Audio Engineering Society

Convention Paper 10660

Presented at the 154th Convention
2023 May 13–15, Espoo, Helsinki, Finland

Audio Capture Using Piezoelectric Sensors on Vibrating Panel Surfaces

Tre DiPassio[1], Michael C. Heilemann[1], Benjamin Thompson[1], and Mark Bocko[1]

[1]*Department of Electrical and Computer Engineering, University of Rochester, USA*

Correspondence should be addressed to Tre DiPassio (`tredipassio@rochester.edu`)

ABSTRACT

The microphone systems employed by smart devices such as cellphones and tablets require case penetrations that leave them vulnerable to environmental damage. A structural sensor mounted on the back of the display screen can be employed to record audio by capturing the bending vibration signals induced in the display panel by an incident acoustic wave - enabling a functional microphone on a fully sealed device. Distributed piezoelectric sensing elements and low-noise accelerometers were bonded to the surfaces of several different panels and used to record acoustic speech signals. The quality of the recorded signals was assessed using the speech transmission index, and the recordings were transcribed to text using an automatic speech recognition system. Although the quality of the speech signals recorded by the piezoelectric sensors was reduced compared to the quality of speech recorded by the accelerometers, the word-error-rate of each transcription increased only by approximately 2% on average, suggesting that distributed piezoelectric sensors can be used as a low-cost surface microphone for smart devices that employ automatic speech recognition. A method of crosstalk cancellation was also implemented to enable the simultaneous recording and playback of audio signals by an array of piezoelectric elements and evaluated by the measured improvement in the recording's signal-to-interference ratio.

1 Introduction

As recent advances in display technology have allowed devices such as smartphones and televisions to become thinner and more lightweight, a need has arisen for audio systems whose form-factors meet the space constraints of these devices. Recent developments have been in the area of audio reproduction, where the display screen itself is used as a loudspeaker. Flat-panel loudspeaker design methods [1, 2, 3, 4] have been adapted to use moving-coil actuators [5, 6, 7, 8], or piezoelectric actuators [9] to induce sound-producing bending waves on the display surface.

In addition to producing sound, the display surface may also be leveraged to capture audio by employing a structural sensor to record the vibration signals induced by an incident acoustic wave. The enclosure of a surface audio system can be fully sealed as the mounting of structural sensors inside the enclosure on the back side of the display would eliminate the need for case penetrations. This gives the proposed interface an advantage over conventional MEMS microphones, which require penetrations in the enclosure for the sensor to record sound - making the device susceptible to environmental damage. Though the resonant modes of the display panel inevitably introduce reverberation into the recorded audio signal, the quality of recorded speech signals has been demonstrated to be sufficient

for speech-to-text software to generate transcriptions with similar accuracies to those produced by speech recordings captured with high-quality measurement microphones [10]. This suggests that structural vibration sensors may be used in place of traditional microphones for smart devices that execute tasks when prompted by key words detected in the transcription of a voice command.

The structural sensor used in [10] was a highly sensitive, low-noise accelerometer. The inclusion of such a sensor in a consumer-grade product would be cost-prohibitive as opposed to using conventional MEMS microphones. Distributed piezoelectric elements come at a substantially reduced cost compared to the previously used accelerometers, and have had widespread use as sensors in applications such as active vibration control [11]. In this work, distributed piezoelectric sensing elements were bonded to the surfaces of panels made of several different materials and were used to record acoustic speech signals. The speech signals were then transcribed using an automatic speech recognition (ASR) system. The accuracy of these transcriptions was compared with those derived from the accelerometers to determine if the quality of a signal recorded by the distributed piezoelectric element is sufficient for use with ASR systems.

In Section 2, we begin with an overview of panel mechanics to provide a mathematical foundation for the ideas presented in the manuscript. In Section 3, we explore the intelligibility of audio recordings made by piezoelectric sensors affixed to various panels. In Section 4, a method for crosstalk cancellation between the panel's inertial exciters and structural sensors is analysed. Conclusions are discussed in Section 5.

