Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

From Pixels to Physicality: Wave-Mediated Interaction above a Screen

Ki-Uk Kyung

kyungku@kaist.ac.kr

KAIST https://orcid.org/0000-0002-2707-8516

Jihwan Park

KAIST

Sunho Park

SK hynix

Jaeseung Han

KAIST

Ji-Sung Kim

KAIST

Article

Keywords:

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11442572/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations:
There is **NO** Competing Interest.

Consent to Participate: Informed consent was obtained from all participants prior to their participation in the study.

Consent to Publish: All participants provided informed consent for the publication of the research results. No personally identifiable information is included in the manuscript.

From Pixels to Physicality: Wave-Mediated Interaction above a Screen

Jihwan Park, Sunho Park, Jaeseung Han, Ji-Sung Kim, & Ki-Uk Kyung*

Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea.

*Corresponding author(s). E-mail(s): kyungku@kaist.ac.kr;
Contributing authors: wise4321@kaist.ac.kr; psho2512@kaist.ac.kr; js99han@kaist.ac.kr; lim950616@kaist.ac.kr;

Abstract

Visuo-physical displays that integrate visual information with physical interaction generally require additional actuation systems. Here, we propose a transparent visuo-physical panel that extends the role of a display front panel beyond its optical function to include acoustic wave propagation and radiation, enabling focused ultrasound above the screen. Unlike conventional mid-air interaction systems composed of hundreds of opaque ultrasonic transducers positioned separately from the display, the panel itself radiates ultrasound into air without obstructing the viewing area. Although waves excited at the panel boundary undergo complex propagation, dispersion, and reflection within the panel before radiating into air, we develop spatiotemporal inverse filtering to focus the radiated ultrasound at a prescribed target point in air. The focal position can be updated more than 100 times per second, and multiple focal points can be generated simultaneously. Acoustic pressures approaching 1 kPa enable object lifting, mid-air haptic feedback, and selective candle extinguishing, demonstrating physical interaction generated directly by the display panel.

Introduction

What if the actions we see on a screen could have physical effects in the world beyond it? Visual displays can depict motion and physical events, but displaying them does not, by itself, generate corresponding forces or sensations beyond the screen. While conventional two-dimensional displays present visual content on a flat screen, recent advances in holographic and light-field displays allow images to appear at different depths in three-dimensional space [1–4]. The ability to generate localized physical effects beyond the screen could open new possibilities for interaction with both conventional and emerging visual displays.

Various technologies have been developed to provide physical feedback directly through display surfaces [5, 6]. Localized vibrotactile feedback [7–9], electrovibration [10, 11], ultrasonic friction modulation [12, 13], and shape-morphing displays [14, 15] can spatially coordinate physical feedback with visual content. However, these interactions remain confined to the display surface. As visual content extends into the space above the display, physical interaction must likewise move beyond the screen surface.

Airborne ultrasound offers a way to generate non-contact physical interaction in three-dimensional space [16, 17]. Focused high-intensity ultrasound produces acoustic radiation force [18], enabling manipulation or levitation of lightweight objects [19, 20] and mid-air tactile stimulation [21, 22]. Phased arrays can electronically steer the acoustic

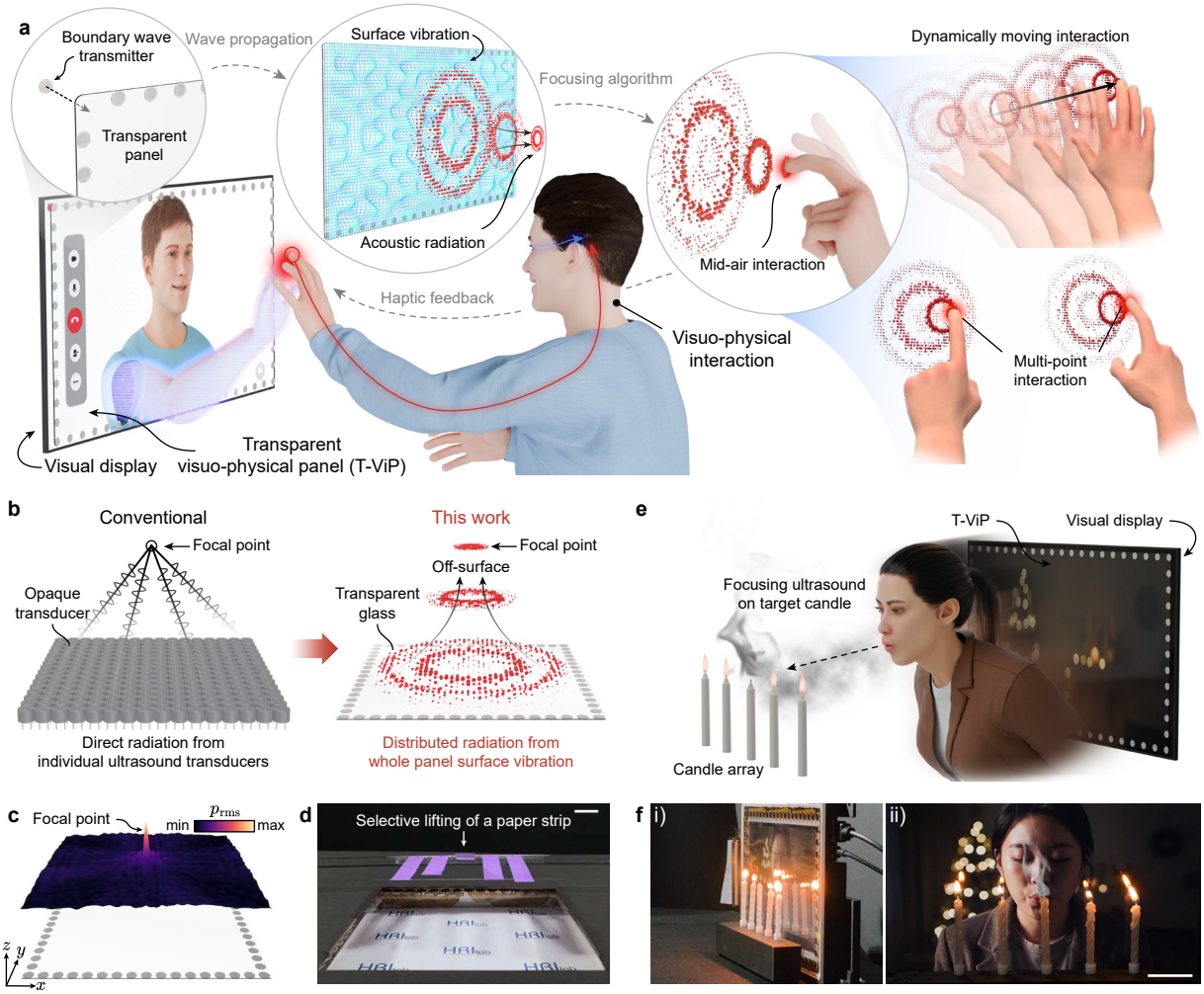


Fig. 1: **Concept and demonstrations of the T-ViP. a** Concept of the proposed T-ViP. Boundary wave transmitters excite bending waves in a transparent panel, whose surface vibration radiates ultrasound to form programmable off-surface focal points. The focal points can be spatially collocated with visual content and controlled to provide dynamic and multi-point physical interaction. **b** Comparison between a conventional ultrasonic phased array and the proposed approach. Conventional systems form the acoustic aperture with individual opaque transducers that directly radiate ultrasound, obstructing the viewing area. In contrast, the T-ViP radiates ultrasound from the vibrating surface of the entire transparent panel, leaving the viewing area unobstructed. **c** Measured acoustic field showing ultrasound focused at a target off-surface focal point. **d** Lifting of a paper strip by the acoustic radiation force generated at the focal point, demonstrating physical interaction without contact. **e, f** Visuo-physical demonstration of spatially collocated visual content and off-surface physical interaction. **e** Conceptual illustration of the integrated system. **f** Demonstration in which focused ultrasound blows out a target candle synchronized with visual content: **i**, oblique view; **ii**, front view. White scale bars in **d** and **f-ii** represent 25 mm.

focus [23, 24] and generate dynamic or multiple focal points [25, 26]. Integrating such arrays with visual displays, however, is difficult because conventional systems use tens to hundreds of opaque transducers [27]. Existing interfaces therefore place arrays beside the display [23, 28], project visual content above the array [24], or rely on additional optical elements such as aerial imaging plates [29] and fog screens [30], adding hardware, space, or optical complexity.

