

All-in-One Augmented Reality Guided Head and Neck Tumor Resection

Yue Yang¹, Matthieu Chabanas¹, Carrie Reale², Annie Benson³, Jason Slagle²,
Matthew Weinger², Michael Topf², and Jie Ying Wu^{1*}

¹ Vanderbilt University Department of Computer Science, Nashville, TN 37235

² Vanderbilt University Medical Center, Nashville, TN 37232

³ Vanderbilt University Department of Biomedical Engineering, Nashville, TN 37235

Abstract. Positive margins are common in head and neck squamous cell carcinoma, yet intraoperative re-resection is often imprecise because margin locations are typically communicated verbally from pathology. We present an all-in-one augmented reality (AR) system that relocalizes positive margins from a resected specimen to the resection bed and visualizes them in situ using HoloLens 2 depth sensing and fully automated markerless surface registration. In a silicone phantom study with six medical trainees, markerless registration achieved target registration errors comparable to a marker-based baseline (median 1.8 mm vs. 1.7 mm; maximum < 4 mm). In a margin relocalization task, AR guidance reduced error from verbal guidance (median 14.2 mm) to a few millimeters (median 3.2 mm), with all AR localizations within 5 mm error. These results support the feasibility of markerless AR margin guidance for more precise intraoperative re-resection.

Keywords: Augmented Reality · Markerless Registration · Head and Neck Cancer · Surgical Margins · Intraoperative Guidance

1 Introduction

Head and neck squamous cell carcinoma (HNSCC) affects hundreds of thousands of patients each year [1] and has one of the highest rates of positive surgical margins among solid tumors [8], contributing to local recurrence and poor outcomes. Intraoperative frozen section analysis is used in over 90% of cases [4], and a positive margin requires the surgeon to re-resect accurately by relocating its position. This is challenging due to complex anatomy and specimen disorientation. In practice, pathologists convey margin locations verbally, which provides only coarse spatial guidance [6]; consequently, re-resections are often imprecise, and only about 20–30% of re-resection specimens contain residual tumor [9].

Augmented reality (AR) head-mounted displays can overlay margin information directly in the surgical field [11]. Prior head-and-neck margin pilots demonstrated feasibility but relied on manual alignment [7, 11] or marker-based registration [2], which can add setup steps, sterile-field clutter, and workflow burden.

* Corresponding author: jieying.wu@vanderbilt.edu

Recent work on deformable AR guidance for head-and-neck resection focuses on tissue deformation but still requires additional registration cues such as fiducials [13]. In contrast, our focus is specimen-to-resection-bed relocalization of small head-and-neck surfaces using only the AR point-of-view time-of-flight depth sensor for sensing.

Here, we present an all-in-one markerless AR margin guidance system that registers the resected specimen back to the resection bed using geometric data and visualizes positive-margin locations in situ. The technical contribution is an automated AR-only workflow that combines curvature-biased sampling, global alignment, automatic coarse-pose ROI cropping, and point-to-plane iterative closest point (ICP) for small specimen-to-bed registration. We evaluate the system on a silicone face phantom with quantitative registration measurements and a user study with five residents and one medical student, comparing AR-guided versus verbal margin relocation.

2 Methods

As shown in Figure 1(a), in a phantom scenario, our system relocalizes a positive margin by scanning the specimen, capturing the resection bed with the HoloLens 2 Articulated HAndTracking (AHAT) depth sensor, performing fully automated markerless registration, and rendering the annotated margin in AR. The AHAT sensor also continuously tracks a stylus for quantitative evaluation.

We define right-handed coordinate frames: Head-Mounted Display (HMD) world W , AHAT depth sensor D , resection bed R , specimen scan S , and tool T . A rigid transform from frame X to Y is ${}^Y T_X \in SE(3)$.

