LeviCursor: Dexterous Interaction with a Levitating Object

Myroslav Bachynskyi
myroslav.bachynskyi@uni-bayreuth.de

Viktorija Paneva
viktorija.paneva@uni-bayreuth.de
Department of Computer Science
University of Bayreuth

Jörg Müller
joerg.mueller@uni-bayreuth.de

Figure 1. LeviCursor enables dexterous interactive control of levitated particles. Users can control particle motion using an optical marker attached to a fingernail. Because we use the optimization-based approach to ultrasonic levitation, particle motion is smooth in any direction. We achieve round-trip latencies of 15 ms, sub-millimeter accuracy, and stability in levitation. Please see the accompanying video for demonstration.

ABSTRACT

We present LeviCursor, a method for interactively moving a physical, levitating particle in 3D with high agility. The levitating object can move continuously and smoothly in any direction. We optimize the transducer phases for each possible levitation point independently. Using precomputation, our system can determine the optimal transducer phases within a few microseconds and achieves round-trip latencies of 15 ms. Due to our interpolation scheme, the levitated object can be controlled almost instantaneously with sub-millimeter accuracy. We present a particle stabilization mechanism which ensures the levitating particle is always in the main levitation trap. Lastly, we conduct the first Fitts' law-type pointing study with a real 3D cursor, where participants control the movement of the levitated cursor between two physical targets. The results of the user study demonstrate that using LeviCursor, users reach performance comparable to that of a mouse pointer.

ACM Classification Keywords

H.5.2 User Interfaces: Input devices and strategies.

Author Keywords

Ultrasonic levitation; user interaction; pointing devices; multimodal displays.

INTRODUCTION

One of the longest-standing visions in Human-Computer Interaction is that of the "Ultimate Display" [21]. This entails a room in which the computer can control the existence of matter. The computer could create chairs, or bullets in such a room, and the virtual and physical worlds would truly be merged.

One approach to creating an "Ultimate Display" has resulted in Programmable Matter [5] and Radical Atoms [8]. Programmable Matter would consist of millions of miniature robots. The main difficulties associated with this "active atom" approach are ensuring sufficient and reliable power supply to the individual units, costs per unit, and miniaturization of the units.

To overcome the limitations of the "active atom" approach, an alternative is to use "passive atoms". In this approach, the "atoms" themselves are passive, but actuation, power supply and intelligence are provided by the environment. This solves the three problems referred to above. This concept was for example employed by Pixie Dust [16].

Pixie Dust uses phased arrays of ultrasonic transducers to generate acoustic standing waves and create a grid of nodes, where small objects can be levitated.

From a Human-Computer Interaction perspective, an important problem of the "Ultimate Display" approach is how to interact, for example, how to move particles interactively. Pixie Dust [16] explores interactive particle control methods, but the use of the classical standing-wave approach for trap generation, introduces limitation in the movement of the particle. While the traps are quite stable, smooth movement is only possible in one dimension between two opposing arrays or

array and reflector. Thus smooth movement in 3D would require six opposing arrays. This would significantly impair the visibility of the levitating display. The Pixie Dust setup consists of four transducer array, allowing for smooth particle movement in 2D. In addition, levitation is only possible within the boundaries of the arrays and is limited to parallel array arrangements.

LeviPath [17] provides an algorithm for moving levitated particles with two opposing arrays, on a 3D grid. The phase values at each step, of approximately 0.2 mm, are precomputed and stored in a table. However, as shown in the video [18], the particles move at a relatively low speed and experience some jittering.

JOLED [19] uses optimization for phase computation, which enables smoother particle movement than the pure standing-wave approach. The JOLED setup is composed of 60 transducers in total. Due to the low number of transducers, real-time particle control using mouse or keyboard is possible. As a consequence, however, the display volume is relatively small.

In general, particle interaction has been implemented at relatively low speeds, due to the risk of particles being dropped during high speeds or high accelerations. In summary, although interaction with levitating particles has been explored by Ochiai et al. [16], Omirou et al. [17] and Sahoo et al. [19], real time gesture interaction with a particle moving along a smooth 3D path with high speed has not been yet achieved.

In this paper, we address the problem of dexterous interactive movement of levitated particles. The main difficulties in achieving this with homogeneous movement along all three dimensions are ensuring: (1) low latency, (2) continuous movement without steps, and (3) stable movement enabling high velocities and accelerations.

We use the optimization-based approach of Marzo et al. [15], avoiding the inhomogeneties of the classical standing-wave levitation. The main limitation of applying the optimization approach in [15] to our setup is that, due to the larger number of transducers, optimization takes about 1 s for each levitation point, thus preventing interactive rates. We solve this problem and achieve (1) low latency by precomputing optimal phases for *all possible levitation points within the entire array at 0.5 mm resolution*, resulting in a round-trip latency of 15 ms. Jumps of the trap, even if just by 0.5 mm, result in noticable jumps followed by oscillations of the particle. We achieve (2) continuous movement by interpolating between the precomputed levitation points at 1 kHz, achieving arbitrarily small step sizes. Optimization creates numerous weaker traps in the vicinity of the main trap. Previously, when placing the particle, one could not be sure that it is actually located in the main trap. Furthermore, over time, the particle might jump to weaker secondary traps, resulting in offset and reduced stability. We achieve (3) stable movement by providing a mechanism to stabilize levitation and ensure the particle is always in the main trap.