2   Mechanics of Vibrating Panel Systems

Consider a thin, isotropic panel with Young's Modulus E, Poisson's ratio v, density ρ, thickness h, and damping factor b. The out-of-plane displacement $w(x,y,t)$ in response to an external load $p(x,y,t)$ at time t and point (x,y) on the panel's surface can be computed following [12] and [13] with,

$$p(x,y,t) = D\nabla^4 w(x,y,t) + \rho h\ddot{w}(x,y,t) + b\dot{w}(x,y,t), \quad (1)$$

where ∇^4 is the biharmonic operator expanded in two dimensions, $\dot{w}$ and $\ddot{w}$ are the first and second time derivatives of w, and D is the bending stiffness of the panel, given by,

$$D = \frac{Eh^3}{12(1-v^2)}. \quad (2)$$

The panel's out-of-plane displacement can be separated into functions of space and time as,

$$w(x,y,t) = \varphi(x,y)e^{j\omega t}. \quad (3)$$

The panel's bending modes are an orthonormal basis for the spatial function $\varphi(x,y)$, which can be expressed via superposition as,

$$\varphi(x,y) = \sum_{r=1}^{\infty} \alpha_r \Phi_r(x,y), \quad (4)$$

where $\Phi_r(x,y)$ is the shape of the panel's r^{th} mode excited with amplitude α_r.

The panels used in the experimental portion of this work are rectangular with dimensions (L_x, L_y), and are fixed to wooden frames using two-part epoxy to create fully clamped edges. For panels with these dimensions and boundary conditions, the resonant frequency ω_r of the rth mode can be found with an approximate formula given in [14]. Fuller [11] shows that α_r at frequency ω can be expressed as,

$$\alpha_r(\omega) = \frac{4}{\rho h(\omega^2 - \omega_r^2 - \frac{j\omega_r\omega}{Q_r})} \int_0^{L_x} \int_0^{L_y} P(x,y,\omega)\Phi_r(x,y)\mathrm{d}y\mathrm{d}x, \quad (5)$$

where $P(x,y,\omega)$ is the magnitude of the pressure distribution at point (x,y) and Q_r is the quality factor of each mode which describes the bandwidth of the mode's resonance. The modal quality factor is shown by Fahy and Gardonio to be inversely proportional to the material's damping factor [13], and is given by,

$$Q_r = \frac{\omega_r\rho h}{b_r}. \quad (6)$$

Table 1: Material and physical properties of the panels used in the experimental portion of this work

Material	E (GPa)	v	ρ (kg m^{-3})	h (mm)	L_x (cm)	L_x (cm)	b_{avg} (kg s^{-1})
Gatorboard	1.5	0.35	222	3	26	36	241.5
Acrylic	3.2	0.35	1180	2	26	36	2172
Gorilla Glass	71.5	0.21	2420	1.2	10.4	11.8	3120
Aluminum	68.9	0.33	2700	1	26	36	10270

The quality factor of each mode is difficult to determine analytically as the effective damping, b_r, varies for each of the panel's bending modes. However, the average damping factor of a panel can be expressed as

$$b_{\mathrm{avg}} = \frac{2\rho h \ln(2)}{t_{1/2}}, \qquad (7)$$

where $t_{1/2}$ is the decay time for the impulse response of the panel to reach one-half amplitude. Following [10], a value for $t_{1/2}$ for each of the panels used in this work was determined experimentally. The material and physical properties of the panels are given in Table 1. The materials were chosen to explore a wide range of damping coefficients. A tablet-sized Gorilla Glass panel was also constructed to address the use case of the proposed interface for mobile devices [3]. Note that the Aluminum panel is constructed of two sheets of aluminum bonded together by a viscoelastic damping material to introduce more damping into the system [15, 16].

3 Intelligibility of Recorded Vibrations

3.1 Experimental Setup

The experimental setup is shown in Fig. 1. Each of the panels listed in Table 1 was placed in a semi-anechoic environment with an affixed 35 millimeter diameter 7BB series piezoelectric element. A KEF LS50 loudspeaker was placed on-axis, one half meter away from the panel. The transfer function $h_1[n]$ from the loudspeaker to the panel's affixed sensor was recorded using a maximum length sequence excitation with a sound pressure level at the panel's position of 71 dBSPL, to approximate the sound pressure level of human speech at this distance [17]. Additionally, a measurement was taken with a calibrated PCB Piezotronics F130F20 free-field microphone [18] as a reference.
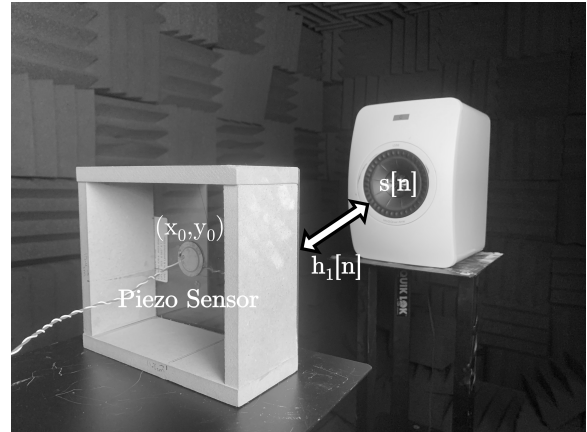

Fig. 1: Experimental setup in which panels equipped with piezoelectric sensors were placed in a semi-anechoic environment. Vibrations are induced on the panels by a loudspeaker at a distance of 0.5 meters.

