

Physics-Based iOCT Sonification for Real-Time Interaction Awareness in Subretinal Injection

Luis D. Reyes Vargas^{*1,5}, Veronica Ruozzi^{*1,5}, Andrea K. M. Ross²,
Shervin Dehghani^{1,5}, Michael Sommersperger^{1,5}, Koorosh Faridpooya³, Mohammad
Ali Nasseri², Merle Fairhurst^{4,6}, Nassir Navab^{1,5}, and Sasan Matinfar^{1,5,6}

¹ Computer Aided Medical Procedures, Technische Universität München, Munich, Germany

² TUM Klinikum Rechts der Isar, Technische Universität München, München, Germany

³ Rotterdam Eye Hospital, Rotterdam, The Netherlands

⁴ Centre for Tactile

Internet with Human-in-the-Loop, Technische Universität Dresden, Dresden, Germany

⁵ Munich Center for Machine Learning, Munich, Germany

⁶ Chair for Social Affective Touch, Technische Universität Dresden, Dresden, Germany

luis.reyes@tum.de

Abstract. Subretinal injection is a delicate vitreoretinal procedure requiring precise needle placement within the subretinal space while avoiding perforation of the retinal pigment epithelium (RPE), a layer directly beneath the target with extremely limited regenerative capacity. To enhance depth perception during cannula advancement, intraoperative optical coherence tomography (iOCT) offers high-resolution cross-sectional visualization of needle–tissue interaction; however, interpreting these images requires sustained visual attention alongside the en face microscope view, thereby increasing cognitive load during critical phases and placing additional demands on the surgeon’s proprioceptive control. In this paper, we propose a structured, real-time sonification framework designed for extensible mapping of iOCT-derived anatomical features into perceptual auditory feedback. The method employs a physics-inspired acoustic model driven by segmented retinal layers from a stream of iOCT B-scans, with needle motion and injection-induced retinal layer displacements serving as excitation inputs to the sound model, enabling perception of tool position and retinal deformation. In a controlled user study ($n=34$), the proposed sonification achieved high retinal layer identification accuracy and robust detection of retinal deformation-related events, significantly outperforming a state-of-the-art baseline in overall event identification (83.4% vs. 60.6%, $p < 0.001$), with gains driven primarily by enhanced detection of injection-induced retinal deformation. Participants also reported higher confidence with the proposed method correlating with correctness, indicating the auditory mapping was perceptually interpretable. Evaluation by experts ($n=4$) confirmed the clinical relevance and potential intraoperative applicability of the method. These results establish structured iOCT sonification as a viable complementary modality for real-time surgical guidance in subretinal injection.

Keywords: Surgical Sonification · Subretinal Injection · Optical Coherence Tomography (OCT) · Auditory Feedback · Microsurgical Guidance

* These authors contributed equally to this work.

1 Introduction

Subretinal injection is a delicate vitreoretinal procedure in which a microsurgical cannula is advanced through the neurosensory retina to deliver therapeutics into the subretinal space. This approach demonstrates significant potential for application in cell- and gene-based therapies [13, 15]. The cannula must first penetrate the Inner Limiting Membrane (ILM) before advancing within the subsequent layers for injection, which in the macular region measures only approximately $250\text{ }\mu\text{m}$ in thickness on average [14]. Immediately beneath the target lies the retinal pigment epithelium (RPE), a layer with very limited regenerative capacity where even minimal overshoot may cause irreversible damage [8, 27]. Because penetration forces are extremely small and masked by friction along the instrument shaft, tactile cues are largely imperceptible, forcing surgeons to rely almost exclusively on visual feedback during insertion [9].

In clinical practice, the en face surgical microscope view serves as the primary guidance modality; however, it provides mainly surface-level information and lacks direct visualization of needle penetration depth relative to the RPE. To address this limitation, intraoperative OCT (iOCT) provides high-resolution cross-sectional views of retinal layers and tool position that are directly relevant for depth control in subretinal injection. In current workflows, iOCT is typically displayed as an overlay on the primary microscope image, requiring simultaneous interpretation of both visual streams. Large prospective studies report that iOCT influences intraoperative decision-making in only a minority of posterior-segment cases ($\approx 30\%$) [6], suggesting that cross-sectional information may not be optimally integrated into routine surgical practice despite its clinical relevance. The addition of this secondary high-bandwidth visual stream may increase cognitive load and promote attentional tunneling during critical phases such as needle insertion [12, 23], when precise depth estimation is essential.

