corresponding to keys 'E', 'D', and 'X' on a handwritten paper keyboard. Figure 5(b) further indicates that the PSD features associated with the finger clicks at the same position have higher correlation than those at different positions. In addition, Figure 6 compares the PSD features of two vibration signals that are received in an active sensing scenario, i.e., when a mug is placed at two different positions with about $10cm$ distance respectively on a table. The results show promisingly distinguishable patterns in the PSDs corresponding to different locations in various frequency bands.

E. Vibration Classification

During the profiling phase, VibSense constructs a set of object profiles with the vibration features (i.e., PSD) by labeling vibration signals collected from various touch-based applications. For example, vibration features are extracted from finger clicks at different locations in *localizing touches*, a smartphone or a cup at a same location in *personal object identification*, etc. In the later identification phase, when there is a vibration signal detected and segmented, VibSense needs to extract the vibration feature from the segmented signal and classify the feature by matching it to the existing object profiles. Specifically, a vibration classifier is built inside of VibSense based on the Support Vector Machine (SVM) using *LIBSVM* [18] and the linear kernel function. The other kernel functions such as Gaussian radial basis kernel, quadratic kernel have been tested and could achieve comparable performance.

For classification, we estimate prediction probabilities for each object profile by combining all pairwise comparisons [19]. An incoming target object with the highest prediction probability for a profile would be identified as the same type. In order to prevent VibSense from mistakenly identifying an unknown target object as a known type, we devise a threshold based approach on top of object classification. After identifying the highest prediction probability for a profile, the classifier compares the probability to a threshold, and only identifies the target object as the type of the profile when the probability exceeds the threshold.

V. Hardware Platform Design

VibSense needs to meet two basic requirements: 1) receiving the vibration of a wide frequency range on the vibration receiver; and 2) precisely regulating the frequency components of the vibrations generated by a vibrator for active sensing.

A. Vibration Receiver

We design and implement two versions of low-cost vibration receivers as shown in Figure 7: the *Static Receiver* is a stand-alone embedded system based on the Arduino platform, which amplifies, digitizes, and stores received vibration signals; while the *Mobile Receiver* is a simplified version that only consists of a vibration sensor and a low-power consumption amplifier, which can be easily connected to mobile devices to facilitate mobile vibration-based sensing applications.

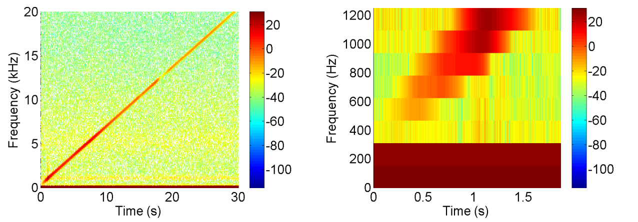
In both versions, we devise a low-cost passive vibration sensor, piezoelectric sensor, to collect vibration signals. Compared to other vibration sensors (e.g., Geophone sensors and capacitive MEMS accelerometers), the piezoelectric sensor has the largest frequency response range and the lowest cost. Moreover, the sensor is so small (i.e., 0.48 square inches in area and 0.3 inches thick) that can be easily attached to any solid surface and integrated with a smartphone, for both passive as well as active sensing.

Mobile Receiver. The mobile receiver is designed for mobile device based applications, for instance, providing a virtual keyboard on a wooden desk rather than just typing on the mobile device's confined touch screen. A low-power consumption operational amplifier *TLC272* is used to provide amplified analog voltage signals from the vibration sensor to a processing device, e.g., a smartphone. The receiver can be easily plugged into the standard audio jack of an off-the-shelf mobile device to sense vibration signals by exploiting the audio components in the mobile device. The sampling rate of this receiver is determined by the Analog to Digital Converter (ADC) used for the audio components in mobile devices, which is typically over 48kHz.

Static Receiver. The static receiver utilizes a rail-to-rail operational amplifier *OPA350* to increase the peak-to-peak voltage of the analog signals obtained from the piezoelectric sensor. The sampling rate of the ADC in the Arduino platform is set to 40kHz so that the receiver can fully recover the vibration signals with the frequency up to 20kHz based on the Nyquist rule.

B. Vibration Transmitter

For active sensing, VibSense utilizes a Linear Resonant Actuator (LRA) [20] based vibrator to regulate both frequen-



(a) Frequency spectrum. (b) Zoom-in spectrum.

