Vision guided learning based bimanual robot sewing

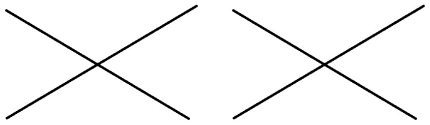
***Abstract*—**

I. Introduction

Stent grafts are tubular shape implants for supporting diseased vessels caused by abdominal aortic aneurysm (AAA). Generally, stent grafts comprise two components: a quasi blood-tight textile tube called graft and a reinforcing metallic ring called stent. Clinically, each stent graft needs to be customised to the patient anatomy, with fenestrations (openings) on the graft body to maintain the patency of important branches to vital organs. Generally, hand sewing techniques are used to attach the reinforcing rings to the fabric and finishing the edge of a fenestration. It is very similar with the suturing technique used in medical filed, in which a circular needle and a needle holder are used. A personalized stent graft with complex shape requires employing with many hundreds or thousands of sutures. Currently, the manufacturing of these device can take 6-12 weeks. The time associates with attaching sutures for manufacturing a custom-made stent graft is long. The burden of assuring the quality of every stitch is also expensive. For patient with complex dilemma, waiting for the delivery of a custom-made stent graft creates unparalleled health risk. The development of an automated or partially automated stent graft sewing technique would be very helpful to change the status. On one hand, automated sewing has been also widely researched in textile industry. Intelligent robotic systems with multi-sensor feedback are built to work in conjunction with a traditional sewing machine. Important topics in this filed includes bimanual robotic sewing [1], fabric tension control, and seam tracking [2, 3]. To cope with environmental changes during the sewing process, various control strategies are implemented, for example, a fuzzy logic controller (Panagiotis, et al. 2006 [4], a hybrid position/force control [1], a leader/follower control strategy [5]. In addition, extensive research has been carried out in the design of sewing heads capable of access the sewn object from one side. For example, KSL Keilmann (Lorsch, Germany) [6] has develop various 3D stitching systems incorporating single sided sewing heads onto KUKA manipulators for sewing fabric-reinforced structure of aircraft parts. On the other hand, as the widely introduction of robotic assisted systems in the field of minimally invasive surgery, research on automated suturing is also widely performed. A suturing task can be divided into two sub-tasks: tissue piercing and knot tying. For each task, research is carried by planning the procedure according to well established manual suture techniques [7] [8] [9] or learning the skills from expert demonstrations [10] [11] [12]. Vision guidance/visual servoing plays a key role in fully automated the suturing skills. In the aspect of positioning the needle to the target point, both the needle posture and the target suturing plane posture need to be measured. Iyer et al. [13] proposed a single arm single camera system auto-suturing system in which the area being sutured on is marked by round markers. With the known geometry of the circular needle and the round marker, the monocular pose measurement algorithm proposed by De Ipna et al. [14] was used for estimating the needle and tissue posture. Another work presented by Staub et al. [15] introduced 3D stereo system and visual servoing technique to improve the accuracy in aligning the needle with target stitching point. Recently, an auto-suturing system with 2D camera guidance and motorized Endo 360 suturing device is presented (Leonard et al. [16]). In this work, a method is presented to track incision contour and automatically distributes equally-spaced stitches along the incision. Comparing the techniques used in industrial sewing and surgical suturing, we found that the latter is more suitable for sewing a stent graft. Firstly, suturing with a circular needle and needle driver is versatile that can do both stent sewing, fenestration finishing and knot tying. In addition, suturing stitch type, like blanket stitch, is stronger than machine made double-thread stitch which easily comes out when one point breaks. So, in this paper we propose a system-level research in which the suturing skills of an expert could be transferred to a robot holding a needle driver and the robot is able to adapt even the needle posture is changed during needle grasping.

