

Robust Guidewire Tracking under Large Deformations Combining Segment-Like Features (SEGlets)

Alessandro Vandini^{a,*}, Ben Glocker^b, Mohamad Hamady^c, Guang-Zhong Yang^{a,*},

^a*The Hamlyn Centre for Robotic Surgery, Imperial College London, SW7 2AZ, London, United Kingdom*

^b*Biomedical Image Analysis Group, Imperial College London, 180 Queens Gate, SW7 2AZ, London, United Kingdom*

^c*Department of Interventional Radiology, Imperial College London, St Mary's Campus, W2 1NY, London, United Kingdom*

Abstract

Robust tracking of interventional tools, such as guidewires and catheters, in X-ray fluoroscopic video sequences has a wide range of clinical applications for endovascular procedures. Thus far, the tracking is usually achieved by finding the optimal displacement of the control points of a spline, which models the guidewire, between consecutive frames. The displacement of the control points is typically driven by a data term and smoothed by a regularization term. In the presence of large deformation and changes in length of the tool, the current tracking methods may fail to recover the guidewire motion. This can occur because of the limitation of the data and regularization terms, and the absence of an explicit solution for coping with elongations of the guidewire. The purpose of this paper is to present an algorithm that can robustly track guidewires under these challenging conditions. The algorithm is based on two main contributions: (a) new robust features termed SEGlets for segment-like features are introduced to overcome the limitations of the current data terms; (b) a tracking formulation based on the generation of tracking hypotheses by organizing the SEGlets in plausible guidewire shapes. The proposed method allows high flexibility of the guidewire between

*Corresponding authors

Email addresses: a.vandini12@imperial.ac.uk (Alessandro Vandini), g.z.yang@imperial.ac.uk (Guang-Zhong Yang)

consecutive frames in contrast to the spline model, which can suffer from the limitations of the regularization terms. Furthermore, the technique models elongations of the guidewire which makes it possible for robust tracking under motion. A tool model which is recursively updated by employing a Kalman filter, is also proposed for modelling the regularization term. A detailed evaluation and a comparative study with three state-of-the-art guidewire tracking methods have been performed to demonstrate the potential clinical value of the technique. The proposed method achieves an overall guidewire tracking precision of 2.40 pixels, tip precision of 25.55 pixels, false tracking rate of 5.73%, missing tracking rate of 9.69%, and F_1 score of 0.92. The implementation of the proposed technique and the three tracking methods will be made publicly available as software libraries.

Keywords:

intraoperative tool tracking, interventional radiology, fluoroscopy, guidewire

1. Introduction

Endovascular procedures are increasingly performed in interventional radiology with the benefits of minimal invasiveness, short recovery time and hospital stay, as well as improving patient safety and prognosis. In this context, robust intraoperative tracking of interventional tools, such as guidewires and catheters in X-ray fluoroscopic video sequences has a wide range of clinical applications. The relatively low image quality of fluoroscopy due to the need of minimising X-ray dosage can adversely affect the visibility of the tools. Figure 1 illustrates an example, showing the fluoroscopic video frames during endovascular procedures for uterine fibroid embolization and angioplasty, for which the visual identification of the guidewires is challenging. To overcome this limitation, tracking algorithms can be used to intraoperatively enhance the visibility of the tools as shown in [Wang et al. (2009); Heibel et al. (2013); Wu et al. (2015)].

Reliable tracking is also fundamental for 3D localisation of the tool [Hoffmann et al. (2012)] in roadmapping frameworks [van Walsum et al. (2005); Bender et al. (2008); Brückner et al. (2009)]. This enables 3D visualization of tools within a preoperative model of the patient's anatomy, therefore facilitating the catheterization procedure. Moreover, tool tracking is employed to perform motion compensation of anatomical roadmaps [Atasoy et al. (2008); Brost et al. (2009, 2012)] which are affected by organ defor-

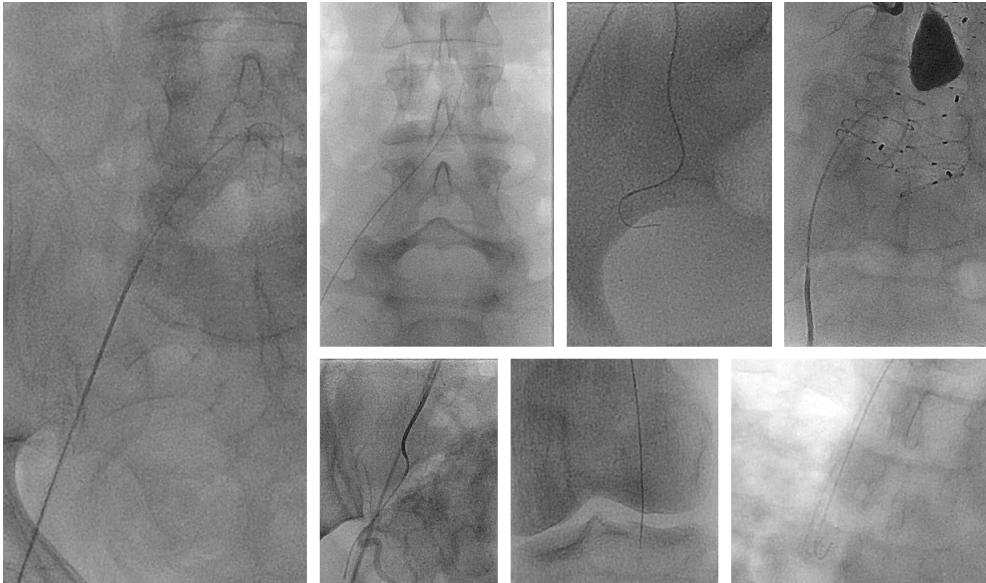


Figure 1: Fluoroscopic frames recorded during endovascular procedures, such as uterine fibroid embolisation and angioplasty, are shown. The visibility of the guidewire is limited due to the low quality of fluoroscopy and the presence of anatomical structures, e.g. spine.

mation and patients' respiratory motion. Motion compensation ensures that roadmaps overlaid on fluoroscopic images are correctly aligned with the underlying anatomy. Similarly, tracking is crucial when robotic endovascular tools are used, e.g. the Magellan catheter system (Hansen Medical, Mountain View, CA). In practice, it is important to track both the tip and body of the catheter as injury of the endothelial lining of the arterial wall may initiate platelet aggregation, causing blood clots to form, and stimulate neointimal hyperplasia (runaway proliferation of cells from the vessel wall into the centre lumen). Excessive pressure could rupture the blood vessel with dire consequences. For these reasons, tracking methods are necessary to perform vision-based shape sensing of the robotic tool [Vandini et al. (2013, 2015)] allowing its closed-loop control via visual servoing and ensuring safe navigation through complex anatomies. Finally, motion analysis of the tools based on tracking can sense essential metrics which are used for assessing the skills of the physicians [Rolls et al. (2013)].

Related Work. Thus far, a common approach to tracking guidewires consists of modelling them with B-splines. Tracking is then achieved by finding the

displacements of the control points of the spline that recovers the guidewire motion between consecutive frames. The optimal displacements minimize an energy function or maximize the posterior probability in a Bayesian framework [Wang et al. (2009)]. The energy function can be defined either in a continuous domain [Baert et al. (2003); Slabaugh et al. (2007)] or in a discrete domain [Heibel et al. (2009, 2013); Honnorat et al. (2011)]. The search for control point displacements is driven by a data term and a regularization term. The data term is an image measurement of the guidewire and can be calculated using filter-based techniques, such as those proposed by [Koller et al. (1995); Frangi et al. (1998); Baert et al. (2003)]. Alternatively, learning-based methods can also be used [Wang et al. (2009); Pauly et al. (2010)]. The regularization term is a function of the spline for avoiding the degeneration of the control points, limiting non-physical behaviours and favoring smooth deformations. This term is usually implemented by introducing length preservation [Baert et al. (2003); Honnorat et al. (2011); Heibel et al. (2009)], or probabilistic priors based on expected guidewire deformations [Wang et al. (2009)].

During endovascular procedures, the projection of the guidewire in fluoroscopic images can be characterized by large displacements and changes in length caused by required and usually abrupt manoeuvres of the physician. In these cases, the current tracking methods may fail to estimate the correct displacement of the control points to recover large guidewire deformations. In particular, the limitations of the data and regularization terms, and the absence of an explicit solution for coping with elongations of the guidewire can cause tracking failure.

Current data terms do not provide a reliable measurement of the guidewire during the whole fluoroscopic sequence. They are usually characterized by a high false positive rate in their measurements when detecting background structures such as bones or other anatomical parts, and image artefacts [Barbu et al. (2007); Pauly et al. (2010); Heibel et al. (2013)]. To circumvent this problem, several methods have been proposed for detecting (semi-)automatically the guidewire. The appearance and shape of the guidewire are learned offline using an hierarchical curve model and Probabilistic Boosting Trees to achieve automatic guidewire segmentation in [Barbu et al. (2007)]. Another promising algorithm for automatic guidewire segmentation is proposed in [Honnorat et al. (2010)]. For this method, the segmentation is achieved by extraction of local segments followed by perceptual grouping. Although these methods show promising results and outperform standard

filter-based techniques [Koller et al. (1995); Frangi et al. (1998); Baert et al. (2003)], their miss/false detection rates are still relatively high. To address this issue, an algorithm that enhances curvilinear structures is proposed in [Bismuth et al. (2012)]. Although the method can improve guidewire segmentation, the high computational costs make it less practical for intraoperative tool tracking. Methods which require user interactions are presented in [Mazouer et al. (2009); Wang et al. (2010)]. They achieve accurate and reliable detection by estimating the best path between the end points of the guidewire provided by the user [Mazouer et al. (2009)] and using a graph-based approach [Wang et al. (2010)]. The integration of these methods in a full tracking framework is limited due to the need of manual interaction. They can be used, however, to provide the position of the guidewire in the first frame of the sequence as initialization tools for the tracking algorithm.

Regarding the limitation of the current regularization terms, they usually penalize large displacements of the spline preferring smoother trajectories. Moreover, they limit changes in the length of the spline in order to avoid its degeneration. These factors constrain the flexibility of the spline in consecutive frames, especially in high dynamic catheterization sequences and when the length of the guidewire is not constant through the sequence.

