Abstract—This paper proposes a new user interface based on a maneuver recognition system, which models the surgeon behavior. This interface includes three different modules: data acquisition and coding, training system and on-line recognition system. The aim is defined as recognizing the surgeon movements while performing a surgical maneuver, by using a 3D surgical tool tracker. The obtained measurements are converted into movement symbols by means of a Wavelet transform and a fuzzy clustering. These symbols are used both for training HMM and for recognizing the current maneuver. The system has been tested in some in-vitro experiments performing a fictitious surgical protocol.

I. INTRODUCTION

MINIMALLY invasive surgery is intended to reduce the incisions that are practiced on the patient, with the aim of improving the postoperative convalescent time and also to diminish any complications. In order to carry out these procedures, a varying number of small incisions are done on the patient, through which the surgeon inserts special long instruments, and an endoscope, which has an incorporated camera [1].

One of the challenges of robotic surgical engineering is focused on substituting the human assistant in the procedures of laparoscopic surgery. In these types of techniques it is critical that the assistant focuses the camera precisely over the area of the operation, this can amount to fatigue or stress influencing, to a great extent, to the quality of the accomplishment of their task. Consequently human error can occur such as the endoscope making contact with the tissue, or where the area required by the surgeon is not focused on, even to unstable images being seen on the monitor through the lack of steadiness of the assistant. It is also necessary to emphasize that these types of techniques require high skill as there are considerable movement limitations, this leaves the surgeon through having to learn eye and hand coordination and having to compensate the loss of sensitivity and 3D vision. Through the use of surgical robots, such disadvantages are hoped to be eliminated in this type of surgical technique.

A human-machine interface is needed in this kind of robotic assistants. In this way, some systems used joystick devices, others made automatic movements to follow the surgeon tool with the endoscope [2], receive the surgeon orders through a gyroscope attached to the surgeon head [3], or are guided with the voice [4]-[6]. 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 [7].

Based on these works, it appears a new concept where the assistant is completely replaced by a robot: solosurgery. It refers to a situation in which the assistant is replaced by a robot in his functions during the surgical intervention, so that the surgeon can operate solo.

In this way, this article proposes a new user interface based on a maneuver recognition system, which models the surgeon behavior. The goal is not only to identify the surgical maneuvers which can be automated, but also to establish models for the different types of procedure, describing the interaction between the robot and the surgeon.

Therefore, section II starts by defining surgical protocols and the best methodology to model them. Sections III and IV state the parameters of the chosen models and which techniques have been used for their development. The experiments design and their environment are exposed on section V. The robotic assistant global system, including this surgeon model, is presented on section VI. Finally, on section VII are resumed some conclusions about results.

II. INTERVENTION MODELS

The objective of this work concentrates in modeling a set of minimally invasive surgery procedures. The methodology in the preparation of these models is based on a first phase study on what are the clinical protocols of the operations [8]. This analysis allows evaluating the workflow of a surgical procedure [9]-[11]. As a consequence, the surgical protocol is divided into simple, easy to evaluate stages; these phases of the operation are called maneuvers. At this stage, the models represent a set of maneuvers that in an ideal form the surgeon and his assistant should complete for the finalization of the operation. Therefore, a general operation can be divided into simple maneuvers [12],[13], that are connected in a systematic and organized way, for example, in the cholecystectomy procedure, a single maneuver is to clip and cut the cystic duct. Within each of these stages of the procedure there always exists a prerequisite that indicates the following maneuver. This condition could be a special given gesture from the surgeon or even the changing of a surgical
instrument. On the other hand, all of the procedures share a common set of basic actions that represent the instrument movements. This way, a maneuver is identified as a specific sequence of these basic actions.

In endoscopic surgery the basic actions are characterized by the movements of the instruments that are limited to four degrees of freedom. This is due to the movement restrictions that are imposed from the point of inserting the instrument into the abdomen of the patient (fulcrum point). In this way the basic actions in the handling of the laparoscopic camera are: left-right, up-down, to rotate on itself and insertion-removal (see Fig. 1). In the case of another endoscopic instrument it is necessary to add other functions as clamp opening and closing.

![Fig. 1. The four free degree of freedom of a laparoscopic tool.](image)

Therefore, the proposed operation model has two defined layers (see Fig. 2). The superior level responds to a deterministic mathematical model, as the order in which the maneuvers \( \beta_0, \beta_1, \ldots, \beta_i \) are fixed in an operation and the relationships between them and the events are known. Each maneuver is defined by an inner layer throughout a stochastic model that is represented by the states \( s_1, s_2, \ldots, s_N \), which are the basic actions. The mathematical representation of the maneuvers should allow them to be analyzed without the necessity of directly observing them, and methodologies with capacity of recognition and prediction should be allowed. The Hidden Markov Models (HMM) fulfill these characteristics, and provide a high flexibility when modeling the behavior of the surgeon in a maneuver. In this way, the actions that the surgeon makes are like a black box for the spectator, where the only information available is speed and position data at each time frame. From these parameters the prediction tools must estimate the current state of the operation. There are works that have used the stochastic model in minimally invasive surgery to analyze the dexterity skills of the surgeons in surgical maneuvers [14]-[18], in other problems of pattern recognition of medical signals [19], and in the automatic recognition of speech [20],[21].

