Needle Steering in 3-D via Rapid Replanning

Sachin Patil Member, IEEE, Jessica Burgner Member, IEEE, Robert J. Webster III Member, IEEE, and Ron Alterovitz Member, IEEE

Abstract—Steerable needles have the potential to improve the effectiveness of needle-based clinical procedures such as biopsy and drug delivery by improving targeting accuracy and reaching previously inaccessible targets that are behind sensitive or impenetrable anatomical regions. We present a new needle steering system capable of automatically reaching targets in 3-D environments while avoiding obstacles and compensating for real-world uncertainties. Given a specification of anatomical obstacles and a clinical target (e.g., from preoperative medical images), our system plans and controls needle motion in a closed-loop fashion under sensory feedback to optimize a clinical metric. We unify planning and control using a new, fast algorithm that continuously replans the needle motion. Our rapid replanning approach is enabled by an efficient sampling-based rapidly exploring random tree (RRT) planner that achieves orders-of-magnitude reduction in computation time compared to prior 3-D approaches by incorporating variable curvature kinematics and a novel distance metric for planning. Our system uses an electromagnetic tracking system to sense the state of the needle tip during the procedure. We experimentally evaluate our needle steering system using tissue phantoms and animal tissue ex vivo. We demonstrate that our rapid replanning strategy successfully guides the needle around obstacles to desired 3-D targets with an average error of less than 3 mm.

Index Terms—needle steering, medical robotics.

I. INTRODUCTION

NEEDLE-based procedures are widely used in minimally invasive clinical procedures for diagnosis and treatment, including biopsy, drug delivery, and radioactive seed implantation for cancer treatment. Performing these procedures using traditional, stiff needles is limited to straight line paths between the needle entry location and target region, which makes it difficult or impossible in some cases to reach clinical targets without puncturing sensitive tissues or colliding with anatomical obstacles. Moreover, the use of stiff needles can result in large targeting errors due to the displacement of the needle from its intended path because of factors such as needle/tissue deformation, uncertain needle/tissue interaction, actuation errors, and noisy sensory feedback [1].

As an alternative to stiff needles, a new class of highly flexible, bevel-tip needles are being developed that enable the needle to move along curved trajectories within tissue when a forward pushing force is applied [8], [40]. These steerable needles offer improved maneuverability within tissue during insertion and greater targeting accuracy. They also facilitate access to previously inaccessible clinical targets while avoiding obstacles such as sensitive anatomical tissues (e.g., vital organs and vessels) and impenetrable structures (e.g., bones). However, guiding a steerable needle around obstacles under image guidance by manipulating the needle at its base requires reasoning in 6D pose space and is not intuitive for a human.

We present a new approach to automatic needle steering to reach targets in 3-D environments while avoiding obstacles and compensating for real-world uncertainties. Our approach uses rapid replanning, a new technique for 3-D needle steering in which a fast motion planner is repeatedly re-executed as the needle is inserted to perform closed-loop planning and control under sensory feedback. In contrast to the standard practice of planning a feasible trajectory and then using a feedback controller for correcting uncertain perturbations, our motion planner is fast enough to correct for perturbations in needle, obstacle, or target motion as they occur. This enables the system to automatically steer the needle along paths that

Fig. 1. Closed-loop needle steering via rapid replanning. Given the current needle tip state, target region, a specification of the anatomy, and characterization of the steerable needle’s properties, our approach uses a fast, randomized motion planner to compute in the available time many feasible motion plans across homotopy classes (top). The method selects the best plan based on a metric such as minimizing path length or maximizing clearance from obstacles (middle). We execute the first control input of the plan and measure the state of the needle tip (bottom). The actual state of the needle tip deviates from the model predicted state because of uncertainty. We repeat the planning process, hence replanning, starting from the actual needle tip state. This approach is made possible by a new, fast planner capable of computing hundreds of feasible plans per second.
avoid obstacles of known location, which is a useful capability to have in a steerable needle system.

We integrate our rapid replanning approach into a system that consists of a bevel-tip steerable needle, a needle steering robot, and an electromagnetic tracker for estimating the needle tip pose in tissue. Given preoperative medical images, the clinician can specify the insertion location and target region as well as sensitive structures such as glands or blood vessels and other obstacles such as bones (Fig. 1). Our rapid replanner then automatically guides the needle around anatomical obstacles to the target region with high accuracy.

Our new rapid replanning approach uses a customized, sampling-based motion planner that speeds up needle steering motion planning to the point that it can be done in real time with typical needle insertion velocities. To enable efficient planning, we leverage several observations and algorithmic advances. First, in contrast to prior motion planning approaches for needle steering, we relax the constant curvature path assumption by planning variable curvature paths and using duty-cycled spinning during insertion [13], [23] to adjust the needle’s net curvature. Second, we propose a new distance metric for incremental expansion of the rapidly-exploring search tree to significantly improve planner performance. These help us achieve orders-of-magnitude reduction in computation time compared to prior sampling-based planners [41] and make the planner suitable for closed-loop needle steering.

In this work, in addition to providing a refined, archival version of our results in [29], we present several important extensions. First, we extend the fast motion planner to create a rapid replanning framework that enables a needle steering system to correct for real-world uncertainties as they occur. Second, we provide experimental results using a new needle steering system that includes preoperative imaging and electromagnetic tracking, demonstrating that the algorithm can work in a practical clinical scenario. In our experiments, the system guided the needle tip in 3-D to targets in phantom and animal tissues ex vivo with errors averaging below 3 mm; for comparison, experienced physicians achieved targeting errors averaging 5.5–6.5 mm when performing procedures using stiff needles [5], [32]. Our experiments demonstrate that our system can achieve targeting accuracy that exceeds current clinical practice while simultaneously enabling avoidance of obstacles.

II. RELATED WORK

Several needle steering techniques have been developed that allow clinicians to adjust the needle path within tissue to improve targeting accuracy. These include bevel-tip flexible needles [40], symmetric-tip needles that can be steered by applying forces at the base [11], [15], curved stylet tips [26], programmable bevel-tip needles [19], and pre-bent concentric tubes [39]. Our emphasis is on bevel-tip flexible needles, but our approach is also applicable to planning and control of needles with stylets and programmable bevel-tip needles.