One approach to overcoming this limitation is to make the acoustic aperture itself transparent. Recent studies have explored transparent micromachined ultrasonic transducers and arrays for integrating mid-air haptic feedback directly with visual displays [31–34]. These approaches provide a potential route to integrating the acoustic aperture into the display viewing area without obstructing the optical path. However, array-based architectures still require active transducer materials, electrodes, electrical interconnects, and supporting structures distributed throughout the

aperture. Achieving sufficient acoustic output and spatial controllability while maintaining high optical transparency over the entire aperture therefore remains challenging.

But what if the transparent panel itself could generate the ultrasound? Here, we propose a transparent visuo-physical panel (T-ViP) that uses the glass panel not simply as an optical window, but as a distributed acoustic aperture. Piezoelectric wave transmitters placed along the boundary excite waves that propagate across the glass, creating a distributed vibration field that radiates ultrasound into the surrounding air (Fig. 1a). In contrast to a conventional ultrasonic phased array, in which discrete transducers directly radiate ultrasound to form the acoustic aperture, the T-ViP uses the transparent panel itself to generate acoustic radiation, leaving the viewing area unobstructed (Fig. 1b).

Controlling such an aperture, however, differs fundamentally from controlling a conventional ultrasonic phased array. In a phased array, each transducer directly radiates ultrasound into air and can be controlled independently. In the T-ViP, by contrast, waves excited at the panel boundary propagate, disperse, reflect, and interfere within the glass before radiating into air. We therefore applied a spatiotemporal inverse filtering (STIF) method that incorporates the electromechanical–acoustic transfer responses between the boundary wave transmitters and a prescribed target point in space [35, 36]. Based on these measured transfer responses, STIF synthesizes broadband driving signals that compensate for the complex wave propagation through the panel and concentrate the radiated acoustic field at the desired spatial location and time.

Using this approach, we generated localized acoustic focal points at prescribed positions in three-dimensional space above the panel, as confirmed by spatial measurements of the acoustic field (Fig. 1c). The focal point could be dynamically moved along prescribed trajectories, and multiple focal points could be generated simultaneously at different locations. The resulting acoustic radiation force was sufficient to lift a paper strip positioned at the focal point, demonstrating non-contact physical interaction above the display surface (Fig. 1d).

We further evaluated whether the localized acoustic field could provide perceivable mid-air haptic feedback. Perceptual experiments confirmed that modulated acoustic stimuli at the focal point produced detectable haptic sensations. Finally, we integrated the T-ViP with a visual display to spatially coordinate visual content and physical effects within the same three-dimensional interaction space (Fig. 1e,f). In one demonstration, focused ultrasound was synchronized with a visual scene of a person blowing, causing a physical candle positioned in front of the display to be blown out (Fig. 1f; Supplementary Video 1). By integrating optical transmission and acoustic wave propagation within the same transparent panel, the T-ViP brings visual content and localized physical interaction into a shared three-dimensional space, providing a unified visuo-physical interface above the display surface.

Results

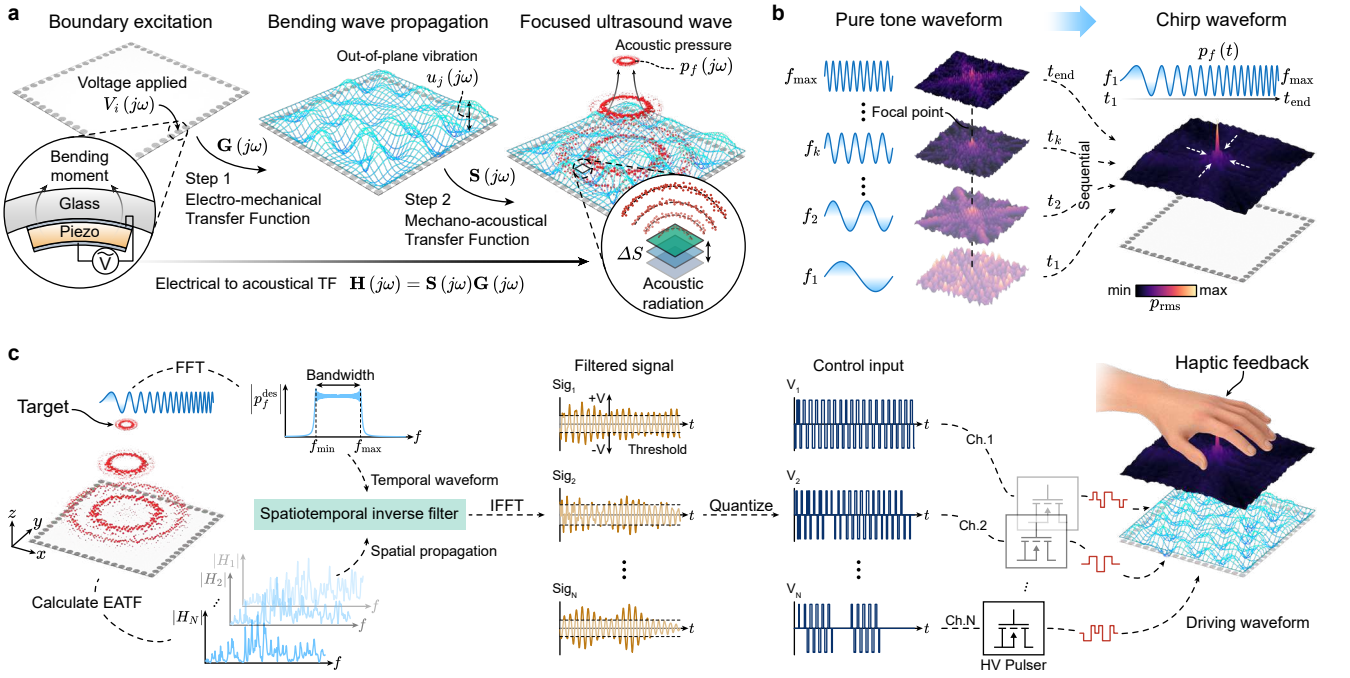


Fig. 2: **Electromechanical–acoustic transfer and STIF for off-surface acoustic focusing. a** Electromechanical–acoustic transfer from boundary excitation to panel vibration and acoustic radiation. A piezoelectric wave transmitter excites bending waves in the glass panel, and the resulting out-of-plane vibration radiates ultrasound into air. **b** Concept of broadband acoustic focusing. Different frequency components are focused at the same target position, while their sidelobe patterns vary with frequency, reducing the prominence of off-target pressure peaks. **c** STIF for acoustic focusing. A prescribed target position and temporal waveform are combined with the EATF in the frequency domain to calculate the input spectrum for each wave transmitter. The resulting signals are transformed to the time domain using an IFFT, quantized into three-level control inputs, and applied to the wave transmitters through an HV pulser.

Design and operation of the T-ViP

The T-ViP comprises piezoelectric wave transmitters arranged along the boundary of a glass panel (Fig. 1a). When a voltage is applied to the wave transmitters, the induced piezoelectric strain generates bending moments in the bonded glass panel, exciting bending waves that propagate across the panel. The resulting distributed out-of-plane vibration radiates ultrasound into the surrounding air (Fig. 2a). Unlike conventional ultrasonic phased arrays, in which individual transducers directly radiate ultrasound into air, the T-ViP converts the electrical inputs applied at the panel boundary into acoustic pressure through the coupled processes of piezoelectric actuation, structural vibration, and acoustic radiation.

The electromechanical transfer matrix $\mathbf{G}(j\omega)$ relates the voltages applied to N_i wave transmitters to the out-of-plane velocities at N_o discrete positions on the glass panel. Each element of $\mathbf{G}(j\omega)$ therefore represents the velocity response at a specific panel position to the voltage applied to a given wave transmitter. The mechanoacoustic transfer matrix $\mathbf{S}(j\omega)$ relates the distributed panel vibration to the acoustic pressure field. The overall electrical-to-acoustic transfer function (EATF) is then expressed as

$$\mathbf{H}(j\omega) = \mathbf{S}(j\omega)\mathbf{G}(j\omega) \tag{1}$$

where $\mathbf{H}(j\omega)$ describes the transfer from the wave transmitter inputs to the acoustic pressure at positions above the panel. The detailed analytical formulations of the electromechanical and mechanoacoustic transfers are provided in Supplementary Note 1 and Supplementary Fig. 1.