III. MODEL THEOREY

This section describes the basic aspects of the deterministic models used to represent the surgical protocol, in addition to the HMM used to model the maneuvers.

A. Deterministic Model for Protocol Intervention

As it has been previously commented on, when a surgical protocol is performed, there are a sequence of maneuvers whose relations are fixed and organized. The parameters that define an intervention protocol are \( DM=\{\Delta, T, \beta_0, \beta_f\} \) where:

\[
\Delta = \{\beta_1, \ldots, \beta_i\}, \text{ is the set of maneuvers of a protocol.}
\]

\[
T(\beta_i)=\beta_{i+1}, \text{ the function with defines the transition from the current maneuver } \beta_i \text{ to the following one defined } \beta_{i+1} \text{ in the surgical protocol.}
\]

\[
\beta_f, \text{ end of the intervention or error situation.}
\]

\[
\beta_0, \text{ initial maneuver.}
\]

B. Stochastic Model for Maneuvers

The Hidden Markov Models is the stochastic model to define the maneuvers. These models have a finite number of states \( N \), where each state represents a basic action and each action is defined by a set of observable outputs. The network that represents the HMM advances in each time interval to a new state, emitting in each of them an observable signal output, according to a probability distribution matrix of emission. The topology of the network is defined by a probability distribution matrix of state changes that marks the relationships between the basic actions. Therefore, the parameters that characterize the HMM of a general maneuver are \( \lambda=(S,E,A,B,\pi) \), where:

\[
S=\{s_1, s_2, \ldots, s_N\}, \text{ is a set of basic actions during one maneuver.}
\]

\[
E=\{e_1, e_2, \ldots, e_K\}, \text{ K observable outputs in a given state, which have been defined as a function of the tip surgical tool position and velocity values.}
\]

\[
A=\{a_{ij}\}, \text{ being } a_{ij}=P[s_{t+1}=s_j \mid s_t=s_i], \text{ probability distribution matrix of the states transition.}
\]

![Fig. 2. Intervention Model.](image)
\[ B = \{ b_j(k) \} \text{ being } b_j(k) = \{ v_k \text{ in } t \mid s_t = s_j \} , \text{ probability distribution matrix of observed outputs in each state } j. \]

\[ \pi = \{ \pi_i \} \text{ where } \pi_i = P \{ s_i \text{ in } t=1 \} , \text{ initial distribution of states.} \]

Once the topology of the maneuver has been defined, two basic problems in the analysis of HMM appear [22]: the problem of the learning and the problem of the inference. Given a sequence of observations and the basic actions, the learning problem of a maneuver consists of estimating the parameters of the model, the matrices \( A \) and \( B \). On the other hand, the problem of the inference consists in obtaining the succession of hidden states that correspond to a sequence of given observations. The learning problem is solved with the algorithm of Baum-Welch [23],[24] and the inference with the algorithm of Viterbi [25]. This way, the next section presents the process on how to get a Hidden Markov Model for the considered maneuvers.

IV. SURGEON BEHAVIOR MODEL

The surgeon behavior model is based on a surgical maneuver library, where each individual component of the library is modeled by means of a HMM. In this way, a single full-connected Markov network typifies the surgeon’s hand movements when a particular surgical maneuver is performed.

The surgeon behavior model building requires a training process with actual data, which are produced by surgeons while they are performing different actions. In this way, it is obtained the HMM parameters which defines the different surgeon maneuvers. This training process is made off-line.

During a surgical procedure, the surgeon behavior model is used by a real-time recognition system, which identifies the current surgical maneuver. The information provided by this system will be employed both to inform the surgical robot assistant about the current surgical protocol stage and to supervise the surgeon actions.

Both, the training process and the real-time recognition system need a previous surgeon data acquisition and coding stage. This stage pursues the conversion of the acquired raw data into the set of outputs HMM symbols. In this way, the next subsections details first the data acquisition and coding stage, second the training process, and finally the real-time recognition system.

A. The Data Acquisition and Coding Stage

The position and orientation of the surgical tools are provided by a 3D localization system. These data must be converted into a set of symbols, which are used, by both the training module and the recognition system. These symbols belong to the set \( E \) of outputs of the HMM used for modeling the components of the surgical maneuver library.

At first, the number of basic actions that compose of all the maneuvers has been identified; each one has a certain label or observable output corresponding to them.

