Issues in Closed-Loop Needle Steering

Article in Control Engineering Practice · March 2017
DOI: 10.1016/j.conengprac.2017.03.004

2 authors:

Carlos Rossa
University of Alberta
38 PUBLICATIONS 102 CITATIONS

Mahdi Tavakoli
University of Alberta
159 PUBLICATIONS 1,300 CITATIONS

Some of the authors of this publication are also working on these related projects:

Haptics to improve task performance in people with disabilities: A review of previous studies and a guide to future research with children with disabilities View project

Teleoperation View project
Issues in Closed-Loop Needle Steering

Carlos Rossa and Mahdi Tavakoli
Department of Electrical and Computer Engineering
University of Alberta, Edmonton, Canada.

Abstract
Percutaneous needle insertion is amongst the most prevalent clinical procedures. The effectiveness of needle-base interventions heavily relies on needle targeting accuracy. However, the needle interacts with the surrounding tissue during insertion and deflects away from its intended trajectory. To overcome this problem, a significant research effort has been made towards developing robotic systems to automatically steer bevel-tipped needles percutaneously, which is a comprehensive and challenging control problem. A flexible needle inserted in soft tissue is an under-actuated system with nonholonomic constraints. Closed-loop feedback control of needle in tissue is challenging due to measurement errors, unmodelled dynamics created by tissue heterogeneity, and motion of targets within the tissue. In this paper, we review recent progress made in each of the complementary components that constitute a closed-loop needle steering system, including modelling needle-tissue interaction, sensing needle deflection, controlling needle trajectory, and hardware implementation.

Keywords: Feedback control, surgical robotics, robotic assistance, steerable needles, sensors.

1. Introduction
Surgical robotics has significantly grown over the past decade to enable the use of robotic systems in various complex medical procedures that are arguably impossible to perform with conventional means. Robotic systems are used to augment and extend the capabilities of surgeons, offering great levels of dexterity and precision in diagnosis and treatment. The goal of surgical robotics is not to replace the surgeon, but rather to extend his/her capabilities. Thus, one often refers to surgical robots as assistants that work in tandem with surgeons [1].

A special subclass of these systems is devoted to minimally invasive surgery and therapy (MIST), where the surgeon inserts the surgical tools into the patient’s body through small incisions or natural orifices. To date, MIST has been deployed in numerous clinical scenarios including treatments for cancers [2–6], radio-frequency and microwave ablation of liver and lung [7], treatments for astrosophageal reflux disease [8], gastric bypass and banding [9], uterine fibroids and prolapse [10], benign cervical disorders [11], mitral valve prolapse and repair [12], atrial septal defect [13], atrial fibrillation [14], kidney disorders [15], and bariatric [16] and prostate surgeries [17]. When compared to open surgery, MIST has been shown to reduce pain and blood loss, lower risk of infections, shorten hospital stay, and quicken recovery time.

Irrespective of the application, precise system performance and patient safety are shared requirements in these systems. Examples of the former include accurate steering of flexible needles during percutaneous soft-tissue insertions subject to tissue inhomogeneity and limited control over the needle trajectory, surgical instrument control under physiological organ motion in surgery on a beating heart [18], image-guided control and motion tracking of medical instruments [19, 20], and optimal trajectory planning for deformable catheters [21]. Regarding patient safety, surgical robots can show a
large variety of extent of automation. Some are held and operated directly by the surgeon and supplement the ability of the surgeon to perform operations inside the patient’s body with superhuman dexterity and precision. Others rather work in tandem with the surgeon and perform functions such as orienting and stabilizing an ultrasound probe or keeping a surgical tool still.

One may surmise that the higher the autonomy granted to the surgical robot, the higher the risk of injuring the patient if the system performance is mediocre or if it becomes unstable [22]. A medical tool operating under feedback control is vulnerable to various sources of disturbances. Amongst other factors, a surgical instrument that interacts with deformable tissue is subject to uncertainties arising from the contact with the tissue [23], measurement noise and delay [24] including image registration errors [25], and poor visualization of the task being performed [26]. Treating these systems from a closed-loop feedback control perspective will allow us to highlight the trade-off that exists between system performance, patient safety, and clinical translation of robotic technologies.

To illustrate the above problem, in this paper we will focus on control issues in percutaneous needle steering; a particularly challenging subclass of MIST. Percutaneous needle insertion has become part of routine clinical practice for tissue sampling, pinpoint drug delivery, permanent brachytherapy, radiofrequency and microwave ablation of liver, lung, and kidney, and regional anaesthesia. The success of these procedures heavily relies on accurate needle placement within an inner body target location. Bevel-tipped needle steering is particularly challenging. Firstly, a flexible needle inserted in soft tissue is an under-actuated system whose equilibrium condition is never reached as it travels in tissue. Secondly, the needle and tissue form a high-dimensional coupled system subject to uncertainties and disturbances arising from tissue heterogeneity and deformation, anisotropy, anatomic organ motion, and target displacement. These observations make the needle steering in soft tissue a challenging control problem.

This paper is not intended to be a traditional survey on surgical robotics. Rather, we will narrow our focus to the different subsystems that are needed for closed-loop feedback control of flexible needles in percutaneous therapy. This survey is based on the author’s extensive work on modelling [27–35], sensing [36–41], control [42–47], and design [48–50] of robotics-assisted needle steering. As a starting point for our discussion, let us consider the fully automated needle steering system depicted in Fig. 1. The issues addressed in this paper arise from each of the subsystems that compose the fully automated closed-loop system i.e., 1) Modelling needle-tissue interaction for trajectory prediction, 2) Sensing needle tip deflection; 3) Model-based and non-model-based controller design; and 4) Collaborative vs. fully automated steering.

The rest of the paper is organized around each of the above points, which will be discussed in details from Section 2 to Section 5, respectively. A discussion on open challenges regarding each of these points will then conclude the paper.

2. Needle-Tissue Interaction Modelling

Here we will consider steerable needles with an asymmetric beveled tip inserted in soft tissue. The needle’s mechanical behaviour during insertion depends on the coupled deformations of both the needle shaft and the surrounding tissue. The interaction can be classified into four distinct phases as illustrated in Fig. 2, i.e., tissue puncturing, tissue cutting, needle-tissue friction, and tissue deformation [51, 52].

Tissue puncturing: Puncturing happens at the initial contact between the needle tip and the tissue. It starts by deforming the tissue and continues until the contact force reaches its maximum and a crack is formed in the tissue surface. Puncturing results in a relatively large force at the needle tip that drops when the needle tip enters the tissue [28, 51, 53].

![Figure 1: Block diagram of feedback control for fully-automated needle steering illustrates the concept of using a measurement of needle tip position to control the system by comparing its output to a desired trajectory.](image-url)
**Tissue cutting:** As the needle tip further advances into tissue, it displaces the immediate surrounding tissue and the crack grows, creating the effect of tissue cutting [28]. Considering the tissue as an elastic medium, tissue compression at the needle tip leads to a distributed load being applied on both sides of the needle tip that, due to the asymmetric bevel tip, results in a net force normal to the needle shaft ($Q$ in Fig. 2) [54].

**Friction:** Friction is applied tangentially to the needle shaft against the motion of the needle (see Fig. 2). Three regimes of interest exist: 1) The static friction while the needle is in steady state, 2) the transition from the steady state to the sliding state, and 3) the velocity-dependent forces as the needle moves [28, 55]. Friction contributes to tissue displacement along the needle shaft but does not have a significant effect on needle deflection [54].