The glass and piezoelectric thicknesses were selected by considering both bending actuation and acoustic radiation. The analytical model showed that bending excitation is maximized when the piezoelectric-to-glass thickness ratio is approximately 0.5 (Supplementary Fig. 2). The glass thickness also determines the coincidence frequency, at which the bending-wave phase velocity matches the speed of sound in air. Above this frequency, the acoustic radiation efficiency of the panel increases (Supplementary Fig. 3). For the 0.4 mm-thick glass panel used here, the calculated coincidence frequency was approximately 29.4 kHz. We therefore used 0.195 mm-thick piezoelectric wave transmitters and selected an operating frequency range of 30 kHz to 80 kHz for subsequent acoustic focusing.

Broadband acoustic focusing with STIF

Acoustic focusing through the T-ViP is complicated by wave propagation within the glass panel. Boundary excitation produces frequency-dependent vibration patterns, and the resulting acoustic waves can interfere constructively not only at the prescribed target position but also at other locations, producing secondary pressure maxima, or sidelobes. These sidelobes reduce the pressure contrast around the focal point and limit the spatial localization of the acoustic field.

To examine this behavior, we numerically analyzed the panel vibration and radiated acoustic field over the 30 kHz to 80 kHz operating range. The analysis used the panel geometry and wave transmitter configuration described in Supplementary Note 2, with the material properties listed in Supplementary Table 1 and the model configuration shown in Supplementary Fig. 4. We first performed single-frequency (pure-tone) focusing using phase conjugation [37]. At each excitation frequency, the phases of the wave transmitter inputs were adjusted so that their acoustic contributions added constructively at the target position (see Supplementary Note 3 for details). Focusing was evaluated at the target positions and focal heights shown in Supplementary Fig. 5 and Supplementary Table 2. A local pressure maximum was formed at the prescribed target position at each tested frequency, accompanied by sidelobes outside the focal region (Supplementary Figs. 6–8). Importantly, both the locations and amplitudes of these sidelobes varied with excitation frequency.

We exploited this frequency dependence to improve spatial focusing using broadband excitation (Fig. 2b). While different frequency components can be focused at the same target position, their sidelobe distributions vary because the panel vibration pattern changes with frequency. Combining these frequency-dependent acoustic fields can therefore produce a more localized RMS pressure distribution by reducing the prominence of persistent sidelobes outside the focal region.

To implement broadband focusing, we used STIF to compensate for the frequency-dependent transfer responses from the wave transmitters to the target position (Fig. 2c) [35]. A linear chirp spanning 30 kHz to 80 kHz was prescribed as the desired pressure waveform at the target position. Using the EATF, the input spectrum for each wave transmitter was calculated as

$$\mathbf{V}^{\mathrm{opt}}(j\omega) = \mathbf{H}^{\dagger}(j\omega)p_f^{\mathrm{des}}(j\omega) \tag{2}$$

where $\mathbf{V}^{\mathrm{opt}}(j\omega)$ is the vector of wave transmitter input spectra, $p_f^{\mathrm{des}}(j\omega)$ is the frequency spectrum of the desired pressure waveform, and $\mathbf{H}^{\dagger}(j\omega)$ denotes the pseudoinverse of the EATF. Regularization was applied to limit excessive input amplitudes at frequencies where the transfer response was weak.

The calculated spectra were transformed into time-domain signals using an inverse fast Fourier transform (IFFT). Previous work has shown that acoustic focusing can be maintained with signal quantization [38]. We therefore normalized and quantized the calculated signals into three voltage levels, $+V_{\max}$, 0, and $-V_{\max}$, enabling fast switching through the high-voltage (HV) pulser. Further details of the STIF procedure are provided in Supplementary Note 4, and the signal-generation process is illustrated in Supplementary Fig. 9.

Numerical simulations showed that STIF generated the prescribed broadband frequency sweep at the focal point (Supplementary Fig. 10). Compared with pure-tone focusing, the simulated broadband RMS pressure fields exhibited a more distinct focal region, as the frequency-dependent sidelobes were distributed over different spatial locations (Supplementary Figs. 11–13).

Acoustic focusing in three-dimensional space

To experimentally validate acoustic focusing with STIF, we fabricated a T-ViP comprising 64 piezoelectric wave transmitters arranged along the boundary of a transparent glass panel measuring 150 mm × 150 mm × 0.4 mm (Fig. 3a). The wave transmitters were connected to external driving electronics through flexible printed circuit boards (FPCBs), and the panel was mounted in a housing that supported its edges (Supplementary Figs. 14 and 15). A multichannel HV pulser was implemented to drive the individual wave transmitters (Supplementary Figs. 16 and 17), together with signal-conditioning circuitry for measuring the input voltage, panel vibration, and acoustic pressure responses (Supplementary Fig. 18). System operation and data acquisition were synchronized using a field-programmable gate array (FPGA)-based control system (Supplementary Figs. 19 and 20). Details of the device fabrication and the driving and measurement hardware are provided in Supplementary Note 5.

To implement STIF on the fabricated T-ViP, we first identified the electromechanical transfer characteristics from broadband measurements of the panel surface velocity (Supplementary Fig. 21). These measurements were combined with the mechanoacoustic model to obtain the EATF used to calculate the STIF driving signals (Supplementary Note 6). The resulting acoustic field was measured by scanning a microphone over a plane parallel to the panel, and the focusing performance was evaluated from the spatial distribution of root-mean-square (RMS) acoustic pressure (Fig. 3a).

