



Adaptive infrared patterns for microscopic surface reconstructions

Srdjan Milosavljevic¹ · Zoltan Bardosi¹ · Yusuf Oezbek¹ · Wolfgang Freysinger¹

Received: 28 October 2023 / Accepted: 18 July 2024 / Published online: 9 October 2024
© The Author(s) 2024

Abstract

Purpose Multi-zoom microscopic surface reconstructions of operating sites, especially in ENT surgeries, would allow multimodal image fusion for determining the amount of resected tissue, for recognizing critical structures, and novel tools for intraoperative quality assurance. State-of-the-art three-dimensional model creation of the surgical scene is challenged by the surgical environment, illumination, and the homogeneous structures of skin, muscle, bones, etc., that lack invariant features for stereo reconstruction.

Methods An adaptive near-infrared pattern projector illuminates the surgical scene with optimized patterns to yield accurate dense multi-zoom stereoscopic surface reconstructions. The approach does not impact the clinical workflow. The new method is compared to state-of-the-art approaches and is validated by determining its reconstruction errors relative to a high-resolution 3D-reconstruction of CT data.

Results 200 surface reconstructions were generated for 5 zoom levels with 10 reconstructions for each object illumination method (standard operating room light, microscope light, random pattern and adaptive NIR pattern). For the adaptive pattern, the surface reconstruction errors ranged from 0.5 to 0.7 mm, as compared to 1–1.9 mm for the other approaches. The local reconstruction differences are visualized in heat maps.

Conclusion Adaptive near-infrared (NIR) pattern projection in microscopic surgery allows dense and accurate microscopic surface reconstructions for variable zoom levels of small and homogeneous surfaces. This could potentially aid in microscopic interventions at the lateral skull base and potentially open up new possibilities for combining quantitative intraoperative surface reconstructions with preoperative radiologic imagery.

Keywords Adaptive pattern · Stereo reconstruction microscope · Random pattern · Bayesian optimizer · ENT procedures

Introduction

In many surgical interventions, specifically microscopic ENT surgeries at the lateral skull base, a surgeon would benefit from additional intraoperative data [1] offered by dense and accurate “real-time” surface reconstruction of the operative site. Registration of these with other pre- or intraoperative medical imagery could be beneficial for intraoperative sur-

geon guidance [2]. In microscopic interventions at the lateral skull base, reconstruction of homogeneous structures like bone or tissue is challenging due varying zoom levels [3], operating room lighting conditions and extreme illumination [4] prone to specularities and shadowing. This significantly affects invariant feature detection, disparity estimation and surface stereo reconstruction [5].

Stereoscopic surface reconstruction is a well-established technology that builds on feature matching between stereo image pairs using detectors like SIFT [6] and SURF [7], detecting invariant object features with inhomogeneous structures and matching blocks of pixels locally (block matching, BM or semi-globally (SGBM [8, 9])). BM yields a significantly higher density, but not necessarily a high accuracy [10]; pairing improper features leads to inaccurate reconstructions unsuitable for further quantification of the amount of resected tissue or for further multimodal registration. The FLANN detector [11], optimized with number

✉ Srdjan Milosavljevic
srdjan.milosavljevic@student.i-med.ac.at

Zoltan Bardosi
zoltan.bardosi@i-med.ac.at

Yusuf Oezbek
yusuf.oezbek@i-med.ac.at

Wolfgang Freysinger
wolfgang.freysinger@i-med.ac.at

¹ Department of Otorhinolaryngology, Medical University of Innsbruck, Innsbruck, Austria

of trees [12], improves the speed of feature detection, but cannot consider all true positive features due to computational tree limits [13]. Brute force approaches that compare all points to each other are prone to false positives and are computationally expensive [14].

To overcome the challenge of (tissue) homogeneity highly discernible patterns are projected on the objects to be reconstructed. Different methodologies are available for stereo reconstructions and even augmented reality [15–17], paradigmatically in ENT procedures at the lateral skull base. However, these methods are static and are not available for variable zoom and small (anatomical) structures, as is the case in this type of surgeries [18].

Enhancing intraoperative setups with proper illumination requires special methods for feature collection and triangulation. Time of flight [19] requires reliable and static optics (without zoom); structure from motion [20] will need some actions of a surgeon to collect valid data, impacting the clinical workflow and extending the procedure time; trifocal tensor [21] approaches provide a robust but bulky setup, too complex for surgical environments.

Adaptive infrared patterns are promising to enhance stereo matching under realistic operating room conditions to enable

dense microscopic reconstructions (MiRe) of homogeneous surfaces [22, 23]. The adaptive pattern reconstruction model, obtained from a Leica surgical stereo microscope, is compared to a model reconstructed from CT images via multimodal registration and point to point comparison. The method is evaluated on a colored realistic 3D print of a human ear for various zoom settings. Different projected patterns and illumination conditions serve assessing reconstruction accuracy aiming at developing additional navigated (ENT) procedures to allow quantitative monitoring surgical procedures of tissue removal during petrous bone surgeries.