Finally, the current tracking formulations do not explicitly model elongations of the guidewire. Although a tip detector is proposed in [Wang et al. (2009)], the learning of a robust prior appearance model of guidewire tips can be challenging. A growing step in direction of the tangent of the spline tip is proposed in [Honorat et al. (2011)] to cope with the elongations of the tool. Guidewires, however, due to their flexibility can reach complex shapes that are difficult to be fully recovered by this simplified growing step.

Contributions. The purpose of this work is to propose an algorithm that can robustly track guidewires in X-ray fluoroscopic video sequences under large deformation due to large displacements and changes in length of the tool. The main contributions of the proposed method are twofold: (a) new segment-like features (SEGlet) designed specifically for tracking guidewires are proposed to overcome the limitations of the current data terms. SEGlets are detected taking into account temporal information of the guidewire segments and background structures, thus leading to a robust feature for guidewire tracking. (b) A new guidewire tracking formulation is introduced based on the generation of tracking hypotheses by organizing the SEGlets in plausible guidewire shapes. This formulation is different to existing algorithms which

are based on the displacement of control points of the spline that models the guidewire. The proposed formulation allows high flexibility of the guidewire between consecutive frames in contrast to the spline model, which can suffer from the limitations of the regularization terms. Furthermore, this formulation explicitly models elongation of the guidewire making it possible for the recovery of these challenging motions. A tool model which is recursively updated by employing a Kalman filter-based framework, is also introduced as a major component of the regularization term used for evaluating the tracking hypotheses. Finally, detailed evaluation of the algorithm and a comparative study with three state-of-the-art guidewire tracking methods [Heibel et al. (2009, 2013); Honnorat et al. (2011)] are performed to demonstrate the potential clinical value of the proposed technique. The implementation of the proposed technique and the three state-of-the-art guidewire tracking methods will be made publicly available as software libraries¹. By making the reference implementations available, we hope to further facilitate collaborative research and objective evaluation of new methods in this area.

2. Material and methods

The proposed method focuses on the tracking of a guidewire by assuming that its position in the first frame of the fluoroscopic sequence is given by a manual or automatic initialization step.

In the following sections, each component of the guidewire tracking algorithm is described in details. The detection of the SEGlets is shown in Sec 2.1 while the algorithm that generates the tracking hypotheses by organizing and interpolating the SEGlets in plausible guidewire shapes is described in Sec. 2.2. Finally, the tool model and the evaluation of the tracking hypotheses that leads to the guidewire detection are shown in Sec. 2.3 and 2.4, respectively.

2.1. Detection of Segment-Like Features (SEGlets)

Guidewires typically appear in fluoroscopic frames as thin and continuous curves with variable thickness and shape of the tip depending on the performed endovascular procedure. Different examples are illustrated in Figure 1. The shape of a guidewire can be modelled as a set of straight and

¹The material is available at the project web page www.imperial.ac.uk/hamlyn-centre/

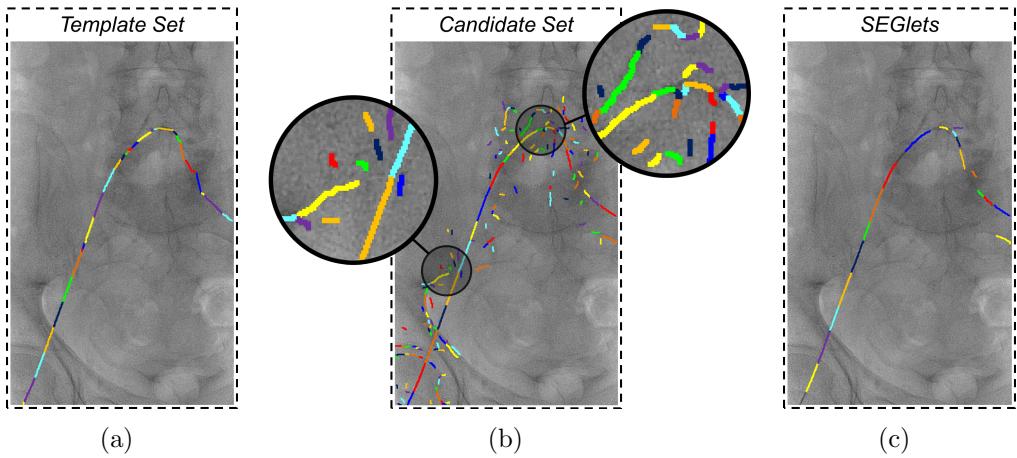


Figure 2: The curve that models the guidewire in the previous frame is divided into a set of straight connected segments called *template set* [See (a)]. Straight segments in the current frame that are within a certain distance ϵ from the guidewire in the previous frame define the *candidate set*, as shown in (b). Since no interpolation or line fitting is performed, some parts of the candidate set may appear as noisy short line segments. Finally, segments from the *candidate set* that are most similar to the segments of the *template set* are selected as SEGlets [See (c)] and organized in plausible guidewire shapes called tracking hypotheses.

connected segments (see Figure 2(a)). Although the guidewire can be characterized by large deformations and changes in length of the tool between consecutive fluoroscopic frames, the main straight segments that model the guidewire are still distinguishable from other background objects. In fact, the background objects are not expected to be characterized by a similar connected and organized line-like shape. These main straight segments are called segment-like features (SEGlets) and they represent a robust and discriminative feature in the proposed tracking framework.

The first step of the proposed algorithm is to segment the image pixels that belong to line-like structures, i.e. the foreground, from the background. To this end, a line-like feature detector, such as [Koller et al. (1995); Frangi et al. (1998); Baert et al. (2003); Honnorat et al. (2011)] is applied to the fluoroscopic image. To achieve a robust separation between the foreground and noise, the latter is assumed to be Gaussian, while the foreground is characterized by strong feature responses that rarely occur in the image. These characteristics are also observed when computing image edges [Canny (1986)]. A global histogram of the feature detector response is calculated and

the low percentile, which in our case is defined by 95% of the distribution, is considered as noise while the rest is segmented as foreground. The use of a single adaptive threshold can lead to disconnected line-like structures due to fluctuations in the feature detector responses. To avoid this problem, a hysteresis thresholding is applied. Pixels with values that are above a high threshold are immediately segmented as foreground while the pixels under a low threshold are segmented as background. For those pixels that lie in between the two thresholds, foreground pixels are searched in their connected area. In case that at least one foreground pixel is found, the pixel is segmented as foreground. The high threshold is defined by the low percentile of the global histogram of the feature detector response and the high and low threshold have a ratio of two². A fast thinning algorithm similar to the one proposed in [Zhang and Suen (1984)] is applied to the segmented image and connected components are labelled from the thinned image.

Straight segments are then detected by clustering pixels of the connected components considering their local curvature. An admissible maximum length for the straight segments is also considered during clustering. This ensures that the guidewire is composed by multiple segments even when it has a very short length. All the straight segments found in the current frame within a distance ϵ from the guidewire in the previous frame define the *candidate set*. An example of a *candidate set* is illustrated in Figure 2(b). The distance ϵ defines the largest possible displacement of a segment between two frames. The guidewire in the previous frame is known and it is modelled as a curve. This curve is divided into a set of straight connected segments called the *template set*, as shown in Figure 2(a). We consider the segment at the tip of the *template set* having constant length v . This avoids the breaking up of the tip in small segments when the detection of the tool in the previous frame is noisy at the tip. A dissimilarity measure between each segment i of the *template set* and every segment j of the *candidate set* is estimated using a cost function f_c . For each segment i of the *template set* the m segments of the *candidate set* that have the lowest dissimilarity measure are defined as SEGlets. The value of m is estimated in order to ensure that the sum of the lengths of the first m candidate segments is equal or greater than the length

²The aforementioned values are in line with the ones used for the computation of image edges [Canny (1986)].

of segment i . The cost function f_c is defined as:

$$f_c = f(i, j) = C_\phi(i, j) + C_{dist}(i, j) + C_{len}(i, j) + C_{BG}(j) \quad (1)$$

which takes into account temporal information of the guidewire segments using orientation, distance and length terms, respectively modelled by $C_\phi(i, j)$, $C_{dist}(i, j)$ and $C_{len}(i, j)$, and background structures via $C_{BG}(j)$. The term $C_\phi(i, j)$ estimates the variation in orientation between the two segments and it is defined as:

$$C_\phi(i, j) = \frac{|\phi_j - \phi_i|}{\frac{\pi}{2}} \quad (2)$$

where ϕ_j and ϕ_i describe the orientation of the candidate segment j and segment i , respectively. The term $C_{dist}(i, j)$ measures the average of Euclidean distances calculated from all the points of the shortest segment between i and j , and the closest points on the other segment. It is defined as:

$$C_{dist}(i, j) = \begin{cases} \frac{1}{\epsilon\kappa} \sum_{s=1}^{\kappa} d_{min}(Seg_i(s), Seg_j), & l_i < l_j \\ \frac{1}{\epsilon f} \sum_{s=1}^f d_{min}(Seg_j(s), Seg_i), & \text{else} \end{cases} \quad (3)$$

where l_i and l_j are the lengths of the segments, $d_{min}(Seg_i(s), Seg_j)$ is the distance between the s^{th} point of the segment i , i.e. Seg_i , to the closest point on the segment j , i.e. Seg_j . The terms κ and f are the number of points of Seg_i and Seg_j , respectively. The term $C_{len}(i, j)$ considers the difference in length between the segment i and j and it penalizes the candidate segment j that is shorter than i . It is defined as:

$$C_{len}(i, j) = \begin{cases} \frac{|l_i - l_j|}{l_i}, & l_i > l_j \\ 0, & \text{else} \end{cases} \quad (4)$$