**Tissue deformation:** The force $Q$ applied at the needle tip makes the needle bend and follow a curved trajectory as it moves. Consequently, the deformed needle shaft compresses the surrounding tissue, which in turn applies forces to the needle shaft and influences the tip trajectory [29]. Tissue reaction forces are applied perpendicularly to the contact surface between the needle shaft and the tissue. Therefore, needle deflection and tissue deformation are coupled effects that influence each other [27, 56].

From a control perspective, the bevelled tip has antagonistic effects: As it facilitates cutting and penetrating the tissue, it also increases the deflection as the needle advances. Thus, twisting the needle base axially changes the direction of the tip force $Q$ and provides steering capabilities. A proper combination of insertion depth and rotation can then be used in order to control the trajectory of the needle tip. A needle with symmetric tip would not provide enough control inputs to compensate for deflection that would arise from interactions with non-homogeneous tissue anyway.

Similar to a bicycle or a car, the forward motion of the needle in tissue is subject to nonholonomic constraints, i.e., the needle cannot instantaneously move in arbitrary directions. The needle is an under-actuated system that is not locally controllable, i.e., any state close to the current state is not reachable in arbitrarily small amounts of time by paths close to the current state [57]. The needle can be manoeuvred within tissue by reorienting the bevel-tip through twists applied to the needle base. The flexible needle shaft bends due to reaction forces from the tissue, causing the needle to follow paths with variable curvatures [58]. To control the states of a flexible needle, mainly only three inputs are available, i.e., needle insertion, needle base twists, and needle base lateral motion. The next subsection review some of the most common models used to represent needle-tissue interactions.

### 2.1. Nonholonomic Kinematics

The simplest and perhaps the most widespread model of needle-tissue interaction is the nonholonomic model first introduced in [59] and derived in its current form in [60]. Essentially, it describes the forward motion of bevelled-tip needles in tissue as a bicycle with a fixed front wheel angle. A simplified 2-dimensional (2D) version is shown in Fig. 3(a). The model is composed of two hypothetical wheels placed at a distance $a$ and $b$ from the needle tip, and oriented by an angle $\phi$ with respect to each other. The steering angle makes it follow a circular path whose radius of curvature $\kappa$ is empirically determined for a given needle and tissue. This arrangement constrains the needle motion to follow a path with a constant curvature, which can be reversed by rotating the needle base axially by 180 degrees.

Fig. 3(b) shows a representation of the needle in 3D space. The position and orientation of the needle tip coincide with that of the moving frame $\{B\}$ attached to the needle tip, with respect to the fixed inertial frame $\{A\}$. In generalized coordinates...
defined as \( p = [x(t), y(t), z(t), \alpha(t), \beta(t), \gamma(t)] \), the nonholonomic model is given by

\[
\begin{bmatrix}
\dot{x} \\
\dot{y} \\
\dot{z} \\
\dot{\alpha} \\
\dot{\beta} \\
\dot{\gamma}
\end{bmatrix} = \begin{bmatrix}
\sin(\beta) & 0 \\
-\cos(\beta) \sin(\alpha) & 0 \\
\cos(\alpha) \cos(\beta) & 0 \\
\kappa \cos(\gamma) \sec(\beta) & 0 \\
\kappa \sin(\gamma) & 0 \\
-\kappa \cos(\gamma) \tan(\beta) & 1
\end{bmatrix} \begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}, \tag{1}
\]

where \( x, y, \) and \( z \) refer to the position of the needle tip, while \( \alpha, \beta, \) and \( \gamma \) are the yaw, pitch and roll of the needle tip, respectively. The system is driven with two control inputs, namely the insertion velocity \( u_1 \) and the axial rotation velocity of the needle base \( u_2 \) written in frame \( \{A\} \). The dot operator \( \dot{} \) represents the first derivative with respect to time \( t \).

For simplicity, the needle tip trajectory is assumed to be the same as the needle shaft, which implies that the tissue is stiff relative to the needle. In practice, however, this assumption does not hold since the needle deflects and compresses the surrounding tissue, which in turn applies forces to the needle affecting its trajectory. To account for tissue displacement, a recent extension to this model has been proposed in [32]. The back wheel is replaced with an omnidirectional wheel that can move sideways, allowing the needle to follow a path with a variable radius of curvature thanks to the additional degree of freedom added by the slippage of the back wheel. As in [59], the model must be calibrated to a given needle and tissue, and for a given insertion velocity. It has been demonstrated that the model accuracy decreases with the insertion velocity due to the increasing tissue cutting force [28].

The principal limitation of the model in (1) is the fact that only the position of the needle tip is estimated and all forces applied by the tissue along the needle shaft cannot be calculated. Yet, this information is critical in order to account for target displacement and other loads applied to the needle during insertion. To address this, finite elements modelling has been proposed.

2.2. Finite Elements Models

Finite Elements Method (FEM) is another common approach taken for simulating needle deflection [61, 62]. Dimaio and Salcudean [61], and Goksel et al. [63] were amongst the first to use FEM to model the needle-tissue interaction. Initially, a FEM model was used to simulate deflection of a needle in free space and to take geometric nonlinearities into account [64]. To model the effects of surrounding tissue, a linear elastic tissue model was used. The geometry of the soft tissue is defined using a mesh composed of 2D or 3D polyhedral elements that are deformed as the needle cuts and advances into tissue [65]. Alterovitz et al. [62, 66] incorporated the effects of the bevelled tip in the 2D FEM model in order to perform motion planning for steerable needles without the need for explicit position feedback [67]. Later, Chentanez et al. [68] expanded the model to 3D. In [69, 70], FEM is also used to estimate needle-tissue contact forces that result from tissue deformation. Similarly, in [71], a FEM-based model path planner takes into account the effects of boundary conditions, elasticity, and nonlinearity, in order to find the best path towards the updated location of a target. Then, a FEM-based feedback controller makes the needle follow the optimal path calculated online [72].

Other applications of FEM include modelling the effect of external forces applied to the tissue in order to shift the target location and improve the needle targeting accuracy [73]. Mallapragada et al. [74] used this concept for breast biopsy procedures. It can also be used to enhance target accessibility by
pushing obstacles and sensitive tissue away from the needle path.

Because of its high flexibility, FEM-based models can effectively describe the behaviour of needles in tissue in the presence of external perturbations. Employing such a comprehensive FEM model can be very time-consuming and not suitable for real-time control. More computationally efficient models may come at the expense of reduced accuracy [63]. Furthermore, certain parameters in FEM simulators may not relate to physical properties that can be experimentally and independently measured. In this regards, mechanics-based models can then be considered.

2.3. Mechanics-based Models

To account for the fact that needle deflection and tissue deformation are coupled effects, researchers have adopted beam theories to develop fundamental mechanics-based models of needles in tissue [31, 64, 75, 76]. Goksel et al. [64] were amongst the first to develop mechanics-based models of a needle in free space subjected to a constant load applied at the tip. Tissue deformation is modelled by contact forces that evolve as the needle bends along its shaft and the applied load.

