OPERATIVE ROOM AWARENESS WITH THE SMART TROCAR : INNOVATIVE LAPAROSCOPIC SURGERY

A Thesis

Presented to

the Faculty of the Department of Computer Science

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Cédric Robinet

August 2015

OPERATIVE ROOM AWARENESS WITH THE SMART TROCAR : INNOVATIVE LAPAROSCOPIC SURGERY

Cédric Robinet

APPROVED:

Dr. Marc Garbey, Chairman
Dept. of Computer Science

Dr. Victoria Hilford
Dept. of Computer Science

Dr. Ahmet Omurtag
Dept. of Biomedical Engineering

Dr. Shishir Shah
Dept. of Computer Science

Dean, College of Natural Sciences and Mathematics

Acknowledgements

I would like to gratefully thank every member of the Atlantis program. First of all, I would like to thank Pr. Marc Garbey for his guidance and patience, every day at Methodist. I also would like to sincerely thank Pr. Christophe Collet who offered me this great experience that will definitively change my professional and personal goals.

I would like to express my gratitude to my committee members Dr. Victoria Hilford, Dr. Ahmet Omurtag, and Dr. Shishir Shah for their cooperation.

I have a special appreciation for the European Union, the Education, Audio-visual and Culture Executive Agency (EACEA), and the Alsace region for their financial backing.

I would also thank all the people who helped me during this exchange program: the previous Atlantis students for transmitting their experience, Mrs. Garbey for the help she always provides to the students, Mrs. Faigh and Mrs. Elder for their great efficiency regarding the visa procedure and the administration, and, of course, all my colleagues for their friendship and their help during this great experience.

Finally, and most importantly, a huge thank you to my familly who supported me during my studies.

OPERATIVE ROOM AWARENESS WITH THE SMART TROCAR : INNOVATIVE LAPAROSCOPIC SURGERY

An Abstract of a Thesis

Presented to

the Faculty of the Department of Computer Science

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Cédric Robinet

August 2015

Abstract

Laparoscopic surgery is a popular alternative to open surgery due to the considerable reduction of recovery time, pain, scaring, and complications. During laparoscopic surgery, operations are performed through small incisions (usually around 1 cm) elsewhere in the body. However, limited access to the operating field, indirect 2D vision of a 3D complex scene, the presence of mirror effect, and operating rooms originally built for open surgery, conspire to make the surgeon's work more difficult and less efficient. We present a new technology, the Smart Trocar, which can compensate for these drawbacks by providing a global positioning system of the laparoscopic tools that are inserted into the body. Our system uses a single wireless camera, which stays outside the body and tracks some features inside the operating room. With this simple and inexpensive device, we can, at any time, and within few millimeters accuracy, both recognize and localize the rigid surgical instruments. Applications of this invention are multiple. For examples, it can be used as a pointer to localize specific part of the organs during a laparoscopic surgery. It can also be used to analyze tools paths and to evaluate surgical dexterity of residents in order to improve their training. We present in this report an analysis of the tool trajectories that allows us to determine the level of skill of a resident based on his results to few laparoscopic exercises. Our study proves the correlation between surgical dexterity and economy of motion.

Contents

1 Introduction **1**

1.1 Presentation of laparoscopic surgery . . . 1

1.2 Advantage for the patient . . . 2

1.3 Challenge for the surgeon . . . 3

2 Background **4**

2.1 Recognition of the laparoscopic tools . . . 4

2.2 Surgical tools localization . . . 5

2.2.1 Localization through computer vision . . . 5

2.2.2 Localization using ultrasounds devices . . . 9

2.2.3 Localization using a tracking system . . . 10

2.3 Laparoscopic training . . . 11

3 Smart Trocar **14**

3.1 The concept . . . 14

3.2 Our bill of specification . . . 20

3.3 Technical development . . . 21

3.3.1 Feature detections . . . 22

3.3.2 Pose estimation problem . . . 27

3.3.3 An alternative solution: pose estimation without poster . . . 30

4 Results and Validation **34**

4.1 Verification 34

4.2 Application to the analyze of tool path 38

4.2.1 Set up of the experience with the surgeons 38

4.2.2 Preliminary results 40

5 Perceptive **45**

5.1 Operating room awareness 45

5.2 Augmented reality 49

5.2.1 Three-dimensional reconstruction of the organs and the tools . 49

5.2.2 Registration of preoperative and intraoperative images: add the Finite Element (FE) model 49

5.2.3 Analysis of the trajectories to prevent dangerous events 52

Bibliography **54**

List of Figures

1.1 A laparoscopic procedure 2

3.1 Our smart trocar, with the GPS part (left) and the recognition part (right) 16

3.2 Geometry behind our GPS 17

3.3 Localization of the tool in the trocar system 20

3.4 Corner detection and selection 24

3.5 Algorithm of features detection and localization 27

3.6 The algorithm of localization 28

3.7 Error function to minimize : left, with no motion restriction; right, with motion restriction 29

3.8 Algorithm of tracking 32

4.1 Validation with a professional camera 35

4.2 Validation with a regular Smart Trocar and two laser beam 37

4.3 Completion time and mistakes for all the participants 42

4.4 Left: FLS box with Smart Trocar; right : one of our exercises: cut between the lines 43

4.5 Example of trajectories, for the different populations 43

4.6 Attention and meditation, provided by the Neurosky system 43

4.7 Different metrics and t-tests 44

5.1 Output of the Smart Trocar system (black) compared to an ideal version procedure (red) 47

5.2 Registration principle in image guided surgery 50

5.3 Geometry of our problem: how to register the 3D surface of the liver 51

Chapter 1

Introduction

1.1 Presentation of laparoscopic surgery

Laparoscopic surgery is a growing market with around 3.5 million operations per year in the US [47]. The main goal of laparoscopic surgery is to reduce the size of the opening in the patient's body in order to help him recover faster. In this kind of surgery, the surgeon make small incisions on the abdomen, instead of making larger incisions, as we do for a conventional open surgery. Next, the surgeon places some plastic pipe (trocar) through those incisions, and uses these trocars as portals for the surgical tools to gain access to organs. Last, but not least, the surgeon also injects some gas within the body, to have a better access to the organ and a better view of the organs and his tools with a small camera called a laparoscope (Figure 1.1).

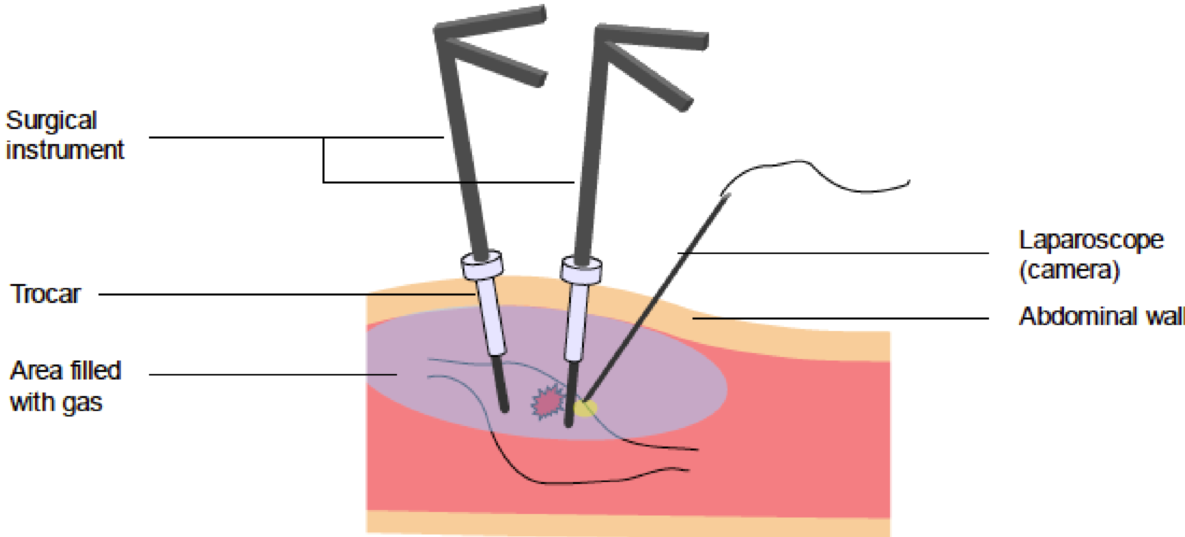

Figure 1.1: A laparoscopic procedure

1.2 Advantage for the patient

The main goal of this kind of surgery is to reduce the risks due to the surgery itself, e.g. :

- With smaller incision, we expect the patient will recover quickly.
- We reduce the risks of complication and infection because the inner organs of the patient are not directly in contact with the OR environment.
- We reduce the time spent in the hospital.
- We create smaller scars.

1.3 Challenge for the surgeon

The main drawback of laparoscopic surgery, in addition to the fact that the procedure is completely different of an open surgery procedure, is the indirect access to the organ and the lack of depth perception. Gould and Frydman [16] present the result of a study made on twenty-one surgeons. It indicates that reverse-alignment, a common problem in laparoscopic procedures because the entry point (the trocar) is a rotation fixed point, has a great influence on the surgeon's efficiency. Indeed, this problem is caused by the creation of a mirror image of the operation field, thus the surgeon has to mentally focus a lot more in order to have the correct image of what he is operating on. Even if our tested sample number is too small to represent the average surgeon, it does at least give us an idea of how far a simple problem of representation can affect the whole procedure. Other properties inherent to the imaging system, like the lack of depth perception, can also be a serious impediment to the smooth execution of the surgery. Unlike during open surgery, during laparoscopy the surgeon has to rely only on what the laparoscope sees. Of course there is first the problem of watching a screen, which is already a huge change in the way of operating procedures, but also the problems of obtaining pertinent depth information based on a 2D endoscopic image.

Chapter 2

Background

2.1 Recognition of the laparoscopic tools

In order to be able to recognize the surgical tools, several techniques have been investigated. The first approach is recognition of features inherent to each kind of laparoscopic tool, such as color, shape, and/or compactness of the tool. These kinds of methods do not seem robust (since surgical tools can look very similar), and are too time-consuming to be used in real time [13][14][18][41]. That is why other approaches, such as adding a specific marker on the surgical tool, have been studied. Some studies [44][45][46] suggest the use of ultrasound markers with a specific signal associated to each tool. Nonetheless, the authors recognize there are many false detections. We can imagine a similar approach based on colored marker, where we add a specific color code on the surgical tool, as described in [22][30]. This method is the one that provides the best robustness.

2.2 Surgical tools localization

2.2.1 Localization through computer vision

While a lot of efficient methods use a combination of two cameras, here we need to create a small device, we will focus here on a 3D reconstruction based on the detection of some points in the image. Yuan [56] explains in his paper how to solve the exterior orientation calibration problem of a single-camera thanks to algebraic considerations (detailed in the next part). He assumes that he knows the coordinates of several feature points of an object, the corresponding coordinates on the image, and the effective focal length. Thanks to a pinhole model of the camera, he poses the problem and gives a general solution to it. Moreover he proves the uniqueness or non-uniqueness depending of the number of points. According to his paper, three coplanar points are enough to solve the exterior orientation calibration problem; however three coplanar points will as provide multiple solutions. It takes four or more coplanar points to give us a unique solution. Four or five non-coplanar feature points should provide more accuracy and more robust solutions than a coplanar case, but they do not provide the uniqueness we desire. That means if we want to assure the uniqueness of the solution, we will have to find more points, which can be more challenging and less accurate. However these results are based on two major assuptions :

1. We can determine the focal length of the camera.

2. The camera works like the pinhole model.

We know this is not realistic for most of the small cameras. Consequently, we have to know if we can correct the distortion before doing any other image treatments.

The problem of using computer vision comes from the robotic field and the need for researchers to create robots that can watch and interpret their environments. Uchida et al. [51] applied this principle to a robot that should be able to assist the surgeon. They put some light points on the mobile part and use a camera to track it. The geometry, such as the angles and distances of the pattern formed by light points, is presumably known. This knowledge allows us to simplify the equations to get the result. Nonetheless, the accuracy of the localization system is above 1 millimeter, and might be worse in real conditions. Moreover, light points could be troublesome for the surgeon, depending on their position with respect to the surgeon's eyes.

