Proceedings of the 2010 3rd IEEE RAS & EMBS
International Conference on Biomedical Robotics and Biomechatronics,
The University of Tokyo, Tokyo, Japan, September 26-29, 2010

An Approach to Robotic Guidance of an Uncalibrated Endoscope in Beating Heart Surgery

Aleksandra Popovic and Paul Thienphrapa

***Abstract*—Minimally invasive cardiac surgery is performed under direct vision provided by an endoscope. During cardiac bypass surgery, counterintuitive control and small field-of-view endoscopic visualization is not sufficient to localize all relevant structures of the heart.**

This work presents a first step towards enhanced endoscopic visualization in cardiac surgery by allowing image based steering of the endoscope without the need of camera calibration. In the envisioned clinical scenario, a surgeon or an endoscopy assistant selects a point or an area in the live endoscope video stream and the robotic system steers the endoscope towards the selected target on a beating heart.

Experiments in a simulated setup have shown that the algorithm successfully performs various feature tracking tasks within the requirements. The results from these experiment led to an improvement of the algorithm to enhance the final stage of feature capturing. With the lessons learned in the simulated setup, next series of experiments have been performed in a clinical-like setting on a beating heart phantom. The experiments show that all reasonable target features can be reached within 1.2 s and 7 pixels of accuracy at a safe and conservative robot speed.

I. Introduction

A. Motivation

Coronary arteries are the major supplier of blood to the heart muscle. Coronary artery disease (CAD) is a result of plaque buildup on an artery wall causing narrowing of the artery and disruption of the free blood flow. CAD caused one out of every five deaths in the United States in 2005 [1]. Coronary artery bypass graft (CABG) surgery is a procedure done to bypass atherosclerotic narrowing in coronary arteries relieving symptoms of the CAD. As of today, CABG remains the most effective therapy for CAD [2]–[4].

Introduction of surgical endoscopes has revolutionized (cardiac) surgery allowing minimally invasive approaches under a direct vision. In many cases, reduced trauma and improved recovery come at the cost of prolonged surgery due to difficulty in handling of long tools fixed at a fulcrum point and unnatural requirements for the hand-eye coordination.

Apart from widely accepted advantage of minimally invasive surgery – reduced trauma to the tissue, another objective in minimally invasive heart surgery is to eliminate the need for Cardiopulmonary Bypass (CPB).The CPB is associated with significant morbidity and mortality through complications such as hemolysis, air embolism, and clotting [5].

The two most reported technical difficulties during minimally invasive procedures are difficult instrument handling during anastomosis on moving tissue and inadequate visualization [6]–[9]. Several solutions have been proposed to tackle the instrumentation challenge, which include of robotic systems (*da Vinci*®, Intuitive Surgical Inc.), mechanical heart stabilizers, and automatic anastomosis tools (e.g. PAS-port®, Cardica Inc.). This work therefore focuses on the visualization issues.

It is widely accepted in the clinical community that small field-of-view counterintuitive endoscopic visualization is not sufficient to localize all relevant structures of the heart during a cardiac bypass surgery [6], [7].

The counterintuitive operation is caused by an unknown mapping between the surgeon's hands and the endoscope view [10]. This mapping is learned during the surgery. Reports show that an intraoperative change in camera position or orientation may cause significant increases in time needed to learn the mapping [10], [11] subsequently increasing duration of the surgery. In CABG, the endoscope is usually inserted into the chest cavity from the right side of the patient or from the posterior-anterior direction. The instruments are inserted from the left side of the patient. The surgeon is standing left from the patient [12]. As an additional difficulty, control of the endoscope and surgical tools is usually distributed over at least two people – a surgeon holding the instruments and an assisting surgeon holding the endoscope. Therefore, there are three relevant coordinate systems: the coordinate system of the endoscopic view, the coordinate system of the surgeon, and the coordinate system of the assistant.

The small field-of-view of standard endoscopes used in cardiac surgery is usually inadequate to visualize important landmarks on the heart [13] while at the same time providing poor image resolution [6]. A usual trade-off is to keep the endoscope as close as possible to the target area to improve spatial resolution. The assisting surgeon performs free-hand endoscope manipulation to detect the position of the current field-of-view relative to prominent anatomic landmarks such as the apex of the heart, the groove between the medial aspect of the left atrial appendage and pulmonary artery, and the ventricular septum [8]. As these landmarks are identified, the location of the target artery can be predicted and the endoscope can be steered towards it. This process is hindered by the above described hand-eye coordination problems.

Visualization issues are reported as one of the most common cause of conversions from minimally invasive to open surgery [8], [9].

This work presents a first step towards enhanced endoscopic visualization in cardiac surgery by allowing image based steering of the endoscope bypassing uncertainties arising from counterintuitive intraoperative setup. In the envisioned clinical scenario, a surgeon or an endoscopy assistant selects a point

A. Popovic is with Philips Research North America, Briarcliff Manor, NY, USA

P. Thienphrapa is with Philips Research North America, Briarcliff Manor, NY, USA and the Laboratory for Computational Sensing and Robotics, Johns Hopkins University, Baltimore, MD, USA

978-1-4244-7709-8/10/$26.00 ©2010 IEEE

or an area in the live endoscope video stream and the robotic system steers the endoscope so that the selected area appears in the middle of the endoscope image. The system is aimed to allow a simpler manipulation of the endoscope using a direct mapping from the endoscope view as seen by the surgeon and the endoscope. This will bypass the difficult mapping required between the view, the endoscope, the surgeon, and the assistant controlling the endoscope and as a result, a solo surgery can be performed. This will also allow the surgeon to explore the heart surface in a more structured fashion.