Significant advancements have been made in modeling bevel-tip steerable needles [8]. A kinematic model generalizing a unicycle was proposed and experimentally validated by Webster et al. [40]. Minhas et al. showed that the curvature of the needle path can be controlled through duty-cycled spinning of the needle during insertion [23]. Swaney et al. [35] proposed a new flexure-based needle tip design that provides enhanced steerability of bevel-tip needles during duty-cycled spinning of the needle, while simultaneously minimizing tissue damage. The mechanics and characteristics of steerable needles have been modeled for tissue ex vivo [24] and in vivo [22].

Motion planning algorithms can be used to compute paths in a robot’s configuration space from a start state to a goal [6], [21]. Motion planning and control for steerable needles in a plane (2-D) has been extensively studied [2]–[4], [17], [30]. Motion planners have been developed for needle steering in 3-D environments with obstacles. Duindam et al. proposed a planner based on inverse kinematics [12], which is fast but offers no completeness guarantees. Sampling-based motion planning algorithms such as rapidly-exploring random tree (RRT) planners, which iteratively explore the robot’s configuration space using a randomized approach, have been effective for a broad range of robotics problems from autonomous vehicles [20] to protein folding [6], [21]. Xu et al. created a variant of RRT for needle steering [41], but the specific approach was too slow for closed-loop implementation. Park et al. proposed a path-of-probability algorithm based on diffusion-based error propagation [27] which considers uncertainty, but this work does not take into account obstacles or noisy sensing. Prior work has also considered controlling steerable needles in 3-D environments to compensate for perturbations during insertion. Hauser et al. [16] proposed a real-time controller which plans helical paths for 3-D needle steering. Seiler et al. [33] proposed a fast trajectory correction method to compensate for uncertainty during insertion. These controllers either do not consider obstacle avoidance or do not provide any guarantees on performance in the presence of obstacles. Van den Berg et al. [37] proposed a framework for planning and LQG-based feedback control of a steerable needle under motion and sensing uncertainty. This framework was extended by Patil et al. [28] for deformable workspaces. Prior LQG-based methods may fail due to control saturation, which is a practical concern for needle steering, and cannot respond in real-time to significant perturbations not in the a priori model.

III. OBJECTIVE

To enable automatic needle steering, our system requires as input a specification of the anatomy. Given registered preoperative volumetric medical images that are standard in clinical care (e.g., CT scans or MRI), the clinician can specify the initial state of the needle tip $X_0 \in SE(3)$, a target $P_{\text{goal}} \subset \mathbb{R}^3$, and obstacles $o_i \in \mathcal{O}$ that include sensitive structures such as glands or blood vessels and other obstacles such as bones. Obstacles are represented as segmented volumes in the images. Our system also requires characterization of the steerable needle. Although rapid replanning could be applied to a variety of steerable needles, we focus on bevel-tip steerable needles [8], which move along an approximately circular arc of constant curvature $\kappa_0$ in the direction of the bevel when inserted into a tissue medium. The needle is controlled by two control inputs: insertion speed $v$ used to insert the needle and
twist speed \( \omega \) applied at the needle base used to reorient the bevel-tip. Our method requires as input the natural maximum curvature of the needle \( \kappa_0 \), and the empirical relationship \( \alpha = h(\kappa) \), \( 0 \leq \kappa \leq \kappa_0 \) that relates the needle’s curvature to the duty cycling factor \( \alpha \) (defined in Sec. IV).

The objective is to automatically steer the needle around clinician-specified anatomical obstacles while optimizing a clinician-specified criteria. The criteria can include metrics such as minimizing insertion length (i.e., minimizing tissue damage) or maximizing clearance from obstacles (i.e., maximizing safety). Our approach, shown in Fig. 2, uses rapid replanning. Given the inputs specified above, the fast planner computes a large number of randomized plans, each defined as a sequence of discrete controls that will steer the needle tip to the target while avoiding anatomical obstacles. From the computed set of plans, the planner selects the best plan based on the clinical criteria. To compensate for uncertain perturbations that occur during needle steering in tissues, the planning process is repeated at frequent intervals in a closed-loop fashion using feedback from the electromagnetic tracking system to sense the needle tip pose at the beginning of each interval and replan a control sequence to reach the target.

At the core of our rapid replanning approach is a fast motion planner for needle steering based on a customized rapidly exploring random tree (RRT) planner described in Sec. V and based on a variable-curvature needle kinematic model described in Sec. IV. The rapid replanning loop in Fig. 2 continues until the target is reached or until the z-coordinate of the needle tip (where the z-axis is the axis along which the needle is inserted prior to entering tissue) is greater than the z-coordinate of the target.

**IV. VARIABLE-CURVATURE NEEDLE KINEMATIC MODEL**

Our planner uses a variable-curvature kinematic model of the motion of the steerable needle’s tip trajectory as the needle is inserted in tissue. The kinematic model is deterministic and does not explicitly consider errors arising from factors such as tissue deformations, actuation errors, and noisy sensing. Our rapid replanning approach will allow us to correct for these errors as they occur during the procedure.

We assume that the needle is flexurally flexible and torsionally stiff, i.e., the shaft exactly follows the needle tip, and the insertions and twists applied to the needle base are directly transmitted to the tip. The motion of the needle is then fully determined by the motion of the needle tip. The state of the entire needle is then described by the needle tip pose, represented as a \( 4 \times 4 \) matrix \( X = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \in SE(3) \), where \( p \in \mathbb{R}^3 \) is the position of the needle tip and \( R \in SO(3) \) is the rotation matrix that encodes the needle tip orientation relative to a world coordinate frame.

We extend the constant curvature unicycle kinematic model of the needle tip proposed by Webster et al. [40] to consider the curvature \( \kappa \) \( (0 \leq \kappa \leq \kappa_0) \) to be an additional control input parameter. Let \( \nu \) be the insertion speed and \( \dot{\omega} \) be the twist speed of the needle. Physically realizing the variable-curvature kinematic model requires that the curvature \( \kappa \) be realizable in terms of insertion speed \( \nu \) and twist speed \( \dot{\omega} \), which are the only two physical control inputs to the system. We later show how the twist speed \( \dot{\omega} \) can be converted to the physical twist speed \( \dot{\omega} \) using duty-cycled spinning of the needle during insertion [23].

