

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**PLANEJAMENTO DE TRAJETÓRIA  
EXPLORATIVA E INFORMATIVA PARA  
MONITORAMENTO DE AMBIENTES  
DESCONHECIDOS COM UAV: UMA  
ABORDAGEM BO-POMDP**

**MARCELA APARECIDA ANICETO DOS SANTOS**

**ORIENTADORA: PROFA. DRA. KELEN CRISTIANE TEIXEIRA  
VIVALDINI**

São Carlos – SP

Outubro de 2023

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**PLANEJAMENTO DE TRAJETÓRIA  
EXPLORATIVA E INFORMATIVA PARA  
MONITORAMENTO DE AMBIENTES  
DESCONHECIDOS COM UAV: UMA  
ABORDAGEM BO-POMDP**

**MARCELA APARECIDA ANICETO DOS SANTOS**

Tese apresentada ao Programa de Pós-Graduação em  
Ciência da Computação da Universidade Federal de  
São Carlos - UFSCar, como parte dos requisitos  
para a obtenção do título de Doutora em Ciência da  
Computação.

Orientadora: Profa. Dra. Kelen Cristiane Teixeira  
Vivaldini

São Carlos – SP

Outubro de 2023



**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

Centro de Ciências Exatas e de Tecnologia  
Programa de Pós-Graduação em Ciência da Computação

---

**Folha de Aprovação**

---

Defesa de Tese de Doutorado da candidata Marcela Aparecida Aniceto dos Santos, realizada em 31/10/2023.

**Comissão Julgadora:**

Profa. Dra. Kelen Cristiane Teixeira Vivaldini (UFSCar)

Prof. Dr. Alexandre Luis Magalhães Levada (UFSCar)

Profa. Dra. Tatiana de Figueiredo Pereira Alves Taveira Pazelli (UFSCar)

Prof. Dr. Rafael dos Santos de Oliveira (CSIRO)

Prof. Dr. Andre Luiz de Oliveira (UFJF)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência da Computação.

Pela minha resiliência e dedicação investidas neste trabalho. A todos que me apoiaram e me incentivam nesta caminhada.

## AGRADECIMENTOS

Em primeiro lugar agradeço a Deus por ter me guiado e dado forças em todos os obstáculos percorridos nessa caminhada.

Agradeço imensamente a minha família, em especial, a minha mãe Doroti por sempre me apoiar em todas as situações. Ela sempre compreensiva quando aos finais de semana a dedicação aos estudos eram necessárias e, portanto, suas orações sempre presentes a mim. A memória de minha avó Ermelinda, que com todo carinho me apoiou em todos os momentos.

Ao meu tio José Luiz, por estar presente em minha vida pessoal e acadêmica desde o Ensino Fundamental, na qual me orientava nas lições da escola.

Aos tios avós Geni e Francisco, que desde a graduação me receberam na cidade de São Carlos – SP e me apoiaram.

Ao meu marido Maykon, pela paciência e incentivo em todos os momentos aqui percorridos. Até mesmo quando eu duvidei ele nunca duvidou do meu potencial.

Ao amigo da família Pe. Alfeu, sempre presente e conselheiro espiritual da palavra de Deus.

À minha orientadora Profa. Dra. Kelen Vivaldini, pelo apoio, incentivo e por todos os ensinamentos conferidos nessa caminhada.

Aos amigos do Laboratório LARIS, Lídia Rocha e Igor Araújo, que contribuíram nas discussões relacionadas a este trabalho e tantas outras discussões enriquecedoras.

Aos amigos, Eduardo Prado, Natália França, Kátia Poloni e Jonas Felinto, que fizeram os meus dias mais felizes quando o descanso era necessário.

Aos amigos, Diogo Melo, Lorrana Morelli, Caroline Raimundo, Gabriel Assis, Richard Valefuogo, Marcella Valefuogo, Jéssica Poellnitz, Welington Puerta, Alan Henrique e Tamiris Marconi que estão presentes em minha vida.

Aos professores do Departamento de Computação – DC e funcionários que contribuíram

para o meu crescimento profissional.

Aos amigos do trabalho do IFSP pelas conversas e incentivo aos estudos.

Aos professores que aceitaram participar da banca de defesa deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Agradeço imensamente!

# RESUMO

Santos, M. A. A (2020). Planejamento de Trajetória Explorativa e Informativa para Monitoramento de Ambientes Desconhecidos com UAV: Uma Abordagem BO-POMDP. 100p Tese – Departamento de Computação, Universidade Federal de São Carlos, São Carlos, 2023.

Os Veículos Aéreos não Tripulados (*Unmanned Aerial Vehicles* - UAVs) vêm sendo usados para diversas aplicações em monitoramento de ambientes complexos e desconhecidos. Um dos desafios está no planejamento de missões em que o UAV visita e explora uma área, analisando-a em tempo real, para que seja possível definir a rota a ser seguida. Essa visita envolve a maximização da área de busca a partir da definição das trajetórias. Isso possibilita a coleta de informações para aquisição de conhecimentos sobre o ambiente fornecendo um mapa do mesmo. Esse tipo de problema é conhecido como Planejamento de Trajetória Informativa (*Informative Path Planning* - IPP) e Exploração Autônoma. Nesse contexto, a Otimização Bayesiana (*Bayesian Optimization* - BO) tem sido usada em conjunto com Processo Gaussiano para coletar informações de um ambiente desconhecido e para realizar a definição das trajetórias. No entanto, para que a definição das trajetórias sejam realizadas de forma sequencial e sejam consideradas restrições específicas, existe a necessidade de um planejador responsável por essas tomadas de decisões e, neste caso, pode-se usar o Processo de Decisão de Markov Parcialmente Observável (*Partially Observable Markov Decision Processes* - POMDP). O POMDP é considerado um método não-míope, pois permite calcular com eficiência a melhor ação, verificando vários passos à frente. Portanto, considerando esses métodos, o objetivo deste trabalho é o desenvolvimento de uma estratégia para planejamento de trajetória explorativa e informativa em ambientes desconhecidos usando UAV. Para essa abordagem foi desenvolvido um sistema para a tomada de decisão sequencial sob incertezas baseado na abordagem BO-POMDP. Testes foram realizados e demonstraram que a abordagem não-Míope para exploração/exploração obteve melhores resultados se comparados com o método Míope e o randômico. Com as métricas RMSE e  $Tr(P)$  usadas para avaliar os modelos, foi inferido que abordagem não-míope obteve menor erro e que a incerteza sobre o ambiente diminuiu ao longo das iterações do algoritmo. Além disso, os resultados demonstraram que o uso da estratégia não-míope levou à redução da trajetória a ser percorrida pelo UAV em um tempo que foi equivalente ao tempo do método míope. Esse fato corrobora com a ideia de que algoritmos não-míopes podem calcular com eficiência a melhor ação, realizando trajetórias menores e, devido ao planejamento visando recompen-

sas futuras, obtendo maior chance de encontrar soluções ótimas globais.

**Palavras-chave:** Planejamento de Trajetória Informativa, Exploração, Monitoramento de Ambiente, Otimização Bayesiana, Processo de Decisão de Markov Parcialmente Observáveis, Processo Gaussiano.



# ABSTRACT

Santos, M. A. A (2020). Explorative and Informative Path Planning for Unknown Environments Monitoring with UAV: a BO-POMDP approach. 100p Thesis - Departamento de Computação, Universidade Federal de São Carlos, São Carlos, 2023.

Unmanned Aerial Vehicles (UAVs) have been used for several applications in monitoring complex and unknown environments. The challenge is to plan missions for the UAV where the vehicle needs to visit and explore an area and analyze it in real time to define the route to be followed. This visit involves maximizing the search area through trajectory definitions, enabling information collection to gain Knowledge about the environment and generate a map. This type of problem is known as Informative Path Planning and Autonomous Exploration. In this context, Bayesian Optimization (BO) has been used together with Gaussian Process to collect information from an unknown environment and define the trajectories. Moreover, However, for the definition of trajectories to be carried out sequentially and specific constraints to be considered, there is a need for a planner responsible for making these decisions and, in this case, the Partially Observable Markov Decision Process (POMDP) can be used. POMDP is considered a non-myopic method as it allows you to efficiently calculate the best action by checking several steps ahead. Therefore, considering these methods, the aim of this work is to develop a strategy for explorative and informative path planning in unknown environments using UAV. For this approach, a system for sequential decision making under uncertainty was developed based on the BO-POMDP approach. Tests were carried out and demonstrated that the non-Myopic approach to exploration/exploitation obtained better results compared to the Myopic and random methods. With the RMSE and Tr(P) metrics used to evaluate the models, it was inferred that the non-myopic approach obtained lower error and that uncertainty about the environment decreases throughout the algorithm iterations. Furthermore, the results demonstrated that the use of the non-myopic strategy led to a reduction in the trajectory to be covered by the UAV in a time that was equivalent to the time of the myopic method. This fact corroborates the idea that non-myopic algorithms can efficiently calculate the best action, taking shorter trajectories and, due to planning aiming for future rewards, obtaining a greater chance of finding global optimal solutions.

**Keywords:** Informative Path Planning, Exploration, Environmental Monitoring, Bayesian Optimization, Partially Observable Markov Decision Processes, Gaussian Processes.

## LISTA DE FIGURAS

|  |    |
|--|----|
| Figura 2.1 Ilustração do funcionamento do algoritmo CMA-ES em 6 gerações.....  | 33 |
| Figura 2.2 Problemas IPP, EA e PR. Adaptado de (SANTOS; VIVALDINI, 2022b).....   | 34 |
| Figura 3.1 A Taxonomia dos métodos analisados para IPP, EA e RP em Monitoramento de Ambientes. Adaptado (SANTOS; VIVALDINI, 2022b) .....                   | 50 |
| Figura 3.2 <i>Framework</i> para Planejamento de Trajetória Informativa .....  | 52 |
| Figura 3.3 <i>Framework</i> para Planejamento de Rota para Classificação Ativa .....   | 53 |
| Figura 3.4 <i>Framework</i> para Exploração Autônoma .....   | 54 |
| Figura 4.1 Arquitetura Proposta para o Planejamento de Trajetória Explorativa e Informativa  | 58 |
| Figura 4.2 Exemplo do funcionamento do algoritmo BO-POMDP - Árvore com 3 níveis ..   | 67 |
| Figura 4.3 Exemplo do funcionamento do algoritmo BO-POMDP - Árvore com 6 níveis ..   | 67 |
| Figura 4.4 Simulador RotorS com o ambiente construído a partir de modelos padrões do Gazebo e um Firefly hex-rotor helicopter. (FURRER et al., 2016) ..... | 71 |
| Figura 5.1 Cenários usados no simulador Gazebo .....   | 73 |
| Figura 5.2 UAV no ambiente gazebo .....  | 74 |
| Figura 5.3 Testes realizados nos 3 cenários para RMSE .....  | 76 |
| Figura 5.4 Testes realizados nos 3 cenários para Tr(P) .....   | 77 |
| Figura 5.5 Gráfico da Abordagem das distâncias obtidas pela abordagem BO-POMDP e   |    |

|   |    |
|---|----|
| pela abordagem Míope .....  | 80 |
| Figura 5.6 Gráfico das distâncias obtidas pela abordagem BO-POMDP e pela abordagem Randômica .....  | 80 |
| Figura 5.7 Gráfico dos tempos obtidos pela abordagem BO-POMDP e pela abordagem Míope .....  | 82 |
| Figura 5.8 Gráfico dos tempos obtidos pela abordagem BO-POMDP e pela abordagem Randômica .....  | 83 |
| Figura 5.9 Em a) tem-se a realização do EIPP usando o algoritmo BO-POMDP com taxa de exploração/exploração em 15%, em b) tem-se a realização do EIPP usando o algoritmo BO-POMDP com taxa de exploração/exploração em 30%, em c) tem-se a execução do algoritmo Míope e em d) tem-se a execução do algoritmo randômico..... | 87 |
| Figura 5.10 Imagens capturadas pelo UAV durante a execução de um planejamento de trajetória explorativa informativa .....   | 88 |

## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 3.1 Os métodos usados para IPP, AE e PR em monitoramento de ambientes. Adaptado (SANTOS; VIVALDINI, 2022b) ..... | 50 |
| Tabela 5.1 Distâncias obtidas em cada execução dos testes no Cenário 1 .....  | 78 |
| Tabela 5.2 Distâncias obtidas em cada execução dos testes no Cenário 2 .....  | 78 |
| Tabela 5.3 Distâncias obtidas em cada execução dos testes no Cenário 3 .....  | 79 |
| Tabela 5.4 Distâncias - Abordagens BO-POMDP e Abordagem Míope .....   | 79 |
| Tabela 5.5 Distâncias - Abordagens BO-POMDP e Abordagem Randômica .....   | 79 |
| Tabela 5.6 Tempo Cenário 1 .....  | 81 |
| Tabela 5.7 Tempo Cenário 2 .....  | 81 |
| Tabela 5.8 Tempo Cenário 3 .....  | 81 |
| Tabela 5.9 Tempos - Abordagens BO-POMDP e Abordagem Míope .....   | 82 |
| Tabela 5.10 Tempos - Abordagens BO-POMDP e Abordagem Randômica .....  | 82 |
| Tabela 5.11 Sumarização das Médias das Distâncias e das Médias dos Tempos .....   | 83 |
| Tabela 5.12 Sumarização das Médias das Distâncias e das Médias dos Tempos .....   | 84 |
| Tabela 5.13 Sumarização das Médias das Distâncias e das Médias dos Tempos .....   | 84 |
| Tabela 5.14 Teste de Wilcoxon - Cenário 1 - Distâncias .....  | 84 |

|             |  |    |
|-------------|--|----|
| Tabela 5.15 | Teste de Wilcoxon - Cenário 2 - Distâncias ..... | 85 |
| Tabela 5.16 | Teste de Wilcoxon - Cenário 3 - Distâncias ..... | 85 |
| Tabela 5.17 | Teste de Wilcoxon - Cenário 1 - Tempo .....      | 85 |
| Tabela 5.18 | Teste de Wilcoxon - Cenário 2 - Tempo .....      | 86 |
| Tabela 5.19 | Teste de Wilcoxon - Cenário 3 - Tempo .....      | 86 |

# GLOSSÁRIO

---

---

**ACO** – *Ant Colony Optimization*

**APF** – *Artificial Potential Field*

**BB** – *Branch Bound*

**BO** – *Bayesian Optimization*

**CMA-ES** – *Covariance Matrix Adaptation Evolution Strategy*

**CS** – *Cubic Spline*

**DEM** – *Digital elevation model*

**DE** – *Differential Evolution*

**DFS** – *Depth-First Search*

**EIPP** – *Explorative Informative Path Planning*

**FGD** – *Functional Gradient Descent*

**GA** – *Genetic Algorithm*

**GP** – *Gaussian Process*

**GSA** – *Genetic Simulated Annealing*

**HC** – *Hill Climbing*

**IPP** – *Informative Path Planning*

**IPSO** – *Improved Particle Swarm Optimization*

**IP** – *Interior Point*

**LARIS** – *Laboratory of Autonomous Robots and Intelligent Systems*

**LR** – *Lagrangian Relaxation*

**MAV** – *Micro Aerial Vehicle*

**MCTS** – *Monte-Carlo Tree Search*

**MDP** – *Markov Decision Process*

**MILP** – *Mixed Integer Linear Program*

**MMAS** – *Max-Min Ant System*

**NBV** – *Next-Best-View*

**NN** – *Nearest Neighbour*

**PF** – *Potential Field*

**POMDP** – *Partially Observable Markov Decision Process*

**PSO** – *Particle Swarm Optimization*

**RMSE** – *Root Mean Squared Error*

**ROS** – *Robot Operating System*

**RRT** – *Rapidly-exploring Random Trees*

**SA** – *Simulated Annealing*

**SBO** – *Sequential Bayesian Optimization*

**SLAM** – *Simultaneous Localization And Mapping*

**STAH** – *Sequential Tasks Allocation*

**STAH** – *Sequential Tasks Allocation Heuristic*

**TSP** – *Travelling Salesman Problem*

**UAV** – *Unmanned Aerial Vehicle*

**UCB** – *Upper Confidence Bound*

**UCT** – *Upper Confidence Bound for Trees*

**VRPTW** – *Vehicle Routing Problem with Time Windows*

**VRP** – *Vehicle Routing Problem*

# SUMÁRIO

|  |             |
|--|-------------|
| <b>AGRADECIMENTOS</b>  | <b>iv</b>   |
| <b>GLOSSÁRIO</b>   | <b>xiii</b> |
| <b>CAPÍTULO 1 – INTRODUÇÃO</b>   | <b>18</b>   |
| 1.1 Contextualização e Motivação . . . . .                                       | 18          |
| 1.2 Hipótese . . . . .   | 21          |
| 1.3 Objetivo . . . . .   | 21          |
| 1.4 Contribuições . . . . .  | 22          |
| 1.5 Estrutura do Trabalho . . . . .  | 22          |
| <b>CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA</b>  | <b>24</b>   |
| 2.1 Processo Gaussiano - GP . . . . .  | 24          |
| 2.2 Otimização Baysiana - BO . . . . .   | 27          |
| 2.2.1 Função de Aquisição . . . . .  | 29          |
| 2.2.2 <i>Upper Confidence Bound</i> - UCB . . . . .                              | 29          |
| 2.3 Sistema para a Tomada de Decisão Sequencial . . . . .                        | 29          |
| 2.3.1 Processo de Decisão de Markov - MDP . . . . .                              | 30          |
| 2.3.2 Processo de Decisão de Markov Parcialmente Observável - POMDP . . . . .    | 30          |
| 2.4 Estratégia Evolutiva de Matriz de Covariância Adaptativa - CMA -ES . . . . . | 31          |
| 2.5 Descrição dos problemas IPP, EA e PR em Monitoramento de Ambientes . . . . . | 33          |



|   |   |           |
|---|---|-----------|
| 2.5.1                                     | Planejamento de Trajetória Informativa . . . . .                      | 34        |
| 2.5.2                                     | Exploração Autônoma . . . . .   | 35        |
| 2.5.3                                     | Planejamento de Rota . . . . .  | 36        |
| 2.5.4                                     | Considerações dos Problemas . . . . .                                 | 36        |
| 2.6                                       | Métricas de Avaliações . . . . .                                      | 37        |
| 2.6.1                                     | Métrica $Tr(P)$ . . . . .   | 37        |
| 2.6.2                                     | Métrica RMSE . . . . .  | 38        |
| 2.7                                       | Considerações finais . . . . .  | 38        |
| <b>CAPÍTULO 3 – REVISÃO BIBLIOGRÁFICA</b> |   | <b>39</b> |
| 3.1                                       | Desafios . . . . .  | 39        |
| 3.2                                       | Planejamento de Trajetória . . . . .                                  | 40        |
| 3.3                                       | Planejamento de Trajetória Informativa . . . . .                      | 42        |
| 3.4                                       | Exploração Autônoma . . . . .   | 45        |
| 3.5                                       | Planejamento de Rota . . . . .  | 47        |
| 3.6                                       | Sumarização dos métodos usados nos problemas IPP, EA e PR . . . . .   | 49        |
| 3.7                                       | Arquiteturas de IPP, PR e Exploração Autônoma para UAVs . . . . .     | 51        |
| 3.8                                       | Considerações Finais . . . . .  | 54        |
| <b>CAPÍTULO 4 – PROPOSTA</b>              |   | <b>56</b> |
| 4.1                                       | Planejamento de Trajetória Explorativa e Informativa - EIPP . . . . . | 56        |
| 4.1.1                                     | Processo Gaussiano - GP . . . . .                                     | 58        |
| 4.1.2                                     | Formulação BO-POMDP . . . . .   | 60        |
| 4.1.3                                     | Função de Recompensa . . . . .  | 62        |
| 4.1.4                                     | Algoritmo proposto para BO-POMDP . . . . .                            | 62        |
| 4.1.4.1                                   | Exemplo do Funcionamento do Algoritmo . . . . .                       | 65        |
| 4.1.4.2                                   | Seleção de Pontos . . . . .   | 69        |

|  |  |           |
|--|--|-----------|
| 4.2  | Ambiente de Simulação . . . . .          | 69        |
| 4.3  | Unmanned Aerial Vehicle - UAV . . . . .  | 70        |
| 4.4  | Considerações Finais . . . . .           | 71        |
| <b>CAPÍTULO 5 – EXPERIMENTOS E ANÁLISES DOS RESULTADOS</b> |  | <b>72</b> |
| 5.1  | Cenários . . . . .                       | 72        |
| 5.2  | Abordagens . . . . .                     | 73        |
| 5.3  | Experimentos . . . . .                   | 74        |
| 5.3.1  | RMSE . . . . .                           | 75        |
| 5.3.2  | Trace of P - Tr(P) . . . . .             | 75        |
| 5.3.3  | Distância . . . . .                      | 78        |
| 5.3.4  | Tempo . . . . .                          | 80        |
| 5.3.5  | Testes de Wilcoxon . . . . .             | 83        |
| 5.3.6  | Exemplo de Execução . . . . .            | 87        |
| <b>CAPÍTULO 6 – CONCLUSÃO</b>                              |  | <b>89</b> |
| 6.1  | Conclusão . . . . .                      | 89        |
| 6.2  | Contribuições . . . . .                  | 90        |
| 6.3  | Limitações e Trabalhos Futuros . . . . . | 91        |
| <b>REFERÊNCIAS</b>   |  | <b>92</b> |
| <b>CAPÍTULO A –I</b>                                       |  | <b>99</b> |

# Capítulo 1

## INTRODUÇÃO

---

---

Neste capítulo será apresentada a contextualização e a motivação deste trabalho. Também serão expostas a justificativa para a escolha dos métodos propostos, a hipótese, o objetivo geral e os objetivos específicos, bem como a contribuição e a organização deste trabalho.

### 1.1 Contextualização e Motivação

UAVs (*Unmanned Aerial Vehicles*) são veículos aéreos controlados por meios eletrônicos e computacionais, capazes de serem guiados de maneira que não necessitem de pilotos embarcados. Dessa forma, podem ser remotamente pilotados (por pilotos certificados), operados de maneira semiautônoma (navegam autonomamente ao longo de uma rota pré-planejada com o monitoramento de operadores) e até mesmo serem totalmente autônomos (possuindo um grau de autonomia para a realização de tarefas como o planejamento de trajetórias, detecção e rastreamento de objetos, bem como tomada de decisões) (MEDEIRO; SILVA, 2010; BERNARDINI; FOX; LONG, 2014; FAA, 2016; VIVALDINI et al., 2018).

Além disso, os UAVs possuem características relacionadas à flexibilidade, segurança, facilidade de operação e baixo custo de propriedade. Essas características vêm facilitando a adoção desses veículos em diversas áreas, como por exemplo, em aplicações de monitoramento de tráfego (KANISTRAS et al., 2013), monitoramento de doenças em eucaliptos (VIVALDINI et al., 2018), monitoramento na agricultura (BARBEDO, 2019; SRIVASTAVA et al., 2019), cenários de busca e resgate (OLEYNIKOVA et al., 2018), inspeção industrial (KWON et al., 2019), áreas de desastre ambiental (MALANDRINO et al., 2019), monitoramento de qualidade da água (Li et al., 2019), missões de vigilância (SCHOPFERER; BENDERS, 2020), dentre outras aplicações de monitoramento como as desenvolvidas nos trabalhos de Blanchard e Sapsis (2022), Fernández et al. (2022), Yuan et al. (2023).