A stent graft is a tubular structure composed of fabric supported by a metal mesh called a stent. It is widely used for a variety of conditions for endovascular intervention, but most commonly is used to reinforce an aneurysm. Clinically, each stent graft needs to be customised to the patient anatomy, with fenestrations (openings) on the graft body to maintain the patency of important branches to vital organs. They often come at a significant cost in addition to long delays in manufacturing, largely due to the labour intensive manual tasks involved, subjecting patients to the risk of rupture during the waiting period and precluding treatment to patients presenting acutely. Improved manufacturing of personalised stentgrafts is therefore a critical unmet clinical demand and robot assisted manufacturing is being pursued.

This paper focus on the key process of the stentgraft manufacturing: sewing the stent to a fabric tube. The shape of the fabric tube is pre-designed for the patient anatomy and pre-manufactured. Unlike normal sewing, to sew the stent requires a “3D sewing” technology. Curved needles are



(a) Curved needle (b) Needle driver griping needle

Fig. 1: Curved needle

commonly used for this task, as it can be pierced in and pierced out from one side of the fabric. We take a similar approach to robotize this task.

Curved needle is widely used in surgery. Needle piercing is an important task for surgery robots.

II. System Overview

This paper describe our system for bimanual sewing with a curved needle. We use two robots to manipulate the needle, i.e. piercing and pass from one robot to another. Two KUKA robots are mounted with needle drivers to grip the needle. In addition to these, a third robot is used to control the fabric tube. An special mandrel is designed to support the fabric tube and bound it tightly with the stent. Holes on the mandrel allow the needle to go through and hence stitch the stent and the fabric together. Our sewing task requires high accuracy and robustness. Each stitch is required to be at the exact place and have the same length. We use a vision system to guide the robot movements in order to maintain the accuracy. The needle position is tracked during the whole task. The robot movements are computed online to deliver the needle to stitch at the correct spot. The robot movements are programmed by human demonstrations.

We adopt an bimanual sewing approach: the curved needle is first carried by one robot (A) to pierce the fabric, and picked up by another robot (B) from the other side. The needle is then pulled out by the robot B and passed back to the robot A. The third robot moves to the next stitch position. These movements complete one cycle and repeat the cycle until the whole stent is sewed.

A. Hardware setup

The motorized needle driver mounted on the robot is shown as Fig.1. This device incorporates a medical used Mayo Hager needle driver. This needle driver is widely used in laparoscopic surgery, the design of which can hold firming a surgical curved needle. The needle driver is motorized for robot to drive it. The design projective is that the needle driver can be controlled without mounting a motor directly on its rotation axis, which may impede the needle driver approaching the sewn object in some direction. This design features two set of constraints/guiding slots working in conjunction with pins. The linear slot lies in the direction along the handle of the needle driver and the constraint slot is coaxial with the needle driver axis; therefore the motor rotation can be mapped to the open and close of the needle driver. To reduce frictions in driving this mechanism, bears are used. One feature of this design is that the two jaws of the needle driver work in an unsynchronized way. In the situation that the needle driver is not fully open, the center line of the grasper is located near one jaw which may create difficulty in approaching a target point and grasp precisely. One method to solve this disadvantage is to replace the linear guiding slot with slot with a critical geometry.

B. Vision System

As mentioned above, the needle is manipulated between two robots. During the sewing task, defect can easily occur by the slippage of the needle on the needle drivers. This usually happens during the passing stage: when one robot passes the needle to another, small displacements of the optimal relative pose between the needle and the needle driver can occur. We use a stereo vision system to monitor the process and measure the displacements. Adaptive robot movements are then generated to cope with these small displacements and deliver the needle to the target.

First, the needle is detected in each stereo image using the needle detection algorithm proposed in [2]. For this purpose, a feature image, i.e. I_H, based on the analysis of the eigenvalues of the Hessian matrix [3] is computed to enhance curvilinear structure in the image. Assuming that a calibrated imaging system is available, the 3D points of the needle defined by its *ideal* pose are projected in the image plane. This is performed in order to include a prior information of needle's shape in the detection algorithm. Although the *ideal* pose of the needle is usually different from its real one due to slippage, it still represents of a good guess of the needle pose. Thus, small straight segments are detected in I_H, and only segments that are close to the projected needle and have similar orientation are considered as needle's parts. Finally, these segments are combined in order to create a continuous curve that represents the detected needle in the images. To improve the detection of the needle, the needle driver is also detected in the images using color-base segmentation in HSV space. This allows the reduction of false positive detections of needle segments which are mainly caused by the presence of the needle driver.

