Vision-based suturing needle tracking with Extended Kalman Filter

M. Ferro[1*], G. A. Fontanelli[2*], F. Ficuciello[2], B. Siciliano[2] and M. Vendittelli[3]

[1]*Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma*
[2]*Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Università di Napoli Federico II*
[3]*Dipartimento di Ingegneria dell'Informazione, Elettronica e Telecomunicazioni, Sapienza Università di Roma*

INTRODUCTION

After the last two decades of lively activity, research in Minimally Invasive Robotic Surgery (MIRS) is focusing on enhancing the degree of autonomy for procedures that are still executed by surgeons, through the introduction of autonomous robot operation modes. These procedures include suturing that, as repetitive and tiring task, usually increases the operation time, costs and complications risks. Automatic suturing presents many challenges related to both perception and manipulation. Left alone the interaction with soft tissues, an automatic suturing procedure has to deal with the difficulty of perception in an unstructured and highly noisy environment for needle and thread detection and tracking. The circular shape of the suturing needle encouraged the development of a family of *model-based* approaches for detection and tracking. Authors in [2] present a RANSAC-based method for needle detection, where 3D needle pose reconstruction is achieved with the use of a stereo camera. However, the method does not run in real time and cannot be used for tracking. In [8], the 3D needle pose is adaptively reconstructed by relying on the observations of needle tip and junction, but tracking is not faced. Finally, [7] presents a method for a colored-needle tracking that involves a partial needle pose reconstruction and the use of markers. However, none of these works takes advantage of the kinematics information available from the robot, that are typically high-frequency and can ease the needle detection and tracking problem.

In this work we propose an approach to needle detection and tracking based on Kalman filtering to combine visual information from a monocular camera with the robot kinematics. Beside providing a fast and reliable needle pose estimation, the proposed method is robust with respect to scene variations as in case of partially needle occlusion or of needle re-grasping operation, as well as external disturbances perturbing the needle pose. In addition, the covariance matrices can be adapted taking into account the particular task that is being performed.

MATERIALS AND METHODS

This work proposes the use of an Extended Kalman Filter (EKF) [5] for the estimation of the needle pose $\mathcal{F}_n$ in Fig. 1a, expressed in the base frame $\mathcal{F}_b$ with origin at the Remote Center of Motion (RCM) of the manipulator holding the needle.

In particular, the filter provides an estimate of the needle pose $\boldsymbol{\zeta} = [\boldsymbol{p}_n, \boldsymbol{q}_n]^T$, being $\boldsymbol{p}_n$ the true needle position, represented by the coordinates of the circle supporting the needle shape, and $\boldsymbol{q}_n$ its quaternion-based true orientation in $\mathcal{F}_b$. The prediction step provides a preliminary estimation of the needle pose through the linear and angular velocities of the gripper provided by the manipulator kinematics. Then, a vision-based 3D pose reconstruction is used in the filter correction step.

Extended Kalman Filter

Determined from robot kinematics the linear and angular velocities $[\boldsymbol{v}_g, \boldsymbol{\omega}_g]^T$ of the gripper in $\mathcal{F}_b$, we consider the following continuous-time process dynamics for the state vector $\boldsymbol{\zeta}$

$$\begin{array}{ll} \dot{\boldsymbol{p}}_n = \boldsymbol{v}_g + \lfloor \boldsymbol{\omega}_{g\times} \rfloor \boldsymbol{r}_{gn} & + \boldsymbol{w}_p \\ \dot{\boldsymbol{q}}_n = \frac{1}{2}\boldsymbol{\Omega}({}^n\boldsymbol{\omega}_g)\boldsymbol{q}_n & + \boldsymbol{w}_q \end{array} \tag{1}$$

