A volumetric display for visual, tactile and audio presentation using acoustic trapping

Ryuji Hirayama, Diego Martinez Plasencia, Nobuyuki Masuda, Sriram Subramanian

Publication date
09-06-2023

Document Version
Accepted version

Citation for this work (American Psychological Association 7th edition)
Hirayama, R., Martinez Plasencia, D., Masuda, N., & Subramanian, S. (2019). *A volumetric display for visual, tactile and audio presentation using acoustic trapping* (Version 1). University of Sussex. https://hdl.handle.net/10779/uos.23472326.v1

Published in
Nature

Link to external publisher version
https://doi.org/10.1038/s41586-019-1739-5

A volumetric display for visual, tactile and audio presentation using acoustic trapping

Ryuji Hirayama, School of Engineering and Informatics, University of Sussex (UK)
Diego Martinez Plasencia, School of Engineering and Informatics, University of Sussex (UK)
Nobuyuki Masuda, Department of Applied Electronics, Tokyo University of Science (Japan)
Sriram Subramanian, School of Engineering and Informatics, University of Sussex (UK)

Science-fiction movies such as Star Wars portray volumetric systems that not only provide visual but also tactile and audible 3D content. Displays, based on swept volume surfaces,[1,2] holography[3], optophoretics[4], plasmonics,[5] or lenticular lenslets[6], can create 3D visual content without the need for glasses or additional instrumentation. However, they are slow, have limited persistence of vision (POV) capabilities, and, most critically, rely on operating principles that cannot also produce tactile and auditive content.
Here, we present for the first time a Multimodal Acoustic Trap Display (MATD): a mid-air volumetric display that can simultaneously deliver visual, auditory, and tactile content, using acoustophoresis as the single operating principle. Our system acoustically traps a particle and illuminates it with red, green, and blue light to control its colour as it quickly scans through our display volume. Using time multiplexing with a secondary trap, amplitude modulation and phase minimization, the MATD delivers simultaneous auditive and tactile content. The system demonstrates particle speeds of up to 8.75m/s and 3.75m/s in the vertical and horizontal directions respectively, offering particle manipulation capabilities superior to other optical or acoustic approaches demonstrated to date. Beyond enabling simultaneous visual, tactile and auditive content, our approach and techniques offer opportunities for non-contact, high-speed manipulation of matter, with applications in computational fabrication[7] and biomedicine[8].

Holographic and lenslet displays rely on a 2D display modulator, constraining the visibility of 3D content to the volume between the observer's eyes and the display surface (i.e. direct line of sight). Volumetric approaches are based on light scattering, emitting, or absorbing surfaces[9]. They offer unconstrained visibility anywhere around the display and can be created using rotating surfaces (active[1] or passive[2]), plasmonics[5,10], air displays,[11] and photophoretic traps[4]. However, none of these approaches rely on operating principles that can also recreate touch and sound. Acoustic levitation displays to date[12–14] have only demonstrated control of a reduced number of points at reduced speeds and do not engage with touch or audible sound.

In contrast, our MATD allows for a volumetric display where, for the first time, users can simultaneously see visual content in mid-air from any point around the display volume and receive auditive and tactile feedback from that volume (as shown in Video SV1).

Our system is based on Acoustic Tweezers, which use ultrasound radiation forces to trap particles[14–17]. Trapping has been demonstrated in media such as air [12,13,18,19] and water [16], and for particle sizes ranging from the micrometre to the centimetre scale. For spherical particles significantly smaller than the wavelength and operating in the far-field regime (i.e. like those used by our MATD), the forces exerted are governed by the gradient of the Gor'kov potential[17]. Several trap morphologies have been demonstrated to date, including twin traps, vortex traps, and bottle beams,[20–22] which can all now be analytically computed with efficiency[22].

Our device (summarized in Figure 1a and detailed in Methods) exploits this by analytically computing a single twin trap or focusing point at a hardware level on an FPGA (Field Programmable Gate Array). This allows for position and amplitude updates of the trap in a volume of 10x10x10cm, at a rate limited only by the transducer frequency. In contrast, Spatial Light Modulators are limited to update rates of hundreds of Hz, while galvanometers are usually limited to ~20kHz. Existing acoustic modulators are limited to hundreds of Hzs[14] and displacement speeds well below 1m/s. Our current MATD implementation enables update rates of up to 40kHz and particle displacement speeds of up to 8.75 m/s and 3.75 m/s in the vertical and horizontal directions respectively. Exploiting such high modulation rates and the mechanical nature of ultrasound, our control techniques (described below and detailed in Methods) allow delivery of tactile and auditive content in addition to 3D POV content.

To create visual content, we levitate a 1 mm radius, white, expanded polystyrene (EPS) particle, as a good approximation to a Lambertian surface. Such particle allows for predictable models of acoustic trapping forces, as well as a simple analytical model to describe perceived colour under controlled illumination (see Methods 2.3). The hardware-embedded computation of the twin trap (see Methods 2.2) provides controlled and fast levitation of our scanning particle, and is synchronized with a diffuse illumination module (RGB LEDs). This allows for a POV display with accurate control of the perceived colour (gamma corrected 2.2), able to deliver 2D or 3D vector contents by POV (Figures 1b, 1c and 1e) or fully rasterized contents (Figure 1d, exposure time 20s), even under conventional indoor illumination conditions (see Video SV4).

Our tests (see Methods 4.2 and 4.3) revealed high scanning speeds and accelerations, well above optical[4] or acoustic[14] setups demonstrated to date. The most critical display parameters are summarized in Table 1, according to the MATD's various modes of operation: single particle with no amplitude modulation (visual content only), single particle with minimum amplitude (worst case displaying visual and audio content), and time multiplexed dual trap with minimum amplitude (worst case delivering all visual, audio and tactile content). Trapping forces and achievable speeds and accelerations vary with the direction of motion of the particle (i.e. highest in the vertical direction). Table 1 provides maximum displacement parameters along the horizontal direction (i.e. worst case with weaker trapping forces), as conservative reference values that allow content reproduction independently of the particle direction.

Parameters in Table 1 are used to compute and plan paths to create POV content visible to the naked eye. Human eyes can integrate different light stimuli under a single percept (i.e. a single shape/geometry) during reduced periods of time (0.1s usually accepted as a conservative estimation, even in bright environments [23]), and thus, our particle needs to scan the content in less than this time (0.1s). Our parameters allow us to determine feasible paths (particle speed, acceleration and curvature within the limits identified), which can be revealed in less than 0.1s exploiting only a fraction of the display's capabilities. The example letter in Figure 1b (traced at 12.5Hz, 1x2cm) requires particle speeds of up to 0.8m/s, while the face and 3D torus knot in Figures 1c (10Hz, 1.8cm diameter) and 1e (10Hz, 2cm side) require speeds of 1.3m/s. Our volumetric contents showed no significant flicker and good colour reproduction, independently of viewer's location (Figures 3a and 3b). Figure 2a shows examples of colour tests performed with vector images (numbers, as in a seven-segment display) and good colour saturation. Brighter images can be obtained by adding extra illumination modules or more powerful LEDs (details in Methods 2.3).

