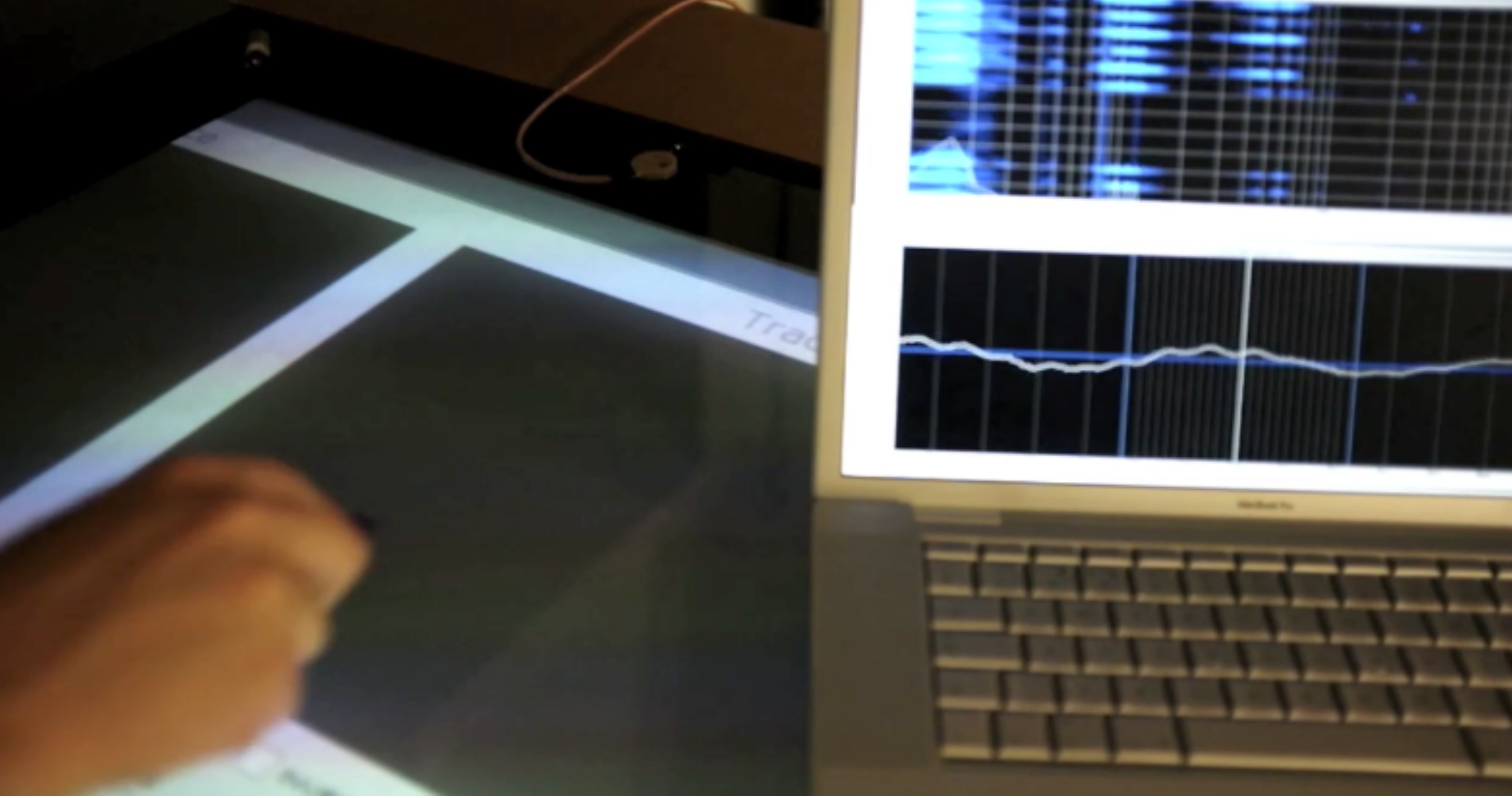


Sound Interfaces

6.810 Engineering Interaction Technologies

Prof. Stefanie Mueller | HCI Engineering Group

what is sound?

if a tree falls in the forest and nobody is there
does it make sound?

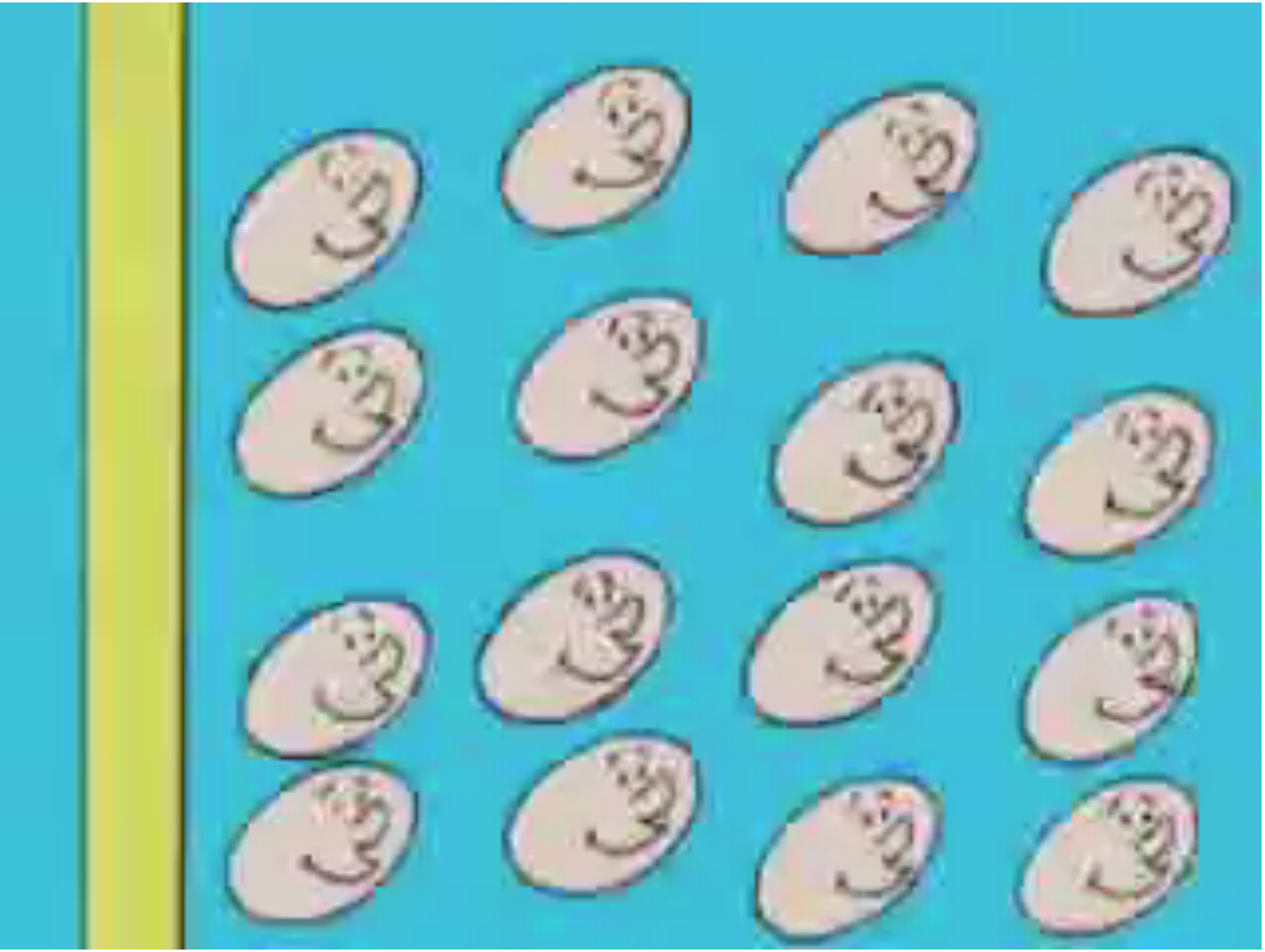
<30 sec brainstorming> [George Berkley, 1710] yes and no!

sound::

- pressure transmitted through a solid, liquid or gas
- mechanical wave (atoms **vibrate)**
- sound = perception of this wave through the ear + brain
- we can only hear specific sounds (20Hz - 20KHz)

example: drum

- drum has flexible membrane
- when being hit, **membrane oscillates**
- pushes atoms in the surroundings -> vibrate
- your ear perceives the vibrations
- and your brain transforms into the notion of ‘sound’



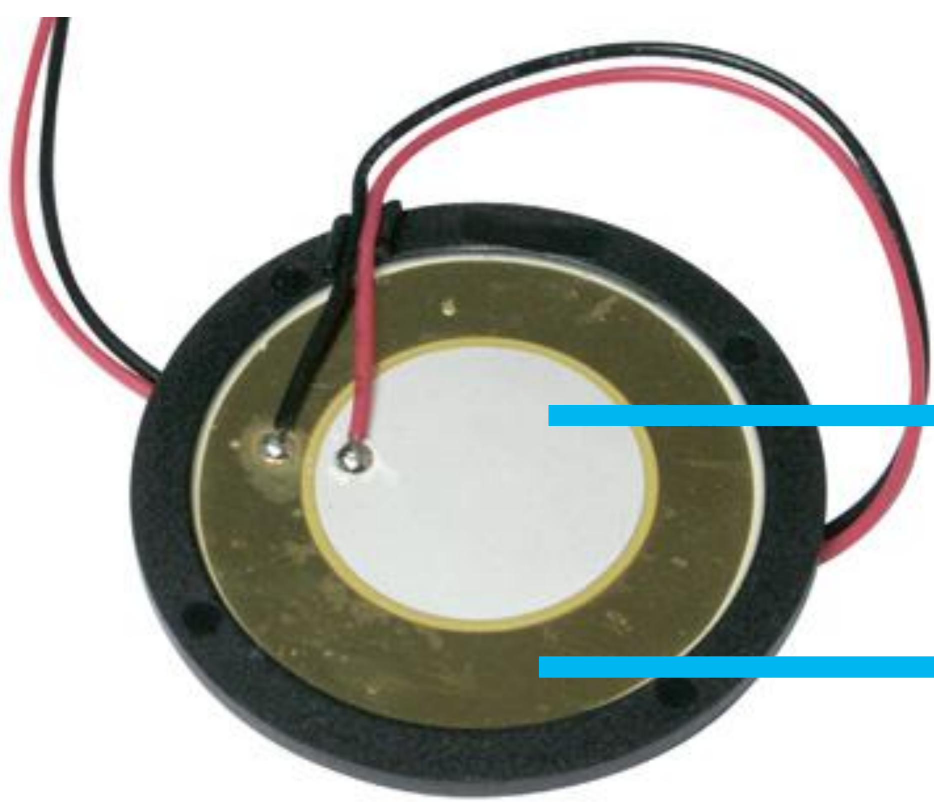
machines don’t perceive sound
they **just perceive the vibration**

example: contact microphone

- also called a piezo (piezo = pressure)
- piezo-electric effect = electricity from pressure

capture sound (use as microphone):
sound in surrounding -> piezo vibrates, creates electricity

create sound (use as speaker):
electricity from computer -> piezo vibrates, creates sound



piezoelectric material
(positively charged)

metal (negatively charged)

2015: MOGEES

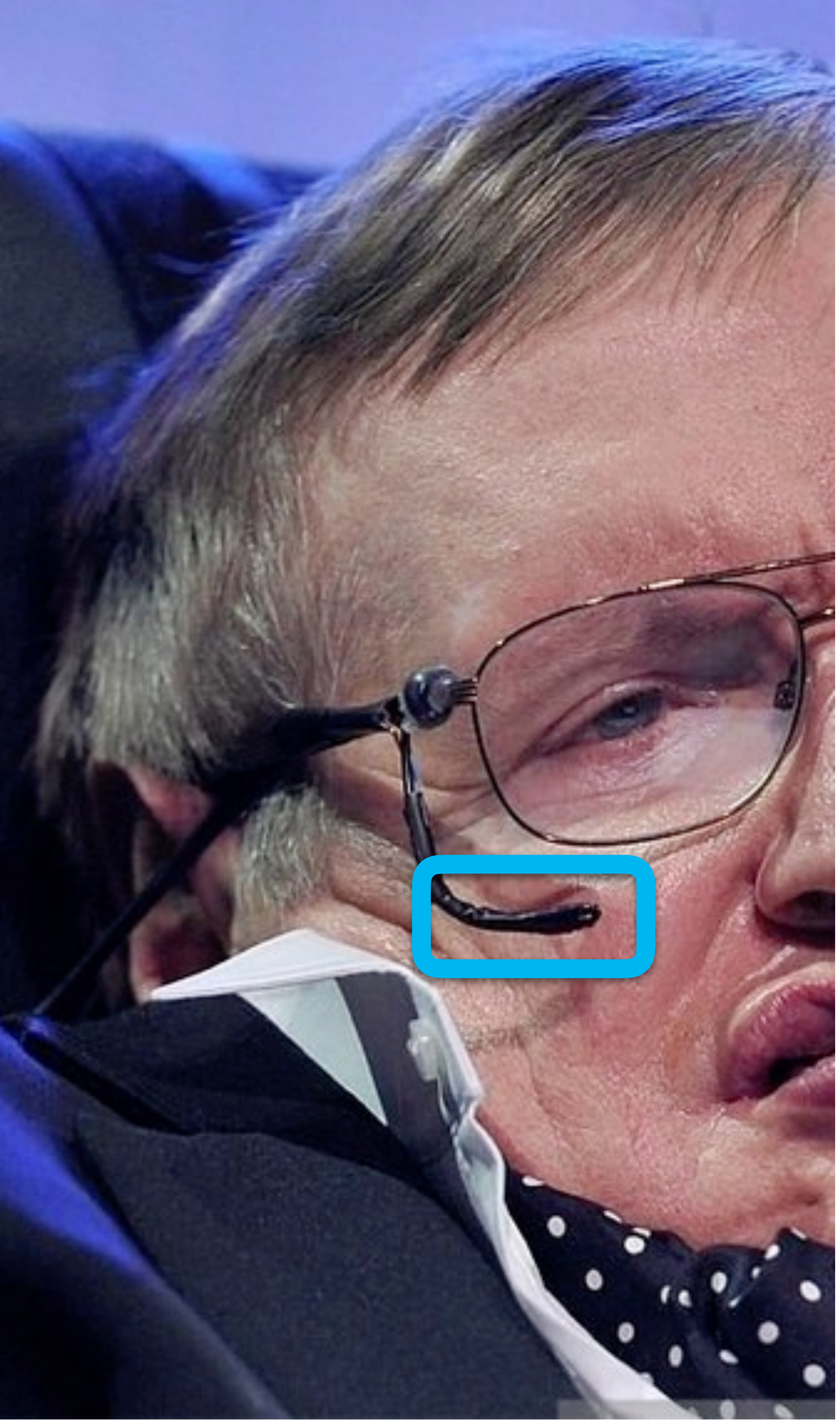
contact mic + gesture recognition software

let’s look at **different categories**
of sound interfaces…

sound interfaces:

	output	input
speech	text-to-speech synthesizers	siri, google glass 'put that there'
non-speech	notifications, earcons	microphones as sensors

#1
speech output



Stephen Hawking used speech synthesizer

- cheek movement to select letters
- cheek movement detected with sensor

1993 GloveTalk gesture interface

1993 GloveTalk gesture interface

Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer

S. Sidney Fels and Geoffrey E. Hinton

Abstract—To illustrate the potential of multilayer neural networks for adaptive interfaces, we used a VPL Data-Glove connected to a DECtalk speech synthesizer via five neural networks to implement a hand-gesture to speech system. Using minor variations of the standard back-propagation learning procedure, the complex mapping of hand movements to speech is learned using data obtained from a single "speaker" in a simple training phase. With a 203 gesture-to-word vocabulary, the wrong is produced less than 1% of the time, and no word is produced about 5% of the time. Adaptive control of the speaking rate and word stress is also available. The training times and final performance speed are improved by using small, separate networks for each naturally defined subtask. The system demonstrates that neural networks can be used to develop the complex mappings required in a high bandwidth interface that adapts to the individual user.

I. Introduction

ADAPTIVE interfaces are natural and important class of applications for neural networks. When a person must provide high bandwidth control of a complex physical device, a compatible mapping between the person's movements and the behavior of the device becomes crucial. With many devices the mapping is fixed and if a poor mapping is used, the device is difficult to control. Using adaptive neural networks, it may now be possible to build device interfaces where the mapping adapts automatically during a training phase. Such adaptive interfaces would simplify the process of designing a compatible mapping and would also allow the mapping to be tailored to each individual user. The key features of

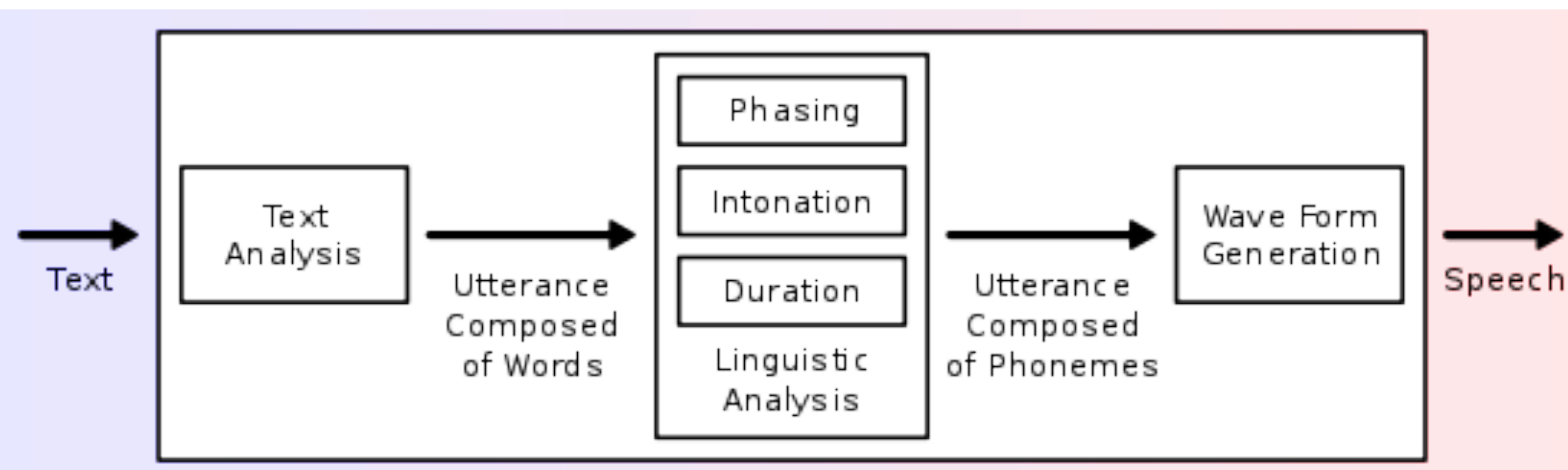
II. Overview of the Glove-Talk System

To demonstrate the usefulness of neural networks for adaptive interfaces, we chose the task of mapping hand-gestures to speech [1]. The hand-gesture data is sensed by a VPL Data-Glove [2] that has two sensors for each finger. The sensors are fiber optic transducers which measure the finger flex angles. There is also a "polhemus" sensor attached to the back of the glove which measures the x, y, z, roll, pitch, and yaw of the hand relative to a fixed source. All 16 parameters are measured every 1/60th second. The speech synthesizer is a DECtalk model DTC01 from Digital Equipment Corporation. This synthesizer can perform text-to-speech synthesis and there is also user control of speaking rate and word stress.

