Tracking and Characterizing Knocks Atop Large Interactive Displays

Joseph A. Paradiso and Che King Leo

Responsive Environments Group
MIT Media Laboratory
Cambridge, MA 02142 USA

{joep,cheking}@media.mit.edu

Abstract

We describe a system that locates and characterizes knocks and taps atop a large sheet of glass using four contact piezoelectric pickups located near the sheet's corners to simultaneously record the structural-acoustic wavefront coming from the impacts. A digital signal processor extracts relevant characteristics from these signals, such as amplitudes, frequency components, and differential timings, which are used to estimate the location of the hit and provide other parameters, including a degree of confidence in the position accuracy, the nature of each hit (e.g., knuckle knock, metal tap, or fist bang – our system responds to any kind of impact), and the strike intensity. As this system requires only simple hardware, it needs no special adaptation of the glass pane, and allows all tracking transducers to be mounted on the inner surface, hence it is quite easy and inexpensive to deploy as a retrofit to existing windows. This opens many applications, such as an interactive storefront, with content controlled by knocks on the display window.

Keywords

touch screen, large interactive display, passive impact tracking, structural acoustics, acoustic characterization

1) INTRODUCTION

A major research thrust in HCI (Human-Computer-Interaction) is centered around investigating large display walls [1]. When they are made interactive, large displays open up entirely new types of group collaboration, in contrast, for example, with video kiosks, which interface mainly with single users. Participants at interactive walls are part user and part performer – in public settings, crowds tend to spontaneously gather to watch, contribute, and suggest choices as somebody interacts with a large display wall. Technologies now used to track user activity atop a very large display, however, generally involve several compromises, such as difficulty in scaling economically to large active areas, robustness to different kinds of user input gesture, or ruggedness, especially for outdoor settings.

Glass is now a very common construction material, often used as clear walls for room dividers or large windows bordering urban buildings. The techniques described in this paper aim at easily enabling these large surfaces to become interactive by tracking the position of knocks atop the glass. For example, information displayed on a projection or monitor on the inside of the glass can be determined by knocking appropriately on the outside. A straightforward application of this niche is an interactive storefront, where passers-by can navigate through information on the store's merchandise or explore special offers by knocking appropriately, even when the store is closed.

This paper describes the hardware, software, and algorithms that we have developed in realizing this capability, and presents application examples from installations that we've mounted in recent years that exploit this system.

2) OTHER APPROACHES

Many techniques have been developed to track the position of hands above interactive surfaces [2]. In general, most approaches used in conventional touch screens [3] don't scale gracefully to large displays. Pressure-sensitive resistive sandwiches, the most common technique, aren't made large

enough to cover very large windows, and because their operational principle requires compression of the sensing surfaces, they would need to be mounted on the window's outside (active) surface, where they would be subject to potential damage over time, especially for outdoor installations. Active acoustic touch screens detect the absorption of ultrasound launched into the outer surface of the glass when a finger is in contact. They require piezoceramic transducers to be mounted at a corner on the outside of the glass and need the glass' surface to be properly patterned with etched reflectors along the edges that direct the acoustic energy towards the middle of the glass plate. These systems also become error-prone as the surface scratches and deteriorates, plus attenuation of ultrasound in the glass can become an issue for large panes. Capacitive techniques are able to measure hands through the glass, but require a matrix of transparent electrodes to be patterned or adhered across the entire sensitive surface. Infrared LED curtains, featuring dense arrays of opposing IR transmitters and receivers measure the hand position when the corresponding optical beams are occluded. Although these systems can be scaled up to a larger screen, their expense increases accordingly, plus, as the IR array must line the perimeter of the screen's outer surface, they can be subject to damage, dirt, and deterioration, especially if they're used outdoors.

Other techniques, such as video tracking [4,5] have been used to make large screens interactive. Although they are steadily improving, vision-based approaches can be slow and are often sensitive to image clutter, target reflectance, and changes in background lighting. Time-of-flight laser rangefinding has also been used to scan the surface of large displays [2] and find hands, but the potentially expensive laser scanner must be mounted outside the window, leading to reliability difficulties for outdoor operation.

3) PASSIVE ACOUSTIC IMPACT TRACKING

Although the original motivation for this system was to track the location of knocks on a virtual fishtank [6], the first implementation of this technique was an augmented ping-pong table [7] that the lead author realized in 1997 in collaboration with colleagues from the Media Lab's Tangible Media Group. Here, four contact electret microphones, mounted at the corners of each contiguous half of the table and conditioned by simple constant-threshold Schmidt Trigger comparitors, detected the impact of the ping-pong ball. A simple 8-bit PIC microcontroller was able to adequately time the leading edge of each microphone signal as it arrived, and the difference in arrival times (processed through a linear least-squares fit) produced the Cartesian coordinates of the impact, essentially in real time. The impact position of the ball was input to a software environment that created a variety of graphical animations projected back onto the ping-pong table, resulting in causal visuals that accompanied the players' activity.

As indicated by the diversity in the waveforms shown in Fig. 1, the knock of a knuckle on glass is markedly different from the impact of a ping-pong ball on wood. The ball provides a consistent, high-amplitude waveform with a steep rising edge, requiring very little signal processing to extract adequate timing. The knuckle knocks, however, are more challenging, as each strike is different (depending on anatomy, intensity, style of knock, etc.), and generally of lower frequency, hence exhibit a much softer edge, which is much more difficult to time. The wave propagation characteristics of the glass add even more complexity. Knocks and taps on a glass surface launch a variety of bulk and plate flexural waves [8], that propagate out from the point of impact - the latter are structural acoustic modes that travel much more slowly than the medium's sound velocity. The propagation is dispersive, however, hence higher frequencies travel faster, causing the wavefront to increasingly slide apart as it travels further through the glass and high-frequency components arrive earlier. Accordingly, different propagation velocities and analysis techniques must be used for knuckle taps and hard-object (e.g., "metal") taps, as the latter

excite modes that produce very different wavefronts, which travel much faster through the glass. These complications preclude the simple constant-threshold discriminator used with the ping-pong table - the signals from knocking on glass need to be digitized and processed algorithmically in order to attain useful accuracies and robustness to a variety of different knocking styles.