The LeviCursor method can be beneficial for studies and applications involving 3D selection with a physical object as cursor, where the correct perception of the 3D targets and the 3D cursor is crucial. It provides a novel method of interacting with tangible interfaces, while opening up new research questions in the HCI community concerning perception, motor control and transfer function of physical cursors which are detached from the user's body. In addition to pointing and selection, precise and accurate manipulation of levitating particles can be used to improve graphical visualizations and animations in mid-air [16], provide better gaming experience in levitation-based games [19] as well as facilitate container-less handling and mixture of sensitive materials, i.e. lab-in-a-drop [3], in favor of preventing contamination.

RELATED WORK

Acoustic Levitation

Acoustic radiation force can be used to counteract gravity and trap millimeter-sized objects in mid-air. This effect is most often achieved by using phased arrays of ultrasonic sound emitters of the appropriate phase and amplitude to create acoustic nodes in mid-air, where particles can be trapped. Acoustic levitation does not require any special (e.g. optical, magnetic, electric etc.) properties of the levitating object. Therefore a variety of objects can be levitated, including solids, liquids and insects [13]. Furthermore, particles of smaller (i.e. Rayleigh particles) [16] and larger (i.e. Mie particles) [14] radius than the wavelength have been levitated.

Moving Levitated Particles

A few methods for achieving controlled movement of levitating particles in the acoustic field have already been developed. *LeviPath* [17] employs an algorithm which combines basic patterns of movement to levitate objects across 3D paths, in a setup consisting of two opposed arrays of transducers. The input path is decomposed into a height variation, controlled by the phase difference between the top and bottom transducer array and a 2D path. The 2D path is then adapted to a possible pattern, obtained by interpolation between adjacent pairs of levitating points. In addition to controlled translational movement in the field, controlled rotations have also been achieved, but with the help of electrostatic forces. In *JOLED* [19] levitating particles of different physical properties are coated with titanium dioxide in order to induce electrostatic charge. This allows for the control of the angular position of the particles by the means of electrostatic rotation. The 3D position of the particles is determined by optimizing the phases of the acoustic arrays.

Interaction with Levitated Particles

For the purpose of contactless manipulation of particles using acoustic levitation, the wearable glove *GauntLev* [11], with integrated ultrasonic transducers, has been designed. The *GauntLev* gloves trap particles either in front of the palm or between a pair of fingers, enabling a set of basic maneuvers such as capturing, transferring and combining levitating particles, a process performed manually or computer assisted. Alternative devices that can be used to manipulate levitated particles which are not attached to the hand are the *Sonic Screwdriver*, a parabolic head with a handle that can generate twin traps and *UltraTongs*, tweezers that generate standing waves [11]. Some of the configurations in [11] and [15]

support one-sided levitation, which provides very good display visibility, but achieving fast and stable levitation is more challenging.

Concerning levitation with static acoustic elements, thus far, only interaction techniques for selection and step translation of particles have been developed. With *Point-and-Shake* [4], users can point a finger to select levitating objects and receive visual feedback in the form of a continuous side-to-side (*shake*) movement. The hand gestures are tracked using a Leap Motion sensor. Interactive ontrol of a single levitated particle using a keyboard, mouse, GUI buttons and a Leap Motion sensor was presented in *LeviPath* [17]. The particle was moved by small steps on a 3D grid. The *Pixie Dust* [16] setup comprises four vertical transducer arrays, facing inwards, which generate a 2D grid of acoustic nodes. Interactive techniques were tested either by using Kinect to detect users' hand gestures, which were then mapped to a particular particle path in the acoustic field (e.g. translating a cluster of particles along one horizontal axis) or by using a pointing touch screen device to assign the trajectories.

Summary
A variety of approaches to ultrasonic levitation have been developed. However, dexterous interaction with levitated objects has not yet been demonstrated. For approaches using only standing waves (in the form of focal lines), the main limitation is that different techniques must be used to move particles in different dimensions. Up to now, this has resulted in less smooth, less agile and often jumpy object motion. Marzo's [15] optimization approach allows for continuous placement of traps at arbitrary locations within the working volume. By displacing these traps with small amounts (approx. 0.1 mm), continuous particle motion can be achieved. In [15] real-time interaction with the system was possible using a keyboard or GUI buttons, however the rates are still too slow for continuous interaction. On larger setups, the optimization would take several seconds for each location. Up to now, this prevented a smooth interactive use of this technique.

Our paper contributes the first implementation of a low-latency, high frame-rate, smooth interactive control of a levitated particle in 3D space, as well as a method that ensures sustained particle positioning in the main trap. In addition, we conduct the first device-mediated Fitts' law study in 3D with a levitated particle as cursor, providing all natural depth cues.

SYSTEM
Our main challenges are to achieve homogeneous movement along all three dimensions with: (1) low latency, (2) continuous movement without steps, and (3) stable movement enabling high velocities and accelerations.

