Surgical Tools Pose Estimation for a Multimodal HMI of a Surgical Robotic Assistant

B. Estebanez, E. Bauzano and V.F. Muñoz

Abstract— **The main objective of this paper is to minimize the occluded areas in order to recognize the navigation of the surgeon's tools for a two-arm autonomous robotic system for laparoscopic procedures. This robotic assistant needs the tracking of the surgeon's surgical gestures in order to recognize the current maneuver and to execute the automated tasks of the robot. The surgical tools pose estimation is carried out by a Multiple Extended Kalman Filter (MEKF), where the movement models of the surgical tools depend on the maneuver which is being developed. This information is obtained by a maneuvers recognition system which is a part of the multimodal human machine interface (HMI) of the robot. The method proposed for reducing shadows has been applied to three in-vitro maneuvers which appear in the majority of the surgical protocols. The experiments show the behavior of this method for different time intervals of the occlusions.**

I. Introduction

ONE ambitious challenge of robotic surgical engineering is to substitute the human assistant on laparoscopic surgery [1]; in this way, the human errors derived from this type of techniques can be reduced and the development of the surgical protocols is improved. Therefore, there are robotic assistants that are able to focus the camera on the area required and others have arms which move additional surgical tools to collaborate with the surgeon ([2]-[3]). Thus, the robotic systems assure more accuracy than human assistants ([4]-[5]) but using them imply the need of a friendly channel of communication with the surgeon.

New interfaces make this communication as comfortable as possible and reduce the current limitations in the communication with surgical robots. In this way, some systems used joystick devices [6], others made automatic movements to follow the surgeon tool with the endoscope [7], they receive the surgeon orders through a gyroscope attached to the surgeon head [8], or are guided with the voice [9]-[10]. There are also works where an artificial vision system is proposed in order to interpret head movements on the surgeon as orders for the robotic assistant [11]. There are orders which are shown by the movements of surgeon's eyes [12] or the gestures of the surgeon's face [11]. Most interfaces developed for surgical robots overload the surgeon with multiple commands and strange devices which make the communication difficult and in consequence, the intervention time can be lengthened [13].

Although accuracy is improved, the use of robots adds new and unnatural tasks to the surgeon. A way for solving this situation is by means of using a HMI that assures an efficient and a more *natural* way of communication [14] and allows to create a more secure and robust system. One of the goals on robotic surgery is the design of a robotic system with an improved capability of reaction and a higher level of intelligence [15]. In this way, a robot assistant would follow the orders of the surgeon and understands the surgical gestures, so the next surgeon's commands can be anticipated. However, some drawbacks appear when trying to recognize the surgeon's gestures: the complexity to attach sensors over the standard surgical tools; the use of 3D markers to the surgeon's tools makes the maneuvers more difficult for the surgeon; and the natural movement during the surgical tasks may locate the tools on occluded areas for the tracking system measurements.

For this reason, the aim of this work is to develop a data acquisition system which minimizes the occluded areas and estimates the location of the surgeon's tools. These estimations allow the continuous knowledge of the surgeon's maneuvers in order to interpret the commands which control the auto-guided navigation of the robot assistant [16]. Therefore, this paper proposes the pose estimation model of the surgeon's tools for different surgical maneuvers, which is an improvement of the multimodal HMI developed for the CISOBOT system at the University of Malaga [16]-[17]. Besides, a double tracking sensor system has been considered to minimize the occluded measures [18]-[19]. Although this configuration improves the visible area for the 3D tracking measurements, even this system may have occluded areas. Thus it has been decided to implement a Multiple Extended Kalman Filter (MEKF) to estimate different surgeon's movements.

After this introduction, section II starts with the description of the robot architecture and its multimodal HMI to understand the problem of occluded areas. Each component of the multimodal HMI is explained in section III, including the pose estimation model. The experimental setup and the results are exposed in section IV. Finally, the conclusions and future works are presented in section V.

Manuscript received March 28, 2011. This work was supported in part by the Regional one P07-TEP-02897.

V. F. Muñoz is with the System Engineering and Automation Department, University of Malaga, Spain (corresponding author to provide phone: +34 952 13 25 40; fax: +34 952 13 14 13; e-mail: vfmm@uma.es).