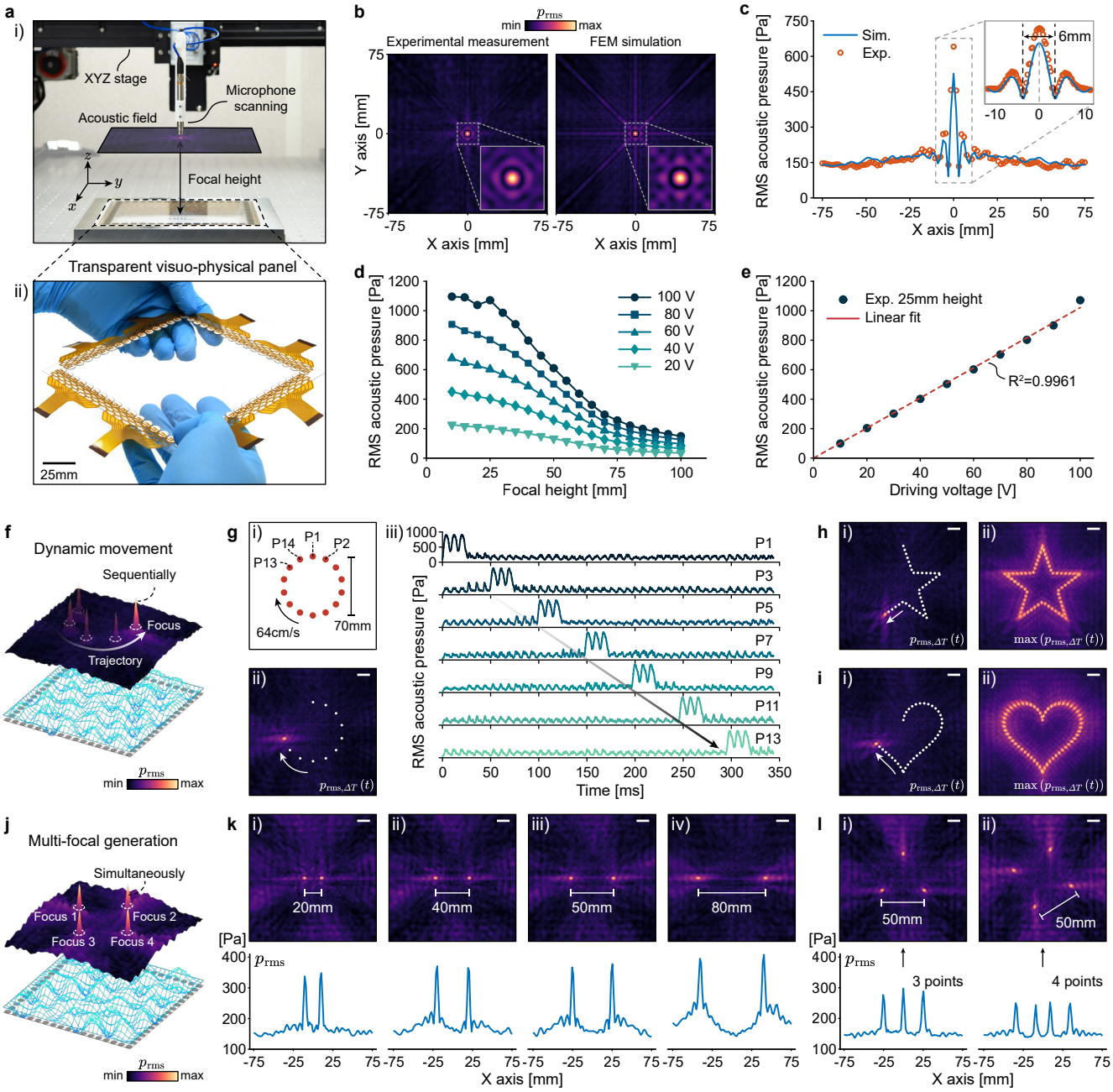


Fig. 3: **Characterization of three-dimensional, dynamic, and multi-point acoustic focusing. a** Measurement system and fabricated T-ViP: i) acoustic field scanning setup; ii) fabricated device. **b** Measured (left) and simulated (right) RMS acoustic pressure fields for a focal point at the panel center, $(x, y) = (0, 0)$, with a focal height of $z = 50$ mm. Insets show enlarged views around the focal point. **c** Corresponding x-axis RMS acoustic pressure profiles, with an enlarged focal profile shown in the inset. The focal diameter is approximately 6 mm. **d** RMS acoustic pressure at the focal point for different focal heights and driving voltages. **e** RMS acoustic pressure at the focal point as a function of driving voltage at a focal height of 25 mm. Dashed line: linear fit ($R^2 = 0.9961$). **f–i** Dynamic focal-point movement. **f** Schematic of sequential focal-point movement. **g** i) Circular trajectory; ii) transient RMS acoustic pressure field; iii) RMS acoustic pressure over time at selected positions. **h,i** Star-shaped (**h**) and heart-shaped (**i**) trajectories: i) transient RMS acoustic pressure field; ii) maximum transient RMS acoustic pressure at each position over time. **j–l** Simultaneous generation of multiple focal points. **j** Schematic of simultaneous multi-point focusing. **k** RMS acoustic pressure fields and profiles for two focal points separated by i) 20 mm; ii) 40 mm; iii) 50 mm; iv) 80 mm. **l** RMS acoustic pressure fields and profiles for three (left) and four (right) focal points. All white scale bars represent 25 mm.

With the focal point prescribed at 50 mm above the center of the panel, the measured RMS acoustic pressure field exhibited a localized pressure maximum at the target position (Fig. 3b). The spatial distribution around the focal point was consistent with the simulated field, and the measured and simulated x-axis RMS acoustic pressure profiles showed similar peak positions and widths (Fig. 3c). The measured focal diameter was approximately 6 mm, and the RMS acoustic pressure at the focal point reached approximately 716 Pa at a driving voltage of 100 V.

We next evaluated whether focal points could be generated at different positions in three-dimensional space. Acoustic fields were measured at nine lateral target positions and at focal heights of 25 mm, 50 mm, and 75 mm. At each focal height, localized pressure maxima were observed near the prescribed target positions, including positions away from the center of the panel (Supplementary Figs. 22–24). These measurements were consistent with the spatial focusing behavior predicted by the simulations (Supplementary Figs. 11–13), demonstrating that the T-ViP can generate localized focal points at prescribed positions in three-dimensional space.

The RMS acoustic pressure at the focal point depended on both focal height and driving voltage (Fig. 3d,e). Under the measurement conditions, the maximum RMS acoustic pressure reached approximately 1096 Pa at a focal height of 10 mm and a driving voltage of 100 V, and generally decreased with increasing focal height (Fig. 3d). The RMS acoustic pressure at the focal point also increased nearly linearly with driving voltage ($R^2 = 0.9961$; Fig. 3e). These results establish spatially controlled acoustic focusing throughout the three-dimensional space above the panel, providing the basis for localized physical interaction.

Dynamic and multi-point physical interaction

We next extended STIF to dynamically move the focal point over time (Fig. 3f–i). STIF driving signals were calculated for a sequence of target positions along a prescribed trajectory and sequentially applied to move the focal point along the path (Fig. 3f; Supplementary Note 6). A circular trajectory with a diameter of 70 mm was first defined at a focal height of 50 mm, with 14 target positions distributed along the trajectory (Fig. 3g). The focal point moved sequentially through these positions at approximately 64 cm/s, completing one revolution in approximately 344 ms. The local maximum of the measured RMS acoustic field followed the prescribed trajectory while maintaining similar pressure levels across the target positions. Dynamic focusing was further demonstrated along more complex star- and heart-shaped trajectories comprising 68 and 59 target positions, respectively (Fig. 3h,i; Supplementary Fig. 25 and Supplementary Video 2). With the focal position updated every 8.2 ms, one complete traversal required approximately 560 ms for the star-shaped trajectory and 480 ms for the heart-shaped trajectory. These results demonstrate rapid and programmable movement of a localized focal point along complex trajectories.

STIF was also extended to multi-point focusing to generate multiple focal points simultaneously (Fig. 3j–l). For multi-point focusing, the desired pressure waveform for a single target was extended to a target vector containing the prescribed waveform for each target position. When identical chirp waveforms were assigned to all targets with the same temporal alignment, the acoustic fields generated at different target positions interfered with one another, producing pronounced sidelobes (Supplementary Fig. 26). To reduce this interference, we introduced temporal staggering by