Como pode-se observar, a área de monitoramento de ambientes, em específico de ambientes desconhecidos, tem sido objeto de diversas pesquisas que envolvem, por consequência, a área de robótica. Assim, para aplicações desenvolvidas nesse contexto, é necessário que os UAVs visitem uma determinada área, analisem o ambiente e decidam qual rota seguir. Desse modo, é possível maximizar a área de busca para realizar a coleta de informações e/ou para obter o conhecimento sobre o ambiente em questão visando minimizar as incertezas ou buscando localizar certos alvos ou padrões. Desse modo, dependendo do tipo de monitoramento, o UAV pode ser equipado com sensores próprios para medir/capturar o alvo/padrões e, com isso, fornecer as informações/observações necessárias para a execução de uma determinada aplicação.

Na literatura, é possível observar que este tipo de situação mencionada acima tem sido abordado em pesquisas da área de robótica que envolvem os temas de: Planejamento de Rota (PR), Exploração Autônoma (EA) e Planejamento de Trajetória Informativa (*Informative Path Planning*- IPP). Estas pesquisas, apesar de distintas, são usadas para resolver a mesma problemática, todavia com enfoques diferentes. O IPP tem como objetivo principal planejar as trajetórias para a coleta de informações com interesse em possíveis alvos ou regiões de interesse. O Planejamento de Rota em ambientes desconhecidos, em alguns casos, é semelhante ao IPP no que diz respeito ao planejamento das trajetórias visando a coleta de informações de uma região/algo de interesse. Já a Exploração Autônoma possui o enfoque na construção do mapa, ou seja, produzir uma representação consistente de um ambiente desconhecido.

No entanto, a EA se assemelha ao IPP no quesito de ambos avaliarem o ganho de informação como estratégia para a realização do planejamento de trajetória informativa. Devido a isso, este trabalho considera o planejamento de trajetória explorativo e informativo, já que, para a realização do planejamento de trajetória informativa é necessário realizar a exploração do ambiente avaliando o ganho de informação para a realização das trajetórias, assim, as trajetórias com os maiores ganhos de informação serão selecionadas. Para mais detalhes sobre IPP, EA e PR ver capítulo 3.

No Planejamento de Trajetória Informativa, destacam-se questões relacionadas às tomadas de decisões autônomas que possibilitem a definição de qual rota um UAV deve seguir considerando a coleta de informações de acordo com o alvo/região de interesse visando a maximização da área de busca e consequentemente minimizando as incertezas sobre o ambiente desconhecido. Algumas pesquisas relacionadas a IPP têm adotado os métodos de Otimização Bayesiana (*Bayesian Optimization* - BO) usando Processos Gaussianos (*Gaussian Process* - GP) (MARCHANT; RAMOS, 2014; VIVALDINI et al., 2018; BLANCHARD; SAPSIS, 2022) para coletar informações sobre o ambiente e definir as trajetórias.

No entanto, para coletar informações sobre o ambiente e definir as trajetórias de forma sequencial diante de incertezas inerentes a um ambiente específico, a formulação BO-POMDP vem sendo usada. Com esta formulação pode-se realizar uma tomada de decisão sequencial calculando com eficiência a melhor ação verificando vários passos à frente.

Exemplos dessas abordagens podem ser encontrados nos trabalhos de Marchant, Ramos e Sanner (2014), Morere, Marchant e Ramos (2017) e Fernández et al. (2022). No entanto, é importante destacar que essas abordagens não incorporam às missões dos UAVs situações em que é necessário realizar um planejamento da trajetória em diferentes altitudes para a tomada de decisão. Em situações onde onde é necessário voar em altitudes mais elevadas pode-se obter uma visão mais ampla sobre o ambiente, enquanto que em outras situações, voar em altitudes mais baixas pode ser necessário para coletar dados de forma mais precisa e eficaz. Essa consideração proporcionaria uma compreensão mais abrangente e precisa do ambiente, conferindo maior flexibilidade à exploração.

No entanto, realizar um planejamento de trajetória em 3D pode ser computacionalmente custoso para abordagens BO-POMDP, pois existe um *trade-off* relacionado ao tempo necessário para monitorar exaustivamente uma grande área para obter o máximo de informações sobre o meio ambiente e buscar uma solução ótima global.

Neste contexto, considerar estratégias para definição da exploração/exploração do algoritmo é fundamental. Isto é, identificar os pontos mais promissores e estabelecer uma taxa adequada de exploração/exploração já nessa etapa pode acelerar a convergência do algoritmo, representando um elemento de relevância crucial, particularmente nas abordagens BO-POMDP.

É importante ressaltar que os planejadores baseados em POMDP são considerados métodos não-míopes, pois podem calcular com eficiência a melhor ação, verificando vários passos à frente obtendo, assim, uma maior chance de soluções ótimas globais. Já os métodos míopes, como pode ser visto nos trabalhos de (HITZ et al., 2017; POPOVIĆ et al., 2017; POPOVIĆ et al., 2020), definem a melhor ação como aquela de maior recompensa imediata e não consideram o planejamento futuro das recompensas, podendo aumentar a chance de se obter soluções que estejam em ótimos locais.

Diferentemente dos autores Marchant, Ramos e Sanner (2014) e Morere, Marchant e Ramos (2017), o estudo conduzido por Popović et al. (2020) incorpora a variação de altitude no planejamento das trajetórias. Adicionalmente, os autores adotam a Estratégia Evolutiva de Matriz de Covariância (*Covariance Matrix Adaptation Evolution Strategy* - CMA-ES) para realizar uma otimização global da trajetória. Isso confere à abordagem uma vantagem significativa em relação às técnicas existentes baseadas em *spline* para obter uma trajetória (HITZ et al., 2017;

MORERE; MARCHANT; RAMOS, 2017), pois os autores incorporam uma inicialização baseada em *Search Grid* como etapa inicial para acelerar a convergência da solução. Entretanto, é importante mencionar que a abordagem proposta é considerada míope, ou seja, a melhor ação é definida com base na recompensa imediata e possuem maior chance de obter soluções ótimas locais.

Desse modo, o foco deste trabalho está no desenvolvimento de uma abordagem não-míope para o Planejamento de Trajetória Explorativa e Informativa (*Explorative Informative Path Planning* - EIPP) em ambientes desconhecidos usando UAV a fim de obter um planejamento que maximizem o ganho de informação com a localização de possíveis alvos ou regiões de interesse e consequentemente minimize as incertezas sobre o ambiente. Para essa abordagem, será desenvolvido um sistema para a tomada de decisão sequencial sob incertezas em ambiente 3D baseado na abordagem BO-POMDP.

## 1.2 Hipótese

Tendo em vista o contexto e a motivação para a realização deste trabalho, a hipótese é a seguinte:

*”No contexto de problemas de monitoramento de ambientes desconhecidos com UAV, um Planejamento de Trajetória Explorativa e Informativa pode ser desenvolvido com o uso da abordagem BO-POMDP a fim de se obter um planejamento que maximize os ganhos de informação em uma dada região de interesse e que, consequentemente, minimize as incertezas sobre o ambiente considerando um espaço contínuo de três dimensões (3D)”.*

## 1.3 Objetivo

Com base na hipótese, o objetivo desse trabalho é:

- Desenvolver uma estratégia para EIPP com o uso da abordagem BO-POMDP a fim de se obter um planejamento que maximize os ganhos de informação em uma dada região de interesse e consequentemente minimize as incertezas sobre o ambiente considerando um espaço contínuo 3D.

## 1.4 Contribuições

Nesta sessão é apresentada uma sumarização das contribuições do presente trabalho. Algumas dessas contribuições foram apresentadas nos artigos Santos e Vivaldini (2022a) e Santos e Vivaldini (2022b) na (artigos disponíveis no Apêndice).

Seguem as contribuições deste trabalho:

- Desenvolvimento de uma abordagem integrada através da formulação matemática BO-POMDP para um sistema de tomada de decisão sequencial sob incertezas a fim de realizar o Planejamento de Trajetória Explorativa e Informativa maximizando os ganhos de informação em uma dada região de interesse e consequentemente minimizando as incertezas sobre o ambiente considerando um espaço contínuo 3D.
- Uma arquitetura modular com a estratégia de explorar o ambiente na medida em que se concentra em áreas/regiões específicas de interesse;
- Investigação do comportamento de abordagens não-míopes em comparação as abordagens consideradas míopes no contexto de Planejamento de Trajetória Explorativa e Informativa;
- A estratégia relacionada a taxa de exploração/exploração para seleção dos pontos a serem avaliados (pontos mais promissores) na abordagem BO-POMDP proposto acelerou a convergência do algoritmo. Já que o problema EIPP possui um *trade-off* relacionado ao tempo necessário para monitorar exaustivamente uma grande área para obter o máximo de informações sobre o meio ambiente e buscar uma solução ótimo global;
- Contribuições na área de robótica no que refere-se a definição dos problemas Planejamento de Trajetória Informativa, Exploração Autônoma e Planejamento de Rotas usando UAVs em monitoramento de ambientes desconhecidos. Os conceitos e diferenças entre esses problemas, bem como a taxonomia do métodos adotados para resolvê-los.

## 1.5 Estrutura do Trabalho

Este trabalho está dividido da seguinte forma: No capítulo 2 é apresentada a fundamentação teórica referente à *Gaussian Process*, *Bayesian Optimization*, *Partially Markov Decision Process* e o algoritmo CMA-ES. Esses tópicos constituem a base para a composição deste trabalho.

No Capítulo 3 é apresentada a revisão bibliográfica contendo uma visão geral do estado da arte. Neste capítulo são destacados os métodos usados para resolver os problemas de Planejamento de Rota, EA e de IPP.

No Capítulo 4 é apresentada a abordagem proposta para o Planejamento de Trajetória Explorativa e Informativa descrevendo a principal contribuição da pesquisa e as etapas do desenvolvimento do projeto, bem como os métodos adotados.

No capítulo 5 são apresentados os experimentos realizados e os resultados obtidos, e por fim, no capítulo 6 é apresentada a conclusão destacando as principais contribuições e limitações da proposta desenvolvida, assim como as contribuições que ela trará para a área e possíveis direções para trabalhos futuros.



# Capítulo 2

## FUNDAMENTAÇÃO TEÓRICA

---

---

Neste capítulo serão apresentadas as fundamentações teóricas referentes aos métodos usados neste trabalho tais como: Processos Gaussianos na Seção 2.1, Otimização Bayesiana na Seção 2.2, Processo de Decisão de Markov Parcialmente Observável na seção 2.3.2, Estratégia Evolutiva de Matriz de Covariância Adaptativa na Seção 2.4. Também será apresentada na Seção 2.5 as definições dos problemas IPP, EA e PR segundo os trabalhos da literatura. Essas definições foram publicadas em Santos e Vivaldini (2022b). Por fim, serão apresentadas as métricas usadas para avaliação da abordagem proposta neste trabalho seção 2.6.

### 2.1 Processo Gaussiano - GP

O processo Gaussiano é usado para modelar correlações espaciais de forma probabilística e não paramétrica (WILLIAMS; RASMUSSEN, 2006). É não paramétrico porque não mantém um modelo explícito do fenômeno subjacente, mas aprende a transformação entre entrada e saída diretamente de amostras de treinamento (observações). É bayesiano no sentido de que coloca distribuições (hipóteses) anteriores nas variáveis observadas e atualiza essas distribuições para produzir uma distribuição posterior conforme novos dados se tornam disponíveis. É uma técnica de regressão porque produz saídas suaves contínuas em resoluções arbitrárias e indica, não apenas a melhor estimativa dada a informação atual, mas também a incerteza inerente a essa estimativa particular (VIVALDINI et al., 2018).

Formalmente um Processo Gaussiano é definido da seguinte forma:

**Definição 1.** *Um Processo Gaussiano é uma coleção de variáveis aleatórias, onde qualquer número finito possui uma distribuição gaussiana conjunta.*

Um GP é completamente definido pela função de média  $m(x)$  e a função de covariância

$K(x, x')$  de um processo real  $f(x)$ , como na Equação 2.1.

$$\begin{aligned} m(x) &= \mathbb{E}[f(x)] \\ K(x, x') &= \mathbb{E}[(f(x) - m(x))(f(x') - m(x')))] \end{aligned} \quad (2.1)$$

Onde  $\mathbb{E}[\cdot]$  é o operador de expectativa,  $x$  e  $x'$  são ambos pontos de entrada, vetores  $D$ -dimensionais compostos pelas coordenadas espaciais correspondente a um local no ambiente.