Materials and methods

The application is written for use on Intel i7, 16GB RAM, Nvidia GTX 1060 powered PC using C++ and incorporating VTK [24], ITK [25], OpenCV [26], PCL [27]. It can read DICOM [28] studies, generate 3D models from MiRe and CT, capture microscope images, perform stereo calibration, control zoom, multimodal registration and generate near-infrared (NIR) pattern via Bayesian optimization. Figure 1 represents the workflow.

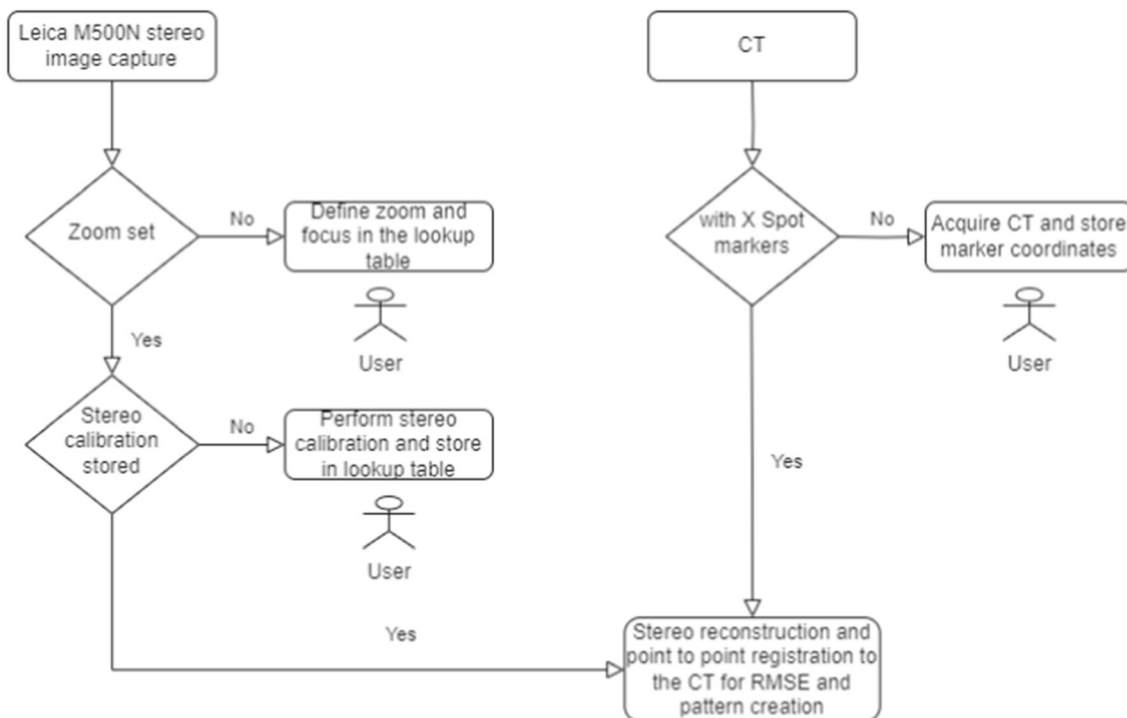


Fig. 1 Flow diagram of the system and user interaction for system initialization: multi-zoom calibration, marker detection and MiRe on the left, CT reconstruction and fiducial detection on the right, for co-

registration of CT images to MiRe by minimizing RMSE (root mean squared error) via an iterative Bayesian optimizer loop. Required user interactions are shown by the mannequins

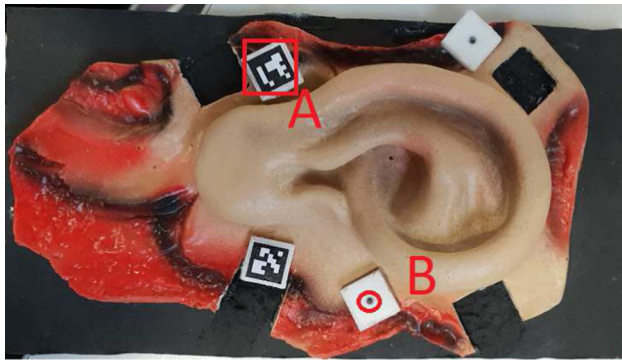


Fig. 2 Real scale ear model with ArUco (A) and X-spot markers (B) in the cube prior to being hidden by the ArUco markers (B)

Surface model and ground truth CT dataset

A real-scale realistic and colored plastic model of an adult external human ear (9×4 cm) was CT scanned as ground truth. This 3D print is suitable for evaluating surface reconstructions as shape, color, occlusion and illumination are demanding. Four 3D-printed holders different in height and inclination of the uppermost surface are placed on the ear and carry custom-made multimodal markers: an ArUco [29, 30] marker for microscope camera pose estimation and an CT X-Spot spherical markers (1.5 mm diameter, titanium CT markers, Beekley Medical) placed 1 mm below the ArUco marker origin (Fig. 2).

A ground truth 3D model was created from CT images with 0.6 mm slice thickness windowed at 2800 HU with the Marching cubes algorithm [31]; it is shown in Fig. 3 where the region of interest (ROI) is shown inside the white cuboid. The cuboid volume was used for the MiRe reconstruction error measurements.

