Sophie Beckmann*, Jean-Claude Rosenthal, Eric L. Wisotzky, Anna Hilsmann, and Peter Eisert

Automatic Registration of Anatomical Structures of Stereo-Endoscopic Point Clouds

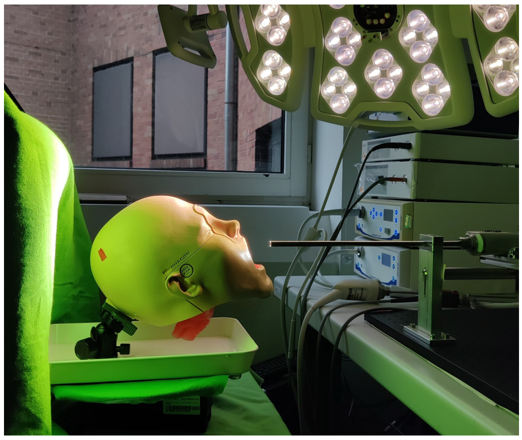
https://doi.org/10.1515/cdbme-2023-1154

Abstract: In image-guided surgery, imaging systems such as (stereo-)endoscopes allow intra-operative 3D reconstructions in form of point clouds. However, endoscopes provide only a narrow field of view, resulting in a confined point cloud. In this paper, we present an analysis and registration pipeline for confined point clouds acquired by stereo endoscopes into a fused representation. For a coarse registration, TEASER is applied, while a refinement is conducted utilizing point-to-plane ICP. The pipeline is tested on two datasets: acquired point clouds of a head phantom using a EinsteinVision® 3.0 endoscope and point clouds from the Stereo Correspondence And Reconstruction of Endoscopic Data challenge. The results for both datasets show that 3D reconstructions of anatomical structures by utilizing stereo-endoscopes and point cloud registration is a promising contactless and radiation-free method. However, non-rigid deformations are not yet incorporated and evaluation of the method compared to reference data is challenging.

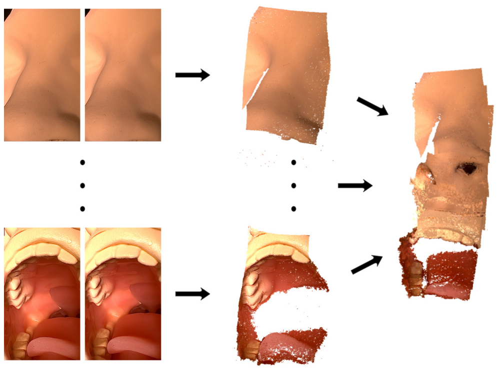
Keywords: endoscopy, stereoscopic imaging, point cloud registration, image-guided surgery, 3D reconstruction

1 Introduction

In image-guided surgery, stereo-endoscopes enable the practitioner to observe anatomical structures with a realistic depth perception [1, 6, 9]. In addition, from a stereo image pair, a 3D point cloud can be reconstructed from which the practitioner can perform an analytical assessment of the anatomy, for example, to manufacture patient-specific implants for nasal septum perforations [6]. However, stereo endoscopes provide only a limited field of view, resulting in a confined point cloud [9]. To produce a more comprehensive representation, several point clouds taken from differing perspectives may be transformed into one global coordinate system by applying point cloud registration. While point cloud registration is frequently used to reconstruct indoor or outdoor scenes, in stereo-endoscopic imaging it is sparsely applied [3]. Hence, within this paper, a pipeline for point cloud registration for stereo-endoscopic images is presented. An overview of the acquisition setup as well as the registration of the captured point clouds is given in Figure 1.



(a) Acquisition



(b) Registration

Fig. 1: Overview of the presented registration pipeline. In **(a)**, the setup for the acquisition of point clouds including a greenscreen, a turn table, a phantom and the EinsteinVision® 3.0 endoscope is shown. Note that for the data within this paper, the phantom head was tilted sideways. In **(b)**, a point cloud is generated from each stereo pair which is registered with the neighboring point clouds.

***Corresponding author: Sophie Beckmann,** Fraunhofer HHI & Humboldt-Universität zu Berlin, Germany, e-mail: sophie.beckmann@hhi.fraunhofer.de
Jean-Claude Rosenthal, Munich Surgical Imaging GmbH, Munich, Germany
Eric L. Wisotzky, Peter Eisert, Fraunhofer HHI & Humboldt-Universität zu Berlin, Germany
Anna Hilsmann, Fraunhofer HHI, Berlin, Germany

2 Methods

The presented pipeline includes the acquisition of point clouds, their analysis and finally, the registration into a joint model. An overview of the analysis and registration pipeline is given in Figure 2.