The AR overlay requires placing the specimen scan in the world frame:

$${}^W T_S = {}^W T_D {}^D T_R {}^R T_S \quad \Leftrightarrow \quad {}^D T_S = \underbrace{{}^D T_R}_{T_{DR}} \underbrace{{}^R T_S}_{T_{RS}}, \quad (1)$$

${}^W T_D$ is provided by HoloLens tracking; ${}^D T_R$ is known from our depth-fusion reference; and the only unknown is ${}^R T_S$, estimated by markerless surface registration. Any annotated margin point $\mathbf{x}_S$ is visualized by $\mathbf{x}_W = {}^W T_S \tilde{\mathbf{x}}_S$.

2.1 Resection-bed and specimen surface acquisition

Resection bed. The surgeon or resident looks at the exposed cavity for ~ 3 s while AHAT depth frames are streamed. Each frame is back-projected to 3D and fused into a fixed resection-bed reference frame R using the HMD pose, followed by voxel downsampling and outlier removal to obtain a denoised bed cloud $\mathcal{P}_R = \{\mathbf{p}_i\}_{i=1}^{N_R}$ with normals $\mathbf{n}_{\mathbf{p}_i}$. Because AHAT depth accuracy can vary with surface properties [3], we performed a one-time local depth correction during setup [14]. We used a tracked stylus to sample the local phantom surface. We then paired the nearest AHAT depth points with the stylus contact points, and a least-squares singular value decomposition rigid correction was applied to the cloud.

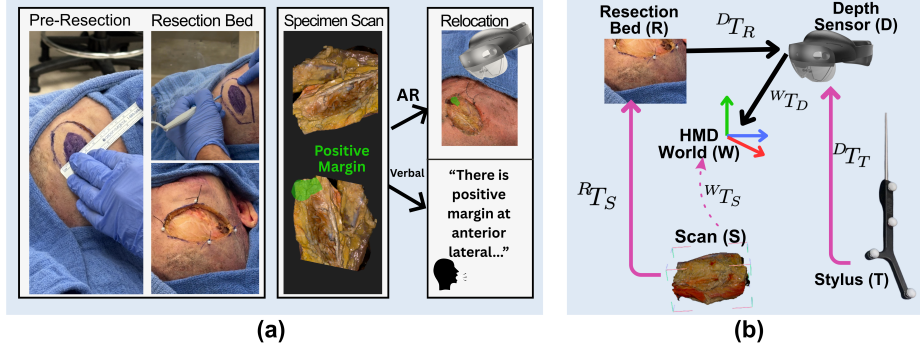


Fig. 1: System pipeline and coordinates: (a) AR-guided vs. traditional margin relocation workflow. (b) System formulation.

Specimen scan. The excised specimen is 3D scanned to obtain a point cloud (or mesh sampled to points) $\mathcal{P}_S = \{\mathbf{q}_j\}_{j=1}^{N_S}$ in frame S , along with a margin annotation defined on the scan.

2.2 Fully automated markerless specimen-to-bed registration

We eliminate the manual ROI selection required by prior work [14] by (i) improving global registration robustness for small targets (2–4 cm specimens) and (ii) automatically constructing a tight refinement ROI from the coarse pose.

Coarse global registration We compute local curvature scores $k(\cdot)$ on $\mathcal{P}_R$ and $\mathcal{P}_S$, and bias keypoint sampling toward landmark-rich regions:

$$\mathbb{P}(\mathbf{p}_i) \propto \frac{k(\mathbf{p}_i)}{\sum_{m=1}^{N_R} k(\mathbf{p}_m)}. \quad (2)$$

We compute Fast Point Feature Histograms descriptors $\phi(\cdot)$ at sampled points and generate putative correspondences by nearest-neighbor matching in descriptor space [10]:

$$\mathcal{C} = \{(i, j) \mid \|\phi(\mathbf{p}_i) - \phi(\mathbf{q}_j)\| < \tau\}. \quad (3)$$

To tolerate the high outlier rates typical for small anatomical surfaces, we estimate R_{TS} using a TEASER++-style truncated least-squares objective with correspondence weights w_{ij} [12]:

$$\min_{\mathbf{R} \in SO(3), \mathbf{t} \in \mathbb{R}^3} \sum_{(i,j) \in \mathcal{C}} w_{ij} \min\left(\|\mathbf{p}_i - (\mathbf{R}\mathbf{q}_j + \mathbf{t})\|^2, \epsilon^2\right). \quad (4)$$