The 3D reconstruction of the needle is performed by triangulating the detected needle points of the stereo image pairs. In the current setup, a section of the needle is occluded, however, in the images due to the presence of the needle driver. To overcome the occlusion and to estimate the new needle pose a discretization of the reconstructed needle and the needle defined by the *ideal* position is performed. Starting from the needle tip, points are sampled along the needle shape at distance equal to the arch length of 1 millimetre generating a set of equidistant 3D points, defined by N_{ide} for the ideal and N_{est} for the reconstructed needle, respectively. Finally, a rigid transformation that best maps the two set of points N_{ide} and N_{est}, i.e. the *new* needle pose, is calculated using ...

C. Learning from human demonstration

Hand sewing is an delicate task and programming the robot to do sewing task is time-stacking. We adopt an learning for human demonstration approach for this task. Our learning starts by demonstrating to the robot multiple times how to make a stitch. A *GaussianMixtureModel* (GMM) [1] is used to encode the sewing motion and the generalised motion is then retrieved via *GaussianMixtureRegression* (GMR). Generally speaking, this learning process involves three phases:

1) 1: Human demonstrate sewing skill
2) 2: Motion segmentation
3) 3: Primitive motions learning

1) Human demonstrate how to sew stent graft: The first step is to recode the stitching motion from human demonstration. Human single side hand sewing motion is composed of a few stages: 1)approaching fabric, 2)piercing, 3)releasing needle, 4)griping the needle tip and 5)pulling the needle out, 6)passing the needle to another hand and 7)picking up the needle head. At the beginning of the task, the needle driver grip firming the needle hand. When the tip of the curved needle pierces out from the bottom of the fabric, the needle driver release the needle. The needle is remained in the same pose by the friction of the fabric. The needle driver is then approach the other end of the needle: the tip, and grip the tip. The needle is hence connected with the needle driver again and being pulled out from the fabric. Once the needle is completely pulled out from the fabric, the needle driver pass it to another driver to re-grip the needle head. A full circle of one stitch is hence done and it is ready to start the next stitch.

We use the kinesthetic teaching method to demonstrate all these stages to the robot. The robot is put in gravity compensation mode and its movement is guided by human. The needle driver open and close is controlled an electronic footpedal. The movement of the robot, as well as the needle driver status, i.e. open and close, are recorded. During the demonstration, when the needle driver is close, we assume the needle is firmly connected with the driver and hence no displacement between the needle and the driver will occur. Hence, during the demonstrations, we presume the relative pose between the needle and the driver is a constant value (Figure 1).

We programme the robot in a learning manner and adopt an object centric approach. In the object centric viewpoint, the centre of an manipulation task the is object movement, i.e. the needle movement. In our task, the needle movements for each stitch should the exactly the same so that the quality of the sewing is maintained.

All the trajectories are recorded in 6 d.o.f with euler angles $\{\alpha, \beta, \theta\}$ representation of the orientation and $\{x, y, z\}$ representing the robot end effector position.

2) Motion segmentation: With all the collected training data (sewing trajectories), we segment each trajectory to reflect the different stages of sewing and learn each stage independently. This segmentation is done based on the relation between the needle and its driver: attached or detached. When the needle is attached to the driver, we take the object centric approach and learn the needle movement so that the needle can repeat the same movement every time. When the needle is detach to the driver, we focus on learning the needle driver trajectory in order to reach the proper location to grip the needle.

Therefore, we use the needle driver open and close events to segment the trajectories (Figure 4). Each segment is then learned as a primitive movement and encoded by a statistical model.

