Uminari: Freeform Interactive Loudspeakers

Yoshio Ishiguro *
Disney Research
Pittsburgh,
4720 Forbes Avenue,
ishiy@acm.org

Ali Israr
Disney Research
Pittsburgh,
4720 Forbes Avenue,

Alex Rothera
Disney Research
Pittsburgh,
4720 Forbes Avenue,

Eric Brockmeyer
Disney Research
Pittsburgh,
4720 Forbes Avenue,

{israr, alex.rothera, eric.brockmeyer}@disneyresearch.com

Figure 1. Freeform Interactive Loudspeakers are flexible, lightweight and used in interactions with everyday objects. (a) The lightweight electrostatic loudspeakers are integrated in a *balloon*. (b) The electrostatic loudspeakers can be made into *large curved wall* surfaces. (c) Our *speaker table* accepts touches and weight on it. (d) *Tangible objects* are made speaker by equipping them with passive layer of conductive paper.

ABSTRACT
We present freeform interactive loudspeakers for creating spatial sound experiences from a variety of surfaces. Surround sound systems are widely used and consist of multiple electromagnetic speakers that create point sound sources within a space. Our proposed system creates directional sound and can be easily embedded into architecture, furniture and many everyday objects. We use *electrostatic loudspeaker* technology made from thin, flexible, lightweight and low cost materials and can be of different size and shape. In this paper we propose various configurations such as single speaker, speaker array tangible speaker and microphone configurations for creating playful and exciting interactions with spatial sounds. Our research of freeform speakers can create new possibilities for the design of various interactive surfaces.

Author Keywords
Interactive architecture; furniture design; interactions; electrostatic loudspeakers; freeform speakers

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces - Auditory feedback.

* Currently at Nagoya University.

ITS 2014, November 16–19, 2014, Dresden, Germany.

ACM 978-1-4503-2587-5/14/11…$15.00
http://dx.doi.org/10.1145/2669485.2669521

INTRODUCTION
Sound and visuals are important ingredients for creating interactive systems that produces an engaging and believable user experience. Traditionally, these systems utilize a high definition visual display alongside multidimensional surround sound systems to render precise spatial sounds collocated with the animated visuals. Current display technologies, however, pose limitations in the overall user experience due to their structural and functional limitations, as well as due to the high cost and power demands associated with acquiring and using such technologies. Our research is inspired by such limitations and we explore technologies that are cost effective, scalable and multi-functional. Overcoming functional limitations in everyday objects is an emerging conceptual framework in the current HCI community [6,7,25].

In recent years, research concepts like *Programmable Matter* [7], *Programmable Physical Architectures* [26], *Programmable Reality* [5] and *Organic User Interfaces* [11] have been introduced that, at least conceptually, extends the flexibility and functionality in real physical objects and morph them into forms that are more suitable for natural and seamless user interactions. For example, display technology has progressed into more flexible, scalable and freeform configurations by use of organic electroluminescent, electrophoretic ink, liquid crystal, LED and beam-steering technologies [11]. These technologies are frequently used in highly immersive virtual reality environments.

On the other hand, freeform and organic functionality in sound reproduction systems has been limited and traditionally large and bulky electromagnetic surround sound speakers are used for 3D sound reproduction (such as

[12]). These surround sound systems are not applicable in interactive and transformable spaces due to the following restrictions:

- Sound localization area is limited
- A large number of speakers are required
- The speakers are placed throughout the environment
- An extensive calibration process is required for a given space configuration

Recently, Ishiguro and Poupyrev presented 3D speakers that utilized 3D printing technology and electrostatic loudspeaker technology to fabricate small objects with loudspeaker function [13]. Due to the layered feature of the electrostatic speaker technology, these objects, such as toys, emit sound all around its external surface.

In this paper, we introduce *freeform electrostatic loudspeakers* (ESL) that can be easily integrated in structures, furniture and everyday objects. A simple ESL [9, 18] consists of two layers of conductive material separated by an insulator medium. An alternating potential difference between the two conductors creates electrostatic coupling between them and moves the air around the layer to produce sound. The ESLs are lightweight, flexible, scalable, cost and power efficient and can be made into almost any shape and size. They have high durability and ideal for use in spatially restrictive and kinetic spaces and structures. Unlike [13], where the focus was to manufacture small toys and objects, the present work utilizes the ESL technology for large-scale objects, such as architecture and furniture. Moreover, we extend the practical use of ESL and introduced programmable and interactive audio environments. We call this system *Uminari.* In Japanese literature, *Uminari* means 'rumbling in the sea'.

The organization of the paper is as follows: In the next section we discuss relevant background related to freeform ESLs. This is followed by the design and evaluation of a series of ESLs made in our facility. A pair of preliminary user evaluation studies follows this. We then explore the application space for these loudspeakers where we have integrated them into architecture, furniture and everyday objects. The paper concludes with a discussion and remarks.

BACKGROUND

Freeform structures, hardware and interactions have become increasingly common, in recent years, which has allowed researchers, designers and engineers to expand their designs with no boundaries [7,14,19,20,22,27]. Furthermore, with advances in rapid prototyping techniques, 3D printing, open source software and novel material research, freeform culture has progressed in to making everyday objects, increasing their functionality and accessibility [4,7]. The challenge, however, is to incorporate functionality beyond the conventional modality of structures; for example, it is not trivial to develop a large erected wall that can be folded to a form that is portable and easily deployed. We are inspired by the physical limitations of objects and structures and equip them with functionality that expands user interactions with such objects and structures.

User interactions on surfaces have been the focus of many studies in HCI where functional capabilities of surfaces were enhanced [2,3,6,8,15]. Incorporating modal functionality not only increases user interactions in these environments but also prompt “surprise” elements in controlled settings, such as in amusement parks, interactive installations, educational classrooms, museum exhibits [2,10, 24].