As in TEASER++, we use translation-invariant pairwise differences to decouple translation from rotation:

$$\Delta \mathbf{p}_{ik} = \mathbf{p}_i - \mathbf{p}_k, \quad \Delta \mathbf{q}_{j\ell} = \mathbf{q}_j - \mathbf{q}_\ell, \quad \Delta \mathbf{p}_{ik} \approx \mathbf{R} \Delta \mathbf{q}_{j\ell} \text{ (inliers)}. \quad (5)$$

This yields a coarse estimate R_{TS}^{coarse} without any user input.

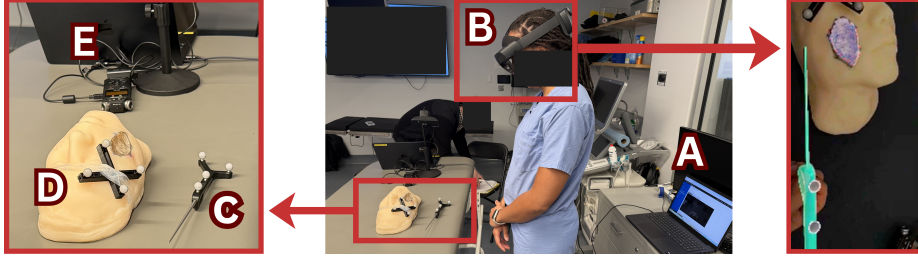


Fig. 2: User study setup: (A) Mixed-reality capture, (B) HMD, (C) tracked stylus, (D) phantom (marker + markerless), (E) close-up mic/camera.

Automatic ROI construction Instead of a user-defined ROI, we improved prior work by [14]. Let $\hat{\mathbf{q}}_j = \mathbf{R}^{\text{coarse}} \mathbf{q}_j + \mathbf{t}^{\text{coarse}}$ be specimen points in R . We form Axis-Aligned Bounding Box (AABB)

$$\mathcal{B} = \text{Dilate}(\text{AABB}(\{\hat{\mathbf{q}}_j\}), m), \quad (6)$$

and automatically crop the bed points for ICP:

$$\mathcal{P}_R^{\text{ICP}} = \{\mathbf{p} \in \mathcal{P}_R \mid \mathbf{p} \in \mathcal{B}\}. \quad (7)$$

This focuses computation on the small target structure and prevents false matches elsewhere in the scene, without any manual selection.

Fine registration Starting from ${}^R T_S^{\text{coarse}}$, we refine with point-to-plane ICP on $\mathcal{P}_R^{\text{ICP}}$ [15]:

$$\min_{\mathbf{R} \in SO(3), \mathbf{t} \in \mathbb{R}^3} \sum_{(\mathbf{p}, \mathbf{q}) \in \mathcal{C}_k} w_{\mathbf{p}\mathbf{q}} \left[(\mathbf{R}\mathbf{q} + \mathbf{t} - \mathbf{p})^\top \mathbf{n}_{\mathbf{p}} \right]^2, \quad (8)$$

with robust weights and outlier rejection. The output is ${}^R T_S^{\text{final}} \equiv T_{RS}$, used in (1) for AR overlay of the specimen and its annotated positive margin.

We further track a stylus/tool instrumented with retro-reflective markers using the STTAR algorithm [5]. STTAR estimates the tool pose in the AHAT sensor frame as $T_{TS} \equiv {}^D T_T$ at runtime. Continuous tool tracking in the world frame is obtained by composing with HMD tracking:

$$T_{TW} \equiv {}^W T_T = {}^W T_D {}^D T_T. \quad (9)$$

STTAR reports sub-millimeter accuracy, enabling reliable quantitative evaluation of margin relocalization.