Fig. 8. Frequency spectrum analysis of the received vibration signals.

cy and amplitude of vibration. The vibrator has a small size of 0.48 square inches. The vibrator has a wide frequency response and low power consumption of 1 watt RMS. The frequency and amplitude of the generated vibration is determined by the frequency and peak-to-peak voltage of an input analog signal. An efficient way to such controllable analog signal is using audio signals that can be easily generated by any off-the-shelf mobile device through its audio jack. In VibSense prototype we choose a class D audio amplifer, *MAX98306*, which can provide about 18dB gain in a wide frequency band with low power consumption. The hardware of the vibration transmitter is shown in Figure 7 (transmitter).

C. Frequency Response in Prototype

Because the vibration frequency is critical to the diversity of vibration features, we conduct an experiment by directly attaching the transmitter with the static receiver to study the frequency response of the prototype. The transmitter generates a 30s analog signal with its frequency linearly sweeping from 0Hz to 20kHz, which includes most natural vibration frequencies. The frequency spectrum of the received vibration signal is shown in Figure 8. We observe that our prototype has a wide frequency response range covering from 300Hz to 20kHz, indicating the prototype can be used to produce and receive vibration signals with a wide range of frequencies. Note that the highest frequency boundary is determined by the ADC's sampling rate. In addition, Figure 8(a) shows that the vibration strength degrades with the increment of frequency (i.e., higher frequencies present lower spectrum power). This suggests us to use a relatively lower frequency range when generating vibration signals in active sensing to cover an extended sensing area and avoid the vibration signal is too weak to be captured by the receiver. Toward this end, we empirically use the frequency band from $300Hz$ to $12kHz$ in active sensing.

VI. Performance Evaluation

A. Experimental Methodology

We evaluate the performance of our VibSense in typical home/office environments with the key applications over a six-month time period.

Keystroke Recognition (Passive Sensing). As shown in Figure 9, we evaluate the performance of localizing touches by identifying finger clicks/keystrokes from three participants on a virtual keyboard (illustrated by a piece of paper on the surface of a wooden desk). In the experiments, only the vibration mobile receiver is connected with a mobile

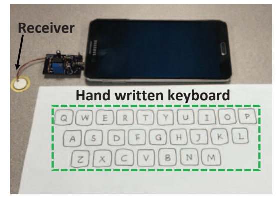


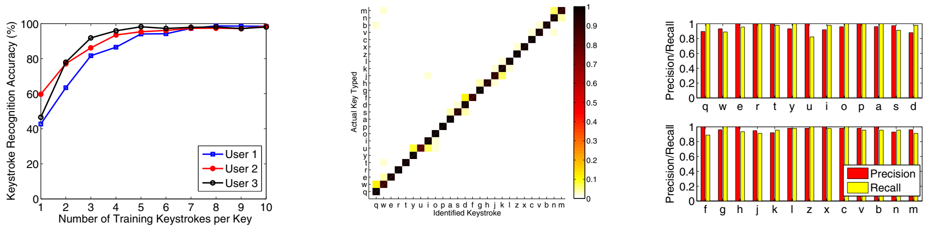
Fig. 9. Experimental setup of localizing touches: keystroke recognition on a paper keyboard.

phone (i.e., Samsung Galaxy Note 3) to perform the passive sensing. The receiver is fixed on the table at a position close to the top-left corner of the virtual keyboard. There are 26 alphabetic keys on three rows in the virtual keyboard. The distance between two rows is about 2cm, and the center-to-center distance between two adjacent keys in the same row is also about 2cm. Participants are asked to randomly type on the virtual keyboard with a natural speed (i.e., $\sim$130 keystrokes/min) in a typical office environment. Each participant types and collects vibration signals 20 times for each key. In total, there are over $1,560$ keystrokes vibration signals are collected from the three participants.