Leica M500N stereo microscope

A Leica M500N surgical stereo operating microscope (Heerbrugg, Switzerland) was equipped with two high-resolution cameras (2456 x 2054 pixels, IDS UI-3080CPM-GL monochrome, IDS GmbH, Obersulm, Germany) connected to the microscope on two beam splitters introduced into the microscope optical path for scene view extraction. These cameras have a maximal NIR sensitivity at 740 nm and provide left and right image pairs of the surgical scene. The NIR adaptive pattern projector is mounted to the body of the microscope to illuminate the microscope's field of view. Microscope images are transferred via USB 3.0, and zoom and focus are controlled by a CAN-bus interface on the planning station (Fig. 4).

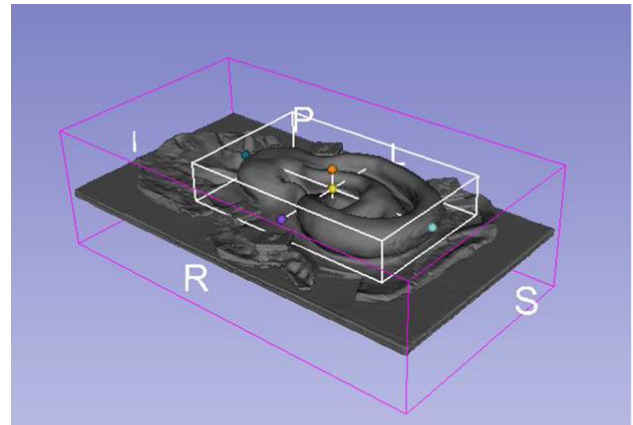


Fig. 3 Segmented ground truth ear CT model with the region of interest contained in the white cuboid. The letters indicate patient coordinates. Image created with 3D-Slicer [32]

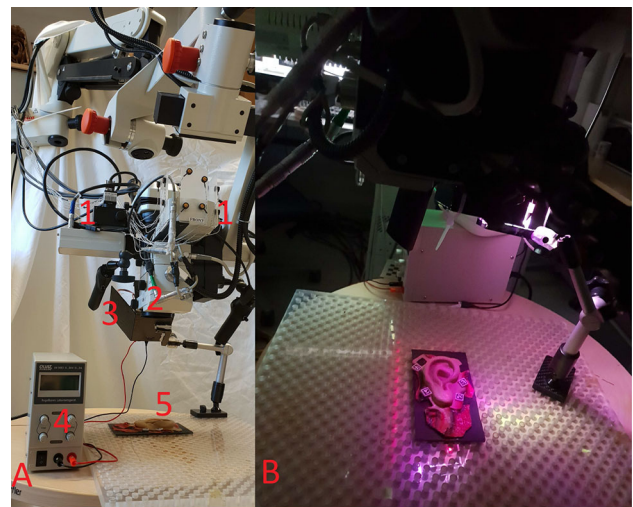


Fig. 4 Entire setup. A all system components 1: NIR cameras, 2: Microscope lens, 3: NIR projector, 4: NIR power source, 5: Ear model; B sample pattern projected on the ear phantom

Microscope optical calibration

The microscope is calibrated at 5 different zoom levels that still allow a focused view on the whole ear model including the markers at the microscope's working distance (below 20 cm) in both camera views. These boundary conditions restricted the use of zoom levels greater than 1.5. Stereo calibration is done with Zhang's method [33] by simultaneously taking image pairs (19 pairs per zoom) of a 9×5 checkerboard (3×3 mm squares) for a wide range of translations/rotations (up to 60°). Intrinsic parameters provide focal length and optical centers of each camera, while extrinsic parameters based on the intrinsic parameters provide relations (rotation and translation) between the two cameras. Both parameters for specific zoom and focus settings are determined and stored in a lookup table. Figure 5 represents

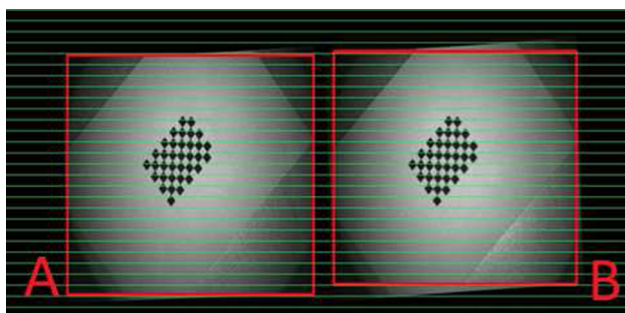


Fig. 5 Rectified left and right camera views after successful calibration

an example of rectified pairs; the image fractions used for disparity calculation are marked with red squares.

Disparity was calculated with SGBM [8]. All parameters (i.e., minimum minus disparity, block size=3) were kept constant for all comparison methods mentioned in “Method comparisons” section. Experimental variations of block size did not improve results, but resulted in inferior detection on reflecting homogeneous surfaces and to disappearance of small details in the reconstructions.

