Scratch Input: Creating Large, Inexpensive, Unpowered and Mobile Finger Input Surfaces

Chris Harrison *Scott E. Hudson*
Human-Computer Interaction Institute
Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213
{chris.harrison, scott.hudson}@cs.cmu.edu

ABSTRACT
We present Scratch Input, an acoustic-based input technique that relies on the unique sound produced when a fingernail is dragged over the surface of a textured material, such as wood, fabric, or wall paint. We employ a simple sensor that can be easily coupled with existing surfaces, such as walls and tables, turning them into large, unpowered and ad hoc finger input surfaces. Our sensor is sufficiently small that it could be incorporated into a mobile device, allowing any suitable surface on which it rests to be appropriated as a gestural input surface. Several example applications were developed to demonstrate possible interactions. We conclude with a study that shows users can perform six Scratch Input gestures at about 90% accuracy with less than five minutes of training and on wide variety of surfaces.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces - Input devices and strategies.

General terms: Design, Human Factors

Keywords: Finger input, gestures, surfaces, acoustic sensing, ad hoc interaction, mobile devices.

INTRODUCTION

The potential benefits of moving computing and communication into aspects of life that transcend the work environment are significant. Offices provide a relatively controlled environment where standing technology infrastructure can be deployed. However, this becomes inaccessible once we leave the office and even when we simply move around within the workplace. Today's powerful mobile computing devices offer one way to overcome this, allowing us to carry parts of our infrastructure with us. However, because they are carried, we prefer them to be small. This pushes us into using tiny displays, buttons, keyboards, and other input methods rather than, for example, making use of large surfaces for input. Even in the home where we *could* deploy computing infrastructure, the cost, difficulty, and intrusion of installation is often prohibitive.

In this paper, we consider a new input technique that allows small devices to *appropriate* existing, large, passive surfaces such as desks and walls, for use as a kind of input device. This *Scratch Input* technique operates by listening to the sound of "scratching" (e.g., with a fingernail) that is transmitted through the surface material. This signal can be used to recognize a vocabulary of gestures carried out by the user. Our sensor is simple and inexpensive, and can be easily incorporated into mobile devices, enabling them to appropriate whatever solid surface they happen to be resting on. Alternately, it can be very easily deployed, for example, to make existing walls or furniture input-capable.

UIST'08, October 19–22, 2008, Monterey, California, USA.

SENSING

Scratch Input takes advantage of particular physical effects in order to detect input on surfaces like tables, walls, and even clothes. Foremost, a fingernail dragged over a textured surface, such as wood, fabric, or wall paint, will produce a sound containing a particularly high frequency component (typically greater than 3000Hz). This high frequency property allows it to be easily separated from other typical house and office noises, for example, voice (90-300Hz), singing (80-1200Hz), typical mechanical vibration (e.g., refrigerator compressors, washing machines), and AC driven lighting, etc. (50 or 60Hz).

Another important property that is exploited is that sound propagates through solid (and liquid) materials much more efficiently than through the air. So while running your fingernail across a surface will produce only a soft audible noise, the signal transmits considerably better through the solid host material. This superior transmission of sound means that a signal is not only transmitted further, but is also better preserved (i.e., less noisy). These two properties work in concert to enable Scratch Input to work reliably across large input areas.

To capture sound transmission through solid materials, we use a modified stethoscope (Figure 1). This is particularly well suited to both amplifying sound and detecting high frequency noises. This is attached to a generic microphone, which converts the sound into an electrical signal. In our particular implementation, the signal is amplified and connected to a computer through the audio-input jack. If mass-produced, this sensor might cost less than one dollar.