As the deflection of the panel induced by incident waves is small compared to the panel's dimensions, the panels vibrate in a linear region of operation [13]. Therefore, the signal model for this experimental setup can be expressed as,

$$z_{(x_0,y_0)}[n] = s[n] \circledast h_1[n], \qquad (8)$$

where $\circledast$ is the convolution operator, and $z_{(x_0,y_0)}[n]$ is the panel's displacement at position (x_0, y_0) at sample n induced by the signal from the loudspeaker, $s[n]$. Consistent with [10], $s[n]$ in this experiment contained recordings of Harvard sentences in accordance with the IEEE Recommended Practice for Speech Quality Measurements [19]. Using (8), convolution can be used to simulate a corpus of 500 Harvard sentences as incident to the panel.

The accuracy with which an ASR system could transcribe the recordings made by the piezoelectric sensor was analyzed using a word error rate (WER) metric. WER is given as a percentage by,

$$\text{WER} = \frac{\text{Insertions} + \text{Deletions} + \text{Substitutions}}{\text{Number of Words in Reference}} * 100\%. \tag{9}$$

The recordings made by each panel were transcribed by IBM Watson's speech-to-text service [20], and compared to the known transcripts of the Harvard sentences to determine the WER. The speech transmission index (STI) was also computed for each panel as an objective measure of speech quality preservation [21, 22].

3.2 Experimental Results

The STI and WER scores obtained from the recordings made by each of the panels are given in Table 2 and Table 3, respectively along with the scores obtained using the reference microphone. As discussed in Section 1, the results are compared to those from [10], which employed a high-cost PCB Piezotronics U352C66 accelerometer [23] affixed to each of the panels. The motivating factor behind this comparison is cost, as the piezoelectric elements are at least two orders of magnitude less expensive than the PCB accelerometers. Note that the results presented from the accelerometer bonded to the Gorilla Glass panel were not presented in the previous work.

All STI scores presented are greater than 0.75, and are therefore deemed excellent in quality by the standard. With regard to the WER metric, the results show a greater reduction in transcription accuracy for recordings made by panels with less material damping. This results from (6), as panels with greater damping experience fewer reverberant high-Q modes that hinder intelligibility. However, at most a 6% increase in WER relative to the reference microphone was observed in recordings made by the piezoelectric sensor on the Gatorboard panel, and as little as a 2% increase was observed when using the highly-damped aluminum panel. This suggests that recordings of speech made by piezoelectric sensors bonded to panels are able to be transcribed by modern ASR systems without a significant reduction in accuracy when compared to recordings made by traditional microphones.

Table 2: Comparison of the STI score achieved by each panel when using a piezoelectric sensor and an accelerometer.

Material	STI Score	
	Piezo sensor	Accelerometer
Gatorboard	0.908	0.918
Acrylic	0.920	0.920
Gorilla Glass	0.922	0.931
Aluminum	0.943	0.983
Ref. Microphone	0.980	

Table 3: Comparison of the WER score achieved by each panel when using a piezoelectric sensor and an accelerometer.

Material	Average WER (%)	
	Piezo Sensor	Accelerometer
Gatorboard	15.3	13.6
Acrylic	14.6	11.5
Gorilla Glass	13.9	11.4
Aluminum	11.3	10.1
Ref. Microphone	9.33	

It is also evident that there is some reduction in quality when replacing the accelerometer with the piezoelectric sensor. However, the WER increased by no more than 3.1% when making this replacement. This magnitude of degradation is also observed when varying the panels' damping factor. For instance, the Gatorboard panel with the accelerometer gave similar WER to the Gorilla Glass panel with the piezoelectric sensor. This suggests that the reduction in WER measured when using piezoelectric sensors can be mitigated by using more highly-damped panels. Therefore, material damping persists as an important design characteristic for these panels where the intelligibility of recorded speech is a functional priority.

