# VibEye: Vibration-Mediated Object Recognition for Tangible Interactive Applications

**Seungjae Oh**
POSTECH
Pohang, Gyeongbuk, Korea
oreo329@postech.ac.kr

**Gyeore Yun**
POSTECH
Pohang, Gyeongbuk, Korea
ykre0827@postech.ac.kr

**Chaeyong Park**
POSTECH
Pohang, Gyeongbuk, Korea
pcy8201@postech.ac.kr

**Jinsoo Kim**
POSTECH
Pohang, Gyeongbuk, Korea
fusion@postech.ac.kr

**Seungmoon Choi**
POSTECH
Pohang, Gyeongbuk, Korea
choism@postech.ac.kr

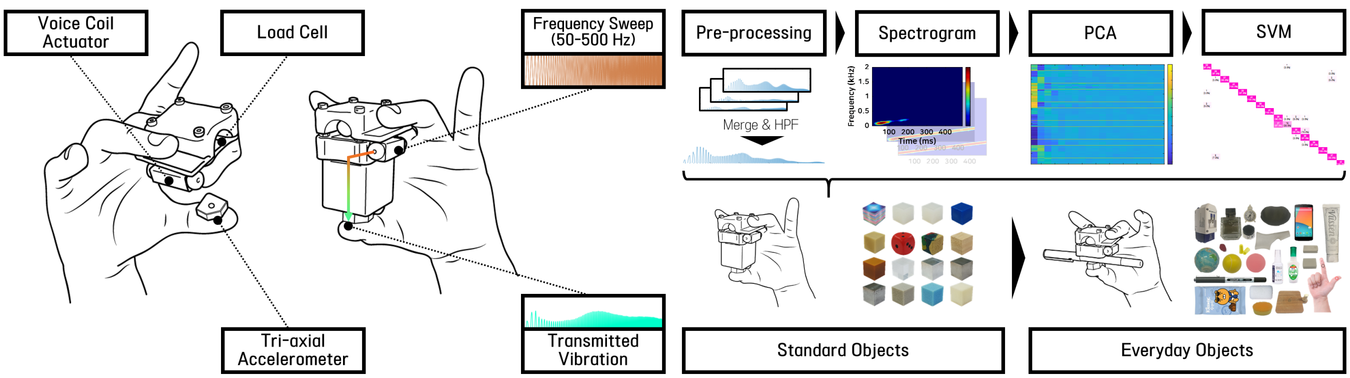


**Figure 1: VibEye: system overview and operation principle (left). Data processing pipeline (right-top). A classification model is built from 16 standard objects and then applied to categorize 25 everyday objects (right-bottom).**

## ABSTRACT

We present VibEye: a vibration-mediated recognition system of objects for tangible interaction. A user holds an object between two fingers wearing VibEye. VibEye triggers a vibration from one finger, and the vibration that has propagated through the object is sensed at the other finger. This vibration includes information about the object's identity, and we represent it using a spectrogram. Collecting the spectrograms of many objects, we formulate the object recognition problem to a classical classification problem among the images. This simple method, when tested with 20 users, shows 92.5% accuracy for 16 objects of the same shape with various materials. This material-based classifier is also extended to the recognition of everyday objects. Lastly, we demonstrate several tangible applications where VibEye provides the needed functionality while enhancing user experiences. VibEye is particularly effective for recognizing objects made of different materials, which is difficult to distinguish by other means such as light and sound.


*CHI 2019, May 4–9, 2019, Glasgow, Scotland UK*

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300906

## CCS CONCEPTS

• **Human-centered computing** → **Haptic devices**; *Virtual reality*; Mixed / augmented reality;

## KEYWORDS

Vibration-based Sensing, Object Recognition, Tangible Interaction, Passive Haptics, Virtual Reality, Augmented Reality

**ACM Reference Format:**
Seungjae Oh, Gyeore Yun, Chaeyong Park, Jinsoo Kim, and Seungmoon Choi. 2019. VibEye: Vibration-Mediated Object Recognition

for Tangible Interactive Applications. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300906

## 1 INTRODUCTION

Living in the era of virtual reality, context-rich digital augmentation is an ultimate goal when coupling digital and physical worlds. Absence of comprehensive sensing methods, however, is one of the key obstacles against expanding the realm of digitally augmented computing. As such, previous studies explored ways for computing devices to recognize the context of use, e.g., location [4, 13], object [18, 37] and its state [8, 43], user's action [17, 18, 26, 41], and material [5, 31, 36–38], via numerous sensing technologies.

When recognizing materials or objects, sensing methods exploit various physical channels, including photic, electromagnetic, acoustic, and vibratory, to amplify differences among a set of materials or objects. This physical channel of sensing determines the class of materials or objects that lead to reliable identification. Light-based sensing yields multi-spectral information that enables recognition of a variety of materials. However, this approach generally requires a flat surface for suitable light reflection. In addition, since only local information is retrieved, light-based sensing does not ensure reliable recognition of the objects with complex material composition or arbitrary shape. Object recognition using electromagnetic signals is effective but only applicable to the objects that emit electromagnetic waves. Acoustic sensing focuses more on detecting the state changes of objects because acoustics waves are sensitive but weak. See Section 2 for further details.

In this work, we are concerned with distinguishing objects on the basis of their *material* differences. As we often shake an object to feel its characteristics, we impose a frequency-modulated mechanical vibration to an object and then measure and analyze the response vibration. This approach allows us to deal with objects that have different mechanical properties, e.g., stiffness, damping, density, and weight. Such mechanical properties define the feel that users perceive when holding objects in hand, and our method is naturally tailored to tangible interaction.

As depicted in Figure 1, our hardware, named VibEye, is designed to support the pinching gesture holding an object between fingers. For accurate and robust object recognition, we process the data in such a way that visualizes the spectral content of the object, which is determined by the material and structural properties, using a spectrogram. Spectrograms enable us to use standard image-based classification methods. We use PCA (Principal Component Analysis) in an unsupervised manner, followed by a C-SVM (Classification Support Vector Machine) classifier. This approach removes the cumbersome feature selection step for classifier design, which was an issue for prior vibration-based techniques (Section 2).

We evaluated the VibEye system with two sets of objects: one consisting of 16 standard objects with the same cubic shape and the other with 25 everyday objects. The former was to stress their differences in the material, while the latter was to explore further applicability of our material-based recognition method to object with different shapes. When tested with 20 users, VibEye recognized the 16 standard objects with high accuracy (92.5%) in spite of uncontrollable hand orientation change and low-frequency motion. We also extend the classifier trained on the standard objects to the recognition of the unseen everyday objects to assess the extent to which material properties are captured in the classifier. The material-based classifier can indeed recognize everyday objects made of similar materials appropriately. Lastly, we showcase two interactive applications utilizing VibEye in virtual and augmented reality, respectively.

The main contributions of our work are with a proposal of a vibration-mediated recognition system of handheld objects emphasizing their material properties and a validation of its performance. VibEye's concept is useful for all tangible applications in real, virtual, and augmented environments.

