Human-Robot Task Handoff: A Probabilistic Modeling Approach Explored through Cooperative Drawing

ABSTRACT

Recent research in human-robot interaction explores the potential for human-machine collaboration in surgical procedures, dividing tasks into manual and automatable subtasks. This paper investigates the task of handoff detection, crucial for the success of robot-assisted surgery, focusing on the creation of a synthetic dataset that can be used for training and benchmarking models for this task. We present a dataset of parabolas, simulating cooperative drawing between a human and a robot, with variations in drawing rates and added noise. The study demonstrates the applicability of HMMs in determining handoff points, laying the groundwork for future research in human-machine collaborative surgery. The dataset, along with the provided code and raw data, are provided as a resource for future research. Finally, we discuss the limitations of the dataset and suggest directions for future research, emphasizing the need for higher-dimensional and real-world datasets.

Introduction

Since its introduction in 1999, the da Vinci surgical system has seen widespread adoption in various surgical procedures, with over 60,000 surgeons worldwide trained on the system and over 10 million surgical procedures completed using it [10]. While historically surgeries were predominantly performed manually by operators, recent research indicates a growing potential for human-machine collaboration in surgical procedures.

In the context of human-machine collaborative surgery, tasks are divided into manual and automatable subtasks. Automation of subtasks is achieved by detecting and learning the operator's surgical performance, allowing for the automation of most subtasks. Operators can intervene during automatic execution to make corrections if necessary. The transition between manual and automated tasks should occur seamlessly and automatically in teleoperation. However, this can be challenging due to the variety of noise sources present in performing and observing surgical maneuvers.

Towards these ideal transitions, in this research we present a dataset of synthetic human-robot handoff events, providing a means for practitioners to design and evaluate models that can be used to automatically identify moments of takeover. By implementing and testing our dataset, these methods may be extended to the conditions of real-world human-robot interaction, especially toward efficient and accurate human-robot surgical cooperation.

Related Research

The handoff task is very important to the success of robot-assisted surgery and has been the subject of many studies. In [1], Hidden Markov Models are used for the novel human-machine collaborative approach for tele-surgery. This is done with the recognition of task completion and temporal curve averaging for learning the executed motions. This is demonstrated using the da Vinci tele-surgical robot. The setup they used is to consider a task with a set of subtasks that are executed by alternating between a manual and an automated subtask. Recognition is done to determine the completion of each manual task to seamlessly transition into automatic control, and then return control to the surgeon. They used two different tasks to demonstrate this: a pin-task where a single instrument displaces three pins to three cardinal locations, and a sut-task where two instruments perform a running suture with three needle insertions.

Along the lines of workflow recognition, specifically for the operating room, Padoy et al. [2] have applied HMMs. Their innovative contribution involves the development of workflow HMMs, a variant of HMMs enhanced with phase probability variables that aptly capture the intricacies of the entire workflow progression. In particular, the primary objective was discerning and monitoring the eleven distinct workflow phases within a surgical operation. Each phase was recognized by WHMMs through the detection and tracking of both human individuals and pertinent objects obtained from the multi-camera system. Later research works have focused on specific procedures for human-machine collaborative surgery. For example, Varier et al. [6] explore similar concepts with collaborative suturing as they use discrete Reinforcement Learning techniques to automate the needle hand-off task. Comparably, Rafii-Tari et al. [7] use HMMs for endovascular intervention in human-robot cooperative systems.

With the modeling of the primitive gestures of pushing, pulling, and twisting with the HMM, they encode a robotic catheterization system based on learning from demonstration to reduce the workload of the operator. Both of these collaborations could be extended to further tasks with the introduction of classification.Hu et al. [8] extend this as they propose a collaborative bimanual peg transfer task. This model enables autonomous and manual control to alternate correspondingly. They hope to explore other tasks such as suturing. Finally, Kaplan et al. [9] describe the benefits of the collaboration with improved task performance, reduced task time and tissue trauma, and fewer times the tissue is palpated. They discuss that collaborative surgery outperforms pure human or pure machine surgery.

These papers have shown the necessity of surgical maneuver classification within human-machine collaborative surgery as it expands the range of possible collaborations.

Methods

Dataset Creation

To test machine learning models on these collaborative movements, we simulate a surgical maneuver with a parabola. We defined a task for a human and robot to complete: "drawing" a parabola cooperatively. The human does the task with some natural noise in the movement of their muscles, so as they draw the "going up" portion of the parabola, there is some expected jitter on the page. The goal is for an algorithm to identify when the human has reached the "peak" of the parabola so that the robot can then take over and finish the drawing.

