scientific reports

Check for updates

OPEN

Sonification as a reliable alternative to conventional visual surgical navigation

Sasan Matinfar[1,4,6✉], Mehrdad Salehi[1,6], Daniel Suter[2], Matthias Seibold[1,3], Shervin Dehghani[1,4], Navid Navab[5], Florian Wanivenhaus[2], Philipp Fürnstahl[3], Mazda Farshad[2] & Nassir Navab[1]

Despite the undeniable advantages of image-guided surgical assistance systems in terms of accuracy, such systems have not yet fully met surgeons' needs or expectations regarding usability, time efficiency, and their integration into the surgical workflow. On the other hand, perceptual studies have shown that presenting independent but causally correlated information via multimodal feedback involving different sensory modalities can improve task performance. This article investigates an alternative method for computer-assisted surgical navigation, introduces a novel four-DOF sonification methodology for navigated pedicle screw placement, and discusses advanced solutions based on multisensory feedback. The proposed method comprises a novel four-DOF sonification solution for alignment tasks in four degrees of freedom based on frequency modulation synthesis. We compared the resulting accuracy and execution time of the proposed sonification method with visual navigation, which is currently considered the state of the art. We conducted a phantom study in which 17 surgeons executed the pedicle screw placement task in the lumbar spine, guided by either the proposed sonification-based or the traditional visual navigation method. The results demonstrated that the proposed method is as accurate as the state of the art while decreasing the surgeon's need to focus on visual navigation displays instead of the natural focus on surgical tools and targeted anatomy during task execution.

Computer-assisted navigation systems provide surgeons with rich and complex multimodal data, enhancing intraoperative diagnosis, decision making, and surgical maneuvers. Despite the high reliability of such systems, they have not yet been fully integrated into the surgical workflow. The dominant way of conveying information in current navigation systems is based on visual displays, a method that assists the surgeon only via the unisensory perceptual channel of vision. This may be explained because we are biologically trained to localize objects, including their semantic meaning, visually, based on a Cartesian grid in a static form[1]. However, in a dynamic interaction with a navigation system, occurring over time, objects' qualities are constantly transforming into new states. This challenges the surgeon's cognition, creating complications, especially in a high-intensity environment such as an operating room. A challenge related to hand–eye coordination is that the surgeon's visual attention has to diverge between navigation displays and the actual operation area, including the surgical tools, targeted anatomy, and the surrounding critical structures. Such complications have not been completely resolved even in more recent augmented reality (AR)-based systems, when overloading multiple virtual visual cues on the display may lead to change or inattentional blindness[2,3].

In cognitive psychological research, it has been shown that multisensory integration facilitates information processing. Multisensory integration, that is, the combination of multiple independent but causally correlated information sources from different senses, including auditory, visual, and haptic, improves performance on a wide range of tasks[4,5]. Research in computer-assisted surgery has not yet fully taken advantage of multisensory feedback and there are unanswered questions in this regard. Although, incorporating navigation data into multiple alternative channels will unload a single modality, creating new possibilities for presenting interaction

[1]Computer Aided Medical Procedures (CAMP), Technical University of Munich, 85748 Munich, Germany. [2]Department of Orthopaedics, Balgrist University Hospital, 8008 Zurich, Switzerland. [3]Research in Orthopedic Computer Science (ROCS), Balgrist University Hospital, University of Zurich, Balgrist Campus, 8008 Zurich, Switzerland. [4]Nuklearmedizin rechts der Isar, Technical University of Munich, 81675 Munich, Germany. [5]Topological Media Lab, Concordia University, Montreal H3G 2W1, Canada. [6]These authors contributed equally: Sasan Matinfar and Mehrdad Salehi. ✉email: Sasan.Matinfar@tum.de

data with computer systems more intuitively. In this article, we have highlighted the importance of alternative perceptual modalities for navigated surgery, investigating potential solutions and discussing the future picture of surgical navigation systems. Human auditory perception, as opposed to visual, is not tied to a spatialized atemporal Cartesian grid. Therefore, sonic qualities such as texture, timbre, and rhythm, which unfold over time, are more efficient and more convenient to embody temporal aspects of objects' qualities. The idea of using sound as a source of information has been well founded in sonification research, which is often defined as the systematic transformation of data relations into perceived relations in an acoustic signal to facilitate communication or interpretation that is reproducible[6–8]. The auditory channel as an alternative perceptual modality to visual feedback has proven to be beneficial in different domains, such as process monitoring, data exploration, and navigation[9,10]. It has been shown that sonification is effective in enhancing athletic performance[11], rehabilitating stroke patients[12], and diagnosing clinically relevant pathology patterns[13]. The challenge of sonification design for surgical navigation is to incorporate the complex dimensionality of the application scenario into an integrated audio stream, that meets the clinician's expectations in terms of reliability, usability, and time efficiency.

We hypothesize that a multisensory-based navigation system improves the surgeon's perception in highly precise interventional tasks. This article, as the first step toward multisensory navigation, introduces a novel standalone four-DOF sonification methodology for the pedicle screw placement task in lumbar spine surgery. To demonstrate the feasibility of the solution, we evaluated the method in a phantom study with 17 orthopedic surgeons in terms of effectiveness, usability, and learnability in comparison with conventional 3D visual navigation as an established method and state of the art with respect to accuracy. Despite the fact that the surgeons have more experience using visual feedback, the study results confirmed the reliability of sonification for surgical navigation tasks and demonstrated the potential behind the core idea of this research.

Clinical motivation

Severe pathological conditions of the spine, including deformity, trauma, degenerative disc disease, and spondylolisthesis, can be treated using the established orthopedic surgical technique called spinal fusion or spondylodesis[14,15]. Spinal fusion implants, which consist of specialized screws that are driven into the pedicles of the respective vertebrae, are used to achieve a fusion between two or more spine segments, thereby immobilizing the respective region and absorbing biomechanical forces. In modern approaches, the surgeon prepares a guiding hole for the smooth insertion of screws, using a surgical awl or by drilling K-wires. To determine the central position of the guiding hole within the pedicle, the surgeon uses bony landmarks for orientation[16,17]. Optimal positioning is crucial for avoiding screw perforation, which can cause serious injury to the spinal cord and its surrounding nerves and vessels. Hence, accurate pedicle screw placement is essential for a surgical outcome, and success depends on the experience and anatomical understanding of the surgeon, especially in severe cases such as scoliosis, kyphosis, or congenital anomalies, where the chance of perforation is even greater[18].

There are three main techniques for pedicle screw placement: freehand, fluoroscopy guidance, and stereotactic navigation[19,20]. The misplacement rate, that is, the rate of screws perforating the pedicle cortex to any degree, in the freehand technique ranges from 5 to 41% in the lumbar spine and from 3 to 55% in the thoracic spine[19]. The high rates of misplaced screws in the freehand approach, various pedicle morphology, and different sizes of the vertebral body motivate computer-assisted systems to improve surgical accuracy[19]. However, there exists some level of disagreement about the necessity of accuracy in pedicle screw placement[21]. A careful analysis of related studies[16–19,21,22] shows that accuracy and safety are dependent on several factors, such as the vertebrae level in question, the definitions of thresholds and safety zones, whether the pedicle cortex has been perforated or not, the applied technique, and the availability of the dataset for comparison studies. There have been studies[22,23] that considered the freehand technique an accurate and safe technique for pedicle screw placement, and many surgeons believe that even when performed slightly inaccurately, such imprecise pedicle screw placement is asymptomatic. However, even those asymptomatic cases can cause implant instabilities, prevent smooth fusion, or expedite adjacent-level degeneration[21,24]. Using conventional fluoroscopy has not entirely solved the problem, as the misplacement rate has been reported as 31.9%[21] and even higher in more challenging cases[22].

Conversely, computer-assisted systems for pedicle screw navigation have been shown to be more accurate, with reduced complications[25–29]. Intraoperative image-guided navigation has evolved in recent years as established approaches such as 2D and 3D fluoroscopic navigation have increased the rate of successful placements respectively to 84.3 and 95.5%, respectively[21]. Furthermore, computer-assisted navigation avoids the use of intraoperative imaging, which reduces the dose of radiation required by conventional fluoroscopy[30–32]. However, while the 3D fluoroscopic navigation system demonstrates the most accurate current solution for pedicle screw placement and is accepted as a standard method according to different *in vivo* studies[19,21], the adoption of such technologies in surgical workflow has been slow, requiring further system improvements[33,34]. In a worldwide survey on the use of navigation in spine surgery, conducted by Härtl et al.[33], although 80% of 677 participants acknowledged the use of navigation systems, they concluded current systems do not meet surgeons' expectations in terms of usability, time efficiency, and integration into the surgical workflow. Participants complained about the complexity of use and the disruption of the surgical workflow as major factors. Additionally, they considered time-consuming training to be a prerequisite factor to support the integration of such systems, and Ryang and colleagues[35] supported this in their study. Current navigation systems predominantly provide surgeons with information through visual displays, increasing the surgeons' cognitive load and complicating hand–eye coordination. Unnaturally, surgeons need to divide their focus of attention between the operation site and navigation displays[36], or their field of view becomes cluttered with multiple holographic cues visualized on head-mounted displays. Visual distraction is problematic for surgeons, considering they need to perceive and process complex structures of navigation data at the highest level of precision in the intensive and stressful situation of a surgical environment[2,3].

Informed consent. The informed consent for publication of identifying information/images in an online open-access publication has been obtained from the study participants.