Personal Object Localization/ Identification (Active Sensing). We conduct experiments by placing personal objects at nine locations (i.e., 3×3 grid) on a middle-size wooden table with the dimension of 120cm $\times$50cm $\times$3cm. In the experiments, six personal objects (including a small empty paper cup of 8 fl oz capacity, an U.S quarter coin, a small apple, an iPhone 5s, an empty glass cup, and a can of coke) are chosen to represent different material, weight, and size. The distance between any two adjacent predefined locations is 5cm. The vibration transmitter and static receiver are attached to the surface of the table. We also adopt two setups with different distances between the transmitter and receiver (i.e., $L = 40$cm and 120cm) to mimic the common sizes of a night table and an office desk, respectively. For each setup, we place the six chosen objects on the nine predefined locations to collect vibration signals for 20 times per location per item. We additionally collect 20 vibration signals when there is no object placed on these nine locations and a mouse (i.e., labeled as *other objects*) at each of the nine predefined locations. In total, there are $1,620$ vibration signals collected for object localization and identification. The sampling rate is also set to 40kHz on the receiver.

B. Performance of Keystroke Recognition

Recognizing Keystrokes from Different Users. We conduct experiments with three participants, and the system trains the classification models for them separately. Figure 10(a) shows the overall accuracy for the keystroke recognition of three participants under different training set size. We observe that the accuracy of all three users increases with the growing size of the training set. And the average accuracy over three users is around 87% and 97% with 3 and 5 training keystrokes per key, respectively. This indicates that VibSense could provide sufficient accuracy to recognize finger clicks at different close-by locations even with only serval training keystrokes per key. Increasing the size of



(a) (b) (c)

Fig. 10. Keystroke recognition on a paper keyboard: (a) keystroke recognition accuracy for different users with different sizes of training set, (b) confusion matrix and (c) precision and recall (training set size is 5 per key).

training set beyond 5 only provides a marginal performance improvement, thus only 3 training keystrokes per key are needed to obtain reasonable performance. Requiring only limited training efforts greatly increases VibSense's usability.

Confusion Matrix of Keystroke Recognition. Figure 10(b) plots the confusion matrix of the keystroke recognition with 5 training signals for each key on the hand written keyboard. We find that there are only few keystrokes that are mistakenly identified as incorrect keys. These mistakenly recognized keystrokes usually correspond to the neighboring keys which have the similar distance to the receiver and vibration propagation path. For example, a few keystrokes of the key u are mistakenly recognized as the key y since they are close to each other with similar distance and path to the receiver attached on the table as shown in Figure 9. These few mistakenly classified keystrokes can be corrected by using a linguistic model.

Precision and Recall for each Key. The precision and recall of identifying keystrokes of each key is shown in Figure 10(c). It combines the results for all three users with 5 training keystrokes per key. Overall, the average precision is about 97% and the average recall is about 96%. These results are a strong evidence that VibSense could accurately localize unknown vibration sources (i.e., finger clicks) in very close proximity.

C. Performance of Object Identification

Object Identification Accuracy. We next evaluate the performance of object identification by placing different personal objects with fixed sizes and shapes (i.e., a glass cup, an iPhone 5, a coin and a paper cup) at the same location on a table across different days. Figure 11 shows the confusion matrix of object identification. VibSense achieves 100% accuracy, and objects which do not belong to any of the selected personal objects are identified as *unknown*. This indicates that VibSense can well capture the vibration changes caused by the characteristics of different objects and distinguish them from each other.

Impact of Object's Weight. We further study the impact of weight to vibration sensing by fixing the material and contacting area of an object while varying its weight. We collect 20 vibration signals when an empty glass is placed in between the vibrator and receiver as the baseline, and collect 20 testing vibration signals each when the same glass contains different amount of water (i.e., 34g, 86g, 159g, 236g, 345g and 414g). We calculate the Euclidean distance of the extracted PSD features between each test and the baseline signal. Figure 12(a) shows the mean and standard deviation of the calculated Euclidean distances, which shows that PSD features change with different object weights, larger weight differences would have stronger effects to the vibration signals.

Impact of Object's Material. Next, we experiment with objects of different materials but the same weight and contacting area. We put water in two cups made of different materials (i.e., glass for $cup1$ and ceramic for $cup2$) to make them have the same weight, and we add a same-size metal piece at the bottom of each cup to make sure their contacting areas are the same. We use 20 vibration signals collected from one of the cups (i.e., $cup1$) as the baseline, and we collect vibration signals of both cups for testing, and calculate the Euclidean distance of the extracted PSD features between the testing and baseline signals. Figure 12(b) shows the box-plot of the Euclidean distance. We observe that the Euclidean distances of the container made of different materials are not overlapped, indicating that the make of an object is also a strong impact factor to vibration sensing. However, the impact of the material is much smaller comparing to that of the object weight.