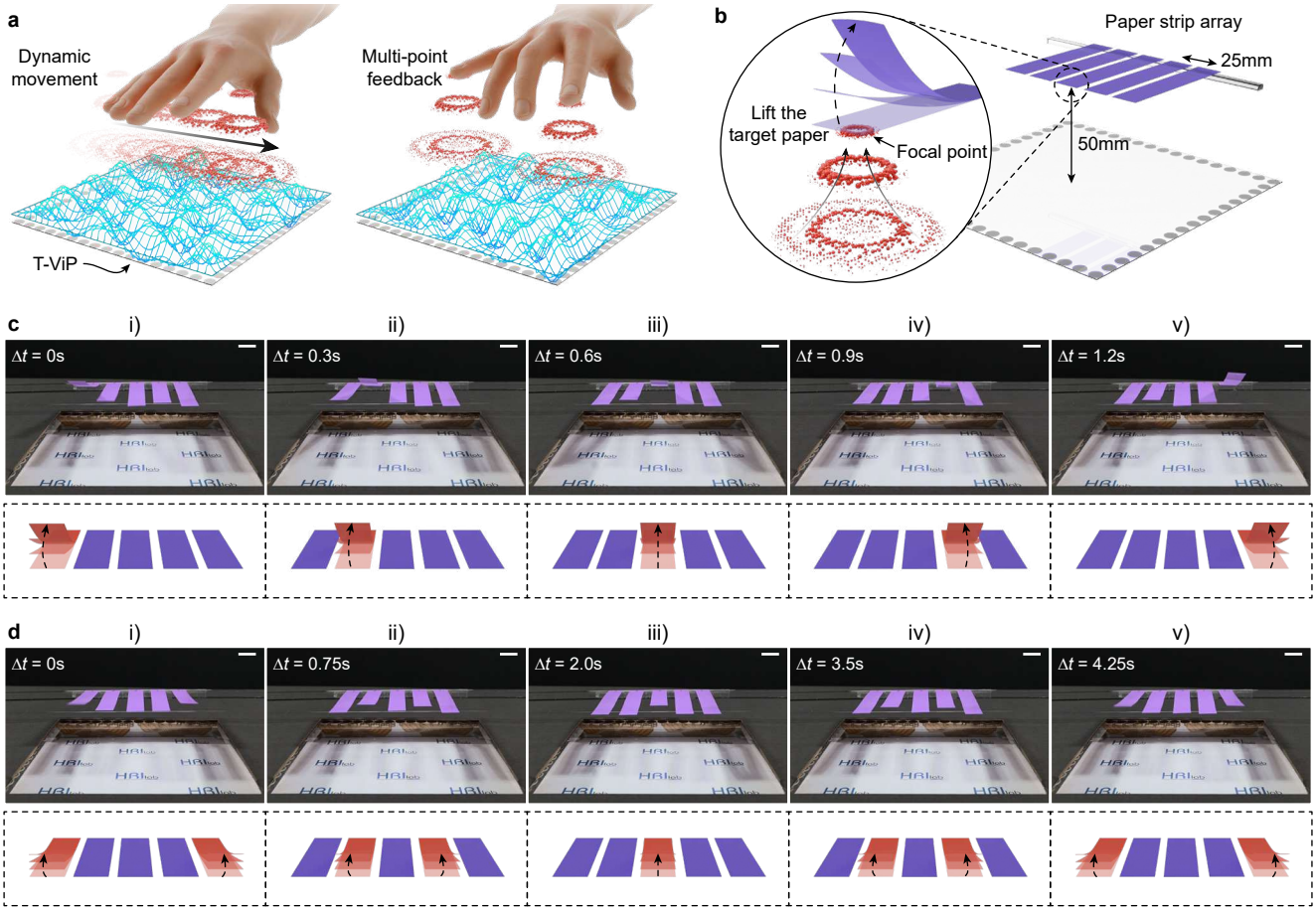


Fig. 4: **Physical interaction enabled by dynamic and multi-focal acoustic focusing. a** Schematic illustrations of dynamic and multi-point physical interaction generated above the T-ViP. **b** Experimental setup for visualizing acoustic radiation force. Paper strips spaced by 25 mm were positioned 50 mm above the panel, and the strip located at the focal point was selectively lifted. **c** Sequential images of dynamic physical interaction. A single focal point was moved across the paper-strip array, sequentially lifting the corresponding strips (Supplementary Video 3). **d** Sequential images of multi-point physical interaction. Two focal points simultaneously lifted two paper strips and were moved between the outer and inner positions (Supplementary Video 3).

assigning different time delays to the target waveforms while maintaining the same frequency band (Supplementary Note 6).

For two-point focusing, the acoustic field was measured at focal-point separations of 20 mm, 40 mm, 50 mm, and 80 mm. Two distinct local pressure maxima were generated near the prescribed target positions at each separation (Fig. 3k). Spatially separated pressure maxima were also obtained for configurations with three and four focal points (Fig. 3l). The RMS acoustic pressure at each focal point decreased as the number of simultaneous focal points increased. Nevertheless, distinct pressure maxima were maintained at the prescribed locations, demonstrating simultaneous multi-point acoustic focusing.

We then examined whether these dynamically moving and multiple focal points could produce spatially controlled physical interaction (Fig. 4). A paper strip positioned 50 mm above the panel was used to visualize the acoustic radiation force. When a focal point was generated at the position of the strip, the strip was lifted by the resulting acoustic radiation force, demonstrating localized non-contact physical interaction above the panel (Fig. 4a,b; Supplementary Video 3).

Dynamic and multi-point focusing were subsequently used to control this physical interaction over space. For dynamic interaction, paper strips were arranged at 25 mm intervals above the panel, and a single focal point was moved sequentially across the array (Fig. 4c). As the focal point moved from right to left and then back from left to right, the corresponding paper strips were lifted one after another (Supplementary Video 3). For multi-point interaction, two spatially separated focal points were generated simultaneously, allowing two paper strips to be lifted at the same time (Fig. 4d). The two focal points were then moved between the outer and inner positions, and the corresponding lifted positions followed their movement (Supplementary Video 3). These results demonstrate that the T-ViP can generate dynamic and multi-point physical interaction at programmable locations in the space above the panel.

Perceptual evaluation and visuo-physical integration

Beyond manipulating physical objects, we asked whether the localized acoustic fields generated by the T-ViP could also be perceived directly by users and spatially integrated with visual content. We first evaluated whether the focal point could provide perceivable mid-air haptic feedback through psychophysical experiments (Fig. 5a–e). Ten participants took part in the experiments, with the stimulated region of the palm positioned approximately 20 mm above the panel surface. To minimize auditory cues during actuation, participants wore noise-canceling headphones throughout the experiments (Fig. 5a).

We first measured the absolute detection threshold at different modulation frequencies to quantify the perceptibility of the localized acoustic field (Fig. 5b,c). In mid-air ultrasound haptics, acoustic radiation force is typically modulated within the vibrotactile frequency range to enhance tactile perception, and perceptual sensitivity can depend on modulation frequency [39]. We therefore varied the modulation frequency from 50 Hz to 300 Hz and measured the absolute detection threshold at each frequency. Ascending and descending staircase procedures were used to determine the threshold for each participant, which was then converted to the corresponding RMS acoustic pressure measured at the focal region. The detection threshold varied with modulation frequency, with the lowest mean threshold of 380.8 Pa measured at 150 Hz. These results demonstrate that the T-ViP can provide perceivable mid-air haptic feedback and that, among the tested conditions, tactile sensitivity was highest at a modulation frequency of 150 Hz.

We next evaluated whether focal-point movement could convey directional information in addition to localized tactile stimulation (Fig. 5d,e). A modulation frequency of 150 Hz, which produced the lowest detection threshold in the preceding experiment, was used to generate four directional patterns: up, down, left, and right. For each direction, the focal point was moved along a 100 mm path. Participants identified the four directions with a mean accuracy of 94.5 %, and responses in the confusion matrix were concentrated primarily along the intended directions. These results show that spatial movement of the focal point can convey directional information through mid-air haptic feedback.