Related studies

Among surgical navigation systems, we focus on AR-based solutions, which include sonification as one of its emerging branches. AR has been shown to be beneficial for surgical applications[37,38], in particular for orthopedic surgery[39]. AR technology has the advantage of superimposing preoperative planning with intraoperative anatomy, which, in the case of visual-centric AR, provides surgical navigation information in the surgeon's field of view. Previous studies have proposed a body of AR-based navigation solutions for pedicle screw placement[40–43]. Similar approaches based on tool-mounted mobile devices have been used to provide information in the line of sight of a surgeon[44–47]. However, all these approaches have utilized visual feedback as the singular feedback modality. As discussed in "Introduction" and "Clinical motivation" Sections, the inherent limitations, such as change or inattentional blindness[2,3] are the motivation behind the research reported here.

Sonification for navigation purposes was initially been designed as a natural application for people with visual impairment[48–53] and has been expanded to more general applications[54–56]. Sonification of one-dimensional data using primitive sound synthesis methods, such as in heart–lung machines, has already been integrated into surgical procedures. Such basic sonification methods do not extrapolate well to more complex multidimensional scenarios, as they lack consideration of psychoacoustics and sound design in their configuration. To address this problem, sonification methods[57–59] have been proposed with more focus on usability and clinical integration, using more flexible and creative sound designs; however, these approaches are unsuitable for presenting precise navigation data.

Sonification methodologies for medical applications have mostly focused on image-guided navigation scenarios. Black et al.[60], in a review paper, named three primary motivations for sonification of surgical navigation: (1) increasing awareness of structures surrounding the tracked instrument, (2) reducing attention to the screen or increasing attention to the patient or test phantom, (3) helping clinicians correctly interpret (multidimensional) navigation data. Wegner et al.[61] recommended different mapping ideas, such as 3D audio spatialization for generalized 3D surgical instrument placement. Sonification in the form of proximity alerts has been proposed for endoscopic cranial base surgery[62], temporal bone drilling[63], protecting facial nerves during otologic surgery[64], guiding cochlear implantation[65], and fluorescence-guided resection of gliomas[66]. More elaborate approaches have been introduced in[67–70] using continuous parameter-mapping sonification for surgical needle guidance in one dimension. Investigation of solutions for one-dimensional distance mapping have been undertaken by Plazak et al.[71], who proposed five different mapping strategies, and Roodaki et al.[72], who introduced a sonification design based on physical modeling sound synthesis that requires minimum training.

Sonification research in recent years has aimed to expand in terms of data dimensionality and degrees of freedom (DOF). Parseihian et al.[54] investigated the efficiency of different sonification strategies in terms of rapidity and precision for a one-dimensional guidance task. Sonification of multidimensional data is challenging[72,73], and researchers have investigated the potential of spatial sound to overcome this challenge for 2D[74] and 3D space[75]. Such approaches have been relatively successful when combined with visual guidance. Spatial sonification as an intuitive and natural method with a high learnability rate is suitable for orientation tasks[54,76]. However, spatial sound does not provide the precise distal and angular resolution required for precise surgical guidance tasks[77]. The resolution of spatial localization is $1° \pm 3°$ along the horizontal axis in front, and becomes less toward the sides. The resolution of estimating distance is decimeters in a short distance area[78]. Conversely, monaural sonification provides flexibility in design, as its efficiency is justifiable because of our inherent perceptual capability, as we can discriminate pitches in a range of 640–4000 steps[78], 120 levels of loudness[78], and 250 levels of sharpness[78,79]. Monaural approaches are efficient regarding dimensionality and resolution; however, they introduce design challenges in terms of intuitiveness and learnability. Sonification methods are proposed for guidance in 2D[55,56] and 3D[76] spaces, providing information such as distance or orientation. These methods employed monaural sonic characteristics such as pitch, amplitude, and timbre.

A review of the state of the art reveals a lack of research on methodologies for surgical tool guidance in two or more dimensions which would be integrable into highly sensitive application scenarios such as pedicle screw placement. In pedicle screw placement, the surgeon aligns the drill with a predefined target trajectory, which can be mathematically defined by two points, the entry and angular target points. Optimal positioning of the tool on these two points requires tool movement in four DOF. To the best of our knowledge, there is no prior study in four-DOF sonification. There are essential questions to address. For example, the approach's effectiveness and usability: to what level of precision and accuracy would sonification provide information along with an appropriate level of immediacy, and simultaneously achieve a satisfactory level of usability required for surgical situations? The majority of methods lack clinical-grade integrable sound designs, and Black et al.[60] described current sonification approaches as being simple. There are a limited number of studies that have compared the effect of sonification to visual feedback. Also, there is a dearth of comprehensive evaluation studies on clinical evaluations and training.

Computer-assisted auditory navigation system

Our approach to providing auditory navigation assistance to surgeons consists of two main components, the navigation and sonification modules.

The navigation module comprises a workstation and an infrared optical tracking camera. The goal is to intraoperatively provide the positioning of the drill controlled by the surgeon. Prior to the operation, the trajectories of the screws are preplanned on the basis of a preoperative computed tomography (CT) volume of the patient, and, to align the preoperative CT with the intraoperative coordinate system of the camera, a registration method
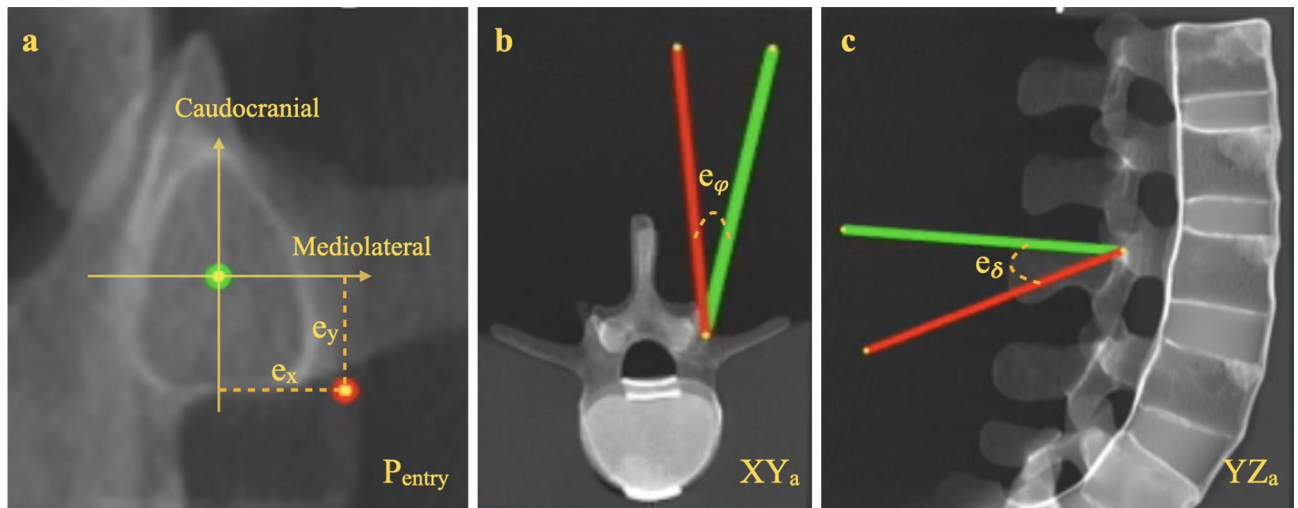
Figure 1. Three cross-sectional views of the CT from the spine phantom model, including the corresponding errors (e_x, e_y, e_ϕ, e_δ). The target and tool are visualized in green and red, respectively. (**a**) Corresponds to the coronal view visualizing, e_x and e_y projected on the P_{entry}; (**b**) represents the axial view visualizing e_ϕ; and (**c**) visualizes the sagittal view including e_δ.

is performed. Intraoperatively, using the camera, markers on the drill sleeve are tracked relative to reference markers on the patient's bed. The real-time position and orientation of the targets relative to the drill tip are sent to the workstation and used to compute error parameters (described in "Error parameters" Section). This information is, in turn, transferred to the sonification module, which generates the output sounds accordingly.

Error parameters. We define the pedicle screw placement as a four-DOF alignment task between the tracked drill sleeve's tip point, T_{tool}, and the preoperative planned trajectory, T_{target}. The first two DOF correspond to the translation of T_{tool}'s tip point projected on the entry point plane P_{entry}. P_{entry} is defined by taking the main direction of the planned trajectory T_{target} as the plane's normal and the planned entry point to the bone as the center of the plane. Hence, both T_{target} and P_{entry} are updated according to each pedicle screw's planned trajectory. The entry point errors e_x and e_y are defined as the distance between the center point of the P_{entry} plane and the projection of the drill sleeve's tip point on P_{entry}. e_x and e_y show the entry point errors in mediolateral and caudocranial directions, respectively.

The remaining two DOF correspond to the orientation mismatch between T_{tool} and T_{target}. This angular error is decomposed into two values, e_ϕ and e_δ, which are Euler angle differences between the projections of the T_{tool} and T_{target} on the axial XY_a and sagittal YZ_a planes, respectively, in the anatomical coordinate system XYZ_a, as illustrated in Fig. 1. The orientation error on the XZ_a plane is negligible because of the symmetry of the tool and the preplanned trajectory. The anatomical coordinate system stays constant throughout the execution of all pedicle screw placements.

Four DOF sonification model. *Interactive alignment model.* The interaction model is designed with two interactive phases, namely, entry point phase (EP), and angle phase (AP), each with two DOF. There are also two static phases, the initial phase (IP) where T_{tool} has not yet entered the entry point working area (W_{EP}), and the final phase (FP) where T_{tool} has reached T_{target}. First, the projections of T_{tool} and T_{target} on the P_{entry} plane have to be aligned (EP); then, the tool orientation is aligned with T_{target} (AP) while the tooltip stays in place. This implies that when the interaction is in the AP, the tooltip has already been aligned to T_{target}. If during the AP the tooltip deviates from T_{target}, the sonification will return to the sound mappings of the EP.