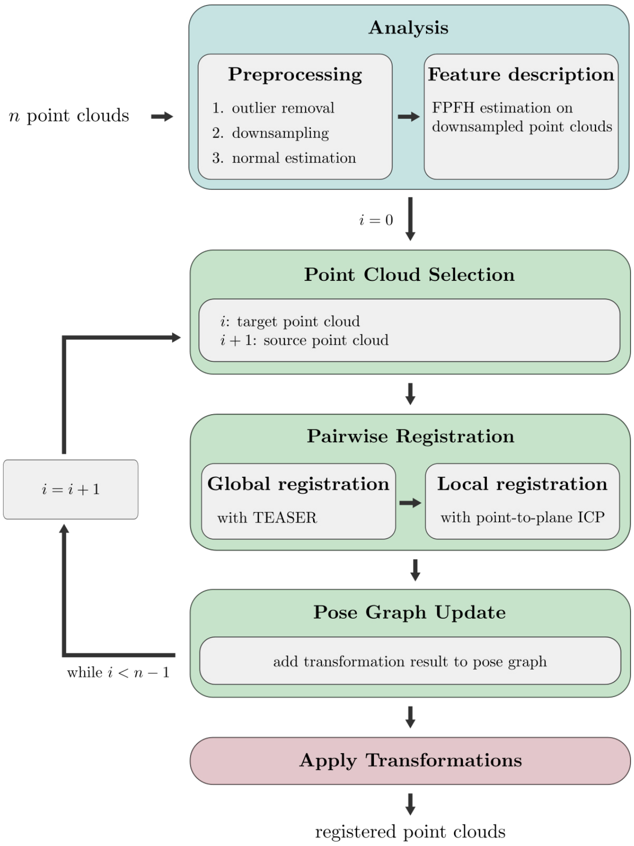


Fig. 2: Presented pipeline for the analysis and registration of point clouds acquired by a stereo-endoscope.

2.1 Acquisition

Stereo images of the throat assistant (PHACON GmbH, Leipzig, Germany) were acquired using the 3D laparoscope EinsteinVision® 3.0 (Aesculap AG, Tuttlingen, Germany). The external surgical light marLED® by KLSmartin was used for optimal illumination of the phantom. To capture the object from different angles, the phantom was placed on the RODEON TurnTable (Dr. Clauß Bild- und Datentechnik GmbH, Zwönitz, Germany) which was rotated by 5° for a total of 30° resulting in seven point clouds.

For the acquisition of point clouds, the STereoscopic ANalyzer (STAN) was utilized [12]. The software takes stereo images and calibration matrices as an input and computes point clouds based on triangulation [6, 12]. The internal and external camera calibration of the stereo-endoscopic system was conducted using a model-based analysis by synthesis technique [2]. To omit saving unwanted background, green screen removal and a distance clipping threshold is implemented in STAN.

In addition to the acquired stereo images, the analysis and registration pipeline was evaluated on point clouds computed by Rosenthal et al. in the course of the Stereo Correspondence And Reconstruction of Endoscopic Data (SCARED) challenge where stereo images of porcine cadavers were captured utilizing a *da Vinci Xi* surgical robot [1].

