Visual Servoing of a Medical Ultrasound Probe for Needle Insertion

Kim Mathiassen[1], Kyrre Glette[2] and Ole Jakob Elle[1]

Abstract— Percutaneous needle insertion guided by ultrasound (US) imaging is routinely performed in hospitals today. Automating these procedures could increase placement accuracy and lower time usage of health care personnel to perform these procedures. Estimation of the needle orientation and position in the ultrasound image is important when automating these procedures. One approach for estimating the needle orientation and position is to have the needle aligned with the image plane of the ultrasound probe. Aligning the needle with the plane is difficult, even with accurate measurements and calibration of both needle and probe. In this paper we propose a visual servoing method to move the ultrasound probe using a robot to align the image plane of the probe with the needle. The method segments the needle and updates a set of visual features based on a model of the needle. A state machine is used to keep track of the alignment process, and different visual features are used to control the probe in the different states. Both simulation using simple synthetic images and experiments in a water tank are conducted to validate the proposed method. The simulation shows that the proposed method manages to align the probe plane with the needle. In the real robot experiment the alignment process is slower, but still aligns the probe plane with the needle. The method is shown to work under simplified conditions, and is a step towards a method that could be used in a clinical setting.

I. INTRODUCTION

Percutaneous needle insertion guided by ultrasound (US) imaging is routinely performed in hospitals today. Percutaneous needle insertion can be used for nerve blockades, fluid drainage, or to obtain a biopsy of a suspected malignant lesion. Another needle insertion procedure is percutaneous tumor ablation, which is a minimally invasive strategy for treatment of malignant tumors [1]. The aim is to destroy the tumors through the application of energy or chemicals. The advantages compared with surgical resection are the potential to destroy only a minimal amount of normal tissue [2] and lower cost [1]. US guidance is the most commonly used modality for thermal ablation [1].

Automation of these procedures by using robots could cut health care cost, by increasing the accuracy and decreasing the execution time of the procedure [3], [4]. Even a small decrease in time or personnel usage would have a large impact, as the procedures are very frequently performed. The automation could also lead to higher accuracy of the placement of the needle, which could lead to a higher success rate and lower mortality. When automating these procedures

it is of paramount importance that the needle is tracked while it is inserted, both to verify that the needle is placed correctly and to detect if the needle is approaching sensitive structures (e.g. major vessels, bones, other organs). Accurate tracking of a needle has been shown in [5], but one drawback with this method is that the needle must be in the plane imaged by the US probe. Therefore a method for aligning the image plane with the needle is needed.

One of the first reported systems for visual servoing control of a US probe was in [6] where a US probe was mounted on a specially made robotic system, tracking the carotid artery. The robot was controlled in 3 degrees of freedom (DoF) using visual feedback.

Extensive contributions on visual servoing of US probes have come from IRISA, INRIA Rennes-Bretagne Atlantique in France. In [7] visual servoing is performed on an egg shaped object where the robot moves the probe to view a specific cross-section of the object. This shape is common in tumors and the method could be used to hold a specific view of a tumor. A drawback with this method is that it requires a pre-operative model of the tumor. In [8], [9], [10] image moments are used to find the cross-section of an object of interest. This work solves the previous drawback as it is a model free method. An approach to have the probe stable at one view even if the patient is moving is given in [11], [12]. Here speckle information if used to track both out-of-plane and in-plane motion of the probe. In [13] the image intensities are used directly to control the probe to track a desired organ by compensating for rigid motion. All the above methods control the probe in 6 DoF.

All the previous mentioned methods are using an eye-in-hand configuration, i.e. a US probe is held by a robot and moved. In [14], [15] an eye-to-hand configuration is used, i.e. another object is moved in the US image. Here a 2 DoF robot inserts a needle into the gallbladder in order to take a biopsy. The gallbladder is tracked using an active contour model and Hough transform was used to track the needle. The robot inserting the needle is rigidly attached to the US probe, thus out-of-plane motion is avoided.

The tip of a flexible needle is tracked in [16] by having the US plane perpendicular to the needle. The visual servoing tracks the needle tip by moving the probe 2 DoF, by always keeping the tip in the image. A major drawback with this approach is that the sensitive structures in the needle path are not imaged.

This paper presents a novel visual servoing method for first locating the needle and then keeping the needle in the plane of a 2D US probe by moving the probe with a robot. Segmenting the needle is necessary to extract features that

[1]K. Mathiassen and O. J. Elle are with the Department of Informatics, University of Oslo, Gaustadalleen 23 B, N-0373 Oslo, Norway and with the Intervention Center Oslo University Hospital, Rikshospitalet, Postboks 4950 Nydalen, N-0424 Oslo, Norway kimmat@ifi.uio.no, oelle@ous-hf.no

[2]Kyrre Glette is with the Department of Informatics, University of Oslo, Gaustadalleen 23 B, N-0373 Oslo, Norway kyrrehg@ifi.uio.no
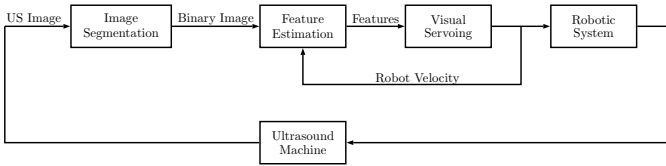
Fig. 1. An overview of the visual servoing control loop. The "Image Segmentation" block is described in Section II-A. The "Visual Servoing" and "Feature Estimation" blocks are described in Section II-C and II-D, respectively. Both these two blocks uses the visual features described in Section II-B. The blocks "Robotic System" and "Ultrasound Machine" represents physical devices, and are not a part of the algorithm.
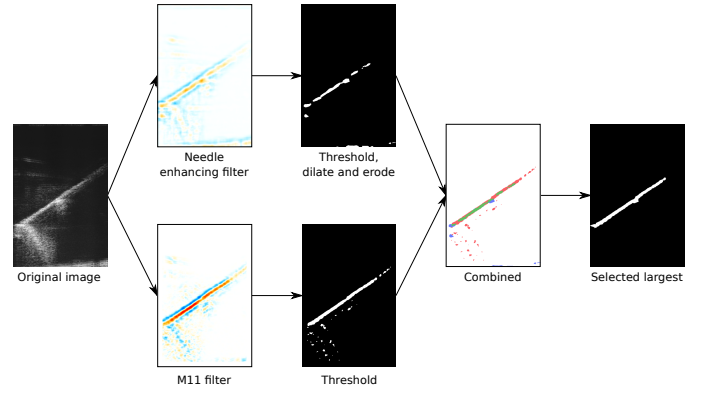

Fig. 2. The original US image is shown to the left. Next are the two filtered images where red represents positive values and blue negative values. In the middle are the thresholded images. In the combined image blue is from the needle enhancing filter, red is from the m_{11} filter and green is combined. To the right is the segmented needle.

will be used in the control. The main focus of the paper is the control approach; therefore we have chosen to simplify the segmentation step. The needle is put in a water tank, instead of using a phantom that resembles real tissue, in order to get a relatively simple segmentation problem. The main innovations are the usage of a state machine to keep track of the alignment process and to select which visual features to use and a method for stochastically resolving ambiguities in the estimated features. To our knowledge there are no methods for aligning a needle with a 2D US image plane. In this paper we use a 2D US system instead of a 3D US system, as 2D US system are still the most commonly used systems today.

