Interdisciplinary Concept of Design, Actuation, Control and Vision Computing of a Laser-Sintered Single-Port Robot

Mattias F. Traeger, Student Member, IEEE, Daniel B. Roppenecker, Brian Jensen, Salman Can, Alois Knoll, Tim C. Lueth, Senior Member, IEEE

Abstract—One standard procedure in minimally invasive surgery is the endoscopic submucosal dissection (ESD), where neoplastic mucosa is being removed. The standard endoscopes and flexible instruments currently used in everyday routine have limited maneuverability and are controlled in an unintuitive manner. To overcome these drawbacks we propose a telemanipulation system capable of independently controlling two standard flexible endoscopic instruments with sufficient triangulation that is comparable to the dexterity of classical open surgery. By using the rapid manufacturing technique selective laser sintering (SLS) of a biocompatible polyamide, we want to examine basic design principles to print our telemanipulation system in one part and to individualize its characteristics for patient specific print-out in the future. In this paper we give an overview of our system from the mechanical design via control and actuation concept to visual environment perception.

I. INTRODUCTION

Minimally invasive surgery arose in the early 90’s and had its first applications in laparoscopic colon resection [1] and laparoscopic cholecystectomy [2], [3]. One camera-optic and usually two instruments with straight and rigid shafts are inserted through small incisions via one trocar each into the abdominal cavity. In order to have sufficient space, the abdominal cavity is inflated with CO₂. Nowadays the challenge is to further reduce the burden on patients to shorten the intervention’s duration and to improve the clinical and aesthetic result. In laparoscopic surgery the next step is the so-called single-port surgery, where the camera and two instruments are inserted together through only one trocar. Besides the laparoscopic interventions another large branch is the flexible endoscopy, in which diseases inside the stomach or the colon are treated. One standard procedure is the Endoscopic Submucosal Dissection (ESD) where neoplastic mucosa is removed [4]. The standard equipment is one flexible gastroscope with one or two flexible endoscopic instruments. The challenge is to develop instruments that offer enough dexterity so the surgeon can reliably manipulate the tissue in an intuitive manner. For this purpose we develop a system produced with selective laser sintering (SLS) of a biocompatible polyamide. The advantage is to produce patient specific instruments at low cost. The challenge is to automate the construction process and to control the laser-sintered structure, where flexure hinges are used instead of common joints. The design concept in Figure 1 shows the interaction between the hardware and software as well as the interaction between the surgeon and the patient. One of the crucial aspects regarding the clinical usability is the sterile barrier between the sterile and non-sterile part of the single-port system, i.e. between the actuators and the laser-sintered overtube. Also mastering the SLS process in order to meet the ISO 10993 requirements (series of standards for evaluating the biocompatibility of a medical device prior to a clinical study) is a necessary for clinical applicability. These different aspects and their challenges are described in the following sections.

![Figure 1: Design Concept of the single-port system. Blue arrows indicate interaction between surgeon, single-port system and patient. Orange arrows indicate interaction between soft- and hardware. Control commands are transmitted via a serial interface (RS422). The system is mainly produced using SLS.](image-url)
II. MECHANICAL DESIGN / SLS / FLEXURE JOINTS

A. Manufacturing with SLS

Selective laser sintering (SLS) is a new additive manufacturing method. 3D-files like STL-files get sliced into layers and can be produced layerwise by SLS.

![Diagram of Manufacturing with SLS](image)

During production a polymer powder, e.g., polyamide (PA 2200) is applied to surface of the building plate. Using a laser, the powder layer is melted at local spots. Afterwards the powder hardens on the lased spots to the desired part. After a short cooling period a new layer of polyamide is provided by the coater. We are using a layer thickness of 0.1 mm. The layer gets heated close to the melting temperature range between 176-181°C. The laser starts to melt the new layer and this process continues layer by layer. Our proposed parts are all realized by the selective laser sintering machine Formiga P 100 (© EOS GmbH, Krailling, Germany) which is a powder based SLS machine. The size of the building plate is 200x250mm (x-y-direction) and 300mm in z-direction. The applied laser is a CO₂ laser with an effective power of 30W (wavelength 10.2 to 10.8 µm). The focused laser spot has a diameter of 0.43-0.45 mm [5]. The powder particle size is in between 56-65 µm. Sintered parts have a density of ~0.93 g/cm³ with a flexural modulus of elasticity of 1500 MPa and a shore hardness D of 75. The sintered parts show anisotropic material behavior (based on the layer-wise production) which means the properties depend on the direction of the parts inside the building area (all data is based on EOS©, machine manual and the datasheet for PA2200). Laser-sintered parts characteristically have a high range of design freedom due to their layer-wise production. Nonetheless there are still some limitations during the design process, e.g., the minimal wall thicknesses of 0.5mm.

