2018 IEEE International Conference on Robotics and Automation (ICRA)
May 21-25, 2018, Brisbane, Australia

An Observer-Based Fusion Method using Multicore Optical Shape Sensors and Ultrasound Images for Magnetically-Actuated Catheters

Alper Denasi, Fouzia Khan, Klaas Jelmer Boskma, Mert Kaya, Christoph Hennersperger, Rüdiger Göbl, Maria Tirindelli, Nassir Navab and Sarthak Misra

***Abstract*— Minimally invasive surgery involves using flexible medical instruments such as endoscopes and catheters. Magnetically actuated catheters can provide improved steering precision over conventional catheters. However, besides the actuation method, an accurate tip position is required for precise control of the medical instruments. In this study, the tip position obtained from transverse 2D ultrasound images and multicore optical shape sensors are combined using a robust sensor fusion algorithm. The tip position is tracked in the ultrasound images using a template-based tracker and a convolutional neural network based tracker, respectively. Experimental results for a rhombus path are presented, where data obtained from both tracking sources are fused using Luenberger and Kalman state estimators. The mean and standard deviation of the Euclidean error for the Luenberger observer is 0.2 ± 0.11 [mm] whereas for the Kalman filter it is 0.18 ± 0.13 [mm], respectively.**

I. INTRODUCTION

Flexible medical instruments such as endoscopes, catheters and needles constitute an important set of tools used in minimally invasive surgery (MIS). The integration of robotic navigation with MIS is becoming increasingly popular due to the advances in sensing and actuation technologies [1]. Remote actuation of flexible catheters has been investigated using the conventional displacement tendons [2] and recently using external magnetic fields [3], [4]. Magnetic actuation is more favorable compared to tendons since it does not suffer from friction and hysteresis (Fig. 1).

Besides the actuation method, acquiring an accurate instrument tip pose is crucial for the success of robotically-guided MIS procedures, since the catheter tip is the end-effector. The tip pose is often acquired using medical imaging modalities such as ultrasound (US), fluoroscopy, computed tomography (CT) and magnetic resonance imaging (MRI). Among these modalities, US has the advantage of being easily accessible, safe to operate and the possibility to image using a hand-held probe. Further, ultrasound provides high acquisition rates, allowing sufficient feedback to allow for closed-loop control. Therefore, a two-dimensional (2D) US probe is used in this work to acquire transverse images of the catheter tip. Nonetheless, acquiring the tip pose using solely the US images is not sufficient. This is due to the fact the visual tracking methods used to estimate the tip pose from the acquired US images are prone to failure. The common reasons for this failure are the low signal-to-noise ratio and an excessive number of image artifacts in practice.

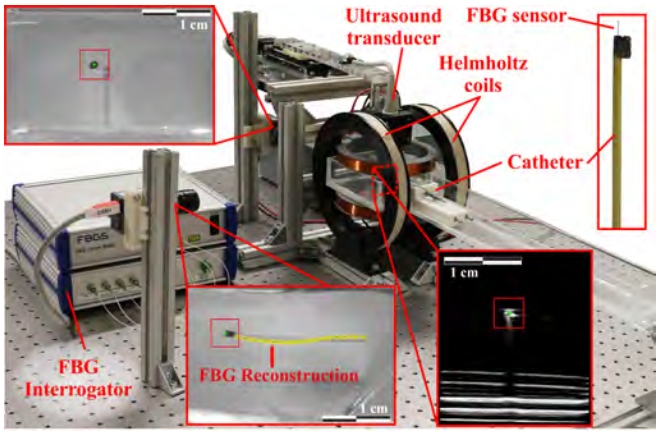


Fig. 1: The magnetically-actuated flexible catheter setup: The magnetic field is generated by two pairs of Helmholtz coils. The flexible catheter has a stack of permanent magnets at its tip which are steered by the coils. The catheter tip is detected using 2D ultrasound images. The shape and tip pose of the catheter are also obtained using multi-core Fiber Bragg Grating sensors.

Besides image-based tracking, catheter poses can also be retrieved by shape sensing techniques to measure the tip pose [5], with common approaches such as fiber-optic-sensor-based and electromagnetic-tracking-based (EM) methods. Among the two methods, EM tracking is very challenging to be used in conjunction with magnetically-actuated catheters. Hence, Fiber Bragg Grating (FBG) sensors, a fiber-optic type of sensor, are used in this work. FBG sensors have been used for various purposes for catheters in medical robotics research [6]. Arkwright *et al.* investigated the measurement of the muscular activity in the gastrointestinal tract using diagnostic catheters with FBG's [7]. Xu *et al.* designed a temperature-insensitive contact force sensor for bi-directional catheters using an FBG pair [8]. Shi *et al.* presented a survey on the state-of-the art of shape sensing techniques for continuum robots used in minimally invasive surgery [5].

In the aforementioned studies, single core optical fibers

A. Denasi, F. Khan, M. Kaya and S. Misra are affiliated with the Surgical Robotics Laboratory, Department of Biomechanical Engineering, MIRA-Institute for Biomedical Technology and Technical Medicine, University of Twente, The Netherlands. A. Denasi, K.J. Boskma and S. Misra are also affiliated with the Department of Biomedical Engineering, University of Groningen and University Medical Centre Groningen, The Netherlands.

C. Hennersperger, R. Göbl, M. Tirindelli, N. Navab are affiliated with Computer Aided Medical Procedures, Technical University of Munich, Germany. C. Hennersperger is also affiliated with Trinity College Dublin, Ireland. N. Navab is also affiliated with Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, USA.

This project (ROBOTAR) has partially received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation programme (Grant Agreement #638428) as well as from the Research and Innovation programme EDEN2020 (Grant Agreement #688279).

are used to reconstruct the 3D shape of the catheters. Recently, Moore *et al.* have proposed using multi-core optical fibers [9]. They investigated reconstructing the shape of a multi-core optical fiber with Fiber Bragg Grating sensors by numerically solving the Frenet-Serret equations. Multi-core optical fibers have advantages over single-core optical fibers. For instance, the multi-core optical fibers occupy less space compared to a combination of multiple single core fibers. Further, the FBG sensors are easier to align on multi-core optical fibers. To the best of the authors' knowledge, multi-core optical fibers have not been used for shape reconstruction of flexible medical instruments. This study investigates the shape reconstruction of magnetically-actuated flexible catheters using multi-core FBGs.

The measurements obtained from multi-core FBGs can be a remedy to the aforementioned possible failure of visual tracking using US images. This is achieved by fusing the data from multi-core FBGs and US images, thereby increasing the success rate of visual tracking. In this study, sensor fusion using state estimators is investigated. The estimation performance of two different methods, namely a high gain Luenberger observer and Kalman filter are compared. A simple model with a suitable disturbance term is used in the design of the state estimators. To the best of the authors' knowledge, such an observer-based fusion algorithm has not been used in medical robotics applications.