Prior work on OCT-based sonification has focused on epiretinal membrane (ERM) peeling [17]. However, its auditory guidance is designed to support positioning before direct tool-tissue interaction and does not account for needle penetration or dynamic tissue deformation. Beyond localizing the needle, subretinal injection requires monitoring interaction-induced structural changes. In particular, successful fluid delivery produces a localized retinal detachment, commonly referred to as a subretinal bleb, whose formation and expansion must be monitored.

To address this gap, we introduce a real-time, deformation-aware sonification framework for subretinal injection. Unlike prior work relying on simulation of preoperative imaging dynamics [22] or focusing on pre-interaction auditory mappings [17], the proposed method encodes interaction-driven retinal deformation directly from intraoperative OCT using a physics-inspired acoustic model anchored to anatomical segmentations. This structured auditory representation enables unified encoding of needle depth, safety boundaries, and fluid-induced tissue deformation within a single extensible framework. In the context of subretinal injection, we focus on signals corresponding to ILM penetration marking retinal entry, proximity to the RPE as a critical safety boundary, and formation of the subretinal bleb confirming successful delivery. The system operates in real time and is compatible with interactive surgical workflows. Code is available at <https://github.com/luisdavid64/ioc-t-subretinal-sonification>.

We evaluate the proposed framework across simulated and ex vivo porcine iOCT sequences and assess its clinical relevance through expert feedback. In a controlled user study with 30 novices and 4 expert participants, the proposed sonification significantly improved identification of ILM penetration, RPE contact, and subretinal bleb formation compared to a parameter-mapping baseline. The largest gains were observed in detecting deformation associated with bleb formation, supporting the framework’s ability to convey dynamic interaction cues through audio alone and suggesting its potential to preserve visual focus during critical insertion phases.

2 Related Work

Medical Sonification. Medical sonification has demonstrated great promise in augmenting clinical procedures through auditory displays [3, 19]. Early work applied to vitreoretinal surgery discretized the anatomical workspace into zones and used musical cues to convey tool-tip position and procedural state [18]. Subsequent approaches moved beyond musical or simple parameter mappings, introducing model-based sound synthesis frameworks that map imaging-derived signals to parameters of a physics-inspired acoustic model, thereby generating structured auditory feedback [16]. However, these methods typically rely on preoperative imaging and therefore cannot accurately reflect dynamic intraoperative anatomical changes.

A more recent study has explored sonification of tissue deformation signals rather than tool localization alone [22], aiming to provide richer insight into anatomical interaction. However, these approaches rely on simulated deformation derived from preoperative imaging, limiting their ability to reflect patient-specific and interaction-driven dynamics encountered intraoperatively.

OCT-specific sonification has been proposed for ERM peeling, mapping A-scans to a physics-inspired acoustic model to provide structural auditory feedback for identifying gaps between the ERM and ILM during membrane localization and peeling initiation [17]. The auditory feedback primarily supports structural assessment prior to direct tool–tissue interaction and does not encode interaction-driven deformation dynamics relevant to subretinal injection.

iOCT Processing for Subretinal Injection. In parallel, robotic and computer-assisted subretinal injection has emerged as an active research area, driven by physiological limitations such as hand tremor and the difficulty of processing iOCT data in real time [1]. This work has enabled key components for iOCT-guided insertion, including tool localization, real-time segmentation of retinal layers and instrument position, and extraction of tool-aligned B-scans from volumetric OCT data [1, 2, 5, 28]. Although these methods improve geometric estimation and visualization, effective intraoperative communication of this information remains challenging.