The Euler-Bernoulli equation describes the relationship between the beam’s deflection \( v \) at a point \( z \) along its shaft and the applied load \( q \) as

\[
\frac{d^2}{dz^2} \left( EI \frac{d^2v}{dz^2} \right) = q, \tag{2}
\]

where \( EI \) is the needle’s flexural rigidity, and \( q \) is a distributed load that acts anywhere along the needle shaft. Integrating both sides of (2) with respect to the position \( z \) gives the shear force acting on the needle. To obtain the shear force, the tip force \( Q \) is added and the result is further integrated until the deflection \( v(t, z) \) is obtained. The key question in mechanics-based modelling is then how to model the distributed load \( q \), and the needle tip force \( Q \).

With regards to the vertical tip force \( Q \), there are two main modelling approaches. Initially, the forces \( F, P \), and \( Q \) were related to the geometry of the bevel edge and to the tissue properties, which suggested a constant, velocity-independent cutting force \( F \) [51, 53, 54, 77, 78]. Nevertheless, it has been shown that model accuracy for a given insertion velocity decreases as the velocity changes [79, 80]. More recent models reported velocity-dependent fracture toughness at the needle tip [28, 81, 82]. Thus, the fact that the tissue acts as a low pass filter in response to fast needle insertion can be used to minimize tissue deformation and damage [83, 84].

The distributed load \( q \) is typically calculated assuming that the tissue is a viscoelastic medium of stiffness \( K \) and viscous coefficient \( C \) that supports the needle shaft (see Figs. 4(a) and 4(b)). Typically, the distributed load \( q \) is assumed to have linear dependence on the magnitude of deflection [28, 31, 54, 61, 78, 85]. The stiffness \( K \) can either be fitted to the model for insertion lengths with constant bevel orientation, or assumed to be depth-dependent [75, 76, 86]. More recently, in [27, 29, 48], the load \( q \) is found by considering the local magnitude of tissue deformation instead of that of needle deflection. The idea is that as the needle tip cuts through tissue, it creates a tunnel corresponding to the historical location of the needle tip (\( v_i(t, z) \) in Fig. 4(a)). The local tissue deformation at point \( z \) is the difference between the position of the needle shaft at that location and the past position of the needle tip, i.e., \( \delta = v(t, z) - v_t(t - \tau, L) \), where \( t \) is time, \( \tau \) is a delay term, and \( L \) is the needle length. Thereby, it yields \( q(z) = K\delta - C\dot{\delta} \).

The partial differential equation (PDE) governing the motion of the needle can be written in the form

\[
EI \frac{d^4v(t, z)}{dz^4} + c_1 \frac{d^2v(t, z)}{dt^2} + P \frac{d^2v(t, z)}{dz^2} = Qu_3(t), \tag{3}
\]

where \( c_1 \) is a constant that depends on the mechanical properties of the needle, and \( u_3(t) \) is a step-type function that changes the sign of \( Q \) depending on the orientation of the bevel tip (rotation of the needle base), i.e., \( u_3 = 1 \) makes the needle deflect downwards, and \( u_3 = -1 \) makes the needle deflect upwards. This formulation automatically accounts for an unlimited number of needle axial rotations.

Note that in (2) the deflection \( v(t, z) \) is both a function of position \( z \) and time \( t \) and thus the PDE given in (3) cannot be solved using conventional methods such as separation of variables [31]. The deflection is then approximated by a linear combination of \( n \) arbitrary candidate shape functions \( W(z) \) representing the modes of vibration of a clamped-free beam (see Fig. 4(c)), that is

\[
v(z, t) = \sum_{i=1}^{n} \Phi_i(t) W_i(z). \tag{4}
\]

Here, \( \Phi_i(t) \) are time-dependent eigenvalues to be determined such that the needle-tissue system
reaches equilibrium, and \( n \) is the number of assumed vibration modes [87].

An energy-based formulation has been used to find the coefficients \( \Phi_i(t) \) in [27, 48]. In [54, 78] the needle shape is approximated by a third order polynomial equation and a similar formulation is derived to find the polynomial coefficients. The model accounts for lateral and axial deflection of the needle, tissue deformation, and force applied at the needle base. Later, the same model was extended to include needle axial rotation during insertion [88], and to model the behaviour of a needle in tissue when a permanent magnetic field applies forces to the needle shaft in an attempt to provide additional control over needle deflection [89].

Since the needle axial rotation is the main control command over deflection, its torsional dynamics can have an effect on steering accuracy [90]. Reed et al. [91, 92] studied the effects of torsional friction and compensated for it by a model-based controller.

Mechanics-based models require mechanical properties of the tissue as inputs, which can be obtained from direct measurement. Tissue parameters such as Young’s modulus, tissue cutting force, and stiffness are commonly assumed to be constant throughout the insertion, or approximated by a series of different local finite homogeneous models [93]. It has been shown such parametric mismatch can have drastic effects on steering accuracy [69, 70]. Yet, most of the models reported above assume a homogeneous medium and model parameters. To account for modelling uncertainties, one may consider adaptive models.

2.4. Adaptive Models

Given feedback from needle deflection, it is possible to evaluate model accuracy and adjust its parameters to best match observed measurements. The motivation behind adaptive models is that uncertainties arising from tissue heterogeneity, tissue deformation, needle buckling, tracks left in the tissue by previous insertions, and other unmodelled factors, can be accounted for to some extent [27, 94].

Adaptive online identification of model parameters has been developed for both mechanics-based and nonholonomic kinematic models. A simple approach involves adding noise to the input parameters of an ideal model [95]. Based on needle tip tracking information obtained from 2D axial ultrasound image slices, Carriere et al. [36] used a particle filter to inform a kinematic model about the current location of the needle tip in order to create adaptive estimates of the radius of curvature \( \kappa \) in (1). It iteratively updates the parameters of (1) for each input ultrasound image frame and updates the predicted needle tip path. Along the same lines, [96] proposes a method to update the needle curvature for use in closed-loop steering. The radius of curvature, initially empirically related to the tissue Young’s modulus of elasticity, is updated online through a linear Kalman filter.

Adaptive models have also been proposed for mechanics-based approaches in order to update physical properties of either the needle or tissue. The authors in [94] devised a model with depth-varying mean parameters that calculates the tissue stiffness \( K \) and viscosity \( C \) effects. As outlined in [27], a needle model that updates the magnitude of the needle-tissue cutting force \( Q(t, L) \) in (3) as the needle is inserted, can account for local variability in the tissue properties.

An alternative approach is to develop data-based methods to build a model of the plant from measurements of needle-tissue interactions [35]. Such models can be a valuable solution to estimate the system output without deep understanding of the system physics. This is, however, beyond the scope of this paper.

3. Sensor for Feedback Control

Closed-loop feedback control requires real-time measurement of needle deflection. Typically, deflection is measured as close as possible to the needle tip using ultrasound images or alternative sensors such as optical fibres. The information is then fed back in the controller. Depending on the model employed, state estimation may also be necessary. In this section, we will review the issues related to sensing needle deflection during insertion.