Finally, the term $C_{BG}(j)$ penalizes candidate segments that belong to background structures. For this purpose, a background map is created on the first frame of the sequence and updated every successive frame using an occupancy grid. Pixels of the segments that are not selected as SEGlets in the current frame add a single contribution to the occupancy grid. The term $C_{BG}(j)$ is defined as:

$$C_{BG}(j) = \frac{1}{f} \sum_{s=1}^f \frac{BG_{map}(Seg_j(s))}{r} \quad (5)$$

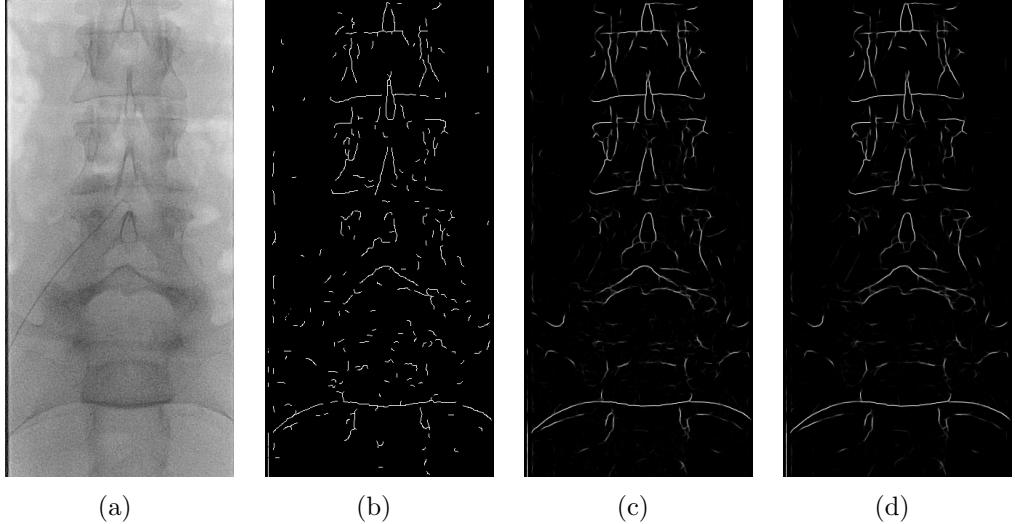


Figure 3: In (b), (c) and (d) the background maps of the first, middle and last frame of a sequence recorded during an endovascular procedure are shown. The brightness of each pixel indicates the likelihood of the pixel being a background structures. The first fluoroscopic frame of the sequence is shown in (a) as reference.

where f is the number of points of the segment j , $BG_{map}(Seg_j(s))$ is the value of the occupancy grid for the s^{th} point of the segment and r is the index of the current frame in the sequence. The background map along a fluoroscopic sequence is shown in Figure 3. Examples of detected SEGlets in a single frame are shown in Figure 2(c).

2.2. Generation of Tracking Hypotheses

Intuitively, SEGlets represent a sparse sample of the guidewire in the current frame, as shown in Figure 4(b). Plausible guidewire shapes can be generated by organizing and interpolating the SEGlets, as illustrated in Figure 4(b). These shapes represent the tracking hypotheses of the guidewire in the current frame and are generated using the following algorithm. First, a connection property is defined; two SEGlets are *connected* if both of the following conditions are satisfied:

- The distance between the SEGlets endpoints is less than a constant κ ;
- The largest difference between the orientation of the line that connects

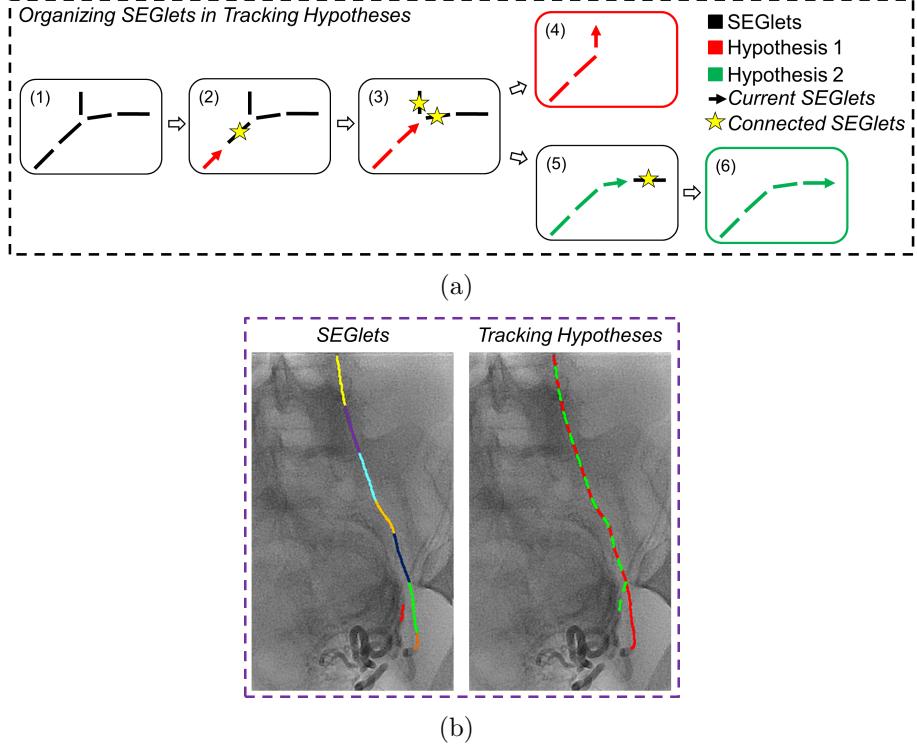


Figure 4: Illustration of the steps of the algorithm used to organize a set of detected SEGlets in tracking hypotheses are shown in (a). Tracking hypotheses resulting from the organization of the SEGlets in plausible guidewire shapes are shown in (a.4) and (a.6). Example of detected SEGlets in a fluoroscopic frame and the tracking hypotheses (depicted in red and green) generated organizing these SEGlets are shown in (b).

the two SEGlets endpoints and the orientations of the two SEGlets is less than 2θ ; and the smallest difference is less than θ .

These conditions guarantee a degree of continuity between SEGlets, which can be tuned with the parameter θ . With a small value of θ , only SEGlets which lie on a straight line will be found *connected* while with higher values more complex and curvilinear shapes can be generated. Finally, the last SEGlet added to a tracking hypothesis is defined as *current SEGlet*.

The algorithm starts with the generation of the first hypothesis by adding to it the SEGlet closest to the “base” of the guidewire tracked in the previous frame, as shown in Figure 4(a.2). The “base” is the endpoint of the guidewire that is usually close to the image border and opposite to the tip of

the guidewire. Next, SEGlets that are *connected* to the *current SEGlet* are searched. If only one SEGlet is found, as shown in Figure 4(a.2), this is added to the current hypothesis, Figure 4(a.3). In case two or more SEGlets are found *connected*, as shown in Figure 4(a.3), one SEGlet is added to the current hypothesis, Figure 4(a.4), while the other SEGlets generate new tracking hypotheses, Figure 4(a.5). Every new hypothesis is generated by adding the set of SEGlets belonging to the current hypothesis together with one of the SEGlets found *connected*. Afterwards, another search is performed for each hypothesis to find the SEGlets *connected* to its *current SEGlet*. The organization is repeated until no SEGlet can be added to any hypothesis, as shown in Figures 4(a.4) and (a.6). An additional segment from the *candidate set* is added to the guidewire “base” if needed in order to avoid shortening of the tool in proximity of the image border. Finally, linear interpolation is performed between the SEGlets of each resulting hypothesis to generate continuous tracking hypotheses. In Figure 4(b) a frame with detected SEGlets and corresponding hypotheses is shown. The algorithm that generates the tracking hypotheses is summarized in pseudo-code in Algorithm 1.

2.3. Tool Model

A tool model is introduced here in order to incorporate temporal information of the guidewire deformations in the fluoroscopic sequences. The model is described by a finite number of landmarks sampled along the guidewire and recursively updated by employing a Kalman filter-based framework (see Figure 5). The prediction phase of the Kalman filter is used to generate an estimate of the guidewire model in the current frame based on past measurements. The estimate, which is defined as \hat{G}_t , is used to calculate a regularization term for evaluating the tracking hypotheses. In our formulation, the regularization term appears as a probabilistic prior of the guidewire deformations in a Bayesian inference framework (see Sec. 2.4).

The tracked guidewire in the previous frame, i.e. G_{t-1} , is discretized in a set of M equidistant points which are called landmarks. The discretization step starts from the tip of the guidewire which consists of the first landmarks of the set. The distance between landmarks is equal to the arc length d_L which describes the resolution of the discretization. The tool model, which consists of a set of landmarks, is updated after every guidewire detection by the Kalman filter considering the measurements provided by the tracking algorithm, i.e. the landmarks of G_{t-1} . For each landmark j of the tool model,

Algorithm 1 Generation of Tracking Hypotheses, i.e. Hps

```
1: Create first hypothesis  $Hps(1)$ ;  
2:  $Hps(1) \leftarrow base$ ;  
3:  $Hps(1)$  current SEGlet =  $base$ ;  
4: while there is a connected SEGlet for at least one hypothesis do  
5:   for each hypothesis  $Hps(i)$  do  
6:     if  $Hps(i)$  current SEGlet is connected to 1 SEGlet then  
7:        $Hps(i) \leftarrow SEGlet$ ;  
8:        $Hps(i)$  current SEGlet =  $SEGlet$ ;  
9:     else  
10:      if  $Hps(i)$  current SEGlet is connected to  $n > 1$  SEGlets then  
11:        Create  $n - 1$  new hypotheses;  
12:        for each new hypothesis  $Hps(j)$  do  
13:           $Hps(j) = Hps(i)$ ;  
14:           $Hps(j) \leftarrow SEGlets(1 + j)$ ;  
15:           $Hps(j)$  current SEGlet =  $SEGlets(1 + j)$ ;  
16:        end for  
17:         $Hps(i) \leftarrow SEGlets(1)$ ;  
18:         $Hps(i)$  current SEGlet =  $SEGlets(1)$ ;  
19:      end if  
20:    end if  
21:  end for  
22: end while  
23: for each hypothesis  $Hps(i)$  do  
24:   Linear interpolation between  $SEGlets$  of  $Hps(i)$   
25: end for
```