Before learning models for each primitive movements, we apply the Dynamic Time Warping (DTW) **??** to align the data across different demonstrations. DTW is a technique that temporally warps the data and find the best match between two time series according to their key features (Figure **??**). In our task, velocity variations do not effect the task quality and hence DTW does not effect our training data.

3) Primitive motion learning: After the we segments the data to a set of primitive movements, we build a model *Omega* to encode each primitive. The same primitive of different trails of the demonstrations are put together as the training data. Each primitive is represented in seven dimension: one temporal value $\{t\}$, three spatial values $h = \{x, y, z\}$ and three orientation values $o = \{\alpha, \beta, \theta\}$. A joint distribution $p\{t, h, o \mid \Omega\}$ is builded by using *GMM*. We choose to use *GMM* because of it's capability of encoding non-linear data and it's robustness of extracting constrains from noise data.

With N Gaussian components, the joint distribution is represented as:

$$\begin{aligned} p(t,h,o \mid \Omega) &= \sum_{n=1}^{N} \pi_n p(t,h,o \mid \mu_n, \Sigma_n) \\ &= \sum_{n=1}^{N} \pi_n \frac{1}{\sqrt{(2\pi)^D \mid \Sigma_n \mid}} e^{-\frac{1}{2}(\{t,h,o\}-\mu_n)^\top \Sigma_n^{-1}(\{t,h,o\}-\mu_n)} \end{aligned} \tag{1}$$

where π_n is the prior of the n^{th} Gaussian component, D the number of variables, and the μ_n, Σ_n the corresponding mean and covariance. For the n^{th} Gaussian component, the mean and covariance μ_n, Σ_n is:

$$\boldsymbol{\mu}_n = \begin{pmatrix} \boldsymbol{\mu}_{t,n} \\ \boldsymbol{\mu}_{h,n} \\ \boldsymbol{\mu}_{o,n} \end{pmatrix} \quad \boldsymbol{\Sigma}_n = \begin{pmatrix} \boldsymbol{\Sigma}_{tt,n} & \boldsymbol{\Sigma}_{th,n} & \boldsymbol{\Sigma}_{to,n} \\ \boldsymbol{\Sigma}_{ht,n} & \boldsymbol{\Sigma}_{hh,n} & \boldsymbol{\Sigma}_{ho,n} \\ \boldsymbol{\Sigma}_{ot,n} & \boldsymbol{\Sigma}_{oh,n} & \boldsymbol{\Sigma}_{oo,n} \end{pmatrix} \tag{2}$$

Each primitive movement is encoded by one model. A smooth generalized trajectory satisfying the constraints encoded with the *GMM* is extracted by using the

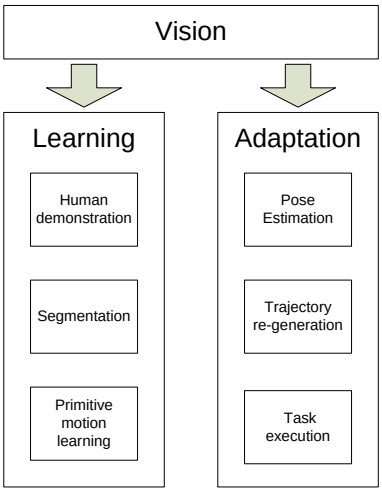


Fig. 2: System overview of bimanual sewing robot

GaussianMixtureRegression (GMR). With the $i-th$ primitive movement model Ω_i, we use a temporal value t to query the trajectory $\{h,o\}$. Here we define:

$$\mu_n = \begin{pmatrix} \mu_n^t \\ \mu_n^{ho} \end{pmatrix} \qquad \Sigma_n = \begin{pmatrix} \Sigma_n^{tt} & \Sigma_n^{t,ho} \\ \Sigma_n^{ho,t} & \Sigma_n^{ho,ho} \end{pmatrix} \tag{3}$$