3.1. Image-based Feedback

Needle steering is often performed under ultrasound (US) image guidance. The literature on segmenting needles from ultrasound images is quite extensive. Here, we will only provide the reader with an overview of the main issues related to ultrasound image-based needle tracking. These are two steps, namely segmenting the needle from the image, and filtering out noise.
Three common imaging modalities are 3D volumetric ultrasound, 2D sagittal (longitudinal) imaging, and 2D transverse imaging (see Fig. 5). Sagittal ultrasound images (Fig. 5(a)) are acquired in a plane parallel to the needle’s shaft and provide a consistent view of the needle from which the needle tip position can be obtained. However, depending on how the needle deflects, only a portion of the needle may be visible in the images (see Fig. 5a). Transverse images (Fig. 5(b)), on the other hand, are obtained in a plane perpendicular to the shaft and show a cross section of the needle, thereby eliminating complications of probe alignment at the cost of only seeing a single cross section of the needle along each transverse image (see Fig. 5b). 3D ultrasound images (Fig. 5c) build a volumetric image of the tissue from crystal arrays pointing in different directions and fired in a particular sequence [97]. Other machines have a single 2D array of crystals that moves within the ultrasound probe in order to construct the 3D image from a set of 2D images [98–100]. The field of view of such 3D transducers is often very narrow and thereby only a small portion of the needle can be visualized at a time. Although 3D imaging is expected to enlarge the sampled volume thereby increasing the accuracy of hitting a target [101], typically only one to two 3D images can be reconstructed per second. This makes the modality not suitable for real-time control.

As an alternative to the above, the 3D path followed by the needle tip can be obtained from a series of 2D transverse images acquired at different depths in tissue (see Fig. 5d). In this case the ultrasound probe moves in synchrony with the needle, such that the same cross section of the needle is visible in the images. Recent techniques make use of motorized probes that move with the needle tip [102], or translate along the shaft once the needle is fully inserted [34].

Another primary limitation in using ultrasound images is the low quality of images that often contain artifacts that are hard to interpret and distinguish from targets. Accurate localization of the needle in such noisy environment requires post-processing and filtering. Kaya et al. [103] used Gabor filtering to estimate the insertion angle and minimize outliers while the needle trajectory is found with polynomial fit estimator. Other researches implemented needle tracking based on a Hough transform [104–107] and Hough circle transform [108] in order to find highly bent needles in the images.

Linear Kalman filter has also been successfully used for needle tracking in ultrasound images to predict where the needle is within a region of interest [33, 109–111]. It is assumed that the needle is being inserted along the $z$ axis as in Fig. 3(b)), and axial ultrasound images are used to capture the needle tip position in the $xy$ plane. If one looks at the needle tip as a point in the $xy$ image, one can make the assumption that the frame-to-frame motion of the needle tip in the imaging plane is slow and so the linear Kalman filter is designed to reduce large quick changes in estimated needle tip position (corresponding to a low needle tip velocity in $xy$). These noise-reduced estimates are then input to a given needle steering controller without any change to the original needle-tissue model. For instance, Waine et al. [33] applied this concept to remove outliers such as those resulting from air bubbles and tissue inhomogeneity, which can often be mistaken for the needle’s cross section when performing insertions into biological tissue.

In most of the above methods, the ultrasound probe must move in synchrony with the needle, which can result in further unwanted deformation of the tissue. Two options are then available. The first involves monitoring the location of the target and the motion of the surrounding tissue in real
time [112–117], and then compensating for any displacement in the control loop [118]. The second option combines ultrasound images with a physical model of the needle in tissue, and relies only on the observation of a portion of the needle shaft for measuring its entire deflection. The latter is elaborated below.

3.2. Feedback from Partial Image Observation

The idea of estimating the needle tip location based on partial image observation has been proposed in order to limit the motion of the ultrasound probe and minimize discrepancies between pre-and intra-operative target locations due to tissue displacement.

In [30], a method is proposed to predict needle deflection based on the observation of deflection from a single transverse image located along the needle shaft. The needle is modelled as a series of springs and rigid bars connected in series. The deflection measurement obtained from the transverse image is then used to determine the model parameters and estimate the entire needle shape. The ultrasound probe can be maintained at one fixed position as the needle is inserted. The method was later adapted to work with sagittal ultrasound images in [34]. In [27], the idea of partial image observation was further extended through a model that adaptively updates the needle-tissue contact forces as a function of the tissue displacement along the needle shaft. An ultrasound probe follows the needle tip and stops at an appropriate position while the needle is still being inserted. The model parameters are then adjusted such that the predicted deflection matches the measurement. Partial observation combined with the linear Kalman filter was used in [33] to determine optimal needle rotation depths that minimize targeting errors.

3.3. State Observers

Consider a typical scenario where feedback of needle tip deflection (i.e., \( x, y, z \) in (1)) is obtained from 2D transverse ultrasound images. Since knowledge about the needle orientation \((\alpha, \beta, \gamma)\) is also necessary for prediction and control, model-based state observers can be employed to estimate non-measurable variables [38, 60, 119].

Kallem and Cowan [120] designed a linear observer for a kinematic model [59] in generalized coordinates such that from the information about needle tip, the needle pitch, roll, and yaw can be determined. This approach only requires the extraction of the needle tip position from images, rather than the entire needle shape, thereby simplifying the image segmentation problem. In [121], the observer is extended to work with a feedback linearisation controller. Similarly, a nonlinear observer that uses Cartesian position measurement data to estimate the orientation of the needle tip in tissue is described in [38]. The zero convergence of the observer error is shown using Lyapunov-based methods. Henverly et al. [122] used a linear model to represent the dynamics of the unmeasured states via state immersion into a finite higher dimensional manifold. The observer estimates the complete needle orientation and also filters noisy position measurements.

Through the following state transformation of variables \( s = [x, \sin(\beta), -\cos(\beta)\sin(\gamma)]^T \), (1) can be re-written in the form \( \dot{s} = As + \phi(u, s) \), which outputs \( y_{mes} = Cs = x \), where \( A \) is constant matrix and \( \phi(u, s) \) is the non-linear component of the transformed system. From the measured deflection \( y_{mes} \), one estimates the states of the transformed system as \( \dot{s} = A\hat{s} + \phi(u, \hat{s}) + \Delta L(\hat{y} - y_{mes}) \), where \( \Delta \) is a diagonal constant matrix, and \( L \) is designed such that \( A + LC \) is Hurwitz.

Disturbance observers can also be employed in the control of systems with external disturbances or model uncertainties [123]. In [124], an FEM model of the tissue calculates the forces acting on the needle by considering them as external disturbances being applied to the needle.

3.4. Alternative Sensors

As an alternative to imaging modalities such as ultrasound [27, 30, 97, 125], X-Ray [126–131], MRI [132–136], or fusion of MRI and ultrasound [137], which are often limited in resolution and sampling rate, image-based feedback can be substituted with soft sensors.

An interesting concept to estimate deflection is to use a continuum robot as a force sensor and then estimate needle deflection using a mechanical model [138]. In [139, 140], Xu and Simaan demonstrated that by sensing loads on a continuum robot, certain components of the force applied at the end-effector can be determined. Rucker et al. [141] extended this approach to estimate forces applied at the tip of a tendon-driven continuum robot using a kinematic model and uncertain pose measurements. Along the same lines, Abolhassani et al. [142] introduced a force-sensor based estimator for needle deflection.
Forces and torque measured at the needle base are related to the loads applied to the needle through (2). The concept is rather simple: Integrating once both sides of (2) gives the shear-force $F(z)$, and the integral of the shear force is the bending moment $M(z)$. When a force sensors is attached to the needle base, $F(z = 0)$ and $M(z = 0)$ are known at any time allowing one to solve for the model parameters $q$ and $Q$. The needle deflection $v(t, z)$ is then calculated by further integrating the bending moment twice. The concept was further explored in [37, 41, 45] in order to adaptively update the shape of loads acting on the needle. A model that accepts needle axial rotations based on this concept is yet to be developed. Longitudinal insertion force data has been also used for identifying tissue layers as the needle is inserted [23], which can help further enhance the force-based estimation proposed in [37, 41, 45, 142].