## 2 RELATED WORK

Recognizing objects of interest affords a computing device opportunities to assume the user's context [4, 36]. Such context-awareness has the potential to enrich interaction by providing relevant functionalities [18, 37]. Further, the material that an object is made of reinforces the context [5, 11]. The recognized information enables the object to function as part of the computing environment. In this work, we focus on how to recognize handheld objects based on their differences in material in order to support tangible interaction in various computing contexts and environments.

### Object Recognition from Mechanical Vibration

A mechanical vibration that propagates through a medium leaves a unique signature. Research on vibration-based sensing has followed two strategies: active object sensing and passive object sensing. The former is to detect the patterns of vibrations emanating from an object that encompasses its own oscillation source, e.g., motor-powered vibrating objects. ViBand [18] measures a vibration transmitted through the human body to recognize the object or hand gesture that has generated the vibration. Vibrosight [41] remotely detects unique vibration patterns of objects in operation with laser vibrometry. While useful, these methods are not applicable to still objects.

The other approach is to apply a structured vibration (e.g., sinusoidal frequency sweep) to a static object and measure the response, and then compare the input and output (I/O) signals for object recognition. This approach works with any objects although it requires an external vibration source. Kunze et al. [13] and Cho et al. [4] find the location of a mobile phone by measuring acceleration (and also sound [13]) using internal sensors after imposing a vibration. They also envision the possibility of material recognition from vibration signatures. In addition, sending a vibration to the hand enables estimation of the grip force [10] and the hand posture [12, 40]. Such hand-transmitted vibrations are generated and measured by surface transducers and contact microphones. VibeBin [43] takes advantage of resonance when the object is exposed to vibration. This system learns the discrete fill-levels of a waste bin and then classifies them using clustering. Our approach is generally in line with the second strategy.

### Data Processing Methods of Vibration-based Sensing

Signal processing of vibration-based sensing methods mostly includes frequency domain analysis for feature extraction. Frequency domain features are calculated using FFT (Fast Fourier Transform) or STFT (Short-Time Fourier Transform) of acceleration measurements [7]. ViBand [18] uses features in four categories (power spectrum, descriptive statistics, 1st derivative, and spectral band ratios). Kunze et al. calculate features in multiple complex categories from a 5-s signal [13]. FingerPing [40] applies a series of chirps (total 0.5 s) and extracts 294 features in nine categories. These methods share the typical issue of complex feature selection and computation. To reduce the feature dimension, VibePhone [4] applies PCA to the FFT results of 1-s vibrations. In contrast, our processing method uses spectrograms and PCA, considering both spectral and temporal information without explicit feature selection and with significant dimension reduction.

### Other Object Recognition Methods

We also review object recognition methods applicable to tangible interaction relying on other physical channels.

*Light.* Light spectra of objects or materials are a very informative source for recognition. Magic Finger [36] integrates multiple image sensors to track and capture the surfaces of different materials touched by a finger. Ohnishi et al. [25] capture images using a wrist-worn camera to identify handheld objects. Researchers also devised several methods to improve the coverage of vision-based object sensing by varying light conditions. Harrison and Hudson present a device that makes use of multiple light sources (RGB, ultraviolet, and infrared) [11]. SpeCam [38] captures multiple images under multi-spectral lighting from a mobile phone display. SpecTrans [31] differentiates transparent surfaces using a mouse laser sensor in a structured light condition. These light spectrum-based methods are likely to work well for tangible interaction, but they are not suitable for querying the invisible aspects of objects, such as weight, stiffness, damping, and internal structure.

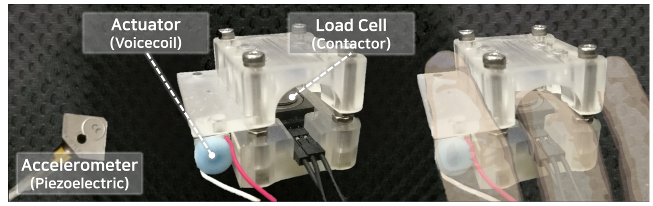


**Figure 2: Hardware design of VibEye.**

*Electromagnetic.* Electric devices emit electromagnetic waves, and sensing them using magnetic sensors provides information about their usage [19, 22, 35]. RadarCat [37] increases the frequency range of electromagnetic signals to gigahertz for less crowded bands. It can classify various kinds of objects such as body parts, transparent materials, and everyday objects. However, objects must be placed on the sensing pad, which makes it difficult to make interaction with the objects.

*Acoustic.* For object identification, acoustic sensing methods generally share a strategy with vibration-based passive object sensing in that both trigger an external signal to an object. However, the frequency ranges of the triggered signals are noticeably different. Unlike vibration-based methods (< 1 kHz), acoustic methods have been developed using a wide and higher frequency range of waves from audible (< 20 kHz) [8, 33] to inaudible (> 16 kHz) [16, 17, 26, 27]. Most of them aim to recognize the state changes of an object rather than the object itself because of weak signal power compared to the other physical channels.

Using an audible sweep sound, EchoTag [33] finds the place on which a mobile phone is, and SoQr [8] estimates the fill-level of a container to which microphone and speaker units are attached. Acoustic sensing in the inaudible range also concentrates on recognizing interaction behaviors with an object, e.g., touch position [26], contact pressure [27], and state changes [16, 17].

## 3 SYSTEM DESIGN AND IMPLEMENTATION

In physics, a mechanical wave propagates through the solid medium while oscillating and transferring energy. The vibration propagation dynamics depends on three mechanical components (mass, spring, and damper), and each of them is a function of vibration frequency. Materials with noticeable differences in such physical characteristics present distinct responses to the imposed vibration.

In our use scenario, a user grasps an object with two fingers, one wearing VibEye and the thumb (Figure 1, left).

