Project SPARS
Sequential Pattern Analysis in Robotic Surgery: Understanding Surgery

Project Participants:

Pierre Jannin, MediCIS, LTSI, Inserm, Université de Rennes (Project leader) (ORCID)

Krystel Nyangoh-Timoh, CHU Rennes: Department of Obstetrics/Gynecology, MediCIS LTSI Inserm, Université de Rennes
Arnaud Huaulme, MediCIS LTSI Inserm, Université de Rennes
Pierre François Marteau, Professor University of South Brittany, EXPRESSION IRISA
Philippe Leray, Professor, University of Nantes, DUKe, LS2N
Christine Sinoquet Associate Professor University of Nantes, DUKe, LS2N

Short abstract: The objective of the project was to study data analysis methods to better understand complex technical human activities, such as surgery. Surgery is a complex activity, that depends on many factors, including the patient and surgeon characteristics. Such complexity and variability explain why there is almost no detailed study of the surgical practice yet. Until now, the surgical procedure performed in the operating room is considered as a whole, as a black-box and is technically described with few words. Analysis usually consisted in comparing impact of different surgical approaches or of different pre-operative clinical patient's parameters on post-operative outcomes. In the SPARS project, we focused on studying innovative analysis approaches based on two dimensions of surgery: the kinematics of the surgical instruments and the sequence of surgical activities described as lists of successive surgical phases and steps. Our analysis strategies relied on a combination of both data and model-driven approaches studied within two workpackages (WP1 & WP2). The proposed methods were applied and validated on diverse simulated datasets including both kinematics and procedural data, including simulated non-realistic datasets, and on-line available reference data sets. Different results were obtained. We showed how new surgical kinematics data can be analysed with new temporal series analysis-based methods for classification based on surgical expertise. We also proposed methods for computing models representing variability and invariance within a population of performed tasks, opening the way for identification of parameters explaining commonality and differences between surgical practices.

1 Life of the project

The project ran through regular monthly meetings (more than 40 during the project) and three full days: two in Rennes and one in Nantes. Several travels between recruited peoples occurred between Rennes and Nantes.

The Spars project benefited from an audio podcast available at https://www.canal-u.tv/chaines/inria/labex-cominlabs/le-projet-spars-analyser-pour-comprendre-la-chirurgie

2 Detailed results per WP

2.1 WP1: Modeling procedural surgical knowledge (Resp. P. Leray)

Despite the widespread use of robot-assisted total laparoscopic hysterectomy (rTLH), substantial variability persists in how the procedure is performed, contributing to inconsistent surgical outcomes and challenges in training. Existing curricula emphasize technical skills but lack structured models that capture the underlying procedural logic and natural variation in technique. Procedural knowledge models represent the variability within the sequences of surgical procedures, described via a hierarchy of surgical phases and steps. Such models are essential to improve reproducibility, strengthen surgical education, and facilitate integration with emerging AI-driven systems. Work package WP1 aimed to studied different approaches, both top-down (T1.1) and bottom-up (T1.2), to represent procedural surgical knowledge models.

T1.1: Top-down design of semantic models of surgical procedures (Resp. P. Jannin)

Using the consensual definitions of the rTLH phases and steps, we first developed an ontology specific for the hysterectomy procedural description: HystOntoSPM. It is based on the upper ontology OntoSPM, an ontology that provides a generic vocabulary of the domain of surgical procedures, focusing on entities such as surgical procedure steps, surgical actions, actors, instruments and roles. It also uses extracts from LapOntoSPM, an ontology developed to standardize knowledge of laparoscopic surgeries. The current version of HystOntoSPM also contains extracts from various well-known ontologies, such Foundational Model of Anatomy (FMA).

Then, we conducted two studies to express surgical procedural knowledge model following a top-down strategy.

The HystOntoSPM ontology was used to build a first version of a so-called generic surgical process model (gSPM) representing the theoretical variability of rTLH. It was based on interviews of a single gynecological surgeon. We validated this model by developing a library of metrics computing the distances between the manually expressed gSPM and the descriptions in terms of phases and steps of 54 clinical surgeries. We demonstrated how such gSPM partially and not exhaustively represents the real clinical practice, demonstrating the need for different gSPMs. This study was submitted to an international journal (in review [1]).