This method has one important constraint not shared by many existing surface input approaches (see for example [2,3,4,9]) - it cannot determine the spatial location of input. Sensing locality is difficult not only because we use a single sensor, but also because Scratch Input is designed to operate in an ad hoc fashion on a range of materials, often with varying sound transmission properties (precluding, for example, the clever acoustic fingerprint approach employed

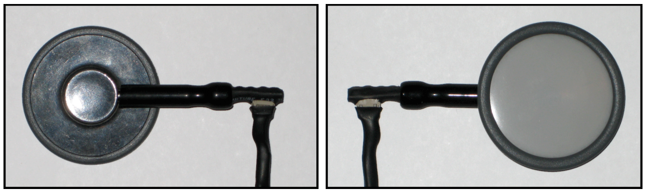

Figure 1. The front and back sides of our sensor, a modified stethoscope.

in [8]). Prior systems that employ multiple microphones to determine position include [10] and the commercial Mimio whiteboard sensor (see mimio.com). Additionally, the ad hoc appropriation of existing surfaces (similar to the use of hands in [1]) stands in contrast to systems like [6,7], which require specifically manufactured textured surfaces.

Scratch Input's non-spatial property gives it a significantly different character from many other surface input techniques and does preclude some uses (e.g., for pointing and cursor control). However, Scratch Input is primarily designed to augment existing, passive surfaces, few of which have visual displays to point at. Instead, Scratch Input is targeted towards simple gestural input, providing the convenience of ad hoc appropriation of surfaces and/or easy and unobtrusive (semi-)permanent installation.

Sensing is surprisingly robust in a wide variety of use contexts. The same sensor can be used for walls, doors, clothes, desks, countertops, cabinets, and many other surfaces. The only notable restriction is that surfaces must be continuous – any gap will prevent the sound from propagating. So, for example, two adjacent tables, even if touching, are not likely to be acoustically coupled. This is not necessarily a negative quality. If there are multiple users working in close proximity, they might not want to send input to each other's workspaces. A good example of this is a classroom, where students have their own desks and carry out independent tasks in parallel.

Although we will only discuss finger input, it should be noted that Scratch Input also works with other implements, most notably styluses like those found on PDAs. Additionally, although not tested, Scratch Input could also be used to augment whiteboards and blackboards, enabling markers and chalk to not only write, but also issue commands.

GESTURE RECOGNITION

Even the simplest of gestures requires a finger to move, accelerate, and decelerate in a particular way. For a straight line, this might start with an accelerating motion followed by a deceleration. These movements interact with the textual features of a surface to produce a particular sound. Specifically, the faster a motion is performed, the higher its amplitude (i.e., volume) and frequency. Conversely, a slower motion produces a lower frequency and lower amplitude signal. This effect can be heard with the naked ear. Complex gestures are ultimately made up of sequences of these base motions, which produces a unique acoustic signature. Figure 2 shows the amplitude response of a line, circle, triangle and square gesture.

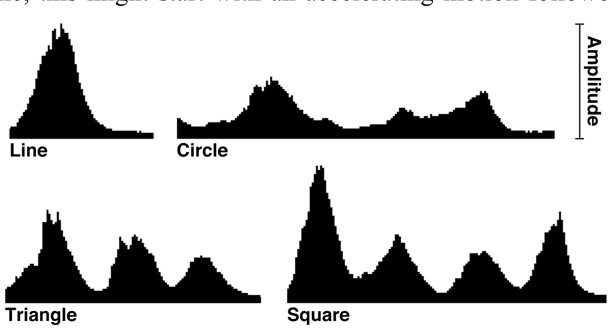

Figure 2. Amplitude profiles for different gestures.

Unlike traditional gestures, which are spatially unique (in two or three dimensions), scratch input gestures must be acoustically unique (essentially one-dimensional). For example, many letters that are written differently, sound very similar (e.g., M/W, V/L, X/T). Scratch Input cannot accurately distinguish between these gestures. Although this limits the expressivity of the input, there is still sufficient power to support dozens of gestures, especially through the use of multi-part inputs (Figure 3 offers some examples).