3 Method

Let $I_t \in \mathbb{R}^{H \times W}$ denote the incoming iOCT B-scan at frame t . Our goal is to map the image stream to a continuous audio signal $y(\tau)$ that reflects subretinal injection dynamics. The framework consists of (i) an initialization phase at $t=0$, in which a

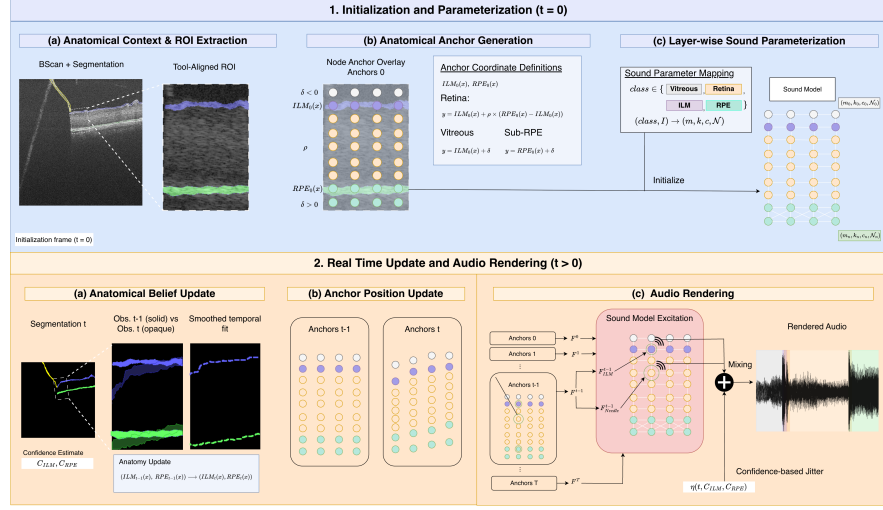


Fig. 1: iOCT Sonification framework

physics-inspired sound model is instantiated and aligned with the segmented retinal anatomy, and (ii) a real-time update and audio rendering loop for $t > 0$.

3.1 Initialization and Parameterization

The underlying sound model is formulated as a two-dimensional mass-spring-damper system that produces audible resonances under external excitation [26], following prior work in medical sonification [17, 22]. The system consists of interconnected masses whose vibrational dynamics are governed by stiffness and damping parameters. During initialization, both the spatial configuration and physical properties are derived from geometric and semantic information extracted from the first segmented B-scan, thereby anchoring the model to the retinal anatomy. These parameters shape the spectral and resonant characteristics of the synthesized sound when nodes are excited, enabling contrastive feedback across tissue layers, implicitly encoding tool-tissue proximity, and intuitively reflecting deformations such as bleb formation.

Anatomical Context and Tool-aligned ROI Extraction We process an initialization frame I_0 acquired during setup, where the needle shaft is visible in iOCT. A coarse segmentation S_0 of the ILM, RPE, and needle is obtained using a pre-trained U-Net, with low-confidence predictions suppressed as in [5]. From S_0 , we estimate the dominant retinal orientation and compute a rotation matrix R to align the tissue horizontally. The needle shaft is modeled by fitting a line ℓ to the segmented needle pixels using a robust Huber regressor [5, 11]. After applying R , a rotated rectangular region of interest (ROI) is extracted around the intersection of ℓ with the anatomical layers and extended to include the vitreous region above the ILM. If direct overlap is

not detected (e.g., due to shadowing), the ROI is estimated from a local neighborhood search or from the needle trajectory alone.

Anatomical Node Generation As shown in Fig. 1b, a uniform grid is overlaid on the ROI to define sound model nodes. Each node is assigned an anatomical class (vitreous, ILM, retina, RPE) via majority voting within its support, with ILM and RPE up-weighted to account for their thin structure.

Node positions are parameterized relative to the segmented ILM_0 and RPE_0 to ensure deformation-consistent alignment. For retinal nodes, axial position is defined by a normalized depth parameter $\rho \in [0,1]$:

$$y = ILM_0(x) + \rho(RPE_0(x) - ILM_0(x)). \quad (1)$$

Nodes outside the retina are defined by offsets δ to the nearest boundary, enabling run-time updates that preserve anatomical consistency of the anchors under deformation.

Layer-wise Sound Parameterization Our $m \times n$ anchor grid defines the masses of a physics-based sound model. To complete the mass-spring-damper system, we assign node-specific physical parameters and define the connectivity between neighboring masses to instantiate spring-damper elements.