The platform presented here will serve in the future as a basis for image-based steering using other imaging modalities, such as intraoperative ultrasound (3D Transesophageal Echocardiogram – TEE) or preoperative Computed Tomography (CT).

B. Background and system requirements

This work employs visual servoing, a robot control method using visual input to set robot movement parameters.

In image-based approaches the control loop of visual servoing is closed through a transformation between the image plane and the robot joint velocities established through a calibration process [14], [15]. The transformation matrix is referred to as *Image Jacobian* or *Interaction Matrix*. An alternative to image-based methods are position-based methods performing explicit 3D reconstruction of feature points from the image [16]. The performance of both methods is highly sensitive to the accuracy of the intrinsic parameters of the camera, e.g. focal length(s) and image optical center.

Clinical environments are not well-suited for accurate camera calibration which is a tedious, error-prone process. The optical system may show a significant drift in its intrinsic characteristics due to heat, humidity, and other environmental disturbances. Extrinsic camera parameters, i.e. transformation between the camera and robot joints can be made reproducible through a unique mounting of the endoscope shaft. Finally, the robot forward kinematics represented through a joint Jacobian – joint velocities as a function of end-effector velocity – is significantly less sensitive to interferences in the operating room, though, a drift in the joint calibration can be induced through incautious handling of the system.

Considering the clinical and technical insights described above, the following requirements for image guided robotic endoscopy system were defined:

- The system should impose a fixed fulcrum point to avoid injuries to the patient's skin and underlying tissue.
- The system should not require a camera calibration. It should work with any kind of endoscope and allow intraoperative replacement of the working endoscope.

Different approaches for solving visual servoing in medical robotics and endoscopic surgery have been proposed. Zhang et al. [17] propose a method involving camera calibration and tool tracking in the endoscope view based on three optical markers. The optical markers can provide depth information given the known distance and thickness. Thus, a full Image Jacobian can be computed. In another system, proposed by Krupa et al. [18], a laser pointer with four laser dots is attached to the endoscope to provide depth information relative to the organ. The *Image Jacobian* is obtained similarly as in Zhang et al. using three markers on the tool. However, the use of artificial markers on the heart surface may not be clinically acceptable. It would require the surgeon to place the markers using endoscopic instruments which might be a time consuming process. Therefore, the time expected to be saved in the learning of mapping might be partially or fully lost due to the time required for the placement of the markers. Also, they might cover some important structures, for example arteries hidden by a layer of fibrous-fatty tissue.

Hynes et al. [19] propose a method for visual servoing without the need for camera calibration used in minimally invasive surgery based on the previous work by Jagersand et al. [20]. This method performs an on-line estimation of *Image Jacobian* using a correction formula. The control loop starts with an initial estimate of the matrix and corrects the values using motion observed by a camera decoupled from the robot. The method requires a random initial estimate of *Image Jacobian*, which can lead to unpredictable erroneous movements [20] unsuitable in sensitive surgical settings.

Hirzinger et al. [21] propose a proportional control scheme for real-time visual servoing of the AESOP robot for laparoscopic surgery. The robot is moved proportionally to the distance between the tracked point and the target location. No optimization of the tracking speed is proposed which might induce oscillations in periodically moving organs such as the heart.

In this work, a practical method for visual servoing with an uncalibrated endoscope is presented. The method simplifies preoperative workflow by avoiding camera calibration and removing uncertainties that might arise from a suboptimal calibration. No artificial markers attached to the heart are used in order to make the proposed set-up more clinically acceptable.

II. Visual servoing

A. Kinematic considerations

The robot used in this system is the IBM/JHU LARS robot (see Acknowledgment). The robot is a 7-DOF (Degree-of-Freedom) arm with three translational DOFs (X,Y,Z), three rotational DOFs (θ, ϕ, ψ) and an insertion DOF along the endoscope shaft. The particular kinematics of the distal 4 DOFs, constructed in a parallelogram-like structure imposes a hardware restriction to possible movements of the mounted end-effector. This structure is referred to as orthogonally decoupled remote center-of-motion (RCM) [22]. The hardware RCM can be positioned so that the motion center coincides with the point where the endoscope enters the patients body (fulcrum) avoiding potentially traumatic injuries to the tissue and ribs. If a hardware RCM is not at the fulcrum, a virtual RCM can be implemented in the kinematic model of the robot.

The RCM is linearly translated as the endoscope is introduced deeper in the body. This motion is analogous to zooming and we consider it convenient for it to be controlled by the user rather than the robot for safety reasons. Considering the

imposed constraints described, the two end-effector degrees-of-freedom (θ and ϕ) are controlled by six robotic joints. The roll of the endoscope is also controlled by the user.

B. Velocity control

The Jacobian J of the LARS robot has been calibrated once and is used as it is. The forward and inverse kinematics, along with the optimization algorithm for the virtual RCM are described in [23]. Implementation available from the *cisst* [24] software package was used.

The inverse kinematics problem of the LARS robot is solved using a constrained optimization framework, with the joint positions acting as variables to be optimized for a given end-effector configuration. The virtual RCM is modeled as a virtual fixture and implemented as a constraint to the optimization problem.

