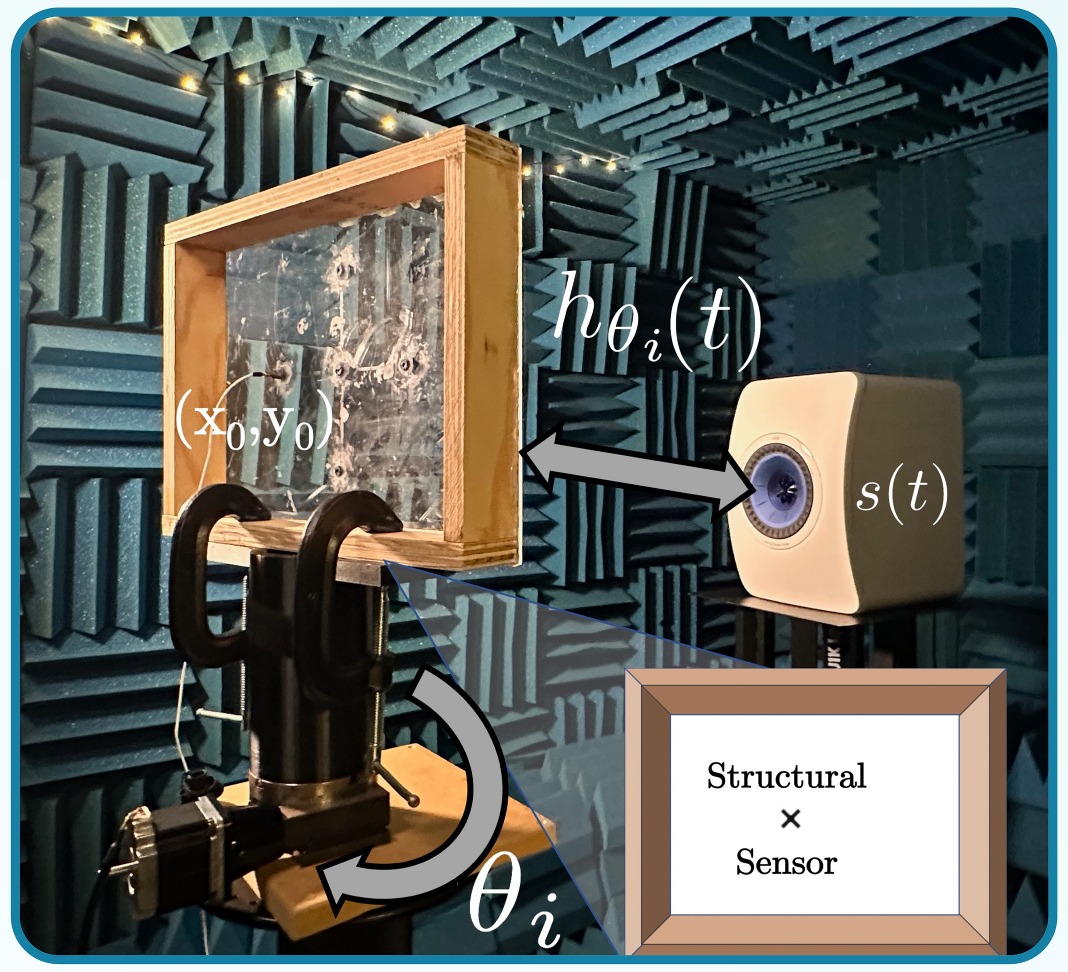
Estimating Acoustic Direction of Arrival using a Single Structural Sensor on a Resonant Surface

Tre DiPassio, Michael C. Heilemann, Benjamin Thompson, Mark F. Bocko
Department of Electrical and Computer Engineering