II. ROBOT ARCHITECTURE AND HMI DESCRIPTION

The CISOBOT robotic assistant (see Fig. 1) consists of two arms: one is responsible for guiding the laparoscopic camera (1); the other one moves an additional surgical tool (2) inside the patient's body (7). The auxiliary instrument aids the surgeon in certain maneuvers of the surgical intervention. The system acts under the surgeon's decisions that are communicated to the robot through gestures of the surgical tools (5) acquired with the 3D tracking system (3), and the voice commands (4) interpreted by a voice recognizer (4). The images inside the abdomen appear in the screen (6).

The 3D tracker system measures the markers attached to the surgeon’s tools. The problem of these markers is that they are only visible when they are facing the 3D tracker sensor; otherwise there are no measures of the surgeon’s tools locations and it is said that the tool is in an occluded area. The auto-guided movement of the CISOBOT robot assistant requires the knowledge of those locations [16]; therefore, this work will focus on the development of a methodology to approach the location of the surgeon’s tools when the 3D tracker measurements are lost in the short time.

For this purpose, the general CISOBOT robot architecture proposed to solve this problem is shown in Fig. 2 in order to explain the different elements which compose the whole system and their relations with the HMI. The main module of this interface is the *Surgeon Model* module which models the surgeon's behavior and recognizes the current *Maneuver* by means of a Hidden Markov Model (HMM) of each of the considered maneuvers [17]. The surgeon model acquires the information of the surgeon's tools positions and orientations by using a *Tracking 3D* system with two sensors, as well as the *Interpreted Voice Maneuvers* which are sent by the *Voice Recognition* module. When the position and the orientation of the surgical tools are not acquired by the tracking 3D, the *Pose Estimation Model* module computes these data by modeling the movements of the surgeon’s tool according to the maneuver facilitated by the surgeon model module. Once the current maneuver is identified, the *Robotic Assistant* receives this information as well as the surgeon’s tools location and, depending on the surgeon’s chosen maneuver, the robot performs the corresponding automated movement.

The proper communication with the robot must be continuous since it is necessary to avoid movement collisions between the surgeon’s tools and the robot’s tools, and to make automatic movements which are secure inside the patient. During the development of the surgical protocol the measure of the surgeon’s tools location may be not read at certain time intervals. Thus the new commands of the surgeon cannot be interpreted, the specific automatic navigation at that moment would be wrong and the process should be aborted. In order to reduce this effect, the data acquisition is recorded with the use of two tracking sensors, whose configuration in the operation room pretends to minimize the occluded areas. In addition to this acquisition system, the pose estimation model exposed in Fig. 2 estimates the missed movements according to the maneuvers which are being carried out.

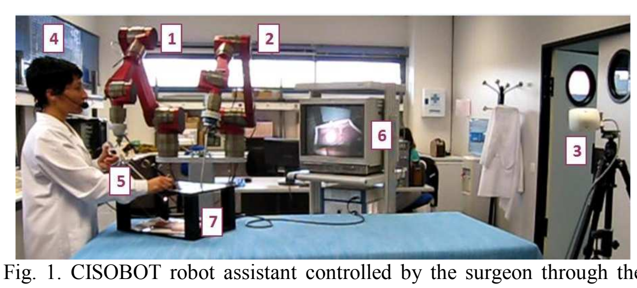


Fig. 1. CISOBOT robot assistant controlled by the surgeon through the multimodal HMI.