Given the control input vector \( \mathbf{u} = [\nu, \dot{\omega}, \kappa] \in \mathbb{R}^3 \), it is convenient to describe the kinematics in terms of the instantaneous twist \( U \in \mathfrak{se}(3) \) expressed in the local coordinate frame attached to the needle tip (Fig. 3), given by [37], [40]:

\[
U = \begin{bmatrix} \dot{\omega} \\ 0 \\ 0 \end{bmatrix}, \quad \dot{\omega} = [\nu \kappa \ 0 \ \dot{\omega}]^T, \quad \mathbf{v} = [0 \ 0 \ v]^T.
\]

where the notation \( [s] \) for a vector \( s \in \mathbb{R}^3 \) refers to the \( 3 \times 3 \) skew-symmetric cross-product matrix. The discrete-time kinematics evolves over time interval \( t \) as:

\[
X_{t+1} = f(X_t, \mathbf{u}, t) = X_t \exp(Ut).
\]

where \( \exp(\cdot) \) denotes the matrix exponential operator. Note that for the special case of \( \kappa = \kappa_0 \), \( \dot{\omega} = \omega \) and Eqn. (2) reduces to the constant curvature kinematic model [40].

Prior work on motion planning for steerable needles in 3D [12], [27], [41] assumes \( \kappa \) is a constant, which severely restricts the range of motion of the needle tip. This makes it difficult for planners to compute a feasible motion plan in...
3-D environments with obstacles, thus sacrificing optimality or completeness. In contrast, our motion planning method assumes a variable-curvature kinematic model that allows us to compute trajectories composed of circular arcs of bounded curvature ($0 \leq \kappa \leq \kappa_0$). This helps us to compute feasible motion plans with sub-second computation time.

We use results from Minhas et al. [23], who demonstrated that any curvature $0 \leq \kappa \leq \kappa_0$ can be approximated by duty cycling the rotation of the needle, i.e., by alternating between (i) insertion without rotation, in which the needle follows a path of maximum curvature ($\kappa = \kappa_0$), and (ii) insertion with rotation, in which the needle moves straight ($\kappa = 0$) by spinning at a constant rate and stopping the spinning such that the tip is at the same axial angle every time. Duty cycling of steerable needles was successfully demonstrated in cadaver brains for neurosurgical procedures [13].

Let the control input $u = [v, \omega, \kappa]^T$ be applied over a time duration $\Delta$. Let $\delta$ be the duration of each duty cycling interval, which is composed of an insert interval of duration $\delta_{\text{ins}}$ and a spin interval of duration $\delta_{\text{spin}}$, as illustrated in Fig. 4. Let $\alpha$ ($0 \leq \alpha \leq 1$) be the proportion of the time spent in spin intervals i.e., $\alpha = \delta_{\text{spin}}/\delta$, where $\delta = \delta_{\text{ins}} + \delta_{\text{spin}}$. The empirical relationship between $\alpha$ is expressed as:

$$\alpha = h[\kappa], \quad 0 \leq \kappa \leq \kappa_0,$$

(3)

where $h[\kappa]$ is dependent on the mechanical properties of the needle and tissue and is determined by fitting a polynomial function to the empirical data gathered during preoperative characterization experiments (Sec. VI-C).

Duty cycling is implemented for needle steering by moving a fixed distance each cycle and spinning with a fixed twist speed $\omega_{\text{spin}}$. Given $\kappa$, we use Eqn. (3) to determine $\alpha$. Since the needle tip arrives at the same axial angle at the end of each spin interval, the duration of the spin interval $\delta_{\text{spin}} = (2k\pi/\omega_{\text{spin}}), k \in \mathbb{Z}$. We then compute the quantities $\delta = (\delta_{\text{spin}}/\alpha)$ and $\delta_{\text{ins}} = (\delta - \delta_{\text{spin}})$. The low level control inputs during a duty cycle interval are given by:

$$v(t) = v, \quad 0 \leq t \leq \Delta/\delta$$

(4)

$$\omega(t) = \begin{cases} \dot{\omega} + \omega_{\text{spin}} & \text{if } j\delta < t \leq j\delta + \delta_{\text{spin}} \\ \dot{\omega} & \text{if } j\delta + \delta_{\text{spin}} < t \leq (j+1)\delta \end{cases},$$

(5)

where $j \in \{0, 1, \ldots, \Delta/\delta\}$ and $\Delta/\delta$ is the total number of duty cycle intervals required to span the duration $\Delta$. Because the needle is axially translating as it rotates, stiction between the needle and tissue is eliminated, which reduces the impact of torsional windup during steering [31].

The needle/tissue parameters, e.g., $\kappa_0$ and $h$, need to be specified before a needle procedure begins. In a clinical setting, we anticipate building a library of needle curvatures indexed by needle type and tissue type that is collected in fresh cadavers or highly similar animal tissues [22]. When performing a procedure on a patient, the library could then be used to select the parameters.

V. RAPID REPLANNING APPROACH

In this section, we present details of the individual components involved in our rapid replanning approach (Fig. 2) for closed-loop needle steering in 3-D environments with obstacles.

The rapid replanning approach requires as input the estimates of the kinematic model parameters and requires a mechanism for unbiased sensing of needle tip pose. While the rapid replanning approach can compensate for perturbations in needle motion during execution, it cannot by itself fully correct for non-zero mean errors that might arise due to factors such as incorrect kinematic model parameters or systemic biases in tip position sensing.

A. Motion Planning

To enable motion planning for a rapid replanning approach, we create a fast motion planner for the steerable needle. We based our planner on a sampling-based rapidly exploring random tree (RRT) [21], which is well suited for the under-actuated, nonholonomic systems like the steerable needle.

The input to the planner is an initial state $X_0$, a target region $P_{\text{goal}}$, and the computation time available for planning $t$. Our algorithm is based on the classic RRT, which proceeds as follows. The planner incrementally builds a tree $T$ over the state space, while satisfying nonholonomic motion constraints of the system and avoiding obstacles in the environment. To expand the tree $T$, a random state $X_{\text{rand}}$ is sampled from the state space. The algorithm identifies a node in the tree $X_{\text{near}}$, that is closest to the sample $X_{\text{rand}}$, as defined by a specified distance metric $\rho[\cdot]$. The algorithm attempts to expand $T$ towards $X_{\text{rand}}$ based upon the best control input $u$ and the resulting state $X_{\text{new}}$ is added to the tree. This process is repeated until either the tree $T$ connects $X_0$ and $P_{\text{goal}}$ or the available computation time is exceeded, in which case the algorithm reports that a solution cannot be found. A feasible plan $\Psi$ is extracted from the tree by traversing it backwards from the goal node to the root.