The presented tracking and sensor fusion algorithm could potentially be used in clinical minimally invasive procedures such as transfemoral aortic valve implantation, atrial fibrillation, and angioplasty. The tip pose estimation accuracy achieved with this method can reduce the duration of the aforementioned surgical procedures and increase their safety.

The rest of the work is organized as follows: Section II describes the shape sensing using multi-core FBGs. This is followed by the tracking algorithms for US images in Section III. The sensor fusion algorithm is introduced in Section IV. Section V reports the results of the experiments. The conclusions are drawn in Section VI and directions for future work are given.

II. SHAPE SENSING USING MULTICORE FIBER BRAGG GRATING SENSORS

This section describes the procedure to reconstruct the 3D shape of the magnetic catheter using sets of four FBG sensors placed along the length of the catheter [10]. The optical fiber used in this study contains four cores, where one of the cores is placed in the center axis of the fiber. The FBG sensors etched on the three cores placed around the center are used to measure the strains whereas the center core can be used for the purpose of temperature compensation. An FBG periodically modulates the refractive index in each core of the fiber. It reflects the light at a specific wavelength, called the Bragg wavelength λ_B. Variations in mechanical strain and temperature result in a change in the Bragg wavelength. The shift in the Bragg wavelength due to applied strain and temperature change is given by the following exact

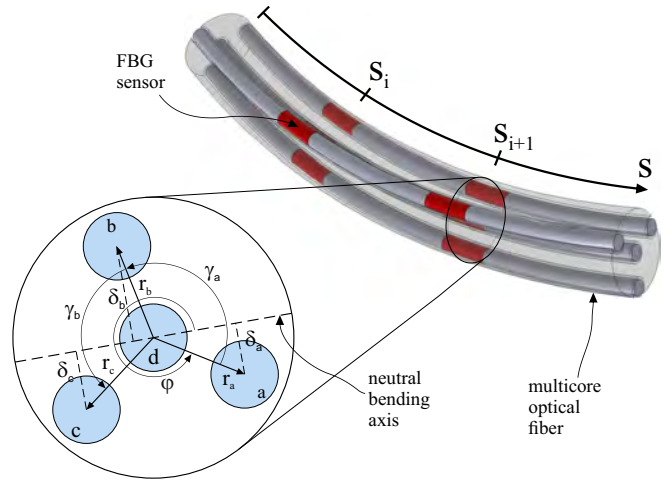


Fig. 2: Section of a multicore fiber with Fiber Bragg Grating (FBG) sensors is shown. The fiber has 4 cores and each core has 32 FBG sensors.

differential form [11]:

$$\frac{\mathrm{d}\lambda_B}{\lambda_B} = (1 - p_e)\,\mathrm{d}\varepsilon + (\alpha_\lambda + \alpha_n)\,\mathrm{d}\mathcal{T}, \tag{1}$$

where p_e, α_λ and α_n are the photoelastic, thermal expansion and thermo-optic coefficients, respectively. Further, ε and $\mathcal{T}$ are the axial strain and temperature, respectively. After integrating both sides of (1) we obtain

$$\ln \frac{\lambda_B}{\lambda_{B,0}} = (1 - p_e)(\varepsilon - \varepsilon_0) + (\alpha_\lambda + \alpha_n)(\mathcal{T} - \mathcal{T}_0), \tag{2}$$

where $\lambda_{B,0}$, ε_0 and $\mathcal{T}_0$ indicate the reference Bragg wavelength, reference strain and reference temperature, respectively. It should be noted that equation (1) is often approximated without performing the above integration procedure (d is replaced by Δ and λ_B by $\lambda_{B,0}$) [10]. However, this introduces errors which increase for increasing strain. If the temperature change is assumed to be negligible (i.e. $\mathcal{T} - \mathcal{T}_0 \approx 0$), then the shift in the Bragg wavelength is due to the applied strain which simplifies (2) as follows:

$$\ln \frac{\lambda_B}{\lambda_{B,0}} = (1 - p_e)(\varepsilon - \varepsilon_0). \tag{3}$$

The axial strain ε measured by the FBG sensors placed on each core can be related to the distance of the fiber to the neutral axis (see Fig. 2) as follows:

$$\varepsilon = \frac{\mathrm{d}s - \mathrm{d}l}{\mathrm{d}l} = \frac{(\rho - \delta)\,\mathrm{d}\theta - \rho\mathrm{d}\theta}{\rho\mathrm{d}\theta} = -\frac{\delta}{\rho} = -\kappa\delta, \tag{4}$$

where κ is the curvature of the flexible catheter. Here, it is assumed that the catheter can be modeled as a beam in pure bending. The curvature and its direction are determined at a particular cross section using the three FBG sensors placed around the center (see Fig. 2) as follows:

$$\begin{aligned}
\varepsilon_a - \varepsilon_0 &= -\kappa\delta_a = -\kappa r_a \sin(\varphi), \\
\varepsilon_b - \varepsilon_0 &= -\kappa\delta_b = -\kappa r_b \sin(\varphi + \gamma_a), \\
\varepsilon_c - \varepsilon_0 &= -\kappa\delta_c = -\kappa r_c \sin(\varphi + \gamma_a + \gamma_b),
\end{aligned} \tag{5}$$

where φ is the angle between r_a and the neutral axis. Further, r_a, r_b and r_c are the distance of the cores a, b and c from the center, respectively. It is assumed that the position and orientation of the FBG sensors are known and constant. The unknown parameters κ, φ and ε_0 can be determined by solving the set of equations (5). The curvature $\kappa(s_i)$ and its direction $\varphi(s_i)$ can be obtained at every location s_i by interpolating the measurements from FBG sensors placed on the catheter. These can then be used to compute the tangent vector of the curve of the catheter [12]. Consequently, the shape and respectively the tip pose of the flexible catheter is determined by numerical integration of the tangent vector.

III. ULTRASOUND TRACKING ALGORITHM

In this section, two algorithms to obtain catheter pose information for given ultrasound images are introduced: i) a template based tracking algorithm, and ii) a tracking algorithm utilizing convolutional neural networks.

A. Template Based Tracking Algorithm

The ultimate goal of the template based tracking method is registering the template image of the catheter, $T(\mathbf{x})$ and current image of the catheter, $I(\mathbf{w}(\mathbf{x},\mathbf{p}))$. Here, $\mathbf{x} \in \mathbb{R}^2$ contains the pixel coordinates $\mathbf{x} = [x,\ y]^T$, and is transformed by a motion model, which is also known as the warping function, $\mathbf{w}(\mathbf{x},\mathbf{p})$. The aim of registration is to calculate the optimal value of the motion parameters, $\mathbf{p}$. The template based tracking algorithm described in this study is adapted from [13] and detailed in the following.