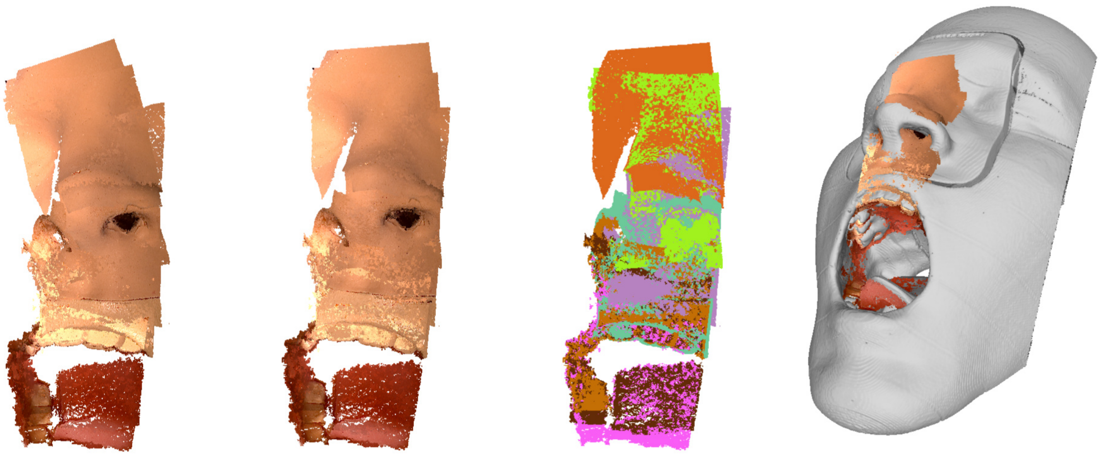
2.2 Point Cloud Analysis

Point cloud analysis includes the removal of outliers, downsampling, normal estimation, geometric feature extraction, as well as correspondence estimation. The analysis is based on Open3D, a library for 3D data processing [11]. A statistical outlier removal is performed. All points that are further away from their 30 nearest neighbors with a standard deviation of 2.0 are deleted.. Point clouds were downsampled for global registration with a voxel size of 1mm. Normal estimation was conducted on both original and downsampled point clouds using Principal Component Analysis (PCA) and based on the surface normals, local geometric feature description Fast Point Feature Histograms (FPFH) were conducted on downsampled point clouds [7].

2.3 Point Cloud Registration

A point cloud registration algorithm targets at finding the optimal transformation of a source point cloud $\mathcal{Q}$ to map onto the target point cloud $\mathcal{P}$ to achieve the best alignment between these point clouds. In this paper, pairwise registration is conducted on neighboring point clouds by firstly applying a global and then a local registration algorithm.

For a coarse alignment, the global correspondence-based registration method Truncated least squares Estimation And SEmidefinite Relaxation (TEASER) was implemented. The method determines scale, rotation, and translation between the target point cloud $\mathcal{P}$ and the source point cloud $\mathcal{Q}$ by employing a non-linear and non-convex least squares method [10]. With this method, the cost is constant for large residuals but is equal to the least squares function for small residuals. For the noise bound, which defines the width of the least-squares function within the mentioned loss function and therefore the amount of inliers, a value of 1mm was selected.



(a) TEASER **(b)** TEASER + ICP **(c)** TEASER + ICP (uniform colors) **(d)** Registration with CT reconstruction

Fig. 3: Results for global and local point cloud registration of seven point clouds captured of the throat assistant. The registered point cloud after applying **(a)** TEASER and **(b-c)** additionally point-to-plane ICP. In **(d)**, the result from **(b)** was registered to a 3D CT reconstruction by selecting five correspondences manually and additionally applying point-to-plane Iterative Closest Point (ICP).

To refine the result of the global registration method, point-to-plane ICP, a method for simultaneous correspondence search and pose estimation, is used [4]. The algorithm uses the surface normals of the target point cloud within the error function and for each point in $\mathcal{Q}$, it tries to find a corresponding point in $\mathcal{P}$ within the radius of a distance threshold. Within this paper, the distance threshold was set to 0.25mm.

3 Results

With the EinsteinVision 3.0 endoscope, seven point clouds of the throat assistant were acquired. The resulting point clouds, both for the global registration using TEASER and the additional local registration using point-to-plane ICP, can be seen in Figure 3. Although global registration results in a recognizable reconstruction of the anatomy, some areas appear to be noisy or distorted specifically at the front teeth and the bridge of the nose. After applying local registration, the teeth exhibit fewer deformities and the nose is less tilted. The manual registration with a 3D Computer Tomography (CT) reconstruction of the phantom shows that the captured point clouds and their registration yields a reasonable result. However, offsets are visible on the tongue, upper molars, and nostrils.

The registration result for the selected SCARED dataset can be seen in Figure 4. The coarse registration results in a recognizable reconstruction, but a change of contrast is visible in the center. With the refinement, the edge at the center of the reconstruction mitigates.

The registration results is evaluated using the fitness F and the inlier Root Mean Square Error (RMSE), where N_{fi} is the size of the inlier correspondence set and N_p the number of points in the target point cloud:

$$F = \frac{N_{fi}}{N_p}, \quad RMSE = \sqrt{\frac{\sum_{i=1}^{N_{fi}} (p_i - q_i)^2}{N_{fi}}} \tag{1}$$

A statistical evaluation for the results of both datasets is given in Table 1. The evaluation indicates that the registration is stable over several pairwise registrations shown by the small standard deviation and difference between mean and median.

Tab. 1: Statistical evaluation of the RMSE and fitness for the captured dataset using the EinsteinVision® 3.0 endoscope and the selected SCARED dataset.