Fig. 3 shows the conversion process from the curves that define the movement of the instruments in to the labels \( e_1, e_2, ..., e_k \). Once the signals of the position and velocity have been acquired, in order to eliminate noise and vibrations, besides characterizing it, the Wavelet transform has been chosen [26]-[28], with the Daubechies functions of order 1. By using this methodology, the distinctive parameters of the acquired signal are obtained. This type of transformation can be considered to be a mathematical tool for the waveform representation (modeling and segmentation), and time-frequency analysis. The signals model is based on its representation from known base functions. Use is made of a form function or a mother function \( \psi(t) \) to represent the signal shape, and of a scale function \( \phi(t) \), that represents the details. Any signal \( f(t) \) can be represented as the following expression (1):

\[
f(t) = \sum_k \sum_j c_{j,k} \phi(t) + \sum_k \sum_j d_{j,k} \psi(t)
\]

(1)

where \( c_{j,k} \) and \( d_{j,k} \) are the elements that compose the vector of characteristics. With the use of the Wavelet transform, a vector of characteristics of the signals is generated, which is a necessary step before the classification phase.

By means of a fuzzy module with memory based on a Mamdani system, which includes one input and six outputs (see Fig. 3), the information of this vector is converted in to a unique label that represents it.

B. Maneuver Training

The object of the training phase is to obtain the parameters of the HMM that defines a certain maneuver, for which it is necessary to have an experimental database to generate the patterns that define these maneuvers. For this procedure the
Baum-Welch algorithm is used.

Fig. 4 details the training process. In the first place, a series of reference maneuvers are defined that are composed of a known determined sequence of basic actions \((S)\). These maneuvers are performed by different surgeons, and the sequence of labels generated by the module of data acquisition and codification are registered \((E)\). The training system establishes the transition state probabilities \(A\), as well as the emission probabilities of a certain output in a determined state, \(B\). The sequence of basic actions and its sequence of observable outputs are generated by a set of movements of the surgical tool. These are used as inputs to train the model with the Baum-Welch algorithm. The initial probability vector \(\pi\) is introduced in a heuristic form.

C. Maneuvers Recognition

The objective of the real time maneuver recognition module consists of identifying a certain movement from the labels that have been generated by the data acquisition and codification module \((E)\). Using this information \((E)\), the library of maneuver models generated in the training step \((A, B)\), and the protocol’s determinist model \((T)\), the recognition module reports on the current surgical maneuver (see Fig. 5).

The maneuver recognition is achieved in two phases (see Fig. 6). The first consists of selecting the most probable maneuver and applying the Viterbi algorithm to each of the \(\beta\) maneuvers from the library with the input labels sequence \(E\). The second stage, with the intention of reducing the uncertainty in the results, as more than just one maneuver has its own high probability, uses the information from the protocol’s surgery determinist model \((T)\) to select the current maneuver, or to report on an error in the protocol.

V. IMPLEMENTATION AND EXPERIMENTS

A set of experiment has been designed in order to validate the proposed the intervention model as a user interface. The experiments are based on surgical instrument movement recognition. These movements are contained in a parallel plane to the pelvitrainer ground, in exception of the introduction and the extraction displacements. In this way, the library of maneuvers is built for modeling the endoscope tip trajectory when follows a triangle, circle or square shaped figures. This simplification of the reality with two similar figures allows testing the behavior of the proposed methodology. In this way, we consider surgical protocols composed of a stream of the named figures, and a single figure is considered a surgical maneuver. The design of the described experiment tries to emulate in-vitro trials of a fictitious surgical intervention.

A. The Experimental System Setup

The Polaris Spectra is the selected device for tracking the position and orientation of several endoscopic tools while the in-vitro trials are performed. This device is a 3D position optical measurement system, which tracks a set of passive markers attached to surgical tools, and it has already been used in other medical applications [29]-[31]. The system provides an API for making the communication with other applications. The designed algorithms for real-time data acquisition are developed with MatLab. In a first stage, the data were acquired during a surgical task and they were analyzed for creating the model of a surgery maneuver.

B. Maneuver Model

Fig. 7 shows in the left-down corner a picture, which illustrates an in-vitro trial by using a pelvitrainer and a surgical tool. Moreover, the scheme in the right-up corner shows the Polaris workspace. In this way, all the surgical tool movement must inside this volume. The four basic actions taking into account for this experiment are represented in Fig. 7 as double-arrowed lines: horizontal straight line, a vertical straight line, a decreasing oblique line or growing oblique line. Moreover, a surgical tool insertion and extraction are added to the named basic actions. An
insertion action will be the initial state of a maneuver and the extraction is defined as the ending state. A surgical maneuver is a trajectory of the tip of the instrument and it is composed of a stream of basic actions, which has a full-connected network topology. The training stage generates a different matrix pair \((A, B)\) for each maneuver. These matrices represent the state transitions and the state observable output probability, and the recognition model will be uses this information for identifying the current maneuver.

