

VARI-SOUND: A Varifocal Lens for Sound

Gianluca Memoli

INTERACT Lab, University of Sussex
Falmer, Brighton, United Kingdom
g.memoli@sussex.ac.uk

Mihai Caleap

University of Bristol
Bristol, United Kingdom
mihai.caleap@bristol.ac.uk

Letizia Chisari

INTERACT Lab, University of Sussex
Falmer, Brighton, United Kingdom
chisariletizia@gmail.com

Bruce W. Drinkwater

University of Bristol
Bristol, United Kingdom
b.drinkwater@bristol.ac.uk

Jonathan P. Eccles

INTERACT Lab, University of Sussex
Falmer, Brighton, United Kingdom
jeccles@sussex.ac.uk

Sriram Subramanian

INTERACT Lab, University of Sussex
Falmer, Brighton, United Kingdom
sriram@sussex.ac.uk

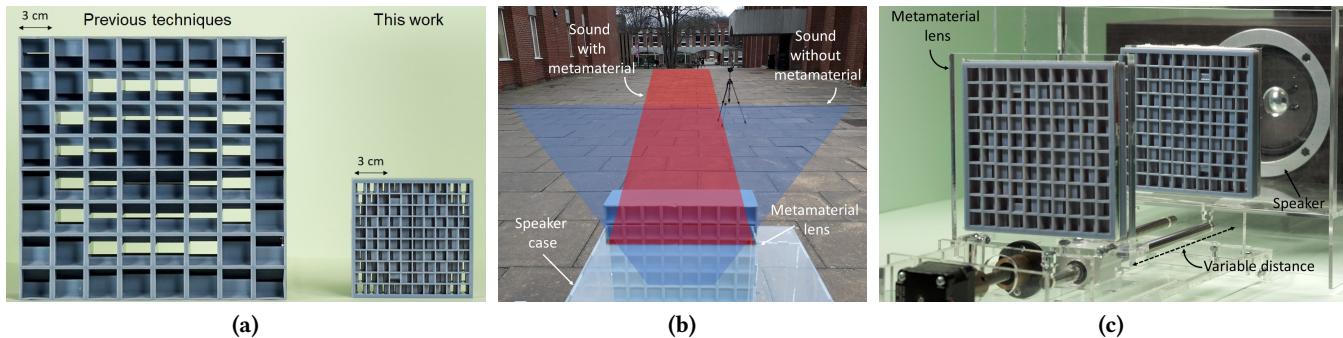


Figure 1: VARI-SOUND is a method to design the equivalent of optical devices for sound, based on acoustic metamaterials. Our contributions include: (a) how to design compact acoustic lenses; (b) how to transform a standard speaker into a directional one; (c) how to build and adjustable focus acoustic lens i.e. the equivalent of a zooming objective for light.

ABSTRACT

Centuries of development in optics have given us passive devices (i.e. lenses, mirrors and filters) to enrich audience immersivity with light effects, but there is nothing similar for sound. Beam-forming in concert halls and outdoor gigs still requires a large number of speakers, while headphones are still the state-of-the-art for personalized audio immersivity in VR. In this work, we show how 3D printed acoustic metasurfaces, assembled into the equivalent of optical systems, may offer a different solution. We demonstrate how to build them and how to use simple design tools, like the thin-lens equation, also for sound. We present some key

acoustic devices, like a “collimator”, to transform a standard computer speaker into an acoustic “spotlight”; and a “magnifying glass”, to create sound sources coming from distinct locations than the speaker itself. Finally, we demonstrate an acoustic varifocal lens, discussing applications equivalent to auto-focus cameras and VR headsets and the limitations of the technology.

CCS CONCEPTS

- Human-centered computing → Sound-based input / output; Laboratory experiments;
- Hardware → Emerging technologies; Emerging tools and methodologies; Emerging interfaces;
- Applied computing → Physics.

KEYWORDS

Metamaterials, Microstructures, Fabrication, Spatial Audio, 3D Printing.

ACM Reference Format:

Gianluca Memoli, Letizia Chisari, Jonathan P. Eccles, Mihai Caleap, Bruce W. Drinkwater, and Sriram Subramanian. 2019. VARI-SOUND: A Varifocal Lens for Sound. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, May 4–9, 2019, Glasgow, Scotland Uk. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3290605.3300713>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI 2019, May 4–9, 2019, Glasgow, Scotland Uk

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-5970-2/19/05...\$15.00

<https://doi.org/10.1145/3290605.3300713>

1 INTRODUCTION

The shaping of acoustic fields is a major technological advance which underpins high-fidelity sound reproduction [60], tactile feedback in consumer electronic devices [6], particle manipulation [10], non-destructive testing [24] and non-invasive therapies (e.g. for essential tremor [43] and cancer [65]). Traditionally, this is achieved by controlling the intensity or the phase of the generating source through phased arrays [36, 57, 58]. Phased arrays offer real-time control of sound, but are often bulky and expensive, with cost and complexity scaling with the number of channels [37]. Despite these limitations, phased arrays are in widespread use. In the human-computer interaction (HCI) community, in particular, they have paved the path towards applications like acoustic levitation [23, 33, 39] and mid-air haptics [6].

The way we manage light, however, is different: in theatres, videogames and virtual reality (VR) audience's immersivity can easily be augmented by using passive devices, like lenses and filters. When the properties of lenses and filters can be changed on demand, we get devices like autofocus cameras, liquid crystal displays and VR headsets. This is still not possible for sound: arrays of speakers are used for applications like surround sound because acoustic lenses [23] are bulky, with a physical size far thicker than the wavelength (i.e. which is ~ 35 cm at 1 kHz, the peak frequency for human hearing perception). This technological difference could explain in part the prevalence of visual (optical) technologies over acoustic ones in our modern, space-hungry world.

Acoustic metamaterials may offer a way forward [11, 31]. These are normal materials (i.e. glass, wood, 3D printer plastic), but engineered to control, direct, and manipulate waves in uncommon ways. However, their use in HCI is currently impeded by three key limitations: acoustic metamaterial devices are thick (e.g. one wavelength in [27, 28, 35, 41, 59, 66]), static and often operate over a limited frequency range (e.g. 10% of the central frequency in [35]).

In this work, we propose a method to design metasurfaces that behave like converging lenses for sound. We address two of the limitations mentioned above, showing how these lenses can be fabricated to be as thin as 1/3rd of their wavelength of operation and how they can be combined in user-controllable devices that can be mechanically adjusted.

In particular, having demonstrated that some design tools commonly used in optics are also valid in acoustics when using metasurfaces, we build the acoustic equivalent of some key optical devices (i.e., a collimator, a magnifying glass, a telescope and a vari-focal lens) and test them at 5-6 kHz, where each lens is 2 cm thick. We discuss real-life applications based on these devices, how to overcome their remaining limitations (e.g. bandwidth) and the potential ways to use them in innovative, sound-based interfaces. We want

to share with the HCI community how we believe acoustic metamaterials may revolutionize the way we think, design and experience sound, hoping to excite others into using them.

Our contributions.

- We show how to design metamaterial lenses for sound that are sufficiently small to be of practical use for the HCI community (Figure 1a).
- Coupling our acoustic lenses with a generic speaker, we demonstrate and test devices equivalent to optical ones i.e. magnifying glasses, telescopes, lighthouses...for sound (Figure 1b).
- We design and test the first dynamic metamaterial device: the equivalent of a zoom objective - i.e. a vari-focal lens, but for sound (Figure 1c).

2 RELATED WORKS

Designing for sound

The design of rooms and public spaces has a binary approach to acoustics: in a cinema or a concert hall there is either the same sound for everyone or none. In practice, current sound design does not have a personal component, without headphones. Light, however, is managed differently: a theatre director or an architect can populate a scene with focused or diffused light, a spotlight that follows a character, alternation of light and shadows. Light engineers can create experiences targeted to different users with light, either by shaping light beams through passive devices (e.g. lenses, holographic filters) or by direct beamforming (e.g. spatial light modulators).

Acoustics, however, is catching up: the quests for 3D spatial audio (which creates the illusion of localised sources for the audience) and localised audio (which uses highly directional sources to deliver different sounds to parts of the audience) are moving our everyday lives towards localised, high quality audio messages beyond headphones. A world like the one imagined by Philip K. Dick in "Minority Report" (later a movie directed by Steven Spielberg), where the main character received personalised, directional audio adverts while walking in the street, does not seem so far away.

Spatial sound

Sound is all around us, with no direction having precedence over the others, but acoustic cues have a profound directional nature. This is evident, for instance, with the sound of a car approaching from behind. Digital content, conversely, is broadcasted from sources (either headphones or loudspeakers) which have a fixed position relative to the audience.

"Spatial sound" is a method of filling this need-gap [49]: playing with piloted delays and intensity differences in the real sources, this technique gives to a specific audience the illusion that sounds come from different locations around.

Using spatial sound techniques, creative designers can now populate scenes with numerous virtual sources, each with its own location and attributes, relying on their position (relative to the audience) for narrative (e.g. the movie “Pearl” [12]) and immersive purposes [63]. They can design soundscapes using object-oriented toolkits, commonly found in packages like Unity 3D (e.g. Season Traveller [47]) or Unreal.