Then, we developed the first consensus-based procedural model for standard rTLH using a multi-round Delphi process involving 35 international expert gynecologic surgeons. Across five rounds, panelists systematically reviewed, refined, and reached consensus on the procedure's phases, steps, and allowable variations. The final model included 34 steps, each defined by consensus, and more than 105,000 procedural variations. The results have been published in an international journal (see [2]).

Recruited peoples:

- Soline Galuret (her), Research Engineer

Publications:

- [1] Galuret, S., Nyangoh Timoh, K., Galuret, S., Hébert, T., Azaïs, H., Barahona, M., Becker, S., ... & Jannin, P. (2025). How much knowledge elicitation represents procedural surgical practice?.
- [2] Nyangoh Timoh, K., Galuret, S., Hébert, T., Azaïs, H., Barahona, M., Becker, S., ... & Jannin, P. (2025). International expert consensus-driven surgical process model for robot-assisted hysterectomy: Delphi study results. *Surgical Endoscopy*, 1-12.

Additional results

- The HystOntoSPM ontology
- The open-data rTLH gSPM available in [2]

T1.2: Hierarchical procedural models of surgery (Resp. P. Leray)

In T1.2, we studied different modeling approaches for automatically compute gSPM from iSPM. It was validated on different datasets, including a data set of cataract surgery.

The PhD work included:

- A systematic review of the state of the art concerning the consideration of variability in SPMs
- A mathematical formalisation of a gSPM and its transformation into models classically used in process modelling
  - taking variability into account with coloured Petri nets
  - taking variability and uncertainty into account using SVGEMs (graphical event models with state variables)
- A proposal of a solution for learning gSPMs from iSPMs, by learning an SVGEM
- The definition of a more general class of SVGEMs taking into account richer temporal dependencies, extension of sampling and learning algorithms for these models.

Recruited peoples:

- Moustafa Said Hawchar (him), PhD student, PhD defense on 19/12/2025

Publications:

- [3] Poster Presentation Probabilistic Graphical Models in Healthcare workshop, 10 September 2024, Nijmegen, Netherlands.
- Planned publications:
  - Literature review at the international journal of Computer Assisted Radiology and Surgery (IJCARS).
  - The new SVGEM model and its sampling algorithm to be submitted at the 35th International Joint Conference on Artificial Intelligence (IJCAI), ou International Conference on Probabilistic Graphical Models 2026 (PGM).
  - Its extension at the International Journal of Approximate Reasoning (IJAR).
  - Formalization and computation of gSPMs to be submitted at the 39th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems IEA/AIE

Additional results

- Software for gSPM and its format transformation in https://gitlab.univ-nantes.fr/leray-p-1/gspm-library
- Extension of the PILGRIM Evential software library with SVGEM