The study of Navarro et al. [29] compensates for this problem. It also uses the correlation in time of two consecutive images and motion analysis in the minimally invasive surgery framework. Among others, the proposed method will provide us the Euclidean angles of surgical tools and their positions. In order to get precise information for the localization, they apply determined movements through a robotic arm with a passive wrist. They choose to select the lines in the images as features of interest. Their technique might give us the 3D position and orientation of surgical tools, but it needs a robotic assistant, and the localization accuracy is highly dependent on the wrist motions.

Tsai and Wang [50] propose a solution to the self-localization of a robot that we can use in our project. Their system consists in three different markers in the room, placed around the robot, and recognized and tracked by a single CMOS (Complementary Metal-Oxide Semiconductor) camera. Thanks to a color analysis (the marker differ only by the color), they extract the markers from the picture and from these landmarks to obtain feature points. Then a triangulation method is applied to localize the robot using the three given feature points. In the end, to minimize the estimation errors, an extended Kalman filtering is used. Though, having known colored markers placed in the room and minimizing the error by Kalman filtering were a good ideas, the results obtained by this method are not sufficient for our problem, as the accuracy is 3.6 centimeters for the shifts and 3.2° accuracy for the rotation angle. Moreover, the context of their study concerns the estimation of three parameters: x and y shifts and one angle of rotation. Other systems involving 3D sensors were imagined to get more accurate results on the localization estimation. Estebanez et al. [15] propose a tracking system composed by two 3D sensors, but this system is space consuming and not robust to occlusions that can appear during surgery.

Even if the many techniques involving external computer vision have been investigated, in the specific context of laparoscopic surgery, we already have a camera, the endoscope, that allows us to use computer vision methods within the patient's body. Dutkiewicz et al. [13] explore a technique of estimating tool positions in the laparoscopic field of view. Thanks to an elaborated image-processing algorithm that consists of mixing the information given by two different kinds of image analysis, they attest that it is possible to extract the orientation of each laparoscopic tool in the laparoscopic images. Their algorithm is first based on a color analysis that is expected to separate the background from the rest, and second on a shape and light distribution analysis that is supposed to give the central line of each object and eliminate any object that shouldn't be considered an instrument. However, as they expressed in [14], this method has major drawbacks: the cumulated rate of false detections (false negatives and false positives) is about 9% and the obtained rate of correctly estimated axis and tip of the tool lies below 80%. This implies that the technique is not accurate enough for practical interventions. Nevertheless, the idea of using an endoscope to localize the tools is good, in essence because if we need to know the absolute position of the tool, we also need their relative positions in reference to the organs.

Using the endoscope field of view in order to localize the surgical tools during the operation is also the idea of Allan et al. [1]. Indeed, they pretend that it is possible to accurately estimate five of the six localization parameters, the last one being the rotation around the tool's shaft's axis. Their method consists first in a probabilistic supervised classification of image pixels, called Random Forest, in order to determine whether the pixels belong to an instrument or not. Then the classified image is taken as initialization of an energy minimization algorithm using the level set technique. To be more precise, this second step initializes the contour with the result of the previous segmentation and by applying an energy function, considering pixel similarity between interior and exterior regions. Using a gradient descent, it tests new sets of the five localization parameters to determine the set that provides the best match between the obtained model and the new real image. The difference of depth between two images is estimated by observing the change in radius of each surgical tool as soon as they are cylindrical and their real radius is known. The results given by this technique should have an error of less than 0.2 millimeters. Even though this method seems to provide idealistic results, it still needs to be improved, as the processing time is really too high to meet real-time requirements. Furthermore, the tests were made on laparoscopic images with a single instrument within their field of view, which is why the accuracy achieved by their algorithm might be lower for real surgery, as it often involves several tools at the same time.

2.2.2 Localization using ultrasounds devices

Devices using ultrasounds are not often used for what we try to do because they are rarely accurate enough. Still, some laparoscopic operations are done with the help of an ultrasound scanner. This imaging modality uses the properties of ultrasounds to show the differences of density inside the human body. This is a property used in ultrasound imaging systems. If we put ultrasound transmitter-receivers all over the surgical field and ultrasound markers on each instrument handles, on the endoscope, and on each trocar, we should be able to localize the tools. That is the idea developed by Tatar et al. [44][45][46]. The markers have to be triggered by their specific radio wave signal, as they include radio frequency receivers that only react to a very specific signal. Then by computing the time-of-flight of the ultrasound, the position of the tools should be retrieved. Although, this method has a lot of weaknesses. As this method involves ultrasounds, that is to say beams that vary according to the density of the medium they go through, in the reality the extraction of localization information from the time-of-flight computation will be a real challenge. However, artifacts caused by multiple reflections of the ultrasounds could corrupt the results so the received signal is really noisy. Moreover it requires modifying laparoscopic tools, trocars, and laparoscopes in order to place the transmitter-receiver devices. On the other hand, the Smart Trocar only modifies the top of the trocar, which is not use during laparoscopic surgery.

We found a method [42] that solves the problem of noise by a smart positioning of the device. Nonetheless, this method gives extremely imprecise results: the accuracy for the rotation is 6°, which can imply an uncertainty of more than 1 cm at the tip of the surgical tool.

2.2.3 Localization using a tracking system

In order to simplify the localization problem, tracking methods are used. Indeed when it comes to real-time situations, the computation time must be as short as possible. Video tracking enables us to take into account the time and spatial correlation that exists between two consecutive images. This information limits the research area for the segmentation part, and also constraints the parameters of position and orientation to a limited field because of the spatial correlation. Kalman filtering is used in most of the publications that involve the tracking of surgical tools [15][41]. Still, according to Yilmaz, Javed, and Shah [54] this filter is used to estimate the state of a linear system where the state is assumed to be distributed by a Gaussian; however the localization issue is a non-linear system. They refer to the work of Rosales and Scarloff [36], who improved this filter to make it possible to estimate 3D relative localization from 2D motion. This filter is called the extended Kalman filter. It might not be robust to occlusions, unlike particle filters that use genetic algorithms to solve occlusion problems, but under certain constraints it gives optimal solutions. It also shows other advantages, like the ability to merge information from different sources.

Nevertheless, easier and faster techniques based on computer vision have been imagined. Several works were based on the possibility offered by the tracking of a colored marker on the tip of each surgical instrument [22][30] by a simple segmentation. In those papers the tracking is made thanks to a color analysis of each laparoscopic image due to the fact that the red color is more likely to belong to the background than to the tools. However, computer vision is not the only method used to track surgical tools.

2.3 Laparoscopic training

In its 11 years of existence, the Fundamentals of Laparoscopic Surgery (FLS) has become a standard in the way of teaching laparoscopy [39]. This consists of practicing a few exercises on an artificial model, inside a box that represents the human body. The task should test some skills useful for this kind of surgery (make a knot, or cut on a specific spot). The test is passed if the resident is able to finish the exercise with no mistakes, in a certain amount of time, according to the Texan Association of Surgical Skills Laboratories. Nonetheless, it seems this method has two main drawbacks:

1. The residents are well trained to do these exercises, and can obtain good results for FLS tasks, without being ready to practice a real laparoscopic operation [23][37].

2. The evaluation of the tasks themselves is not accurate enough. The evaluation of the accuracy (number of mistake) is both subjective (based on the judgement of the surgeons) and qualitative (pass or fail).

The only objective and quantitative metric currently used is the time of completion, which is limited as an indicator of dexterity: the fastest surgeons are not always the best. Indeed, Bann shows in [4] that, after a certain level, the experienced surgeons progress less in completion time, but are more economical in quantity of motion and suggest the use of this second metric as a second indicator of dexterity.

Since it is recognized that FLS is an efficient and cost effective way to teach laparoscopic surgery, a lot of work has been done to try to provide more information to the residents, and find new indicators to quantitatively evaluate the performances. Most of this work is using additional systems to obtain the trajectories of the surgical tools [12][26] or the torque [25]. Nevertheless, there is no consensus about the use of these new data to improve training of laparoscopic surgery. Besides, those methods are often invasive and not comfortable for the resident. Bann and McBeth put the sensors directly on the hand of the practitioner while Lum needed an entirely passive robot attached to the tool to measure the torque. Finally, some studies show that the stress is a key factor that can explain decreasing performance when the residents go from a virtual model to a real Operating Room (OR) [34][35][57]. While taking stress into account is a great step, the studies are also limited. The NASA-TLX group uses a self-evaluation stress system, using a questionnaire, when the group of Pavlidis uses thermal imaging, that is expensive and takes time to setup.

For all these reasons, it is important to encourage the development of the FLS, which offers a cheap and efficient way to train residents, but we need to develop new metrics to objectively quantify the dexterity of the practitioner, and to take into account the mental workload by measuring the alertness and the stress. To be useful, our system should not disturb the surgeons, and it has to be easy and fast to setup.

Chapter 3

Smart Trocar

As we exposed previously, laparoscopic surgery is a constant challenge even for skilled surgeons, mainly because of the indirect vision and the resulting lack of depth perception. To compensate for these drawbacks and make laparoscopic surgery safer, the Smart Trocar should provide the surgeon information about the global position of the surgical tools, the kind of tools we are using, and their motions along the surgery.

3.1 The concept

Our device provides useful information for surgical application as long as the accuracy on the position is less than 1 mm along the three axis (x,y,z) and the accuracy on the orientation is less than $0.5°$ around those axis. With the need for this high level of accuracy, we cannot use a direct 3D sensor system like in [15]. For the same reason, we cannot use the ultrasonic sensor; all the studies quoted in this thesis show very complex and expensive processes for a small number of positive results. We consequently decided to use a system based on computer vision and triangulation. A stereoscopic vision system is a system using two different cameras separated by a known distance. By triangulation, we can easily determine the depth of the object present in the two vision fields of the camera. The reverse principle of stereovision used the same idea but inversing the camera and the object, i.e. in this case, we have one single camera and two objects separated by a well known distance. By the same kind of triangulation, we can determine the distance between the camera and the two objects. The choice of using one single camera is justified both by consideration of price and space; we want to create a device as cheap and as small as possible.

We imagined the system on Figure 3.1, consisting of a small camera, fixed on the top part of the trocar, tracking some markers inside the operating room. By adding specific markers on the tools, we are able to recognize them, and by adding different markers on the ceiling at a known position, we are able to know the position of the camera, and consequently its motion.

The recognition part requires simple image processing algorithms. We put a circular unicolor/bicolor marker on each laparoscopic tool. With only n colors, we can create $\frac{n \cdot (n+1)}{2}$ different markers (e.g. with a palette of only six different colors, we can create twenty-one different markers, which is usually more than the number of tools used during an operation). The idea of the algorithm is to first detect the marker on the tool (using a circular Hough Transform detection), and then identify the color and associate it to the corresponding tool, thanks to a hue analysis.