A commercial system developed in France by a company called "Intelligent Vibrations" [9] appears to operate on a similar principal. Their implementation detects and analyzes the first symmetric Lamb plate mode (S_0), which exhibits limited dispersion. As this is an ultrasonic excitation, the user is required to hit the glass with hard fingernail taps, requiring a very constrained "fingerflick" gesture. To reject symmetric modes and other background, this approach requires all transducers to be differential, hence paired PZT pickups are used, mounted facing each other on both the inner and outer surfaces of the window. Our implementation needs transducers only on the inside of the glass, and as we process audio-band signals, it responds to any kind of knock – an important property, as requiring people to use such a highly-constrained gesture tends to hamper public acceptance and accordingly lower usability. Implementing this feature, however, required considerable effort in compensating for the different kinds of excitation waveforms from the various types of knocks (Fig. 1) and countering potentially large effects from dispersion in the glass medium, as described in Section 5. Another French startup called "Sensitive Object" have recently announced a system that tracks impacts on glass surfaces using only one contact transducer [10]. Their approach requires each surface to be calibrated by knocking across a grid of reference locations, creating data that generates a map of the plate's acoustical characteristics that can be exploited to resolve the impact location.

4) HARDWARE SETUP

Figure 2 shows the layout for our acoustic tap tracker system. The best performance was obtained when mounting the glass plates securely along all four edges (hence avoiding rattling), with

each of the four pickups located near a corner of the plate, where the broadest frequency response was exhibited for uniformly distributed knocks. Although the contact microphones that we originally used [11] were made from strips of simple PVDF (polyvinylidine fluoride) piezoelectric foil laminated onto the surface of the glass, after testing a wide variety of inexpensive transducers, we shifted to piezoceramic devices made by Panasonic (the EFV-RT series, actually designed as headset speakers for telephones), which provided the best sensitivity across the full bandwidth of various knocks and taps (e.g., 100 Hz - 5 kHz). Figure 3 shows photos of both transducers attached to the glass (the electronics for the Panasonic microphone are shielded from interference by a copper housing). Both systems employ a gain-of-11 voltage amplifier based around a noninverting OP162 opamp mounted at the transducer (which is shunted by a 1 MΩ resistor to limit low frequency response), enabling a long cable to be driven to the signal conditioning electronics, which filter out unwanted noise while applying additional gain. Although we generally adhere the microphones to the glass surface with a standard cyanoacrylate adhesive, long-term installations prefer a more permanent epoxy joint. An inexpensive digital signal processor (the ADMC 401 from Analog Devices, designed for motor control applications) extracts relevant parameters from the digitized waveforms and ships them across a serial link to a standard PC, which classifies the knock, estimates its coordinates, and runs appropriate interactive content. Modern PC's are now fast enough to do all needed processing as a 4-channel digital audio application, eliminating the need for the DSP in future implementations.

Although these piezoceramic transducers respond well to knocks and metallic hits, they don't give much signal when the window is banged with a fist. This primarily introduces low frequency vibration, down below 50 Hz, where the pickups and front-end amplifier aren't as sensitive. Accordingly, we often attach another transducer to the window to detect these "bash" events; it is an inexpensive electrodynamic cartridge, with the diaphragm epoxied to the window glass. As such bashes produce very low frequencies, they don't tend to provide good timing resolution and tracking (the impact

position is estimated from the small amount of coincident higher-frequency signal picked up by the piezoceramic microphones), hence one transducer, placed near the edge of the glass where it is out of the way, is sufficient to identify a bash. As this pickup works as an accelerometer mechanically tuned to low frequencies, it responds very strongly to bash events (and very weakly to knuckle taps and metal hits), and detects them with essentially no ambiguity, as illustrated in the next section.

Because the piezoceramic transducers are strongly adhered to the glass, they are essentially contact pickups and give very little response to signals coming from the air and not generated by taps on the glass. Certain sounds, however, such as a very loud clap or "snap" produced near the window (especially on the side where the microphones are mounted), couple enough energy into the system to trigger a false event. To prevent this from occurring, we have provided for another transducer in our system; this is a simple crystal microphone, not adhered to the window, but listening to ambient sounds made in its vicinity. The peaked high-frequency response of the crystal microphone makes it respond strongly to the sharp sounds that can instigate false triggers, hence any events that exhibit significant amplitude from this channel are vetoed as external sounds, not assumed to be generated in the glass.

Although the signals from the four piezoceramic transducers of Fig. 2 appear to provide sufficient information to detect bashes and veto claps on their own, this would require considerable additional signal processing and tweaking - the additional two transducer channels make this determination much more simple and robust. Our transducer packages are also sensitive to induced EMI pickup from very close cell phones, particularly GSM devices. Although better shielding could remedy this, a simple algorithm running in the DSP easily identifies and vetos these pulses via their the steep rising edge and flat baseline.

The DSP provides 26 MIPs worth of processing and samples all 6 input channels into 12 bits at 50 kHz. It continuously samples the input signals, triggering when any (from the contact pickups or the

bash microphone signals – the "clap" microphone only produces a veto) rise above a preset threshold. We currently retain 8 ms worth of data from the 4 piezoceramic transducers and the clap veto transducer (5 ms before the trigger and 3 ms afterwards, nicely capturing the incoming wavefront before the glass begins steady-state modal oscillation). As its low-frequency signal comes somewhat later, we retain 20 ms from the bash transducer (5 ms before trigger and 15 ms afterwards). In order to avoid multiple re-triggering as the glass rings down from a strong impact while still retaining a prompt response to new hits, we employ a dynamic trigger threshold that is raised immediately after an impact is detected and progressively lowered to its quiescent value.