The granularity of speech can be used to define a spectrum of possible methods for mapping from hand-gestures to speech. At the finest granularity, rapid finger movements could play the role of movements of the speech articulators, or they could represent some other parameterization of the speech wave such as the frequencies and amplitudes of the first four formants plus the pitch, the degree of voicing, and the amplitude of the nasal formant. This gives the user an unlimited vocabulary and analog control over the quality of the speech, but the finger movements must be extremely fast and they must be recognized very rapidly to produce real-time speech without a noticeable lag. In the middle of the spectrum, a brief movement or hand configuration could represent a diphone or syllable. At the other end of the spectrum, a complete hand-gesture

challenges:

- concatenating sounds
- emphasis on words
- sentence melody
- tone, pauses



draw-to-**sing** synthesis

HandSketch Bi-Manual Controller

Investigation on Expressive Control Issues of an Augmented Tablet

Nicolas D'Alessandro
Laboratoire de TCTS - FRIA/FNRS
Faculté Polytechnique de Mons
31, Boulevard Dolez - Mons, Belgium
nicolas.dalessandro@fpms.ac.be

Thierry Dutoit
Laboratoire de TCTS
Faculté Polytechnique de Mons
31, Boulevard Dolez - Mons, Belgium
thierry.dutoit@fpms.ac.be

ABSTRACT

In this paper, we present a new bi-manual gestural controller, called HandSketch, composed of purchasable devices : pen tablet and pressure-sensing surfaces. It aims at achieving real-time manipulation of several continuous and articulated aspects of pitched sounds synthesis, with a focus on expressive voice. Both prefered and non-prefered hand issues are discussed. Concrete playing diagrams and mapping strategies are described. These results are integrated and a compact controller is proposed.

Keywords

Pen tablet, FSR, bi-manual gestural control.

1. INTRODUCTION

As explained in [11], the development of gestural controllers for sound synthesis is today a bit less a pioneer's business. Indeed, plenty of sound synthesis modules are now available in usual and inexpensive computers. Integrated toolkits, with various sensors, user friendly analog-to-digital converters and configuration softwares, can be purchased at affordable prices. Moreover, several environments for mappings and synthesis implementation, such as Max/MSP, and powerful transmission protocols (e.g. OSC) are well supported and integrated on usual plateforms.

Thus, in several contexts, the matter is less having tech-

finger-based gestures. Both approaches are joined together in order to propose the "HandSketch" controller. A typical playing situation is illustrated in Figure 1.

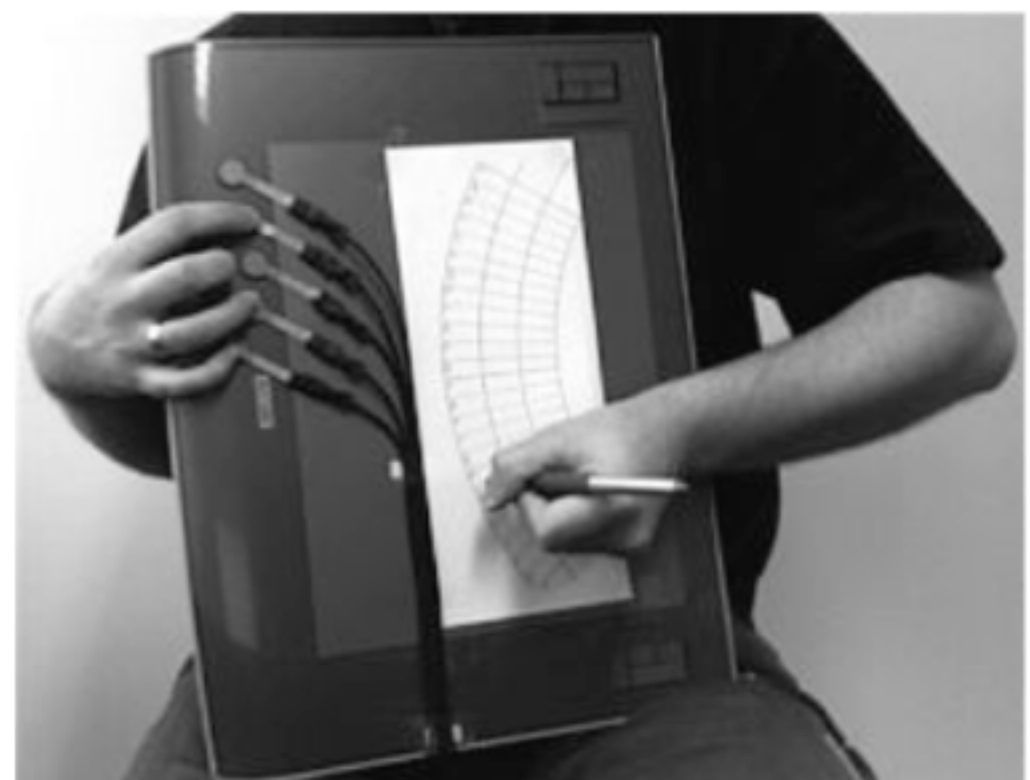


Figure 1: Representation of the "HandSketch" controller in use : Wacom™ tablet, with radial pen diagram, and 8 FSRs (Force Sensing Resistors).

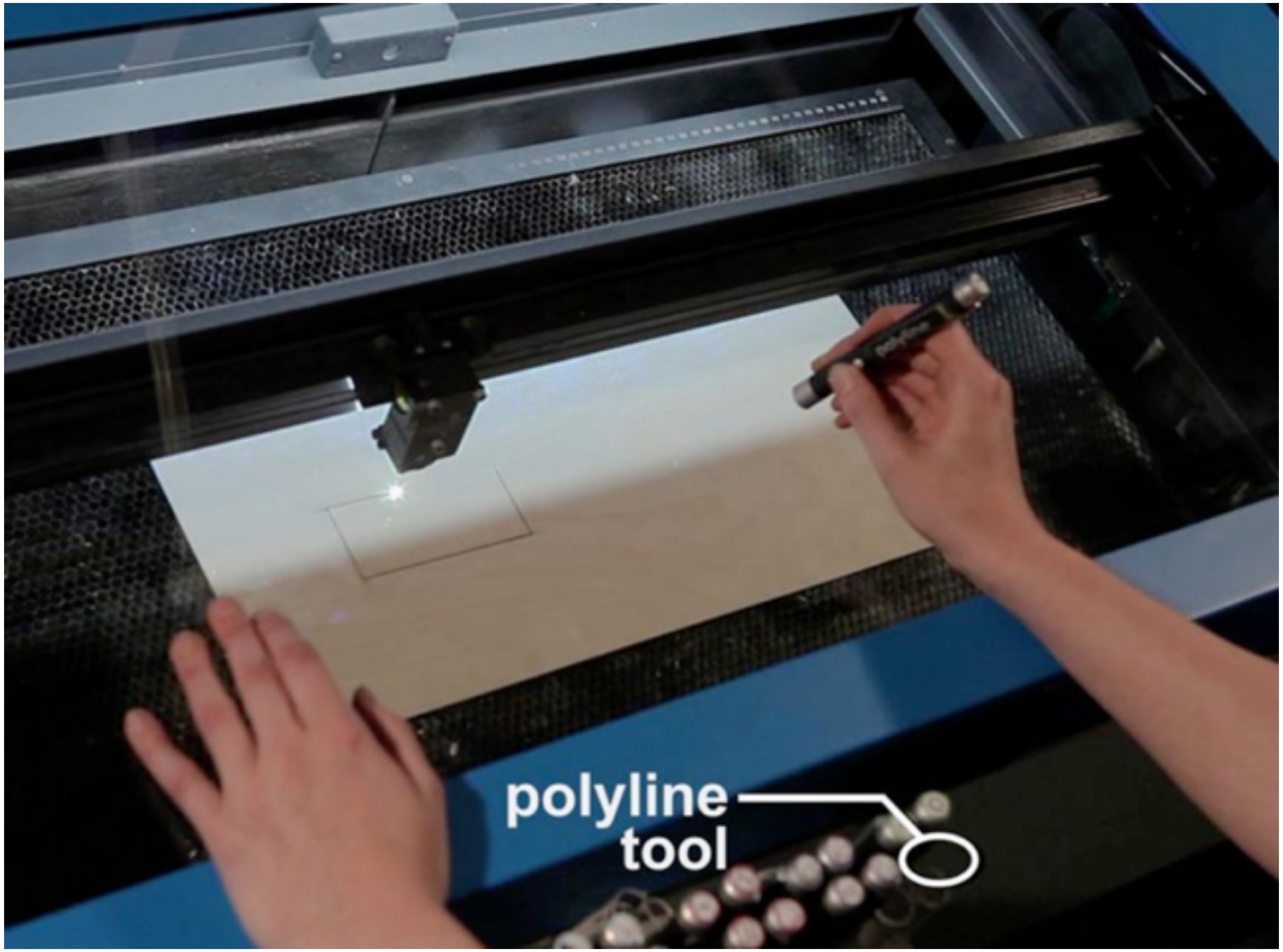
has anyone used speech output before?

<raise you hand>

hard to master,
but **easy to get started** with…

in terminal:
> say “hello”

great for **debugging spatial UI** when you have no text-output



sound interfaces:

	output	**input**
speech	text-to-speech synthesizers	siri, google glass ‘put that there’
non-speech	notifications, earcons	microphones as sensors

#2
non-speech output

we get a lot of **feedback** from out computers
… often in the form of..

earcon:

- auditory equivalent of an icon (eye + con)

can you come up
with other examples?

<30s brainstorming>

Mapping	Visual	Auditory
Symbolic	 for stop	**Sirens** for approaching ambulance
	 for peace	**Applause** for approval
Metaphorical	 for organization of relationships	**Hiss** for snake
	 for horses	**Pitch** for falling
Nomic	 for scissors	**Mailbox sound** for arriving mail
	 for file	**Hit wood or metal sound** for size of object

- **arbitrary**
- rely on social conventions

- **some similarities**
- but not exactly as real-world

- meaning depends on **real-world physics**

Windows
(symbolic)

Mac trash
(nomic)

(be careful with mapping)

HUMAN-COMPUTER INTERACTION, 1986, Volume 2, pp. 167-177

Auditory Icons: Using Sound in Computer Interfaces

William W. Gaver
University of California, San Diego

ABSTRACT

There is growing interest in the use of sound to convey information in computer interfaces. The strategies employed thus far have been based on an understanding of sound that leads to either an arbitrary or metaphorical relation between the sounds used and the data to be represented. In this article, an alternative approach to the use of sound in computer interfaces is outlined, one that emphasizes the role of sound in conveying information about the world to the listener. According to this approach, auditory icons, caricatures of naturally occurring sounds, could be used to provide information about sources of data. Auditory icons provide a natural way to represent dimensional data as well as conceptual objects in a computer system. They allow categorization of data into distinct families, using a single sound. Perhaps the most important advantage of this strategy is that it is based on the way people listen to the world in their everyday lives.

HUMAN-COMPUTER INTERACTION, 1989, Volume 4, pp. 11-44

Earcons and Icons: Their Structure and Common Design Principles

Meera M. Blattner, Denise A. Sumikawa, and **Robert M. Greenberg**
Lawrence Livermore National Laboratory and
The University of California, Davis

ABSTRACT

In this article we examine *earcons*, which are audio messages used in the user–computer interface to provide information and feedback to the user about computer entities. (Earcons include messages and functions, as well as states and labels.) We identify some design principles that are common to both visual symbols and auditory messages, and discuss the use of representational and abstract icons and earcons. We give some examples of audio patterns that may be used to design modules for earcons, which then may be assembled into larger groupings called *families*. The modules are single pitches or rhythmicized sequences of pitches called *motives*. The families are constructed about related motives that serve to identify a family of related messages. Issues concerned with learning and remembering earcons are discussed.

sound interfaces:

	output	**input**
speech	text-to-speech synthesizers	siri, google glass ‘put that there’
non-speech	notifications, earcons	microphones as sensors

#3
speech input

1980! ‘Put that there’

"Put-That-There": Voice and Gesture at the Graphics Interface

Richard A. Bolt

Architecture Machine Group
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

Recent technological advances in connected-speech recognition and position sensing in space have encouraged the notion that voice and gesture inputs at the graphics interface can converge to provide a concerted, natural user modality.

The work described herein involves the user commanding simple shapes about a large-screen graphics display surface. Because voice can be augmented with simultaneous pointing, the free usage of pronouns becomes possible, with a corresponding gain in naturalness and economy of expression. Conversely, gesture aided by voice gains precision in its power to reference.

Key Words: Voice input; speech input; gesture; space sensing; spatial data management; man-machine interfaces; graphics; graphics interface.

Category Numbers: 8.2, 6.9.

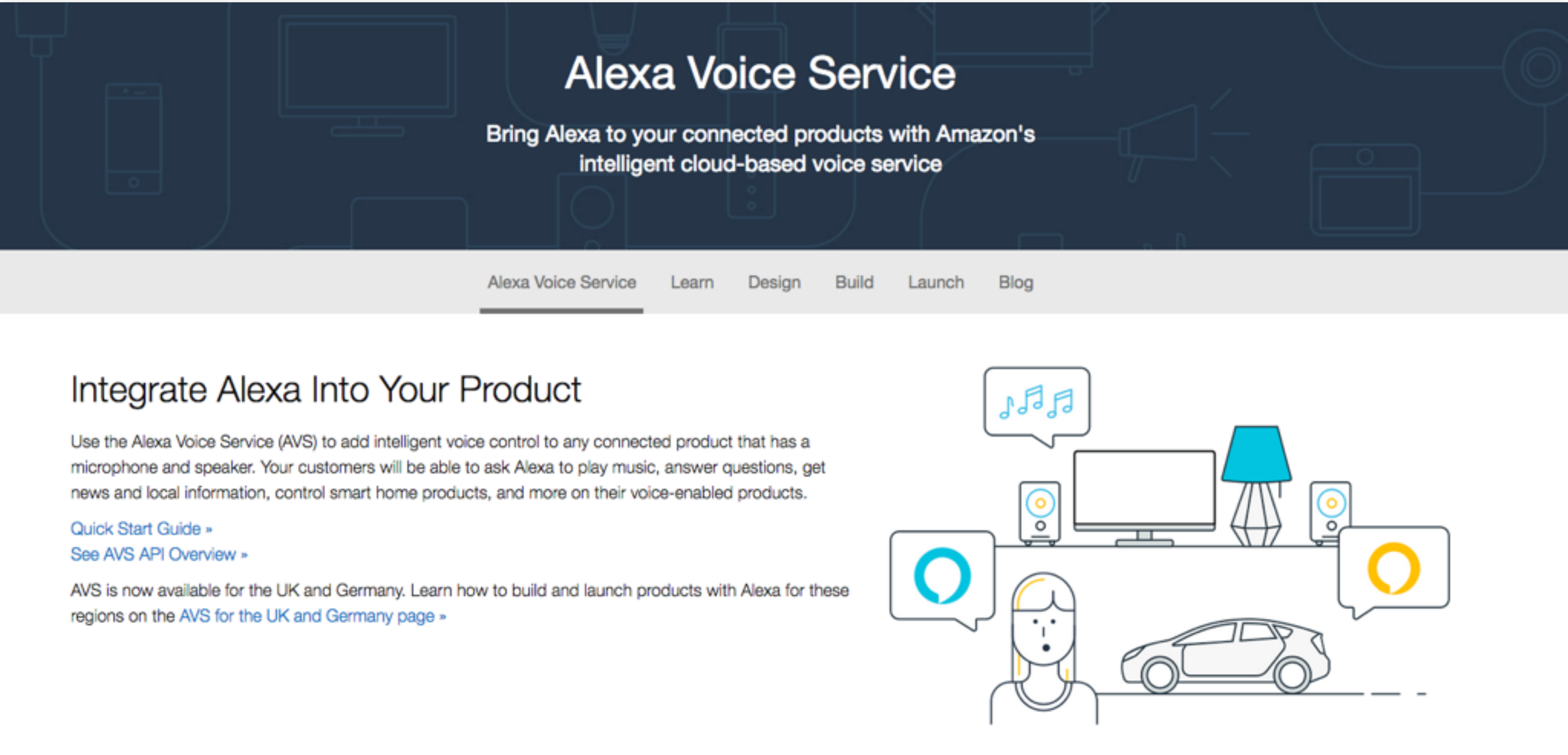
benefits of speech input:

- voice commands are **useful when hands-busy**
- voice commands make interactions very **“human”-like**

(“create yellow circle here” rather than
“click” > “menu: select shape” > “circle” > “click)

Google's electronic eyewear gets 'OK Glass' voice commands

Hoping to carve out a new type of personal computing, Google shows off how to use its computerized eyewear to search, navigate, chat, and take photos.