Figure 2b shows the MATD's ability to create additive and grayscale colours whereas Figures 1d, 2c, and 3c show examples of raster colour content in 2D and 3D similar to those created by Smaley et al.[4], using particle speeds of up to 0.6, 0.2 and 0.9m/s, respectively. The effects of particle scattering properties (i.e. perceived colour around it), particle speed (i.e. illuminance affected by path length) and human response (i.e. non-linear luminance response) must be considered for accurate colour reproduction (see Methods 2.3).

Mid-air tactile feedback at controlled locations (e.g. user's hand) is created by using a secondary focusing trap and custom multiplexing policy (*position* but not *amplitude* multiplexing with phase difference minimization; details in Methods 3.1). Well-differentiated tactile feedback was delivered using only a 25% duty cycle for tactile content. Thus, 75% of the cycles could still be used to position the primary trap, and the tactile content results in a minimum loss of scanning speed. For our experiments, we chose a 250Hz modulation frequency, avoiding the 2kHz–5kHz primary range of human auditive perception[24] (minimize parasitic noise), but remaining well within the optimum perceptual threshold of skin Lamellar corpuscles for vibration[25]. The 10kHz update rate for tactile stimulation is sufficient for spatio-temporal multiplexing strategies to maximize fidelity of mid-air tactile content[26]. Our results (see Methods 4.6) show accurate positioning and focusing of the tactile points and sound pressure levels of >150dB, well above the threshold of 72dB levels required for tactile stimulation[27] (illustrated in Video SV5).

Audible sound is created by ultrasound demodulation using upper sideband amplitude modulation[28] of the traps. Our sampling at 40kHz encodes most of the auditive spectrum (44.1kHz), and the high power transducer array produces audible sound even from a relatively small modulation index ($a = 0.2$), while still modulating particle positions and tactile points at the 40kHz rate. Figure 2a shows three examples of visual content with simultaneous audible content of 60dB. For simultaneous auditive and tactile stimulation, we combine the 40kHz multifrequency audio signal with the tactile modulation signal (250Hz), maintaining the sampling frequency of the individual signals and reducing losses in audio quality (Video SV1). The MATD supports two modes for audio generation (see Methods 4.5). The first mode uses the trapped particle as a scattering media implicitly providing spatialized audio[29] (i.e. sound coming from the content displayed), but our experience indicates such directional cues are weak (most sound coming from the centre of our working volume). The second mode uses the secondary trap to steer sound towards the user, resulting into a stronger directional component and higher sound levels. However, the use of directional audio currently comes at the expense of not simultaneously delivering tactile feedback (simultaneous visual, tactile and directional audio would require multiplexing of three traps, one for each modality).

Our current instantiation of the MATD was created using low-cost, commercially available components, making it easy to reproduce but also introducing limitations. Our tests were performed at the transducers' voltage allowing for continued usage (12Vpp). Tests at higher voltages (15Vpp, duration less than one hour) indicate that increasing the transducer's power can result in better performance parameters (e.g. max horizontal speed 4m/s) and more complex content. Increased power would also allow operation of the MATD at a 50% duty cycle, further reducing audio artefacts (see Figure S7d). Similarly, transducers operating at higher frequencies (i.e. 80kHz) can also improve audio quality and, combined with a reduced transducer pitch, would improve the spatial resolution of the levitation traps (more accurate paths of the scanning particle).

The MATD has demonstrated the possibility to manipulate particles by retaining them in a dynamic equilibrium (rather than a static one, as most other levitation approaches, see Methods 4.2), enabling the high accelerations and speeds observed. The use of models accurately predicting the dynamics of the particle (i.e. in terms of acoustic forces, drag, gravity and centrifugal forces, but also considering interference from secondary traps and transient effects in the transducers' phase updates) would allow for better exploitation of the observed maximum speeds and accelerations, enabling larger and more complex visual content. Alternatively, they could instead allow for a more efficient use of the acoustic pressure, providing similar speeds and accelerations to the ones provided by our current MATD, but allocating a lower duty cycle for the primary trap. This power could then be dedicated for stronger tactile content or to support more simultaneous traps (e.g. the three traps required for the simultaneous visual, tactile and directional audio scenario).

More advanced illumination approaches (e.g. using galvanometers[4] or beam steering mechanisms[11]) would allow for focused light and brighter displays. The use of several illumination modules around the display would allow for more control on the visual properties of the content displayed. For instance, four illumination modules, one at each corner of the MATD, would allow us to only illuminate the outside part of the globe in Figure 3c. The hidden parts of the globe would only be minimally visible, independently of the user location.

Combining a denser illumination array (e.g. a ring of light sources) and the predictable light scattering pattern of our particle, the final scattered field from the particle can be computed as the linear combination of the scattered field from each light source. This could be used, for instance, to create visual content approximating various material properties (e.g. make content look metallic or matte), simulating different lighting conditions or even delivering different contents in different viewing directions.

The presence of the user's hands can distort the acoustic field due to scattering from the hand's surface. The power and top-down arrangement of our array allow stable operation as the user's hand approaches from the sides or front (see Video SV4). Placing the hand below or above the location of the primary trap (occluding one array) is much more likely to produce failures (i.e. scanning particle being dropped). Close proximity of the secondary trap to the primary trap can also distort the trapping of the scanning particle. We successfully reproduced curvature tests at maximum speed with the tactile point at 2cm from the circle, suggesting that while tactile feedback cannot be reproduced directly on top of visual content (avoiding scattering or directly colliding with the scanning particle), tactile feedback can be created in close proximity to it.

Our study demonstrates an approach to create volumetric POV displays with simultaneous delivery of auditive and tactile feedback, exceeding the capabilities of alternative optical approaches[4]. Polarization based photophoretic approaches[30] could potentially match the potential for particle manipulation (i.e. speeds and accelerations) demonstrated in this study, but they would still not be able to engage with sound and touch. The MATD prototype demonstrated hence brings us closer to volumetric displays providing a full sensorial reproduction of virtual content. Beyond opening a new venue for multimodal 3D displays, our device and techniques enable positioning and amplitude modulation of acoustic traps at the sound-field frequency rate (i.e. 40kHz), providing also an interesting experimental setup for chemistry or lab-on-a-chip applications (e.g. multi-particle levitation and mode oscillations demonstrated in Figure S10 and Video SV6).