Based on anatomical class labels and local intensity statistics within the ROI, we define a hand-crafted mapping

$$\mathbf{M}: (\text{class}, I) \rightarrow (m, k, d, \mathcal{N}), \quad (2)$$

which assigns mass m , stiffness k , and damping d to each node.

Nodes are connected in a regular lattice defined by a neighborhood order $\mathcal{N} \in \{1, 2\}$, indicating first- or second-order ordinal proximity. First-order connectivity links each node to its immediate neighbors along the Cartesian axes, i.e., $(i \pm 1, j)$ and $(i, j \pm 1)$. Second-order connectivity extends coupling to the diagonal neighbors $(i \pm 1, j \pm 1)$.

For each neighboring pair $(i, j) \in \mathcal{N}$, spring stiffness and damping are defined symmetrically as

$$k_{ij} = \frac{1}{2}(k_i + k_j), \quad d_{ij} = \frac{1}{2}(d_i + d_j). \quad (3)$$

3.2 Real-time Anatomical Update ($t > 0$)

Following initialization, the sound model is updated online using the incoming iOCT B-scan stream. At fixed temporal intervals, each frame is segmented to obtain updated ILM_t , RPE_t , and needle estimates with associated confidence measures. To ensure robustness under partial occlusion and shadowing, ILM and RPE trajectories are modeled using cubic spline fitting with confidence-weighted regularization [4, 25] applied along the lateral axis. This lightweight temporal modeling enables smooth extrapolation across missing or low-confidence regions while remaining compatible with real-time processing constraints.

The updated anatomical estimates are then used to recompute node positions according to the relative parameterization introduced in the previous section. In

particular, each node preserves its relative depth parameters (ρ, δ) , while its absolute position is updated based on the current $ILM_t(x)$ and $RPE_t(x)$ geometry. This allows the anchors to follow tissue deformation over time.

3.3 Excitation Protocol and Audio Rendering

The sound model is driven by a set of excitation mechanisms designed to reflect both tool–tissue position and ongoing anatomical changes, while also conveying the reliability of the underlying visual estimates.

Tool-driven Excitation. Excitations are derived from the estimated needle tip position. As the needle advances or interacts with tissue, forces are applied to the nearest node and propagated through the spring–damper lattice. Excitation magnitude is scaled by local stiffness [22], ensuring anatomically consistent response across tissue types. This produces auditory events reflecting needle proximity and contact.

Anatomy-driven Excitation. Surgical events such as tissue penetration or fluid injection induce localized deformation of the retinal layers, which manifests as differential axial motion between the ILM and RPE. Rather than directly measuring deformations in noisy iOCT scans, we estimate a coarse deformation by comparing temporal changes in the positions of these anatomical layers in the vicinity of the needle tip.

At each time step t , we define a lateral window $\mathcal{W}(x_t)$ centered at the current injection site (approximated by the needle tip position x_t). This focuses the analysis on tool–tissue and fluid injection–related interaction regions rather than global anatomical drift. Within this window, we compute the local ILM–RPE separation from the segmentations.

$$d_t(x) = RPE_t(x) - ILM_t(x), \quad x \in \mathcal{W}(x_t). \quad (4)$$

The robust temporal change in separation is then defined as

$$\Delta d_t = P_{95}(d_t(x) - d_{t-1}(x)), \quad x \in \mathcal{W}(x_t), \quad (5)$$

where $P_{95}(\cdot)$ denotes the 95th percentile, used to suppress outliers due to segmentation noise. Finally, we define the deformation-driven excitation proxy

$$f_{ILM} = \min(2, \max(0, \Delta d_t)), \quad (6)$$

which retains responsiveness to positive separation increases corresponding to injection-induced retinal elevation while suppressing compression. The resulting signal is used to modulate excitation strength in the acoustic model. To prevent excessive excitation, f_{ILM} is bounded to a predefined range.

Confidence-modulated rendering. Confidence estimates are used to modulate a stochastic post-processing component in the sound synthesis $\eta(t, C_{ILM}, C_{RPE})$, such that lower confidence results in increased temporal jitter.