Figure 4 shows data from one transducer triggered by the signal from another for a series of knuckle taps running between them atop a 1 cm-thick piece of glass; the pickups are separated by 0.9 meters. The progressive time delay, upon which this technique is based, is quite evident, as is the dispersive nature of the glass medium.

5) DATA PROCESSING

Rather than attempt to remove the effects of dispersion through signal processing, we have implemented a set of fast heuristics that align the captured waveforms adequately to resolve the knock position with sufficient accuracy for simple "point-and-click" applications on glass plates measuring over 2 x 2 meters in area. As detailed below, our impact-tracking algorithm can be described in four sections: (1) tap-type determination, signal filtering and normalization; (2) relative time-of-flight analysis between diagonally adjacent sensors; (3) realizing tap location; (4) transformation to absolute screen coordinates.

An analysis of the frequency response for impacts originating from different materials and sources (see Fig. 1) revealed that the EFVRT receiver was cleanly picking up two primary bands of frequencies – one near 1 kHz, and another near 3 kHz. This bi-modal frequency distribution was

characteristic over all surfaces that we worked with (¼" tempered and 1-cm safety glass), with the higher-frequency peak related to metallic-like objects, and the lower-frequency peak corresponding to impacts from softer, human tissue. Hence, our algorithm used two propagation velocities, choosing one or the other according to the dominant frequency of the sampled waveforms. As these velocities were characteristic for different windows, a calibration procedure, where a set of both kinds of taps are made at known locations, is performed at setup to determine the wavefront speed for both types of impact (metal taps tend to travel 2-3 times faster than knuckle taps). A third, much lower-frequency excitation with most components below 200 Hz originates from fist bashes on the glass, but these signals are not used for position tracking analysis.

Although one could discern knuckle taps from metal taps by examining the frequency distribution obtained from a FFT, we accumulate a coarse but adequate estimate of frequency by counting the number of times the sensor waveforms significantly cross zero across a fixed interval. As seen in Figure 5, which shows this quantity for several knuckle taps and metal taps, the two distributions are nicely separated and this distinction works well. In order to make this critical choice even more reliable, we also use the differential timing (e.g., the propagation velocity) between sensor pairs in this discrimination – this correctly classifies cases that have components at both frequencies, such as the upper knuckle knock shown at bottom in Fig. 1 (as we exploit only the first arrival, we process these events as metal taps).

Once the impact type is classified and a propagation velocity selected, the DSP derives features used for differential timing. Before these were extracted, the wave data from each transducer was lightly low-pass filtered, absolute-valued, and normalized in the DSP. Our initial strategy [2] triggered the data acquisition system when the amplitude detected at any transducer rose well above background, then walked back down the attack of the stored knock transient to determine the point at which the signal emerged from the noise floor. Although this attained a basic degree of performance, it could be

quite erratic, as baseline noise and the low-amplitude high-frequency signals arriving earlier through dispersion could often have considerable (and variable) influence on the timing. Our next efforts [11] digitized the signals with a data acquisition card, allowing us to explore more complex algorithms under MATLAB. We attained much more robust performance by cross-correlating the signals across pairs of sensors and extracting the differential time from the correlation peak. Because of the distortion encountered when propagating through the glass, however, the correlation could often become ambiguous, producing two or more significant peaks. In these cases, the redundancy in the system (provided by the additional sensor - only 3 are needed to specify position in a plane) and data from a calibration procedure were used to select the appropriate peak. Details are provided in Refs. [11,12].

Different procedures were used to time knuckle knocks and metallic taps in our final implementation [12]. The relative time-of-flight calculation for metallic taps again employed a dual threshold, which robustified against artifacts from saturation, reflections, and dispersion. After finding the spot where each signal first surpassed ¼ of its maximum sampled height, the algorithm backtracked 50 steps before proceeding forward to find the first point that rose above a lower threshold. This method worked well with metallic taps, as their sharp rising edges allowed the lower second threshold to consistently select the same part of the initial rising edge. The wider, more gently sloping waveform peaks produced by knuckle taps, however, caused problems with this approach and produced poor position resolutions. Accordingly, another method was developed that incorporated a few new steps. First, it was assumed that only the first few peaks in the signals would be examined, thereby selecting the section of the sampled waveforms with prompt data, free of subsequent reflections and other slowly arriving frequencies. Second, a chi-squared difference was calculated between the normalized pairs of peaks from each diagonal sensor pair. Finally, a cross-correlation was done between the pair of peaks that garnered the best chi-square statistic, and the maximum of this function was used to calculate the

relative time-of-flight of between the sensor pair. This method tended to match portions of the waveforms from different transducers that had similar frequency characteristics.

On average, when used for calculating knuckle-based taps, the chi-square-driven cross-correlation method provided significantly better precision 80% of the time when compared to the rising edge method. The derived coordinates can sometimes shift discretely, however, if the wrong peaks were picked for correlation. To smear any systematics, the knuckle knocks are processed by both methods, and both estimates are averaged when their resulting coordinates are within $d = 10$ cm of each other. When d is larger, it is assumed that the wrong peaks were selected for correlation, hence we use the rising edge estimate alone. This results in an improved precision and accuracy over using only one of the two methods for knuckle knocks.