During SLS different plastics and even metals can be handled. For our purposes we used the material: PA 2200 a certification for biocompatibility (according to DIN EN ISO 10993-1) for raw powder and a mixture of raw and recycled powder (ratio 50:50) is provided by EOS. Currently only the powder is certified for biocompatibility, sintered parts are still pending certification, as this is typically the domain of the manufacturer of the medical product. Besides the biocompatibility the material is also approved for food contact according to EU Plastics Directive 2002/72/EC.

B. Post-processing after manufacturing:

After an adequate cooling period the parts can be removed from the machine (to avoid delay and stress inside the parts). After a manually cleaning of the parts to remove loose powder, different finishing (post-processing) treatments can be considered:

- Pressurized air (pressure depends on the rigidity of the parts, e.g., up to eight bar for strong parts and max. two bars for fragile parts)
- Blast cleaning (lower pressure than by using only pressurized air)
- Use of an ultrasonic bath (cleaned water)
- Vibratory grinding
- Surface polishing
- Other treatments: e.g., flocking, coloring, enameling, bonding, powder coating, and others (based on data from EOS©).

C. Structure of the Multi-Arm Snake-Like Robot

Our robot system as proposed in [5] and [6] consists of an overtube frame for a flexible endoscope and a tip with two hollow Multi-Arm Snake-Like Robot kinematics [6]. The kinematics are produced monolithically by selective laser sintering. The endoscope is inserted into the overtube and follows the endoscope’s motions in a passive manner.

III. ACTUATION AND STERILE COUPLING

The purpose of an ESD is the resection of neoplastic mucosal tissue in one piece. Therefore reliable manipulation of mucosal tissue is essential. In contrast to classical open surgery the surgeon faces the problem of limited dexterity regarding the control of the instruments in minimally invasive surgery. Especially during ESD, where only one single instrument can be used for active manipulation [7]. The ideal solution would be a system that allows the surgeon to control two instruments independently in order to oppose them. This configuration is the so-called triangulation [7, 8].
A. Kinematic Concept

In our previous works we tested the possibilities of SLS and presented some examples of flexible structures, forming a hollow arm to insert endoscopic instruments [5], [6], [9]–[13]. The structure basically consisted of a helical spiral with a hole in its middle to insert endoscopic instruments (Figure 4). To improve the stiffness and the triangulation, we developed a new kinematic that consists of flexure hinge segments, as shown in Figure 8. These segments are actuated via push rods and Bowden cables and each arm has two push rods to mainly bend the arms outwards and one Bowden cable to pull the tip towards the middle. The Bowden Cables and push rods are guided via medical tubings with an inner diameter of 0.55 mm and an outer diameter of 0.86 mm. They are actuated by linear motors with a stroke of 30 mm each. With this approach we had been able to triangulate the instruments and perform a pick and place task (Figure 8-a) and demonstrate the ability to apply sufficient forces by lifting weights (Figure 8-b, c). The usable workspace dimension is shown in Figure 7 and has the dimensions

- Single arm: 80x70x50 mm³ (WxHxD)
- Cooperative area: 60x50x30 mm³ (WxHxD)

Due to the use of SLS we can easily adapt the workspace to specific situations.

B. Sterile Coupling

A crucial aspect regarding the clinical use might be the need to have a sterile barrier between the sterilized manipulator that is inserted into the patient and the electronics and actuators needed to make it move. For this purpose we chose to transmit the motor’s movements via a magnetic coupling to the Bowden cables and push rods, as in Figure 5. With this approach it is possible to easily insert a sterile foil between the non-sterile and the sterile part of the system.

Figure 4: Early prototype with helical structure actuated with cables. One forceps and one cauterizing hook inserted. The third arm in the middle can be used for a second camera or another instrument (Taken from [5], [11]).

