

PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

Methods for a fusion of optical coherence tomography and stereo camera image data

Jan Bergmeier, Dennis Kundrat, Andreas Schoob, Lüder Alexander Kahrs, Tobias Ortmaier

Jan Bergmeier, Dennis Kundrat, Andreas Schoob, Lüder Alexander Kahrs, Tobias Ortmaier, "Methods for a fusion of optical coherence tomography and stereo camera image data," Proc. SPIE 9415, Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling, 94151C (18 March 2015); doi: 10.1117/12.2082511

SPIE.

Event: SPIE Medical Imaging, 2015, Orlando, Florida, United States

Methods for a Fusion of Optical Coherence Tomography and Stereo Camera Image Data

Jan Bergmeier, Dennis Kundrat, Andreas Schoob, Lüder Alexander Kahrs, Tobias Ortmaier
Leibniz Universität Hannover, Institute of Mechatronic Systems, Appelstr. 11a, Hannover,
Germany, 30167

ABSTRACT

This work investigates combination of Optical Coherence Tomography and two cameras, observing a microscopic scene. Stereo vision provides realistic images, but is limited in terms of penetration depth. Optical Coherence Tomography (OCT) enables access to subcutaneous structures, but 3D-OCT volume data do not give the surgeon a familiar view. The extension of the stereo camera setup with OCT imaging combines the benefits of both modalities. In order to provide the surgeon with a convenient integration of OCT into the vision interface, we present an automated image processing analysis of OCT and stereo camera data as well as combined imaging as augmented reality visualization. Therefore, we care about OCT image noise, perform segmentation as well as develop proper registration objects and methods. The registration between stereo camera and OCT results in a Root Mean Square error of 284 μm as average of five measurements. The presented methods are fundamental for fusion of both imaging modalities. Augmented reality is shown as application of the results. Further developments lead to fused visualization of subcutaneous structures, as information of OCT images, into stereo vision.

Keywords: optical coherence tomography, stereo camera, stereoscopic vision, surface reconstruction, multi-modal registration, augmented reality

1. INTRODUCTION

The use of stereo cameras in combination with surgical microscopes and endoscopes introduces depth perception into visualization of surgical scenes on monitors and enables further processing of the image data. A recent approach combines a custom laser setup with a stereo camera for surveillance.¹ Based on stereo camera image data, three-dimensional surface reconstructions of observed surgical fields are realized. However, the depth resolution of the stereo camera is insufficient for detailed reconstruction of submillimetric structures, like laser incisions.

Another three-dimensional imaging modality, which has already been used in endoscopes² as well as in surgical microscopes,³ is Optical Coherence Tomography (OCT). As non-invasive method with high depth resolution and penetration depth up to 3 mm in e.g. biological tissue, it can be prospective application for fusion with a stereo camera.

Figure 1 compares a 3D reconstructed, micro-scaled ablation pattern (typical several hundred μm) obtained from stereo camera and OCT data. The shape of the ablation profiles on the surface, reconstructed from stereo camera images (see Figure 1 (left)), is not very distinctive, due to low depth resolution of the stereo reconstruction and dependency on complete illumination of the incisions. Contrary, the segmented surface from OCT image data of the same ablation profile is shown right hand in Figure 1, exhibiting more distinctive shape of the ablation patterns while picturing a smaller area of the pattern. In our previous work,¹ we used OCT to evaluate laser ablations in wooden, soft tissue and bone samples, which shows a general applicability of OCT for evaluation of micro incisions.

The simultaneous use of two imaging devices with two screens is non-intuitive and manual analysis of large data sets - in particular the OCT volume data - is demanding for the user. This contribution proposes an integration of OCT information into stereo camera images. Consequently, the multimodal, three-dimensional,

Address all correspondence to: Jan Bergmeier, E-mail: jan.bergmeier@imes.uni-hannover.de,
Tel: +49 (0)511-762-17839, Fax: +49 (0)511-762-19976