The *GMR* estimate the conditional expectation value as $\hat{\mu}_{ho}$ with variance $\hat{\Sigma}_{ho}$:

$$\hat{\mu}^{ho} = \sum_{n=1}^{N} \beta_n \hat{\mu}_n \qquad \hat{\Sigma}^{ho,ho} = \sum_{n=1}^{N} \beta_n{}^2 \hat{\Sigma}_n \tag{4}$$

where

$$\hat{\mu}_n = \mu_n^{ho} + \Sigma_n^{ho,t}(\Sigma_n^{tt})^{-1}(t-\mu_n^{ho}) \tag{5}$$

$$\hat{\Sigma}_n = \Sigma_n^{ho,ho} - \Sigma_n^{ho,t}(\Sigma_n^{tt})^{-1}\Sigma_n^{t,ho} \tag{6}$$

and

$$\beta_n = \frac{\pi_n p(t|\mu_n^t, \Sigma_n^{tt})}{\sum_{n=1}^{N} \pi_n p(t|\mu_n^t, \Sigma_n^{tt})} \tag{7}$$

D. Task execution

The relative posture between the needle driver and the needle may change during task execution. This situation happen frequently during needle regrasping procedure, in which the needle may not in its original place during demonstration. To keep the learned needle trajectory and perform fabric piercing precisely, the robot end-effector trajectory needs to be modified in order to adapt to the new needle posture, so each time before performing fabric piercing, needle pose estimation is performed using the stereo vision system and the relative transformation between the ideal needle posture and actual needle posture is calculated. Using the hand-eye calibration matrix, a new robot end-effector trajectory can be achieved.

1) 1: Needle pose re-detection
2) 2: Trajectory adaptation

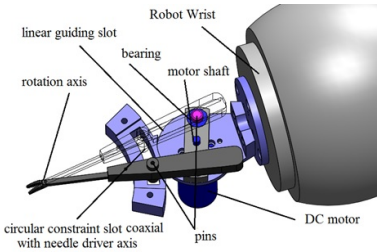


Fig. 3: Motorized needle driver

TABLE I: Overall Shape and Tip Reconstruction Errors ($\mu \pm \sigma$)

Experiments	Shape Errors [mm]	Tip Errors [mm]
Experiment1	1.49 ± 0.84	1.25 ± 0.67
Experiment2	$\mathbf{0.83} \pm 0.34$	0.96 ± 0.44

E. Vision

The shape reconstruction error, $Dist(S_{gt}, S_{est})$, between the estimated shape, S_{est}, with respect to the ground truth shape, S_{gt}, is defined in [3] as:

$$Dist(S_{gt}, S_{est}) = \frac{1}{w+f}\left(\sum_{i=1}^{w} d_{min}(S_{gt}(i), S_{est}) + \sum_{j=1}^{f} d_{min}(S_{est}(j), S_{gt})\right) \tag{8}$$

where $d_{min}(S_{gt}(i), S_{est})$ is the distance between the i^{th} point of S_{gt} to the closest point on S_{est}, while w and f are the number of points of S_{gt} and S_{est}, respectively.

III. Experiments

A. System setup

The system is consisted of a robot mounted with a motorized needle driver, a robot mounted with a stent graft sewing mandrel, a fixed position needle driver, a curved needle and a stereo camera (Figure **??**). The system workflow is as below:

1) 1: Needle driver holding needle root
2) 2: Vision system detect the needle pose relative to the needle driver
3) 3: New robot trajectory is generated according to the needle pose
4) 4: Needle driver approaching mandrel
5) 5: Needle piercing into the fabric and the tip piercing out of fabric
6) 6: Needle driver releasing the needle root, approaching the needle tip
7) 7: Needle driver griping the needle tip and piercing the whole needle out of fabric
8) 8: Needle driver bringing the needle to the second needle driver
9) 9: The second needle driver griping the middle of the needle, the first needle driver griping the root of the needle

The stereo system is consisted of two identical Logitch HD cameras. The cameras are fixed on a tripod and are about