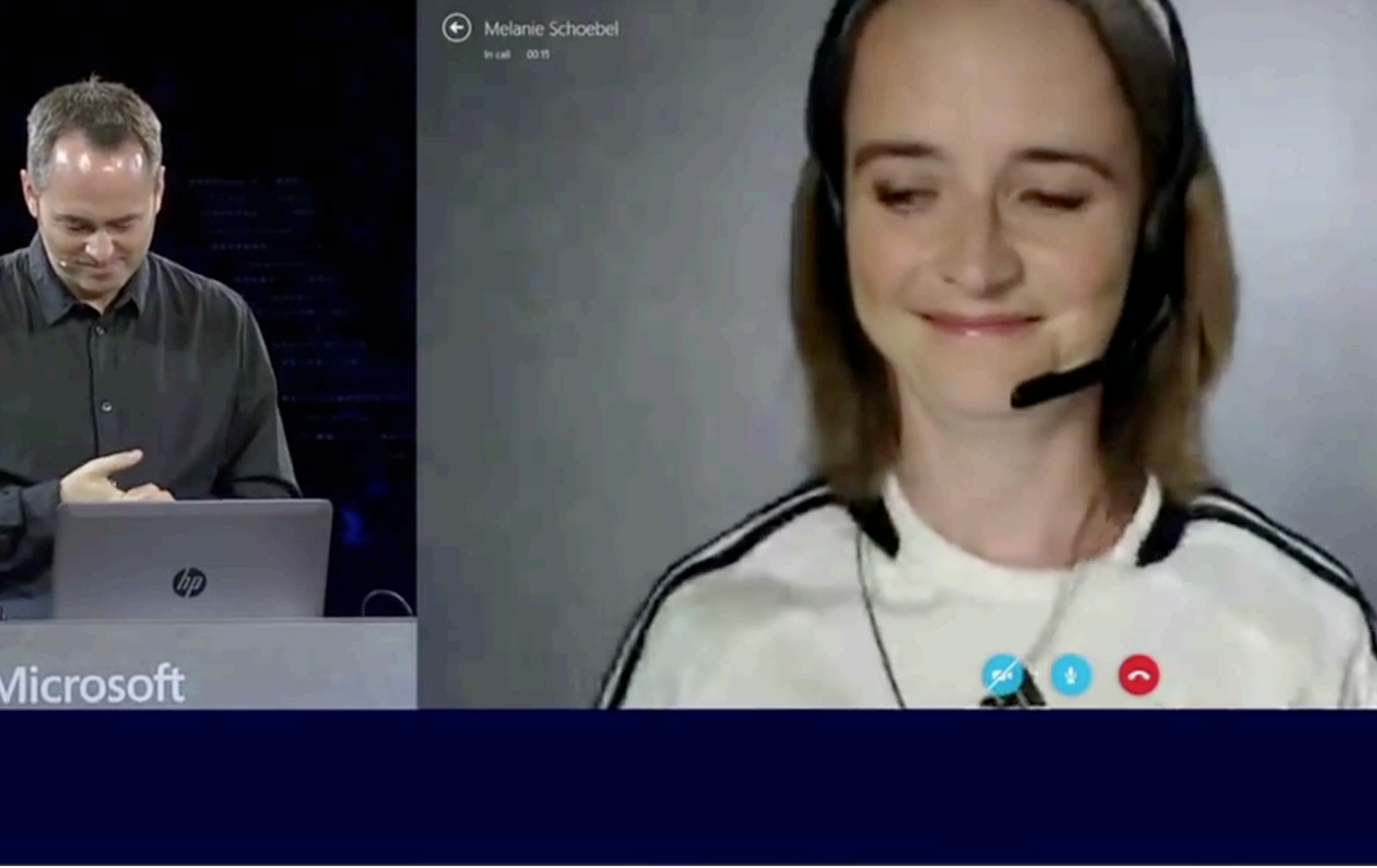


easy to integrated today!
has anyone used this for their projects / classes?

<raise you hand>

live translation - holy grail

- **speech-to-text**
- translate
- text-to-speech



2014: Microsoft Skype Translator demo

live translation - holy grail

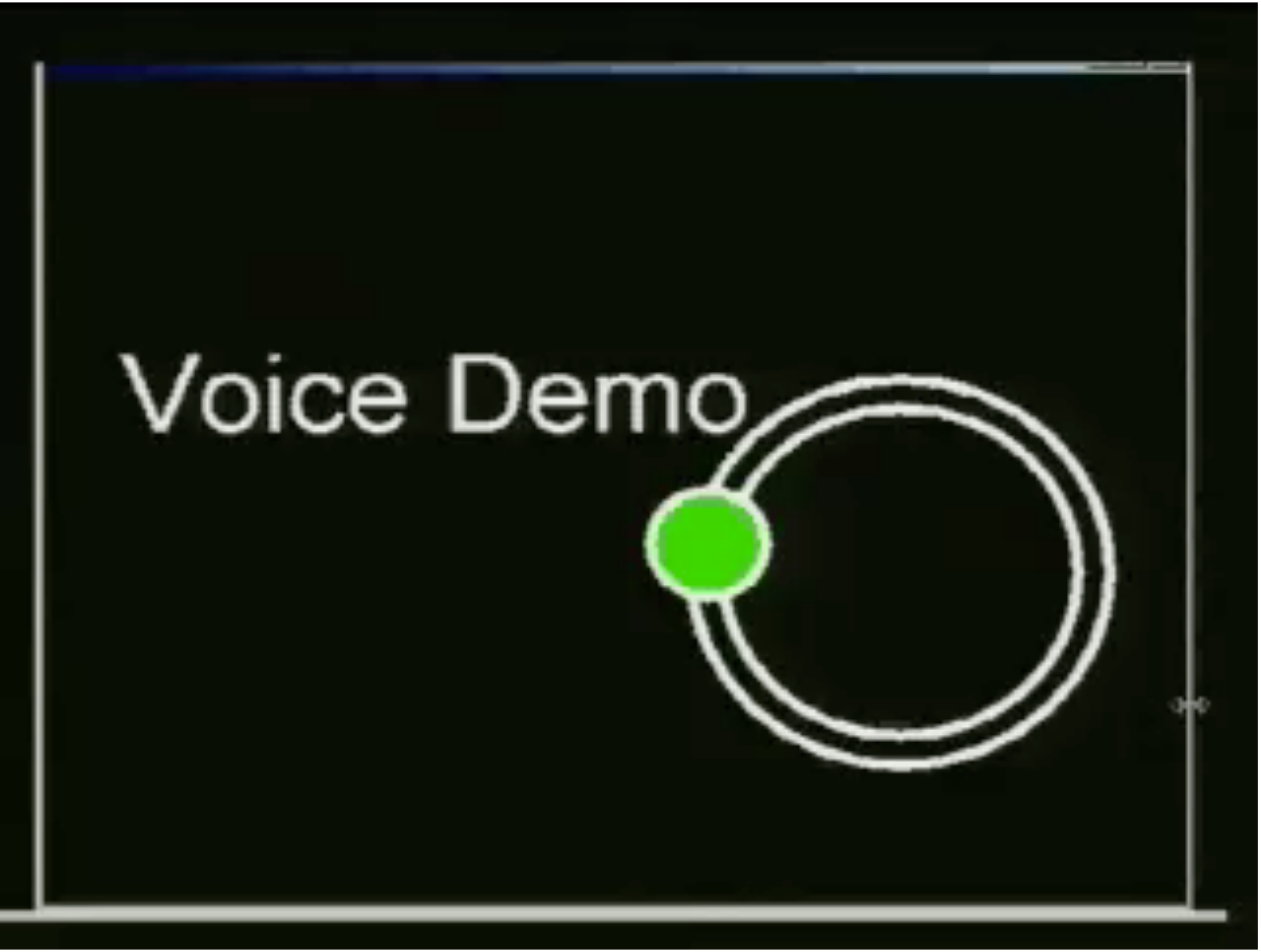
- **speech-to-text**
- translate
- text-to-speech

sound interfaces:

	output	input
speech	text-to-speech synthesizers	siri, google glass ‘put that there’
non-speech	notifications, earcons	microphones as sensors

#4
non-speech input

command computer with voice… but not speech..



2001: voice as sound:
more **immediate feedback** in contrast to put-that-there

Voice as Sound: Using Non-verbal Voice Input for Interactive Control

Takeo Igarashi *John F. Hughes*
Computer Science Department, Brown University
115 Waterman Street, Providence, RI 02912, USA
takeo@acm.org, jfh@cs.brown.edu

ABSTRACT

We describe the use of non-verbal features in voice for direct control of interactive applications. Traditional speech recognition interfaces are based on an indirect, conversational model. First the user gives a direction and then the system performs certain operation. Our goal is to achieve more direct, immediate interaction like using a button or joystick by using lower-level features of voice such as pitch and volume. We are developing several prototype interaction techniques based on this idea, such as "control by continuous voice", "rate-based parameter control by pitch," and "discrete parameter control by tonguing." We have implemented several prototype systems, and they suggest that voice-as-sound techniques can enhance traditional voice recognition approach.

KEYWORDS: Voice, Interaction technique, direct manipulation, entertainment.

INTRODUCTION

Typical voice-based interfaces focus primarily on the verbal aspects of human speech. Speech recognition engine turns speech into words or sentences, and the system performs appropriate actions based on recognized texts. One of the limitations of these approaches is that the interaction turnaround is long. The user must complete a word and wait for the recognition results. While this is reasonable for complicated tasks like flight reservation, it is inappropriate for direct, low-level controls such as scrolling. This paper proposes the use of non-verbal features in speech, features like pitch, volume, and continuation, to directly control

Figure 1. The user controls the application directly using continuous voice, and the system provides immediate feedback.

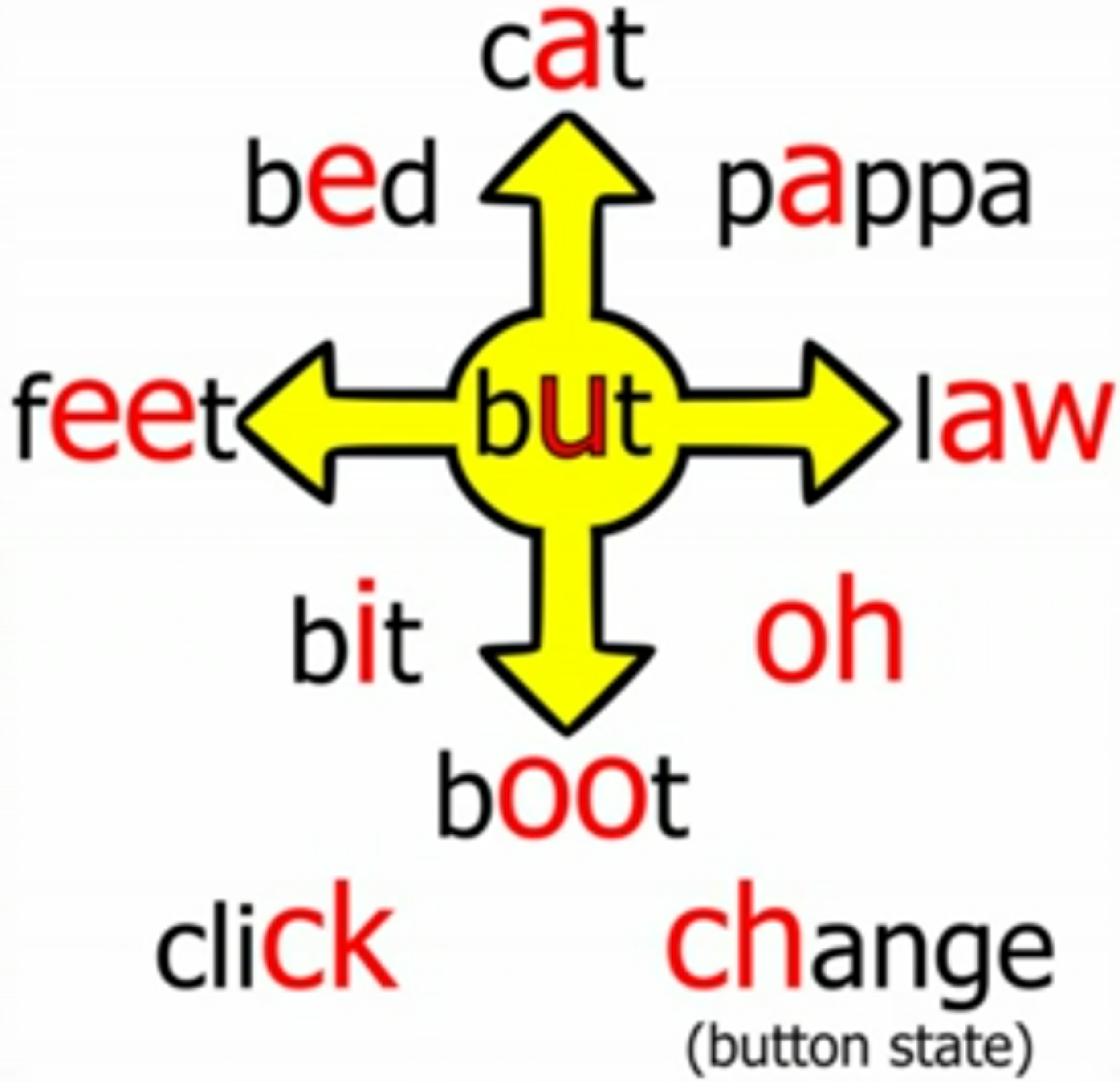
her utterances, and adjusts its behavior accordingly [7]. Goto *et. al.* described a voice-completion interface [1] that detects each filled pause during an utterance as a trigger and automatically completes the utterance like tab-key completing commands in Unix shells.

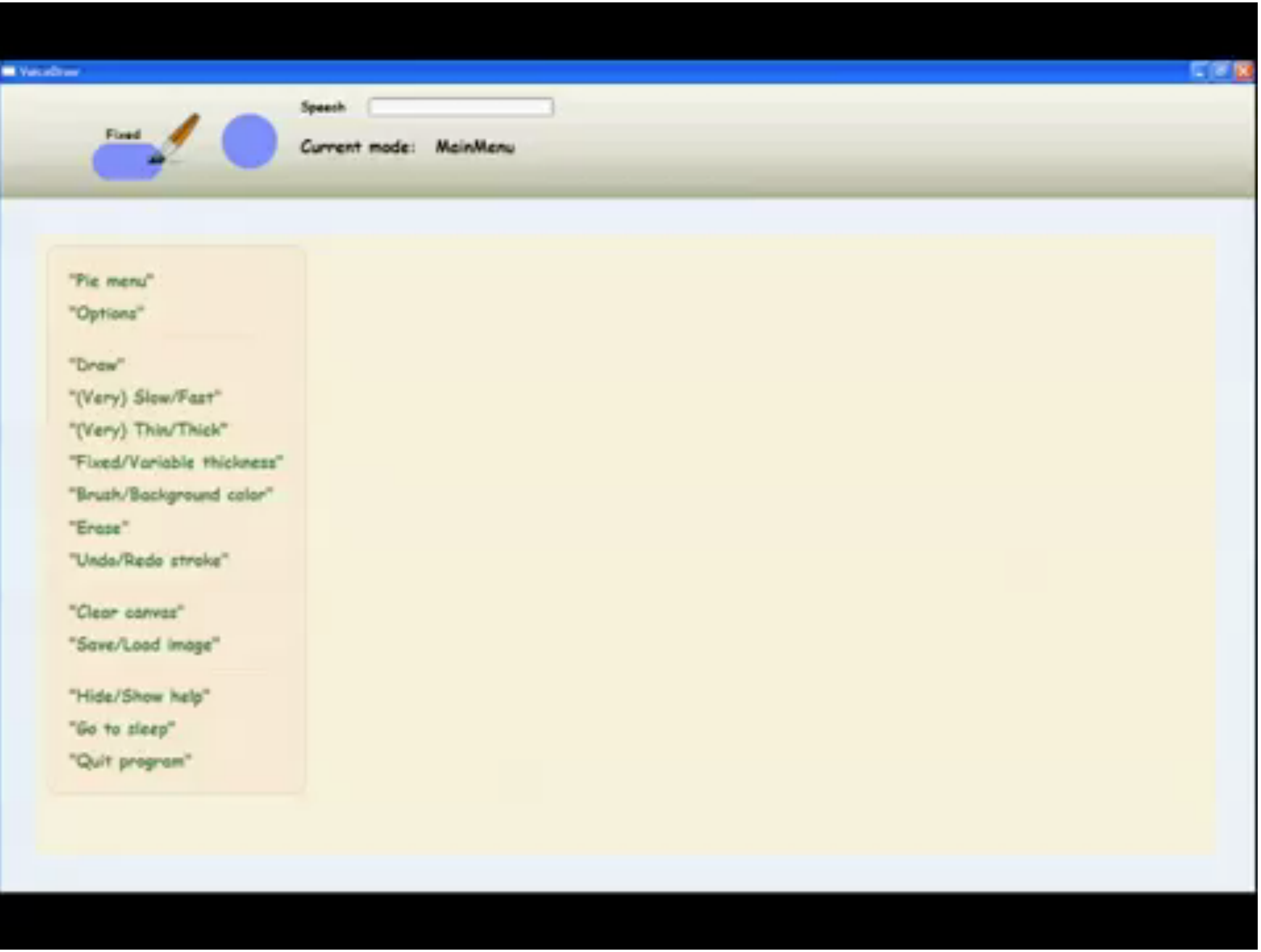
The SUITEKey system is a speech interface for controlling a virtual keyboard and mouse for motor-disabled users [6]. When the user says "move mouse down … stop", the mouse pointer moves downward during the pause. Our techniques extend this work by introducing additional interaction techniques for voice-based direct manipulation.

INTERACTION TECHNIQUES

Control by Continuous Voice

In this interface, the user's voice works as an on/off button.


cat
bed
pappa
feet
but
law
bit
oh
boot
click
change
(button state)



2007: voice draw

VoiceDraw: A Hands-Free Voice-Driven Drawing Application for People with Motor Impairments

Susumu Harada
Computer Science and Engineering
Box 352350
University of Washington
Seattle, WA 98195 USA
harada@cs.washington.edu

Jacob O. Wobbrock
The Information School
Box 352840
University of Washington
Seattle, WA 98195 USA
wobbrock@u.washington.edu

James A. Landay
Computer Science and Engineering
Box 352350
University of Washington
Seattle, WA 98195 USA
landay@cs.washington.edu

ABSTRACT

We present VoiceDraw, a voice-driven drawing application for people with motor impairments that provides a way to generate free-form drawings without needing manual interaction. VoiceDraw was designed and built to investigate the potential of the human voice as a modality to bring fluid, continuous direct manipulation interaction to users who lack the use of their hands. VoiceDraw also allows us to study the issues surrounding the design of a user interface optimized for non-speech voice-based interaction. We describe the features of the VoiceDraw application, our design process, including our user-centered design sessions with a "voice painter," and offer lessons learned that could inform future voice-based design efforts. In particular, we offer insights for mapping human voice to continuous control.

Categories and Subject Descriptors

H.5.2 [**Information interfaces and presentation**]: User Interfaces – *Voice I/O*.

General Terms

Design, Human Factors.

Keywords

Voice-based user interfaces, speech recognition, drawing, painting, computer art, continuous input, motor impairments.