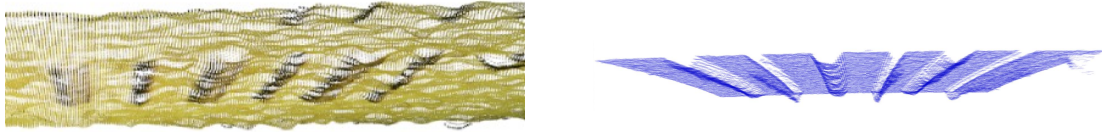


Figure 1. Laser ablation profiles on a wooden probe, reconstructed and segmented from stereo camera (left) and OCT volume (right). OCT workspace is a cutout of the stereo camera workspace with lateral dimension of 10×10 mm.

rigid registration problem between the reconstructed surface of the stereo camera and the OCT volume data is solved. The purpose of the registration is the determination of the transformation ${}^{\text{OCT}}T_{\text{cam}}$, which allows for mapping between stereo camera image data and OCT image data. As result, image data is superimposed.

2. METHODS

2.1 Experimental setup

The experimental setup consists of a spectral domain OCT (customized GANYMEDE Spectral Domain, Thorlabs Inc, Newton, New Jersey, USA) and a stereo camera rig (customized SOMIKON USB 2.0 stereo camera setup). Both are mounted to cover the same workspace (see Figure 2 (left)). The OCT scan range is set to $15 \times 15 \times 2.76$ mm³ with a resolution of $256 \times 256 \times 1024$ voxels, resulting in a depth resolution of $2.69 \mu\text{m}$. The stereo camera consists of two low cost camera modules with a diameter of 6 mm and a resolution of 640×480 pixels. The camera is positioned at a distance of 40 mm above the scene surface, resulting in a theoretical maximum depth resolution of $140 \mu\text{m}$.

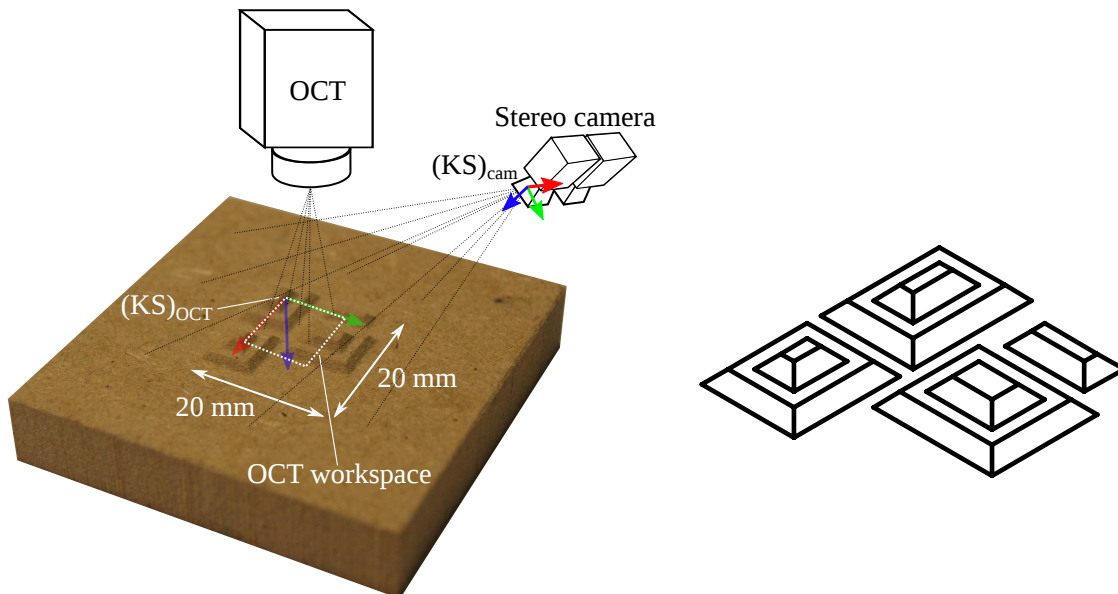


Figure 2. Experimental setup with OCT and stereo camera, observing the reference structure (left) and CAD contour of reference structure with truncated pyramids (right).