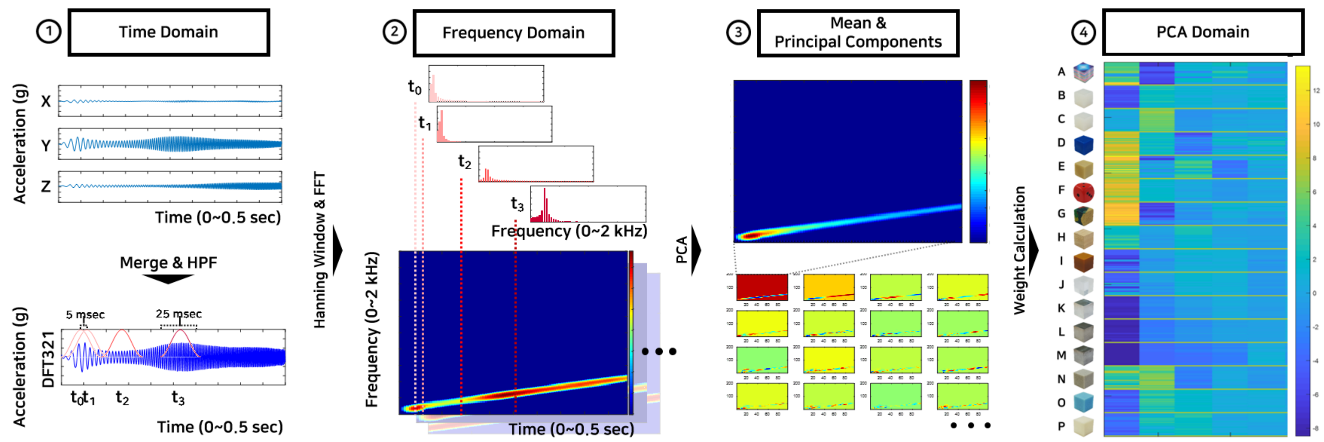


**Figure 3: Computational procedure for signal processing and object recognition.**

Then VibEye generates a short vibration, and it is transmitted to the thumb *through the object* and measured by an accelerometer fastened on the thumb's pad. From this I/O pair, we can estimate the vibration propagation dynamics of the object, and it may provide reliable information for object recognition. The vibration propagation dynamics is generally nonlinear, and its identification using the system theory is a quite complicated problem. Exploiting the fact that we trigger the object dynamics using a known input, we have designed a simple and effective data processing procedure. The input vibration also delivers confirmation feedback notifying proper and stable grasping for tangible interaction.

## Hardware Design

VibEye is designed to support object grasping by pinching. Our 3D-printed prototype has a cube-like structure with a hole into which the middle finger is inserted like a thimble (Figure 2). Top and bottom parts are made separately and assembled using bolts and nuts. For vibration generation, a voice coil actuator (Tactile Labs; Haptuator MM3C-HF), powered by an audio amplifier (Texas Instruments; TPA6211A1), is attached to the bottom part. A small compression load cell (TE Connectivity; FS2050-000X-1500-G) is also put inside on which the first phalanx of the middle finger is placed. This load cell is to measure active pressing force for contact detection and vibration trigger control. For the latter, we generate a vibration only when the measured pressure reaches 0.2 kgf for the regulation of operating condition. The vibration that has propagated through the object is sensed by a high-performance triaxial accelerometer (Kisler; Type8765A250M5) attached to the thumb's pad. All the sensors and actuators are connected to a data acquisition unit (National Instruments; USB-6251) with 20-kHz sampling rate.

## Data Processing Pipeline

In general, the vibration propagation dynamics of an object is nonlinear. However, for a short period of time, it can be approximated by using linear dynamics. Our approach is based on this general linearization strategy.

VibEye produces a sinusoidal vibration with the frequency varying linearly from 50 Hz to 500 Hz as an input to the object. This method is called sinusoidal frequency sweep, and it is widely used for system identification to trigger the response of a system in the frequency range of interest [20]. The frequency range was experimentally optimized for object recognition. The signal duration is 0.5 s, which is transient enough not to interfere with the user's gross motion and also sufficiently long for faithful dynamics reconstruction. The output vibration from the object is measured by the accelerometer. This I/O vibration pair is fed to a series of signal processing operations (implemented with MATLAB).

The vibration output is preprocessed as follows. First, we combine the three orthogonal signals from the triaxial accelerometer to make it invariant to the hand's orientation changes. We use the DFT321 method that maximizes both temporal and spectral similarity between the three individual signals and the merged signal [15]. This step may substantially improve the reliability of accelerometer information compared to the previous methods using individual signals separately [18]. Second, we apply a high-pass filter (HPF) to remove the effects of low-frequency hand motion. The bandwidth of human voluntary motion is very low, and that of the wrist motion for daily activities is about 5 Hz [23]. For HPF, we use a Chebyshev type-2 filter that has a flat passband with stopband frequency 5 Hz, passband frequency 50 Hz, and stopband attenuation level 100 dB.

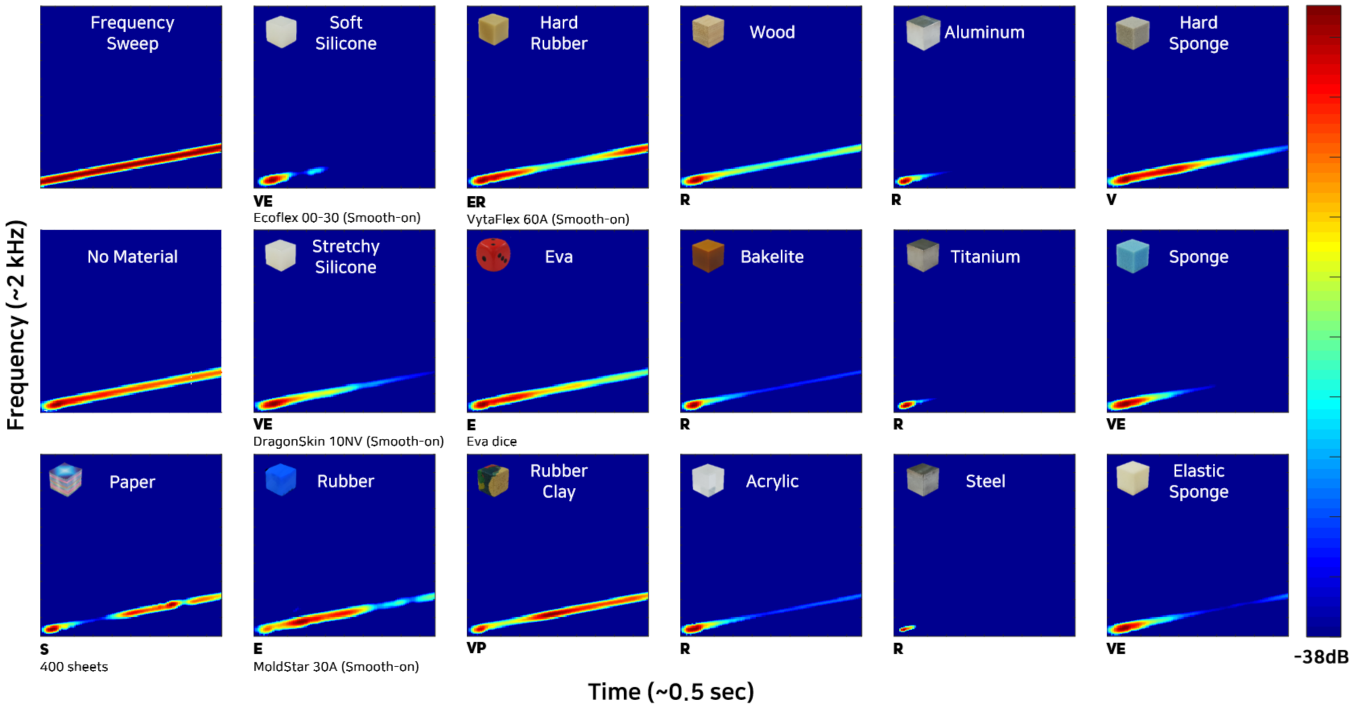


**Figure 4: Spectrograms of 16 standard objects. Objects are marked with material properties (R: rigid, E: elastic, V: viscous, P: plastic, and S: stacked). The cutoff level $\delta$ was -38 dB. The cubic objects that we molded from liquid materials are specified with the material manufacturers and models.**

Next, we compute the spectrogram of the preprocessed output signal, following the procedure shown in Figure 3. The output signal is 0.5-s long, and it is segmented to a sequence of 96 short signals by sliding a 0.025-s long Hanning window. Each segment overlaps 80% with the neighbors to preserve spectral continuity. Then, we apply 2048-DFT to every segment and stack the results along the time to construct a single image of spectrogram (96 by 1025). This raw spectrogram is filtered by removing the values lower than a cutoff level $\delta$ to suppress noise and transients, and then the outcome is normalized. The final spectrogram is unique to the object, like a *signature*, since we use the same input vibration to all objects (see Figure 4). This method allows us to recover the mechanical characteristics of the object sufficient for recognition, without identifying the full nonlinear dynamics of the object.