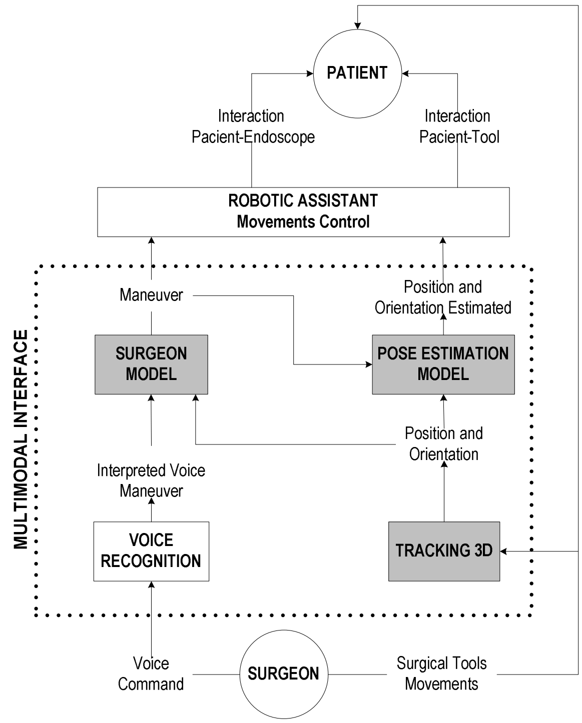


Fig. 2. Architecture of the HMI integrated into a two-arm robot assistant.

The whole system has been described in order to understand the goal of this work. Next section is focused on the development of the pose estimation model and its interrelation with the surgeon model module.

III. POSE ESTIMATION MODEL

This section describes how the position and orientation of the surgeon’s tools are acquired through the tracking 3D module. Firstly, the surgeon model is explained in order to know how this element recognizes the current maneuver. As the surgeon model requires the location of the surgeon’s tools at all times, secondly this section presents a double 3D tracking sensor system to avoid the occlusions as much as possible. However, even with two trackers there may appear occluded areas; therefore, a pose estimation model with the use of Multiple Extended Kalman Filters has been developed to estimate the location of the surgeon’s tools during the occlusion.

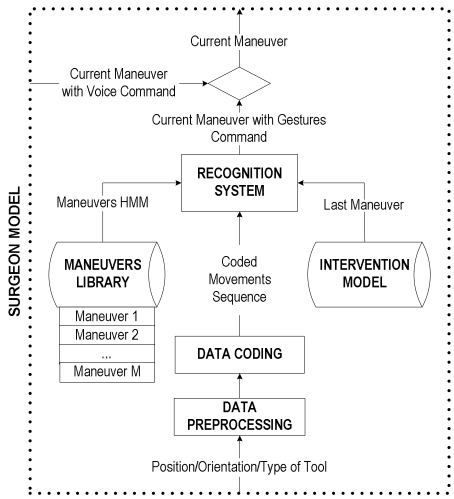


Fig. 3. Description of the Surgeon Model module.

A. Surgeon Model Module

The surgeon model module is composed of the *intervention model*, the *maneuvers library*, the *recognition system*, the *data preprocessing* and *data coding modules* (see Fig. 3). The standard operation of the surgeon model is described as follows: during a surgical intervention, the data preprocessing module filters and fuses the acquired signals of the tools position/orientation as well as the type of tool and data coding module codes this information. Afterwards the recognition system identifies the current maneuver using both the intervention model and the maneuvers library [17].
The surgeon model system communicates the current maneuver of the surgical protocol to the robot in order to plan future actions. More specifically, the main elements of the surgeon model are detailed as follows:

1) Intervention Model: The intervention model has been created from the study of the surgical protocols and its workflow, which is divided into maneuvers connected in a systematic and organized way [10],[20]. In this way, each maneuver has an ending condition which indicates the beginning of the next one, for example, a special movement or a change of the surgical tool.

2) Maneuvers Library: The maneuvers library has to contain the different maneuvers which appear in the majority of protocols of a surgical intervention [10],[20]-[22]. Each maneuver can be divided as a sequence of elemental movements called basic actions. The mathematical models chosen to represent the relation between these basic actions are Hidden Markov Models (HMM) because they provide high flexibility for modeling the surgeon's behavior from the real data of the surgical tasks and give the recognition function [17]. Thus, each maneuver is represented by HMM which model the surgeon's behavior during this task.

3) Data Preprocessing and Coding Modules: These modules are required for both, reducing the noise levels and extracting the kinematics features of the surgical tools movements. After that, the velocities, angles and distance between both surgical tools tips are coded for expressing these parameters in an understandable language for mathematical method of HMM and both the training process and the recognition system of the maneuvers models.