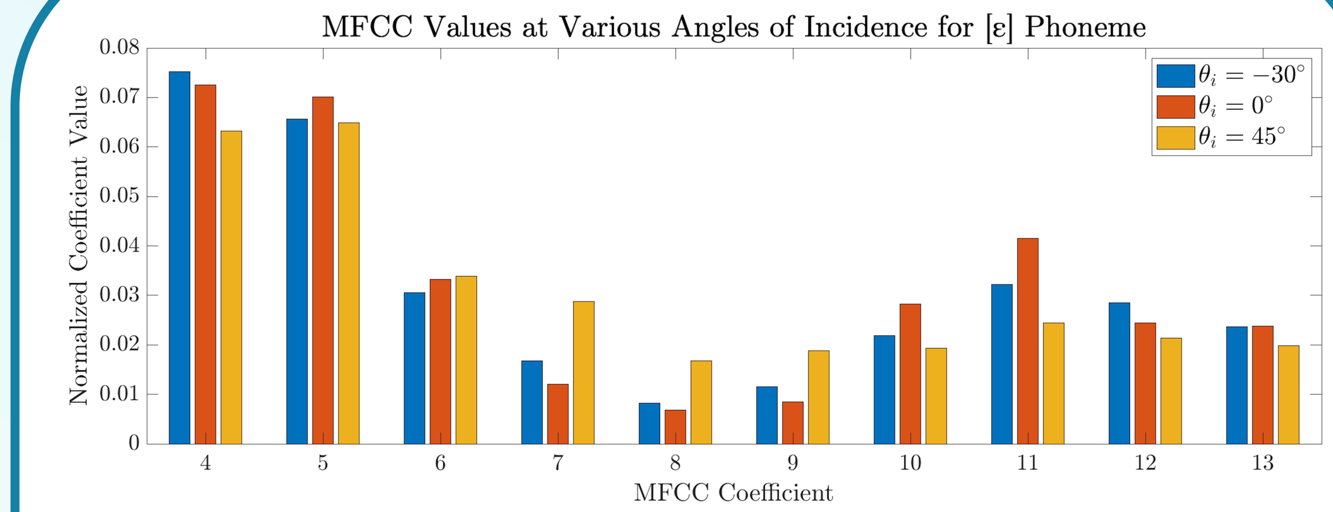
Abstract

The direction of arrival (DOA) of an acoustic source is a signal characteristic used by smart audio devices to enable signal enhancement algorithms. Though DOA estimations are traditionally made using a multi-microphone array, we propose that the resonant modes of a surface excited by acoustic waves contain sufficient spatial information that DOA may be estimated using a singular structural vibration sensor. In this work, sensors are affixed to an acrylic panel and used to record acoustic noise signals at various angles of incidence. From these recordings, feature vectors containing the sums of the energies in the panel's isolated modal regions are extracted and used to train deep neural networks to estimate DOA. Experimental results show that when all 13 of the acrylic panel's isolated modal bands are utilized, the DOA of incident acoustic waves for a broadband noise signal may be estimated by a single structural sensor to within 5° with a reliability of 98.4%. The size of the feature set may be reduced by eliminating the resonant modes that do not have strong spatial coupling to the incident acoustic wave. Reducing the feature set to the 7 modal bands that provide the most spatial information produces a reliability of 89.7% for DOA estimates within 5° using a single sensor.

Experimental Setup



Direction of Arrival Estimation using Harmonic Properties



Selected **mel-frequency cepstral coefficients (MFCCs)** extracted from a recording of an elastic panel's vibrational response to acoustic waves containing the **speech sound "eh"** incident at −30°, 0°, and 45°. The recording of the panel's vibrations was made using a single structural vibration sensor. This figure demonstrates that the MFCCs are dependent on the incident angle of the acoustic wave. A neural network may therefore utilize an MFCC vector to create decision boundaries and estimate the DOA of the excitation using information from a single structural vibration sensor.

Material	# Sensors	Reliability (%) to within: ±5°	±10°	±20°
Acrylic	1	99.8	99.9	99.9
	3	100	100	100
	5	100	100	100

Tabulated is the reliability of the DOA estimates made by a recurrent neural network trained with **MFCC feature vectors** extracted from recordings of a panel's vibrational response to incident **broadband noise bursts**. The recordings were made using 1, 3, and 5 structural vibration sensors. The results demonstrate that MFCC feature vectors can be employed to estimate DOA using a single structural vibration sensor.