We overcome these challenges by: (1) Precomputation of optimal transducer phases for *all possible levitation points within the entire array at 0.5 mm resolution*. (2) Phase interpolation. (3) A particle stabilization mechanism to ensure that the particle is always in the main trap.
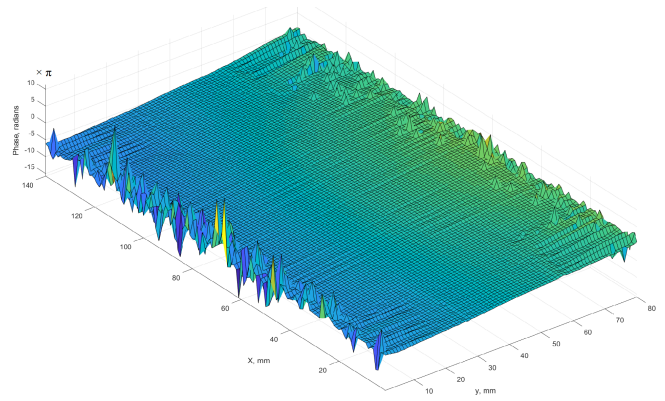

Figure 2. Evaluation of phase smoothness within the sound-field volume. The surface represents phases of the transducer 197 at the plane $Z = 9$ **mm.**

Precomputation of Optimal Transducer Phases
The main limitation of using the optimization-based approach to render interactive levitating interfaces, is that optimization can take several seconds for each new point. We update the levitation points at 1 kHz, rendering this approach to be unfeasible. [15] presents an approach to precomputing discrete animation paths which can then be played back. We extend this approach to precompute *all levitation points in the entire levitation volume at* 0.5 *mm discretization*. Our levitation volume measures 140 mm width * 80 mm height * 90 mm depth. At 0.5 mm resolution, this results in approximately 8 mio. levitation points. For each of these points, the 252 transducer phases have to be optimized. We optimize each point using 20000 iterations of BFGS. We use Armijo line search with coefficient $\alpha = 0.8$ to determine the step size. This takes about 20 seconds per point. The entire calculation takes approx. 44800 hours (> 5 years) of computation time. Since calculation on a workstation is not feasible, we resort to using a computer cluster. We stored the result in a lookup table with a size of 8 GB in RAM.

Because we interpolate the phases between levitation points, it is very important that the phases for neighboring points are smooth. Unfortunately, the optimization problem inherently contains many local optima. Ideally, neighboring points should use the "same" local optimum, and avoid jumping to a distant one, as such a transition would render the interpolated data inconsistent and lead to unpredictable behavior of the levitated particle. After evaluating diverse approaches to achieving this, we propose the following strategy. First, the center of the levitation volume is optimized from random starting phases. Any subsequent point is optimized using the phase values of a neighboring point for starting phases. After the center point, we optimize progressively in the height dimension (up and down). To ensure smoothness, we optimize with 0.1 mm resolution. From this line, we optimize the entire width of the array with 0.1 mm resolution. This results in an optimized plane at depth 0. From this plane, we optimize in the depth dimension at 0.5 mm resolution. This procedure results in very smooth transducer phases between neighboring points (see Figure 2). Any remaining non-smoothness is mostly in the height dimension.

Phase Interpolation

In order to achieve sub-millimeter precision in manipulation of the sound field, we use trilinear interpolation between the eight neighboring points from the lookup table. We first evaluate the acceptability of such interpolation by numerically computing smoothness within the whole sound-field volume. We consider the transition between two neighboring points as smooth, if the differences between the phase values of each transducer are not larger than π radians. The majority (96.2%) of the phase transitions within the sound field volume are smooth and far smaller than π. However, there is still a small fraction of non-smooth transitions, which needs to be investigated. We inspect spatial properties of the transition smoothness, and in particular those of non-smooth transitions, using visualizations of the transducer phases over multiple slice surfaces within the volume (Figure 2). As can be observed, the phases are smooth close to the center of the volume, and become non-smooth closer to the boundaries, in particular in the proximity of the transducers. Based on our observations, we configured the trilinear interpolation so that it is applied if the neighborhood of the point is smooth, and it is replaced by the nearest neighbor values if the neighborhood is non-smooth. The particle movement is less smooth (0.5 mm steps) when entering a non-smooth region, but the general stability of the particle movement is increased.

Particle Stabilization

One major problem with ultrasonic levitation is placing the particles. When a focus point is generated by the optimizer, weaker secondary traps also appear in the acoustic field. These secondary traps can levitate particles, but are prone to disappear and drop them once the primary trap is moved. Since the acoustic field cannot be seen with the naked eye, one can not distinguish between different traps. Consequently, placing the particle in the main trap is not a trivial task. Furthermore, after some time, the particle may jump to a secondary trap.

We stabilize the particle, i.e. reassure it is located in the primary trap, both when the particle is first placed into the acoustic field as well as during direct interaction. When placing the particle, we optimize the field for a levitation point at the origin. We place the particle in the acoustic field, using a piece of an acoustically-transparent fabric. Then we turn on the transducers, which causes levitation of the particle in

some secondary trap. We determine the actual particle position using the motion capture system and generate a primary trap at the actual particle position. During interactive control of the particle, excessively large jumps of the primary trap can cause the particle to jump into a secondary trap. Therefore, we interpolate the primary trap position towards the target indicated by the user, while ensuring that the primary trap never moves more than 0.2 mm between frames in the regions with interpolation. In Figure 3, a levitating particle moving towards a new target position is shown. In the subsequent frame, a new primary trap is generated in the direction of the target, at a distance of 0.2 mm. This procedure contributes substantially to the stability of the levitated particle.

HARDWARE

Our acoustic levitator comprises two 9×14 arrays of *muRata MA40S4S* transducers. The transducers are cylindrical and have a 10 mm diameter and a 7 mm height. The ultrasonic transducers are equally spaced at a distance of 0.3 mm from each other and have maximum input voltage of 20 Vpp. Each emits a sound wave of frequency $f = 40$ kHz (wavelength $\lambda = 8.6$ mm), which is inaudible to humans. The two arrays are mounted horizontally, facing each other, at a distance of 80 mm. We developed an aluminum rail system, which allows for easy adjustment of the distance between the arrays.