where $\lfloor *_\times \rfloor$ denotes the skew-symmetric operator, $\boldsymbol{r}_{gn} = \boldsymbol{p}_n - \boldsymbol{p}_g$ is the relative position of the needle with respect to the gripper, expressed in $\mathcal{F}_b$, ${}^n\boldsymbol{\omega}_g$ is the angular gripper velocity expressed in $\mathcal{F}_n$, and $\boldsymbol{w} = [\boldsymbol{w}_p, \boldsymbol{w}_q]^T \sim \mathcal{N}(0, \boldsymbol{W})$ is the process noise. Details on the form of $\boldsymbol{\Omega}$ can be found in [6]. On the other hand, the measurement model of the filter employs the visual-based pose measurements extracted from the monocular camera: first, a detection algorithm computes the ellipse-shaped projection of the needle on the image plane. Then, the 3D pose is reconstructed from the size and projective reasonings ([3]). So, The measurement model is

$$\boldsymbol{y} = \boldsymbol{\zeta} + \boldsymbol{m} \tag{2}$$

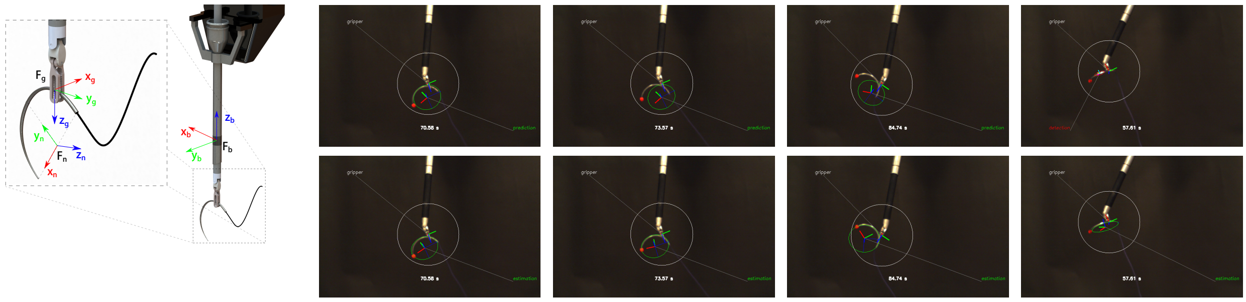
where $\boldsymbol{m} \sim \mathcal{N}(0, \boldsymbol{M})$ is the measurement noise. The error-state vector is defined as

$$\tilde{\boldsymbol{\zeta}} = \begin{bmatrix} \tilde{\boldsymbol{p}} & \delta\tilde{\boldsymbol{\theta}} \end{bmatrix}^T \tag{3}$$

where $\tilde{\boldsymbol{p}} = \boldsymbol{p} - \hat{\boldsymbol{p}}$ is the position error and $\delta\tilde{\boldsymbol{\theta}}$ is the 3×1 small-angle approximation vector of the quaternion orientation error $\delta\tilde{\boldsymbol{q}} = \boldsymbol{q} \otimes \hat{\boldsymbol{q}}^{-1}$, defined to avoid numerical instability issues related to the unit norm constraint on the quaternion vector. Expressing the process dynamics (1) and the measurement model (2) with $\boldsymbol{f}$ and $\boldsymbol{h}$, respectively, the corresponding jacobian matrices, $\boldsymbol{F}$ and $\boldsymbol{H}$, are constant with respect to (3)

$$\boldsymbol{F} = \begin{bmatrix} \lfloor \boldsymbol{\omega}_{g\times} \rfloor & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \lfloor {}^n\boldsymbol{\omega}_{g\times} \rfloor \end{bmatrix} \quad , \quad \boldsymbol{H} = \boldsymbol{I}_6 \ . \tag{4}$$

*Joint first authorship.



(a) (b) (c) (d) (e)

Fig. 1: (1a): Frames of interest in our discussion. (1b)-(1c)-(1d): Prediction failure scenario compared with the vision-based corrected estimation. (1e): Detection failure scenario compared with the estimation. The white circle represents the image area in which the needle is assumed to be found, based on its radius and the depth of the gripper with respect to the camera.