Section II describes the proposed method. Section III accounts for the experiments used to validate the proposed method, while Section IV shows the results of the experiments. Section V discusses the results of the experiments and the limitations of the method. Final conclusions are drawn in Section VI.

II. ALGORITHMS

The different parts of the visual servoing method are shown in Fig. 1 and presented in this section.

A. Image segmentation

To segment the needle two filters are used. The filtered images are thresholded and then they are combined. This is shown in Fig. 2. The first filter enhances the needle. The filter uses a quadratic kernel of 29x29 px with a circle in the center, with a diameter equal to the needle diameter (19 px). The pixels in the circle have equal values and have a sum of 2.25. The border pixels have equal value and a sum of -2.25. The resulting image is thresholded at $t_1 = 50$ (of a maximum of 255). A fixed threshold is used to avoid a high increase in the threshold value when there is nothing of interest in the US image, and thus avoiding to segment background noise. Then a morphological closing operation with a circular structuring element with a diameter of 9 px is used in four iterations in order to connect together the segmented parts of the needle.

The second filter uses the image moment m_{11} on a region around each pixel in the original image. The image moment m_{11} emphasizes line structures that are angled in the image. As the needle needs to be inserted from the surface which the

probe is in contact with, the needle will always be angled. The filtered image is given below

$$I_m(x,y) = \frac{1}{4(2s+1)^2} \sum_{i=-s}^{s} \sum_{j=-s}^{s} xyI(x+i, y+j) \quad (1)$$

where (x, y) is a pixel position in the original image, s is the size of filter kernel, $\frac{1}{4(2s+1)^2}$ is a scaling factor and the double sum is the image moment m_{11}. The above equation has been implemented using dynamic programming, which reduced the complexity from a fourth order to a third order polynomial. In this paper we use $s = 9$, which means that a 19x19 region is covered for each pixel. The filtered image is thresholded at $t_2 = 50$ (of a maximum of 255).

In the next step the two segmented images are combined using a logical *OR* operation. In the combined image each segmented area must have at least one pixel from each of the segmented images, or that area will be dropped. After this step the largest segmented area is selected as the needle.

B. Needle model

Two models have been derived in order to control the robot. In the first model the needle intersects the image plane of the US probe, and thus is unaligned with the plane. Then a second model is used once the needle lies within the plane.

1) Unaligned: Fig. 3 shows the needle plane and the needle intersecting the plane. It is assumed that the segmented image contains the cross-section of the needle, although this is an approximation. The cross-section of the needle forms an ellipse in the image plane. The parameter α is the angle between the image plane and the needle axis. In order to align the needle this value should be zero. While the probe is moved to be aligned with the needle it should not disappear from the image. This is ensured by also controlling the center of the needle $\boldsymbol{p}_c$. The last parameter that should be controlled is the insertion angle of the needle θ. Assuming that the needle ellipse is segmented correctly one can calculate the four above parameters by using the area of the ellipse.
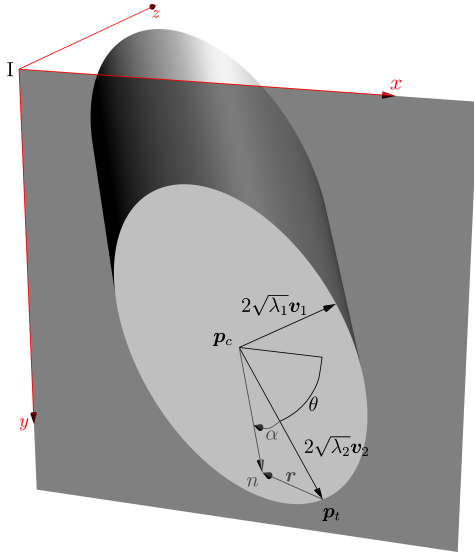
Fig. 3. A needle cross-section for modeling the unaligned needle. $\boldsymbol{p}_c$ is the center of the needle. $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are the directions of the minor and major semi-axis, while $2\sqrt{\lambda_1}$ and $2\sqrt{\lambda_2}$ are the length of the minor and major semi-axis, respectively. $\boldsymbol{n}$ is the needle direction in space and $\boldsymbol{v}_2$ corresponds to the needle direction in the image plane. θ is the angle between the x-axis and the needle direction in the image plane, and α is the angle between the image plane and the needle direction $\boldsymbol{n}$. r is the needle radius. $\boldsymbol{p}_t$ is the location where the major semi-axis intersects the ellipse, and is used in the *Aligned* model as an estimate for the needle tip location.

The position of the center of the needle $\boldsymbol{p}_c = [x_c \ y_c]^T$ can be found by calculating the centroid of the ellipse, using image moments [8]

$$x_c = \frac{m_{10}}{m_{00}} \qquad y_c = \frac{m_{01}}{m_{00}} \qquad (2)$$

where m_{ij} are image moments. The image moments are calculated for the segmented area selected in the previous section. The eigenvalues of the covariance matrix

$$cov(I) = \begin{bmatrix} \frac{m_{20}}{m_{00}} - x_c^2 & \frac{m_{11}}{m_{00}} - x_c y_c \\ \frac{m_{11}}{m_{00}} - x_c y_c & \frac{m_{02}}{m_{00}} - y_c^2 \end{bmatrix} \qquad (3)$$

are proportional to the major and minor axis of the ellipse. The orientation of the ellipse (which corresponds to the insertion angle) is given by

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \qquad (4)$$

where μ_{11} is the off-diagonal elements of $cov(I)$ and μ_{20} and μ_{02} are the diagonal elements of $cov(I)$.