1. INTRODUCTION

Creative self-expression and artistic endeavors can play a vital role in enhancing people's quality of life, including those with various types of disabilities [22]. Despite the challenges that motor impairments pose to an individual's ability to manipulate physical art mediums such as paint brushes or drawing pencils,

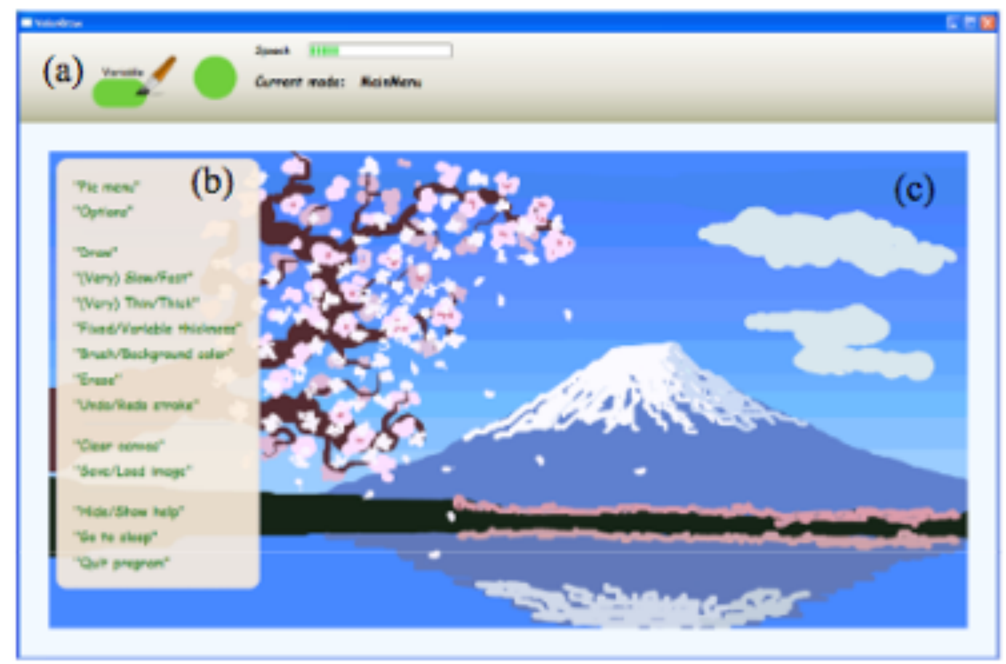


Figure 1: A screenshot of the VoiceDraw application[1] showing **(a)** the status bar, **(b)** help overlay, and **(c)** canvas area. The first author created this painting using only his voice in about 2.5 hours.

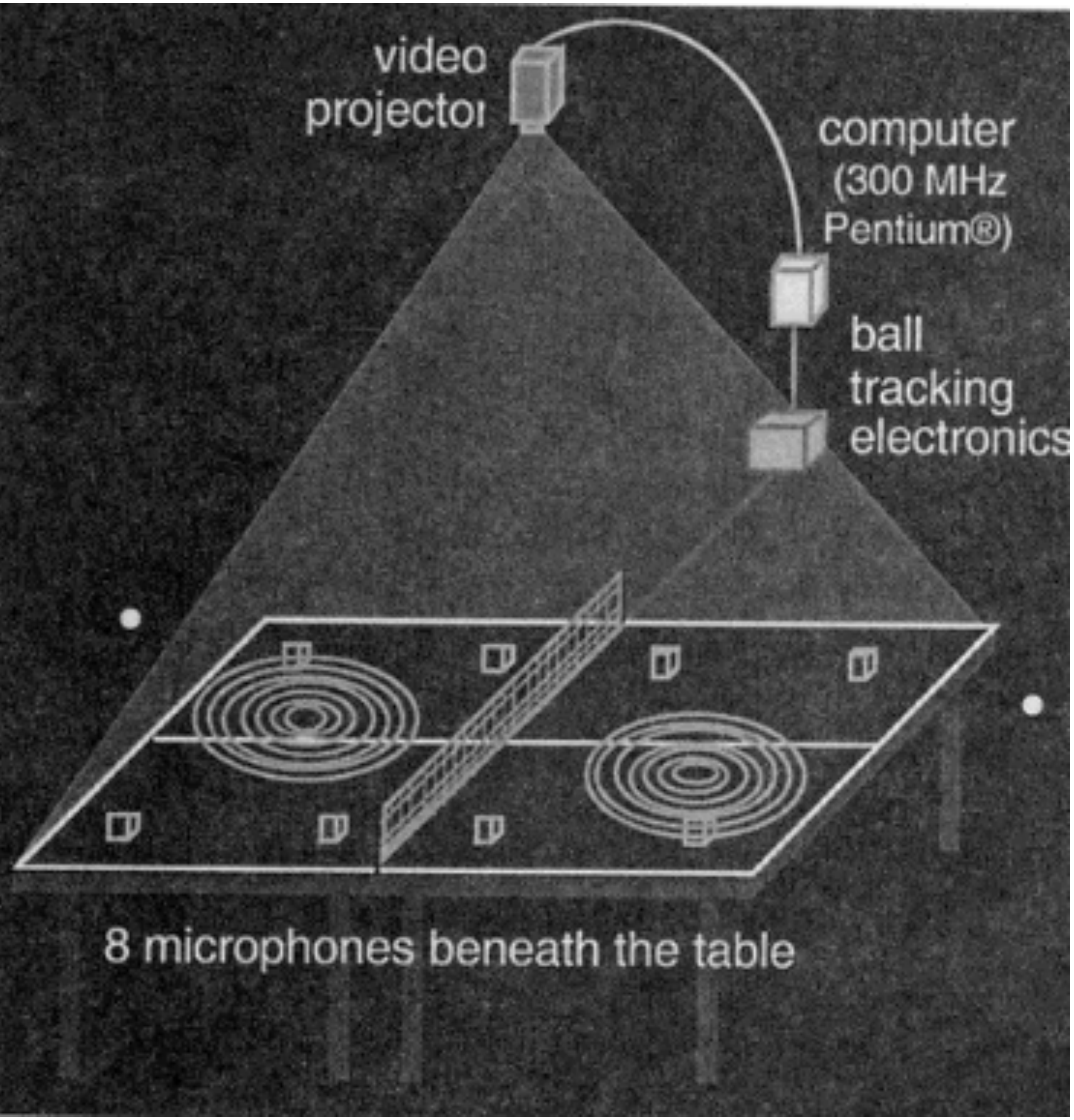
However, for those with moderate to severe motor impairments, manipulation of physical tools may be difficult or impossible. Even those with some ability to manipulate physical artistic media may find the process arduous enough to be a barrier to engaging in creative activity.

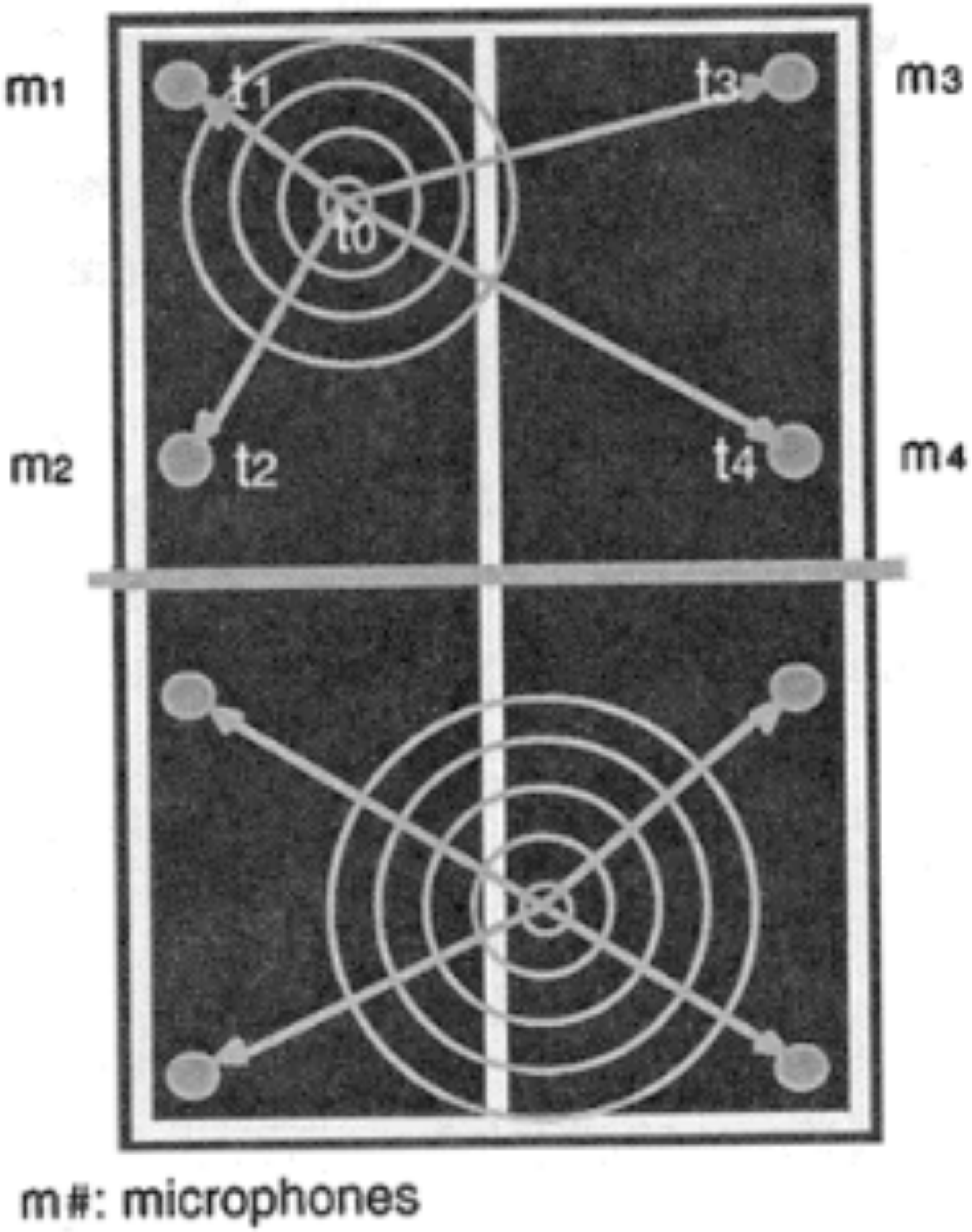
Computer applications hold promise for enabling such individuals with limited motor abilities to engage in creative activities with reduced overhead of manipulating physical tools. Painting programs on a computer can simulate physical brush strokes or even provide artistic effects not possible in the physical domain.

A challenge that limits the realization of this potential is that

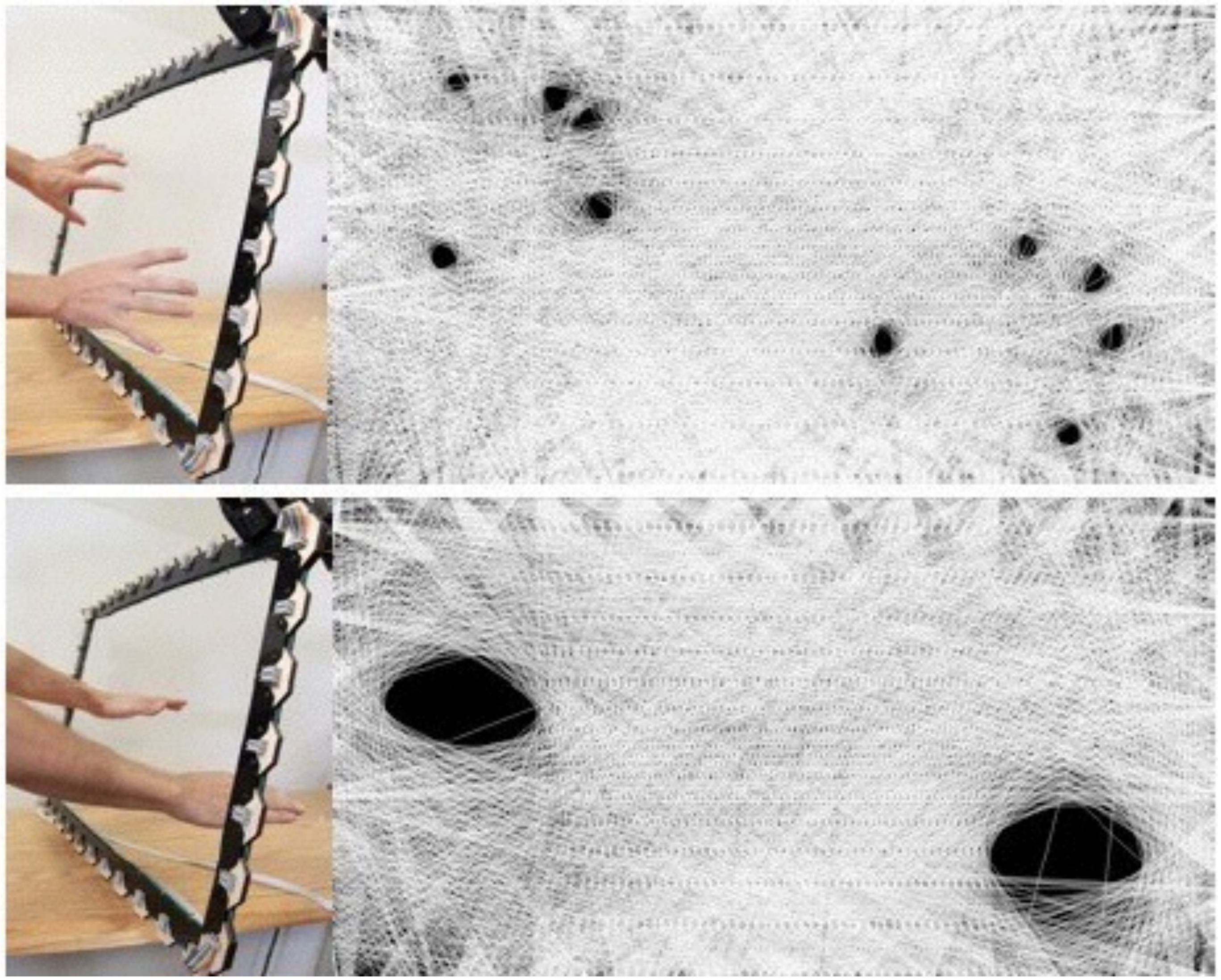
command computer with
sounds from the environment…

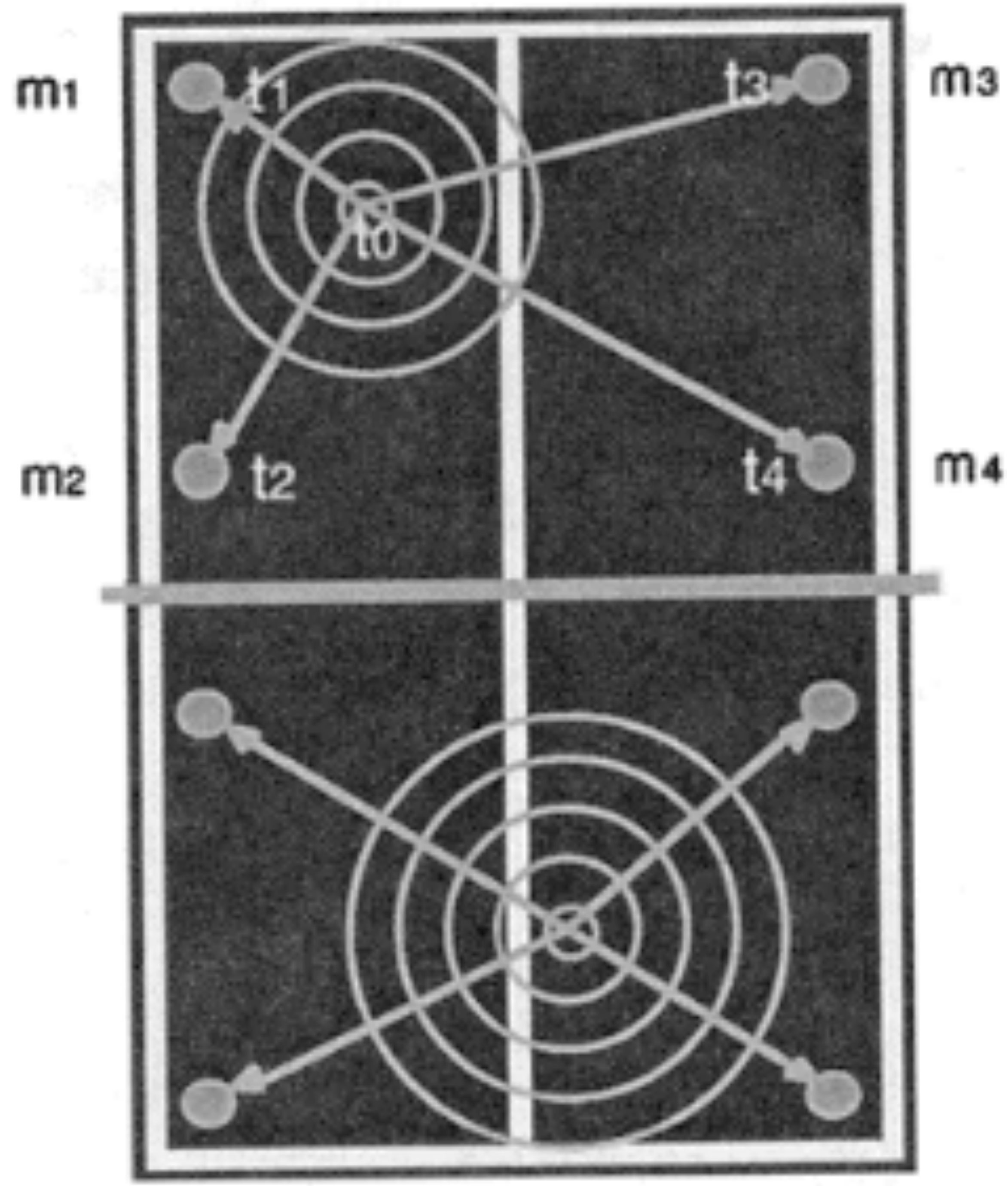
1999: pingpongplus
a ‘touch screen’ made from microphones





measure time signal needs to travel to each microphone





m#: microphones

similar triangulation
as seen in multitouch lecture

PingPongPlus: Design of an Athletic-Tangible Interface for Computer-Supported Cooperative Play

Hiroshi Ishii, Craig Wisneski, Julian Orbanes, Ben Chun, and Joe Paradiso*
Tangible Media Group
*Physics and Media Group
MIT Media Laboratory
20 Ames St., Cambridge, MA 02139, U.S.A.
{ishii, wiz, joules, benchun, joep}@media.mit.edu

ABSTRACT

This paper introduces a novel interface for digitally-augmented cooperative play. We present the concept of the "athletic-tangible interface," a new class of interaction which uses tangible objects and full-body motion in physical spaces with digital augmentation. We detail the implementation of PingPongPlus, a "reactive ping-pong table", which features a novel sound-based ball tracking technology. The game is augmented and transformed with dynamic graphics and sound, determined by the position of impact, and the rhythm and style of play. A variety of different modes of play and initial experiences with PingPongPlus are also described.

Keywords

tangible interface, enhanced reality, augmented reality, interactive surface, athletic interaction, kinesthetic interaction, computer-supported cooperative play.