Fiber Bragg grating is another attractive alternative to sensing in the biomedical field due to advantageous properties [143] such as lower noise when compared to imaging. It has been demonstrated in [144–146] that an optical fiber embedded into the needle can be used for direct measurement of needle deflection, and even for three-dimensional reconstruction of the needle shape [147].

4. Needle Steering Controller Design

The next subsystem of our needle steering control scheme is the controller itself that, given the current information about needle deflection and, in some cases, the estimated future deflection, calculates the necessary steering actions that bring the needle on the desired trajectory.

Needle steering controllers can be classified into three main categories according the control objective, as illustrated in Fig. 6. The first and simplest controller aims at navigating the needle tip to a desired point in tissue. We will henceforth refer to this control problem as weak regulation. Two typical applications of weak regulation are tissue sampling (biopsy) [148–150] and percutaneous ablation [151, 152], where the needle tip must move to a desired location regardless of what trajectory it takes.

The second category comprises controllers whose control objective is to minimize the needle deflection at all depths. We will refer to this category as regulation. Applications of regulation can be found in transperineal prostate brachytherapy where the needle is controlled to follow a path as close as possible to a straight line such that strands of radioactive seeds can be deposited along the insertion path [153–155]. The last category is the tracking problem, where the needle tip follows a pre-defined trajectory that is not necessarily a straight line. It is typically employed in cases where the needle must be manoeuvred to avoid anatomical obstacles such as muscles, bones, or vessels [66, 156–158]. In the following, we will address each control objective.

4.1. Weak Regulation

A weak regulator is often a model-based predictive controller. It is composed of a needle-tissue model and a solver that minimizes a cost function. The cost function typically relates different steering actions, i.e., twists or lateral motion of the needle shaft at different depths, to the model-predicted targeting error. The objective of the controller is to reach a desired target with a minimum amount of control actions such that tissue trauma is minimized. A popular choice for this purpose is the quadratic cost function

$$\Gamma(u) = \|x - x_{ref}\|^2 + \|y - y_{ref}\|^2 + \Lambda\|u_2\|^2,$$  (5)

where $\| \cdot \|$ denotes the Euclidean norm, $(x_{ref}, y_{ref})$ is the reference trajectory, that is, a single point in
case of weak regulation, and $\Lambda$ is a weighing parameter that penalizes the control action.

The simplest weak regulator selects a single rotation depth that minimizes $\Gamma$, amongst a set of discrete rotation depth candidates, ranging from the current depth to the depth of the target [45]. When more than one rotation is allowed, optimization of (5) needs a multi-variable and interactive solver. For instance, in [42], Rapidly-Exploring Random Tree (RRT) is used. RRT incrementally grows a tree of feasible control actions, and provides a quick high dimensional search subject to different optimization constraints [159, 160]. In al. [44], a nonlinear Model Predictive Controller (MPC) iteratively optimizes (5) over a receding control horizon. The main difference between these approaches relies in the cost function and the selected iterative solver. Approaches such as [44] penalize both the predicted targeting error and steering actions, whereas [33, 36, 45] define the control objective as having the smallest deviation from the target and do not restrict the control action ($\Lambda = 0$).

Other ways to minimize (5) involve solving for inverse kinematics models and create an optimal off-line path planning [19, 161]. A variety of algorithms are available for solving such optimization problems involving rigid bodies and articulated rigid bodies with kinematic and dynamic constraints [162]. Several motion planning algorithms have been used [163–165], including planning of 3D paths considering motion and sensing uncertainty. Researches have also combined online feedback obtained from needle tracking in ultrasound images such that the optimal depth of needle rotation can be updated online [33, 91, 166, 167].

Generally speaking, the weak regulation control problem heavily relies on model accuracy. Alternative control methods are addressed in the next subsections.

### 4.2. Regulation

Regulation can be seen as a particular case of trajectory tracking where the reference trajectory is a straight line connecting the needle’s entry point to the target point deep inside the tissue. The regulation error is, therefore, the measured deflection at all depths (see Fig. 6). Regulation problems are generally based on controllers such as sliding mode control, although predictive controllers have also been employed in an attempt to penalize the control action and minimize tissue trauma.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Control objective</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak regulation</td>
<td>Reach a desired deflection at a given depth</td>
<td>Transperitoneal brachytherapy</td>
</tr>
<tr>
<td>Regulation</td>
<td>Minimize deflections at all depths</td>
<td>Biopsy Ablation</td>
</tr>
<tr>
<td>Tracking</td>
<td>Follow a depth dependent deflection profile</td>
<td>Obstacle avoidance</td>
</tr>
</tbody>
</table>

Figure 6: Classification of needle steering controllers according to different control objectives.

A common and intuitive approach in regulation is the continual duty-cycled rotation of the needle base [168–172]. When the needle is inserted without any change to the orientation of the bevel angle (no axial rotation of the needle base), the needle follows a trajectory with natural curvature $\kappa$. As the needle is inserted with constant twists of its base, at a rate relatively larger than its insertion velocity $u_1$, a trajectory close to a straight line can be achieved [170]. By combining periods of needle base rotation ($T_{rot}$) with periods of non-rotation ($T$), any curvature $\kappa_d$ given by

$$\kappa_d = \kappa \left( 1 - \frac{T_{rot}}{T_{rot} + T} \right),$$

ranging from the natural curvature $\kappa$ to zero curvature can be achieved [169]. The major limitation of this method arises from the tissue trauma and drilling effect generated by such periods of constant rotation.

Feedback linearisation of (1) is another common approach taken to regulate the needle to a single plane, namely the $y, z$ plane [120]. Let $r = [r_1, r_2, r_3]^T = [x, \beta, \gamma]^T$ denote the state vector of the reduced order system and $r = 0$ be the desired equilibrium state. Through the state transformation $s = [r_1, \sin r_2, \kappa \cos r_2 \sin r_3]^T$ and $v = -\kappa^2 \sin r_2 + \kappa \cos r_2 \cos r_3 u_1$, the equation in the feedback linearised form is

$$\dot{s} = As + Bu_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} s + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_1,$$

and $u_2 = Cs = [1, 0, 0]^T s$. This leads to a fully controllable and observable system. Recently, it
has been shown that a similar controller can be derived using an integrator-backstepping approach [46]. Let \( \dot{r} = [\sin \beta, \cos \gamma, \tan \gamma \cos \beta] \mathbf{u}_1 + [0, 0, 1]^T \mathbf{u}_2 \), and using the change of variable \( \xi_1 = \sin \beta \) and \( \xi_2 = \sin \gamma \), one can rewrite \( \dot{r} \) in a strict feedback form. By controlling the needle to navigate around singularity points while switching regularly between two 2D integrator-backstepping controllers, it is possible to achieve 3D steering.

Inverse kinematics has been applied to regulation in [19, 173], where a closed-loop control using X-ray imaging as feedback is designed based on a FEM model and a low-dimensional linear system of equations. Alternatively, a dynamic model of the needle-tissue system was proposed in [174] and an inverse dynamics controller was designed. It showed that parametric mismatch can have drastic negative effects on the system behaviour and accuracy.