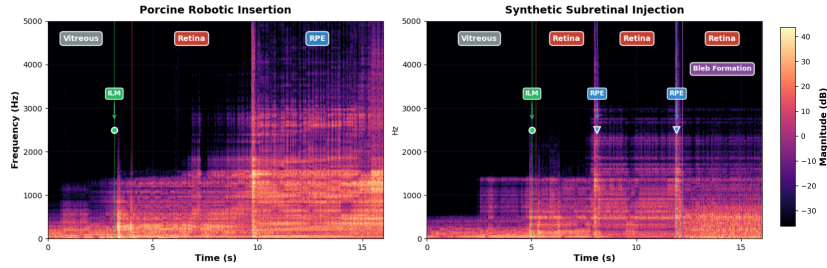


Fig. 2: Spectrograms of sonified iOCT signals from (left) ex vivo porcine robotic insertion and (right) a synthetic subretinal injection sequence. In the synthetic example, bleb formation is characterized by a broadband increase in spectral energy.

4 Experiments and Result

The sound model was implemented using the miPhysics library⁷, and the physical parameter mapping M was designed based on psychoacoustic principles [24].

Qualitative Results We qualitatively evaluate the method on a publicly available porcine insertion [20] and a representative synthetic sequence exhibiting bleb formation (Fig. 2). We provide corresponding supplementary videos for these two cases (S01 and S02). In both cases, detected events align closely with annotated ground truth, indicating accurate temporal localization. We additionally provide an example with vivid confidence modulation (S03).

Real-time System – Runtime Performance We demonstrate the full system operates in real time within an interactive simulated subretinal injection environment (supplementary video S04). We evaluate runtime on a CPU-bound system across 10 insertion sequences at 512×512 resolution. The full pipeline requires 27.8 ± 1.5 ms per frame (mean $\pm$ SD), corresponding to approximately 36 FPS. Segmentation and spline fitting dominated computation time, while all other components required < 1 ms per frame.

User Study We conducted a controlled within-subject study comparing the proposed sonification with a parameter-mapping baseline on realistic simulated subretinal injection sequences. Thirty-five participants were recruited; one was excluded ($< 25\%$ accuracy in both methods), resulting in 34 analyzed participants, from which 30 are novices and 4 are clinical experts. Each completed 20 trials (10 per method), identifying ILM contact, RPE contact, and bleb onset while rating confidence (1–5). Trials included single- and multi-event scenarios requiring correct temporal ordering. Method order was randomized, and trials were stratified by difficulty and randomized within strata. Differences (Δ) are reported as proposed minus baseline and tested using the

⁷ https://github.com/mi-creative/miPhysics_Processing

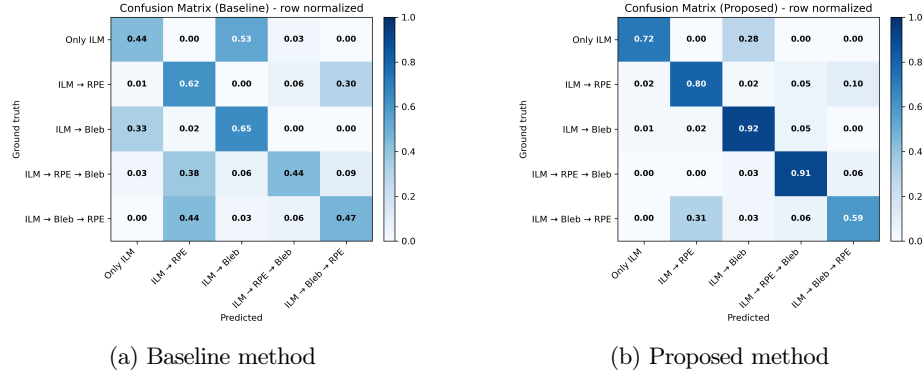


Fig. 3: Row-normalized confusion matrices ($n=34$). The baseline shows entanglement between contact-only and deformation states, while the proposed method improves separation of bleb-related events, reflected by stronger diagonal concentration.

