# The Sound of One Hand: A Wrist-mounted Bio-acoustic Fingertip Gesture Interface

**Brian Amento, Will Hill, Loren Terveen**
AT&T Shannon Laboratory
180 Park Avenue
Florham Park, New Jersey USA
{brian, willhill, terveen}@research.att.com

## ABSTRACT

Two hundred and fifty years ago the Japanese Zen master Hakuin asked the question, "What is the Sound of the Single Hand?" This koan has long served as an aid to meditation but it also describes our new interaction technique. We discovered that gentle fingertip gestures such as tapping, rubbing, and flicking make quiet sounds that travel by bone conduction throughout the hand. A small wristband-mounted contact microphone can reliably and inexpensively sense these sounds. We harnessed this "sound in the hand" phenomenon to build a wristband-mounted bio-acoustic fingertip gesture interface. The bio-acoustic interface recognizes some common gestures that state-of-the-art glove and image-processing techniques capture but in a smaller, mobile package.

### Keywords

Human-computer interaction, gestural interfaces, gestures, fingers, wrist, acoustics, mobile devices.

## INTRODUCTION

Small, wearable digital devices are becoming increasingly popular with people who want mobile access to voicemail, email, headlines, stock quotes, calendars, and music. Cell phones, PDAs, and music delivery systems like the Walkman and Rio exemplify this trend. Such devices provide users with linear media streams to view and listen to. In general, users do not want to access these media streams sequentially; instead, they want to access them in a non-linear fashion (e.g. skipping ahead, backing up, replaying). Therefore, in our research we are considering ways to provide *non-linear access to linear media* (i.e. random access control of a sequential stream).

## HOW IT WORKS

We use a small wristband-mounted piezo-electric microphone to sense sounds internal to the hand produced by gentle fingertip gestures (Figure 1). The piezo-electric material senses vibration in human skin over the ulnar and radial styloid (wrist) bones or any anterior point between them over the skin of the wrist.

Users performs simple movements with their fingers (e.g., tapping, rubbing, flicking, snapping), internal sound at the fingertips is conducted by the bones of the hand and wrist through to the skin below the wristband microphone. The microphone is designed to reject airborne sound: even extremely loud airborne sounds will not register at all.

*CHI 2002*, April 20-25, 2002, Minneapolis, Minnesota, USA.
ACM 1-58113-454-1/02/0004.

Figure 1: We use piezo-electric microphones, mounted on a wristband to sense bone-conducted sounds from gentle fingertip gestures.

Our current prototype includes a wrist-mounted microphone (Figure 2) and a controller that uses a series of filters and state machines to classify quantized voltages coming from the microphone into instances of pre-existing gesture classes (tap, rub, flick and double-tap). Classified gesture signal instances are mapped to application control actions.

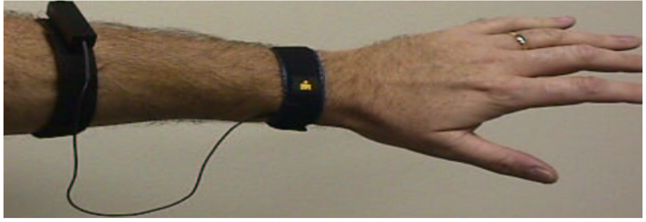

Figure 2: Experimental wristband contact microphone. The forearm strap steadies the cable attaching the microphone to the processor.

## PREVIOUS WORK

Our first effort to recognize fingertip gestures from an 'invisible" wrist mounted device took the form of a myo-electric interface sensing small muscle-generated electric signals [2]. This method proved much less accurate than the bio-acoustic technique. Our insight for bio-acoustics of the hand originated from an article describing how Thornbugs use vibrational channels to communicate by drumming on stems and branches.

## "NATURAL" FINGERTIP GESTURES

Buxton, W., E. Fiume, et al. [1] points out the importance of selecting gestures carefully. It was quite important for us to aim our fingertip gesture controller at sensing and classifying gentle, low-strength gestures, requiring only the most dexterous, repeatable, commonplace finger movements. We wanted to avoid gestures that cause stress in the hand and wrist, when repeated frequently. Therefore we focus on naturally occurring gestures like tap, rub, and flick.

## GESTURE SIGNAL CLASSIFICATION

We have built two gesture classifiers in Matlab to test applications and devices using our finger gesture interaction technique. The classifier samples audio from the contact microphone at a rate of 8000 samples per second. The voltages are quantized into 10 distinct levels based on the maximum in each sample period. This quantized voltage is then used as input to a finite state machine to determine which gesture was performed. The model includes intermediate states to ensure that background noise will not trigger gestures. We also include a series of buffer states that allow continuous gestures. This classifier is surprisingly accurate, considering its simplicity, and we have used it to generate a number of prototype applications.