Figure 5: Control concept (from left to right): The user controls the motors via a linear potentiometer. The signal is transmitted via an ADC to a microcontroller that generates the PWM signal for the motors. The motors’ movement is transmitted via a sterile barrier (red zigzag line) to the manipulator’s flexible arms to control the instruments.

Figure 6: Setup of the first working laser sintered prototype with push rods. Two endoscopic instruments (2) are inserted through the overtube (3) into the flexible arms (1). The arms are driven by push rods and Bowden cables (4) that are connected via a magnetic coupling (red zigzag lines) with linear actuators (5). The actuators are controlled manually via linear potentiometers (6), one for each actuator.

Figure 7: Workspace simulation of the laser-sintered, spine-inspired flexible arms: a) view from the side, b) view from top.

Figure 8: Early prototype for endoscopic interventions (total diameter 19 mm) with a third arm for a camera and the spine-inspired kinematic structure actuated with push rods demonstrating a) triangulation with two flexible forceps, b) radial weightlifting (200 g) and c) axial weightlifting (300 g).
IV. CONTROL OF THE SINGLE-PORT SYSTEM

The control architecture for a single-port system comprising all entities including the high-level control, vision and sensory inputs was proposed in [14]. The described structure, which was initially implemented for a laparoscopic single-port system, has been adapted to enable the control for our flexible NOTES and single-port platform. The concept provides the interaction between the user and the telemanipulation system by using a modular architecture including a high- and low-level control, a communication protocol for the individual entities and visual processing to control the system. It gets e.g. the reference position from the input device and provides the desired position to the telemanipulator or the simulation. The communication for the low-level control is realized by using a serial interface (e.g. RS422). A suitable protocol for the communication between the other entities will be defined during the further development.

A. Simulation of the Surgical Scenario

The simulation of the platform with the manipulators as well as the flexible NOTES and single-port application scenario with the surgical environment has been implemented in order to evaluate the control, manipulability and system interaction parallel to the hardware development. An early model of the single-port system is shown in Figure 9. It has two flexible arms mounted on the bending section of a flexible endoscope and is visualized inside the abdominal cavity of a human body. The flexible bending sections of the laser-sintered manipulators are simplified by a periodical anchoring of cylindrical tubes with hinges of 1.5 mm in length for a simulation of the control concept. It is assumed that the curvature of the bending section is consistent and each hinge bends in equal amount. With different number and arrangement of the tube elements, various manipulators with multiple bending segments and lengths can be realized. In accordance with the requirements of the physicians, the length and stiffness of the flexible arms can be adjusted to the patient’s specific anatomical situation. In Figure 9 the left arm is shorter with a ratio of 2:3 to the right arm. The third arm for an additional camera is not taken into account in this simulation. In the current state, the instruments are attached at the manipulator’s tip and enables one rotary and the opening/closing of each forceps. Further details of the simulated manipulators can be obtained in [11]. Although the conventional flexible endoscope is controlled manually by the physician, a bendable section was implemented to simulate the control of the endoscope and the attainable workspace of the system.

B. Human Machine Interface (HMI)

As illustrated in Figure 10, the front and rear bending sections of the manipulator in the simulation has one DOF in horizontal direction and the front bending section has one additional DOF in vertical direction. The flexible instruments and the endoscope are inserted through the system’s frame. Using the approach to simplify the flexible hinges of the laser sintered kinematics, a joint-space control of the manipulators was implemented in the current solution in order to test whether the intended control concepts are easy to understand and intuitive for the physicians. Initial tests with common input devices such as joystick and space mouse showed the functionality of the implemented simulation and control. However, they were not enough intuitive to control the manipulators. Physicians preferred their accustomed input devices, such as scissor handles of rigid, laparoscopic instruments. We modified a joystick by integrating a laparoscopic instrument handle (Figure 10). The front bending section is controlled with the two rotary axes of the joystick and the rear bending section with buttons that are mounted on the handle. The instruments can be rotated and opened/closed by the rotation or opening/closing of the grip. Initial evaluation showed that this simple, prototypical input device seems suitable and will be optimized and enhanced in further research.

Figure 10: Joint-space control with a modified joystick as input device.

V. VISUAL ENVIRONMENT PERCEPTION

The human machine interface of the flexible NOTES and single-port system is augmented with geometric environment information extracted from the endoscopic stereo camera, as shown in figure 12. A primary goal is to mitigate the surgeon’s limited field of view and compensate for the lack of haptic feedback by maintaining a 3D model of the patient’s anatomy, while simultaneously tracking the manipulators and estimating their pose relative to the tissue surfaces, thereby increasing the operating precision and safety of the single-port system.