The first method of delivering spatial sound is through headphones. The main advantage here is that the sound can be adjusted to move with the listener, which is ideal for sound-based navigation systems (e.g. gpsTunes [54] and SWAN [15]) and one-to-one interactions with learners (e.g. AudioChile [50]) or visually-impaired listeners [55]. The user, however, needs to be completely removed from the real space, with interactions (e.g. comments among co-workers attending the same VR meeting) only possible in the virtual world.

The second method, based on deploying arrays of speakers where needed, is at the basis of the surround sound systems in cinemas and home theatres (e.g. Dolby Digital, Sony SDDS). These applications marry with the room in creating the audio experience, which can affect multiple users, but the “surround” effect is restricted to a narrow listening area (i.e. a “sweet spot”) and sometimes does not provide for up-down cues [1]. This is not sufficient for audio-visual VR [21]. A way forward is offered by Wave Field Synthesis [2], which aims at local reproduction of “virtual sources” near the listener using speakers in a linear array (e.g. modern soundbars). This method leads to better immersivity (as it reacts to where the listener is), but the finite distance between the speakers and their physical size introduce effects of spatial aliasing.

Directional sound

An alternative towards personalised multi-user audio experiences, which minimises spatial aliasing, comes from directional speakers [3]. Directional speakers have been used in HCI for multi-sensorial art displays (e.g. Tate Sensorium [38]) or to send audio messages to a specific individual or region of space (e.g. Holographic Whisper [40] or Project Telepathy [5]). The most common type (e.g. Acouspade, SoundLazer) exploits an array of ultrasonic transducers to produce a highly directional carrier wave, which is then modulated with audible signals [13]. Other speakers (i.e. Holosonics) produce the ultrasonic beam through a vibrating plate. All these speakers exploit the nonlinear effects of air to produce audible sounds [45] and can be used as audio spotlights [62], but steering the beam requires physically moving the speaker.

Audio spotlights are not low-cost technology (costs are typically around 2,000 USD for a decent-sized system) and come with a main limitation i.e. the bandwidth of the sound they can transmit. This is typically ~10–20% of the driving ultrasonic frequency of operation – e.g. 4 kHz for a parametric speaker operating at 40 kHz – and typically lacks the

low-frequency components: it is therefore sufficient to pass public announcements and badly reproduced music [22], but not high-quality sounds. Furthermore, as most applications involving phase control of multiple speakers, audio spotlights are prone to thermal losses and therefore have a poor response to high sound volume.

Acoustic lenses

Acoustic lenses appeared for the first time in 1930s, as a by-product of the Bell labs. They used to be inserted in loudspeakers to spread their high-frequency directional emission, but gradually fell out of favor through the 1970s and 1980s. Lens designs at the time were in fact fragile and cumbersome i.e. an accident-waiting-to-happen for the music industry, always on the move. Today, however, the technology is re-emerging: acoustic lenses can be found to help beam-shaping in ultrasonic transducers and in some high-end home audio systems (e.g. Bang & Olufsen). These lenses, however, are still designed to be much larger than the wavelength (i.e. 1 metre at 340 Hz): their use is therefore confined to the higher part of the acoustic spectrum. More compact solutions, usable at lower frequencies, are highly desirable.

Acoustic metamaterials

Acoustic metamaterials [11, 31] are normal materials (i.e. made of glass, wood, paper, plastic, LEGO bricks), but with an internal structure engineered to control, direct, and manipulate waves in uncommon ways [16, 17, 35, 41]. They are made of a collection of sub-wavelength structures (i.e. “unit cells”), each capable of locally manipulating the impinging wave in terms of its phase and/or intensity [11, 35].

There are many unit cell designs in the literature, either based on labyrinthine structures [59], helical structures [66], space-coiling [27, 28], multi-slits [34], or Helmholtz resonators. Selecting the right one depends on design constraints, like the available space to deploy them, the manufacturing techniques and the desired frequency response. This is because most unit cells typically operate over a small bandwidth¹ $\Delta f_{2\text{dB}}$ on both sides of their design frequency f_0 (with $\Delta f_{2\text{dB}} \leq 0.2f_0$). In addition, unit cells’ dimensions increase as f_0 gets lower, with thicknesses² as large as ~ 1 m at 340 Hz. Of these two limitations, we believe size to be the most crucial for the HCI community.

Narrow-band devices, in fact, are sufficient for many applications. In audio reproduction, for instance, it is normal to cascade speakers optimised for different frequency ranges to cover the audible spectrum, even if well known songs

¹The bandwidth is here defined as the frequency range $\Delta f_{2\text{dB}}$ over which the impinging sound loses no more than 2 dB going through the metamaterial (i.e. transmission $\geq 80\%$).

²Here and in the following, “thickness” is the dimension of the unit cell along the direction of propagation of the sound.

like “*Nessun dorma*” often span little more than 1 octave³. In buildings and factories, with human hearing perception peaked at 1-2 kHz, controlling the delivery of one octave may be sufficient to quench annoyance due to an air conditioning unit or to deliver personalised alarm messages. Furthermore, the literature already presents acoustic metamaterials with larger bandwidths [20, 26, 29, 52], obtained combining unit cells of different sizes. Li et al. [26], for instance, used basic 3D-printed cuboids as unit cells (i.e. Acoustic Voxels), scaled and joined to create wind instruments spanning ~2 octaves. Additional solutions will be discussed in section 6.

Unit cells are then assembled into larger 3D structures, with a bottom-up “metamaterial design” concept recently used to create mechanical actuators (i.e. Metamaterial Devices [16, 17]) and user-controllable 3D-printed structures (e.g. Coded Skeleton [18]). Particularly interesting for beam-shaping applications is the case of *metasurfaces*: closely-packed structures of phase shifters whose thickness is smaller than or comparable to the wavelength of operation and can therefore be considered to be 2D [25, 56, 64].

Once a metamaterial device is designed, however, its function is fixed, while speaker arrays can change on demand. This is another key limitation for HCI applications, which has been addressed so far using hybrid systems (devices where phased arrays are used to manipulate shapes created using metasurfaces [37]) or motorised belts (to position metasurfaces in front of a directional speaker and steer its emission [19]). Both these solutions, however, rely on a phased array to begin with. Different dynamical metamaterials have been proposed (e.g. [7, 46]), with some solutions as large as a room [30], but none has the simplicity of a system of lenses. Here, we demonstrate a method to design compact acoustic devices, inspired by optics, that can be used with standard, low cost, computer-sized audio speakers.

3 METASURFACE DESIGN

There are four key steps in designing a metasurface: 1) choosing its function (i.e. what it does to the input field); 2) transforming this information into an analogic phase/ intensity distribution on the metasurface (Figure 2a); 3) selecting the unit cells to use; 4) fabricate the metasurface, taking into account constraints in terms of space and frequency response (Figure 2b). In this section, we show how to design acoustic metasurfaces to be used in transmission. Some of our more general considerations, however, apply also to reflecting metasurfaces and other types of waves [8, 9, 14].

³To visualise the frequency response $\Delta f_{2\text{dB}}$ of a metamaterial, the reader should keep in mind the 12 keys on a piano keyboard that form an octave. For reference, a typical piano has 7 octaves and a minor third, from $A_0 = 27.5$ Hz to $C_8 = 4186$ Hz, while most human voices span a maximum of 3 octaves.

Steps 1 & 2: From the desired field to a phase distribution

Assigning a function to a metasurface means deciding how the distribution of the acoustic pressure will look like after passing through it, both in terms of geometry and of intensity distribution. Li et al. [28] suggest to address this step as a problem of acoustic ray tracing [51]: the desired far-away field gets back-propagated to the metasurface itself, thus giving a required phase/intensity distribution at its exit. This will be the phase/intensity distribution to be encoded by the unit cells on the input wave. More details on this step for a generic field have been given elsewhere [37], so in this work we will only discuss the case of a converging lens.

A converging lens is characterised by two quantities: its focal length and its physical extension (i.e. how many unit cells it contains). Once the desired focal length f is set along the axis of the lens (\hat{z}), the phase distribution $\varphi(x, y)$ on the metasurface (assumed to be in the $z = 0$ plane) is obtained by imposing that all the contributions from the unit cells arrive in phase at $(0, 0, f)$. In this work, we choose to design our lenses using a parabolic phase profile [42]:

$$\varphi(r) = \varphi_0 - A^2(x^2 + y^2) \quad (1)$$

where $\varphi(x, y)$ is local phase (assigned to a unit cell), A is a constant (related to the local curvature of the phase profile), λ_0 is the design wavelength and φ_0 is an arbitrary constant. In optics, this choice leads to more compact lenses (i.e. GRIN lenses [61]) and, as shown in Figures 2a and 3, allows the parameter A to be easily related to the “curvature” of the lens: a larger value of A corresponds to a more focusing lens.

Step 3: Unit cell selection

Once $\varphi(x, y)$ is known, the designer has to choose the unit cells to implement it. As mentioned above, many designs are available in the literature, but all of them have one point in common: the smaller the frequency, the larger is the cell.