INTRODUCTION

When an expert plays ping-pong, a well-used paddle becomes *transparent,* and allows a player to concentrate on the task – playing ping-pong. The good fit of grasp is vital to making a paddle transparent [10]. To achieve a "good fit," a user has to choose a paddle of the right size, right form, and right weight for his or her hand and style of play. To achieve a "better fit," the user has to *customize* the tool by scraping the edge of the paddle with a knife and

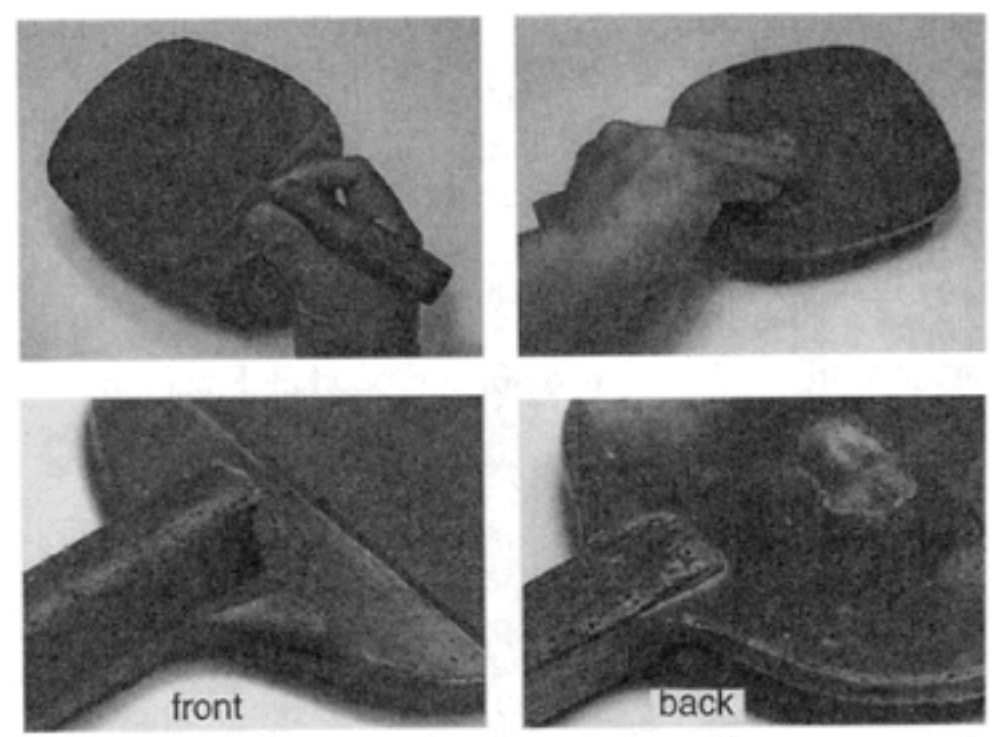
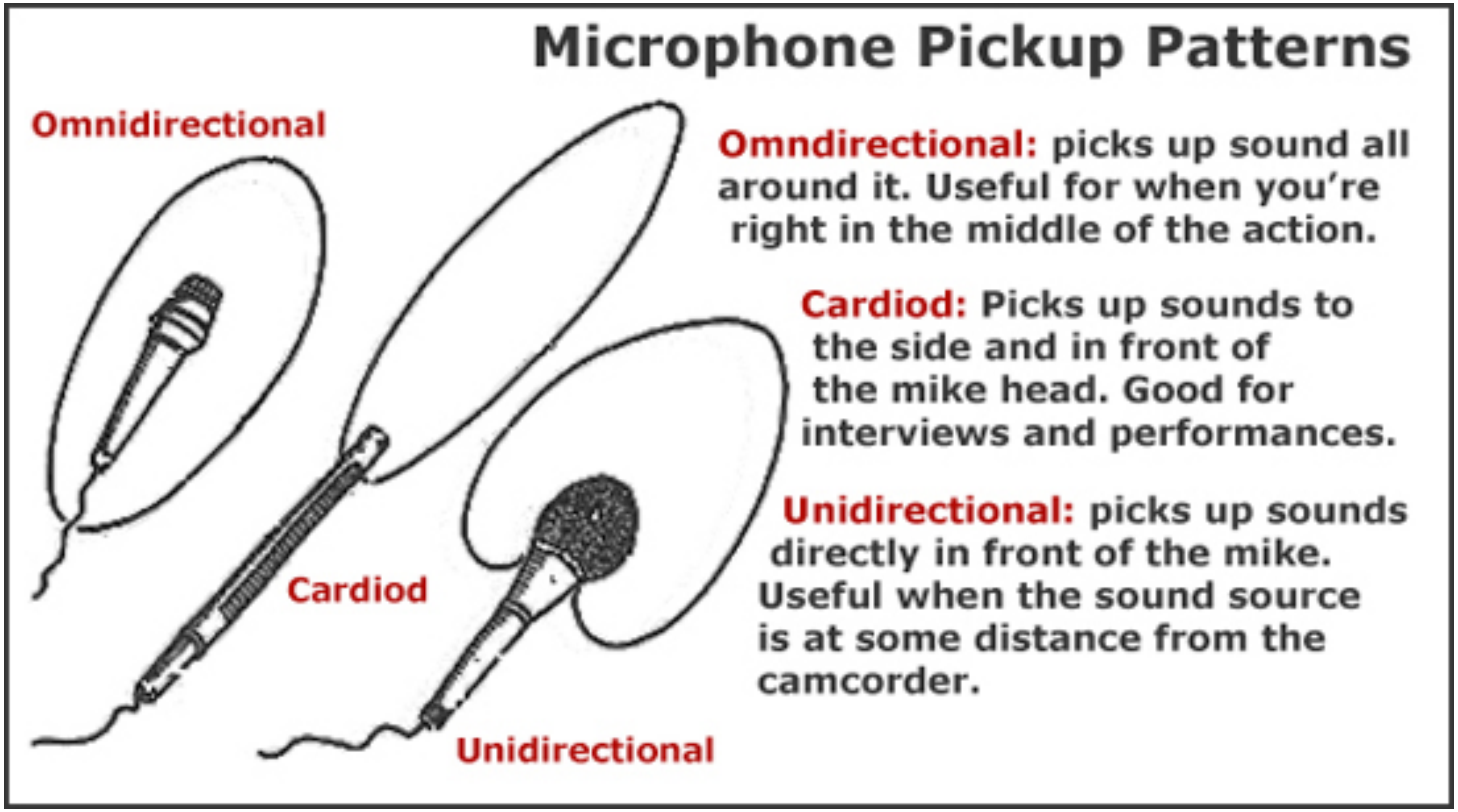


Figure 1 Traces of grasping hand left on the well-used ping-pong paddle

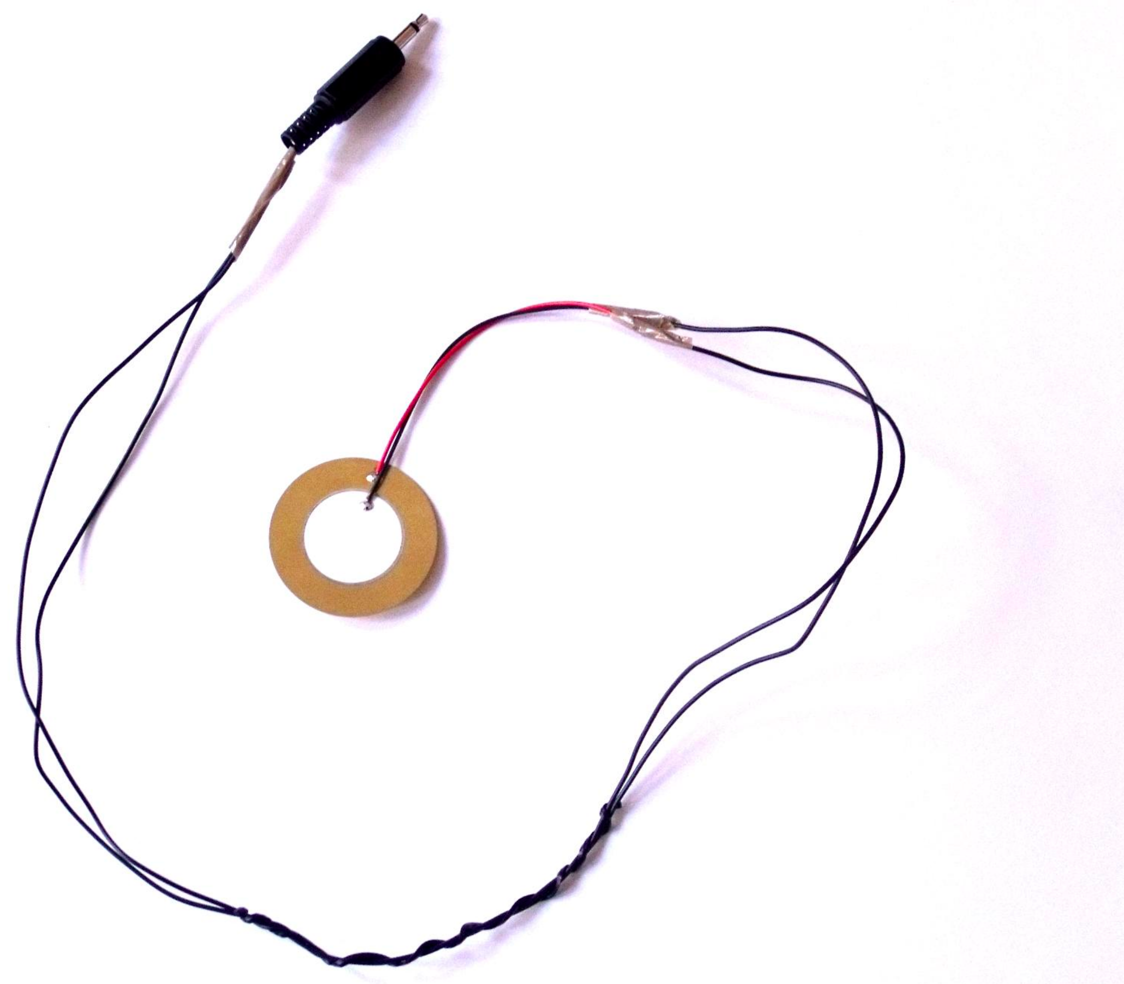
Moreover, the full-body motion, speed, and rhythm of a ping-pong game make the interaction very engaging and entertaining. Kinesthesia is one of the keys of what makes ping-pong enjoyable.

Modem graphical user interface (GUI) technologies provide very limited, generic physical forms (e.g. mouse, keyboard, and monitor) and allow limited physical motions (only



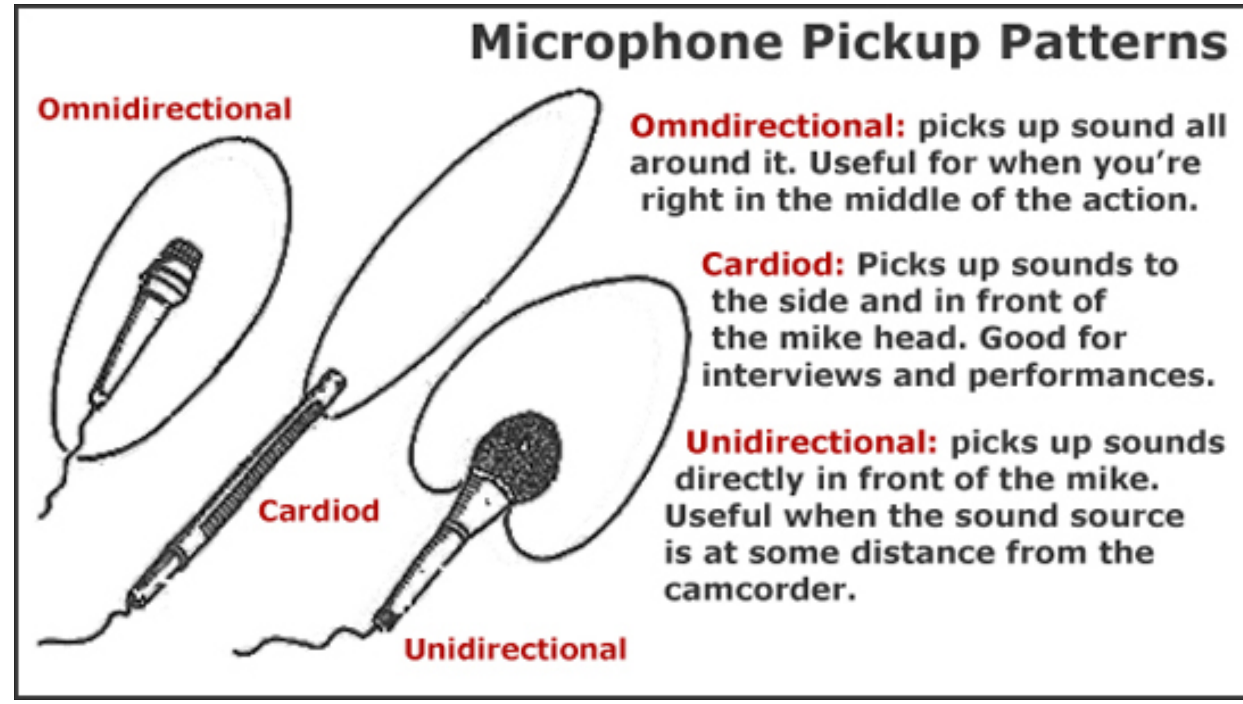
which of these
would you use for **making a touch sensor?**

<30s brainstorming>



contact microphone:
vibration through surfaces

(if you speak into this,
nothing happens)



other microphone:
vibration through air

none of these good
for touch sensor

2002: before any major touch screen technology

Passive Acoustic Knock Tracking for Interactive Windows

Joseph A. Paradiso, Che King Leo, Nisha Checka, Kaijen Hsiao
Responsive Environments Group
MIT Media Laboratory
1 Cambridge Center, 5FL
Cambridge, MA 02142 USA
{joep,cheking,nchecka,kjhsiao}@media.mit.edu

ABSTRACT
We describe a novel interface that locates and characteriz s knocks and taps atop a large glass window. Our curre t setup uses four contact piezoelectric pickups located ne r the sheet's corners to record the acoustic wavefront comi g from the knocks. A digital signal processor extrac s relevant characteristics from these signals, such s amplitudes, frequency components, and different l timings, which are used to estimate the location of the t and provide other parameters, including the rough accura y of this estimate, the nature of each hit (e.g., knuckle knoc , metal tap, or fist bang), and the strike intensity. This syste n requires only simple hardware, needs no special adaptati n of the glass pane, and allows all transducers to be mount d on the inner surface, hence it is quite easy to deploy as a retrofit to existing windows. This opens many applicatio , such as an interactive storefront, with projected conte t controlled by knocks on the display window.

Keywords
touch screen, interactive surface, acoustic tracking

INTRODUCTION
Glass is now a very common construction material, often used as clear walls for room dividers or large windows enclosing urban buildings. The techniques described in this paper enable these surfaces to become interactive. For example, information displayed on a projection or monitor on the inside of the glass can be navigated by knocking

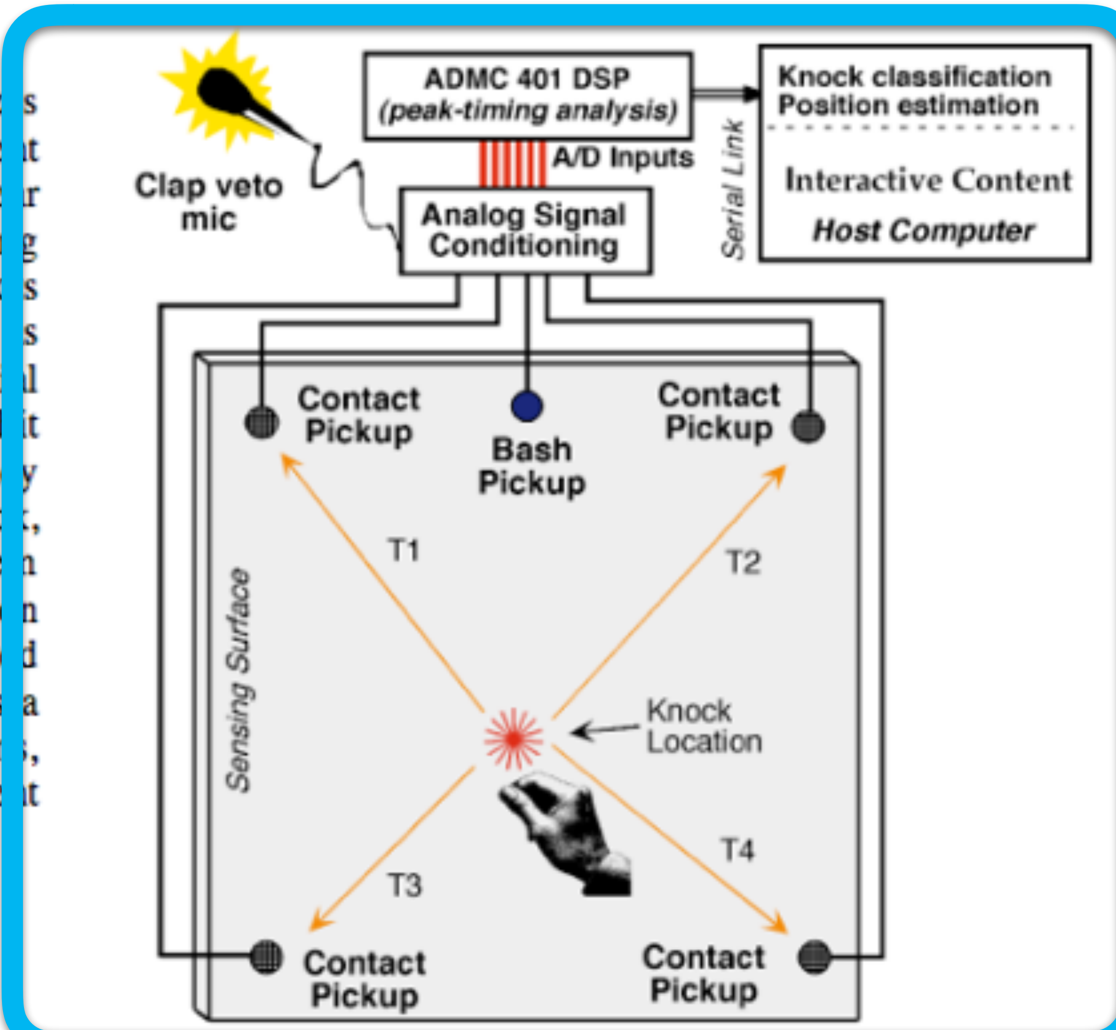


Figure 1: Hardware configuration for knock tracker system

Capacitive and active acoustic techniques require the glass to be patterned with transparent electrodes or waveguides, which can be expensive and problematic over large areas.