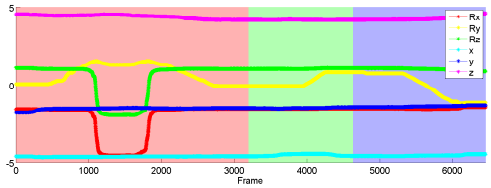


Fig. 4: Segmentation result of human demonstration. The red, green and blue patches label the three segments of the motion

10*cm* distance from each other. Stereo calibration and points triangulation are done by use the OpenCV. The calibration accuracy is measured by using the triangulation results to measure distance between two feature points on the camera view. The error is 0.89 *mm*.

The camera frame is registered to the robot frame by hand eye calibration. During the calibration, a key dot pattern is fixed on a know position of the robot end effector, whose origin is aligned with the end effector origin. The robot moves the key dots around and records the end key dot pattern positions in the robot frame, as well as in the camera frame. The rigid transformation between these two set of positions are computed by using the singular value decomposition technique. This transformation is hence the transformation from the robot frame to the camera frame. We mount the motorized needle driver on the end effector and register the tip pose to the robot. With the result of the hand eye calibration, the needle driver pose in the camera frame is computed.

The needle is initially grip at the very end of the needle driver and we assume that only small displacement of the needle pose will occur during the sewing task. The needle driver tip position is hence used as a prior of the needle position.

B. Human demonstration

For teaching robot the sewing task, we carry out four demonstrations. All demonstrations starts from the same position and sew the same slot on the mandrel. To control the quality of the stitches, across all demonstrations the needle pierces in at the same location and pierces out at the same location. At the beginning of each demonstration, the needle is placed at the same place and normal to the needle driver. Hence the needle pose in the robot frame can be computed accurately.

The demonstrations are segmented into three primitive movements, according to the needle drive open and close even. Figure 4 shows one segmentation results. Figure 6 shows the demonstrated needle driver trajectories in 3D.

C. Learning

GMM is used to learn model for each phase. Figure **??** shows a 2D projection of the build model of each phase. It can be seem from the model that the three phases have different charactoristics. Phase one has small variance from the beginning to the end, as all the movements start from the same point and pierce into the same location. The piercing movements are the same in order to produce similar stitches. Phase two has larger variance compare to phase one, as the needle is detached with the robot and the robot movement has less constraints. Phase three has small variance at the beginning, when the robot needs to pull out the needle from the same location, and has large variance once the needle is pulled out from the fabric. These show that the GMM can effectively capture the constraints at each phase and hence generate proper trajectories for the robot to complete the task.

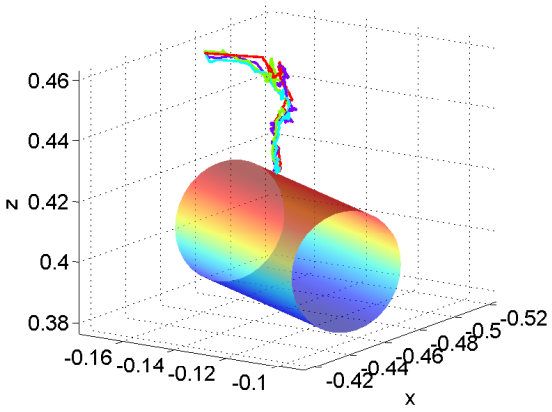
D. Task execution

REFERENCES

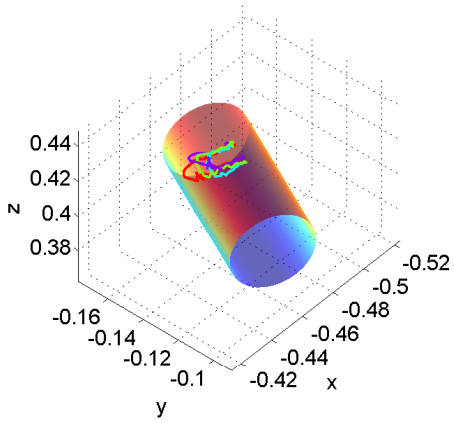
[1] D.A. Cohn, Z. Ghahramani, and M.I. Jordan. Active learning with statistical models. *Journal of Artificial Intelligence Research*, 4:129–145, 1996.