To minimize the effects of parameter mismatch, sliding model control for needle steering has been proposed [47, 175]. Based on (1), a sliding mode control law independent of any model parameter is formulated in [47] in 2D, and in 3D in [175], which allows for the model to reach any desired trajectory within a specified error. Considering only the deflection on the \( y, z \) plane, a common choice for the sliding surface is

\[
s = c_2 \frac{de}{dt} + c_3 e
\]

where \( e = y_{\text{ref}} - y \) is the error from the desired trajectory, and \( c_2 \) and \( c_3 \) are positive defined constants. In a regulator, \( y_{\text{ref}} = 0 \) \( \forall t \). When \( s \) exceeds a predefined threshold the needle is rotated by 180 degrees. The switching threshold is defined such that the orientation of the needle tip, i.e., \( \alpha \), remains bounded. Such controllers can be extended to follow any desired trajectory \( x_{\text{ref}}, y_{\text{ref}}, \) which brings us to the trajectory tracking problem.

### 4.3. Trajectory Tracking

Trajectory tracking schemes are designed to use physician-selected and patient-specified parameters to define a path towards the target given the feasible needle insertion points, and locations of anatomical obstacles. Typically a trajectory tracking problem is implemented in two steps. The first step is preoperative. Based on the open-loop model, a motion planner determines a feasible path towards the target and the subsequent sequence of actions. The second part is intraoperative. The planned trajectory and/or the steering actions are updated online based on feedback of the needle tip [60].

Tracking was initially proposed as a path planning algorithm in a 2D space with obstacles [66, 176], and in an obstacle-free 3D environment [95]. Reed et al. [167] combined a 2D planner with image feedback, a state observer, and the feedback linearized controller discussed in the previous section in order to compensate for out-of-plane deviations of the needle. In [170, 172], it is demonstrated that duty-cycle spinning can also be extended to follow arbitrary paths other than a straight line.

Since the 3D nonholonomic model has no closed-form solution in its original form [59], and, as we will see in Section 6, it is not possible to control such a 3D nonholonomic model using a smooth continuous control law, the control action is typically discretised and a multidimensional optimization solves for (5).

RRT optimization has been used in [42, 156, 177] to deal with motion planning under nonholonomic constraints with discrete control input. Levenberg-Marquardt optimization is applied in [178] with a discrete controller whilst the path planning is solved algebraically. Path planning for such discrete control action is cast as a nonlinear optimization problem, which is solved via an analytic relation between position of the needle and control input. Researchers have also considered the effects of motion uncertainty and used Markov decision process to find optimal paths [62, 179], or replanned the reference trajectory online given the updated location of targets [111, 166].

In order to increase manoeuvrability around obstacles, modified needles such as notched [49], and flexure-based steerable needles [180] have been proposed. The idea is to modify the needle’s flexural rigidify locally such that it follows a curvature that facilitates steering in constrained environments. Alternatively, pre-curved needles [181] or those with actuated tip [182] can also enhance steerability. These are only a few examples of regulation problems; many others can be found in the literature.

## 5. Needle Steering Robots

The last component of the closed-loop needle steering scheme is the steering device itself. Let us now turn our focus back to the starting point of our
Table 1: The different varieties of robotic needle steering according to their levels of automation

<table>
<thead>
<tr>
<th>Automation level</th>
<th>Operation modality</th>
<th>Surgeon’s action</th>
<th>Machine action</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>fully manual</td>
<td>insertion/steering</td>
<td>none</td>
</tr>
<tr>
<td>1</td>
<td>assisted manual</td>
<td>insertion/steering</td>
<td>sensory feedback</td>
</tr>
<tr>
<td>2</td>
<td>semi automated</td>
<td>insertion only</td>
<td>steering only</td>
</tr>
<tr>
<td>3</td>
<td>fully automated</td>
<td>none</td>
<td>insertion/steering</td>
</tr>
</tbody>
</table>

Discussion, i.e., the system depicted in Fig. 1. Typically, two main control actions are used to steer the needle in combination with insertion, namely axial rotation of lateral forces applied at the needle base. The system we have considered so far is fully automated, meaning that the device performs both insertion and steering actions. There exists other levels of automation, i.e., semi-automated systems, where insertion is performed by the surgeon while steering occurs automatically (see Fig. 7(a)), and assisted manual steering, where the subsystems previously presented simply increase the surgeon’s awareness about the procedure without explicitly intervening in it (see Fig. 7(b)). The role of the surgeon and of the robot in each of these scenarios is summarized in Table 1.

5.1. Automation Level 3 - Fully automated steering

In this category the robot performs the insertion and all the steering actions [115, 116, 135, 183–194] (as depicted in Fig. 1). Once the surgeon defines the insertion location, the desired target, and anatomical obstacles, the robotic system calculates a feasible path and steers the needle to the target.

Examples of fully automated needle steering systems include the backdrivable 5-DOF needle manipulator by Bassan et al. [193], designed to orient the needle base and perform both insertion and rotation in prostate brachytherapy. A similar concept was adopted in [115, 189] and combined with intraoperative prostate tracking. Wei et al. [116] used a 6-DOF industrial robotic arm to orient and insert the needle while images of the target are acquired in 3D via a static volumetric ultrasound probe. MRI guidance [135, 184], and elastography [188] have also been integrated with steering devices.

These robots replace the surgeon and are intended to make the motions and manoeuvres very precise, which may lead to better targeting accuracy when compared to traditional manual needle steering [44]. However, integrating these systems into current clinical practice is challenging and most often several modifications to the operating room are necessary. Anecdotally, one of the earliest medical robots and in fact the first ever to remove tissue from a patient falls in this category [195]. It was used in transurethral resection of prostatic hyperplasia.

Fully-automated needle steering may represent a risk for the patient if the system becomes unstable. To manage this limitation, one can consider sharing needle insertion and steering actions between a physician and a robot, which leads to the second category of steering robots.

5.2. Automation Level 2 - Semi automated steering

In this category the robotic system acts as a needle holder that either rotates the needle axially or moves its base laterally with the physician being in charge of insertion [186, 196–201].

The latter concept has been introduced in [198]. The 4-DOF robot translates a needle guide in the $x, y$ plane allowing for precise needle insertion along the $z$ direction. It can also rotate the guide about the $x$ and $y$ axes, providing control over the needle insertion point and angle. Similarly, Fichtinger et al. [202] developed a planar 2D needle holder that provides planar motion, through which the physician manually inserts the needle into the patient, thus retaining full control and natural haptic sensing. In [199], the needle manipulator has 5-DOF, allowing for angled insertions. Other applications of co-manipulation of either needles and/or ultrasound probes can be found in biopsy [201], and teleoperated schemes [134, 136, 203–208].

Moving away from fixed robotic structures, researchers have also considered hand-held devices [43, 48, 200, 209]. Reducing the complexity of the robotic scheme not only facilitates implementation in a clinical scene, but also offers move dexterity and freedom to the surgeon. In [209], a hand-held
needle steering device actuates a stylet placed inside the shaft, changing the needle’s natural curvature to achieve a desired steering direction. In [43, 48], a hand-held apparatus for accurate steering in prostate brachytherapy is proposed to automatically rotate the needle as the surgeon manually inserts it. Such a system is compatible with contemporary operating room settings, leaving current practice intact.

