

Acoustruments: Passive, Acoustically-Driven, Interactive Controls for Handheld Devices

Gierad Laput^{1,2} Eric Brockmeyer² Scott E. Hudson^{1,2} Chris Harrison^{1,2}

¹Carnegie Mellon University HCI Institute and ²Disney Research Pittsburgh

{gierad.laput, scott.hudson, chris.harrison}@cs.cmu.edu

eric.brockmeyer@disneyresearch.com

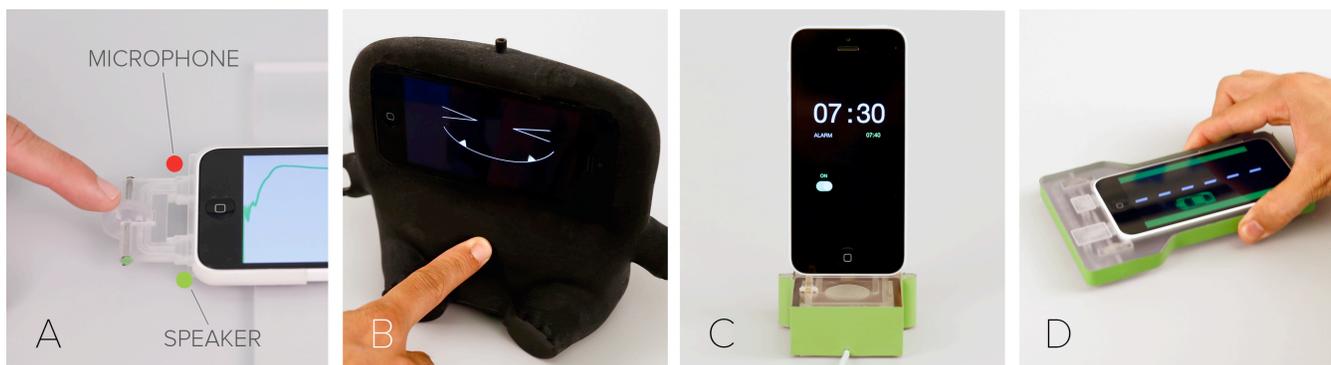


Figure 1. Like musical instruments, we introduce structural elements along the speaker-microphone pathway to characteristically alter the acoustic output. Using an expansive vocabulary of design primitives, we construct physical mechanisms (A), and on top of that, build end-user applications such as an interactive doll (B), an alarm clock (C), or an interactive toy car (D).

ABSTRACT

We introduce *Acoustruments*: low-cost, passive, and powerless mechanisms, made from plastic, that can bring rich, tangible functionality to handheld devices. Through a structured exploration, we identified an expansive vocabulary of design primitives, providing building blocks for the construction of tangible interfaces utilizing smartphones' existing audio functionality. By combining design primitives, familiar physical mechanisms can all be constructed from passive elements. On top of these, we can create end-user applications with rich, tangible interactive functionalities. Our experiments show that *Acoustruments* can achieve 99% accuracy with minimal training, is robust to noise, and can be rapidly prototyped. *Acoustruments* adds a new method to the toolbox HCI practitioners and researchers can draw upon, while introducing a cheap and passive method for adding interactive controls to consumer products.

Author Keywords

Acoustic sensing; mechanisms and controls; fabrication; mobile and handheld devices;

ACM Classification Keywords

H.5.2. [Information interfaces and presentation]: User interfaces – Input devices and strategies.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CHI 2015, April 18 - 23 2015, Seoul, Republic of Korea
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3145-6/15/04 \$15. <http://dx.doi.org/10.1145/2702123.2702416>

INTRODUCTION

Smartphones and other handheld devices are increasingly being employed in interactive applications that extend beyond their conventional touchscreens. For example, tangibles allow users to interact with mobile devices using physical objects both on the screen and around the device. Similarly, there is a growing class of auxiliary devices that require a smartphone to be plugged in or docked, transforming an otherwise simple object into something with rich and dynamic interactivity, as well as wireless connectivity if needed. These include appliances, such as alarm clocks and speakers, to children's toys, audio mix decks, and even robots (see e.g., [48]). The latter devices can be made less costly by relying on the "smarts" from an expensive, general-purpose computing device.

However, these auxiliary devices still require numerous components, including mechanical mechanisms, wires, ICs, PCBs and sometimes batteries. This dramatically increases manufacturing cost and reduces physical robustness. In this work, we introduce *Acoustruments*: low-cost, passive, and powerless mechanisms, made from plastic, that can bring rich, tangible functionality to handheld devices. The operational principles were inspired by wind instruments (e.g., slide whistles, ocarinas and flutes), which produce expressive musical output despite a simple physical design. In general, a sustained source of sound is injected into one end and various physical elements are altered to produce distinctive outputs. For example, a flute has a series of holes along its main axis that can be covered, while a trombone varies its pitch by altering the size of an in-lined cavity.

In a similar manner, we can create an “instrument” that attaches to a smartphone. Specifically, one end of an enclosed tube is connected to the speaker, which emits a continuous ultrasonic sweep (Figure 2). The other end is directed into the smartphone’s microphone, which can monitor the output. Like musical instruments, we can introduce structural elements along this pathway that can characteristically alter the acoustic properties of the output, which in our case, we use for interactive control.

Through a structured exploration, we identified an expansive vocabulary of design primitives, providing the building blocks for the construction of tangible interfaces utilizing smartphones’ existing audio functionality. By combining design primitives, familiar physical controls, such as knobs, valves, rotary encoders, and sliders, can all be constructed from passive elements. On top of these, we can create end-user applications, ranging from games to appliances. Using technologies like 3D printing, we show that rich physical controls can be rapidly prototyped, adding a new method to the toolbox HCI practitioners and researchers can draw upon. In addition, through mass production techniques (e.g., injection molding and training on a master mold), Acoustruments introduces a cheap and passive method for the construction of consumer “pluggables.”