We collect spectrograms for various objects. Then, object recognition corresponds to classifying an input spectrogram into the spectrograms of the objects. To this end, we apply PCA to the spectrograms. PCA extracts the most discriminant features from the images in an unsupervised manner, and use the results for classification [34]. Therefore, we simply use a single image for classification, without defining and calculating many explicit features from the results of DFT as in the previous work [13, 18, 40]. As features for SVM, all the resulting PCs (principal components) of the spectrograms are fed to C-SVM (implemented with the LIBSVM library [1]) with a linear kernel to train a classifier.

The above procedure is simple and requires very little effort for tuning. The only critical parameter is the cutoff level $\delta$ for spectrogram computation, and the tuning results are easily distinguished by looking at the images.

## 4 SYSTEM EVALUATION

### Sample Sets

*Standard Objects.* We prepared 16 cubes as shown in the insets of Figure 4. Their mechanical responses are independent of the vibration stimulation orientation (top-bottom, left-right, or front-back). We carefully selected 16 different materials to cover the diverse range of elasticity, viscosity, and weight of everyday objects. The edge length was 35 mm for all cubes, except for two slightly larger cubes of eva foam and paper. Using a variety of materials is important not only for the evaluation of object recognition performance, but also for affording rich haptic sensations to users who use the cubes as props—the 16 cubes work as basic building blocks for the feel of interaction. We call them *standard* objects, and their spectrograms are shown in Figure 4.

*Everyday Objects.* We also collected 25 *everyday* objects of complex shapes and material compositions (Figure 5). They represent casual objects that can be used for tangible interaction with natural affordance. They also encompass the various material properties of the standard objects and are sufficiently small to be held by pinching. Some of them are rigid and heavy (perfume, mobile phone, and multi-plug), rigid and light (spray, cosmetics, small table clock, wooden plate, and pen), packaged (wet glue, wet tissue, and toothpaste), heterogeneous (wash sponge and wrist pad), or homogeneous (wet sponge, small polyurethane (PU) ball, large PU ball, synthetic rubber ball, jelly, elastic eraser, hard eraser, ear plug, and wristband). The rest are three human finger parts (the first phalanx of the thumb and the first and second phalanx of the index finger). Here we consider a scenario of using the thumb and index finger of the dominant hand for interaction while wearing VibEye in the non-dominant hand for object recognition.

Figure 5: Twenty-five everyday objects.

### Recognition Performance

We measured the response of each standard object 20 times using VibEye worn in *one* user's fingers and then processed the measured vibrations as described in Section 3. This step resulted in 320 spectrograms, and they were used as input images to PCA. The PCA results were reasonable (see the right panels in Figure 3). Nearly 0.1 million pixel information in each spectrogram is represented by 319 PCs. On average, the first PC explained 65.4% of the variance, and 90.7% of the variance was accounted for by the first three PCs.

We then ran cross-validation tests using C-SVM classifiers and obtained an accuracy of 96.9% and 96.3% for 5-fold and 10-fold cross validation, respectively. This high performance instantiates the effectiveness of our method. Hence, we proceeded to build a classifier to be used in our user study with many users, also including the 25 everyday objects. For that, we divided the dataset into two for training and testing while varying the proportion and tested the resulting classifiers with different users. The most balanced classifier was found at the 1:1 proportion with 94.4% accuracy, and this one was used for the user study.

The computation performance of our object recognition method is also appropriate for tangible interaction. A single execution takes only 31 ms on average in Matlab.

## 5 USER STUDY

We conducted a user study in order to test the performance of VibEye with actual users. The classifier used was built from the data of one user, as described in Section 4. The first goal was to test whether the classifier can recognize the standard objects held in other users' hands. The second goal was to observe how the standard object classifier reacts to unseen everyday objects. Some of the 25 everyday objects and some of the 16 standard objects are composed from similar materials. If there is an everyday object consistently recognized as a certain standard object, it means that we can use the classifier to recognize that everyday object on the basis of material. This user study was approved by the Institutional Review Board (IRB) at the authors' institution (PIRB-2017-E068).

### Methods

We recruited 20 right-handed participants (16 males and 4 females; 19–39.5 years old with M 23.5 and SD 4.18) for this user study. Each participant was paid 10 USD.

Participants put the VibEye device on their non-dominant hands. All participants finished two sessions, first with the 16 standard objects and the other with the 25 everyday objects. Also, each session had three repetitions. The presentation order of the objects within each repetition was fully randomized. The first repetitions were regarded as practice, and their data were not included in data analysis.

During the experiment, participants followed instructions shown on a monitor screen. The instructions were simple: hold an object by pinching and apply gentle pressure between the middle finger and the thumb so that it remains in the displayed force range (0.2 ± 0.02 kgf). After the contact pressure had stayed within the force range for 500 ms, a vibration was generated for frequency sweep. The accelerometer on the thumb picked up the vibration that had propagated through the object. Participants were asked to align the object's center of mass between the two grasping fingers while avoiding touching the object with other fingers or any other things around with the object. There were 48 trials for the session with the standard objects and 75 trials for the everyday objects. The experiment took less than an hour.

### Results

*Standard Objects.* A confusion matrix for the standard objects is shown in Figure 6, where a number in the $(i, j)$ cell indicates the percentage of the $i$-th object classified as the $j$-th object. It is evident that the diagonal terms are dominant