A. Stereoscopy

Stereoscopic cameras have been previously deployed with several laparoscopic robotic surgery systems, such as the daVinci (Intuitive Surgical, Inc, Sunnyvale, CA, USA). Initially aiming to improve the surgeon’s depth perception capabilities, some success has been demonstrated for algorithmically reconstructing the depth of the operating scene [15], [16]. Although stereo reconstruction is rich field of research, the typical NOTES and single-port MIS scenarios pose difficulties for many otherwise well performing reconstruction algorithms due to the high self-similarity and
the many smooth curved surfaces, which can be compensated for by using a robust regularization strategy. The stereo reconstruction algorithm employed for the single-port system has been previously covered in detail [17] and is briefly reproduced here for completeness. Stereo reconstruction for the single-port systems involves two calibrated and rectified cameras in horizontal alignment. In this stereo setup corresponding images, \( l_l \) and \( l_r \): \( \Omega_l \rightarrow \mathbb{R}^2 \) and \( \Omega_r \subseteq \mathbb{R}^2 \), from the left and right camera respectively are related by the constraint \( l_l(x, y) = l_r(x - d, y) \) for the projection of an arbitrary 3D scene point that is visible in both the left \( l_l \) and right \( l_r \) images with a disparity \( d \). Since we assume no prior knowledge about the 3D geometry of the scene we seek a mapping \( d: \Omega \rightarrow \mathbb{R}_+ \) that maps image points to disparities and \( n: \Omega \rightarrow \{ x \in \mathbb{R}^2 : |x| \leq 1 \} \) a normal map that encodes for each image point its corresponding 3D scene point’s surface orientation, which implicitly relates neighboring image point disparity values. We evaluate the quality of a given mapping pair \( d, n \) using an energy function formulation

\[
E(d, n) = E_{\text{data}}(d, n) + \lambda E_{\text{smoothness}}(d, n)
\]

that is comprised of a data term \( E_{\text{data}}(d, n) \) that encodes the dissimilarity between corresponding stereo image points and a smoothness term \( E_{\text{smoothness}}(d, n) \) that favors configurations where neighboring image points have similar normal and disparity values. The reconstruction algorithm thus searches for the disparity \( d \) and normal \( n \) maps that minimize the energy function. The data term is implemented using a patch match stereo formulation. The smoothness term captures dissimilarity between neighboring points with the robust Huber norm together with a weighting function to allow for depth discontinuities at image locations with strong gradients. Because this energy function is difficult to optimize in its original form, we approximate it using quadratic relaxation that decouples the data and smoothness terms, where both convex sub problems can be solved efficiently via alternation. Figure 11 illustrates the resulting depth calculation.

\textbf{B. Visual SLAM}

There has been much research related to 3D mapping and localization in general robotics applications, but there are a variety of challenges in adapting such visual processing approaches to single-port scenarios. Most visual SLAM approaches are based upon the assumption that the elements in the scene are rigid, that the illumination of scene is constant, and that there plentiful discriminative feature points distributed throughout the scene, all of which do not necessarily hold for single-port scenarios.

We chose a visual SLAM approach similar to that outlined in [18]. In this approach an Extended Kalman Filter (EKF) formulation is used to maintain and update the SLAM system state. The system state vector \( x_t = \{ C_t, P_1, P_2, P_3 \} \) is composed of the camera state \( C_t = \{ t, r, v, \omega \} \), itself consisting of translation \( t \) and rotation vector \( r \), encoded using Rodrigues formula, relative to the world frame, and angular \( \omega \) and linear velocity \( v \) vectors, as well as a set of tracked 3D feature points \( P_i \) from the scene. Updates to the system are processed using each frame pair from the stereo camera individually. After a bootstrapping process, where an initial set of 3D scene points are estimated using matches between the stereo images, subsequent stereo images are aligned to the 3D scene using a 2D-3D reciprocal feature matching strategy where the predicted current state \( x_{t-1} \) is used to restrict the feature correspondence search windows. A sliding window technique is used to govern adding or removing 3D points to the system state and improve robustness towards scene deformation. If a 3D scene point was not in the inlier set for a sliding window of the last \( w \) frames it is removed from the state. The sliding window size \( w \) is adjusted with each system update so that the total inlier rate \( \rho \) over all frames in the window is above a given threshold. Each time the sliding window is shrunk new 3D scene points are added to the state. After each system update a point cloud from the corresponding dense stereo reconstruction is fused to an external map using the current camera pose, providing for improved visualization of surgical site.