2.3 Evaluation protocol and metrics

As shown in Figure 2, the system was evaluated on a silicone head-and-neck phantom with an artificial cheek tumor resected to create a cavity and spec-

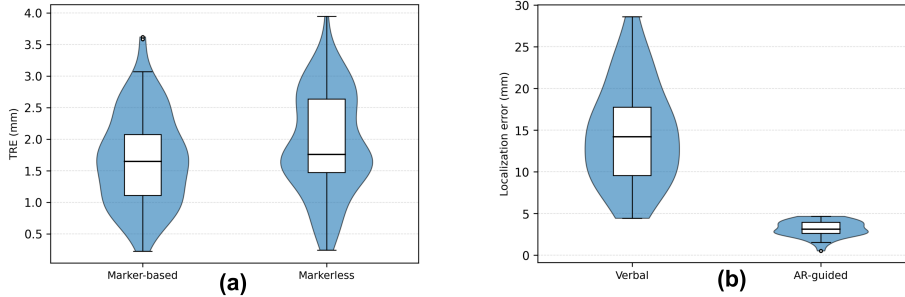


Fig. 3: (a) TRE for marker vs. markerless registration. (b) Margin relocation error for verbal vs. AR guidance.

imen. A structured-light scan provided the specimen model for AR visualization. Participants wore a Microsoft HoloLens 2 and completed two tasks under counterbalanced marker-based and markerless registration conditions. The marker-based baseline used a rigid retro-reflective marker tracked by AR [5]; the markerless condition estimated T_{RS} using the automated pipeline in Sec. 2.2.

Target locations were predefined ink landmarks on the phantom and annotated in the structured-light scan. Similar distractor marks were included to blind participants to the ground-truth locations during verbal localization. Expert stylus touches were used only to digitize reference landmark coordinates and reduce tracking noise.

In the point-matching task, participants touched 19 AR targets on the resection surface with a tracked stylus. Target registration error (TRE) was the Euclidean distance to the corresponding reference landmark. In the margin-relocation task, participants localized six simulated positive margins using either surgeon-scripted verbal guidance simulating the current clinical standard or AR guidance showing a 5 mm green target. We report median, interquartile range, RMSE, maximum/Hausdorff error, and 5 mm coverage (Cov_5). Registration conditions were compared with exact paired Wilcoxon signed-rank tests and rank-biserial effect sizes. Margin-relocation performance was compared using paired t -tests on participant-level medians with Cohen’s d_z . NASA-TLX, SUS, and AR-specific ratings were exploratory subjective outcomes.

3 Results

Registration accuracy Participants included five ENT trainees (PGY-1 to PGY-5+) and one medical student ($M_{age} = 30.3$ years, range: 27–33; 4/6 female). All five trainees reported low prior VR/AR experience, whereas the medical student had previously used AR in a work scenario. In the point-matching task, each participant localized 19 targets, yielding 114 point targets per registration method. Figure 3(a) shows the TRE distributions for marker-based and markerless registration, with detailed participant-level metrics reported in

Table 1: **Participant-level and group-level performance.** Distances are in mm and Cov_5 is in %.

(a) Registration (point matching): marker-based vs. markerless.

User	Marker-based reg.					Markerless reg.				
	Med	[Q1,Q3]	RMSE	Haus	Cov_5	Med	[Q1,Q3]	RMSE	Haus	Cov_5
P1	2.4	[1.8,2.7]	2.4	3.6	100.0	2.5	[2.1,2.8]	2.5	3.5	100.0
P2	1.3	[1.1,1.6]	1.4	2.1	100.0	1.7	[1.6,2.1]	1.9	2.8	100.0
P3	1.0	[0.8,1.4]	1.1	1.7	100.0	1.4	[1.1,1.6]	1.4	2.1	100.0
P4	1.9	[1.4,2.0]	1.8	2.6	100.0	2.9	[2.5,3.1]	2.9	3.9	100.0
P5	1.1	[0.7,1.6]	1.2	2.1	100.0	1.6	[1.2,1.7]	1.6	2.7	100.0
P6	2.4	[2.0,2.8]	2.5	3.6	100.0	1.7	[0.9,2.3]	1.9	3.2	100.0
All	1.7	[1.1,2.1]	1.8	3.6	100.0	1.8	[1.5,2.6]	2.1	3.9	100.0

(b) Guidance (margin relocalization): verbal vs. AR.