The relation between a velocity in robot configuration space $v_r = [v_\theta, v_\phi]^T$ and a point velocity in the image coordinate system $v_i = [v_x, v_y]^T$ can be expressed[1] as:

$$\begin{bmatrix} v_\theta \\ v_\phi \end{bmatrix} = \begin{bmatrix} \frac{\frac{b_x^2}{x^2}}{\sqrt{1-\frac{b_x^2}{x^2}}} & 0 \\ 0 & -\frac{\frac{b_y^2}{y^2}}{\sqrt{1-\frac{b_y^2}{y^2}}} \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix}, b_x = \frac{d}{lf_x}, b_y = \frac{d}{lf_y} \tag{1}$$

where l is the length of endoscope from the fulcrum point to the tip, d is distance from the endoscope tip to the heart surface, and f_x and f_y are focal lengths of the endoscopic system in x and y axis respectively. None of these values are known without the calibration.

We propose an optimization technique to control v_θ and v_ϕ from v_x and v_y without explicitly measuring or estimating the *Image Jacobian*.

The first step in the optimization scheme is to map the spherical cap parametrized with θ and ϕ to a 2D plane using a spherical projection. This projection is 1-1 un-scaled warping. The resulting rectangle represents the full configuration space of the end-effector (endoscope) with respect to two parameters of the end-effector. In the next step, the 2D rectangle of point velocity (image) can be mapped to the 2D rectangle of the configuration space.

The first assumption introduced is a linear mapping between angular velocities of the robot and point displacement in the image. The motivation for mapping the position but not velocity resides in the fact that the expected goal of visual servoing is to track a user-selected feature point so that, after the robot task is completed, the point is in the geometrical center of image (reference point).

The velocity setting loop is shown in Figure 1. As the feature point moves away from the center of the image it produces a 2D velocity vector ($\vec{v}$). Since the scale between the robot and the image is not known, vector $\vec{v}$ can be normalized and transferred to the θ-ϕ plane resulting in vector

[1]Assuming that the optical and geometric centers of the image are co-located.

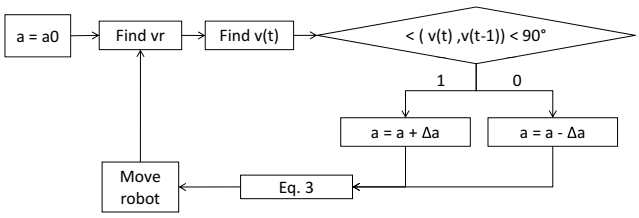

Fig. 1. Velocity setting loop

$\vec{v_r}$. The projected values on θ and ϕ axes define a unit robot displacement in a given direction ($\vec{v_r}$). Since $\vec{v_r}$ is normalized, velocities in θ and ϕ direction would be ≤ 1 rad/s. Depending on other unknown parameters (depth, focal length, speed of the point) that speed would usually be either insufficient to reach the point, or too fast, so it would overshoot the point. Therefore, a scale has to be adopted to account for these uncertainties.

Adaptation of the mapping scale can be considered an optimization problem with the objective to control the robot at a high velocity far from the feature to assure high tracking speed and to control the robot at a low velocity in the proximity of the feature to assure stability and prohibit overshoot. Four numerical components have been used for solving this optimization problem:

- A proportional factor representing Euclidean distance of the feature point to the center of the image (p).
- A PD controller implemented with two numerical factors K_p and K_d.
- A dynamic scale factor a.

These factors are combined in a control loop as:

$$[v_\theta(t), v_\phi(t)]^T = a \cdot (K_p \cdot p(t) \cdot \vec{v_r}(t) + K_d \cdot \frac{dp(t)}{dt} \cdot \vec{v_r}(t)), \tag{2}$$

where $[v(t)_\theta, v(t)_\phi]^T$ is a velocity vector at time t. Therefore, velocities vectors, assuming a unit of time, is:

$$\begin{bmatrix} v_\theta(t) \\ v_\phi(t) \end{bmatrix} = a \begin{bmatrix} K_p \cdot p_x(t) + K_d \cdot (p_x(t) - p_x(t-1)) \\ K_p \cdot p_y(t) + K_d \cdot (p_y(t) - p_y(t-1)). \end{bmatrix} \tag{3}$$

The scale factor a is a dynamic optimization factor used to facilitate learning of the current setup and compensate for changes in the setup and target velocity. The value is recomputed in each optimization step, i.e. in each frame of the endoscopy stream. The following section describes the setting of the scale factor in more detail.

C. Scale factor

This section describes the heuristics used to set the scale factor for the control loop (Fig. 1). As the robot moves along a defined direction, the position of the tracked point is updated due to two reasons: 1) the point is moving in the endoscope coordinate system due to the endoscope movement, 2) the point is moving in the world coordinate system (e.g. through the beating of the heart or breathing) and thus in the endoscope coordinate system as well. A new position of the point updated at every frame is evaluated against an accuracy threshold.

If the point displacement is smaller than a given threshold, the robot is not moved and the loop continues waiting for larger displacement (this helps maintain a stable image in the presence of noise or tremor). If the point has moved between two given frames, the new unit direction vector is evaluated against the previous one. There are two possible scenarios: the new vector is in the same general direction as the previous one (the vectors are in the same quadrants) or the new vector is pointing to a new direction. In the first case, it can be assumed that the robot is moving in the correct direction and that the robot did not overshoot the point. Thus, the direction is updated, the value of a increased by predefined amount Δa, and tracking continued. If the direction is changed by more than $\pi/2$ it is assumed that the endoscope over-shot the target and a is decreased by Δa to facilitate convergence. This method performs well in presence of feature detection noise, because the direction is evaluated against a reference point (e.g. the geometrical center of the image) and it does not change if the feature point is oscillating.

An advantage of this control loop is that the robot is learning an optimal speed factor (a) in real-time.

An additional constraint in the system is a maximum speed set in each axis for safety reasons ($v_\theta^{MAX} = v_\phi^{MAX} = v^{MAX}$).