4 Crosstalk Cancellation

In applications where the display panel is employed as a duplex audio interface to simultaneously record audio with sensors and reproduce audio with actuators, the recorded signal will contain a mixture of vibration induced by both the actuators and the user's
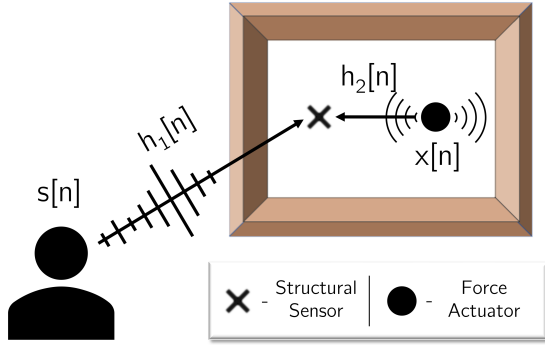
Fig. 2: Signal model for the recorded audio stream with contributions from both the panel's affixed actuator and the user's speech.

speech. A subtraction method for cancelling an inertial exciter's contribution to the audio stream recorded by an accelerometer bonded to the panel was demonstrated in [10]. In this work, the subtraction method will be evaluated for audio streams recorded by the low-cost piezoelectric sensors. The method will also be evaluated on panels driven by piezoelectric actuators.

4.1 Experimental Setup

The signal model for an audio stream recorded by an affixed structural sensor is shown in Fig. 2, where $s[n]$ is the acoustic speech signal, $h_1[n]$ is the transfer function from the speaker to the sensor, $x[n]$ is the signal being played by the actuator and $h_2[n]$ is the transfer function from the actuator to the sensor. Since the panel is operating in a linear deflection region, the recorded audio stream $z_{(x_0,y_0)}[n]$ can be expressed as,

$$z_{(x_0,y_0)}[n] = s[n] \circledast h_1[n] + x[n] \circledast h_2[n]. \quad (10)$$

As the actuator and sensor are bonded to the same panel, the audio stream will contain a larger contribution from the actuator than from the acoustic speech signal when both signals originate with the same power, as was the case for this experiment.

An estimate of the contribution from the recorded speech can be obtained using subtraction as,

$$z_{(x_0,y_0)}[n] - x[n] \circledast h_2[n] = s[n] \circledast h_1[n]. \quad (11)$$

A purely convolution approach is not possible for this experiment, as the subtraction of a simulated vibrational contribution would be trivial. Instead, a KEF LS50 loudspeaker was placed in the far-field of each panel, where $s[n]$ consisted of Harvard sentences while a different signal was simultaneously played through the actuator affixed directly to the panel.

The actuators played white noise, classical music, and synthesized speech to replicate common signals played by a smart speaker, which act as interference to the recorded speech signal. From (11), the desired signal after subtraction contains only the user's speech. Therefore, a signal-to-interference ratio (SIR) can be reported as,

$$\text{SIR (dB)} = 10\log_{10}\left(\frac{P_s}{P_x}\right), \quad (12)$$

where P_s and P_x are the power of the signals in the recorded audio stream from the incident acoustic waves and the induced panel vibrations from the actuator, respectively. The performance of each panel will be evaluated by the improvement in the SIR of the audio stream after the subtraction method is applied.

4.2 Crosstalk from Intertial Exciters

In a first experiment, an inertial exciter was affixed to each of the non-glass panels and used to reproduce the interference signal. These actuators can be modeled as point forces on the panel at frequencies where the bending wavelength is larger than the contact area. The magnitude of the force distribution $F(x,y,\omega)$ from an actuator located at point (x_i, y_i) can be expressed as,

$$F(x,y,\omega) = F(\omega)\delta(x - x_i)\delta(y - y_i), \quad (13)$$

where $F(\omega)$ is the force applied by the actuator at frequency ω. Combining (13) and (5) give the response of each mode $\alpha_r(\omega)$ as,