Wilcoxon signed-rank test. **The baseline** builds on pitch-based anatomical encoding strategies from prior surgical sonification frameworks [18,21], where static pitch represents anatomical zones and pulse-rate modulation proportional to ILM–RPE distance provides proximity feedback [19]. Parameterization follows established psychoacoustic principles [7], and prior work has demonstrated its effectiveness in supporting perceptual discrimination of anatomical regions. Bleb formation is conveyed indirectly through distance-dependent modulation rather than explicit deformation encoding.

User Study Results The proposed sonification significantly improved overall event identification accuracy compared to the baseline (83.4% vs. 60.6%; $\Delta = 22.8$ percentage points, 95% CI [15.7, 30.1], $p < 0.001$). Component-wise analysis showed that gains were primarily driven by improved bleb identification (63.6% vs. 85.5%; $\Delta = 21.8$ percentage points, $p < 0.001$), while ILM exhibited a minor numerical increase (+0.9%) and RPE a small decrease (-1.6%), neither statistically significant. As illustrated in Fig. 3, the proposed method reduces structured confusions between contact-only and deformation states related to bleb formation. Participants reported significantly higher confidence on average under the proposed sonification (4.20 vs. 3.73; $\Delta = +0.47$, 95% CI [0.23, 0.71], $p < 0.001$), with the median confidence increasing from 4 to 5. Confidence was positively correlated with correctness in both conditions (Spearman $\rho = 0.27$ for both methods; both $p < 0.001$), indicating that subjective certainty reflected actual performance.

Expert Feedback Two vitreoretinal surgeons and two residents ($n = 4$) from the main cohort provided additional interview feedback. In this group, recognition accuracy increased descriptively from 65% (baseline) to 75% (proposed), with comparable gains in bleb detection. Experts highlighted the method’s clinical relevance and its potential to reduce reliance on visual iOCT interpretation in favor of the optical microscope, potentially alleviating visual load [23].

5 Discussion and Conclusion

We presented a real-time sonification framework for subretinal injection that converts iOCT B-scans into structured auditory feedback. Anchored in segmented retinal anatomy, a physics-inspired sound model encodes interaction-driven cues directly from intraoperative imaging rather than static proximity mappings. The user study demonstrated significantly improved identification of clinically relevant events, with the largest gains in bleb detection. These gains likely stem from the model’s physical formulation, which integrates deformation signals observable in iOCT, whereas the baseline relies on rigid parameterizations that do not accommodate such interaction cues. The framework supports real-time workflows and is extensible to additional interaction structures. Reported sound roughness by a few participants motivates future work in auditory and perceptual refinement [10] to further enhance usability.

Acknowledgments. This work was primarily supported by the German Research Foundation (DFG) under Project No. 560596849. Further support was provided by the German Research Foundation (DFG) under Germany’s Excellence Strategy – EXC 2050/2 – Project No. 390696704 – Cluster of Excellence “Centre for Tactile Internet with Human-in-the-Loop” (CeTI) at Technische Universität Dresden. The authors would like to thank SynthesEyes for supporting this work through data provision and technical collaboration. Shervin Dehghani and Michael Sommersperger were partially supported by the Bavarian Research Foundation (BFS) under Grant No. AZ-1569-22.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Arikan, D., Zhang, P., Sommersperger, M., Dehghani, S., Esfandiari, M., Taylor, R.H., Nasser, M.A., Gehlbach, P.L., Navab, N., Iordachita, I.I.: Real-time deformation-aware control for autonomous robotic subretinal injection under iOCT guidance. 2025 IEEE International Conference on Robotics and Automation (ICRA) pp. 10531–10537 (2024), <https://api.semanticscholar.org/CorpusID:273963767>
2. Arikan, D., Zhang, P., Sommersperger, M., Dehghani, S., Esfandiari, M., Taylor, R.H., Nasser, M.A., Gehlbach, P.L., Navab, N., Iordachita, I.I.: Towards motion compensation in autonomous robotic subretinal injections. 2025 International Symposium on Medical Robotics (ISMR) pp. 66–72 (2024), <https://api.semanticscholar.org/CorpusID:274306241>
3. Black, D., Hansen, C., Nabavi, A., Kikinis, R., Hahn, H.K.: A survey of auditory display in image-guided interventions. *International Journal of Computer Assisted Radiology and Surgery* **12**, 1665–1676 (2017), <https://api.semanticscholar.org/CorpusID:11833443>
4. de Boor, C.: A practical guide to splines. In: *Applied Mathematical Sciences* (1978), <https://api.semanticscholar.org/CorpusID:122101452>
5. Dehghani, S., Sommersperger, M., Zhang, P., Martin-Gomez, A., Busam, B., Gehlbach, P.L., Navab, N., Nasser, M.A., Iordachita, I.I.: Robotic navigation autonomy for subretinal injection via intelligent real-time virtual iOCT volume slicing. 2023 IEEE International Conference on Robotics and Automation (ICRA) pp. 4724–4731 (2023), <https://api.semanticscholar.org/CorpusID:255998244>