The transitions between these phases and states are carried out using a threshold mechanism, with two control parameters, d and θ. d is the 2D Euclidean distance between the projections of T_{tool} and T_{target} on P_{entry}, and θ is the 3D Euler angular distance between T_{tool} and T_{target}. The user interaction with the sonification model starts when the tooltip enters the W_{EP}, which is a circle on the P_{entry} plane with radius r_{EP} around the target entry point. Furthermore, we define the angular working area W_{Ang}, which includes all T_{tool}s with Euler angular distance less than θ_{Ang} from the T_{target}; i.e., $\theta < \theta_{Ang}$. The alignment task is accomplished when T_{tool} is aligned in all four DOF at T_{target} (Fig. 2).

In interactive phases, EP and AP, we define two thresholds, namely, the target and transition zones. The transitions to a next step, that is, from EP to AP and from AP to FP, are executed only when the tool reaches inside the transition zone. When T_{tool} exits the target zone, the alignment returns to a previous step, that is, from FP to AP or from AP to EP. In these cases, the user needs to reach the transition zone to be able to proceed to the next step. The threshold mechanism with the space between the target and transition zones enables us to smooth out the interaction with the system, avoiding unwanted transitions due to slight hand tremors of the surgeon or optical tracking jitter (Fig. 3).
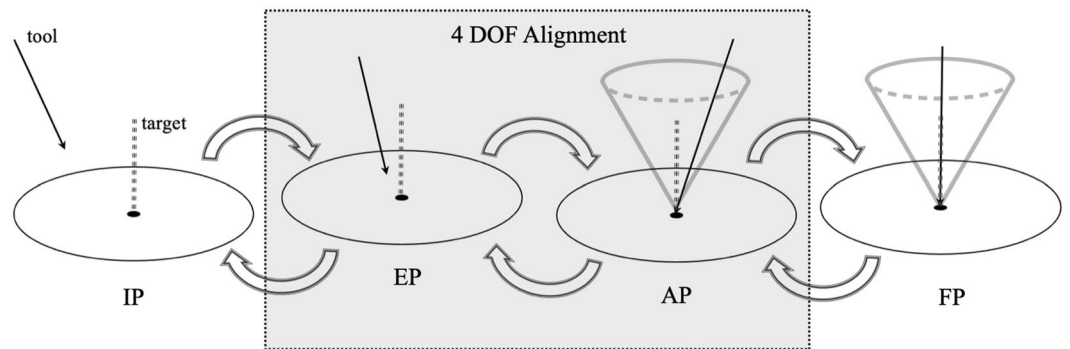
4

Figure 2. Four DOF alignment model with four phases, initial phase (IP), entry point phase (EP), angle phase (AP), and final phase (FP). EP and AP are two interactive phases with continuous mappings, whereas IP and FP are the static phases with constant mappings.
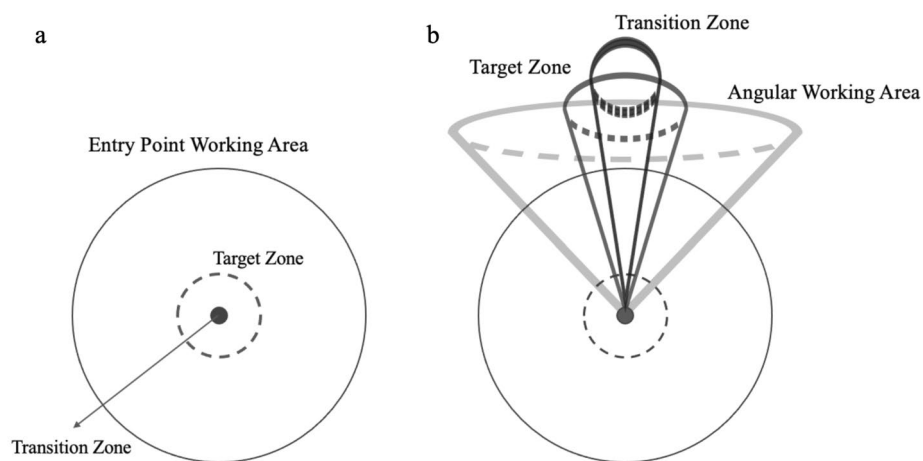

Figure 3. Illustration of the thresholds for transition between phases. (**a**) the circles demonstrate thresholds for the transition between IP, EP and AP; (**b**) the cones represent the thresholds for transition between AP and FP.

Mapping to acoustic features. Sonification mapping is based on a continuous stream of pulse tones generated using the well-known FM synthesis method[80]. In FM synthesis, a modulating oscillator modulates the frequency of the carrier-oscillator, which is an efficient way of producing complex sounds with multiple harmonics with only two oscillators. The input data to the sonification function are the 4D vector $(e_x, e_y, e_\phi, e_\delta)$ as its parameters are described in "Error parameters" Section. These components control the fundamental frequency of the synthesizer and the pulse rate of the pulsing stream. Human auditory system can interpret fundamental frequency and pulse rate separately, as they are orthogonal sound features. Perceptual orthogonality means, two simultaneously sonified quantities are perceived separately by the individual. In particular, when FM synthesizer parameters change, it can unambiguously be interpreted in its corresponding sound attribute. As a result of the complex, nonlinear processing of the auditory system, all physical quantities of the sound field are practically capable of directly affecting all perceptual attributes of sound, which makes it extremely challenging to attain perceptual orthogonality[76,78]. Depending on the alignment phase, the system controls which parameters of the input vector should be used for parameter mapping. In EP, e_x and e_y are used to map to a fundamental frequency and pulse rate, respectively, whereas in AP, e_ϕ and e_δ are used. The mapping of the pulse rate is interpolated linearly; however, exponential interpolation is used for the fundamental frequency, as the human auditory system perceives pitch in an exponential manner.

Because both EP and AP phases use the same implementation of the synthesis function, we apply different ranges for the fundamental frequency of the FM synthesis to create higher contrast between the two alignment phases. An experienced surgeon can generally approximate the target entry point using anatomical landmarks, but finding the target angle in a 3D environment is considerably more challenging. Therefore, we set the frequency range in AP (2 octaves) to be larger than in EP (1 octave) in order to achieve a higher resolution in AP. Furthermore, high-pitch fluctuations of the sound in such micro-temporal interactive tasks are likely to cause fatigue, so a larger range is allocated to lower pitches. To facilitate learning, the range of pulse intervals in both EP and AP are identical. The lower bound is selected because in our design the values smaller than 0.1 sec can not be perceived as discrete pulses, thus the parameter changes are not distinguishable. The higher bound is chosen as a trade-off between delay in the interaction and having a larger mapping range. In the IP and FP, the sonification

Phase	Fundamental frequency	Pulse interval
EP	$e_x \xrightarrow{exp}$ [880, 1760] Hz	$e_y \xrightarrow{lin}$ [0.35, 0.1] sec
AP	$e_\phi \xrightarrow{exp}$ [110, 440] Hz	$e_\delta \xrightarrow{lin}$ [0.35, 0.1] sec
IP	(123.47, 155.56, 185, 246.94) Hz	0.66 sec
FP	(440, 523.25, 659.26, 880) Hz	1.5 sec

Table 1. Parameters for the FM synthesis mapping functions with the input data $e = (e_x, e_y, e_\phi, e_\delta) \in [0, 1]$ for entry point phase (EP), angle phase (AP), initial phase (IP), and final phase (FP).
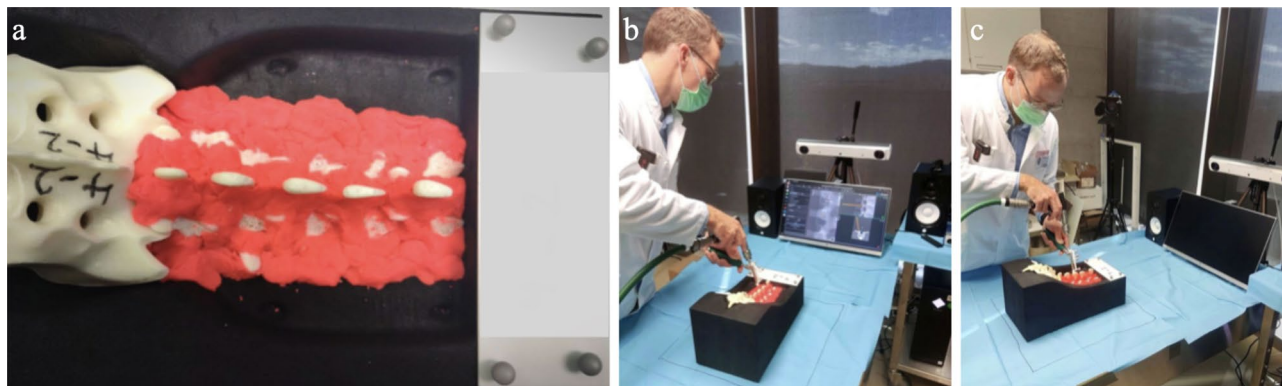

Figure 4. The experiment setup, (**a**) the phantom covered with Play-Doh, (**b**) task assisted with visualization, (**c**) task assisted with sonification.

is limited to musical major and minor chords, respectively, both pulsing at a constant rate with different values, as listed in Table 1. In each interactive phase, when the T_{tool} reaches the T_{target} value in only one dimension, an earcon is played to facilitate the process of finding the target in the second dimension. The so-called optimum earcons consist of two sequential notes with a slight difference, depending on which target dimension has been reached. The optimum earcon for e_x and e_ϕ is the same, whereas a slightly different earcon is used for e_y and e_δ. To make the transitions clear, two additional earcons were designed, consisting of eight sequential notes in ascending order for the EP to AP transition and in descending order for the AP to EP transition, representing the movement forward and backward in the procedure. Earcons' parameters, including the IP and FP chords, are selected considering the fact that they can conveniently be distinguished and learned by the user.