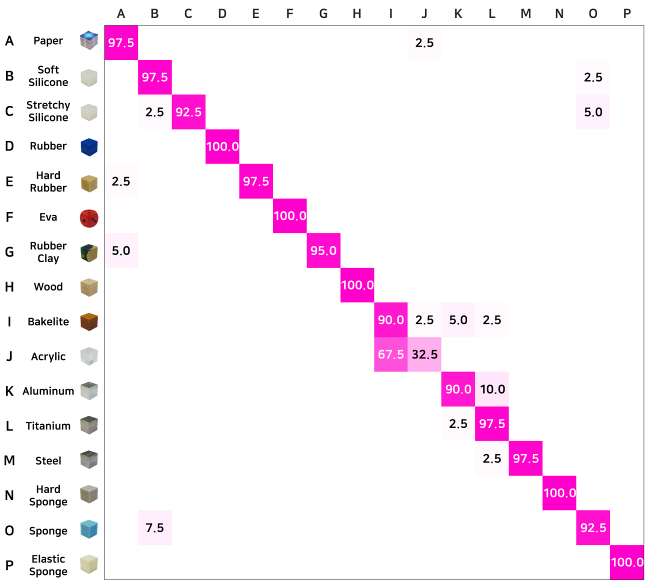


**Figure 6: Confusion matrix for recognizing standard objects.**

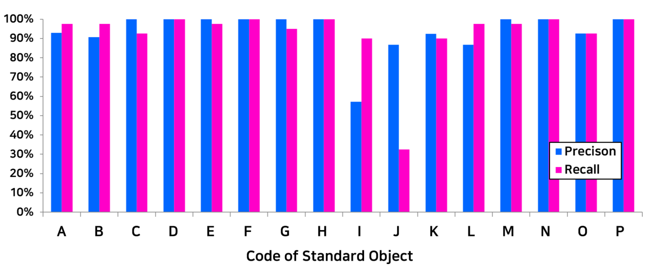


**Figure 7: Precision and recall for standard objects.**

for almost all objects. The average accuracy over all standard objects and participants was 92.5%. This result validates the effective of our one-person classifier in recognizing the standard objects held by many different users.

The most noticeable errors were present with the acrylic cube that was recognized as the bakelite cube by 67.5% (Figure 6). These two rigid plastic cubes have similar mechanical properties (acrylic: density 1.19 kg/m$^3$, weight 50.8 g and bakelite: 1.3 kg/m$^3$, 59.9 g). Our method was unable to distinguish such subtle differences. When the two cubes' data are aggregated, the accuracy is improved to 96.9% (SD 3.8%).

Individual recall and precision values are shown in Figure 7 for all the standard objects. On average, precision was 93.7% (SD 11.0%), and recall was 92.5% (SD 16.4%)[1]. The average f-score was 92.0%. These statistics re-confirm the effectiveness of VibEye. Here again the bakelite (code I) and the acrylic cube (J) caused the performance drop. Combining the two

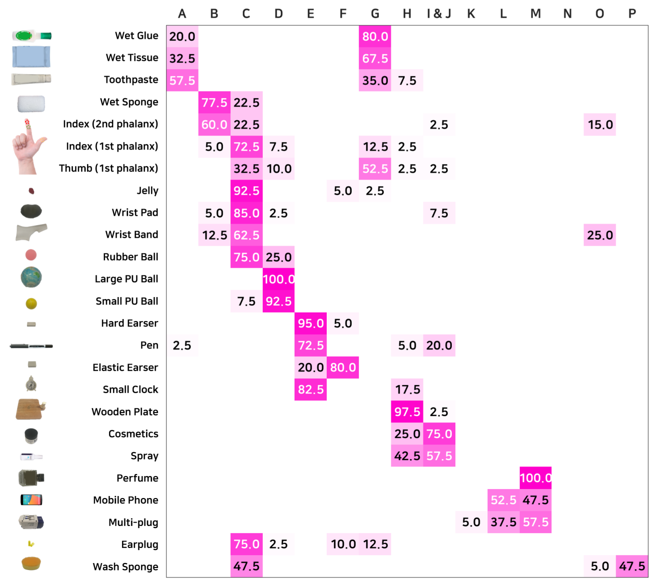


**Figure 8: Confusion matrix for categorizing 25 everyday objects to 16 standard objects.**

cubes's data improves the statistics to 97.0% (SD 4.4%), 96.9% (SD 3.2%), and 96.9%, respectively.

*Everyday Objects.* Figure 8 shows a confusion matrix of the everyday object classification results. Note that the two standard objects with very similar materials (bakelite and acrylic) are aggregated as plastics for simplicity.

We illustrate how to interpret the results in Figure 8 using examples. The cell at which "wet glue" and "G" are crossed represents that wet glue was recognized as rubber clay (code G; see Figure 6) by 80%. Similarly, the cell of "wet glue" and "A" tells that wet glue was also identified as paper (code A) by 20%. Examining the individual cells this way allows us to summarize prominent mappings from everyday object to standard object, as shown in Table 1. If an everyday object was classified as a standard object with over 70% of ratio, that pair is enrolled in the table with their common mechanical properties. Eighteen everyday objects (out of 25) were paired.

Some of the everyday objects are elastic, and most of them are classified into one of the two silicones (B and C), rubber (D), and eva (F), as shown in the top group of Table 1. The specific relations are in agreement with their level of elasticity. In the middle group, wet (liquid) glue shows slow restoration after deformation due to its viscous content, and it is 80% classified as the most viscous standard object, rubber clay (G). The everyday objects in the bottom group are all rigid. They are paired with rigid standard objects with a reasonable distribution. Weight is an influential factor to our

[1]The average recall is identical to the average accuracy because the numbers of presenting each object were all the same.

**Table 1: Prominent mappings from 18 everyday objects to standard objects (from Figure 8).**

| Everyday Objects | Standard Object | Features |
|---|---|---|
| Wet sponge | B (soft silicone) | Elastic and viscous |
| Jelly; wrist pad; index (1st Phalanx); rubber ball; earplug | C (stretch silicone) | Very elastic and viscous |
| Two PU valls | D (rubber) | Highly elastic |
| Elastic eraser | F (eva) | elastic and little viscous |
| Wet glue | G (rubber clay) | viscous |
| Wooden plate | H (wood) | Wooden and rigid |
| Hard eraser; pen; small clock | E (hard rubber) | Light and rigid (Little elastic or buzzing) |
| Cosmetics | I & J (plastics) | Light and rigid |
| Mobile phone; perfume; multi-plug | L & M (metals) | Heavy and rigid |

classifier model, and heavy objects over 100 g (mobile phone, multi-plug, and perfume) are categorized to the standard objects made of metals (titanium (L) and steel (M)). Light and rigid objects are mapped to hard rubber (E) if they exhibit some noticeable properties, e.g., bouncing a bit when thrown (hard eraser) or making buzzing noise from assembled parts (pen and small clock). Other light rigid everyday objects are classified to wood (H) or plastics (I & J), mostly depending on their materials.

## 6 DISCUSSION

### Summary of Results

Our work is based on a simple idea: when a user holds an object between fingers, we can recognize the object by triggering a short vibration and sensing the transmitted vibration changed by the object's mechanical properties. Our method empowers ordinary objects to work as props of rich haptic sensations, contexts, and affordance for tangible interaction in any real, virtual, or augmented environments. For proof of the concept, we designed and implemented a simple prototype named VibEye and effective, efficient, and easy-to-use object recognition algorithms capitalizing on spectrograms and PCA.