We utilize electrostatic speaker technology to enhance structures and everyday objects with sound reproduction capabilities. The technology is well explored and published [9,18,21] and has been applied to commercial speaker systems. Recently, Ishiguro and Poupyrev [13] combined these techniques with 3D rapid printing technology to fabricate the loudspeakers as small and complex shaped objects. This research showed high flexibility and scalability of electrostatic loudspeakers embedded in toys, ornaments, and small decorative pieces.

The traditional method for enhancing the functional capabilities of objects with a sound source is to embed a traditional speaker in the object, such as in *Acoustable* [1]. In this case, large cavities are designed inside the object to place the speaker and power system, and shape of the object is designed to optimize the sound quality. The entire object resonates and produces sound as if it were coming from the object. However, the shape and size of the speaker highly depend on the size, shape, form and functionality of the object. Moreover, sound localization is limited and sound quality is greatly affected when the system is used with other objects.

Another way to add sound reproduction capabilities to objects is to utilize the parametric speaker technology [16,28]. This technology utilizes ultrasound to modulate with the sound signal, producing highly directional sound at the source. This modulated sound is aimed towards objects equipped with a layer of special reflective material, which absorb the ultrasound carrier and reflects the sound as if it was reproduced at the object instead of at the source.

Other alternatives for freeform speaker technology include the use of piezoelectric elements and electro active polymers in the speakers [23]. These materials are lightweight and can be fabricated into any shape and size. However, the raw material and fabrication techniques are quite expensive, and in the case of for large size speaker application, not practical.

DESIGN OF FREEFORM LOUDSPEAKER

Principles of ESL Technology

The basic principles of electrostatic sound reproduction are simple and were explained in detail in [9, 13]. In a simple construction, an ESL consists of a thin conductive diaphragm and an electrode plate, separated by an insulation layer as shown in Figure 2. When the high alternating voltage difference is applied between the diaphragm and the electrode, then the electrostatic attraction forces are generated between these conductive layers and produce relative motion between them. The thin diaphragm pulls towards the electrode plate and displaces the air along its motion. If the high voltage is replaced by amplified audio signal, then the displaced air would create an audible signal and the diaphragm actuated with electrostatic force would become a speaker.

It has been demonstrated that ESLs allow for high quality sound reproduction, outperforming traditional electro-magnetic speakers in both high and medium sound frequency ranges [18,21].

Configurations of ESL

In this section we describe practical configurations of ESLs. These configurations will be used in applications described in the next section.

Single-Side Configuration

The single-side configuration is the simplest form of ESL and consists of two layers (a diaphragm and an electrode) separated by an insulator (Figure 2). In the ESL speakers reported here we connect the ground to the diaphragm and the audio signal to the electrode (Figure 3b). To create different potential between two conductive materials (diaphragm and body), they both had different potentials from the ground as shown in figure 3a. However, this is dangerous for the human body which normally has potential same as ground. It is important that interactive speakers embedded in objects are safe and can be touched by users. Therefore, the grounded electrode is placed on top and protects the user touching the speaker from the high-voltage audio source, making it safe to handle and manipulate the object with embedded speakers. This becomes particularly important in the interactive applications. This configuration can be used as surface of tabletops, walls, toys, floating objects, hanging center pieces, and many more. In this configuration, either the diaphragm or the electrode of both could be free to move and the high voltage audio signal can be either applied to the diaphragm or to the electrode, allowing ESLs to be utilized in variety of situations. Moreover, the ESL can be used as a microphone by replacing the voltage driver with a pre-amp circuitry.

Array Configuration

An extension of the *single-side* is the *array* configuration as shown in Figure 3c. The array allows ESLs to be placed spatially, where each loudspeaker is controlled independently. Multiple sound patterns can be channeled through individual loudspeakers creating a moving, expressive and dynamic sound system. All loudspeakers in the array can have a common ground plane. Therefore, a single diaphragm surface could be used with multiple loudspeaker electrodes.

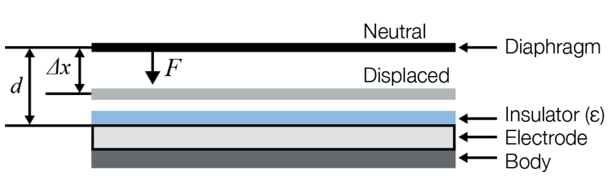

Figure 2. Construction of an electrostatic loudspeaker

Passive Loudspeaker Configuration

Figure 3d shows the *passive loudspeaker configuration*. In this configuration, high voltage electrodes and the ground electrodes are arranged on the same surface, and an object with passive diaphragm (conductive sheet) is placed on the surface. The surface itself cannot produce sound, however, the conductive surface of the object creates potential difference between the diaphragm and the high voltage electrode as well as between the diaphragm and the ground electrode. Changing the potential difference changes the electrostatic potential of the diaphragm. Therefore, the passive object behaves as a switch for sound reproduction.

Advantages of Electrostatic Loudspeaker

Ease of construction

One of the most important features of the ESL is ease of its construction. Instead of winding coils in a specific manner around a magnet for electromagnetic speakers and complicated fabrication methods for piezoelectric speakers [17], the ESL can be made in the nominal laboratory and/or workshop settings using everyday components.

Size and shape

Due to the ease of construction and cost, the ESLs can be made in different sizes using the same construction procedure. Unlike piezoelectric and electromagnetic speakers, where the cost of magnet, coils and piezoelectric material increases exponentially with size, the cost is linearly related to the size of the loudspeaker. Moreover, the same electronic driver is used for small and large loudspeakers. There is no added cost or power requirements for the driver, which is not the case with other speaker technologies. And the ESLs can be made in different shapes, such as curved, spiral, circular, cylindrical, etc.