A major problem when using transducer arrays for levitation is that the arrays heat up fast, leading to destruction within a few minutes. We solved this problem with a cooling system that generates an air stream on the back of the array PCBs, without leaking an air stream into the levitation volume. This allows us to operate the arrays continuously. We use expanded polystyrene beads of small diameter (approx. 2 mm) as levitating particles, due to their low density.

For driving the transducer arrays, we use the logic board of the *Ultrahaptics* [1] Evaluation Kit. We connected the board to both transducer arrays, leading to on-board synchronization of both arrays. The logic board is connected to a driving PC using USB.

We track the particle position and index finger of the user using optical motion capture (*OptiTrack*). We use a small velcro-attached retro-reflective marker with a diameter of 9 mm, placed directly on top of the user's fingertip. We use six *Prime 13* infrared cameras capturing 240 FPS. Three cameras observe the levitation volume from the side, while three additional cameras track the user's finger from above. The cameras are connected via Ethernet to a second PC that drives the motion capture system and our levitation software.

SOFTWARE

Our precomputation software is based on the system implemented by Marzo et al, which is generously shared in [12]. Based on this, we developed a program for phase optimization that is suited for execution on a computer cluster. We slice the workload into 88000 task description files using a script. Worker nodes read these files and generate a result file.
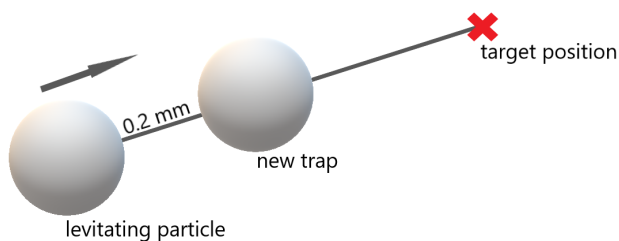

Figure 3. When the particle is moving towards a new target position, it never takes steps larger than 0.2 mm per frame, to ensure that it stays in the primary trap.

---

[1]http://ultrahaptics.com

The interactive hardware and software has to operate in real-time. We use two workstations to operate the system, so as to reduce latencies. The first workstation operates in high-performance mode and runs the *OptiTrack Motive* motion-capture system. Particle and fingertip are tracked and streamed via NatNet to a custom Java program running on the same machine. The Java program performs particle stabilization and computes the particle motion. This program reads in the results files from the cluster computation at startup, so as to generate the lookup table. It looks up the necessary transducer phases in this table. Finally, it performs phase interpolation and sends the resulting transducer phases to a C++ program on a second workstation.

The second workstation is tuned to run the C++ application which receives the transducers' states through a UDP socket. The C++ program caches the phases locally and uses the Ultrahaptics Low-level SDK to stream the phases to the Ultrahaptics logic board. To ensure smooth levitation, the C++ software needs to respond to a callback from the Ultrahaptics driver at 1 kHz with a latency of a few milliseconds at maximum. This workstation runs only the critical operating system processes with low priority on one half of the CPU cores, as defined by an affinity mask. The real-time priority and the other half of the cores (non-hyperthreaded) are dedicated to the C++ application. The machine runs in high-performance mode with CPU sleep states and SpeedStep disabled. Both workstations are connected via Ethernet using a local Gigabit switch. The experiment is controlled and logged using the Java program on the first workstation, which also computes particle motion.

TECHNICAL EVALUATION

To evaluate velocities and stability, similarly to LeviPath [17], we performed an experiment in which we moved a particle back and forth within the levitation volume along a 7 cm straight path. We repeated the movement five times at each velocity and recorded the number of successes and failures. When the particle correctly completed the full movement along the given path, success was registered. A failure was noted when the particle fell off or switched to a secondary trap during the movement. We started with a velocity of 0.2 m/s, gradually increasing it by steps of 0.2 m/s up to 1.2 m/s, where failure was observed in all five trials.

As can be seen in Figure 4, our system achieved particle velocity of 0.8 m/s with a 100% success rate, thereafter the success rate decreased almost linearly and eventually reached 0 at a velocity of 1.2 m/s. From this experiment we can conclude a lower bound on the maximum velocity of 0.8 m/s. We observed, however, that most of the failure cases consisted of the particle dropping either at the beginning or at the end of the movement. This indicates that the limiting factor is not the velocity, but the acceleration. In fact, we believe that by providing more dynamically consistent control it should be possible to achieve even higher particle velocities. For example, our system was able to achieve velocities close to 1.5 m/s, however in this case the particle was shooting out of the end-trap. In the future, we want to conduct experiments where the maximum reachable velocity and acceleration in
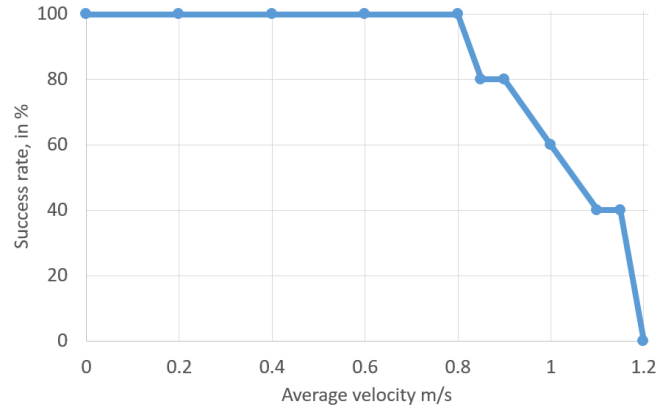

Figure 4. Success rate with respect to average particle velocity.

the right-most part of Figure 4 (0.8 to 1.2 m/s), are explored separately.