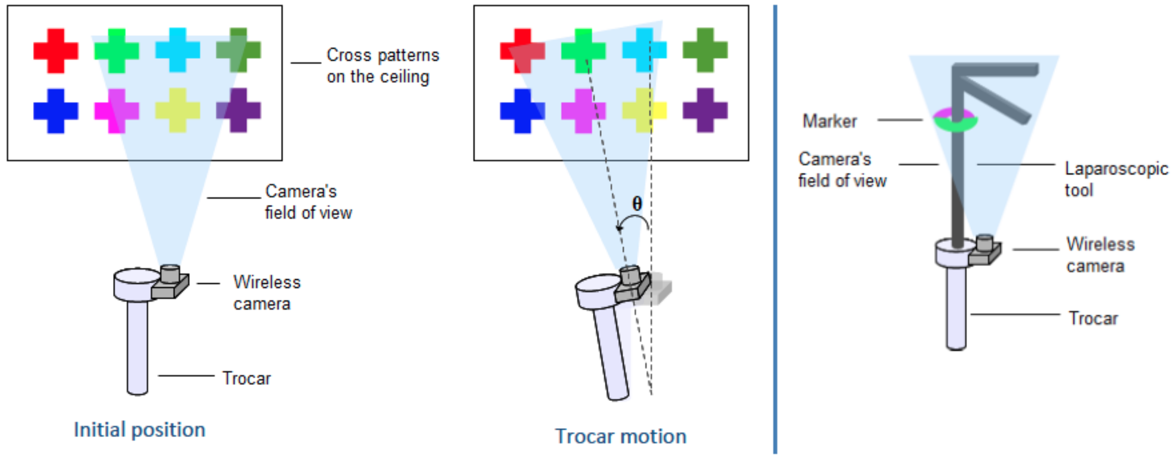

Figure 3.1: Our smart trocar, with the GPS part (left) and the recognition part (right)

The idea behind the GPS system is a little more complex. Let us assume that we have painted a cross on the ceiling, and that this cross is captured by the video stream from the Smart Trocar. We can extract from the video the trajectory of four points, for example the middle of the end point of the cross. With the following well-known math, we can deduce the position of the camera from the position of those four points, and we can also guarantee the uniqueness of the solution. [56]

Our marker is fixed at the position {X-1,X0,X+1} on the ceiling. Its position is known before the motion, that is to say in the blue coordinate system. In his initial state, the system can be represented by the following equations, where x and X are respectively the coordinates of the point in the image and in the real space, col is the number of columns in the image, θ_C is the opening angle of the camera, H is the distance between the trocar and the ceiling, and L_x is the real part of the ceiling

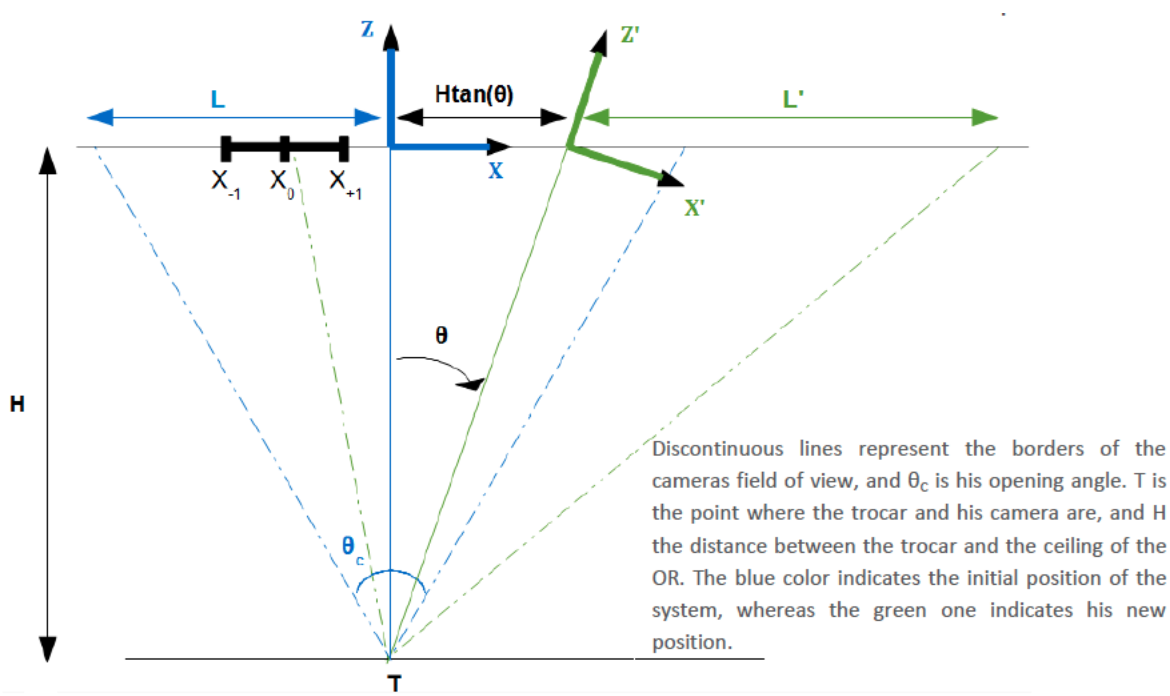

Figure 3.2: Geometry behind our GPS

viewed by the camera.

$$L_x = \tan(\frac{\theta_C}{2}) \cdot H$$

$$x \cdot L_x = col \cdot X$$

Now we can also express the image coordinate x' for the new real position X' of the trocar. The system of equations below enables us to achieve this goal. So on the X axis we will have:

$$X' = \cos(\theta) \cdot (-(H + dZ) \cdot \tan(\theta) + X + dX)$$

$$Z'_x = \sin(\theta) \cdot (-(H + dZ) \cdot \tan(\theta) + X + dX)$$

$$L'_x = \tan(\frac{\theta_C}{2}) \cdot (\frac{H}{\cos(\theta)} + Z'_x)$$

$$x' \cdot L'_x = col \cdot X'$$

On the Y axis we will face a similar problem, and for a motion described by a translation dZ on the Z axis, a second translation dY on the Y axis, and a rotation ϕ around the X axis, we will have the following system of equations (on the Y axis):

$$Y' = \cos(\phi) \cdot (-(H + dZ) \cdot \tan(\phi) + Y + dY)$$

$$Z'_y = \sin(\phi) \cdot (-(H + dZ) \cdot \tan(\phi) + Y + dY)$$

$$L'_y = \tan(\frac{\phi_C}{2}) \cdot (\frac{H}{\cos(\phi)} + Z'_y)$$

$$y' \cdot L'_y = raw \cdot Y'$$

This is an ideal case that is far from the reality. We have to keep in mind that the laparoscopic tools and the surgeon himself will be in the field of the camera. To be sure to always have the coordinate of 4 points in the image, we add other cross markers, of different colors, separated by a well-known distance. Now, we have to add a color recognition part on our algorithm to identify which cross the points belong to.

Our system provides us the position and orientation of the camera, which is attached to the trocar. Since we know the exact position of the camera (including the rotation angle along z-axis) and since we know the geometry of our Smart Trocar, it is easy to obtain the position of the entire trocar. To completely solve our system and determine the position of the tool, we have to solve the additional unknown:

1. How much of the tool is inserted inside the trocar?

2. The inclination of the tool might not be exactly along the axis of the trocar.

3. The rotation of the tool on its axis.

In order to localize the tip of the instrument, we need to know how deep the tool is inserted, i.e. how close the marker is from the camera. By determining the size of the marker, we can know how deep the surgical tool is inserted.

The second unknown is due to the inclination of the tool inside the trocar: indeed, even if the diameter of the trocar is designed to fit exactly with the one of the surgical tools (in this case, the inclination of the tool is the one of the trocar), it can happen that the surgeon changes the surgical tool without changing the trocar. That is why the diameter of the trocar can be a few millimeters bigger than the diameter of the tool and allow the tool to move along the transversal directions. If we name ΔD the difference of diameter between the trocar and the surgical instrument, and L the length of the trocar, the angle α of the tool, in reference to the trocar axis is:

$$\tan(\alpha) = \frac{2 \cdot \Delta D}{L}$$

When $\Delta D = 3$ mm and $L = 150$ cm, we obtain an angle $\alpha = 2.2°$. Even if this case is uncommon, it cannot be neglected and we now need to think of a way to know the inclination of the tool, in reference to the trocar axis. This can be solved again thanks to the circular marker mounted on the tool (see Figure 3.3). Since we know the exact position of the trocar axis and of the camera, we are able to determine where the center of the marker should be, if α was equal to zero. So, the angle α only depends on the position of the center of the marker and its radius, that we can both measure on the image.

The last unknown is the rotation of the tool on itself, which is simply the rotation of the tool marker. It can be measured thanks to the two colors on the pattern.

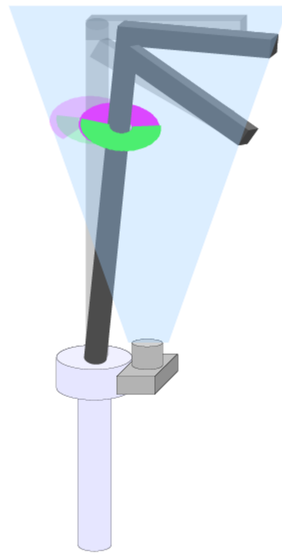

Figure 3.3: Localization of the tool in the trocar system

3.2 Our bill of specification

We want to meet the following requirements:

- Shift accuracy : 1mm
- Angular accuracy: 0.5°
- GPS robust to partial occlusion of the field of view of the camera
- GPS robust to lightness change

- Wide angles tolerated : more than 40° between the two extreme positions
- Real time
- No additional work for the surgeon
- Safe and non-invasive for the patient
- Compatible with the existing technology (no need a new trocar)
- Easy and fast to mount
- Cheap
- Small
- Wireless transmission of the signal
- Enough battery to support one operation

3.3 Technical development

In this project, we focus on the localization of the trocar. The algorithm is divided into two main parts; the first one uses image processing tools to recognize the feature points in the image, and the second one solves the estimation pose problem, based on these points.

3.3.1 Feature detections

A preliminary step is to correct distortion; indeed, we need to know the exact position of the feature points, and we assume a pinhole model. The camera we use causes a radial distortion. As we always use the same camera, we can determine its parameters once, and implement the corresponding transformation [6][10].

Now, for the detection of the feature points, the most natural approach was to detect the crosses. Different methods close to the machine-learning field were investigated and quickly abandoned because of the low quality of their results on our images e.g. the unsupervised learning methods like k-means clustering, or pattern recognition. As we want to detect a known shape, one of the first approaches we imagined was to use a generalized Hough transform [3]. Indeed, this method is robust to noise, lightness change, and partial occlusion of the field of view. But this method also requires a very high computational time for determing general shape. We might reduce the computational time by choosing a simpler marker (a circle or a line instead of a cross) or by indicating where we guess the cross should be (using the time correlation of the images). Still, if we want to create a real-time system, it would be better to use something else.

We looked for another shape-detection method that seemed interesting because it is robust to affine deformations and to the blur. Suk and Flusser [43] claim that there is a kind of feature that is insensitive to a centrosymmetric and energy-preserving PSF type of blur and to affine deformations. This is good but not good enough, since this implies conditions of focus we are not sure to provide. For the triangulation part, our camera must not have autofocus (otherwise, we change our optical system in an unknown way). But the trocar will still have a little motion on the z axis, thus we need a method that can deal with these little blurred images. Vu and Manjunath [52] present an interesting method using prior shape knowledge to localize an object even in case of partial occlusion. This method might provide some results for blurred images, but the publication does not give enough data for this kind of noise. Besides, the algorithm seems very slow.

Since it was difficult to find an accurate and fast shape detector, we decided to detect our feature using a corner detector algorithm. The extraction of the corner is made by the Matlab corner detector, using the Minimum Eigenvalues Method [38]. This method is fast and accurate; the position of the detected corners was detected, on average, with an error of less than 2 pixels. To make this method more robust to change of light and light contrast, we pre-treated the image with an histogram equalizer. The advantage of this method is that it provides accurate results; the drawback is it detects false positive (e.g. corners that are not on the crosses). We can see the results provided by the algorithm on figure 3.4 (in red, the corner we want to remove, in blue, the corner we want to keep). The corners selection can be done easily if we know the approximate position of the cross, using the following criteria:

1. When a corner is detected twice (this is due to the thickness of the black edge) we remove the internal one (the closest to the center of the cross).

2. When the distance of the corner to the center is not inside a certain range, this

is an internal corner, or it is outside the cross mask and we remove it.