6. Ehlers, J.P., Modi, Y.S., Pecun, P.E., Goshe, J.M., Dupps, W.J., Rachitskaya, A.V., Sharma, S., Yuan, A., Singh, R.P., Kaiser, P.K., Reese, J.L., Calabrese, C., Watts, A., Srivastava, S.K.: The discover study 3-year results: Feasibility and usefulness of microscope-integrated intraoperative oct during ophthalmic surgery. *Ophthalmology* **125** 7, 1014–1027 (2018), <https://api.semanticscholar.org/CorpusID:3781305>
7. Fastl, H., Zwicker, E.: Psychoacoustics: Facts and models (1990), <https://api.semanticscholar.org/CorpusID:144689350>
8. George, S.M., Lu, F., Rao, M., Leach, L.L., Gross, J.M.: The retinal pigment epithelium: Development, injury responses, and regenerative potential in mammalian and non-mammalian systems. *Progress in Retinal and Eye Research* **85** (2021), <https://api.semanticscholar.org/CorpusID:233354736>
9. Gupta, P.K., Jensen, P.S., de Juan, E.: Surgical forces and tactile perception during retinal microsurgery. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention* (1999), <https://api.semanticscholar.org/CorpusID:12455806>
10. Hermann, T., Weger, M.: Data-driven auditory contrast enhancement for everyday sounds and sonifications. *Proceedings of the 25th International Conference on Auditory Display (ICAD 2019)* (2019), <https://api.semanticscholar.org/CorpusID:197859996>
11. Huber, P.J.: Robust methods of estimation of regression coefficients. *Statistics (Berlin, DDR)* **8**(1), 41–53 (1977). <https://doi.org/10.1080/02331887708801356>
12. Hughes-Hallett, A., Mayer, E.K., Marcus, H.J., Pratt, P., Mason, S., Darzi, A., Vale, J.: Inattention blindness in surgery. *Surgical Endoscopy* **29**, 3184–3189 (2015), <https://api.semanticscholar.org/CorpusID:21429272>
13. Irigoyen, C., others: Subretinal injection techniques for retinal disease: A review. *Journal of Clinical Medicine* **11**(16), 4717 (2022). <https://doi.org/10.3390/jcm11164717>, <https://www.mdpi.com/2077-0383/11/16/4717>
14. Jo, Y.J., Heo, D.W., Shin, Y.I., Kim, J.Y.: Diurnal variation of retina thickness measured with time domain and spectral domain optical coherence tomography in healthy subjects. *Investigative Ophthalmology & Visual Science* **52**(9), 6497–6500 (2011). <https://doi.org/10.1167/iovs.11-7403>
15. Ling, S., Yang, S., Hu, X., Yin, D., Dai, Y., Qian, X., Wang, D., Pan, X., Hong, J., Sun, X., Yang, H., Paludan, S., Cai, Y.: Lentiviral delivery of co-packaged cas9 mrna and a vegfa-targeting guide rna prevents wet age-related macular degeneration in mice. *Nature Biomedical Engineering* **5**, 1–13 (02 2021). <https://doi.org/10.1038/s41551-020-00656-y>
16. Matinfar, S., Dehghani, S., Salehi, M., Sommersperger, M., Navab, N., Faridpooya, K., Fairhurst, M., Navab, N.: From tissue to sound: A new paradigm for medical sonic interaction design. *Medical Image Analysis* **103**, 103571 (2025)
17. Matinfar, S., Dehghani, S., Sommersperger, M., Faridpooya, K., Fairhurst, M., Navab, N.: Ocular stethoscope: Auditory support for retinal membrane peeling. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2024), <https://api.semanticscholar.org/CorpusID:273232345>
18. Matinfar, S., Nasser, M.A., Eck, U., Kowalsky, M., Roodaki, H., Lohmann, C., Maier, M., Navab, N.: Surgical soundtracks: automatic acoustic augmentation of surgical procedures. *International Journal of Computer Assisted Radiology and Surgery* **13** (07 2018). <https://doi.org/10.1007/s11548-018-1827-2>
19. Matinfar, S., Salehi, M., Suter, D., Seibold, M., Navab, N., Dehghani, S., Wanivenhaus, F., Furnstahl, P., Farshad, M., Navab, N.: Sonification as a reliable alternative to conventional visual surgical navigation. *Scientific Reports* **13** (2022), <https://api.semanticscholar.org/CorpusID:250144608>
20. Pannek, S., Dehghani, S., Sommersperger, M., Zhang, P., Gehlbach, P.L., Nasser, M.A., Iordachita, I.I., Navab, N.: Exploring the needle tip interaction force