For a nonholonomic system like the steerable needle, finding the best control input to a sampled state requires solving a difficult two-point boundary value problem of connecting two states in $SE(3)$. Prior RRT-based needle steering planners [41] avoid this by performing deterministic or uniform random sampling of control inputs to determine the best control input $[v, \omega]^T \in \mathbb{R}^2$ that leads the needle tip to a new state $X_{\text{new}}$ closest to $X_{\text{rand}}$. Since these methods also assume the constant curvature kinematic model, the limited range of motion of the needle tip requires a large number of control samples to make progress towards the sampled state. This results in wasted computational effort and is a major computational bottleneck.
To enable efficient planning, we customize the classic RRT algorithm for steerable needles by leveraging several observations and algorithmic improvements. We consider variable-curvature kinematics (Sec. IV) and introduce a new distance metric $\rho[\cdot]$. We present each step of the algorithm (outlined in Alg. 1) in detail below. In the available computation time, we compute many feasible bounded-curvature ($0 \leq \kappa \leq \kappa_0$) trajectories through 3-D environments with obstacles. The individual function definitions in Algorithm 1 are as follows:

**random_point_in_\mathbb{R}^3():** To avoid solving the SE(3) two-point boundary value problem or performing random sampling of control inputs, we sample a random point $p_{\text{rand}} \in \mathbb{R}^3$ in the workspace as opposed to sampling a random state in $SE(3)$. The sampled point can then be connected to a given state $X_{\text{near}} = \begin{bmatrix} R_{\text{near}} & p_{\text{near}} \\ 0 & 1 \end{bmatrix}$ directly using a circular arc parameterized by $[l, \phi, r]^T$, where $l$ is the arc length, $\phi$ is the change in orientation of the needle tip coordinate frame $X_{\text{near}}$ around the $z_{\text{near}}$-axis, and $r$ is the arc radius (Fig. 3). Let $[x, y, z]^T = R_{\text{near}}^T(p_{\text{rand}} - p_{\text{near}})$ be the coordinates of $p_{\text{rand}}$ in the local coordinate frame of $X_{\text{near}}$. The parameters of the circular arc are then given by:

$$r = \frac{x^2 + y^2 + z^2}{2\sqrt{x^2 + y^2}} \quad (6)$$

$$\phi = \arctan(x, -y) \quad (7)$$

$$l = r\theta = r \cdot \arctan 2(z, r - \sqrt{x^2 + y^2}) \quad (8)$$

To accelerate motion planning for steerable needles, we incorporate two forms of biasing that empirically result in significant performance gains. First, we bias the growth of the tree $T$ towards the target region $P_{\text{goal}}$ by sampling from $P_{\text{goal}}$ with a higher probability than the rest of the workspace. If this bias is large, the planner behaves like greedy best-first search [21]. Second, whenever a new node $X_{\text{new}}$ is added to the tree, the planner attempts to connect $X_{\text{new}}$ to a randomly sampled point in $P_{\text{goal}}$.

**control_inputs():** Given a circular arc parameterized as $[l, \phi, r]$ and a given time interval $\Delta$, we derive the augmented control input vector required to compute the new state of the needle tip $X_{\text{new}}$. First, we reorient the needle tip by $\phi$ radians such that the circular arc is contained in the plane defined by the $y$-$z$ axes in the reoriented local coordinate frame $X_{\text{near}}$, which is obtained by applying a rotation of $\phi$ radians around the $z$-axis to the current state $X_{\text{near}}$. We then compute the augmented control input $u = [v, \dot{\omega}, \kappa]^T$ that steers the needle tip along a circular arc of length $l$ and radius $r$ using the relations: $v = l/\Delta$, $\dot{\omega} = 0$, and $\kappa = 1/r$. We compute $X_{\text{new}}$ by applying $u$ to the reoriented frame $X_{\text{near}}$ for a time duration $\Delta$ according to Eqs. (1) and (2).

**nearest_neighbor():** The efficiency with which the RRT algorithm is able to explore the state space is highly sensitive to the distance metric $\rho[\cdot]$ used to compute the nearest node in the tree. In the presence of nonholonomic constraints, widely used metrics like the Euclidean distance are a very poor approximation of the true distance between points in the constrained state space. The performance of the RRT planner degrades as a result of repeated attempts at extending the same nodes in the tree without making sufficient progress [34]. We introduce a new distance metric customized for steerable needles that accounts for the needle’s nonholonomic constraint as well as the buckling of the needle in soft tissue. Since the needle has a maximum curvature $\kappa_0$, not all sampled points will be reachable from a given state because of the nonholonomic constraints of the needle. The reachable set from a state $X_{\text{near}} = \begin{bmatrix} R_{\text{near}} & p_{\text{near}} \\ 0 & 1 \end{bmatrix}$ consists of all points that can be connected to $p_{\text{near}}$ by a circular arc that has a radius $r \geq 1/\kappa_0$ and is tangent to the $z_{\text{near}}$-axis of the local coordinate frame. This definition of the reachable set also directly relates to the distance metric $\rho[\cdot]$ that is used to select the tree node that is nearest to the sampled point $p_{\text{rand}}$. Accordingly, we define the distance metric $\rho[X_{\text{near}}, p_{\text{rand}}]$ as the length of such a circular arc connecting $p_{\text{rand}}$ and $X_{\text{near}}$ if $p_{\text{rand}}$ is in the reachable set of $X_{\text{near}}$ and infinity otherwise i.e.,

$$\rho[X_{\text{near}}, p_{\text{rand}}] = \begin{cases} l(\equiv r\theta) & \text{if } r \geq 1/\kappa_0 \land \theta \geq 0 \\ \infty & \text{otherwise} \end{cases} \quad (9)$$

This strategy restricts the search domain to only those nodes that are within the reachable set of the nearest node $X_{\text{near}}$, thus increasing the likelihood of coverage of the state space [34].