The same applies to the 16 unit cell designs proposed by Memoli et al. [35]: rectangular cuboids, $\sim 4.3 \times 4.3 \times 8.6$ mm in size, designed to have maximum transmission (~97% of the input sound) at $f_0 \pm \Delta f_{2\text{dB}} = 40 \pm 1$ kHz. A simple way to use these designs at a different frequency \bar{f} is to scale each cuboid until its thickness is equal to the new wavelength $\bar{\lambda} = c_0/\bar{f}$ (where $c_0 \sim 343$ m/s is the speed of sound in air). At the new frequency, each of the 16 scaled unit cells will encode the same phase delay (between 0 and 2π) and have the same transmission and bandwidth they had at f_0 . With this method, Jackowski-Ashley et al. [19] manipulated the emission of a parametric speaker operating at 64 kHz.

The cell design in [35], however, offers a possibility not fully explored by the authors: a cuboid designed for f_0 also has the same transmission at other frequencies (see Figure

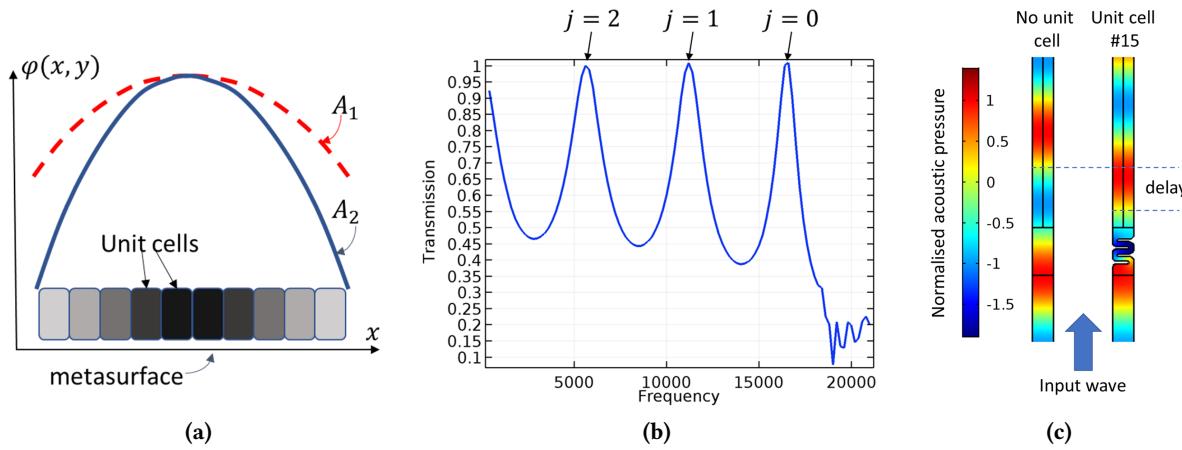


Figure 2: Key moments of the design process used in this work: (a) comparison of two different phase profiles, with $A_1 < A_2$, highlighting how the unit cells encode the phase distribution $\varphi(x, y)$ (each color corresponds to a different phase); (b) COMSOL simulation for the transmission of unit cell #15 from [35], scaled so that its base is 10.4 mm, highlighting the resonances described by equation (2); (c) working principle of a Type B unit cell, highlighting the delay the input wave experiences in its presence (COMSOL simulation at 5.6 kHz, base size 10.4 mm); .

2b). These frequencies will be given by:

$$f_j = f_0 - j \cdot c_0 / L_{eff} \quad (2)$$

where $j = 0, 1, 2 \dots N$ are integers, L_{eff} is a design parameter of the specific unit cell⁴ and $N = \text{round}(L_{eff}/\lambda_0)$ is the (integer) number of times L_{eff} contains the wavelength. It is therefore possible to operate the unit cells at one of the frequencies f_j (see Figure 2c), maintaining a similar transmission to the one in f_0 , but considering that the phase encoded at f_j is different from the one at f_0 : a look-up table is necessary. In particular, our simulations show that the maximum phase these sub-resonant structures can achieve is $2\pi/2^j$.

Step 4: Lens fabrication

In this work, we operate at $f_0 = 5,600$ Hz (i.e. a frequency close to F8, with a wavelength of ~ 6 cm). We selected this frequency due to restrictions in our manufacturing capabilities (i.e. it was the largest size that could be made on our 3D-printer) but, since everything can be scaled, this choice does not limit the conclusions. We use two type of lenses:

Type A lenses. Obtained by scaling the unit cells from Memoli et al. [35], so that their first resonance ($j = 0$) is 5.6 kHz and the thickness is equivalent to λ_0 (i.e. 60 mm). Each of this lenses is made by a 8×8 array of unit cells and is $240 \times 240 \times 60$ mm in size (see Figure 1a, left). Their bandwidth is the same as in [35] i.e. $2 \cdot \Delta f_{2dB} \sim 0.05 \cdot f_0$ or 2 piano keys.

⁴Related to the length of the maze-like structure, L_{eff} sets by how much the sound is delayed going through the cell [31]

Type B lenses. Obtained by scaling the unit cells from Memoli et al. [35], so that their second resonance - i.e. f_2 in equation (2) - is equal to 5,600 Hz. Each of this lenses is made by 10×10 array of unit cells and is 104×104 mm in size and 20.8 mm (i.e. $\sim \lambda_0/3$) in thickness (see Figure 1a, right).

Type 2 lenses also have a larger bandwidth: our COMSOL simulations (see e.g. Figure 2b) show that the width of the peak with $j = 2$ is larger than the one for $j = 0$: by definition, $2 \cdot \Delta f_{2dB} \sim 0.28 \cdot f_0$ i.e. 5 piano keys. The major disadvantage with Type B lenses is that, since the 16 cell designs now only span a limited part of the phase space, only a limited number of focal lengths can be realised with a lens of a fixed size. Simulations in Figure 3 show, for instance, that in the case of a 10×10 Type B lens with radial phase profile from equation (1), the maximum focal length is 57 mm. To achieve larger focal lengths, it is necessary either to extend the lens (e.g. to 12×12 unit cells) or to use the techniques in section 5.

4 BASIC TOOLS FOR DEVICE DESIGN

In this section, we find that the basic design tools available in optics are also valid in acoustics, when metamaterial lenses are involved. This discovery simplifies the realisation of metamaterial based devices (section 5) and leads to solving some of the limitations of the metamaterials described so far.

The thin lens approximation

Optical devices typically require a lens to be placed between the source and the receiver, until the former is imaged on the latter. For most applications, mutual distances are calculated

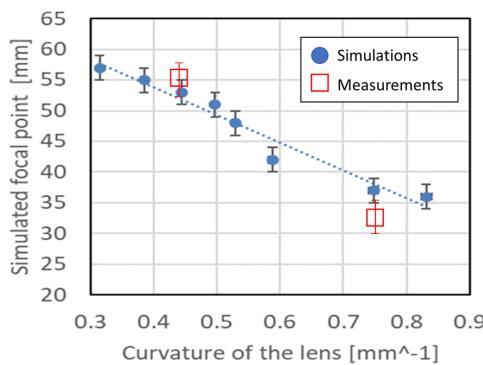


Figure 3: Simulated focal lengths for a 10×10 unit cells, 104 mm Type B lens used at 5.6 kHz ($j = 2$), as a function of the lens curvature A . Focal lengths are reported as a distance from the center of the lens. For each value of the curvature, COMSOL simulations were used to predict the resulting field with an input plane wave at normal incidence. The graph also reports the measured focal lengths of the two Type B lenses we realised for this study (see below).

using the thin-lens equation [4]:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} \quad (3)$$

where f is the focal length of the lens, p is the distance between the source and the lens, q the distance between the lens and the image of the source. Equation (3) is based on the same hypotheses used to design the metasurface and should directly apply when the thickness is much smaller than the wavelength [4], but its validity has not been tested before.

To do this, we run the following experiment: (a) we design a lens and simulate (COMSOL) its focal length (Figure 3); (b) we mount the set-up in Figure 4a measure the acoustic sound pressure (SPL) in 1/3rd octaves at different distances from a speaker emitting a tone at 5.6 kHz, without the lens; (c) we repeat the measurements, after inserting the lens in the acoustic path, at a distance p from the speaker; (d) we look for the “image” of the speaker, defined as the position where the intensity changes the most due to the lens and record the distance as q ; (e) we plot $1/q$ vs. $1/p$ (see Figure 4b) and fit with a line of angular coefficient -1 from equation (3) to find the focal length f .

As shown in Figure 4a (inset), the selected type B lens was held in a mask during the tests, at different distances from the speaker. Measurements at large distances (Figure 4a) were conducted in a non-reverberant environment outdoors, using a Norsonic 121 class I sound-level meter and a $\frac{1}{4}$ " microphone (Norsonic, Nor-1225). More detailed 2D scans of the area in front of metasurfaces (120×300 mm) were taken using a modified 3D-printer, a $1/8$ " microphone

(B&K, model 4138) and an amplifier (B&K, Nexus conditioning amplifier) directly connected to a Picoscope and then to the PC. The 5.6 kHz signal originated from a .wav file prepared in-house (using Audacity) and was amplified using a car stereo amplifier.