The major and minor semi-axis of the ellipse in Fig. 3 are equal to $2\sqrt{\lambda_2}$ and $2\sqrt{\lambda_1}$, respectively, where λ_1 and λ_2 are the eigenvalues of $cov(I)$. The directions of the major and minor semi-axis are given by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, which are the eigenvectors of $cov(I)$ and have unit length. The minor semi-axis ($2\sqrt{\lambda_1}$) corresponds to the radius of the needle. The major semi-axis ($2\sqrt{\lambda_2}$) equals the distance between the points $\boldsymbol{p}_c$ and $\boldsymbol{p}_t$. By using the triangle formed by $\boldsymbol{v}_2$, r and $\boldsymbol{n}$ in Fig. 3 we get α as

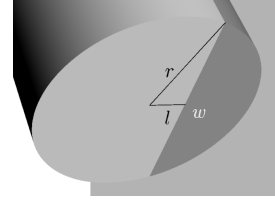$$\alpha = \arcsin\sqrt{\frac{\lambda_1}{\lambda_2}} \qquad (5)$$

Fig. 4. A needle cross-section for model of the aligned needle. r is the radius of the needle, l is the length from the needle center to the image plane and w is width of the needle in the image plane.

2) Aligned: When the needle is aligned to the image plane, a new model is needed. The visual feature θ is kept, as we still wanted to control the insertion angle. Now with the needle aligned to the plane, we would like to control the z-direction of the US image to ensure that the needle lies in the middle of the plane. To do this the offset l in Fig. 4 is modeled by using the radius r of the needle and the width w of the needle in the US image. This yields

$$l = \sqrt{r^2 - \frac{w^2}{4}} \qquad (6)$$

where r is assumed to be known. The segmented needle is approximated to be an ellipse. The parameter w is estimated as the area of the segmented needle (m_{00}) divided by the approximated length of the needle, which is $4\sqrt{\lambda_2}$ as this is the length of the major axis. This yields

$$w = \frac{m_{00}}{4\sqrt{\lambda_2}} \qquad (7)$$

The above formula yields more stable results than using $4\sqrt{\lambda_1}$, which is the length of the minor axis.

The last visual feature is the position of the needle tip. The tip position is approximated to be

$$\boldsymbol{p}_t = \boldsymbol{p}_c + 2\sqrt{\lambda_2}\boldsymbol{v}_2 \qquad (8)$$

where $\boldsymbol{p}_c$ is the centroid of the segmented needle, $2\sqrt{\lambda_2}$ is the length of major axis and $\boldsymbol{v}_2$ is the direction of the major axis.

C. Visual servoing

In order to control the probe using a visual servoing scheme the interaction matrices for the states are needed [17]. The interaction matrix describes the relationship between the visual features $\boldsymbol{s}$ and the velocity of the US probe $\mathbf{v} = [\boldsymbol{v}^T \ \boldsymbol{\omega}^T]^T$ with $\boldsymbol{v}$ being the linear and $\boldsymbol{\omega}$ angular velocity of the probe. The relationship is given by $\dot{\boldsymbol{s}} = \boldsymbol{L}\mathbf{v}$.

The set of visual features which are controlled depend on the state of the controller. The control state machine is given in Fig. 5, and this section describes all the states. All the states use the control law [17]

$$\mathbf{v} = -\boldsymbol{K}\boldsymbol{L}^+(\boldsymbol{s} - \boldsymbol{s}^*) \qquad (9)$$

where $\boldsymbol{s}^*$ is the desired visual feature and $\boldsymbol{K}$ is the gain matrix, specifying a gain for each feature on its diagonal.

First we will find the interaction matrices for the visual features, and in the later section we will modify the interaction matrices to obtain the wanted behavior in each

1 - Location unknown

$v_z = \pm 0.01$

$m_{00} > 0$

2 - Estimate direction

$\begin{bmatrix} \dot{p}_c \end{bmatrix} = \begin{bmatrix} L_z \end{bmatrix} \mathbf{v}$

$z_1 = [\tilde{\alpha}_1 \ \tilde{\theta}_1 \ \tilde{p}_c]$

$\frac{\lambda_2}{\lambda_1} < kr_0$

3 - Unaligned

$\begin{bmatrix} \dot{\alpha} \\ \dot{p}_c \end{bmatrix} = \begin{bmatrix} L_\alpha \\ L_{p_c} \end{bmatrix} \mathbf{v}$

$z_1 = [\tilde{p}_c]$

$v_z < 0.001$

$\frac{\lambda_2}{\lambda_1} \le r_\lambda$

4 - Near aligned

$\begin{bmatrix} \dot{\alpha} \\ \dot{\theta} \\ \dot{p}_c \end{bmatrix} = \begin{bmatrix} L_\alpha \\ L_\theta \\ L_{p_c} \end{bmatrix} \mathbf{v}$

$z_1 = [\tilde{\alpha}_2 \ \tilde{\theta}_2 \ \tilde{p}_c]$

$\frac{\lambda_2}{\lambda_1} > r_\lambda$

$S_I = 0$

5 - Aligned

$\begin{bmatrix} \dot{l} \\ \dot{\theta} \\ \dot{p}_t \end{bmatrix} = \begin{bmatrix} L_l \\ L_\theta \\ L_{p_t} \end{bmatrix} \mathbf{v}$

$z_2 = [\tilde{l} \ \tilde{\theta}_2 \ \tilde{p}_c]$
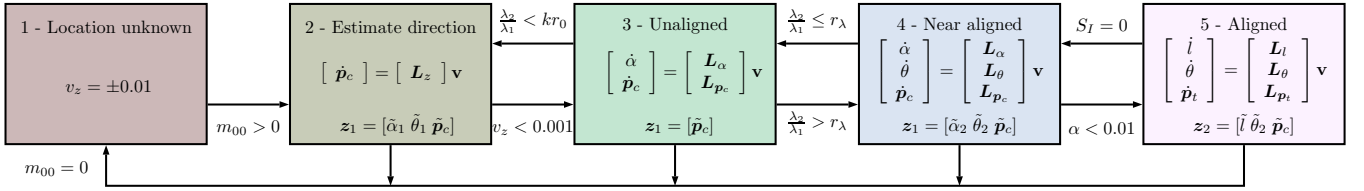
$\alpha < 0.01$