III. Experiments on a virtual phantom

This section describes a series of experiments on a virtual phantom to demonstrate validity of the proposed approach.

A. Setup

A rigid forward-view monocular endoscope (Richard Wolf Inc.) with 170mm length, 10 mm diameter and 1/32mm focal length, attached to a 30 Hz CCD (Charge-Coupled Device) camera operating at NTSC (National Television System Committee) image resolution is mounted on the LARS robot's end-effector. A RCM point is set at 100mm off the tip of the endoscope. The inertia of the robot was measured at 0.250s (represent the latency of motion after a non-zero velocity is set).

A red laser dot was projected onto a blue screen used as a background. The dot is segmented using an RGB thresholding algorithm and the dot center is computed from principal shape moments of the dot contour. The laser dot serves as a measurement reference. The robot can be guided so that the laser dot appears to move towards virtually defined targets. This particular imaging setup was used to decouple the effects of image processing computations from the performance of the robot control algorithms.

Data sets of virtual targets moving on a flat surface have been used as input to the visual servoing system. This approach allows the use of a flat blue screen with virtual functions implementing a 3D-like movement. For example, 1.5 periods of a sinusoidal speed profile on a flat surface are analogous to a constant velocity movement of a point along a hemisphere. More details will be given in later sections of the text.

This setup allows for flexible testing in a simple environment without oversimplifying the problem under consideration.

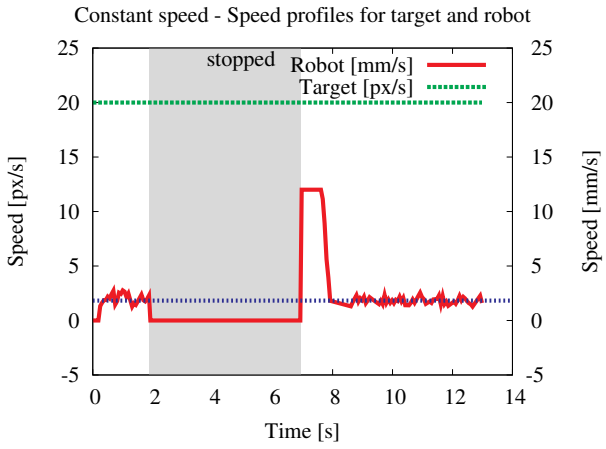

Fig. 2. Result of the experiment of tracking a target point traveling at a constant speed along a straight line

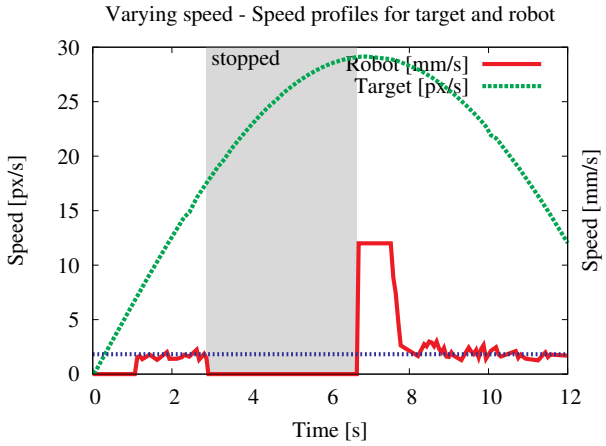

Fig. 3. Result of the experiment of tracking of a target point traveling at a varying speeds along a straight line

B. Experiments

1) Straight line: This set of experiments is performed to evaluate tracking of a point moving along a straight line at different speed patterns.

The empirical values for this set of experiments are set to $a = 1$ with increment/decrement step of $\Delta a = \pm 0.1$, $K_p = 0.1$, $K_d = 0$. The maximal speed of the robot is set at 0.2 rad/s. The system was operated at 30Hz.

The first experiments are done for a constant speed. The experiment has two segments. In the first segment, the robot is tracking a constantly moving point. In the second segment, the robot is stopped and the target continues moving away. After 5 s, the robot is restarted (Fig. 2).

After issuing the restart command, movement of the robot is observed after a period of about 0.3 s. That is in accordance with the latency of the system. The robot reaches the target after approximately 1 s, overshoots it (Fig. 2), and stabilizes at approximately 5 px from the target. The speed stabilizes at approximately $\pm$ 1 mm/s around the expected speed. This can be filtered-out using two methods: 1) selecting a larger

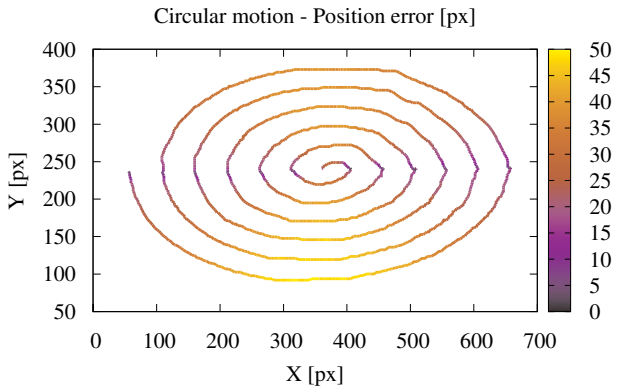

Fig. 4. Position of the target for the circular motion experiment (X-Y axes) and position error of the robot with respect to the target (color coding).