Figure 4b reports the results for two acoustic metasurfaces, designed to have different focal lengths ($A_1 = 0.44 \text{ mm}^{-1}$ for $f_1 = 53 \text{ mm}$ and $A_2 = 0.75 \text{ mm}^{-1}$ for $f_2 = 38 \text{ mm}$). We found a very good agreement between design and experiments for $f_{sim}=53 \text{ mm}$, but 16% difference for the smaller focal length. In the optical case, the accuracy decreases when the focal length gets smaller, as lenses cannot be considered “thin”: we will therefore limit this study to $f > 53 \text{ mm}$.

Having verified the thin-lens equation leads naturally to designing systems of two lenses, like telescopes and microscopes. The focal length F of such a system is given by [4]:

$$1/F = 1/f_1 + 1/f_2 - D/(f_1 \cdot f_2) \quad (4)$$

where f_1 and f_2 are the focal lengths of the two lenses and D is the distance between the two lenses. This equation has been largely exploited to design zoom objectives and, by the HCI community, in designing the first models of VR headsets (see references in [53]). Once F is known, equation (3) can be applied, in the form $1/P + 1/Q = 1/F$ where all the distances (i.e. P , Q and F) are measured from two known reference planes, called “principal planes” [4].

To test the validity of equation (4) in the acoustic case, using metasurfaces, we simulated a system of two lenses (both with $f = 53 \text{ mm}$) using finite elements (COMSOL Multiphysics) and adjusted the mutual distance D , recording the position of the image relative to the last lens. Figure 5 shows that this quantity, known as the “Back Focal Length”, increases with the distance between the two lenses⁵ as predicted from equation (4). Experimentally, we realised the two lenses and positioned the first at 50 mm from the speaker. We then changed D , noting the position of the maximum intensity. Good agreement was found with equation (4).

5 DEVICE REALISATION

Here, we exploit the capabilities acquired in the previous sections to prototype three lens-based acoustic devices. For these, we describe conceptual application scenarios, to be evaluated in future works through user studies.

Acoustic collimator

The possibility of designing an acoustic lens with a selected focal length (and the validity of equation (3)) allowed us to build a collimator: a system that corrects the geometric divergence of a source, so that the output sound is spatially contained in a directional beam (Figure 1b). Collimators are

⁵Intuitively, when the lenses are further apart, the focal point is further away.

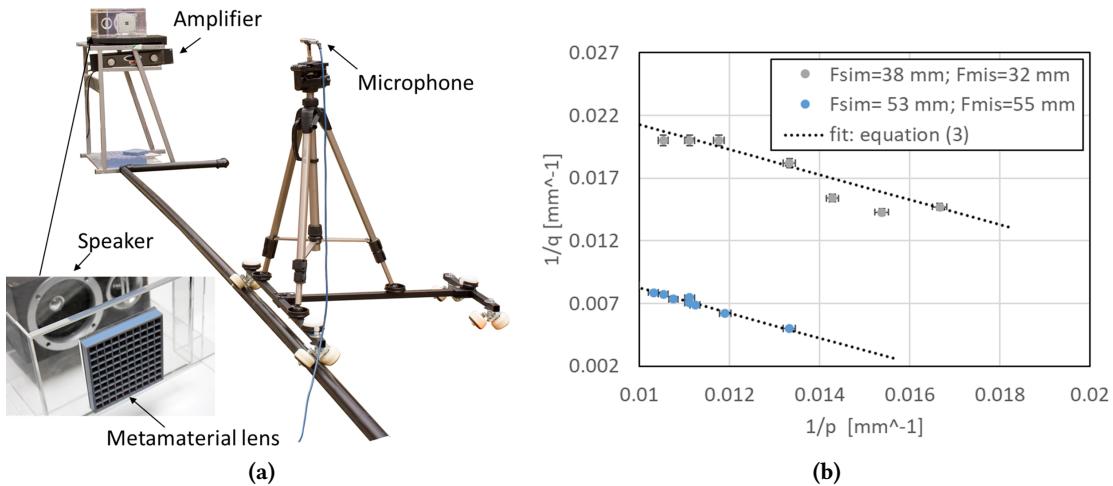


Figure 4: Measurements for testing the thin-lens equation: (a) outdoor set-up for positioning the microphone at fixed distances from the loudspeaker; (b) results for two different lenses, comparing f_{sim} (directly obtained from COMSOL simulations) with f_{mis} (obtained by fitting measurements with (3)). Reporting the results in terms of $1/p$ and $1/q$ allows to obtain the focal length from a linear fit (from equation (3)).

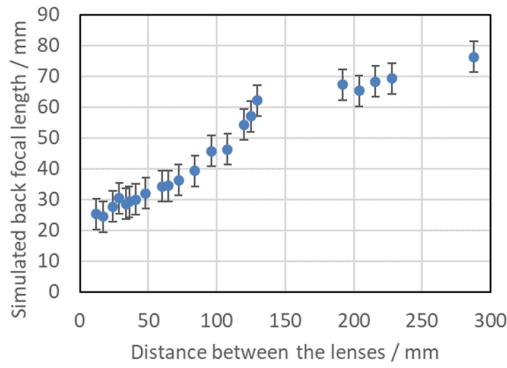


Figure 5: Results for the telescope configuration: dependence of the back focal length on the mutual distance between the lenses (COMSOL simulations for two Type B lenses ($f = 53 \text{ mm}$)).

used in optics after the lamp in a slide projector (i.e. to make rays parallel) or in lighthouses (i.e. to project the light over larger distances). They are also used in theatres for producing spotlights. In these applications, a converging lens is positioned at a distance from the source equal to its focal length, transforming the impinging wave into a parallel beam.

In our realisation, we used a Type A lens ($f = 150 \text{ mm}$) positioned at $150 \pm 2 \text{ mm}$ from the speaker. Figure 6a shows that, while the decrease with distance typical of a spherical source is maintained (i.e. -6 dB when doubling the distance), the

acoustic pressure measured at different distances is consistently larger than the one in absence of the lens. The angular emission, sampled at 4.24 m (Figure 6b), shows that the angle of divergence of the speaker (defined as the width 10 dB below the peak) was reduced from $60 \pm 1^\circ$ to $27 \pm 1^\circ$.

The metamaterial lens modified the emission of our low-cost speaker, making it comparable with the top-of-the-range audio spotlight by Holosonics™, which has an emission angle of $\sim 30^\circ$ [44, 48]. In our tests outdoors (Figure 1b), we also found that the sound, which could only be heard up to 10 m away without the lens, was perceived up to 40 m away when the collimator was present (test conducted with passer-bys, in a local green area). Theoretical predictions based on equation (3) suggest that the divergence angle could be further reduced by finer adjustments to the mutual position of the speaker and the lens (e.g. provided by an automated positioning system). As detailed in section 2, the HCI community has used expensive audio spotlights for many applications: our method hints to the same applications, in a more cost-effective way. In addition, as shown by Norasikin et al. [37], metamaterials can be used to give non-conventional shapes to sound, and this opens to additional usage scenarios, like:

Personalised experiences in shared spaces. These include sending specific sounds only to parts of an audience (see Figure 6c); having different acoustic cues depending on the location in a space (e.g. with large real-walking virtual reality experiences confined to limited tracking volumes [32], a few metamaterial corrected speakers may be able to map whole virtual mansions); creating different sound zones in cars or