$m_{00} = 0$

Fig. 5. Control state machine showing the visual servoing features and matrices used for controlling the features in each state. Which features that are measured in each state are also shown by the z_1 and z_2 vectors, where z_1 is used as a measurement update for an extended Kalman filter.

state. The in-plane motion of a point is given in [8], and covers the relationship between $\dot{p}_c$ and v_x, v_y and ω_z. The relationship between $\dot{p}_c$ and v_z is straightforward to find using trigonometry. We use the needle direction $\mathbf{n}$ to derive the relationships for remaining degrees of freedom. The needle direction $\mathbf{n}$ is the first column vector of $\mathbf{R}_z(\theta)\mathbf{R}_y(\alpha)$, where $\mathbf{R}_y$ and $\mathbf{R}_z$ are simple rotation around the y and z axes, respectively. This yields $\mathbf{n} = [c_\alpha c_\theta, \ c_\alpha s_\theta, \ -s_\alpha]^T$, where s_x and c_x denotes $\sin x$ and $\cos x$. Then the $\mathbf{n}$ vector is decomposed into the xz-plane and yz-plane, denoted $\mathbf{n}_{xz}$ and $\mathbf{n}_{xy}$. The angle ϕ is defined as the angle between the x-axis and $\mathbf{n}_{xz}$ while the angle ψ is the angle between the y-axis and $\mathbf{n}_{yz}$. Using the dot and cross product on these vectors yield

$$\tan \phi = -\frac{\tan \alpha}{\cos \theta} \qquad \tan \psi = \frac{\tan \alpha}{\sin \theta} \qquad (10)$$

Using the definition of angular velocity yields $\omega_x = \frac{v_z^\omega}{y_c}$ and $\omega_y = \frac{v_z^\omega}{x_c}$, where v_z^ω is the velocity in the z-direction caused by the rotation. Using the equation above, and the relations $\tan \phi = \frac{v_z^\omega}{v_x^\omega}$ and $\tan \psi = \frac{v_z^\omega}{v_y^\omega}$ we find the relationships between $\dot{p}_b$ and ω_x and ω_y. This yields the interaction matrix

$$\mathbf{L}_{p_c}^\star = \begin{bmatrix} -1 & 0 & \frac{\cos \theta}{\tan \alpha} & -\frac{y_c \cos \theta}{\tan \alpha} & -\frac{y_c \cos \theta}{\tan \alpha} & y_c \\ 0 & -1 & \frac{\sin \theta}{\tan \alpha} & \frac{x_c \sin \theta}{\tan \alpha} & \frac{x_c \sin \theta}{\tan \alpha} & -x_c \end{bmatrix} \quad (11)$$

To find the interaction matrix for θ and α with regards to ω_x and ω_y we take the time derivative of (10) and solve for $\dot{\theta}$ and α. Only θ is affected by ω_z, and neither θ nor α is affected by translations. This yields

$$\mathbf{L}_{\alpha,\theta}^\star = \begin{bmatrix} 0 & 0 & 0 & c_\alpha^2 s_\theta^3 + s_\alpha^2 & -c_\alpha^2 c_\theta^3 + s_\alpha^2 & 0 \\ 0 & 0 & 0 & -\frac{c_\theta}{t_\alpha}(s_\theta^2 + t_\alpha^2) & -\frac{s_\theta}{t_\alpha}(c_\theta^2 + t_\alpha^2) & -1 \end{bmatrix} \quad (12)$$

The last visual feature l is the displacement in the z-direction when the needle is aligned, thus the interaction matrix is

$$\mathbf{L}_l^\star = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \quad (13)$$

1) Needle location unknown: In this initial state the probe is moved back and forth in the z-direction to locate the needle. The needle is detected when the needle mass m_{00} is greater than zero, and then the state changes to the *Estimate direction* state. This state is not visual servo controlled, as there are no features in the image. If the needle is lost ($m_{00} = 0$); the state machine returns to the initial state.

*Denotes that the matrix is an interaction matrix.

2) Estimate direction: When the needle is detected, an initial estimate of the needle orientation is needed. This state estimates the orientation by moving the probe in the z-direction, and using the center of the needle in the image p_c as a visual feature. Equation (11) is modified to only allow movement in the z-direction

$$\mathbf{L}_z = \begin{bmatrix} 0 & 0 & \operatorname{sign} \alpha \cos \theta & 0 & 0 & 0 \\ 0 & 0 & \operatorname{sign} \alpha \sin \theta & 0 & 0 & 0 \end{bmatrix} \quad (14)$$

where $\operatorname{sign} \alpha$ returns the sign of α. Note that $\frac{1}{\tan \alpha}$ is approximated by $\operatorname{sign} \alpha$ to avoid $\mathbf{L}_z$ becoming singular, assuming $\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

When the center of the needle p_c is close to the desired location in the image p_c^* the state changes to the *Unaligned* state. The velocity in the z-direction is used as a measure for closeness, as it is proportional to the position error. In the estimation of α and θ it is required that the probe moves a certain distance between the frames. Therefore the velocity is used to trigger the state change, rather than the position.

3) Unaligned: As the orientation of the needle was found in the previous state, the probe can be rotated towards the needle. Although θ also has an initial estimate we choose to only control p_c and α in this state, because the angle between the needle and the image plane is large. The large angle makes the estimates of α and θ from (5) and (4) inaccurate.

The motion of p_c is limited to in-plane motions, yielding[8]

$$\mathbf{L}_{p_c} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & y_c \\ 0 & -1 & 0 & 0 & 0 & -x_c \end{bmatrix} \quad (15)$$

For the visual feature α we choose to rotate around the point p_c and use the vector $\mathbf{v}_1$ as rotation axis (i.e. $\dot{\alpha} = [(\mathbf{v}_1^\times \mathbf{p}_c)^T \ \mathbf{v}_1^T]\mathbf{v}$), rather than using (12). This yields

$$\mathbf{L}_\alpha' = \begin{bmatrix} 0 & 0 & -v_{1y}x_c + v_{1x}y_c & v_{1x} & v_{1y} & 0 \end{bmatrix} \quad (16)$$

where v_{1x} and v_{1y} are the x- and y-component of $\mathbf{v}_1$. The benefit from this approach is that the rotation only affects the parameter α. The downside is that rotation around the x-axis sometimes is unwanted. If only rotation around the y-axis is wanted the following matrix may be used

$$\mathbf{L}_\alpha = \begin{bmatrix} 0 & 0 & -x_c \cos \theta & 0 & -\cos \theta & 0 \end{bmatrix} \quad (17)$$

The above matrix will also affect θ, which is undesirable in a control perspective.