Power source and near-infrared (NIR) projector

NIR light sources for the pattern projector were evaluated with a photo-diode power sensor (Thor labs S121C, USA) at 740 nm and a bandwidth filter (≥ 700 nm) in an operating room (OR). No other significant light sources were found at this wavelength, with environmental NIR light being $< 10\mu\text{W}$ at 740 nm. No significant NIR absorption was measured for the microscope, too. Thus, NIR light at 740 nm was useful for the experiment with the microscope. A temperature controlled LED with custom-made active cooling, powered with a stabilized power supply (QUAT Power LN-3003, Pforring, Germany), overheating protection provided 3W of infrared illumination.

The custom-built NIR source illuminates the adaptive pattern projector, a customized APEMAN, DLP (Digital Light Processing), (resolution 800×600 pixels, Apeman International Co., Ltd., Shenzhen, China). All other optical components of the DLP were removed to avoid NIR absorption. NIR LED light intensity output of the projector was measured 800 mW, sufficient for adaptive pattern projection and detection (Fig. 6).

Rigid body registration and closest point RMSE

ArUco markers are detected in the MiRe reconstruction as points for the rigid body registration to the X Spot markers detected in the CT to set the ground for the further evaluation of the reconstruction quality. The registration is performed with an algorithm from Horn [34] for absolute orientation,

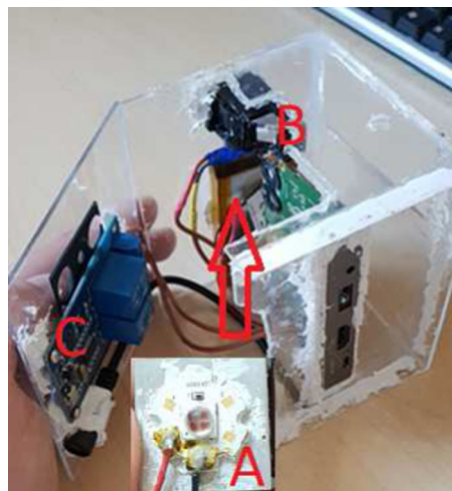


Fig. 6 NIR LED with active cooling (A), former DLP optics housing to fit the NIR LED (B) and relay system (C)

implemented to find the transformation from MiRe to CT and prepare two models for the closest point RMSE as an ultimate reconstruction error, explained in “Adaptive pattern generator with Bayesian optimizer” section .

Adaptive pattern generator with Bayesian optimizer

The adaptive pattern consists of an arrangement of dots in a rectangular area. Dot size (s) and distance (d) are to be optimized for each zoom level, representing significant effort if performed manually with respect to the procedure time, due to different possibilities of dot size and distance to generate the pattern and minimize the reconstruction error. To ease this, a Bayesian optimizer [35] was used. A simple Gaussian process [36, 37] surrogate model with an RBF+Scaling kernel was used to model the mean squared reconstruction error over the reference object. For each iteration, the goal was to find an optimal setting for d and s that minimizes the reconstruction error to our reference CT object providing to the optimizer parameters, $x:= [d, s]$ where dot size, s , from 2 to 16 pixels, and distance, d , from 1 to 10 pixels in the pattern. At each evaluation, a MiRe reconstruction was made with the predefined zoom level and NIR projection created with the testing parameters x .

The centers of the ArUco markers were located in the stereo reconstruction (MiRe). With the known relation of the ArUco to the radiolucent marker in the base of the artificial ear, the stereo reconstruction was registered to the 3D reconstructed CT surface, see “Rigid body registration and closest point RMSE” section, for further evaluation and RMSE extraction. Random 5000 points subset was chosen from the reconstructed point cloud ($t_s, s \in [1..5000]$), and for each t_s in the reconstructed point set, the closest point of the CT point set was chosen as a correspondence (marked

as r_s). The reconstruction error was calculated as the RMSE between t_s and r_s , respectively.

$$\sum_{s=1}^{5000} \left(\frac{|t_s - r_s|}{5000} \right)$$

The Bayesian optimization process was initialized with 10 random evaluations. The acquisition function was chosen to be the expected improvement formulation at each step until the end of the iteration.

The optimized dot distance d and size s values were obtained and stored in a lookup table for all 5 zoom levels investigated before applying changes to the 3D model, as this would impact the MiRe dataset and therefore comparison to the CT.

Method comparisons

The effects of the different object illumination on the MiRes were investigated in the following experiments. The artificial external ear was placed in the view of the optically calibrated microscope, and the illumination conditions were set using different methods:

- Standard OR environmental light, no microscope light, no extra additional light source.
- Standard microscope illumination.
- Projection of random NIR patterns [18, 38].
- Projection of Bayesian optimized adaptive NIR patterns. [18, 38].

Experiments were performed at 5 zoom levels in which the whole ear model was visible and in focus with 20 cm distance between an ear placed on a table and the microscope objective. The model was manually rotated relative to the microscope to detect possible angulation dependent reconstruction errors due to potential over-, under illumination, occlusions and specularities. Illumination power in the visible and IR domain was constant for all tests. RMSE mean and standard deviation were calculated between CT model and MiRe based on randomly selected, normally distributed 5000 closest points inside the ROI for 5 zoom levels, 4 methods and 10 reconstructions per setting, yielding 200 surface reconstructions as shown in Table 2.