REFERENCES

1. Sawalha, L. *et al.* A large 3D swept-volume video display. *IEEE/OSA J. Disp. Technol.* (2012). doi:10.1109/JDT.2012.2183339
2. Langhans, K., Oltmann, K., Reil, S., Goldberg, L. & Hatecke, H. FELIX 3D display: Human-machine interface for interactive real three-dimensional imaging. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2005). doi:10.1007/11590361_3
3. Blanche, P. A. *et al.* An updatable holographic display for 3D visualization. *IEEE/OSA J. Disp. Technol.* (2008). doi:10.1109/JDT.2008.2001574
4. Smalley, D. E. *et al.* A photophoretic-Trap volumetric display. *Nature* (2018). doi:10.1038/nature25176
5. Ochiai, Y. *et al.* Fairy Lights in Femtoseconds. *ACM Trans. Graph.* (2016). doi:10.1145/2850414
6. Gao, X. *et al.* High brightness three-dimensional light field display based on the aspheric substrate Fresnel-lens-array with eccentric pupils. *Opt. Commun.* (2016). doi:10.1016/j.optcom.2015.10.040
7. Foresti, D. *et al.* Acoustophoretic printing. *Sci. Adv.* (2018). doi:10.1126/sciadv.aat1659
8. Watanabe, A., Hasegawa, K. & Abe, Y. Contactless fluid manipulation in air: Droplet coalescence and active mixing by acoustic levitation. *Sci. Rep.* (2018). doi:10.1038/s41598-018-28451-5
9. Smalley, D. E. *OSA Display Technical Group Illumiconclave I. Optical Society of America* (2015).
10. Saito, H., Kimura, H., Shimada, S., Naemura, T. & Kayahara, J. Laser-plasma scanning 3D display for putting digital contents in free space National Institute of Advanced Industrial Science and Technology. *SPIE-IS&T* (2008).
11. Perlin, K. & Han, J. Y. Volumetric display with dust as the participating medium. (2006).
12. Sahoo, D. R. *et al.* JOLED: A mid-air display based on electrostatic rotation of levitated Janus objects. in *UIST 2016 - Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (2016). doi:10.1145/2984511.2984549
13. Norasikin, M. A. *et al.* SoundBender. in *The 31st Annual ACM Symposium on User Interface Software and Technology - UIST '18* 247–259 (2018). doi:10.1145/3242587.3242590
14. Marzo, A. & Drinkwater, B. W. Holographic acoustic tweezers. *Proc. Natl. Acad. Sci.* (2018). doi:10.1073/pnas.1813047115
15. Brandt, E. H. Acoustic physics. Suspended by sound. *Nature* (2001). doi:10.1038/35097192
16. Wu, J. Acoustical tweezers. *J. Acoust. Soc. Am.* (1991). doi:10.1121/1.400907
17. Bruus, H. Acoustofluidics 7: The acoustic radiation force on small particles. *Lab on a Chip* (2012). doi:10.1039/c2lc21068a
18. Omirou, T., Marzo, A., Seah, S. A. & Subramanian, S. Levipath: Modular acoustic levitation for 3D path visualisations. in *Conference on Human Factors in Computing Systems - Proceedings* **2015–April**, (2015).
19. Memoli, G. *et al.* Metamaterial bricks and quantization of meta-surfaces. *Nat. Commun.* **8**, (2017).
20. Baresch, D., Thomas, J. L. & Marchiano, R. Observation of a Single-Beam Gradient Force Acoustical Trap for Elastic Particles: Acoustical Tweezers. *Phys. Rev. Lett.* (2016). doi:10.1103/PhysRevLett.116.024301

21. Whymark, R. R. Acoustic field positioning for containerless processing. *Ultrasonics* (1975). doi:10.1016/0041-624X(75)90072-4

22. Marzo, A. *et al.* Holographic acoustic elements for manipulation of levitated objects. *Nat. Commun.* (2015). doi:10.1038/ncomms9661

23. Bowen, R. W., Pola, J. & Matin, L. Visual persistence: Effects of flash luminance, duration and energy. *Vision Res.* (1974). doi:10.1016/0042-6989(74)90079-0

24. Tubbs, R. S. *Essentials of Audiology, Third Edition*. *Otology & Neurotology* (2010). doi:10.1097/MAO.0b013e3181c99550

25. Makous, J. A critical band filter in touch. *J.* ... (1995). doi:7722630

26. William Frier, Dario Pittera, Damien Ablart, Marianna Obrist, S. S. Sampling Strategy for Ultrasonic Mid-Air Haptics. in *ACM Human Factors in Computing Systems* (ACM, 2019). doi:10.1145/3290605.3300351

27. Carter, T., Seah, S. A., Long, B., Drinkwater, B. & Subramanian, S. UltraHaptics: Multi-Point Mid-Air Haptic Feedback for Touch Surfaces. *ACM Symp. User interface Softw. Technol.* (2013). doi:10.1145/2501988.2502018

28. Gan, W. S., Yang, J. & Kamakura, T. A review of parametric acoustic array in air. *Applied Acoustics* (2012). doi:10.1016/j.apacoust.2012.04.001

29. Ochiai, Y., Hoshi, T. & Suzuki, I. Holographic Whisper: Rendering Audible Sound Spots in Three-dimensional Space by Focusing Ultrasonic Waves. *CHI '17 Proc. 2017 CHI Conf. Hum. Factors Comput. Syst.* (2017). doi:10.1145/3025453.3025989

30. Shvedov, V., Davoyan, A. R., Hnatovsky, C., Engheta, N. & Krolikowski, W. A long-range polarization-controlled optical tractor beam. *Nat. Photonics* (2014). doi:10.1038/NPHOTON.2014.242

Extended data is available for this paper at www.nature.com/nature.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgement We would like to thank Dr. Rafael Morales Gonzalez and Mr. Ben Danesh-Kazemi from the Sussex University, who helped with building the FPGA array, and Luis F. Veloso from Sussex University, who helped with the supplementary video. We acknowledge funding from the EPSRC project EP/N014197/1, "User Interaction with self-supporting free-form physical objects", EU FET-Open project Levitate (grant agreement No 737087), the Royal Academy of Engineering Chairs in Emerging Technology Scheme (CiET1718/14), the Rutherford fellowship scheme, Japan Society for the Promotion of Science Grant-in-Aid No. 18J01002 and Kenjiro Takayanagi Foundation.

Author contributions D.M.P and S.S. conceived the idea; R.H. and D.M.P implemented the system and gathered experimental data demonstrating the idea. Data analysis were led by R.H., with contributions from all authors. D.M.P led the optimization design with contributions from R.H. and S.S. R.H. optimised the firmware code with contributions from N.M. R.H. wrote the paper, with contributions from all authors.

Author information:

- **Competing interests** The authors declare no competing interests.

- **Reprints and permissions information** is available at http://www.nature.com/reprints.
- **Correspondence and requests for materials** should be addressed to R.H.

TABLES:

Table 1: Main parameters MATD

	Visual only	Visual and audio	Visual, audio and tactile
Highest speed recorded (v_{max})	3.75 m/s	3.375 m/s	2.5 m/s
Highest acceleration recorded (a_{max})	141 m/s^2	122 m/s^2	62 m/s^2
Highest speed for corner features (v_{corner})	0.75 m/s	0.5 m/s	0.375 m/s
Highest image framerate until now	12.5Hz	10.0Hz	10.0Hz
Colour	24bpp	24bpp	24bpp

FIGURE LEGENDS:

Figure 1. Main elements in the MATD. (A) A geometrical description of the visual and tactile stimuli, along with sound, are used as input. The system multiplexes the position of levitation and tactile traps. A quick scanning levitated particle and RGB illumination provide visual content (POV method); modulated acoustic pressure provides tactile feedback and amplitude modulation provides audible sound. (B, C) Example POV images (visible by the naked eye) scanned at 12.5Hz and 10Hz (Video SV1). (D) Multicolour 2D raster image (exposure time 20s, peak speed 0.6m/s); (E) Example 3D POV content (3:2 torus knot) scanned at 10Hz (Video SV2).