Comparison study.   To compare the sonification and the conventional visual navigation methods, we conducted an experiment with 17 orthopedic surgeons, 4 senior experts, and 13 assistant surgeons. In the study, participants performed the pedicle screw placement procedure on phantoms. We used phantom models of the lower lumbar spine (manufactured by Synbone AG, Zizers, Switzerland) consisting of vertebrae L1–L5. The phantoms incorporate facet joints and discs, which create more realistic, intervertebral movement. To simulate the surrounding anatomical landmarks similar to the real surgical environment, we covered the phantoms with Play-Doh to hide the deeper and medial areas around the drilling surface, as shown in Fig. 4. Each surgeon drilled 20 pedicle screws on two phantoms with an alternating order between auditory and visual navigations. Our primary measures were the entry point distance error and angular error between the executed and preplanned trajectories. For the procedure with conventional 3D visual navigation, participants performed the four-DOF alignment based on three cross-sectional CT slices from three views. The coronal view visualizes e_x and e_y on the P_{entry} plane, aligned to the 3D anatomical coordinate system. The axial view visualizes e_ϕ on the XY_a plane. The sagittal view corresponds to e_δ on the YZ_a plane (Fig. 1). For the visual model, similar to the sonification model, tracking markers are used to track the drilling sleeve's position relative to a reference marker fixed on the phantom's bed. The real-time processing of the tracking data is performed by the workstation and transferred to the visualization module, which renders the image on a visual display. Figure 4 shows the experimental environment.

Starting with a preoperative CT of one of the phantoms, a senior spine surgeon planned 10x lumbar pedicle screws on L1–L5. The preplanned trajectories were aligned to each phantom before starting the trials using a landmark registration method. For the landmark registration, eight points were collected on the most lateral section of each transverse process on L1–L4. L5 was excluded because we observed slight variations among L5 levels in different phantoms; therefore, a higher error for L5 evaluation would be expected.

We used the fusionTrack 500 real-time optical tracking system (Atracsys) and passive infrared markers for tool tracking. The tracking targets on the drill sleeve (3.2 mm, No. 03.614.010, Synapse System) and the phantom's bed were designed with four passive spheres on each. Pivot calibration[81] was performed on the drill sleeve target to transform the tracking coordinates to the center of the drill sleeve's tip. The real-time processing of tracking

data (at 50 Hz) was implemented using ImFusionSuite [ImFusion GmbH, Munich, Germany—https://www.imfusion.com] software. The generation of sounds was implemented with the SuperCollider3 [https://supercollider.github.io/] software platform for audio synthesis. The communication between the ImFusionSuite and SuperCollider modules was established using the OSC networking protocol[82]. Finally, the generated audio signal was sent to a pair of two-way bass reflex studio monitors to be played for the surgeons.

The working area's radius, r_{EP}, was set to 20 mm, and the working area's angle, θ_{ANG} was set to 30°. The target zone's thresholds for both alignment phases (EP and AP) were set to 2 mm and 1.5°, and the transition zones' thresholds were set to 0.5 mm and 0.375°. Choice of these parameters was based on a pilot experiment with an expert spine surgeon, and the optimum values depend on the accuracy of the tracking system, registration, and calibrations.

Each participant was presented with a short introduction about the method ($\approx$ 5 min). The trials consisted of two phases, a training and an execution phase. In the training phase, the participants were asked to conduct 10 alignment tasks with the aid of sonification, on L1–L5 on both sides of the phantom. In the execution phase, they were asked to conduct the alignment and drilling on two phantoms, resulting in 20 executions on the same vertebrae levels. The executions were divided into four sequences, and each sequence was assisted with either visualization (V) or sonification (S). We randomized the order of the sequences between subjects as V, S, V, S or S, V, S, V. Each subject started from either the left or right side of the first phantom and the opposite side of the second phantom, again in a uniformly randomized order.

The secondary outcome measures were the alignment time and the participants' cognitive load. The alignment time is considered the duration between two events, namely, the alignment start and the drilling starting points. This was performed by the trial examiner, pushing a button for each event to record their timestamps. The cognitive load was assessed by asking the participants to respond to a questionnaire, including the NASA Task Load Index (NASA-TLX), a subjective workload assessment measure, and four additional questions provided by the authors. The additional questions were as follows: Q1: Which method helped you better to find the target entry point? Q2: Which method helped you better to find the target angle? Q3: How do you evaluate the overall usability of both systems? Q4: Which navigation feedback method would you like to use in the future?

Ethics statements. This study does not fall within the scope of the Human Research Act (HRA). According to the clarification of responsibility approved by the Institutional Ethics Committee of Canton of Zurich, Switzerland (BASEC-Nr. Req-2021-00820), authorization from the ethics committee was not required.

Evaluation and results

Evaluation. We compared the preoperative planned trajectories with the postoperative CT of the drilled phantoms. To detect the exact drilled path, cylindrical graphite sticks with the same diameter as the drill (3 mm) were inserted into the phantoms before taking the postoperative CT. The average length of the graphite sticks was 5 cm. The centers of the first and last disks of each graphite were manually labeled for every drilled screw. We also marked the actual point where the drill had entered into the bone phantom. The actual entry point might be slightly deeper in the bone phantom compared with the planned trajectory because part of the bone surface was removed by surgeons in order to create a flat surface on the pedicle to stabilize the drill sleeve and prevent sliding; a similar procedure is performed during real surgery.

The preoperative and postoperative CT volumes were registered using an image-based registration method. To minimize the movement and possible deformation between vertebrae, we did not remove the Play-Doh before taking the postoperative CT, which caused different appearances in between CT volumes. To resolve this issue, we masked the image-based registration within a 3-mm area around the segmented vertebrae surface. The registration algorithm was manually initialized within its capture range, and a nonlinear optimizer with the LC^2[83] similarity metric was used to register the volumes.

Results. The results of the post-CT analysis revealed a total mean error of 1.82 mm $\pm$ 0.89 mm for the entry point and 1.75° $\pm$ 1.01° for the angle as deviation from the planned trajectories (CT error, $n = 336$). Conversely, the system-generated data, which were used to generate both visual and auditory feedback modalities, resulted in a mean error of 0.82 mm $\pm$ 0.46 mm and 0.88° $\pm$ 0.47° (feedback error, $n = 323$). We estimated our system error (tracker, registration, and calibration) by subtracting pair-wise samples of the errors (CT and feedback, $n = 314$) to a mean of 0.98 mm $\pm$ 0.77 mm and 0.82° $\pm$ 0.92°. The mean CT error for visualization ($n = 167$) was 1.67 mm $\pm$ 0.87 mm and 1.78° $\pm$ 1.04° and for sonification ($n = 167$) 1.96 mm $\pm$ 0.88 mm and 1.69° $\pm$ 0.96°. The details of the error over the expertise groups and the spinal levels are presented in Figs. 5 and 6 and Table 2.

We expect < 3-s error margin in the recording process for the completion time of each alignment. The mean alignment time for visualization was 33.5 s $\pm$ 16.1 s and for sonification 44.1 s $\pm$ 21.6 s. The details of the alignment time of the first and second executions on the same level are shown in Fig. 7.