- with retinal tissue deformation in vitreoretinal surgery. 2024 IEEE International Conference on Robotics and Automation (ICRA) pp. 16999–17005 (2024), <https://api.semanticscholar.org/CorpusID:271798975>
21. Roodaki, H., Navab, N., Eslami, A., Stapleton, C., Navab, N.: Sonifeye: Sonification of visual information using physical modeling sound synthesis. *IEEE Transactions on Visualization and Computer Graphics* **23**(11), 2366–2371 (2017). <https://doi.org/10.1109/TVCG.2017.2734327>
 22. Ruozzi, V., Matinfar, S., Schütz, L., Wiestler, B., Redaelli, A., Votta, E., Navab, N.: Biosonix: Can physics-based sonification perceptualize tissue deformations from tool interactions? In: Oguz, I., Zhang, S., Metaxas, D.N. (eds.) *Information Processing in Medical Imaging (IPMI 2025)*. Lecture Notes in Computer Science, vol. 15830, pp. 19–33. Springer Nature Switzerland, Cham (2025). https://doi.org/10.1007/978-3-031-96625-5_2
 23. Schütz, L., Dehghani, S., Sommersperger, M., Faridpooya, K., Navab, N.: The impact of intraoperative optical coherence tomography on cognitive load in virtual reality vitreoretinal surgery training. *Scientific Reports* **15**(1) (Dec 2025). <https://doi.org/10.1038/s41598-025-07670-7>, publisher Copyright: © The Author(s) 2025.
 24. Serafin, S., Franinović, K., Hermann, T., Lemaitre, G., Rinott, M., Rocchesso, D.: *The sonification handbook* (2011), <https://api.semanticscholar.org/CorpusID:31502560>
 25. Unser, M.A.: Splines: a perfect fit for signal and image processing. *IEEE Signal Process. Mag.* **16**, 22–38 (1999), <https://api.semanticscholar.org/CorpusID:62688047>
 26. Villeneuve, J., Leonard, J.: *Mass-interaction physical models for sound and multi-sensory creation : Starting anew* (2019), <https://api.semanticscholar.org/CorpusID:204780346>
 27. Zhao, C., Boles, N.C., Miller, J.D., Kawola, S., Temple, S., Davis, R.J., Stern, J.H.: Development of a refined protocol for trans-scleral subretinal transplantation of human retinal pigment epithelial cells into rat eyes. *Journal of visualized experiments : JoVE* **126** (2017), <https://api.semanticscholar.org/CorpusID:13764597>
 28. Zhou, M., Guo, X., Grimm, M., Lochner, E., Jiang, Z., Eslami, A., Ye, J., Navab, N., Knoll, A., Nasser, M.A.: Needle detection and localisation for robot-assisted subretinal injection using deep learning. *CAAI Trans. Intell. Technol.* **10**, 703–715 (2025), <https://api.semanticscholar.org/CorpusID:280322493>