Results are reproduced from our recent publication in the Journal of Sound and Vibration, linked here

Motivation

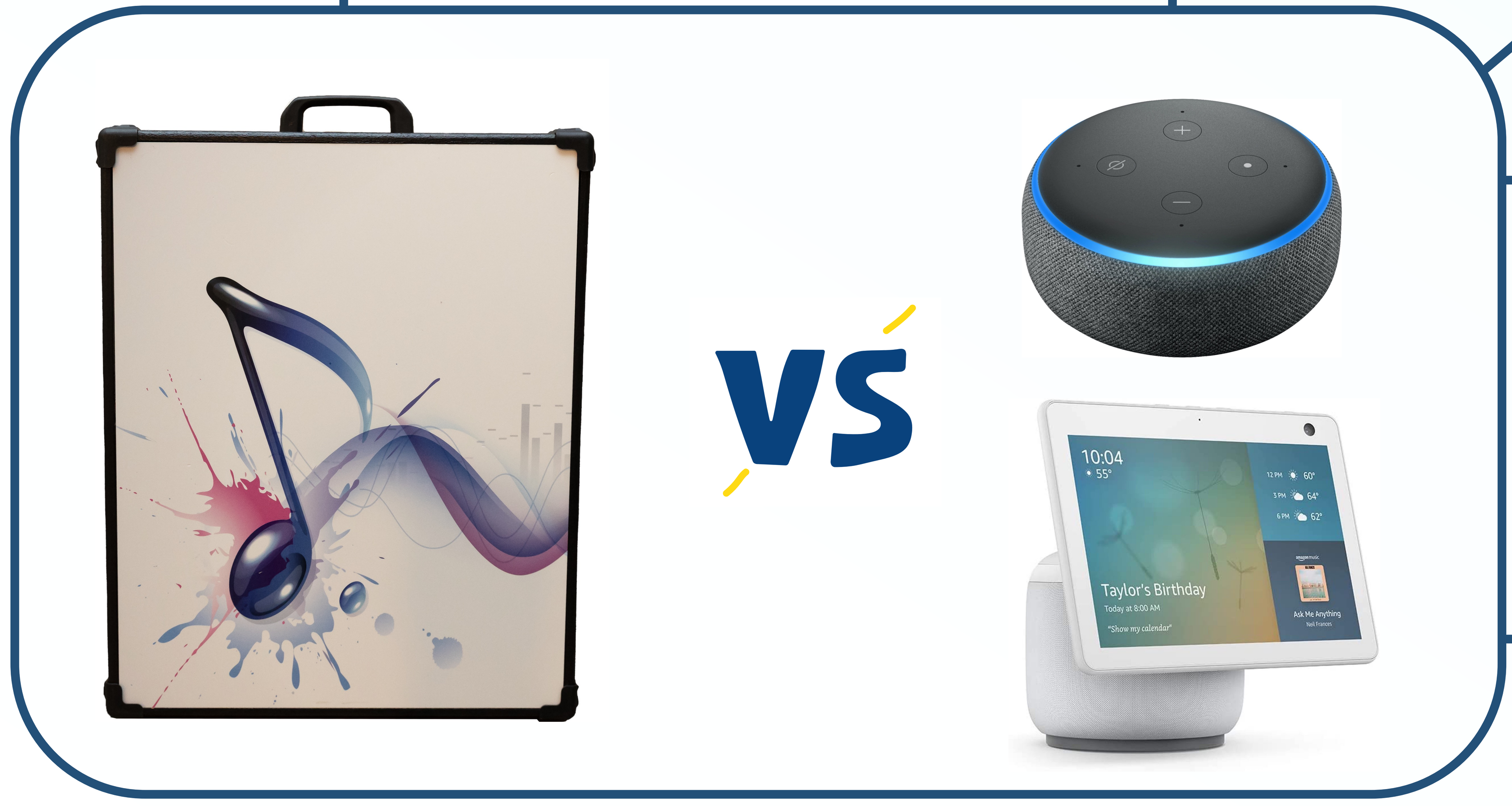
Smart acoustic surfaces allow for **seamless integration** of a smart speaker into existing environments, as any surface (such as picture frames and artwork) can be used

Mounting sensors internally to the display **eliminates the need for case penetrations,** improving the device's water resistance and durability