RELATED WORK

Our work touches upon several areas including tangible interfaces for mobile devices, fabrication and rapid prototyping, 3D printing, and mobile sensing. We now summarize key related work in these respective domains.

On-Screen and Around-Device Tangibles

Researchers have explored techniques for augmenting interactions on desktops, handhelds, and tabletops with external physical controls using several sensing and fabrication techniques. Huwang et. al. [21] and Liang [25] utilize magnetically-driven tangibles as passive controls for devices equipped with a magnetometer. Capacitively-modulated widgets, such as CapStones and Zebra Widgets [9], Slap-Widgets [46], and CapWidgets [23], utilize the device’s touchscreen for on-screen tangible interaction. Similarly, camera-based systems [3,44] and fabricated optical elements offer configurable interactive physical controls: in Lumino 3 [3], Baudisch et. al. augment tabletop interaction with fiducially-marked “blocks” assembled using glass fibers; Willis [47] and Brockmeyer et. al. [7] leverage 3D-printed optical elements for interactive sensing, and display.

3D-printing of Musical instruments

3D-printed musical instruments, earlier thought to be unwieldy to fabricate, are now beginning to emerge. For instance, Thingiverse [49] currently hosts a large collection of user-designed 3D-printable models of instruments, including whistles, ocarinas and acoustic guitars. Similarly, after building a series of 3D-printed guitars, Olaf Diegel [50] has constructed one of the earliest 3D-printed saxophones. The

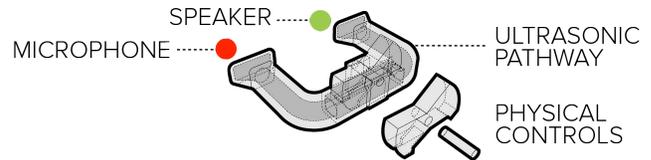


Figure 2. Physical controls manipulate the traversing ultrasonic signal from the speaker to the microphone.

level of control and customization from CNC-based techniques offer a novel approach for constructing tangible acoustic interfaces [10], whose applications can be extended to interactive physical controls.

Rapid Fabrication and Prototyping

Researchers have investigated techniques for rapidly prototyping interactive physical objects across numerous contexts. In Phidgets [16], Greenberg proposed a hardware and software ecosystem of “physical widgets” inspired by real-world mechanisms. Printed Optics utilizes light pipes to add sensing and display capabilities to 3D-printed objects [47], and Hudson [20] introduces techniques for embedding conductive thread and electronics in objects printed on felt material. Techniques for augmenting inanimate prototypes with interactive capabilities have also been explored, including wireless input components [2], mounted capacitive pads [38], and cameras [36]. These approaches complement techniques for rapidly prototyping object structure [26,27]. Savage et. al. [37] leveraged routed pipes as a *post hoc* process for embedding input and output functionalities to 3D printed objects, and provide a tool for automatic routing. Although we utilize related primitives—pipes and tubes—our technique does not require additional electronics for input control.

“Pluggables” and Smartphone-Driven Devices

There is an emerging class of use-cases where mobile devices are being “plugged-in” into objects and environments, provisioning the “smarts” that augment their capabilities. As such, smartphones transform an otherwise simple or static object into something with rich interactive functionality. Although researchers have traditionally leveraged mobile devices as data acquisition interfaces [24,34], “pluggables” have seen increasing interest in consumer applications. They can be simple products such as alarm clocks, speaker boxes, dolls, and children’s toys (apptoyz.com), or sophisticated tools such as card-readers (squareup.com), appliances (thefusionphone.com), DJ mixers, and robots [48]. However, these pluggables still require extra components and electronics, which increase cost and complexity.

Related Sensing Techniques

Modal Analysis [5] and Swept Frequency Acoustic Interferometry (SFAI) are well known in fields such as materials [39] and structural analysis [15]. Acoustic sensing has been used in many HCI pursuits, and approaches can be grouped into three major categories. Acoustruments makes use of fingerprint-based techniques, which employ an uncon-

trolled, but characteristic signal to distinguish between states. Identification-based techniques can be active (i.e., emit a signal), as in the case of Acoustruments and Touch & Activate [30] (see also [28,40]) or passive (i.e., receive a signal), such as Scratch Input [18], Skinput [19], and BLUI [31]. Another, more structured approach is to use acoustic time difference of arrival (TDoA) or time of flight (ToF), most commonly encountered in the form of “sonar”, which has seen extensive use in the HCI literature [22,32,41]. Finally, sensing acoustic Doppler shifts has also proven useful, as shown in SoundWave [17] and Doplink [1] (see also [33,45] for an overview).

IMPLEMENTATION

As noted earlier, our approach explicitly routes an enclosed, *pipe*-like pathway from the speaker to the microphone, enabling direct transmission of high-fidelity ultrasonic signals (Figure 1A). Sound reaching the microphone is altered by various physical elements along the path. That is, the enclosure alters the traversing acoustic signal depending on its shape, material, and other physical conditions. Specifically, carefully designed cavities, holes, and partial obstructions produce distinct spectral changes in the final signal.

For reference, our prototype uses an iPhone 5C, with physical dimensions 124.4mm x 59.2mm x 8.97mm, audio input/output sampling rate of 44.1kHz, and 1.3GHz dual-core Apple A6 processor. We utilize the collocated speaker and microphone at the bottom of the handset, with a physical separation of approximately 26.2mm.

Fabrication

Our fabrication approach was motivated by two goals. First, we wanted to be able to create pipes that would transmit sound from the speaker to the microphone with as little unintended loss as possible. In addition, and more importantly, we wanted to explore how the internals of the pipes affect acoustic response. Although we used 3D printing as a means for rapidly iterating on our ideas, it is important to note that the same structures can be manufactured using high-volume, low-cost fabrication approaches, most notably injection molding, but also casting and machining.

We constructed our prototypes using an Objet260 Connex 3D printer, using VeroClear-RGD810 UV-cured photopolymer. This printer is capable printing multiple materials, and we took advantage of this capability in two instances (see next section). For the most part, our structures were printed as a single material. We used Rhino and its Grasshopper visual language as our authoring software, enabling parameterized modeling of expansive pipe configurations.