To simulate the varied speeds of surgeons performing a maneuver, each parabola had a rate randomly selected from between 1 to 10 units per second. Each point represents the passage of approximately 1 second. Therefore, fewer points represent a faster drawing while more points represent a slower drawing. For example, in Figure 1a, it is observable that the rate of the human is slower as there is more passage of time as it goes up, compared to Figure 1b which has a fast rate, leaving fewer points being observed..

Additionally, to simulate the unknown variables of performing a maneuver such as the surgeon's hands shaking, being disturbed by an outside force, or inexperience, we added noise to these parabolas in both the x and y directions. The noise added was determined by adding or subtracting a random value from 0 to 100 for the y value and a random value of 0 to 5 for the x value.

Figure 2 demonstrates the average drawing of the parabola with the average rate and noise. The average rate was determined by compiling all of the rates to get an average of 4.89 and an average time of 20.4499s. Figure 3 demonstrates the averages of all the parabolas together. Figure 3 takes the averages of the parabolas with noise, disregarding the different rates, and compiling them into one graph of a parabola traversed at a rate of 1.

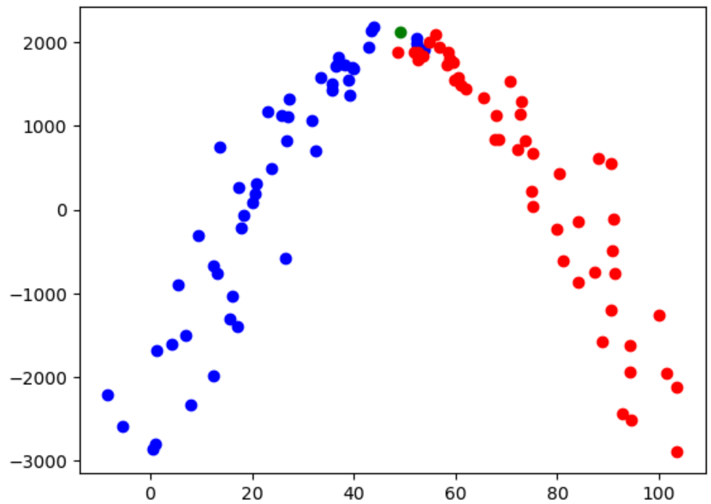

Figure 1a. Sample Parabola with a rate of 1. The blue points represent the human drawing the increasing section. The green point at the peak represents the point of handoff between the human and the robot. The red points represent the robot finishing the drawing by going down.

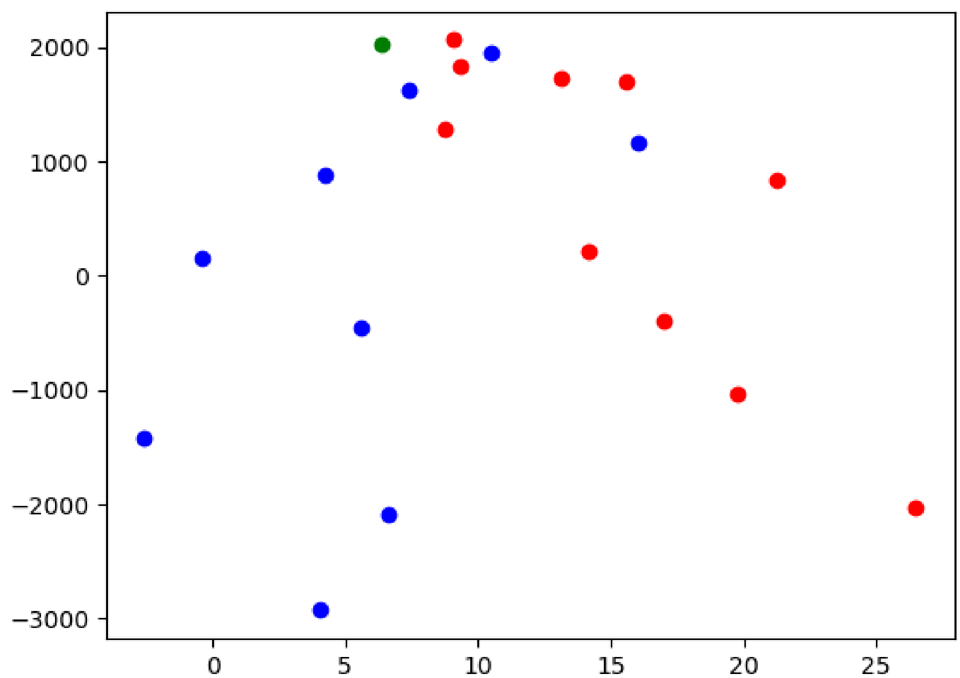

Figure 1b. Sample Parabola with a rate of 5. The blue points represent the human drawing the increasing section. The green point at the peak represents the point of handoff between the human and the robot. The red points represent the robot finishing the drawing by going down.