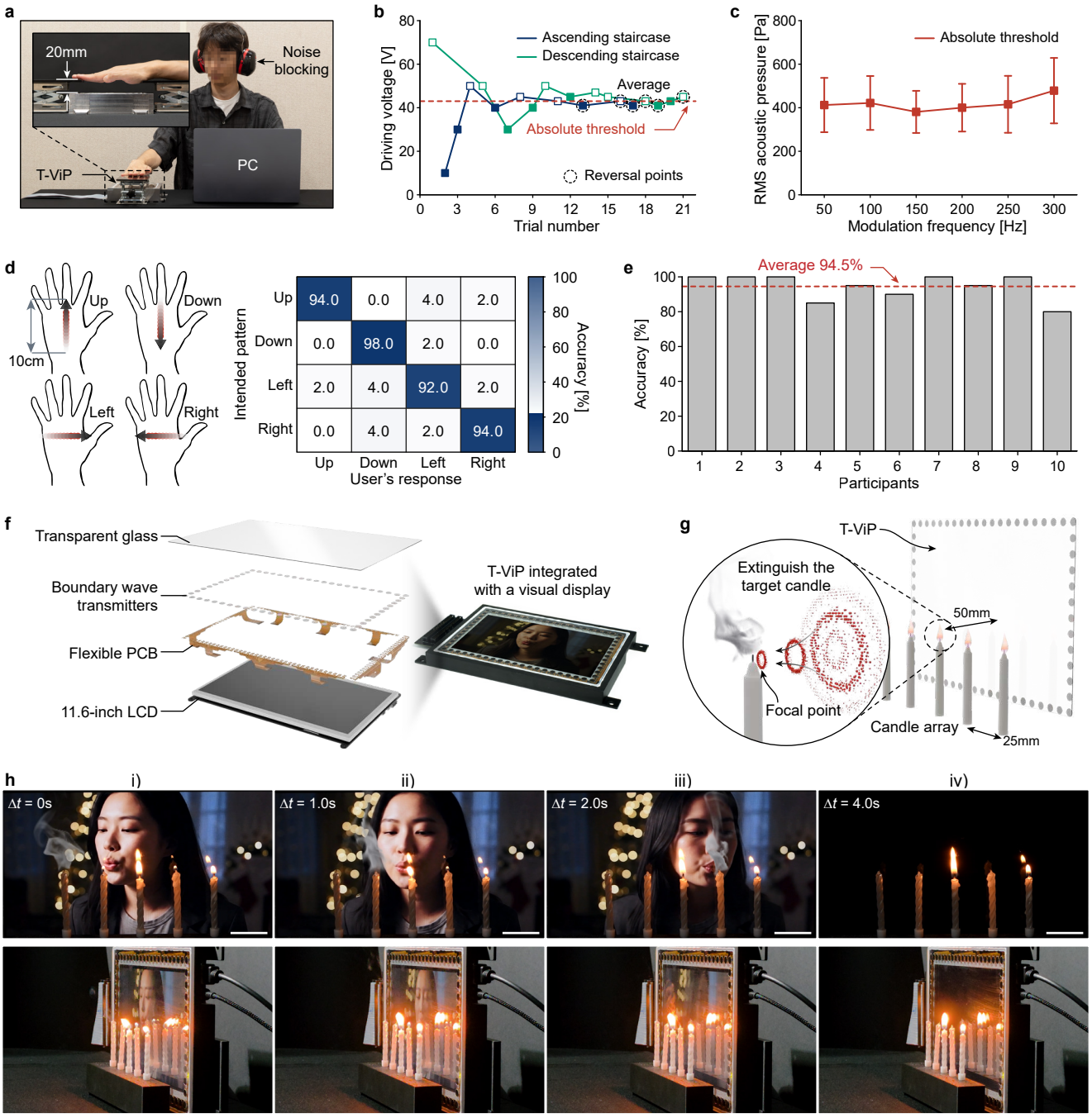


Fig. 5: **Perceptual evaluation and visuo-physical integration of the T-ViP. a** Experimental setup for evaluating mid-air haptic feedback generated by the T-ViP. Participants positioned a hand above the panel while auditory cues were blocked using headphones. **b** Representative ascending and descending staircase procedure used to determine the absolute detection threshold in terms of driving voltage. Reversal points from the two staircases were used to estimate the threshold. **c** Measured absolute detection threshold expressed as RMS acoustic pressure for modulation frequencies ranging from 50 Hz to 300 Hz. Data show the mean and variability across participants. **d** Evaluation of spatially dynamic haptic patterns corresponding to four directions (up, down, left and right), together with the confusion matrix between the intended patterns and participants' responses. **e** Direction-recognition accuracy for individual participants, with an average accuracy of 94.5 %. **f** Exploded schematic of the visuo-physical display integrating the T-ViP with an 11.6-inch LCD. **g** Experimental configuration for spatially aligning visual content and physical interaction. A focal point was generated at the position of a selected candle in front of the display. **h** Sequential images showing selective physical interaction with the target candle spatially aligned with the visual content. All white scale bars represent 25 mm.

Finally, we integrated the T-ViP with a visual display to spatially align visual content and off-surface physical interaction within the same three-dimensional space (Fig. 5f–h). The T-ViP, with 64 piezoelectric wave transmitters arranged along the panel boundary, was combined with an 11.6-inch LCD, allowing focal points to be generated in front of the display without obstructing the central viewing area (Fig. 5f; detailed structure in Supplementary Fig. 27). To demonstrate this spatial alignment, a candle array was positioned in front of the display, and a focal point was generated at positions corresponding to the visual content (Fig. 5g). By moving the focal point to a selected candle, the corresponding candle could be selectively blown out, demonstrating spatial alignment between the visual event and the physical interaction (Fig. 5h; Supplementary Video 1). These results show that the T-ViP can preserve the display viewing area while bringing visual content, mid-air haptic feedback, and localized physical interaction into a shared three-dimensional interaction space.

Conclusion

The central finding of this work is that a transparent display panel can serve not only as a medium for visual information, but also as a wave-propagation and radiation medium for generating spatially controlled physical effects in the surrounding three-dimensional space. Using piezoelectric wave transmitters arranged along the panel boundary and spatiotemporal inverse filtering, we controlled the complex wave propagation within the panel to form localized acoustic focal points at prescribed target positions without obstructing the viewing area. This approach enables spatially controlled mid-air interaction without distributing hundreds of acoustic emitters across the viewing aperture, as in conventional ultrasonic phased arrays.

The T-ViP generated focal points at different positions in three-dimensional space, moved them dynamically along prescribed trajectories, and formed multiple focal points simultaneously. The resulting acoustic fields exerted sufficient radiation force to produce localized physical interaction, enabling objects to be selectively lifted and moved through dynamic and multi-point focusing. When modulated within the tactile frequency range, the same focal fields provided perceivable mid-air haptic feedback and conveyed directional information through focal-point movement. Integration of the T-ViP with a visual display further demonstrated that visual events and physical effects can be spatially aligned within the same interaction space while preserving the central viewing area.

Building on these capabilities, the T-ViP provides a route toward a unified display architecture in which visual content and spatial physical interaction are realized through the same transparent interface. Further advances in acoustic output and focusing efficiency, and focal point update rate could enable physical interactions across larger volumes, with greater force, faster motion, and more complex spatial patterns. Ultimately, this architecture could extend the interaction space of a display from the screen surface into three-dimensional space above it, allowing digital content to generate physical effects at the same spatial locations where it is visually presented.

Methods

Fabrication of the T-ViP

The T-ViP consisted of a 150 mm × 150 mm × 0.4 mm borosilicate glass panel (D263, SCHOTT) with 64 piezoelectric wave transmitters arranged along its boundary. Each wave transmitter (B59070Z0285D120, TDK) was 7 mm in diameter and 0.195 mm thick and was bonded to the glass panel at a center-to-center spacing of 8 mm using an epoxy adhesive (Loctite EA0151, Henkel). The wave transmitters were connected to the external driving electronics through flexible printed circuit boards (FPCBs), and the assembled panel was mounted in a stainless-steel housing that supported the panel edges. Further details of the fabrication and assembly are provided in Supplementary Note 5 and Supplementary Figs. 14 and 15.

Numerical modeling of the T-ViP

The electromechanical response of the T-ViP was modeled using COMSOL Multiphysics 6.0, and the resulting panel vibration was used to calculate the radiated acoustic field in MATLAB R2025b. The finite-element model reproduced the geometry and mounting configuration of the fabricated T-ViP, including the glass panel, piezoelectric wave transmitters, and epoxy bonding layers. A 2.5 mm-wide region along the panel boundary was fixed to represent the mounting condition. The material properties used in the simulation are summarized in Supplementary Table 1.

Frequency-domain analysis was performed from 30 kHz to 80 kHz in increments of 100 Hz. Each wave transmitter was sequentially driven with a sinusoidal voltage of 100 V peak while the remaining transmitters were grounded. The complex out-of-plane velocity was evaluated at 5625 uniformly distributed points over the panel surface with a spatial interval of 2 mm. The voltage-normalized responses for all transmitter inputs were assembled into the electromechanical transfer matrix $\mathbf{G}(j\omega)$.

The acoustic pressure radiated from the vibrating panel was calculated from the surface velocity using the Rayleigh integral, yielding the mechanoacoustic transfer matrix $\mathbf{S}(j\omega)$. The overall EATF, represented by $\mathbf{H}(j\omega)$, was then obtained according to Eq. (1). Further details of the electromechanical–acoustic formulation and numerical implementation are provided in Supplementary Notes 1 and 2.

Multichannel driving and data acquisition system

A multichannel driving and data acquisition system was developed to independently drive the 64 piezoelectric wave transmitters and synchronously measure the driving signals and system responses. Eight 8-channel HV pulser modules provided a total of 64 independent driving channels. Each channel consisted of a high-speed gate driver (MD1822K6-G, Microchip) and an HV MOSFET array (TC8220K6-G, Microchip) and generated three-level voltage outputs of $+V_{\max}$, 0, and $-V_{\max}$, with $V_{\max}$ up to 100 V. A dead time of approximately 180 ns was introduced between switching transitions to prevent simultaneous conduction of the high-side and low-side MOSFETs. Driving waveforms were generated on a host PC, transmitted to the control system via UDP communication, and synchronously reproduced across all 64 channels.

For synchronized measurement of the driving voltages and the corresponding panel-vibration and acoustic pressure responses, three 8-channel simultaneous-sampling ADC modules (ADS8568, Texas Instruments) were used, enabling simultaneous acquisition of up to 24 analog signals. The driving-voltage signals and response signals from a laser Doppler vibrometer (LDV) and microphone were acquired using a common timing reference. The measured data were collected by the control system and transmitted to the host PC via UDP communication.

The overall driving and acquisition process was controlled using a Zynq system-on-chip (XC7Z020-1CLG484C, AMD). The programmable logic performed timing-critical operations, including pulse generation, ADC sampling, and high-speed data transfer, whereas the processing system managed UDP communication with the host PC and coordinated overall system operation. Further details of the pulser circuits, signal-conditioning electronics, and control architecture are provided in Supplementary Note 5 and Supplementary Figs. 16–20.

Experimental identification of the EATF

The electromechanical transfer response of the fabricated T-ViP was experimentally identified by sequentially exciting each piezoelectric wave transmitter with a broadband pseudorandom binary sequence and measuring the resulting panel surface velocity using an LDV. The driving voltage and corresponding surface-velocity response were acquired synchronously for transfer-function estimation. The panel surface was scanned over a 75×75 grid with a spatial interval of 2 mm, and the transfer functions between the measured input voltages and surface velocities were estimated over the 30 kHz to 80 kHz frequency range.

The resulting voltage-to-velocity responses were assembled into the experimental electromechanical transfer matrix $\mathbf{G}(j\omega)$. This matrix was combined with the mechanoacoustic model to obtain the EATF used for STIF. Further details of the excitation signal and transfer-function estimation are provided in Supplementary Note 6.