Fourteen individuals (3 experts and 11 assistant surgeons) returned the questionnaires. The small sample size ($n = 14$) made it necessary to determine the distribution of the variables Visualization (V) and Sonification (S) before selecting an appropriate statistical method. A Shapiro-Wilk test was performed and did not reveal any evidence of non-normality ($V : W = 0.97, p\ value = 0.88$; $S : W = 0.91, p\ value = 0.17$). On the basis of this outcome, we analyzed the NASA-TLX data, using a t-test for independent samples with unequal variances. The results indicated a P value of 0.59. Therefore, we failed to reject the null hypothesis of having equal means in the samples. Accordingly, we found no significant differences regarding the cognitive load between sonification and visualization. The responses to Q1 and Q2 (which method better helped to find the target entry point and the target angle, respectively) had the same proportions; i.e., 10 individuals voted for visualization, three
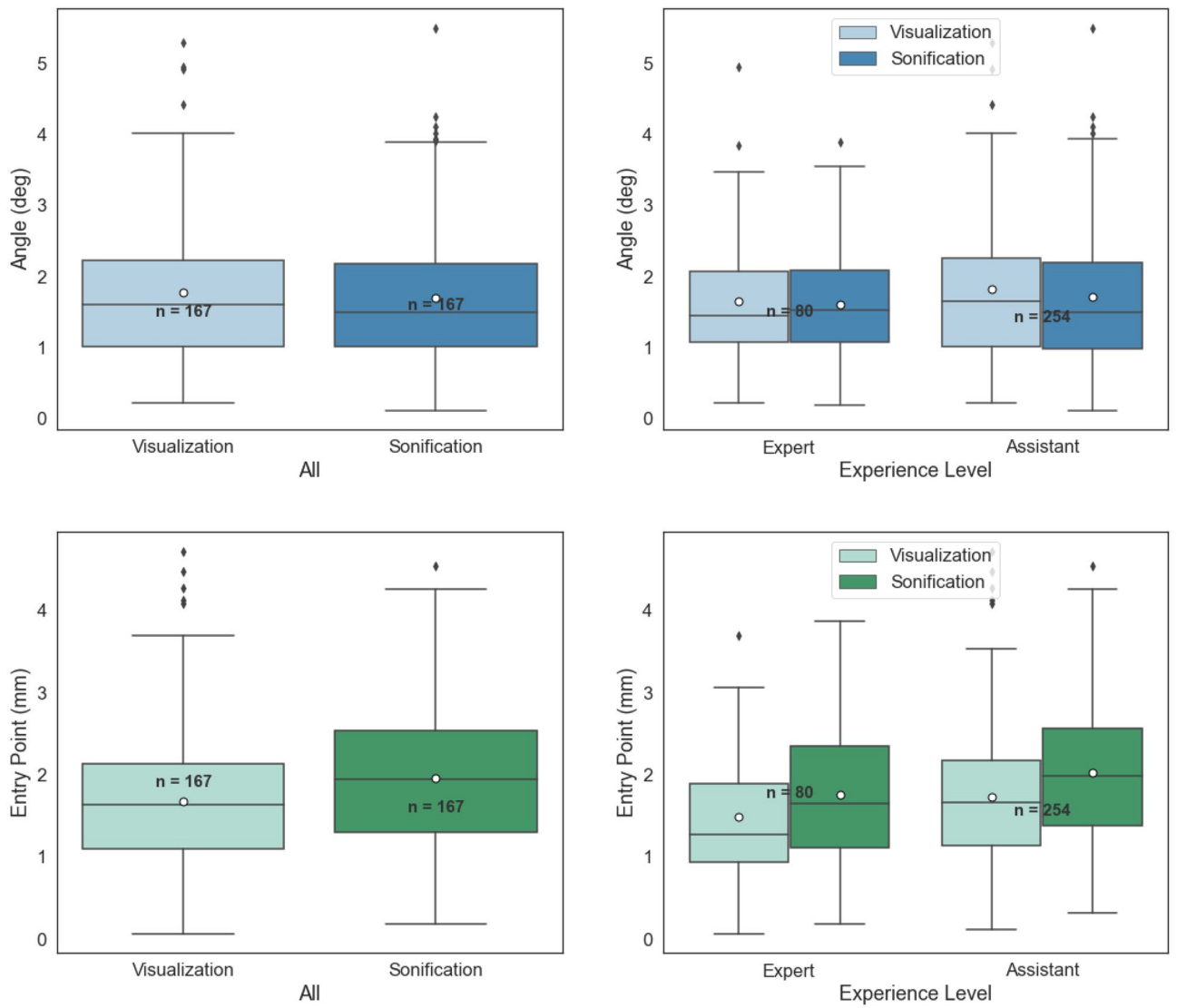
Figure 5. The mean (CT) error of angle (Top), entry point (Bottom) over expertise groups. The white dots in the middle of box plots represent the mean.
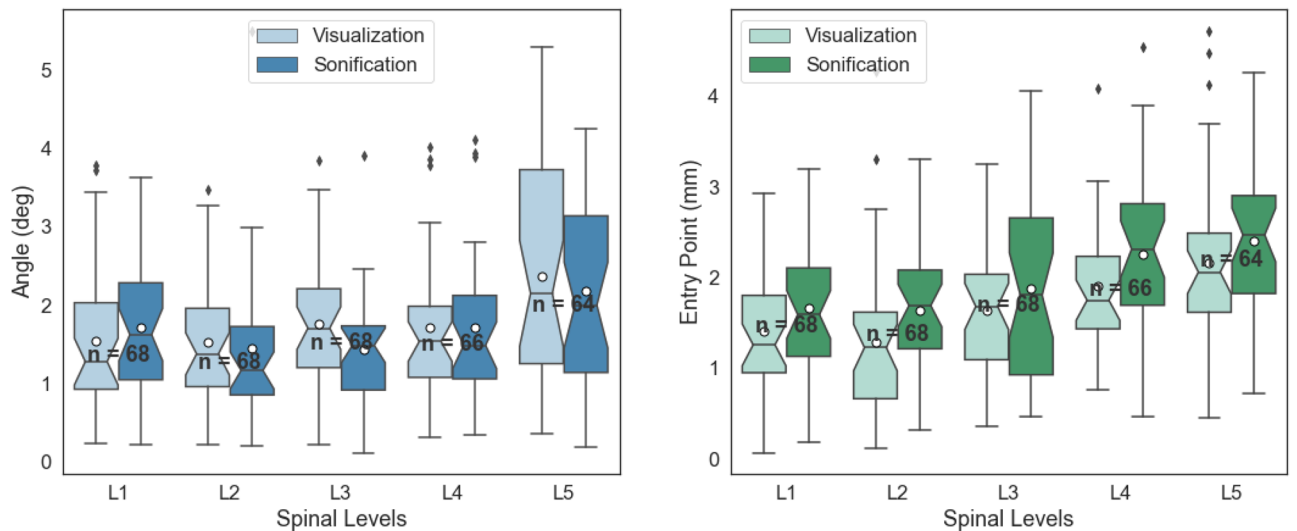

Figure 6. The mean (CT) error of angle on the left and entry point on the right, per spinal levels. The white dots in the middle of box plots represent the mean.

Level			L1	L2	L3	L4	L5
CT error	Visualization	EP (mm)	1.4	1.29	1.63	**1.91**	**2.16**
		ANG (°)	1.55	1.52	1.76	**1.72**	**2.37**
	Sonification	EP (mm)	1.67	1.64	1.87	**2.26**	**2.4**
		ANG (°)	1.71	1.45	1.43	**1.71**	**2.18**
feedback error	Visualization	EP (mm)	0.51	**0.54**	**0.63**	**0.71**	**0.96**
		ANG (°)	1.02	**0.81**	**1.02**	**0.84**	**1.07**
	Sonification	EP (mm)	0.77	0.84	**1.01**	**0.99**	**1.26**
		ANG (°)	0.88	0.78	**0.76**	**0.8**	**0.82**

Table 2. The mean error of angle (ANG) and entry point (EP) over the vertebrae levels L1–L5. The bold cells satisfy the safety requirement suggested by Rampersaud et al.[84].

for sonification, and one believed there was no difference between both methods. In response to Q3 (overall usability), 6 individuals responded that visualization was more usable than sonification, 5 were of the opinion that both methods were equally usable, three believed the sonification method was better than visualization in terms of usability, and no one chose the option "none of the methods are acceptable". Finally, in response to Q4 (which method would you like to use in the future), the majority of respondents (12 individuals) preferred a system that combines both methods, one voted for each of visualization and sonification, and no one chose for the option "none of them".

Discussion

We proposed four-DOF sonification as a novel method for pedicle screw placement, investigating an alternative method toward multisensory assistive technology in the surgical context. The challenge was to design a clinically compatible and accurate system that simultaneously fulfills usability requirements and is competitive with the more conventional visual peer. The results of the comparison study against the state of the art, as demonstrated in "Evaluation and results" Section, offer clear support of the idea behind this research.

Accuracy. Many clinical and anatomic studies have considered the accuracy of pedicle screw placement as the rate of successful screw placements. A successful screw placement has often been referred as the one fully contained in the pedicle cortex without any degree of perforation. The violating degrees of misplacement have been defined as: < 2 mm (Grade A), 2–4 mm (Grade B), and > 4 mm (Grade C)[19,21,85]. Considering this definition, the accuracy of pedicle screw placement using 2D and 3D visual navigation has been reported at 84.3% and 95.5% respectively[21]. To determine whether a particular system will enable the safe performance of the task, we need to specify the safety requirements, as well. The clinical safety requirements are dependent on the type of procedure and the patient's anatomy. The margin of error for a given pedicle is dependent on different factors, such as the size of the screw and the critical dimensions of the pedicle, such as isthmus. Rampersaud et al. [84] proposed a mathematical analysis method for calculating safe margins for the pedicle screw placement task: the maximum entry point and angular error tolerances for L1–L5, given 6.5-mm pedicle screws, are 0.65–3.8 mm and $2.1 - 12°$, respectively. The entry point error was defined as the distance between the actual screw insertion point and the ideal starting point for the screw (at the central axis of the pedicle) and the angular error as the angular deviation between the screw trajectory and the ideal trajectory (parallel to the central pedicle axis). For the same pedicle, the error tolerances increase when using a smaller-diameter screw. Similar to this approach, we calculated the error based on the deviations of the actual drilling trajectories from the preplanned targets (the ideal trajectories).

The overall accuracy of our navigation setup needs to be sufficiently appropriate for the pedicle screw placement task such that we can conduct a valid comparison between the sonification and visualization methods. The accuracies for L4 and L5 in both modalities satisfy the accuracy requirements suggested by Rampersaud et al.[84], as highlighted in Table 2. Conversely, the results for L1–L3 do not fully meet these requirements (assuming a 6.5-mm-diameter screw). However, during the actual procedure, the surgeon first drills a guiding hole, with 3-mm diameter in our case, and then inserts a wider screw, with 6.5-mm diameter, which enables the surgeon to manually refine the trajectory based on haptic feedback and the mechanical constraints of the pedicle wall. Therefore, the practical safety thresholds would provide slightly higher tolerance than the suggested thresholds of Rampersaud's. Moreover, as the state-of-the-art navigation method for pedicle screw placement has not yet provided a 100% success rate, we conclude that the navigation setup has provided an acceptable range of accuracy to compare the sonification and visualization methods. The evaluation of the sonification condition's error (1.96 mm, 1.69°) indicated a similar accuracy to the visualization condition (1.67 mm, 1.78°), both demonstrating a better result in comparison with those in[40] (3.35 mm, 2.74°) and[43] (2.77 mm, 3.38°). Considering the estimated system error (0.98 mm, 0.73°), which includes registration errors, calibration errors, and tracking data noise, we assume the lower error boundary of 0.65 mm and 0.84° for visualization and 0.96 mm and 0.79° for sonification.