One of the most onerous steps in our fabrication process was support material removal. Our 3D printer scaffolds support material on regions with sufficient overhang and, in most cases, it can be removed through a waterjet process i.e., dissolving support material by directing a high-pressure water stream, much like a fire hose. Further, remaining

support material is soaked into an ultrasonic cleaner with a sodium hydroxide aqueous solution. However, for more sophisticated configurations such as pipes with extended branching, these methods are insufficient. To minimize support material removal, we designed our models by making them modular (i.e., detachable parts), or by making escape holes to provide the waterjet and the ultrasonic cleaner with more surface area to pass through. In extreme cases, we avoided using support materials altogether by bisecting our models and re-assembling them post-hoc.

Swept-Frequency Acoustic Sensing

Our sensing technique emits rapid ultrasonic frequency sweeps while simultaneously recording sound. As the signal traverses through the speaker-microphone pathway, certain frequencies resonate or are absorbed, creating a unique spectral signature. We emit a continuous 100ms linear sweep from 16.50kHz to 22.05kHz. We experimented with different variants of sweep lengths and interpolations (e.g., 40ms logarithmic sweeps), and can say anecdotally that linear sweeps between 50ms and 100ms seem to provide the right balance between latency and signal fidelity. We note that these values should be treated as a lower bound, since the sweep rate and frequency range can be drastically increased as the sampling rate for handheld devices continues to improve in the future.

Incoming audio is recorded as linear PCM, and stored into a circular buffer. The captured signal is transformed into the frequency domain using a Fast Fourier Transform (FFT). In turn, spectrum bands higher than 15.50kHz are saved, and all values in this range are downsampled into 100 bins. In addition to raw FFT, we also compute the first and second derivatives of the downsampled signal, along with spectral band ratios, max index, min index, and center of mass. These features are passed to a support vector machine (SVM) optimized for either classification or regression for discrete and continuous values, respectively.

We forked a variant of libSVM [51] and made it run on iOS, which enabled real-time training and testing directly from the mobile device. For classification tasks, we trained a Sequential Minimal Optimization-based Support Vector Machine (SMO SVM) using default optimization parameters (kernel=polynomial, C=1000, gamma=0, e=0.001). For regression, we optimized a nu-Support Vector Regression model (nu-SVR) using an RBF kernel, along with optimization parameters identical to above. Of note, the system needs to train the models before it can perform real-time classification and regression.

Our technique shares similarities to the touch sensing principle embodied in Touch & Activate [30], which was inspired by Touché [35]. However, our approach utilizes commodity sensors and does not require additional electronics. Further, all computation is performed on the device, which means our approach can be deployed to existing devices as a downloadable application or software update.

IN-LINE DESIGN PRIMITIVES

Physical configurations along the speaker-microphone pathway influence the frequency response properties of the transmitted acoustic signal. In order to develop a vocabulary of discriminable in-line design primitives for this purpose, we first looked at related literature in both instrument design [8,12,13,42], and also more generally, the physics of sound [4,6,11,14]. We then iteratively prototyped, testing hundreds of single purpose prototypes to better understand how to control and isolate the acoustic effects. This process enabled us to build a comprehensive vocabulary of physical primitives and associated design recommendations that characteristically alter the acoustic properties of sound propagating through a controlled pathway.

Holes

When a flutist applies a jet stream of input air across the mouthpiece, the column of air vibrates with respect to the configuration of keyholes along its body. The human vocal system functions similarly (e.g., try speaking while covering your nostrils). Simply put, when holes are closed or opened, they indirectly modify the volume of air the sound effectively resonates within, thereby characteristically altering frequencies of the output sound. In a similar fashion, holes along the speaker-microphone pathway influence the acoustic properties of the traversing ultrasonic signal. Like those in a flute, holes offer a rich palette for input control.



Figure 3. Tubes with two (top) and five holes, along with their respective ultrasonic responses between 16kHz to 22.05kHz.

Our experiments revealed that hole diameters between 3.0mm and 7.0mm affect the acoustic signal more discriminatively (for reference, these holes were fabricated on a 3D-printed pipe with a rectangular pipe (3.5mm by 9.0mm), extending 88mm from the speaker to the microphone). Further, diameters less than 3.0mm had no significant effect. Additionally, distance between holes influences the resonance of the acoustic signal, and our experiments show that holes should be at least 7.5mm apart in order for them to render significant effects. This distance is also effective at preventing the fat-finger problem by appropriate spacing of the openings. Finally, there is an upper bound on the number of holes that can be introduced along the pipe; in our experiments, more than 5 holes reduced the ability for any one hole to sufficiently affect the acoustic signal, which we suspect is due to energy loss from majority of the signal escaping into open air.

Tube Length

Similar to holes, changing the tube length affects the overall volume of vibrating air along the pathway. For example, in trombones, lengthening its main slide expands the air column, thereby producing lower frequency notes. Similarly, adjusting the tube length between the speaker and the microphone distinctively affects the acoustic properties of the traversing signal. As tube length increases, the strength of the acoustic signal reaching the microphone experiences characteristic attenuation. Tubes can be turned or twisted to accommodate for limitations on physical space. From our experiments, we learned that the number of turns does not significantly affect the acoustic properties of the signal, as long as the overall length of the tube remains constant.



Figure 4. Short (top) and long tubes.

Blockages

Full or partial blockages in the pathway disturb airflow, which in turn alters the properties of the resulting output signal. As a blockage narrows the aperture along the pipe, it increases the speed of air motion. Of note, a fully blocked pathway does not necessarily prohibit the ultrasonic signal from reaching the microphone; the signal still resonates across the enclosure, but its acoustic profile is markedly distinct relative to an unobstructed signal.



Figure 5. Open (top) and blocked tubes.