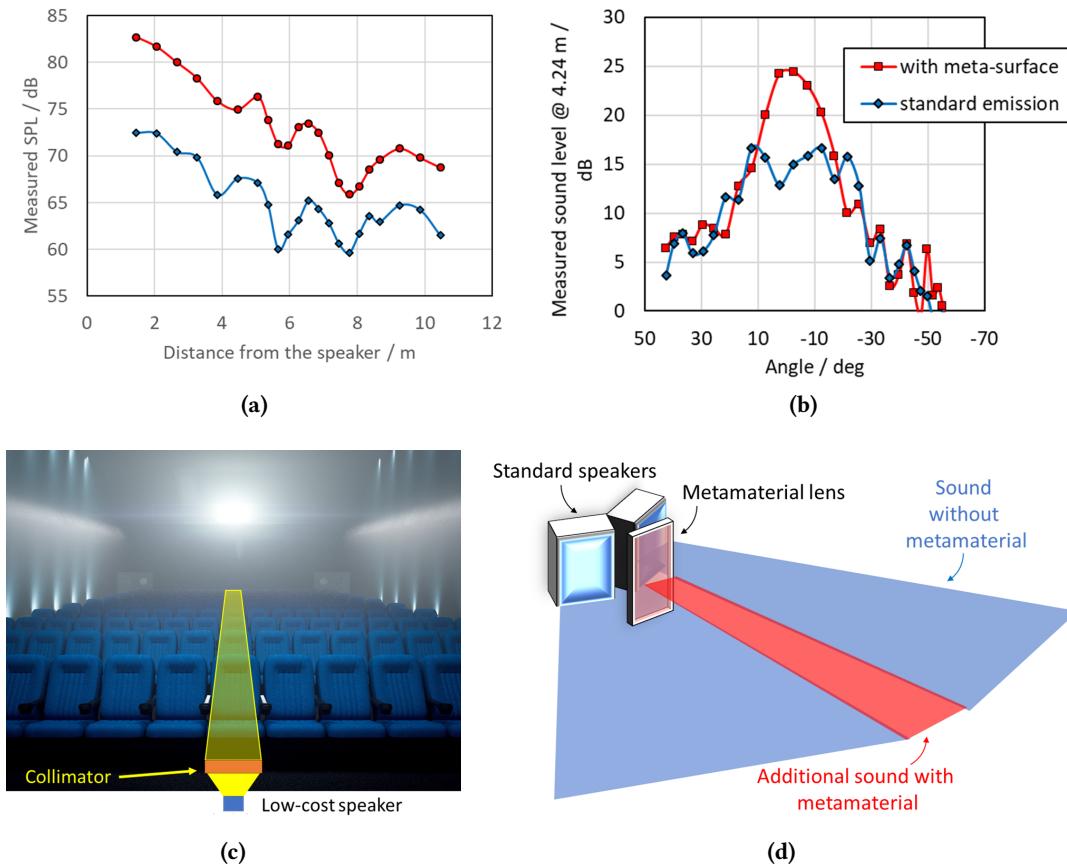


Figure 6: Effect of a collimator on the emission from a computer speaker, when a lens is positioned at the focal length from it: (a) intensity at different distances and (b) angular emission at 4.24m from the speaker. Also reported are two conceptual drawing detailing the use of a collimator (c) in an auditorium and (d) to correct the emission of two speakers mounted back to back, to cover holes in their emission.

on a sofa (so that passengers and driver can listen to different sounds); sending sound behind corners.

Increase the spatial performance of sound systems. At concerts, or in cinemas, speakers are arranged to minimize the spatial changes in level and tonality, but there is always a minority of the audience who does not have an optimal acoustic experience. Figure 6d reports the example of two speakers mounted in a symmetrical coupled point source arrangement, highlighting the gap in front of them: a dedicated speaker, made directional by a collimating lens, may be used to fill such gaps. Similar considerations can be applied to smart speakers, like Google Home or Amazon Echo, whose 360° emission is due to an array of speakers.

Modify the spatial sensitivity of acoustic sensors. Collimators can also be used in detection, transforming generic acoustic sensors into highly directional ones.

Acoustic magnifying glass

The art of glass blowing introduced the concept of magnifying glasses in the 13th century: an application where the distance between the observer and the sample to be observed is fixed, and the user positions the lens in between until the image is “in focus”. In our realisation (Figure 7a), we positioned the speaker and the microphone along a line and inserted a lens in between, adjusting its position until the signal on the microphone was maximised.

According to equation (3), for a lens of fixed focal length f there is only one optimal distance from the source \bar{p} where this happens. In this configuration, the lateral extension of the image (a.k.a. the magnification M) also depends on the value of \bar{p} : real images with $|M| > 1$ are formed when $f < \bar{p} < 2f$, while $|M| < 1$ when $\bar{p} > 2f$ (i.e. the magnification depends on the position of the lens: the farther is the lens, the smaller the image).

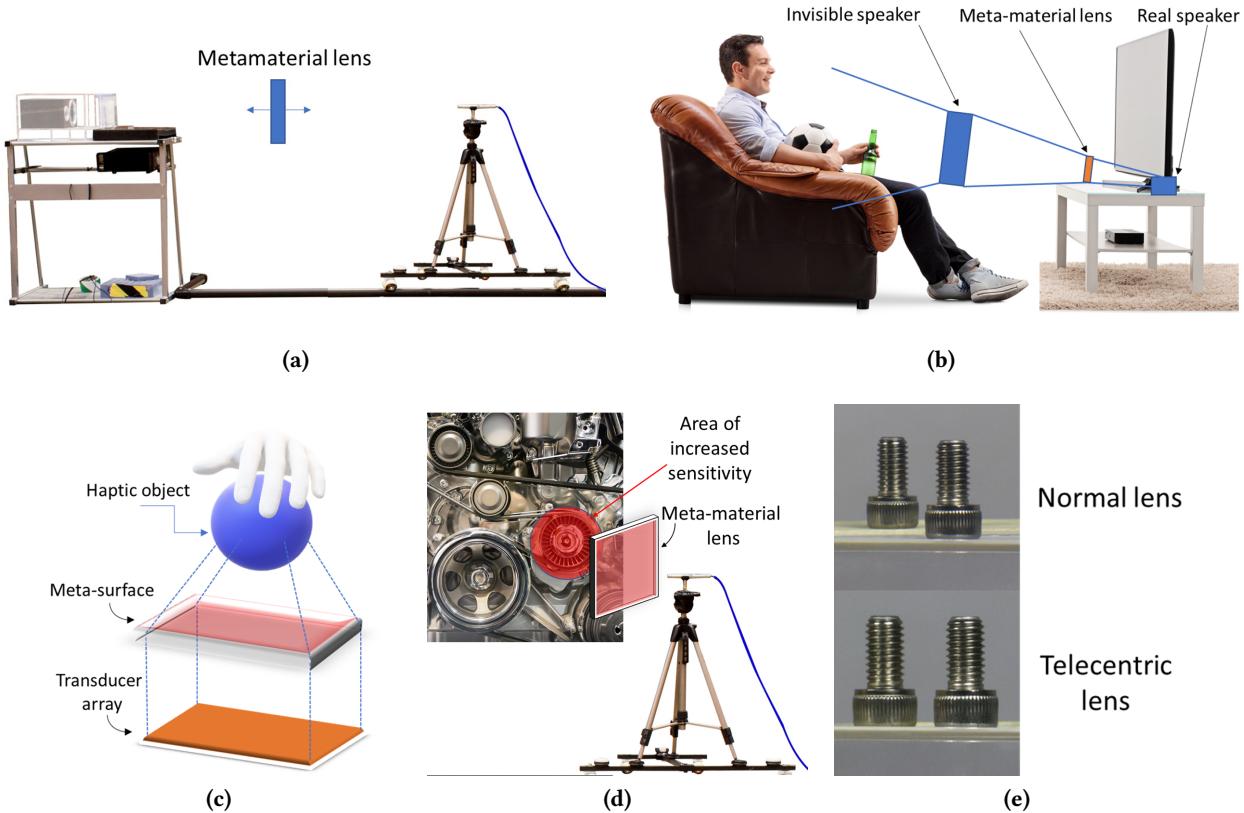


Figure 7: Applications of a single metamaterial lens: (a) our realisation of the device, using a low-cost speaker; (b) an acoustic lens creating acoustic objects for end-users; (c) an acoustic lens used to extend the range of mid-air haptics; (d) an acoustic lens used to increase sensitivity to certain parts of a machinery; (e) two equally-sized screws separated by 100 mm and observed two different types of optical lenses.

The possibility of creating the real image of an acoustic source and to vary its size according to the position of a lens (between the source and the listener) allows:

Modifying the apparent position of the source. One possibility would be to create the image of a speaker in front of the user (imagine him/her seating on a sofa, with the original speaker where the television is) and thus the feeling that the sound is coming from a localised source (see Figure 7b). This effect has some similarity to spatial sound (1D, so far), and should be evaluated by user studies in comparison to that.

Extending the range of haptic devices. The HCI community is familiar with using ultrasonic transducer arrays for mid-air haptics [6] and levitation [33], even behind objects [37]. These effects, however, lose in definition as the distance from the source increases. Using an appropriate lens may help moving these effects at larger distances, so that the source array may be located far from the end-user (see Figure 7c).

Similar methods have been used in holographic trapping [42] and there is no reason why they should not be applicable to 3D shapes made of sound.

Augmenting reception. A lens may also be used on the receiver side, to modify the spatial performance of a microphone. As shown in Figure 7d, using the lens to image certain areas instead of others may help detecting alarming noises from certain parts of a machine (e.g. the cog that tends to break all the time, or the hissing sound of a gas leak) or from selected areas of the house (e.g. a burglar breaking a window instead of a dog barking in the background). Similar arrangements may make personal assistants (e.g. Amazon Echo) sensitive only to orders issued in certain areas of the house, with advantages over background noise.

The change of magnification with distance may be an issue when the source is very small or very far from the detector. In addition, sources located at the periphery of the lens may appear distorted, when imaged. A solution to these

two issues, often used in optics (e.g. machine vision), is offered by telecentric systems [4]: an additional entrance pupil, positioned at the focal length of the lens, allows to select only the central rays and maintain the image size with object displacement, provided the object stays within the “depth of field”. In the acoustic case, a telecentric system would allow the user to listen to two different sources (e.g. two audience members in an auditorium) with the same intensity, even if they are located at different distances, without changing the lens. In a conference, it may reduce the need to fetch around the microphone during questions time.