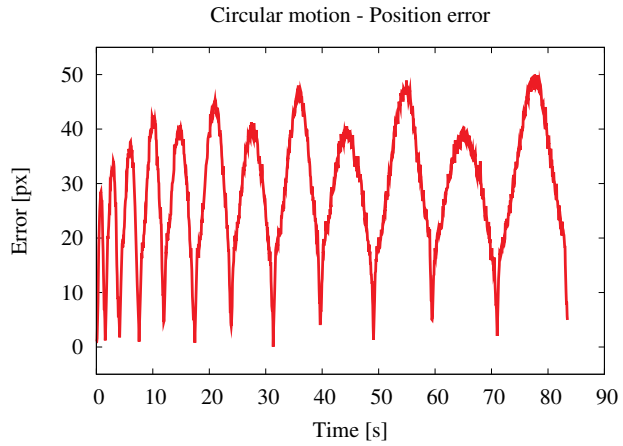

Fig. 5. Position error from Fig. 4 as a fuction of time.

accuracy circle, and 2) imposing the minimal speed to prevent the robot from slowing down too much in proximity of the target. These insights will be used while setting-up the heart phantom experiment.

The second experiment is done for a varying target speed. The same motion pattern was used. Resulting profiles are similar to those of the previous experiment (Fig. 3).

2) Circular motion: In the second set of experiments the target is moving on a circle with a constant angular velocity (Fig. 4). Referring back to the algorithm implementation, two axes are controlled separately. Therefore, this movement is analogous to a combined movement on a damped sinusoidal path with a damped speed (Fig. 5).

The first apparent irregularity that can be observed is the fact that errors are more prominent along the x-axis. That axis is dominant because that is where the main motion is occurring, due to asymmetry of the spiral. A second important insight arises from the periodicity of error in time. The period of the error signal is changing indicating that catching-up with the varying speed of the feature is more critical than its change in direction. This is an important realization that will lead to

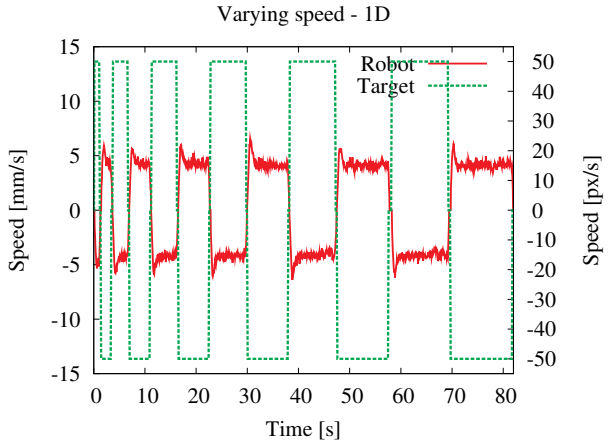

Fig. 6. Speed response of the robotic system for a pulse speed profile of the target point.

improvement of the algorithm, described later. Finally, top-bottom asymmetry can be explained though asymmetry in the spiral (the bottom part is longer) so when the point gets to that part, it is the first time the algorithm is encountering that speed and it requires more time to compensate.

3) Profile motion: The final experiments are done for a pulse speed profile (Fig. 6) of a changing period. This test was done to evaluate for abrupt changes in direction. The results show stabilization after 1s and a stable speed at $\pm$1mm/s.

C. Insights obtained from virtual target experiments

The following insights are used to design the experimental setup for the beating heart phantom.

1) The results demonstrate a stable operation in the presence of both varying speed and changing direction in linear and non-linear target trajectories.
2) The change in target speed appears to be more prone to errors in the robotic position than the target's change in direction. Since direction setting is more responsive, oscillatory motion with direction setting might be acceptable.
3) Very rare occasions of a large overshoot indicate that $K_p = 0.1$ might be too conservative.
4) To improve stability, minimal speed should be introduced in a similar manner as maximal speed.
5) The algorithm tends to lag behind the target.

The most important change introduced as a consequence of these insights is improvement in the target catching phase. As explained before, as the distance between the robot point and the target point is decreasing, the speed decreases accordingly. An improvement is introduced in the last segment. When p drops below a threshold, K_d is set to zero and a is put in the increase mode. Therefore, in the last few frames, the robot is speeding up again to catch the point. Insight 2 is in favor of this improvement, since it relaxes the restriction to not overshoot.

IV. Experiments on a Beating Heart Phantom

In the envisioned clinical scenario, the commands for defining the tracked feature or point are issued on a screen showing

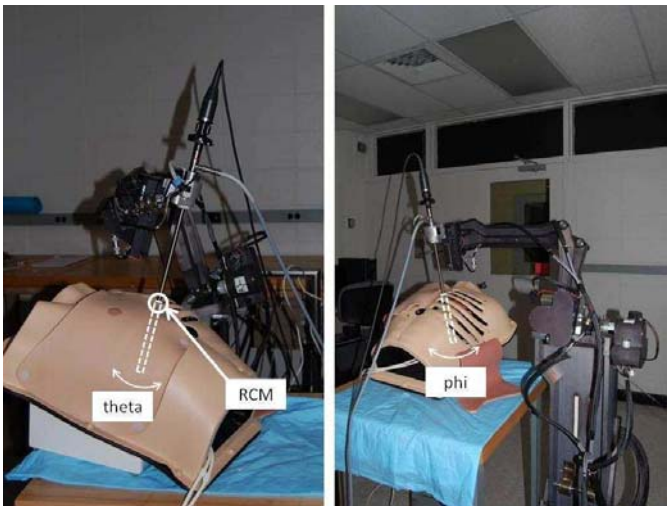

Fig. 7. The setup of the robotic system for the phantom experiments showing coordinate system and permitted motion around the RCM.