User	Verbal guidance					AR guidance				
	Med	[Q1,Q3]	RMSE	Haus	Cov_5	Med	[Q1,Q3]	RMSE	Haus	Cov_5
P1	11.0	[9.4,11.7]	11.2	14.7	0.0	2.5	[2.1,2.8]	2.6	3.6	100.0
P2	13.3	[9.5,16.4]	13.9	19.0	0.0	3.3	[2.7,3.9]	3.3	4.0	100.0
P3	12.6	[6.5,18.0]	14.1	21.4	33.3	3.0	[2.7,3.6]	3.3	4.7	100.0
P4	17.3	[16.0,23.0]	20.3	28.6	0.0	4.4	[4.1,4.4]	4.2	4.6	100.0
P5	9.4	[8.8,13.3]	11.4	16.7	0.0	3.6	[2.9,3.9]	3.4	4.6	100.0
P6	21.9	[17.1,22.5]	20.5	24.9	0.0	2.7	[2.6,2.9]	2.8	3.9	100.0
All	14.2	[9.6,17.8]	15.7	28.6	5.6	3.2	[2.6,3.9]	3.3	4.7	100.0

Table 1. Across all targets, marker-based and markerless registration produced similar median TREs (1.7 mm and 1.8 mm, respectively), and both methods achieved 100% 5 mm coverage. A paired Wilcoxon signed-rank test showed no significant difference between registration methods, $p = .31$. These results indicate that the automated depth-only markerless registration achieved accuracy comparable to the fiducial marker-based baseline in this facial specimen setup.

Margin relocalization task performance Margin relocalization was evaluated using six simulated positive-margin targets per participant, yielding 36 localization attempts per condition. Figure 3(b) shows the localization error distributions for conventional verbal guidance and AR-guided localization, with detailed participant-level metrics reported in Table 1. AR guidance substantially improved margin relocalization accuracy compared with verbal guidance. Across all attempts, the median localization error decreased from 14.2 mm with verbal guidance to 3.2 mm with AR guidance. AR guidance improved participant-level median error for every participant, with a mean paired reduction of 11.0 mm. This improvement was statistically significant, $p = .002$. All AR-guided attempts were within 5 mm of the target ($\text{Cov}_5 = 100\%$), whereas only 5.6% of verbal-guided attempts were within 5 mm. Thus, AR guidance reduced a clinically