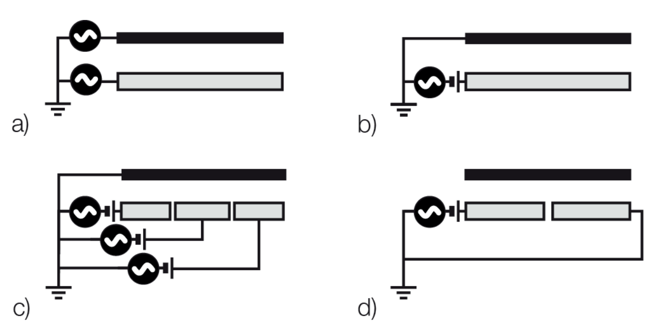

Figure 3. Four configurations of ELSs.
(a) Conventional speakers, (b) single-sided, (c) arrayed, and (d) passive electrostatic loudspeakers.

Weight
ESLs are inherently light and could be made with conductive paper and sheets, such as Mylar and Indium tin oxide (ITO) coated polyethylene terephthalate (PET). Other conductive materials such as Graphene and Carbon nanotube (CNT) structure can also be used for the construction of transparent ESLs. Due to its lightweight characteristics, ESLs are ideal for floating and hanging objects, such as balloons, lampshades, and for pages of storybooks.

Power requirement
The overall loudness depends on the potential difference across the two surfaces, and not the current. Therefore, very minute current is required to generate sound (~ 1.5 mA). In electromagnetic speakers, current is a basis factor to generate torque to move the mass in order to generate sound, which could require higher amperage especially in large size speakers.

Cost
The minimal construction of an ESL requires two sheets of thin conductive material, wires, tape and electronic driver. These items are quite cheap and easily available in a typical general or hardware store.

SONIC PERFORMANCE OF FREEFORM ESL

The performance of ESL depends on many factors, mainly shape and size of the speaker surface. In this section, we measure the performance of freeform ESLs for their reliability and use in a variety of interactive applications. Specifically, we are interested in measuring the detailed frequency response, effect of size on uniformity and directivity of sound, and effect of loading placed on the diaphragm of the ESL. We also measure the performance of ESL microphone in a normal setting. These measurements are necessary for understanding the behavior of ESL with a variety of sound and interactions, and serve as a reference for designers extending the freeform nature of ESL.

Two test speakers are assembled by stretching a thin aluminized mylar sheet on top of a perforated stainless steel sheet. The dimensions of the smaller loudspeaker are 15 cm long by 15 cm wide and that of the larger are 35 cm long by 19 cm wide. These speakers are housed in an acrylic frame as shown in Figure 6. All measurements are taken in a quiet room with background noise level below 45 dB SPL. For testing ESL microphone, a small panel equipped with a simple microphone pre-amp circuitry is also produced and sound recordings are evaluated.

The audio signal is generated through a pc output through an audio interface and passed through a transistor based custom driver, which amplifies the input signal from ~1V_{p-p} to ~1000 V_{p-p}. The positive terminal of the output is connected to the fixed metallic mesh (electrode) and the ground terminal is connected to the aluminized mylar. The sound is measured with a digital sound pressure level (SPL) meter (Extech Instruments Corp., model 407730, Nashua, USA) place 15 cm away from the surface of the speaker.

Sound Quality

The sound quality is measured by taking the frequency response of the larger speaker. Instead of running a typical frequency sweep, we use single frequency sinusoids to measure gain, distortion, harmonics and residual background noise levels. Figure 4 shows the Fast Fourier Transform (FFT) spectrum of the measured sound signal at 0.5 kHz, 1 kHz, 2 kHz and 4 kHz. The magnitude of the FFT is normalized by number of samples. Similar to the measurements made in previous literature (such as [21]) distortion and harmonics of 500 Hz sinusoids were significant and could be sensed by a normal ear. At frequencies 1000 kHz and higher, signal distortion were low and very low harmonic overtone.

Sound Uniformity and Sound Directivity

One important feature of the electrostatic loudspeaker technology is that the sound radiates uniformly throughout the surface of the speaker, which is not possible with point sound source of electromagnet speakers. Figure 5 shows the measurements taken from the SPL meter at seven locations (-200mm, -100mm, -50mm, 0mm, 50mm, 100mm and 200mm) along the surface of the two electrostatic loudspeakers. The plotted data is normalized at 0 SPL of 0 mm measurements.

For the smaller loudspeaker, the maximum sound level was recorded at the center of the speaker and the sound intensity is maintained along most of the width, before it steeply

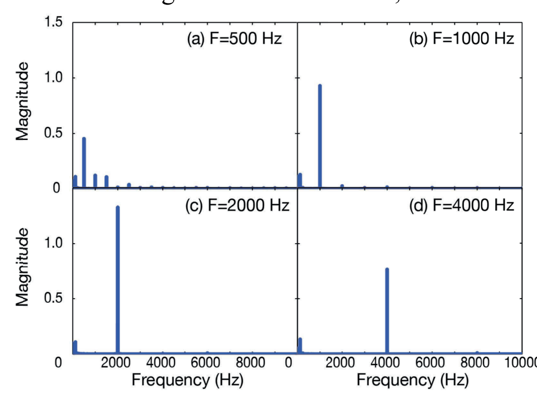

Figure 4. FFT spectrum of electrostatic loudspeakers at (a) 500 Hz, (b) 1000 Hz, (c) 2000 Hz and (d) 4000 Hz.