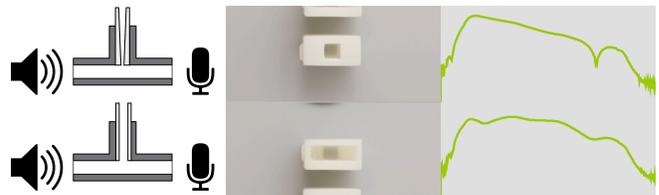


Figure 6. Insertable blocks with narrow and wide cavities.

Cavities

Cavities are closed volumetric shapes along the pipe that alter acoustic resonance in distinctive ways. Its shape—such as a hemisphere, a tapered cylinder, or an irregular hollow chamber—produces characteristic acoustic resonance, much like the phenomena embodied in a *Helmholtz*

resonator [43]. Cavities can also be made soft, in which case, they become squeezable, allowing cavity properties to be directly changed by user input.

Tube Softness

Modifying the tube material also affects the acoustic profile of the signal. More precisely, soft tubing e.g., latex rubber is more flexible, which means it can be bent, squeezed, or pinched at several segments along its axis, thereby modifying how acoustic signals propagate along the pathway.



Figure 7. An example of soft tubing using latex material.

Tube Gap

When a physical discontinuity is introduced along a tube (Figure 8), the enclosed vibrating air column is forced to disperse out from one end of the aperture, into open air, and back into the continuing end. Unlike holes, gaps externally diffuse sound energy, and the traversing signal reaching the microphone suffers from distortion. Additionally, if gaps are wide enough, physical objects can be used to block or alter the flow of the diffused signal (e.g., a piece of soft cloth or a finger along the path), allowing for interesting input control modalities.



Figure 8. Physical gaps along the speaker-microphone pathway allow the audio signal to disperse into open air.

Wall Thickness

A tube's interior thickness affects how sound waves reflect and travel internally. In general, tubes with thicker walls tend to be more rigid and reflect sound with less dampening. In our experiments, the range of variations we explored (thicknesses between 0.5mm and 2.0mm) did not produce acoustic differences large enough to be highly discriminatory. Thus, we selected a wall thickness of 1.0mm, which was economical and provided sufficient structural durability.

Tube Diameter

Airflow oscillating through a tube is directly related to its diameter, which in turn affects its acoustical output (Figure 9). Tighter tubes make the conduit resonate at higher frequencies. We built majority of our models using tubes with a diameter of 3.5mm. Although other variations were explored, in general, narrow tubes of this size offered sufficient sensitivity and were conveniently space-efficient.

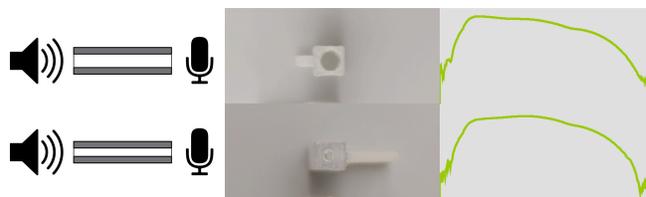


Figure 9. Cross-section of tubes with different diameters.

Structured Filters

It is possible to characteristically attenuate the traversing acoustic signal by constructing tubes such that they function as basic filters (e.g., a low-pass filter). Figure 10 depicts the design of a primitive band pass filter which was derived from Munjal's [29] exploration of the acoustics of ducts and mufflers. In this example, the filter's center frequency is controlled by the distance between two *baffles* (Figure 10, bottom left), which is half that frequency's wavelength.



Figure 10. Engineered baffles along the tubular pathway (bottom left) serve as primitive physical acoustic filters. Here, baffle distance affects the filter's center frequency.

Tube Interior Texture

The texture of a tube's interior characteristically dampens and excites certain frequency components of the traversing acoustic signal. They are a superset to the structured filters mentioned earlier, but far less constrained. Textures can range from simple symmetric elements—flat, sawtooth ridges, or sinusoidal contours—to more irregular shapes, like ruffles, waves, or diminutive bristles.

COMBINING PRIMITIVES

Although in-line design primitives by themselves offer enough expressivity for rich input control, they can also be combined into various configurations, providing even greater possibilities. In this section, we discuss several techniques for combining two or more primitives along the conduit between the speaker and the microphone, all while preserving maximum signal discriminability.

Parallel Routing

Two or more enclosed tubes can run in parallel across the speaker and the microphone, effectively enabling several primitives to affect the acoustic properties of the swept frequency signal. For example, in Figure 12 two parallel tubes run between the speaker and the microphone, controlled by two valves. The first tube is longer than the second, with holes placed along its path (as part of the valve). In this specific multi-tube combination, two design primitives—



Figure 12. Primitives can be combined through parallel routing. Here, two combined valves offer increased input.

tube length and holes— are being varied, producing several combinatorial states for input control.

Branching

Tubes can branch into two or more paths. As such, this configuration splits the traversing signal into multiple routes, thereby strategically isolating the effect of in-line primitives. Other configurations are possible, e.g., merging forked paths into a single conduit (Figure 11B), or branching paths into “dead-end” outlets (e.g., Figure 11A).

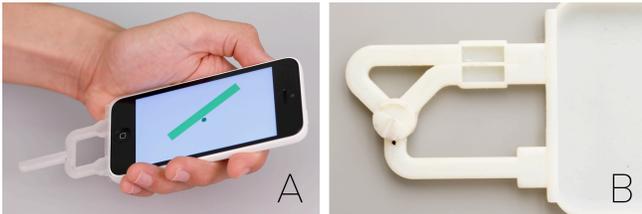


Figure 11. Branching Examples. They can terminate, as seen in this tilt sensor (A), or they can merge back into a single acoustical conduit (B).

Modular Insertions

It is possible to integrate in-line primitives into a conduit, *post-hoc*. Specifically, “generic tubes” can be fabricated with slots that can accommodate “insertable” modules (Figure 13), and in this fashion, components can be mixed and re-used. This introduces interesting combinations, and it de-couples the design of tubes and primitives into modular parts. This technique also provides a mechanism for identifying one of several different objects, which could be inserted by the user.