It is important to prevent buckling of the needle shaft, which may occur during insertion because of reaction forces from the tissue. This implies that not all points in the reachable set can be physically accessed by the steerable needle from some poses. In our experiments, we have observed that the needle starts to buckle roughly when the needle tip heading is greater than $\pi/2$ radians from its initial orientation. We preclude such points from being added to the tree by setting the distance to these points to infinity.

**collision_free():** To enable obstacle avoidance, only collision-free arcs are added to the tree. We check if the circular arc connecting $X_{\text{near}}$ and $p_{\text{rand}}$ is collision free by approximating it as a sequence of line segments and checking if all the segments are collision free. Since the obstacle definitions are obtained from segmentation of 3-D scans, the

---

**Algorithm 1**

\[\Psi \leftarrow \text{needle_RRT_planner}(X_0, P_{\text{goal}}, \Gamma)\]

1. $T \leftarrow \text{initialize_tree}(X_0)$
2. $\tau \leftarrow 0$
3. while ($T \cap P_{\text{goal}} = \emptyset$ $\land$ $\tau < \Gamma$) do
   4. $p_{\text{rand}} \leftarrow \text{random_point_in}_\mathbb{R}^3()$
   5. $X_{\text{near}} \leftarrow \text{nearest_neighbor}(p_{\text{rand}}, T)$
   6. $u \leftarrow \text{control_inputs}(X_{\text{near}}, p_{\text{rand}}, \Delta)$
   7. $X_{\text{new}} \leftarrow f[X_{\text{near}}, u, \Delta]$
   8. if collision_free($X_{\text{near}}, X_{\text{new}}, u, \Delta$) then
      9. $T \leftarrow \text{add_vertex}(X_{\text{new}})$
     10. $T \leftarrow \text{add_edge}(X_{\text{near}}, X_{\text{new}}, u, \Delta)$
    11. end if
   12. if $p_{\text{new}} \in P_{\text{goal}}$ then
     13. $\Psi \leftarrow \text{extract_plan}(T, X_{\text{new}})$
   14. end if
   15. $\tau \leftarrow \text{update_time}()$
16. end while
17. return $\Psi$
obstacle meshes are likely to be non-manifold. We use the SOLID library [38] for detecting collisions with arbitrary, polyhedral obstacles at interactive rates.

extract_plan(): When the position \( p_{\text{new}} \) of a newly added state \( X_{\text{new}} \) is found to lie in the target region \( P_{\text{goal}} \), the RRT planner terminates. By traversing the tree \( T \) backwards from the goal state to the root, we obtain a trajectory composed of piecewise circular arcs of bounded curvature \( (0 \leq \kappa \leq \kappa_0) \). We extract a motion plan \( \Psi \) comprised of a discrete sequence of control inputs, in terms of the insertion speed \( v(t) \) and twist speed \( \omega(t) \), that guide the needle to the target along the computed trajectory.

For each circular arc parameterized by a triplet \([l, \phi, r]\) in the trajectory, we first reorient the needle tip by \( \phi \) radians by applying a control input \( \omega = \omega_{\text{spin}} \) for a duration of \( \phi/\omega_{\text{spin}} \). We then compute the factor \( \alpha \) based on the curvature \( \kappa = 1/r \) using Eqn. (3). Given the control input \( u = [l/\Delta, 0, 1/r]^T \), we compute the controls \([v(t), \omega(t)]^T\) for traversing the circular arc in a plane using Eqns. (4) and (5).

The RRT planning algorithm provides a theoretical probabilistic completeness guarantee [21], i.e., if a solution exists, the probability of finding it approaches one as time goes to infinity. However, our rapid replanning method assumes bounded computation time for each replanning interval. Hence, like all RRT implementations on real robots, the planner does not provide a completeness guarantee in practical usage. However, as seen in the results in Sec. VII, our approach of restarting the RRT planner while computing multiple plans in each interval can be effective.

B. Clinical Metrics

When selecting a plan, we consider the following clinically motivated criteria \( c[\Psi] \):

- **Minimizing the total needle insertion length (shortest path)**, i.e., minimizing \( c[\Psi] = \int_0^T v(t)\,dt \). This metric is relevant to procedures in vital organs such as the brain where limiting tissue damage is important [14]. Shortest trajectories, however, often pass in close proximity to obstacles, thereby increasing the likelihood of collisions.

- **Maximizing the minimum clearance from obstacles (maximum clearance)**, i.e., maximizing:

\[
c[\Psi] = \max_{0 < t \leq T} \min_{\forall o_i \in O} d[p_t, o_i], \tag{10}
\]

where \( d[p_t, o_i] \) is the distance of the needle tip \( p_t \) from obstacle \( o_i \in O \). Trajectories that have a greater minimum clearance from obstacles are safer because they are less likely to collide with anatomical obstacles when deviations occur. Such trajectories, however, tend to be longer, thereby increasing the amount of tissue cut during the procedure. This metric could be useful when obstacle avoidance is critical but other tissue damage is manageable, e.g., in liver or muscular tissue.

Since during execution it is important to avoid critical structures, we allow the clinician to specify a safety buffer \( \epsilon \), which requires the motion planner to only return plans that pass at least \( \epsilon \) distance away from each obstacle. Setting an appropriate value of \( \epsilon \) is particularly important when using the shortest path metric. A good value for \( \epsilon \) is the maximum error in tip location that can occur in a single replanning interval due to uncertainty, which would guarantee avoidance of obstacles during execution. We enforce the safety buffer in our RRT-based planner by artificially enlarging all obstacles by a predefined safety buffer \( \epsilon \) using Minkowski sums [38].

The correct choice of the clinical criterion will vary by specific procedure, and we will assume that the clinician will select \( c[\Psi] \) based on the requirements of the procedure. To compute a plan that optimizes \( c[\Psi] \) as much as possible in the allowable computation time, we use our fast, randomized planning algorithm to compute hundreds of different feasible motion plans in a second and then select the plan that performs best under the selected criterion.

We note that any sampling-based motion planner, including our method, cannot guarantee that a globally optimal solution will be found in a finite time interval. Methods like RRT* [18] can compute optimal motion plans as computation time is allowed to increase, but cannot guarantee optimality in finite time and will not be efficient for needle steering due to their requirement of a solver for two-point boundary value problems. Our method will explore the steerable needle’s state space and repeatedly generate independent paths in search of a higher quality solution, and the best found path will progressively improve over the duration of the time interval. A further advantage of our approach is that it is trivially parallelizable, allowing for plan quality to improve as the number of cores in modern multi-core architectures increases.