A heat map was generated in the Z-direction to visualize differences in structure recognition and depth estimation of the 3D reconstructions obtained using proposed methods.

For extra validation purposes, the stereo reconstructed 3D surfaces of the ear by the adaptive pattern were compared visually to a 3D surface provided by Carl Zeiss Optotechnik GmbH, Neubeuern, Germany.

Table 1 Bayesian optimizer values per zoom level; the upper and lower lines in the columns show the **best** and the worst results for the adaptive patterns, respectively

| Zoom | Dot distance (pixel) | Dot size (pixel) | RMSE (mm) |
|------|----------------------|------------------|------------|
| 1.1× | 10 | 1 | 1.0 |
| | 6 | 2 | 0.7 |
| 1.2× | 10 | 1 | 1.0 |
| | 14 | 6 | 0.7 |
| 1.3× | 9 | 1 | 0.9 |
| | 16 | 7 | 0.5 |
| 1.4× | 16 | 1 | 1.0 |
| | 14 | 5 | 0.5 |
| 1.5× | 5 | 1 | 0.9 |
| | 15 | 7 | 0.5 |

Results

Results were collected for the described methods using environmental light showing acceptable overall results as presented in Table 2, having still significant outliers showing bad disparities as in Fig. 7, part B, caused by the light strength and the source position, but also due to the lack of the salient features to compare homogeneous surfaces [39].

Quite on the contrary. We observed a lack of information when using the microscope's halogen illumination, see Fig. 7, part A. This was not an issue for the other light sources used as their intensity could be down-regulated very quickly and switched to another approach, as proposed in the discussion.

Random NIR pattern was based on a diffractive optics showing promising results as presented in Table 2, but it was impossible to control the distribution or the size of the features, which brought the noise to the reconstruction as shown in Fig. 7, part C.

For the adaptive NIR pattern, Table 1 presents Bayesian optimizer's worst and best predictions for pattern dot size and separation per zoom level. Table 1 shows better results on higher zoom levels and in Fig. 7, part D. The time to generate MiRe was around 2 s, while CPU time to obtain adaptive pattern parameters, co-registration and calculating the reconstruction error was below 20 s.

First the Bayesian optimizer's worst and best predictions for pattern dot separation and dot size per zoom level are presented, see Table 1. The time to generate MiRe was around 2 s, while CPU time to obtain adaptive pattern parameters, co-registration and calculating the reconstruction error was below 20 s.

Table 2 represents sum of 10 mean and standard deviation RMSE per zoom level for proposed methods.

Figure 7 represents the heat map of the proposed methods, showing clear distinction in quality. A significant lack of

Table 2 Zoom levels with mean and standard deviation RMSE given in millimeters

| Zoom | Mean \pm SD (mm) | | |
|------|--------------------|---------------|---------------|
| | Adaptive | Random | Environmental |
| 1.1 | 0.7 \pm 0.4 | 1.9 \pm 1.8 | 1 \pm 0.9 |
| 1.2 | 0.7 \pm 0.4 | 1.7 \pm 1.7 | 1.2 \pm 1 |
| 1.3 | 0.5 \pm 0.3 | 1.8 \pm 1.8 | 1.1 \pm 1.0 |
| 1.4 | 0.5 \pm 0.3 | 1.5 \pm 1.6 | 1.8 \pm 1.5 |
| 1.5 | 0.5 \pm 0.3 | 1.6 \pm 1.7 | 1.7 \pm 1.4 |

structures can be seen in panels a and b. SGBM parameters were varied without significantly improving RMSE (Fig. 8).

Discussion

Fast, dense and accurate reconstructions of anatomical regions such as helix, antitragus, and antihelix are possible with Bayesian optimized NIR patterns. Adaptive NIR infrared patterns for microscopic surface reconstructions largely reduce homogeneities of the surgical site (skin, blood, bone, etc.) in the stereo image pairs captured by the stereo microscope [40]. Projection in near-infrared wavelengths allows to overcome some of the critical challenges in intraoperative stereo image reconstructions such as environmental lighting conditions and the intense surgical light sources used to illuminate the surgical site [41].

Further, the adaptive pattern addresses intraoperative usability by eliminating the need for moving the microscope, optimizing zoom due to lack of structures caused by the wrong pattern shape at the given zoom level, illumination caused by bad lightning at the given distance, repositioning

the patient or changing reconstruction parameters manually to obtain accurate reconstructions of the surgical field.

The results confirm that, not surprisingly, strong environmental/surgical lighting in the operating room does significantly impact MiRe. In some regions, depth information was lost, while in other regions wrongly calculated, leading to lack of reconstruction details, in Fig. 7. The environmental light approach provided an inaccurate disparity map composed of false depth information, leading to inaccurate 3D point definition, as some parts were without useful textures, therefore homogeneous.