D. Performance of Object Localization

Localization Accuracy. Figure 13(a) shows the localization accuracy of six different objects under different number of training vibration signals when placed at 9 positions. We observe that heavier objects obtain better localization accuracy, and the localization accuracy increases with the growing number of training signals. In particular, for heavier objects such as glass cup, phone, and coke can, VibSense localizes them with accuracy higher than 86% when only one training data is used, and reaches 100% accuracy when the number of training signals is greater than four. Whereas for lighter ones such as the coin and paper cup, the average accuracy of localization reaches 60% and above when more than six training vibrations are used. This is encouraging as it shows that VibSense is capable of localizing various personal objects. Even for smaller and lighter objects like coins and paper cups, VibSense can achieve acceptable accuracy with a few training vibration signals.

Impact of Distance between the Vibrator and Receiver. Figure 13(b) compares the object localization accuracy using

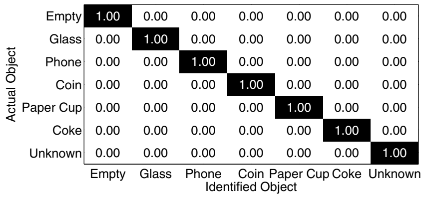
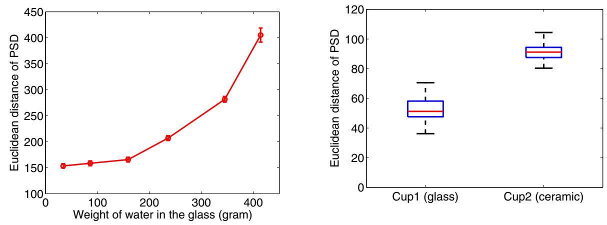
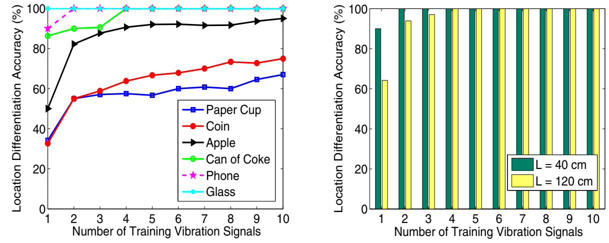


Fig. 11. Object identification: confusion matrix of identifying 5 objects placed at the same location.



(a) (b)

Fig. 12. Personal object identification: (a) impact of object weight and (b) impact of object shape/materials on the extracted vibration features.



(a) Localization of 6 different objects.

(b) Localization of a glass cup with different distances between vibrator and receiver.

Fig. 13. Fine-grained object localization accuracy.

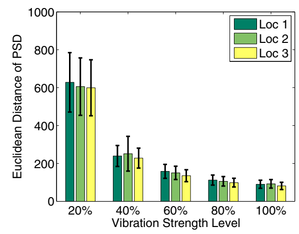


Fig. 14. Localization of a glass cup: impact of vibration signal strength.

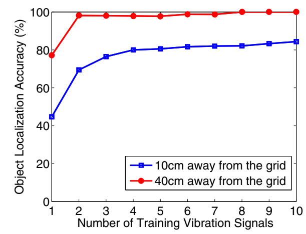


Fig. 15. Object localization with temporary presence of other objects.

a glass cup when L is set to $40cm$ and $120cm$, respectively. The accuracy performance under $L = 40cm$ is better than that under $L = 120cm$ when the number of training signals is small, whereas the localization accuracy reaches 100% under both setups when the training signals exceed four.

Impact of Vibration Signal Strength. Finally, we study the impact of different vibration signal strengths on the vibration feature consistency. We regulate the vibration signal strength by changing the amplitude of the input AC signals for vibrator from 20% to 100%. For each vibration strength level, we collect 20 pre-defined vibration signals when a glass cup is placed at three different locations of the table. At each vibration strength level, we calculate the Euclidean distance between the features extracted from any two collected vibration signals at a specific location. Figure 14 shows the mean and standard deviation of the vibration features. The results show that the stronger the strength of the vibration signals, the more stable and consistent the vibration features become (i.e., the smaller the Euclidean distances are) when an object is placed at the same location.