The first step in assessing accuracy was to conduct t-tests for independent samples, with the null hypothesis of equal means. Since the system feedback supports users in finding the target, we do not assume a non-normal distribution for alignment error. According to the Shapiro-Wilk test with the CI ($P < 0.05$), the experts' samples for entry point in both V and S ($n = 38$) indicated no evidence of non-normal distribution. In contrast,
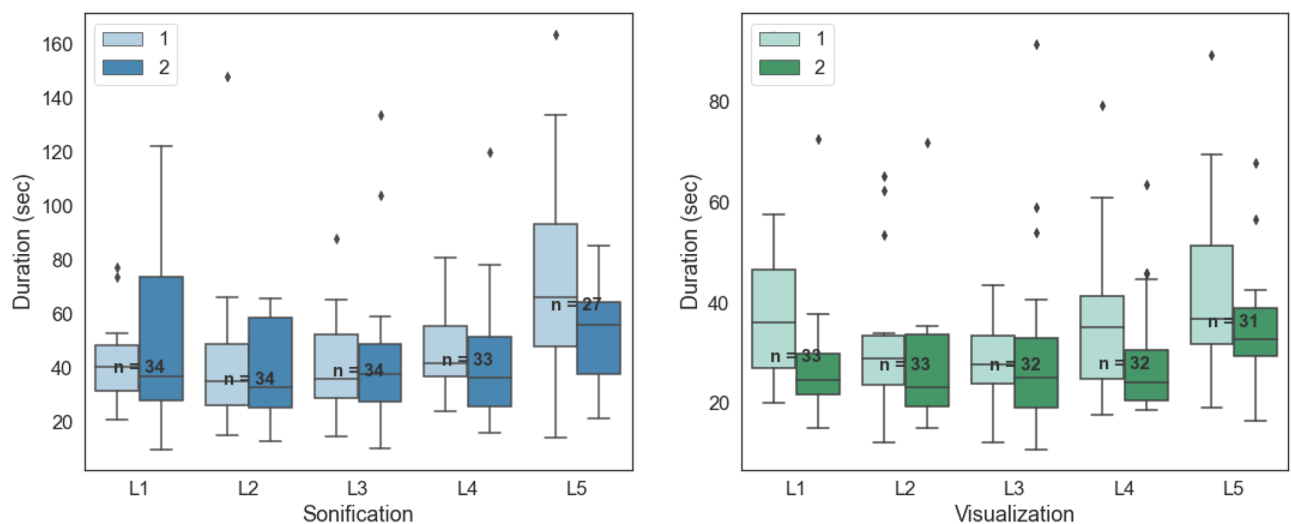
Figure 7. Alignment time of the 1st. and 2nd. executions over the spinal levels.

Group	All ($n = 155$)		Experts ($n = 38$)		Assistants ($n = 117$)	
Error (CI $p < 0.05$)	EP	ANG	EP	ANG	EP	ANG
CT	±0.46 mm	±0.23°	±0.51 mm	±0.29°	±0.5 mm	±0.28°
Feedback	±0.39 mm	±0.2°	±0.45 mm	±0.24°	±0.41 mm	±0.23°
Error (CI $p < 0.025$)						
CT	±0.49 mm	±0.26°	±0.55 mm	±0.35°	±0.54 mm	±0.31°
Feedback	±0.41 mm	±0.21°	±0.48 mm	±0.27°	±0.42 mm	±0.24°

Table 3. The least EI of the TOST to reach the CI $P < 0.05$, and $P < 0.025$ (the adjusted CI using the Bonferroni correction method due to the multiple comparisons problem), over the expertise groups, including both CT and feedback errors for entry point (EP) and angle (ANG).

the Shapiro-Wilk test showed a deviation from normality among novices for entry point and angle errors in V and for angle error in S. One possible explanation is that novices exhibit relatively unstable performance, which results in more extreme values. Furthermore, the sample size of novices ($n = 117$) is large enough that makes the Shapiro-Wilk detect the smallest deviation from normality, with a p value of < 0.05, regardless of whether the variable is generally expected to be normally distributed. Considering these two facts, we conducted a Lilliefors test (a normality test based on the Kolmogorov-Smirnov test), which did not demonstrate evidence of non-normality for novices. The error over all participants does not follow a normal distribution because it includes the samples from both expertise groups with different average of accuracy. However, because the levene test failed to reject any variance differences, which means all other t-test assumptions are satisfied in this case and due to the advantages of parametric tests, we conducted t-tests over all participants as well. The corresponding results of the t-tests for the entry point over experts and angle over both expertise groups failed to reject the null hypothesis. It means, the results of the t-tests did not demonstrate any significant difference between both methods in these cases, which is consistent with the research's hypothesis.

Investigating to what extent both methods have a similar effect, we applied the equivalence test for two independent samples, which is the two one-sided t-test (TOST). TOST works on an equivalence interval (EI) with lower and upper limits ($-\Delta_L, \Delta_U$), and two composite null hypotheses H0-1: $\Delta \leq -\Delta_L$ and H0-2: $\Delta \geq \Delta_U$. If both hypothesis tests can statistically be rejected, we can conclude that the difference between sonification and visualization samples, Δ, falls within the EI $-\Delta_L < \Delta < \Delta_U$, which is considered equivalent. The results indicate statistical significant equivalence of both methods within the EI ± 0.46 mm, $\pm 0.23°$ ($P < 0.05$, $n = 155$). Adding the upper limit of the resultant EI to the sonification error, we can estimate errors of 2.42 mm and 1.92°, which is still comparable with the state-of-the-art visual navigation[40,41]. Details of the EI for different expertise groups for actual and feedback errors are presented in Table 3. In general, we observed a larger EI for the entry point compared with the angle, which is because the thresholds for entry point 2 mm were set larger, compared with the angle 1.5°. Considering that the threshold mechanism in sonification does not allow the user to obtain any feedback after the threshold level, discussion about accuracy after this level would not be relevant, and the sonification outcome could present a random effect. The thresholds were empirically set during the pilot study as a compromise between accuracy and user's convenience. Having a system design with more accurate tracking, the thresholds can also be reduced, leading to a more accurate result for sonification.

10

Group	All ($n = 77$)	Experts ($n = 19$)	Assistants ($n = 58$)
Visualization	< 0.001	< 0.001	< 0.001
Sonification	0.18	< 0.001	0.6

Table 4. The *P* values of the Wilcoxon signed-rank tests on the alignment time of the 1st. and 2nd. executions.

According to the resultant equivalent intervals and the safety requirements presented in the literature[84], we conclude that the sonification method is significantly as accurate as the visualization method for the pedicle screw placement task, which supports the research's primary hypothesis.

Learning curve.   To determine the training effect in both expertise groups, we compared completion time and errors, as functions of performance, on two consecutive executions on the same vertebrae level. To confirm the learnability, we have to determine whether the mean duration of the second execution decreased compared with that of the first execution, without significant decrease in accuracy. Hence, we conducted Wilcoxon signed-rank tests on the alignment time of both executions. The *P* values for each expertise group are shown in Table 4. Moreover, to determine any decrease in accuracy, we performed Wilcoxon signed-rank tests ($\alpha < 0.05$) on both entry point and angle errors, which failed to reject the null hypothesis in terms of equal means for both samples. The mean differences between both executions are shown in Table 5.

The mean difference between both errors (Table 5) and the results of the Wilcoxon signed-rank tests indicate that the accuracy remained consistent during both executions for all expertise groups. As presented in Tables 4 and 5, the experts demonstrated a significant decrease in time for visualization and sonification. We can observe this pattern in the less experienced group only for visualization ($P < 0.001$ in Tables 4 and 5). Considering the error consistency and the observed time patterns, we can conclude high learnability for the both modalities for the experts; however, this was demonstrated by the assistants only for the visualization condition.

Our interpretation is that the experience level of the expert group enabled them to focus more on learning the untrained auditory navigation method. Conversely, the assistant surgeons required more of their cognitive processing capacity for executing the screw drilling task, and therefore, they had a negative learning rate in the sonification condition. Because the visual navigation was more familiar for both groups, they could improve their speed on this modality. However, further research is required to accurately evaluate the effect of training for the sonification method and for developing a full picture of its learning curve.

Multisensory processing and research outlook.   Although the plurality of the questionnaire respondents preferred visualization (Q1–Q3), the absolute majority (12 out of 14) imagined that a desirable future system would combine the advantages of both modalities. Such responses from the field's experts are absolutely in accordance with the equivalent results of the accuracy, NASA-TLX assessments, and principles of multisensory perception. Multisensory solutions result in increased performance and recall, in particular, in intense and complex sensory scenarios[86–88]. Research questions that could be asked in future studies include, e.g., *To what extent can each modality convey complex information accurately? When is it better, or preferable, for the perceptual modalities to be presented in a complementary fashion, and in which situations do they have to provide redundant contextual information? How do our decisional resources respond to each perceptual cue?* Future research in computer-assisted surgery can focus on investigating the possible answers to such fundamental questions in the application field of surgical navigation, as the foundation is well established in cognitive science[89,90].

Further, previous sonification research[50–56,76] has investigated different sonification strategies for navigation, providing the preliminary basis for further research. Future studies should take the cons and pros of sonification paradigms into account. For instance, spatial sonification as an intuitive and natural method with a high learnability rate is suitable for orientation tasks[54,76]; however, we cannot not disregard its limitations with respect to resolution[78]. Additionally, spatial sonification may cause localization anomalies such as front–back confusion, a vague distance and elevation perception, and orientation errors[91,92]. On the other hand, monaural sonification as a candidate approach provides efficiency[78,79] and flexibility in design. Nonetheless, it requires a more prolonged learning phase, which can also depend on the design concept and parameterization.

Four-DOF sonification is a useful tool for scenarios with complex dimensionality and accuracy challenges, as demanded by surgical applications. We divided multiple dimensions into subsets and controlled switching between them using a threshold mechanism. This idea is expandable to contexts with higher dimensionality.