Scratch Input can also support gestures for continuous control of actions like scrolling, seeking, and volume adjustment. For example, scrolling down a menu could be done with a "menu start" gesture, which then trails into a continuous straight line. When the desired item is reached, the user simply stops moving his or her finger. (This approach is conceptually related to the "pigtail" gestures described in [5] and should have similar advantages.) A similar effect can be achieved with a repeating up and down gesture or a circling motion. Additionally, Scratch Input uses the amplitude of the signal (correlated with finger velocity) to provide variable rate control. This allows users to scroll quickly or slowly. This not only allows for precise adjustment, but also is intuitive, as real world magnitude controls offer the same flexibility (e.g. dials, sliders, levers).

Early prototype gesture recognizers employed dynamic time warping and naive Bayes classification. These were ultimately discarded in favor a lighter-weight approach - our final recognizer uses a shallow decision tree primarily based on peak count and amplitude variation. This implementation can easily run in real-time on a low-powered mobile device. Several methods are employed to reduce false positives, including gesture rejection, noise suppression, amplitude-independent peak detection, and input windowing. For our simple, proof-of-concept gesture set, we were able to achieve high accuracy rates using only amplitude response. However, a more sophisticated recognition engine could incorporate other dimensions, such as frequency and duration, and likely be able to support considerably more gestures and at higher accuracies.

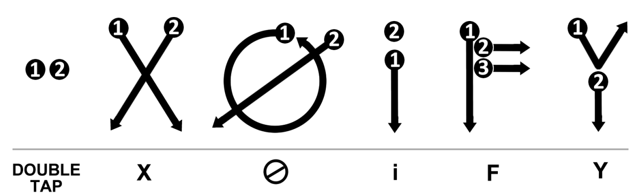

Figure 3. Distinct multi-part gestures composed of taps, lines and circles.

Our recognizer also handles finger taps. Although not "scratches", they are a convenient, easily preformed, and accurately classified type of input. They are characterized by a short burst of high frequency and high intensity sound.

As with any acoustic-based system, environmental noise is problematic. Scratch Input sidesteps much of this problem by operating solely above 3KHz. Also, environmental noises (e.g., voice, music, footsteps) tend not to be transmitted with sufficient volume to surfaces like walls. Nonetheless, the technique is susceptible to interference in certain deployments, and recognizers will need to employ sophisticated methods to robustly handle false signals.

EXAMPLE APPLICATIONS

During initial experimentation, we primarily concentrated on two classes of surfaces: tables and walls. These were the strongest candidates for large, easily appropriated or retrofitted input surfaces. Part of our exploration of these surfaces involved producing and testing a proof-of-concept application for each use context, the outcome of which we present below. We also briefly investigated device enclosures and fabric (e.g., clothes) as input surfaces.

Tables and Mobile Devices

Tables, usually made of dense materials like wood, have excellent transmission characteristics. We found that for typically sized tables (1 to 3m wide and 1 to 2m deep), our sensor can be placed anywhere on the surface and receive input from any part. A dozen different tables were tested, with almost all yielding excellent results. The two exceptions were a glass and highly glossy wood table, both of which are almost texture-less and produced too weak of a signal to be detected. Furthermore, it appears that objects resting on the table (e.g., mugs, lamps, pens, papers, computer monitors) do not affect sound transmission quality.

Our sensor could be easily incorporated into mobile devices, such as cell phones, PDAs, and laptops. This would allow them to receive Scratch Input on whatever surface they are resting on. This essentially provides a mobile, ad-hoc input surface wherever the user sets down the device.

Consider the example of a cell phone augmented with our sensor sitting on a table. If there was incoming call, the user could silence the ring by simply writing a gesture anywhere on the table's surface. On the other hand, if the user wished to take the call, they could issue a gesture that answered using the speakerphone. This has the beneficial property of not requiring the user to move their visual attention away from their current task or to reach for the device to interact with its buttons.

We simulated this interaction by placing our sensor on the surface of a table. We then rested a cell phone on top of the sensor to mimic the correct pressure that would be applied. Our computer-based recognizer understood three gestures, switch to normal mode (a back and forth motion with at least five passes), switch to silent mode (the letter 'S'), and answer using speakerphone (an abstract 'A' gesture, drawn like an upside-down 'V'). This simple, yet useful gesture vocabulary proved to be very accurate.