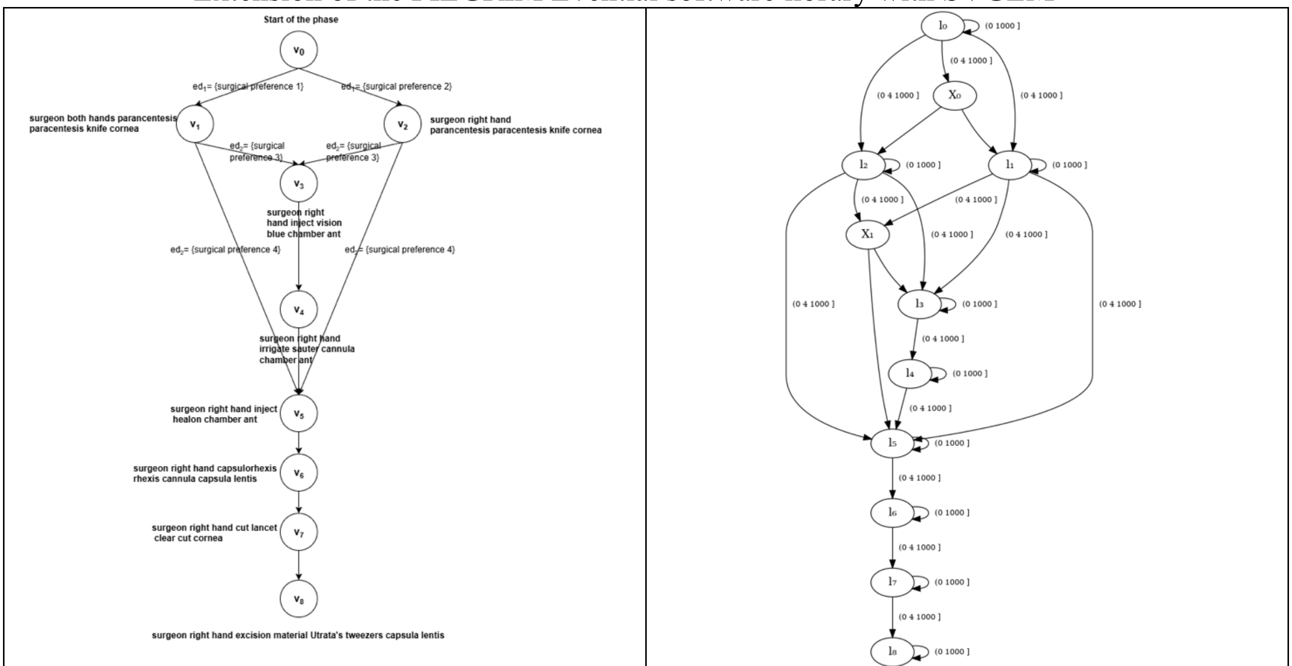

Figure 1: gSPM of cataract surgery and corresponding SVGEM

2.2 *WP2: Trajectory data analysis (Resp. C. Sinoquet)*

WP2 focused on the low-level analysis of time series (*i.e.*, surgical instrument trajectories) through two tasks: T2.1: Clustering and classification of trajectories (Resp. C. Sinoquet) and T2.2: Measure of the distance and variance of a trajectory with respect to an "average trajectory" (Resp. P. F. Marteau).

T2.1: Clustering and classification of trajectories

Clustering and averaging:

Our work in T2.2 focused on averaging time series after clustering. Accordingly, in T2.1 we addressed the *Partitioning of surgical trajectories* using k-means and k-medoids methods, which both rely on time-series averaging. Such averaging requires a dissimilarity measure (e.g., DBA), and the literature offers many options, including kernel metrics. The large number of possible combinations , (k-means / k-medoids / kernel k-means) × dissimilarity, makes it difficult to select an appropriate clustering method.

Our survey of the literature showed that no existing work compares multivariate dissimilarity measures for clustering tasks (most studies being univariate and focused on classification). We therefore produced the first multivariate clustering benchmark restricted to methods available in the AEON and tslearn libraries. Additionally, comparative studies rarely include kernelized dissimilarities; we incorporate, for the first time in such a review, the KDTW measure, developed by P.-F. Marteau, member of the SPARS project, and its associated averaging method, TEKA.

Using three clustering strategies (k-means, k-medoids, kernel k-means) and seven dissimilarities (DTW, MSM, Shape-DTW, SBD, GAK, Soft-DTW, KDTW), we compared the performance of 11 clustering approaches.

In a second study, we focused on four averaging methods (DBA and two variants, TEKA). Using known partitions of multivariate time series, we computed consensus representatives for groups of time series. We then derived statistics on the deviation of each time series from the consensus of its group. The metric used to quantify this deviation was a time-series dissimilarity measure.

These two studies have been presented in a state-of-the-art review article that includes method comparisons. The article will be submitted to an international journal.

Classification

A second focus of WP2 was time series classification. We used the JIGSAWS robotic surgical trajectory reference dataset and improved over the state of the art using the ROCKET classification method. Data augmentation further enhanced performance for one of the two augmentation strategies tested.

