
Christoph Staub, Takayuki Osa, Alois Knoll and Robert Bauernschmitt

Abstract—Despite the fact that minimally invasive robotic surgery provides many advantages for patients, such as reduced tissue trauma and shorter hospitalization, complex tasks (e.g. tissue piercing or knot-tying) are still time-consuming, error-prone and lead to quicker fatigue of the surgeon. Automating these recurrent tasks could greatly reduce total surgery time for patients and disburden the surgeon while he can focus on higher level challenges. This work tackles the problem of autonomous tissue piercing in robot-assisted laparoscopic surgery with a circular needle and general purpose surgical instruments. To command the instruments to an incision point, the surgeon utilizes a laser pointer to indicate the stitching area. A precise positioning of the needle is obtained by means of a switching visual servoing approach and the subsequent stitch is performed in a circular motion.

Index Terms—robot surgery, minimally invasive surgery, tissue piercing, visual servoing

I. INTRODUCTION

Minimally Invasive Surgery (MIS) has become commonplace for an ever-growing number of procedures. While patients profit from faster recovery time, less pain and tissue trauma through the application of small incisions or ports in the body compared to large cuts in conventional “open” surgery, surgeons have to cope with reduced dexterity and perception. To some extend awkward instruments with less degrees of freedom (DoF) and missing force feedback or flattened 2D visual feedback make delicate maneuvers such as knot-tying very time-consuming. With the introduction of robotic telemanipulators as the daVinci™ machine [1] minimally invasive instruments become more enhanced. In [2] we introduced an experimental system for minimally invasive surgery that incorporates force feedback. Nevertheless, one of the major issues of robot-assisted MIS is significantly reduced working speed, most noticeable during complex tasks which renders these type of operations stressful for both patients and surgeons.

Surgical suturing is probably one of the most difficult and time consuming tasks for surgeons in laparoscopic surgery [3]. Given that suturing occurs frequently during interventions, several research groups work on automating this task. Previous works have focused on miscellaneous aspects of suturing in robot-assisted minimally invasive surgery: Kang et al. [4] devised a specialized stitching device which is capable of creating a knot. Wakamatsu et al. [5] proposed a robotic setup for knot tying, but they did not evaluate the task under realistic surgical circumstances, yet (e.g. small-scale knots performed under trocar kinematics and original suture material). Hynes et al. [6] employed visual servoing in order to command the robot movements, but the desired points have to be selected manually in the images. In [7] and [8] we demonstrated an approach to tie a suture knot autonomously using general purpose laparoscopic instruments. Situated learning and scaffolding was realized by means of a skill transfer architecture, which is based on the abstract descriptions of tasks. Its instantiation is carried out by means of fluid dynamics. A research group of the John Hopkins Haptics Lab [9] proposed general assistance using guidance virtual fixtures to assist the surgeon to move towards a desired goal. The framework could be used to compute joint motions required for the stitching task. Nageotte et al. [10] introduced a method for computing optimal needle-paths with a kinematic analysis of the stitching task and a regular 4DoF needle-holder. An initial entry point and an exit point is selected manually and the corresponding instrument pose and piercing angles are calculated while tissue deformations can be predicted. In doing so, it is assumed that the needle already intersects with tissue. They approach visually-servoed path following in order to drive the instruments along the calculated 3D trajectory.

A major contribution to the success rate of automating tasks in unknown environments is given by the overall system calibration. Unfortunately, in many MIS systems a multitude of different error sources contribute to a comparatively bad accuracy. First of all, the absolute accuracy of typically used robots is limited to 2mm, depending on the individual robot. In addition, the exact location of the base point of the robots is not known exactly (particularly the coplanarity of the robots’ bases cannot be guaranteed). The major drawback, however, is due to the high elasticity of the carbon fiber instrument shafts. The possible play of the end effectors employed in our system is quiet large due to the carbon fiber material (±1.5cm) and contributes most to the overall error. Image-based visual servoing (IBVS) aims to overcome these problems, including mechanical play and intrinsic camera errors. Endoscopic cameras often suffer from inherent barrel distortions and are delicate to calibrate.

The focus of this paper is on automated tissue piercing. The surgeon may choose a piercing point using a laser pointer mounted on one of the surgical instruments. Afterwards, a switching visual servoing control is used to maneuver the
pinpoint of the needle precisely to the assigned incision point to perform the stitch. A circular movement is executed after the needle was correctly aligned.