1) Motion Model: One of the most significant layers of visual tracking is determining the motion of a flexible catheter in the image plane. The motion model describes the transformation between the template and current images. The motion of a flexible catheter in the US image plane can be decomposed into translation and scaling transformations. These transformations can be described by an affine motion model, $\mathbf{w}(\mathbf{x},\mathbf{p})$, given as follows:

$$\mathbf{w}(\mathbf{x},\mathbf{p}) = \begin{bmatrix} 1+p_1 & p_3 \\ p_2 & 1+p_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} p_5 \\ p_6 \end{bmatrix}, \tag{6}$$

where $\mathbf{p} \in \mathbb{R}^6$ is the 6-DOF motion parameter vector. Further, the motion model (6) can be used to track the tip position of the flexible catheter.

While tracking the tip of the catheter, transient variations of its pixel values may occur. In order to estimate the catheter tip position with a high accuracy, the tracker must compensate for these intensity fluctuations. A scale-offset model (α, β) can compensate these variations [14]. Then, the compensated current image $I^*(\mathbf{w}(\mathbf{x},\mathbf{q}))$, is written as:

$$I^*(\mathbf{w}(\mathbf{x},\mathbf{q})) = (\alpha+1)I(\mathbf{w}(\mathbf{x},\mathbf{p})) + \beta. \tag{7}$$

Using the affine motion and scale-offset models, eight parameters $\mathbf{q} \in \mathbb{R}^8$ are calculated for the transformation from the template to the current image during the visual tracking, where $\mathbf{q} = [\mathbf{p}^T, \alpha, \beta]^T$. In the rest of this section, the calculation of $\mathbf{q}$ vector using the SSD cost function is explained.

2) SSD-based Visual Tracking: The sum of squared differences (SSD) between $I^*(\mathbf{w}(\mathbf{x},\mathbf{p}))$ and $T(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}$ can be computed as follows:

$$SSD(\mathbf{q}) = \sum_{\mathbf{x}} [I^*(\mathbf{w}(\mathbf{x},\mathbf{q})) - T(\mathbf{x})]^2. \tag{8}$$

The main goal of SSD based visual tracking is to find the vector $\mathbf{q}$ that minimizes the SSD value between $T(\mathbf{x})$ and $I^*(\mathbf{w}(\mathbf{x},\mathbf{q}))$. For this purpose, forward or inverse compositional methods [15], first order optimization, or efficient second order minimization (ESM) methods [16] can be used. In this study, an ESM method is used since it is more robust to noise and its convergence rate is higher as compared to other optimization methods. Using ESM, the vector $\mathbf{q}$ is iteratively computed by accumulating $\Delta\mathbf{q}$ in each iteration ($\mathbf{q} \leftarrow \mathbf{q} + \Delta\mathbf{q}$). $\Delta\mathbf{q}$ is computed using ESM as follows:

$$\Delta\mathbf{q} = -2(\mathbf{J}(\mathbf{q}_0) + \mathbf{J}(\mathbf{q}_c))^{\dagger}(I^*(\mathbf{w}(\mathbf{x},\mathbf{q})) - T(\mathbf{x})), \tag{9}$$

where $\dagger$ is the Moore-Penrose pseudoinverse of a matrix. Further, $\mathbf{J}(\mathbf{q}_0)$ and $\mathbf{J}(\mathbf{q}_c)$ are the Jacobian of $I^*(\mathbf{w}(\mathbf{x},\mathbf{q}))$ [14], [16]. Before the iteration loop starts, $\mathbf{J}(\mathbf{q}_0)$ is precomputed. During the iterations, first $\mathbf{J}(\mathbf{q}_c)$ is computed and then $\Delta\mathbf{q}$ is calculated using (9). Iterations last until the number of iterations reach a predefined maximum iteration number or sum of $\Delta\mathbf{q}$ vector elements are smaller than the predefined threshold.

3) Pyramidal Implementation: In our experiments, it was observed that the motion of the flexible catheter can be very large in the US image plane. In this situation, the displacement of the catheter between the previous and the current US frames can be large. Hence, the number of iterations to calculate the motion parameters between the template and current images increases dramatically. Further, the catheter can be so far away from the search region in the image plane that tracking might fail. In order to increase the convergence rate and robustness of the tracking, pyramidal implementation of the proposed template-based tracking method is applied [17]. During the tracking, optimization methods are first applied to the coarsest level of pyramid and then, results are transferred to the next finer level of the pyramid. The number of pyramid layer for the frames in our data set and imaging system was selected as two. If the number of pyramid layers in the experiments was more than two, significant texture loss was observed because our maximum template size is 40×40 pixels.

4) Template Update Strategy: During the tracking, the template images are updated with a drift correction strategy to minimize the registration error and prevent template drifts [18]. This strategy consists of two registration steps. In the first step, the template and current images are registered. In the second step, the output image obtained in the first step is registered with the master template. The master template is the first appearance of the flexible catheter in the US image. After the second step, the template image is updated with registration output obtained in the second step.

B. Convolutional Neural Network Based Tracking

In this section, a tracking approach utilizing a Convolutional Neural Network (CNN) is presented. Fully convolutional neural networks have been shown to be well suited for the imaged-based estimation of object positions. When applied to the task of human pose estimation, the works of Newell et al. [19] and Wei et al. [20] demonstrated improved tracking performance compared to naive joint coordinate regression. In these works, fully convolutional neural networks are trained to predict a dense probability map, providing the probability for each pixel to contain the object of interest.

Following this approach, we propose a CNN architecture based on the work of Milletari et al. [21], originally described for segmentation of magnetic resonance images. As shown in Fig. 3, the network architecture consists of an encoding and a decoding branch, both organized in different levels. At each level of the encoding part, the input data are processed with a set of convolutional layers, the residuals are computed and downsampled by means of an additional convolutional layer, which stride size is set to two in both x and y direction. When moving toward deepest levels in the encoding branch, the resolution of the features maps decreases, while the number of filters in each convolutional layer increases. At the last level of the encoding branch, the resolution is lowered to 23×24, while the number of channels is 256.

The decoding branch of the structure is employed to retrieve the original resolution. As the encoding branch, it is organized in different levels. At each level of the decoding branch, the input data is concatenated with the output of the correspondent level in the encoding branch forming a so called skip-connection, and subsequently processed with a set of convolutional layers. Residuals are then computed and upsampled by means of a deconvolution layer. When moving from lower to higher levels in the right branch, the resolution of the features maps increases again, while the number of channels decreases. The output of the last deconvolution is processed with two convolutional layers, which reduces the number of channels of the output features map to one. Though the whole network $\tanh(\cdot)$ activation is applied after the convolutional layers, as it sped up training compared to the original ReLU.