We trained an actual classifier of the 16 standard objects on the data measured from one user. All the standard objects were of the same shape (cube), so the classifier learned their material differences. The classifier showed high cross-validation performance. When tested with 20 other users, the one-person classifier showed quite high performance (accuracy 92.5% and precision 93.7%). This is evidence of the capability of our method handling individual differences, e.g., different hand size and weight, and time-varying factors, e.g., hand orientation change and low-frequency motion. In another experiment where the 25 everyday objects with different shapes and materials were used, we conducted the experiment with the classifier built on the standard objects. This was to see how the material-based classifier reacts to the everyday objects of different shapes but similar materials. The one-person material-based classifier could find many good matches between the everyday and standard objects. Therefore, depending on objects used, material-based classifiers may suffice for accurate and robust object recognition, without training classifiers every time for new sets of objects.

When very high recognition accuracy is crucial, we can build individual models for each set of objects. Such classifiers are trained on both the materials and shapes of objects. For example, we trained a C-SVM classifier to the 20 participants' data of the 25 everyday objects with a cutoff level $\delta = -60$ dB. The recognition accuracy was as high as 93.1%. Moreover, we tested our method with the four sets of objects shown in Figure 9. They were liquid body products in soft tubes, candies in hard containers, stacked papers, and empty or full spray bottles. The recognition accuracy was nearly 100% for all the four sets. These four sets of objects are very difficult to recognize with other methods using vision or sound.

### Hardware Improvement

Our prototype of VibEye is a bit bulky and not designed optimally for ergonomics; we rather focused on the function for proof of the concept. As such, its design can be improved in many ways by using other actuators and sensors. For example, there exist flexible thin-film vibration actuators that can be attached to the skin [28, 29]. Likewise, commercial film-type pressure sensors (e.g., those from Tekscan, Inc.) have been available for a long time. For accelerometers, tiny ones have recently been developed (approximate size 1 mm $\times$ 1 mm $\times$ 0.8 mm [24]). Using such advanced actuators and sensors is likely to allow for a design of smaller, lighter, and more ergonomic hardware. However, the requirement of vibration transmission and sensing between fingers might still have fundamental interference with the normal use of fingers during interaction. This issue can be addressed more carefully after we build an improved hardware of VibEye.

Though the highest frequency we use is 500 Hz, the vibration output can have higher frequency due to nonlinearity or harmonics (e.g. candies inside a container). Therefore, we set our sampling rate to 20 kHz to ensure reliable measurement. However, the spectrograms in Figure 4 did not show significant energy above 500 Hz. Hence, 5–10 kHz sampling rates seems to be sufficient for spectrogram construction[2].

[2]In practice, sampling rate for digitizing must be higher than 10-20 times of the highest frequency of interest of an analog signal [9]. Otherwise, important information is lost during analog-to-digital conversion.

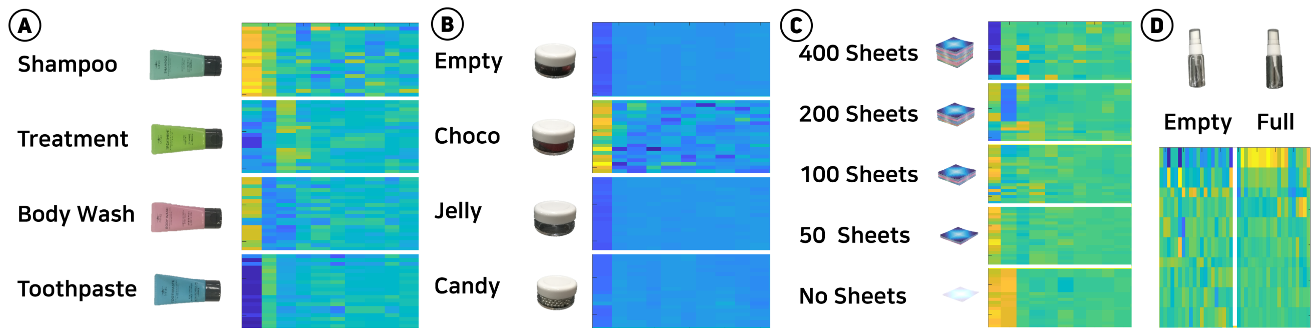


**Figure 9: Another four sets of everyday objects and their PCA results (20 repetitions each): (A) Liquid body products in soft tubes, (B) candies in hard containers, (C) stacked papers, and (D) a spray bottle (empty or full).**

### Vibration as a Sensory Cue

VibEye makes sound and vibration for object recognition, and they are perceptible. For some applications, we can *design* sound and vibration so that users accept them as adequate sensory feedback while the vibration still includes sufficient spectral information for building a spectrogram. There exists a plethora of literature even for vibrotactile effect design (e.g., see reviews in [6, 21, 32]). Our current implementation presents pleasant sensations to users. However, our method is inappropriate for the applications that disallows any sound or vibration other than those resulted from natural contact. It is a fundamental limitation of our approach that leverages structured mechanical vibrations.

## 7 APPLICATIONS

VibEye allows tangible props to self-illustrate computing contexts in a simple and direct manner. The unlocked possibilities enable the design of digital environment that encompasses the afforance and experiences of the real world. VibEye also enables the unique haptic properties of real world objects, leading to rich tangible interaction. Delicate and assorted feels from the objects can be mapped to the functions and characteristics of user interface. Under these visionary design criteria, we have designed and implemented two applications, and they are presented in the rest of this section.

### 3D Modeling Interface for VR

Using blocks for 3D modeling is a popular paradigm (e.g., Minecraft and isometric toolkits in Unity). We can instill tangibility into such block-based modeling by using our 16 standard cubes as props (Figure 10, left). A user can design a 3D model by manipulating the cubes of desired materials with VibEye, and the corresponding virtual blocks attain those materials. This function affords an authoring environment of highly congruent manipulation and representation. The 16 tangible cubes not only behave as haptic proxies [3, 42], but also provide diverse haptic sensations that are copied to the 3D model. Additionally, measuring the user's pinching force enables to distinguish whether the grasping is a mere grip, a request to identify the cube, or a command to insert the corresponding virtual block into the model.

Our 3D modeling application can be used with a tool that transforms multi-block models to CAD models [14]. Then the result can support material-rich 3D printing; its needs have been consistently raised in the literature [2]. Hence, our tangible 3D modeling interface has implication to 3D printing of diverse viscoelastic materials [39] and estimating the responses of elastic objects to be printed [30].

### Drawing Interface for AR

In this application, a user puts on an optical see-through HMD and draws in a 3D space while holding an object in the hand wearing VibEye (Figure 10, right). The hand is tracked by a 3D tracker, and its trajectory is colored using AR. So the user can see both the real environment and the virtual drawing. Objects can be anything, e.g., a pen, sponge, finger, and eraser, but those with natural affordance and salient material properties are more adequate. Such tangible objects not only indicate functions (determined by the object identity), but also the textures of virtual drawing (represented by the object texture). Using this tangible input, we transfer the visual and haptic analog experiences to the digital domain, also enabling other users to see and feel the saved drawings. In particular, depending on the haptic interface used, we can render different haptic properties such as texture, elasticity, and friction. For example, imagine a user holding a smartphone so that the user can perceive the haptic textures of drawing through vibration feedback.