Spatiotemporal inverse filtering

STIF was used to generate a prescribed broadband acoustic pressure waveform at a target position in space [35]. A linear chirp spanning the 30 kHz to 80 kHz operating frequency range was prescribed as the desired pressure waveform. For single-point focusing, the wave-transmitter voltage spectrum required to reproduce the desired pressure spectrum $p_f^{\mathrm{des}}(j\omega)$ was calculated from the EATF using a regularized minimum-norm solution. Tikhonov regularization was introduced to improve numerical stability and limit excessive input amplitudes at frequencies where the transfer response was weak. The resulting optimal voltage spectrum was calculated as

$$\mathbf{V}^{\mathrm{opt}}(j\omega) = \mathbf{H}^{\mathrm{H}}(j\omega)\frac{p_f^{\mathrm{des}}(j\omega)}{\|\mathbf{H}(j\omega)\|_2^2 + \lambda_{\mathrm{reg}}} \tag{3}$$

where $\mathbf{H}^{\mathrm{H}}(j\omega)$ denotes the Hermitian transpose of $\mathbf{H}(j\omega)$. The regularization parameter was defined as

$$\lambda_{\mathrm{reg}} = \epsilon_{\mathrm{reg}} \max_{\omega} \|\mathbf{H}(j\omega)\|_2^2 \tag{4}$$

with $\epsilon_{\mathrm{reg}} = 1 \times 10^{-3}$. The calculated voltage spectra were transformed into time-domain waveforms using an IFFT. Each waveform was then independently normalized by its maximum absolute amplitude before being quantized into three-level driving signals of $+V_{\max}$, 0, and $-V_{\max}$ for implementation with the multichannel HV pulser. The normalized quantization threshold was set to $V_{\mathrm{th}} = 0.05$.

For multi-point focusing, STIF was extended to generate focal points simultaneously at multiple target positions. When identical chirp waveforms were prescribed at multiple targets with the same temporal alignment, interference between the corresponding acoustic fields produced pronounced sidelobes. To reduce this inter-target interference, a different temporal delay was assigned to the desired chirp waveform at each target position. For M focal points, the delay applied to the mth target waveform was defined as

$$\Delta t_m = (m-1)\frac{T_{\mathrm{c}}}{M}, \qquad m = 1, 2, \ldots, M \tag{5}$$

where T_c denotes the chirp duration. The target chirps therefore retained the same frequency band while being uniformly time-staggered over one chirp duration. For the numerical simulations, the chirp duration was set to $T_c = 10\,\mathrm{ms}$ based on a frequency resolution of $\Delta f = 100\,\mathrm{Hz}$. For the experimental implementation, $T_c = 8.2\,\mathrm{ms}$ was used based on the frequency resolution of the experimentally identified transfer response, $\Delta f \approx 122.1\,\mathrm{Hz}$. Further details of the STIF formulation and signal-generation procedure are provided in Supplementary Note 4, while the experimental implementation and multi-focal STIF formulation are described in Supplementary Note 6.

Acoustic field measurement

The airborne acoustic field generated by the T-ViP was measured using a microphone (46DD, GRAS) and a signal conditioner (442C04, PCB Piezotronics). The microphone was mounted on a motorized XYZ stage and scanned over an x–y plane parallel to the glass panel. The acoustic field was measured over a range of $\pm 75\,\mathrm{mm}$ in both the x- and y-directions relative to the panel center, with a spatial interval of 1.5 mm. At each measurement position, the time-domain acoustic pressure signal was acquired, and the RMS acoustic pressure was calculated to characterize the spatial distribution of the acoustic field.

Single-point focusing was evaluated at the same nine lateral target positions used in the numerical simulations, with focal heights of 25 mm, 50 mm, and 75 mm. The RMS acoustic pressure at the focal point was also measured at different focal heights and driving voltages to characterize the acoustic output of the T-ViP.

To evaluate the temporal evolution of dynamic focusing, a moving time window was applied to the time-domain acoustic pressure signal acquired at each spatial measurement position. The transient RMS acoustic pressure at time t was calculated as

$$p_{\mathrm{rms},\Delta T}(t) = \sqrt{\frac{1}{\Delta T}\int_{t}^{t+\Delta T} p^2(\tau)\, d\tau} \tag{6}$$

where ΔT denotes the time window used to calculate the RMS acoustic pressure. The same calculation was applied at each spatial measurement position to reconstruct the transient RMS acoustic field over time. Further details of the time-window conditions are provided in Supplementary Note 6.

Psychophysical experiments

Psychophysical experiments were conducted to evaluate the perception of mid-air haptic feedback generated by the T-ViP. Ten participants (four females and six males) took part in the experiments, and all reported no history of sensory impairment in their hands. During the experiments, each participant's hand was positioned above the glass panel such that the target palmar region was located at the acoustic focal point, 20 mm above the panel. A height-adjustable jig and an acrylic mount were used to maintain a consistent hand position and height. To mask auditory cues generated by the actuation, participants wore noise-canceling headphones playing pink noise and entered their responses directly through a PC. All experimental protocols were approved by the Korea Advanced Institute of Science and Technology Institutional Review Board (KAIST IRB-2026-161).

The absolute detection threshold was measured using an interleaved 1-up/1-down adaptive staircase procedure. Two staircase sequences were alternated throughout the experiment. Sequence A started at a driving voltage of 70 V, which was decreased when the participant detected the stimulus, whereas sequence B started at 10 V, which was increased when the stimulus was not detected. The step size was progressively reduced to 20 V, 10 V, 5 V, and 2 V as reversals accumulated. Each staircase sequence was terminated after three reversals, and the absolute detection threshold at each modulation frequency was defined as the average of the six reversal voltages obtained from the two sequences. Each stimulus was presented for 1 s, followed by an inter-stimulus interval of 5 s.

Detection thresholds were measured at modulation frequencies ranging from 50 Hz to 300 Hz. The threshold voltage obtained at each modulation frequency was converted to the corresponding RMS acoustic pressure threshold using acoustic pressure measurements at the focal region.

A separate directional-pattern identification experiment was conducted to evaluate the perception of dynamically moving mid-air haptic feedback. The modulation frequency was set to 150 Hz, which produced the lowest mean detection threshold in the preceding experiment, and four movement directions were tested: up, down, left, and right. The same ten participants took part in this experiment, with the center of the palm aligned with the geometric center of the stimulation path at a height of 20 mm above the glass panel.

Each directional path had a total length of 100 mm and consisted of 26 sequential focal points. Ultrasonic stimuli modulated at 150 Hz were sequentially generated at these focal positions to produce a haptic pattern moving along the prescribed direction. Each directional pattern was presented for 0.7 s. Each participant completed 20 trials, with each of the four directions presented five times in randomized order. After each trial, participants selected one of the four directions corresponding to the perceived direction of motion.

Visuo-physical display integration

To spatially align visual content and off-surface physical interaction within the same display area, the T-ViP was integrated with an 11.6-inch LCD. For this demonstration, a larger T-ViP was fabricated using a 260 mm × 171 mm transparent borosilicate glass panel with 64 piezoelectric wave transmitters arranged along its boundary. The wave transmitters were connected to the external multichannel driving electronics through FPCBs and a signal-connector PCB, while the LCD was positioned directly behind the glass panel. Visual content was provided from an external PC via HDMI, and the acoustic focal position was controlled independently through the multichannel driving system. The detailed configuration of the integrated visuo-physical display is shown in Supplementary Fig. 27.

For the visuo-physical demonstration, a candle array was positioned 50 mm in front of the glass panel, with adjacent candles separated by 25 mm. The focal coordinates were assigned to the candle positions corresponding to the visual targets, and the focal point was moved among these positions in coordination with the displayed visual events. The maximum driving voltage was set to 100 V throughout the demonstration. This configuration enabled spatial alignment between visual content and localized physical interaction in front of the display.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The codes that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgement

This work was supported by Samsung Science and Technology Foundation under Project Number SRFC-IT2102-04.

Author contributions

J.P. and K.-U.K. conceived the research and developed the underlying ideas. J.P. developed the analytical model of the T-ViP, performed the numerical analyses, and designed the overall system and control methodology. J.P. and S.P. fabricated the devices, performed the acoustic measurements and psychophysical experiments, and analyzed the resulting data. J.P., S.P., J.H. and J.-S.K. conducted the physical interaction demonstrations and prepared the visualizations. K.-U.K. supervised the research and provided guidance throughout the study. J.P. and K.-U.K. wrote the manuscript, with review and editing by all other authors.