We also evaluated the total latency of the system using a high frame-rate camera. We setup a motion capture marker-based event (marker crossing a plane) and a response of the levitation system (dropping the currently levitated particle). The camera observed the space where both event and response were generated and recorded the corresponding segments. We repeated the experiment three times and tallied the number of frames between the marker event and the system response. For a system to be perceived as real-time in pointing tasks, the total latency has to be below 20 ms [9]. In our experiment, in all three cases the latency between the event and the response was less than 17 ms, with the average value being 15 ms, which is below the threshold value perceptible for users.

USER STUDY

As suggested in the introduction, key application areas allowed by LeviCursor are physical 3D pointing, including 3D pointing with tangibles, and aimed movement user studies providing all natural depth cues of the cursor and the targets. LeviCursor allows user studies of mediated 3D pointing to investigate effects of latency, control-to-display ratio or the transfer function on the pointing process, accuracy, speed, physical ergonomics, cognitive load, movement dynamics, velocity and acceleration profiles etc. There are multiple user studies investigating pointing movements in 3D space, however in contrast to LeviCursor they provide either limited cues for depth perception e.g. using volumetric display [7] or virtual reality [22], or they do not allow for any transfer function, for example non-mediated 3D pointing [1]. We demonstrate applicability of LeviCursor to pointing tasks by running a short user study of 3D aimed movements.

The task was a variation of Fitts' *serial pointing task* adapted to 3D. It is very difficult to place physical targets for levitating particles. The targets should disturb physical particle motion, the sound field, and the motion capture system as little as possible. We decided to use needles painted with black matte color to show the center of the targets. The actual targets were internally represented as spheres around the needle tips and were registered using the motion capture system. We

Figure 5. Participants had a retroreflective marker attached to their right index finger and were seated on a chair in front of the levitation apparatus. With their fingertip, they were able to control the levitating particle in front of them and complete the given pointing tasks. The two targets within the levitation volume are marked with red.

used three target sizes of: 2 mm, 4 mm and 8 mm radius. The distance between the targets was 68 mm. The target size conditions for each user were randomized. The task of the user was to move the particle between the two targets as quickly as possible. The motion capture system was tracking the position of the particle with respect to both targets. When the particle entered the target, a confirmation tone sounded and a success was registered.

We recruited 8 participants (mean age 30.5 years, std. dev. 5.6, 4 male, all normal or corrected to normal eyesight, all right-handed). Participants sat on a chair in front of the apparatus (see Figure 5). A retroreflective marker of 9 mm diameter was attached to the index finger of their right hand. The particle was placed in the levitation volume by the experimenter. Participants could control particle motion in 3D with their fingertip, using a control-to-display ratio of 3. Participants were allowed to explore the particle motion for approx. 30 s. We asked participants to place the particle as accurately at each of the needle tips as they could, in order to calibrate the target location according to their perceptions of the target. Afterwards, participants were asked to move between the targets as quickly as possible. After performing 50 aimed movements, the experiment was shifted to the next target size-condition.

During the experiment, our software was continuously recording the 3D position of the particle, the real-time timestamps and the timestamps when the user reached each target and was notified by the sound. After the experiment, the participants were informally interviewed concerning their experience with LeviCursor.

Analysis

We applied Fitts' law analysis, as is typical for the HCI field [10]. While there exist multivariate models of pointing [7], for spherical targets they are equivalent to Fitts' law. We use Fitts' law in the Shannon formulation

$$MT = a + b \times \log_2\left(\frac{D}{W} + 1\right),$$

where MT is the movement time, D the amplitude, W the target width, a and b are free regression coefficients. Following the recommendations of [10], instead of D and W, we use effective target width W_e, based on the standard deviation of the end-points (σ) as

$$W_e = 4.133\sigma$$

and the effective amplitude D_e as the distance between the corresponding effective target centroids:

$$D_e = \sum_{i=1}^{N} \frac{D_i}{N},$$

where D_i is the amplitude of individual aimed movement and N is the number of movements terminating within the effective target. We group the data into six ranges according to the ID. We average IDs and MTs within each group and then fit a Fitts' law model as a first-degree polynomial optimally representing the data in the least-squares sense. We evaluate goodness-of-fit using the coefficient of determination (R^2). To evaluate the performance of the users using LeviCursor, we compute the average effective throughput

$$TP_{ea} = \frac{1}{P}\sum_{i=1}^{P}\left(\frac{1}{C}\sum_{j=1}^{C}\frac{ID_{eij}}{MT_{ij}}\right),$$

where P is the number of participants and C is the number of conditions, as well as maximum effective throughput

$$TP_{emax} = \max_{i=1}^{P}\left(\max_{j=1}^{C}\frac{ID_{eij}}{MT_{ij}}\right).$$

Results

The experimental data can be modeled successfully by Fitts' law with R^2 of 0.92, as can be seen in Figure 6. The participants achieved an average throughput of 4.93 $bits/s$ and a maximum throughput of 8.69 $bits/s$. These values are comparable to the throughput of the mouse [20]. Furthermore,
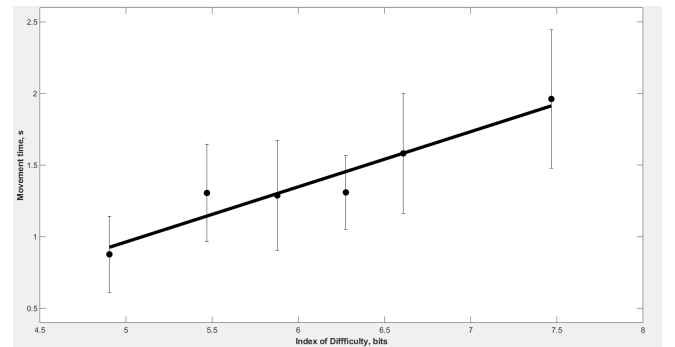

Figure 6. Fitts' law model representing the data of all participants.

they are only slightly below the throughput of uninstrumented mid-air pointing (average $TP = 5.48\ bits/s$ [1]).