MECHANISMS

In-line design primitives form the building blocks for creating acoustically-driven interactive elements. They can be used, and in some cases combined, to create familiar human interface controls or even sensors. As a proof of concept, we selected nine illustrative examples, discussed below and depicted in Figure 14.

Rocker Switch

A rocker switch is single-pole single-throw switching mechanism supporting “ON” or “OFF” states (e.g., a light switch), and it leverages a *blockage* in-line primitive (Figure 2). To construct a rocker switch, we place a hole along the tube’s main axis and attached a lever element that



Figure 13. Modular insertions (left and middle) allow primitives to be easily integrated into a pathway *ad-hoc* (right).

pivots from the hole’s midpoint. When one side of the lever is pushed inward (OFF), the output signal is blocked. Conversely, when the opposite side is pushed, the acoustic signal is allowed to pass through (ON).

Ball Valve

A ball valve functions as a binary switch, which supports a turning actuation mechanism. Valves in our case function much like ball valves in water pipes; they are either positioned “ON” or “OFF.” Figure 14 illustrates the mechanism behind our ball valve. When in the “ON” position, a hole on the valve shaft aligns with the tube carrying the acoustic signal, permitting sound to pass through.

Multi-Turn Knob

Multi-turn knobs are identical to single-handle spigots. An adjustable knob controls fluid flow (e.g., water, air), and in our case, it restricts how much of the audio signal can pass through its constriction point. Our multi-turn knobs use a rotatable threaded bolt and a matching nut affixed to a tube. In our prototype, we constructed our threads with a pitch spacing of 1.3 mm, permitting $3\frac{1}{2}$ full rotations. Unlike switches and valves, knob actuation is more appropriately represented as a range of continuous values, rather than discrete states, and so we employed regression-based machine learning to model this behavior.

Slider

Sliders utilize *tube length* as an in-line design primitive. We model our slider mechanism on a slide whistle; when a piston moves along a tube, it adjusts the air cylinder’s length, thereby characteristically altering its output sound. Similarly, as seen in Figure 14, we fork a terminating *branch* (i.e., a dead-end) along the main conduit, and we insert a cylindrical shaft that slides along that branch’s main axis. Fitted rails help constrain the shaft’s min and max position, and its movement is modeled as a continuous value.

Rotary Encoder

A rotary encoder mechanism senses the changes in angular position of a rotating body. We construct a discrete-step rotary encoder by leveraging multiple *parallel* tubes with *blockage* primitives. Specifically, we fabricated a wheel element that rotates along the major axis of tubes routed in parallel (e.g., two parallel tubes, Figure 1A and 14). We then precisely position holes on the wheel; as it rotates, the blockage mechanisms encoded on the wheel permit one or more tubes to participate in altering the acoustic signal. Conveniently, this scheme offers expansive combinatorial states, which would be more than enough for approximating kinematic values such as rotational speed and direction.

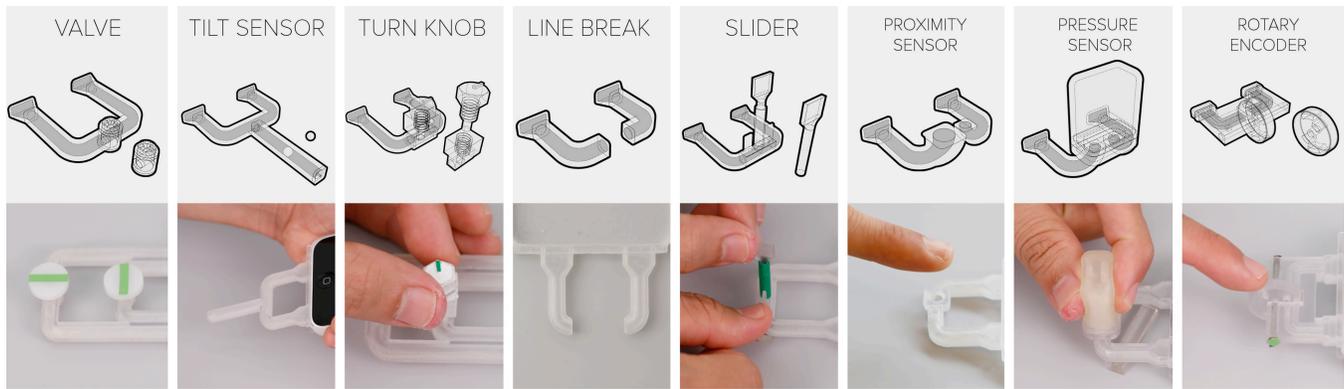


Figure 14. Using design primitives, we can construct several mechanisms for interactive physical control. Examples include valves, tilt sensors, knobs, line break sensor, sliders, switches, pressure sensors, proximity sensors, and rotary encoders.

Proximity Sensor

We leverage the *tube gap* in-line primitive to construct a proximity sensor. Specifically, we introduce a single gap along the tube—creating two apertures funneled outward—enabling sound to disperse into open air. When a finger or an object approaches the gap’s aperture, portions of the sound energy reflect and enter back into the other end of the gap. The strength of the acoustic signal due to proximate reflections can be utilized to model continuous values for proximity sensing.

Line Break Sensor

Like proximity sensors, we utilize *tube gap* as in-line primitive for constructing a line break sensing mechanism. However, instead of the apertures directing outward, the apertures in a line break sensor face each other. In a line break sensor, the mechanism is either “ON” or “OFF.” When “ON”, sound disperses from one aperture, flows into open air, and funnels primarily back into the other end of the gap. When an object or a finger is placed between the apertures, the traversing sound is *blocked*—an indication that the connection is “OFF.”

Tilt Sensor

We built a tilt sensor by embedding a ball that rolls along an elongated branch. When tilted, gravity forces the ball to slide towards one end of the branch, altering airflow and thereby creating a distinctive acoustic output. Our tilt sensor can detect two discrete states (i.e., tilt forward or backward), but can also be utilized to sense continuous values when trained accordingly.