Acoustic telescope

The introduction of the refracting telescope, in the 16th century, allowed scientists and explorers to monitor objects at much larger distances. Unique telescope designs have been developed over the centuries, in the effort of increasing the field of view and the magnification, but all are based on the combination of two lenses at a mutual distance D .

In our realisation (see Figure 1c), we demonstrated a Keplerian telescope (1611) for sound, based on two convergent lenses positioned at a (variable) distance D (see Figure 8). The lenses, that we picked of the same focal length for simplicity ($f = 53$ mm), were mounted on a rail and their mutual distance could be adjusted (with 1 mm precision) using an Arduino Nano and a stepper motor.

Telescopes solve the main limitation of Type B lenses: the focal lengths that cannot be achieved with one lens will be achieved with two, at an appropriate distance D . Applications then include those of magnifying glasses, but without the need of deploying the lens in the field. A telescope, in fact, is a vari-focal lens: the distance D can be changed to create a lens of the desired focal length. One user scenario would be listening to a source among others (e.g. zooming on a single person in a crowd - see Figure 8b - to either deliver or receive acoustic messages). Other applications of our telescope include acoustic displays (e.g. dancers following an acoustic spot in a disco like a cat would follow a laser pointer) or music with a dynamic spatial component.

Modern cameras, however, all feature auto-zooming objectives. A similar solution for sound would be capable of following a source in the field of view and, in VR, the same speaker could be used to deliver location-specific sounds to multiple users moving in the same virtual world, without headphones. As a first step towards an auto-zoom lens (see supplemental video), we positioned a receiving microphone at an unknown position in front of the speaker and adjusted the distance between the lenses (with 1 mm precision) using an Arduino Nano and a stepper motor until the signal on it was maximised (i.e. using the reading of the sound-level meter as feedback for the positioning system). We found that, when the speaker was imaged on the microphone, the meter

measured 7 dB more than the level it would have measured in the same position without metamaterials. User studies and an automatic feedback loop will be needed to decide whether this increase is sufficient, especially in otherwise noisy environments.

6 DISCUSSION

In the previous sections we have described many conceptual applications of converging metasurfaces, highlighting how our prototype devices can be used to advance their realisation. Here we summarise the advantages of our approach, and describe how remaining limitations can be overcome.

Bandwidth. As mentioned earlier (section 2), metamaterials have a limited bandwidth (5 piano keys for our Type B lenses). While this may be already sufficient for delivering alarms and conveying personal audio messages, the human audible range cover 11 octaves: larger bandwidths are highly desirable for consumer audio. This is a hot research topic: more and more solutions are appearing on the horizon.

In this work, we have shown how using unit cells at sub-resonance frequencies extends the bandwidth. This observation hints to a “rule-of-thumb”, connecting the transmission of single unit cells and their bandwidth, suggesting that larger bandwidths may be obtained already at the single cell level by accepting a certain amount of transmission loss⁶. With visual LCD displays only requiring a 20% transmission, this may be a minor issue for some applications.

The possibility of using smaller Type B cells, however, can be exploited also by designing multi-frequency structures that occupy the same space as a Type A cell (e.g. a unit cell made of Type B cells, mounted in a 2×2 array, like the RGB crystals forming pixels in a LCD display). This is the route followed by Jiménez et al. [20], who have reached a very high absorption over a broad frequency band in deep-subwavelength thickness panels by stacking side-by-side unit cells of different size, with close frequency response.

Finally, the problem of a limited bandwidth can be tackled at the device level. Achromatic lenses, for instance, are realised by stacking two different lenses (in the direction of propagation), so that two wavelengths (typically red & blue) focus on the same plane. Having proven that the thin-lens equation also applies to acoustics, when metamaterials are involved, similar solutions may be imagined for sound⁷.

⁶Type B lenses, for instance, cover 1 octave for applications where a 6 dB loss is acceptable.

⁷Cascading different acoustic filters, each with different frequency response, is not a novel idea: the first sound-level meters used just this method to apply the A-weighting and take into account perception.



Figure 8: Acoustic telescope configuration, using two converging lenses of the same focal length: (a) side view and (b) a potential application, connecting acoustically with a single person in a crowd. Front view in Figure 1.

Comparison with speaker arrays. Acoustic metamaterials which shape sound are not a competing, but a complementing approach compared with traditional sound emitters [37]. Metamaterials are smaller, cheaper and easier to manufacture than phased arrays: they can even be fabricated in recyclable materials. Metamaterial devices lead to less aberrations than speaker arrays, even over limited bandwidths.

Metamaterials solutions, however, are static.

Hybrid systems [37], mechanically actuated metamaterial devices (like our zooming lens or [19]) and fully active metamaterials [30] will be the future, but require further study. Our varifocal lens, for instance, could be integrated with a motion tracking system (e.g. LeapmotionTM or a RGB-D camera) to keep the object is “in focus” while it moves. The limitation due to the 1D mechanical actuation will be easily exceeded thanks to the integration of multiple actuators: not only on the second lens that makes up the telescope system, but also on the speaker-first lens component.

Combination of metasurfaces. As shown by Memoli et al. [35] and reinforced in this work, acoustic metasurfaces may be layered into unique devices to obtain a detailed shaping of the sound, with each layer adding complexity and functionalities to the end-user experience. This work is only a first step into multi-layered acoustic metamaterials, that needs to be further addressed in future studies.

Scalability and Compactness. Scalability is an important issue, as the size of a metasurface may vary from small to very large. While we have demonstrated the possibility of making compact, sub-wavelength thick metasurfaces, the challenge of meeting user-defined space constraints is still open. Due to their compactness metamaterial structures could be incorporated into speakers or headphones for example, paving the way for multiple applications.

7 CONCLUSIONS

Designing experiences based on sound delivery is challenging: not only sound is invisible, so that “holes” in the delivery cannot be easily spotted, but it is difficult to have an “eagle view” i.e. the perception of how the sound is “felt” at all points of the space at the same time. While metamaterials cannot shape sound on demand yet, our approach adds simple, but powerful tools to the ones available to acoustic designers, presenting them in a way that builds on centuries of visual design in the HCI community.

Next steps include on the technical side more complex shapes (e.g. Zhu et al. [66]), larger bandwidths and on-demand sound control (i.e. a “space sound modulator”). In terms of experiential design, user studies (in collaboration with musicians and psychologists) will be needed to determine the potential effects of personal sound on audiences.

Our prototypes, while simple, lower the access threshold to designing novel sound experiences: devices based on acoustic metamaterials will lead to new ways of delivering, experiencing and even thinking of sound.

ACKNOWLEDGMENTS

GM, LC and JPE acknowledge funding from the Engineering and Physical Sciences Research Council (EPSRC-UKRI) through grant EP/S001832/1. SS acknowledges funding from the Royal Academy of Engineering through their Chair in Emerging Technology grant CiET1718/14. MC acknowledges funding from the Royal Academy of Engineering through grant The authors would also like to thank the anonymous referees for the useful comments.

REFERENCES

- [1] Armando Barreto, Kenneth John Faller, and Malek Adjouadi. 2009. 3D Sound for Human-computer Interaction: Regions with Different