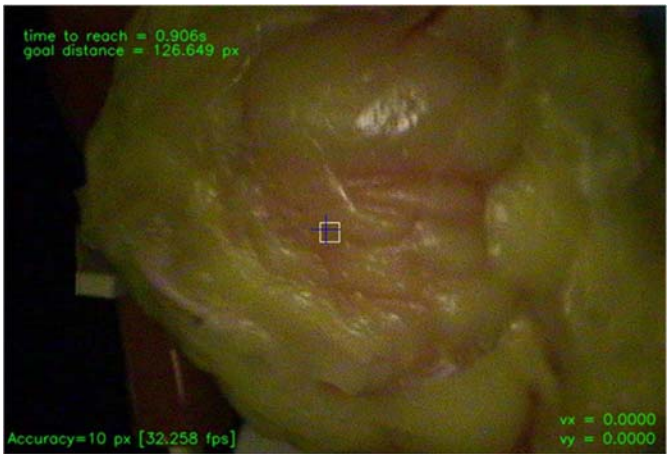

Fig. 8. Screen-shot of graphical user interface for feature (crosshair) tracking showing accuracy window (the square in the middle of the image.)

the endoscopy stream, either by using a mouse, a touch screen or any other input device. The goal is to simplify handling of an endoscope by removing hand-eye coordination from the work-flow.

A. Feature tracking

A video processing algorithm for robust automatic tracking of image features in an endoscopy stream has been used [25]. The algorithm implements a two step process, including pyramidal Lucas-Kanade optical flow followed by adaptive tracking modified to address both periodic motion of the heart and robot motion relative to the heart. More details on the algorithm can be found in [25]. The algorithm tracks a feature point at the native frame-rate of the endoscope system (30 fps).

B. Setup

The same endoscope used in the previous experiment is mounted on the LARS robot. The distance between the endoscope tip and the phantom surface was varied around distances usually observed in CABG surgery (30-50 mm), focusing on the area around the Left Arterial Descending (LAD) artery which is a typical target for the CABG surgery.

An additional master-slave control of the robot was implemented using a 3D mouse input device (Space Navigator™, 3Dconnexion) mapping all 7 DOFs. This allowed direct manipulation of the endoscope between the experiments. The mapping is provided by the *cisst* software library.

The Chamberlain Group Robotic Beating Heart Trainer model with a highly detailed exterior and movement of a live human heart has been used. It was operated at 60 beats per minute (bpm).

Fig. 7 shows the system setup.

A graphical user interface is implemented to allow selection of feature points by the user (Fig. 8). It shows the video stream at the native frame rate with information about current speeds in two axes, distance from the target to the center of the image, and the time needed to reach the target overlaid onto the video frames. A square around the center of the video stream shows the accuracy window. The selected feature is shown with a crosshair.

Learning from the virtual phantom experiments improvements have been implemented, and algorithm parameters are set as shown in Table I.

Parameter	value
a	1.0
Δa	0.1
K_p	0.2 $\frac{\text{mm}}{\text{s}\cdot\text{px}^2}$
K_d	0.1 $\frac{\text{mm}}{\text{s}\cdot\text{px}^2}$
Accuracy area	7 px
v^{MAX}	0.3 rad/s
v^{MIN}	0.001 rad/s

TABLE I
PARAMETER VALUES FOR BEATING HEART PHANTOM EXPERIMENTS

C. Results

A total of 23 experimental trials were performed. In each trial, an anatomical landmark was selected manually on the heart surface using the video stream window interface. The selection of a target automatically triggers visual servoing. For each trial, time-to-reach and distance from the geometrical center of video to the feature were recorded. Time-to-reach is defined as a period between initial feature selection and the moment in which a feature is stabilized within the accuracy window. In order to avoid rating an overshoot as a hit, a single trial was declared a success only if the system stabilized within that heart cycle. If a feature is lost, the trial was considered not valid. Out of 23 runs, 20 were marked as a success (87%). All three failures were due to instability in the image processing algorithm. In all three failure cases, retriggering of the task by selecting the same feature led to a success.

Fig. 9 shows the measured time-to-reach as a function of the initial distance from the feature. A line fitted using least square fitting has an offset of 0.238s which is consistent with previously measured latency of the robot at 0.250s. Indicative

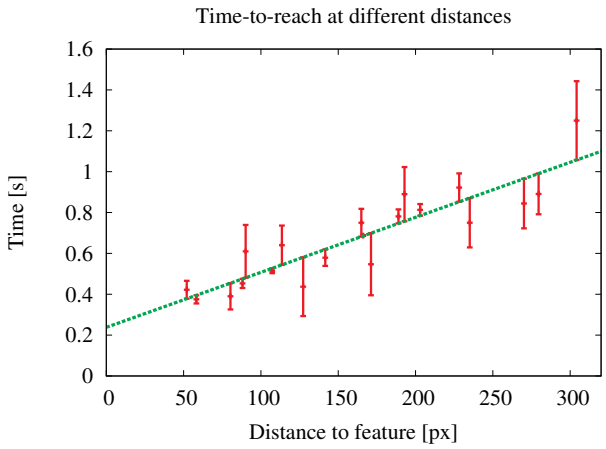

Fig. 9. Time-to-reach a feature as a function of distance. The line represents the least square fitting of the times measured in the experiments. Error bars are showing fitting error in each observed point.