Pressure Sensor

To construct a pressure sensor, we employ an in-line *cavity* fabricated with *soft material*. When pressure is applied (e.g., via squeezing), the cavity undergoes a distinct deformation, thereby altering the acoustic resonance of the output sound. Our pressure sensor can detect bilateral deformations, and we model its behavior as a mechanism with continuous values.

EVALUATION

As seen in our video figure, sensing is robust and rapid across a range of sensing types and combinations. However, as further validation, we empirically tested the performance of four mechanisms of varying complexity – two controls and two sensors. Further, within each category, we use one example of discrete and one example of continuous sensing.

Training

Before any experiments could be run, a classifier first had to be trained. For the two binary sensing examples, 25 instances of each class were captured and used for training. An SMO-based SVM was used for training and classification (via libSVM ported on iOS with default parameters, see Implementation for details). For the two continuous sensing examples, we used a nu-Support Vector Regression model (nu-SVR, RBF kernel), which produces continuous output. It is impossible to train on every possible value, so instead the full sensor range was divided into five equal parts for capturing training data, with ten training instances acquired for each. In both the discrete and continuous cases, only 50 instances were used for model training – a small number by machine learning standards (but as we will see, sufficient for accurate sensing). Lastly, the tests below were evaluated using live prediction (i.e., no post hoc analysis).

Accuracy

First, we tested a simple binary button (Figure 2 mechanism). We ran 40 trials (20 trials per ‘ON’ or ‘OFF’ condition), and collected five data samples per trial, generating 200 sample points across all trials. To evaluate accuracy, we turned the rocker switch on or off according to a target configuration requested by our evaluation app (e.g., ‘switch ON’). Thereafter, we record the system’s prediction. Across 200 samples, the mean accuracy was 99.50% (SD=0.07%).

Increasing in complexity, we next tested the rotational accuracy of a turn knob. Like its real-world counterpart, we tested its angular position between 0°-240° using 36 trials across nine equally spaced rotational positions (which included positions where the system was untrained). Five data

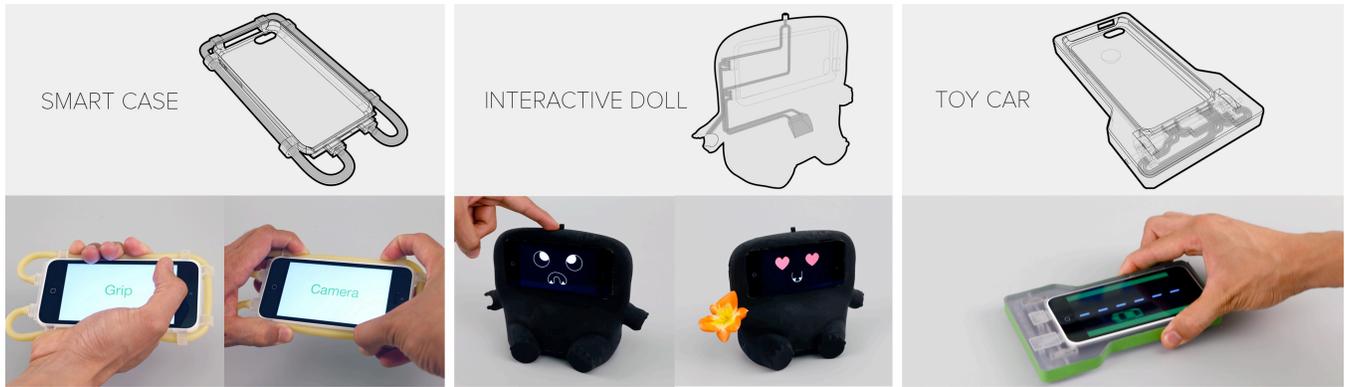


Figure 15. With Acoustruments, end-user applications can be built by combining mechanisms and design primitives.

samples were collected for each trial, generating a total of 180 data samples. Across nine configuration steps, the knob’s mean absolute error was 7.8° ($SD=6.0^\circ$). This means that even with just 10 training instances across five positions, our turn knob was correct within $\pm 7.8^\circ$, sufficient to accurately predict nine knob positions between 0° - 240° .

A line break sensor is an example of very basic, discrete classification. Similar to the evaluation we ran for the rocker switch, we collected a total of 200 sample points across 40 trials, splitting them into 20 trials per condition. To evaluate accuracy, the line break sensor was ‘opened or blocked’ (i.e., placing a finger within the aperture’s proximity) to meet the specified target condition. To prevent configuration bias, we return the line break sensor to its original state before collecting samples for subsequent trials (i.e., setting it back to ‘open’ whenever it is ‘blocked’). Across 200 samples, the line break sensor misclassified one data point, garnering an accuracy of 99.50% ($SD=0.07\%$).

Finally, we tested our continuous pressure sensor with a rotatable clamp, which allowed for pressure to be varied and controlled. The clamp and pressure sensors were mounted on fixed positions, and pressure was varied by incrementally turning the clamp. In a similar fashion to the knob mechanism, our pressure sensor was trained with 10 instances across five equally spaced variations (between

0.0 - 4.0), and tested for 36 trials across nine equally spaced clamp rotations (with five samples collected for each trial). Across trials, the mean absolute error was ± 0.29 ($SD=0.21$).

Noise Robustness

Since our technique is acoustically driven, we investigated its behavior and robustness to noise across different environments. Using our alarm clock example application (see below), we conducted hour-long tests in two locations, one in an office environment (~ 57 dB ambient noise) and the other in a noisy café (~ 72 dB). The alarm clock was previously trained in a lab setting, using a relatively small number of 50 training instances for each of its three states (ON, OFF, SNOOZE). In each environment, we toggled the alarm clock between ‘ON’ and ‘OFF’ states for 30 minutes each, and recorded any misclassifications. We collect 10 data samples per second, generating a total of 36K readings per environment. Overall, the alarm clock misclassified 4 readings, garnering an overall false positive rate of less than 0.006%. We found no significant difference in noise robustness results between the two environments.

