USING NEEDLE DETECTION AND TRACKING FOR MOTION COMPENSATION IN ABDOMINAL INTERVENTIONS

Peng Wang[a], Marcus Pfister[b], Terrence Chen[a], Dorin Comaniciu[a]

a). Siemens Corporate, Corporate Research, 755 College Road East, Princeton NJ, U.S.A.
b). Siemens Healthcare, Siemensstr. 1, Forchheim, Germany

ABSTRACT

In this paper, we present a method of using the needle detection and tracking to compensate breathing motion in 2D fluoroscopic videos. The method can robustly detect and tracking needles, even with the presence of image noises and large needle movements. The method first introduces an offline learned needle segment detector that detects needle segments at individual frames. Based on detected needle segments, a needle is interactively detected at the beginning of an intervention, and then is automatically tracked based on a probabilistic tracking framework. A multi-resolution kernel density estimation is applied to handle large needle movements efficiently and effectively. Experiments on phantom and clinical sequences demonstrate that the method can successfully track needles in fluoroscopy, and can provide motion compensation for abdominal interventions.

Index Terms— needle, tracking, detection, fluoroscopy

1. INTRODUCTION

Many applications in image guided interventions require a fast and robust method to compensate breathing motions. For example, the predominant breathing motion need to be compensated in order to accurately register a 3D model, which may be acquired from pre-operational data, to 2D fluoroscopy that is acquired real-time [1]. In abdominal interventions, needle movements, when without pulling, is mainly caused by breathing motions. This paper presents a method of using needle detection and tracking in fluoroscopy to compensate 2D breathing motions in abdominal interventions.

Although many existing methods have been developed for different types of object tracking tasks [2], the needle tracking remains a challenging problem that needs specialized solutions due to some unique characteristics of interventional procedures. First, the image quality of 2D fluoroscopy is usually poor because of desirable low radiation. Traditional edge and ridge detectors will produce many false detections, while missing thin needles. Second, since a needle is a thin one-dimensional structure, it is sometimes difficult to distinguish from cluttered background and other line structures in 2D fluoroscopic images, as shown in Fig. 1. The methods that use regional features such as holistic intensity, textures, and color
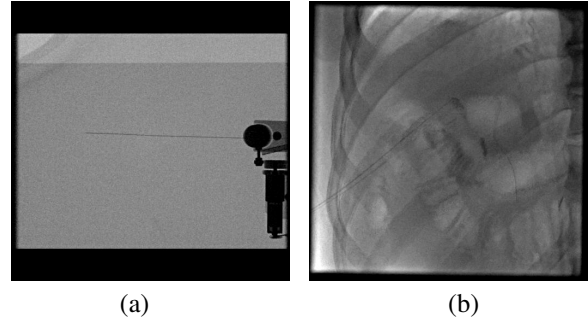

Fig. 1. Needles for interventions. (a): a phantom needle; (b): a fluoroscopic image containing a needle in a clinical case

histogram [2], cannot detect and track the thin needle well. Third, the need of real-time speed in abdominal interventions poses additional challenges to the needle tracking.

The existing tracking methods for image guided interventions can be categorized into electromagnetic based [3] and image based methods[1, 4]. The electromagnetic tracking can provide 3D positions at near real-time with good accuracy, but the high expense and needs of additional devices limit its applications. Image based methods, on the other hand, provide an inexpensive alternation without major modifications in an imaging procedure. However, traditional image based tracking methods, such as the template matching [2, 1] and active contour based methods [5], have difficulties in handling the low image quality and large breathing motions that are are typically seen in clinical cases.