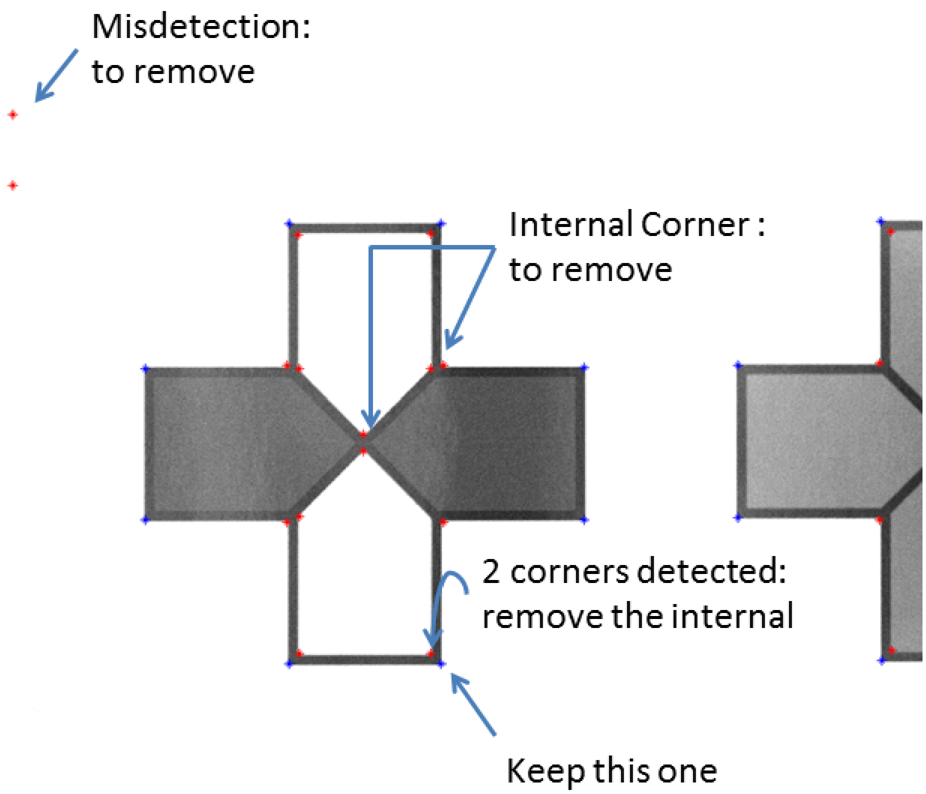

Figure 3.4: Corner detection and selection

But to apply this simple method, we need to know the position of the approximate position of the cross, and its center. This can be done in a very fast way, without using the complex shape detector we detailed before. To do this, we will use the difference of intensity between the cross and the background. Indeed, the crosses are on a white background and can be easily detected. Even if this method seems very fast and efficient, it has two main drawbacks:

1. The position of the crosses may not be accurate.

2. We will have some false positive (all visible objects with a low intensity, e.g. surgeons, tools).

Nonetheless, when we have detected our possible crosses, we can define criteria to eliminate the false positive: e.g., since we know the approximate height of the ceiling, we can define a specific range for the surface and diameter; every object outside these ranges is eliminated. This method has shown good results and that is the one we chose.

At that point of the algorithm, we have been able to detect the external corners of the crosses, within the accuracy offered by Matlab corner detector function. The features we decided to use in our algorithm are not those points, but rather the middle of the ends of the cross. This can be done by taking the middle of the segment between two corners. This choice of features can be understood in terms of accuracy; i.e. taking the middle of those two points may be more accurate than the corners themselves. Now, in order to identify which feature we have detected, we have to know which cross it belongs to, and on which position on the cross it occurs (left, right, up, or down). The first problem can be solved by detecting the color of the cross, and the second by knowing both the position of the point in reference to the center of the cross, and the rotation of the poster inside the image.

For the color detection, Coughlan et al. [8][9] presented a method for the recognition of a colored circular object in an image using color gradients. In our case, the pattern had already been detected and the only thing that can trouble our color recognition is change of light. This method should be independent of light change.

One other proposed method is the one of Kloss et al. [21]. It uses a learning algorithm that is a color contrast fusion algorithm. But this algorithm does not meet our main requirement i.e. the robustness to light change. Toti et al. [49] also present colored markers analysis based on HSV image. The idea of using HSV image is good because the color information is in the hue values, whereas the lightness acts mainly on the value. Then, they check for each detected object, which colors are present or not. This is exactly what we want to do and we use this publication to write our color detection code; for every cross, we look at the hue of every pixel and identify the two dominant colors. Because our crosses are made with at most two colors among those 6 (red, green, blue, yellow, cyan, magenta) which have distinguished hues values, the recognition of the cross is always possible.

Now that we have detected on which cross is our feature, we have to know if the feature is on the left, right, top, or bottom of the cross, which is trivial if we know the inclination of the poster in the image. This inclination is known for every frame; it simply corresponds to the rotation of the trocar within its axis, which should not change a lot between two frames. Besides, this angle can be computed using the crosses. Since we have the poster map, and the position and color of the crosses, with only two points (e.g. the center of two crosses) we are able to find the angle we are looking for. This redundant information can also be used as a way to check if the color has been correctly detected.

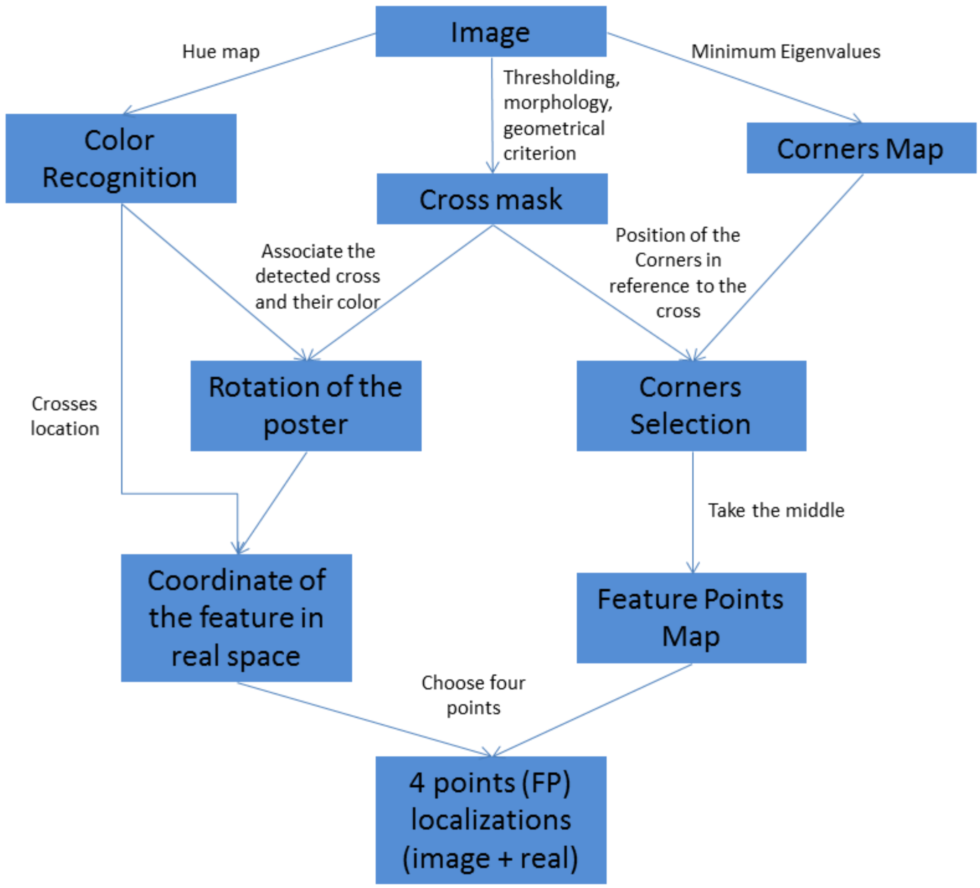

Figure 3.5: Algorithm of features detection and localization

3.3.2 Pose estimation problem

Thanks to the first part, we are now able to detect the coordinates of some points, with accuracy around one pixel, and its corresponding coordinate in real space, even with bad lighting and partial occlusion. We select four of these points to run our algorithm (as long as one cross is visible, the algorithm detects at least four points). Since the equations developed in the previous part gives the image coordinate if we have the camera coordinate, we can run a simulation that provides for every rotation and translation possible of the camera, the expected position of the points in the image. In order to find the set of points that fits best with the one we measured, we define an error function and find its minimum. Our error function is the quadratic error along x (respectively along y) of the distance between the points measured in the image and the result of the simulation.

$$\epsilon_x = \sum_{i=1}^{4}(x_i^{mesured} - x_i^{simu})^2$$

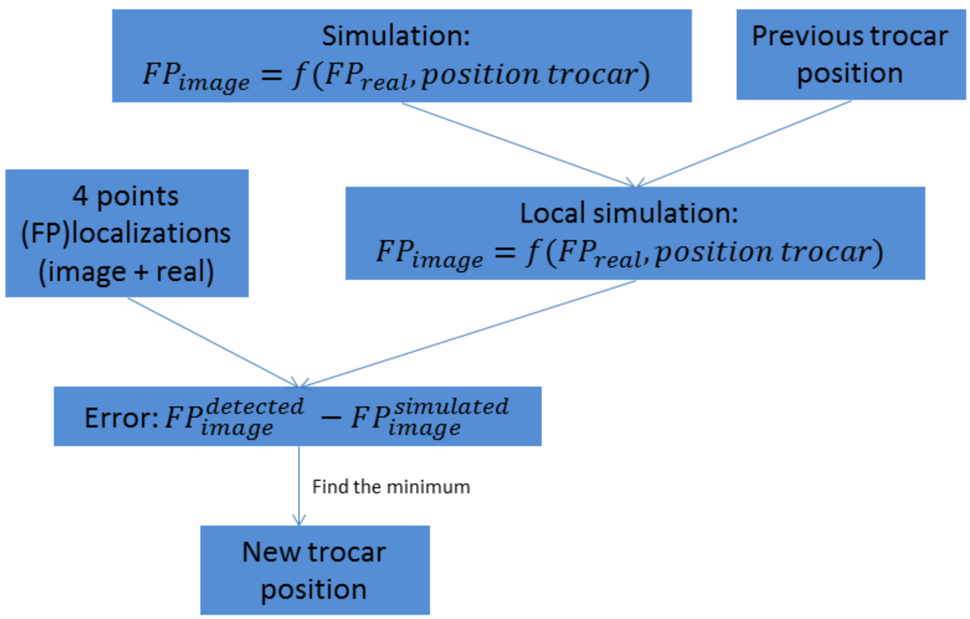

Figure 3.6: The algorithm of localization

The first results with only one degree of displacement were very encouraging. However, there was a problem as soon as we tried to estimate the parameters for a motion with more than one degree of freedom. Indeed, there were too many chances to find a local minimum instead of the global minimum of the error function (see

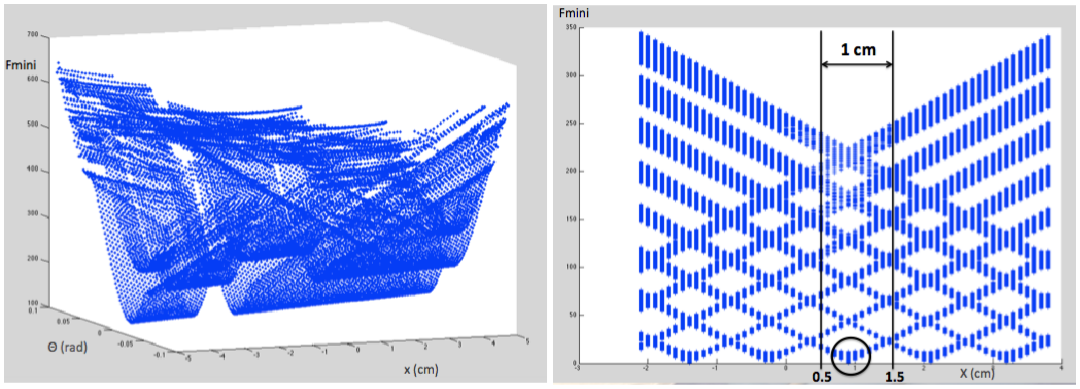

Figure 3.7: Error function to minimize : left, with no motion restriction; right, with motion restriction

Figure 3.7 left). The paper of Yuan [56] gave us the answer: in order to solve the exterior calibration problem of a single camera we needed to use four collinear points to solve the system with existence and uniqueness of the solution. At this point we entered into a dynamic problem and not a static one. Indeed, at the end what we wanted was to analyze videos. In a video, two successive images are linked, which means that we can use the parameters found in the images n-1 to compute the parameters of image n. By implementing that and putting a limitation on the movement of the trocar between two images, we achieve to finally have a unique minimum (see Figure 3.7 right). This implies a limitation on the surgeons' movements between two images of 0.5 cm for the translations and 4° for the rotations. Because the camera takes 30 images per second and the movement of a surgeon is not supposed to be very fast, we should never reach this limitation.