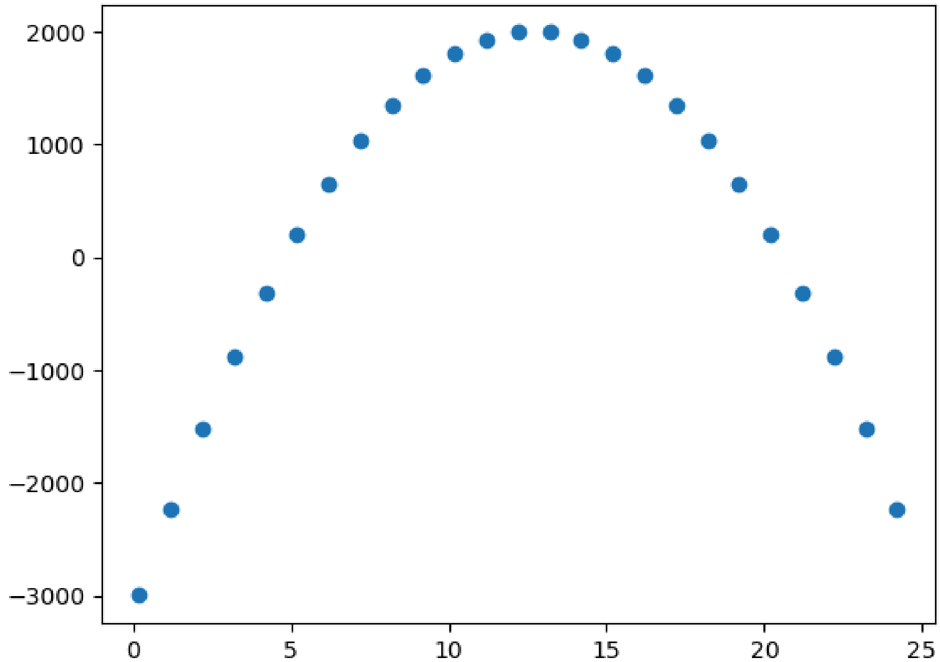

Figure 2. Expectation Parabola. This parabola is the ideal expectation of one drawn at the average rate with the average noise of all the parabolas.