Group	All	Experts	Assistants
Visualization	5.96 sec,	4.97 sec,	6.29 sec,
	0.15 mm, −0.11°	0.072 mm, 0.07°	0.17 mm, −0.17°
Sonification	1.58 sec,	7.07 sec,	−0.26 sec,
	−0.09 mm, −0.04°	0.02 mm, 0.02°	−0.12 mm, −0.05°

Table 5. The mean of differences for entry point (CT) error, angle (CT) error, and alignment time between the 1st. and 2nd. executions for each expertise group.

Nonetheless, issues such as intuitiveness and the learning process have to be considered. Finally, we should consider that monaural sonification is a rather new approach, evolving in terms of dimensionality and interaction design. Even though the presented study exhibited promising results, future research should investigate the effect of enhanced learning phases on performance.

Conclusions

In this article, we have outlined the problem of complex data perception in high-intensity environments, such as the operating room, and highlighted the importance of multisensory processing and development of creative solutions to overcome the information overload issue. To investigate the effects of multisensory processing, we conducted a study with 17 medical professionals in a lab environment using a spinal bone phantom and compared two different techniques for surgical navigation assistance. We proposed the four-DOF sonification method – as a stand-alone audio-based solution – for navigating pedicle screw placement in spinal fusion surgery and compared the method with state-of-the-art visual navigation. Four-DOF sonification did not demonstrate any statistically significant differences in performance compared with the visual navigation. Considering the resultant equivalence intervals and the safety requirements, we can conclude that the proposed method can be reliably used for the pedicle screw placement task. However, the results of the secondary metrics such as cognitive load, usability, and learnability, despite not reaching statistical significance, provided evidences which have led to valuable discussions that could open new paths for further research for interdisciplinary teams of biomedical engineers, cognitive psychologists, sonification designers, and medical experts. The novel design concept of the method supports the idea for accurate sonification of high-dimensional data within a complex interactive task scenario. This study is the first step toward enhancing our understanding of perceptual multisensory processing in the surgical context ("Supplemantary infromation").

Data availability

The datasets generated during and/or analysed during the current study including the code for data analysis are available from the corresponding author on reasonable request. The SuperCollider script for sonifying the error parameters is documented and provided in the supplementary material.

References

1. Middlebrooks, J. C. & Green, D. M. Sound localization by human listeners. *Annu. Rev. Psychol.* **42**, 135–59 (1991).
2. Dixon, B. J. *et al.* Surgeons blinded by enhanced navigation: The effect of augmented reality on attention. *Surg. Endosc.* **27**, 454–461 (2013).
3. Mack, A. Inattentional blindness: Looking without seeing. *Curr. Dir. Psychol. Sci.* **12**, 180–184 (2003).
4. Alais, D., Newell, F. & Mamassian, P. Multisensory processing in review: From physiology to behaviour. *Seeing Perceiving* **23**, 3–38 (2010).
5. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**, 411–417 (2008).
6. Kramer, G. An introduction to auditory display. *Audit. Displ. Sonification Audification Audit. Interfaces* 1–77 (1994).
7. Dombois, F., Eckel, G., Hermann, T., Hunt, A. & Neuhoff, J. G. *The Sonification Handbook* (2011).
8. Hermann, T. Taxonomy and definitions for sonification and auditory display (International Community for Auditory Display, 2008).
9. Hermann, T., Hunt, A. & Neuhoff, J. G. *The Sonification Handbook* (Logos Verlag, Berlin, 2011).
10. Franinovic, K. & Serafin, S. *Sonic Interaction Design* (MIT Press, Cambridge, 2013).
11. O'Brien, B., Hardouin, R., Rao, G., Bertin, D. & Bourdin, C. Online sonification improves cycling performance through kinematic and muscular reorganisations. *Sci. Rep.* **10**, 1–11 (2020).
12. Raglio, A. *et al.* Hand rehabilitation with sonification techniques in the subacute stage of stroke. *Sci. Rep.* **11**, 1–11 (2021).
13. Kather, J. N. *et al.* Polyphonic sonification of electrocardiography signals for diagnosis of cardiac pathologies. *Sci. Rep.* **7**, 1–6 (2017).
14. Rajaee, S. S., Bae, H. W., Kanim, L. E. & Delamarter, R. B. Spinal fusion in the united states: Analysis of trends from 1998 to 2008. *Spine* **37**, 67–76 (2012).
15. Mirza, S. K. & Deyo, R. A. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine* **32**, 816–823 (2007).
16. Verma, K., Boniello, A. & Rihn, J. Emerging techniques for posterior fixation of the lumbar spine. *JAAOS J. Am. Acad. Orthop. Surg.* **24**, 357–364 (2016).
17. Barnes, A. H. *et al.* Biomechanical pullout strength and stability of the cervical artificial pedicle screw. *Spine* **34**, E16–E20 (2009).
18. Tian, W. *et al.* Technical committee on medical robot engineering of Chinese society of biomedical engineering; technical consulting committee of national robotic orthopaedic surgery application center. Guideline for thoracolumbar pedicle screw placement assisted by orthopaedic surgical robot. *Orthop. Surg.* **11**, 153–159 (2019).
19. Gelalis, I. D. *et al.* Accuracy of pedicle screw placement: A systematic review of prospective in vivo studies comparing free hand, fluoroscopy guidance and navigation techniques. *Eur. Spine J.* **21**, 247–255 (2012).
20. Yu, T. *et al.* Accuracy of pedicle screw placement in posterior lumbosacral instrumentation by computer tomography evaluation: A multi-centric retrospective clinical study. *Int. J. Surg.* **43**, 46–51 (2017).
21. Mason, A. *et al.* The accuracy of pedicle screw placement using intraoperative image guidance systems: A systematic review. *J. Neurosurg. Spine* **20**, 196–203 (2014).
22. Modi, H. N., Suh, S. W., Fernandez, H., Yang, J. H. & Song, H.-R. Accuracy and safety of pedicle screw placement in neuromuscular scoliosis with free-hand technique. *Eur. Spine J.* **17**, 1686–1696 (2008).
23. Avila, M. J. & Baaj, A. A. Freehand thoracic pedicle screw placement: Review of existing strategies and a step-by-step guide using uniform landmarks for all levels. *Cureus* **8**, e501 (2016).
24. Açıkbaş, S., Arslan, F. & Tuncer, M. The effect of transpedicular screw misplacement on late spinal stability. *Acta Neurochir.* **145**, 949–955 (2003).
25. Verma, R., Krishan, S., Haendlmayer, K. & Mohsen, A. Functional outcome of computer-assisted spinal pedicle screw placement: A systematic review and meta-analysis of 23 studies including 5,992 pedicle screws. *Eur. Spine J.* **19**, 370–375 (2010).

