RoboSurgery

Matteo Nunziante*

January 13, 2025

Abstract

Keywords: keyword1, keyword2, keyword3

*matteo.nunziante@studenti.units.it

Contents

1 Abstract

2 Introduction

Lung cancer remains a leading cause of cancer-related deaths worldwide, with surgical resection offering the best chance for cure in early-stage cases. However, the intraoperative environment presents significant challenges, including dynamic anatomical changes, limited visibility, and high stakes for patient safety. Assistive technologies capable of real-time decision-making and adaptability could greatly enhance surgical outcomes. In this context, this thesis explores the potential of Partially Observable Markov Decision Processes (POMDPs) and Reinforcement Learning (RL) as frameworks for developing intelligent, adaptive, and safe intraoperative tools.

3 Problem Statement

Over the last decade, 3D models have entered oncologic surgery as a means to achieve better outcomes in renal and hepatic surgery [1, 2]. Nevertheless, the integration of 3D models into the operative field has been lacking due to three main reasons. Firstly, proper model alignment with the intraoperative anatomy has proven to be a major challenge due to shifting of organs during surgery and different patient positioning in surgery ver- sus during computed tomography [2, 3]. Secondly, automated organ registra-tion in a continuously moving surgical video has been another major challenge for many years [4]. Thirdly, 3D model overlay obscures the surgical field, including sharp surgical instruments which are manipulated, hence creating a possible hazardous situation rather than facilitating surgery. The latter occlusion problem has been a longstanding study topic [5] which, if solved, would further advance various surgical domains and applications [6]. Already in 2004, Fischer et al. [7]

4 Related Work

- Instrument occlusion
- DL recognition of anatomical structures
- Elastic Regristration

While robotic systems have revolutionized minimally invasive surgery, they lack the ability to adapt to unexpected intraoperative changes autonomously. POMDPs, with their capacity to model uncertainty, and RL, with its learning-based adaptability, present a promising synergy for addressing these challenges. However, their application to real-time intraoperative decision-making remains unexplored.

5 POMDP

Formal description of Partially Observable Markov Decision Process as in [3]

Formally, a POMDP is a 7-tuple $(S, A, T, R, \Omega, O, \gamma)$, where:
S is a set of states,
A is a set of actions,
T is a set of conditional transition probabilities between states,

$R : S \times A \rightarrow \mathbb{R}$ is the reward function.
Ω is a set of observations,
O is a set of conditional observation probabilities,
$\gamma \in [0, 1)$ is the discount factor.

There are at least two causes for the intractability of POMDPs: 1) state space size, and 2) policy size. State-of-the-art POMDP methods yield good policies even for POMDP problems with hundreds of thousands of states [7], [8] by trying to limit policy search to state space parts that are reachable and relevant for finding good policies. However, in complex real-world problems the state space can be still much larger. In POMDPs with discrete variables, the state space size grows exponentially w.r.t the number of state variables. In order to make POMDPs with large state spaces tractable, there are a few approaches: compressing probability distributions into a lower dimension [12], using factored probability distributions [13], [14], or using particle filtering to represent probability distributions [9], [10]. Particle filtering is particularly attractive, because an explicit probability model of the problem is not needed. In fact, in order to cope with a complex state space, we use particle filtering in the online POMDP method presented in more detail in Section 3.2.

6 Variational Bayes

7 Reinforcement Learning

7.1 PPO

8 Codebase

This codebase provides a framework for simulating a surgical robot navigating a deformable gridworld. It includes:

- An `ObservableDeformedGridworld` class for modeling environment states, actions, and transformations.
- A C++ module that can be wrapped with pybind11 for Python integration.
- Randomized sampling of deformations (stretch and shear) and an obstacle-rich environment.
- An infrastructure for step-by-step interaction and reward evaluation, aligning with a POMDP structure.

8.1 State Space

We are interested in the estimation of the deformaition parameters of an object in a 2D space.

$$s = (\theta)$$

θ represents the deformation tensor.

8.2 Action Space

The action space is the set of all possible actions that the robot can take, Forwards, Backwards, Left, Right.

8.3 Observation Space

The observation space is the set of all possible observations that the robot can make, i.e. the presence or absence of an obstacle in its field of view.

8.4 Bayesian estimation

rendere il belief indipendente dalla posizione e orientazione del robot, dipenderò unicamente dalla deformazione

$$b(s) = b(theta) * \delta(xyphi = real(xyphi))$$

real(xyphi) = x* from now on x = xyphi

$$\begin{aligned}
b'_{a,o}(x', \theta') &= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot \sum_{x,\theta} p(x', \theta'|x, \theta, a) b(x, \theta) \\
&= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot b(x' - a, \theta') \\
&= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot \delta(x' - a = x^*) b(\theta') \\
&= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot b(\theta') \\
&= \eta \cdot \delta(x' = x^* + a) \cdot b'_{a,o}(\theta')
\end{aligned}$$

dunque a seconda della azione scelta i non zero saranno i 4 x raggiungibili con a

$$p(o|b,a) = \sum_{x,\theta} p(o|x,\theta)b(x.\theta)$$
$$= \sum_{\theta} p(o|x^*,\theta)b(\theta)$$

9 Experiments

10 Results

References

[1] Aurore Loisy and Christophe Eloy. Searching for a source without gradients: how good is infotaxis and how to beat it. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 478(2262), June 2022.

[2] Daniel Russo and Benjamin Van Roy. Learning to optimize via information-directed sampling, 2017.

[3] Matthijs T. J. Spaan. *Partially Observable Markov Decision Processes*, pages 387–414. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.