5.3. Automation Level 1 - Assisted manual steering

In the third class of automation, the physician performs both the insertion and steering actions. This category combines models and controllers with a communication medium designed to provide relevant information about required steering manoeuvres. Hence, the surgeon is in full control of the procedures and may or may not perform the steering actions calculated by the control scheme (see Fig. 1). One can classify these systems into two main subcategories, namely, visual or tactile feedback devices.

In the first category, control actions are transmitted to the surgeon visually, for instance through augmented reality [210–212]. The device projects onto the patient’s skin reconstructed images of the inner body acquired from different medical imaging modalities, adding an extra layer of visual information on top of the perception of the real world in real time, making many surgical tasks simpler and safer for the surgeon. Therefore, it enhances the surgeon’s ability to visualize needles and anatomical structures within the patient’s body. Specific applications for needle guidance include arthrography [213, 214] ultrasound guided needle placement training [215], surgical laparoscopy [216], magnetic resonance guided biopsy [217], liver puncture [218] and ablation [219, 220], and computed tomography [221–223].

In the second subcategory, information is given to the surgeon in the form of tactile haptic feedback [50, 224, 225]. In [50], a wristband composed of several vibrating motors conveys haptic patters to inform about required steering manoeuvres such as needle rotation, needle base manipulation, acceleration, withdrawal etc. In [224], haptic feedback on the arm provides intuitive movement cues to assist a human during needle insertion into the chest. Although implementing such assistants would only require minor modifications to the operating room, the outcomes still depend heavy on the operator’s ability to perform steering actions.

6. Issues in Closed-Loop Needle Steering

The process of designing a closed-loop steering system involves the very same steps found in control systems generally. A typical scenario is as follows: study of the systems to be controlled, modelling, simplification, specification of performance, controller design, hardware implementation, controller tuning, and performance evaluation. Let us now discuss how each of these points fit in the context of needle steering schemes.

**Studying the system to be controlled:** In this step, one analyses all aspects of the medical procedure for which the system shall be designed. At this state, one must specify what types of sensors and actuators will be used and where they will be placed. Some applications only accept medical imaging as sensors, which comes at the cost of low sampling rate, poor image quality (particularity in the case of ultrasound), acquisition and processing delays, image registration uncertainty, and poor visualization of the task being performed. Sensorizing needles directly is the most reliable way of acquiring feedback on the tip position with the inconvenience
Table 2: Non-exhaustive summary of documented needle steering controllers and systems

<table>
<thead>
<tr>
<th>Primary author</th>
<th>Model type&lt;sup&gt;a&lt;/sup&gt;</th>
<th>2D or 3D</th>
<th>Position feedback</th>
<th>Control objective</th>
<th>Autom. level</th>
<th>Control action</th>
<th>Target error&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith (2001)</td>
<td>n.a.</td>
<td>3D US images</td>
<td>regulation</td>
<td>3</td>
<td>rotation</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Fichtinger (2002)</td>
<td>n.a.</td>
<td>3D CT images</td>
<td>regulation</td>
<td>2</td>
<td>shaft motion</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Golzman (2004)</td>
<td>FEM</td>
<td>2D camera</td>
<td>tracking</td>
<td>3</td>
<td>base motion</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Schneider (2004)</td>
<td>n.a.</td>
<td>2D US images</td>
<td>weak reg.</td>
<td>2</td>
<td>rotation</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>DiMaio (2005)</td>
<td>FEM</td>
<td>2D camera</td>
<td>tracking</td>
<td>3</td>
<td>base motion</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Fichtinger (2005)</td>
<td>n.a.</td>
<td>3D CT images</td>
<td>weak reg.</td>
<td>1</td>
<td>base motion</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>Okazawa (2005)</td>
<td>(2)</td>
<td>2D camera</td>
<td>tracking</td>
<td>2</td>
<td>curvature</td>
<td>≈1.0</td>
<td></td>
</tr>
<tr>
<td>Phee (2005)</td>
<td>FEM</td>
<td>3D US images</td>
<td>regulation</td>
<td>2</td>
<td>base motion</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>Webster (2006)</td>
<td>(1)</td>
<td>2D camera</td>
<td>model&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3</td>
<td>rotation</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td>Minhas (2007)</td>
<td>(1)</td>
<td>2D camera</td>
<td>model&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3</td>
<td>rotation</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Bassan (2009)</td>
<td>(1)</td>
<td>3D US images</td>
<td>regulation</td>
<td>3</td>
<td>rotation</td>
<td>1.45</td>
<td></td>
</tr>
<tr>
<td>Dehghan (2009)</td>
<td>FEM</td>
<td>3D US images</td>
<td>tracking</td>
<td>3</td>
<td>base motion</td>
<td>1.40</td>
<td></td>
</tr>
<tr>
<td>Kallem (2009)</td>
<td>(1)</td>
<td>3D camera</td>
<td>regulation</td>
<td>3</td>
<td>≈1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kokes (2009)</td>
<td>n.a.</td>
<td>3D MRI</td>
<td>regulation</td>
<td>1</td>
<td>rotation</td>
<td>2.54</td>
<td></td>
</tr>
<tr>
<td>Maghsoudi (2012)</td>
<td>(3)</td>
<td>2D none</td>
<td>weak. reg.</td>
<td>3</td>
<td>n.a.</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>Abayazid (2013)</td>
<td>(2)</td>
<td>2D US images</td>
<td>weak reg.</td>
<td>3</td>
<td>rotation</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>Rucker (2013)</td>
<td>(1)</td>
<td>3D magnetic</td>
<td>weak reg.</td>
<td>3</td>
<td>rotation</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Adebar (2014)</td>
<td>(1)</td>
<td>3D US images</td>
<td>tracking</td>
<td>3</td>
<td>rotation</td>
<td>1.56</td>
<td></td>
</tr>
<tr>
<td>Patil (2014)</td>
<td>(1)</td>
<td>3D US images</td>
<td>tracking</td>
<td>3</td>
<td>rotation</td>
<td>2.38</td>
<td></td>
</tr>
<tr>
<td>Vrooijink (2014)</td>
<td>(1)</td>
<td>3D US images</td>
<td>tracking</td>
<td>3</td>
<td>rotation</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Moreira (2015)</td>
<td>(1)</td>
<td>3D US images</td>
<td>tracking</td>
<td>3</td>
<td>rotation</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>Fallahi (2016)</td>
<td>(1)</td>
<td>2D US images</td>
<td>regulation</td>
<td>3</td>
<td>rotation</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Khadem (2016)</td>
<td>(1)</td>
<td>2D US images</td>
<td>tracking</td>
<td>2</td>
<td>rotation</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>Khadem (2016)</td>
<td>(3)</td>
<td>2D US images</td>
<td>regulation</td>
<td>3</td>
<td>rotation</td>
<td>1.45</td>
<td></td>
</tr>
<tr>
<td>Waine (2016)</td>
<td>(2)</td>
<td>2D US images</td>
<td>regulation</td>
<td>3</td>
<td>rotation</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Rossa (2016)</td>
<td>(2)</td>
<td>2D US images</td>
<td>weak reg.</td>
<td>2</td>
<td>rotation</td>
<td>0.44</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Needle-tissue model formulation: (1) kinematic as in Eq. (1), (2-3) mechanical as in (2) or (3)

<sup>b</sup>Error is given in millimetres

<sup>c</sup>Paper presents a model only

of having to modify the clinical practice.