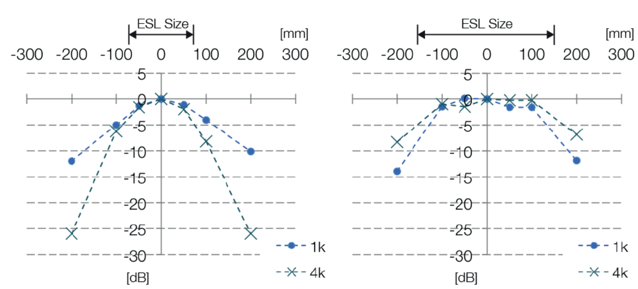

Figure 5. Measurement of Sound Pressure Levels along the surface of a 150 mm wide loudspeaker (*left*) and 350 mm wide loudspeaker (*right*).

reduced towards the edges. Similarly, the sound intensity of the larger loudspeaker was not changed across the width of the loudspeaker and the sound intensity significantly decreased beyond its width. This shows that the electrostatic loudspeakers radiate sound uniformly along its surface and therefore are a good candidate for maintaining sound directivity.

Effect of Loads

A normal electromagnetic speaker generates sound by moving a coil and diaphragm. Any load or obstruction placed on the diaphragm hinders the motion and as a consequence the sound intensity is significantly compromised. The ESL has minimal effect of loading placed on its diaphragm. Figure 6 shows the effect of loading on the diaphragm of the larger electrostatic loudspeaker. The overall shape of the frequency spectrum was maintained, however the intensity level was reduced by ~5 dB when a load of 500 g was placed on the loudspeaker.

Measurements of ESL Microphone

The performance of ESL microphone is assessed by comparing the frequency response of the sound recorded by an ESL microphone and by a conventional microphone (*Electro-Voice*, model 635N/D-B). Figure 7 shows the frequency response of both microphones when a white noise (top) and a whistle (bottom) are played on a conventional electromagnetic speaker placed in between the two microphones, 20 cm from each microphone. The level of white noise is 77 dB and that of whistle is 95 dB measured with the SPL meter 20 cm away from the speaker. The gain of the microphones is adjusted to record comparable levels.

Figure 7 shows that the spectrum of two microphones matches in the mid-frequency range; however, the low-frequency sound (<1000 Hz) is dampen by the ESL microphone. Moreover, the ESL recording has a uniform noise across the spectrum. The sound quality of ESL microphone can be improved by software sound processing and modified electronics.

PRELIMINARY INTERACTION USER STUDIES

The sonic performance of the freeform ESL used in various applications is described in the previous section. In this section, we evaluate the performance of electrostatic loudspeaker and microphone in two user studies. Both studies are designed to evaluate functional reliability of ESL while a human user interacts with them. In the first preliminary study, we compare the performance of ESL with conve ntional stereo speakers in sound localization. In the second study, we evaluate the performance electrostatic microphone in understanding speech.

Study 1: Evaluation of Electrostatic Loudspeaker

A typical application of ESL is to embed them in large walls with projected imagery. Any non-collation of imagery and sound will deteriorate the user experience and absorption with the content. Therefore, in the first study we measure the localization performance of an ESL and compare it with conventional stereo speakers.

Procedures

We mount an ESL (102 cm wide by 45 cm high) on a wall and attach conventional speakers (*Bose*, model Computer MusicMonitor speakers) on the two extreme sides of the ESL. Participants stand in front of the center of the ESL and listen to the sound from ESL and stereo speakers. In a trial, a sound is produced either through the ESL or through stereo speakers and participants indicate if the “perceived sound source is directly in front of them”. They are asked to stare straight at the center of the ESL and respond by saying ‘yes’, if they feel that the sound source is on the surface they are looking at, or by ‘no’, if otherwise. A 3.7 second animated sound is used throughout the experiment. The stereo speakers are balanced equally to simulate the sound source at the midline between the two conventional speakers. Three participant locations are tested. These locations are 30, 152 and 274 cm in front of the ESL along the midline of the ESL.

Ten participants (5 males, 19 to 56 years old) take part in the study. None of the participant report auditory impairments. They are recruited from our research facility and are not paid for their participation. Each participant is tested for 60 trials (3 distances x 2 sound sources x 10 repetitions) that lasts ~10 minutes. Participants also try a few training trials before starting the experiment.

Results

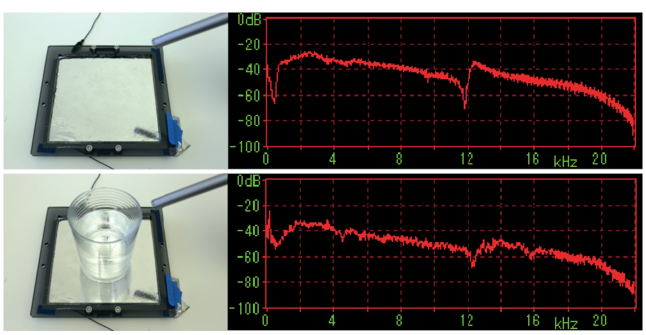

Figure 6. Frequency response measurements of an electrostatic loudspeaker in the free unloaded state (*top*) and in the loaded state (*bottom*)

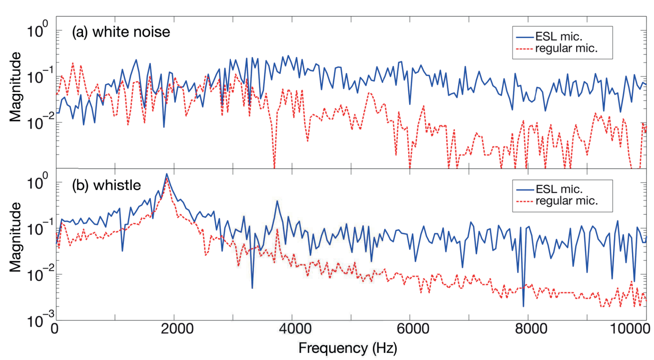

Figure 7. Spectrum of (a) white noise and (b) whistle recoded with ESL and standard microphones.