II. MATERIALS AND METHODS

We first describe the hardware that was used for the experiments. Afterwards, we will describe the stitching task in detail. The crucial part will be the automatic guidance of the robot by means of visual servoing in a way that the needle aligns with the tissue surface and the selected incision point.

A. Robotic System

We have conducted our experiments with a system for robot-assisted minimally invasive surgery. Hardware and software of the system itself have already been introduced to the research community [7]. Therefore, we constrict the following description to an extent necessary for understanding the subsequent sections. As illustrated in Figure 1, the slave manipulator of the system consists of four ceiling-mounted robots which are attached to an aluminum gantry. The robots have six degrees of freedom and are equipped with either a 3D endoscopic stereo camera or with minimally invasive surgical instruments which are originally deployed by the daVinci™ system. The surgical instruments have 3DoF. A micro-gripper at the distal end of the shaft can be rotated and adapted if pitch and jaw angles is possible. Fast and easy interchangeability is ensured by a magnetic clutch which releases in case of forces at the instruments exceeding a certain limit (e.g. during a collision). To measure forces during operation the instruments are augmented with strain gauge sensors. The master-side manipulator is mainly composed of a 3D display, some foot switches for user interaction (like starting and stopping the system or executing the piercing process) and of the main in-/output devices, two PHANToM™ haptic displays. On one hand, the devices are used for 6DoF control of the slave manipulator and on the other hand provide a 3DoF force feedback derived from the measurements at the instruments. The control software of the system realizes trocar kinematics, whereby all instruments will move about a fixed fulcrum after insertion into the body.

B. Description of the Stitching Task

Due to a more complex handling of minimally invasive systems, an intrinsic issue of these procedures is that their completion takes significantly more time and suffers from an increased rate of errors. To disburden the surgeon during knot-tying, which takes up to a minute compared to seconds in open surgery, we proposed a method to automate a suture knot [8]. So far, the surgeon has to align the needle and to pierce the tissue by himself, before the autonomous part can be executed. The intension of this work is to extend the existing framework to automatically align a circular needle with the tissue surface and to perform a subsequent penetration.

Suturing can be subdivided into independent steps: First, the needle has to be brought into the human body and grasped by the surgical instrument. The next step comprises the selection of a suitable entry point of the needle on the surface of the tissue. The surgeon aligns the needle according to the mentally selected point and pierces the tissue in a circular movement. After the tip of the needle bites the tissue for a second time one can pull out the needle by grasping it close to the tip.

For this work, we assume that the robot is holding the needle perpendicular to the jaws of the forceps as illustrated in Figure 2(a). Up to now, this posture is initialized manually and allows the jaws to hold the needle as secure as possible, minimizing the risk of a slip during the piercing process. The surgeon uses a laser pointer to mark a cut-in point $I$. The laser pointer can be carried by any of the three available surgical instruments, while the forceps of the instrument can still be operated regularly. Piercing the tissue in a circular way is one of the most important factors to preserve tissue and minimize deformations. The ideal motion of the needle should follow the arc of the needle as long as the tip is inside the tissue to avoid tearing at the incision point. The automation of the stitching task means two reliefs to the surgeon. On one hand, he can focus on high level tasks while the system executes recurrent tasks autonomously. On the other hand, the visualization of the incision point with a laser provides a decision support in order to monitor the actions performed by the system.

Earlier experiments showed that the contact between the needle and the tissue interface cannot be predicted using force information derived at the instrument tip solely. The contact was recognized too late and the position of the needle
in the brackets was already dislocated. A circular tissue penetration got impossible. This is mainly because of the soft material property of tissue and the associated delayed sensor response. In the next section, we propose a switching visual servoing scheme which facilitates a precise positioning of the needle.

C. Feature Extraction

As a first step, image feature points have to be identified and tracked to visually servo a robot. Image processing is used to identify the tip of the circular needle. Previous works have tried to track the position and orientation of the entire surgical instrument. Several methods have been proposed, among them the detection of the shaft’s geometrical appearance in the image (see [11], [12]), tracking special orientation markers (e.g. gray encoders [6]) or using colored patches [13]. Also specialized devices equipped with LED markers and a projective laser to identify the tool orientation have been proposed [14]. Nageotte et al. [15] tried to localize an entire circular needle including position and orientation information. A comparatively simple method is applied to find the tip of the needle in our approach. The needles are segmented in HSV (Hue, Saturation, Value) space that allows for an extraction of the pinpoint of either green or white needles. The needles were colored ahead. In a similar way, the center of the laser projection is detected. Improving this simple image processing method is subject of current development. Our intension is to use a tracking algorithm that combines both color and geometrical information. The identified pinpoints are colored greenish in Fig. 3(a).