Other techniques, such as video tracking [2] have been used

also works with feet

Let's Kick It: How to Stop Wasting the Bottom Third of Your Large Scale Display

Ricardo Jota
Dept. of Computer Science
University of Toronto
Toronto, Canada
jotacosta@dgp.toronto.edu

Pedro Lopes
Hasso Plattner Institute
Potsdam, Germany
pedro.lopes@hpi.uni-potsdam.de

Daniel Wigdor
Dept. of Computer Science
University of Toronto
Toronto, Canada
dwigdor@dgp.toronto.edu

Joaquim Jorge
VIMMI / Inesc-ID
IST
Lisbon, Portugal
jaj@inesc-id.pt

ABSTRACT
Large-scale touch surfaces have been widely studied in literature and adopted for public installations such as interactive billboards. However, current designs do not take into consideration that touching the interactive surface at different heights is not the same; for body-height displays, the bottom portion of the screen is within easier reach of the foot than the hand. We explore the design space of foot input on vertical surfaces, and propose three distinct interaction modalities: hand, foot tapping, and foot gesturing. Our design exploration pays particular attention to areas of the touch surface that were previously overlooked: out of hand's reach and close to the floor. We instantiate our design space with a working prototype of an interactive surface, in which we are able to distinguish between finger and foot tapping and extend the input area beyond the bottom of the display to support foot gestures.

Author Keywords
Large-scale display, foot interaction, kick, floor input.

ACM Classification Keywords
H.5.2. [Information interfaces and presentation]: User Interfaces. – Graphical user interfaces.

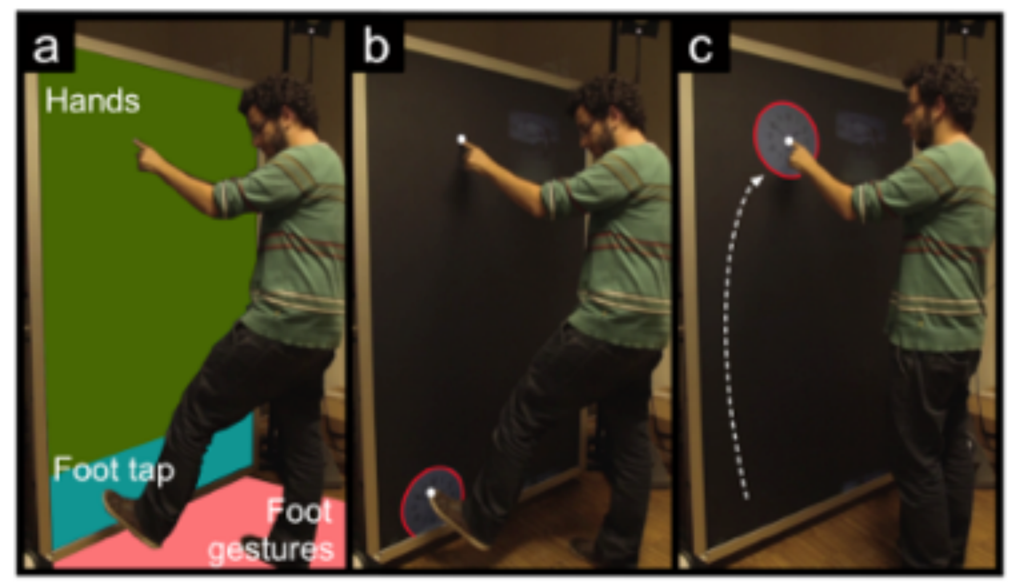
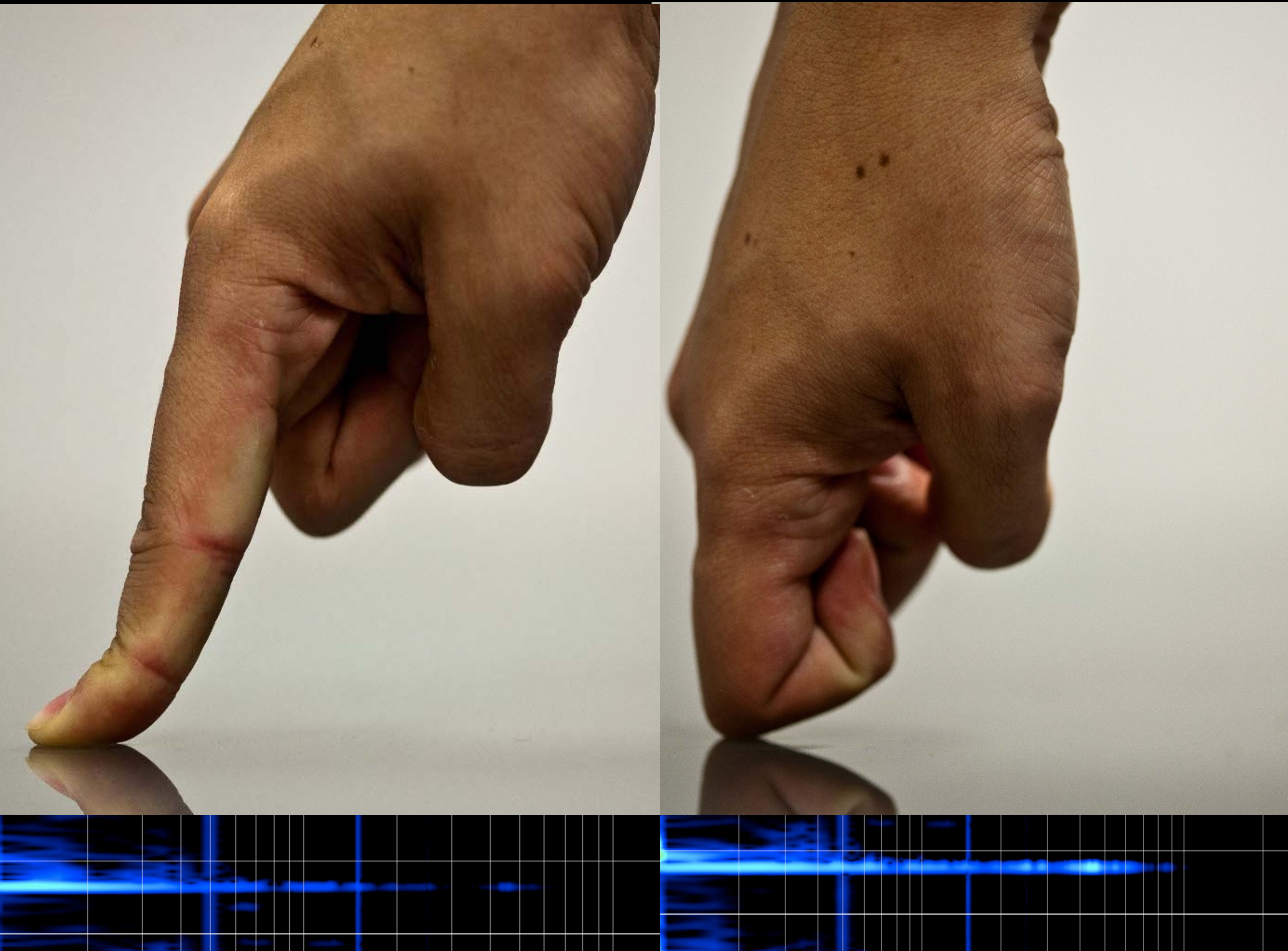


Figure 1. (a) We propose three types of input, each ergonomically fitting one interactive region: hand interaction, foot taps, and foot gestures; (b) and (c) shows an example technique: kicking an object in the lower portion of the display causes it to pop-up to the finger position.

INTRODUCTION
Researchers foresee a future in which all walls, windows, and doors – indeed, all vertical surfaces – hold the potential to

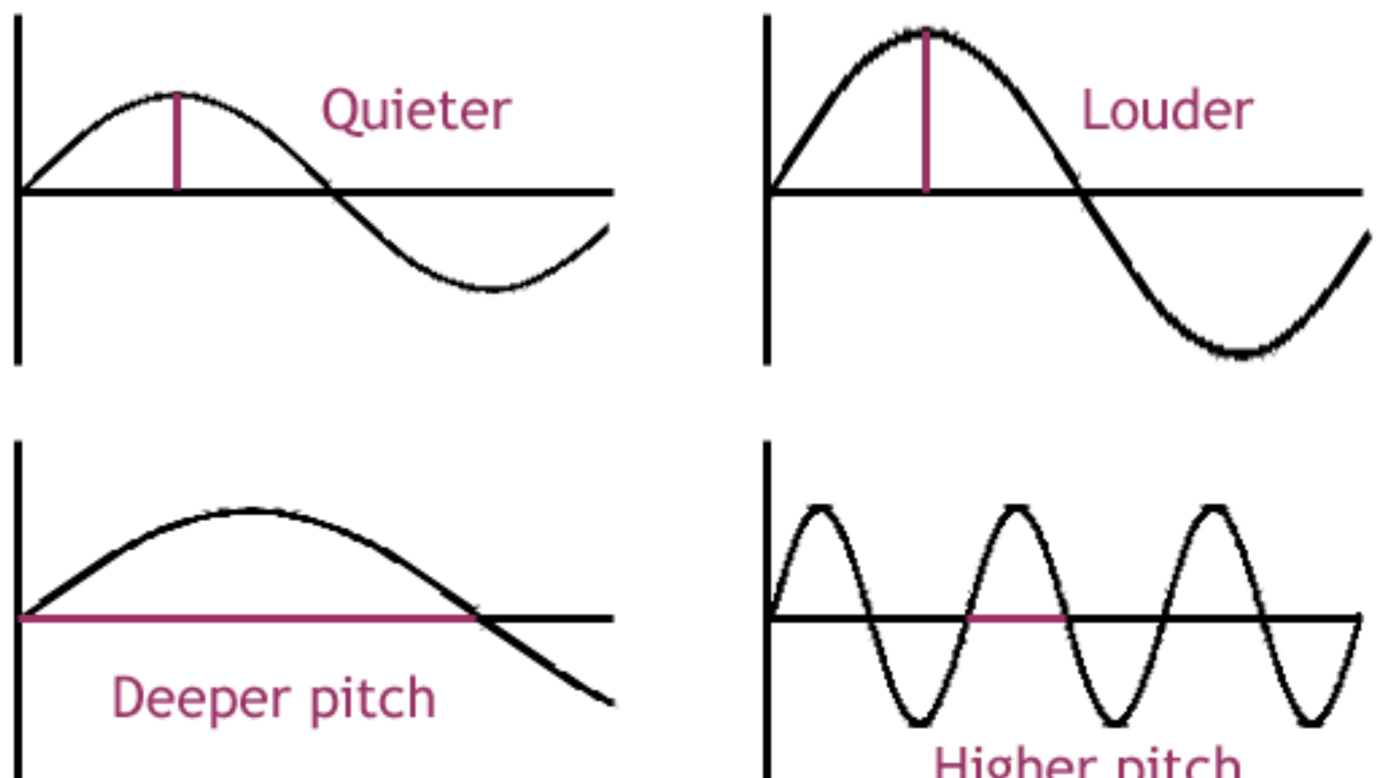
In the present work, we address exactly this issue: "how to facilitate interaction with the lower region of touch displays, which is beyond hand-reach?". To explore this concern, we developed a vertical touch surface, sensitive to both feet and finger input, which is depicted in Figure 1. In our early explorations, it became immediately apparent that, for adult users, the area towards the bottom of the screen was far more easily reached with the feet than with the hands; we often found ourselves kicking at items on the

microphones are still used today
as part of touch sensing in mobile phones.

acoustic features:

acoustic signature of the touch (frequency, amplitude)



Augmenting Touch Interaction Through Acoustic Sensing

Pedro Lopes *Ricardo Jota* *Joaquim A. Jorge*
INESC-ID, IST, Technical University of Lisbon
Rua Alves Redol, 9, 1000-029 Lisboa, Portugal
{pedro.lopes,jotacosta}@ist.utl.pt, jaj@inesc.pt

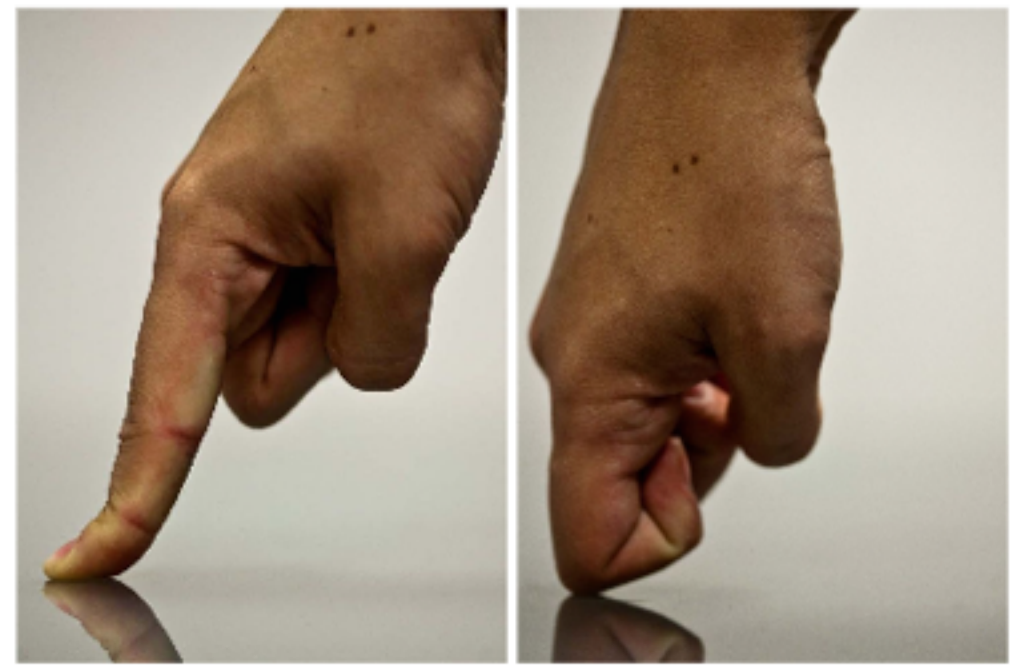
ABSTRACT
Recognizing how a person actually touches a surface has generated a strong interest within the interactive surfaces community. Although we agree that touch is the main source of information, unless other cues are accounted for, user intention might not be accurately recognized. We propose to expand the expressiveness of touch interfaces by augmenting touch with acoustic sensing. In our vision, users can naturally express different actions by touching the surface with different body parts, such as fingers, knuckles, fingernails, punches, and so forth - not always distinguishable by touch technologies but recognized by acoustic sensing. Our contribution is the integration of touch and sound to expand the input language of surface interaction.

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

General terms: Design, Human Factors

Keywords: Touch, Acoustics, Interaction

INTRODUCTION
The introduction of touch technology strongly influenced direct input interfaces. Indeed, in certain scenarios, input devices such as mouse or pens are being replaced by touch, which now supports multiple fingers [2] and gesture recognition [10]. While diversifying the input language, these advances are limited to the hand position and shape. We argue that user intention cannot be fully understood, if touch location and shape are the only cues captured. The action of touching a surface generates a contact that can be sensed

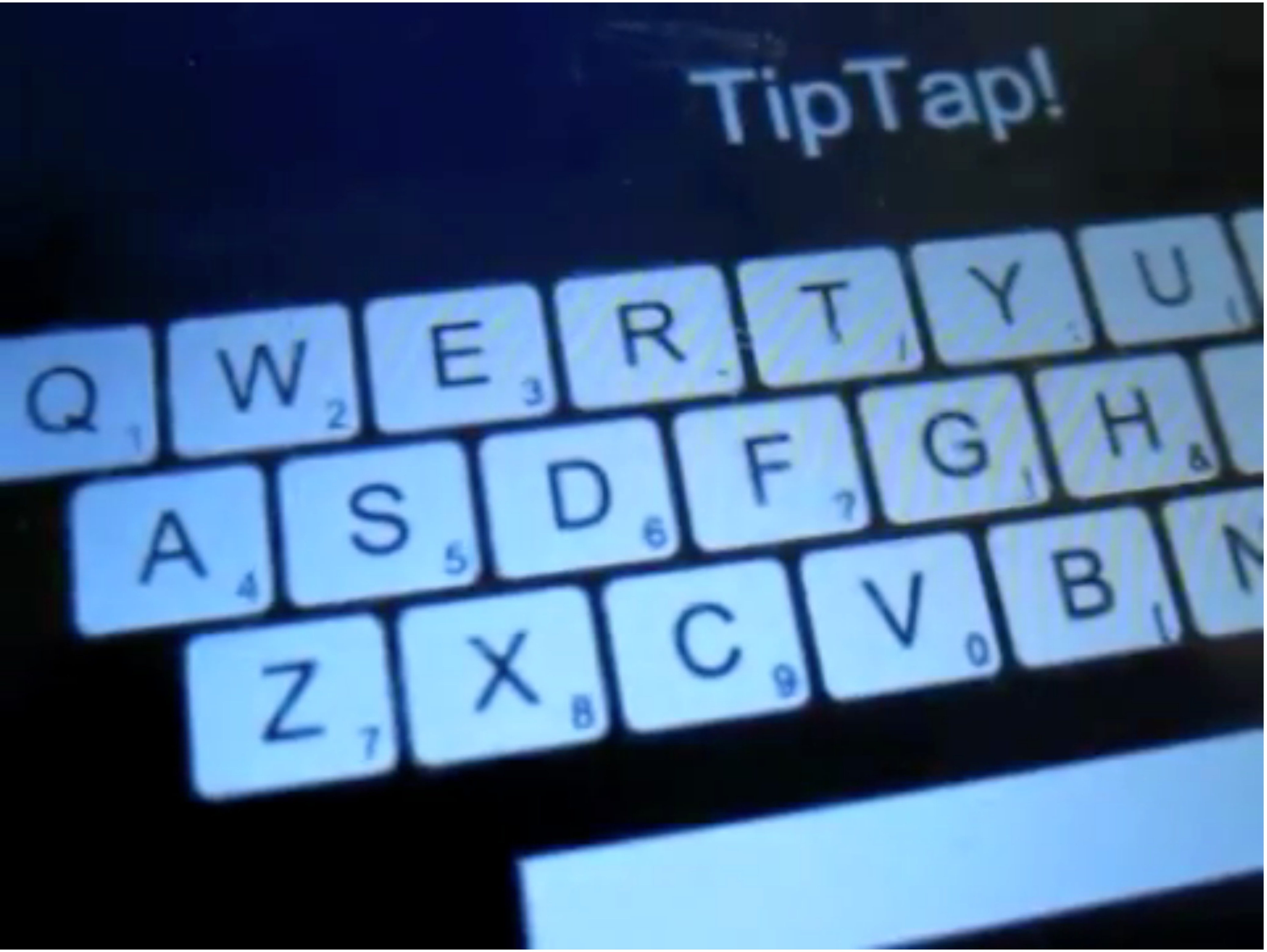


(a) Finger Touch (b) Knuckle Touch

(c) Finger Contact (d) Knuckle Contact

Figure 1: Most technologies cannot distinguish between contacts with similar signatures.

contribution of our research is not to replace available touch technology, but rather to provide additional cues to be integrated in current technologies. Thus, expanding the input language of multi-touch technologies with acoustic gestures, and extending the interaction space to surrounding regions, such as bezel and casing, that commonly do not have any



implicit **mode switching** -> less graphical UI

TapSense: Enhancing Finger Interaction on Touch Surfaces

Chris Harrison *Julia Schwarz* *Scott E. Hudson*

Human-Computer Interaction Institute and Heinz College Center for the Future of Work
Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213
{chris.harrison, julia.schwarz, scott.hudson}@cs.cmu.edu

ABSTRACT

We present *TapSense*, an enhancement to touch interaction that allows conventional surfaces to identify the type of object being used for input. This is achieved by segmenting and classifying sounds resulting from an object's impact. For example, the diverse anatomy of a human finger allows different parts to be recognized – including the tip, pad, nail and knuckle – without having to instrument the user. This opens several new and powerful interaction opportunities for touch input, especially in mobile devices, where input is extremely constrained. Our system can also identify different sets of passive tools. We conclude with a comprehensive investigation of classification accuracy and training implications. Results show our proof-of-concept system can support sets with four input types at around 95% accuracy. Small, but useful input sets of two (e.g., pen and finger discrimination) can operate in excess of 99% accuracy.