The entire adaptive pattern process was automatized without affecting the clinical workflow or the outcome of an intervention. The application with a calibrated microscope may be run fully automatically intraoperatively. Further optimization or adaptation to other stereo techniques using machine learning and synthetic data for disparity calculation might improve results [42]. This could improve assessment of disparity in regions which lack information. GPU CUDA [43] parallel processing could allow “real-time” quantitative evaluations and surgeon support.

Manual rotation of an object represents the weak point of the system due to the necessity to involve the user in the data collection process to assess the accuracy and perform calibration of the setup if needed. In the future, this could be resolved with a setup providing controlled linear and rotational motion of the object in the surgical scene. There are already approaches to perform the stereo rig calibration, but on the opposite camera side [44].

The adaptive pattern approach has shown its potential. However, the current setup of a standard surgical microscope does not allow changing its construction details (i.e., the stereo camera baseline). Our approach to project NIR patterns onto the surgical site can easily be further optimized by directly injecting the NIR patterns in the illumination

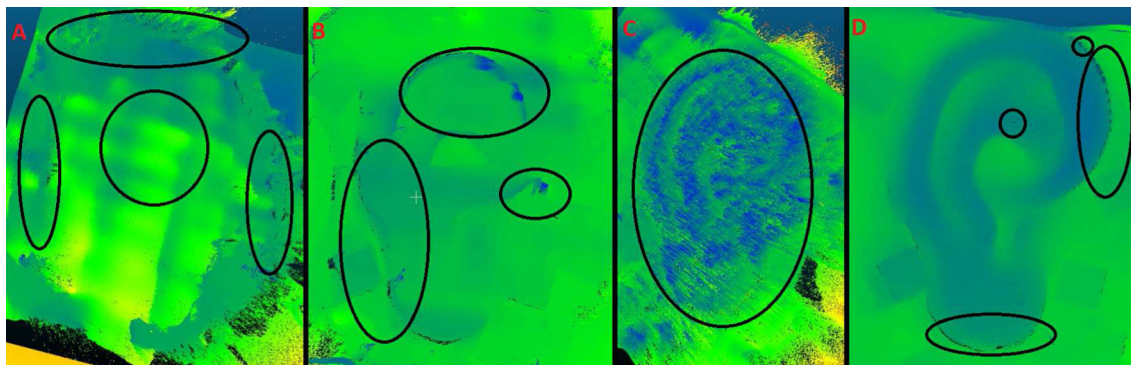


Fig. 7 Ear model reconstruction/structures for 4 different methods at zoom level 1.5. Ellipses show areas with lack of correct depth information: **A** microscope light; **B** environmental (OR) light; **C** random/diffractive pattern; and **D** adaptive pattern

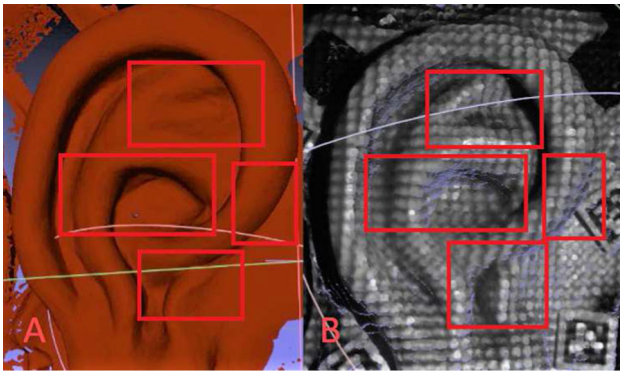


Fig. 8 Visual comparison of surfaces from **A** Carl Zeiss Optotechnik 3D surface and **B** adaptive pattern MiRe showing some advantages especially in the regions of concha and incisura

pathway of a surgical microscope with minor technological efforts. This would lead to an optimal illumination of the surgical site at all zoom levels, reducing shadowing as the optical path is used directly. If necessary, microscope optics would benefit from NIR cutoff filter implementation in the oculars. Such improvements, however, are beyond the scope of this work.

The surgical microscope in use was a standard medical device that was extended by standard extraction cameras. It was calibrated for 5 zoom levels, which could keep the anatomical region (external ear) completely in the field of view. This is in general fine for predefined zooms levels in a laboratory setting; for intraoperative use, however, when the whole range of zoom levels is accessed or the microscope optics are changed by other means: simple camera realignment, beamsplitter reinsertion, etc., the current calibration might not prove adequate. For real surgical use, extended camera calibration would be necessary to allow for extrapolating a current zoom/focus setting from an extended lookup table. For research purposes, however, this would be beyond the scope of the current project. One might envision other approaches such as feature detection and pairwise correspondence matching which would allow calibration “on the fly” [45], and such features might even be projected from an adaptive pattern [46] or on screen visualized patterns [47, 48].