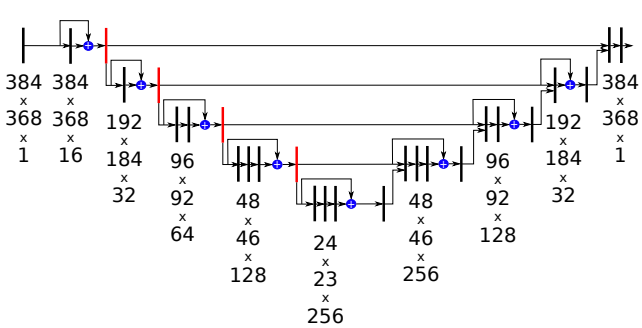


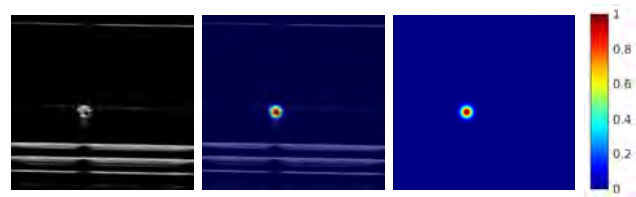
Fig. 3: Architecture of the convolutional neural network, with the encoding branch left and the decoding branch of inverse structure right.

During network training, the network output is compared to the ground truth images to compute the loss. Ground truth images consist in heatmaps, where decreasing intensities are associated to pixels with increasing distance from the catheter tip. The heatmap intensity at pixel (x, y) provides the probability that the catheter's tip is located at (x, y), given the Ground Truth tip location (x_t, y_t),

$$\mathrm{HM}(x, y) = \mathcal{P}((x, y)|(x_t, y_t)), \tag{10}$$

where $\mathrm{HM}(x, y)$ is the heatmap intensity value at pixel (x, y).

Under the assumption that $\mathcal{P}((x, y)|(x_t, y_t))$ follows a Gaussian distribution, ground truth heatmaps can be computed as the 2D Gaussian centered in (x_t, y_t).



(a) Input (b) Overlay (c) Heatmap

Fig. 4: Convolutional neural network based tracking results of a flexible catheter in a 2D transverse ultrasound image.

An exemplary output of the network can be seen in Fig. 4. The Loss function used for training is defined as

$$Loss = \sum_{k \in B} \left[\frac{1}{2} \sum_i (\mathrm{HM}_i^k - P_i^k)^2 + \omega R^k \right], \tag{11}$$

where B is the minibatchsize, HM_i is the ith pixel of the ground truth heatmap, P_i is the ith pixel of the predicted map, R^k is a regularization term and ω is the weight associated to the regularization term. In particular, R is defined as:

$$R = \sum_i \left[\left(\frac{\partial P_i}{\partial x} \right)^2 + \left(\frac{\partial P_i}{\partial y} \right)^2 \right] \tag{12}$$

During network validation, the catheter tip position is given by the position of the peak in the output map:

$$(x_t^p, y_t^p) = \arg\max_{i,j} P_{i,j} \tag{13}$$

IV. SENSOR FUSION ALGORITHM

In this section, a simplified kinematics based model is used to estimate the tip position of the magnetic catheter in 2D by fusing ultrasound and FBG measurements. This is followed by the application of a Luenberger state observer and a Kalman filter. In the following, it is assumed that the dynamics of the catheter can be approximated by a linear time-invariant model. Further, the influence of deviations of the actual system from the linear system are modeled with a disturbance term. The dynamics of a continuous-time LTI (linear time-invariant) system is expressed as follows:

$$\dot{\mathbf{x}}_p = \mathbf{A}_{pc}\mathbf{x}_p + \mathbf{B}_{pc}\mathbf{u} + \mathbf{B}_{pc}\mathbf{d} \tag{14}$$

$$\mathbf{y} = \mathbf{C}_{pc}\mathbf{x}_p + \mathbf{D}_{pc}\mathbf{u} + \mathbf{v} \tag{15}$$

where $\mathbf{x}_p \in \mathbb{R}^{n_p}$ and $\mathbf{u} \in \mathbb{R}^m$ are the state of the plant and the control input, respectively. Further, $\mathbf{d} \in \mathbb{R}^{n_d}$, $\mathbf{v} \in \mathbb{R}^l$ and $\mathbf{y} \in \mathbb{R}^l$ are the disturbance, the measurement noise and the measured output, respectively. The measurement noise is assumed to be zero-mean, Gaussian and white. The measurement data from different sensors can be fused by collecting them together as follows:

$$\mathbf{y} = \begin{bmatrix}\mathbf{y}_{sensor,1}^T & \dots & \mathbf{y}_{sensor,l}^T\end{bmatrix}^T \tag{16}$$

where each $\mathbf{y}_{sensor,i}$ for $i \in \{1, \dots, l\}$ represents a measurement from a particular sensor. The dynamics of the disturbance term given in (14) is expressed as follows:

$$\dot{\boldsymbol{\eta}}_d = \mathbf{A}_{dc}\boldsymbol{\eta}_d + \mathbf{B}_{dc}\mathbf{w} \tag{17}$$

$$\mathbf{d} = \mathbf{C}_{dc}\boldsymbol{\eta}_d, \tag{18}$$

where $\boldsymbol{\eta}_d \in \mathbb{R}^{n_\eta}$ and $\mathbf{w} \in \mathbb{R}^{n_w}$ are the state of the disturbance dynamics and an external driving signal, respectively. The signal $\mathbf{w}$ can either be deterministic but unknown or stochastic with the assumption of being zero-mean, Gaussian and white. If the signal $\mathbf{w}$ is deterministic but unknown and bounded, then a Luenberger state observer can be designed. But, if the signal $\mathbf{w}$ is stochastic with the assumption of being zero-mean, Gaussian and white, then a Kalman filter can be designed. The plant dynamics given by (14)-(15) and the disturbance dynamics given by (17)-(18) can be combined into the following augmented system

$$\dot{\mathbf{x}} = \mathbf{A}_c\mathbf{x} + \mathbf{B}_c\mathbf{u} + \mathbf{B}_{wc}\mathbf{w} \tag{19}$$

$$\mathbf{y} = \mathbf{C}_c\mathbf{x} + \mathbf{D}_c\mathbf{u} + \mathbf{v}, \tag{20}$$

with

$$\mathbf{A}_c = \begin{bmatrix}\mathbf{A}_{pc} & \mathbf{B}_{pc}\mathbf{C}_{dc} \\ \mathbf{0}_{n_d \times n_p} & \mathbf{A}_{dc}\end{bmatrix}, \mathbf{B}_c = \begin{bmatrix}\mathbf{B}_{pc} \\ \mathbf{0}_{n_d \times m}\end{bmatrix}, \mathbf{B}_{wc} = \begin{bmatrix}\mathbf{0}_{n_p \times n_w} \\ \mathbf{B}_{dc}\end{bmatrix},$$

$$\mathbf{C}_c = \begin{bmatrix}\mathbf{C}_{pc} & \mathbf{0}_{l \times n_\eta}\end{bmatrix}, \quad \mathbf{D}_c = \mathbf{D}_{pc}, \quad \mathbf{x} = \begin{bmatrix}\mathbf{x}_p^T & \boldsymbol{\eta}_d^T\end{bmatrix}^T \tag{21}$$

and $\mathbf{0}$ a zero matrix of appropriate size. For systems with a low sampling rate, the LTI system given by (19)-(20) can be discretized exactly at the sampling instants using the exact discretization method. The system is discretized with a sampling time, $T_s \in \mathbb{R}_{>0}$ which corresponds to reciprocal of the frames per second (1/fps) of the imaging system. After discretizing (19) and (20), the resulting system of difference equations are given as follows:

$$\mathbf{x}\left[(k+1)T_s\right] = \mathbf{A}_d\mathbf{x}\left[kT_s\right] + \mathbf{B}_d\mathbf{u}\left[kT_s\right] + \mathbf{B}_{wd}\mathbf{w}\left[kT_s\right] \tag{22}$$

$$\mathbf{y}\left[kT_s\right] = \mathbf{C}_d\mathbf{x}\left[kT_s\right] + \mathbf{D}_d\mathbf{u}\left[kT_s\right] + \mathbf{v}\left[kT_s\right]. \tag{23}$$