ACM Classification: H.5.2 [Information interfaces and presentation]: User Interfaces - Graphical user interfaces; Input devices and strategies.

General terms: Human Factors

Keywords: Acoustic classification, tabletop computing, interactive surfaces, tangibles, tools, pens, stylus, finger, multi-user, touchscreen, collaborative, input.

INTRODUCTION

and nail (Figure 1 and 2). The latter is especially valuable on mobile devices, where input bandwidth is limited due to small screens and "fat fingers" [16]. For example, a knuckle tap could serve as a "right click" for mobile device touch interaction, effectively doubling input bandwidth. Right-click-like functionality is currently achieved on touch surfaces with fairly unintuitive and un-scalable chording of fingers and tap-and-hold interactions. Finally, our approach requires no electronics or sensors to be placed on the user.

RELATED APPROACHES

Many technologies exist that have the ability to digitize different types of input. There are two main touch sensing approaches: active and passive.

The key downside of active approaches is that an explicit object must be used (e.g., a special pen), which is implemented with electronics (and batteries if not tethered). For example, pens augmented with infrared light emitters on their tips can be used on the commercially available Microsoft Surface [15]. There have also been efforts to move beyond pens, including, e.g., infrared-light-emitting brushes for painting applications [27]. Current systems generally do not attempt to discriminate among different pens (just perhaps pen from finger input). Variably-modulated infrared light enables identification, but requires specialized hardware. Additionally, ultrasonics can be used for input

Qeexo Closes $4.5M Series B Round

2016-01-21

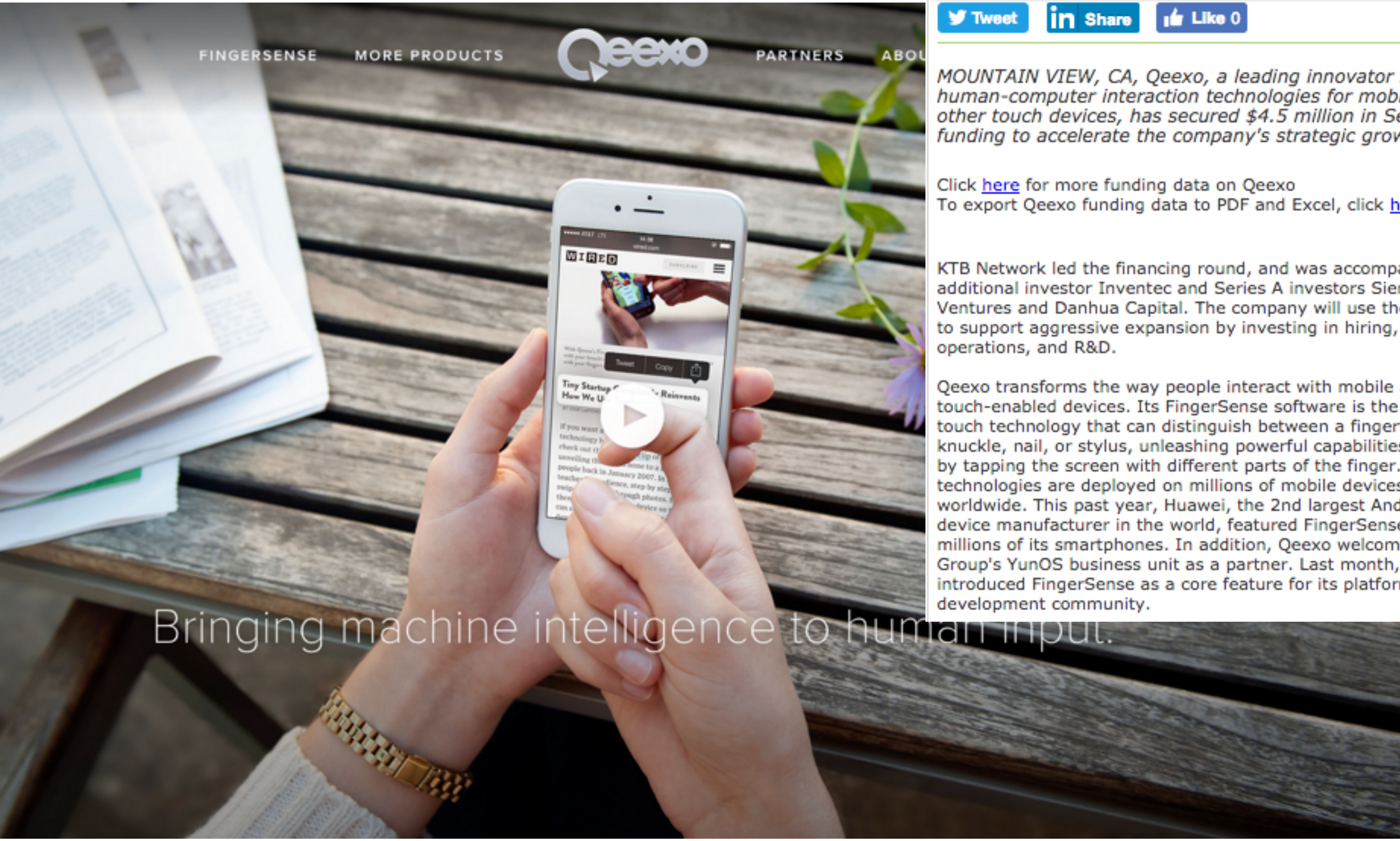
Tweet | in Share | Like 0

MOUNTAIN VIEW, CA, Qeexo, a leading innovator in human-computer interaction technologies for mobile and other touch devices, has secured $4.5 million in Series B funding to accelerate the company's strategic growth.

Click here for more funding data on Qeexo
To export Qeexo funding data to PDF and Excel, click here

KTB Network led the financing round, and was accompanied by additional investor Inventec and Series A investors Sierra Ventures and Danhua Capital. The company will use the funding to support aggressive expansion by investing in hiring, operations, and R&D.

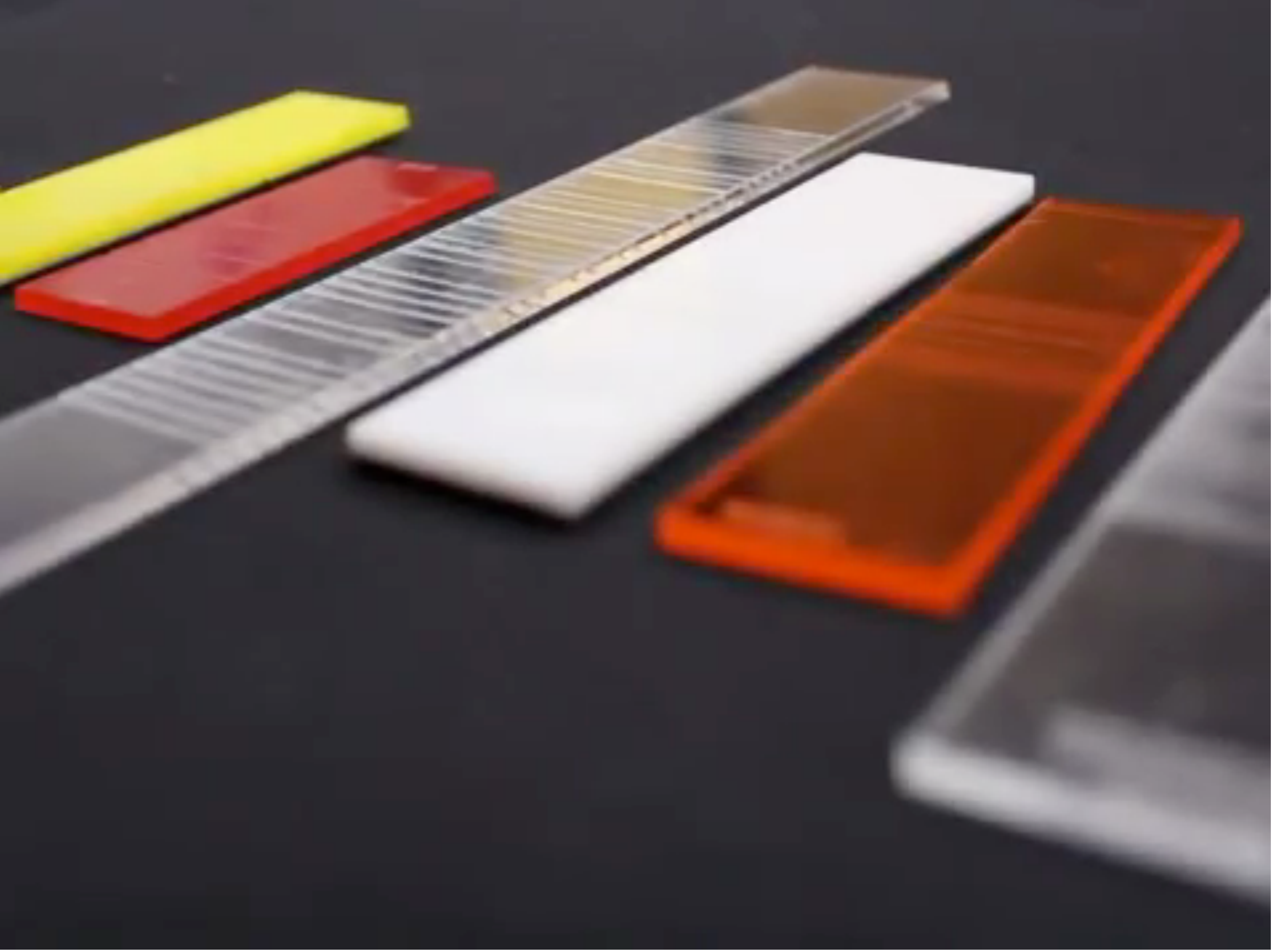
Qeexo transforms the way people interact with mobile and other touch-enabled devices. Its FingerSense software is the only touch technology that can distinguish between a fingertip, knuckle, nail, or stylus, unleashing powerful capabilities simply by tapping the screen with different parts of the finger. Qeexo technologies are deployed on millions of mobile devices worldwide. This past year, Huawei, the 2nd largest Android device manufacturer in the world, featured FingerSense on millions of its smartphones. In addition, Qeexo welcomed Alibaba Group's YunOS business unit as a partner. Last month, YunOS introduced FingerSense as a core feature for its platform and development community.



2011 paper, startup 2014.

more acoustic gestures

not just good for touch screens,
but also for **markers!**



2012 acoustic barcodes

Acoustic Barcodes: Passive, Durable and Inexpensive Notched Identification Tags

Chris Harrison **Robert Xiao** **Scott E. Hudson**

Human-Computer Interaction Institute and Heinz College Center for the Future of Work
Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213
{chris.harrison, brx, scott.hudson}@cs.cmu.edu

ABSTRACT

We present acoustic barcodes, structured patterns of physical notches that, when swiped with e.g., a fingernail, produce a complex sound that can be resolved to a binary ID. A single, inexpensive contact microphone attached to a surface or object is used to capture the waveform. We present our method for decoding sounds into IDs, which handles variations in swipe velocity and other factors. Acoustic barcodes could be used for information retrieval or to triggering interactive functions. They are passive, durable and inexpensive to produce. Further, they can be applied to a wide range of materials and objects, including plastic, wood, glass and stone. We conclude with several example applications that highlight the utility of our approach, and a user study that explores its feasibility.

ACM Classification: H.5.2 [Information interfaces and presentation]: User Interfaces - Graphical user interfaces; Input devices and strategies.

General terms: Design, Human Factors

Keywords: Sound, vibration, microphones, identification, ID, tags, markers, classification, location, interaction techniques, ubiquitous and pervasive computing.

INTRODUCTION

Our world is increasingly "tagged" to facilitate information retrieval and to trigger interactive functions. For example, UPC optical barcodes are ubiquitous on consumer goods and QR Codes are being used in smartphone applications.

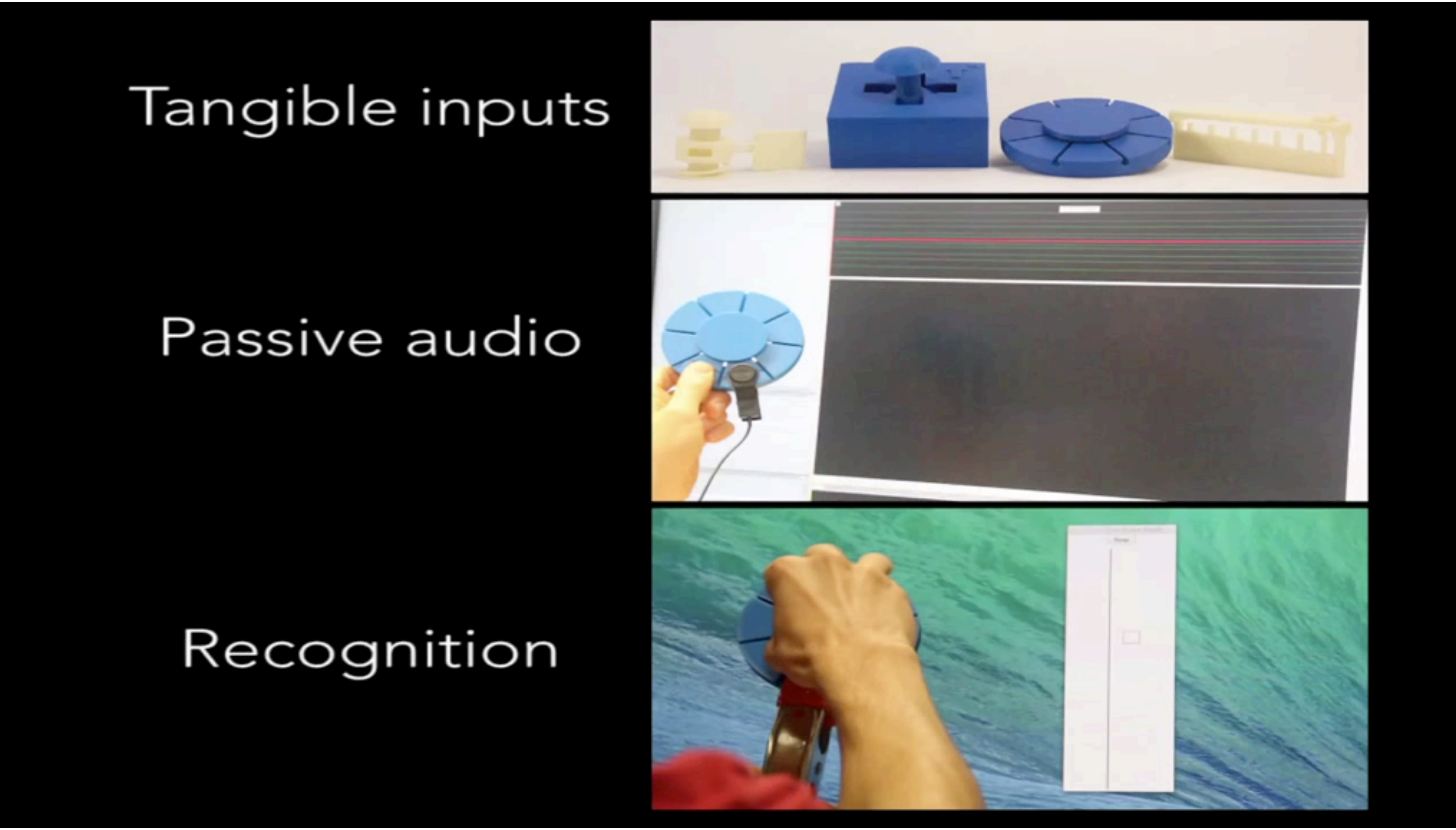
ver, tags are passive, durable, and inexpensive to mass-produce. We used a $6 microphone for sensing, which can monitor roughly 10m^2 of surface area, for example, a large whiteboard or table. Surfaces and objects can be augmented and easily retrofitted with acoustic barcodes. On some surfaces, tags can be made invisible. Overall, they can be smaller and subtler than visual markers (e.g., fiducial markers [10]). Finally, acoustic barcodes can be incorporated in a wide variety of materials.

RELATED WORK

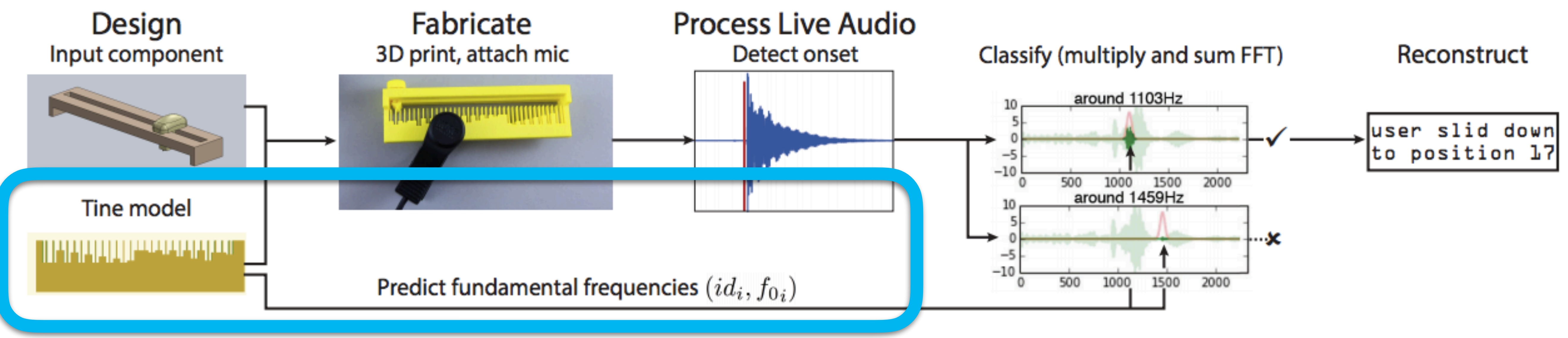
Researchers have explored many approaches for encoding unique identifiers. For example, visual schemes are popular, including 1D barcodes [11], 2D barcodes [19], and fiducial markers [10,17]. Identity can also be time-encoded using infrared light [18]. RFID tags use radio waves (electromagnetic radiation) for identification. Data can also be magnetically encoded, for example, the black strips on credit cards.