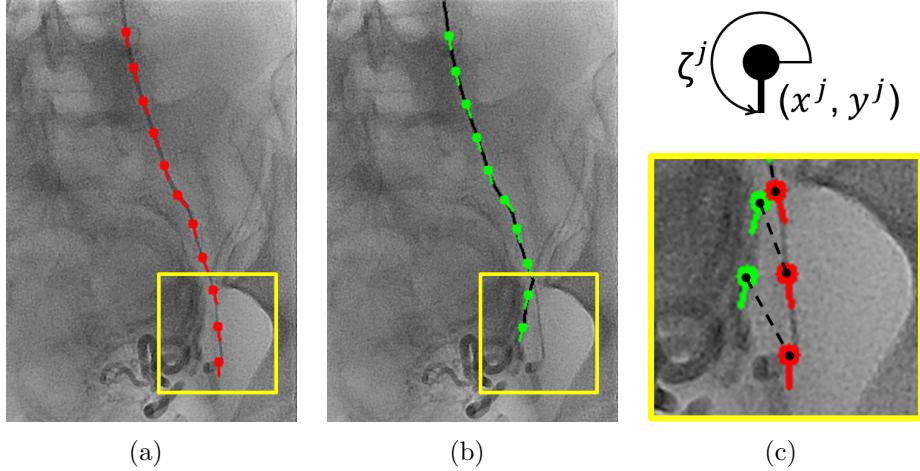


Figure 5: The prediction of the tool model in the current frame, namely \hat{G}_t , is shown in (a) while a tracking hypothesis H_{pt} of the current frame (in black) and its landmarks (in green) are shown in (b). An illustration of a guidewire landmark, which is described by its position (x^j, y^j) , and orientation ζ^j is reported in (c) together with a zoom-in that shows the matching between few landmarks of \hat{G}_t and H_{pt} used to calculate the distance D .

where $j = 1, \dots, M$, a Kalman filter state \mathbf{x}^j is defined as:

$$\mathbf{x}^j = [x^j \ y^j \ v_x^j \ v_y^j \ \zeta^j \ \omega^j]^T \quad (6)$$

where (x^j, y^j) is the 2D position of the landmark j , v_x^j and v_y^j are the velocity of the x^j and y^j components, ζ^j is its local orientations as shown in Figure 5(c), and ω^j is the angular velocity. A constant-velocity model is assumed to describe the motion of the landmarks and the discrete system dynamic model of the Kalman filter is defined as:

$$\mathbf{x}_t^j = \mathbf{A}\mathbf{x}_{t-1}^j + \gamma_t \quad (7)$$

with \mathbf{A} being the system matrix and γ_t the process noise which is described by a zero-mean Gaussian distribution with covariance Q_t . The measurement used at the correction stage are provided by (x^j, y^j) , and local orientations ζ^j of the landmark j of the last tracking results, namely G_{t-1} . Therefore, the output model is defined as:

$$\mathbf{z}_t^j = \mathbf{H}\mathbf{x}_t^j + \psi_t \quad (8)$$

where \mathbf{H} and ψ_t are the measurement matrix and noise, respectively. The measurement noise is described by a zero-mean Gaussian distribution with covariance R_t . In our Kalman filter-based framework, the matrices \mathbf{A} and \mathbf{H} have the following structures:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & T & 0 & 0 & 0 \\ 0 & 1 & 0 & T & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

where T is the sample period of the C-arm X-ray fluoroscopy.

2.4. Evaluation of Tracking Hypotheses

The tracking hypotheses are evaluated in a Bayesian inference framework which has demonstrated to be robust for tracking guidewires in fluoroscopic sequences [Wang et al. (2009)]. Considering a Markovian representation of the guidewire motion, the posterior probability of the guidewire at frame t is defined as:

$$P(Hp_t|I_t) \propto P(Hp_t)P(I_t|Hp_t) \quad (9)$$

where the image observation at frame t is I_t . The hypothesis that maximizes the posterior probability, namely the tracked guidewire at frame t , is $\hat{H}p_t$ and is defined as:

$$\hat{H}p_t = \arg \max_{Hp_t} P(Hp_t|I_t) \quad (10)$$

The prior component $P(Hp_t)$ is the regularization term which measures the distance between the tracking hypothesis Hp_t and the estimate of the tool model in the current frame, namely \hat{G}_t , calculated by the prediction phase of the Kalman filter. The discretization of the guidewire in landmarks is also performed for Hp_t (see Figure 5(b)) and the cardinality of its set of landmarks is equal to L . The prior component $P(Hp_t)$ is defined as:

$$P(Hp_t) = \frac{1}{\sigma_p \sqrt{2\pi}} \exp\left(\frac{-|D(Hp_t, \hat{G}_t)|^2}{2\sigma_p^2}\right) \quad (11)$$

where D is a function that considers both Euclidean and orientation distances between each landmark of Hp_t and \hat{G}_t . The function D is defined as:

$$D(Hp_t, \hat{G}_t) = \frac{1}{F} \sum_{s=1}^F \frac{1}{2} \left[\frac{|\zeta^s(Hp_t) - \zeta^s(\hat{G}_t)|}{\pi} + \frac{\|p^s(Hp_t) - p^s(\hat{G}_t)\|_2}{\epsilon} \right] \quad (12)$$

where $F = \min(M, L)$ and p^s is the 2D position of the landmark s , i.e. (x^s, y^s) .

The likelihood term $P(I_t|Hpt)$ provides the likelihood of the tracking hypothesis Hpt at frame t . This term is composed by two image measurements of the guidewire and is defined as:

$$P(I_t|Hpt) = P^v(I_t|Hpt)P_v + P^c(I_t|Hpt)P_c \quad (13)$$

where P_v and P_c are the prior components of the measurements and are set to 0.5. The term $P^v(I_t|Hpt)$ is the first image measurement and describes the lineness profile along the guidewire. It is defined by a Gaussian distribution, i.e. $N(u, \sigma)$. The distribution is initialized in the first frame of the fluoroscopic sequence and updated online for every frame of the sequence as:

$$u_t = (1 - \lambda)u_{t-1} + \lambda u_t^d \quad (14)$$

$$\sigma_t = (1 - \lambda)\sigma_{t-1} + \lambda\sigma_t^d \quad (15)$$

where u_t^d and σ_t^d are calculated considering the lineness of the pixels of the tracked guidewire at frame t , and λ is the learning parameter. The lineness value of each pixel in the frame is calculated using the line-like feature detector.

The second image measurement is $P^c(I_t|Hpt)$ which is defined as the Spline Local Binary Pattern (SLBP) measurement model proposed in [Cao and Wang (2012)]. The SLBP is an intensity measurement robust to low signal-to-noise ratio of fluoroscopic sequences. The SLBP intensity pattern is defined as:

$$IP(Hpt) = (bin_1^1, bin_1^2, bin_1^3, \dots, bin_F^1, bin_F^2, bin_F^3) \quad (16)$$

where bin_i^1 , bin_i^2 and bin_i^3 are defined as:

$$bin_i^1 = \begin{cases} 1, & I_i^{pn} > I_i^{nn} \\ 0, & \text{else} \end{cases} \quad bin_i^2 = \begin{cases} 1, & I_i^{nn} > I_i^c \\ 0, & \text{else} \end{cases} \quad bin_i^3 = \begin{cases} 1, & I_i^{pn} > I_i^c \\ 0, & \text{else} \end{cases} \quad (17)$$

The terms I_i^{pn} , I_i^{nn} and I_i^c are the mean intensity profiles centred at the i^{th} landmark of Hpt along the positive norm, the negative norm and the tracking hypothesis curve associated to Hpt , respectively. The image measurement derived from SLBP is defined as:

$$P^c(I_t|Hpt) = \frac{|IP(Hpt) \cdot IP(G_{t-1})|}{|IP(Hpt)| \cdot |IP(G_{t-1})|} \quad (18)$$

3. Experimental Evaluation

To demonstrate the clinical value and contributions of the proposed tracking algorithm an extensive comparative study with three state-of-the-art guidewire tracking methods [Heibel et al. (2009, 2013); Honnorat et al. (2011)] is conducted. The tracking methods are evaluated on a large clinical dataset which was acquired during several endovascular procedures. Details on the clinical dataset and the implementation of the algorithms are reported in Sec. 3.1 and 3.2, respectively.

3.1. Clinical Data

The evaluation of the proposed tracking method is conducted on a challenging yet common clinical dataset composed of 12 fluoroscopic sequences recorded during uterine fibroid embolisation, angioplasty of the iliac, popliteal and superficial femoral arteries, catheterization of the common iliac artery and right innominate vein, cannulation iliac artery, aorta IADSA and diagnostic angiogram of the iliac artery. The use of the dataset for this study was approved by the Imperial College Joint Research Compliance Office (JRCO) and the interventional C-arm used during surgery was a Philips Allura Xper. We considered sequences where the tool’s endpoint was located at the image border and the anatomical areas present in the dataset include aortic, iliac, femoral, popliteal territories as well as SVC, uterine artery and innominate veins. Regarding the interventional tools used during these procedures, a range of 0.018-inch guidewire and microguidewire, 0.035-inch “J” guidewire and 0.035-inch terumo guidewire are present. The image sizes of the sequences is 1024×1024 . The ground truth of the position of the guidewire in each frame of the dataset was manually annotated and used subsequently for algorithm evaluation. We considered a subsequence of the full recording for which the ground truth could be obtained. In addition, we discarded the parts during which the guidewire was completely retracted into the catheter or during the setup of the C-arm. Parts of the catheter that contains the guidewire are also visible in several sequences, such as 1, 2, 4, 8 and 12 of the dataset. Due to the difficulty in annotating separately these parts from the guidewire³ and their limited presence, they were considered as guidewire in our evaluation. Sequences 9 and 10 contain a 5 French pigtail catheter and

³The catheters used have similar dimensions than the guidewires making it difficult to visually distinguish the two ends of the tools in every frame of the sequences.

2.7 French microcatheter respectively without guidewires. The total number of frames of the dataset is 752.

Other characteristics of each sequence derived from the ground truth position of the guidewire are given in Table 1 including MeanD, i.e. the mean displacement of the tip of the guidewire between two consecutive frames along the whole sequence. The dataset is characterized by large and abrupt displacements of the guidewire as it can be observed from the values of MeanD and the maximum displacements of the tip of the guidewire between two consecutive frames (MaxD). The maximum displacement in the dataset occurs on sequence 12 with 85.7 pixels. The dataset presents large changes in length of the guidewire in the sequences as it can be observed by the maximum variation in length of the guidewire between two consecutive frames (MaxVL) with a maximum value of 130.8 pixels for sequence 12, and the minimum (MinL) and maximum lengths (MaxL) of the guidewire in the sequence. Anatomical structures and image artefacts caused by cardiac and respiratory motion are also present in the sequences of the dataset as it is shown in Figure 1. Finally, the dataset presents a large variety of guidewires and anatomical areas as previously been described. All these factors make these sequences a challenging and clinically relevant dataset for evaluating guidewire tracking methods.