EXAMPLE APPLICATIONS

To illustrate the potential of our approach, we created four example applications, each consisting of multiple mechanisms operating in concert (see Figure 15 and Figure 16).

Alarm Clock. In our first example, we built a docking station that operates as an alarm clock equipped with interactive physical controls. Our system can sense when users dock their mobile device, and it launches a companion alarm clock application. A rocker switch mechanism turns the alarm on or off, and a large soft cavity functions as the snooze button. When the alarm goes off, periodic beeps are played on the speakers. While this is happening, our system can still normally emit and sense ultrasonic signals.

Smart Case. In this example, we built a smartphone case with sensing mechanisms wrapped around its exterior. This enables the case to sense dynamic configurations (e.g., grasps or gestures), allowing for interesting interaction modalities. For example, it can sense whether the phone is



Figure 16. Our alarm clock application utilizes several mechanisms for interactive control, including a rocker switch and a pressure sensor.

placed on a tabletop or hand carried. Additionally, it can support unique hand gestures such as gripping or a camera gesture, which can be used for launching system-wide apps.

Interactive Doll. We built an interactive “pluggable” doll using pressure and proximity sensors, modular insertions, and several branching combinations. These mechanisms provide an otherwise static doll with rich interaction capabilities. For instance, it can respond when its tummy is poked or when its antenna is being bent, or it can sense and change moods when objects are inserted in one of its hands.

Car Toy. Finally, we fabricated a “pluggable” toy car using two rotary encoders as a differential-drive wheel mechanism, and a set of engineered baffles to more effectively discriminate the rotational configuration of each wheel. This allows the car to sense speed, forward or backward movement, and steering direction. With these capabilities, the toy car itself can be used as a powerless tangible controller for a racing game, or as an interactive story prop.

DISCUSSION

One of the major limitations of our approach is cross talk. As more structural elements are introduced, signal bandwidth becomes overloaded since cross talk distorts a design primitive’s overall effect in altering the acoustic signal. However, we note that due to the hardware limitations of our device, our effective sweeping range is quite narrow (~5K). Devices with higher sampling rates, as well as signal modulation techniques could help mitigate cross-talk effects. Since we utilize the speaker for sensing, our technique impedes with how users regularly use their phone for audio output. Our sensing accuracy is sub-optimal when audio is playing on the speaker (i.e., since sound from the speakers creates unwanted harmonic effects to the signal traversing along the tubes). One possible solution is to incorporate silence detection and interweave ultrasonic sweeps during quiet portions of playback (e.g., as seen in our video figure, where we demonstrate this approach in our alarm clock application).

Our method is fairly noise robust as the tube acts like a closed system, significantly dampening external noise (and thus curtailing the effect of ambient ultrasound). It also helps that our approach is active (and thus loud relative to external noises). In specific situations, our method can be susceptible to noise when the tube is overly exposed to open air. Also, forcefully hitting the enclosure (e.g., banging) will introduce unwanted vibro-acoustic interference.

Although no formal tests were performed, we can anecdotally report that the acoustic signatures of similar 3D-printed models are also similar. As long as the physical configurations are the same, the acoustic output will be identical. As a consumer product, Acoustruments can be pre-trained more exhaustively from thousands of data points, increasing robustness and eliminating calibration.

CONCLUSION

Acoustruments are passive, acoustically driven mechanisms that provide rich, tangible functionality to handheld device interaction. By combining design primitives, familiar physical mechanisms, and even end-user applications, can all be constructed from passive elements made of plastic. Using technologies like 3D printing, we show that rich physical controls can be rapidly prototyped, providing new methods for experimentation by HCI practitioners. Further, our approach can be extended to traditional fabrication techniques, such as injection molding, milling and machining, which can further drive down cost and improve ease of deployment in consumer products.

ACKNOWLEDGEMENTS

We would like to thank Mo Mahler and Tyler Porten for creating Yakamo, the character for our interactive toy doll.

REFERENCES

1. Aumi, M.T.I., Gupta, S., Goel, M., Larson, E., Patel, S. DopLink: using the doppler effect for multi-device interaction. In *Proc. UbiComp '13*.
2. Avrahami, D., Hudson, S.E. Forming interactivity: a tool for rapid prototyping of physical interactive products. In *Proc. DIS '02*.
3. Baudisch, P., Becker, T., Rudeck, F. Lumino : Tangible Blocks for Tabletop Computers Based on Glass Fiber Bundles. In *Proc. CHI '10*.
4. Benade, A.H. Fundamentals of musical acoustics. *Oxford Press, 1976*.
5. Brandt, A. Noise and Vibration Analysis. *Wiley, 2011*.
6. Briggs, G.A. Musical instruments and audio. *Wharfedale Wireless Works, Idle; Yorkshire, 1965*.
7. Brockmeyer, E., Poupyrev, I., Hudson, S. PAPHON: designing curved display surfaces with printed optics. In *Proc. ACM UIST '13*.
8. Carse, A. Musical wind instruments. *Da Capo, 1975*.
9. Chan, L., Müller, S., Roudaut, A., Baudisch, P. CapStones and ZebraWidgets: sensing stacks of building blocks, dials and sliders on capacitive touch screens. In *Proc. UIST '12*.
10. Crevoisier, A., Polotti, P., Politecnico, I. Tangible Acoustic Interfaces and their Applications for the Design of New Musical Instruments. In *Proc NIME '05*.
11. Culver, C.A. Musical acoustics. *McGraw-Hill, 1956*.
12. Daubeny, U. Orchestral wind instruments, ancient and modern. *Books for Libraries Press, 1970*.
13. Debost, M. The simple flute : from A to Z. *Oxford University Press, 2002*.
14. Fletcher, N.H., Rossing, T.D. The physics of musical instruments. *Springer, 1998*.