In this paper, we present a novel method that addresses the above mentioned challenges, and demonstrates its robustness through extensive experiments of using both phantom and real clinical data. The method includes two major components: learning based needle detection, and multi-resolution probabilistic needle tracking. At each individual frame, the learning based method detects needle segments, based on which the needle is interactively detected at the beginning of interventions, and is continuously tracked throughout the intervention under a multi-resolution probabilistic tracking framework. Finally, the motion parameters obtained by the needle tracking provide motion compensation for abdominal interventions. The presented method is a general framework, and can be applied to other tasks, e.g., catheter tracking. Ex-
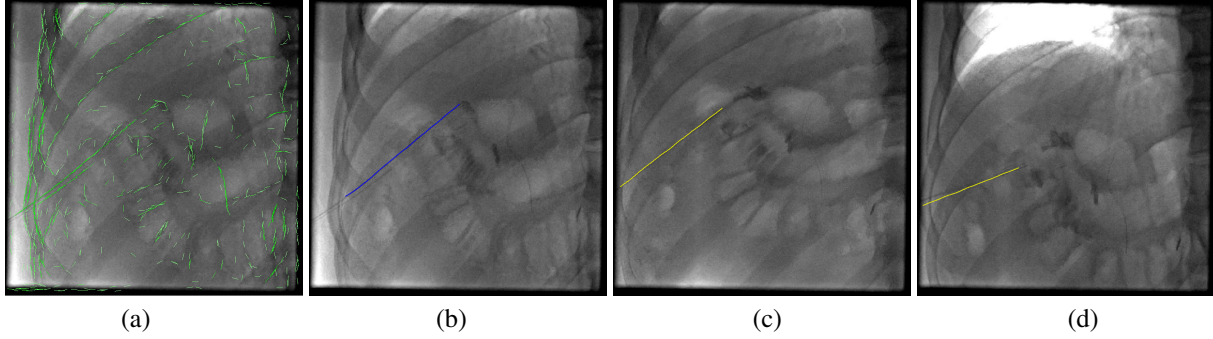
Fig. 2. Steps in the needle detection and tracking. (a): the needle segment detections at the first frame, shown in green; (b): interactive detection results, shown in blue; (c):a tracked needle shown in yellow; (d) a tracked needle at another frame.

periments on phantom and real clinical data demonstrate that the presented method can effectively and efficiently track needles, even with the presence of image noises and large needle movements.

2. METHOD

2.1. Overview

The workflow of the presented method is illustrated in Fig. 2 and Fig. 3, and is briefly summarized here. As shown in Fig. 3, the learning based detection serves two purposes in our method: it first provides features for the interactive needle detection to initialize the needle tracker at the first frame, and then continuously provides primitive features for the subsequent tracking. It has been shown in many applications that the learning based methods are more effective than those based on low level filtering methods, and are well suited to the image guided interventions [4, 6].
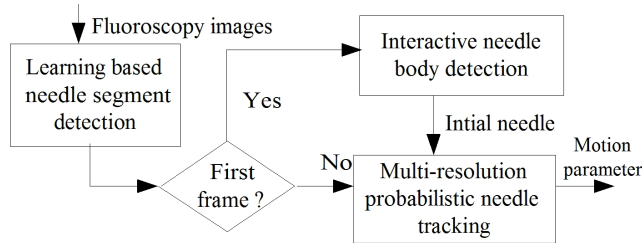

Fig. 3. Illustration of the needle detection and tracking method

The multi-resolution needle tracking continuously identifies the position of the needle, to provide the breathing motion compensation for abdominal interventions. A kernel density smoothing method has been incorporated into a probabilistic framework formalized for the needle tracking. By varying the kernel bandwidth in the kernel based density smoothing, the needle tracking can effectively track large needle movements from coarse to fine, while maintaining satisfactory accuracy.

2.2. Needle Detection

2.2.1. Learning based needle segment detection

The purpose of needle segment detection is to identify if an image patch belongs to a part of needle. It serves as a fundamental step toward a whole needle body detection, and also provides primitive features for the subsequent needle tracking. Compared to conventional edge or ridge detection methods based on image derivatives, the learning based detectors are constructed based on machine learning techniques from a set of collected data, which includes both objects (positive samples) and background (negative samples), and can be easily adapted to specific detection tasks. The advantages make the learning based detectors well suited to image guided interventions.

Our method uses the probabilistic boosting tree (PBT) [7] to build the needle segment detector. PBT is a supervised learning method that extends the AdaBoost [8] into a tree structure with a better generalization capability in object detection tasks. For the consideration of speed, computationally efficient Haar features [9] are extracted from images to be used in the PBT classifier. For training, many images containing thin line structures that are similar to needle are collected. Please refer to literature [8, 7, 9] for the details about AdaBoost, PBT, Haar features, and training procedures.