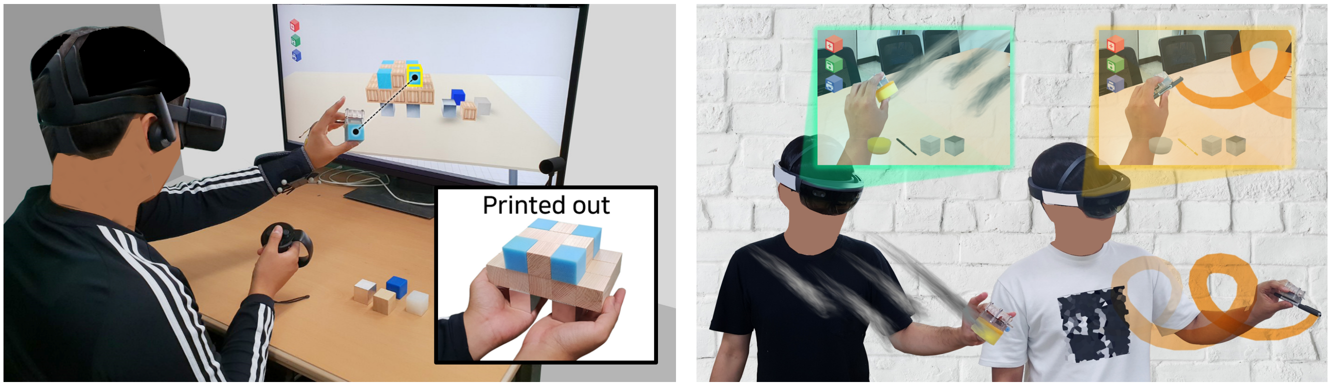


**Figure 10: Applications of VibEye: (left) VR 3D modeling and (right) AR drawing tool.**

## 8 CONCLUSIONS

In this work, we have presented VibEye, a system for vibration-mediated object recognition. VibEye is simple: its hardware requires only a vibration emitter and a sensor, and its software processes the data using well-defined image-based methods. Essentially, VibEye transforms the object recognition problem to an image classification problem. We have validated the effectiveness of VibEye in several ways, using the cross-validation results for the standard objects of the same shape but different materials, and recognition performance for other users' data and other unseen various everyday objects' data. Also demonstrated are the two tangible applications that capitalize on the advantages of VibEye.

We envision tightly-coupled virtual and real environments that are seamlessly controlled by tangible objects. We hope that the concepts embodied by VibEye could pave the way.

## ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF-2017R1A2B4008144) and the ‘‘ICT Consilience Creative program'' (IITP-2018-2011-1-00783) supervised by the IITP(Institute for Information & communications Technology Promotion).

## REFERENCES


[1] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* 2, 3, Article 27 (May 2011), 27 pages. https://doi.org/10.1145/1961189.1961199

[2] Xiang 'Anthony' Chen, Stelian Coros, and Scott E. Hudson. 2018. Medley: A Library of Embeddables to Explore Rich Material Properties for 3D Printed Objects. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 162, 12 pages. https://doi.org/10.1145/3173574.3173736

[3] Lung-Pan Cheng, Eyal Ofek, Christian Holz, Hrvoje Benko, and Andrew D. Wilson. 2017. Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3718–3728. https://doi.org/10.1145/3025453.3025753

[4] Jungchan Cho, Inhwan Hwang, and Songhwai Oh. 2016. VibePhone: efficient surface recognition for smartphones using vibration. *Pattern Analysis and Applications* 19, 1 (feb 2016), 251–265. https://doi.org/10.1007/s10044-015-0460-8

[5] Youngjun Cho, Nadia Bianchi-Berthouze, Nicolai Marquardt, and Simon J. Julier. 2018. Deep Thermal Imaging: Proximate Material Type Recognition in the Wild Through Deep Learning of Spatial Surface Temperature Patterns. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 2, 13 pages. https://doi.org/10.1145/3173574.3173576

[6] Seungmoon Choi and Katherine J. Kuchenbecker. 2013. Vibrotactile Display: Perception, Technology, and Applications. *Proc. IEEE* 101, 9 (2013), 2093–2104. https://doi.org/10.1109/JPROC.2012.2221071

[7] Waltenegus Dargie. 2009. Analysis of Time and Frequency Domain Features of Accelerometer Measurements. In *Proceedings of 18th International Conference on Computer Communications and Networks (ICCCN'09)*. IEEE, 1–6. https://doi.org/10.1109/ICCCN.2009.5235366

[8] Mingming Fan and Khai N. Truong. 2015. SoQr: Sonically Quantifying the Content Level Inside Containers. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 3–14. https://doi.org/10.1145/2750858.2804264

[9] Gene F. Franklin, J. David Powell, and Michael L. Workman. 1990. *Digital Control of Dynamics Systems* (second ed.). Addison-Wesley Publishing Company, Reading, MA, U.S.A.

[10] Nobuhiro Funato and Kentaro Takemura. 2017. Grip Force Estimation by Emitting Vibration. *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17* October, 141–142. https://doi.org/10.1145/3131785.3131829

[11] Chris Harrison and Scott E. Hudson. 2008. Lightweight Material Detection for Placement-aware Mobile Computing. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST '08)*. ACM, New York, NY, USA, 279–282. https://doi.org/10.1145/1449715.1449761

[12] Hiroyuki Kato and Kentaro Takemura. 2016. Hand Pose Estimation Based on Active Bone-conducted Sound Sensing. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 109–112. https://doi.org/10.1145/2968219.2971403

[13] Kai Kunze and Paul Lukowicz. 2007. Symbolic Object Localization Through Active Sampling of Acceleration and Sound Signatures. In

*Proceedings of the 9th International Conference on Ubiquitous Computing (UbiComp '07)*. Springer-Verlag, Berlin, Heidelberg, 163–180. http://dl.acm.org/citation.cfm?id=1771592.1771602

[14] Ben Lafreniere and Tovi Grossman. 2018. Blocks-to-CAD : A Cross-Application Bridge from Minecraft to 3D Modeling. In *Proceedings of the 31th Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. ACM, New York, NY, USA.

[15] Nils Landin, Joseph M. Romano, William McMahan, and Katherine J. Kuchenbecker. 2010. Dimensional reduction of high-frequency accelerations for haptic rendering. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 6192 LNCS. Springer, Berlin, Heidelberg, 79–86. https://doi.org/10.1007/978-3-642-14075-4_12

[16] Gierad Laput, Eric Brockmeyer, Scott E. Hudson, and Chris Harrison. 2015. Acoustruments: Passive, Acoustically-Driven, Interactive Controls for Handheld Devices. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2161–2170. https://doi.org/10.1145/2702123.2702414

[17] Gierad Laput, Xiang 'Anthony' Chen, and Chris Harrison. 2016. SweepSense: Ad Hoc Configuration Sensing Using Reflected Swept-Frequency Ultrasonics. In *Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI '16)*. ACM, New York, NY, USA, 332–335. https://doi.org/10.1145/2856767.2856812

[18] Gierad Laput, Robert Xiao, and Chris Harrison. 2016. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 321–333. https://doi.org/10.1145/2984511.2984582

[19] Gierad Laput, Chouchang Yang, Robert Xiao, Alanson Sample, and Chris Harrison. 2015. EM-Sense: Touch Recognition of Uninstrumented, Electrical and Electromechanical Objects. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 157–166. https://doi.org/10.1145/2807442.2807481

[20] Lennart Ljung. 1999. *System Identification: Theory for the User* (second ed.). Prentice Hall, Upper Saddle River, NJ, USA.