Temporary Presence of Other Objects. The current system is designed for identifying/localizing a single object on a surface. The temporary presence of additional objects could alter vibration features from the trained ones but in a preliminary experiment we find the effect to be pronounced only when the other objects are very close. In this experiment, we repeat the object location differentiation experiment with the glass cup, considering six possible different locations (i.e., 2×3 grid and 5cm between adjacent locations). This time we put a Samsung Note 4 mobile phone on the table during testing. As shown in Figure 15, when the phone was 10cm away, accuracy decreased noticeably from close to 100% to about 80%. When the phone was moved about 40 cm away, accuracy again approached 100%, however. This suggests that only other objects in close proximity would have a noticeable effect and this effect might be reduced with further filtering or directional sensing techniques.

VII. Related Work

In general, we separate the research on pervasive sensing into *touchless* and *touch-based* sensing depending on whether the sensing modalities require human or object touching or not. The touchless approaches rely on sensing modalities that do not require "touching", such as radio frequency (RF), acoustic sound, and visible light, whereas touch-based approaches are developed based on "touching" including capacity, pressure and vibration.

Active studies [1–4, 21] have been driven by the emerging trend of mobile devices with RF module (e.g., WiFi) to detect and track hand gestures (e.g., WiSee [1]), movements (e.g., WiDraw [3]), human daily activities [2, 4] and even recognize keystrokes from a nearby keyboard [21]. Additionally, approaches based on acoustic signal sensing have drawn considerable attentions recently [5–7, 22]. EchoTag [6] enables phones to tag and remember indoor locations by sending/sensing a sound signal with a phone's speakers/microphones. Yun *et al.* [5] track hand movements to realize mouse functions in the air by sending inaudible sound pulses. Moreover, dual-microphones of a phone are used to recognize keystrokes on a nearby paper keyboard [7, 22]. However, the stability of acoustic-based user interaction is vulnerable to various ambient noises. Furthermore, visible light is proposed to be utilized to localize fingers and reconstruct 3D human skeleton postures in real time [8, 9]. The visible light based approaches can only be implemented in a particular optical environment that is easy to be interfered. The aforementioned touchless sensing techniques are mostly sensitive to interference resulted from environmental changes and ambient noises.

Different from touchless sensing, touch-based sensing (e.g., capacitive touch sensing, resistance and pressure sensing) requires users touching a medium, thus being more applicable for applications requiring security and robustness. Recently, Touche [10] utilizes the difference of body electrical conductivity to recognize the complex configuration of the human hands and body by using special circuits. Karata and Gruteser [11] create multi-key conductive-ink touch interfaces which can be printed on a paper. Additionally, UnMousePad [16] leverages pressure-sensing sensors to distinguish multiple fingertip touches while simultaneously tracking pens. However, electrical conductivity based interaction systems can only be implemented on electrical conductive surfaces, which is not always applicable in daily lives. And, pressure sensor based techniques have the limitation of sensing the slow changes in pressure distributions of objects placed on its sensing panel.

Comparing to existing approaches, vibration based sensing is resistant to radio interferences and various environmental sounds. In previous studies, vibration has been used to communicate information [20, 23]. In terms of object sensing, Toffee [12] leverages multiple piezoelectric sensors to determine the direction of touches on a surface with respect to a device based on the acoustic time-of-arrival correlation. Touch & Activate [13] actively generates acoustic signals and records the sound patterns to identify how the user touches an object by using a pair of vibration speaker and piezo-electric microphone. Toffee relies on four well-separated sensors for determining directions, whereas Touch & Activate focuses on active sensing with both vibration speaker and sensor mounted on the same small object. In our work, the proposed VibSense enables both passive and active sensing leveraging a single receiver (and a single vibrator for active sensing). It can be easily deployed and integrated with existing mobile devices. VibSense provides an extended sensing area to demonstrate the power of vibration sensing in a broad array of applications.