Extreme illumination conditions are regularly encountered in microscopic surgery including under- and over-illumination of the scene. This is well compensated by the human eye and modern video imagery, but can severely affect stereo reconstruction results due to the lack of depth information in such regions [39]. The adaptive pattern was projected from a micromirror-based DLP projector with a light source that does not interfere with OR environmental light, being visible by the two detection cameras of the microscope and the surgeon oculars in reduced intensity, with the improvement suggestions in the discussion. This device projected adaptive dot patterns, generated with a Bayesian optimizer

based on RMSE between CT (ground truth) and MiRe closest points, to increase the amount of salient features and to decrease reconstruction inaccuracies by modification of the pattern parameters (dot size and dot distance), making it suitable for multiple zoom levels. The present results suggest that this approach can bring improvements in the reconstruction of demanding homogeneous anatomical regions while addressing different environmental scenarios without the necessity for surgeon input, therefore not impacting the default clinical workflow and providing a good base for quantification of a resected tissue.

The adaptive patterns presented here open up the possibility for further automatic evaluation of disparity maps to find regions with higher errors or missing depth information. This could be addressed, e.g., by locally refining the adaptive pattern to yield optimal stereo reconstructions of the surgical field. ArUco/X spot markers could be replaced by other multimodal ones or by software solutions as in [49].

Further clinical optimization close in cooperation with surgeons will provide a useful clinical tool at the end.

Conclusion

The proposed method allows creation of stereo microscopic reconstructions of areas of relevance for microscopic surgeries at the lateral skull base. It overcomes the challenges of variable zoom levels, homogeneous surfaces and environmental illumination conditions via a customized NIR setup and a real-time intraoperative Bayesian optimized pattern exploiting the reconstruction error between co-registered ground truth and microscopic stereo reconstructions. To this end, however, introduction of the NIR light parallel to the microscope illumination path is a prerequisite to combine the complete surgical view with the exploitable NIR illuminated views. This was, however, out of the scope of this project. In the future, this approach could allow new approaches to monitor the progress of a surgery quantitatively in preoperative imagery.

Acknowledgements We would like to thank to the Austrian National Bank, Anniversary Fund, for funding this research under Project Number: 16154.

Funding Open access funding provided by University of Innsbruck and Medical University of Innsbruck.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as