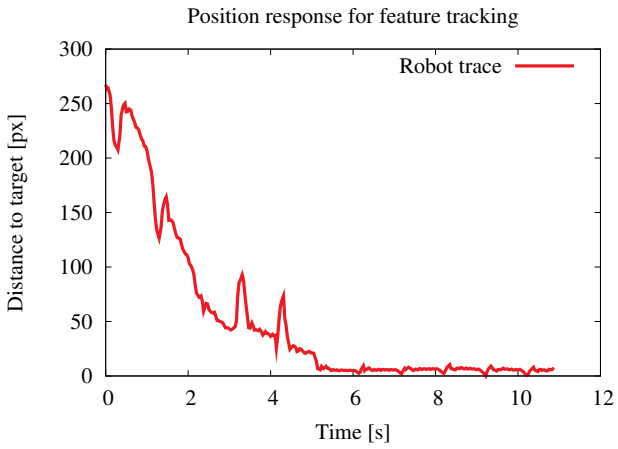

Fig. 10. Distance from the robot to the tracked feature in one run

of the system speed is the slope which is 385 px/s. In contrast to classical visual servoing where a full calibration and depth are known, this algorithm performs, by design, in exactly the same way independent of feature depth. Therefore, these values are valid for any clinically relevant physical setup. A boundary condition where this algorithm might break is in proximity to the physical limitations of the robot. Given that this robot is controlled at 10% of its nominal maximum speed, any reasonable surgical setup would be well below these conditions.

The tracking data demonstrates that all tested features were reached within 1.2s (Fig. 9). A strong fitting error at the end of the data range indicates that the imposed maximum speed limits the performance at longer distances. Since time-to-reach is in the approximate range of the heart cycle, the moment in which the feature has been selected plays an important role in the robot tracking speed.

In a second series of experiments, the maximal speed is reduced to 2 mm/s and the heart phantom is controlled at 60 bpm. The main motivation for this experiment is to observe the response of the algorithm in presence of a periodic noise. Lower speeds might also be more beneficial for the safety of the procedure as collisions at 2mm/s have a reduced potential to harm the patient and operating room staff. Fig. 10 shows the distance from the feature as a function of time for a single run in this setup. As expected, time-to-reach is significantly higher, at about 6s. In the first segment, as the robot is moving at a high speed, the robot is fast enough to compensate for heartbeats observed as a relatively smooth transition (0-2s). The smooth tracking breaks down as the speed of the robot is not enough to compensate ($<$ 2s). An interesting side-effect is observed. The algorithm shows a tendency to follow a a straight line path (thus regular periods in the trace $>$ 3s). For these set of experiments, it means that the feature will be caught at approximately the same segment of the heart cycle. Finally, in the last segment ($>$ 6s), a significantly smaller amplitude of the period signal is observed. The system is attempting to stabilize the feature at a point on the screen resulting in a damped signal.

V. Discussion

A practical approach to visual servoing with an uncalibrated image sensor and a virtual RCM is presented. Robot control using an image-based velocity compensation loop allows guidance without knowing the intrinsic parameters of the camera, relative distance from the robot to the tracked point, or absolute distance the tracked point has traveled. The endoscope can be mounted at any depth relative to the robot adapter. As a result, the closed loop approach reduces accuracy requirements of both robot kinematics and transformation between the endoscope and the robot. The system can accommodate to clinically relevant intraoperative changes in the setup – adjustment of the endoscope zoom, exchange of the endoscope (often done in CABG surgery to switch between oblique and forward viewing endoscopes [8]), and modifications in the endoscope mounting.

The algorithm requires preoperative setting of various control parameters (a, K_p, K_d) which might not be desired. However, the optimizing nature of the algorithm allows a significantly larger flexibility in the parameters as compared to camera calibration in the conventional image-based visual servoing. An additional impractical constraint imposed in this system is a fixed orientation between the camera coordinate system relative to the robot coordinate system (similar as in [21]) allowing a direct translation of the direction vector. This problem can be solved either through a unique mechanical coupling of the camera or automatic image re-orientation using robot encoder values from the rotational joints.

VI. Conclusion

Advanced techniques of conventional/traditional visual servoing do not translate well to real-time systems used in dynamic surgical environments due to system setup and calibration requirements. A design of a robotic control system for automatic servoing of the endoscope towards landmarks selected in an endoscopy video stream is presented. The purpose of the system is to improve the handling of the endoscope

by bridging hand-eye coordination problems with automatic operation for endoscopic cardiac surgery. The results of experiments using simulated targets yielded valuable insights on the practical application of visual servoing to complex surgical environments. The results of phantom experiments indicate the promise of the approach proposed for use in live minimally invasive cardiac surgery. The system will serve as a platform for future integration of other imaging modalities.

ACKNOWLEDGMENT

Authors thank ERC CISST - Computer Integrated Surgical Systems and Technology, Johns Hopkins University, for loaning LARS robot used in these experiments.

Authors thank Haytham Elhawary (Philips Research North America) for critically reading the manuscript.

REFERENCES

[1] *2005 Heart and Stroke Statistical Update. Dallas, Tex: American Heart Association; 2005.*

[2] P. W. Serruys, M.-C. Morice, A. P. Kappetein, A. Colombo, D. R. Holmes, M. J. Mack, E. Sthle, T. E. Feldman, M. van den Brand, E. J. Bass, N. V. Dyck, K. Leadley, K. D. Dawkins, F. W. Mohr, and S. Y. N. T. A. X. Investigators, "Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease.," *N Engl J Med*, vol. 360, pp. 961–972, Mar 2009.

[3] M. A. Hlatky, D. B. Boothroyd, K. A. Melsop, M. M. Brooks, D. B. Mark, B. Pitt, G. S. Reeder, W. J. Rogers, T. J. Ryan, P. L. Whitlow, and R. D. Wiens, "Medical costs and quality of life 10 to 12 years after randomization to angioplasty or bypass surgery for multivessel coronary artery disease.," *Circulation*, vol. 110, pp. 1960–1966, Oct 2004.