D. Robot Control

The procedure explained in section II-B described tissue piercing with a circular needle and general purpose laparoscopic instruments. Next, we show how the robot can be controlled to align the needle with the incision point \(I\) and pierces the tissue. The task can be decoupled into two subproblems: (a) alignment of the needle and (b) piercing of the tissue in a circular way. In doing so, the compliance of the trocar constraint has to be ensured during MIS.

Visual servoing is a popular approach to guide a robotic appendage (e.g. an instrument or a needle) using visual feedback from a camera system. A good survey is given by Chaumette et al. [16], [17]. In many cases of visually-servoed systems the movement of an object is expressed in a Cartesian reference frame. For this reason, many systems dealing with moving objects use position-based visual servoing. In general, the accuracy of image-based methods for static positioning is less sensitive to calibration than comparable position-based methods [18], [19]. A practical difficulty during the needle alignment with \(I\) lies, however, in the fact that the instruments are not necessarily in the field of view of the endoscopic camera. The instrument first has to be driven to an initial position which is already close to the target and observable by the camera. For this reason a switching approach is employed in our system. After PBVS the control scheme is changed to an image-based approach and the residual error can be minimized. As depicted in Figure II-B the final alignment of the needle is accomplished in three steps. At the beginning, the needle is not mandatory in the sight of the endoscope. Figure 2(a) shows the state after a rough alignment to the Cartesian coordinate of \(I\), achieved by PBVS. Afterwards the image-based positioning was performed and the needle already has bitten the tissue (Fig. 2(c)). After a circular motion the needle bites the tissue for a second time and the process is completed (Fig. 2(d)).

4587
1) Position-Based Servoing: the Cartesian coordinate of point \( I \) is calculated (remember, the system is equipped with a stereo endoscope) and the needle is placed within a certain error to it. The position of the needle is extracted in image coordinates and projected to the world frame by a mapping determined during camera calibration. The camera is treated as a 3D positioning sensor. The robot motion is then performed in an open-loop manner and no error function is optimized. Initializing the image-based servoing at a starting point which is already close to the target also shows positive effect on the convergence of the error during the image-based alignment. Now, the current position of the needle is slightly displaced by an error \( e \) (about 3 − 10 mm) from its final position. This error does not only result from the camera parameters, but is mainly introduced by the overall calibration of the kinematic chain mentioned in the introduction (Section I).

2) Image-Based Visual Servoing\(^1\): the positioning method introduced above allows for a positioning of the surgical instruments within a certain error in Cartesian space. So as to further minimize this error an image-based visual servoing approach is employed. Given a target pose or position that the robot is to reach, image-based visual servoing aims to minimize an error function \( e(t) \), defined by

\[
e(t) = s(t) - s^*.
\]

where \( s^* \) represents the target pose and \( s(t) \) characterizes the measured pose, in our case the tip of the tracked needle. The image-based control problem is to find a mapping from this error function to a commanded robot motion. In many IBVS scenarios the camera is directly attached to the robot which is to be controlled (eye-in-hand configuration) and therewith the velocity \( \xi \) of the camera is calculated. In our system setup, the instrument and the endoscope is carried by two different robots and the calculated velocity \( \xi \) has to be transformed to the robot which is equipped with the instrument (cp. 8). Only a single image feature point, the tip of the needle is tracked in both left and right camera images provided by the stereo endoscope. The vector \( s = (x_L, x_R)^T = (u_L, v_L, u_R, v_R)^T \) comprises these coordinates:

\[
s(t) = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}
\]

Its derivative \( \dot{s}(t) \) is referred to as image feature velocity. It is linearly related to the camera velocity \( \xi = [v \omega]^T \), which is composed of linear velocity \( v \) and angular velocity \( \omega \). The relationship between the time variation of the feature vector \( s \) and the velocity in Cartesian coordinates \( \xi \) is then established by

\[
\dot{s} = L_s \xi,
\]

where \( L \) is the interaction matrix or image Jacobian related to \( s \). The interaction matrix \( L_s \) related to a single image coordinate \( x = (u, v) \) is written as [16]:

\[
L_x = \begin{bmatrix}
-\frac{1}{2} & 0 & \frac{u}{2} & \frac{uv}{2} & -(1 + u^2) & v \\
0 & -\frac{1}{2} & -\frac{v}{2} & 1 + v^2 & -uv & -u
\end{bmatrix}
\]