$$\alpha_r(\omega) =$$
$$\frac{4}{\rho h L_x L_y(\omega^2 - \omega_r^2 - \frac{j\omega_r\omega}{Q_r})} \times$$
$$\int_0^{L_x}\int_0^{L_y} F(\omega)\delta(x - x_i)\delta(y - y_i)\Phi_r(x,y)\mathrm{d}y\mathrm{d}x$$
$$= \frac{4F(\omega)\Phi_r(x_i,y_i)}{\rho h L_x L_y(\omega^2 - \omega_r^2 - \frac{j\omega_r\omega}{Q_r})}. \quad (14)$$
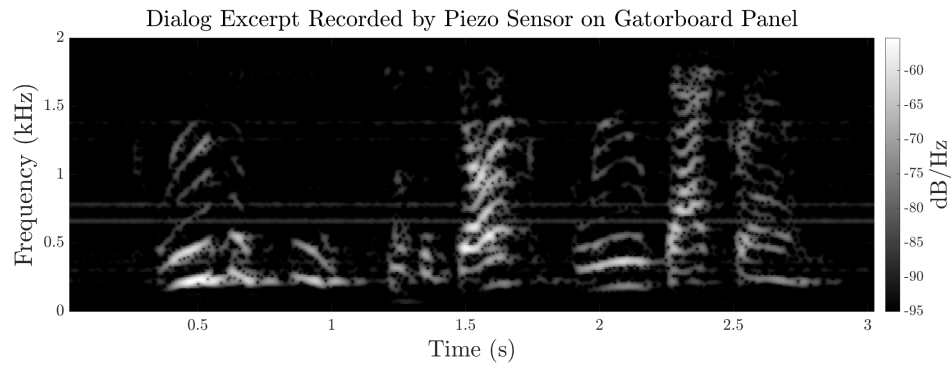
Fig. 3: Spectrogram of speech recorded by the piezoelectric sensor on the Gatorboard panel without an interference signal being played.
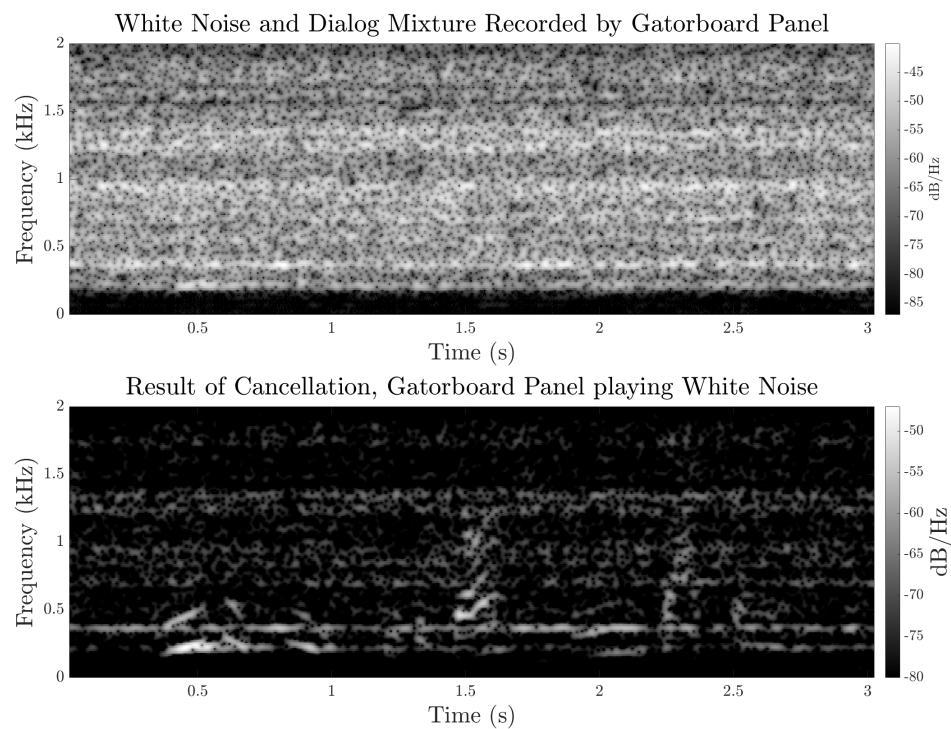

Fig. 4: Spectrogram of dialog recorded while white noise was played by the exciter before and after cancellation.

Table 4: Increase in SIR in the piezoelectric sensor's audio stream after cancelling each type of interference signal being played by the inertial exciter.

Material	SIR Increase (dB)			
	White Noise	Classical Music	Alexa Speech	Average
Acrylic	15.8	19.5	17.9	17.6
Aluminum	34.1	29.6	29.8	31.2
Gatorboard	17.2	19.2	17.4	17.9

Table 5: Increase in SIR in the piezoelectric sensor's audio stream after cancelling each type of interference signal being played by the piezoelectric actuator.

Material	SIR Increase (dB)			
	White Noise	Classical Music	Alexa Speech	Average
Acrylic	17.2	12.9	11.2	13.8
Aluminum	18.5	20.9	17.5	19.0
Gatorboard	12.0	11.2	10.1	11.1
Gor. Glass	16.5	16.9	13.9	15.7

The response of the panel to an interial exciter and therefore the coloration of the interference signal recorded by the sensor can be fully modeled with (4) and (14).