Different surgeon executes 120 times each considered maneuver (circle, square or triangle) in order to build the surgery library. These generated sequences are used in the training process by means of Baum-Welch algorithm. This algorithm adjusts the transitions between states and it returns the best solution, which represent the given maneuver. Remark, that the surgeon can start the movement in anywhere and follows a counterclockwise displacement of the surgical tool tip, however the figure must keeps its original orientation with respect to the sensor reference frame. Fig. 8 shows the actual maneuver execution of a circle. The continuous line is the trajectory followed by the tool tip, the bold dark points are the fulcrum estimations, and the star shape points are the tool handle positions.

The built surgical library is tested by means of 300 random figures. Test results appear in the following table I.

<table>
<thead>
<tr>
<th>Trajectory</th>
<th>% Real Correct Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circle</td>
<td>98.35 %</td>
</tr>
<tr>
<td>Square</td>
<td>100.00 %</td>
</tr>
<tr>
<td>Triangle</td>
<td>95.87 %</td>
</tr>
</tbody>
</table>

Results described in table I shows a high hit recognition level, but the performed experiment it only consider isolated maneuvers. The next experiment will test sequence of figures as examples of surgical protocols. This new trial is oriented to protocol real time recognition, and the supervision of the execution of a given protocol. In this situation, the recognition system uses both the deterministic model and the stochastic HMM models. The table II presents the percentages of hits protocol recognition in the execution of 150 trials of the sequence of maneuvers shown in the first column. The second column shows the ideal correct answer ratio computed as the product of the ratio of the single figures shown at the table I, and the third one the actual result of the trials for the given surgical protocol. This second experiment shows that the recognition hits decrease a small quantity, but it keeps an acceptable level.

**TABLE II**

<table>
<thead>
<tr>
<th>Surgery Protocol</th>
<th>% Ideal Correct Answer</th>
<th>% Real Correct Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>□□□□</td>
<td>90.39 %</td>
<td>85.00 %</td>
</tr>
<tr>
<td>□□□□□□</td>
<td>91.20 %</td>
<td>82.50 %</td>
</tr>
<tr>
<td>□□□△</td>
<td>94.29 %</td>
<td>81.00 %</td>
</tr>
</tbody>
</table>

VI. FUTURE WORKS

The work presented in this paper is a component of a control architecture for a two arms autonomous robotic system. This is a current running research project whose goal is to propose a new solution for the solo-surgery problem. One of the two mentioned arms is responsible for guiding the laparoscopic camera and the other is devoted to move an additional instrument. The functional scheme of this robotic assistant is shown in Fig. 9.

The mentioned diagram details the different functional modules included in the control architecture. The two-manipulator arms appear at the top of the scheme, and they are designed for interacting with the patient body. In order to register the interaction forces applied to the patient’s body, they are equipped with force sensor attached to the wrist. These forces, as well as the arms joint position and the laparoscopic image are obtained by the sensor subsystem, which interprets and combines this information and provides it to the spherical control, the trajectory and interaction control and the planner. These modules implement different levels of control strategies. The planer module needs a movement reference, which is supplied by the human-machine interface, in order to calculate the goal position inside of the abdomen. The intervention model module
contains the protocols and the modeled surgical maneuvers described in this paper. The planner uses this information for recognition purposes and for generating the next robot movement. With this information, the Cartesian trajectory control module plans the on-line trajectory that the tip of the tools must follow in order to achieve the desired reference. However, these Cartesian trajectories need an inner control loop, which adjusts the movement of the manipulator arms to the holonomic constraints imposed by the fulcrum point. The spherical control module fulfills this action, by using the joint positions and contact forces feedback. The spherical control module allows the Cartesian control strategy to be independent of the arms kinematics. Finally, the supervisor module ensures the correct operation of the system.

VII. CONCLUSIONS

This paper has described a methodology for designing a user interface based on a maneuver recognition system which models the surgeon behavior. This procedure is composed of three stages. The first one is devoted to convert and coding the data acquired by a 3D tracker. Then, this information is used for two purposes: i) off-line training for constructing a maneuver library and ii) on-line maneuver identification. Finally, this last action has been extended in order to recognize a given surgical protocol and supervise it. In order to test the proposed system, a set of geometrical movements has been selected as a representation of surgical maneuvers. In a real situation, the intervention model has to include instrument-organs interactions forces and consider information about other different type of surgical tools. Finally, this paper shows the feasibility of the proposed methodology for surgeon instrument movement recognition, and in next works will be expanded in order to include other useful information about the surgical procedure (i.e. 3D movements of the surgical tools and contact forces).

REFERENCES