Assim, podemos escrever uma Processo Gaussiano representado pela na Equação 2.2.

$$f(x) \sim GP(m(x), K(x, x')) \quad (2.2)$$

Na função  $K$ , conhecida por Kernel,  $x$  é o *input* de um espaço dimensional  $\mathbb{D}$  tal que  $x \in \mathbb{R}^{\mathbb{D}}$ .

Além disso, pode ser assumido um modelo de ruído aditivo para observações ruidosas  $y$  de  $f$ , onde  $\varepsilon$  é a observação ruidosa considerado independente de acordo com  $\mathbb{N}(0, \sigma_n^2)$ , um ruído Gaussiano independente.

$$y_i = f(x_i) + \varepsilon_i \quad (2.3)$$

Considere um conjunto de  $n_*$  pontos de consulta (teste)  $x_*$  especificados arbitrariamente no ambiente e coletados de  $\mathbb{D} \times n_*$  na matriz de consulta  $X^*$ . O objetivo é avaliar a distribuição Gaussiana condicional, dada por:

$$f_* | X, y, X_* \sim N(\mathbb{E}[f_*], V[f_*]), \quad (2.4)$$

onde  $V[\cdot]$  é o operador de variância e  $f_*, f(X_*)$  representa os valores da função latente correspondentes aos pontos de consulta em  $X_*$ . A condicional pode ser encontrada aplicando as equações preditivas clássicas para regressão GP, que são definidas da seguinte forma:

$$\mu = \mathbb{E}[f_*] = m(X_*) + K(X_*, X) \times [K(X, X) + \sigma_n^2 I_n]^{-1} \times (y - m(X)), \quad (2.5)$$

$$P = V[f_*] = K(X_*, X_*) - K(X_*, X) \times [K(X, X) + \sigma_n^2 I_n]^{-1} \times K(X, X_*)^T \quad (2.6)$$

onde  $I_n$  é a matriz identidade  $n \times n$  e  $\sigma_n^2$  é um hiperparâmetro que representa o ruído de observação da variância. A matriz  $n_* \times n$ ,  $K(X_*, X)$  denota os termos de correlação cruzada entre a consulta e os pontos observados, e  $K(X, X)$  e  $K(X_*, X_*)$  são os  $n \times n$  e os  $n_* \times n_*$  matrizes de covariância conjunta para os pontos de entrada e consulta observados, respectivamente. As variáveis  $\mu$  e  $P$  são introduzidas acima para representar a média GP e a covariância para fins de conveniência.

A função de covariância, ou kernel,  $K()$  determina as propriedades de generalização do modelo GP e é usada para avaliar as dependências entre as medições. É escolhido para descrever as características da função latente  $f$  em relação a uma aplicação particular. Para descrever fenômenos ambientais, existem várias funções de covariância bem conhecidas, comuns em análises geoestatísticas e estatísticas espaciais, por exemplo, o Exponencial Quadrado, o Quadrático Racional ou a família de funções Matérn na equação a seguir.

$$K_{mat3}(x, x') = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left( 1 + \frac{\sqrt{2\nu d}}{l} \right)^\nu K_\nu \left( \frac{\sqrt{2\nu d}}{l} \right) \quad (2.7)$$

onde  $d$  é a distância euclidiana entre os pontos de entrada  $x$  e  $x'$ ,  $l$  e  $\sigma_f^2$  são os hiperparâmetros da função de covariância que representa o *length scale* e *signal variance*, respectivamente.  $\Gamma()$  é a função Gamma,  $K_\nu()$  é a função de Bessel modificada de segunda ordem  $\nu$ , na qual  $\nu$  é um parâmetro positivo que determina a suavidade da covariância. Em particular, este trabalho aplica a função de covariância Matérn amplamente utilizada com  $\nu = 3/2$ :

$$K_{mat3}(x, x') = \sigma_f^2 \left( 1 + \frac{\sqrt{3}}{l} \right) \exp \left( -\frac{\sqrt{3}}{l} \right) \quad (2.8)$$

Os parâmetros livres da função de covariância, chamados hiperparâmetros  $\theta$ , controlam as relações entre pontos de dados dentro do GP. Os valores dos hiperparâmetros são otimizados para corresponder às propriedades de  $f$  treinando em vários mapas obtidos anteriormente na resolução necessária. A método comum é minimizando o Log Negativo da função de Verossimilhança Marginal (*Negative Log of the Marginal Likelihood* - NLML):

$$\log p(y|X, \theta) = -\frac{1}{2} y^T (K(X, X) + \sigma_n^2 I_n)^{-1} y - \frac{1}{2} \log(K(X, X) + \sigma_n^2 I_n) - \frac{n}{2} \log 2\pi \quad (2.9)$$

Mais informações e discussões podem ser encontrados em (WILLIAMS; RASMUSSEN, 2006).

## 2.2 Otimização Bayesiana - BO

A Otimização Bayesiana é uma estratégia de otimização que trata de funções desconhecidas onde as avaliações são custosas e ruidosas. Essa estratégia é aplicada em situações em que não se tem uma expressão definida da função objetivo, mas pode-se obter a expressão, por meio de observações (possivelmente ruidosas) desta função em valores amostrados (BROCHU; CORA; FREITAS, 2010).

Formalmente a Otimização Bayesiana é uma técnica de otimização que visa encontrar o ótimo  $x \in \mathbb{R}^D$  de uma função objetivo desconhecida  $f : \mathbb{R}^D \rightarrow \mathbb{R}$  reunindo observações ruidosas. Assim, a equação 2.10 representa esse processo:

$$\hat{x} = \arg \max_x f(x) \quad (2.10)$$

Observações ruidosas são consideradas como resultado de um aditivo *Gaussian noise* na função de avaliação. A  $i$ -ésima observação é definida como  $y_i = f(x_i) + \varepsilon$ , onde  $\varepsilon \sim N(0, \sigma_n^2)$  é o ruído associado a cada observação independente.

O algoritmo BO utiliza a *priori* Processo Gaussiano e evidências (função de aquisição) para definir uma distribuição *posteriori* ao longo do espaço de funções. O modelo Bayesiano permite uma forma sofisticada para que as informações a *priori* (parâmetros iniciais da função de covariância de um GP e a mesma) possam descrever os atributos da função objetivo, tal como a suavidade ou as posições mais prováveis de serem escolhidas (o máximo da função), mesmo quando a própria função objetivo não é conhecida (BROCHU; CORA; FREITAS, 2010).

Sendo assim, o modelo GP é atualizado com o par de dados  $(x_i, y_i)$  toda vez que uma observação ruidosa  $y_i$  é feita no local  $x_i$ . A busca para obter uma nova observação é guiada por uma função de aquisição  $h(x)$ . Em cada iteração no algoritmo BO, a localização que avalia  $f$  é determinada encontrando 2.11.

$$\hat{x} = \arg \max_x h(x) \quad (2.11)$$

Logo, o componente chave para a eficácia do BO é o uso de uma função de aquisição  $h(x)$  que oriente a busca pelo ótimo selecionando os locais onde as amostras são coletadas com base na posterior em cada iteração.

As funções de aquisições mais comuns são *Probability of Improvement* (JONES, 2001), *Expected Improvement* (JONES, 2001) e *Upper Confidence Bound - UCB* (Cox; John, 1992).

No Algoritmo 2.1 temo-se o pseudocódigo do BO clássico (MARCHANT; RAMOS, 2014).

---

**Algoritmo 2.1:** BO- Discreto
 

---

**Entrada:**  $f, h$   
**Saída:**  $\hat{x}, f(\hat{x})$

```

1 início
2   for  $i = 1, 2, 3, \dots$  iterações do
3     Encontre  $x_i = \arg \max_x h(x)$ ;
4      $y(i) \leftarrow f(x_i)$ ;
5     conjunto de treinamento  $(x_i, y_i)$ ;
6     Atualizar GP;
7   end
8 fim
```

---

Em seguida temos o pseudocódigo do algoritmo 2.2 BO contínuo (MARCHANT; RAMOS, 2014). Note que  $\mu$  é a média de uma distribuição Gaussiana e  $\beta^*$  é um conjunto de pontos a serem encontrados. O algoritmo 2.1 - BO clássico é discreto, ou seja, apenas o destino final de cada iteração é levado em consideração. No entanto o algoritmo 2.2 considera um conjunto de pontos o que representa em problemas de planejamento de trajetória os pontos de início e fim da mesma.

---

**Algoritmo 2.2:** BO (Trajetórias Contínuas)
 

---

**Entrada:**  $f, h, C$   
**Saída:**  $\hat{x}, f(\hat{x})$

```

1 início
2   for  $i = 1, 2, 3, \dots$  iterações do
3     Encontre  $\beta^* = \arg \max_{\beta} (C(\mu, \beta) | h)$ ;
4      $\{x_i, y_i\} \leftarrow C(\mu, \beta^*)$ ;
5     conjunto de treinamento  $(x_i, y_i)$ ;
6     Atualizar GP;
7   end
8 fim
```

---

### 2.2.1 Função de Aquisição

As funções de aquisições são projetadas para orientar a busca por melhores locais de amostragem na otimização bayesiana. Portanto, é crucial selecionar uma boa função de aquisição para obter melhor desempenho e eficiência da amostra. Normalmente, as funções de aquisição são projetadas para refletir a probabilidade da função objetivo ser alta em uma determinada localização. Essa probabilidade pode ser estimada usando informações posteriores completas do modelo probabilístico aprendido no BO. Existem diversas funções de aquisição populares, tais como, *Probability of Improvement - PI*, e uma das mais populares *Upper Confidence Bound - UCB*.

### 2.2.2 Upper Confidence Bound - UCB

A função de aquisição UCB descrita em (MARCHANT; RAMOS, 2014; MORERE; MARCHANT; RAMOS, 2017) é frequentemente usada pela sua simplicidade e facilidade de uso. Ela considera apenas a média e o desvio padrão de um modelo probabilístico subjacente em um determinado local de entrada:

$$UCB(x) = \mu(x) + k\sigma(x) \quad (2.12)$$

onde  $k \geq 0$  também é um parâmetro de *trade-off* de exploração-exploração. Valores altos de  $k$  resulta em mais exploração em torno de áreas menos amostradas, enquanto valores mais baixos favorecem exploração perto de locais previamente amostrados.

## 2.3 Sistema para a Tomada de Decisão Sequencial

Os Processos de Decisão de Markov (Markov Decision Process - MDP) e os Processos de Decisão de Markov Parcialmente Observáveis (Partially Observable Markov Decision Process - POMDP) são estruturas utilizadas para a tomada de decisão sequencial, usados nas áreas de planejamento de Trajetória Explorativa e Informativa. Ambos os *frameworks* definem uma forma de modelar as interações entre um agente e elaboram um plano de ações visando alcançar estados mais promissores, ou seja, calculando com eficiência a melhor ação verificando vários passos à frente.

### 2.3.1 Processo de Decisão de Markov - MDP

Os agentes podem interagir com um ambiente (domínio) realizando ações e podem observar o estado do ambiente. A execução de uma ação em um estado específico resulta na mudança de estado do ambiente, chamada de transição. Cada ação executada em um determinado estado é associada a uma recompensa que reflete o quão útil a ação é para atingir um objetivo. Os agentes agem de acordo com uma política  $\pi$ , que determina as ações a serem executadas em qualquer estado do ambiente.

Um MDP é totalmente definido pela tupla  $\langle S, A, T, R, \gamma \rangle$ , com

- $S$  é um conjunto de estados;
- $A$  é um conjunto de ações;
- $T : S \times A \times S \rightarrow [0, 1]$  é uma função de transição interpretada como a probabilidade de transição de um estado  $s'$  quando executamos a ação  $a$  no estado  $s$ , isto é,  $T(s, a, s') = p(s'|s, a)$ ;
- $R : S \times A \rightarrow \mathbb{R}$ : é uma função de recompensa que define a recompensa de executar a ação  $a$  no estado  $s$ , ou seja,  $R(s, a)$ ;
- $\gamma \in [0, 1)$  é o fator de desconto.