Figure 2. Colour reproduction of the MATD display. (A) Example POV content (visible by naked eye) with simultaneous sound (see Video SV2), showing highly saturated colours. (B) Additive colour reproduction of the CIE colour space and grayscale (exposure 8s, peak scanning speeds 0.4m/s for CIE and 0.1m/s for grayscale, non POV). (C) Raster image with simultaneous tactile stimuli (exposure 8s, peak scanning speed 0.2m/s, non POV).

Figure 3. Rendering of volumetric contents. (A, B) Example pyramid visible from all angles around the display (4cm side, 2s exposure (non POV), scanning speed 0.5m/s). (C) Example 3D raster image with rich colour information (6.4cm diameter, 20s exposure (non POV), peak scanning speed 0.9m/s).

METHODS

1. Experimental Setup Overview

Experiments were performed using two opposed arrays of 16x16 transducers, aligned on top of each other and with a separation of 23.4 cm (see Figure S1). We used Murata MA40S4S transducers (40kHz, 1cm diameter (~1.2 λ), 12Vpp, delivering ~1.98Pa at 1m distance) for the two arrays and high intensity RGB LEDs (OptoSupply, OSTCWBTHC1S) to illuminate the bead.

A Waveshare CoreEP4CE6 Field Programmable Gate Array (FPGA) board was used to receive updates from the CPU (3D position, RGB colour, phase and amplitude), using 10 bits to encode each XYZ position (0.25mm resolution), 24 bits for colour (RGB) and 8 bits for the amplitude and phase of the trap, requiring 18 bytes for each update (9 bytes per array of transducers). Communication was implemented using a UART protocol at 12Mbps allowing for 40k updates per second. The following sections provide details on the relevant aspects of our setup, such as operational modes, technical characterization, multiplexing strategies and experimental tests.

2. Driving parameters:

2.1 Transducer's operation (phase and amplitude control).

Transducers were driven using a 12Vpp square wave signal at 40kHz, producing a sinusoidal output due to the narrowband response of the transducers used. Phase delays were implemented by temporal shifting of the 40kHz square wave (see Figure S2a), while amplitude control was implemented by reducing the duty cycle of the square wave (i.e. reduce duration of the high period, as in the lower row in Figure S2a). Complex amplitude of the transducers did not vary linearly with duty cycle (i.e. see Figure S2b, a control signal with 25% duty cycle does not result in half the amplitude of a control signal using 50% duty cycle). We measured this mapping by using one transducer and a microphone placed 4 cm in front of it. We used a GW INSTEK AFG-2225 signal generator to drive the transducer (i.e. square wave, varying phases and duty cycle, as per Figure S2a), and a Brüel & Kjær 4138-A-015 microphone connected to a PicoScope 4262 to measure the differences between the received and reference signals. This allowed us to assess the sinusoidal response of our transducers (no harmonics introduced due to the square wave used to drive them, see Figure S2c), and also allowed us to register how amplitude changed with duty cycle. We experimentally matched duty cycle to effective amplitude as in equation (1), with overall behaviour as shown in Figure S2b.

$$A_t = \sqrt{\sin^2\left(\frac{duty}{100}\pi\right)} \qquad (1)$$

We stored this function as a look-up table in the FPGA (mapping amplitude to duty cycle) for efficient computation of the updates at the required rate (40kHz). This resulted in a modulator providing 64 levels of phase (resolution π/32 radians) and 32 levels of amplitude resolution.

2.2 Embedded computation of twin levitation traps and focusing points.

The computation of focus points and twin levitation traps is embedded into the FPGA. For a focus point at position p and with phase ϕ_p, the phase of each transducer (ϕ_t) was discretised as follows:

$$\phi_t = \left(-\frac{32}{\pi} \cdot k \cdot \boldsymbol{d}\,(p, p_t) + \phi_p\right) \, mod \, 64 \qquad (2)$$

Where k represents the wave number for the frequency used (k=2 π/λ ≈ 726.4 rad/m), p_t represents the position of each transducer and $\boldsymbol{d}$ represents the Euclidean distance function.

Twin traps were computed by combining a high intensity focus point (as in equation (2)) and a levitation signature. Levitation signature was implemented by adding a phase delay of π radians to the transducers in the top array as used by Marzo et al.[14], producing traps maximizing vertical forces. Transducer positions and discretized phase delays relative to distance were stored in two look-up tables in the FPGA, simplifying the computation of the focus point and levitation signature.

2.3 Illumination control

We used one illumination module placed to the top right corner of our MATD prototype, implemented with high intensity RGB LEDs (OptoSupply, OSTCWBTHC1S). The LEDs were driven as per the manufacturers' parameters (I = 150 mA; V = 2.5 V (R) and 3.3 V (G/B)), resulting in luminous flux values of 22 lm (red), 35 lm (green) and 12 lm (blue).

The resulting perceived luminance of the particle (e.g. a point in our visual content) for an observer around the MATD can be analytically approximated from the definition of the Bidirectional Reflectance Distribution Function (BRDF) as shown in equation (3), and it only depends on the angle α between the observer, the particle and the light. The white and diffuse surface of our particles allows us to approximate its BRDF as a Lambertian surface. The small diameter of the particle compared to the distance to the light source allows us to assume incoming illuminance is almost constant across the illuminated surface of the particle, as well as a constant incoming direction (i.e. light source approximated as a directional light). Similarly, the large distance to the observer (compared to the particle diameter) allows us to assume that the direction of the rays from the particle to the observer are also parallel. The perceived luminance is then the summation of the luminances scattered towards the observer direction from each fraction of the sphere illuminated by the source and visible to the observer, as in equation (3). Here, dE_i represents the differential of incoming illuminance hitting the particle; dL represents the differential in luminance towards the observer at each point of the particle's surface; dS represents the differential of surface and θ and ϕ represent spherical coordinates:

$$dL_{obs}(\alpha, dE_i) = \frac{\int_{\theta=0}^{\pi}\int_{\phi=\alpha-\pi/2}^{\pi/2} dL(\alpha,\theta,\phi,dE_i)\cdot dS}{\int_{\theta=0}^{\pi}\int_{\phi=\alpha-\pi/2}^{\pi/2+\alpha} dS} = \frac{dE_i}{4\pi}\cdot\left(1-\sin\left(\alpha-\frac{\pi}{2}\right)\right) \qquad (3)$$

Finally, incoming illuminance (amount of perceived radiant energy emitted per unit area and unit time) needs to be corrected for the ratio of time per second that the particle will be actually present across each discretized part of the visual content. Non-linear human response to luminance (e.g. Steven's power law) needs to be considered and we used a Gamma correction method (γ=2.2), similar to the one used in CRT monitors, to correct for these effects.