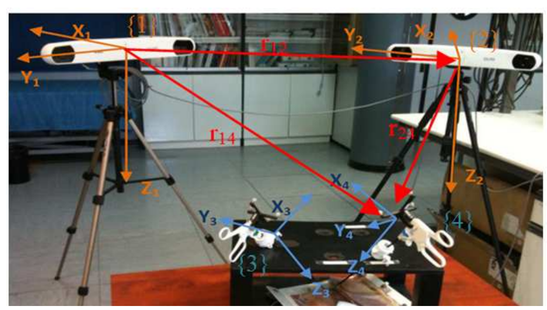
4) Recognition System: This module uses the Viterbi algorithm to compare the coded movements sequence with all the maneuver models available in the maneuvers library. With this information and the last maneuver obtained from the intervention model, the algorithm finds the probable maneuver developed by the surgeon [17].

B. Tracking 3D

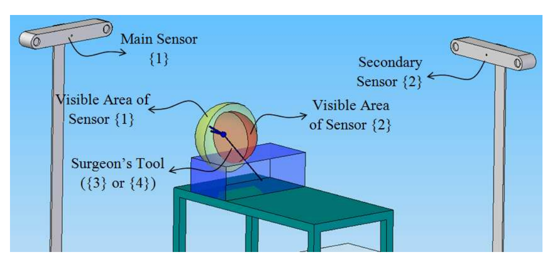
The data acquisition system has two optical sensors for tracking the markers attached to the surgical tools (Fig. 4a). Tracker labeled as {1} is the main sensor for recovering the surgeon's tools measures. The secondary sensor {2} is used for supplying data when the main sensor {1} cannot receive a valid measure of the surgeon's tools.

$$\begin{aligned} {}^{\{1\}}\vec{r}_{left\,\{3\}} &= {}^{\{1\}}\vec{r}_{offset\,\{3\}} + {}^{\{1\}}R_{\{2\}} \cdot {}^{\{2\}}\vec{r}_{left\,\{3\}} \\ {}^{\{1\}}\vec{r}_{right\,\{4\}} &= {}^{\{1\}}\vec{r}_{offset\,\{4\}} + {}^{\{1\}}R_{\{2\}} \cdot {}^{\{2\}}\vec{r}_{right\,\{4\}} \end{aligned} \qquad (1)$$

Expression (1) relates the measures of the secondary sensor {2} with the main sensor system reference {1}. In this way, location of tool {3} measured by the main sensor {1} and denoted ${}^{\{1\}}\vec{r}_{left\{3\}}$ can be expressed as the sum of: the distance between the main and secondary sensors ${}^{\{1\}}\vec{r}_{offset\{3\}}$; and the distance between the secondary sensor and the tool {3} ${}^{\{2\}}\vec{r}_{left\{3\}}$ referred to the main sensor by means of the orientation matrix ${}^{\{1\}}R_{\{2\}}$. Similar conclusions can be obtained for the tool {4}.



a) Reference systems for both 3D tracker sensors and tools.



b) Increased visible area of the surgeon's tool with two sensors.

Fig. 4. Two 3D tracker sensors for the Data Acquisition system.

Both sensors are located in a way that can minimize the occluded areas when the surgeon moves the tools during the surgical tasks because of the increased visible area (see Fig. 4b). However, although this sensors system can reduce some occlusions in the measurements, this assumption cannot be guaranteed at all times. Therefore, it is necessary to make a pose estimator which indicates where the tool is located.

C. Pose Estimation Model

The development of a pose estimation model is necessary to minimize the occluded areas during the surgical maneuvers. In order to create the movement models of the surgeon's tools, the modeling of the maneuvers with a HMM already described in subsection III.A allows extracting the different trajectories and velocities of the surgical tools tips.