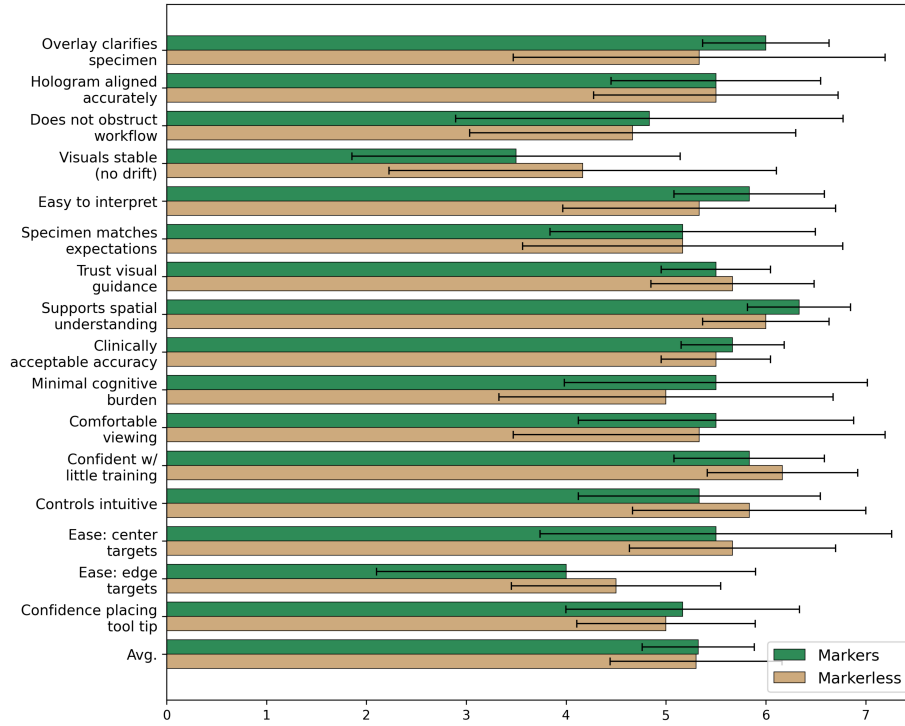


Fig. 4: AR task-specific ratings

ambiguous verbal relocation task to consistently sub-5-mm target localization in this phantom setting.

Subjective questionnaires We used the six standard NASA-TLX subscales (Mental, Physical, Temporal, Performance, Effort, Frustration). Avg.Score was the mean of the six subscales. Exact paired Wilcoxon signed-rank tests were used to compare Verbal vs. AR guidance within each registration condition ($n = 6$). In the **Markers** condition, AR reduced mental demand compared with verbal guidance ($M = 2.33$, $SD = 1.75$ vs. $M = 4.50$, $SD = 2.07$), $p = .031$. Other TLX subscales and Avg.Score were lower with AR but did not reach significance. In the **Markerless** condition, workload showed a similar trend for Avg.Score ($M = 2.31$, $SD = 1.00$ vs. $M = 3.50$, $SD = 1.11$), $p = .063$, but no subscale was significant. System Usability Scale (SUS) total was similar between registration modes ($M = 73.75$, $SD = 13.49$ for Markers; $M = 70.42$, $SD = 14.00$ for Markerless), $p = .438$, with no significant item-level differences (all $p \geq .159$). In the semi-structured interview, one participant commented on the HMD form factor in relation to existing surgical equipment: “people put our loupes on and off in the OR and I feel like that’s even harder to position correctly. And so I

think it'd be very easy for someone to put it [HMD] on your head and take it off when scrubbed-in."

Figure 4 summarizes the AR task-specific questionnaire, including the overall average across items (AR Avg). Paired Wilcoxon signed-rank tests were used for each item, with Holm adjustment within the questionnaire family. No item or overall average differed significantly between Markers and Markerless conditions (all $p \geq .157$). AR Avg was nearly identical across conditions ($M = 5.32$, $SD = 0.56$ for Markers; $M = 5.30$, $SD = 0.86$ for Markerless), suggesting that perceived AR task experience was similar with and without fiducial markers.

4 Discussion and Conclusion

We demonstrated an all-in-one AR guidance system for head-and-neck tumor resection that automatically registers a specimen scan to the resection bed and visualizes positive-margin locations in situ. In a controlled phantom study, markerless registration achieved approximately 2 mm TRE, comparable to a marker-based baseline, while AR guidance reduced margin relocalization from centimeter-level error under verbal guidance to a few millimeters. Our system combines lightweight point cloud surface acquisition, automated markerless specimen-to-bed registration, and AR visualization without external RGB-D cameras, fiducial markers, or manual ROI selection. However, our study did not model tissue deformation, bleeding, or specimen shrinkage, and iterative resections were not tested. AHAT depth accuracy and HMD form factor may also vary with surface type and view angle. Finally, the user study included six trainees. Thus, we show that markerless registration can obtain performance comparable to marker-based method, paving the way for future clinical adoption.

5 Acknowledgment

This work was supported by the National Institutes of Health (NIH) under grant number 1R01EB037685-01. We thank Dr. Michael I. Miga for his contribution.

6 Disclosure of Interests

The authors have no competing interests in the paper.