The quantities described above are currently all calculated by the DSP, and the relevant parameters are sent over a serial link to the attached PC, which then determines the impact position. In our earlier work [2,11], we derived the (x,y) impact coordinates by running the timing results through a third-order polynomial that was determined via a linear least-squares fit to data collected on a regular grid. To avoid this lengthy calibration process (which was prone to overfitting), we have subsequently developed a deterministic algorithm [12]. As the position ambiguity curves for any pair of sensors form a hyperbola [13], we use a binary search to calculate the closest intersection between a pair of hyperbolas, each derived from the data produced by two diagonally-opposed sensors. The amount of miss between the hyperbolas reflects the quality of the data, hence the resolution of the position estimate. We adjust this hyperbola calibration to fit the characteristics of specific windows when the system is installed.

The peak amplitude is also captured for each sensor across the data acquisition interval. For the four contact pickups at the corners, this quantity reflects the hit's intensity. Although some approaches

use differential amplitude to determine position of impact (e.g., [14]), exploiting the progressive attenuation of the wavefront as it travels through the glass, we have found that the timing data is much better behaved. The peak amplitudes of the two sensors not used for tracking are consulted to determine a "bash" event (if the bash sensor amplitude dominates), or a background sound "clap" veto (if the clap sensor amplitude dominates).

6) PERFORMANCE

We have used this system thus far with two different thicknesses of glass: a 1/4" thick pane of tempered window glass and a 1-cm thick pane of shatterproof room-divider or shop window glass. In the 1-cm pane, we observe a propagation velocity of 1200 m/s for the knuckle tap and 2000 m/s for the metal tap; the speeds are roughly 50% slower in the thinner glass. Our present systems have been used with an active area between the sensors of up to 2 by 2 meters. The results presented here in Figure 6 are all taken with the 1/4" glass, with sensors spaced at the corners of a 1.6 x 1.4-meter rectangle.

Figure 6 shows the (x,y) reconstructed points for knuckle knocks at 5 locations (100 knocks at each site) at left, and the analogous data for metal taps at right (the "+" symbols in the corners denote the sensor positions). We see average resolutions of $\sigma = 2.5$ cm for the knuckle taps and $\sigma = 2.7$ cm for the metal taps at each position. The metal tap results tend to be somewhat less accurate because of the higher propagation velocity, while the knuckle knocks can show inaccuracy due to errors in choosing the proper correlation peak. Because of the higher velocity, results with the thicker, 1-cm glass tend to be somewhat less precise, with σ approaching 3-4 cm. By using the tweaked hyperbolic fits, the coordinate accuracy was quite good, with the centroids of the distributions in Figure 6 appearing within a few cm of the knock and tap positions. Occasionally, depending on the type of glass and the style of knock, the pickup waveforms can be inconsistent, resulting in an ambiguous condition. In these cases (generally 1% or lower), the event is rejected, and no coordinate is produced.

As we have been running this system in public settings for a few years, it has been installed on many different windows. In one installation using 1-cm glass, the window extended for over a meter above and below the area enclosed by the sensors, as plotted in Fig. 6 (it was grabbed along its sides by a rubber bumper to quickly damp resonances and avoid rattling). Accordingly, users were able to tap well above and below the sensors. The differential timings are able to determine this condition, however, and still produce coordinate estimates. As the time differences in the vertical direction show little change with vertical position in these regions, the resolution here was poor; the horizontal coordinate was still usable, however, reaching the order of $\sigma = 5.5$ cm when tapping a foot above or below the sensor-bounded perimeter.

The trigger thresholds were placed sufficiently low to respond to a soft knock and not activate sporadically with room noise. When sharp room sounds did occur, the "clap" veto signal successfully prevented the system from falsely responding. As poor pickup shielding in our prototype hardware made the system susceptible to interference from nearby cellphones, we implemented a simple waveshape identification algorithm in the DSP that successfully prevented this situation from triggering the system. The "bash" detector reliably discriminated between knocks and even modest fist "bangs", as illustrated in Figure 7, which shows the bash sensor saturating for a modest bang but showing very little signal for a knock. As our algorithms only used the first few milliseconds of transient signal, taps made at roughly the same time would tend to produce the coordinates of the earlier impact.

As seen in the following section, we have run this system in a variety of settings and, especially for the outdoor installation, across a wide range of temperatures (as we ran through a New York winter, temperature at the outer surface ran from 0-20º C). No effect was seen in system performance across this range, however, and if velocity drift were to become significant, data from the fourth, "redundant" transducer could be used for self-calibration.

The system responded quite quickly. The DSP generally produced parameters within 50 ms of an impact, and a PC (a 1 GHz Pentium 3) produced (x,y) coordinates circa 15 ms later, yielding a net system latency within 65 ms. Tighter coding could speed this up considerably, and the DSP could potentially be replaced with a simple 4-channel audio digitizer, with all processing happening in the PC.

7) APPLICATIONS

As this system has evolved, we have explored its use with increasingly sophisticated applications. All involved graphics projected onto a screen behind the glass; the user interacts with the information by knocking, tapping, and banging on the front surface. In all cases, the pickups were mounted on the rear of the glass.

Our first application [2] was a simple diagnostic that plotted circles centered at the location of the knock, with radius proportional to the estimated coordinate accuracy. After perfecting our hardware and analysis with this tool, we built a more user-relevant application, illustrated in operation in Figure 8. This is a "knock-driven" browser [11] demonstrated in our lab, where a user navigates group projects by knocking on the corresponding picture, launching relevant web pages and/or playing associated video clips. The targets (bounded by the interpicture gaps) were sufficiently far apart to make this system quite usable, even with a coarse resolution of several centimeters.

Our next deployment, also shown in Figure 8, was somewhat more sophisticated. It was an installation at the Ars Electronica Center that ran for three years after debuted in the 2001 Ars Electronica Festival as the "Responsive Window" [15]. A "holoscreen" [16] is mounted behind the 1-cm glass and located within the perimeter enclosed by the transducers. The holoscreen is a holographic diffuser that gives a rather bright and "ghostly" appearance to the graphics. Nothing is projected (and the screen is mildly translucent) until the glass is knocked, at which point the images appear. The graphics that we ran were essentially a simple knock-driven drawing program that illustrates the full

responsiveness of the system. One could outline and extrude an object by knocking about, with different effects and events caused by knock amplitude, bangs, and tap characteristics (e.g., metal vs. knuckle).