According to the informal interviews, the users experienced the interaction as exciting. It was described, for example, as "Jedi using the force", and they felt "in control of the particle". Some of them mentioned a common problem in mid-air interaction - tension and fatigue in the shoulder, known as the "Gorilla arm".

We find it promising that even though LeviCursor has different physical properties than a virtual mouse-controlled cursor on a desktop, it can provide comparable interaction behavior and performance. This demonstrates that using our method, users can exercise dexterous control over levitated particles.

This was, however, a preliminary study to test a new concept. We plan to conduct bigger studies with more participants in future work.

DISCUSSION
From the results of both the technical evaluation and the user study, we can clearly see that the proposed method for interactive control of levitated particles is an effective tool for applications which require pointing in the real 3D space. While this is the first paper which demonstrates such smooth and dexterous control of levitated particles, the method also has multiple limitations and large potential for further improvement. Below we describe the limitations and as future work we plan to explore new approaches to workaround the main limitations.

Limitations
The limitations of LeviCursor can be split into two parts - first the limitations inherited from the underlying levitation algorithm [15] and second the limitations of the current algorithm.

The inherited limitations of the method relate to the optimization approach, namely we can levitate a single particle of size smaller than half of the wave length, preferably spherical (although we have also levitated flat and ellipsoidal particles), made of low-density materials. Levitation of multiple particles should be made possible by changing the objective function for the optimization or using a method similar to [15]. Although ultrasound technology has passed safety tests and is cleared for commercial use for haptic and parametric audio devices (e.g. *Ultrahaptics*, *Ultrasonic Audio* etc.), there are still concerns about the effects of high-intensity ultrasound on humans. As a cautionary measure, we provided the participants of the user study with earmuffs.

The approach described in this paper also has multiple limitations, in particular: scalability with respect to the acoustic volume and the computational power necessary for precomputation, flexibility of the ultrasound array setup, extensive hardware both for ultrasound levitation and for motion tracking, and in the current implementation with optical motion tracking - color of the particle and the surroundings. The scalability is limited by the size of the lookup table and the necessary precomputation time. The required memory and computational time scale linearly in each dimension. While in this paper we work with a levitating interface of relatively small acoustic volume, the state of the art hardware and software can allow significantly larger setups, for example current supported size of main memory (2TB by Windows) allows a levitation volume of 1.2 m^3 while keeping the entire table in RAM. Considering that the current lookup table is computed by a cluster within few hours, it should be possible to compute the table for the above mentioned movement volume in reasonable time. In regard to flexibility, it is necessary to recompute the lookup table for each ultrasound array setup, which takes significant computation time. Apart from sophisticated ultrasound hardware, the current approach also requires optical motion capture hardware. The optical motion capture cameras need to be positioned in a way, that allows the levitated particle to be visible across the entire volume of the levitating display. As an additional requirement, the particle has to provide high visual contrast in comparison to the surrounding hardware, in the optimal case it should be retroreflective.

Future work
There are multiple potential improvements to the current approach of interactive control of levitated particles, as well as extensions and additional applications.

As a main direction of our future work we plan to apply other algorithms for levitation which can work in real time instead of the lookup table, namely we plan to work on using holographic acoustic elements (focus point and signature) [15] for computation of the levitation trap in real time. Up to this point we have tried the focus and signature approach, but it was less stable than the optimized phases from the lookup table.

Next, we would like to explore levitation with multiple particles as well as interactive control of them.

Lastly, we would like to identify and test additional realms that can benefit from the LeviCursor method.

CONCLUSION
In this paper, we presented LeviCursor, a method for interactively moving a 3D physical pointer in mid-air with high agility. The method allows a levitated particle to move continuously in any direction. We addressed the three problems of low latency, continuous movement without steps, and stable movement enabling high velocities and accelerations. We contribute three solutions for solving these problems. The first is a complete precomputation of all transducer phases, achieving a round-trip latency of 15 ms. The second is a 3D interpolation scheme, allowing the levitated object to be controlled almost instantaneously with sub-millimeter accuracy. Lastly, we presented a particle stabilization mechanism which ensures that the particle is always in the main levitation trap. This interactive system has been validated by a user study. The results of the study showed that interaction with LeviCursor can be successfully modeled by Fitts' law, with throughput that is comparable to interaction with mouse pointers.

REFERENCES

1. Myroslav Bachynskyi, Gregorio Palmas, Antti Oulasvirta, and Tino Weinkauf. 2015. Informing the Design of Novel Input Methods with Muscle Coactivation Clustering. *ACM Trans. Comput.-Hum. Interact.* 21, 6, Article 30 (Jan. 2015), 25 pages. DOI: http://dx.doi.org/10.1145/2687921

2. Henrik Bruus. 2012. Acoustofluidics 7: The Acoustic Radiation Force on Small Particles. *Lab on a Chip* 12, 6 (2012), 1014–1021.