The matrices in (22)-(23) are given by

$$\mathbf{A}_d = \exp\left(\mathbf{A}_c T_s\right), \quad \mathbf{B}_d = \int_0^{T_s} \exp\left(\mathbf{A}_c \tau\right)\mathbf{B}_c d\tau, \tag{24}$$

$$\mathbf{B}_{wd} = \int_0^{T_s} \exp\left(\mathbf{A}_c \tau\right)\mathbf{B}_{wc} d\tau, \ \mathbf{C}_d = \mathbf{C}_c, \ \mathbf{D}_d = \mathbf{D}_c \tag{25}$$

where $\exp(\cdot)$ is the matrix exponential. If $\mathbf{A}_c$ is nonsingular, the integral terms in (24) and (25) are computed as follows:

$$\mathbf{B}_d = \mathbf{A}_c^{-1}\left(\mathbf{A}_d - \mathbf{I}\right)\mathbf{B}_c. \tag{26}$$

However, if $\mathbf{A}_c$ is singular, $\mathbf{B}_d$ can be computed as follows [22]:

$$\begin{bmatrix}\mathbf{A}_d & \mathbf{B}_d \\ \mathbf{0} & \mathbf{I}\end{bmatrix} = \exp\left(\begin{bmatrix}\mathbf{A}_c & \mathbf{B}_c \\ \mathbf{0} & \mathbf{0}\end{bmatrix} T_s\right). \tag{27}$$

For clarity of notation, the sampling time variables in (22)-(23) can be dropped to obtain

$$\mathbf{x}(k+1) = \mathbf{A}_d\mathbf{x}(k) + \mathbf{B}_d\mathbf{u}(k) + \mathbf{B}_{wd}\mathbf{w}(k), \tag{28}$$

$$\mathbf{y}(k) = \mathbf{C}_d\mathbf{x}(k) + \mathbf{D}_d\mathbf{u}(k) + \mathbf{v}(k) \tag{29}$$

for $k = 1, 2, 3, \dots$. In the following, the specific model for (19)-(20) used to estimate the tip positions of the flexible catheter in 2D is explained. Consider the x- and y- tip position coordinates of the catheter denoted by $p_x, p_y \in \mathbb{R}$ and the corresponding velocities denoted by $v_x, v_y \in \mathbb{R}$. We consider the case when there are no control inputs, thus $\mathbf{u} = \mathbf{0}$ holds. Consequently, the system dynamics is given as follows:

$$\underbrace{\begin{bmatrix}\dot{p}_x \\ \dot{p}_y \\ \dot{v}_x \\ \dot{v}_y\end{bmatrix}}_{\dot{\mathbf{x}}_p} = \underbrace{\begin{bmatrix}0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0\end{bmatrix}}_{\mathbf{A}_{pc}} \underbrace{\begin{bmatrix}p_x \\ p_y \\ v_x \\ v_y\end{bmatrix}}_{\mathbf{x}_p} + \underbrace{\begin{bmatrix}0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1\end{bmatrix}}_{\mathbf{B}_{pc}} \underbrace{\begin{bmatrix}d_x \\ d_y\end{bmatrix}}_{\mathbf{d}} \tag{30}$$

$$\mathbf{y} = \begin{bmatrix}\mathbf{y}_{US} \\ \mathbf{y}_{FBG}\end{bmatrix} = \underbrace{\begin{bmatrix}1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0\end{bmatrix}}_{\mathbf{C}_{pc}} \underbrace{\begin{bmatrix}p_x \\ p_y \\ v_x \\ v_y\end{bmatrix}}_{\mathbf{x}_p} + \underbrace{\begin{bmatrix}\mathbf{v}_{US,x} \\ \mathbf{v}_{US,y} \\ \mathbf{v}_{FBG,x} \\ \mathbf{v}_{FBG,y}\end{bmatrix}}_{\mathbf{v}} \tag{31}$$

where the matrices given by (14)-(15) are used. Further, $\mathbf{y}_{US}$ and $\mathbf{y}_{FBG}$ denote the position measurements obtained from the ultrasound images and FBG sensors, respectively. If the tracking of the catheter in 3D coordinates is required, then the system dynamics can be extended by including the states p_z and v_z to the state vector and d_z to the disturbance vector. The system dynamics described by $\mathbf{A}_{pc}$ in (30) with $\mathbf{d} = 0$ is commonly used in vision applications in the literature [23]. However, in order to have a more general yet simple model, a disturbance term should be included. There are different ways to model the disturbance term depending on the application. In order to keep the formulation sufficiently general a polynomial function is selected [24]. The disturbance term can be locally represented by an $(n-1)^{\text{th}}$ degree family of Taylor polynomial function of time as follows:

$$d_j(t) = \sum_{i=0}^{n-1} d_{j,i}t^i + d_{j,r}(t) \tag{32}$$

where $j \in \{x, y\}$, $d_{j,i} \in \mathbb{R}$ and $d_{j,r}(t) \in \mathbb{R}$ are the coefficients of the polynomial and a residual term, respectively. It is assumed that the residual term, $d_{j,r}(t)$, is such that its time derivatives for $i \geq n$ satisfy $|d_{j,r}^{(i)}(t)| \leq \gamma_{i-r} \approx 0$, thus they are all uniformly absolutely bounded and small enough to be negligible [24]. This implies that the residual term is slowly varying with respect to time. The disturbance dynamics for

a polynomial of degree n is given by

$$\underbrace{\begin{bmatrix} \dot{\boldsymbol{\eta}}_{d,1} \\ \dot{\boldsymbol{\eta}}_{d,2} \\ \vdots \\ \dot{\boldsymbol{\eta}}_{d,n-1} \\ \dot{\boldsymbol{\eta}}_{d,n} \end{bmatrix}}_{\dot{\boldsymbol{\eta}}_d} = \underbrace{\begin{bmatrix} \mathbf{0}_2 & \mathbf{I}_2 & \dots & \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 & \mathbf{0}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 & \mathbf{I}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 & \mathbf{0}_2 \end{bmatrix}}_{\mathbf{A}_{dc}} \underbrace{\begin{bmatrix} \boldsymbol{\eta}_{d,1} \\ \boldsymbol{\eta}_{d,2} \\ \vdots \\ \boldsymbol{\eta}_{d,n-1} \\ \boldsymbol{\eta}_{d,n} \end{bmatrix}}_{\boldsymbol{\eta}_d} + \underbrace{\begin{bmatrix} \mathbf{0}_2 \\ \mathbf{0}_2 \\ \vdots \\ \mathbf{0}_2 \\ \mathbf{I}_2 \end{bmatrix}}_{\mathbf{B}_{dc}} \underbrace{\begin{bmatrix} d_{x,r}^{(n)} \\ d_{y,r}^{(n)} \end{bmatrix}}_{\mathbf{w}},$$