Figure 8 also shows another "knock-screen" application, where the picture of our collaborating artist's desk is projected on the glass; when images of relevant objects on the desk are knocked, a corresponding video stream is launched. If the user knocks during the video, a relevant still image will briefly flash up and rotate, centered at the knock location and rotating faster with increasing knock intensity. Bangs create a shower of images that fly away from the rough impact location. The data of Fig. 6 were taken with this installation, which was exhibited during the winter of 2001-2002 at the Kitchen Gallery in New York City.

Finally, Fig. 8 also shows a photo of an installation called the "Innovation Corner", which is still running at the Motorola iDEN facility in Plantation, Florida. Here, users can explore corporate information by knocking on the corresponding graphical icon.

Figure 9 shows our device in the window of an American Greetings store near Rockefeller Center in Manhattan, where it successfully ran during the Christmas and Valentine's Day seasons (their peak business period) last winter [17]. Passers-by could start informative videos or play simple games by appropriately knocking.

Another installation of this system, shown in Figure 10, ran at the ACM SIGGRAPH conference in San Antonio during July of 2002 [18]. Here, we ran the system on a large window (2 x 2 meters, ¼" glass), where it drove a complex interactive art visualization [19] that evolved in intricate ways with each knock. As schematicized at right in Fig. 10, this system also used a simple pair of Doppler radars [20] to measure people moving in front of the screen. The radars have an onboard processor that extracts three features reflecting the net amount of motion, speed, and direction. The radars are immune

to changes in light conditions or optical characteristics of clothing – unlike video imagers, they see directly through noncounductive walls and penetrate clothing, sensing the skin directly. The radars hence open up a degree of noncontact interaction as people approach the wall – in this case, motion in front of the screen generated global, nonspecific behavior (e.g., rolling, scrolling, activity creation) in the graphics in accordance with the motion characteristics. Knocking created more specific and highly localized phenomena. The type of knock also affected the graphics accordingly; hard metallic knocks made more "brittle", streaming foci than knuckle knocks, and fist bashes created images that essentially "exploded" near the knock position.

8) CONCLUSIONS

We have demonstrated a technique of easily retrofitting common windows for contact interactivity by measuring the position of a knock or tap, determining the type of impact, and estimating the impact intensity. Because of the complicated nature of the various impacts and nonideal propagation characteristics of the glass, this is not a precise pointing device, as it yields resolutions on the order of σ = 2-4 cm across roughly two meters. Its accuracy is adequate, however, for applications that involve coarse choices, such as hitting buttons when navigating content at interactive storefronts. Better signal analysis could extract more reliability and accuracy from this system; some possibilities could involve locating the arrival transients with a wavelet analysis or compensating for the dispersion in the glass by using coarse estimated range information to remove the expected dispersion from the signals (note that the amount of dispersion present in the signal also provides exploitable information about the range to the impact [21]). For larger windows, more transducers can be added, providing increased redundancy in the measurements and keeping the maximum propagation distance small. Also, more sensitive pickup sensors could be investigated, such as wideband accelerometers, providing reliable response across larger areas. We have brought this system out to the public on several occasions, where it is

received with interest because of its novelty - once users recover from their initial confusion with touch screen operation (one must hit here rather than touch), they generally adapt well to it. For ease in usability, we do not require the user to constrain the knock to a particular gesture (e.g., fingernail flick) – our system has been designed to work with essentially any kind of impact. Rear-mounting of the sensors will probably not function with air-insulated double-paned windows, as the rear surface is acoustically isolated from the front, where the impact occurs.

Video clips showing this system in action can be downloaded at:

http:// www.media.mit.edu/resenv/Tapper

9) ACKNOWLEDGMENTS

10) REFERENCES

[1] Funkhouser, T., and Li K., (eds.). "Onto the Wall: Large Displays," Special issue of *IEEE Computer Graphics and Applications*, 20(4) (Jul/Aug 2000).

[2] Paradiso, J.A., Hsiao, K., Strickon, J., Lifton, J. and Adler, A., "Sensor Systems for Interactive Surfaces," IBM Systems Journal, Vol. 39, No. 3&4, October 2000, pp. 892-914.

[3] Quinnell R.A., "Touchscreen technology improves and extends its options," EDN, vol.40, no.23, 9 Nov. 1995, pp.52, 54-6, 58, 60, 62, 64.

[4] Martin, D., "The Future for Interactive Surfaces," in the Proc. of the UbiComp 2002 Workshop on Collaboration with Interactive Walls and Surfaces, Gothenburg Sweden, September 29, 2002. Available at: http://www.ipsi.fraunhofer.de/ambiente/collabtablewallws/papers/D-122-Martin%20-%20Workshop%203.doc.

[5] Rekimoto, J., Matsushita, N., "Perceptual Surfaces: Towards a Human and Object Sensitive Interactive Display. In Workshop on Perceptual User Interfaces (PUI-97), Banff, Published by Microsoft, October 1997, pp. 30-32.

[6] Galyean, T., et al., "Virtual FishTank," SIGGRAPH 98, Conference Abstracts and Applications, ACM Press, New York, 1998, p. 116.

[7] Ishii, H., Wisneski, C., Orbanes, J., Chun, B., and Paradiso, J., "PingPongPlus: Design of an Athletic-Tangible Interface for Computer-Supported Cooperative Play," in Proceedings of Conference on Human Factors in Computing Systems (CHI '99), Pittsburgh, Pennsylvania, May 15-20, 1999, ACM Press, NY, pp. 394-401.