O fator de desconto  $\gamma \in [0, 1)$  diminui a importância das consequências da ação a longo prazo. Os MDPs dependem da suposição de Markov, que afirma que a distribuição nos estados futuros depende apenas do estado atual e, portanto, o estado atual contém todas as informações do ambiente necessárias para agir de maneira ideal.

Resolver um MDP é equivalente a encontrar a política ótima  $\pi$  começando do estado  $s_0$ :

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{T, R, \pi} \left[ \sum_{i=0}^{\infty} \gamma^i r_i \right] \quad (2.13)$$

### 2.3.2 Processo de Decisão de Markov Parcialmente Observável - POMDP

O Processo de Decisão de Markov Parcialmente Observável (POMDP) é uma estrutura para a tomada de decisão sob incerteza (SMALLWOOD; SONDIK, 1973). Ao contrário dos processos de decisão de Markov, o agente não pode observar diretamente seu estado atual. Em vez

disso, deve agir confiando na crença do estado subjacente, construída a partir de observações. Portanto, é importante para o agente manter uma distribuição de probabilidade sobre o conjunto de estados possíveis (MORERE; MARCHANT; RAMOS, 2017).

Um POMDP é totalmente definido pela tupla  $\langle S, A, T, R, \Omega, O, \gamma \rangle$ , com:

- $S$ : Conjunto de estados  $\{s_1, s_2 \dots s_n\}$ .
- $A$ : Conjunto de ações  $\{a_1, a_2 \dots a_m\}$ .
- $T : S \times A \times S \rightarrow [0, 1]$  é uma função de transição interpretada como a probabilidade de transição de um estado  $s'$  quando executamos a ação  $a$  no estado  $s$ , isto é,  $T(s, a, s') = p(s'|s, a)$ .
- $R : S \times A \rightarrow \mathbb{R}$ : é uma função de recompensa que define a recompensa de executar a ação  $a$  no estado  $s$ , ou seja,  $R(s, a)$ .
- $\Omega$ : Conjunto finito de observações  $\{\omega_1, \omega_2 \dots \omega_n\}$ .
- $O : S \times A \times \Omega \rightarrow [0, 1]$ : é uma função de observação que representa a probabilidade de observar quando a ação  $a$  foi executada e levou ao estado  $s$ , isto é  $O(o, a, s) = p(o|a, s)$ .
- $\gamma \in [0, 1]$  é o fator de desconto

Resolver um POMDP é equivalente a encontrar a política ótima  $\pi^* : \Omega \rightarrow A$ . A política ótima é definida como aquela que maximiza a soma infinita esperada de recompensas  $r_t$  a partir do estado de crença  $b_o$ . Segue formalmente na Equação 2.14.

$$\pi^* = \arg \max_x \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t r_t^\pi | b_0 \right] \quad (2.14)$$

onde  $r_t^\pi$  é a recompensa dada por seguir a política  $\pi$  no tempo  $t$ . Neste contexto, o foco é planejar as funções de transição e recompensa.

## 2.4 Estratégia Evolutiva de Matriz de Covariância Adaptativa - CMA -ES

O CMA-ES (Covariance Matrix Adaptation Evolution Strategy) é um tipo particular de estratégia para otimização numérica. As estratégias de evolução são métodos estocásticos, conhecidos

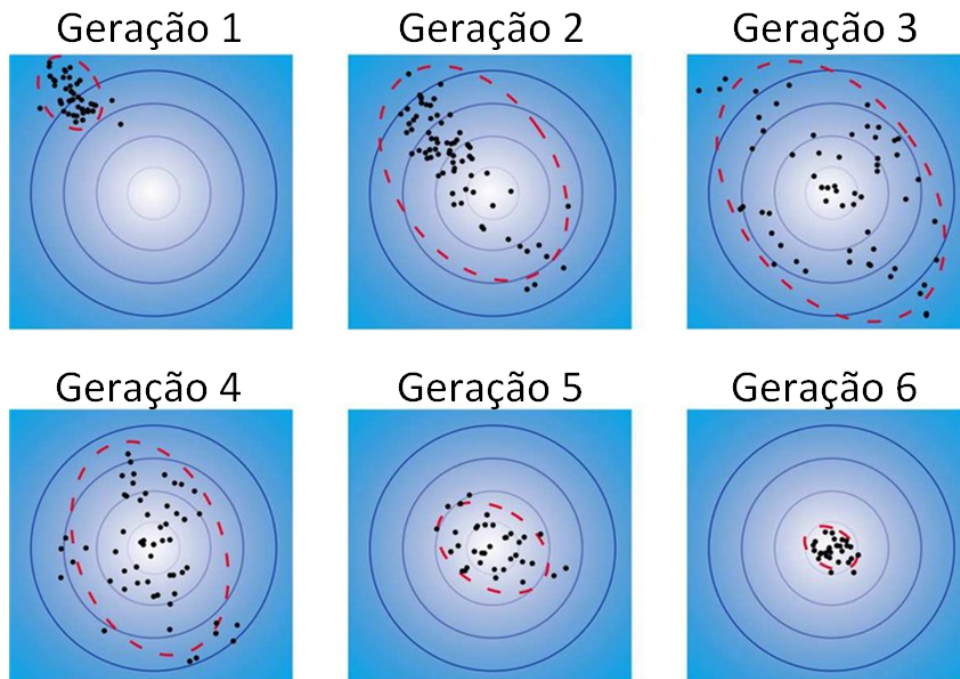


também como problemas de otimização numérica não-lineares ou problemas de otimização não-convexos contínuo. Eles pertencem à classe dos algoritmos evolutivos e da computação evolutiva (HANSEN, 2023).

Um algoritmo evolutivo é amplamente baseado no princípio da evolução biológica, ou seja, a interação repetida de variação (via recombinação e mutação) e seleção: em cada geração (iteração) novos indivíduos (soluções candidatas, denotadas como  $x$ ) são gerados pela variação, de forma estocástica, dos indivíduos parentais atuais. Em seguida, alguns indivíduos são selecionados para se tornarem pais na próxima geração com base em seu valor de aptidão ou função objetivo  $f(x)$  (HANSEN; MÜLLER; KOUMOUTSAKOS, 2003).

Assim, ao longo da sequência de gerações, indivíduos com cada vez melhores  $f$ -valores são gerados. Em uma estratégia de evolução, novas soluções candidatas são amostradas de acordo com uma distribuição normal multivariada em  $\mathbb{R}^n$ . A recombinação equivale a selecionar um novo valor médio para a distribuição. A mutação equivale a adicionar um vetor aleatório, uma perturbação com média zero. As dependências pareadas entre as variáveis na distribuição são representadas por uma matriz de covariância. A adaptação da matriz de covariância (CMA) é um método para atualizar a matriz de covariância dessa distribuição.

Na figura 2.1 tem uma execução de otimização usando CMA-ES com adaptação de matriz de covariância em um problema bidimensional simples. O cenário de otimização esférica é representado com linhas sólidas de valores  $f$  iguais. A população (pontos) são maiores do que o necessário, no entanto, mostra claramente como a distribuição da população (linha pontilhada) muda durante a otimização. Neste problema, a população concentra-se sobre o ótimo global dentro de algumas gerações, pois a medida que as gerações progredem, o algoritmo se aproxima do ótimo global ao mesmo tempo que direciona a busca ao longo do caminho para o ótimo global.

**FIGURA 2.1 – Ilustração do funcionamento do algoritmo CMA-ES em 6 gerações.**

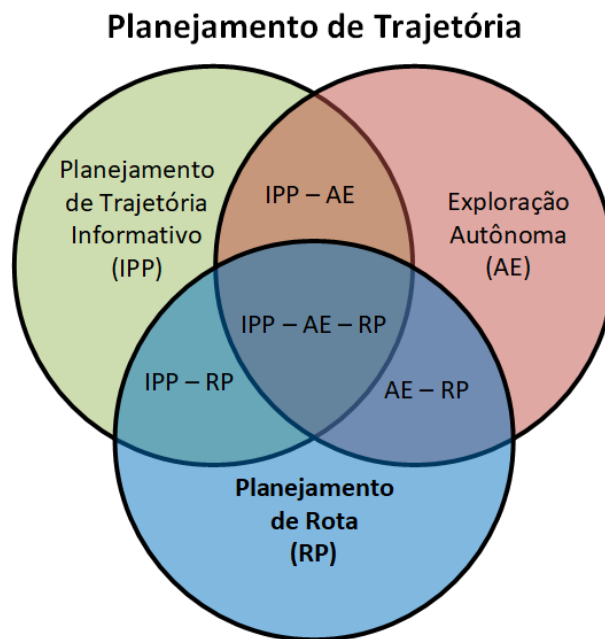
Fonte: Adaptado de Tan et al. (2019)

Vale destacar que o CMA-ES foi aplicado com sucesso para resolver problemas de alta dimensão, não lineares e em domínios contínuo, inclusive para geração de trajetória (HITZ et al., 2017; POPOVIĆ et al., 2020).

## 2.5 Descrição dos problemas IPP, EA e PR em Monitoramento de Ambientes

Como forma de analisar os problemas e fornecer uma definição para a área de robótica, esta análise contribui para a compreensão dos problemas de Planejamento de Trajetória Informativa, Exploração Autônoma e Planejamento de Rotas utilizando UAVs em ambientes desconhecidos. Vale ressaltar que os problemas foram analisados com base nas definições dadas pelos trabalhos desenvolvidos na área de robótica. Desta forma, serão apresentados os conceitos e diferenças entre eles.

Conforme mostrado na figura 2.2 o Planejamento de Trajetória Informativa, a Exploração Autônoma e o Planejamento de Rota estão destacados. A interseção entre eles formam subcategorias tais como: IPP - EA (BLANCHARD; SAPSIS, 2022), IPP - PR (VIVALDINI et al., 2018;



**FIGURA 2.2 – Problemas IPP, EA e PR. Adaptado de (SANTOS; VIVALDINI, 2022b).**

HARI et al., 2021), EA - PR (WANG et al., 2020) e IPP-EA-PR. Essas subcategorias são a união dos Problemas de IPP, EA e PR.

### 2.5.1 Planejamento de Trajetória Informativa

Baseado nos trabalhos da literatura (BINNEY; SUKHATME, 2012; MARCHANT; RAMOS, 2014; LIM; HSU; LEE, 2016; HITZ et al., 2017; POPOVIĆ et al., 2019), IPP é caracterizado pela falta de conhecimento sobre o ambiente a priori. Neste caso, é necessário obter um mapa do ambiente ou um grafo.

De acordo com (BINNEY; SUKHATME, 2012), se o cientista especificar manualmente a trajetória exata do robô ao coletar as medições do sensor, o problema será relativamente simples. Assim, o robô decidir de forma autônoma qual caminho seguir durante a coleta de medições, com base em um modelo probabilístico de um conjunto de dados a ser estudado é conhecido como IPP. Outra definição é que a tarefa de escolher trajetórias para maximizar o ganho de informação é conhecida como planejamento de trajetória informativa e é um conceito fundamental de monitoramento.(MORERE; MARCHANT; RAMOS, 2017). No mesmo sentido do trabalho anterior (POPOVIĆ et al., 2019) (POPOVIĆ et al., 2019) define IPP.

Outra definição para IPP é que planejar a rota que um veículo deve seguir de forma que um determinado objetivo seja maximizado é considerado IPP, pois, o IPP maximiza as informações coletadas considerando um alvo em uma Região de Interesse (RI) (ERGEZER; LEBLEBICIOĞLU,

2021).

Outros autores afirmam que o IPP é um problema de otimização *NP-hard* e possui um *trade-off* relacionado ao tempo necessário para monitorar exaustivamente uma grande área para obter o máximo de informações sobre o meio ambiente e buscar uma solução ótima global (LIM; HSU; LEE, 2016; POPOVIĆ et al., 2017; POPOVIĆ et al., 2019).