VI. EXPERIMENTAL SETUP

We describe our needle steering system, shown in Fig. 5, and our experimental setup.

A. System Components

**Bevel-tip steerable needle:** We use needles fabricated from nitinol. In our experiments, we used two needles with tube outer diameters of 0.92 mm and 0.88 mm, henceforth referred to as Needle 1 and Needle 2, respectively. To enable steering at tight curvatures, the needles (1) incorporate a hand-machined bevel tip, and (2) are pre-bent just behind the bevel tip [30].
Fig. 6. (Left) Sim-Test tissue phantom. (Right) Porcine tissue ex vivo.

**Needle steering robot:** We use the needle steering robot design proposed by Das et al. [9] to actuate the needle. The needle is inserted through a hole in the front plate of the robot. Buckling of the needle during insertion is prevented using an external telescoping sheath [40]. Our planner, implemented in C++, runs on a PC and sends the control inputs to the robot controller via PCI bus.

**Electromagnetic tracking system:** Accurate needle steering requires sensing the state of the needle tip position and orientation. Approaches for accurate state estimation include using stereo cameras [30], [40] or fluoroscopic images [22], but these approaches either cannot be used in opaque media such as the human body or can result in high radiation exposure to the patient for longer procedures.

We use an electromagnetic tracking system (Aurora® v1, Northern Digital Inc., Canada) [25] for tracking the needle tip pose. Embedded within the tip of the tube is a 5-DOF magnetic tracking coil, the position and orientation of which (other than the roll about the needle axis) can be measured by the system. In our setup we used the 5 DOF sensor because its diameter is only 0.5 mm, enabling it to fit inside both steerable needles unlike the 6 DOF sensor which has a 0.8 mm diameter. We estimate the roll of the needle using encoders on the servo motor that applies axial twists at the needle base. Electromagnetic tracking is a cost effective and non-invasive method for reliably sensing the state of the needle tip in opaque tissue. The manufacturer specifications for the standard deviation of the error in sensing the position along any given axis is 0.7 mm and in sensing an angle is 0.2° [25].

**B. Tissue Sample Materials**

**Tissue phantom:** We first evaluate our approach using a tissue phantom composed of an animal-protein-based gel marketed as the Simulated Muscle Tissue Ballistic Test Media (Sim-Test) from Corbin, Inc. [7], which was used in prior needle steering experiments [40]. We cast the Sim-Test material, diluted with water by a 5:1 ratio, into a cuboidal block of approximate dimensions 11 cm × 7 cm × 15 cm for our experiments (Fig. 6).

**Porcine tissue:** We also evaluate our approach in fresh porcine tissue ex vivo. In our experiments we used porcine loin tissue of approximate dimensions of 10 cm × 5 cm × 19 cm. It was inhomogeneous and comprised of both muscular and fatty tissue types (Fig. 6).

**C. Needle Characterization**

The approach presented in Sec. V requires that we characterize the maximum curvature of the needle \( \kappa_0 \) and the empirical relationship \( h[\kappa] \) between the curvature \( \kappa \) and the duty cycling factor \( \alpha \). We empirically determined that \( h[\kappa] \) is dependent on the mechanical properties of the needle and the tissue and is not necessarily linear as demonstrated by prior work with duty-cycled needle steering in a gelatinous phantom [23].

To construct the relationship \( h[\kappa] \), we varied the value of \( \alpha \) between 0 and 1 in increments of 0.1. We then computed the duration of the duty cycling interval \( \delta \) for a time interval \( \Delta = 1 \text{ sec} \) (Sec. IV). Given a fixed insertion speed \( v_{\text{ins}} \) and twist speed \( \omega_{\text{spin}} \), we commanded the actuators during each duty cycling interval with control inputs computed by substituting \( v = v_{\text{ins}} \) in Eqn. (4) and \( \omega = 0 \) in Eqn. (5).

The application of these controls causes the needle tip to traverse a circular arc of variable curvature \( \kappa \) in a plane. We performed repeated insertions of both needles for up to 10 cm in both the Sim-Test tissue phantom and porcine tissue ex vivo. We computed a best-fit polynomial curve with a fixed maximum degree (= 3) that minimized the sum of the squared errors of the data points from the curve. This curve defines the relationship \( \alpha = h[\kappa] \). An important point to note is that the smaller the distance \( v_{\text{ins}} \delta \) traveled by the needle tip in every duty cycling interval, the better the approximation of \( \kappa \). However, we empirically observed that for an insertion distance per duty cycling interval of less than 0.5 cm, the effect of inserting the needle without spinning was negligible i.e., the effective curvature was close to 0. This is important because it physically limited the interval lengths at which we could replan during closed-loop steering to at least 0.5 cm.

To determine the effective curvature \( \kappa \) of the planar arc, we recorded the state of the needle tip \( X_t = \begin{bmatrix} R_t & p_t \end{bmatrix} \) after the end of each duty cycling interval for \( N \) such intervals. We observed that the needle tip deviated from the plane because of initialization errors and other sources of uncertainty. To robustly estimate \( \kappa \), we fit a circle to the set of 3-D points given by \( p_t \in \mathbb{R}^3, t = 0, \ldots, N \). We first computed a best-fit plane that minimized the sum of the squared orthogonal distances from each point to the plane by performing principal component analysis (PCA) on the set of points. We then projected the points onto the first two principal components that span the plane and then robustly fit a circle to the set of projected 2-D points [36]. The curvature \( \kappa \) was obtained by taking the reciprocal of the radius of this circle.

Fig. 7 shows the relationship \( \alpha = h[\kappa] \) for Needle 1 and Needle 2 in Sim-Test tissue phantom and porcine tissue. Needle 1 achieved a maximum curvature \( \kappa_0 = 0.11 \text{ cm}^{-1} \) in Sim-Test (Fig. 7(a)). Needle 2 had a lesser outer diameter (0.88 mm) and achieved a maximum curvature of \( \kappa_0 = 0.15 \text{ cm}^{-1} \) in Sim-Test and a maximum curvature of \( \kappa_0 = 0.073 \text{ cm}^{-1} \) in porcine tissue (Figs. 7(b) and 7(c)). Fig. 7 also shows the best-fit curves for \( h[\kappa] \) for each of the needle-tissue combinations considered. In particular, we found that any value of \( \alpha > 0.5 \) for duty-cycled insertion in porcine tissue resulted in a 0 effective curvature, which explains the lack of empirical data points in Fig. 7(c).