JIGSAWS allows evaluation on 15 surgical gestures, 3 expertise levels, and six degrees of expertise (regression). Compared to prior work on this benchmark, we show improvements.

This study is described in an article soon to be submitted to the AIME international journal.

T2.2: measure of the variance of a trajectory with respect to an "average trajectory"

In prior work, we designed a kernelized dissimilarity measure for time series, KDTW, along with its associated averaging algorithm, the iterative Time Elastic Kernel Averaging (TEKA). TEKA introduces a probabilistic framework for time series averaging, where pairwise alignments are modeled as outputs of a stochastic automaton. Accordingly, TEKA's alignment process is based on the forward–backward algorithm used in Hidden Markov Models. Forward and backward probabilities are then used to compute the expected values of the observations of a time series s aligned with a given sample ct of the consensus c at timestamp t. Symmetrically, it is possible to compute the expected timestamps of the observations of s aligned with ct. These expectation computations form the core of the TEKA algorithm.

We leveraged previous work on separating shape and phase, based on the earlier paradigm, to quantify deviations from the consensus, using a shape dissimilarity score as well as a temporal dissimilarity score. In plots showing the shape dissimilarity score versus the temporal dissimilarity score for the JIGSAWS benchmark, the three levels of expertise (novice, intermediate, expert) are clearly highlighted. In other plots showing the mathematical expectation of the alignment times of observations from series s of a given class (e.g., intermediate expertise) with the sample c(t) of the consensus from the expert class, we can identify slowdowns or accelerations in the trajectories. The results are being consolidated for an upcoming manuscript submission.