- Limitations in Elevation Localization. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '09)*. ACM, New York, NY, USA, 211–212. <https://doi.org/10.1145/1639642.1639680>
- [2] A. J. Berkhout, D. de Vries, and P. Vogel. 1993. Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America* 93, 5 (1993), 2764–2778. <https://doi.org/10.1121/1.405852>
- [3] Marinus M. Boone, Edwin N. G. Verheijen, and G. Jansen. 1996. Virtual Reality by Sound Reproduction Based on Wave Field Synthesis. In *Audio Engineering Society Convention 100*. <http://www.aes.org/e-lib/browse.cfm?elib=7624>
- [4] K. Born and W. Wolf. 1970. *Principles of Optics*. Pergamon Press.
- [5] Anne-Claire Bourland, Peter Gorman, Jess McIntosh, and Asier Marzo. 2018. Project Telepathy. *Interactions* 25, 5 (Aug. 2018), 16–17. <https://doi.org/10.1145/3241945>
- [6] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. 2013. UltraHaptics: Multi-point Mid-air Haptic Feedback for Touch Surfaces. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. ACM, New York, NY, USA, 505–514. <https://doi.org/10.1145/2501988.2502018>
- [7] Jordan Cheer, Stephen Daley, and Cameron McCormick. 2017. Feed-forward control of sound transmission using an active acoustic metamaterial. *Smart Materials and Structures* 26, 2 (2017), 025032. <http://stacks.iop.org/0964-1726/26/i=2/a=025032>
- [8] Andrea Colombi, Victoria Ageeva, Richard J. Smith, Adam Clare, Rikesh Patel, Matt Clark, Daniel Colquitt, Philippe Roux, Sébastien Guenneau, and Richard V. Craster. 2017. Enhanced sensing and conversion of ultrasonic Rayleigh waves by elastic metasurfaces. *Scientific Reports* 7 (2017), Issue 1. <https://doi.org/10.1038/s41598-017-07151-6>
- [9] Andrea Colombi, Daniel Colquitt, Philippe Roux, Sébastien Guenneau, and Richard V. Craster. 2016. A seismic metamaterial: The resonant metawedge. *Scientific Reports* 6 (2016). <https://doi.org/10.1038/srep27717>
- [10] Charles R. P. Courtney, Christine E. M. Demore, Hongxiao Wu, Alon Grinenko, Paul D. Wilcox, Sandy Cochran, and Bruce W. Drinkwater. 2014. Independent trapping and manipulation of microparticles using dexterous acoustic tweezers. *Applied Physics Letters* 104, 15 (2014), 154103. <https://doi.org/10.1063/1.4870489>
- [11] Steven A. Cummer, Johan Christensen, and Andrea Alù. 2016. Controlling sound with acoustic metamaterials. *Nature Reviews Materials* 1 (2016), 16001. <https://doi.org/10.1038/natrevmats.2016.1>
- [12] Cassidy Curtis, David Eisenmann, Rachid El Guerrab, and Scot Stafford. 2016. The Making of Pearl, a 360° Google Spotlight Story. In *ACM SIGGRAPH 2016 VR Village (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 21, 1 pages. <https://doi.org/10.1145/2929490.2956565>
- [13] Woon-Seng Gan, Jun Yang, and Tomoo Kamakura. 2012. A review of parametric acoustic array in air. *Applied Acoustics* 73, 12 (2012), 1211 – 1219. <https://doi.org/10.1016/j.apacoust.2012.04.001>
- [14] Stanislav B. Glybovski, Sergei A. Tretyakov, Pavel A. Belov, Yuri S. Kivshar, and Constantin R. Simovski. 2016. Metasurfaces: From microwaves to visible. *Physics Reports* 634 (2016), 1 – 72. <https://doi.org/10.1016/j.physrep.2016.04.004>
- [15] Simon Holland, David R. Morse, and Henrik Gedenryd. 2002. AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface. *Personal Ubiquitous Comput.* 6, 4 (Jan. 2002), 253–259. <https://doi.org/10.1007/s007790200025>
- [16] Alexandra Ion, Johannes Frohnhofer, Ludwig Wall, Robert Kovacs, Mirela Alistar, Jack Lindsay, Pedro Lopes, Hsiang-Ting Chen, and Patrick Baudisch. 2016. Metamaterial Mechanisms. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 529–539. <https://doi.org/10.1145/2984511.2984540>
- [17] Alexandra Ion, Ludwig Wall, Robert Kovacs, and Patrick Baudisch. 2017. Digital Mechanical Metamaterials. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 977–988. <https://doi.org/10.1145/3025453.3025624>
- [18] Miyu Iwafune, Taisuke Ohshima, and Yoichi Ochiai. 2016. Coded Skeleton: Programmable Deformation Behaviour for Shape Changing Interfaces. In *SIGGRAPH ASIA 2016 Emerging Technologies (SA '16)*. ACM, New York, NY, USA, Article 1, 2 pages. <https://doi.org/10.1145/2988240.2988252>
- [19] Louis Jackowski-Ashley, Gianluca Memoli, Mihai Caleap, Nicolas Slack, Bruce W. Drinkwater, and Sriram Subramanian. 2017. Haptics and Directional Audio Using Acoustic Metasurfaces. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, New York, NY, USA, 429–433. <https://doi.org/10.1145/3132272.3132285>
- [20] Noé Jiménez, Vicent Romero-García, Vincent Pagneux, and Jean-Philippe Groby. 2017. Rainbow-trapping absorbers: Broadband, perfect and asymmetric sound absorption by subwavelength panels for transmission problems. *Scientific Reports* 7 (2017). Issue 1. <https://doi.org/10.1038/s41598-017-13706-4>
- [21] Toshiyuki Kimura, Munenori Naoe, Yoko Yamakata, and Michiaki Katsumoto. 2009. Subjective Effect of Synthesis Conditions in 3D Sound Field Reproduction System Using a Few Transducers and Wave Field Synthesis. In *Proceedings of the 3rd International Universal Communication Symposium (IUCS '09)*. ACM, New York, NY, USA, 215–220. <https://doi.org/10.1145/1667780.1667825>
- [22] Jussi Kuutti, Juhana Leivo, and Raimo E. Sepponen. 2014. Local Control of Audio Environment: A Review of Methods and Applications. *Technologies* 2, 1 (2014), 31–53. <https://doi.org/10.3390/technologies2010031>
- [23] R. J. Lalonde, A. Worthington, and J. W. Hunt. 1993. Field conjugate acoustic lenses for ultrasound hyperthermia. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 40, 5 (1993), 592–602. <https://doi.org/10.1109/58.238113>
- [24] C. J. L. Lane, A. K. Dunhill, B. W. Drinkwater, and P. D. Wilcox. 2010. The inspection of anisotropic single-crystal components using a 2-D ultrasonic array. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 57, 12 (December 2010), 2742–2752. <https://doi.org/10.1109/TUFFC.2010.1748>
- [25] Fabrice Lemoult, Nadège Kaina, Mathias Fink, and Geoffroy Lerosey. 2016. Soda Cans Metamaterial: A Subwavelength-Scaled Phononic Crystal. *Crystals* 6, 7 (2016). <https://doi.org/10.3390/cryst6070082>
- [26] Dingzeyu Li, David I. W. Levin, Wojciech Matusik, and Changxi Zheng. 2016. Acoustic Voxels: Computational Optimization of Modular Acoustic Filters. *ACM Trans. Graph.* 35, 4, Article 88 (July 2016), 12 pages. <https://doi.org/10.1145/2897824.2925960>
- [27] Yong Li, Xue Jiang, Rui-qi Li, Bin Liang, Xin-ye Zou, Lei-lei Yin, and Jian-chun Cheng. 2014. Experimental Realization of Full Control of Reflected Waves with Subwavelength Acoustic Metasurfaces. *Phys. Rev. Applied* 2 (Dec 2014), 064002. Issue 6. <https://doi.org/10.1103/PhysRevApplied.2.064002>
- [28] Yong Li, Bin Liang, Zhong-ming Gu, Xin-ye Zou, and Jian-chun Cheng. 2013. Reflected waveform manipulation based on ultrathin planar acoustic metasurfaces. *Scientific Reports* 3 (2013). <https://doi.org/10.1038/srep02546>
- [29] Chenkai Liu, Jie Luo, and Yun Lai. 2018. Acoustic metamaterials with broadband and wide-angle impedance matching. *Phys. Rev. Materials* 2 (Apr 2018), 045201. Issue 4. <https://doi.org/10.1103/PhysRevMaterials.2.045201>
- [30] Guancong Ma, Xiyi Fan, Ping Sheng, and Mathias Fink. 2018. Shaping reverberating sound fields with an actively tunable metasurface. *Proceedings of the National Academy of Sciences* 115,