3.3.3 An alternative solution: pose estimation without poster

We imagine another approach to obtain the motion of the trocar that has been developed as an alternative to the first one for two reasons:

1. we needed a fast and efficient algorithm to obtain the tool trajectories, and
2. we wanted to stop the use of the poster on the ceiling, which can be a constraint for a bigger development of the Smart Trocar.

The exact knowledge of the position of the point in the real space (i.e. the knowledge of the poster) was crucial in our algorithm, and it seems impossible to use the same kind of approach to solve our pose problem. Nonetheless, we know that the motion of the trocar is mainly a rotation, along x and y (if z is the direction to the ceiling). Besides, there are markers on the ceiling that we are able to identify and track, even if we do not know their coordinates (air vent, light). If we are able to determine the motion, on the video, of a fixed object in the operating room, we will deduce the motion of the camera, thus, the motion of the trocar. In order to do this, we need to be able to detect accurately some feature points in the image, and then to recognize those points from a frame to the next one and match those points to determine its motion of the camera. We also have to be sure that the object that carries those points are fixed in the operating room (so the motion of those points corresponds to the motion of the camera); otherwise, we need to suppress the outliers (feature point detected on a moving object, e.g. the surgeon). To be sure that the detected points will be on the background (ceiling, air vent), we ask the user to define a region of interest to localize the background and we track features inside this region.

To identify the feature points, we need a fast and robust algorithm that can detect enough points in the image: the first we imagine is the one we presented in the first part, to detect the corners. The literature also suggests other detectors such as Self-Invariant Feature Transform (SIFT)[24], Speeded Up Robust Features (SURF) [5], or Harris detector [17]. We found some studies that compared those different algorithms [2][27]. Those studies suggested that the most adapted algorithm depends of our input images, and to provides the biggest number of features and correct matching. A good feature is repeatable, i.e. we can detect it on different images, even under different viewing conditions. It is distinctive enough from its neighborhood to be recognized, and small enough to be accurately localized. In our case, under bad light condition, SURF was barely able to detect 10 points. Harris was on the average less good than Minimum Eigenvalue (fewer points, and poor robustness to small-scale change). SIFT was the most time consuming which may be a problem for real-time applications.

Now that we have detected the feature points, we have to do the matching. There are two main approaches characterized by two algorithms: the use of a feature vector or directly looking for the best geometric transform. The SURF descriptor attributes a vector to each feature that characterizes its immediate neighborhood. These descriptors are used to pair points between images and find the best match. The other method, illustrated by the Kanade-Lucas-Tomasi (KLT) features a tracker algorithm, which tries to find the best geometric transformation. In our case, the second approach seems to be the most adapted for two reasons: first, the descriptor of SURF might not be very robust since we are looking at a very homogeneous and periodic surface with few elements to distinguish a specific point. Second, the geometric transformation should be exactly the same for all the detected points (since the only motion is the motion of the camera), so the algorithm should be able to find them. We have to note that the algorithm uses only the best matches to determine the geometric transformation between two frames; consequently, it is robust to few outliers.

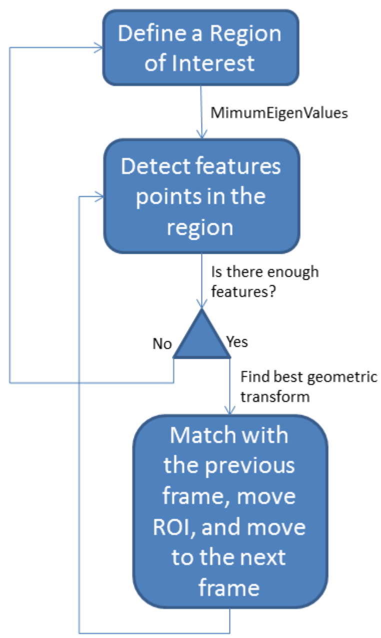

Figure 3.8: Algorithm of tracking

The advantage on this method based on a KLT tracker instead of a pose estimation problem is that it is based on well-known fully automatic algorithms, and does not need any marker. In addtion, the KLT method is cost effective. This tracker is efficient for small displacement, so we need to analyze around 30 frames per second, which is too much if we want algorithms that work in real time. Besides, our method only gives us the displacement into the image plane, which can be related to the displacement into the (x,y) plane, but cannot give the depth information. Some papers show that we can obtain the depth with these kind of methods, and this might be a future development [38][48].

Chapter 4

Results and Validation

4.1 Verification

The navigation system has been tested in an old operating room at the Houston Methodist Hospital, where we mounted the poster with the markers on the ceiling. Working in an operating room allows us to test our devices in realistic conditions, in terms of lightness, objects on the ceiling, and geometry of the room. Our camera is mounted on a home-made wood-platform which is possible to move along the six degrees of freedom (see Figure 4.1). The goal is to be able to move the camera in a measurable way between two shots. To do so, two laser pointers with measurement abilities are fixed on the platform too. These lasers give us an accurate measurement between the laser and the walls of the operating room. Thanks to this system, it is easy for us to measure (by using simple trigonometry and geometry of the room) the translation and the rotation that we apply to the platform between two shots. As we are looking for high accuracy results, this is very important for us to validate the results that we find with our algorithm.

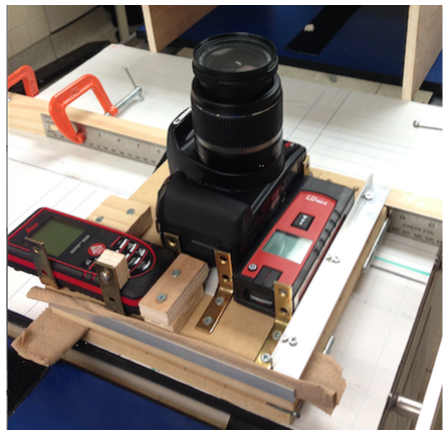

Figure 4.1: Validation with a professional camera

We first tested our algorithm with a professional camera, i.e. a camera without autofocus, a small opening angle, no obstructions, and no distortions on the images. Once our prototype was set up, we started to take pictures of the ceiling with a measured movement between each one. The fact that the camera was moving made the position of the crosses from the pattern moves from one picture to the next one. Our algorithm detects this movement and measures it to then reconstructs the rotation and translation associated to it. As we explained in the previous part, the motion between two frames should be small enough to allow our algorithm to converge; this is not a problem in the final application, since we are working on video. We tested our algorithm on different kinds of movements, such as rotations and translations, as it can experience during a real use of a trocar, allowing a motion of 6 mm and 4° between two frames. Repeating this experience a couple of times, we obtain an average error of 0.1° for the rotation and 0.5mm for the translation.

Now that we proved that our algorithm is working on fixed images, we also wanted to show its reliability on movies, because these are the data that we record from the camera once installed on the trocar during a real procedure. The goal is to keep this excellent level of accuracy in general conditions with a wireless camera. Also, the fact that we now work on a movie and not anymore on unrelated images brings us a new problem: the validation part here is more difficult and the movement between two frames is less controlled. To measure it, we use a similar system as previously, but additional lasers are placed on the trocar in such a way that they are visible in the field of view of the camera, which allows us to track the laser beam frame after frame to measure the motion of the trocar [49]. Working in a dynamic situation allows us to use the position of the trocar in the previous frame to estimate the position of the next one. Now, we can reduce the area of research to an interval centered on this position, which reduce drastically the computation time and allows us to approach real time condition. As we said before, thanks to the work of Yuan [56], we know that our algorithm will converge to the unique solution, if we stay under the range of $4°$ for the rotation and 5mm for the translation.

We repeat this experience in a different room, with different lightness and partial obstruction of the field of view of the camera, to obtain the final conclusion:

- *Accuracy*: the average accuracy on the rotation angle is $0.2°$ and 0.4 mm.
- *Tolerated motion between frames*: $4°$ of rotation and 5 mm of translation.
- *Range*: the total rotation tested is from $-20°$ to $20°$.
- *OR geometry*: the distance from the trocar to the ceiling varies from 110 cm

Figure 4.2: Validation with a regular Smart Trocar and two laser beam

to 170 cm with no significant influence on the accuracy.

- *Lightness*: the prototype has been tested with less lightness than usually in an OR: the camera compensates this lack of light and the accuracy is the same.

- *Obstruction*: the navigation system stays in the range of error fixed as long as two crosses are visible, i.e. when the features are separated of 13 cm or more one to the other, along x and along y. This correspond to an allowed obstruction of 80% of the field of view.

- *Computational time*: the total computational time is around 5 frames per second. Most of the time is needed to localize the feature in the image. This time can be reduced by introducing parallel programing and using a compiled programming language like C++ instead of Matlab.

4.2 Application to the analyze of tool path

4.2.1 Set up of the experience with the surgeons

As we have shown in part 2, laparoscopic training can be improved by providing real feedback to the residents. One of the most promising way to do this is to analyze the trajectories of the surgical tools and to compare them with the ideally expected motion. As we concluded in part 2, the system used to record the trajectories has to be cost effective, fast, easy to use, and should not constrain the surgeons in their motions or disturb them in any way. The Smart Trocar is the perfect answer to these constraints and can be used to record the tools trajectories of surgeons for training purposes. Since the literature also shows strong importance of mental engagement in the success or failure of the laparoscopic exercises, it would be interesting to also access some cognitive metric.

Taking into account all these parameters, we imagine an experience that aims to determine how the tools trajectories evolve with experience and dexterity, and how can we characterize and measure those differences, in order to objectively evaluate the ability of the practitioner and provide participants a feedback to help improve their performances. In order to do this, we asked volunteers to do different exercises on a Fundamental of Laparoscopic Surgery (FLS) model while we were recording the tools trajectories. In this work, we also aimed to analyze the importance of brain activity during the surgical tasks. We wanted to assess at what extent the ElectroEncephaloGram (EEG) signals reflect the trend of tasks, and if it can be correlated with performance and tools trajectories. The EEG signal was recorded with the Neurosky Mindwave commercial system (single dry electrode on the frontal area) and the eSense meter for attention and meditation were used to monitor alertness level. Attention and meditation are two functions of time, that can vary from 1 to 100.

We divided our subjects in two groups: beginners and surgeons. The beginners never practiced FLS exercises before. The surgeons were comprised of two populations: senior surgeons (considered as professionals) and surgical residents (considered as intermediates). Surgeons are considered experts in laparoscopic surgery, and have experience with FLS box. Thanks to their collaboration, we analyzed eight surgeons (three skilled surgeons and five residents) and ten beginners. To compensate those small samples, we asked everyone to do the exercises twice. All the subjects use the same FLS box, and practice in the same experimental conditions: we record the tool trajectories, the time, the EEG and the number of errors. In this thesis, we present the results we obtained analyzing the EEG of the participants when the subjects realized a specific FLS task that consist of a precise circular cutting of a piece of gauze (Figure 4.4). According to the FLS program, the proficiency levels is reached when the surgeons are able to realize the task two consecutive times in less than 98 seconds each. An error (cut outside the two circles) is not allowed. We provide the results of the participants in the following tables.

4.2.2 Preliminary results

The analysis of those trajectories can provide an objective and quantitative tool to evaluate dexterity of a practitioner. Following the work exposed in [4][12][26], we imagine different metrics to analyze the trajectories, based on the observed trajectories, and the EEG (the exact definition of these metrics is provided later):

- Length of the trajectories (mm).
- Average radius of curvature.
- Velocity.
- Area covered by the trajectory.
- Proportion of high frequency in the spectrum (measure of shiver).
- Average distance to the center of motion.
- Completion time.
- Correlation between attention and meditation (EEG).