[4] K. Shirai, A. J. Lansky, R. Mehran, G. D. Dangas, C. O. Costantini, M. Fahy, S. Slack, G. S. Mintz, G. W. Stone, and M. B. Leon, "Minimally invasive coronary artery bypass grafting versus stenting for patients with proximal left anterior descending coronary artery disease.," *Am J Cardiol*, vol. 93, pp. 959–962, Apr 2004.

[5] A. T. Cheung, G. E. Cruz-Shiavone, Q. C. Meng, A. Pochettino, J. A. Augoustides, J. E. Bavaria, and E. A. Ochroch, "Cardiopulmonary bypass, hemolysis, and nitroprusside-induced cyanide production.," *Anesth Analg*, vol. 105, pp. 29–33, Jul 2007.

[6] H. Shennib, A. Bastawisy, M. J. Mack, and F. H. Moll, "Computer-assisted telemanipulation: an enabling technology for endoscopic coronary artery bypass.," *Ann Thorac Surg*, vol. 66, pp. 1060–1063, Sep 1998.

[7] M. J. Mack, "Minimally invasive cardiac surgery.," *Surg Endosc*, vol. 20 Suppl 2, pp. S488–S492, Apr 2006.

[8] V. Falk, S. Jacobs, and F.-W. W. Mohr, "Closed-chest, robotically assisted CABG," *Multimedia Manual of Cardiothoracic Surgery*, vol. 315, p. 96, 2006.

[9] V. Falk, S. Jacobs, J. F. Gummert, T. Walther, and F. W. Mohr, "Computer-enhanced endoscopic coronary artery bypass grafting: the da Vinci experience.," *Semin Thorac Cardiovasc Surg*, vol. 15, pp. 104–111, Apr 2003.

[10] J. G. Holden, J. M. Flach, and Y. Donchin, "Perceptual-motor coordination in an endoscopic surgery simulation," *Surg Endosc*, vol. 13, pp. 127–132, 1999.

[11] G. R. H. M. and van Veelen M. A., "Assessment of ergonomics in laparoscopic surgery," *Minimally Invasive Therapy and Allied Technologies*, vol. 10, pp. 175–179, 2001.

[12] H. A. Tabaie, J. A. Reinbolt, P. Graper, T. F. Kelly, and M. A. Connor, "Endoscopic coronary artery bypass graft (ECABG) procedure with robotic assistance," *The Heart Surgery Forum*, vol. 2, pp. 310–317, 1999.

[13] R. P. Budde, T. C. Dessing, R. Meijer, P. F. Bakker, C. Borst, and P. F. Grndema, "Robot-assisted 13 MHz epicardial ultrasound for endoscopic quality assessment of coronary anastomoses," *Interactive Cardiovascular and Thoracic Surgery*, vol. 3, pp. 616–620, 2004.

[14] Y.-H. Liu, H. Wang, C. Wang, and K. K. Lam, "Uncalibrated visual servoing of robots using a depth-independent interaction matrix," *IEEE Transactions of Robotics and Automation*, vol. 22, no. 4, pp. 804–817, 2006.

[15] E. Malis, "Visual servoing invariant to changes in camera-intrinsic parameters," *IEEE Transactions of Robotics and Automation*, vol. 20, no. 1, pp. 72–81, 2004.

[16] W. J. Wilson, C. C. W. Hulls, and G. S. Bell, "Relative End-Effector Control Using Cartesian osition Based Visual Servoing," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 684–696, 1996.

[17] X. Zhang and S. Payandeh, "Application of visual tracking for robot-assisted laparoscopic surgery," *Journal of Robotic Systems*, vol. 19, no. 7, pp. 315–328, 2002.

[18] A. Krupa, J. Gangloff, C. Doignon, M. F. de Mathelin, G. Morel, J. Leroy, L. Soler, and J. Marescaux, "Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing," *IEEE Transactions of Robotics and Automation*, vol. 19, pp. 842–853, Oct 2003.

[19] P. Hynes, G. I. Dodds, and A. J. Wilkinson, "Uncalibrated visual-servoing of a dual-arm robot for mis suturing," in *Proc. First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics BioRob 2006*, pp. 420–425, 2006.

[20] M. Jagersand, O. Fuentes, and R. Nelson, "Experimental evaluation of uncalibrated visual servoing for precision manipulation," in *Proc. IEEE International Conference on Robotics and Automation*, vol. 4, pp. 2874–2880, Apr. 20–25, 1997.

[21] G.-Q. Wei and G. Hirzinger, "Real-time Visual Servoing for Laparoscopic Surgery," *Engineering in Medicine and Biology*, vol. 16, no. 1, pp. 40–45, 1997.

[22] R. H. Taylor, J. Funda, B. Eldridge, S. Gomory, K. Gruben, D. LaRose, M. Talamini, L. Kavoussi, and J. Anderson, "A telerobotic assistant for laparoscopic surgery," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, pp. 279–288, May 1995.

[23] A. Kapoor, M. Li, , and R. Taylor, "Constrained control for surgical assistant robots," in *Proc. IEEE International Conference on Robotics and Automation*, pp. 231–236, May 15-19 2006.

[24] " https://trac.lcsr.jhu.edu/cisst," *cisst package*, 2010.

[25] A. Popovic, "Real-time feature tracking for image-guided robotics in cardiac endoscop," in *International Journal of Computer Assisted Radiology and Surgery*, vol. 5, pp. 142–143, Springer, 2010.