long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Sieber D, Erfurt P, John S (2019) The OpenEar library of 3D models of the human temporal bone based on computed tomography and micro-slicing. *Sci Data* 6:1038
- Agus M, Giachetti A, Gobetti E, Zanetti G, Zorcolo A, John N, Stone R (2002) Mastoidectomy simulation with combined visual and haptic feedback. *Stud Health Technol Inform* 85:17–23
- Psang DL, Sung K (2007) Comparing two new camera calibration methods with traditional pinhole calibrations. *Opt Express* 15:3012–3022
- Akkoyun F, Ozcelik A, Arpacı I, Erçetin A, Gucluer S (2023) A multi-flow production line for sorting of eggs using image processing. *Sensors* 23:117
- Yang Z (2021) Hierarchically adaptive image block matching under complicated illumination conditions. *J Comput Methods Sci Eng* 21:1455–1468
- Lowe D (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vision* 60:91–110
- Bay H, Tuytelaars T, Van Gool L (2006) SURF: speeded up robust features. *ECCV* 1:404–417
- Hirschmuller H (2008) Stereo processing by semiglobal matching and mutual information. *IEEE Trans Pattern Anal Mach Intell* 30:328–341
- Milosavljevic S, Freysinger W (2016) Quantitative measurements of surface reconstructions obtained with images a surgical stereo microscope. *CURAC* 47–52
- Verma R, Verma AK (2014) A comparative evaluation of leading dense stereo vision algorithms using OpenCV. *Int J Eng Res Technol* 2:196–201
- Muja M (2009) Fast approximate nearest neighbors with automatic algorithm configuration. *VISAPP* 5:18
- Bentley J (1975) Multidimensional binary search trees used for associative searching. *Commun ACM* 18:509–517
- Avrithis Y, Emiris IZ, Samaras G (2016) High-dimensional visual similarity search: k-d generalized randomized forests. *Comput Graph Int*, 25–28
- Ozturk A, Cayiroglu I (2022) A real-time application of singular spectrum analysis to object tracking with SIFT. *Eng Technol Appl Sci Res* 12:8872–8877
- Colchester AC, Zhao J, Holton-Tainter KS, Henri CJ, Maitland N, Roberts PT, Harris CG, Evans RI (1996) Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging. *Med Image Anal* 1:73–90
- Thomas DG, Doshi P, Colchester A, Hawkes DJ, Hill DL, Zhao J, Maitland N, Strong AJ, Evans RI (1996) Craniotomy guidance using a stereo-video-based tracking system. *Stereotact Funct Neurosurg* 66:81–83
- Edwards PJ, Hawkes DJ, Hill DL, Jewell D, Spink R, Strong A, Gleeson M (1995) Augmentation of reality using an operating microscope for otolaryngology and neurosurgical guidance. *J Image Guid Surg* 1:172–178
- Wang Y (2018) *Microscopy and microanalysis. disparity surface reconstruction based on a stereo light microscope and laser fringes*, vol. 24. Cambridge University Press, pp 503–515
- Pycinski B, Czajkowska J, Badura P, Juszczyk J, Pietka E (2015) Time-of-flight camera, optical tracker and computed tomography in pairwise data registration. *PLoS ONE* 11:1–20
- Lin J, Clancy N, Hu Y, Qi J, Tatla T, Stoyanov D, Maier-Hein L, Elson D (2017) Endoscopic depth measurement and super-spectral-resolution imaging. *LNIP* 10434:39–47
- Bishop C (2006) *Pattern recognition and machine learning*. Springer, Berlin
- Xu Y, So Y, Woo S (2022) Plane fitting in 3D reconstruction to preserve smooth homogeneous surfaces. *Sensors* 22:9391
- Fleigab OJ, Devernay F, Scarabinad JM, Jannin P (2001) Surface reconstruction of the surgical field from stereoscopic microscope views in neurosurgery. *Int Congr Ser* 1230:268–274
- Schroeder W, Martin KM, Lorensen WE (2006) *The visualization toolkit: an object-oriented approach to 3D graphics*. Kitware <https://vtk.org/vtk-textbook>. Accessed 17 April 2024
- Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, Aylward S et al (2002) Engineering and algorithm design for an image processing API: a technical report on ITK-the insight toolkit. *Stud. Health Technol. Inform.* 85:586–592
- Bradski G (2008) *The OpenCV library*. Dr. Dobb's J Softw Tools. <https://opencv.org/>. Accessed 17 April 2024
- Rusu RB, Cousins S (2011) 3D is here: point cloud library (PCL). In: *IEEE international conference on robotics and automation*, pp 1–4
- Parisot C (1995) The DICOM standard. *Int J Cardiac Imag* 11:171–177
- Romero-Ramirez F, Muñoz-Salinas R, Carnicer R (2018) Speeded up detection of squared fiducial markers. *Image Vis Comput* 76:38–47
- Garrido-Jurado S, Muñoz Salinas R, Madrid-Cuevas FJ, Medina-Carnicer R (2016) Generation of fiducial marker dictionaries using mixed integer linear programming. *Pattern Recogn* 51:481–491
- Lorensen WE, Cline EH (1987) Marching cubes: a high resolution 3d surface construction algorithm. *ACM Comput Graph* 21:163–169
- Kikinis R, Pieper SD, Vosburgh K (2014) 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. *Intraoper Imaging Image-Guided Ther* 19:277–289
- Zhang Z (2000) A flexible new technique for camera calibration. *IEEE Trans Pattern Anal Mach Intell* 22:1330–1334
- Horn B (1987) Closed-form solution of absolute orientation using unit quaternions. *Josa* 4:629–642
- Gonzalez J, Osborne M, Lawrence N (2016) GLASSES: relieving the myopia of Bayesian optimisation. *AISTATS*, 790–799
- Quadrianto N, Kersting K, Xu Z (2011) Gaussian process. In: *Encyclopedia of machine learning*. Springer, pp 428–439
- Rasmussen CE, Williams CKI (2005) *Gaussian processes for machine learning*. MIT Press, Cambridge
- Jeong J, Shin H, Chang J, Lim E, Choi S, Yoon K, Cho J (2013) High-quality stereo depth map generation using infrared pattern projection. *ETRI J* 35:1011–1020
- Giacomini BN, Roncella A, Thoeni R (2021) Influence of illumination changes on image-based 3d surface reconstruction. *Int Arch Photogramm Remote Sens Spatial Inf Sci* 43:701–708
- Reiter A, Sigaras A, Fowler D, Allen PK (2014) Surgical structured light for 3D minimally invasive surgical imaging. In: *International conference on intelligent robots and systems*, 1282–1287
- Lin Q, Cai K, Yang R, Chen H, Wang Z, Zhou J (2016) Development and validation of a near-infrared optical system for tracking surgical instruments. *J Med Syst* 40:1–14

42. Song X, Yang G, Zhu X, Zhou H, Ma Y, Wang Z, Jianping S (2022) AdaStereo: an efficient domain-adaptive stereo matching approach. *Int J Comput Vision* 130:226–245
43. Cook S (2012) *CUDA programming: a developer's guide to parallel computing with GPUs*. Newnes, Waltham
44. Cavagna A, Feng X, Melillo S, Parisi L, Postiglione L, Villegas P (2021) CoMo: a novel comoving 3D camera system. *IEEE Trans Instrum Meas* 70:1–16
45. Vasconcelos F, Barreto JP, Boyer E (2018) Automatic camera calibration using multiple sets of pairwise correspondences. *IEEE Trans Pattern Anal Mach Intell* 40:791–803
46. Mahdy YB, Hussain KF, Abdel-Maji MA (2013) Projector calibration using passive stereo and triangulation. *Int J Future Comput Commun* 5:385–390
47. Tan L, Wang Y, Yu H, Zhu J (2017) Automatic camera calibration using active displays of a virtual pattern. *Sensors* 17:685
48. Xiao S, Tao W, Zhao H (2016) A flexible fringe projection vision system with extended mathematical model for accurate three-dimensional measurement. *Sensors* 16:612
49. Yang J, Li H, Jia Y (2013) Go-ICP: solving 3D registration efficiently and globally optimally. *Int Conf Comput Vis* 11:2241–2254

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