Para realizar o IPP é necessário calcular a rota para obter o máximo de ganho de informações sobre o ambiente, não se preocupando a princípio com o movimento que o UAV fará para realizar a trajetória, mas sim com qual rota será percorrida para coletar as informações sobre o meio ambiente. Isto deve ser feito monitorando o máximo de área possível e maximizando assim os ganhos de informações coletadas do meio ambiente.

## 2.5.2 Exploração Autônoma

Da mesma forma que o IPP, a EA se caracteriza pela falta de conhecimento sobre o meio ambiente.

Assim, de acordo com os trabalhos da literatura, tratam o problema de planejamento de trajetórias em ambientes desconhecidos como uma EA que visa produzir uma representação consistente do ambiente. A EA também envolve a tomada de decisões complexas, selecionando as trajetórias que um robô seguirá para maximizar as informações coletadas sobre o ambiente e também minimizar a incerteza geral do modelo (FRANCIS et al., 2019; MORERE et al., 2020).

Outros autores enfatizam que a EA é uma pré-condição importante para a construção de um mapa de um ambiente desconhecido com um robô para fornecer os dados de interesse. No entanto, realizar esta tarefa de forma eficiente em ambientes de grande escala ou de alta dimensão ainda é um desafio. (WANG et al., 2020; LI; SAVKIN; VUCETIC, 2020).

De acordo com (FRANCIS et al., 2019) a Exploração Autônoma pode ser vista como um aprendizado ativo que visa minimizar a incerteza e produzir mapas de alta fidelidade (Bourgault et al., 2002; STACHNISS, 2009), onde a exploração requer a solução simultânea de mapeamento, planejamento de trajetória e localização.

Assim, a EA é uma tarefa de alto nível dedicada à construção do modelo de um ambiente desconhecido no qual o veículo frequentemente toma decisões para selecionar as trajetórias que um robô seguirá para maximizar as informações coletadas sobre o ambiente e também minimizar a incerteza do modelo (WANG et al., 2020).

### 2.5.3 Planejamento de Rota

O Planejamento de Rota usando UAV pode ser considerado uma variante do problema clássico da literatura que é o Problema de Roteamento de Veículos (*Vehicle Routing Problem - VRP*<sup>1</sup>).

No domínio de roteamento de veículos há várias pesquisas sendo desenvolvidas (EKSI- OGLU; VURAL; REISMAN, 2009; BRAEKERS; RAMAEKERS; NIEUWENHUYSE, 2016) e usam a formulação de VRP para vários problemas de roteamento em geral (THIBBOTUWAWA et al., 2020).

No contexto de Planejamento de Rota usando UAV, a tomada de decisão é voltada a encontrar rotas ótimas em ambientes desconhecidos visando monitorar o ambiente a fim de maximizar a área visitada aumentando o conhecimento sobre o ambiente ou minimizando incertezas (VIVALDINI et al., 2018). Vale ressaltar que essa descrição se assemelha com o IPP.

Outra definição é que em PR considerando ambientes desconhecidos, o alvo é conhecido, mas a posição desses alvos é desconhecida. Portanto uma área deve ser explorada sem saber a localização dos alvos e portanto, é necessário que o UAV tome a decisão de identificar esses alvos e planejar sua rota, esta abordagem é conhecida como planejamento de rota para um classificação ativa. Assim, para a tomada de decisão é necessário verificar as informações e extrair os dados relevantes para a otimização de um planejamento de rotas eficiente de acordo com cada aplicação (VIVALDINI et al., 2018).

Outros autores consideram o planejamento de rotas como uma parte importante do sistema de planejamento de missão de UAVs (MENG et al., 2016). E que existem muitos fatores incertos no ambiente para a missão do UAV ou o UAV está voando em um ambiente completamente desconhecido (ZHU et al., 2010).

### 2.5.4 Considerações dos Problemas

De acordo com os trabalhos analisados na literatura, percebe-se que os problemas IPP, PR e EA apresentados tratam o Planejamento de Trajetória Informativa como um problema de tomada de decisão autônoma do robô aéreo para saber qual rota seguir para coletar informações sobre ambientes desconhecidos .

Porém, existem algumas diferenças nas definições apresentadas na literatura, o IPP e o PR

---

<sup>1</sup>O VRP busca encontrar rotas ótimas na medida em que visita um conjunto de locais com um ou mais veículos (THIBBOTUWAWA et al., 2020).

em ambientes desconhecidos têm como objetivos principais a coleta de informações para o planejamento de trajetórias. Já a EA tem como foco a construção do mapa, ou seja, produzir uma representação consistente de um ambiente desconhecido.

Percebe-se que os três problemas consideram que a trajetória que o UAV deve fazer para executar a missão deve ser feita de forma a cobrir a maior área possível para maximizar os ganhos de informações coletadas sobre um ambiente desconhecido e minimizar as incertezas do modelo. Para isso, são utilizadas diferentes restrições, como tempo máximo de monitoramento, bateria, distância, prevenção de colisão, suavização de curvas, entre outras.

## 2.6 Métricas de Avaliações

Foram avaliadas e testadas diversas métricas da literatura e com base nos trabalhos de (MORERE; MARCHANT; RAMOS, 2017; VIVALDINI et al., 2018; POPOVIĆ et al., 2019) que avaliam as abordagens no contexto de IPP foram escolhidas as métricas. Neste contexto, as métricas a serem empregadas para avaliar e comparar as abordagens míope e não míope no contexto do Planejamento de Trajetória Explorativa e Informativa (EIPP) são as seguintes:

1.  $Tr(P)$  - (Matrix Trace of  $P$ )
2.  $RMSE$  - (Root Mean Squared Error)

### 2.6.1 Métrica $Tr(P)$

Segundo Fedorov (2013), a Teoria do Projeto Experimental Ótimo (*TOED - Theory of Optimal Experimental Design*) é aplicada para medir a incerteza escalar associada a  $P$  (Eq. 2.6) de forma mais eficiente, ou seja, reduzem os custos computacionais de experimentação, permitindo que modelos estatísticos sejam estimados com menos execuções experimentais.

Outros trabalhos (SIM; ROY, 2005; CARRILLO et al., 2018; POPOVIĆ et al., 2020) também consideram o critério de informação A-ótimo  $\hat{H}(X)$ , que leva em conta a soma dos autovalores  $\lambda_i$  de  $P$ , ou equivalentemente, a variância total de seus estados, conforme dado pelo traço da matriz:

$$\hat{H}(X) = \hat{H}(P) = Tr(P) = \sum_{i=1}^D \lambda_i \quad (2.15)$$

O critério A-ótimo, busca minimizar o traço da matriz de informação. Este critério resulta

na minimização da variância média das estimativas dos coeficientes de regressão. No contexto do presente trabalho o  $Tr(P)$  está associado com a matriz de covariância do GP.

### 2.6.2 Métrica RMSE

A métrica RMSE é calculada conforme a equação 2.16 onde  $n$  é o número de observações (amostras),  $y_i$  é o valor real observado para amostra  $i$  e  $\hat{y}_i$  é o valor previsto da amostra  $i$ .

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.16)$$

## 2.7 Considerações finais

Neste capítulo foram apresentados os conceitos fundamentais que envolvem os Processos Gaussianos, a Otimização Bayesiana com destaque para os algoritmos discretos e contínuos. Abordou-se também os Sistemas de tomada de decisão MDP e POMDP e por fim, o CMA-ES a estratégia de otimização para a geração das trajetórias e, por fim, as métricas que serão usadas para avaliação da abordagem BO-POMDP foram apresentadas.

Os conceitos abordados nesse capítulo formam a base para o desenvolvimento do projeto da abordagem BO-POMDP e da estratégia CMA-ES a fim de se obter um planejamento que maximize os ganhos de informação em uma dada região de interesse e conseqüentemente minimize as incertezas sobre o ambiente considerando um espaço contínuo de três dimensões (3-D)

# Capítulo 3

## REVISÃO BIBLIOGRÁFICA

---

---

A área de monitoramento para aquisição de dados eficientes em ambientes desconhecidos vem sendo estudo de diversas pesquisas na área de robótica. De acordo com a literatura, pode-se verificar que os problemas de Planejamento de Trajetória Informativa (*Informative Path Planning* - IPP), Planejamento de Rota e Exploração Autônoma abordam a mesma problemática, porém com enfoques diferentes. Assim, com o intuito de explorar de forma abrangente o problema de planejamento de trajetória informativa e explorativa, esta revisão foi dividida da seguinte forma: na Seção 3.1 apresenta-se uma breve descrição sobre os desafios em planejar missões para aquisição de dados eficientes. Uma visão geral do estado da arte destacando os principais métodos desenvolvidos para resolver os problemas de Planejamento de Trajetória é apresentada na Seção 3.2. O Planejamento de Trajetória Informativa é apresentado na Seção 3.3. A Exploração Autônoma é apresentada na Seção 3.4 e o Planejamento de Rota é apresentado na Seção 3.6. Na Seção 3.7 abordam-se as arquiteturas de UAVs usadas para a realização do Planejamento de Trajetória Informativa, Exploração Autônoma e Planejamento de Rota. Por fim, na Seção 3.8 são feitas considerações a respeito dos métodos utilizados em cada um dos problemas destacados.

### 3.1 Desafios

Na literatura, existem diferentes propostas de pesquisa sobre os problemas de Planejamento de Trajetória Informativa, Exploração Autônoma e Planejamento de Rota. Porém, ainda persistem alguns desafios de como planejar missões para obter dados/informações de maneira eficiente em ambientes complexos e desconhecidos. Para resolver esses problemas é necessário tomar decisões para várias operações executadas por UAVs. Essas decisões requerem a construção de um mapa do ambiente da missão, para que os UAVs estejam cientes de suas localizações ou



próximos de seus alvos/objetivos. Mantendo o foco nos pontos acima mencionados, esta revisão analisa várias técnicas de planejamento de trajetória de UAVs usadas nos últimos anos. O objetivo das técnicas pode não ser apenas encontrar um caminho ideal e mais curto, mas também pode estar relacionado ao fornecimento do mapa do ambiente.

## 3.2 Planejamento de Trajetória

Majeed e Lee (2018) propuseram um algoritmo de planejamento de trajetória que encontra rotas sem colisão, usando UAVs em ambientes tridimensionais com obstáculos fixos. O algoritmo proposto reduz significativamente o tempo de computação na busca por trajetórias usando os métodos *Space Circumscription* e *Sparse Visibility Graph* no processo. Os resultados da simulação realizados em vários mapas e a comparação com outros métodos existentes, *ApVL Algorithm* e *Bounded Space Algorithm* mostram a eficácia do algoritmo proposto.

Li et al. (2020) desenvolveram um planejamento de trajetória em ambientes dinâmicos. Neste trabalho, o problema de planejamento de trajetória refere-se ao processo pelo qual um robô planeja um caminho seguro de acordo com a percepção do ambiente pelos sensores e ao mesmo tempo conclui sua tarefa com eficiência. Na abordagem proposta é usado um *Particle Swarm Optimization* (PSO) combinado com *Cubic Spline* (PSO-Sp) para resolver o problema de planejamento de trajetória do robô em um ambiente dinâmico e incerto. Os resultados experimentais mostram que a estratégia de posicionamento inercial proposta pode reduzir significativamente o caminho em um ambiente dinâmico e acelerar o processo de busca. Comparado com os métodos clássicos A\* (A-Star), *Artificial Potential Field* (APF), *Ant Colony Optimization* (ACO) e fuzzy, o *Differential Evolution*, *Genetic Algorithm* GA-Sp e PSO-Sp são mais eficientes no ambiente dinâmico incerto. Comparado com o GA-Sp e o PSO-Sp, o *Improved Particle Swarm Optimization* (IPSO-Sp) proposto obteve melhor desempenho ao encontrar o caminho mais curto, de forma estável. Esse era o principal objetivo para melhorar o PSO padrão. Vale destacar que, neste trabalho, o ambiente é conhecido, porém os obstáculos dinâmicos não.