- 26 (2018), 6638–6643. <https://doi.org/10.1073/pnas.1801175115> arXiv:<https://www.pnas.org/content/115/26/6638.full.pdf>
- [31] Guancong Ma and Ping Sheng. 2016. Acoustic metamaterials: From local resonances to broad horizons. *Science Advances* 2, 2 (2016). <https://doi.org/10.1126/sciadv.1501595>
- [32] Sebastian Marwecki and Patrick Baudisch. 2018. Scenograph: Fitting Real-Walking VR Experiences into Various Tracking Volumes. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. ACM, New York, NY, USA, 511–520. <https://doi.org/10.1145/3242587.3242648>
- [33] Asier Marzo, Sue Ann Seah, Bruce W. Drinkwater, Deepak Ranjan Sahoo, Benjamin Long, and Sriram Subramanian. 2015. Holographic acoustic elements for manipulation of levitated objects. *Nature Communications* 6 (2015). <https://doi.org/10.1038/ncomms9661>
- [34] Jun Mei and Ying Wu. 2014. Controllable transmission and total reflection through an impedance-matched acoustic metasurface. *New Journal of Physics* 16, 12 (2014), 123007. <http://stacks.iop.org/1367-2630/16/i=12/a=123007>
- [35] Gianluca Memoli, Mihai Caleap, Michihiro Asakawa, Deepak R. Sahoo, Bruce W. Drinkwater, and Sriram Subramanian. 2017. Metamaterial bricks and quantization of meta-surfaces. *Nature Communications* 8 (2017). <https://doi.org/10.1038/ncomms14608>
- [36] S. C. Mondal, P. D. Wilcox, and B. W. Drinkwater. 2005. Design and Evaluation of Two Dimensional Phased Array Ultrasonic Transducers. *AIP Conference Proceedings* 760, 1 (2005), 899–905. <https://doi.org/10.1063/1.1916769>
- [37] Mohd Adili Norasikin, Diego Martinez Plascencia, Spyros Polychronopoulos, Gianluca Memoli, Yutaka Tokuda, and Sriram Subramanian. 2018. SoundBender: Dynamic Acoustic Control Behind Obstacles. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. ACM, New York, NY, USA, 247–259. <https://doi.org/10.1145/3242587.3242590>
- [38] Marianna Obrist, Grace Boyle, Marcel van Brakel, and Frederik Duerinck. 2017. Multisensory Experiences & Spaces. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, New York, NY, USA, 469–472. <https://doi.org/10.1145/3132272.3135086>
- [39] Yoichi Ochiai, Takayuki Hoshi, and Jun Rekimoto. 2014. Pixie Dust: Graphics Generated by Levitated and Animated Objects in Computational Acoustic-potential Field. In *ACM SIGGRAPH 2014 Posters (SIGGRAPH '14)*. ACM, New York, NY, USA, Article 83, 1 pages. <https://doi.org/10.1145/2614217.2614219>
- [40] Yoichi Ochiai, Takayuki Hoshi, and Ipppei Suzuki. 2017. Holographic Whisper: Rendering Audible Sound Spots in Three-dimensional Space by Focusing Ultrasonic Waves. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4314–4325. <https://doi.org/10.1145/3025453.3025989>
- [41] Johannes T.B. Overvelde, Twan A. de Jong, Yanina Shevchenko, Sergio A. Becerra, George M. Whitesides, James C. Weaver, Chuck Hoberman, and Katie Bertoldi. 2016. A three-dimensional actuated origami-inspired transformable metamaterial with multiple degrees of freedom. *Nature Communications* 7 (2016). <https://doi.org/10.1038/ncomms10929>
- [42] Giuseppe Pesce, Giorgio Volpe, Onofrio M. Maragó, Philip H. Jones, Sylvain Gigan, Antonio Sasso, and Giovanni Volpe. 2015. Step-by-step guide to the realization of advanced optical tweezers. *J. Opt. Soc. Am. B* 32, 5 (May 2015), B84–B98. <https://doi.org/10.1364/JOSAB.32.000B84>
- [43] Rory J. Piper, Mark A. Hughes, Carmel M. Moran, and Jothy Kanadasamy. 2016. Focused ultrasound as a non-invasive intervention for neurological disease: a review. *British Journal of Neurosurgery* 30, 3 (2016), 286–293. <https://doi.org/10.3109/02688697.2016.1173189> PMID: 27101792. arXiv:<https://doi.org/10.3109/02688697.2016.1173189> PMID: 27101792.
- [44] F. Pokorny and F. Graf. 2014. Akustische Vermessung parametrischer Lautsprecherarrays im Kontext der Transauraltechnik (in German). In *Annual meeting of the German Acoustical Society (DAGA'14)*. http://pub.dega-akustik.de/DAGA_2014/data/articles/000129.pdf
- [45] F. Joseph Pompei. 1998. The Use of Airborne Ultrasonics for Generating Audible Sound Beams. In *Audio Engineering Society Convention 105*. <http://www.aes.org/e-lib/browse.cfm?elib=8327>
- [46] Bogdan-Ioan Popa, Durvesh Shinde, Adam Konneker, and Steven A. Cummer. 2015. Active acoustic metamaterials reconfigurable in real time. *Phys. Rev. B* 91 (Jun 2015), 220303. Issue 22. <https://doi.org/10.1103/PhysRevB.91.220303>
- [47] Nimesha Ranasinghe, Pravar Jain, Nguyen Thi Ngoc Tram, Koon Chuan Raymond Koh, David Tolley, Shienny Karwita, Lin Lien-Ya, Yan Liangkun, Kala Shamaiah, Chow Eason Wai Tung, Ching Chiuan Yen, and Ellen Yi-Luen Do. 2018. Season Traveller: Multisensory Narration for Enhancing the Virtual Reality Experience. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 577, 13 pages. <https://doi.org/10.1145/3173574.3174151>
- [48] Jaime Reis. 2016. Short overview in parametric loudspeakers array technology and its implications in spatialization in electronic music. In *International Computer Music Conference 2016*. Michigan Publishing, 241–248. <https://quod.lib.umich.edu/i/icmc/bbp2372.2016.047/1>
- [49] Francis Rumsey. 2013. *Spatial Audio (Music Technology) 3rd ed.* Focal Press.
- [50] Jaime Sánchez and Mauricio Sáenz. 2005. 3D Sound Interactive Environments for Problem Solving. In *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '05)*. ACM, New York, NY, USA, 173–179. <https://doi.org/10.1145/1090785.1090817>
- [51] Lauri Savioja and U. Peter Svensson. 2015. Overview of geometrical room acoustic modeling techniques. *The Journal of the Acoustical Society of America* 138, 2 (2015), 708–730. <https://doi.org/10.1121/1.4926438>
- [52] Gang Yong Song, Qiang Cheng, Bei Huang, Hui Yuan Dong, and Tie Jun Cui. 2016. Broadband fractal acoustic metamaterials for low-frequency sound attenuation. *Applied Physics Letters* 109, 13 (2016), 131901. <https://doi.org/10.1063/1.4963347> arXiv:<https://doi.org/10.1063/1.4963347>
- [53] Robert E. Stevens, Thomas N. L. Jacoby, Ilinca S. Aricescu, and Daniel P. Rhodes. 2017. A review of adjustable lenses for head mounted displays. , 10335 pages. <https://doi.org/10.1117/12.2276677>
- [54] Steven Strachan, Parisa Eslambolchilar, Roderick Murray-Smith, Stephen Hughes, and Sile O'Modhrain. 2005. GpsTunes: Controlling Navigation via Audio Feedback. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services (MobileHCI '05)*. ACM, New York, NY, USA, 275–278. <https://doi.org/10.1145/1085777.1085831>
- [55] Yutaka Takase and Shoichi Hasegawa. 2012. Presentation of Directional Information by Sound Field Control. In *Proceedings of the 3rd Augmented Human International Conference (AH '12)*. ACM, New York, NY, USA, Article 32, 3 pages. <https://doi.org/10.1145/2160125.2160157>
- [56] Kun Tang, Chunyin Qiu, Manzhu Ke, Jiuyang Lu, Yangtao Ye, and Zhengyou Liu. 2014. Anomalous refraction of airborne sound through ultrathin metasurfaces. *Scientific Reports* 4 (2014). <https://doi.org/10.1038/srep06517>
- [57] D. H. Turnbull and F. S. Foster. 1991. Beam steering with pulsed two-dimensional transducer arrays. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 38, 4 (July 1991), 320–333. <https://doi.org/10.1109/58.84270>
- [58] Shi-Chang Wooh and Yijun Shi. 1998. Influence of phased array element size on beam steering behavior. *Ultrasonics* 36, 6 (1998), 737 – 749. [https://doi.org/10.1016/S0041-624X\(97\)00164-9](https://doi.org/10.1016/S0041-624X(97)00164-9)

- [59] Yangbo Xie, Wenqi Wang, Huanyang Chen, Adam Konneker, Bogdan-Ioan Popa, and Steven A. Cummer. 2014. Wavefront modulation and subwavelength diffractive acoustics with an acoustic metasurface. *Nature Communications* 5 (2014). <https://doi.org/10.1038/ncomms6553>
- [60] Jun Yang, Woon-Seng Gan, Khim-Sia Tan, and Meng-Hwa Er. 2005. Acoustic beamforming of a parametric speaker comprising ultrasonic transducers. *Sensors and Actuators A: Physical* 125, 1 (2005), 91 – 99. <https://doi.org/10.1016/j.sna.2005.04.037>
- [61] Chunfang Ye and Robert R. McLeod. 2008. GRIN lens and lens array fabrication with diffusion-driven photopolymer. *Opt. Lett.* 33, 22 (Nov 2008), 2575–2577. <https://doi.org/10.1364/OL.33.002575>
- [62] Masahide Yoneyama, Jun-ichiro Fujimoto, Yu Kawamo, and Shoichi Sasabe. 1983. The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design. *The Journal of the Acoustical Society of America* 73, 5 (1983), 1532–1536. <https://doi.org/10.1121/1.389414>
- [63] Seraphina Yong and Hao-Chuan Wang. 2018. Using Spatialized Audio to Improve Human Spatial Knowledge Acquisition in Virtual Reality. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion (IUI '18 Companion)*. ACM, New York, NY, USA, Article 51, 2 pages. <https://doi.org/10.1145/3180308.3180360>
- [64] Nanfang Yu, Patrice Genevet, Mikhail A. Kats, Francesco Aieta, Jean-Philippe Tetienne, Federico Capasso, and Zeno Gaburro. 2011. Light Propagation with Phase Discontinuities: Generalized Laws of Reflection and Refraction. 334 (2011), 333–337. <https://doi.org/10.1126/science.1210713>
- [65] Bertram Yuh, An Liu, Robert Beatty, Alexander Jung, and Jeffrey Y C Wong. 2016. Focal therapy using magnetic resonance image-guided focused ultrasound in patients with localized prostate cancer. *Journal of therapeutic ultrasound* 4 (2016). Issue 8. <https://doi.org/10.1186/s40349-016-0054-y>
- [66] Xuefeng Zhu, Kun Li, Peng Zhang, Jie Zhu, Jintao Zhang, Chao Tian, and Shengchun Liu. 2016. Implementation of dispersion-free slow acoustic wave propagation and phase engineering with helical-structured metamaterials. *Nature Communications* 7 (2016). <https://doi.org/10.1038/ncomms11731>