Fig. 5 presents the proposed solution to estimate the location of the surgical tools. The pose estimation model receives the current maneuver identified from the surgeon model, whereas the position and the orientation of each tool are acquired by the Tracking 3D system. The current maneuver consists of a series of *Characteristic Movements* for each of the surgical tools. Each of these characteristic movements can be represented by functions denoted as *Estimation Functions*. The *Position* $\vec{r}$ and *Orientation* $\vec{\theta}$ *Estimated* of each surgical tool can be modeled by means of the estimation functions f^{pos} and f^{ori} (2).

$$\begin{aligned} \vec{r} &= f^{pos}\left(e_N, \vec{r}_0, P_m, t\right) \\ \vec{\theta} &= f^{ori}\left(e_N, \vec{\theta}_0, P_m, t\right) \end{aligned} \quad (2)$$

Both expressions (2), that may be non-linear, depends on: the last position $\vec{r}_0$ or orientation $\vec{\theta}_0$ measured by the 3D tracker sensors; different median parameters P_m depending on the function f; the current time t; and the actual characteristic movement e_N, where N denotes the number of characteristic movements recorded on the library.

The number of characteristic movements N may differ from the quantity of estimation functions f^{pos} and f^{ori}. For example, the insertion and extraction are different characteristic movements that may share the same function of a linear movement. According to the maneuver developed, the functions on (2) allow the estimation of the future locations of the surgeon's tools by means of an Extended Kalman Filter (EKF) [23]. The set of all these EKFs is

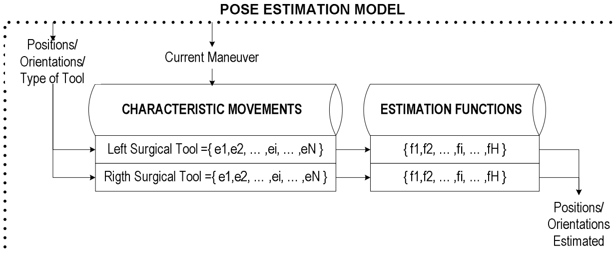


Fig. 5. Pose Estimation Model with all possible movements of each surgeon's tool.

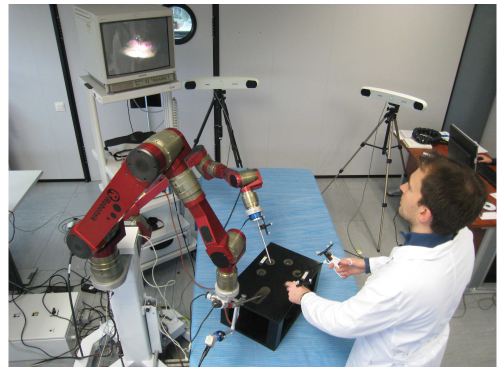


Fig. 6. CISOBOT experimental setup with double 3D tracking system.

denoted as the Multiple Extended Kalman Filter (MEKF) algorithm. Therefore, a maneuver consists of a sequence of basic actions which are expressed as a sequence of characteristic movements' trajectory.

IV. Implementation and Experiments

This section describes the experiments considered to validate the proposed methodology for estimating the surgical tools location. The implementation appears in Fig. 6 and shows the CISOBOT with two 3D tracking sensors which are devoted to measure the surgeon's tools locations thanks to the markers attached to them. The system is controlled by the HMI architecture explained in section II. The recognition system has been trained according to the movements developed by a specific surgeon. Different surgeons should make the same maneuver with different trajectories. Therefore, the recognition system must be trained for each surgeon independently.

Firstly, the experimental setup is introduced to explain the different maneuvers which illustrate the behavior of the pose estimation model already shown in section III.C. Secondly, some experimental results are presented to validate the characteristic movements considered for each surgical tool.

A. Experimental Setup

The pose estimation model proposed in this robotic system is tested for three surgical maneuvers: suture, transporting and cutting. These three maneuvers have been divided into characteristic movements in order to extract their corresponding equations according to expression (2). The movement model depends on the basic action which is being developed when the occluded area is generated. Thus, Table I shows the different functions which represent the related characteristic movement that may appear on the selected maneuvers and are used for estimating the surgeon's tools locations with MEFK. All maneuvers have an insertion at the beginning and an extraction of the surgical tools when the maneuver finishes. The middle part is different for each maneuver because the interrelation between two surgical tools is different. In this way, the characteristic function f_1 on Table I corresponds to a movement which follows the direction of the median on the velocity $\vec{v}_m$ related to the last pose seen by the tracker $\vec{r}_0$. When the surgical tool is still, the location is supposed to remain at the last pose $\vec{r}_0$ like expressed in function f_2. Finally, the stitching movement of the surgical tools is modeled by the function f_3 of sine type with the amplitude A_l (computed at the last maximum/minimum of the measures) and frequency ω_l (elapsed time since the position had the same value).