VIII. Conclusion

We propose *VibSense* to explore the limit of vibration-based sensing when supporting a broad array of touch-based applications. Through sensing physical vibrations from either unknown sources (passive sensing) or a vibrator (active sensing), VibSense works with extended surface areas through a single sensor. We push the limits of vibration-based sensing by applying VibSense to key applications including keystroke recognition on ubiquitous surfaces for mobile devices, personal object localization and identification. Such an approach is robust to environmental interferences from acoustic or radio-frequency noise. The extensive experiments demonstrate VibSense successfully pushes further the limits of vibration sensing to extended surface areas with only a single receiver, making vibration-based sensing a suitable candidate to achieve high accuracy in localizing touches and fine-grained object identification/localization through both passive and active sensing. We note that there are still many other factors such as different surface materials/sizes/thicknesses affecting the system's trained models, which are left for our future work to investigate.

IX. Acknowledgement

This work was partially supported by the National Science Foundation Grants CNS-1409767, CNS-1514436, CNS-1409811 and CNS-1618019.

References

[1] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *ACM MobiCom*, 2013, pp. 27–38.

[2] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures," in *ACM MobiCom*, 2014, pp. 617–628.

[3] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, "Widraw: Enabling hands-free drawing in the air on commodity wifi devices," in *ACM MobiCom*, 2015, pp. 77–89.

[4] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *ACM MobiCom*, 2015, pp. 65–76.

[5] S. Yun, Y.-C. Chen, and L. Qiu, "Turning a mobile device into a mouse in the air," in *ACM MobiSys*, 2015, pp. 15–29.

[6] Y.-C. Tung and K. G. Shin, "Echotag: Accurate infrastructure-free indoor location tagging with smartphones," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (ACM MobiCom)*, 2015, pp. 525–536.

[7] J. Wang, K. Zhao, X. Zhang, and C. Peng, "Ubiquitous keyboard for small mobile devices: harnessing multipath fading for fine-grained keystroke localization," in *ACM MobiSys*, 2014, pp. 14–27.

[8] C. Zhang, J. Tabor, J. Zhang, and X. Zhang, "Extending mobile interaction through near-field visible light sensing," in *ACM MobiCom*, 2015, pp. 345–357.

[9] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, "Human sensing using visible light communication," in *ACM MobiCom*, 2015, pp. 331–344.

[10] M. Sato, I. Poupyrev, and C. Harrison, "Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects," in *ACM CHI*, 2012, pp. 483–492.

[11] Ç. Karataş and M. Gruteser, "Printing multi-key touch interfaces," in *ACM UbiComp*, 2015, pp. 169–179.

[12] R. Xiao, G. Lew, J. Marsanico, D. Hariharan, S. Hudson, and C. Harrison, "Toffee: enabling ad hoc, around-device interaction with acoustic time-of-arrival correlation," in *ACM MobileHCI*, 2014, pp. 67–76.

[13] M. Ono, B. Shizuki, and J. Tanaka, "Touch & activate: adding interactivity to existing objects using active acoustic sensing," in *ACM UIST*, 2013, pp. 31–40.

[14] A. Abdullah and E. F. Sichani, "Experimental study of attenuation coefficient of ultrasonic waves in concrete and plaster," *The International Journal of Advanced Manufacturing Technology*, vol. 44, no. 5-6, pp. 421–427, 2009.

[15] D. M. Howard and J. Angus, *Acoustic and Psychoacoustics: Second edition*. Focal Press, 2001.

[16] I. Rosenberg and K. Perlin, "The unmousepad: an interpolating multi-touch force-sensing input pad," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 3, p. 65, 2009.

[17] R. A. Russell, "Object recognition by a smart tactile sensor," in *Proceedings of the Australian Conference on Robotics and Automation*, 2000, pp. 93–8.

[18] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[19] T.-F. Wu, C.-J. Lin, and R. C. Weng, "Probability estimates for multi-class classification by pairwise coupling," *The Journal of Machine Learning Research*, vol. 5, pp. 975–1005, 2004.

[20] N. Roy, M. Gowda, and R. R. Choudhury, "Ripple: Communicating through physical vibration," in *USENIX NSDI*, 2015, pp. 265–278.

[21] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using wifi signals," in *ACM MobiCom*, 2015, pp. 90–102.

[22] J. Liu, Y. Wang, G. Kar, Y. Chen, J. Yang, and M. Gruteser, "Snooping keystrokes with mm-level audio ranging on a single phone," in *ACM MobiCom*, 2015, pp. 142–154.

[23] N. Roy and R. R. Choudhury, "Ripple ii: Faster communication through physical vibration," in *USENIX NSDI*, 2016, pp. 671–684.