3.2. Implementation Details

A comparative study between the proposed method (GSEG), the discrete tracking method (DT) [Heibel et al. (2009)], the interventional tool tracking method (ITT) [Heibel et al. (2013)], and the geometric-iconic (G-I) method presented in [Honnorat et al. (2011)] was conducted. G-I is a robust method for tracking guidewires which couples an iconic spline tracking formulation [Heibel et al. (2009, 2013)] with landmark detection along the guidewire. In previous studies, G-I outperformed DT thanks to the introduction of the landmark detection which makes this tracking formulation more robust to local minima and allows to preserve the global configuration of the spline [Honnorat et al. (2011)]. ITT which consists of an extension of DT, was found to perform on average as good as the robust guidewire tracking method (RGWT) proposed in [Wang et al. (2009)] where learning is involved [Heibel et al. (2013)]. G-I, contrary to DT and ITT methods, has a growing step in order to cope with elongations of the guidewire. The growing is performed in direction of the tangent of the spline tip after the optimal displacements of the control points is found and pixels that present a strong response to

Table 1: The mean (MeanD) and the maximum displacements (MaxD) of the tip of the guidewire between two consecutive frames, the minimum (MinL) and maximum lengths (MaxL) of the guidewire in the sequence, the maximum variation in length of the guidewire (MaxVL) between two consecutive frames are shown for each sequence of the dataset. The metrics are calculated using the ground truth of the guidewire and are in image pixels (px). The number of frames for each sequence is also reported in the last column of the table.

Seq	MeanD (px)	MaxD (px)	MinL (px)	MaxL (px)	MaxVL (px)	Frames
1	2.7	19.6	223.1	357.0	20.3	102
2	3.4	10.8	246.2	270.8	13.2	50
3	3.4	9.2	286.4	422.4	10.2	50
4	3.5	11.2	283.3	462.3	10.5	59
5	4.3	11.0	685.5	809.6	15.2	50
6	5.8	31.4	303.9	607.5	29.1	67
7	14.4	44.3	103.4	547.4	53.7	46
8	14.4	68.4	274.2	902.9	63.7	66
9	2.3	9.2	256.7	427.7	12.4	82
10	5.6	19.1	366.6	608.9	18.6	56
11	5.7	22.0	419.3	589.0	24.3	56
12	11.7	85.7	114.8	471.8	130.8	68

the line-like feature detector are added to the spline model [Honnorat et al. (2011)]. Our clinical dataset used to evaluate the tracking algorithms consists of several sequences that are characterized by large elongations of the guidewire, such as sequence 1, 5, 6, 7, 10, 11 and 12 (see Table 1). DT and ITT methods are not able, however, to recover these elongations of the tool due to the absence of a growing step for the spline model. To make the comparison between the methods more fair and interesting, an extension of DT and ITT is introduced here by integrating the growing step of G-I in DT and ITT methods. Therefore, we refer with DT* and ITT* to the extended version of DT and ITT, respectively, enhanced by the growing step. The implementation of DT, DT*, ITT, ITT* and G-I methods is a combination of the implementation provided by the original authors or re-implementation in close communication with these authors. It provides a unified interface in C++ for DT, DT*, ITT, ITT* and G-I methods since they are all based on a discrete optimization scheme. All these are available as the supplement material of this paper.

A filter based on the Hessian matrix [Baert et al. (2003)] is used as line-like feature detector for all tracking methods. This choice is justified by the need to measure the performance of the methods depending mostly on the

tracking formulation rather than the different feature detector used. The line-like features in the image are detected by analysing the value of the eigenvalues of the Hessian matrix, which is calculated as:

$$\mathbf{H} = \begin{pmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{pmatrix} \quad (19)$$

where L_{xy} is the result of the convolution of a scaled Gaussian derivative defined as:

$$L_{xy} = L * \frac{\partial^2}{\partial x \partial y} G(X, \sigma_H) \quad (20)$$

where $G(X, \sigma_H)$ is calculated as:

$$G(X, \sigma_H) = \frac{1}{2\pi\sigma_H^2} e^{\frac{-x^2}{2\sigma_H^2}} \quad (21)$$

and L is the fluoroscopic image. The eigenvalues of the matrix are estimated as:

$$\lambda_{1,2} = \frac{1}{2} \left(L_{xx} + L_{yy} \pm \sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2} \right) \quad (22)$$

The largest absolute eigenvalue should usually be positive in pixels that belong to dark line-like objects in the image, as the case of the guidewire in fluoroscopic frames. For these experiments, σ_H was set to 1.70 for DT, DT*, ITT, ITT* and G-I methods.

The parameters of the GSEG, DT, DT*, ITT, ITT* and G-I methods were empirically tuned using the first 8 sequences of the clinical dataset. The parameters of GSEG for all the sequences were set empirically as $\epsilon = 60$, $\kappa = 100$, $\theta = 50$, $\lambda = 0.5$ and $\sigma_p = 0.3$. The last section of the guidewire, which includes the tip, is characterized by the largest deformations making it the detection of the tip difficult. To consider this, for the guidewire tip segment only the 3 best candidates were chosen to be SEGlets and the following parameters were used for their connectivity with the *current SEGlet*: $\kappa_{tip} = \frac{\kappa}{3}$ and $\theta_{tip} = \frac{\theta}{2}$. In order to speed up the computational time by reducing the cardinality of the *candidate set*, a minimum admissible length of straight segments that could be considered in the *candidate set* was applied and σ_H was set to 1.80. The minimum admissible length, i.e. ι , was chosen to be 6 pixels to ensure a good compromise between resolution of the SEGlets and overall segmentation of the guidewire. The parameters of GSEG are summarized in Table 2.

Regarding DT, DT*, ITT and ITT* methods, the feature image parameters β and γ were set as 1.0 and 2.5, respectively, and the λ of DT and DT* was set to 2.0. The rest of the parameters of the four methods were set as the author suggested [Heibel et al. (2009, 2013)]. The G-I's parameters, such as *memory*, γ and the threshold used for tip growing were set as 1, 0.6 and 0.6 respectively. The choice of the value of *memory* was justified by the high dynamism of sequences used for the evaluation. The other parameters were set as reported in [Honorat et al. (2011)]. The number of displacements of the sparse pattern of DT, DT*, ITT and ITT* and G-I methods, namely the cardinality of the set of labels, was set to 97.⁴ This large number of labels was found necessary to allow the spline to cope with the large deformations of the guidewire in the clinical dataset. The following number of labels guarantees also a maximum displacement of the control points equal to ϵ of GSEG, namely 60 pixels, so that GSEG, DT, DT*, ITT, ITT* and G-I have the same deformation range. The initialization of the tracking algorithms is performed by providing the position of the guidewire, which can be obtained efficiently, in the first frame of each sequence. Here, the same initialisation is used for all the tracking methods.

The guidewire tracking algorithms run on a desktop PC with the following specifications: i7-2600 at 3.40 GHz and 16 GB of RAM. The average computational time of GSEG for a single frame is 0.79 sec, while for DT* is 1.95 sec, for DT is 1.84 sec, for ITT* is 2.02 sec, for ITT is 1.94 sec and for G-I is 2.66 sec, with an unoptimised C++ implementation. The computational times was estimated on the first 8 sequences and do not consider the time for calculating the line-like feature detector which is on average 0.22 sec.

4. Results

The performance of the proposed tracking method on the clinical data is measured using several quantitative error metrics. An extensive evaluation of the SEGlets is also conducted to demonstrate the robustness of the proposed features for guidewire tracking. The error metrics are defined in Sec. 4.1

⁴We found in our experiments that a larger set of 321 labels as suggested in [Heibel et al. (2013)] with a 1 pixel label spacing is yielding slightly better accuracy (the overall average improvement for the DT*, ITT* and G-I methods was about 3% in the first 8 sequences). The ranking of the overall score for all methods was unchanged while the computational time increased significantly for DT*, ITT* and G-I.

Table 2: The parameters of the algorithm are listed in the following. The “Imager Pixel Spacing” is 0.371\0.371 for sequence 2.

Symbol	Description	Method Section	Value
ϵ	max displacement guidewire between frames	detection SEGlets	60 pixels
ι	min admissible length, straight segments	detection SEGlets	6 pixels
v	length tip segment, <i>template set</i>	detection SEGlets	4ι pixels
κ	max distance SEGlets	generation hypotheses	100 pixels
κ_{tip}	max distance SEGlets, tip	generation hypotheses	$\frac{\kappa}{3}$ pixels
θ	max change in orientation SEGlets	generation hypotheses	50 degree
θ_{tip}	max change in orientation SEGlets, tip	generation hypotheses	$\frac{\theta}{2}$ degree
λ	learning parameters Gaussian distributions	evaluation hypotheses	0.5
σ_p	standard deviation prior component	evaluation hypotheses	0.3

while the evaluation of the SEGlet is reported in Sec. 4.2. The results of the proposed method (GSEG) compared to the discrete tracking method (DT) [Heibel et al. (2009)], its extension (DT*), the interventional tool tracking method (ITT) [Heibel et al. (2013)], its extension (ITT*), and the geometric-iconic method (G-I) [Honnorat et al. (2011)] are shown in Sec. 4.3.

4.1. Error Metrics

The error metrics used to evaluate the performance of the tracking methods are the guidewire tracking precision, the missing and false tracking rates, F_1 score, the tip precision and tip error [Wang et al. (2009)].

The guidewire tracking precision is the mean of the distances of each point on the tracked guidewire to its closest point on the ground truth for each frame of the sequence. The formula of the distance, also given in [Heibel et al. (2013)], is the following:

$$d(C_T, C_G, s) = \min_t (\|C_T(s) - C_G(t)\|_2) \quad (23)$$

where C_T is the estimated guidewire and C_G its ground truth.

False tracked points are defined as the points on the tracked guidewire that have a distance from their closest point on the ground truth greater than a threshold. Points on the ground truth that have a distance from their closest point on the tracked guidewire greater than a threshold are missed tracked points. The percentages of these quantities over all amounts of tracked guidewire points and ground truth points in each frame of the

sequence are the false and missing tracking rate, respectively. We used a threshold of 3 pixels as suggested in [Wang et al. (2009); Heibel et al. (2009, 2013)] to calculate the false and missing tracking rate in our evaluation.