		EinsteinVision® 3.0	**SCARED**
RMSE	**median**	0.127	0.130
	mean	0.123 ± 0.006	0.131 ± 0.011
Fitness	**median**	0.523	0.633
	mean	0.527 ± 0.042	0.616 ± 0.130

4 Discussion and Conclusion

In this paper, a pipeline for the analysis and registration of point clouds acquired by stereo-endoscopes was presented.

(a) TEASER

(b) TEASER + point-to-plane ICP

Fig. 4: Results for global and local point cloud registration of seven point clouds captured of the throat assistant. The registered point cloud after the application of **(a)** TEASER and **(b)** additionally point-to-plane ICP.

The results show that the registration of multiple point clouds is a feasible option to extend a single shot. While the coarse registration using TEASER results in an identifiable reconstruction, the refinement utilizing point-to-plane ICP enhances the result at edges, for example at the front teeth, and improves the fading of neighboring point clouds.

Additional quantitative evaluation using e.g. a patch-wise RMSE or a targeted registration error and further experiments evaluating the reconstruction reproducibility and the influence of overlap would be beneficial for further assessment.

Further improvements would include a non-rigid transformation for the registration and the automatic registration between the stereo-endoscopic reconstruction and the CT reconstruction. However, non-rigid transformations are specifically challenging. In addition, since a point cloud sampled from a CT reconstruction is considerably larger than the registered stereo point clouds, prior segmentation of the CT and/or a keypoint detector might be useful.

Acknowledgment: The authors would like to thank ImFusion, Munich, Germany, for providing the ImFusion Suite software for the visualization and processing of 2D/3D medical CT data.

Author Statement

Research funding: This research was funded by the German Research Foundation (DFG), CRC 1404: "FONDA: Foundations of Workflows for Large-Scale Scientific Data Analysis" and the German Federal Ministry of Education and Research (BMBF) under the grant number 13GW0581C (EndoScan3D). Conflict of interest: Authors state no conflict of interest. Informed consent and ethical approval is not applicable for this paper.

References

[1] Allan M, Mcleod J, Wang C, Rosenthal JC, Hu Z, Gard N, et al. Stereo correspondence and reconstruction of endoscopic data challenge. arXiv preprint, 2021, arXiv:2101.01133

[2] Eisert P. Model-Based Camera Calibration Using Analysis by Synthesis Techniques. In Proc. Int. Workshop on Vision, Modelling and Visualization (VMV), 2002, pp. 307-314.

[3] Huang X, Mei G, Zhang J, Abbas R. A comprehensive survey on point cloud registration. arXiv preprint, 2021, arXiv:2103.02690.

[4] Low, KL. Linear least-squares optimization for point-to-plane icp surface registration. Chapel Hill, University of North Carolina, 2004, 4(10):1-3.

[5] Rosenthal JC, Wisotzky EL, Eisert P, Uecker FC. Endoscopic Single-shot 3D Reconstruction of Oral Cavity, 41st Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019.

[6] Rosenthal JC, Wisotzky EL, Matuschek C, Hobl, M, Hilsmann A, Eisert P, Uecker FC. Endoscopic measurement of nasal septum perforations. HNO Journal, 2022, 70(S1):1-7.

[7] Rusu RB, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration. In 2009 IEEE international conference on robotics and automation, 2009, pp. 3212-3217.

[8] Wisotzky EL, Rosenthal JC, Eisert P, Hilsmann A, Schmid F, Bauer M, Schneider A, Uecker FC. Interactive and multimodal-based augmented reality for remote assistance using a digital surgical microscope. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 2019, pp. 1477-1484.

[9] Wisotzky EL, Rosenthal JC, Wege U, Hilsmann A, Eisert P, Uecker FC. Surgical guidance for removal of cholesteatoma using a multispectral 3D-endoscope. Sensors, 2020, 20(18):5334.

[10] Yang H, Shi J, Carlone L. Teaser: Fast and certifiable point cloud registration. IEEE Transactions on Robotics, 2020, 37(2):314-333.

[11] Zhou QY, Park J, Koltun V. Open3D: A modern library for 3D data processing. arXiv preprint, 2018, arXiv:1801.09847.

[12] Zilly F, Müller M, Eisert P, Kauff P. The stereoscopic analyzer - an image-based assistance tool for stereo shooting and 3D production. In 2010 IEEE Int. Conf. on Image Processing, 2010, pp. 4029-4032.