The state changes to the *Near aligned* state when α is small. This is determined by the ratio between λ_2 and λ_1, and when the ratio is above r_λ the state changes.

4) Near aligned: In this state α, θ and p_c are controlled. When the ratio between λ_2 and λ_1 is large the estimation of both α and θ becomes more stable. This enables direct estimation of all the controlled features in this state.

The matrices for controlling α and p_c are given in the previous section. The matrix for controlling θ depends on matrix used for controlling α. If L'_α is used the matrix is given as [8]

$$L'_\theta = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \tag{18}$$

If L_α is used, the rotation around the y-axis also affects θ. Taking the time derivative of (10) and inserting for $\dot\alpha$ yields an additional term, and the control matrix becomes

$$L_\theta = \begin{bmatrix} 0 & 0 & 0 & 0 & -\tan\alpha\sin\theta & -1 \end{bmatrix} \tag{19}$$

When the α becomes close to zero, the state changes to the *Aligned* state. If for some reason the ratio between λ_2 and λ_1 becomes lower than r_λ, the state changes to the *Unaligned* state in order to align the needle closer to image plane.

5) Aligned: In this state the needle is aligned with the image plane, and the visual feature l from the aligned needle model is used. The interaction matrix in (13) is modified to handle changing directions of l. The new matrix is

$$L_l = \begin{bmatrix} 0 & 0 & -d_l & 0 & 0 & 0 \end{bmatrix} \tag{20}$$

where d_l is either 1 or -1, depending on the sign of l, and is estimated along with the features.

The second new feature in this state is the position of the needle tip p_t. As this feature is based on p_c by an offset, the same interaction matrix is used, thus $L_{p_c} = L_{p_t}$. This is an approximation, as the time derivative of the term $2\sqrt{\lambda_2}v_2$ in (8) is not zero. The last controlled feature is θ, using the interaction matrix from (18).

When the needle is aligned the segmented needle is always touching at least one side of the US image. S_I denotes how many sides the needle is touching, and if this value drops to zero the state changes to the *Near aligned* state.

D. Visual feature estimation

The visual features estimation is described in this section, and an overview of the features estimated in each state is given in Fig. 5.

1) Stochastic ambiguity resolver: Both α from (5) and and l from (6) are ambiguous. α as both λ_1 and λ_2 are always positive interdependently of the sign of α, and l as the square root always yields a positive value, while l may be both positive and negative. Therefore Bayesian inference is used to find the sign of α and l, denoted d_α and d_l and has either the value 1 or -1. We generalize the method by using f as either α or l, d as the direction and the error $e = f - f^*$. We approximate the Bayesian inference using the following equation

$$P(d|\dot e) \approx \frac{P(\dot e|d)P(d) + N}{\sum_d P(\dot e|d)P(d) + N} \tag{21}$$

where N is a small value added in order to even out the probabilities for the two directions over time, if there is no

change in f. $P(d)$ is modeled to be $P(d|\dot e)$ from the previous time step. If $\dot e$ is negative, the feature error is decreasing and it is likely that the direction is correct. Therefore $P(\dot e|d)$ is given a high probability P_H if d equals the current direction $\hat d$, and a low probability P_L if not. If $\dot e$ is positive the situation is reversed, and the feature error is increasing. Then if $d = \hat d$, $P(\dot e|d)$ is given a low probability P_L, and if not it is given a high probability P_H.

The probabilities P_H and P_L are modeled as

$$P_H = 1 - \tfrac{1}{2}e^{-k|\dot e|} \qquad P_L = \tfrac{1}{2}e^{-k|\dot e|} \tag{22}$$

A high change in the error e over time indicates a high probability that the movement is either correct or incorrect. With slower change the probability is less. With no change it is impossible to know if the movement is in the correct direction or not, and $P_H = P_L$.

The current direction $\hat d$ changes when the probability of the direction being wrong ($P(d \neq \hat d|\dot e)$) is greater than 0.6. This is to avoid the direction oscillating when the probabilities become close to 0.5. The method is initialized with the positive direction having a probability of 0.75 and the negative direction a probability of 0.25.

2) Extended Kalman filter: For the states 2-4 an extended Kalman filter is used, as it is not possible to extract all the features from the image at all times. The extended Kalman filter uses a matrix L depending on the control state as control input, given as $\hat s_{k+1} = \hat s_k + T L(\hat s_k, \mathbf v_k)\mathbf v_k$ where the state $\hat s_k = [\hat\alpha_k \ \hat\theta_k \ \hat p_c^T]^T$ contains the estimated visual features, and T is the period.

3) Estimate direction: In the *Estimate direction* state the initial estimate of α and θ needs to be found. By using the probe movement one can determine these features. The movement in the z-direction of the probe is measured by the robot, and the change in position is denoted Δz. The change in the in-plane center of the needle position p_c is denoted Δx and Δy. The position change is calculated using the current position and the position five images ago, for increased robustness. Using the position changes yields the estimates

$$\tilde\alpha_1 = \arctan\frac{\Delta z}{\sqrt{\Delta x^2 + \Delta y^2}} \qquad \tilde\theta_1 = \arctan\frac{\Delta y}{\Delta z} \tag{23}$$

Although $\arctan$ is used to find α, the sign is not necessarily correct, as the denominator is always positive. Given our definition of α multiplying $\tilde\alpha_1$ with $-\text{sign}\,\tilde\theta_1\,\text{sign}\,\Delta y$ will yield the correct sign.

The extended Kalman filter is using $L = [\mathbf 0_{2\times6}^T \ L_z^T]^T$ where $\mathbf 0_{2\times6}$ is a 2×6 matrix of zeros and L_z is from (14). The Kalman filter is updated using $\tilde\alpha_1$, $\tilde\theta_1$ and p_c from (2).

4) Unaligned: In this state the extended Kalman filter uses the interaction matrix $L = [L_\alpha^T \ L_\theta^T \ L_{p_c}^T]^T$, and the measurement update is p_c. Note that in this state both α and θ are estimated solely based on the extended Kalman filter, and not updated based on the images.

