Sarascope 2.0: A Novel Endoscopic System for Robotic Minimally Invasive Surgery

Andrea Roberti[1], Nicola Piccinelli[1], Michele Sandrini[1], Matteo Piano[1] and Riccardo Muradore[1]

[1]*Department of Engineering for Innovation Medicine, University of Verona, Italy*
name.surname@univr.it

Introduction

In Robotic Minimally Invasive Surgery (RMIS), accurate intra-operative depth perception is critical for complex maneuvering and is a prerequisite for advanced computer-assisted interventions, such as Augmented Reality (AR) [1] and autonomous task execution [2]. While stereoscopic endoscopes address the limitations of monocular systems, they are severely constrained by the anatomical access ports (trocars). Conventional stereo endoscopes must fit both optical channels within a 10 mm or 12 mm diameter shaft, resulting in a narrow stereo baseline (typically <8 mm). This intrinsic limitation leads to poor triangulation accuracy, especially at increasing working distances. Several research efforts have explored methods to overcome this limitation. Early work in computer vision investigated the concept of variable-baseline stereo, demonstrating that increasing the distance between viewpoints significantly improves depth estimation accuracy, particularly in mid- to long-range scenarios [3]. However, translating these principles into endoscopic systems remains challenging due to strict size and access constraints. Recent developments in both industry and academia aim to overcome fixed-baseline limitations in endoscopy. The Vicarious Surgical platform utilizes a multi-camera, distributed viewpoint architecture to broaden 3D perception. Similarly, Salcudean et al. (UBC) [4] have developed 'pickup' stereoscopic cameras that can be repositioned intra-operatively to optimize sensing baselines and reconstruction accuracy. Building upon previous work in endoscopy [5], we introduce Sarascope 2.0, shown in Figure 1, a completely new concept focused on overcoming the baseline limitation. Unlike standard fixed-geometry endoscopes, Sarascope 2.0 features a rigid shaft with a deployable distal tip (Figure 2). Through a novel mechanical "tilt" actuation, the distal cameras rotate outwards upon reaching the surgical site, achieving a wide extended baseline of the stereo cameras of 2.5 cm. This significant increase in baseline facilitates high-fidelity 3D reconstruction previously unattainable with standard laparoscopic tools. In this paper, we present the system architecture of Sarascope 2.0, the calibration methodology developed and experimental results validating the quality of the resulting dense 3D point clouds and reconstruction.

Fig. 1 Sarascope 2.0.: on the right the tip with cameras.

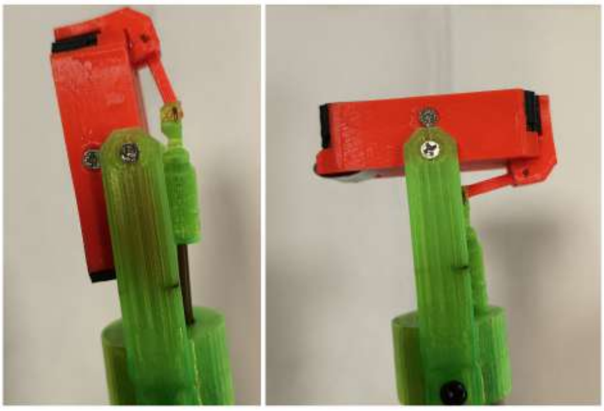

Fig. 2 Deployable distal tip of the Sarascope 2.0.

System Description and Software Architecture

Sarascope 2.0 is a compact and portable stereo endoscope designed to overcome baseline limitations while remaining compatible with robotic surgical platforms. The system consists of three main components: (i) a proximal electronics module, (ii) a rigid shaft, and (iii) a deployable distal flange integrating two RGB cameras and an illumination unit. To comply with trocar constraints, the distal flange remains closed during insertion and is mechanically deployed once inside the surgical cavity. A tilt actuation mechanism rotates the cameras outward, achieving a stereo baseline of 2.5 cm. This wider baseline improves triangulation accuracy, particularly at increasing working distances. Actuation is provided by an MG996R servo motor located in the proximal module, which drives a rack-and-pinion system. Motion is transmitted to the distal tip via an internal rod running through the shaft, which also houses the camera wiring. The mechanical design ensures repeatable deployment and stable positioning of the stereo pair during operation. The flange includes dimmable LEDs for adjustable illumination. The system is controlled by a Raspberry Pi 5, which handles both device actuation and image processing within a modular C++/ROS2 framework. Synchronized image acquisition from the stereo pair is performed in real time, enabling downstream 3D reconstruction. Camera calibration is performed using Zhang's method [6] to estimate intrinsic parameters and distortion coefficients for each sensor. Stereo calibration is then applied to compute the relative pose between the cameras, defining the effective baseline. Rectified image pairs are processed using standard ROS2 stereo pipelines to compute disparity maps and generate dense 3D point clouds in real time.