OCT acquires a longitudinal depth profile alongside a light beam, called A-scan, by measuring runtime of emitted light, reflected and scattered at boundary layers. This runtime corresponds to the distance of boundary layers with respect to the OCT, the intensity gives information about the refractive index of the layer. By lateral deflection of the OCT beam, a three-dimensional image can be concatenated. OCT gives spatial volume information without any texture. Fan distortion and additional effects result in recording errors during OCT imaging. For detailed information refer to Ortiz et al.⁴

2.2 Design of reference structure

For calculation of the registration between OCT and stereo camera ${}^{\text{OCT}}\mathbf{T}_{\text{cam}}$, a reference structure with identifiable points or structures is required. The main difficulty arises from different resolutions of the imaging devices. In particular, the low axial scan depth of the OCT of about 2.8 mm limits the possible pattern height. Slight inclines at step structures ensure a good visibility of the contour in the OCT image data, because edges in axial direction of the OCT beams often lead to weak visibility of the structure surface, resulting into gaps in the pictured surface. Furthermore, the sample must not be too shallow to be reconstructed with sufficient accuracy from the stereo camera data, due to the limited depth resolution of the stereo camera. Therefore, a structure of truncated pyramids, with a height of 2 mm and an extend of 20×20 mm, has been designed. The developed reference structure is shown in Figure 2, including the experimental setup with the OCT and the stereo camera, imaging an overlapping workspace on the reference structure.

2.3 Stereomatching-based surface reconstruction

Our stereomatching algorithm⁵ creates a surface reconstruction, based on stereo camera data. A local census-based disparity computation is applied to determine corresponding points in both camera images. Hence, a disparity map is calculated, which comprises the distance between corresponding points in left and right camera image. Consequently, for each pixel a three-dimensional point on the recorded object surface can be created. For display and further processing, all coordinates of all points are stored as point cloud (see Figure 3 (left)).

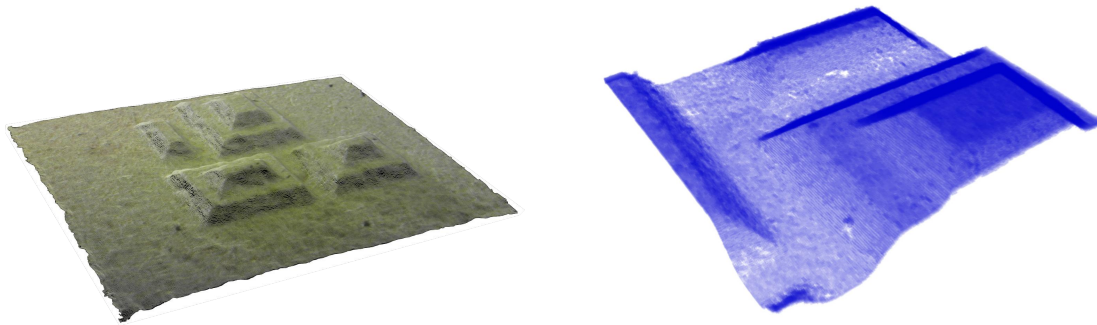


Figure 3. Extracted reference structure surfaces from stereo camera (left) and OCT image data (right). OCT workspace is a cutout of the stereo camera workspace with lateral dimension of 10×10 mm.

2.4 Surface segmentation in OCT image data

OCT data are represented as a 8-bit volume dataset. In order to identify relations between OCT and stereo camera image data, the surface has to be segmented. Image noise, which is typical for an OCT, has to be considered and removed. The segmentation is carried out as multi-step process, consisting of a binary thresholding method, a connected components labeling,⁶ followed by a relabeling. Figure 4 shows the segmentation process, illustrated on 2D images for better understanding, but performed on the whole 3D volume. In the first step, raw image data (see Figure 4 (left)) is separated in two groups by a gray scale threshold. The selection of the threshold is critical, because the gray scales of the object surface are close to the gray scales of the image noise. This is the reason why image noise cannot be eliminated during binarisation. A binarised image (see Figure 4 (center)) is the base for applying the connected components method. Connected components labeling searches for connected regions in image data⁷ and assigns each connected region its own size and label. This enables the possibility to sort all regions according to their size. Eliminating all regions below a certain size during relabeling takes care of the image noise as well as performs the segmentation of the object volume in OCT image data (see Figure 4 (right)). This is possible, because image noise remains as small connected regions. The final step is to extract the surface of the segmented volume. Therefore, we iterate through the relabeled image in A-scan direction of the OCT. Thus, the first points on the object surface from the OCT's point of view are extracted as segmented object surface. Applying this approach to our sample yields the surface segmentation of our aforementioned reference structure, shown right hand in Figure 3.