\textbf{C. Instrument Detection and Tracking}

A variety different approaches have been applied to task of surgical instrument detection, ranging from pure feature based approaches to heuristic based image processing techniques [19], [20], with recent work showing the potential of probabilistic segmentation approaches for instrument detection and pose estimation [21]. For general-purpose multi-class object segmentation and classification tasks conditional random field (CRF) formulations in particular have been shown to have excellent performance [22] and are well suited to applications in NOTES and single-port scenarios. Given a graph \( \mathcal{G}(\mathcal{J}) = (\mathcal{V}(\mathcal{J}), \mathcal{E}(\mathcal{J})) \) over an image \( \mathcal{I} \) where \( \mathcal{V}(\mathcal{J}) = \{ X_1, \ldots, X_N \} \) is a set of \( N \) vertices representing random variables over all the image regions and \( \mathcal{E}(\mathcal{J}) \) the set of edges between these adjacent regions, the pair \( (X, \mathcal{J}) \) constitutes a conditional random field if

\[
P(X_i | \mathcal{J}, X_j, \forall j \neq i) = P \left( X_i | \mathcal{J}, X_j, j \in \mathcal{N}(X_i) \right)
\]

the Markov property holds for all \( X_i \) conditioned on \( \mathcal{J} \) and its neighbors in the graph \( \mathcal{N}(X_i) \). Segmentation and detection is defined as assigning a class label \( l_i \in \{1, \ldots, K\} \) to each

![Figure 11: a) Left stereo frame; b) superpixel segmentation; c) instruments detection and segmentation with a CRF; d) dense stereo disparity image from the stereo algorithm, lighter colors are closer to the camera.](image-url)

![Figure 12: Laser sintered manipulator with two instrument arms and an additional arm in between with slots for two micro cameras.](image-url)
random variable \( X_i \). The conditional density over a label assignment to all nodes \( \mathcal{L} \), is
\[
P(\mathcal{L}|\mathcal{I}; w) \propto \exp \left( \sum_{i} \Psi_u(X_i, l_i, \mathcal{I}; w) + \sum_{(i,j) \in E(\mathcal{G})} \Psi_p(X_i, X_j, l_i, l_j; w) \right)
\]
with \( \Psi_u \) a unitary potential encoding the particular image regions label preference, \( \Psi_p \) is pair wise potential enforcing similarity between adjacent image regions, and \( w \) model parameters encoding prior knowledge about the single-port scene. Our goal is to find a maximum a posteriori (MAP) label assignment for a given image and parameters \( \mathcal{L}^* = \arg \max_{\mathcal{L}} P(\mathcal{L}|\mathcal{I}; w) \). Direct MAP estimation in this form is intractable, instead approximate inference using mean-fields is applied. To further reduce the computational complexity, segmentation is not performed on a per pixel basis, but on super pixels, which are small image regions grouped together by local color similarity [23], as show in figure 11. The resulting segmentation is then fused with the dense stereo and SLAM state to localize the instruments in the surgical site.

VI. CONCLUSION

We propose a fully flexible robotic surgery system capable of meeting the needs of single-port and NOTES interventions. Our design features a completely modular concept. The mechanical design of the system makes use of novel designs involving a selective laser sintering process featuring fully flexible and bio-compatible structures. The actuation strategy ensures optimal use of the fully flexible structures while at the same time maintaining a sterile coupling between the drive system and the manipulator. The control concept of the system introduces strategies well suited for single-port and NOTES interventions including the design of an intuitive human machine interface for surgeons. Visual processing modalities specific to the surgical environment enable a more effective use of the surgical system.

ACKNOWLEDGMENT

The authors like to thank MicroLumen Inc. (Oldsmar, FL, USA) for providing the tubings, and the MiMed-workshop team for its continuing support, especially Gerhard Ribnitzky and Markus Wühr. This work was supported in part by the German Science Foundation (DFG) under Grant LU604/31-1.

REFERENCES