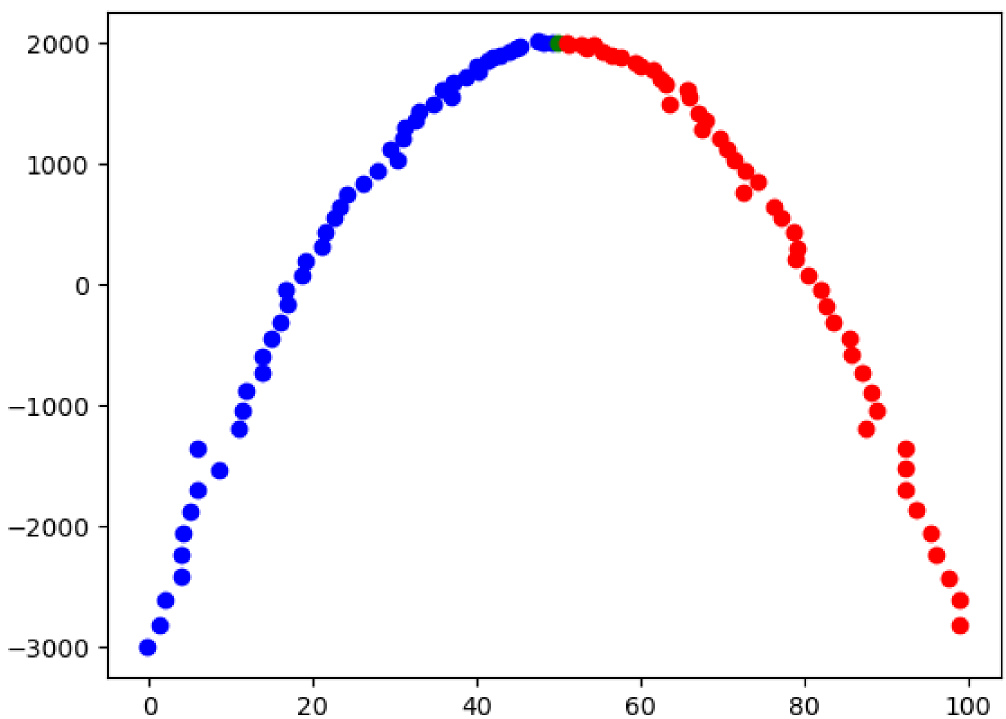

Figure 3. Average of the Parabolas. This parabola takes the average of all the y values and the x values without the differing rates to represent the overall trajectory of the parabolas. This figure follows the same color legend as Figure 1.

Results

We created 1000 parabolas of varying rates from 1 through 10. Statistics on the thousand parabolas are described in Table 1. To compare model results on a benchmark set of data, we also provide raw data for this set of 1000 parabolas at xnoisy.p, ynoisy.p We make our generating script available at https://github.com/lucasc073/GMM1/blob/main/ParabolasWithNoise.ipynb

On average, out of 1000 formulated graphs, we observe 20 dots in each parabola. Towards interpreting the parabolas, it is recommended to assume the observing AI system makes observations at sample rate T, and the average person draws their portion of the parabola in t = 10 seconds (so the entire parabola is drawn in 20 seconds). In this average case, for example, since the interpreting system "sees" M (=20) dots in the parabola, that means M samples were taken by the perception system and it took (M/T) seconds for the person to draw the parabola. Each sample provides an opportunity for the observing system to guess the state of the drawing progress.

Table 1. Compilation of all the average values from the dataset.

Average			
Average Peak	**Average Rate**	**Averages Graph Peak**	**Average Time**
(11.856, 418.320)	4.991 units/second	(49.983, 2000.950)	20.036s

Discussion

Relationship to HMM Approaches

Hidden Markov Modeling was first proposed by Baum L.E [5] and is used by augmenting a Markov chain. A Markov chain is a model that describes state transitions and stochastic processes that describe the statistical correspondence of sequences of random variables, and states, each of which can take on values from some set. A Markov chain makes a very strong assumption to predict the future in the sequence based on the current state. The states before the current state have no impact on the future except via the current state.

Hidden Markov Models have many uses as shown in [3], where the authors researched different models and methods that would be useful for early warnings such as for an infectious disease. The researchers observed how Hidden Markov Models can automatically and flexibly adjust the trends, seasonal, covariant, and distributional elements. Therefore, they have been used in many studies on time series surveillance data; for example, Le Strat and Carrat used a univariate HMM to handle influenza-like time series data in France. Additionally, Madigan indicated that HMM needed to include spatial information based on existing states [3].

Here, we describe the preliminary definitions that one could use to approach the handoff determination task using HMMs, towards future research utilizing this dataset. From [1], "formally, the resulting HMM is a sextuplet $\lambda=(N,A,O,B,\pi,\phi)$ where N is the number of states $\{x_i: 1 \leq i \leq N\}$ in the model, A the transition probability matrix between the states, modeling the topology, and O the space of observations. B is the observation model, indicating for any observation $o \in O$ and state x the probability $B_x(o)=P(o|x)$ that o can be observed by x. π is a probability distribution over the initial states. ϕ denotes the phase probability variables, indicating the probability $\phi_x(p)=P(p|x)$ of each state x to belong to a phase $p \in L$." We provide further clarity here to each of these components:

- N, the number is states, is determined as follows: "The number of states is defined based on the average length of the sub-sequences." Since we see on average 20 dots in each parabola, the average length of the sub-sequences is 20 for the total parabola, and therefore 10 for the human portion.

- The transition probabilities are initialized such that the expected duration corresponds to the duration of the human phase. This expected duration is 25 seconds to reach the peak of the parabola. M, the number of distinct observation symbols per state, in this case, is 2: whether the drawer is human or AI.

- We can use a value of 0.39 for the initial transition probabilities (A), found by simulating the Hidden Markov Model process and determining the probability that would allow the AI to traverse all 20 states in 50 seconds when transitions are sampled at a rate of 1 Hz.

- We can model $Bx(o) = P(o|x)$ using a mixture of Gaussians, where the Gaussians for each state are initialized using k-means, with k = number of states (that is, each state is approximated as a Gaussian of possible locations).