The F_1 score is a metric usually employed as overall score for comparison studies of tracking algorithms [Smeulders et al. (2014)] since it combines *recall* and *precision* in one single score. The F_1 score is defined as:

$$F_1 = 2 \cdot \frac{recall \cdot precision}{recall + precision} \quad (24)$$

In case of guidewire tracking performance, the *recall* can be defined as the fraction of guidewire points correctly detected over the total number of points of the ground truth, namely can be written as a function of the missing tracking rate:

$$recall = \frac{100 - missing\%}{100} \quad (25)$$

The same consideration can be done for the *precision*, which can be defined as the fraction of guidewire points correctly detected over the total number of points detected. Thus, the *precision* can be written as a function of the false tracking rate:

$$precision = \frac{100 - false\%}{100} \quad (26)$$

Finally, the tip error is the distance between the tip of the tracked guidewire and the ground truth of the tip in a particular frame. The tip precision of a sequence is the mean of tip errors for all the frames in the sequence.

4.2. SEGlets

Extensive evaluation of the SEGlets is conducted on the dataset described in Sec. 3.1 to demonstrate the robustness of the proposed features for guidewire tracking. The metrics used to measure the performance of this feature are the missing and false detection rates. They are calculated similarly to the false and missing tracking rate; however, instead of using points on the tracked guidewire, the points detected as SEGlets are used.

The false and missing detection rates of the SEGlets are reported in Table 3 for each sequence of the dataset. The mean values of these metrics considering all the sequences on the dataset except sequence 9 where tracking failures occurred are 5.59% and 15.51%, respectively for the false and

missing detection rates. The reported false and missing detection rates for the guidewire segmentation method presented in [Barbu et al. (2007)] are 10% and 22%, while for the method presented in [Honnorat et al. (2010)] are 13.2% and 25.8%. Although it is not possible to perform a fair comparison between SEGlets and these automatic guidewire segmentation methods since the error metrics were estimated using results from different datasets and these methods cannot rely on temporal information, SEGlets indicatively reduces the false and missing detection rates by 44% and 30% for the method presented in [Barbu et al. (2007)] and by 58% and 40% for the one proposed in [Honnorat et al. (2010)]. These results highlight the robustness of SEGlets as feature for guidewire tracking, and the importance of considering temporal information of the tool segments and background structures to improve guidewire segmentation.

Table 3: The mean of the false and missing detection rates of the proposed SEGlets features are shown for each sequence of the dataset.

Seq	False (%)	Missing (%)
1	4.04	14.69
2	5.54	14.22
3	6.04	12.11
4	6.40	4.54
5	2.42	12.68
6	3.21	19.40
7	9.21	16.06
8	4.45	16.21
9	46.44	36.64
10	7.47	12.91
11	2.69	17.86
12	10.00	29.90

4.3. Tracking

A comparative study between the proposed method (GSEG), the discrete tracking method (DT) [Heibel et al. (2009)], its extension (DT*), the interventional tool tracking method (ITT) [Heibel et al. (2013)], its extension (ITT*) and the geometric-iconic method (G-I) [Honnorat et al. (2011)] was conducted on the clinical dataset described in Sec. 3.1. The mean, standard deviation and median of the guidewire tracking precision in pixels of GSEG,

DT^* , ITT^* and G-I methods are reported in Table 4 for each sequence of the dataset while the mean of the false and missing tracking rates, and the F_1 score are shown in Table 5. Looking only at the mean of the guidewire tracking precision (see Table 4 and Figure 6(a)) it can be concluded that GSEG and G-I reach similar tracking performances in the dataset. This metric represents, however, an asymmetric distance of the tracked guidewire to the ground truth annotations and it does not provide any information of the distance of the ground truth to the tracked guidewire. If we consider the sequence 1, 5, 6, 7, 10 and 12 where GSEG has a lower tracking precision than DT^* , ITT^* and G-I methods (see Table 4), GSEG reaches the lowest missing tracking rates achieving a great improvement in the overall detection of the guidewire compared to DT^* , ITT^* and G-I methods (see Table 5). This is not evident, however, in the guidewire tracking precision, which is proportional to the false tracking rate rather than the missing tracking rate. In other words, DT^* , ITT^* and G-I methods generate a tracked guidewire that is closer to the ground truth with respect to the one from GSEG for the aforementioned sequences but they do not detect the guidewire as completely as the GSEG does. This factor is considered, however, by the F_1 score which combines the missing and false tracking rates in an overall score that measures the performance of the tracking methods. This metric is reported in Table 5 and Figure 6(b). The proposed GSEG outperforms overall DT, DT^* , ITT, ITT^* and G-I methods resulting in the highest F_1 score for the majority of the sequences. From Figure 6(b) it is also possible to observe that DT^* and ITT^* outperform the original methods DT and ITT pointing out the need of a growing step for the spline model which can cope with elongations of the guidewire.

If we consider the overall mean of the false tracking rates for the 12 sequences, the use of GSEG decreases the false tracking rate by 21% with respect to DT^* , 23% with respect to ITT^* , and 5% with respect to G-I. Moreover, the use of GSEG decreases the missing tracking rate by 38% compared to DT^* , 46% compared to ITT^* , and 8% compared to G-I considering the overall mean of the missing tracking rates.

Regarding the performance of the methods to track the tip of the tool, the tip precision, min and max tip errors of GSEG, DT^* , ITT^* and G-I methods are shown in Figure 7 for each sequence of the dataset. The tip errors for the last frame of each sequence are reported in Figure 8. These results demonstrate that the proposed method has a better accuracy to track the tip of the guidewire along the sequences compared to DT^* , ITT^* and

Table 4: The mean, standard deviation and median of the guidewire tracking precision are shown for the proposed GSEG, the extended discrete tracking method DT*, the extended interventional tool tracking method ITT* and the geometric-iconic method G-I for the 12 sequences. The metrics reported are in image pixels (px).

Seq	Mean \pm Std (px)				Median (px)			
	GSEG	DT*	ITT*	G-I	GSEG	DT*	ITT*	G-I
1	0.91 \pm 0.57	0.65 \pm 0.10	0.70 \pm 0.11	0.71 \pm 0.18	0.73	0.66	0.69	0.68
2	1.13 \pm 0.76	0.75 \pm 0.21	1.08 \pm 0.47	0.98 \pm 0.34	0.88	0.70	0.94	0.95
3	1.04 \pm 1.05	2.60 \pm 3.39	2.68 \pm 3.44	1.71 \pm 1.39	0.60	1.13	1.09	1.24
4	0.56 \pm 0.26	0.76 \pm 0.30	2.39 \pm 3.13	0.77 \pm 0.29	0.50	0.70	0.86	0.71
5	0.77 \pm 0.46	0.49 \pm 0.12	0.45 \pm 0.11	0.46 \pm 0.11	0.60	0.46	0.42	0.43
6	0.97 \pm 0.15	0.72 \pm 0.10	1.24 \pm 0.77	0.72 \pm 0.09	0.97	0.71	0.80	0.72
7	1.53 \pm 2.38	0.94 \pm 0.53	0.80 \pm 0.31	0.75 \pm 0.26	0.82	0.77	0.72	0.67
8	1.02 \pm 0.57	38.05 \pm 45.30	28.71 \pm 36.00	41.06 \pm 41.41	0.82	18.93	12.76	31.55
9	15.82 \pm 15.87	17.19 \pm 19.10	18.84 \pm 18.23	10.55 \pm 17.94	8.73	4.07	16.88	1.58
10	1.70 \pm 1.61	0.66 \pm 0.11	0.64 \pm 0.10	0.67 \pm 0.12	0.87	0.64	0.63	0.65
11	0.87 \pm 0.31	2.85 \pm 1.58	0.70 \pm 0.55	0.59 \pm 0.12	0.76	3.65	0.60	0.57
12	2.44 \pm 4.06	1.06 \pm 0.57	1.10 \pm 0.36	1.18 \pm 1.13	0.69	0.87	1.02	0.87

G-I. In fact, GSEG improves the tip precision by 52% compared to DT*, 57% compare to ITT*, and 35% compare to G-I if we consider the overall mean of the tip precision for the 12 sequences. Considering the overall mean of the tip errors for the last frame, GSEG decreases the tip errors by 62% with respect to DT*, 66% with respect to ITT*, and 32% with respect to G-I.

Finally, it can be observed that although sequence 1 is characterized by the lowest MeanD (see Table 1), DT*, ITT* and G-I have relatively high missing tracking rate as shown in Table 5. This can be explained by the difficulty of the growing strategy to cope with the elongations of the guidewire that do not follow straight trajectories, as it can be observed in Figure 10 where the tool reaches a hook-like shape. Contrary, in sequence 11 G-I recovers the elongation of the guidewire where its motion is characterized by a straighter trajectory. In addition, in sequence 8 the DT*, ITT* and G-I methods do not follow the guidewire during its fast pulling back (see Figure 10) resulting on high values of tracking precision, false tracking rates and tip errors. These are largely due to the limitation of the regularization term that does not allow abrupt shortenings of the spline that models the tool. Finally, the GSEG, DT*, ITT* and G-I methods failed to track the tool in sequence 9 where the presence of another tool rendered these methods to fail. During

Table 5: The mean of the false and missing tracking rates, and the F_1 score are shown for the proposed GSEG, the extended discrete tracking method DT*, the extended interventional tool tracking method ITT* and the geometric-iconic method G-I for the 12 sequences. The false and missing tracking rates are in percentages (%) while the F_1 score has a value between 0-1 where 1 is the best and 0 the worst value.

Seq	False (%)				Missing (%)				F_1 Score (0-1)			
	GSEG	DT*	ITT*	G-I	GSEG	DT*	ITT*	G-I	GSEG	DT*	ITT*	G-I
1	2.62	0.20	0.31	0.77	10.16	23.33	21.62	16.99	0.935	0.867	0.878	0.904
2	2.41	1.66	3.86	3.79	2.40	1.51	1.38	1.21	0.976	0.984	0.974	0.975
3	3.17	7.37	7.61	6.13	0.80	0.80	0.87	0.45	0.980	0.958	0.956	0.966
4	1.04	1.91	6.71	1.79	1.13	0.95	1.51	0.10	0.989	0.986	0.958	0.990
5	1.57	0.92	0.59	0.56	0.53	2.01	2.05	2.37	0.990	0.985	0.987	0.985
6	0.67	0.00	3.72	0.06	13.12	36.29	38.41	19.72	0.927	0.778	0.751	0.890
7	4.17	3.24	1.54	1.17	16.16	22.78	31.72	19.75	0.894	0.859	0.806	0.886
8	2.93	30.01	26.41	35.29	8.61	8.49	8.90	9.36	0.941	0.793	0.814	0.755
9	36.87	31.44	36.14	19.27	27.24	23.09	25.84	12.29	0.676	0.725	0.686	0.841
10	4.53	0.17	0.00	0.01	2.21	22.50	15.73	8.37	0.966	0.873	0.915	0.956
11	1.17	8.08	0.59	0.23	10.46	14.89	14.97	8.58	0.940	0.884	0.917	0.954
12	7.55	2.14	2.24	3.10	23.52	30.98	52.08	27.58	0.837	0.809	0.643	0.829

the error calculation of sequence 9 we considered its end the frame where these methods fail.