A spectrogram of a speech segment recorded by the piezoelectric sensor affixed to the Gatorboard panel without an interference signal being played is shown in Fig. 3. This spectrogram of speech in isolation becomes the target spectrogram when applying the cancellation method to a recording made in the presence of an interference signal from the actuator. Spectrograms showing this speech segment in a mixture with the white noise interference signal before and after cancellation are shown in Fig. 4, and the post-cancellation SIR improvement among all panels are listed in Table 4.

After cancellation, the SIR increased an average of 17.6 dB on the acrylic panel, 31.2 dB on the aluminum panel, and 17.9 dB on the Gatorboard panel. Note that the sensor on the highly-damped aluminum panel shows greater reliability in cancelling the interference signal. This may be due to the long decay times of the high-Q modes that are present on the lesser-damped panels, which may cause signals to smear into adjacent processing frames. Withstanding the apparent effect of damping, the reported SIR improvements provide experimental evidence of the feasibility of suppressing an interference signal being played by an inertial exciter when recorded by a piezoelectric sensor.

4.3 Crosstalk from Piezoelectric Exciters

In a second experiment, a rectangular piezoelectric element of dimensions 2.0 cm by 1.0 cm was affixed to each of the panels and used to drive the interference signal. If the piezoelectric element is bonded off of the neutral axis of the panel, it will exert bending moments on the panel at the edges of the actuator when a voltage is applied across the actuator. The response of each panel mode is scaled based on coupling between the actuator dimensions and the modal half-wavelength [24]. Since the piezoelectric element's dimensions are small compared to the dimensions of the panel, the force exerted on the low-frequency modes of the panel is small compared to the force exerted on the high-frequency modes [3]. The result is a weaker bass response in the reproduced audio compared with an inertial exciter acting in the same location.

A spectrogram of a speech segment recorded by the piezoelectric sensor affixed to the aluminum panel without an interference signal being played is shown in Fig. 5. Spectrograms showing this speech segment in a mixture with a white noise interference signal played through a piezoelectric actuator, are shown in Fig. 6 before and after cancellation. Table 5 lists the post-cancellation SIR improvement among all panels. Notice that the noise power in Fig. 6 is weighted toward high-frequencies in the uncancelled recording due to the effects of the piezoelectric actuator.

Experimental results again suggest that the subtraction method is viable for suppression of the interference signal recorded by the piezoelectric sensor. After cancellation, the SIR increased an average of 13.8 dB on the acrylic panel, 19.0 dB on the aluminum panel, 11.1 dB on the Gatorboard panel, and 15.7 dB on the Gorilla Glass panel. It is again apparent that the highly-damped aluminum and Gorilla Glass panels provide a better medium for interference cancellation. The
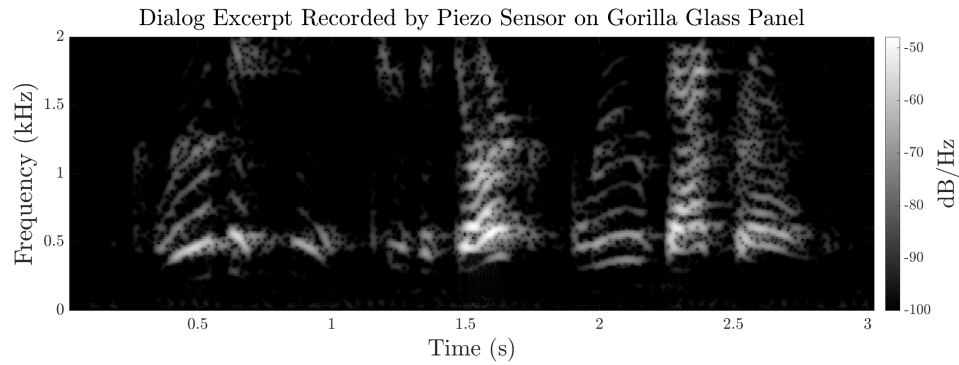
Fig. 5: Spectrogram of speech recorded by the piezoelectric sensor on the Gorilla Glass panel without an interference signal being played.
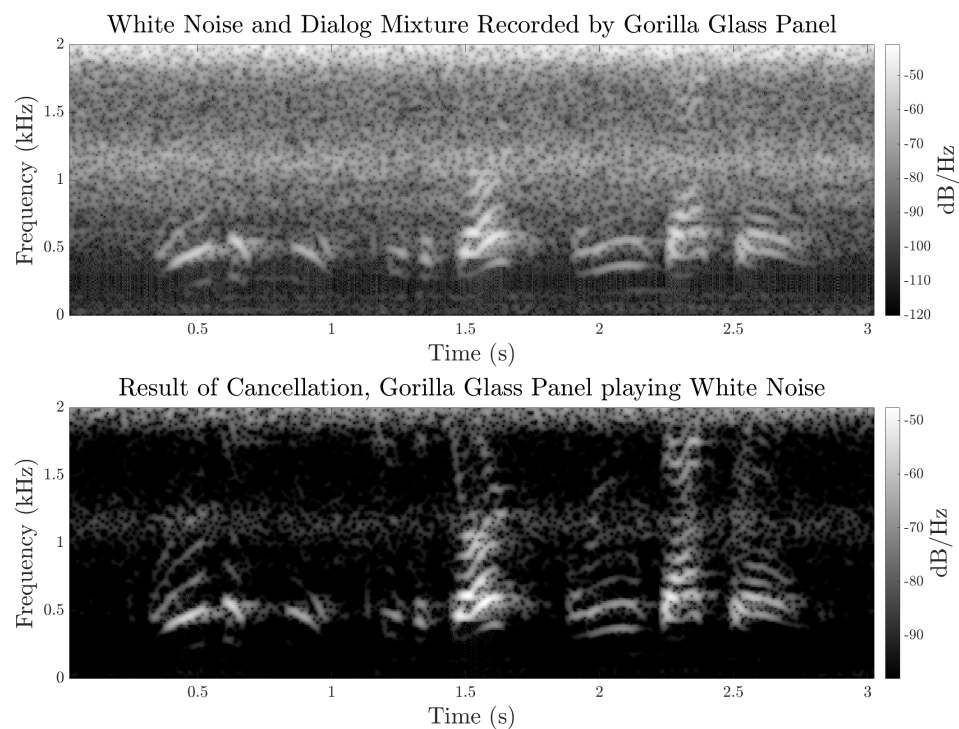