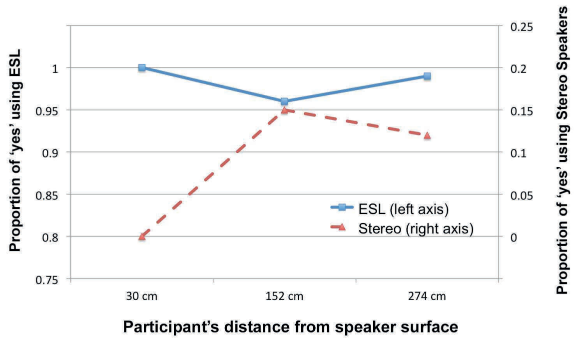

Figure 8. Sound source localization performance of ESL and conventional stereo speaker

Figure 8 shows proportion of 'yes' (p('yes')) when the sound was produced by ESL and by conventional stereo speakers as a function of the location of participants. At 30 cm separation (very close to the speaker), participants always responded 'yes' to ESL and never to the stereo speakers, indicating that the ESL provides the best collocation performance of projected imagery at close proximities. At distances 152 cm and 274 cm, the rating of ESL was high (>95% of the time); however p('yes') using stereo speakers also started to improve. Pairwise T-tests showed that the localization performance was significantly different between 30 cm and 152 cm (p=0.04) and 30 cm and 274 cm (p=0.03), however it was not different between 152 cm and 274 cm (p=0.4). In conclusion, the ESL is better than the stereo speakers in localizing the sound source on the surface, and therefore enhances the interactions with the projected content on the surface.

Study 2: Evaluation of Electrostatic Microphone

An application of the electrostatic microphone is to capture the sound from the environment and use it for sensing or communication purposes. In this preliminary study, we evaluate the microphone configuration of the ESL for speech communication.

Procedures

Two non-native (one male and one female) speakers utter five non-sense consonant-vowel (CV) syllables and are recorded using the ESL based microphone. These syllables are /ba/, /pa/, /be/, /ve/ and /he/. The first two syllables differ in *voicing* and other three syllables differ in *place* and *manner* of articulation. This speech corpus is used to evaluate how well the phonetics can be recorded through the microphone.

Twelve participants (8 males, ages 22 to 38 years old) listen to each of the ten syllables. Half of the participants are asked to guess the syllable (unprompted) and the other half are given possible five responses (prompted). The order of test syllables is randomized for each participant. The mono-channel microphone recording is saved as stereo signal and played only once though a headphone worn by participants.

		Response				
		/ba/	*/pa/*	*/be/*	*/ve/*	*/he/*
Stimulus	*/ba/*	1.00	0.00	0.00	0.00	0.00
	/pa/	0.50	0.50	0.00	0.00	0.00
	/be/	0.00	0.00	0.92	0.08	0.00
	/ve/	0.00	0.00	0.00	1.00	0.00
	/he/	0.00	0.00	0.00	0.00	1.00

Table 1. Stimulus-response confusion matrix of prompted speech identification using ESL microphone

Results

Participants correctly identified unprompted syllables 43 out of 60 times. In many cases, /ve/ was identified as /we/ and /pa/ uttered by the male speaker was always responded as /ba/. For prompted syllables, participants correctly identified the syllable 88% of time. The stimulus-response matrix is shown in Table 1. All /pa/'s uttered by the male speaker were identified as /ba/, and participants made only one error response (/be/ identified as /ve/). In general, the ESL microphone transmits sound and speech signal reliably with the exception of /b/ from /p/ that could be due to the distortion in low-frequency signal by the ESL microphone.

DESIGN SPACE AND APPLICATIONS

In previous sections, we describe the design freedom and characteristics of freeform ESLs, which allows them to be used in wide range of applications. Due to the shape, size and flexibility in these ESLs, we focus the applications to architecture, furniture and everyday objects that have been made and installed for in-house testing and evaluation. In each subsection below, we also highlight design and interactions associated with each application.

Architectures

Large Curved Walls

The freeform ESLs are very good candidate to embed in large and complex shaped walls because it has high flexibility of shape and it is scalable. We integrated the ESLs into a wall setting as shown in Figure 9. The requirements for the installations were: (i) An immersive multisensory experience, (ii) on a large high definition curved display, and (iii) spatial media to make animations feel real and collocate the sound with moving visuals.

We fabricated an array of six ESL panels and mounted them on a curved frame on the wall. The array was covered with a single sheet of white paper, which was used as a surface for projected graphics. A high definition projector was mounted on the ceiling that projected animated video sequences on the wall as shown in Figure 9. Instead of replicating our electronic driver for six channels, we designed a multi-channel ESL driver, which consist of multiple high voltage amplifiers using a single high voltage power supply. A single 1000 V (1.25mA) high voltage supply (EMCO, model QH10) was mounted on the motherboard that provided the high voltage reference to up to eight channels. The sound experience was designed in max/msp that channeled six outputs through a MOTU USB audio interface (model UltraLite-mk3 Hybrid). We divided the display into six sections and used sound fading techniques to create seamless sound movements related to the motion of animated objects.

We demonstrated the installation to technical staff and guests visiting the facility to gather feedback. All guests agreed that the installation was immersive as the sound was collocated with the projected imagery and moved with the animated imagery along the wall. They were also intrigued by the flexibility of the ESL in limited space. Guests frequently moved close to the display and touched it. In order to avoid passing high voltages (current was limited to less than 1.25 mA) to guests, the ground potential electrode (metal mesh panels) was placed on the outside and the diaphragms connected to the high voltage audio signal were placed in between the wall and ground potential electrode. Therefore guests could not touch the diaphragm.