$$\underbrace{\begin{bmatrix} d_x \\ d_y \end{bmatrix}}_{\mathbf{d}} = \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 & \mathbf{0}_2 \end{bmatrix}}_{\mathbf{C}_{dc}} \boldsymbol{\eta}_d \tag{33}$$

where $\mathbf{0}_2$ and $\mathbf{I}_2$ are zero and identity matrices, respectively. Further, in (33) each element of the state vector $\boldsymbol{\eta}_d$ satisfies $\boldsymbol{\eta}_{d,i} \in \mathbb{R}^2$.

Both the Luenberger observer and Kalman filter uses the model of a given system together with a suitable update term to estimate state variables that cannot be measured. In order to apply them to the augmented system dynamics (22)-(23), the system should be observable. This is satisfied if the observability matrix $\mathcal{O}^T = \begin{bmatrix} \mathbf{C}_d^T & \mathbf{A}_d^T \mathbf{C}_d^T & \dots & (\mathbf{A}_d^{n_p+n_d-1})^T \mathbf{C}_d^T \end{bmatrix}$ has full rank, i.e. $\operatorname{rank}(\mathcal{O}) = n_p + n_d$. The Luenberger observer is typically a copy of the augmented system dynamics (22)-(23) with a correction term

$$\widehat{\mathbf{x}}(k+1) = \mathbf{A}_d \widehat{\mathbf{x}}(k) + \mathbf{B}_d \mathbf{u}(k) - \mathbf{L}\left(\widehat{\mathbf{y}}(k) - \mathbf{y}(k)\right), \tag{34}$$

with the observer state $\widehat{\mathbf{x}}(k)$ and feedback gain matrix $\mathbf{L}$. Using (28), (29) and (34) and defining the observation error as $\mathbf{e}(k) = \widehat{\mathbf{x}}(k) - \mathbf{x}(k)$, the error dynamics is given as:

$$\mathbf{e}(k+1) = (\mathbf{A}_d - \mathbf{L}\mathbf{C}_d)\,\mathbf{e}(k) + \mathbf{L}\mathbf{v}(k) - \mathbf{B}_{wd}\mathbf{w}(k). \tag{35}$$

Here, the feedback gain matrix is designed such that if the system is observable, the eigenvalues of the nominal system (i.e. for $\mathbf{v} = 0$ and $\mathbf{w} = 0$) can be placed at arbitrary locations, for instance using Ackermann's formula. The nominal system (i.e. for $\mathbf{v} = 0$ and $\mathbf{w} = 0$) is asymptotically stable if all of the eigenvalues are inside the unit disk. The stability of the error dynamics (35) can be shown using a variation of the input-to-state stability theorem.

The Kalman filter consists of two phases at each sampling loop: prediction and update. In both phases next to the state estimate, the covariance of the observation error is estimated. In the prediction phase the estimate of the state is obtained using the discretized system dynamics (28)-(29) as follows:

$$\widehat{\mathbf{x}}(k+1 \mid k) = \mathbf{A}_d \widehat{\mathbf{x}}(k \mid k) + \mathbf{B}_d \mathbf{u}(k) \tag{36}$$

where $\widehat{\mathbf{x}}(k \mid k)$ is the estimate of $\widehat{\mathbf{x}}$ at time instant k given observations up to and including at time k. The predicted estimate of the error covariance matrix $\mathbf{P}(k+1 \mid k) = \operatorname{cov}\left(\mathbf{x}(k+1) - \widehat{\mathbf{x}}(k+1 \mid k)\right)$ is computed as follows:

$$\mathbf{P}(k+1 \mid k) = \mathbf{A}_d \mathbf{P}(k \mid k)\,\mathbf{A}_d^T + \mathbf{Q} \tag{37}$$

where $\mathbf{Q} \in \mathbb{R}^{(n_d+n_\eta)\times(n_d+n_\eta)}$ represents the covariance matrix of the process noise. In the update phase the state estimate can be corrected with an innovation term using the measurements

$$\widehat{\mathbf{x}}(k+1 \mid k+1) = \widehat{\mathbf{x}}(k+1 \mid k) + \mathbf{K}(k+1)\left[\mathbf{y}(k) - \mathbf{C}_d \widehat{\mathbf{x}}(k+1 \mid k)\right], \tag{38}$$

where $\mathbf{K}(k+1)$ is the optimal Kalman gain. The gain is given by

$$\mathbf{K}(k+1) = \mathbf{P}(k+1 \mid k)\,\mathbf{C}_d^T \left[\mathbf{C}_d \mathbf{P}(k+1 \mid k)\,\mathbf{C}_d^T + \mathbf{R}\right]^{-1}, \tag{39}$$

and $\mathbf{R} \in \mathbb{R}^{l\times l}$ is the covariance matrix of the measurement noise. On this foundation, the estimate of the error covariance matrix can be finally updated

$$\mathbf{P}(k+1 \mid k+1) = \left[\mathbf{I}_{(n_d+n_\eta)} - \mathbf{K}(k+1)\,\mathbf{C}_d\right] \mathbf{P}(k+1 \mid k) \tag{40}$$

where $\mathbf{I}_{(n_d+n_\eta)}$ is the identity matrix. The Luenberger and the Kalman state estimators are used to fuse US and FBG measurements in the next section.

V. EXPERIMENTS

In this section, the experimental setup is briefly presented. This is followed by the experimental results comparing the tracking and fusion algorithms introduced in Sections III and IV.