We show on Figure 4.7 the results with the most significant metrics. We also ran some t-test to determine if the difference between the group was significant or not. The choice of a t-test is justified both by the small amount of data and by the non-Gaussian aspect of these distributions. Even if these results are very encouraging and confirm our literature review, they are preliminary; we only used 31 trajectories, obtained on 17 people (beginners and experts). Nonetheless, we can imagine that a combination of these metrics is able to accurately evaluate the dexterity of a resident.

Definition of the metrics: Let's name T the trajectory, $curv(t)$ the radius of curvature at the point t, med the meditation, att the attention, and t the time.

$$Area = (\max T_x - \min T_x) \cdot (\max T_y - \min T_y)$$

$$Energy = \| \frac{1}{N} \cdot T(f)_{HF} \|$$

$$Curvature = \mu \mathrm{max}(curv(t), th)$$

$$Symilarity(Att, Med) = \int_t | \ (med(t) - \mu_{med}) - (att(t) - \mu_{att}) \ | \ dt$$

Trial	Category	Time (s)	Mistakes
1	Surgeon	94	0
		97	0
2	Surgeon	120	0
		114	0
3	Surgeon	73	1
		83	0
4	Resident	122	0
		127	1
5	Resident	77	0
		67	0
6	Resident	51	1
		128	0
7	Resident	553	2
		461	1
8	Resident	108	0
		90	1
9	Novice	109	2
		124	2
10	Novice	287	1
		183	1
11	Novice	153	0
		228	1
12	Novice	242	0
		248	1
13	Novice	80	1
		398	2
14	Novice	50	0
		167	0
15	Novice	98	1
		190	1
16	Novice	83	2
		134	0
17	Novice	506	2
		261	2
18	Novice	173	0
		189	1

Figure 4.3: Completion time and mistakes for all the participants

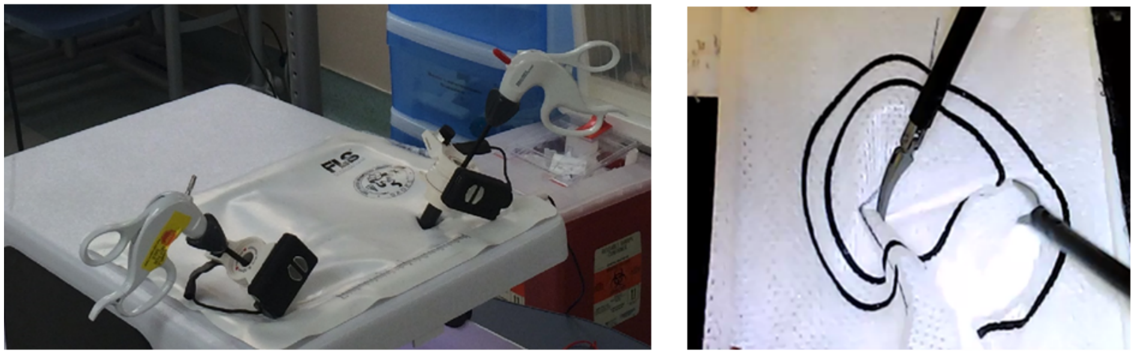

Figure 4.4: Left: FLS box with Smart Trocar; right : one of our exercises: cut between the lines

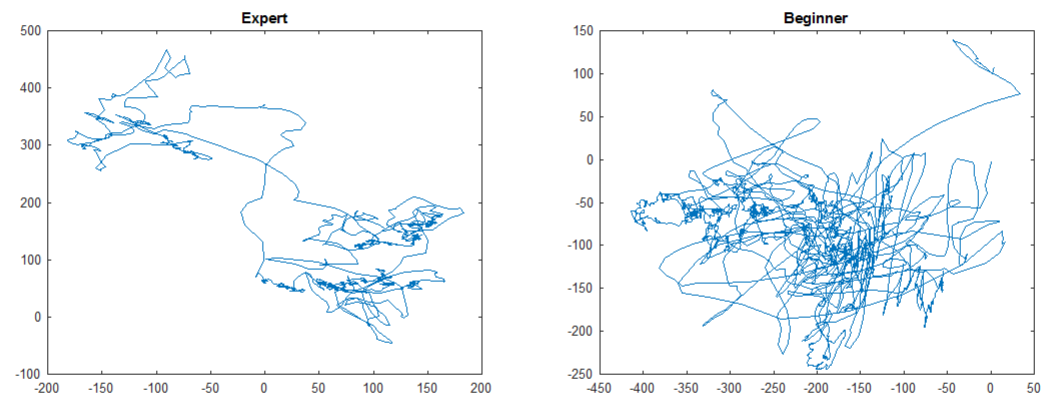

Figure 4.5: Example of trajectories, for the different populations

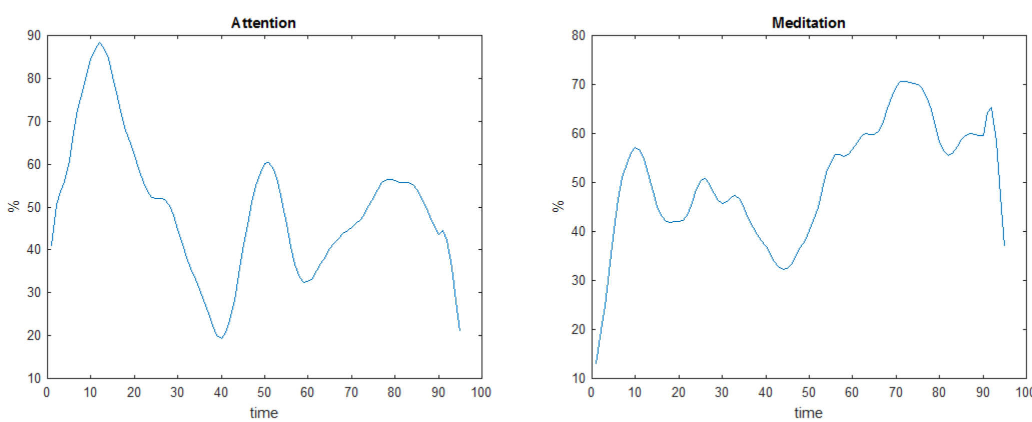

Figure 4.6: Attention and meditation, provided by the Neurosky system

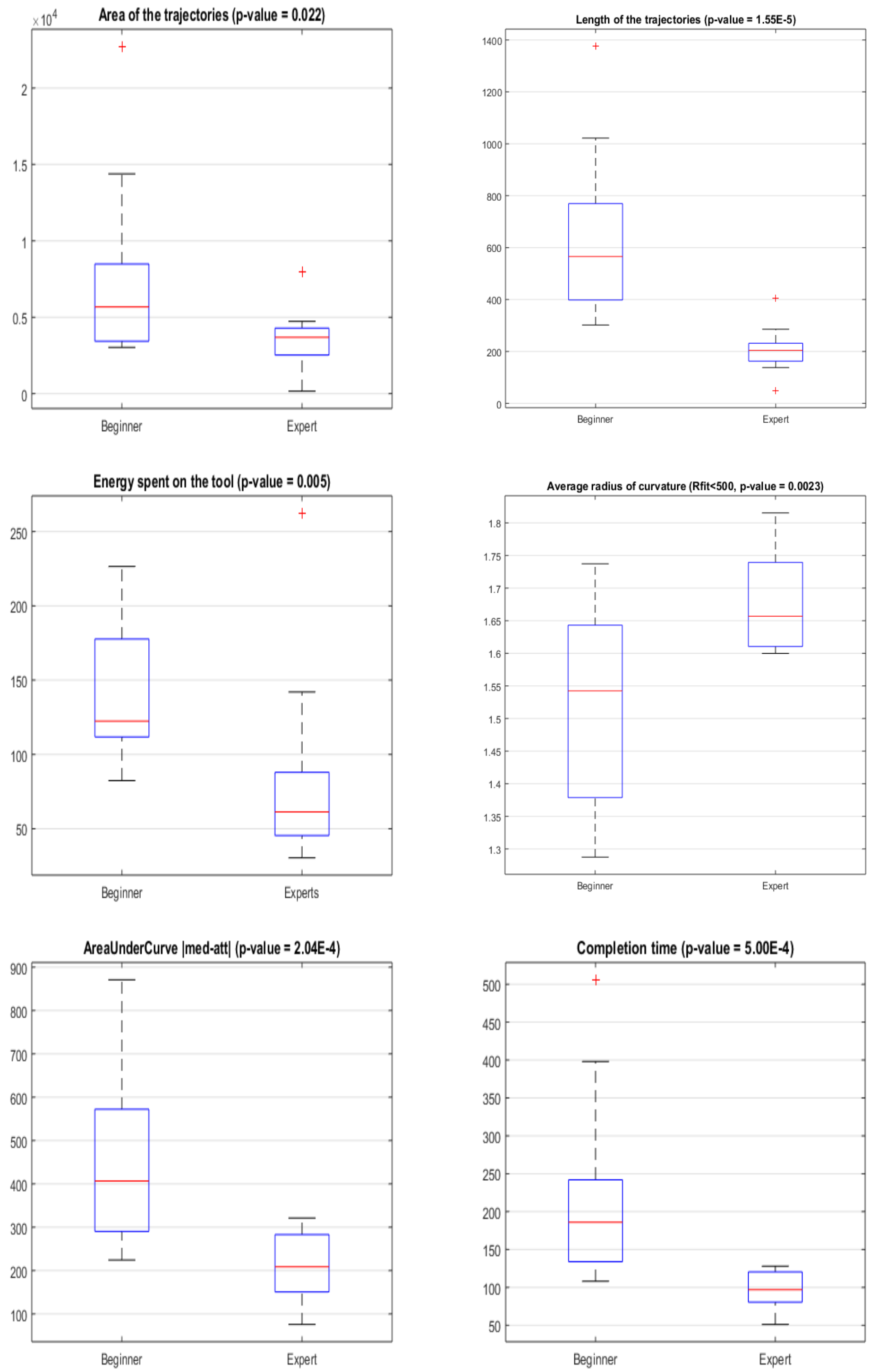

Figure 4.7: Different metrics and t-tests

Chapter 5

Perceptive

5.1 Operating room awareness

Operating room awareness is one of the main goals we want to achieve and many systems have been developed for this purpose. Most of them deal with the recognition of surgical procedures. Padoy et al. [33] presented a method to identify the phase of a surgical operation based on the synchronization of signals recorded during the operation. This technique is very simple because it only needs to recognize some key events in the surgical procedure to recognize the step the surgeon is at. Doryab et al. [11] use the same kind of device and go one step further by creating a recommender learning system that uses the recorded parameter to tell the doctors and nurses what they should do according to the ongoing step of the surgery. This system is based on a reinforcement learning algorithm with a patient safety-based calculation of the utilities. A similar application can be done with the Smart Trocar.

Since our technology can record the type of instrument used during the operation [49], it can draw a timeline of the surgical operation, and track some key events such as the use of a specific instrument that allows us to determine the current step of the operation, prepare the next one, and help to improve operating room management by informing the coordinator of the operating room system when an operation is nearly complete. Since the different steps of a laparoscopic operation are relatively systematic and predictable, a change on the timeline provided by the Smart Trocar allows the operating room coordinator to know quickly that there is an issue in the corresponding operating room (see Figure 5.1). In this example, the Smart Trocar shows irregularities in the procedure due to a disfunctionement of the clip applier. At the end, the Smart Trocar can automatically generate notifications of when an operation starts and ends, send advance warnings when the operation will be done, detect singularities during the operation, detect when a particular stage of the surgery takes longer than usual, and improve readiness of the staff by providing information on the next step while automatically provide summarized reports. This application is priceless for a management hospital team, when we know that minimally invasive surgery procedures are a growing market that already represents 3.5 million procedures per year in the USA, that operating room time costs about $100 a minute in North America, and that current reports show operating room efficiency is not better than 70% in ideal cases [47].