Fig. 6: Spectrogram of dialog recorded while white noise was played by piezoelectric actuator before and after cancellation.

experimental results for crosstalk cancellation on the tablet-sized Gorilla Glass panel suggest that a surface-based duplex audio interface may be possible for smart display devices using array of piezoelectric sensors and actuators.

Comparing Table 4 with Table 5, the amount of reported interference suppression is reduced when a piezoeletric actuator is used. This may be due to the frequency response of the actuator, which is visibly characteristic of a high-pass filter as shown in Fig. 6. It is possible that, even though the power contained in $x[n]$ was equal between the two experiments, the signal power when coupled to the panel via $h_1[n]$ may not be sufficient to achieve full cancellation. It may be necessary in future work to develop a method of evaluation that accounts for the relative power of $s[n] \circledast h_1[n]$ and $x[n] \circledast h_2[n]$, instead of $s[n]$ and $x[n]$ directly.

5 Conclusions

The results presented in this work demonstrate the potential of using piezoelectric sensors and actuators with sub-\$1 price points in place of high-cost accelerometers and inertial exciters for surface-based audio interfaces.

Though recordings made using piezoelectric sensors demonstrated a reduction in transcription accuracy when compared to those made by accelerometers, the magnitude of this reduction was similar to that observed when switching from a highly-damped panel material such as viscoelastically damped aluminum to a lesser-damped Gatorboard material. Therefore, it may be possible to compensate for the reduction in accuracy introduced by the piezoelectric sensor by increasing the damping in the panel material. As this experiment was carried out in a semi-anechoic environment, it will be necessary in future work to evaluate the transcription accuracy of recordings made with structural sensors in more typical, noisy environments.

For situations where the panel surface is simultaneously employed to record and reproduce audio, a crosstalk cancellation method based on subtraction was presented. Tables 4 and 5 show the viability of suppressing the interference signal that is played by the actuators bonded to the panel. An improvement in SIR was observed when using both inertial exciters and piezoelectric actuators.

Since smartphones and tablets represent a potential host for this technology, one of the experimental panels was made of Gorilla Glass - a material commonly used for the screens of these devices. As with the other panels used in these experiments, recordings made with the tablet-sized Gorilla Glass panel were able to be accurately transcribed by the ASR system, and crosstalk cancellation was successful with an array of piezoelectric sensors and actuators. The results from this work suggest that it is possible to construct duplex audio interfaces in a surface-based form factor using low-cost piezoelectric elements.

6 Acknowledgement

References

[1] Heilemann, M. C., Anderson, D. A., Roessner, S., and Bocko, M. F., "The evolution and design of flat-panel loudspeakers for audio reproduction," *Journal of the Audio Engineering Society*, 69(1/2), pp. 27–39, 2021.

[2] Bank, G. and Harris, N., "The distributed mode loudspeaker-theory and practice," in *Audio Engineering Society Conference: UK 13th Conference: Microphones & Loudspeakers*, Audio Engineering Society, 1998.