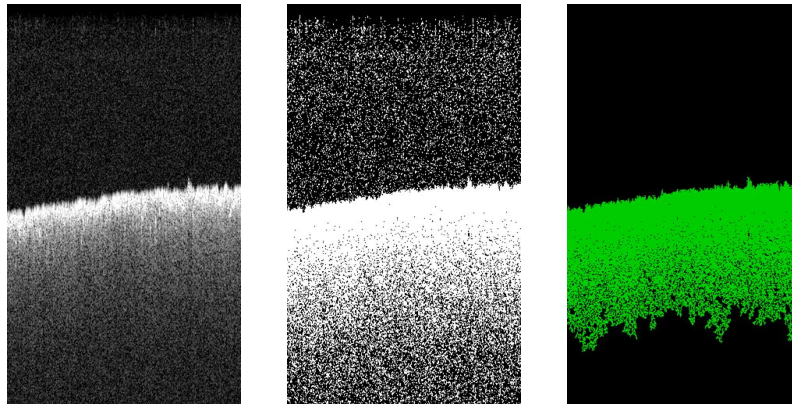


Figure 4. 2D views of surface segmentation steps in OCT image data from left to right: Raw OCT image, binary OCT image and relabeled region.

2.5 Registration of three-dimensional structured surfaces

After preprocessing, two surface point clouds are available to perform a registration, applying the Iterative Closest Point (ICP)⁸ algorithm. ICP provides a robust local convergence and good performance. As local method, it depends on a preceding coarse registration, which is performed feature-based and fully automated. An initial downsampling-step speeds up the calculation process. The image space is partitioned to a grid of voxels. All cloud points within a voxel are consolidated and approximated with the voxel's center of gravity. The employed features are Fast Point Feature Histograms (FPFH),⁹ a speed-improved version of the Point Feature Histograms (PFH).¹⁰ PFH are meaningful, position-independent, local features, describing the geometric surface characteristics around a point on a surface. Correspondences between the two point clouds can be identified, using the features calculated from surrounding points and their normals, by comparing the FPFH. Significance of the features is evaluated with a Persistence Analysis,¹⁰ which is used to determine unique and descriptive features. A mean histogram is calculated from all feature histograms. Each FPFH is compared to the mean histogram by calculating the Kullback-Leibler (KL) divergence as similarity measure. Feature histograms, generating a KL divergence greater than the standard deviation of an assumed gaussian distribution of the mean histogram are considered as unique features. Subsequently, correspondence estimation is performed between the extracted, maximal descriptive features. This results in an assignment of points that have been identified as same features on both point clouds. Unsuitable correspondences are sorted out by a Random Sample Consensus (RANSAC) method, estimating a model of the correspondences and identifying outliers.¹¹ Based on these correspondences, the optimal transformation is determined. The data processing sequence of the coarse registration is visualized in Figure 5.

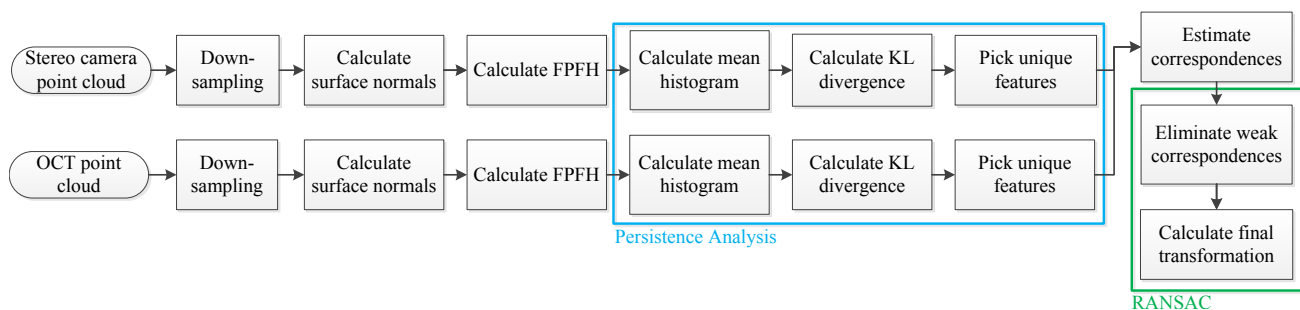


Figure 5. Sequence of feature-based coarse registration.