3. Operating configurations of the MATD, multiplexing strategy and local phase updates.

3.1 Operational modes and multiplexing strategies for single or dual traps

The hardware can provide individual phase and amplitude updates at 40kHz and time multiplexing to simultaneously create several levitation traps (Figure S10a). However, our MATD prototype only requires the use of up to two time-multiplexed traps: a primary twin trap and a secondary focus point; according to two main operating configurations:

- Single trap mode: Only the primary twin trap is present (100% duty cycle, 40K updates per second), and loaded with an Extended Polystyrene (EPS) particle of ~1mm radius. This levitation trap is used to scan the volume which, synchronized with our illumination modules, provides the visual component of the display. Audible sound is generated by sampling the intended 40kHz audio signal (e.g. from a file), which is then used to modulate the amplitude of the transducers in our array.

A single sided band modulation method (modulation index *a*=0.2) is used, resulting in audible sound of >60 dB (i.e. in the level of a conventional human conversation, see Methods 4.5.). We modulate amplitude while the particle is levitated, in order to create audible sound at the levitation point. More specifically we use an upper sideband modulation (see equation (4)), which avoids harmonics distortion and allows for simultaneous levitation and audible sound (see Video SV1).The modulated signal was computed as:

$$A_{SSB} = \sqrt{\left(1 + ag(t)\right)^2 + \left(a\hat{g}(t)\right)^2} \quad (4)$$

Where $g(t)$ represents the audio signal required to be created at time t, $\hat{g}(t)$ represents a Hilbert transform of $g(t)$ and a represents the modulation index. The signal was sampled at 40kHz and the resulting amplitude (A_{SSB}, from equation (4)) sent to the FPGA together with the remaining required parameters for the current update (i.e. position, colour and phase), implicitly retaining the synchronization between the visual (position and colour) and tactile content with the audio.

- Dual trap mode: This mode is used for cases where tactile feedback needs to be delivered (e.g. only in the presence of the user's hand). In this case, the primary trap can be setup as above, but it needs to be multiplexed with a secondary trap, which creates the tactile stimulation. Two main parameters need to be considered for this multiplexing: *amplitude multiplexing* and *position multiplexing*.

First, *amplitude multiplexing* relates to the recreation of tactile textures, which involves a modulation frequency which can be detected by skin's Lamellar corpuscles (we used an example modulation frequency of 250 Hz). A naïve approach would be to multiplex between the amplitude of the tactile signal (250Hz) and the auditive signal (multiple frequencies), at the expense of limiting the frequency of each individual signal. We instead combine both the tactile and audible signals into a single signal at 40kHz, thus avoiding *amplitude multiplexing* (see Methods 4.4).

Second, the location of the levitation and tactile traps also requires multiplexing, which we refer to as *position multiplexing* to reflect the fact that the traps are created at different spatial locations. Unlike *amplitude multiplexing*, *position multiplexing* only affects the phases of the transducers, and it cannot be avoided in such dual trap scenarios. In our MATD system, we allocate 75% of the updates (3 contiguous updates or 75 µs; update rate 30kHz) to recreate the levitation trap, and 25% for the tactile stimulation (1 update or 25 µs; update rate 10kHz).

This high frequency changes of location (i.e. 10k changes between the tactile and the levitation trap per second) introduce sudden changes in the transducers phases, which might force them to operate at sub-optimal frequencies. To alleviate this, the phase of the next update (ϕ_p, in equation (2)) is set to the value that minimizes the summation of absolute phase differences between the current transducer phase distribution and the previous one.

3.2 Experimental conditions tested

The inclusion of the features above (amplitude modulation for sound and multiplexing in dual trap mode) has implications for the performance of the system. During our tests, we explored three fixed experimental conditions, characterising operating performance of the MATD in both optimistic and worst-case scenarios:

i) **Optimistic single trap mode (OSTm),** with only the main trap and fixed maximum amplitude (A_{SSB}=1);

ii) **Pessimistic single trap mode (PSTm),** with only the main trap and minimum amplitude (A_{SSB}=0.83, equivalent to using the silent section of an audio file);

iii) **Pessimistic dual trap mode (PDTm)**, with both traps (75% duty cycle for the primary trap; 25% for the secondary trap) and minimum amplitude (A_{SSB}=0.83). The location of the secondary (tactile) trap was fixed, horizontally placed at the edge of the array and at a height equal to the centre of the array.

4. Technical characterization: Particle control, visual, audio and tactile modalities.

4.1 Preliminary characterization: particle sizes and update rates

Particle sizes influence the performance of the MATD, due to differences in weight and drag effects. From a selection of highly spherical EPS particles of varying sizes (seven categories, ranging from 1-4mm diameter), we initially assessed each particle for sphericity defects and then used a measuring setup to characterize them.

Our setup (see Figure S3a) uses a Logitech HD Pro c920 camera located 24 cm above a 10x6cm measuring bed. Our software automatically detects the measuring bed and uses a homography to correct for perspective distortion. This allowed for a corrected pixel accuracy of <0.1mm. We then computed circularity as ratio of perimeter and area (*circularity = 4·π·area/perimeter²*), accepting only particles with circularity >0.9. Each particle was dropped on the bed 5 times (to capture different angles of the bead) and only accepted if the circularity test was successful across all 5 measurements. Our software also returned the diameter of the particle, which we used to classify them in 7 binned categories (from 1mm to 4mm diameter, ±0.2mm tolerance for each category). Twenty particles were collected for each category and used during our tests.

We used these initial sets of particles to choose an optimum particle size for our MATD. Figure S3b shows preliminary speed tests (experimental procedure, as described in Methods 4.2), identifying maximum horizontal displacement speed for each category. This initial assessment shows an optimum peak speed for particle diameters between 1.5 and 2.5mm. Although various sizes could successfully be used to create volumetric representations with the MATD (Figure S3c), we chose the curated set of 2mm diameter particles for our remaining experiments. The particle size distribution and sphericity of the set of selected particles is shown in Figure S3d. Particle density and speed of sound in EPS were approximated as 19 kg/m^3 and 900 m/s respectively.

Finally, we also explored the effects of update rate of the MATD on achievable particle speeds. More specifically, we performed speed tests (procedure as in Methods 4.2) along the vertical direction, identifying maximum particle speeds for a range of update rate frequencies of the MATD between 156Hz and 40kHz. Our results are summarized in Figure S3e, illustrating the benefits of the high update rate used by the MATD (higher update rates allow higher particle speeds), and how our PDTm mode could not be supported at rates below 2.5kHz (i.e. operated at 2.5kHz, the 3:1 time multiplexing rate from our PDTm required 400 µs every 1600 µs to create the tactile point, a time during which the levitated particle would fall).

4.2 **Linear Speed Tests**

Trapping forces are dependent on direction due to the type of levitation trap we use. Our trap maximizes vertical trapping forces, while forces along the horizontal plane are weaker, which affects the accelerations and speeds that can be imparted on the particle in each of these directions. This section describes our exploration of the speeds that can be achieved with the MATD. Particularly, we made use of our chosen particles (~2mm) and performed tests characterizing maximum displacement speeds for each of our 3 experimental conditions (OSTm, PSTm and PDTm) for particles moving along three directions: along the vertical axis Y (both in the upwards and downwards directions) and the horizontal axis X. Given our MATD setup, axis X and Z are equivalent (e.g. 90-degree rotation). Speed results along Z are similar to X and not reported here.