[8] Cremer, L., Heckl, M., and Ungar, R.R., <u>Structure-Borne Sound</u>, Second Edition, Springer-Verlag, New York, 1990.

[9] Devige, F. and Nikolovski, J-P., "Accurate Interactive Acoustic Plate," US Patent Application No. US2003/0066692 A1, April 10, 2003.

[10] Eisenberg, A., "What's Next: Knock 3 Times on the Ceiling (to Turn on the DVD Player)," *New York Times*, Circuits, p. G8 July 1, 2004.

[11] Checka, N., "A System for Tracking and Characterizing Acoustic Impacts on Large Interactive Surfaces," MS Thesis, MIT Department of EECS and MIT Media Lab, May, 2001.

[12] Leo, Che King, "Contact and Free-Gesture Tracking for Large Interactive Surfaces," Meng Thesis, MIT Dept. of EECS and MIT Media Lab, May 2002.

[13] Reynolds, M*., et. al*., "An Immersive, Multi-User, Musical Stage Environment," ACM Computer Graphics, SIGGRAPH 2001 Proceedings, ACM Press, NY, 2001, pp. 553-560.

[14] Mathews, M.V., "The Conductor Program and Mechanical Baton," in Mathews, M.V., and Pierce, J.R., eds., <u>Current Directions in Computer Music Research</u>, MIT Press, Boston, MA, 1989, pp. 263-282.

[15] Paradiso, J., "The Responsive Window," in Take Over, Proc. of the 2001 Ars Electronica Festival, Springer Verlag, Linz, Austria, September 2001, pp. 261-263.

[16] See: http://www.holopro.de/

[17] Wilson, C., "If You're Game, Tap on the Window," USA Today, December 21, 2001, p. 1D.

[18] Paradiso, J.A., Leo, C., Yu, N., Downie, M., "The Interactive Window," in SIGGRAPH 2002, Conference Abstracts and Applications, ACM Press, 2002, p. 79.

[19] Downie, M., "'Loops' - A Digital Portrait," SIGGRAPH 2002 Conference Abstracts and Applications, ACM SIGGRAPH Press, p. 209.

[20] Paradiso, J., "Several Sensor Approaches that Retrofit Large Surfaces for Interactivity," presented at the UbiComp 2002 Workshop on Collaboration with Interactive Tables and Walls, Gothenburg, Sweden, September 29, 2002. See:

http://ipsi.fhg.de/ambiente/collabtablewallws/papers/ParadisoUbicompSurfaces.pdf

[21] Ing, R.K. and Fink, M., "Time-Reversed Lamb Waves," in the IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 45, No. 4, July 1998, pp. 1032-1043.
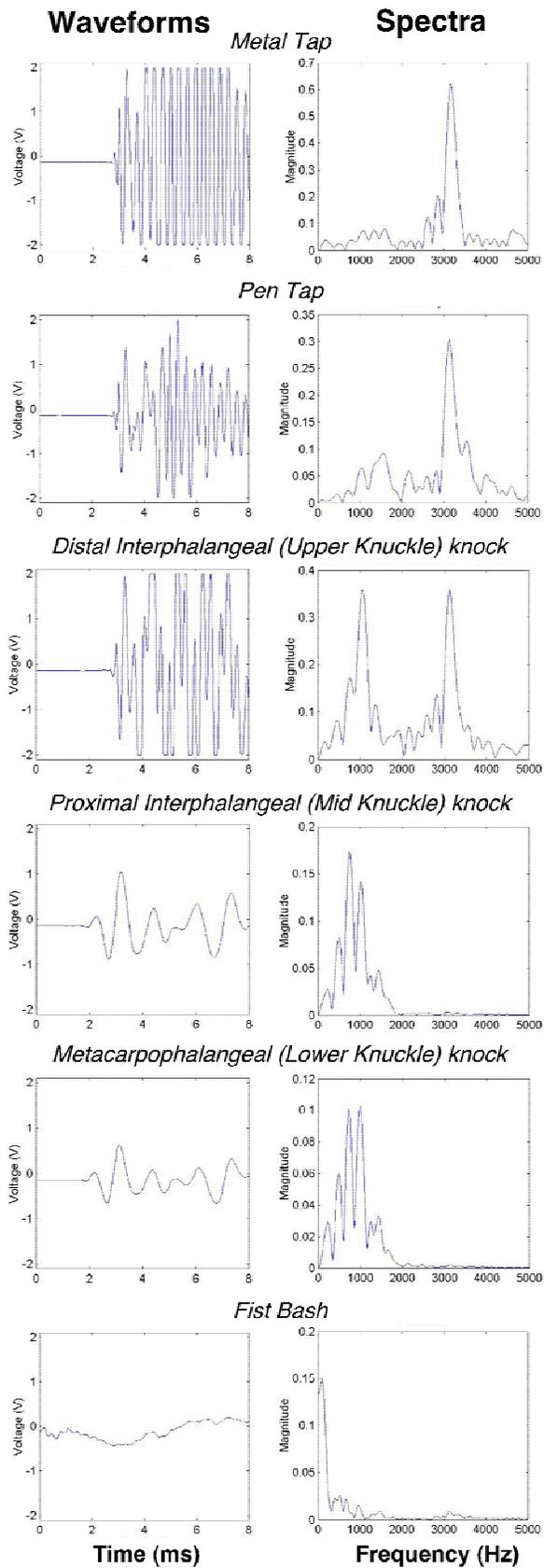
Figure 1: Piezo pickup signals (left) and corresponding spectra (right) in response to a variety of impacts atop a sheet of ¼" tempered glass.
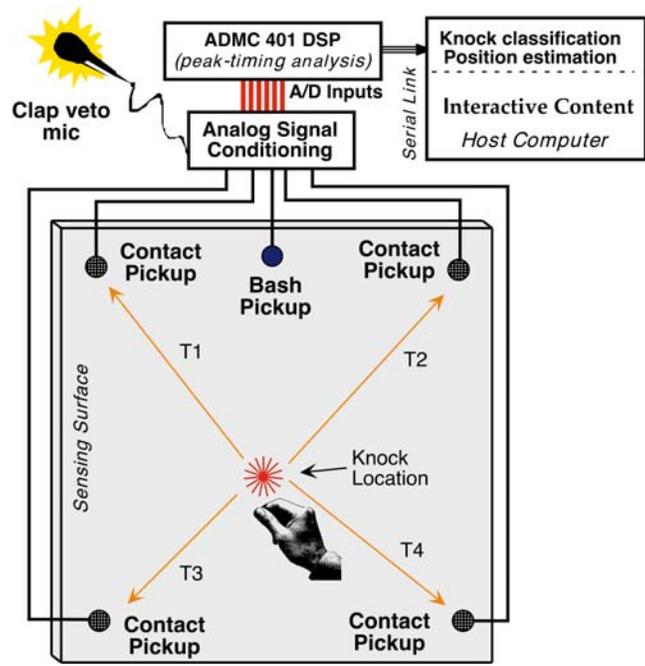
**Figure 2: Hardware configuration for knock tracker system**
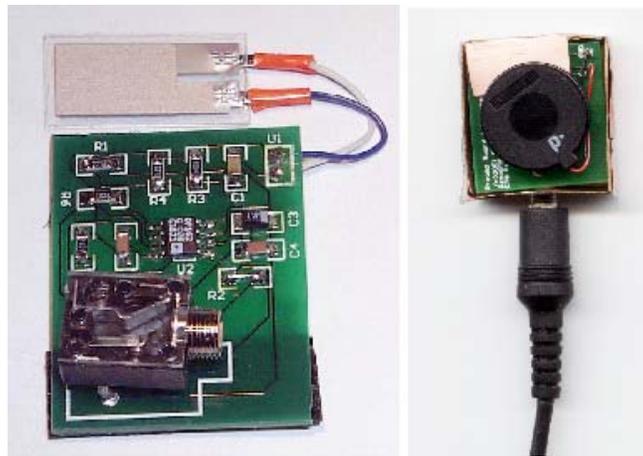
Figure 3: Contact pickups - original PVDF w. preamp (left) and current piezoceramic unit w. assembly (right)
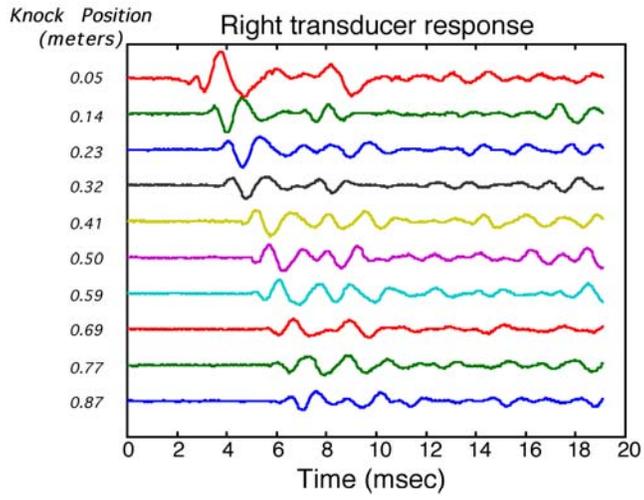
Figure 4: Increased differential delay as knock moves away from triggered pickup
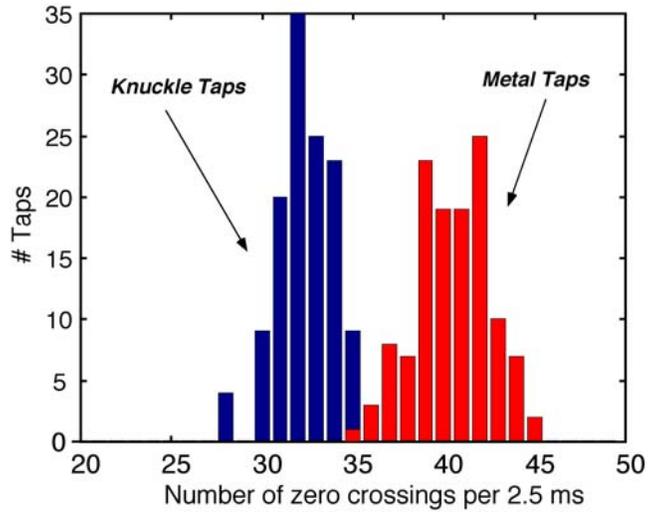