TABLE I
ESTIMATION FUNCTIONS

Function	Equation
f_1	$\vec{r}_i = \vec{r}_0 + \vec{v}_m \cdot \Delta t$
f_2	$\vec{r}_i = \vec{r}_0$
f_3	$\vec{r}_i = \vec{r}_0 + A_l \sin(\omega_l \Delta t)$

Fig. 7 shows an example of the X component of the three maneuvers considered before for the movement of the surgeon's right hand tool. The different characteristic movements on each maneuver are labeled in a bar downside the graph as e_m. Fig. 7 also presents the occluded areas where the pose estimation model is going to compute the surgeon's tools location on next section.

In this way, Table II allocates the proper function behavior to the tool handled according to the current maneuver and the characteristic movement model. The separation of each characteristic movement is detected by the surgeon model explained in section III.A.

TABLE II
MANEUVERS MODELED BY ESTIMATION FUNCTIONS

Maneuver	Characteristic Movements	Left Hand	Right Hand
Suture	e_1	f_1	f_1
	e_4	f_3	f_3
	e_3	f_1	f_1
Transporting	e_1	f_1	f_1
	e_2	f_1	f_1
	e_3	f_1	f_1
Cutting	e_1	f_1	f_1
	e_5	f_2	f_1
	e_3	f_1	f_1

B. Experimental results

Once the maneuvers and the characteristic movements considered for this experiment have been defined, the pose estimation model is tested in the occluded areas marked in Fig. 7. As the movements of the right hand for the transporting and cutting maneuvers are modeled with the same characteristic movement, this section will just focus on the transporting as well as the suture maneuvers.

The choice of the estimation function for the MEKF depends on the maneuver recognized by the surgeon model when the surgeon's tools were measured for the last time. In this way, Fig. 8a and Fig. 8d show the predicted locations for the surgeon's right tool when the maneuver is correctly

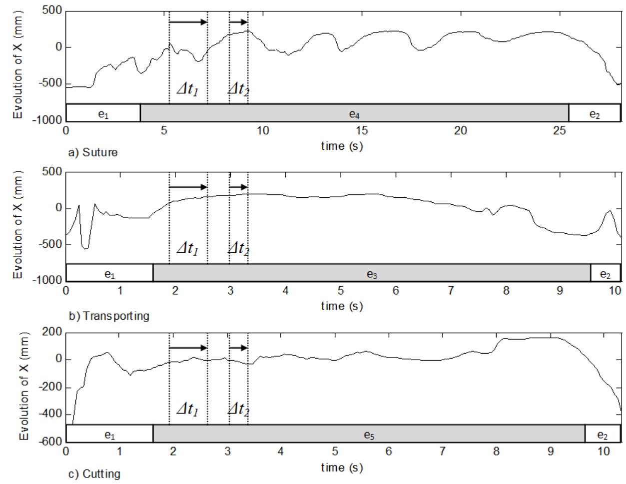


Fig. 7. Evolution of X component (mm) in three different maneuvers: a) Suture, b) Transporting and c) Cutting.

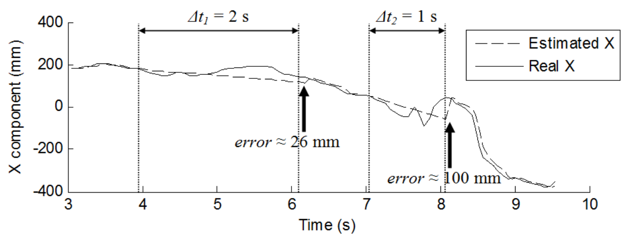
identified whereas Fig. 8b and Fig. 8c represent the estimation with bad maneuver recognition. In both experiments, two occluded areas of duration Δt_1 (2 seconds) and Δt_2 (1 second) have been considered in order to show the increment on the location error with time.