Figure 6 shows the surface of the stereo reconstruction (blue) and the segmented surface from the OCT volume (orange), respectively. The determined correspondences (red) are used to calculate the coarse alignment of the OCT point cloud (green). The computed transformation ${}^{\text{OCT}}\mathbf{T}_{\text{cam,coarse}}$ serves as a starting value for further registration with ICP, which is applied on the fully occupied point cloud for highest accuracy. For correspondence

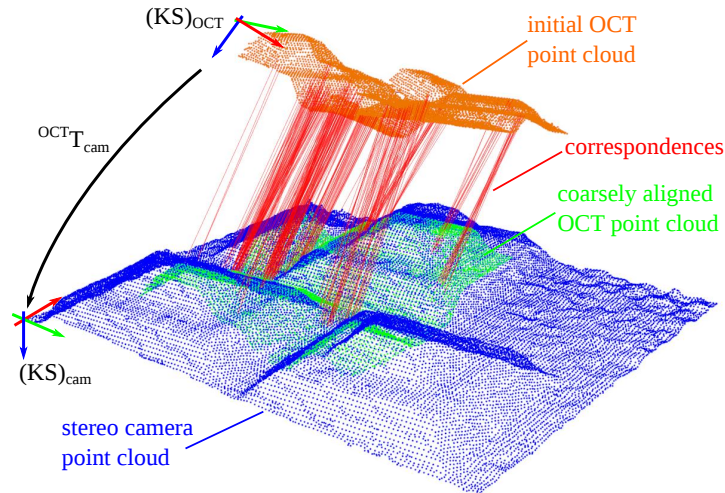


Figure 6. OCT and stereo camera point clouds with feature-based correspondences (lines).

estimation, the nearest neighbor approach is employed. The implementation was done using the Point Cloud Library (PCL)¹² in C++.

2.6 Experimental verification

Quantitative evaluation of the methods is performed after registration of stereo camera and OCT. The reconstructed and the segmented surface are registered as point clouds. The Euclidean distances between corresponding points are calculated after coarse registration and after the last registration step. Root Mean Square (RMS) serves as error metric. Additionally, the CAD model of our reference structure is neglecting manufacturing tolerances and thus used as reference. In total, five data sets from different views are captured, each consisting of a stereo camera and an OCT recording. The views were chosen in order to take multiple reasonable perspectives into account. Distance from the scene to OCT and camera was similar for each measurement, to remain in camera focus and in scanning range of the OCT. The camera angle was limited by a convenient mount of the stereo camera in the experimental setup. Therefore, the sample was rotated around its vertical axis, while OCT and camera remained in a fixed position. Figure 7 shows point clouds of all five recordings for camera and OCT after registration. Hence, the utilized OCT workspace and the camera angle for each recording are visible.

3. RESULTS

Table 1 shows the RMS of the experimental evaluation of coarse and final registration for all five data sets. In addition, the mean values over all five recordings are listed. Minimal and maximal errors are highlighted green and orange, respectively, the mean values yellow.

Measurement	OCT-Cam	OCT-Cam	OCT-CAD	OCT-CAD	Cam-CAD	Cam-CAD
	coarse [mm]	final [mm]	coarse [mm]	final [mm]	coarse [mm]	final [mm]
1	0.430	0.176	0.204	0.152	0.493	0.419
2	0.215	0.107	0.106	0.052	0.591	0.458
3	0.216	0.120	0.545	0.084	0.357	0.242
4	0.590	0.606	0.620	0.591	1.046	0.894
5	0.786	0.409	0.478	0.411	0.536	0.534
Mean	0.447	0.284	0.391	0.258	0.605	0.510