**Figure 5: Histogram of inferred knuckle and metal tap frequency**
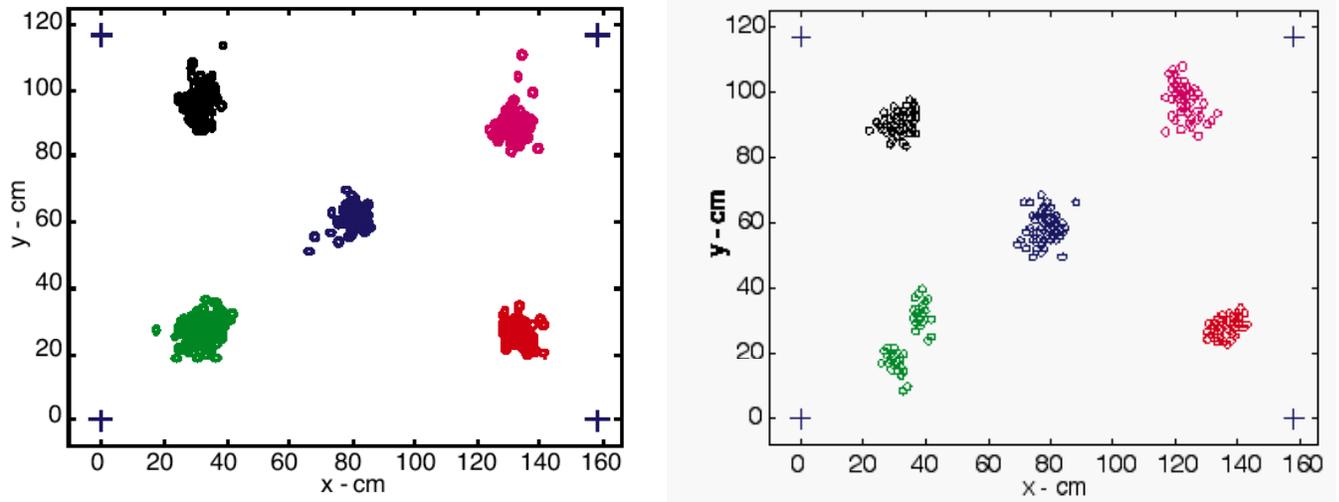
Figure 6: Estimated positions for knuckle taps (left) and metal taps (right) at 5 locations on ¼" tempered glass
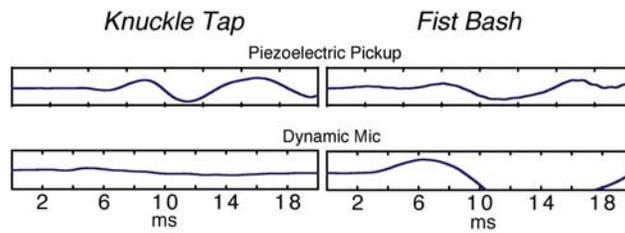

Figure 7: Piezo & dynamic pickup responses to tap and bash
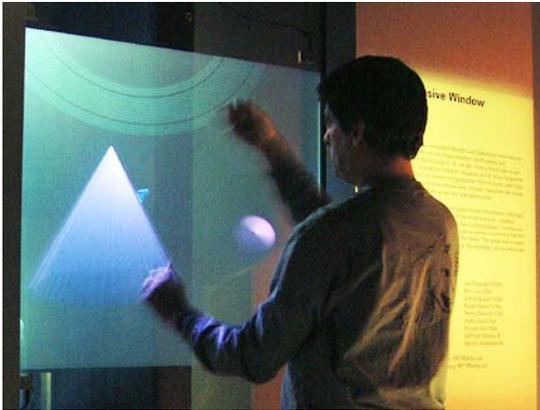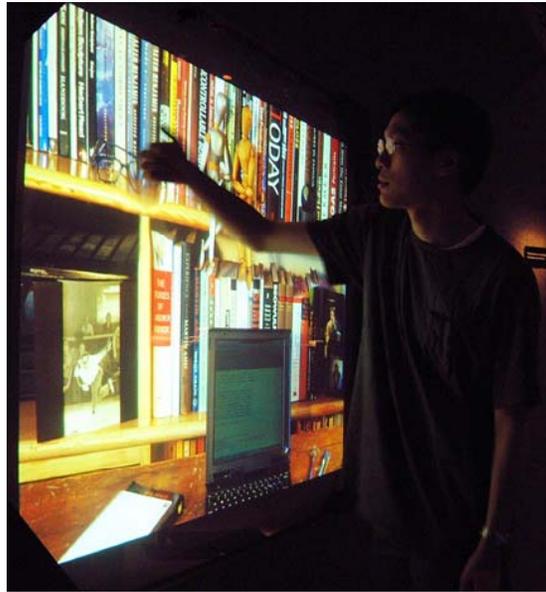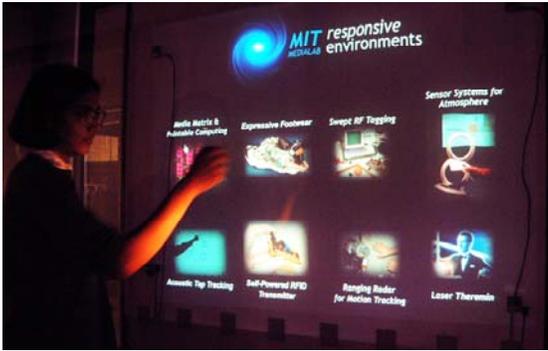
Figure 8: Interactive window installations at the MIT Media Lab (top left), the Ars Electronica Center (lower left), the Kitchen Gallery (upper right) and Motorola (lower right)

**Figure 9: Interactive window browsing at an American Greetings store near Rockefeller Center**
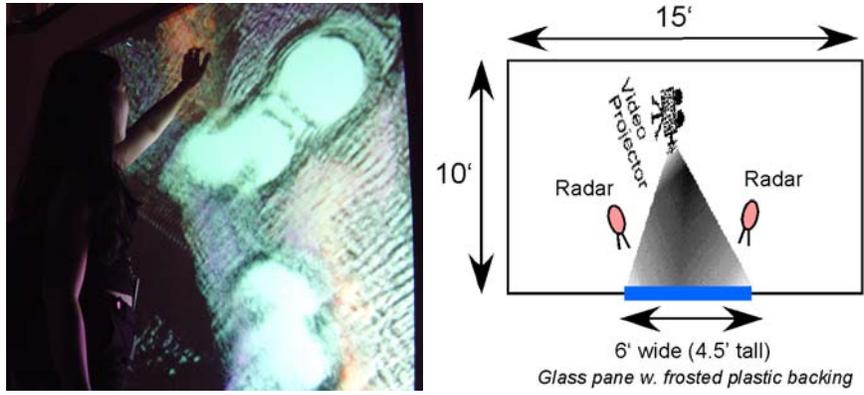
Figure 10: System deployed at SIGGRAPH 2000 with microwave motion sensors to detect noncontact motion in front of screen