Additionally, it is not unusual for multiple devices to reside on a single work surface, such as a desk at home or work. This means that gestures issued to the surface will be broadcast to all devices resting on it. Although devices will likely want to support their own unique gestures (so you could gesture to a particular device or class of devices), it is also interesting to consider universal gestures. For example, imagine we have a cell phone and laptop sitting on a common table. If a user did not want to be disturbed, a single gesture could be issued that turns the cell phone to silent, logs out of instant messenger, and closes the email client.

Walls

Walls, being both ubiquitous and large, are a strong candidate for input. In a typical wood frame house with painted drywall, our sensor had an effective range of about 8m. Thus, a single sensor could provide a wall-based input surface 16m wide from floor to ceiling, an area of approximately 40 square meters. Tests also revealed that the signal propagates strongly around corners, and, although with some loss of signal, successfully around door openings. This means a few dozen easily installed sensors could be used to outfit an entire house. However, this property also means adjacent rooms (with common walls) may not be acoustically isolated, making use in shared environments (e.g., an office) potentially problematic. It should also be noted that repeated "scratching" of walls without proper surface treatment could produce marks after extended use.

We built a simple, proof-of-concept audio player that lets users issue commands on their home walls. Users were able to pause and play with a double tap. A double swipe gesture toggled between volume and seek modes. Users could single tap to toggle between increasing or decreasing the volume and seeking forwards or backwards (depending on the current mode). A continuous circling motion was used as magnitude control. This is used for seeking and adjusting the volume.

Device Enclosures

Most electronics are encased in a plastic shell. We can augment these devices with Scratch Input by coupling a microphone to this surface. Some devices already have this feature, including cell phones and many laptops. To verify this effect, we captured audio samples of Scratch Input on a cell phone and a laptop, both of which performed well.

As noted previously, sound only transmits through contiguous materials. For example, with the flip phone we tested, this meant that Scratch Input was limited to the bottom half (i.e., the half with the microphone). Regardless, this opened the possibility for several new input areas, including the bottom and sides of the device, as well as the large area covering the battery on the backside the phone. This was also true of the Apple MacBook we tested. In this case, the microphone was located above the display, limiting input to the bezel around the LCD screen and the rear of the monitor. However, microphones are sufficiently inexpensive that it would be trivial to add additional sensing locations. For example, the palm-rest area of a laptop could be turned into a large, non-spatial trackpad.

We produced two simple Scratch Input applications for use with a laptop enclosure. The first application allowed users to move forwards and backwards in their PowerPoint presentation. The second application allowed users to navigate their file system. Like a mouse, a single tap was used to select a file or directory, while a double tap was used to open. Users could also scroll by dragging their finger in a continuous line down the side of their display.

Fabric

Mobile devices often reside in people's pockets. In order to perform even simple actions, like turning a cell phone to silent, requires "getting out" the device. This frequently requires a great deal of visual attention, especially if menus have to be navigated. This can be obtrusive in some social contexts. Scratch Input, however, would allow commands to be issued by simply writing a gesture on or near the pocket area. Since the human body mostly composed of liquid, it can transmit sound short distances without too much degradation. In simple tests, we found that a sensor placed in the pocket facing towards the leg could detect input from roughly the belt line down to the knee. The chief obstacle with fabric-based input is the high level of noise, especially during activities like walking. However, it may be possible to engineer gestures that are sufficiently unique that false positives remain low.

EVALUATION

To gauge the accuracy of our proof-of-concept Scratch Input gesture set and recognizer, we recruited 15 participants (9 female) with a mean age of 31. The experiment was conducted in participants' offices or homes in order to simultaneously evaluate the robustness of Scratch Input on a realistic variety of real world input surfaces. Participants were paid ten dollars for their involvement.

Participants were given a brief description of Scratch Input and shown the sensor. The sensor was then placed on a surface, off to one side, and weighted down with a cell phone. A laptop was situated in front of them, but sufficiently far away such that there was at least one square meter of input area.