3. Edward T. Chainani, Woo-Hyuck Choi, Khanh T. Ngo, and Alexander Scheeline. 2014. Mixing in Colliding, Ultrasonically Levitated Drops. *Analytical Chemistry* 86, 4 (2014), 2229–2237. DOI: http://dx.doi.org/10.1021/ac403968d PMID: 24460103.

4. Euan Freeman, Julie Williamson, Sriram Subramanian, and Stephen Brewster. 2018. Point-and-Shake: Selecting from Levitating Object Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 18, 10 pages. DOI: http://dx.doi.org/10.1145/3173574.3173592

5. Seth Copen Goldstein, Jason Campbell, and Todd C. Mowry. 2005. Programmable matter. *Computer* 38, 6 (May 2005), 99–101. DOI: http://dx.doi.org/10.1109/MC.2005.198

6. Lev Petrovich Gor'kov. 1962. On the forces acting on a small particle placed int an acoustic field in an ideal liquid. *Sov. Phys. Dokl.* 6, 9 (1962), 773.

7. Tovi Grossman and Ravin Balakrishnan. 2004. Pointing at Trivariate Targets in 3D Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 447–454. DOI: http://dx.doi.org/10.1145/985692.985749

8. Hiroshi Ishii, Dávid Lakatos, Leonardo Bonanni, and Jean-Baptiste Labrune. 2012. Radical Atoms: Beyond Tangible Bits, Toward Transformable Materials. *interactions* 19, 1 (Jan. 2012), 38–51. DOI: http://dx.doi.org/10.1145/2065327.2065337

9. Ricardo Jota, Albert Ng, Paul Dietz, and Daniel Wigdor. 2013. How Fast is Fast Enough?: A Study of the Effects of Latency in Direct-touch Pointing Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 2291–2300. DOI: http://dx.doi.org/10.1145/2470654.2481317

10. I. Scott MacKenzie. 1992. Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human-Computer Interaction* 7, 1 (1992), 91–139. DOI: http://dx.doi.org/10.1207/s15327051hci0701_3

11. Asier Marzo. 2016. GauntLev: A Wearable to Manipulate Free-floating Objects. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3277–3281. DOI: http://dx.doi.org/10.1145/2858036.2858370

12. Asier Marzo. 2017. Ultraino. (2017). https://github.com/asiermarzo/Ultraino.

13. Asier Marzo, Adrian Barnes, and Bruce W. Drinkwater. 2017. TinyLev: A multi-emitter single-axis acoustic levitator. *Review of Scientific Instruments* 88, 8 (2017), 085105. DOI:http://dx.doi.org/10.1063/1.4989995

14. Asier Marzo, Mihai Caleap, and Bruce W. Drinkwater. 2018. Acoustic Virtual Vortices with Tunable Orbital Angular Momentum for Trapping of Mie Particles. *Phys. Rev. Lett.* 120 (Jan 2018), 044301. Issue 4. DOI: http://dx.doi.org/10.1103/PhysRevLett.120.044301

15. Asier Marzo, Sue Ann Seah, Bruce W. Drinkwater, Deepak Ranjan Sahoo, Benjamin Long, and Sriram Subramanian. 2015. Holographic Acoustic Elements for Manipulation of Levitated Objects. *Nature communications* 6 (Oct. 2015), 8661.

16. Yoichi Ochiai, Takayuki Hoshi, and Jun Rekimoto. 2014. Pixie Dust: Graphics Generated by Levitated and Animated Objects in Computational Acoustic-potential Field. *ACM Trans. Graph.* 33, 4, Article 85 (July 2014), 13 pages. DOI: http://dx.doi.org/10.1145/2601097.2601118

17. Themis Omirou, Asier Marzo, Sue Ann Seah, and Sriram Subramanian. 2015a. LeviPath: Modular Acoustic Levitation for 3D Path Visualisations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 309–312. DOI: http://dx.doi.org/10.1145/2702123.2702333

18. T. Omirou, A. Marzo, S. A. Seah, and S. Subramanian. 2015b. LeviPath: Modular Acoustic Levitation for 3D Path Visualisations. (2015). https://www.youtube.com/watch?v=ahE37cvr1uo.

19. Deepak Ranjan Sahoo, Takuto Nakamura, Asier Marzo, Themis Omirou, Michihiro Asakawa, and Sriram Subramanian. 2016. JOLED: A Mid-air Display Based on Electrostatic Rotation of Levitated Janus Objects. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 437–448. DOI: http://dx.doi.org/10.1145/2984511.2984549

20. R. William Soukoreff and I. Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies* 61, 6 (2004), 751 – 789. DOI:http://dx.doi.org/https://doi.org/10.1016/j.ijhcs.2004.09.001 Fitts' law 50 years later: applications and contributions from human-computer interaction.

21. Ivan E. Sutherland. 1965. The Ultimate Display. *Multimedia: From Wagner to Virtual Reality* (1965), 506–508.

22. R. J. Teather and W. Stuerzlinger. 2011. Pointing at 3D targets in a stereo head-tracked virtual environment. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)*. 87–94. DOI: `http://dx.doi.org/10.1109/3DUI.2011.5759222`

APPENDIX

METHODS

We use the same method proposed by Marzo et al. [15] to model acoustic levitation. In this section, we provide an overview of this method for the reader.