Linear paths of 10 cm were used for these tests, with the particles starting at 5cm to the left and stopping at 5 cm to the right of the centre of the MATD (i.e. or 5 cm above/below the centre, for the vertical tests). Particles started at rest and were constantly accelerated to reach maximum speed at the centre of the array. They were then constantly decelerated until brought back to rest at a position 10 cm away from the starting position (e.g. see Video SV3). We used a static camera (CANON, EOS 750D) placed 12 cm in front of the MATD (see Figure S4a) and removed all light. We used a long exposure shot to record our trials and made use of our RGB illumination system, to illuminate (i.e. colour code) the evolution of the bead along its path at steps of 1ms (e.g. see Figure S4b and S4c).

While exploring potential maximum linear speeds (v_{max}), we followed a bisection method (initial boundaries v_1=0, v_2=16m/s). We performed 10 tests at each velocity, and only considered the test successful (i.e. and tested the higher semi-interval) if 9/10 repetitions were successful. We stopped after three consecutive tests were failed, and we only report the highest successful speed observed. This same test procedure (bisection search, 9/10 success rate required, stopping criteria: 3 consecutive failures) was used in all subsequent experiments in this section (i.e. acceleration, radius of curvature and corner speeds).

Figure S5 summarizes the results of the maximum linear speeds (v_{max}) obtained for each condition (OSTm, PSTm and PDTm), for particles travelling along the horizontal direction (Figure S5a), as well as travelling in the vertical direction (Figures S5b and S5c). In the higher part of plots in Figure S5, the solid black lines represent the speed of the levitation trap, while the coloured lines show examples of actual particle velocities as captured during the tests. As expected, maximum displacement speeds are influenced by the mode of operation used. While the decrease in maximum speed is small when audio is included (OSTm vs PSTm), the effect is much larger when tactile effects are introduced as the acoustic power is split between two traps (i.e. time multiplexing for the PDTm mode). Also, linear speeds are much higher along the vertical axis (particularly when going downwards, due to the effect of gravity), when compared to horizontal displacements. This is because our setup with top and bottom arrays and the twin traps used create trapping forces around the levitation trap that are much stronger along the vertical direction (see Figure S4d), allowing for higher accelerations.

The paths observed in Figure S5 show expected correlations between particle velocities (top), particle to trap distances (middle) and accelerations (bottom). Points of zero Δp (i.e. no net force being applied to the particle) correspond with maximum/minimum points in each velocity plot (i.e. derivative equal to zero), and the sign of Δp aligns with the monotonicity of velocity plots, increasing when delta p is negative or decreasing otherwise. Similar correlations can be observed between Δp (middle) and acceleration plots (bottom). Accelerations remain positive when Δp is negative and vice-versa (i.e. trap as a restorative force, following the distribution in Figure S4d), and prominent features in both plots match well (e.g. maximum, minimum, roots).

As shown in the middle part of the plots in Figure S5, it is worth noting that the particle almost always remained at a few millimeters from the place where the actual levitation trap was placed (Δp), being subject to high acceleration rates. This observation is important to understand the behaviour of the MATD in comparison to other levitators.

A particle placed exactly at the centre of the levitation trap ($\Delta p=0$) receives a zero net force contribution, making it stable at that position, but also providing no acceleration. This is ideal for levitators designed for precise (but slow) particle manipulation. Also, such levitators usually operate at much lower update rates (i.e. hundreds of hertz), so when the position of the trap is moved, the particle has enough time to transition to the new trap location. As the particle approaches the centre of the trap, the acceleration received will decrease. If the duration of each update is long enough, the particle will go past the centre of the trap and start receiving negative forces (decelerating), getting engaged in a oscillatory motion until it stabilizes (nearly) at the centre of the trap. As such, modulators with a slow update rate can result in uneven accelerations of the particle or make it difficult for the particle to retain its momentum (accumulate speed) between updates.

The particles manipulated by the MATD do not reach such a static equilibrium after each update. Instead, they need to remain at a distance from the centre of the levitation trap (Δp), so as to receive force and hence be accelerated. This behaviour can be understood in terms of the derivative of the Gor'kov potential at the points around the trap. Figure S4d shows how such forces evolve for points around a trap, as analytically derived considering our particular trap (twin trap), particle (radius ~1 mm, density ~19 kg/m^3, speed of sound in EPS 900 m/s), setup (top and bottom arrays of 16x16 transducers, each modelled using a piston model[22]) and assuming 346 m/s and 1.18 kg/m^3 as the speed and density of air.

As shown in the top of Figure S4d, restorative forces along the horizontal axis peak at distances of nearly $\pm$3.5mm from the centre of the trap, closely matching the distances at which our particles were detected during our horizontal speed tests. A similar behaviour can be observed for the vertical tests. In these cases, the peaks of the restorative forces along the vertical direction (see Figure S4d, bottom) are at distances $\pm$1.5 mm, again matching our observed displacements.

The fact that the trap and the particle did not always remain at those peak distances (i.e. $\pm$3.5mm and $\pm$1.5mm) seems to indicate that even higher speeds should be achievable for both horizontal and vertical displacements. This, however would require a more complex control mechanism to determine the location of the levitation trap, accurately predicting the current location of the particle at each point in time (considering the acoustic force along with drag, gravity and centrifugal forces) and positioning the trap accordingly (e.g. 3.5mm ahead of the particle for maximum horizontal acceleration). Other factors, such as the temporal changes in complex amplitude (and hence force) related to the simultaneous creation of audible sound; or the multiplexing and interference from the secondary trap should also be considered for such a model.

4.3 Acceleration, sharp corners and minimum radius of curvature:

The creation of content for the MATD was approached through the definition of closed and smooth parametric curves, illuminated with varying RGB colours at different points of the path. For content to be visible by the naked eye, such closed curves need to be traversed by the particle in less than 0.1 s [23], which becomes a constraint influencing the particle manipulation required, that is, the speeds and accelerations that need to be imparted at each point along the curve to reveal it within 0.1 s.

While maximum displacement speeds (v_{max}, as identified in Methods 4.2) are a relevant constraint to plan/design such paths, other parameters (i.e. maximum particle acceleration, feasible radius of curvature vs speed and maximum speed at corner features) are equally relevant and were explored next. Again, our characterization follows a conservative philosophy, identifying maximum/minimum values for horizontal displacements (i.e. with weakest trapping forces) and the final parameters obtained for each of our experimental conditions are summarized in Figure S6.

Maximum acceleration per condition

Some contents do not (or cannot) make use of maximum speeds, but they would benefit from increased accelerations. The accelerations identified in Methods 4.2 could be limited as a result of the high particle speed v_{max} used. For instance, drag forces increase with speed and could be one element limiting the maximum feasible acceleration in those tests. Similarly, high speed particle displacements involve more frequent and larger changes to the phase of each transducer, making them operate at frequencies different than 40kHz, and resulting in decreased performance (i.e. emitted pressure).