Schopferer e Benders (2020) propuseram uma estrutura de planejamento de trajetória em cenários de baixa altitude e longo alcance. Um banco de dados georreferenciados com mapas de terreno, obstáculos e restrições de espaço aéreo foram usados. Uma função de custo multiobjetivo, considerando tempo de voo, altitude e risco operacional foi utilizada. A abordagem proposta combina o método probabilístico *Road map*, pré-processamento de amostras aleatórias, com o conhecido *A\* graph search algorithm*. Os resultados de *benchmark* apresentados para cenários de complexidade realista indicam que a localização de trajetetos no grafo de

busca pode ser realizada em poucos segundos - mesmo para voos de longo alcance por várias centenas de quilômetros. A abordagem heurística adotada para quantificar o risco com base na proximidade de objetos relevantes e na violação da camada de altitude admissível, embora permita comparar os algoritmos propostos em cenários de complexidade realista, não é aceitável para casos de uso práticos. Em vez disso, o custo deve ser modelado após uma avaliação de risco adequada que leve em consideração outros aspectos que não foram considerados.

Krell et al. (2019) desenvolveram um sistema de navegação autônoma de robôs em ambientes desconhecidos usando PSO. Os autores definem a navegação autônoma de robôs como uma integração de vários subsistemas, como por exemplo, o mapa do ambiente, a localização do robô no mapa e o planejamento do movimento (*Motion Planning*). Assim, foi construído o sistema de navegação, de modo que cada subsistema realiza uma tarefa específica sendo integrado para alcançar a missão do robô. O PSO foi usado para otimizar a trajetória do robô minimizando a distância de deslocamento. O simulador Gazebo foi usado para testar a resposta do sistema sob vários ambientes, representando uma solução para o problema de otimização. O sistema proposto manteve uma navegação robusta e evitou os obstáculos em diferentes ambientes desconhecidos.

Biswas, Anavatti e Garratt (2019) apresentaram um planejamento de trajetória baseado em PSO em ambientes de terreno desconhecido. A abordagem adotada integra duas estratégias, uma evita obstáculos e a outra executa o reconhecimento do terreno. Essas estratégias são integradas no PSO para realizar o planejamento de trajetória. Os resultados da simulação mostram que o algoritmo proposto é capaz de planejar com eficiência um caminho ideal em ambientes simples e complexos.

Apesar dos trabalhos Krell et al. (2019), Biswas, Anavatti e Garratt (2019) e Li et al. (2020) não usarem UAVs em suas aplicações, os autores apresentam abordagens de planejamento de trajetória que também poderiam ser adotadas para robôs aéreos.

Vale ressaltar também que a ideia de apresentar trabalhos relacionados ao planejamento de trajetória visa dar uma visão geral do estado da arte e apresentar os métodos que vêm sendo usados e como os autores desenvolvem a problemática considerando as diversas restrições. Já que, o proposto trabalho combina conceitos de exploração do ambiente e planejamento da trajetória que é o responsável por executar a rota informada pelo planejamento de trajetória informativa. Assim, considerando a visão geral do problema teremos um embasamento mais adequado para as soluções propostas.

Outro destaque é que nem todos trabalhos citados acima consideram o ambiente desconhecido, como é o caso dos trabalhos Schopferer e Benders (2020), Li et al. (2020).

### 3.3 Planejamento de Trajetória Informativa

Na busca por soluções para o Planejamento de Trajetória Informativa (IPP), a literatura apresenta uma variedade de métodos propostos. Entre eles, destacam-se os algoritmos baseados em Grafos, os quais são empregados para identificar trajetórias com o intuito de obter informações valiosas sobre o ambiente. Singh et al. (2009), Binney, Krause e Sukhatme (2010), Binney e Sukhatme (2012).

Binney e Sukhatme (2012) usaram o método *Branch and Bound* para resolver o IPP. O algoritmo usa a monotonicidade da função objetivo para fornecer uma aceleração dependente da função objetivo versus a *Brute Force Search*. Foram apresentados os resultados que sugerem que ao maximizar a redução da variância em um modelo de processo gaussiano, a aceleração do algoritmo é significativa. Para validão do algoritmo foram analisados o tempo de execução em diferentes cenários.

Vale destacar que em situações em que há um grande conjunto de dados, os métodos baseados em grafos são limitados quanto a resolução dos problemas. Nestas situações, esses métodos possuem crescimento exponencial o que dificulta a busca por uma solução.

Uma outra proposta para o planejamento de trajetória informativa foi desenvolvido por Marchant e Ramos (2014), sendo aplicada ao monitoramento de ambiente usando BO. A abordagem usa um algoritmo em camadas, onde a primeira camada otimiza o fenômeno ambiental para encontrar áreas de interesse (por exemplo, alta poluição, alta temperatura). Essa camada usa um modelo espaço-temporal incremental do fenômeno, dado por um Processo Gaussiano (WILLIAMS; RASMUSSEN, 2006; VASUDEVAN et al., 2009) anterior, que leva em consideração a incerteza e os valores previstos propagados no tempo. A segunda camada é usada para encontrar o melhor conjunto de parâmetros que determinam um caminho contínuo no qual o robô percorre. A validação do algoritmo foi realizada em um ambiente de larga escala para monitorar a concentração de ozônio nos EUA e em um robô móvel que monitora a dinâmica das mudanças de luminosidade.

Marchant, Ramos e Sanner (2014) propuseram uma abordagem utilizando *Sequential BO problem* (SBO) juntamente com *Partially Observable Markov Decision Process* (POMDP). Esta abordagem consiste em determinar uma solução para uma tomada de decisão não míope/gulosa que maximize uma função de recompensa e leve em consideração uma crença de um processo espaço-tempo desconhecido e o estado de um robô atuando com um sensor. Foi formulado uma solução para SBO com informações de POMDP usando uma versão modificada do algoritmo *Upper Confidence bound for Trees* - (UCT) para *Monte-Carlo Tree Search* - (MCTS), que é

uma forma escalonável de encontrar decisões assintoticamente ótimas para trabalhar com estado contínuo e espaços de observação.

Como continuidade do trabalho desenvolvido por Marchant, Ramos e Sanner (2014), outra abordagem usando BO-POMDP foi proposta por Morere, Marchant e Ramos (2017). A melhoria desta abordagem esta na reformulação da função de recompensa do POMDP para equilibrar a exploração (áreas de alto gradiente) e a exploração (áreas de alta incerteza). Isso significa que incentivar o robô a *exploit* (áreas de alto gradiente) é fundamental para obter um monitoramento de alta qualidade, pois os mapas reconstruídos geralmente apresentam maiores erros em áreas complexas e variáveis. Em experimentos simulado mostrou-se o comportamento da nova função de recompensa, que favorece áreas variáveis em relação às constantes, resultando em melhor precisão de mapeamento e recompensa acumulada do que os métodos anteriores. Comparações de reconstrução de terreno 3D mostram que BO-POMDP atinge menos erros do que seu equivalente míope. Por fim, os autores mostram os primeiros resultados práticos de uma formulação BO-POMDP, usando um *ARDrone quad-copter* para realizar o monitoramento em ambiente real. Embora os experimentos em ambiente real não tenham produzido um mapa reconstruído muito preciso, o problema da reconstrução precisa das bordas é desafiador ao lidar com poucos e dados escassos.

Uma abordagem que podemos destacar é a desenvolvida por Popović et al. (2020), pois são estratégias adaptativas e não adaptativas. Abordagens não adaptativas exploram um ambiente usando uma sequência de ações pré-determinadas para execução da rota (GALCERAN; CARRE-RAS, 2013). Já abordagens adaptativas permitem que as rotas mudem conforme as informações são coletadas, tornando-as adequadas para o planejamento com base em interesses específicos (LIM; HSU; LEE, 2016; HITZ et al., 2017; POPOVIĆ et al., 2019, 2020) .

Popović et al. (2016) apresentam uma estrutura de planejamento de trajetória informativa para classificação ativa usando UAVs. O algoritmo desenvolvido usa uma combinação de seleção de ponto de vista global e otimização evolutiva, usando o algoritmo *Covariance Matrix Adaptation Evolution Strategy* (CMA-ES), para refinar a trajetória planejada no espaço 3D contínuo e satisfazer restrições dinâmicas. A abordagem foi avaliada na aplicação de detecção de plantas daninhas para agricultura de precisão em ambiente simulado comparando os métodos *RIG-tree* e *Coverage Path Planners*, analisando os efeitos de diferentes estratégias de planejamento. Os resultados mostram que o algoritmo proposto cria mapas com entropia mais baixa em comparação com *lawnmower* no mesmo período de tempo .

Popović et al. (2017) propuseram uma extensão do trabalho anterior com o foco nas estratégias de mapeamento multirresolução para o planejamento de trajetória informativa no mo-

nitoramento de terrenos agrícolas usando UAVs. A estratégia explora a correlação espacial codificada em um modelo de Processo Gaussiano como um precedente para a fusão de dados bayesianos com sensores probabilísticos. Isso permite incorporar vários modelos de sensores dependentes para imagens aéreas e realizar atualizações de medição em tempo constante. Os mapas resultantes foram usados para planejar trajetórias ricas em informações no espaço 3D contínuo por meio de uma combinação de *Grid Search* e otimização evolutiva. Simulações realizadas mostram que o planejador obteve um desempenho melhor do que os métodos, *lawn-mower RIG-tree*, com reduções médias de erros de até 45% em comparação com a cobertura tradicional do *lawnmower*.

Hitz et al. (2017) propuseram uma abordagem para o planejamento de trajetória informativa que. Ao invés de construir gradualmente uma solução em um grafo, os autores usam um conjunto de caminhos completos no espaço contínuo, juntamente com um método de otimização evolutiva (CMA-ES) para maximizar o ganho de informações. Além disso, foi apresentado um esquema de replanejamento para adaptar o caminho de acordo com os dados medidos, permitindo focar automaticamente nas regiões de interesse, garantindo ao mesmo tempo que o robô explore suficientemente a área em estudo. O método foi extensivamente testado tanto em simulações quanto em ambiente real. Os autores destacam especificamente a eficácia do planejador para monitorar microorganismos tóxicos em um lago pré-alpino e para identificar pontos quentes dentro de sua distribuição.

Como continuação dos trabalhos desenvolvidos por Popović et al. (2016) e Popović et al. (2017), Popović et al. (2020) apresentam uma estrutura geral de planejamento de trajetória informativa para monitorar ambientes usando UAV. A abordagem é capaz de mapear variáveis discretas e contínuas em um terreno usando dados de resolução variável recebidos de sensores probabilísticos. Durante uma missão, os mapas do terreno construídos on-line são usados para planejar trajetórias ricas em informações em um espaço 3D contínuo, otimizando as soluções iniciais obtidas por *Coarse Grid Search*. Os autores compararam os métodos *Lattice Graph*, CMA-ES, *Interior Point-(IP)*, *Simulated Annealing-(SA)* e BO. Com as simulações realizadas foi possível observar que a abordagem proposta com o CMA-ES é mais eficiente que os métodos existentes. Também foi feita uma aplicação em tempo real em um cenário de mapeamento fotorrealista usando um conjunto de dados disponíveis.

## 3.4 Exploração Autônoma

No mesmo sentido do Planejamento de Trajetória Informativa, a Exploração Autônoma possui como principal característica a falta de conhecimento sobre o ambiente e no quesito de ambos avaliarem o ganho de informação como estratégia para a realização do planejamento de trajetória informativa.

Alguns autores tratam o planejamento de trajetória em ambientes desconhecidos como uma exploração autônoma que tem por objetivo produzir uma representação consistente do ambiente. A exploração autônoma também envolve tomada de decisões complexas, selecionando as trajetórias que um robô deve seguir de forma a maximizar as informações coletadas sobre o ambiente e também minimizar a incerteza geral do modelo (FRANCIS et al., 2019; MORERE et al., 2020).