5) Near aligned: The method in Section II-D.1 is in this state used to resolve the ambiguity of α. The method needs $\dot\alpha$, which is estimated as $\dot\alpha = \frac{\hat\alpha_k - \hat\alpha_{k-1}}{T}$. Using the current direction $\hat d_\alpha$ one gets the estimated angle $\tilde\alpha_2 = \hat d_\alpha\alpha$.
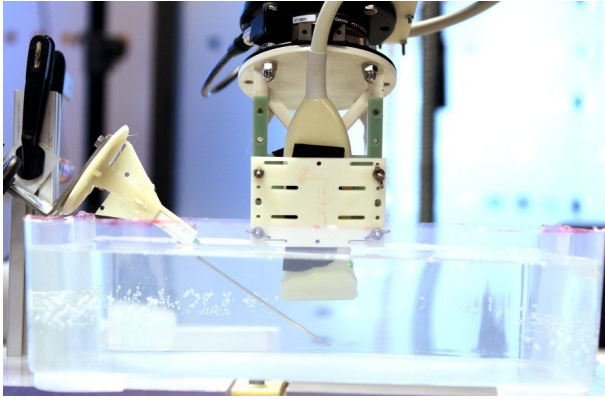
Fig. 6. The experimental setup consists of a water tank with a biopsy needle. The robot is holding the US probe in the water tank.

The extended Kalman filter uses the same $\boldsymbol{L}$ matrix as in the *Unaligned* state, but the update is using $\tilde{\alpha}_2$, θ from (4) and $\boldsymbol{p}_c$. If the segmented needle is touching one or more of the image borders the needle model is invalid, and no measurement updates are done.

6) Aligned: In order to solve the ambiguity of l, $\dot{l}$ is needed. Taking the time derivative of (6) yields $\dot{l} = -\frac{1}{4}\frac{w\dot{w}}{l}$ where w is obtained from (7), l from (6) and $\dot{w}$ is estimated as $\frac{w_k - w_{k-1}}{T}$. Using the current direction $\hat{d}_l$ the estimate of l is $\tilde{l} = \hat{d}_l l$

The estimated visual features $\tilde{l}$, θ and $\boldsymbol{p}_t$ (from (8)) are used directly in this state, as they are extracted from the image at all times. They are given in $\boldsymbol{z}_2$ in Fig. 5 to indicate which features that are obtained from the US image, although they are not used in any stochastic filter.

III. EXPERIMENTS

A. Simulation using synthetic images

In the simulation 332x502 8-bit gray scale synthetic US images were created, using a pixel size of 0.1 mm/px. The position in the image that intersects the center of the needle was given the value 255. The gray value was scaled down depending on the distance from the needle center. The needle was modeled to have a radius of 1 mm, and the pixel value at the needle border was set to 15. The rest of the image was black (i.e. having a value of zero). The image and needle size resembles the ones used in the experiment using a physical robot, presented in the next section. The desired visual features $\boldsymbol{s}^*$ were $\alpha^* = 0°$, $\theta^* = 35°$, $\boldsymbol{p}_c^* = [150\ 350]^T$px, $l^* = 0$ and $\boldsymbol{p}_t^* = [200\ 400]^T$px. The gains used were 0.5, $[0.2\ 1\ 1]$, $[0.2\ 0.2\ 0.4\ 0.4]$ and 0.3 for the states *Estimate direction*, *Unaligned*, *Near aligned* and *Aligned*, respectively. The vectors represent the diagonal elements of the gain matrix $\boldsymbol{K}$.

In the simulation ten initial needle poses were used, and the image delay were ranging from no delay to four frames delay. The gain was multiplied by the vales 1, 2, 4, 8 and 16 compared to the values used in the physical experiment.

B. Experiments using a physical robot

The proposed control approach was verified using a robot; the experimental setup is shown in Fig. 6. The robot holds an ultrasound probe, and a biopsy needle was fixed to the tank with the needle submerged. The probe was positioned close to the needle, and the proposed method was started. The desired visual features and gains were the same as in the simulation.

The robot used was an UR5 from Universal Robots (Denmark). The US Machine was a System Five from GE Vingmed (Norway). The US image stream was acquired using a video VGA2Ethernet frame grabber from Epiphan (USA) at a rate of 20 frames per second. An Angiotech ProMag 14 GA x 10 cm biopsy needle from Medical Device Technologies Inc. (USA) was used. The needle was sanded using a sanding paper for better visibility.

The images acquired from the US machine were 332x502 8-bit gray scale images. A FLA 1A probe was used with the RF program. The US machine settings were: frequency 10.0 MHz, depth 6 cm, compression 16, dynamic range 6, rejection 0, power -2 dB and fps 32.7.

IV. RESULTS

A. Simulation using synthetic images

The errors of the visual features are shown in Fig. 7. All the feature errors approaches zero, except α which has an offset of -0.87°. Different controller gain and image delay is explored in Fig. 9, and it shows the average time used from the moment the method enters the *Estimate direction* state until the needle is 1 mm from the target location.

B. Experiment using a physical robot

The errors of the visual features during the experiment are shown in Fig. 8. It is worth noting that α changed to a negative value twice. First between 6.9 s and 7.4 s, and then also between 10.2 s and 21.1 s. Between 32.2 s and 57.1 s the method uses only the extended Kalman filter to predict the visual features. At 57.1 s the state machine enters the *Aligned* state, and the visual features are again extracted from the image. The median time used for each control iteration was 45.15 ms, where 45 ms was used on the segmentation and 0.15 ms on the visual servoing control.

V. DISCUSSION

A. Simulation using synthetic images

The simulation creates a very simplified synthetic US image, and the simulation is only meant to validate the theoretical framework of the control scheme. Fig. 7 shows the ideal execution of the visual servoing method.