Here we explored if higher accelerations were then feasible for lower target linear speeds. The experimental procedure followed for this test was similar to the previous speed test, but the maximum target speeds were limited to the $0.5 \cdot v_{max}$, $0.8 \cdot v_{max}$ and v_{max} values identified for each condition. Our tests (see Figure S6) revealed that maximum acceleration achievable was not affected (i.e. increased) by the target speed used (i.e. accelerations observed for all OSTm, PSTm and PDTm modes matched the accelerations identified in Methods 4.2), which seems to indicate that the observed upper limit of accelerations was not related to the particle speed used, but rather due to the trapping force exerted by the MATD.

Maximum speed at corner features

We tested the maximum speed at which the particle could execute a complete change of direction (v_{corner}), such as those required to render corners or sharp features (see Video SV3). The general experimental procedure was again similar to Methods 4.2 (i.e. measuring setup, bisection search, 9/10 success rate required, stopping criteria: 3 consecutive failures). The design of each trial, however, was modified to test if the levitated particle could perform a complete change in direction for a given speed. For each speed tested, the particle started again 5 cm to the right of the centre of the array, accelerating linearly at $0.5 \cdot a_{max}$ until the test speed was reached, and performing a complete 180 degree turn when it arrived at 5cm to the left of the array. The maximum speeds obtained for each condition were 0.75 (OSTm), 0.5 (PSTm) and 0.375 (PDTm) m/s, as reported in Table 1.

Radius of curvature vs speed

Figure S6d shows the maximum displacement speed that can be achieved for a particle moving along a circular path of different radii (1 to 6 cm). The experimental procedure again followed the method used for the other tests (i.e. bisection search, acceptance criteria). For each radius and speed tested, the particles started at rest and were accelerated at $0.5 \cdot a_{max}$ until the test speed was reached, moving along a horizontal circle of the desired radius (see Video SV3). As expected, our results show a decrease in maximum linear speed as the radius reduces (i.e. introducing higher centripetal forces). A reduction is also observed for the highest radius tested (12 cm diameter), as such circle spans across the limits of our operational volume, where it receives less acoustic radiation from the transducers.

4.4 Audio generation and quality

We explored the quality of the audio generated by the MATD, as well as the artefacts introduced due to multiplexing in the dual trap mode. The audio signal used in all these tests was a chirp signal with frequency increasing quadratically from 100Hz to 20kHz (spectrogram shown in Figure S7a, left).

To characterise the performance of our single trap mode, we trapped one particle and used our chirp audio signal to modulate the amplitude of our transducers (as shown in Methods 2.1. and Figure S2). We recorded the sound generated with an audio-technica PRO35 microphone (spectrogram of recorded sound shown in Figure S7b, left), revealing accurate representation of the input signal with some degradation due to harmonics.

To explore the effects of *amplitude* and *position multiplexing* (see Methods 3.1)*,* we repeated the experiment above for two simultaneous (time multiplexed) traps and two input audio signals. We used the same chirp signal for a channel and a 250-Hz sinusoidal signal (spectrogram shown in Figure S7a, centre) to recreate the tactile texture. This represents the case when a primary trap is used to trap a particle (visual and auditive feedback), while the second trap is used to create tactile feedback on the user's skin.

Figure S7b shows the results of mixing both audio and tactile signals either by *amplitude multiplexing* (time multiplexing the amplitude of each signal at 20kHz), or by combination into a single 40kHz (signals added in the frequency domain, as in Figure S7a, right). Our tests show improvements in reconstructed audio in the second case (Figure S7b, right), discouraging the use of naïve *amplitude multiplexing* (Figure S7b, centre).

The use of *position multiplexing* (i.e. focusing the acoustic power at the location of the levitation trap for 75 μs, and then refocussing it at the location of the tactile trap for 25 μs) cannot be avoided if simultaneous tactile and audio-visual content is to be delivered. *Position multiplexing* introduces frequency aliasing at the 10kHz multiplexing rate (as well as harmonic frequencies), as a result of acoustic pressure being focalised at different locations. Our tests show how our multiplexing approach (using *position* multiplexing with combined 40kHz signal, see Figure S7c, right) reduces audible artefacts when compared to the use of both *amplitude* and *position* multiplexing (Figure S7c, left), particularly for harmonics and how our approach minimizes the artefacts present in the human primary auditory range (i.e. 2kHz – 5kHz [24]).

This study also illustrates the need for high update rates for an MATD modulator (i.e. beyond enabling higher particle speeds, as shown in Figure S3e). Our multiplexing schedule involves a multiplexing rate of 10kHz, creating aliasing effects also at harmonic frequencies (i.e. 20kHz). A modulator with a lower rate would create artefacts at many more frequencies, spread across the auditory range (e.g. a modulator at 10kHz would require a multiplexing rate of 2.5kHz, introducing artefacts around 2.5kHz, 5kHz, 7.5kHz, etc.). It is also worth noting that the aliasing effects in our prototype (around 10kHz) are related to the multiplexing schedule used (75% for levitation, 25% for tactile), which in turn is related to the power constraints of our current prototype. Increased transducer power, allowing for effective levitation at a 50% duty cycle (50% for levitation, 50% for tactile feedback) would avoid most of these artefacts, by shifting them around a primary 20kHz frequency. Figure S7d shows a test performed using such configuration (50% duty cycle), with reduced artefacts and with our method (Figure S7d, right) still providing better quality.

4.5 Audio modes supported

The MATD supports two different modes to create audio: a *scatter mode* (Figure S8a), providing non-directional sound but compatible with simultaneous visual and tactile content; and a *directional mode* (Figure S8b), implemented by using the secondary trap to steer the sound on the direction of the user but not allowing simultaneous tactile points (i.e. only visual content and directional audio).

We measured the audible sound generated by each of the two approaches, using a 2kHz audible signal as the audible output. Our measuring setup is comprised of a modified 3D printer (OpenBuilds Sphinx 55), where the extruder has been removed and replaced by a calibrated microphone (i.e. Norsic Environmental Analyser 121, shown in Figure S8c). Our software controls the position of the microphone with 0.1 mm accuracy by issuing G-Code commands over a serial port connection. Displacements of the microphone were followed by 1s pauses (after the end of the motion), to avoid interference due to vibrations. We also configured the microphone to measure sound only in the one third octave band of 2kHz around our intended audible signal (i.e. unconstrained measurements would also capture harmonics, resulting in higher but misleading dB results).

Each of these audio modes (*scatter* and *directional*) were tested for two cases: one measuring audible response when only audio is delivered; and another one when both audio and tactile feedback are delivered. For the *directional mode* (which cannot support all three modalities simultaneously) the second case is representative of situations when the primary trap is used for directional audio generation and the secondary one to create tactile feedback.
Figures S8d and S8e show the results of our tests for horizontal and vertical scans around the MATD volume. Results show audible levels of sound at all points around the display (74±12 dB for the non-directional *scatter mode* and 72±13 dB for the *directional mode*). Points of higher intensity can be found at some points around the MATD, which are to be expected as a result of constructive interference. In the directional case, high pressure levels of 103 dB can be observed around the intended targeted point, which then continue to propagate forwards along the direction between each transducer array and the focussing point. In all cases, the inclusion simultaneous tactile and audio information results in only a small reduction on the intensity of audible sound (66±11 dB and 63±12 dB for the non-directional and the directional methods).