A. Experimental Setup

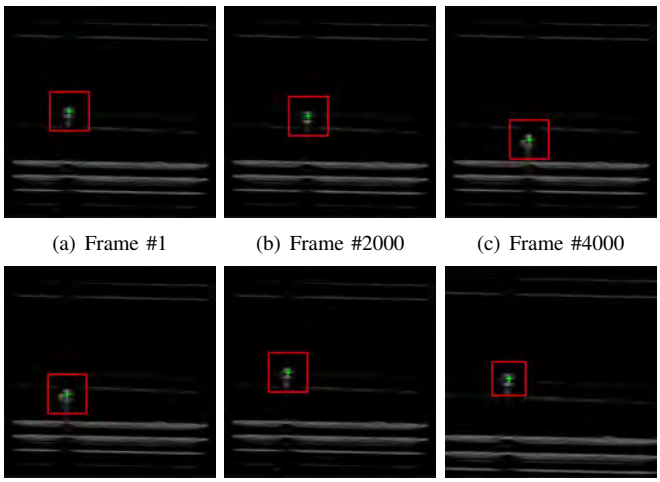
The setup consists of two pairs of Helmholtz coils (Teltron Helmholtz Coils S, 3B scientific, Hamburg, Germany) that can generate a homogeneous magnetic field along the x- and y-axis of the system (see Fig. 1). The details of the physical properties of the coils can be found in [4]. The catheter is 55 mm long and is made of a flexible hollow PVC tube with an outer diameter of 2 mm and an inner diameter of 1.2 mm. It is steered by the Helmholtz coils using a stack of 4 cylindrical Neodymium N48 (Supermagnete, Gottmadingen, Germany) magnets (2 mm diameter, 1 mm height) attached to the distal end of its shaft. The tip position of the catheter in 2D is obtained using transverse US images and FBG strain measurements. The transverse B-mode US images are acquired using a 14 MHz US transducer (L14-5/38, Ultrasonix, Richmond, Canada) in which the radial cross-section of the distal end of the shaft is visualized as a circular or oval shape. The strain measurements are obtained using a multicore FBG fiber (FBGS International NV) which has 4 cores where each core has 32 FBG sensors. The FBG sensor data is acquired using an FBG-Scan 804D interrogator. The accuracy of the visual tracking for the US tracker is measured using stereo cameras (Sony XCD-SX90, lenses Pentax 8.5mm, focus 0.2 - infinity, Cosmicar/Pentax 12 mm, focus 0.2 - infinity). The software to acquire the FBG data, US and stereo camera images is written in C++ on Linux Ubuntu. The sampling frequency of the measurements is set to 14 Hz.

B. Experimental Results

In this section, the results related to the template based tracking algorithm described in Section III-A are given. This is followed by the results for the fusion of US and FBG data.

1) Visual Tracking Results: The magnetic catheter is steered in a water tank using a pair of Helmholtz coils and its tip is imaged using a 2D ultrasound probe transversely. The US images are acquired with a rate of 14 fps.

The results of template tracking are shown in Fig. 5. It can be observed from Fig. 5 that the catheter pixel intensity changes over time. Also, a small number of pixels are available in the US image plane. Under these circumstances, the catheter tip is successfully tracked in long term using the scale-offset and affine motion models.



(a) Frame #1 (b) Frame #2000 (c) Frame #4000

(d) Frame #6000 (e) Frame #8000 (f) Frame #10240

Fig. 5: Template based tracking results of a flexible catheter in 2D transverse ultrasound images. *Please refer to the accompanying video.*

For the training of the proposed CNN tracking approach, two additional US sequences with different catheter motions were used and the network was trained to reproduce the results of the template tracker. The sequences were randomly separated into training, test, and validation sets. Training was performed with stochastic gradient descent with iterative learning rate reduction once there was no improvement in testing accuracy. On the validation set we observed an average error of 0.085mm. This error represents the difference between the network output and the ground truth. It is independent of the resolution of the US probe and it can even be lower. Although, practically such an error level would not be meaningful. When applying this network to the previously unseen sequence shown in Fig. 7, however, we observed an average error of 1.411mm. This implies, that for a more robust network more data would have to be collected, covering different motion patterns and visual appearances of the catheter.

2) Sensor Fusion Results: The accuracy of the sensor fusion method is evaluated using the ground truth data obtained from the stereo camera setup. The template based tracker described in Section III-A is used to obtain the ground truth data with respect to the world coordinate system.

TABLE I: Performance criteria for the fusion algorithms

	SSD	mean	std	min	max	NCC	SNR
Template-based							
Luenberger	1565.9	0.20	0.11	0.0013	0.88	0.9955	48.8
Kalman	1450.6	0.18	0.13	0.0027	0.95	0.9959	49.3
CNN-based							
Luenberger	3569.5	0.45	0.41	0.0005	2.24	0.9803	42.5
Kalman	3004.2	0.38	0.37	0.0043	2.44	0.9864	44.1

First of all, the stereo cameras (see Fig. 1) are calibrated using the method described in [25]. Then, the tip coordinates of the catheter obtained using visual tracking and FBGs are expressed with respect to the an inertial coordinate frame attached to the US. In order to obtain the ground truth position of the catheter, its tip is tracked in stereo camera images using the SSD based tracking method. Thus, the complexity of the algorithm is reduced.

The magnetic catheter is commanded to follow different geometric paths such as circle, rhombus and figure eight paths using the controller described in [4]. The Luenberger state observer and Kalman filter described in Section IV is used to fuse the data from US and FBG. The closed-loop observer poles for the Luenberger state observer are selected as $[0.9, 0.9, 0.9, 0.81, 0.81, 0.81, 0.729, 0.729, 0.729]$. Further, the process noise covariance is obtained empirically as $\mathbf{Q} = 5 \times \mathbf{I}_9$. The measurement noise covariance is $\mathbf{R} = \text{diag}[0.0099, 0.0059, 0.0370]$. The order of the disturbance model for both estimators is selected as $n_d = 1$.

The results for the rhombus path are shown as an example in Fig. 6 for the template based tracker and in Fig. 7 for the CNN tracker, respectively. Further, quantitative performance criteria including the sum of squared differences (SSD), normalized cross correlation (NCC), the maximum and minimum errors, the mean and the standard deviation of the Euclidean distance of each sample and signal-to-noise ratio (SNR) are given in Table I. The error is computed between the ground truth and the fused positions. It can be observed from these results that the Kalman filter slightly performs better as compared to the Luenberger observer. This is more pronounced in the presence of larger US tracking errors, as is the case with the employed CNN tracker.

VI. CONCLUSIONS

This study presents a sensor fusion method for magnetically-actuated flexible catheters. The algorithm makes use of two different source of measurements, images acquired from US and strains obtained from multicore FBGs. The tip positions for the tip of the catheter are obtained using two different visual tracking algorithms, the template-based and convolutional neural network based methods. The data obtained from both sources are fused using Luenberger and Kalman state estimators. The mean and standard deviation of the Euclidean error for the Luenberger observer is 0.2 ± 0.11 [mm] whereas for the Kalman filter it is 0.18 ± 0.13 [mm], respectively. In the future work, the accuracy of the tracking and sensor fusion will be improved.