Variable \( z \) represents the depth of the point relative to the camera frame. There exist different ways to approximate the value of \( z \), e.g. via triangulation in a stereo setup or via pose estimation. Most of the existing methods assume an calibration of the camera parameters, even if the impact of the calibration quality in IBVS does not affect the servoing process as much as in PBVS methods. Many systems even assume a constant depth and therewith a constant image Jacobian. Instead of a triangulation, \( z \) is estimated via the kinematic chain of the system. All necessary transformations have been established via calibration in advance. The interaction matrix is updated on-line and the method can also be transferred to different camera configurations. Using (1) and (3) we obtain the relationship between the camera velocity and the time variation of the error as \( \dot{e} = L_e \xi \) (where \( L_e = L_s \)) and the final control law is

\[
\xi = \lambda L_e^+ e
\]

where \( \lambda \) is a positive gain factor and \( L_e^+ \) the pseudo-inverse of \( L_e \).

As mentioned above, a single visual feature \( s \) is tracked in the left and right images and (3) can be rewritten as

\[
\begin{bmatrix}
\dot{x}_L \\
\dot{x}_R
\end{bmatrix} = \left[ \begin{array}{c}
L_L \\
L_R
\end{array} 
\right] R_L V \xi_L = L_a
\]

The spatial motion transform \( V \) to transform velocities expressed in the right camera frame \( R \) to the left camera frame \( L \) is given by

\[
V = \left[ \begin{array}{cc}
R_{LR} & S(t) R_{LR} \\
0 & R_{LR}
\end{array} \right]
\]

where \( S(t) \) is the skew symmetric matrix associated to the linear transformation vector \( t \) and where \( (R, t) \) is the transform from the left to the right camera frame. During the alignment with point \( I \) five degrees of freedom have to be controlled: the linear velocities \( u_x, v_y, v_z \) are derived from the visual servoing process and the angular velocities \( \omega_x, \omega_y \) are governed by the trochacantr (cp. Section II-D.3). The calculated velocities correspond to the camera robot, but the robot that is equipped with the surgical instrument is to be controlled. The velocities can be transformed by

\[
V = \left[ F_1 V \cdot F_1 V \cdot F_2 V \cdot F_2 V \cdot V \right]
\]

where \( I \) is the tip of the surgical instrument, \( R_1 \) the frame of the endoscope robot, \( R_2 \) the frame of the instrument robot and \( F_1, F_2 \) the corresponding flanges respectively.

The calculated velocities are governed by an adaptive gain factor \( \lambda \) which allows adaption of the motion speed. If the distance between the current and the desired position exceeds a certain value the speed is considered to be constant.
Whenever the distance between both points decreases, the movement is retarded:
\[
\lambda = \begin{cases} 
\frac{1}{\|T^e - T^0\|} \lambda_0 & , \|L^+ e < a\| \\
\lambda_0 & , \|L^+ e > a\| 
\end{cases}
\] (9)

Variable \(\lambda_0\) is a constant coefficient and \(a\) is the threshold.

3) Trocar Kinematics: in MIS it is indispensable to fulfill the trocar constraint and to ensure no endue movement at the incision point. The Cartesian coordinate of the fulcrum is well-known in our system setup and is considered to be fixed. Constraining 6DoF of the robot to 4DoF at the trocar \(T\), all motions perpendicular to the instrument shaft in \(T\) are interdicted and 2 degrees of freedom are determined by the port. The velocities at the trocar point \(T^I \xi = (\overrightarrow{T^I v}, \overrightarrow{T^I \omega})^T\) and the velocities of the instruments tip \(I^V = (\overrightarrow{I^V v}, \overrightarrow{I^V \omega})^T\) are related as follows:
\[
\begin{bmatrix}
\overrightarrow{T^I v} \\
\overrightarrow{T^I \omega}
\end{bmatrix} = \begin{bmatrix}
T^I R & S(T^I t) \cdot T^I R \\
0 & T^I R
\end{bmatrix} \begin{bmatrix}
\overrightarrow{I^V v} \\
\overrightarrow{I^V \omega}
\end{bmatrix}
\] (10)