Results and Discussion

To quantitatively assess the accuracy of the 3D reconstruction, a custom cubic support was designed in Autodesk Fusion 360 and fabricated using a 3D printer (Figure 3.a). The reference cube has a nominal edge length of 4 cm. The object was acquired using the Sarascope 2.0 system, and a dense 3D point cloud was reconstructed following the proposed stereo pipeline. Figure 3.b shows the resulting mesh, highlighting the overall geometric consistency of the reconstruction. The reconstructed cube dimensions were estimated by measuring the edge lengths along multiple directions, yielding a mean value of 3.979 cm with a standard deviation of 0.031 mm, which is very close to the nominal value. Additionally, a planar surface of the cube was extracted and used for plane fitting, resulting in a root mean square (RMSE) error of 1.051 mm. These results indicate low geometric noise and consistent surface reconstruction. To evaluate the impact of the increased stereo baseline, we refer to the well-established relationship in stereo vision stating that depth estimation error grows quadratically with the distance from the camera and is inversely proportional to the stereo baseline. This implies that increasing the baseline directly improves depth accuracy, particularly at larger working distances. In this context, the proposed 2.5 cm baseline provides a clear improvement over conventional stereo endoscopes, which are typically limited to baselines below 8 mm. The experimental results are consistent with this behavior: the reconstructed mesh (Figure 3.b) exhibits sharp edges and coherent planar surfaces, indicating improved triangulation accuracy enabled by the wider baseline. These observations support the effectiveness of the proposed deployable mechanism in enhancing depth perception compared to standard stereo configurations.

To further assess performance in a more realistic scenario, we conducted a qualitative experiment using a cholecystectomy phantom. The resulting 3D reconstruction demonstrates the ability of the system to capture complex anatomical structures, preserving coherent geometry despite increased visual complexity compared to the reference cube. These results highlight the potential of Sarascope 2.0 in realistic surgical environments. Figure 4 shows an example of the reconstructed phantom.

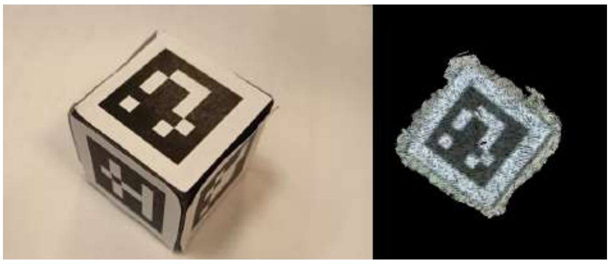

Fig. 3 Qualitative results of the 3D reconstruction of the reference cube.

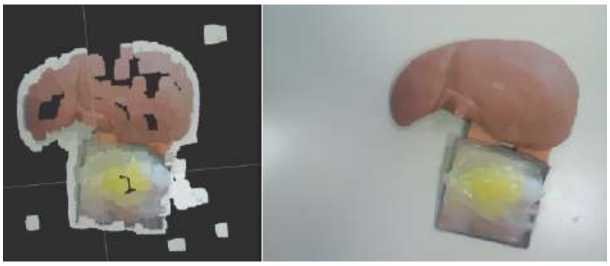

Fig. 4 Qualitative 3D reconstruction of a cholecystectomy phantom.

The current prototype has an outer diameter of 2 cm, primarily constrained by the flat ribbon cables. Future iterations will focus on reducing the diameter to comply with standard surgical requirements and addressing sterilization constraints. Although the evaluation was conducted under controlled, static conditions, the results demonstrate strong potential for intra-operative 3D perception. Future work will investigate robustness in dynamic and realistic surgical environments, including deformable tissues and challenging lighting conditions, as well as applications in augmented reality guidance and autonomous robotic assistance.

REFERENCES

[1] J. Seetohul, M. Shafiee, and K. Sirlantzis, "Augmented reality (ar) for surgical robotic and autonomous systems: state of the art, challenges, and solutions," *Sensors*, vol. 23, no. 13, p. 6202, 2023.

[2] N. Preda, F. Ferraguti, G. De Rossi, C. Secchi, R. Muradore, P. Fiorini, and M. Bonfè, "A cognitive robot control architecture for autonomous execution of surgical tasks," *Journal of Medical Robotics Research*, vol. 1, no. 04, p. 1650008, 2016.

[3] D. Gallup, J.-M. Frahm, P. Mordohai, and M. Pollefeys, "Variable baseline/resolution stereo," in *2008 IEEE conference on computer vision and pattern recognition*. IEEE, 2008, pp. 1–8.

[4] G. Samei, S. E. Salcudean, C. N. Schneider, and R. Rohling, "A stereoscopic "pickup" camera for the da vinci surgical system," *International journal of computer assisted radiology and surgery*, vol. 10, pp. 745–754, 2015.

[5] A. Roberti, N. Piccinelli, F. Falezza, G. De Rossi, S. Bonora, F. Setti, P. Fiorini, and R. Muradore, "A time-of-flight stereoscopic endoscope for anatomical 3d reconstruction," in *2021 International Symposium on Medical Robotics (ISMR)*. IEEE, 2021, pp. 1–7.

[6] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2002.