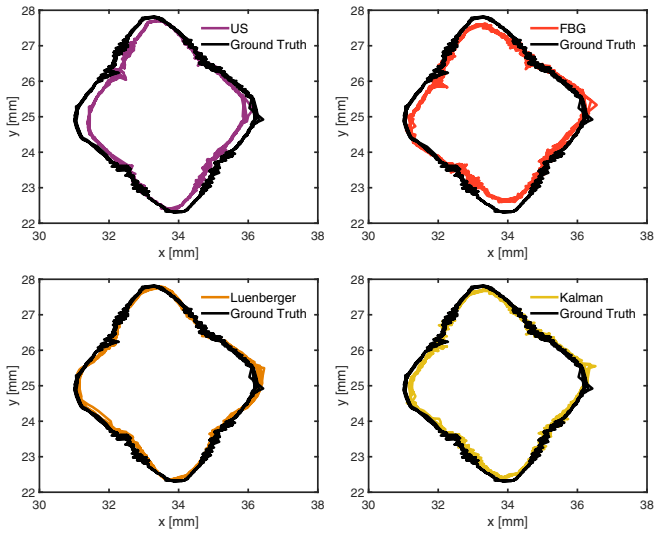


Fig. 6: The data fusion results for the catheter tip using Luenberger and Kalman state estimators are shown. Further, the ultrasound position obtained from the template based tracker and Fiber Bragg Grating sensor are also plotted. The ground truth obtained from the stereo cameras are also given.

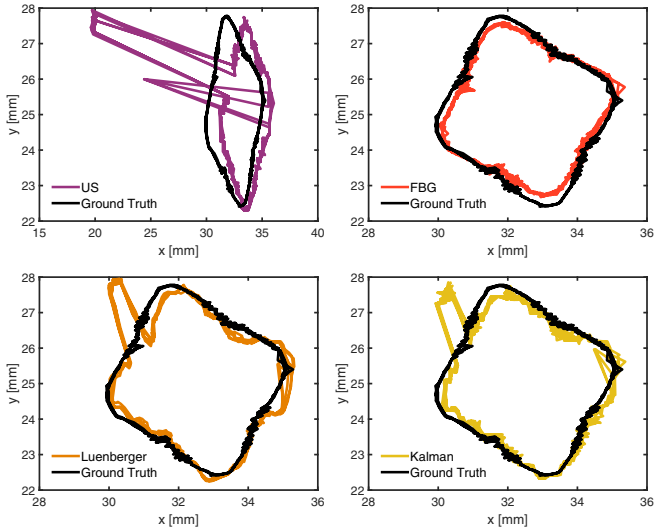


Fig. 7: The data fusion results for the catheter tip using Luenberger and Kalman state estimators are shown. Further, the ultrasound position obtained from the convolutional neural network based tracker and Fiber Bragg Grating sensor data are also plotted. The ground truth obtained from the stereo cameras are also given.

References

[1] V. Vitiello, S. L. Lee, T. P. Cundy, and G. Z. Yang, "Emerging robotic platforms for minimally invasive surgery," *IEEE Reviews in Biomedical Engineering*, vol. 6, pp. 111–126, 2013.

[2] G. J. Vrooijink, A. Denasi, J. G. Grandjean, and S. Misra, "Model predictive control of a robotically actuated delivery sheath for beating heart compensation," *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 193–209, 2017.

[3] J. Sikorski, I. Dawson, A. Denasi, E. E. G. Hekman, and S. Misra, "Introducing bigmag-a novel system for 3d magnetic actuation of flexible surgical manipulators," in *IEEE/ICRA International Conference on Robotics and Automation*, May 2017, pp. 3594–3599.

[4] K. J. Boskma, S. Scheggi, and S. Misra, "Closed-loop control of a magnetically-actuated catheter using two-dimensional ultrasound images," in *6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, June 2016, pp. 61–66.

[5] C. Shi, X. Luo, P. Qi, T. Li, S. Song, Z. Najdovski, T. Fukuda, and H. Ren, "Shape sensing techniques for continuum robots in minimally invasive surgery: A survey," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1665–1678, Aug 2017.

[6] V. Mishra, N. Singh, U. Tiwari, and P. Kapur, "Fiber grating sensors in medicine: Current and emerging applications," *Sensors and Actuators A: Physical*, vol. 167, no. 2, pp. 279 – 290, 2011, solid-State Sensors, Actuators and Microsystems Workshop.

[7] J. W. Arkwright, N. G. Blenman, I. D. Underhill, S. A. Maunder, N. J. Spencer, M. Costa, S. J. Brookes, M. M. Szczesniak, and P. G. Dinning, "A fibre optic catheter for simultaneous measurement of longitudinal and circumferential muscular activity in the gastrointestinal tract," *Journal of Biophotonics*, vol. 4, no. 4, pp. 244–251, 2011.

[8] L. Xu, L. A. Feuerman, J. Ge, K. R. Nilsson, and M. P. Fok, "Temperature-insensitive contact force sensing in bi-directional catheter using fiber bragg grating pair," *IEEE Sensors Journal*, vol. 17, no. 16, pp. 5118–5122, Aug 2017.

[9] J. P. Moore and M. D. Rogge, "Shape sensing using multi-core fiber optic cable and parametric curve solutions," *Opt. Express*, vol. 20, no. 3, pp. 2967–2973, Jan 2012.

[10] R. J. Roesthuis, M. Kemp, J. J. van den Dobbelsteen, and S. Misra, "Three-dimensional needle shape reconstruction using an array of fiber bragg grating sensors," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 4, pp. 1115–1126, Aug 2014.

[11] J. V. Roosbroeck, J. Vlekken, E. Voet, and M. Voet, "A new methodology for fiber optic strain gage measurements and its characterization," in *SENSOR+TEST Conferences 2009*, 2009, pp. 59 – 64.

[12] F. Khan, R. J. Roesthuis, and S. Misra, "Force sensing in continuum manipulators using fiber bragg grating sensors," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Accepted 2017.

[13] M. Kaya, E. Senel, A. Ahmad, and O. Bebek, "Visual tracking of biopsy needles in 2D ultrasound images," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2016.

[14] M. Hwangbo, J.-S. Kim, and T. Kanade, "Inertial-aided klt feature tracking for a moving camera," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 1909–1916.

[15] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International journal of computer vision*, vol. 56, no. 3, pp. 221–255, 2004.

[16] R. Richa, R. Sznitman, and G. Hager, "Robust similarity measures for gradient-based direct visual tracking," The Johns Hopkins University, Tech. Rep., 2012.

[17] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001.

[18] I. Matthews, T. Ishikawa, and S. Baker, "The template update problem," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 810–815, 2004.

[19] A. Newell, K. Yang, and J. Deng, *Stacked Hourglass Networks for Human Pose Estimation*. Cham: Springer International Publishing, 2016, pp. 483–499.

[20] S. E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 4724–4732.

[21] F. Milletari, N. Navab, and S. A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, Oct 2016, pp. 565–571.

[22] R. A. DeCarlo, "Some techniques for computing the matrix exponential and its integral," in *Linear Systems: A State Variable Approach with Numerical Implementation*. Prentice-Hall, 1989, ch. 14.

[23] M. Azizian and R. Patel, "Data fusion for catheter tracking using kalman filtering: applications in robot-assisted catheter insertion," p. 796413, 2011.

[24] H. Sira-Ramírez and V. F. Batlle, "Robust Σ-Δ modulation-based sliding mode observers for linear systems subject to time polynomial inputs," *International Journal of Systems Science*, vol. 42, no. 4, pp. 621–631, 2011.

[25] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov 2000.