**VII. Experimental Evaluation**

We evaluated our new needle steering system in tissue phantoms and porcine tissue ex vivo to demonstrate the ability...
to steer needles to targets with clinically acceptable accuracy while avoiding obstacles. For all the experiments described below, we considered a spherical target region of 1 mm and measured the targeting accuracy of the needle tip by computing the distance between the center of this spherical target region and the final needle tip position after insertion. We executed the motion planner on an Intel® i7 3.33 Ghz PC. We set the replanning interval $\Delta$ to 1 sec and allocated 1 sec of computation time per replanning step, which is a sufficiently short time interval for clinical applications that require needle insertion depths of $\approx$ 10–15 cm.

### A. Evaluation in Tissue Phantoms

We first evaluated our needle steering system in the Sim-Test tissue phantom described in Sec. VI-B. We chose 10 random target regions in the workspace at distances ranging from 9 cm to 11.5 cm from the face of the cuboidal block through which the needle is inserted as shown in Fig. 8. To evaluate the accuracy of the proposed system, we performed 3 insertions for each of the 10 targets under closed-loop rapid replanning using Needle 1 (0.92 mm diameter) and the shortest path metric. The mean targeting error for each of the scenes was 1.24 mm ($\pm$ 0.71 mm), 1.29 mm ($\pm$ 0.79 mm), 1.12 mm ($\pm$ 0.90 mm), and 1.25 mm ($\pm$ 0.84 mm), respectively. Even with obstacles that restrict the navigable space in the environment, our approach successfully steered the needle to the target region without collisions in any of the insertions.

We also evaluated the impact of the selected metric (i.e., shortest path or maximum clearance) on target accuracy and obstacle avoidance for each scene. We used Needle 2 (0.88 mm diameter) and chose 3 out of the 10 target regions in the workspace. For the shortest path criterion, we enlarged all obstacles by a safety buffer of $\epsilon = 5$ mm. We performed 3 insertions for each of the two criteria using our rapid replanning approach. Fig. 9 shows the means and standard deviations of the targeting error for each metric. The maximum mean closed-loop rapid replanning targeting error was 1.7 mm for the shortest path criterion and 1.66 mm for the clearance criterion. To illustrate the impact of uncertainty, we also ran
We compare the targeting error using closed-loop steering and open-loop execution for each of the two metrics. Our closed-loop replanning approach significantly outperforms open-loop plan execution. Error bars indicate one standard deviation of targeting error over repeated trials.

![Targeting Error (Tissue Phantom)](image)

For the open-loop insertions, the mean targeting errors were as high as 10 mm for the shortest path criterion and 9.1 mm for the clearance criterion. Our closed-loop, rapid replanning approach reduces targeting errors compared to open-loop execution for both criteria.

We also evaluate the impact on path quality of the planning time in a replanning interval. In Fig. 10, we illustrate for one target from each scene the impact of planning time on path length when using the shortest path criteria, and the impact of planning time on the minimum clearance from an obstacle when using the maximum clearance criteria. As the planning time increases, the method generates a larger number of feasible motion plans, averaging over 500 plans at 1 second, over which to select the best plan. The motion planner is able to find plans that perform better under the chosen metric as planning time rises, but the improvement in quality diminishes as computation time approaches 1 second.

![Impact of planning time on path quality](image)

the system using an open-loop plan for each target and scene. For the open-loop insertions, the mean targeting errors were as high as 10 mm for the shortest path criterion and 9.1 mm for the clearance criterion. Our closed-loop, rapid replanning approach reduces targeting errors compared to open-loop execution for both criteria.

We also evaluate the impact on path quality of the planning time in a replanning interval. In Fig. 10, we illustrate for one target from each scene the impact of planning time on path length when using the shortest path criteria, and the impact of planning time on the minimum clearance from an obstacle when using the maximum clearance criteria. As the planning time increases, the method generates a larger number of feasible motion plans, averaging over 500 plans at 1 second, over which to select the best plan. The motion planner is able to find plans that perform better under the chosen metric as planning time rises, but the improvement in quality diminishes as computation time approaches 1 second.

C. Evaluation in Porcine Tissue

We also evaluated our rapid replanning approach in porcine tissue samples ex vivo as shown in Fig. 6. We created two scenes with virtual obstacles (shown in Fig. 11). The two scenes are similar to scenes constructed earlier (Fig. 8) and are modified to take into account the different dimensions of the workspace. We use two cylindrical obstacles in Scene #1 and two box-like obstacles in Scene #2, which create a narrow passage and require the needle to traverse two-bend trajectories around the obstacles to reach the target regions.

We used Needle 2 (0.88 mm diameter) for this set of experiments and considered 3 randomly chosen targets in the workspace shown in Fig. 11. We evaluated the system for both the shortest path and maximum clearance criteria in each of these scenes using 3 insertions per target for each criterion. For the shortest path criterion, we enlarged all obstacles by a safety buffer of 5 mm. As before, we also execute the system using an open-loop motion plan for each target for comparison.

Fig. 12 shows the mean targeting error and standard deviations of the targeting error for each of the two criteria for steering using our closed-loop rapid replanner and using an open-loop plan. The mean targeting error for the shortest path criterion for both scenes was 3.6 mm (± 1.85 mm) for our closed-loop rapid replanner and 10 mm (± 2.6 mm) for open-loop steering. The targeting errors are larger than in Sim-Test phantom tissue because of the anisotropic nature of interaction between needle and tissue and heterogeneity of the tissue sample. In spite of the slightly larger errors, the targeting errors using our approach are within clinically acceptable thresholds and are significantly smaller than open-loop steering. The mean targeting error for the clearance criterion for both scenes was 2.6 mm (± 1.2 mm) for our closed-loop rapid replanner and 15.6 mm (± 3 mm) for open-loop steering. Two of the open-loop insertions collided with the virtual obstacles. In contrast, our rapid replanning approach steered the needle safely to the target region. In terms of the metrics, we found that the maximum clearance criterion worked better than the shortest path criterion because of the narrow passage in the environment, which is further constricted by imposing an artificial safety buffer in case of the shortest path criterion.