When the surgeon is carrying out a transporting maneuver and the surgeon model is correctly detected, Fig. 8a shows that the estimated location given by MEKF follows a line like expressed by the estimation function f_1 according to Table II. On the other hand, Fig. 8b represents a failure prediction of the same transporting maneuver. It can be noticed that the location error at the end of both intervals Δt_1 and Δt_2 is lower when the maneuver is correctly detected.

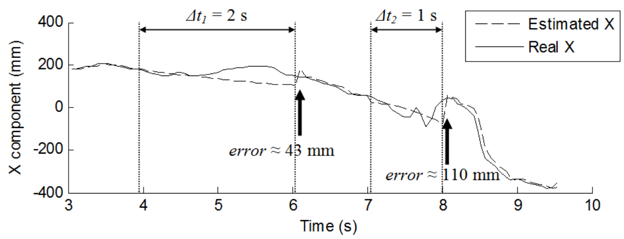
Related to the suture maneuver, Fig. 8c shows the estimation of the MEKF when the maneuver has not been correctly identified. Thus, the estimation function is a line that diverges to the real location of the surgeon's tool, even with short intervals of the occluded area. On the other hand, if the maneuver is correctly determined Fig. 8d represents the estimated trajectory through the estimation function f_3 previously explained in Table I and the error is not as high as with the other estimation.

V. CONCLUSIONS AND FUTURE WORKS

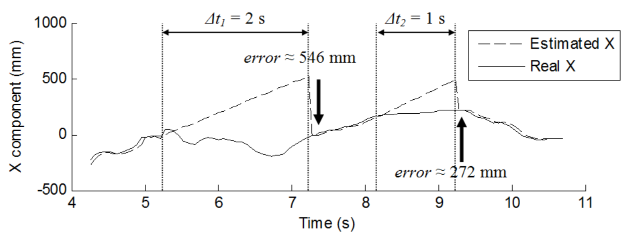
This paper has shown a methodology for estimating the location of the surgeon's tools when they cannot be measured by a 3D tracker system. The pose estimation model through MEKF algorithm has been tested on in-vitro experiments and the results can conclude that this work is valid to give an approximation on the surgeon's tools location when they are in an occluded area during a short time. However, this assumption can be guaranteed only if the maneuver is correctly detected. During the development of a maneuver, the probability of a successful detection is higher when the surgical task is finishing. Thus, the reliability of this work depends on two factors: time duration of occluded areas and the correct identification of a maneuver.



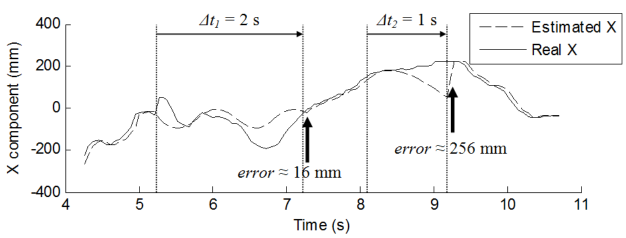
a) Transporting with occluded areas and f_1 as the Estimation Function.



b) Transporting with occluded areas and f_3 as the Estimation Function.



c) Suture with occluded areas and f_1 as the Estimation Function.



d) Suture with occluded areas and f_3 as the Estimation Function.

Fig. 8. Right hand estimation of X component (mm) for each correct and wrong identified maneuver.

These limitations cannot be blamed on the MEKF methodology because the maneuver identification is based on a non-stochastic model like the HMM. As future works some improvements can be made on the precision of the selected estimation functions which model the characteristic movements. Another interesting field consists on adding a learning system so the MEKF may change its set of characteristic movements depending on the surgeon who is developing the maneuvers. This way, the pose estimation model could be used for a general purpose robot in order to improve the accuracy on the surgeon's tools locations.

REFERENCES

[1] F. Gharagozloo, et al., "Robotic Surgery", *Mc Graw Hill Medical,* Ed. 2009.

[2] Y. Wang, et al., "Automated endoscope system for optimal positioning", *Patente number US5815640,* United States.

[3] M. Ghodoussi, et al., "Robotic surgery-the transatlantic case", *IEEE Conference on Robotics and Automation*, May 2002, pp.1882-1888.