2.2.2. Interactive needle body detection

The needle tracking can be initialized either by automated detection, or interactive detection. To ensure the accuracy and robustness, a user-constrained interactive detection method [10] is used. In the interactive detection, at least two clicks, one at each end of a needle, are required to initialize a needle. The detected segments are used to construct a graph, based on which the dynamic programming is applied to search the best path between two user clicks. In case the initial detection given two points are not satisfactory, additional user clicks are provided to constrain the algorithm and to obtain refined detection results.

2.3. Multi-resolution Needle Tracking

2.3.1. A Probabilistic Tracking Framework

Most tracking tasks can be formalized in a probabilistic framework, i.e., the Bayesian inference framework [2]. Following the same principle, we formalize the needle tracking as to maximize the posterior probability of a tracked needle given 2D fluoroscopic images. In this framework, a needle hypothesis Γ_t at the t-th frame is deformed from the needle at the $t-1$-th frame:

$$\Gamma_t(\mathbf{u}) = T(\Gamma_{t-1}, \mathbf{u}) \tag{1}$$

where T is a needle shape transformation function, and $\mathbf{u}$ is its parameter. Based on a Markov assumption of $P(\Gamma_t|\Gamma_{t-1}, \mathbf{Z}_{t-1}) = P(\Gamma_t|\Gamma_{t-1})$, the posterior probability $P(\Gamma_t|\mathbf{Z}_t)$ is given in Eqn. (2).

$$P(\Gamma_t|\mathbf{Z}_t) \quad \propto \quad P(\Gamma_t)P(\mathbf{Z}_t|\Gamma_t(\mathbf{u})). \tag{2}$$

$\mathbf{Z}_t$ denotes the fluoroscopic image at time t. The tracked needle $\hat{\Gamma}_t$ is then estimated as the needle candidate that maximizes the posterior probability, i.e., $\hat{\Gamma}_t = \underset{\Gamma_t}{argmax} P(\Gamma_t|\mathbf{Z}_t)$.

In Eqn. (2), $P(\Gamma_t)$ is a prior probability, which can be propagated from previous tracking results. We model the prior probability as:

$$P(\Gamma_t) = \frac{1}{\sqrt{2\pi}\sigma_\Gamma} exp(\frac{-|D(\Gamma_t, \Gamma_{t-1})|^2}{2\sigma_\Gamma^2}), \tag{3}$$

where $D(\Gamma_t, \Gamma_{t-1})$ is the average of the shortest distances from points on $\Gamma_t(\mathbf{u})$ to the shape template $\Gamma_{t-1}(\mathbf{x})$. A large kernel size σ_Γ is chosen to allow for a large needle movement between two frames.

We define the measurement model $P(\mathbf{Z}_t|\Gamma_t(\mathbf{u}))$ as:

$$P(\mathbf{Z}_t|\Gamma_t(\mathbf{u})) = \frac{1}{|\Gamma_t(\mathbf{u})|} \sum_{\mathbf{x}_i \in \Gamma_t(\mathbf{u})} P(\mathbf{Z}_t|\mathbf{x}_i) \tag{4}$$

where $P(\mathbf{Z}_t|\mathbf{x}_i)$ is the measurements at individual points on a needle, and $|\Gamma_t(\mathbf{u})|$ is the needle curve length. We adapt the learning based detection results to define the measurements at individual needle segments. It has been shown [11] that the numeric output of an AdaBoost classifier, denoted as $f(\mathbf{x}_i)$ given a image patch at the position $\mathbf{x}_i$, can be interpreted into a probabilistic measurement:

$$P(\mathbf{Z}_t|\mathbf{x}_i) \quad \propto \quad \frac{e^{f(\mathbf{Z}_t, \mathbf{x}_i)}}{e^{-f(\mathbf{Z}_t, \mathbf{x}_i)} + e^{f(\mathbf{Z}_t, \mathbf{x}_i)}}. \tag{5}$$

2.3.2. Multi-resolution Needle Tracking

The needle tracking aims at recovering the motion parameter of a needle between two successive frames. In the needle tracking, the parameter $\mathbf{u}$ in Eqn. (1) contains global translation, rotation and scale changes, i.e., $\mathbf{u} = (c, r, \theta, s_c, s_r)$,

where c, r and θ are the translation and rotation parameters, and s_c and s_r are the scaling factors. Therefore, the motion parameters are estimated by maximizing the posterior probability under the parametric motion model:

$$\hat{\mathbf{u}}_t = \underset{\mathbf{u}}{arg} \hat{\Gamma_t(\mathbf{u})} = \underset{\mathbf{u}}{argmax} P(\Gamma_t(\mathbf{u})|\mathbf{Z}_t) \tag{6}$$