Morphing Structures

Our freeform ESLs are light and generate directional and uniform sound along the surface of diaphragm. We used these speakers for transforming room environments as shown in the creative concept in Figure 10. In addition to the features of ESLs described above, we also used them to work as a microphone with modifications to the driver. In this case, the same ESL unit was used as a sound reproduction as well as the sound recoding device. The requirements of the installation were: (i) A room transforms into different configurations, (ii) each panel has a speaker and a microphone, (iii) sound is directional and targets specific guest from the audience, and (iv) transforming panels must be lightweight.

Figure 10 shows the moving speaker structure. The loudspeaker panels were suspended by thin strings, whose lengths were controlled by a series of servomotors. By varying the length of strings, the artist could vary the height and alignment of each panel. A user whispered to the ESL and the sound was played back through the speaker array. The recorded sound was processed in a max/msp patch and passed through the speaker array to create animated sound experience with dynamically morphing structures. The installation was used for internal testing and measurements.

Furniture

Interactive Tabletops

In this installment, we integrated an array of nine freeform ESLs on the top surface of a wooden table and enhanced tabletop activity between users. Eight speakers are placed radially across the table and the ninth circular speaker is embedded in the center as shown in Figure 11. The

Figure 9. Freeform ESLs are arranged to form a curved projection surface. The sound and animated graphics are collocated using the arrayed configuration.

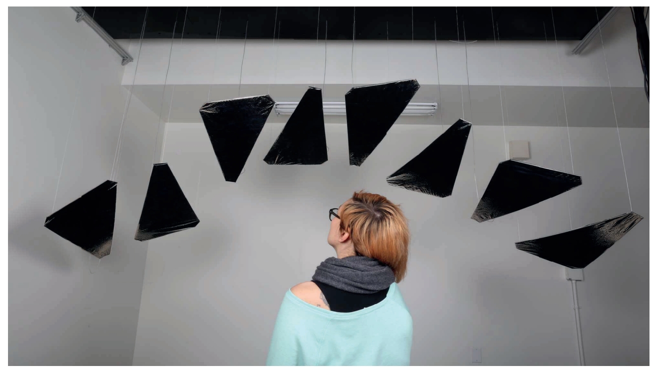

Figure 10. A user interacts with moving ceiling panels that can record users' sound and play them at different locations in the room.

centerpiece was also mounted with a subwoofer to produce low-frequency sound. The setup was similar to that used in the large curved wall installation above. A projector was mounted on the ceiling to project animated images on the table. The resulting installation was used in an interactive dinning situation where dynamic sounds were produced while users had a dinner served on the table.

Six highly skilled concept designers and media artists, who evaluated the performance of ESL and provided recommendations. The main feedback was that the installation had a surprise element and could not be achieved only with normal loudspeakers. The sound localization was effective but sound levels were low and distorted. We modified the amplifier driver noting that the amperage would be lost in transistor circuitry. The larger amperage (10 mA) improved both the quality and perceived loudness of the sound. Due to the minimal effect of loading on ESL, the food served on the table did not affect the sound quality. Guests also spilled water on the table that did not affect the sound and safety quality.

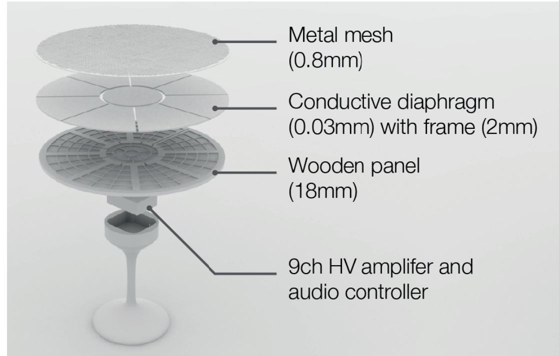

Figure 11. The surface of a large round table is embedded with an array of freeform electrostatic loudspeakers.

Cushions

By using soft conductive cloth - cushions, pillows, mattresses and similar soft materials can also reproduce sound. Figure 12 shows an ESL pillow. The pillow is made by wrapping two layers of conductive cloth around a pillow. Each conductive layer acts like a diaphragm. A normal non-conductive fabric covers the, thus isolating conductive diaphragm from direct user contact. Low sound levels were reproduced by the soft pillow speaker because the distance between the two diaphragms was not close enough to reproduce enough audible sound pressure level. However, when a user put the head on the pillow, then the diaphragms are pressed against each other and produce light perceivable sound. This sound level was sufficient to utilize the pillow, because the head (and consequently the ears) would be directly placed on the pillow.

Other Everyday Objects

Toys and Tangibles

In all previous cases, we implemented the active mode of the freeform ESL, that is, the entire speaker configuration was integrated onto the surface, for example, on the wall surface or on the table surface. In some situations, interactions were desired with passive tangible objects. Here an object equipped with a passive diaphragm layer is placed on a fixed electrode pattern embedded on a surface. Figure 13 shows a table with a grid pattern made from conductive tape and the tangible object. The tangible object only creates sound when placed on the table and can be used as toys for playful user experience.

Floating Speakers

The lightweight nature of ESL allows them to be a candidate for floating lightweight objects, such as balloons, as shown in Figure 1. In this case, the high voltage electrode was placed inside the balloon and the outer diaphragm layer was grounded. The quality of sound reproduced with these floating loudspeakers depends on the pressure of air inside the balloon. An inflated balloon would generate louder sound than a half inflated balloon.