15. Gloth, G., Sinapius, M. Analysis of swept-sine runs during modal identification. In *Proc. Mechanical Systems Signal Processing '04*.
16. Greenberg, S., Fitchett, C. Phidgets: easy development of physical interfaces through physical widgets. In *Proc. UIST '01*.
17. Gupta, S., Morris, D., Patel, S., Tan, D. SoundWave: Using the Doppler Effect to Sense Gestures. *UIST '12*.
18. Harrison, C. and Hudson, S.E. Scratch input: creating large, inexpensive, unpowered and mobile finger input surfaces. In *Proc. UIST '12*.
19. Harrison, C., Tan, D., Morris, D. Skinput: appropriating the body as an input surface. In *CHI '10*.
20. Hudson, S.E. Printing teddy bears: a technique for 3D printing of soft interactive objects. In *Proc. CHI '14*.
21. Hwang, S., Ahn, M., and Wohn, K. MagGetz: Customizable Passive Tangible Controllers on and Around Conventional Mobile Devices. In *UIST '13*.
22. Ishii, H., Wisneski, C., Orbanes, J., Chun, B., Paradiso, J. PingPongPlus: design of an athletic-tangible interface for computer-supported cooperative play. In *UIST '07*.
23. Kratz, S., Westermann, T., Rohs, M., Essl, G. CapWidgets: tangible widgets versus multi-touch controls on mobile devices. In *Proc. CHI '11*.
24. Kuo, Y.-S., Verma, S., Schmid, T., Dutta, P. Hijacking power and bandwidth from the mobile phone's audio interface. In *Proc. ACM DEV '10*.
25. Liang, R., Chan, L., Kuo, H.T.H. GaussBricks: Magnetic Building Blocks for Constructive Tangible Interactions on Portable Displays. In *Proc. CHI '14*.
26. Mueller, S., Kruck, B., Baudisch, P. LaserOrigami: laser-cutting 3D objects. In *Proc. CHI '13*.
27. Mueller, S., Mohr, T., Guenther, K., Frohnhofen, J., Baudisch, P. faBrickation : Fast 3D Printing of Functional Objects by Integrating Construction Kit Building Blocks. In *Proc. CHI '14*.
28. Mujibiya, A., Cao, X., Tan, D.S., Morris, D., Patel, S.N., Rekimoto, J. The sound of touch: on-body touch and gesture sensing based on transdermal ultrasound propagation. In *Proc. ITS '13*.
29. Munjal, M.L. Acoustics of ducts and mufflers with application to exhaust and ventilation system design. *Wiley 1987*.
30. Ono, M., Shizuki, B., Tanaka, J. Touch & activate: adding interactivity to existing objects using active acoustic sensing. In *Proc. UIST '13*.
31. Patel, S.N., Abowd, G.D. Blui: low-cost localized blowable user interfaces. In *Proc. UIST '07*.
32. Priyantha, N.B., Chakraborty, A., Balakrishnan, H. Cricket location-support system. In *Proc. MobiCom '00*.
33. Raj, B., Kalgaonkar, K., Harrison, C., Dietz, P. Ultrasonic Doppler Sensing in HCI. In *Pervasive'12*.
34. Robinson, A., Verma, S., Dutta, P. AudioDAQ: turning the mobile phone's headset port into a universal data acquisition interface. In *Proc. IPSN '12*.
35. Sato, M., Poupyrev, I., Harrison, C. Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. In *Proc. CHI '12*.
36. Savage, V., Chang, C., Hartmann, B. Sauron: embedded single-camera sensing of printed physical user interfaces. In *Proc. UIST '13*.
37. Savage, V., Schmidt, R., Grossman, T., Fitzmaurice, G., Hartmann, B. A Series of Tubes: Adding Interactivity to 3D Prints Using Internal Pipes. In *Proc. UIST '14*.
38. Savage, V., Zhang, X., Hartmann, B. Midas: Fabricating Custom Capacitive Touch Sensors to Prototype Interactive Objects. In *Proc. ACM UIST '12*.
39. Sinha, D., Springer, K., Han, W., Lizon, D., Houlton, R. Swept-frequency acoustic interferometry technique for noninvasive chemical diagnostics. In *Proc. 3rd Int. Conf. On-Site Analysis '95*.
40. Takemura, K., Ito, A., Takamatsu, J., Ogasawara, T. Active bone-conducted sound sensing for wearable interfaces. In *Proc. UIST '11*.
41. Tarzia, S.P., Dick, R.P., Dinda, P.A., Memik, G. Sonar-based measurement of user presence and attention. In *Proc. Ubicomp '09*.
42. Tromlitz, J.G., Powell, A. The keyed flute. Clarendon Press ; *Oxford University Press, 1996*.
43. Von Helmholtz, H. and Ellis, A.J. On the sensations of tone as a physiological basis for the theory of music. *Dover Publications, 1954*.
44. Watanabe, C., Cassinelli, A., Watanabe, Y., Ishikawa, M. Generic method for crafting deformable interfaces to physically augment smartphones. In *Proc. CHI EA '14*.
45. Watanabe, H., Terada, T., Tsukamoto, M. Ultrasound-based movement sensing, gesture-, and context-recognition. In *Proc. ISWC '13*.
46. Weiss, M., Wagner, J., Jansen, Y., Jennings, R., Khoshabeh, R., Hollan, J.D., Borchers, J. SLAP Widgets: Bridging the Gap Between Virtual and Physical Controls on Tabletops. In *Proc. CHI '14*.
47. Willis, K., Brockmeyer, E., Hudson, S., Poupyrev, I. Printed optics: 3D printing of embedded optical elements for interactive devices. In *Proc. UIST '12*.
48. Romo Robot. <http://www.romotive.com/>.
49. Thingiverse. <http://www.thingiverse.com/>.
50. The First 3D-Printed Sax. <http://gizmodo.com/the-first-3d-printed-saxophone-sounds-surprisingly-dece-1616272037>.
51. LibSVM. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>