26. Tian, N.-F. *et al.* Pedicle screw insertion accuracy with different assisted methods: A systematic review and meta-analysis of comparative studies. *Eur. Spine J.* **20**, 846–859 (2011).
27. Kosmopoulos, V. & Schizas, C. Pedicle screw placement accuracy: A meta-analysis. *Spine* **32**, E111–E120 (2007).
28. Gebhard, F., Weidner, A., Liener, U. C., Stöckle, U. & Arand, M. Navigation at the spine. *Injury* https://doi.org/10.1016/j.injury. 2004.05.009 *(2004).*
29. Luther, N. *et al.* Comparison of navigated versus non-navigated pedicle screw placement in 260 patients and 1434 screws. *J. Spinal Disord. Tech.* **28**, E298–E303 (2015).
30. Kraus, M. D., Krischak, G., Keppler, P., Gebhard, F. T. & Schuetz, U. H. Can computer-assisted surgery reduce the effective dose for spinal fusion and sacroiliac screw insertion?. *Clin. Orthop. Relat. Res.* **468**, 2419–2429 (2010).
31. Narain, A. S. *et al.* Radiation exposure and reduction in the operating room: Perspectives and future directions in spine surgery. *World J. Orthop.* **8**, 524 (2017).
32. Gebhard, F. T. *et al.* Does computer-assisted spine surgery reduce intraoperative radiation doses?. *Spine* **31**, 2024–2027 (2006).
33. Härtl, R. *et al.* Worldwide survey on the use of navigation in spine surgery. *World Neurosurg.* **79**, 162–172 (2013).
34. Helm, P. A., Teichman, R., Hartmann, S. L. & Simon, D. Spinal navigation and imaging: History, trends, and future. *IEEE Trans. Med. Imaging* **34**, 1738–1746 (2015).
35. Ryang, Y.-M. *et al.* Learning curve of 3d fluoroscopy image-guided pedicle screw placement in the thoracolumbar spine. *Spine J.* **15**, 467–476 (2015).
36. Herrlich, M. *et al.* Instrument-mounted displays for reducing cognitive load during surgical navigation. *Int. J. Comput. Assist. Radiol. Surg.* **12**, 1599–1605 (2017).
37. Sielhorst, T., Feuerstein, M. & Navab, N. Advanced medical displays: A literature review of augmented reality. *J. Disp. Technol.* **4**, 451–467 (2008).
38. Navab, N., Blum, T., Wang, L., Okur, A. & Wendler, T. First deployments of augmented reality in operating rooms. *Computer* **45**, 48–55 (2012).
39. Jud, L. *et al.* Applicability of augmented reality in orthopedic surgery-a systematic review. *BMC Musculoskelet. Disord.* **21**, 1–13 (2020).
40. Ma, L. *et al.* Augmented reality surgical navigation with ultrasound-assisted registration for pedicle screw placement: A pilot study. *Int. J. Comput. Assist. Radiol. Surg.* **12**, 2205–2215 (2017).
41. Müller, F. *et al.* Augmented reality navigation for spinal pedicle screw instrumentation using intraoperative 3D imaging. *Spine J.* **20**, 621–628 (2020).
42. Dennler, C. *et al.* Augmented reality-based navigation increases precision of pedicle screw insertion. *J. Orthop. Surg. Res.* **15**, 1–8 (2020).
43. Liebmann, F. *et al.* Pedicle screw navigation using surface digitization on the microsoft hololens. *Int. J. Comput. Assist. Radiol. Surg.* **14**, 1157–1165 (2019).
44. Kassil, K. & Stewart, A. J. Evaluation of a tool-mounted guidance display for computer-assisted surgery. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 1275–1278 (2009).
45. Bellego, G. L., Bucki, M., Bricault, I. & Troccaz, J. Using a smart phone for information rendering in computer-aided surgery. In *International Conference on Human-Computer Interaction* 202–209 (Springer, 2011).
46. Mullaji, A. B. & Shetty, G. M. Efficacy of a novel iPod-based navigation system compared to traditional navigation system in total knee arthroplasty. *Comput. Assist. Surg.* **22**, 10–13 (2017).
47. Brendle, C. *et al.* Can a hand-held navigation device reduce cognitive load? a user-centered approach evaluated by 18 surgeons. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* 399–408 (Springer, 2020).
48. Nagarajan, R., Yaacob, S. & Sainarayanan, G. Role of object identification in sonification system for visually impaired. In *TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region* Vol. 2 735–739 (IEEE, 2003).
49. Giudice, N. A. & Legge, G. E. Blind navigation and the role of technology. In *The Engineering Handbook of Smart Technology for Aging, Disability, and Independence* Vol. 8 479–500 (Wiley, Hoboken, 2008).
50. Hu, W. *et al.* A comparative study in real-time scene sonification for visually impaired people. *Sensors* **20**, 3222 (2020).
51. Skulimowski, P. *et al.* Interactive sonification of u-depth images in a navigation aid for the visually impaired. *J. Multimodal User Interfaces* **13**, 219–230 (2019).
52. Brock, M. & Kristensson, P. O. Supporting blind navigation using depth sensing and sonification. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication* 255–258 (2013).
53. Ahmetovic, D. *et al.* Sonification of rotation instructions to support navigation of people with visual impairment. In *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom)* 1–10 (IEEE, 2019).
54. Parseihian, G., Gondre, C., Aramaki, M., Ystad, S. & Kronland-Martinet, R. Comparison and evaluation of sonification strategies for guidance tasks. *IEEE Trans. Multimed.* **18**, 674–686 (2016).
55. Ziemer, T., Black, D. & Schultheis, H. Psychoacoustic sonification design for navigation in surgical interventions. In *Proceedings of Meetings on Acoustics 173EAA* Vol. 30 050005 (Acoustical Society of America, 2017).
56. Ziemer, T., Schultheis, H., Black, D. & Kikinis, R. Psychoacoustical interactive sonification for short range navigation. *Acta Acust. Acust.* **104**, 1075–1093 (2018).
57. Matinfar, S. *et al.* Surgical soundtracks: Towards automatic musical augmentation of surgical procedures. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* 673–681 (Springer, 2017).
58. Matinfar, S. *et al.* Surgical soundtracks: Automatic acoustic augmentation of surgical procedures. *Int. J. Comput. Assist. Radiol. Surg.* **13**, 1345–1355 (2018).
59. Matinfar, S. *et al.* Sonification for process monitoring in highly sensitive surgical tasks. In *Proceedings of the Nordic Sound and Music Computing Conference 2019 (Nordic SMC 2019)* (2019).
60. Black, D., Hansen, C., Nabavi, A., Kikinis, R. & Hahn, H. A survey of auditory display in image-guided interventions. *Int. J. Comput. Assist. Radiol. Surg.* **12**, 1665–1676 (2017).
61. Wegner, K. Surgical navigation system and method using audio feedback. In *International Conference on Auditory Display'98* 1–10 (1998).
62. Dixon, B. J. *et al.* Augmented real-time navigation with critical structure proximity alerts for endoscopic skull base surgery. *Laryngoscope* **124**, 853–859 (2014).
63. Voormolen, E. H. *et al.* Validation of exposure visualization and audible distance emission for navigated temporal bone drilling in phantoms. *PLoS ONE* **7**, e41262 (2012).
64. Cho, B., Matsumoto, N., Komune, S. & Hashizume, M. A surgical navigation system for guiding exact cochleostomy using auditory feedback: A clinical feasibility study. *BioMed Res. Int.* https://doi.org/10.1155/2014/769659 *(2014).*
65. Cho, B. *et al.* Warning navigation system using real-time safe region monitoring for otologic surgery. *Int. J. Comput. Assist. Radiol. Surg.* **8**, 395–405 (2013).
66. Utsuki, S. *et al.* Auditory alert system for fluorescence-guided resection of gliomas-technical note-. *Neurol. Med. Chir.* **48**, 95–98 (2008).
67. Hansen, C. *et al.* Auditory support for resection guidance in navigated liver surgery. *Int. J. Med. Robot. Comput. Assist. Surg.* **9**, 36–43 (2013).

68. Black, D., Al Issawi, J., Hansen, C., Rieder, C. & Hahn, H. K. Auditory support for navigated radiofrequency ablation. In *CURAC* 30–33 (2013).
69. Bork, F. *et al.* Auditory and visio-temporal distance coding for 3-dimensional perception in medical augmented reality. In *2015 IEEE International Symposium on Mixed and Augmented Reality* 7–12 (IEEE, 2015).
70. Black, D. *et al.* Auditory feedback to support image-guided medical needle placement. *Int. J. Comput. Assist. Radiol. Surg.* **12**, 1655–1663 (2017).
71. Plazak, J., Drouin, S., Collins, L. & Kersten-Oertel, M. Distance sonification in image-guided neurosurgery. *Healthc. Technol. Lett.* **4**, 199–203 (2017).
72. Roodaki, H., Navab, N., Eslami, A., Stapleton, C. & Navab, N. Sonifeye: Sonification of visual information using physical modeling sound synthesis. *IEEE Trans. Visual Comput. Graph.* **23**, 2366–2371 (2017).
73. Barrass, S. A perceptual framework for the auditory display of scientific data. *ACM Trans. Appl. Percept. (TAP)* **2**, 389–402 (2005).
74. Vasilijevic, A., Jambrosic, K. & Vukic, Z. Teleoperated path following and trajectory tracking of unmanned vehicles using spatial auditory guidance system. *Appl. Acoust.* **129**, 72–85 (2018).
75. Lokki, T. & Grohn, M. Navigation with auditory cues in a virtual environment. *IEEE Multimed.* **12**, 80–86 (2005).
76. Ziemer, T. & Schultheis, H. Psychoacoustical signal processing for three-dimensional sonification (Georgia Institute of Technology, 2019).
77. Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, 1997).
78. Zwicker, E. & Fastl, H. Facts and models. *Phys. Today* **64** (1999).
79. Pedrielli, F., Carletti, E. & Casazza, C. Just noticeable differences of loudness and sharpness for earth moving machines. *J. Acoust. Soc. Am.* **123**, 3164 (2008).
80. Chowning, J. M. The synthesis of complex audio spectra by means of frequency modulation. *J. Audio Eng. Soc.* **21**, 526–534 (1973).
81. Yaniv, Z. R. & Holmes III, D. R. *Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling* Vol. 9036 (2014).
82. Wright, M., Freed, A. *et al.* Open soundcontrol: A new protocol for communicating with sound synthesizers. In *ICMC* (1997).
83. Wein, W., Brunke, S., Khamene, A., Callstrom, M. R. & Navab, N. Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Med. Image Anal.* **12**, 577–585 (2008).
84. Rampersaud, Y. R., Simon, D. A. & Foley, K. T. Accuracy requirements for image-guided spinal pedicle screw placement. *Spine* **26**, 352–359 (2001).
85. Gertzbein, S. D. & Robbins, S. E. Accuracy of pedicular screw placement in vivo. *Spine* **15**, 11–14 (1990).
86. Van der Burg, E., Olivers, C. N., Bronkhorst, A. W. & Theeuwes, J. Pip and pop: Nonspatial auditory signals improve spatial visual search. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 1053 (2008).
87. Ngo, M. K. & Spence, C. Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Atten. Percept. Psychophys.* **72**, 1654–1665 (2010).
88. Barutchu, A., Sahu, A., Humphreys, G. W. & Spence, C. Multisensory processing in event-based prospective memory. *Acta Physiol.* **192**, 23–30 (2019).
89. Ernst, M. O. & Di Luca, M. Multisensory perception: From integration to remapping. In *Sensory Cue Integration* 224–250 (Oxford University Press, Oxford, 2011).
90. Helbig, H. B. & Ernst, M. O. Haptic perception in interaction with other senses. In *Human Haptic Perception: Basics and Applications* 235–249 (Springer, Berlin, 2008).
91. Rumsey, F. Spatial audio: Binaural challenges. *J. Audio Eng. Soc.* **62**, 798–802 (2014).
92. Bertet, S., Daniel, J., Parizet, E. & Warusfel, O. Investigation on localisation accuracy for first and higher order ambisonics reproduced sound sources. *Acta Acust. Acust.* **99**, 642–657 (2013).

Author contributions

S.M. and M.S. proposed and implemented method, evaluated the data and analysed the results, wrote the main manuscript text, and prepared all figures and tables. D.S. and MT.S. supported the experiments from the clinical side. S.D. consulted and reviewed the mathematical formulation of the method. NV.N. consulted the work from the sound design, perceptual, and cognitive aspects. M.F. and F.W. supervised the work and experiment design from the clinical side. N.N. and P.F. supervised the work from the technical side. All authors reviewed the manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-32778-z.

Correspondence and requests for materials should be addressed to S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.