We have explored the effect of higher gains and image delay using simulations (see Fig. 9). Increasing the image acquisition delay caused a slight increase of the completion time. Multiplying the gain by two approximately halves the completion time. Increasing both the gain and the image delay makes the method unstable, and unable to complete the task. For instance, with the gain multiplied by two and
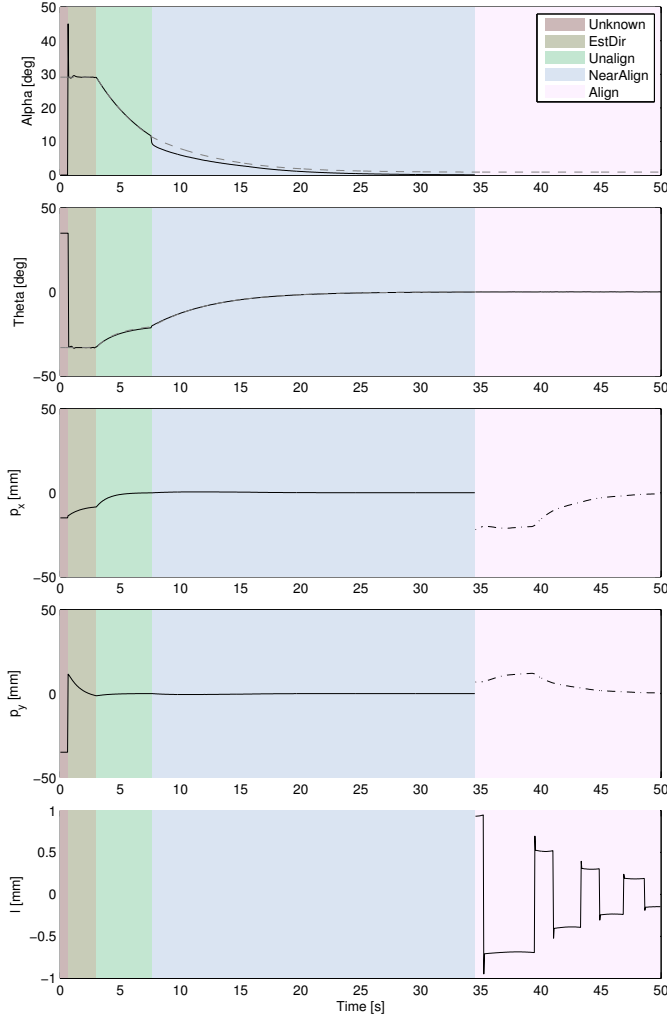
Fig. 7. The figures show the visual feature errors during one simulation. $\boldsymbol{p}_c$ and $\boldsymbol{p}_t$ share graph and are drawn by solid and dotted line, respectively. For α and θ the true values are shown as gray dotted lines. The states are shown in color in the background, and the same colors as in the state diagram are used.

Fig. 8. The figures show the visual feature errors during the experiment. $\boldsymbol{p}_c$ and $\boldsymbol{p}_t$ share graph and are drawn by solid and dotted line, respectively. The states are shown in color in the background, and the same colors as in the state diagram are used.

a four image delay only two of ten initial poses find the needle tip. With a gain multiplier of 16 the average time for completing the task is 3.0 s and 4.3 s for zero and one image delay, respectively. With more delay the method is unable to complete the task. This shows that the gain could be set much higher if the image acquisition delay was reduced.

B. Experiment using a physical robot

The experiment for validating the proposed method is not as challenging as using the method on a real tissue. The water makes the needle stand out in the US image, and the segmentation of the needle is not very difficult. Another advantage of using water is that there are almost no cases where the segmentation step finds something that is not the needle. The focus of this paper is to validate the control approach, not the segmentation of the needle. And the aim of the experiment is to show that the control approach is feasible under fairly ideal conditions. As the segmentation
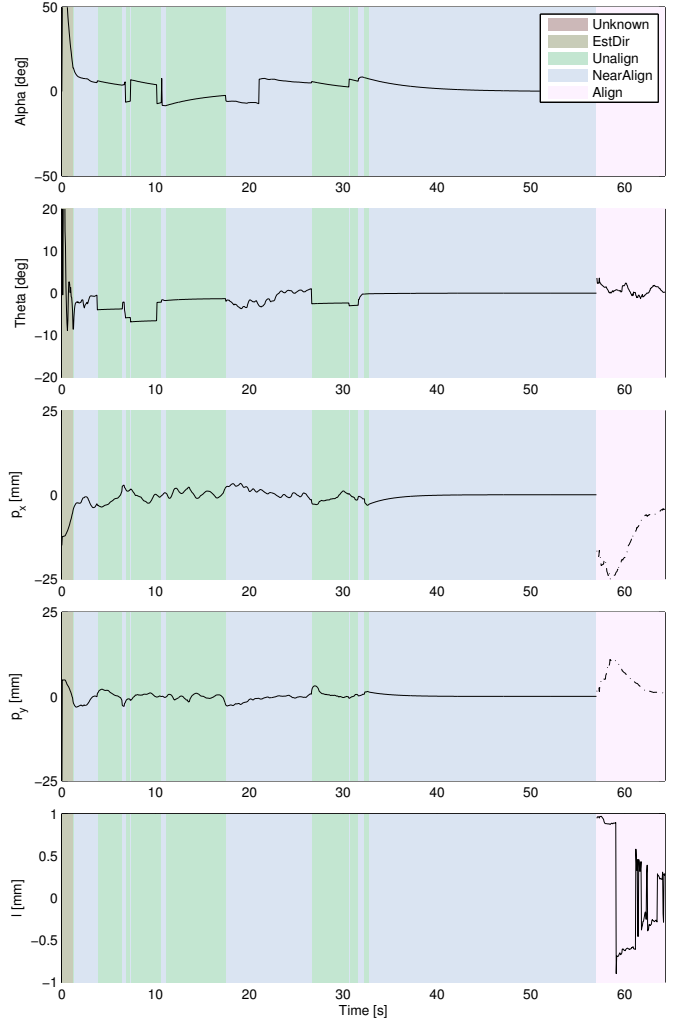
and control are two separate steps, the segmentation step should be improved when using a more challenging phantom.

The experiment uses almost 60 s to align the needle, which is too long. In the experiment the frame grabber introduced a delay of approximately four frames (200 ms). This delay deteriorates the performance of the control, and it is necessary to have a low gain to avoid unstable control. If the delay in the image acquisition can be reduced, higher gains may be used. Simulations indicate that the gain could be multiplied by between 8 and 16, thus reducing the time to between approximately five to ten seconds.

In the video attachment one can see an appearing and disappearing artifact on the needle. The appearance and disappearance and of the artifact causes the ratio between λ_2 and λ_1 to change rapidly, and is one likely cause for the oscillation between *Unaligned* and *Near aligned* states.