Additional information

Supplementary information

The online version contains supplementary material available at https://doi.org/xx.xxxx/xxx.

Competing interests

The authors declare no competing interests.

Reprints and permission

Reprints and permission information is available online at http://npg.nature.com/ reprintsandpermissions/

References

[1] Moon, S. *et al.* Switchable 2d–3d display through a metasurface lenticular lens. *Nature* **652**, 1181–1187 (2026).

[2] Ma, W., Zhao, Z., Zhao, C., Ouyang, W. & Zhong, H.-S. Glasses-free 3d display with ultrawide viewing range using deep learning. *Nature* **648**, 76–83 (2025).

[3] Wang, D. *et al.* Color liquid crystal grating based color holographic 3d display system with large viewing angle. *Light Sci. Appl.* **13**, 16 (2024).

[4] An, J. *et al.* Slim-panel holographic video display. *Nat. Commun.* **11**, 5568 (2020).

[5] Yang, T.-H. *et al.* Recent advances and opportunities of active materials for haptic technologies in virtual and augmented reality. *Adv. Funct. Mater.* **31**, 2008831 (2021).

[6] Basdogan, C., Giraud, F., Levesque, V. & Choi, S. A review of surface haptics: Enabling tactile effects on touch surfaces. *IEEE Trans. Haptics* **13**, 450–470 (2020).

[7] Hudin, C., Lozada, J. & Hayward, V. Localized tactile feedback on a transparent surface through time-reversal wave focusing. *IEEE Trans. Haptics* **8**, 188–198 (2015).

[8] Woo, J.-H. & Ih, J.-G. Vibration rendering on a thin plate with actuator array at the periphery. *J. Sound Vib.* **349**, 150–162 (2015).

[9] Park, J., Han, J., Yang, T.-H. & Kyung, K.-U. Programmable localized surface haptic feedback achieved through seamless transparent panel. *IEEE Trans. Ind. Electron.* **72**, 14236–14247 (2025).

[10] Bau, O., Poupyrev, I., Israr, A. & Harrison, C. Teslatouch: Electrovibration for touch surfaces. In *Proc. 23nd Annual ACM Symposium on User Interface Software and Technology*, 283—292 (ACM, 2010).

[11] Park, B. *et al.* Flexible and transparent electrovibration-based haptic display with low driving voltage. *ACS Appl. Mater. Interfaces* **16**, 55864–55872 (2024).

[12] Biet, M., Giraud, F. & Lemaire-Semail, B. Squeeze film effect for the design of an ultrasonic tactile plate. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **54**, 2678–2688 (2007).

[13] Song, L. *et al.* Highly transparent lead-free piezoelectric haptic device. *Commun. Mater.* **6**, 91 (2025).

[14] Shan, B. *et al.* Fully transparent haptic interface for high-resolution tactile feedback on touchscreens. *Adv. Sci.* **12**, e11874 (2025).

[15] Firouzeh, A., Mizutani, A., Groten, J., Zirkl, M. & Shea, H. Poptouch: A submillimeter thick dynamically reconfigured haptic interface with pressable buttons. *Adv. Mater.* **36**, 2307636 (2024).

[16] Rakkolainen, I., Freeman, E., Sand, A., Raisamo, R. & Brewster, S. A survey of mid-air ultrasound haptics and its applications. *IEEE Trans. Haptics* **14**, 2–19 (2020).

[17] Georgiou, O., Frier, W., Freeman, E., Pacchierotti, C. & Hoshi, T. *Ultrasound mid-air haptics for touchless interfaces* (Springer, 2022).

[18] Hamilton, M. F. & Blackstock, D. T. *Nonlinear acoustics* (Academic Press, San Diego, 1998).

[19] Marzo, A. *et al.* Holographic acoustic elements for manipulation of levitated objects. *Nat. Commun.* **6**, 8661 (2015).

[20] Hirayama, R., Martinez Plasencia, D., Masuda, N. & Subramanian, S. A volumetric display for visual, tactile and audio presentation using acoustic trapping. *Nature* **575**, 320–323 (2019).

[21] Iwamoto, T., Tatezono, M. & Shinoda, H. Non-contact method for producing tactile sensation using airborne ultrasound. In *Haptics: Perception, Devices and Scenarios. EuroHaptics*, 504–513 (Springer, 2008).

[22] Reardon, G., Dandu, B., Shao, Y. & Visell, Y. Shear shock waves mediate haptic holography via focused ultrasound. *Sci. Adv.* **9**, eadf2037 (2023).

[23] Hoshi, T., Takahashi, M., Iwamoto, T. & Shinoda, H. Noncontact tactile display based on radiation pressure of airborne ultrasound. *IEEE Trans. Haptics* **3**, 155–165 (2010).

[24] Carter, T., Seah, S. A., Long, B., Drinkwater, B. & Subramanian, S. Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces. In *Proc. 26th Annual ACM Symposium on User Interface Software and Technology*, 505–514 (ACM, 2013).

[25] Long, B., Seah, S. A., Carter, T. & Subramanian, S. Rendering volumetric haptic shapes in mid-air using ultrasound. *ACM Trans. Graph.* **33**, 1–10 (2014).

[26] Shen, Z., Vasudevan, M. K., Kučera, J., Obrist, M. & Martinez Plasencia, D. Multi-point stm: Effects of drawing speed and number of focal points on users' responses using ultrasonic mid-air haptics. In *Proc. 2023 CHI Conference on Human Factors in Computing Systems*, 1–11 (ACM, 2023).

[27] Suzuki, S., Inoue, S., Fujiwara, M., Makino, Y. & Shinoda, H. Autd3: Scalable airborne ultrasound tactile display. *IEEE Trans. Haptics* **14**, 740–749 (2021).

[28] Young, G. *et al.* Designing mid-air haptic gesture controlled user interfaces for cars. *Proc. ACM. Hum. Comput. Interact.* **4**, 1–23 (2020).

[29] Monnai, Y. *et al.* Haptomime: mid-air haptic interaction with a floating virtual screen. In *Pro. 27th annual ACM symposium on User interface software and technology*, 663–667 (ACM, 2014).

[30] Sand, A., Rakkolainen, I., Isokoski, P., Raisamo, R. & Palovuori, K. Light-weight immaterial particle displays with mid-air tactile feedback. In *2015 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, 1–5 (2015).

[31] Halbach, A. *et al.* Display compatible pmut array for mid-air haptic feedback. In *2019 20th International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSDUCERS & EUROSENSORS XXXIII)*, 158–161 (IEEE, 2019).

[32] Chare, C. *et al.* Polymer-based piezoelectric ultrasound transducer arrays on glass demonstrating mid-air applications. In *2020 IEEE International Ultrasonics Symposium (IUS)*, 1–4 (IEEE, 2020).

[33] Chare, C. *et al.* High performance large-area polymeric pmut phased arrays in air. In *2022 IEEE International Ultrasonics Symposium (IUS)*, 1–4 (IEEE, 2022).

[34] Laughlin, E. A. *Transparent Capacitive and Piezoelectric Micromachined Ultrasonic Transducers for Tactile Feedback with 3D Displays.* Master's thesis, Brigham Young University (2021).

[35] Tanter, M., Aubry, J.-F., Gerber, J., Thomas, J.-L. & Fink, M. Optimal focusing by spatio-temporal inverse filter. i. basic principles. *J. Acoust. Soc. Am.* **110**, 37–47 (2001).

[36] Hudin, C. & Dhiab, A. B. Spatio-temporal control of vibrotactile stimulation on oled displays with inverse filtering. *IEEE Trans. Haptics* **19**, 746–761 (2026).

[37] Fink, M. Time reversal of ultrasonic fields. i. basic principles. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**, 555–566 (1992).

[38] Derode, A., Tourin, A. & Fink, M. Ultrasonic pulse compression with one-bit time reversal through multiple scattering. *J. Appl. Phys.* **85**, 6343–6352 (1999).

[39] Raza, A., Hassan, W., Ogay, T., Hwang, I. & Jeon, S. Perceptually correct haptic rendering in mid-air using ultrasound phased array. *IEEE Trans. Ind. Electron.* **67**, 736–745 (2019).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryVideo2.mp4
- SupplementaryInformation.pdf
- SupplementaryVideo1.mp4
- SupplementaryVideo3.mp4