OCT to camera registrations are shown in Figure 7 for all five cases. Table 1, in combination with Figure 7, illustrates the relationship between recorded fragment of the reference structure and registration result. Aligned point clouds of OCT after coarse (green) and final registration (red) on the camera point clouds (blue) are

shown. In the five measurements, different parts of the reference structure have been recorded and used for registration. Measurements 2 and 3 both utilize a centered region of the structure, resulting in a low registration error. Contrary, measurement 4 only employs the information of one bigger and the small pyramid, producing the highest registration error. Leaving measurements 1 and 5 as average results, both utilizing more structural information as measurement 4 but less than measurement 2 and 3. Comparing measurement 1 and 5, the increased employment of pyramid structures in measurement 1 results in a lower registration error. The number of red lines reveal the number of applied correspondences between extracted unique features, indicating the number of meaningful features that have been found in the OCT workspace. Most features can be found at edges of the truncated pyramids. Therefore, key to successful feature-based registration is imaging an area of the reference structure, containing as many as possible edges. Comparably, ICP shows lower registration errors, if given better coarse alignment of point clouds and more diverse spatial information.

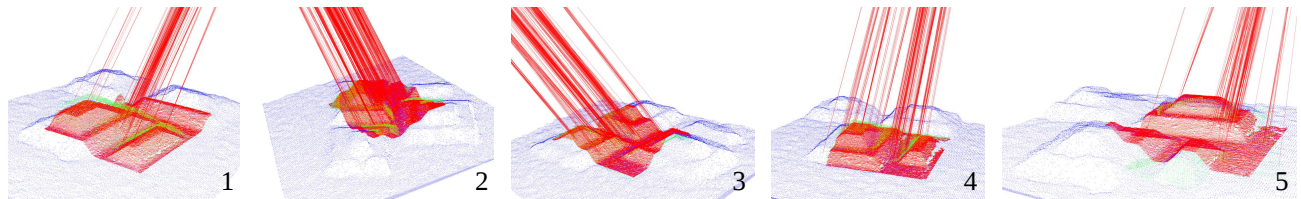


Figure 7. Alignment of OCT and camera point clouds after coarse and final registration for measurements 1-5 from left to right.

The included OCT error is determined to 40-50 μm , using a procedure developed by Díaz Díaz et al.¹³ The range of the results is given with respect to different sphere localization algorithms (hough transformation, cross correlation and least-squares).

Camera calibration was performed with a back projection RMS error of 0.512 pixels. The average computational times were 2.8 s for segmentation, 73.74 s for coarse and 7.5 s for the final registration (Intel®Core™i7-3770K 3.5GHz, 16GB RAM, Linux Ubuntu 12.04 LTS).

Two examples for augmented reality were created: The depth information, obtained from the sample surface in the OCT data, were colored related to their OCT distance and mapped to the corresponding camera points, shown left hand in Figure 8. The colors assign distance values to the camera pixels, ranging from minimal (blue) to maximal distance (red). Colored depth information can be used for a rapid evaluation of distances in surgical scenes. Additionally, we used the color information of the stereo camera to assign texture to the segmented OCT surface, which can be seen right hand in Figure 8. This provides a highly accurate surface representation from OCT data, enhanced with texture information from stereo camera data.

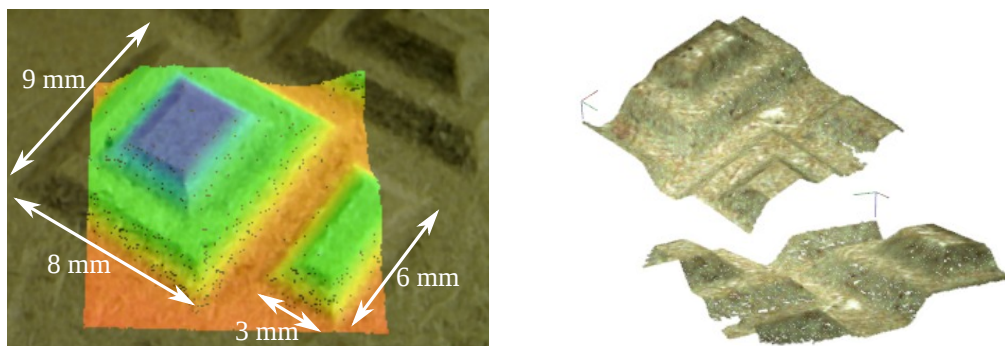


Figure 8. Augmented reality examples: Overlay of a colored depth map onto the camera image (left) and texturized OCT point cloud from two different perspectives (right).