[21] Karen E. MacLean. 2008. Foundations of Transparency in Tactile Information Design. *IEEE Transactions on Haptics* 1, 2 (2008), 84–95. https://doi.org/10.1109/TOH.2008.20

[22] Takuya Maekawa, Yasue Kishino, Yasushi Sakurai, and Takayuki Suyama. 2013. Activity Recognition with Hand-worn Magnetic Sensors. *Personal Ubiquitous Comput.* 17, 6 (Aug. 2013), 1085–1094. https://doi.org/10.1007/s00779-012-0556-8

[23] Kenneth A Mann, Frederick W Wernere, and Andrew K Palmer. 1989. Frequency spectrum analysis of wrist motion for activities of daily living. *Journal of Orthopaedic research* 7, 2 (1989), 304–306. https://doi.org/10.1002/jor.1100070219

[24] EPS News. 2015. mCube Intros Tiny MEMS Accelerometer. (Nov. 2015). Retrieved January 7, 2019 from https://epsnews.com/2015/11/04/mcube-intros-tiny-mems-accelerometer/

[25] Katsunori Ohnishi, Atsushi Kanehira, Asako Kanezaki, and Tatsuya Harada. 2016. Recognizing Activities of Daily Living with a Wrist-Mounted Camera. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 3103–3111. https://doi.org/10.1109/CVPR.2016.338

[26] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. 2013. Touch & Activate: Adding Interactivity to Existing Objects Using Active Acoustic Sensing. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. ACM, New York, NY, USA, 31–40. https://doi.org/10.1145/2501988.2501989

[27] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. 2015. Sensing Touch Force Using Active Acoustic Sensing. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. ACM, New York, NY, USA, 355–358. https://doi.org/10.1145/2677199.2680585

[28] Won-Hyeong Park, Eun-Jae Shin, Sungryul Yun, and Sang-Youn Kim. 2018. An Enhanced Soft Vibrotactile Actuator Based on ePVC Gel with Silicon Dioxide Nanoparticles. 11, 1 (2018), 22–29. https://doi.org/10.1109/TOH.2018.2808176

[29] Won-Hyeong Park, Tae-Heon Yang, Yongjae Yoo, Seungmoon Choi, and Sang-Youn Kim. 2015. Flexible and Bendable Vibrotactile Actuator Using Electro-conductive Polyurethane. In *Proceedings of the IEEE World Haptics (WHC '15)*. IEEE, 165–170. https://doi.org/10.1109/WHC.2015.7177708

[30] Michal Piovarči, David I. W. Levin, Jason Rebello, Desai Chen, Roman Ďurikovič, Hanspeter Pfister, Wojciech Matusik, and Piotr Didyk. 2016. An Interaction-aware, Perceptual Model for Non-linear Elastic Objects. *ACM Trans. Graph.* 35, 4, Article 55 (July 2016), 13 pages. https://doi.org/10.1145/2897824.2925885

[31] Munehiko Sato, Shigeo Yoshida, Alex Olwal, Boxin Shi, Atsushi Hiyama, Tomohiro Tanikawa, Michitaka Hirose, and Ramesh Raskar. 2015. 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 (CHI '15)*. ACM, New York, NY, USA, 2191–2200. https://doi.org/10.1145/2702123.2702169

[32] Oliver Schneider, Karon MacLean, Colin Swindell, and Kellogg Booth. 2017. Haptic Experience Design: What Hapticians Do and Where They Need Help. *International Journal of Human-Computer Studies* 107 (2017), 5–21. https://doi.org/10.1016/j.ijhcs.2017.04.004

[33] Yu-Chih Tung and Kang G. Shin. 2015. EchoTag: Accurate Infrastructure-Free Indoor Location Tagging with Smartphones. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 525–536. https://doi.org/10.1145/2789168.2790102

[34] M.A. Turk and A.P. Pentland. 1991. Face recognition using eigenfaces. In *Proceedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '91)*. IEEE Comput. Sco. Press, 586–591. https://doi.org/10.1109/CVPR.1991.139758

[35] Edward J. Wang, Tien-Jui Lee, Alex Mariakakis, Mayank Goel, Sidhant Gupta, and Shwetak N. Patel. 2015. 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 (UbiComp '15)*. ACM, New York, NY, USA, 15–26. https://doi.org/10.1145/2750858.2804271

[36] Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic Finger: Always-available Input Through Finger Instrumentation. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 147–156. https://doi.org/10.1145/2380116.2380137

[37] Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. 2016. RadarCat: Radar Categorization for Input & Interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 833–841. https://doi.org/10.1145/2984511.2984515

[38] Hui-Shyong Yeo, Juyoung Lee, Andrea Bianchi, David Harris-Birtill, and Aaron Quigley. 2017. SpeCam: Sensing Surface Color and Material with the Front-facing Camera of a Mobile Device. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, New York, NY, USA, Article 25, 9 pages. https://doi.org/10.1145/3098279.3098541

[39] Jonas Zehnder, Espen Knoop, Moritz Bächer, and Bernhard Thomaszewski. 2017. Metasilicone: Design and Fabrication of Composite Silicone with Desired Mechanical Properties. *ACM Trans. Graph.* 36,

6, Article 240 (Nov. 2017), 13 pages. https://doi.org/10.1145/3130800.3130881

[40] Cheng Zhang, Qiuyue Xue, Anandghan Waghmare, Ruichen Meng, Sumeet Jain, Yizeng Han, Xinyu Li, Kenneth Cunefare, Thomas Ploetz, Thad Starner, Omer Inan, and Gregory D. Abowd. 2018. FingerPing: Recognizing Fine-grained Hand Poses Using Active Acoustic On-body Sensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 437, 10 pages. https://doi.org/10.1145/3173574.3174011

[41] Yang Zhang, Gierad Laput, and Chris Harrison. 2018. Vibrosight: Long-Range Vibrometry for Smart Environment Sensing. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. ACM, New York, NY, USA, 225–236. https://doi.org/10.1145/3242587.3242608

[42] Yiwei Zhao, Lawrence H. Kim, Ye Wang, Mathieu Le Goc, and Sean Follmer. 2017. Robotic Assembly of Haptic Proxy Objects for Tangible Interaction and Virtual Reality. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, New York, NY, USA, 82–91. https://doi.org/10.1145/3132272.3134143

[43] Yiran Zhao, Shuochao Yao, Shen Li, Shaohan Hu, Huajie Shao, and Tarek F. Abdelzaher. 2017. VibeBin: A Vibration-Based Waste Bin Level Detection System. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 122, 22 pages. https://doi.org/10.1145/3132027