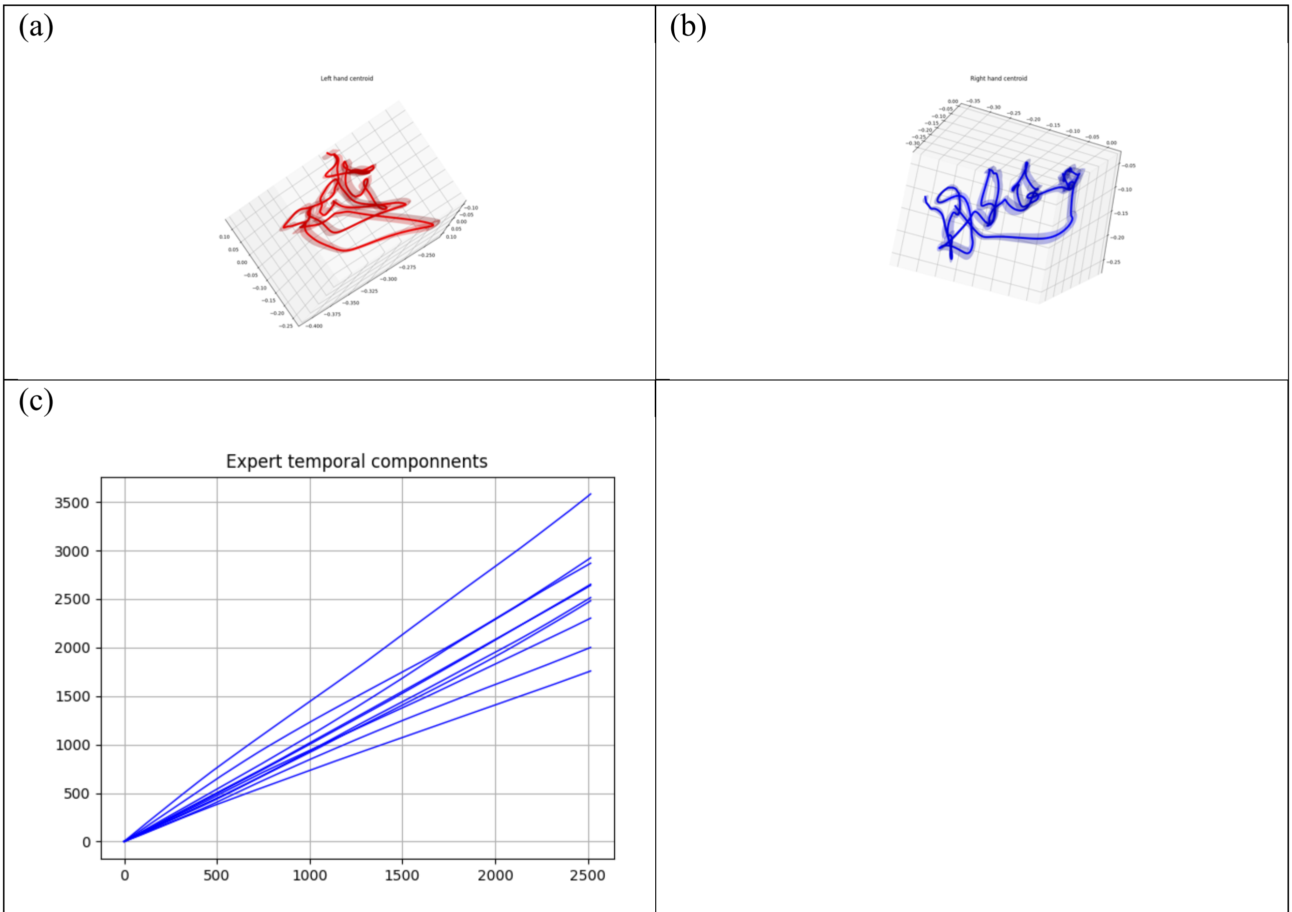

Figure 2: Phase (c) and Shape (a,b) separation approach: shape model for left (a) and right (b) hands of experts performing a suturing task

Recruited peoples:

- Hafssa Boujida (her), research engineer

Publications:

- Planned publications: Two scientific papers have been completed and will be submitted, and results for a third have been collected.

Additional results

- By the end of WP2, several time-series processing pipelines are ready for application to clinical data. These pipelines were transferred from the research engineer in Nantes to the team in Rennes. The TEKA-based method has provided a novel practical extension to existing work on this original averaging approach.

2.3 *WP3: Validation and Evaluation (Resp. P. Jannin)*

In Work package WP3, we developed a method for displaying gSPM with the objective of having an educational tool.

Generic Surgical Process Models (gSPM) models the sequence of a surgical procedure in a standardized manner. These models are often depicted as graphs, where nodes represent individual steps and edges define their order. In complex surgeries such as the rTLH (and its 105,000 variations), the resulting graphs can grow large and densely interconnected, making them difficult to read and interpret. A dedicated visualization aid then becomes essential to simplify navigation, highlight key information, and support a clear understanding of the procedure. We designed SPMapp Browser, an interactive graphical navigation tool for the step-by-step exploration of generic gSPM, guiding the user through the procedure node by node. By considering the user's navigation history, the interface displays only the next reachable steps, thereby streamlining the visualization and enhancing comprehension of the graph's structure. To reinforce its educational value, SPMapp Browser offers additional resources that enrich the procedural exploration. Selecting a step reveals complementary content: a definition, an illustration, and/or a video demonstrating of the surgical action behind the step. We also integrated an optional view of the complete gSPM, continuously updated to show the path followed by the user. This feature helps learners grasp the overall logic of the procedure and, if desired, interpret the graph independently outside the application. SPMapp Browser was developed as an open-source, browser-native, and multi-platform solution, requiring no installation, ensuring an optimal accessibility for all users.