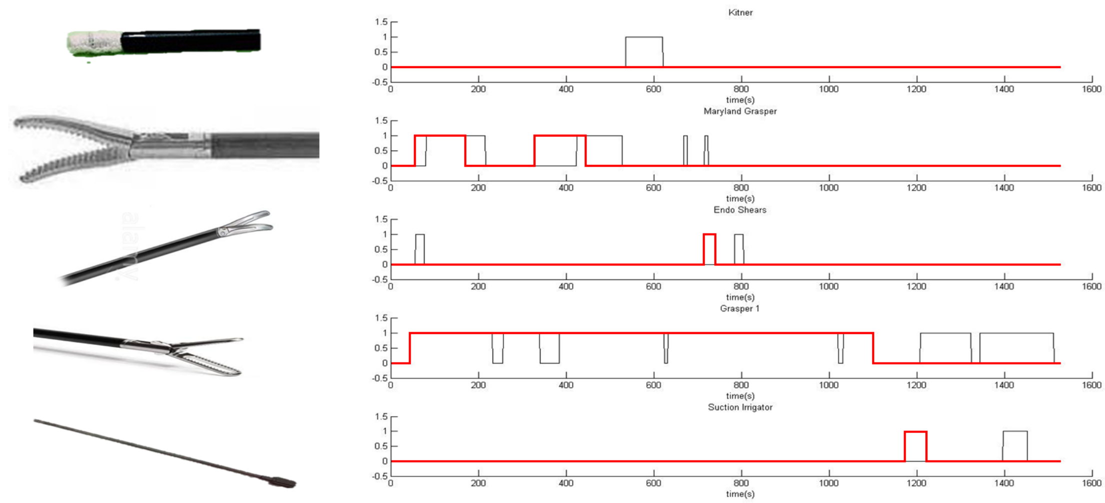

Figure 5.1: Output of the Smart Trocar system (black) compared to an ideal version procedure (red)

To summarize, the Smart Trocar delivers the following services:

- Recognition of specific types of instrument (see Figure 5.1).
- Accurately detects duration and port used for the insertion or removal of instruments.
- Reconstructs the motion of these instruments in space and time.
- Provides a time series of events during the surgery from first to last use of laparoscopic instruments.
- Provides retrospective analyses of cases to allow identification of step(s) of the operation that may be improved to safely increase efficiency.
- Allows the generation of statistical models of standardized, large volume procedures to provide a real-time notification of step(s) or parts of the operation

that fall outside of expected ranges, thus improving safety and efficiency

From this information, we can generate various intelligent feedback mechanisms such as:

- Automatically generate notifications regarding when an operation starts and ends.
- For large volume surgeries such as gastric bypass or cholecystectomy:
  - Send an advanced notification about when the surgery will be completed.
  - Detect singularities in the surgery.
  - Detect when events are not done in order.
  - Detect when a particular stage of the operation takes longer than usual.
  - Improve readiness of the staff team by providing details on the next step of the procedure.
  - Automatically generates reports of the operation.

5.2 Augmented reality

5.2.1 Three-dimensional reconstruction of the organs and the tools

The Smart Trocar gives us the position and orientation of the surgical tools inside the body. If we combine it with a 3D endoscope, and if we are able to register these 3D maps, we can obtain a 3D reconstruction of both the tools and the organs. We can then localize the tool in reference to a specific organ, even when they are outside the field of view of the endoscope (since we have their position thanks to our 'GPS' system), and we can compensate the lack of depth perception characteristic of minimally invasive surgery. The registration between the coordinate system of the Smart Trocar and the one of the endoscope can be done if we are able to associate three of the points from one system to the other. Since the Smart Trocar is, in essence, able to localize the tip of a surgical tool, in space, we can use a tool such as a pointer to obtain the coordinate of the points we need in order to solve the registration problem.

5.2.2 Registration of preoperative and intraoperative images: add the Finite Element (FE) model

Preoperative surgical planning coupled with accurate image-guided surgical therapy for liver resection would be a valuable tool for increasing the possibility of favorable outcomes for patients. Surgical planning is based on preoperative scans, segmented and digitalized in order to obtain an accurate 3D model of the organ with sub-surfacique data. We identify on these scans the important parts such as tumors, blood vessels, etc. We use this segmentation to generate a biomechanical finite element model that contains valuable information for the surgeon. During the operation we have intraoperative data: e.g. a surface acquired as a point cloud representation using tracked surface, or/and subsurface, acquired through intraoperative ultrasound. This information is less accurate than the preoperative data, and a lot of work has been done in order to fit the finite element model (preoperative) and the actual position of the organ(intraoperative). (Image 5.2 from [53])

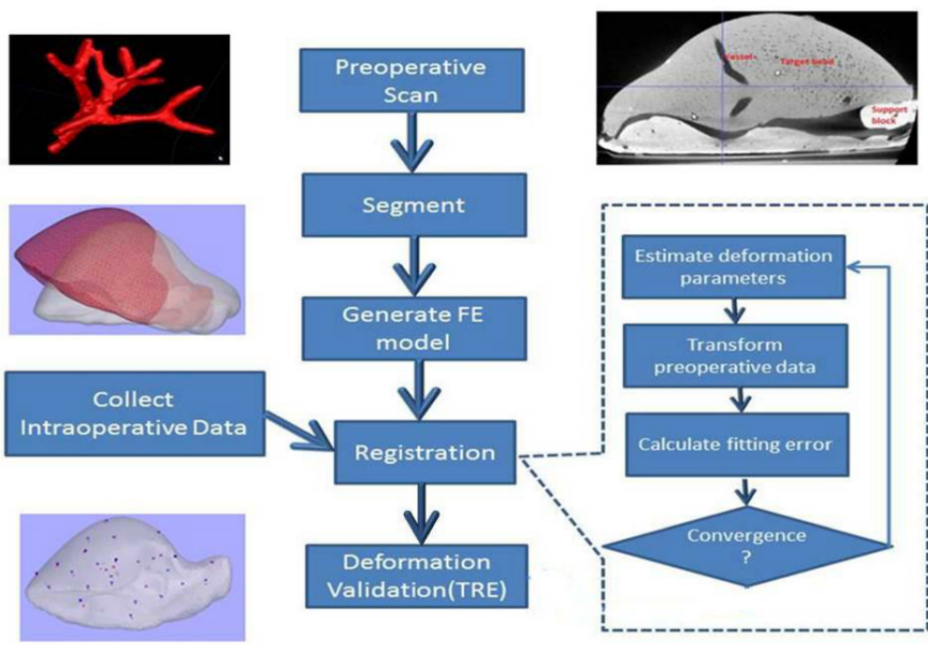

Figure 5.2: Registration principle in image guided surgery

The Smart Trocar can help to semi-automatically register the preoperative and intraoperative data. Indeed, some markers on the preoperative data, such as a ligaments, can be localized on the endoscopic view by using the known coordinates surgical tool's tip. As we can see on figure 5.3, if we use tool 1 as a pointer (touching the feature on the liver), and if the field of view of tool 2 (endoscope) is centered on the same point, we are able to measure the distance O2 and the position of the tip of tool 1, thanks to our knowledge of the inclination and length of the tools, and the distance m1 and m2. By using the Smart Trocar, surgical tools with an appropriate setup can be used as laparoscopic navigation pointer: palpate and locate blood vessels, tumors, and organs, even when they are not visible on the intraoperative surfacique data.

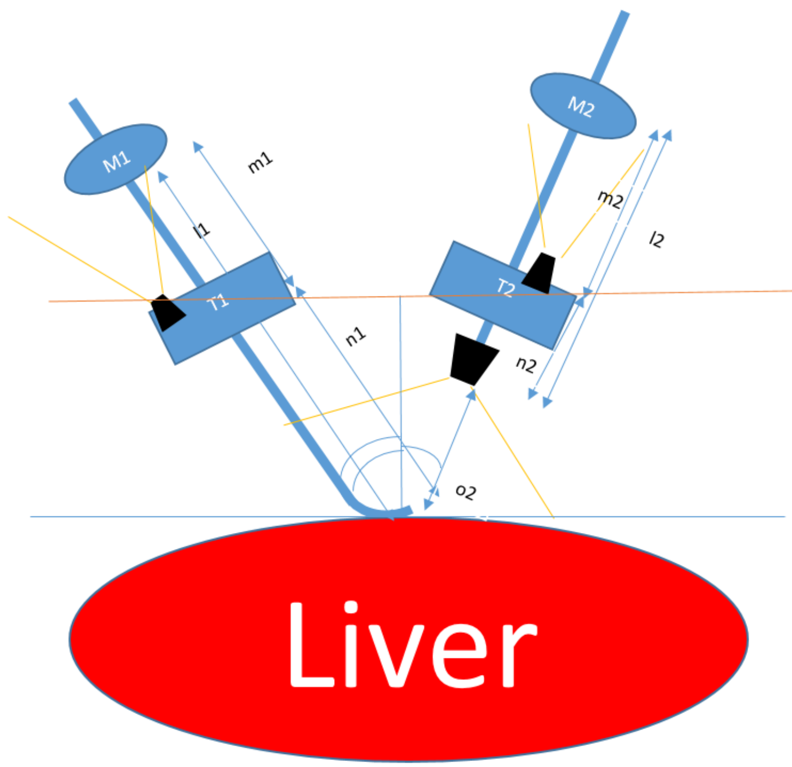

Figure 5.3: Geometry of our problem: how to register the 3D surface of the liver

5.2.3 Analysis of the trajectories to prevent dangerous events

Not only can the Smart Trocar be used as a 'GPS' to track the surgical tools, and as a pointer to localize important organic structures, but it can also analyze the trajectories of the surgical tools. We previously showed that it can analyze the trajectories to determine the dexterity of the surgeon, but we can also analyze the trajectories in real-time, to track some specific events, and help surgeons to achieve their task by:

- Showing the best trajectory to reach the target avoiding important dangerous strucures.
- Sending a reminder when dangerous/unexpected situation are detected, such as:
  - Entrance in a no fly zone.
  - Retractor/Grasper that has not moved for a while and may compress tissues.
  - High velocity or unusual trajectories.

Conclusions

At the end of the day, the Smart Trocar project tackles the issues of context-aware systems for operating rooms. Our goal is to compensate the laparoscopy drawbacks by building a new channel of information designed to help the surgeon. There are many projects that try to address the problems caused by indirect vision and limited access to the operating field during laparoscopic surgeries. Most of them see the localization of surgical instruments as an important tool to overcome the difficulties encountered by the surgeon during procedures. But most of the systems that have been designed until now suffer from big disadvantages, either because they are too sensitive to noise, not accurate enough, too space consuming, or too expensive. Our system, on the other hand, is simple, cheap, and accurate enough for surgical needs. Moreover, it is non-invasive, and can be mounted on classic trocars. Not only can our technology be used to assess surgical dexterity and improve surgical training, but it can also improve operating room management by providing a timeline of events. On top of this, we imagine that our Smart Trocar will allow us to use surgical tools as a pointer to localize important features, and improve safety and efficiency of minimally invasive surgery.

Bibliography

[1] Allan M., Ourselin S., Thompson S., Hawkes D.J., Kelly J., and Stoyanov D., “ Toward detection and localization of instruments in minimally invasive surgery.” *Biomedical Engineering, IEEE Transactions*, 60(4):1050-1058, 2013.