4. CONCLUSION

The presented workflow is the basis for an integration of OCT into existing stereo imaging systems, such as stereo endoscopes² or surgical microscopes.³ Combination of OCT and stereo vision systems benefits from the advantages of both. Optical Coherence Tomography produces useful three-dimensional image data with high depth resolution and penetration in biological tissue and is therefore well suited for applications in the field of microsurgery. Stereo vision enables depth perception in a broad field of view. Hence, OCT yields high accuracy and access to subcutaneous image information, while stereo cameras provide textural information as well as a familiar view. Incorporating the known position relationship between the imaging modalities, imaging information can be fused to create a combined visualization, like augmented reality. Information, being extracted from the image data of the OCT (e.g. thickness of a tissue layer or distance to a subcutaneous structure), are superimposed onto camera images. By merging two imaging modalities, the user obtains additional information of the surgical scene, which cannot be determined with conventional methods. This fused presentation of image information enhances the user's perception and decreases the surgeon's cognitive load. In addition, measurement and monitoring is feasible for microsurgical scenarios. As one of those, laser microsurgery¹ in combination with OCT enables very accurate ablations. Future work will treat OCT data processing of subcutaneous structures, as well as their integration and visualization augmented into stereo vision.

ACKNOWLEDGMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) grant OR 196/17-2.

REFERENCES

- [1] Schoob, A., Kundrat, Kleingrothe, L., Kahrs, L. A., Andreff, N., and Ortmaier, T., "Tissue surface information for intraoperative incision planning and focus adjustment in laser surgery," *Int J Comput Assist Radiol Surg* (**online first**), 1–11 (2014).
- [2] Adler, D. C., Chen, Y., Huber, R., Schmitt, J., Connolly, J., and Fujimoto, J. G., "Three-dimensional endomicroscopy using optical coherence tomography," *Nat Photonics* **1**, 709–716 (2007).
- [3] Lankenau, E., Klinger, D., Winter, C., Malik, A., Müller, H. H., Oelckers, S., Pau, H.-W., Just, T., and Hüttmann, G., "Combining Optical Coherence Tomography (OCT) with an Operating Microscope," *Advances in Medical Engineering, Springer Proceedings in Physics* **114**, 343–348 (2007).
- [4] Ortiz, S., Siedlecki, D., Remon, L., and Marcos, S., "Optical coherence tomography for quantitative surface topography," *Appl. Opt.* **48**(35), 6708–6715 (2009).
- [5] Schoob, A., Podszus, F., Kundrat, D., Kahrs, L. A., and Ortmaier, T., "Stereoscopic Surface Reconstruction in Minimally Invasive Surgery using Efficient Non-Parametric Image Transforms," *Proceedings of the 3rd joint workshop on new technologies for computer/robot assisted surgery (CRAS)*, 26–29 (2013).
- [6] Rosenfeld, A., "Connectivity in Digital Pictures," *J ACM* **17**(1), 146–160 (1970).
- [7] Samet, H., "Connected Components Labeling using Quadrees," *J ACM* **28**(3), 487–501 (1981).
- [8] Besl, P. J. and McKay, N. D., "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**(2), 239–256 (1992).
- [9] Rusu, R. B., Blodow, N., and Beetz, M., "Fast Point Feature Histograms (FPFH) for 3D Registration," *IEEE Conference on Robotics and Automation*, 3212–3217 (2009).
- [10] Rusu, R. B., Blodow, N., Marton, Z. C., and Beetz, M., "Aligning point cloud views using persistent feature histograms," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3384–3391 (2008).
- [11] Fishler, M. A. and Bolles, R. C., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Commun ACM* **24**(6), 381–395 (1981).
- [12] Rusu, R. B. and Cousins, S., "3D is here: Point Cloud Library (PCL)," *IEEE International Conference on Robotics and Automation (ICRA)*, (2011).
- [13] Díaz Díaz, J., Rau, T., Majdani, O., and Ortmaier, T., "Suitability of optical coherence tomography for an intra-operative navigation system by estimation of the fiducial localization error," *Int J Comput Assist Radiol Surg* **7**(1), 440–441 (2012).