Most related to our technique are acoustic or tactile coding schemes. Braille [2] is a human-readable tactile encoding. The Edison Phonograph [4] is an analog system for recording and playing back sound by means of grooves cut in tinfoil. Moving from analog to digital: the Cricket System [15] uses coded ultrasound pulses to locate and identify users in an instrumented room. A listening device is attached to the user's hand and calculates its position in the room by measuring its distance to a set of fixed, coded emitters.

Another approach is acoustic-feature-driven classification.



2015: 3D printed passive sliders, dials



no training data needed,
they use the **3D model geometry** to do the prediction!

Lamello: Passive Acoustic Sensing for Tangible Input Components

Valkyrie Savage*†, Andrew Head†, Björn Hartmann†, Dan B Goldman*, Gautham Mysore*, Wilmot Li*
* Adobe Research, † UC Berkeley EECS
{valkyrie,andrewhead,bjoern}@eecs.berkeley.edu, {dgoldman,gmysore,wilmotli}@adobe.com

ABSTRACT
We describe *Lamello*, an approach for creating tangible input components that recognize user interaction via passive acoustic sensing. Lamello employs comb-like structures with varying-length tines at interaction points (e.g., along slider paths). Moving a component generates tine strikes; a real-time audio processing pipeline analyzes the resultant sounds and emits high-level interaction events. Our main contributions are in the co-design of the tine structures, information encoding schemes, and audio analysis. We demonstrate 3D printed Lamello-powered buttons, sliders, and dials.

Author Keywords
3D Printing; Sound; Tangible Interaction

ACM Classification Keywords
H.5.2 User Interfaces: Prototyping

INTRODUCTION
Tangible input components have advantages over "flat" interfaces like touch screens. They are critical for eyes-free interaction and muscle memory, and can improve task speed and precision [9]. Such devices typically comprise integrated assemblies of electronics, enclosures, and microcontroller code.

Recently, researchers have begun to explore acoustically sensing interactions – such as scratching – with digitally-

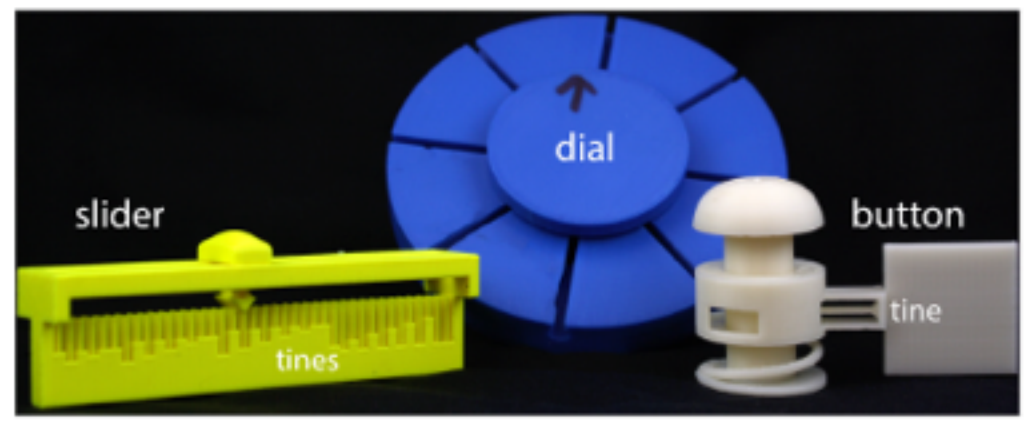


Figure 1. Passive tangible inputs that can be sensed acoustically.

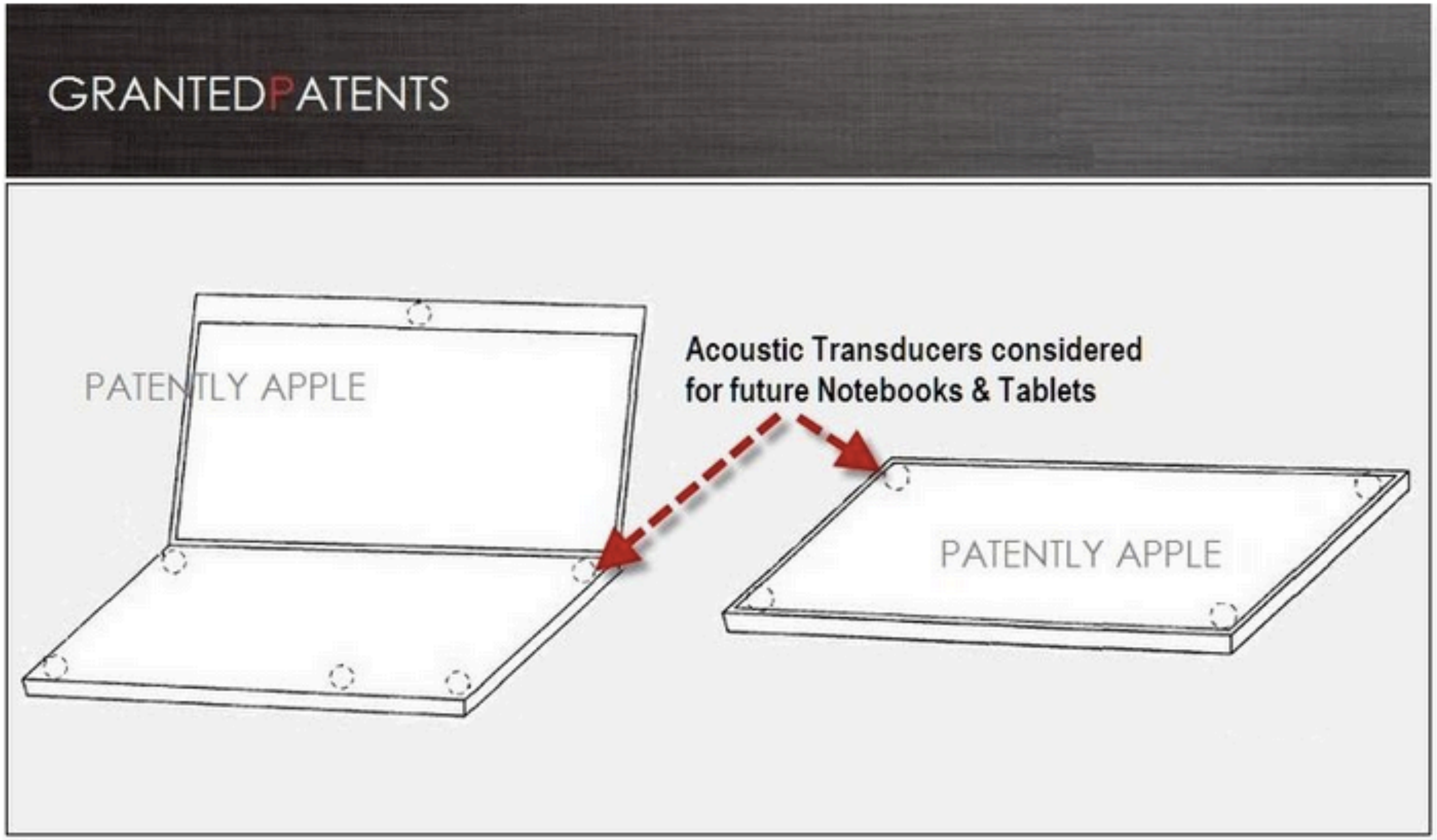
training examples. The decoded high-level events can then be forwarded to interactive applications. The name "Lamello" is derived from the Lamellophone family of instruments, which create sound through vibrating tongues of varying lengths.

Recognizing mechanically-generated sound for input has important limitations – only movement generates sound, so steady state cannot be sensed – but also appealing characteristics: Components can be fabricated from a single material (e.g., 3D printed ABS plastic), and "wiring" only requires attaching a microphone. In this paper, we provide design and fabrication guidelines, and demonstrate several components that use the Lamello approach. Our evaluation shows that training-free recognition is possible, though our recognizer only has useful accuracy for a subset of tested tines.

the future…

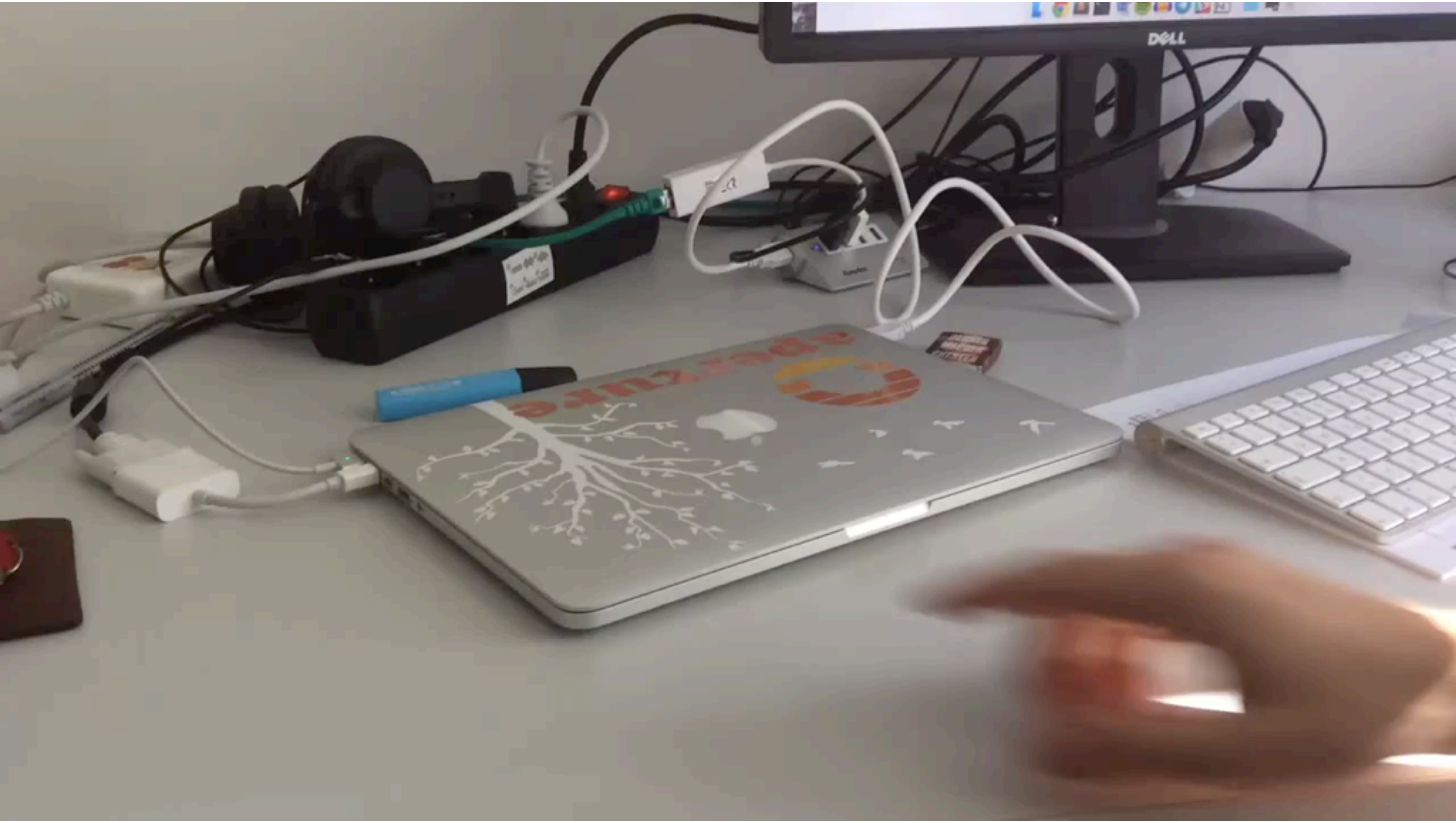
May 14, 2013

Apple Granted 35 Patents Covering Future iDevices that will respond to Acoustic Commands like Taps, Scratches & More



The US Patent and Trademark Office officially published a series of 35 newly granted patents for Apple Inc. today. In this particular report we first cover Apple's granted patent which discusses how acoustic transducers could one day be configured to accept acoustic commands that consist of tapping, scratching and other interactions with a surface of an iPad, iPhone or MacBook Pro. Our report concludes with a list of the remaining patents that were granted to Apple today.

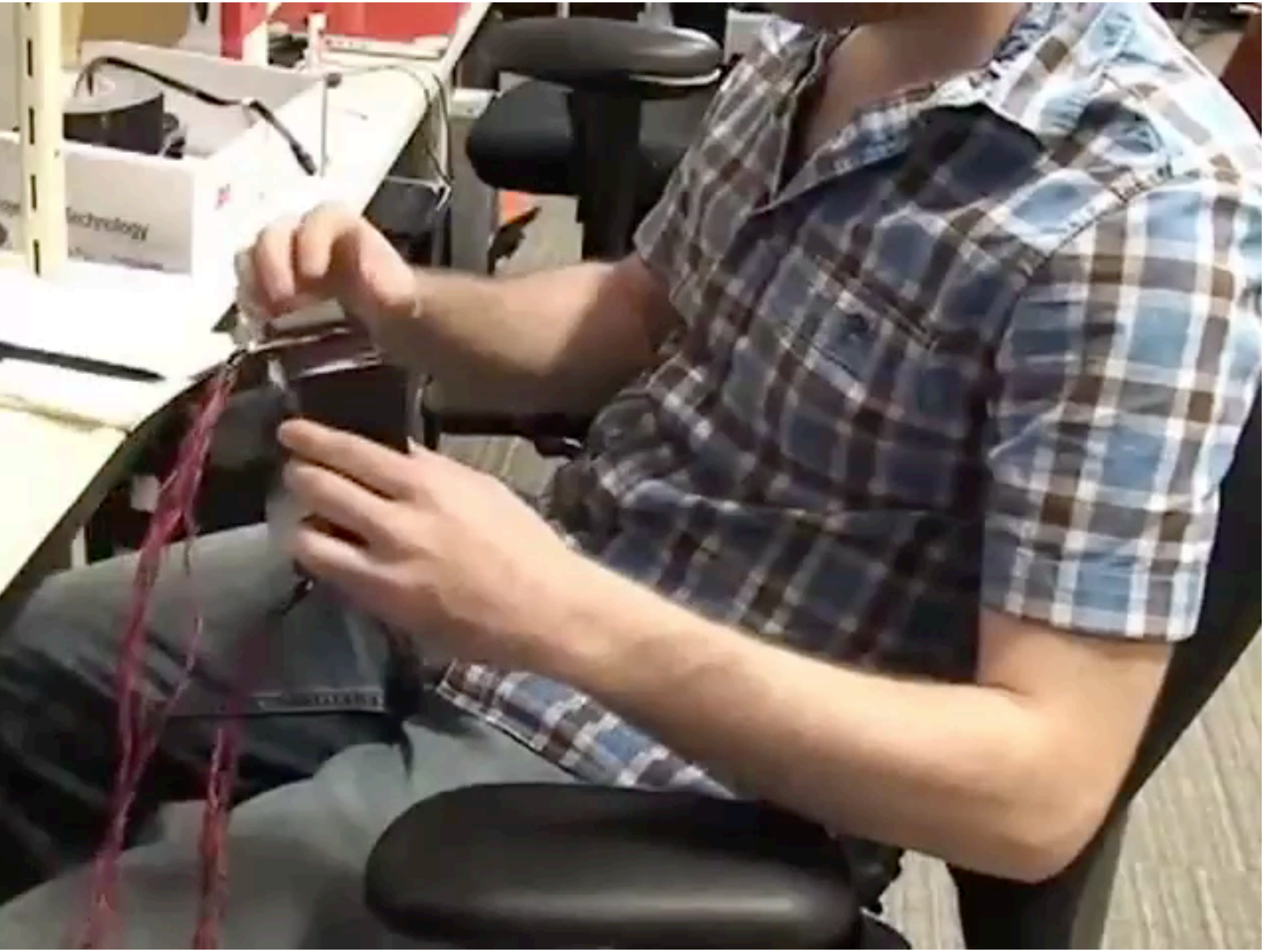
Apple Granted Patent: Electronic Devices used as Acoustic Input Devices


DELL

but there’s more…

touch input anywhere!

2010 Skinput: input on 3D deformable surface



2010 Skinput: input on 3D deformable surface

Skinput: Appropriating the Body as an Input Surface

Chris Harrison[1,2], Desney Tan[2], Dan Morris[2]

[1]Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
chris.harrison@cs.cmu.edu

[2]Microsoft Research
One Microsoft Way
Redmond, WA 98052
{desney, dan}@microsoft.com

ABSTRACT

We present *Skinput*, a technology that appropriates the human body for acoustic transmission, allowing the skin to be used as an input surface. In particular, we resolve the location of finger taps on the arm and hand by analyzing mechanical vibrations that propagate through the body. We collect these signals using a novel array of sensors worn as an armband. This approach provides an always available, naturally portable, and on-body finger input system. We assess the capabilities, accuracy and limitations of our technique through a two-part, twenty-participant user study. To further illustrate the utility of our approach, we conclude with several proof-of-concept applications we developed.

Author Keywords

Bio-acoustics, finger input, buttons, gestures, on-body interaction, projected displays, audio interfaces.

ACM Classification Keywords

H.5.2 [User Interfaces]: Input devices and strategies; B.4.2 [Input/Output Devices]: Channels and controllers

General terms: Human Factors

INTRODUCTION

Devices with significant computational power and capabilities can now be easily carried on our bodies. However, their small size typically leads to limited interaction space (e.g., diminutive screens, buttons, and jog wheels) and consequently diminishes their usability and functionality. Since we cannot simply make buttons and screens larger without losing the primary benefit of small size, we consider alternative approaches that enhance interactions with small mobile systems.

propriated surfaces with them (at this point, one might as well just have a larger device). However, there is one surface that has been previous overlooked as an input canvas, and one that happens to always travel with us: our skin.

Appropriating the human body as an input device is appealing not only because we have roughly two square meters of external surface area, but also because much of it is easily accessible by our hands (e.g., arms, upper legs, torso). Furthermore, proprioception – our sense of how our body is configured in three-dimensional space – allows us to accurately interact with our bodies in an eyes-free manner. For example, we can readily flick each of our fingers, touch the tip of our nose, and clap our hands together without visual assistance. Few external input devices can claim this accurate, eyes-free input characteristic and provide such a large interaction area.

In this paper, we present our work on *Skinput* – a method that allows the body to be appropriated for finger input using a novel, non-invasive, wearable bio-acoustic sensor.

The contributions of this paper are:

1) We describe the design of a novel, wearable sensor for bio-acoustic signal acquisition (Figure 1).
2) We describe an analysis approach that enables our system to resolve the location of finger taps on the body.

2008 Scratch Input: walls are all connected

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 environ-

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.

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

conclusions

sound interfaces:

	output	**input**
speech	text-to-speech synthesizers	siri, google glass ‘put that there’
non-speech	notifications, earcons	microphones as sensors

end.