Exhaustively searching over all possible parameters $\mathbf{u}$ is time consuming, and does not satisfy the real-time requirements in interventions. We here present a multi-resolution scheme, which is based on a kernel-based measurement smoothing method, to expedite the needle tracking. In the kernel-based estimation, measurements are made at a set of samples $\mathbf{x}_j^s$, instead of a whole parameter space. In the method, $\mathbf{x}_j^s$ are those points that are detected as needle segments. We can conveniently assume the Markov conditional independence that the observations at sampling points $\mathbf{x}_j^s$ are independent with the un-sampled points $\mathbf{x}_i$, i.e., $P(\mathbf{Z}_t|\mathbf{x}_i, \mathbf{x}_j^s) = P(\mathbf{Z}_t|\mathbf{x}_j^s)$. Therefore, the kernel-based measurement estimation is represented as Eqn.(7):

$$P(\mathbf{Z}_t|\mathbf{x}_i) = \sum_j P(\mathbf{Z}_t|\mathbf{x}_j^s) G_\sigma(\mathbf{x}_j^s, \mathbf{x}_i), \tag{7}$$

where $P(\mathbf{x}_j^s|\mathbf{x}_i) = G_\sigma(\mathbf{x}_j^s, \mathbf{x}_i)$ is a Gaussian kernel with a bandwidth σ. The kernel-based measurement estimation can obtain smooth measurements in a neighborhood, and reduce computations.

The multi-resolution tracking can be efficiently implemented using variable bandwidths in the kernel-based measurement smoothing. For example, the translations searching is performed at multiple resolutions with searching intervals decreased from coarse to fine, and the corresponding bandwidth in Eqn. (7) varies accordingly. At each resolution, the bandwidth is set the same as the translation searching step. At coarse resolutions, larger kernel bandwidths can help avoid missing tracking; and at fine resolutions, using smaller kernel bandwidths can obtain accurate tracking results. Usually with 3 or 4 resolutions, the tracking converges to the true needle positions. As the result, the needle tracking can run around 8 to 10 frames per second at a Core 2 Duo 2.0GHz computer.

3. RESULTS

The presented tracking method is evaluated on a set of fluoroscopic sequences, including both phantom and real clinical cases. The testing sets contain 3 phantom sequences (totally 395 frames), and 10 clinical sequences (totally 1312 frames). Some tracking results are shown in Fig. 4. The pixel size of the sequences is between 0.15mm and 0.21mm per pixel. To establish a ground truth for evaluation, we carefully annotate the needle body in those sequences. Some sequences contain catheters, instead of needles. Since our method can be directly applied to catheter tracking, these sequences are
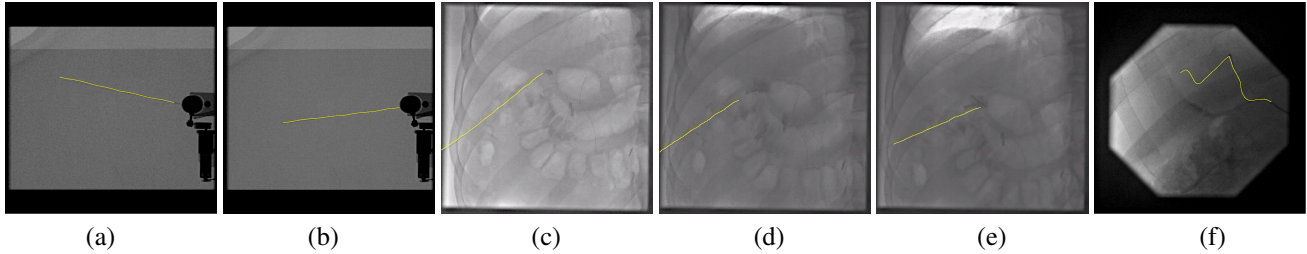
Fig. 4. Some needle tracking results:(a)-(b): tracking rotated a phantom needle; (c)(d)(e): tracking a needle under rotations and translations; (f): extending the needle tracking to catheter tracking

also included in the evaluations. To quantitatively evaluate the performance of needle tracking, we measure both **tracking pixel error**, and **frame tracking rate**. The tracking pixel error is defined as the shorted distance from the points on the tracked needle to the corresponding annotated ground truth. The frame tracking rate is used to evaluate the percentage of frames where the needle tracking provides accurate motion compensations. Here, a frame is asserted to be successfully tracked when the corresponding pixel error is smaller than a threshold (e.g. 10 pixels in the experiment).