Analysis of Tracking Failure. To illustrate how tracking may fail for the proposed method, a qualitative evaluation of an example case is shown in Figure 9(a). In this particular case, the tracking hypotheses broke down due to an extremely large distance between two adjacent SEGlets. The distance involved is greater than κ , i.e. parameter that describes the maximum distance between two SEGlets, leading to tracking failures in several frames until the method was eventually able to recover the whole shape of the guidewire thanks to its explicit tip growing strategy. The issue can be avoided by introducing an “on the fly” adaptive estimation of those parameters responsible for the generation of the tracking hypotheses, e.g. κ and θ . The tuning can be driven by the score of the hypotheses as well as introducing constraints on the maximum change in length of the tool which can potentially overcome such extreme cases. As shown in Figure 9(b), the tracking result is improved by increasing the value of κ (in this case by 20 pixels).

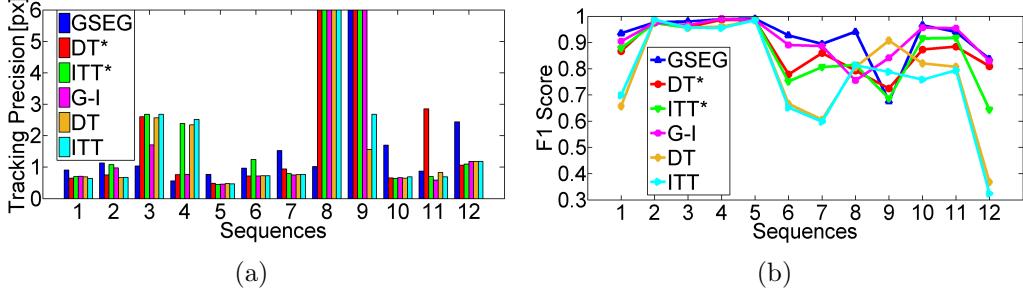


Figure 6: The mean of the tracking precision (a) and F_1 score (b) are shown for the proposed GSEG, the discrete tracking method DT, its extension DT*, the interventional tool tracking method ITT, its extension ITT* and the geometric-iconic method G-I for the 12 sequences. For sequences 8 and 9 the tracking precision of all methods are not fully displayed due to tracking failure. The F_1 score has a value between 0 and 1 where 1 is the best and 0 the worst performance.

5. Conclusions

In this paper, we have introduced a novel method for robustly tracking guidewires under large deformations. The proposed SEGlets have shown to be reliable for guidewire tracking, overcoming some of the limitations of the current data terms, as evident from the promising performance metrics compared to current state-of-the-art. The results also highlight the importance of considering temporal information of the tool segments and background structures to improve guidewire segmentation.

The proposed guidewire tracking formulation based on the organization of the SEGlets in plausible guidewire shapes and a novel tool model appears to be robust and flexible in presence of high dynamic catheterization sequences recovering great elongations and shortenings of the tool in contrast to the spline model which suffers from the limitations of the current regularization terms that constrains the flexibility of the spline. Furthermore, the proposed formulation explicitly models elongation of the guidewire allowing the recovery of tool motions that do not follow straight trajectories. Finally, SEGlets provide a versatile tool for tracking curvilinear objects in fluoroscopy images since they can be implemented using any line-like feature detectors and employed as a robust data term in other algorithms.

The proposed tool model can be enhanced further by integrating common guidewire trajectories that can be learned offline, allowing extensive mod-

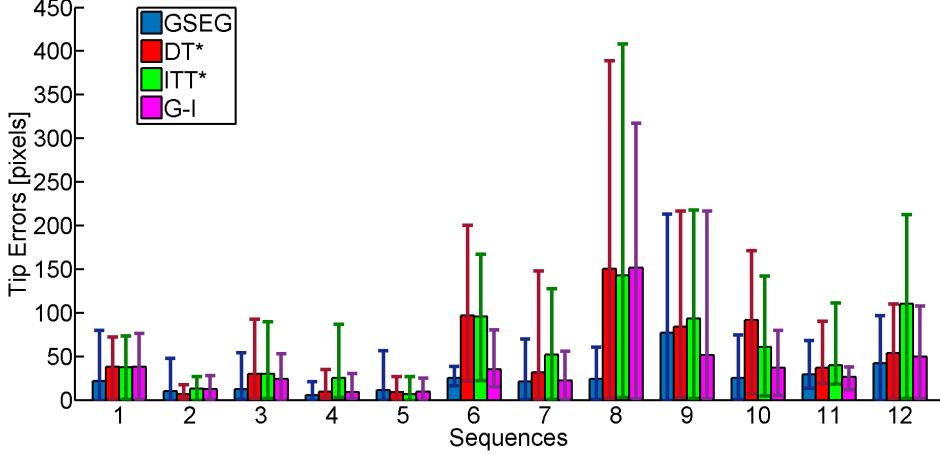


Figure 7: The tip precision, min and max tip errors for the proposed GSEG, the extended discrete tracking method DT*, the extended interventional tool tracking method ITT* and the geometric-iconic method G-I for the 12 sequences. The metrics are in image pixels.

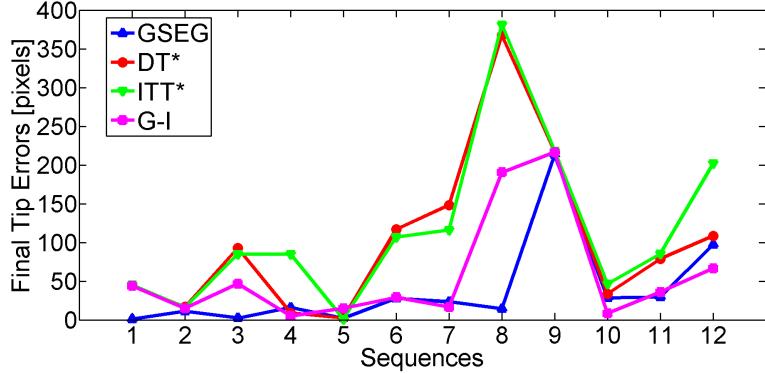


Figure 8: The tip errors for the proposed GSEG, the extended discrete tracking method DT*, the extended interventional tool tracking method ITT* and the geometric-iconic method G-I for the last frame of each sequence. The tip errors are in image pixels.

elling of the tool dynamics. In fact, endovascular procedures share similar workflows and particular guidewire manipulations appear frequently across different sequences, such as tool pull-back/pushing manoeuvres. Those could be captured in a statistical motion model that can be incorporated as a prior in the tool model. Furthermore, automatic guidewire detection can be incorporated into the tracking algorithm as an initialization or reinitialization tool, especially for those clinical applications where tracking initialization is

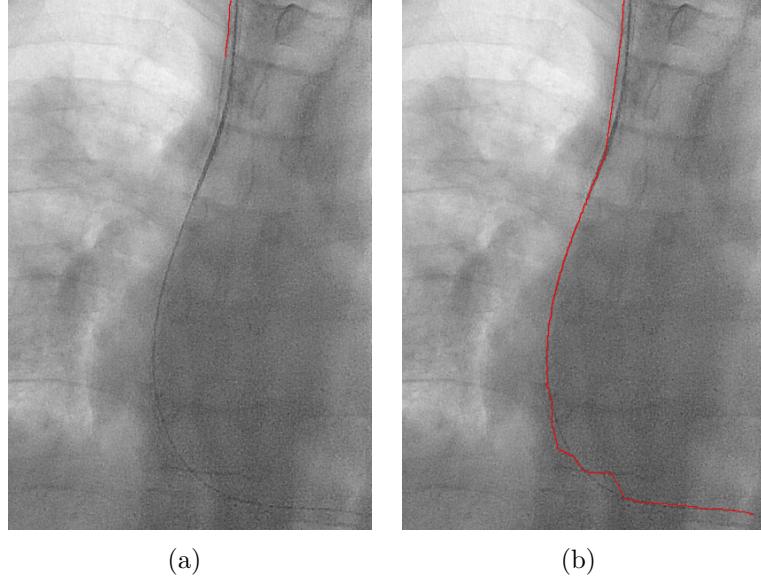


Figure 9: A tracking failure which occurred on 1 additional sequence is shown in (a). In this case, the organisation of the SEGlets in tracking hypotheses is early interrupted. The tracking result is improved by increasing the value of κ by 20 pixels as shown in (b).

required often, e.g., when a procedure is interrupted.

In case a fluoroscopic sequence is characterized by a high level of noise and the guidewire is only partially visible in the images, a robust detection and organization of the SEGlets can be difficult to achieve due to a high fragmentation of the segments. Furthermore, the presence of excessive motion and deformation at the “base” of the guidewire, which represents the starting point for the organization of the SEGlets, can lead to inaccurate tracking hypotheses. Under these challenging scenarios, hybrid approaches by integrating the proposed method with spline model-based methods can be investigated. This would combine the flexibility of the proposed method and its robustness with a stronger prior given by model-based approaches. Finally, in case the endpoint of the tool does not reach the image border, a “base” detection step could be explored in order to further improve the robustness of the method.

It should be noted that the current implementation of the algorithm has not been optimized for high speed computation yet. However, many components of the algorithm can be parallelized, such as the detection of the

SEGlets and the evaluation of the hypotheses. Thus, the computational time of the algorithm, which is around 1-2 frames per second without considering the detection of the line-like features, would highly benefit by parallel computing solutions with either CPU or GPU-based architectures. These solutions can be explored for real-time clinical applications. Current frame rates of fluoroscopy are around 7 to 15 frames per second so this is easily achievable.