[2] Babbar G., Bajaj P, Chawla A., and Gogna M., “A comparative study of image matching algorithms.” *International Journal of Information, Technology and Knowledge Management.* 2(2): 337-339, 2010

[3] Ballard. D.H., “Generalizing the Hough transform to detect arbitrary shapes.” *Pattern Recognition*, 13(2):111-122, 1981

[4] Bann S, Khan M, Darzi A, “Measurement of surgical dexterity using motion analysis of simple bench tasks,” *World, Journal of Surgery*, 27(4):390-394, 2003

[5] Bay H., Ess A., Tuytelaars T., and Van Gool. L., “SURF: Speeded up robust features.” *Computer Vision and Image Understanding (CVIU)*, 110(3):346-359, 2008

[6] Brauer-Burchardt C, and Voss K., “Automatic lens distortion calibration using single views”, *Mustererkennung*, 13:187-194, 2000

[7] Brown D.C., “Close-range camera calibration”, *Photogrammetry Engeneering,*, 37(8):855-866, 1971

[8] Coughlan J. and Manduchi. R., “Color targets: Fiducials to help visually impaired people find their way by camera phone.” *EURASIP Journal on Image and Video Processing*, 2007:10-32, 2007

[9] Coughlan J., Manduchi M., Mutsuzaki, and Shen. H., “Rapid and robust algorithms for detecting colour targets.” *In 10th Congress of the International Colour Association,* AIC Colour, 5:959-962, 2005.

[10] Devernay F., and Faugeras 0., "Automatic calibration and removal of distortion from scenes of structured environments", *SPIE,* 2567:62-72, 1995.

[11] Doryab A., Togelius J., and Bardram. J., " Activity-aware recommendation for collaborative work in operating rooms." *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces,* 21(3):301-304, 2012.

[12] Dosis A., Aggarwal R., Bello F., Moorthy K., Munz Y., Gillies D., Darzi A., "Synchronized video and motion analysis for the assessment of procedures in the operating theater," *Archive of Surgery*, 140(3):293-302, 2005

[13] Dutkiewicz P., Kielczewski M., and Kowalski. M., "Visual tracking of surgical tools for laparoscopic surgery." *Proceedings of the Fourth International Workshop*, RoMoCo'04 Conference, 23-28, 2004.

[14] Dutkiewicz P., Kietczewski M., Kowalski M., and Wroblewski. W. " Experimental verification of visual tracking of surgical tools." *Proceedings of the Fifth International Workshop*, RoMoCo'05 Conference, 237-242, 2005.

[15] Estebanez B., del Saz-Orozco P., Rivas I., Bauzano E., Munoz V., and Garcia-Morales I., " Maneuvers recognition in laparoscopic surgery: Artificial neural network and hidden markov model approaches." *Biomedical Robotics and Biomechatronics (BioRob),* 2012 4th IEEE RAS and EMBS International Conference, 1164-1169. IEEE, 2012.

[16] Gould J.C. and Frydman J. "Reverse-alignment surgical skills assessment." *Surgical Endoscopy,* 21(4):669-671, 2007.

[17] Harris C., and Stephens. M.J, "A combined corner and edge detector." *Proceedings of the 4th Alvey Vision Conference.*, 147-152, 1988

[18] Hsu J.and Payandeh. S., "Toward tool gesture and motion recognition on a novel minimally invasive surgery robotic system." *Proceedings 2006 IEEE International Conference*, ICRA 2006 Conference, 631-636, 2006.

[19] Juan L. and Gwun. O., "A comparison of sift, pca-sift and surf." *International Journal of Image Processing (IJIP)*, 3(4):143-152, 2009.

[20] Khan N.Y., McCane B., and Wyvill B. "Sift and surf performance evaluation against various image deformations on benchmark dataset. In Digital Image Computing Techniques and Applications (DICTA)," 2011 *International Conference*, pages 501-506. IEEE, 2011.

[21] G. K. Kloss, H. Shin, and N. H. Reyes. "Dynamic colour adaptation for colour object tracking." *In Image and Vision Computing New Zealand,* 2009. IVCNZ'09. 24th International Conference, pages 340-345. IEEE, 2009.

[22] S.-Y. Ko, J. Kim, W.-J. Lee, and D.-S. Kwon. "Surgery task model for intelligent interaction between surgeon and laparoscopic assistant robot." *International Journal of Assitive Robotics and Mechatronics,* 8(1):38-46, 2007.

[23] Korndorffer JR Jr, Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ. "Simulator training for laparoscopic suturing using performance goals translates to the operating room." *J American College of Surgery*, 201:23-29, 2005

[24] Lowe D. "Object recognition from local scale-invariant features". *Proceedings of the International Conference on Computer Vision 2.*, 1150-1157, 1999

[25] Lum M, Friedman D, Sankaranarayanan G, King H, Fodero K, Leuschke R, Hannaford B, Rosen J, Sinanan M, "The RAVEN: Desing and Validation of a Telesurgery System," *The International Journal of Robotic Research*, 28(9):1183-1197, 2009

[26] McBeth B., Hodgson A., Nagy A, Qayumi K., "Quantitative Methodology of Evaluating Surgeon Performance in Laparoscopic Surgery", *IOS Press,* 2002

[27] Mikolajczyk, K., T. Tuytelaars, C. Schmid, A. Zisserman, T. Kadir, and L. Van Gool. "A Comparison of Affine Region Detectors." *International Journal of Computer Vision.*, 65(1-2):43-72, 2005

[28] F. Miyawaki, K. Masamune, S. Suzuki, K. Yoshimitsu, and J. Vain. "Scrub nurse robot system-intraoperative motion analysis of a scrub nurse and time based model for surgery." *Industrial Electronics, IEEE Transactions* 52(5):1227-1235, 2005.

[29] A. A. Navarro, A. Hernansanz, E. Villarraga, X. Giralt, and J. A. Lopez. "Automatic positioning of surgical instruments in minimally invasive roboticsurgery through vision-based motion analysis." *IROS,* 2(3-4):202-207, 2007.

[30] A. Nishikawa, S. Asano, R. Fujita, T. Yohda, F. Miyazaki, M. Sekimoto, M. Yasui, S. Takiguchi, and M. Monden. " Robust visual tracking of multiple surgical instruments for laparoscopic surgery." *CARS,* 1256:1372-1372, 2003.

[31] P. Ordonez, P. Kodeswaran, V. Korolev, W. Li, O. Walavalkar, B. Elgamil, A. Joshi, T. Finin, Y. Yesha, and I. George. "A ubiquitous context-aware environment for surgical training." *Mobile and Ubiquitous Systems: Networking*

54 and Services, 2007. MobiQuitous 2007. Fourth Annual International Conference on, 1-6. IEEE, 2007.

[32] E. Oyallon and J. Rabin. "An analysis and implementation of the surf method, and its comparison to sift." *Image Processing On Line,* 1:1-13, 2013.

[33] N. Padoy, T. Blum, S.-A. Ahmadi, H. Feussner, M.-O. Berger, and N. Navab. "Statistical modeling and recognition of surgical workow." *Medical Image Analysis,* 16(3):632-641, 2012.

[34] Pavlidis I, Tsiamyrtzis P, Shastri D, Wesley A, Zhou Y, Lindner P, Ruddharaju P, Joseph R, Mandapati A, Dunkin B, Bass B, "Fast by nature - How stress patterns define human experience and performance in dexterous tasks," *Nature - Scientific Reports,* 305:1-9, 2012

[35] Prabhu A, Smith W, Yurko YY, Acker C, Stefanidis D. "Increased stress levels may explain the incomplete transfer of simulator-acquired skill to the operating room." *Surgery*, 147:640-645, 2010

[36] R. Rosales and S. Sclaro. "3d trajectory recovery for tracking multiple objects and trajectory guided recognition of actions." *In Computer Vision and Pattern Recognition,* IEEE Computer Society Conference on., 2(79):1481-1497, 1999.

[37] Seymour NE, Gallagher AG, Roman SA, et al. "Virtual reality training improves operating room performance: results of a randomized, double-blinded study." *Annals Surgery*, 236:458-463; discussion 463-464, 2002

[38] Shi, J., and C. Tomasi, "Good Features to Track," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 593-600, 1994

[39] Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayez R, Fried G. "Fundamentals of Laparoscopic Surgery simulator training to proficiency improves laparoscopic performance in the operating room a randomized controlled trial. " *American Journal of Surgery,* 199:115-120, 2010.

[40] M. Smereka and I. Duleba. "Circular object detection using a modified hough transform." *International Journal of Applied Mathematics and Computer Science,* 18(1):85-91, 2008.

[41] C. Staub, C. Lenz, G. Panin, A. Knoll, and R. Bauernschmitt. "Contour-based surgical instrument tracking supported by kinematic prediction." *In Biomedical Robotics and Biomechatronics (BioRob),* 2010 3rd IEEE RAS and *EMBSInternational Conference*, pages 746-752. IEEE, 2010.

[42] J. Stoll, H. Ren, and P. E. Dupont. "Passive markers for tracking surgical instruments in real-time 3-d ultrasound imaging." *Medical Imaging,* IEEE Transactions on, 31(3):563-575, 2012

[43] T. Suk and J. Flusser. "Combined blur and afine moment invariants and their use in pattern recognition." *Pattern Recognition,* 36(12):2895-2907, 2003.

[44] F. Tatar, J. Mollinger, J. Bastemeijer, and A. Bossche. "Time of flight technique used for measuring position and orientation of laparoscopic surgery tools.", *IEEE Sensors Conference,* 3:1480-1483, 2004.

[45] F. Tatar, J. Mollinger and A. Bossche, " Ultrasound system for measuring position and orientation of laparoscopic surgery tools," *Proceedings of IEEE Sensors,* 2:987-990, 2003

[46] F. Tatar, J. Mollinger, R. Den Dulk, W. van Duyl, J. Goosen, and A. Bossche. "Ultrasonic sensor system for measuring position and orientation of laproscopic instruments in minimal invasive surgery.", *Microtechnologies in Medicine*, Biology 2nd Annual International IEEE-EMB Special Topic Conference on, pages 301-304. IEEE, 2002.

[47] Thomson Reuters InPatient and OutPatient Database 2009, available on http://www.tdrdata.com/

[48] C. Tomasi and T. Kanade, "Detection and Tracking of Point Features", *Technical Report CMU-CS-91-132*, Carnegie Mellon University, April 1991.

[49] G. Toti, M. Garbey, V. Sherman, B. L. Bass, and B. J. Dunkin. "A smart trocar for automatic tool recognition in laparoscopic surgery." *Surgical Innovation,* 22(1):77-82, 2015.

[50] C.-C. Tsai and T.-Y. Wang. "Small humanoid robot localization by landmarks." *Intelligent Control and Automation,* 2008. WCICA 2008. 7th World Congress on, pages 6584-6589. IEEE, 2008.

[51] T. Uchida, H. Koyama, T. Komeda, M. Miyagi, and H. Funakubo. "Measuring method of three-dimensional position and orientation of a moving object for medical robots." *Systems, Man, and Cybernetics,* 1996., IEEE International Conference on, volume 2, pages 887-892. IEEE, 1996.

[52] N. Vu and B. Manjunath. "Shape prior segmentation of multiple objects with graph cuts." *Computer Vision and Pattern Recognition,* 2008. CVPR 2008. IEEE Conference on, pages 1-8. IEEE, 2008.

[53] Y. Wu, D. C. Rucker, R. H. Conley, T. S. Pheiffer, A. L. Simpson, S. K. Geeverghese, M. I. Miga, "Registration of Liver Images to Minimally Invasive Intraoperative Surface and Subsurface Data." *SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, Conference*, 2014

[54] A. Yilmaz, O. Javed, and M. Shah. "Object tracking: A survey." *Association for Computing Machinery, Computing Surveys,* 38(4), article 13, 2006.

[55] K. Yoshimitsu, T. Tanaka, K. Ohnuma, F. Miyawaki, D. Hashimoto, and K. Masamune. "Prototype development of scrub nurse robot for laparoscopic surgery." *In International Congress Series,* volume 1281, pages 845-850. Elsevier, 2005.

[56] J.-C. Yuan. "A general photogrammetric method for determining object position and orientation." *Robotics and Automation,* IEEE Transactions on, 5(2):129-142,1989.

[57] Yurko Y, Scerbo M, Prabhu A, Acker C, Stefanidis D, "Higher Mental Workload is Associated With Poorer Laparoscopic Performance as Measured by the NASA-TLX Tool," *Journal of the Society for Simulation in Healthcare,* 5(5):267-271, 2010