[3] Heilemann, M. C. and DiPassio, T., "Piezoelectric Actuators for Flat-Panel Loudspeakers," in *Audio Engineering Society Convention 153*, Audio Engineering Society, 2022.

[4] Anderson, D., "Introduction to modeling and analysis of small, piezoelectrically excited bending-wave loudspeakers," *Proceedings of Meetings on Acoustics*, 43(1), p. 030002, 2021, doi:https://doi.org/10.1121/2.0001520.

[5] Choi, Y., Park, K., Lee, S., and Kim, K., "Panel vibration type display device for generating sound," 2017, US 2017/028246A1.

[6] Lee, S., Choi, Y., and Ham, S., "Display Apparatus," 2019, US 2019/0310684A1.

[7] Choi, Y., Oh, C., Park, K., and Lee, S., "Organic Light Emitting Display Device Including a Sound Generating Apparatus," 2017, US 9,818,805B2.

[8] Lee, S., Park, K., Choi, Y., Kim, K., and Bae, M., "Actuator Fixing Device and Panel Vibration Type Sound-Generating Display Device Including the Same," 2019, US 2019/0215607A1.

[9] Kim, T.-H. and Park, G.-C., "Display Device," 2019, US 2019/0163234A1.

[10] DiPassio, T., Heilemann, M. C., and Bocko, M. F., "Audio Capture Using Structural Sensors on Vibrating Panel Surfaces," *Journal of the Audio Engineering Society*, 70(12), pp. 1027–1037, 2022.

[11] Fuller, C., Elliott, S., and Nelson, P., *Active Control of Vibration*, Associated Press, 1996.

[12] Cremer, L., Heckl, M., and Petersson, B., *Structure-Borne Sound: Structural Vibrations and Sound Radiation at Audio Frequencies*, Springer Berlin Heidelberg, 2005.

[13] Fahy, F. J. and Gardonio, P., *Sound and structural vibration: radiation, transmission and response*, Elsevier, 2007.

[14] Mitchell, A. and Hazell, C., "A simple frequency formula for clamped rectangular plates," *Journal of Sound and Vibration*, 118(2), pp. 271 – 281, 1987, ISSN 0022-460X, doi:https://doi.org/10.1016/0022-460X(87)90525-6.

[15] Kerwin Jr, E. M., "Damping of flexural waves by a constrained viscoelastic layer," *J. Acoust. Soc. Am.*, 31(7), pp. 952–962, 1959.

[16] Ungar, E. E., Ross, D., Kerwin, J., and Edward, M., "Damping of flexural vibrations by alternate viscoelastic and elastic layers," Technical report, Bolt Beranek and Newman Inc. Cambridge, MA, 1959.

[17] French, N. R. and Steinberg, J. C., "Factors governing the intelligibility of speech sounds," *The journal of the Acoustical society of America*, 19(1), pp. 90–119, 1947.

[18] Inc, P. P., *Model 130F20 ICP® Electret Array Microphone Installation and Operating Manual*, 2016 [Online], https://www.pcb.com/contentstore/docs/pcb_corporate/vibration/products/manuals/130f20.pdf.

[19] "IEEE Recommended Practice for Speech Quality Measurements," *IEEE Transactions on Audio and Electroacoustics*, 17(3), pp. 225–246, 1969, doi: https://doi.org/10.1109/TAU.1969.1162058.

[20] IBM, *Watson Speech to Text*, [Online], https://www.ibm.com/cloud/watson-speech-to-text.

[21] Houtgast, T. and Steeneken, H. J., "The modulation transfer function in room acoustics as a predictor of speech intelligibility," *Acta Acustica United with Acustica*, 28(1), pp. 66–73, 1973, doi: https://doi.org/10.1121/1.1913632.

[22] Schroeder, M. R., "Modulation transfer functions: Definition and measurement," *Acta Acustica united with Acustica*, 49(3), pp. 179–182, 1981.

[23] PCB Piezotronics Inc, *Model 352C66 High sensitivity, miniature (2 gm), ceramic shear ICP accel., 100 mV/g, 0.5 Installation and Operating Manual*, 2002 [Online], https://www.pcbpiezotronics.fr/wp-content/uploads/352C66.pdf.

[24] Dimitriadis, E. K., Fuller, C. R., and Rogers, C. A., "Piezoelectric Actuators for Distributed Vibration Excitation of Thin Plates," *J. Vib. Acoust.*, 113(1), pp. 100–107, 1991, doi:https://doi.org/10.1115/1.2930143.