We choose six simple gestures for inclusion in the experiment: single tap, double tap, single swipe, double swipe, triple swipe, and quad swipe. The latter four gestures were equivalent to the letters I, V, N and W. These were selected because they represented increasingly complex gesture sequences already familiar to our subjects, allowing us to test a hypothesis that accuracy rates would fall as gesture complexity rose.

First participants were shown what the gestures look like graphically with a brief PowerPoint presentation (swipe gestures were shown as letters). Then, participants were given a maximum of five minutes to practice inputting these gestures. The laptop ran a full-screen, real-time version of our gesture recognizer to show participants if their input was being classified correctly. During this period, the experimenter provided advice and/or demonstrated gestures in order to help train the participant.

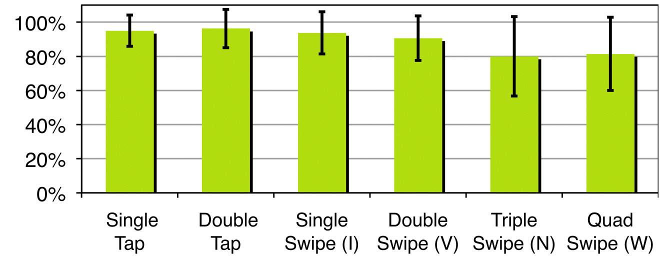

Figure 4. Average accuracy for the six test gestures.

Following this training period, a new program was launched that displayed which gesture the participant needed to perform. Each gesture was requested five times, for a total of 30 trials. The sequence was randomized for each participant to compensate for any order effects.

Results indicate participants were able to achieve an average accuracy of 89.5%. As hypothesized, accuracy suffered as gesture complexity grew (Figure 4). Gestures with two of fewer motions achieved accuracies in excess of 90%.

CONCLUSION

We presented Scratch Input, an acoustic-based finger input technique that can be used to create large, inexpensive and mobile finger input surfaces. This can allow mobile devices to appropriate surfaces on which they rest for gestural input. Our investigations revealed that Scratch Input is both easy to use and accurate on a variety of surfaces.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the Intel Research Council, General Motors, and the National Science Foundation under Grant IIS-0713509.

REFERENCES

1. Amento, B., Hill, W., and Terveen, L. The sound of one hand: a wrist-mounted bio-acoustic fingertip gesture interface. In *Proceedings of CHI '02 Ext. Abstracts*, pp. 724-725.
2. Buxton, W., Hill, R., and Rowley, P. Issues and techniques in touch-sensitive tablet input. In *SIGGRAPH '85*, pp. 215-224.
3. Dietz, P. and Leigh, D. DiamondTouch: a multi-user touch technology. In *Proceedings of UIST '01*, pp. 219-226.
4. Han, J. Y. 2005. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proc. UIST '05*, pp. 115-118.
5. Hinckley, K., Baudisch, P., Ramos, G., and Guimbretiere, F. Design and analysis of delimiters for selection-action pen gesture phrases in scriboli. In *Proc. of CHI '05,* pp. 451-460.
6. Kim, J., Sunwoo, J., Son, Y., Lee, D., Lee, D., and Cho, I. A gestural input through finger writing on a textured pad. In *Proceedings of CHI '07 Ext. Abstracts*, pp. 2495-2500.
7. Murray-Smith, R., Williamson, J., Williamson, J., Hughes, S., and Quaade, T. Stane: synthesized surfaces for tactile input. In *Proceedings of CHI '08*, pp. 1299-1302.
8. Patel, S. N. and Abowd, G. D. Blui: low-cost localized blowable user interfaces. In *Proceedings of UIST'07*, pp. 217-220.
9. Paradiso, J. A., Hsiao, K., Strickon, J., Lifton, J., and Adler, A. Sensor systems for interactive surfaces. *IBM Systems Journal*, 39, 3-4 (Jul. 2000), 892-914.
10. Paradiso, J., Leo, C., Checka, N., and Hsiao, K. Passive acoustic sensing for tracking knocks atop large interactive displays. In *Proc. of IEEE Sensors Conference*, 2002, pp. 521-527.