Then, the EKF state estimate is readily computed.

RESULTS

To evaluate the robustness of needle pose estimation with respect to perturbations due to the needle-tissues interactions, we use the simplified experimental setup shown in Fig. 1. Typical vision-related challenges (e.g., shadows, sparkling metal surfaces, small-sized objects) are not considered to focus on the geometric part of the pose reconstruction. So, an RGB segmentation procedure and a least-square fitting are sufficient to extract the projection of the needle on the image plane [4]. In addition, the needle tip has been colored to ease the detection of the projective point required for the 3D pose reconstruction.

The target experimental system is a da Vinci Research Kit robot (DVRK) [1]. Fig. 1 shows some preliminary results to prove the advantages in employing both robot kinematics and visual information for the needle pose estimation. During grasping, the needle is assumed as a rigid body attached to the gripper, and its pose can be predicted through robot kinematics, provided an initial estimate. However, since the needle-gripper transformation is not rigid, external disturbances (e.g., contact with tissues, slippages) can alter its pose, as shown in Fig (1b)-(1c)-(1d), where the needle pose has been explicitly changed. Robot kinematics can not cope with these disturbances, the prediction fails and propagates the error on the next iterations (top). The vision-based correction of our filter allows to detect the needle movements due to disturbances, and adjust the estimation accordingly (bottom). On the other hand, Fig (1e) shows a scenario where the vision-based detection fails, because the projected ellipse of the circular needle is degenerate on the image plane of the camera (top). However, kinematics information provided by the robot allows to maintain a stable estimation of the needle, even when this is not clearly visible. The figures are extracted from the videos that can be found at the link `http://www.diag.uniroma1.it/~labrob/research/ekfNeedleTracking.html`.

CONCLUSION AND DISCUSSION

In this work we propose an EKF-based approach to estimate the pose of the suturing needle during robot-assisted laparoscopic procedures. The filter fuses the kinematic information from the robot proprioceptive sensors with the visual information provided by a monocular endoscopic camera. The use of robot kinematics allows to restrict the region of interest in processing the image, thus speeding up computations, and renders the estimation robust with respect to visual occlusions. On the other hand, visual information allows to catch possible unmodeled motions of the needle, e.g., slippage or movements due to the interaction with tissues. Future work include development of robust image processing algorithms for the detection of the needle in realistic experimental setups. In vitro and ex-vivo experiments are planned to validate the approach in increasingly challenging conditions.

REFERENCES

[1] da Vinci research kit (DVRK) wiki web page.

[2] J. Chang. Robust needle recognition using Artificial Neural Network (ANN) and Random Sample Consensus (RANSAC). *Applied Imagery Pattern Recognition Workshop (AIPR), 2012 IEEE*, pages 1–3, 2012.

[3] D. L. De Ipina, P. R. S. Mendonca, and A. Hopper. TRIP: A low-cost vision-based location system for ubiquitous computing. *Personal and Ubiquitous Computing*, 6(3):206–219, 2002.

[4] A. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least square fitting of ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):476–480, 1999.

[5] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.

[6] N. Trawny and S. I. Roumeliotis. Indirect Kalman Filter for 3D Attitude Estimation A Tutorial for Quaternion Algebra Indirect Kalman Filter for 3D Attitude Estimation A Tutorial for Quaternion Algebra 1 Elements of Quaternion Algebra 1.1 Quaternion Definitions. 2005.

[7] C. Wengert, L. Bossard, C. Baur, G. Székely, and P. C. Cattin. Endoscopic navigation for minimally invasive suturing. *Computer aided surgery : official journal of the International Society for Computer Aided Surgery*, 13(5):299–310, 2008.

[8] F. Zhong, D. Navarro-alarcon, Z. Wang, Y.-h. Liu, T. Zhang, H. M. Yip, and H. Wang. Adaptive 3D Pose Computation of Suturing Needle Using Constraints From Static Monocular Image Feedback. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5521–5526, 2016.