[2] Hedyeh Rafii-Tari, Alessandro Vandini, Lin Zhang, Archie Hughes-Hallett, and Guang-Zhong Yang. Vision-guided learning by demonstration for adaptive surgical robot control. In *Hamlyn Symposium*, pages 39–40, 2015.

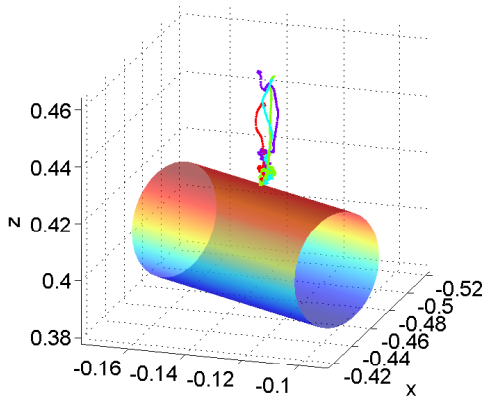
[3] T. van Walsum, S. A M Baert, and W.J. Niessen. Guide wire reconstruction and visualization in 3DRA using monoplane fluoroscopic imaging. *IEEE Trans. Med. Imag.*, 24(5):612–623, 2005.



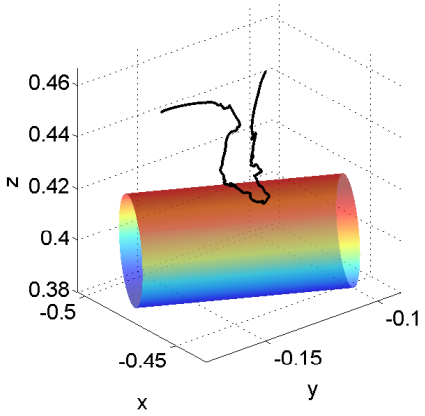
(a) Demonstrations of phase 1



(b) Demonstrations of phase 2

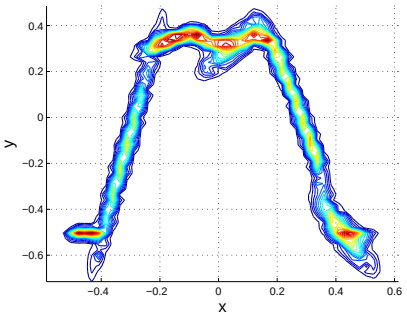


(c) Demonstrations of phase 3

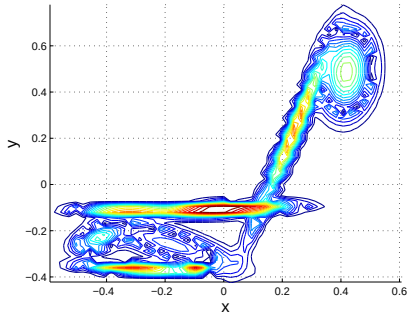


(d) Learnt trajectory of phase 1,2 and 3

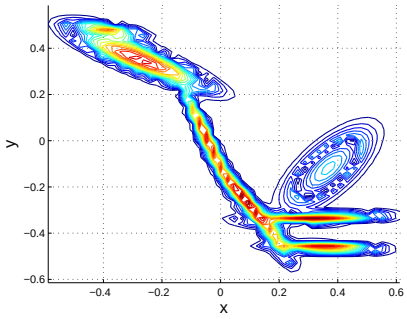
Fig. 5: Needle driver trajectories of human demonstrations and the learnt result. The cylinder represents the mandrel



(a) Phase 1



(b) Phase 2



(c) Phase 3

Fig. 6: 2D representation of the learnt models of different phases.