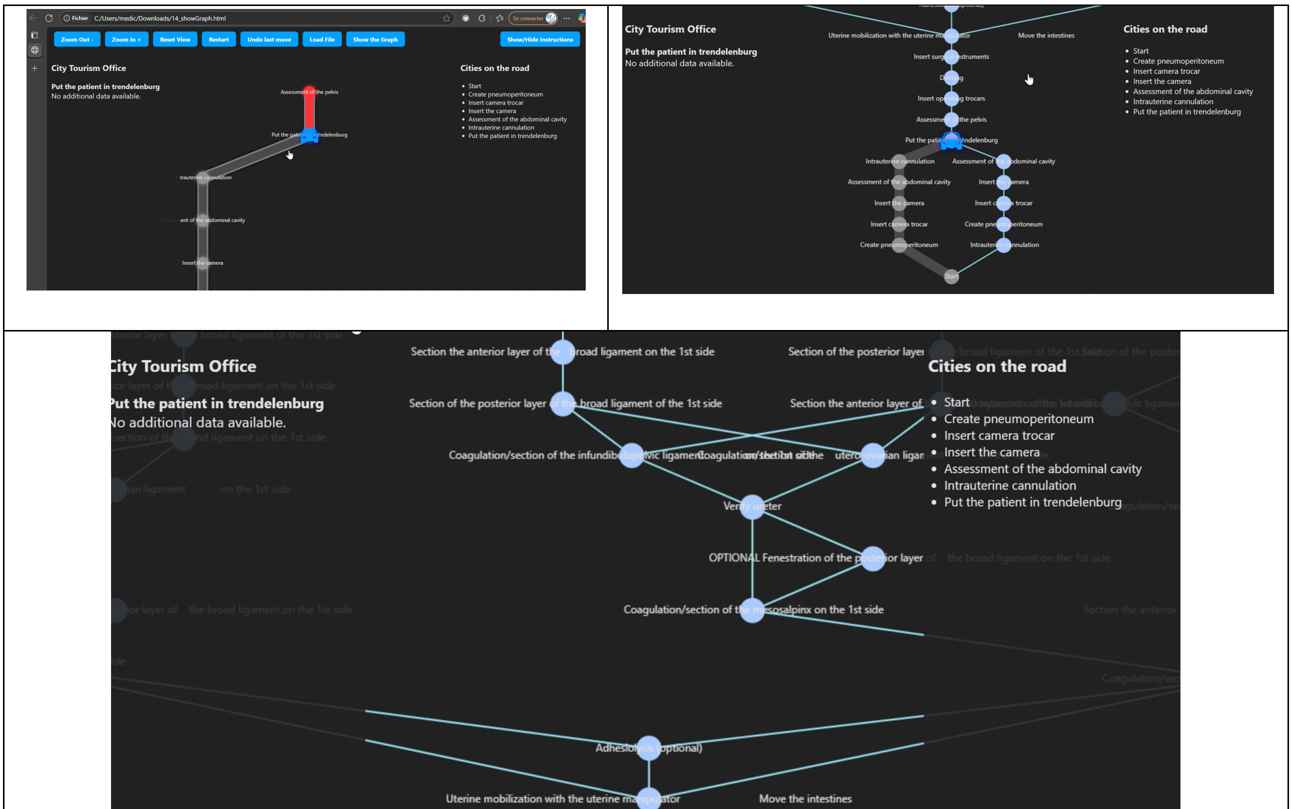

Figure 3: SPMapp Browser

3 Conclusion & Perspectives

The project was able to address the main methodological challenges: studying complementarity of top-down and bottom-up approaches for modeling procedural knowledge, leveraging of state of the art of temporal series analysis approaches for kinematic data. Several results of the project open the page for new research questions and further funded projects.

WP1 works on ontology and top-down design of semantic models of surgical procedures will be adapted in the context of endometriosis within a national PEPR women health project, as well as within a 4-year ANR PRCI project. This project started in 2024 in collaboration with the Chinese University of Hong-Kong focus on the analysis of rTLH procedural and videos data. The project take benefit of the results from WP1, including the ontology and the top-down generated gSPM in order to provide a priori knowledge and constraints to facilitate automatic recognition of surgical phases and steps from surgical videos.

WP1 works on hierarchical procedural models of surgery was the core aspect of a letter of intent submitted to 2025 ANR call on procedural modeling with LS2N and LTSI. The letter was rejected. We plan to resubmit the project.

WP2 works on kinematic data modeling and analysis will be pursued by a PhD student (Ilyes Akli) starting in October 2025. The objectives of the PhD are to adapt the developed methods to be able to deal with the complexity of clinical data.

WP3 developed tools will serve as a basis for an extension towards kinematic data visualization which will be used in many following projects.