[4] N.V. Tsekos, et al., "MRO-guided robotics at the U of Houston: Evolving Methodologies for interventions and surgeries", *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009, pp. 5637-5640.

[5] Y. Güven, et al., "Medical user interface for orthopedical surgical robotic system", *15th 5ational Biomedical Engineering Meeting*, 2010, pp. 16-22.

[6] S.H. Zhang, et al., The human machine interface implementation for the robot assisted endoscopic surgery system", *IEEE International Workshop on Robot and Human Interactive Communication*, 2002, pp. 442-447

[7] A. Casals, et al., "Vision Guided Robotic System for Laparoscopic Surgery", *IFAC International Congress on Advanced Robotics*, 1995 Barcelona, Spain, p. 33-36.

[8] P.A. Finlay, et al., "Controlling the Movement of a Surgical Laparoscope", *IEEE Engineering in Medicine and Biology,* May/June 1995, pp. 289-291.

[9] J. Sackier, et al. , "Voice Activation of a Surgiral Robotic Assistant", *The American Journal of Surgery*, October 1997 ,Vol. 174, No. 4, pp. 406-409.

[10] Seong-Young Ko, et al., "Intelligent Interaction between Surgeon and Laparoscopic Assistant Robot System", *2005 IEEE International Workshop on Robots and Human Interactive*, August 2005, p. 60–65.

[11] A. Nishikawa, et al., "FAce MOUSe: A Novel Human–Machine Interface for Controlling the Position of a Laparoscope", *IEEE Transactions on Robotics and Automation*, October 2003, Vol. 19, No. 5, pp. 825-841.

[12] D. P. Noonan, et al., "Gaze contingent articulated robot control for robot assisted minimally invasive surgery", *IEEE Transactions on Intelligent Robots and Systems*, 2008, pp. 1186-1191.

[13] R. Kumar, et al., "Experiments with Steady Hand Robot in Constrained Compliant Motion and Path Following", *Proccedings of the 1999 IEEE Internacional Workshop on Robot and Human Interaction,* Pisa, Italy, September 1999, pp. 92-97.

[14] K. Wasen, et al., "Person-friendly Robot Interaction: Social, Psychological and Technological Impacts in Health Care Work", *IEEE International Workshop on Robots and Human Interactive Communication,* 2005, pp. 643-648.

[15] P. Dario, et al.," Smart Surgical Tools and Augmenting Devices", *IEEE Transactions on Robotics and Automation,* Vol. 19, No. 5, October 2003, pp. 782-792.

[16] E. Bauzano, et al., "Auto-Guided Movements on Minimally Invasive Surgery for Surgeon Assistance", *International Conference on Intelligent Robots and Systems,* October 2010, Taipei, Taiwan, pp. 1843-1848.

[17] B. Estebanez, et al., " Maneuvers Recognition System for Laparoscopic Surgery", *19th International Workshop on Robotics in Alpe-Adria-Danube Region,* June 2010, Budapest, Hungary, pp. 87-92.

[18] A. Tobergte, et al., " Towards Accurate Motion Compensation in Surgical Robotics", *IEEE International Conference on Robotics and Automation*, 2010, Anchorage, Alaska, pp. 4566-4572.

[19] Q. Gan, et al., "Comparison of two measurement fusion methods for Kalman filter based multisensor data fusion", *IEE Transactions on Aerospace and Electronic Systems,* 37(1), pp. 273-280.

[20] T. Blum, et al."Workflow mining for visualization and analysis of surgeries", *Int J CARS 2008,* July 2008, pp. 379-386.

[21] J. Rosen, et al., "Generalized Approach for Modeling Minimally Invasive Surgery as a Stochastic Process Using a Discrete Markov Model", *IEEE Transactions on Biomedical Engineering,* March 2006, Vol. 53, No. 3, pp. 399-413.

[22] G. Megali, et al., "Modelling and Evaluation of Surgical Performance Using Hidden Markov Models", *IEEE Transactions on Biomedical Engineering,* October 2006, Vol. 53, No. 10, pp. 60–65.

[23] Gary Bishop Greg Welch. *An introduction to the kalman filter.* SIGGRAPH 2001 Course 8.