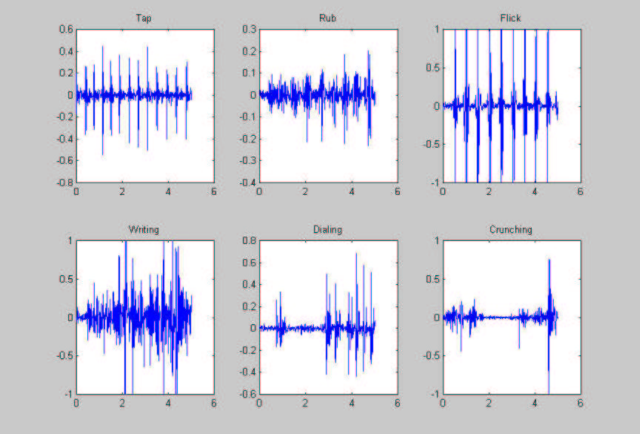

**Figure 3: Sample audio signals. Top row: 6 seconds of repeated tap, rub and flick gesture signals. Bottom row: 6 seconds of writing with a pen, dialing a phone and grabbing a handful of potato chips.**

Further analysis of the audio signals (Figure 3) reveals noticeable differences in the audio, which indicates that more sophisticated classifiers might enable higher accuracy for a wider range of recognized gestures. A tap results in a thin line extending to mid range voltage which is distinct from the continuous low range signal with thicker peaks that represents a rub. We have implemented an improved classifier using Hidden Markov Models [3]. The classifier is trained with multiple repetitions of each signal until an accurate model of each gesture is obtained. The noise signals can also be used in training so that the classifier will be able to easily reject various signals as non-gestures. After the signal is conditioned, it can be fed into the classifier that will produce confidence levels for each trained gesture.

## WRIST DISPLAY PROTOTYPE

One application prototype we have developed is a gesture-controlled watch/pager. The Timex Internet Messenger watch is capable of receiving and displaying small amounts of information such as sports scores, stock quotes and appointments. When a page arrives, the user must press the read button to navigate the pages stored on the watch.

Using a relay switch connected to our computer classifier through the serial port, we can simulate button presses and map them to finger gestures. The piezo-electric microphone is fastened to the underside of the wristwatch and connected to the classifier. In general this relay device we have designed gives us the ability to control any wearable device that uses physical contact buttons in its user interface. This bulky setup is a proof of concept. We intend a more seamless micro-miniature integration into existing devices. For the watch, the microphone can become part of the watch body and an on-board processor can be included to classify the audio signals. Then finger gestures could directly control the various software features of the watch without relays and external devices.

The gesture control for this device has been kept intentionally simple. A tap is mapped to the main button that checks incoming messages and advances through existing old messages. The other button controls rewinding a message and escaping back to the main time view and is accessed using a double tap. Using our version of the watch a user positions the watch face to read messages and then just taps index finger and thumb to advance through messages. The wristwatch emits audio feedback to the user when a gesture is sensed and the mapped action is initiated.

While this particular application of the finger gesture interface is straightforward, it illustrates the benefits of this technique. Instead of requiring two hands to operate a messaging device, users can now receive and read pages using a single hand freeing their other hand for another activity like holding a package or a child's hand.

## FUTURE WORK

Now that we have shown that our finger gesture technique is useful as a control mechanism and accurate enough for usable prototypes, we can continue this research in numerous directions. We will begin by building improved gesture classifiers and developing additional applications using finger gesture control. After that, we intend two different classes of user study. First, we will compare our finger gesture interface to current cellphone, PDA and desktop control interfaces to determine the usefulness of this interaction technique. Second, we will investigate how users exploit this technology and the social impact of having ubiquitous, private control techniques to mobile devices. It is clear that this fingertip technique will not be as efficient as common desktop control mechanisms such as a mouse, but it will be efficient in mobile situations where unholstering wearable devices is now the only means of control. We envision future wearable configurations to include both a wristband-mounted display/controller for viewing brief headers for news, email and voicemail and a belt-mounted device for viewing full content. If a header seen on a wrist display is relevant enough, the user can unholster his or her PDA/cellphone to access the content using a device with more sophisticated display and interaction capabilities.

## REFERENCES

1. Buxton, W., E. Fiume. (1983) Continuous hand-gesture driven input. *Proc. of Graphics Interface '83*: 191-195.
2. Ju, P.M. (2000) Classification of Finger Gestures from Myoelectric Signals, *Masters Thesis*, MIT EE Dept.
3. Rabiner, L.R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proc. IEEE,* Vol. 77, 257-28.