Assuming a straight shaft, \(T^I R\) is the identity matrix and \(t = (0, 0, d)^T\) with \(d\) being the insertion depth of the instrument. Since only the \(z\)-direction (the direction of the shaft) is free to move, the linear velocity at the insertion point is donated by \(I^v = (0, 0, I^v z)^T\). Solving (10) yields to
\[
I^v z = -\frac{I^v y}{d} \quad \text{and} \quad I^v y = \frac{I^v z}{d}
\] (11)

E. Stitching

So far, all conditions to perform the needle stitch have been considered. The stitching task itself is not conducted under visual guidance but rather commanded by a generated trajectory. Under the constraint of trocar kinematics the lost degrees of freedom are regained by the control of the surgical instruments (e.g. the instruments have a wrist at the distal end). In fact, the simulation and control software of the system allows us to maneuver an instrument in Cartesian space and both the control sequences for the servo motors which are incorporated into the instrument and the robot commands are generated simultaneously. The corresponding coordinate system related to an instrument is depicted in Fig. 3(b). As already stated above, the needle is inserted into the gripper manually, perpendicular to the \(x\)-axis of the instrument frame. As the posture is known now, it can also be oriented perpendicular to the surface of the tissue. As the micro-grippers have a rough surface the needle does not move anymore after it is inserted into the gripper. If this pose cannot be maintained, slippage during the stitch is possible and no circular movement can be guaranteed. To perform the stitch itself, a rotation that is displaced by the radius of the needle is performed around the \(x\)-axis of the instrument.

III. Experimental Results

In order to demonstrate the ability of the system to perform high precision positioning tasks and to perform a circular stitch, several experiments have been conducted (see video attachment). A calibration trihedron was mounted on a robot and was driven to 3D Cartesian coordinates extracted from the stereo system to assess the accuracy of the PBVS. The achieved precision is (as expected) comparatively low. The error ranged between 3 and 10mm and is basically influenced by the issues mentioned in the introduction (aberration of the instruments, magnetic clutch, etc.). Nevertheless, the accuracy is adequate to initialize the instruments in the field of view of the camera. The image-based part of the control scheme was verified with a simple positioning task. Different positions that are uniformly distributed over the entire image were marked. Using the same initial position for all points, the pinpoint was driven subsequently to all labels. The distance between the target position and the final position of the needle did usually not exceed 1mm after the IBVS was completed. The convergence of the error in image coordinates is depicted in Fig. 5(a) and Fig. 5(b) shows the characteristics of the generated trajectory (from simulations and from real data). The actual piercing process was performed with both phantom tissue and a piece of meat. While the laser projection was well visible on the phantom tissue, the surface of the meat caused some diffusions. Besides, the texture of the meat turned out to be too soft...
and the needle could not get enough grip to cut the flesh during some experiments. A force controlled regularization of the penetration depth of the needle into the tissue could be a solution.

IV. CONCLUSION AND PERSPECTIVE

An experimental system for robotic-assisted minimally invasive surgery has been presented which is used for automated tissue piercing. The incision point was identified by an laser projection. A switching visual servoing scheme was used to precisely position a needle on the surface of the tissue. The position-based part of the approach was able to maneuver the instruments to the target position within a certain error, mainly caused by calibration imprecisions. After the change to image-based control, the remaining error was further minimized to an acceptable value of about $1\text{mm}$ in average. We assumed an obstacle-free environment, but for the application of autonomous executed surgical skills the recognition of the working space is crucial. Considering the presented approach, image processing and tracking is under ongoing work and has to be improved against noise and the loss of image features. This also affects the tracking of the tissue surface that may deform. Modeling deformations during incision processes can be realized i.e. by means of XFEM simulation [21]. Also the utilization of better servomotors for the instruments would probably enhance the execution of the circular motion. A visual pose estimation of the needle could counteract a slip in the holder. Also data fusion with force information gathered at the tip of the instruments is in progress to stop the piercing process if the measured forces exceed a certain limit to prevent tissue damage.

V. ACKNOWLEDGMENTS

This work is supported by the German Research Foundation (DFG) within the Collaborative Research Center SFB 453 on “High-Fidelity Telepresence and Teleaction”.

REFERENCES