References

1. Barsouk, A., Aluru, J.S., Rawla, P., Saginala, K., Barsouk, A.: Epidemiology, risk factors, and prevention of head and neck squamous cell carcinoma. *Medical Sciences* **11**(2), 42 (2023)
2. Duan, H., Yang, Y., Niu, W.L., Anders, D., Dreisbach, A.M., Holley, D., Franc, B.L., Perkins, S.L., Leuze, C., Daniel, B.L., et al.: Localization of sentinel lymph nodes using augmented-reality system: a cadaveric feasibility study. *European Journal of Nuclear Medicine and Molecular Imaging* **52**(10), 3643–3652 (2025)

3. Li, H., Yan, W., Liu, D., Qian, L., Yang, Y., Liu, Y., Zhao, Z., Ding, H., Wang, G.: Evid surgical guidance with retro-reflective tool tracking and spatial reconstruction using head-mounted augmented reality device. *IEEE Transactions on Visualization and Computer Graphics* **31**(9), 6497–6511 (2024)
4. Long, S.M., Mclean, T., Valero Mayor, C., Fitzgerald, C.W., Feit, N.Z., Katabi, N., Xu, B., Cohen, M.A., Ganly, I., Ghossein, R.A., et al.: Use of intraoperative frozen section to assess final tumor margin status in patients undergoing surgery for oral cavity squamous cell carcinoma. *JAMA Otolaryngology–Head & Neck Surgery* **148**(10), 911–917 (2022)
5. Martin-Gomez, A., Li, H., Song, T., Yang, S., Wang, G., Ding, H., Navab, N., Zhao, Z., Armand, M.: Sttar: surgical tool tracking using off-the-shelf augmented reality head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics* **30**(7), 3578–3593 (2023)
6. Nakhleh, R.E.: Quality in surgical pathology communication and reporting. *Archives of pathology & laboratory medicine* **135**(11), 1394–1397 (2011)
7. Necker, F.N., Chang, M., Leuze, C., Topf, M.C., Daniel, B.L., Baik, F.M.: Virtual resection specimen interaction using augmented reality holograms to guide margin communication and flap sizing. *Otolaryngology–Head and Neck Surgery* **169**(4), 1083–1085 (2023)
8. Orosco, R.K., Tapia, V.J., Califano, J.A., Clary, B., Cohen, E.E., Kane, C., Lippman, S.M., Messer, K., Molinolo, A., Murphy, J.D., et al.: Positive surgical margins in the 10 most common solid cancers. *Scientific reports* **8**(1), 5686 (2018)
9. Prasad, K., Sharma, R., Habib, D., Sinard, R., Mannion, K., Rohde, S., Langerman, A., Nettekville, J., Rosenthal, E., Lewis, J., et al.: How often is cancer present in oral cavity re-resections after initial positive margins? *The Laryngoscope* **134**(2), 717–724 (2024)
10. Rusu, R.B., Blodow, N., Beetz, M.: Fast point feature histograms (fpfh) for 3d registration. In: 2009 IEEE international conference on robotics and automation. pp. 3212–3217. IEEE (2009)
11. Tong, G., Xu, J., Pfister, M., Atoum, J., Prasad, K., Miller, A., Topf, M., Wu, J.Y.: Development of an augmented reality guidance system for head and neck cancer resection. *Healthcare Technology Letters* **11**(2-3), 93–100 (2024)
12. Yang, H., Shi, J., Carlone, L.: Teaser: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics* **37**(2), 314–333 (2020)
13. Yang, Q., Li, F., Xu, J., Liu, Z., Sridhar, S., Jin, W., Du, J., Heiselman, J., Miga, M., Topf, M., et al.: Augmented reality-based guidance with deformable registration in head and neck tumor resection. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 45–54. Springer (2025)
14. Yang, Y., Necker, F., Leuze, C., Chen, M., Finegersh, A., Lee, J., Divi, V., Daniel, B., Hargreaves, B., Wu, J.Y., et al.: Markerless augmented reality registration for surgical guidance: A multi-anatomy clinical accuracy study. *arXiv preprint arXiv:2511.02086* (2025)
15. Zhang, J., Yao, Y., Deng, B.: Fast and robust iterative closest point. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **44**(7), 3450–3466 (2021)