4.6 Tactile generation and Quality

We reused the measuring setup described in Methods 4.5 to scan the sound pressure level (SPL (dB)) generated by our MATD when delivering tactile sensations (see Figure S9a), by replacing the microphone by a calibrated Brüel & Kjær 4138-A-015 microphone connected to a PicoScope 4262, and using the PicoScope SDK to retrieve measurements. We measured SPL generated by our system for a single tactile point at the centre of the array under three conditions, always using the multiplexing schedule described for the dual trap mode.

In the first condition, only the tactile content was delivered (i.e. the array created a tactile point during the 25% duty cycle allocated for the secondary trap, and no output was produced by the array during the remaining 75% percent of the time).

For the second and third conditions, we reused the content displayed in the second part of Video SV2 and Figure S9b, with the scanning bead (primary trap delivering visual content) placed 5cm to the front and left of the tactile point. As a difference, the second condition used a 250Hz signal for side band modulation, representing the case when only visual and tactile content are presented. The third condition, however, included the combined signal (i.e. audio with a 2kHz, combined with 250Hz signal) to represent the case where all visual, tactile and auditive content is presented.

In order to assess the effects that a user hand could have (i.e. due to hands occluding part of the transducers or to scattering on the user's hand), we measured the field both in the presence and absence of a silicone hand (Figure S9c). When the silicone hand was present, the tactile point was created on the surface of the bottom part of the index's fingertip. In all three conditions (visual only; visual and tactile; and multimodal), a horizontal and vertical plane of 10x10cm was scanned, measuring SPL levels at a resolution of 1 mm. Our results from these scans for the three conditions tested (tactile only, tactile and audio and multimodal) are presented in Figures S9d and S9e. It must be noted that the presence of the hand prevented measuring across the entirety of the plane (see white regions in Figure S9e), but the areas within ±3cm around the fingertip could still be reached, covering an area 8 times larger than the width of the focusing point(~7mm⌀). Also, given the thickness of our scanning microphone (3.5mm) and irregularities on the surface of the hand, we could not measure exactly the surface of the hand and the scans presented in Figure S9e are taken at the plane Y=-4mm.

Results show that the device provided accurate positioning and focussing of the acoustic pressure around the central point (where tactile feedback is presented) in all three cases and both in the presence and absence of the hand. Vertical scans show a repeated pattern of lobes, consistent with the interference of the acoustic radiation emitted from the top and bottom arrays. Some differences can be found between the tactile only condition (first column) and the other two cases, as a result of the effects of the primary trap (visual content). However, the effects around the tactile point are small, the sharpness of the tactile point is maintained and there is very little variation across all three cases. Maximum pressure levels are found at the centre of the tactile points (157.0dB, 158.6dB and 158.5dB, in Figure S9d; 154.7dB, 155.0dB and 154.6dB, for Figure S9e), and are always well above the thresholds of 78dB required for perceivable tactile feedback[27]. It must be noted that the presence of a second-high pressure area to the bottom left of the images in the second and third conditions is the result of the primary trap used to deliver the visual content.

EXTENDED DATA LEGENDS:

Extended Data Figure 1: Overview of our MATD prototype.

Extended Data Figure 2: Phase and amplitude control of the transducers used. (A) Square wave input from the FPGA, used to drive the transducer's phase and amplitude, by controlling their phase delays and duty cycles; (B) Non-linear correlation between transducers' pressure and duty cycle as per measurements (dots) and as per our analytical approximation (line); (C) Sinusoidal responses measured from the transducers, when driven by the square waves shown in (A).

Extended Data Figure 3: Preliminary characterization of particle sizes and update rates. (A) Camera setup to measure sphericity and diameter of the beads; (B) Maximum linear speeds for different particle sizes; (C) POV representation using different particle diameters; (D) Particle size distribution and sphericity of the 2mm diameter particles used; (E) Maximum linear speeds along the vertical (downward) path for different update rates and for each mode (OSTm, PSTm and PDTm).

Extended Data Figure 4: Speed measurement setup. (A) A camera takes a long exposure photograph of the moving bead, which is illuminated by the LED at steps of 1ms; (B, C) The captured images of the horizontal and vertical linear speed test of three different conditions (OSTm, PSTm and PDTm); (D) Approximation of horizontal and vertical radiation forces exerted on a particle located around a levitation trap, as analytically approximated from Gor'kov potential.

Extended Data Figure 5: Plots of the speed, distances between the acoustic trap and levitated particle (Δp) and accelerations, as measured during our speed tests along the horizontal (A), upward (B) and downward (C).

Extended Data Figure 6: Summary of the particle control performance tests of the MATD for each of the experimental conditions tested. (A-C) Maximum linear speeds and accelerations for each mode (OSTm, PSTm and PDTm). Please note paths denote the speed of the levitation trap, not observed particle trajectories; (D) Maximum linear speeds achieved by particles following circular paths of increasing radii, for each mode (OSTm, PSTm and PDTm).

Extended Data Figure 7: Spectral analysis of the audio response in the MATD. (A) Signals used for input: chirp (left), 250Hz (tactile, centre) and signals combined in frequency domain (right); (B) Output from the system when only sound is created (left) and when multiplexed with tactile content using *amplitude* multiplexing (centre) and using combined signals (right); (C) Effects of *position* multiplexing on an *amplitude* multiplexed signal (left) and our combined signal (right) for a 75-25% duty cycle; (D) Effects of *position* multiplexing when applied to 50-50% duty cycle signals.

Extended Data Figure 8: Audio modes supported by the MATD. (A, B) Illustration of the two different modes (*scatter mode* and *directional mode*) and how sound tests were conducted; (C) Audio measurement setup; (D, E) Measured sound pressure level (SPL) distribution of the modes. The SPL distributions were measured in two conditions, sound only and sound + tactile feedback, across horizontal and vertical planes.

Extended Data Figure 9: Characterization of tactile feedback. (A) Measuring setup used; (B) Visual content used, together with the tactile point; (C) Measuring setup with a silicone hand (KI-RHAND, from Killer Inc Tattoo); (D) Results of our horizontal and vertical scans of the SPL (dB) for each of our conditions while delivering only tactile feedback, tactile and visual content, and all three modalities (tactile, visual and audio); (E) Results from our vertical and horizontal scans in the presence of a hand, for all three conditions.

Extended Data Figure 10: Other applications of the MATD: (A) Simultaneous levitation of 6 EPS particles in a diamond pattern (16.7% duty cycle for each particle, maximum number of particles levitated to date); (B, C) Frequency modulation at 148Hz to produce resonant oscillations (n=2) for a 2mm water droplet, captured from a side.

Data availability

The data that support the plots within this paper and other findings of this study are available in the main text and Extended Data Figures. Additional information is available from the authors upon reasonable request.

Code availability

Custom C++ code used for controlling our MATD during our tests is available on GitHub for anyone under the Creative Commons Attribution-Noncommercial-Sharealike license.