Table 1. Quantitative evaluation on clinical data

Initialization	frame tracking rate	Pixel error		
		median	mean	std
Full	96.3%	1.31	2.96	5.16
Partial	94.6%	1.83	3.07	4.71

The quantitative evaluation results on clinical sequences are summarized in Table 1. The needle tracking is evaluated under two experiment settings. In the first experiment, the tracker is initialized at the first frame with the annotation to simulate the situation where the interactive detection is carefully done. As shown in the Table 1, with full initialization, the median tracking error is around 1.3 pixels, and the mean error is around 3 pixels. More than 96% of frames are successfully tracked.
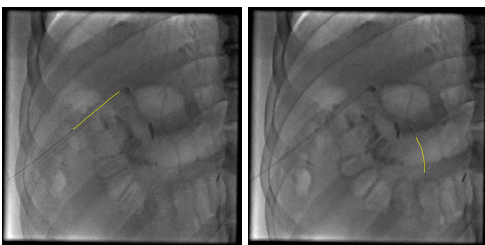

Fig. 5. Partial initializations to test tracking robustness

At the second experiment, a user is asked to randomly initialized the needle at a random length. The user repeated the random initialization 20 times for each clinical sequence, therefore totally 200 tracking results with different initializations have being evaluated. This experiment is to evaluate the tracking performance with partial initializations. As shown in Table 1, under partial initialization, the median tracking error

is around 1.8 pixels, the mean error remains around 3 pixels, and more than 94% frames are successfully tracked.

4. CONCLUSION

This paper presents a framework of needle tracking in fluoroscopy for motion compensation in abdominal interventions. Our method has applied a learning based method to detect needle segments, and to track needle motion from coarse to fine. The experimental results demonstrate that the tracking method has a great potential in clinical applications for breathing motion compensation.

5. REFERENCES

[1] Selen Atasoy, Martin Groher, Darko Zikic, Ben Glocker, Tobias Waggershauser, Marcus Pfister, and Nassir Navab, "Real-time respiratory motion tracking: roadmap correction for hepatic artery catheterizations," 2008, vol. 6918, pp. 691815.1–691815.9, SPIE.

[2] Alper Yilmaz, Omar Javed, and Mubarak Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, pp. 13, 2006.

[3] Zhang H, Banovac F, Lin R, Glossop N, Wood BJ, Lindisch D, Levy E, and Cleary K., "Electromagnetic tracking for abdominal interventions in computer aided surgery," *Comput Aided Surg.*, vol. 11, no. 3, pp. 127–36, 2006.

[4] Peng Wang, T. Chen, Y. Zhu, S. Kevin Zhou, and D. Comaniciu, "Robust guidewire tracking in fluoroscopy," in *CVPR*, 2009.

[5] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1987.

[6] A. Barbu, Vassilis Athitsos, Bogdan Georgescu, Stefan Boehm, Peter Durlak, and Dorin Comaniciu, "Hierarchical learning of curves application to guidewire localization in fluoroscopy," in *CVPR*, 2007.

[7] Zhuowen Tu, "Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering," in *ICCV*, 2005, pp. 1589–1596.

[8] Yoav Freund and Robert E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *European Conference on Computational Learning Theory*, 1995, pp. 23–37.

[9] Paul Viola and Michael Jones, "Robust real-time object detection," *Interntional Workshop on Statistical and Computational Theories of Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[10] P. Mazouer, T. Chen, Y. Zhu, Peng Wang, P. Durlak, J-P Thiran, and D. Comaniciu, "User-constrained guidewire localization in fluoroscopy," in *Medical Imaging: Image Processing*, 2009, Proc. SPIE.

[11] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *The Annals of Statistics*, vol. 28, no. 2, pp. 337–374, 2000.