**Modelling the system to be controlled:** The second step involves identifying needle-tissue interaction models that are suitable as a basis for robust control design. There are several aspects to be considered in this regards. Firstly, such a model must be controllable, observable, and provide good control performance. Secondly, it must be fast enough to be run in real time. Thirdly, the model must be a good approximation of the real system and therefore needs to limit the upper bound on the mismatch between the plant and the identified model. Notwithstanding, none of the models we have discussed in Section 2 gather all of these features concurrently, not only because uncertainties such as tissue heterogeneity cannot be accurately accounted for, but also because FEM models are not suitable for real-time control, mechanics-based models are generally limited to 2D and in most cases are not useful for control, and no closed form solution can be obtained for the nonholonomic model expressed in 3D space nor can the tissue be modelled.

Let \( p_0 \) denote an equilibrium solution of (1) corresponding to \( u_1 = u_2 = 0 \). The following observations can be made about the model in (1):

1. The systems is nonholonomic, since the distribution closure is not involutive. Using successive Lie brackets, it can be shown that the system has nonholonomy degree of 4 and the rank of the system accessibility distribution is 6.

2. The system is driftless and affine in the inputs. Considering the accessibility rank, the system is strongly accessible controllable at \( p_0 \).

3. Based on Brockett’s theory [229, 230], a necessary condition for the existence a continuously differentiable control law that asymptotically sta-
the orientation of the needle tip \([46, 47]\), the de-\nddle tip, measurement of the insertion velocity \([29]\), and/or the insertion force \([54]\). Some of these measure-
ments, however, cannot be obtained in certain ap-
lications of MIST. Other considerations concern
robustness to intrinsic measurement noise, and con-
trol constraints. Depending on the desired perfor-
mane and available measurements, one shall judi-
ciously decide between predictive, sliding mode,
duty-cycling approaches, or other methods.

Choosing the type of controller: When deciding
the type of controller to use, one must consider
what signals it needs. For instance, some controller
might require in addition to the position of the nee-
dle tip, measurement of the insertion velocity \([29]\),
the orientation of the needle tip \([46, 47]\), the de-

Simplifying the model: Needle steering models
are typically nonlinear, and in most of the cases,
a closed form solution cannot be obtained unless
some assumptions are made such as local linearisation,
and small deflections \([29, 46, 47]\). This ob-
servation, however, goes against the previous point
regarding model mismatch. The simplified model
must still provide good accuracy while allowing for
good control performance.

Specifying the desired performance: When referring
to performance, one has to specify clearly what
it means. In addition to desired outcomes such as
robustness, disturbance rejection, fast convergence
of tracking error, and accuracy, specific outcomes
for needle based interventions involve minimizing
any control effort that creates tissue trauma. Yet,
it is clear that in regards to tissue trauma, those
control objectives are contradictory, which reveals
another trade-off in the design process.

Choosing the type of controller: When deciding
the type of controller to use, one must consider
what signals it needs. For instance, some controller
might require in addition to the position of the nee-
dle tip, measurement of the insertion velocity \([29]\),
the orientation of the needle tip \([46, 47]\), the de-

Designing the controller: The next natural step
consists in integrating the different parts that com-
pose the needle steering system and implementing
the controller. Depending on the hardware char-
acteristics, control constrains must be addressed at
this state such as imposing an upper bound on the
insertion velocity and total rotation.

Designing the hardware and implementing the
controller: We have classified the steering robots
into three main categories depending on the degree
of autonomy granted to it. A given controller re-
quires a specific category of robots. For instance,
a controller that modulates the needle insertion ve-

celosity should be used with a robot that takes over
insertion \([29]\). Design considerations are also found
in the range of Cartesian motion required to steer
the needle, maximum insertion force and/or veloci-
ty, needle rotation capability, etc. It is also impor-
tant to consider factors such as mechanical, elec-
trical, and software failure, sterilization, operation,
and safety \([22]\). The latter is one of the key issues
in designing a medical robot and implementing a
controller. It can be addressed in many different
ways ranging from the design of actuators to the
use of redundant sensors.

Tuning the controller: Controller tuning refers
to the selection of parameters to ensure satisfac-
tory response. Adjusting the controller and the
model parameters to a given tissue is essential to
achieve good control performance. Choosing tun-
ing that is too slow will result in slow response and
convergence and the controller will not handle dis-
turbances from the tissue. Choosing tuning that is
too aggressive will create overshoots or lead the
system to become unstable. As an example, one
can consider the nonholonomic model that requires
the radius curvature followed by the needle tip that
varies from tissue to tissue and depends on the in-
sertion depth \([47]\). This directly influences the op-
timal control gains impacting the control effort and
tissue trauma \([46]\).

Evaluating system performance: Several indices
characterize the controller and hardware perfor-
Figure 8: The trade-off between safety, accuracy, and clinical compatibility for different levels of automation granted to needle steering robots. Automation level 0 corresponds to fully manual needle insertion and level 3 is fully automated.

performance. Typically it involves comparing the behaviour to some standard that, in the case of a medical robot, can be subjective. The most common measures of performance can be attributed to process variance (repeatability in reaching the target), set-point accuracy (accuracy in reaching the target), and minimization of the control effort (related to the effect on the surrounding tissue). If the resultant closed-loop system does not meet the specified standard, the design steps presented above can be reconsidered iteratively.

The highlighted trade-off in the design of steering robots translate into Fig. 8. Finding a good equilibrium between accuracy, patient safety, and compatibility with the clinical scene is still an open challenge that one can address in each of the design steps considered here. A variety of combinations of needle-tissue interaction models, deflection measurement modalities, and steering controllers have been proposed in order to best respond to the design requirements listed above, some of which can be seen in Table 2.

7. Concluding Remarks

In this paper, we reviewed progress made in closed-loop needle steering control. We addressed various aspects that must be considered when implementing such systems, namely modelling of needle-tissue interaction, sensing needle deflection during insertion, implementing the controller, developing the hardware and evaluating the obtained performance.

Despite significant research effort on the subject during the past 15 years that has led to a solid average in reaching targets of about 1.22 mm (see Table 2), to date there is no commercially available solution for robotics-assisted needle steering. The limitations highlighted in this paper suggest that automated needle steering will not be widespread in the clinical scene until technological and clinical limitations related to modelling and control of needle steering are overcome. Hence, robust robotics-assisted needle steering remains to date an open challenge with considerable room for improvement.

One classic example that clearly highlights the potential benefits of accurate needle steering can be found in prostate brachytherapy. Due to the currently limited accuracy of manual steering (about 5 mm [233]), brachytherapy has been limited to primarily overtreating the entire prostate even for patients with only localized prostate cancer. Improving needle targeting accuracy by means of robotic assistance can result in enhanced treatment of localized prostate cancer and, in addition, make this treatment applicable to other clinical situations [234, 235]. Accurate needle steering can also lead to more precise tissue biopsy, pinpoint drug deliver, improved ablation of lung, liver and kidney, access to deep zones in the brain, amongst many other benefits in general applications of MIST.

References


[107] W. Qiu et al., “Needle segmentation using 3D quick randomized hough transform,” in International Con-


[172] A. Majewicz et al., “Design and evaluation of duty-cycling steering algorithms for robotically-driven steerable needles,” in *IEEE International Conference on