One shortcoming with the method is that it is not possible to update the estimated α in the *Unaligned* state and in the *Near aligned* state when the needle is touching one of the
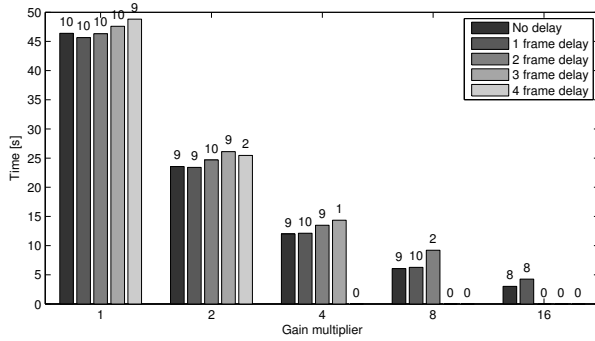
Fig. 9. Average time taken from the moment the method enters the *Estimate direction* state until the needle is 1 mm from the target location. Ten different initial needle poses are simulated, and the diagram reports the average time used grouped by gain and image delay. The numbers of successful completions are reported on top of the bars (ten is the maximum).

image borders. In the *Unaligned* state α is not extracted from the images as the angle is too large to get accurate measurements based on the needle model. This is generally not a problem as the feature was estimated in the *Estimate direction* before entering the *Unaligned* state. For the *Near aligned* state the needle model is not valid when some parts of the needle are outside the image. This typically occurs when the needle and image plane is close to being aligned.

The method may process an image stream with a frame rate of up to 22 frames per second. Nearly all of the time is used on segmentation of the US image, and if a higher frame rate is required the segmentation should be optimized.

VI. CONCLUSION

This paper has demonstrated a method for visual servoing of a US probe to align the image plane with a needle. Simulation has shown that the theoretical framework and state machine function as intended under ideal conditions. An experiment using a robot showed that the method works, although less optimally than in the simulation. Especially the state changes between the *Unaligned* and *Near aligned* states leave room for improvement, as the probe rotates in the wrong direction for some period of time. Even though the method has some drawbacks, we have shown that the method potentially could solve the problem of aligning an image plane with a needle.

Future work includes conducting experiments where the conditions are closer to a clinical setting and having the needle move slightly during the visual servoing. In addition a faster image acquisition is required for faster convergence of the visual features.

ACKNOWLEDGMENT

REFERENCES

[1] S. N. Goldberg, G. S. Gazelle, and P. R. Mueller, "Thermal ablation therapy for focal malignancy," *American Journal of Roentgenology*, vol. 174, no. 2, pp. 323–331, feb 2000.

[2] M. Ahmed, C. L. Brace, F. T. Lee, and S. N. Goldberg, "Principles of and advances in percutaneous ablation," *Radiology*, vol. 258, no. 2, pp. 351–369, 2011.

[3] M. Bonfe, F. Boriero, R. Dodi, P. Fiorini, A. Morandi, R. Muradore, L. Pasquale, A. Sanna, and C. Secchi, "Towards automated surgical robotics: A requirements engineering approach," in *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on*. IEEE, 2012, pp. 56–61.

[4] R. Muradore, P. Fiorini, G. Akgun, D. E. Barkana, M. Bonfe, F. Boriero, A. Caprara, G. D. Rossi, R. Dodi, O. J. Elle, F. Ferraguti, L. Gasperotti, R. Gassert, K. Mathiassen, D. Handini, O. Lambercy, L. Li, M. Kruusmaa, A. O. Manurung, G. Meruzzi, H. Q. P. Nguyen, N. Preda, G. Riolfo, A. Ristolainen, A. Sanna, C. Secchi, M. Torsello, and A. E. Yantac, "Development of a cognitive robotic system for simple surgical tasks," *International Journal of Advanced Robotic Systems*, vol. 12, 2015.

[5] K. Mathiassen, D. Dall'Alba, R. Muradore, P. Fiorini, and O. J. Elle, "Real-time biopsy needle tip estimation in 2d ultrasound images," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, May 2013, pp. 4363–4369.

[6] P. Abolmaesumi, S. Salcudean, W.-H. Zhu, M. Sirouspour, and S. Di-Maio, "Image-guided control of a robot for medical ultrasound," *IEEE Trans. Robot. Automat.*, vol. 18, no. 1, pp. 11–23, 2002.

[7] W. Bachta and A. Krupa, "Towards ultrasound image-based visual servoing," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* Institute of Electrical & Electronics Engineers (IEEE), 2006, pp. 4112–4117.

[8] R. Mebarki, A. Krupa, and F. Chaumette, "Image moments-based ultrasound visual servoing," in *2008 IEEE International Conference on Robotics and Automation*. Institute of Electrical & Electronics Engineers (IEEE), may 2008, pp. 113–119.

[9] ——, "Modeling and 3d local estimation for in-plane and out-of-plane motion guidance by 2d ultrasound-based visual servoing," in *2009 IEEE International Conference on Robotics and Automation*. Institute of Electrical & Electronics Engineers (IEEE), may 2009.

[10] ——, "2-d ultrasound probe complete guidance by visual servoing using image moments," *IEEE Trans. Robot.*, vol. 26, no. 2, pp. 296–306, apr 2010.

[11] A. Krupa, G. Fichtinger, and G. D. Hager, "Full motion tracking in ultrasound using image speckle information and visual servoing," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*. Institute of Electrical & Electronics Engineers (IEEE), apr 2007, pp. 2458–2464.

[12] ——, "Real-time motion stabilization with b-mode ultrasound using image speckle information and visual servoing," *The International Journal of Robotics Research*, vol. 28, no. 10, pp. 1334–1354, 2009.

[13] C. Nadeau and A. Krupa, "Intensity-based direct visual servoing of an ultrasound probe," in *2011 IEEE International Conference on Robotics and Automation*. Institute of Electrical & Electronics Engineers (IEEE), may 2011, pp. 5677–5682.

[14] J. Hong, T. Dohi, M. Hashizume, K. Konishi, and N. Hata, "An ultrasound-driven needle-insertion robot for percutaneous cholecystostomy," *Physics in Medicine and Biology*, vol. 49, no. 3, pp. 441–455, jan 2004.

[15] J.-S. Hong, T. Dohi, M. Hasizume, K. Konishi, and N. Hata, "A motion adaptable needle placement instrument based on tumor specific ultrasonic image segmentation," in *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002*. Springer Science + Business Media, 2002, pp. 122–129.

[16] G. J. Vrooijink, M. Abayazid, and S. Misra, "Real-time three-dimensional flexible needle tracking using two-dimensional ultrasound," in *2013 IEEE International Conference on Robotics and Automation*. Institute of Electrical & Electronics Engineers (IEEE), may 2013, pp. 1680–1685.

[17] F. Chaumette and S. Hutchinson, "Visual servo control, part i: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.