We also evaluate the impact on path quality of the planning time in a replanning interval for the porcine tissue scenes. In Fig. 13, we illustrate for one target from each scene the
Fig. 12. We compare the targeting error using closed-loop, rapid replanning steering and open-loop execution for the two proposed metrics in porcine tissue. Our approach significantly outperforms open-loop execution. Error bars indicate one standard deviation of the targeting error.

Fig. 13. Impact of planning time on path quality for the shortest path metric (top) and maximum clearance metric (bottom) in the porcine tissue scenes.

impact of planning time on path length when using the shortest path criteria. We also illustrate the impact of planning time on the minimum clearance from an obstacle when using the maximum clearance criteria. As with the Sim-Test scenes, the motion planner finds plans that perform better under the chosen metric as planning time rises, but the improvement in quality diminishes as computation time approaches 1 second.

D. Evaluation in Anthropomorphic Liver Phantom

We apply our approach to an example scenario motivated by the task of ablating a tumor in the liver while avoiding the hepatic veins. We built the anthropomorphic liver phantom that models the hepatic veins based on the hepatic vein anatomical model provided by Desser et al. [10]. In this experiment, the obstacles are physically embedded in a tissue phantom. We modeled the major hepatic veins (middle, left, and right) and the inferior vena cava using hollow tubing (see Fig. 14) so that the veins would be visible on preoperative CT images. We constructed a tumor from plastic that was roughly spherical and 5 mm in diameter. The tumor model was coated with calcium sulfate to assure visibility in the CT images. We placed the model veins, model tumor, and fiducial markers for registration in a box which we filled with Sim-Test to create the anthropomorphic tissue phantom.

After the phantom was constructed out of Sim-Test material, diluted with water by a 5:1 ratio, we used a portable flat-panel CT scanner to obtain preoperative images of the environment (see Fig. 14(b)). We specified 5 insertion locations on the surface of the box and specified 5 different target sites on the tumor for ablation. We also segmented the major vessels and the tumor from the CT scans to obtain obstacle meshes for planning. We used Needle 2 (0.88 mm diameter) for this set of experiments. For each pair of insertion location and target region, we performed closed-loop steering using our rapid replanning approach using the maximum clearance criterion.

We did not perform open-loop steering in this experiment to avoid damaging the needle in case it collided with the model veins during the procedure.

In each instance, our rapid replanning approach successfully steered the needle to the target region on the tumor surface while avoiding the hepatic veins, with an average error of 2.38 mm (± 1.02 mm) over up to 15.5 cm insertion length.

VIII. Conclusion

We presented a new approach to automatic needle steering to reach targets in 3-D environments while avoiding obstacles and compensating for real-world uncertainties. Our approach relies on rapid replanning, a new technique for 3-D closed-loop needle steering that is based on a fast RRT motion planner for steerable needles that uses variable-curvature kinematics and a novel distance measure for planning. This planner allows us to compute many feasible motion plans per second, of which the best plan is chosen for execution based on clinically motivated metrics. Our approach eliminates the need for a separate feedback controller by accounting for perturbations as they occur while simultaneously enabling obstacle avoidance.

We experimentally evaluated our approach by performing procedures in tissue phantoms and porcine tissue ex vivo. Our experimental results demonstrate that our rapid replanning approach successfully guides the needle to desired targets while avoiding obstacles with an average error of less than 3 mm, which is within clinically acceptable thresholds and better than the accuracy achieved by trained clinicians. The method currently relies on a relatively simple variable curvature kinematic model, which suggests that a simple kinematic model coupled with rapid replanning may be sufficient for many clinical applications.

In future work we plan to investigate options to further improve accuracy. One avenue is to incorporate more detailed models of the needle’s kinematic behavior, including needle torsion models (e.g., extensions of [31]) while ensuring motion planning is still sufficiently fast for rapid replanning. Another avenue is to investigate methods for handling constant biases that might arise due to incorrect sensor calibration (which creates offsets in the estimated tip pose) or incorrect estimation
Fig. 14. We applied our needle steering system with rapid replanning to an example scenario motivated by the task of ablating a tumor in the liver while avoiding the hepatic veins. (a) We constructed an anthropomorphic liver phantom that includes the major hepatic veins in the liver (right) based on an anatomical model provided by Desser et al. (Fig. 1 in [10]). The model was built to scale to match human liver dimensions and is shown next to a geometrically correct human liver model manufactured based on segmented CT images of a human patient. (b) We placed the model in a container that was filled with Sim-Test material to create the liver phantom for experiments. We used a portable flat-panel CT scanner to obtain preoperative images of the environment while the electromagnetic tracking system provided measurements of the position and orientation of the needle tip during the procedure. (c) We specified the insertion location and target region and annotated segmented structures such as veins that needed to be avoided. We illustrate feasible motion plans (shown in green) computed at time step 1. (d) Via rapid replanning, our planner successfully guided the needle (reconstructed from CT scans after the procedure) between the middle and left hepatic veins to reach the target on the surface of the tumor.

of the needle curvature. The rapid replanning approach by itself cannot fully correct for these types of systemic errors. For problems in which systemic errors may occur, we will investigate automatically learning uncertain system parameters during execution using a Kalman filter or related framework. For example, the system could estimate the needle curvature online by fitting a curve to the most recent measurements of the needle tip position and then use the latest estimated needle curvature in each replanning interval.

In future work we also plan to evaluate the rapid replanning approach for specific clinical applications. New procedures in the brain or in the abdominal cavity may require specialized designs of asymmetric-tip needles [13], [22], and the rapid replanning approach should be applicable to these needle designs since the underlying kinematic model is similar. For these applications, our rapid replanning approach could enable accurate targeting while automatically avoiding anatomical obstacles such as sensitive or impenetrable structures.

ACKNOWLEDGMENT

The authors thank Phil Swaney and Hunter Gilbert for their help with the experiment setup and Wen Sun for his help in evaluating the motion planner. This research was supported in part by the National Institutes of Health (NIH) under award #R21EB011628 and by the National Science Foundation (NSF) under awards #IIS-0905344 and #IIS-1149965.

REFERENCES

IEEE TRANSACTIONS ON ROBOTICS 12