- The probability distribution over initial states (π) is trivial in this case since we assume the human will begin their portion at state 1. The phase probability variables are p(transition) = 0.39 and p(stationary) = 0.61 initially, as this would provide the expected 25-second traversal to draw the first half of the parabola.

- In this sample case, ϕ represents that s = 1 through 10 all belong to the human phase, and s = 11 through 20 all belong to the robot phase.

With these HMM parameters defined, one could train the HMM using Baum–Welch or another algorithm, then using an observed sequence, predict the point at which the handoff should occur.

Conclusion

In this article, we introduced the Human-Machine Drawing Tradeoffs, a set of noisy parabolas to test hand-off estimation for a simulated human-machine collaborative procedure on drawing. The dataset has a human's "drawing" up the first half of the curve and a designated handoff point that machine learning models, such as the HMM, would need to detect to take over to finish the curve.

With the noise of the dataset, the handoff timings vary sometimes significantly, especially when drawn at a faster rate, making detection challenging. This dataset can provide training and benchmark testing for models for handoff estimation, with a variety of different future applications, such as human-robot surgical operations, and other human-robot collaborations.

Specifically towards the training of HMMs, these handoffs are complicated to coordinate as the number of states any single instance transitions through differs with the different rates. Determining that the human is approaching the peak requires the recognition of the states prior along with the rate. Overall, the dataset can be helpful for testing methods of handoff state estimation (whether the robot reliably detects a handoff), leading to smoother human-robot collaboration on cooperative tasks.

Limitations

While this dataset provides a low-dimensional and noisy environment to explore and test a variety of algorithms for identifying key points in the human-robot simulated collaboration, real-world data in this domain is subject to other sources of noise (such as measurement noise), and exists in a higher-dimensional space which may lead to further confounding of positions. Accordingly, a 3D version of this dataset, and one modeled from real surgical tasks, would be a helpful direction for future research.

Acknowledgments

The author would like to acknowledge their mentor, Ross Greer.

References

[1] Padoy, Nicolas, & Hager, G. D. (2011). Human-machine collaborative surgery using learned models. *2011 IEEE International Conference on Robotics and Automation*. https://doi.org/10.1109/icra.2011.5980250

[2] Padoy, N., Mateus, D., Weinland, D., Berger, M.-O., & Navab, N. (2009). Workflow monitoring based on 3D motion features. *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*. https://doi.org/10.1109/iccvw.2009.5457648

[3] Lan, Y., Zhou, D., Zhang, H., & Lai, S. (2017). Development of early warning models. *Early Warning for Infectious Disease Outbreak*, 35–74. https://doi.org/10.1016/b978-0-12-812343-0.00003-5

[4] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," in Proceedings of the IEEE, vol. 77, no. 2, pp. 257-286, Feb. 1989, doi: 10.1109/5.18626 .

[5] Leonard E. Baum. Ted Petrie. "Statistical Inference for Probabilistic Functions of Finite State Markov Chains." Ann. Math. Statist. 37 (6) 1554 - 1563, December, 1966. https://doi.org/10.1214/aoms/1177699147

[6] Varier, V. M., Rajamani, D. K., Goldfarb, N., Tavakkolmoghaddam, F., Munawar, A., & Fischer, G. S. (2020, August). Collaborative suturing: A reinforcement learning approach to automate hand-off task in suturing for surgical robots. In 2020 29th IEEE international conference on robot and human interactive communication (RO-MAN) (pp. 1380-1386). IEEE.

[7] Rafii-Tari, H., Liu, J., Payne, C. J., Bicknell, C., & Yang, G. Z. (2014). Hierarchical HMM based learning of navigation primitives for cooperative robotic endovascular catheterization. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014: 17th International Conference, Boston, MA, USA, September 14-18, 2014, Proceedings, Part I 17 (pp. 496-503). Springer International Publishing.

[8] Hu, Z. J., Wang, Z., Huang, Y., Sena, A., y Baena, F. R., & Burdet, E. (2023). Towards Human-Robot Collaborative Surgery: Trajectory and Strategy Learning in Bimanual Peg Transfer. IEEE Robotics and Automation Letters.

[9] Kaplan, K. E., Nichols, K. A., & Okamura, A. M. (2016, May). Toward human-robot collaboration in surgery: Performance assessment of human and robotic agents in an inclusion segmentation task. In 2016 IEEE International Conference on Robotics and Automation (ICRA) (pp. 723-729). IEEE.

[10] Intuitive. Robotic-assisted surgery with da Vinci systems. https://www.intuitive.com/en-us/patients/da-vinci-robotic-surgery/