In conclusion, the contribution of our work is demonstrated in terms of improvement or comparable results of the false and missing tracking rates, F_1 score, and tip precision respect to three state-of-the-art guidewire tracking methods [Heibel et al. (2009, 2013); Honnorat et al. (2011)], as assessed by using a clinical dataset composed by 12 interventional sequences. Although good performance has been shown in multiple clinical applications, the evaluation of the proposed method in the presence of diverse fluoroscopic sequences and interventional tools should be further investigated.

Acknowledgment

The authors would like to thank Dr Nicolas Honnorat for providing a C++ implementation of [Honnorat et al. (2011)] and helping during the implementation of the guidewire tracking software library.

References

- Atasoy, S., Groher, M., Zikic, D., Glocker, B., Waggershauser, T., Pfister, M., Navab, N., 2008. Real-time respiratory motion tracking: roadmap correction for hepatic artery catheterizations, in: SPIE Medical Imaging, pp. 691815–691815–9.
- Baert, S.A.M., Viergever, M., Niessen, W., 2003. Guide-wire tracking during endovascular interventions. *IEEE Trans. Med. Imag.* 22, 965–972.
- Barbu, A., Athitsos, V., Georgescu, B., Boehm, S., Durlak, P., Comaniciu, D., 2007. Hierarchical learning of curves application to guidewire localization in fluoroscopy, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8.
- Bender, F., Groher, M., Khamene, A., Wein, W., Heibel, T., Navab, N., 2008. 3D dynamic roadmapping for abdominal catheterizations, in: Medical Image Computing and Computer-Assisted Intervention, pp. 668–675.

- Bismuth, V., Vaillant, R., Talbot, H., Najman, L., 2012. Curvilinear structure enhancement with the polygonal path image - application to guide-wire segmentation in X-ray fluoroscopy, in: Medical Image Computing and Computer-Assisted Intervention, pp. 9–16.
- Brost, A., Liao, R., Hornegger, J., Strobel, N., 2009. 3-D respiratory motion compensation during EP procedures by image-based 3-D lasso catheter model generation and tracking, in: Medical Image Computing and Computer-Assisted Intervention, pp. 394–401.
- Brost, A., Wimmer, A., Liao, R., Bourier, F., Koch, M., Strobel, N., Kurzidim, K., Hornegger, J., 2012. Constrained registration for motion compensation in atrial fibrillation ablation procedures. *IEEE Trans. Med. Imag.* 31, 870–881.
- Brückner, M., Deinzer, F., Denzler, J., 2009. Temporal estimation of the 3D guide-wire position using 2D X-ray images, in: Medical Image Computing and Computer-Assisted Intervention, pp. 386–393.
- Canny, J., 1986. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 8, 679–698.
- Cao, Y., Wang, P., 2012. An adaptive method of tracking anatomical curves in X-ray sequences, in: Medical Image Computing and Computer-Assisted Intervention, pp. 173–180.
- Frangi, A., Niessen, W., Vincken, K., Viergever, M., 1998. Multiscale vessel enhancement filtering, in: Medical Image Computing and Computer-Assisted Intervention, pp. 130–137.
- Heibel, H., Glocker, B., Groher, M., Paragios, N., Komodakis, N., Navab, N., 2009. Discrete tracking of parametrized curves, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1754–1761.
- Heibel, H., Glocker, B., Groher, M., Pfister, M., Navab, N., 2013. Interventional tool tracking using discrete optimization. *IEEE Trans. Med. Imag.* 32, 544–555.
- Hoffmann, M., Brost, A., Jakob, C., Bourier, F., Koch, M., Kurzidim, K., Hornegger, J., Strobel, N., 2012. Semi-automatic catheter reconstruction

- from two views, in: Medical Image Computing and Computer-Assisted Intervention, pp. 584–591.
- Honnorat, N., Vaillant, R., Paragios, N., 2010. Guide-wire extraction through perceptual organization of local segments in fluoroscopic images, in: Medical Image Computing and Computer-Assisted Intervention, pp. 440–448.
- Honnorat, N., Vaillant, R., Paragios, N., 2011. Graph-based geometric-iconic guide-wire tracking, in: Medical Image Computing and Computer-Assisted Intervention, pp. 9–16.
- Koller, T., Gerig, G., Szekely, G., Dettwiler, D., 1995. Multiscale detection of curvilinear structures in 2-D and 3-D image data, in: IEEE International Conference on Computer Vision, pp. 864–869.
- Mazouer, P., Chen, T., Zhu, Y., Wang, P., Durlak, P., Thiran, J.P., Comaniciu, D., 2009. User-constrained guidewire localization in fluoroscopy, in: SPIE Medical Imaging, pp. 72591K–72591K.
- Pauly, O., Heibel, H., Navab, N., 2010. A machine learning approach for deformable guide-wire tracking in fluoroscopic sequences, in: Medical Image Computing and Computer-Assisted Intervention, pp. 343–350.
- Rolls, A., Riga, C., Bicknell, C., Stoyanov, D., Shah, C., Van Herzele, I., Hamady, M., Cheshire, N., 2013. A pilot study of video-motion analysis in endovascular surgery: development of real-time discriminatory skill metrics. *Eur. J. Vasc. Endovasc. Surg.* 45, 509–515.
- Slabaugh, G., Kong, K., Unal, G., Fang, T., 2007. Variational guidewire tracking using phase congruency, in: Medical Image Computing and Computer-Assisted Intervention, pp. 612–619.
- Smeulders, A., Chu, D., Cucchiara, R., Calderara, S., Dehghan, A., Shah, M., 2014. Visual tracking: An experimental survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 36, 1442–1468.
- Vandini, A., Bergeles, C., Lin, F.Y., Yang, G.Z., 2015. Vision-based intraoperative shape sensing of concentric tube robots, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, [In Press].

- Vandini, A., Giannarou, S., Lee, S.L., Yang, G.Z., 2013. 3D robotic catheter shape reconstruction and localisation using appearance priors and adaptive C-arm positioning, in: Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions, pp. 172–181.
- van Walsum, T., Baert, S.A.M., Niessen, W., 2005. Guide wire reconstruction and visualization in 3DRA using monoplane fluoroscopic imaging. *IEEE Trans. Med. Imag.* 24, 612–623.
- Wang, P., Chen, T., Zhu, Y., Zhang, W., Zhou, S., Comaniciu, D., 2009. Robust guidewire tracking in fluoroscopy, in: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 691–698.
- Wang, P., Liao, W.s., Chen, T., Zhou, S., Comaniciu, D., 2010. Graph based interactive detection of curve structures in 2D fluoroscopy, in: *Medical Image Computing and Computer-Assisted Intervention*, pp. 269–277.
- Wu, X., Housden, J., Ma, Y., Razavi, B., Rhode, K., Rueckert, D., 2015. Fast catheter segmentation from echocardiographic sequences based on segmentation from corresponding X-ray fluoroscopy for cardiac catheterization interventions. *IEEE Trans. Med. Imag.* 34, 861–876.
- Zhang, T.Y., Suen, C.Y., 1984. A fast parallel algorithm for thinning digital patterns. *Commun. ACM* 27, 236–239.

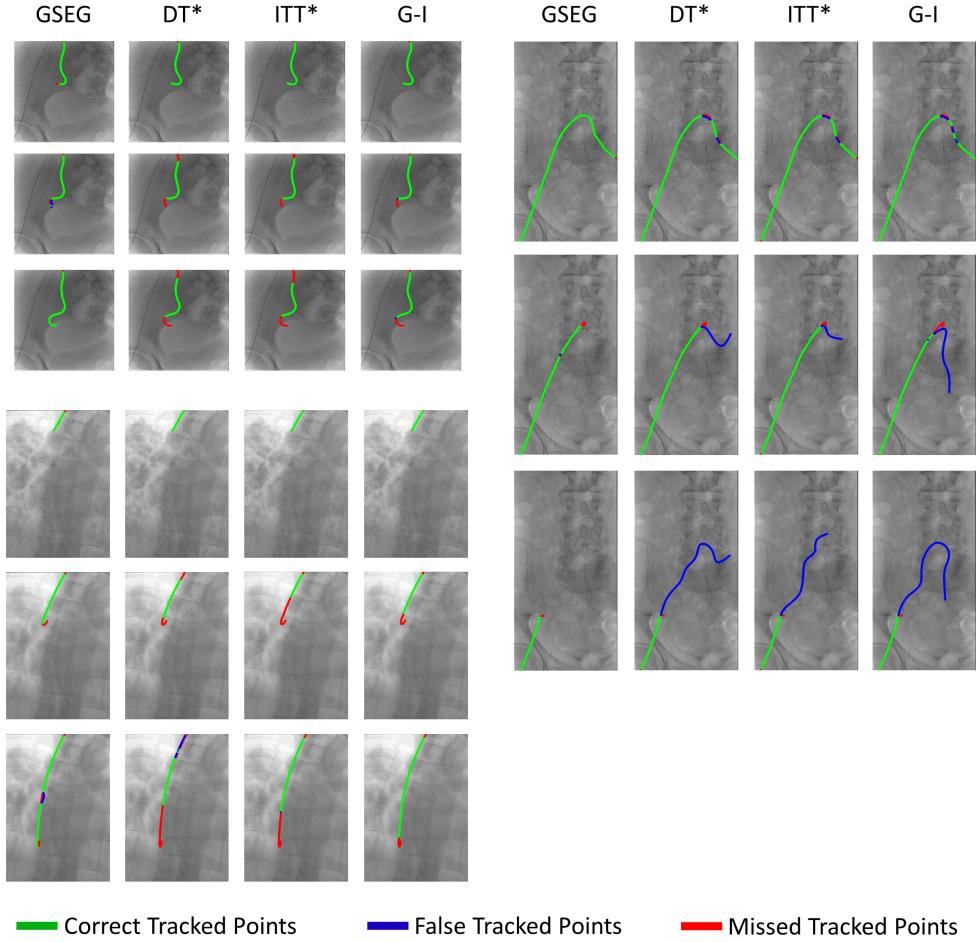


Figure 10: Qualitative tracking results for the proposed GSEG, the extended discrete tracking method DT*, the extended interventional tool tracking method ITT* and the geometric-iconic method G-I for the sequence (starting from the top left corner and proceeding counter-clockwise) 1, 7, and 8 are reported. For each method only the results from the second, middle and last frame of the sequence are shown.