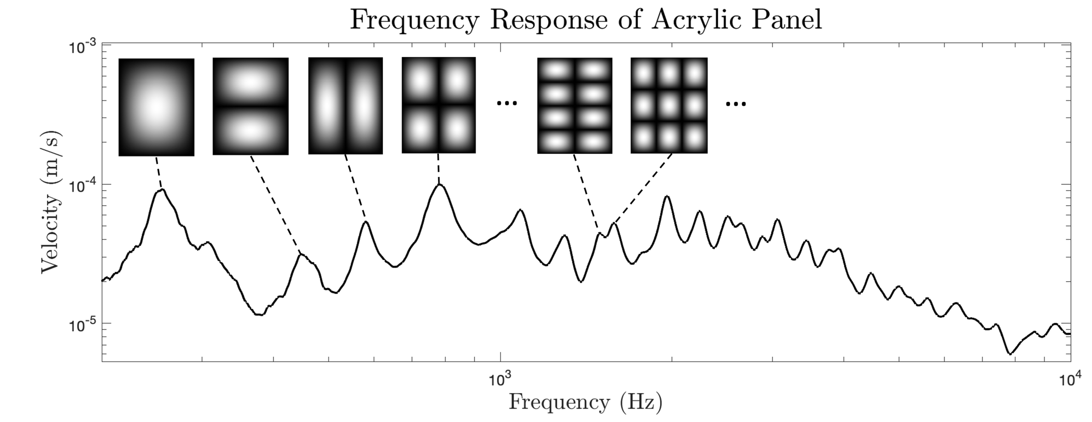
Extended surfaces allow for **signal processing advantages**, as sensors can be placed further apart than the standard 1-4 cm on modern smart devices

By coupling to a modal surface, direction of arrival estimation and beamforming can occur with as few as one sensor, which can **lower manufacturing cost**

Smart devices with display panels can include **higher-fidelity audio without changing form factor** by using their screen's surface for audio reproduction



Experimental Results using a Resonance-Informed Feature Set



The panel's spatially averaged velocity response measured using a scanning laser vibrometer. Peaks in the response occur at frequencies where one or more of the panel's bending modes resonate. The relative amplitudes of these peaks are determined by the spatial coupling between the mode shape, and the incident angle of the acoustic source signal. Modes whose amplitudes have a greater variance with incident angle provide the most information about the direction of the source, while modes whose amplitudes remain constant with incident angle provide relatively little information about the direction of the source. From this, a hierarchy may be determined that ranks each panel mode based on its ability to convey directional information.

f_c (Hz)	Δf (Hz)	Mode(s)	Rank
256.5	60.0	(1,1)	5
454.9	115	(1,2)	T6
582.7	88.6	(2,1)	2
784.2	146	(2,2)	T6
1168	312	(2,3)	1
1287	210	(3,2)	T6
1564	340	(2,4),(3,3)	3
1962	224	Unclear	N/A
2233	317	(4,3)	4
2554	666	Unclear	N/A
3028	707	Unclear	N/A
3430	506	Unclear	N/A
3832	900	Unclear	N/A

The center frequencies f_c, bandwidths Δf, and the excited modes in each of the isolated resonant bands shown in the figure on the left. The rank is determined by the ability of the incident wave's angle to affect the excitation level of each mode

The **resonance-informed filter bank** was employed to train a recurrent neural network to estimate the direction of arrival of a **broadband noise burst.** This feature set is more compact than the MFCC vectors, which require the application of up to 40 filters. Additionally, the spatial information contained in the panel's isolated resonant bands is directly computed. The reliability of the DOA estimates made by the network to within a defined angular tolerance are tabulated at right. To further abridge the feature set, modal bands are removed from the feature vector in the leftmost columns in order by the least excitation variance due to the angle of incidence, and removed in order by the most angular variance in the *italicized rightmost columns.*

# Bands	Reliability (%) to within: ±5°	±10°	±20°	±5°	±10°	±20°
13	98.4	100	100			
12	97.8	100	100	*98.1*	*100*	*100*
11	96.2	99.8	100	*97.7*	*99.9*	*100*
10	96.6	99.8	100	*96.6*	*99.9*	*100*
9	94.5	99.3	100	*91.7*	*99.1*	*99.5*
8	92.1	99.3	99.8	*0.880*	*98.4*	*99.4*
7	89.7	98.3	99.8	*83.6*	*95.8*	*97.4*

Bands removed by **least** excitation variance | Bands removed by **most** excitation variance