DISCUSSION AND CONCLUDING REMARKS

In this paper, we have explored programmable matter to reproduce sound for everyday interactions, by integrating loudspeakers into surrounding architectures, furniture, tangible objects, soft materials, floating and many other everyday objects. We have modified the current electrostatic speaker technology and configured them to produce durable, safe and interactive environments. The resulting freeform electrostatic loudspeakers are flexible, lightweight, scalable and inexpensive.

In a series of experimental measurements, we determined that ESL has following characteristics:

- Low frequency sounds are distorted higher than the high frequency sounds.
- The sound radiates out of the entire surface of the diaphragm.
- The sound is highly directional
- The effect of loading on the diaphragm is minimal.

The construction of freeform ESL is simple and requires components that are easily available in a typical workshop setting. These components are selected on the basis of applications they were used in; for example, lightweight and paper-like Aluminized Mylar sheet was used in balloon and a solid metallic mesh was used in the wall. Other materials such as copper plates, conductive ink and transparent ITO could also be utilized.

Figure 12. The freeform electrostatic loudspeakers are flexible and used in a soft cushion.

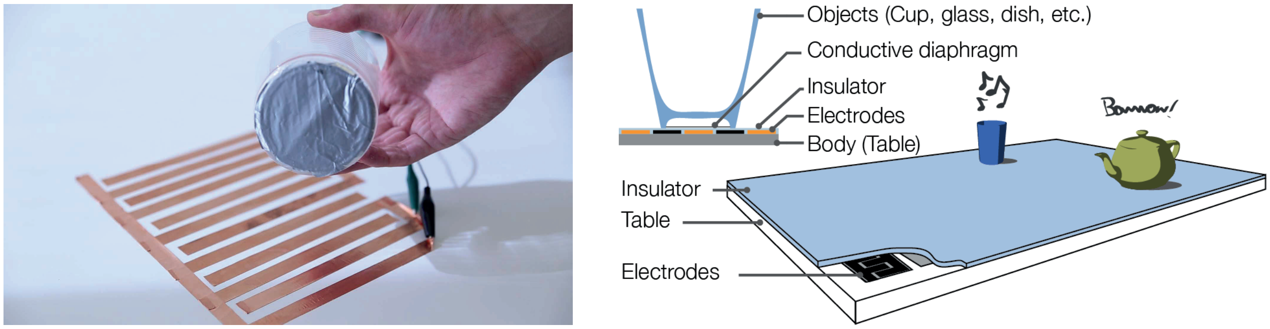

Figure 13. Tangible and everyday objects become interactive speakers by placing them on an electrode grid.

Another advantage of the freeform ESL is that it can be configured to many possible configurations with slight modification in the structure. Because of the electrostatic coupling generated between the two conductive surfaces and the relative motion between these surfaces produces sound, any one of the surface could be used as a diaphragm and any surface could be grounded while the other is connected to the high voltage. The quality of sound does not vary with changing the polarity of the speaker.

One important feature of the freeform ESL is that the entire diaphragm surface vibrates and reproduces sound. This is different than the conventional electromagnetic speakers, in which the coils around the magnet vibrates and the motion is translated to the diaphragm. This provides two advantages in favor of ESL.

1) The loading on the loudspeaker diaphragm only hinders the movement of the surface directly under the loading. Other parts of the speaker will still vibrate and generate sound. Unlike conventional speaker where the loading affects the entire sound generation process, the sound quality in electrostatic loudspeakers is slightly depreciated with the loading.
2) The sound emits uniformly and a user walking along a wall equipped with an ESL would feel that the sound source is moving along him/her. This is not possible with conventional speakers as the perceived sound intensity decreases as the user move away from the speaker and vice versa. Therefore, electrostatic loudspeakers also enhance the sound listening experience.

Safety Issues

In practice, we have used a low-power, high-voltage direct current power supply to amplify the sound signal up to 1000 V. The output current, however, is limited to 1.25 mA. Therefore, when a user touches the high voltage diaphragm only, they do not feel an electrical shock. But if the user touches both the high voltage diaphragm and the ground electrode, it can be painful and dangerous.

One way to prevent direct user contact with the high voltage source is to isolate the electrodes by placing them inside the objects and structures that are not accessible during interactions. In all configurations proposed in this paper, (except for tangible passive loudspeakers) we have placed the high voltage electrode inside the structure and covered it by the ground electrode. Moreover, these electrodes were insulated to reduce any mishap due to accidental contact between the skin and electrode. The resulting ESL are safe but care must be taken during fabrication, testing and installation of ESLs.

In future we will add isolation breakdown detection circuits for ensuring that these speakers are more safe and interactive.

Limitations

Current freeform ESLs are handmade and there are variations in between the speakers. These variations are mainly due to not maintaining uniform tension along the surfaces and maintaining uniform distance between the two electrode surfaces. This results in poor sound quality and inconsistent sound production throughout the surface. However, these procedural variations can be reduced in a calibrated production process.

The freeform ESLs could produce sound levels up to 80 dB. Replacing the high voltage source to higher voltage can produce higher sound levels. However, this will also increase the safety concern and increase the thickness of the insulation protective layer, and ultimately the size and weight of each speaker.

The inherent nature of the freeform ESL is that it distorts low frequency content of the incoming sound. Therefore, low frequency sound should be filtered out in the pre-processing stage to eliminate distortions and artifacts. The ESL can be combined with a low frequency subwoofer to play the entire spectrum of sound signal. It should be noted that high frequency sound is more important in sound localization, therefore augmenting the low frequency subwoofer would only complement the sound perception and the main interaction could be still done with the ESL.