The acoustic radiation force F on a small (radius $<<$ wavelength), spherical particle in an invicid medium is given by the gradient of the Gor'kov potential U [6] [2]

$$F = -\nabla U. \tag{1}$$

The equation

$$U = k_1(|p|^2) - k_2(|p_x|^2 + |p_y|^2 + |p_z|^2), \tag{2}$$

describes the Gor'kov potential [15]. The complex modulus denoted by $|\cdot|$ is defined as $|\beta + i\gamma| = \sqrt{\beta^2 + \gamma^2}$. Equation (2) consists of two parts; the first is a complex pressure part, which shows that particles move from areas of high modulus of the complex pressure p, to areas of low pressure modulus and the second - a velocity part, written in terms of the spatial pressure derivatives, describing how particles are drawn to areas of large modulus of the velocity gradient. The two constants are given by:

$$k_1 = \frac{1}{4}V\left(\frac{1}{c_m^2 p_m} - \frac{1}{c_p^2 p_p}\right),$$

$$k_2 = \frac{3}{4}V\left(\frac{\rho_m - \rho_p}{\omega^2 \rho_m(\rho_m + 2\rho_p)}\right),$$

where V is the volume of the (spherical) levitating particle and ω the wave frequency. c_m and c_p denote the speed of sound through the medium and the particle respectively. The density of the medium is given by ρ_m, and of the particle by ρ_p. In our case, the medium is air and the particle is an expanded polystyrene bead, so we have: $c_m = 343\frac{m}{s}$, $c_p = 2400\frac{m}{s}$, $\rho_m = 1.2\frac{kg}{m^3}$ and $\rho_p = 25\frac{kg}{m^3}$ at a room temperature of $20°C$. From (2), it is clear that for determining the Gor'kov potential, it is necessary to know the pressure field.

Acoustic levitation traps are regions in the field where the acoustic radiation forces converge. Consequently, objects placed in the levitation traps remain suspended in mid-air. Strong and stable acoustic traps can be created using optimization [15]. Maximizing the converging radiation forces is equivalent to maximizing the Laplacian of the Gor'kov potential, given by:

$$\nabla^2 U = U_{xx} + U_{yy} + U_{zz}, \tag{3}$$

with the notation $U_a = \frac{\partial U}{\partial a}$ and $U_{aa} = \frac{\partial^2 U}{\partial a^2}$.

The pressure inside the traps tends to be very high, which can create disturbances for the levitating object. In order to avoid such disturbances, in addition to maximizing the Laplacian, following [15], we minimize the pressure as well.

The two arrays of transducers we employ in this study emit acoustic waves with constant amplitude and frequency. For simplicity, we assume the transducer to be a piston source, and use a model that neglects reflections and nonlinear effects, which is also in the interest of fast computation.

The complex acoustic pressure of the jth transducer in the array can be written as [15]

$$p^j = e^{i\phi^j}M^j, \tag{4}$$

where ϕ is the phase shift and M a complex number, specific to a transducer and a given point in space. Due to the rule of linearity of differentiation, $p_x^j = e^{i\phi^j}M_x^j$ also holds. We calculate M^j by the means of the example of Marzo et al. [15], assuming a circular piston source and using a single frequency far-field model.

$$M^j = P_0 J_0(kr sin\theta_j)\frac{1}{d_j}e^{ikd_j}, \tag{5}$$

where P_0 is a constant determined by the transducer power, J_0 is a zeroth-order Bessel function of the first kind, k is the wave number, r is the radius of the piston source, θ_j is the angle between the transducer normal and the focus point and d_j the distance between the jth transducer and the focus point.

The total acoustic pressure field generated by N transducers, assuming linear superposition of waves, is given by the sum of the pressures $p = \sum\limits_{j=1}^{N} p^j$, generated by individual transducers. As before, by linearity of differentiation, it holds that $p_x = \sum\limits_{j=1}^{N} p_x^j$. Going back to (3), it is clear that now, the Laplacian of the Gor'kov potential can be expressed as a function of only one variable - the phase shift: $\nabla^2 U = f(\phi_1, ..., \phi_N)$. Hence, to produce a specific pattern in the acoustic field, the phase shift for each individual transducer needs to be calculated.

Objective Function

Numerical optimization methods make it possible to choose phase shifts for the individual transducers that best fulfill our predefined requirements. To this end, following [15], we define a function which represents our problem objective - to minimize the pressure and maximize the Laplacian of the Gor'kov potential at a given point in space. This function is used as a criterion by the optimization procedure to select better rather than poorer solutions. To obtain a levitation trap at the point $\vec{q}$, we minimize the objective function

$$O(\phi_1, ..., \phi_N; \vec{q}) = |p(\vec{q})|^2 - \nabla^2 U(\vec{q}). \tag{6}$$

Adding weights to control the relative strength of the trap in a particular direction and to balance the contributions of different terms, results in

$$O(\phi_1, ..., \phi_N; \vec{q}) = w_p |p(\vec{q})|^2$$
$$- w_x U_{xx}(\vec{q}) + w_y U_{yy}(\vec{q}) + w_z U_{zz}(\vec{q}). \tag{7}$$

We apply equal weights to each direction of propagation, hence generating a vortex trap. An overview of different types of levitation traps is found in [15].

BFGS Optimization

The Broyden-Fletcher-Goldfarb-Shanno algorithm is an iterative method for solving nonlinear optimization problems. It belongs to the group of quasi-Newton methods, which have the advantage that the Hessian matrix is not evaluated at each step. Instead, an approximation generated by analyzing the successive gradient vectors is used, making the process more time-effective. This is favorable for solving our optimization problem, as we have to deal with a very large state space, consisting of 252 phase values. Thus, similarly to [15], we employ BFGS optimization to minimize (7).