The ESL structure can be used as a microphone by replacing the high-voltage driving amplifier with a pre-conditioning amplifier. In addition to the sound input, the ESL microphone could also be used to detect locations and interactions of users in front of the ESL. Moreover, ESL can also be used for haptic feedback by optimizing the low-frequency (< 500 Hz) response. These functionalities of ESL require further exploration and will be investigated in future.

In spite of limitations in the current freeform ESL, we have integrated them in to a large variety of user interaction scenarios. We believe that these loudspeakers will expand in to new modality for programmable matter and interactive surfaces.

REFERENCES

1. http://www.acoustable.com
2. Bau, O., Poupyrev, I., Israr, A., & Harrison, C. (2010). TeslaTouch: electrovibration for touch surfaces. In Proc. UIST'10 (pp. 283-292). ACM.
3. Benko, H., Wilson, A. D., & Balakrishnan, R. (2008). Sphere: multi-touch interactions on a spherical display. In Proc. UIST'08 (pp. 77-86). ACM.
4. Chua, C. K., Leong, K. F., & Lim, C. S. (2010). *Rapid prototyping: principles and applications*. World Scientific.
5. Coelho, M., Poupyrev, I., Sadi, S., Vertegaal, R., Berzowska, J., Buechley, L., Maes, P., & Oxman, N. (2009). Programming reality: from transitive materials to organic user interfaces. In CHI'09 Extended Abstracts (pp. 4759-4762). ACM.
6. Follmer, S., Leithinger, D., Olwal, A., Hogge, A., & Ishii, H. (2013). inFORM: dynamic physical affordances and constraints through shape and object actuation. In Proc. UIST'13 (pp. 417-426). ACM
7. Goldstein, S. C., Campbell, J. D., & Mowry, T. C. (2005). Programmable matter. *Computer*, *38*(6), 99-101.
8. Han, J. Y. (2005). Low-cost multi-touch sensing through frustrated total internal reflection. In Proc. UIST'05 (pp. 115-118). ACM.
9. Hanna, C.R. Theory of the electrostatic loud speaker. The Journal of the Acoustical Society of America 2(2) (1930), 143–149.
10. Harrison, C., & Hudson, S. E. (2008). Scratch input: creating large, inexpensive, unpowered and mobile finger input surfaces. In Proc. UIST'08 (pp. 205-208). ACM.
11. Holman, D., & Vertegaal, R. (2008). Organic user interfaces: designing computers in any way, shape, or form. *Communications of the ACM*, *51*(6), 48-55.
12. IOSONO GmbH, www.iosono-sound.com.
13. Ishiguro, Y., & Poupyrev, I. (2014). 3D printed interactive speakers. In Proc. of ACM conference on Human factors in computing systems (pp. 1733-1742)..
14. Ishii, H., Wisneski, C., Orbanes, J., Chun, B., & Paradiso, J. (1999). PingPongPlus: design of an athletic-tangible interface for computer-supported cooperative play. In Proc. of the SIGCHI conference on Human Factors in Computing Systems (pp. 394-401). ACM.
15. Kim, S. C., Israr, A., & Poupyrev, I. (2013). Tactile rendering of 3D features on touch surfaces. In Proc. UIST'13 (pp. 531-538). ACM.
16. Kurihara, K., & Tsukada, K. (2012). SpeechJammer: A System Utilizing Artificial Speech Disturbance with Delayed Auditory Feedback. arXiv preprint arXiv:1202.6106.
17. Lee, C. S., Kim, J. Y., Lee, D. E., Joo, J., Wagh, B. G., Han, S., ... & Koh, S. K. (2003). Flexible and transparent organic film speaker by using highly conducting PEDOT/PSS as electrode. Synthetic metals, 139(2), 457-461.
18. Lloyd, B., Goldman, R.B., and Roop, R.W. Design and Performance of a High-Frequency Electrostatic Speaker. The Journal of the Acoustical Society of America 27(6) (1955), 1128–1133.
19. Malone, E. and Lipson, H. Freeform fabrication of ionomeric polymer-metal composite actuators. Rapid Prototyping Journal 12, 5 (2006), 244–253.
20. Malone, E. and Lipson, H. Multi-material freeform fabrication of active systems (2008). ESDA’08, ASME, pp. 1–9.
21. Medley, A. P., Billson, D. R., Hutchins, D. A., & Davis, L. A. J. (2006). Properties of an electrostatic transducer. The Journal of the Acoustical Society of America, 120(5), 2658-2667.
22. Palette Kits, http://www.palettegear.com/
23. Paradiso, J. A. (1996). The interactive balloon: Sensing, actuation and behavior in a common object. IBM Systems Journal, 35(3.4), 473-487.
24. Poupyrev, I., Schoessler, P., Loh, J., & Sato, M. (2012). Botanicus Interacticus: interactive plants technology. In ACM SIGGRAPH 2012 Emerging Technologies (p. 4). ACM.
25. Rekimoto, J., & Saitoh, M. (1999). Augmented surfaces: a spatially continuous work space for hybrid computing environments. In Proc. CHI' 99 (pp. 378-385). ACM.
26. Rekimoto, J. (2012). Squama: modular visibility control of walls and windows for programmable physical architectures. In *Proc. of the International Working Conference on Advanced Visual Interfaces* (pp. 168-171). ACM.
27. Scott, P. J., & Jiang, X. (2014). Freeform surface characterisation: theory and practice. In Journal of Physics: Conference Series (Vol. 483, No. 1, p. 012005). IOP Publishing.
28. Sugibayashi, Y., Kurimoto, S., Ikefuji, D., Morise, M., & Nishiura, T. (2012). Three-dimensional acoustic sound field reproduction based on hybrid combination of multiple parametric loudspeakers and electrodynamic subwoofer. Applied Acoustics, 73(12), 1282-1288.