Outros autores enfatizam que a exploração autônoma é uma das principais pré-condições para a construção de um mapa de um ambiente desconhecido com um robô e fornecer os dados de interesse de um determinado ambiente. No entanto, esta tarefa por vezes pode ser não-trivial e computacionalmente cara em ambientes de grande escala ou dimensões elevadas (WANG et al., 2020; LI; SAVKIN; VUCETIC, 2020).

Em contexto de mapeamento de ambientes, a exploração é o processo de produção de mapas de alta fidelidade, onde a exploração requer a solução simultânea de mapeamento, planejamento de trajetória e localização (STACHNISS, 2009; Bourgault et al., 2002). Segundo Francis, Ott e Ramos (2020), este é um problema complexo principalmente devido a dimensionalidade do espaço de solução.

Francis et al. (2019) propuseram uma abordagem para a exploração autônoma segura e para construção de mapas com base no método de BO. Este método encontra caminhos contínuos ideais em vez de locais de detecção discretos que satisfazem restrições de movimento e segurança do UAV. Ao equilibrar a função de recompensa e o risco associados a cada caminho, o otimizador minimiza o número de avaliações de funções que são custosas computacionalmente. A eficácia da abordagem foi validada através de uma série de experimentos simulados e com um robô real comparando com outras técnicas de exploração. Os resultados experimentais mostram que o método ofereceu desempenho robusto e consistente em todos os testes obtendo um desempenho melhor ou tão bom quanto o estado da arte.

Francis, Ott e Ramos (2020) propuseram uma abordagem baseada na *Functional Gradient Descent* (FGD) para otimizar eficientemente os caminhos exploratórios em mapas de ocupação contínua. O FGD estocástico foi adotado para superar as limitações dos métodos padrão de

FGD, a fim de garantir a convergência.

Esse processo permite a otimização de todo o caminho, resultando em caminhos suaves contínuos que maximizam a qualidade geral do mapa, mantendo o robô a salvo de colisões. A partir dos experimentos foi verificado a habilidade do método proposto em encontrar caminhos suaves e seguros, além disso os resultados foram comparados com métodos de exploração, tais como, *Rapidly-exploring Random Trees (RRT) planner* (YANG; GAN; SUKKARIEH, 2013b) e *Frontier* (YAMAUCHI, 1997).

Rossi e Brunelli (2015) desenvolveu uma abordagem baseado em um algoritmo de monitoramento para localização de vazamento de gás. Foi modelado o perfil das medidas de concentração de gás, usando um modelo de distribuição gaussiana 2D e a busca foi realizada aplicando uma estratégia de exploração baseada no algoritmo *hill-climbing*. A estratégia de localização da fonte de gás otimiza a velocidade do robô aéreo enquanto minimiza o consumo de energia do sistema de monitoramento. Assim, para avaliar o desempenho dessa estratégia foi feita uma simulação e fornecidos um projeto de exploração espacial para UAV movidos a energia solar.

É interessante observar no trabalho de Rossi e Brunelli (2015), que apesar dos autores considerarem o algoritmo proposto como uma estratégia explorativa que visa a construção de um mapa do ambiente para localização de vazamento de gás. O mesmo poderia ser apresentando como Planejamento de Trajetória Informativa já que as informações coletados do ambiente como missão principal é a detecção de gás em ambiente desconhecido e para que isso aconteça de forma autônoma um mapa é construído on-line para que o UAV se localize no ambiente.

Outra abordagem para a exploração eficiente de UAV foi proposta por Wang et al. (2020), onde foi construído um *road map* de forma incremental junto com o processo de exploração que exhibe explicitamente a estrutura topológica do ambiente em 3D. Ao simplificar o ambiente, *road map* pode fornecer com eficiência o ganho de informação e o custo para uma região candidata a ser explorada, que são duas quantidades para avaliação *Next-Best-View (NBV)*, sugerindo assim a eficiência para a determinação do NBV. Além disso, foi feito um planejador local baseado no método de *Potential Fields* que conduz o robô para a área rica em informações durante o processo de navegação. O *framework* proposto e seus módulos foram verificados em vários ambientes 3D usando o simulador Gazebo, os quais apresentam suas características distintas na seleção de NBV e melhor desempenho na eficiência de exploração com o método proposto do que o método RRT.

## 3.5 Planejamento de Rota

O Planejamento de Rota usando UAV pode ser considerado uma variante do problema clássico da literatura que é o Problema de Roteamento de Veículos (*Vehicle Routing Problem* - VRP<sup>1</sup>). Nesse contexto, a tomada de decisão voltada a encontrar rotas ótimas em ambientes desconhecidos visa monitorar o ambiente a fim de maximizar a área visitada aumentando o conhecimento sobre o ambiente ou minimizando incertezas (VIVALDINI et al., 2018). Alguns métodos usados para o planejamento de rota são a Otimização Bayesiana (*Bayesian Optimization* - BO) usando Processos Gaussianos (*Gaussian Process* - GP) (VIVALDINI et al., 2016, 2018), *Rapidly-exploring Random Tree* (RRT) (MENG et al., 2016), Recozimento Simulado (*Simulated Annealing* - SA) (CAPRAZ; AZYIKMIS; OZSOY, 2020), dentre outros. Lavallo e Kuffner (2000) propuseram um método para o planejamento de rota baseado no RRT, que constrói iterativamente um grafo entre um ponto de origem e um objetivo, criando um caminho para a navegação.

Já os autores Yang, Gan e Sukkarieh (2013a) desenvolveram um algoritmo que combina o RRT com mapas de ocupação gerados por um GP para exploração de ambiente desconhecidos. O método consiste em um planejador de rota que coleta informações sobre uma área de busca, focando em partes com maior incerteza e seguindo uma geometria desorganizada.

Weinstein e Schumacher (2007) formulou um modelo de *Mixed Integer Linear Program* (MILP) para o problema de alocação de tarefas com base no *Vehicle Routing Problem with Time Windows* (VRPTW), adicionando várias restrições para representar uma variedade de cenários focados na necessidade militar para inteligência precisa, vigilância e reconhecimento. O foco foi agrupar os alvos e implementar essas informações no MILP para atribuir de maneira otimizada as rotas dos UAVs em direção aos alvos, considerando um único depósito e muitos locais de alvos. Foi comparado tempo de solução para minimizar três funções de custo diferentes: distância total de todas as rotas, menor rota e tempo total da realização das tarefas.

Karakaya (2014) apresentou um planejamento de rota com UAV usando o algoritmo *Max-Min Ant System* (MMAS) modificado que calcula distâncias mínimas cobrindo um número maior de alvos em um intervalo de vôo predefinido. O método proposto foi comparado com uma heurística do vizinho mais próximo *Nearest Neighbour* (NN) e os resultados mostram um aumento de até 10% no número de alvos cobertos.

Uma técnica de planejamento de rota de UAV baseada no algoritmo *Genetic Simulated Annealing* (GSA) foi aplicada em Meng e Xin (2010). *Digital Elevation Model* (DEM) foi

---

<sup>1</sup>O VRP busca encontrar rotas ótimas na medida em que visita um conjunto de locais com um ou mais veículos (THIBBOTUWAWA et al., 2020).



usado para produzir uma superfície de vôo suave, enquanto o algoritmo de GSA planeja a rota do UAV nessa superfície. O GSA foi aplicado com sucesso e os resultados simulados demonstraram a eficácia desse algoritmo.

Turker, Sahingoz e Yilmaz (2015) usa o SA para planejar a rota de um UAV em um cenário 2D com obstáculos aleatórios. Os resultados simulados demonstraram que o SA gera soluções aceitáveis, capaz de evitar obstáculos e seguir a rota global planejada.

Meng et al. (2016) propuseram um algoritmo baseado no *Rolling Rapidly-Exploring Random Tree* para o problema de planejamento de rota de UAVs em ambiente desconhecido. De acordo com as informações atuais do ambiente, o planejamento de rota local é realizado ao mesmo tempo, as novas informações do ambiente são detectadas e o próximo estágio da rota é gerado. Os resultados da simulação mostram que o método proposto pode atender à demanda do planejamento de rotas on-line em ambiente desconhecido.

Kim et al. (2017) desenvolve uma abordagem para o planejamento de entrega e retirada de medicamentos auxiliado por UAV. As rotas são pré-determinadas e as ordens de entrega e retirada são conhecidas (destinos). O modelo de Planejamento Operacional foi proposto para um cronograma ideal de voo de UAVs para cada centro, e o método de análise de custo-benefício foi introduzido como critério de tomada de decisão. Uma análise computacional foi realizada para comparar o desempenho do problema usando o método *Partition* e o algoritmo *Lagrangian Relaxation* (LR), que produziu um desempenho melhor que o modelo sem esses componentes.

Vivaldini et al. (2016) propuseram uma metodologia de planejamento de rotas com classificação ativa para UAVs, a fim de aumentar o conhecimento das áreas visitadas e minimizar as incertezas na classificação de árvores doentes. Foram avaliados cinco algoritmos de planejamento de rotas diferentes, nomeadamente BO contínuo, BO discreto, pontos aleatórios, trajetória pré-estabelecida e RRT. Em todos os métodos, um modelo de Regressão Logística fornece as imagens classificadas e um Processo Gaussiano interpola as informações no mapa. Ao contrário dos métodos tradicionais de planejamento de rota, o BO contínuo com classificação ativa não tem todos os alvos na rota pré-estabelecidos, mas gera novos alvos ao longo do caminho. A principal vantagem do BO contínuo é a combinação de planejamento de rota (construção de uma rota entre os pontos de origem e objetivo) e classificação ativa (que permite a escolha de objetivos com base em como eles afetam a incerteza do ambiente).

Park, Lee e Morrison (2017) usa um modelo MILP e uma heurística de alocação de tarefas sequenciais *Sequential Tasks Allocation Heuristic* (STAH) no planejador da missão para determinar o serviço essencial da alocação de tarefas do UAV. Um cliente seleciona a rota desejada, limitada a 4 locais específicos, e as rotas e os trabalhos divididos para cada local são

predeterminados, com o UAV calculando sua trajetória. O modelo MILP oferece uma solução ideal apenas para modelos pequenos devido a um tempo computacional lento, enquanto as aproximações heurísticas podem determinar soluções não ideais com muito mais eficiência.

Como extensão de Vivaldini et al. (2016), Vivaldini et al. (2018) desenvolveram um framework de planejamento de rotas para classificação ativa usando UAVs visando maximizar as informações coletadas dentro de uma determinada distância, limitada pela autonomia de voo. Os autores propuseram BORRT, onde BO escolhe os pontos de destino e o RRT sugere a trajetória entre estes pontos. O classificador de Regressão Logística foi usado para classificar o solo e árvores doentes/saudáveis e um GP foi usado para interpolar essas informações, produzindo um mapa de navegação. Ao contrário dos métodos tradicionais de planejamento de rotas, que possuem metas pré-estabelecidas, o *framework* de classificação ativa proposto pode se adaptar ao fluxo constante de novas informações. A partir de sua posição atual (GPS e dados inerciais) e usa o BO contínuo para buscar o melhor ponto de destino a ser visitado de acordo com seu objetivo atual (ou seja, exploração), para maximizar a quantidade de informações coletadas em cada etapa. A partir desses pontos, o algoritmo RRT identifica as trajetórias consideradas relevantes a serem percorridas. Resultados experimentais mostram que a abordagem proposta se compara favoravelmente a outros métodos de planejamento de rotas tradicionalmente usados.

É interessante destacar que o trabalho de Vivaldini et al. (2016) e Vivaldini et al. (2018) poderiam ser apresentados como IPP já que para tomada de decisão referente a qual rota percorrer para coletar informações do ambiente foi feito de forma a percorrer o máximo de área possível para maximizar os ganhos de informações coletados sobre um ambiente.

### **3.6 Sumarização dos métodos usados nos problemas IPP, EA e PR**

Na Tabela 3.1 apresenta uma sumarização dos métodos usados em problemas de IPP, EA e PR juntamente com as aplicações de cada trabalho e o ano em que foram publicadas.

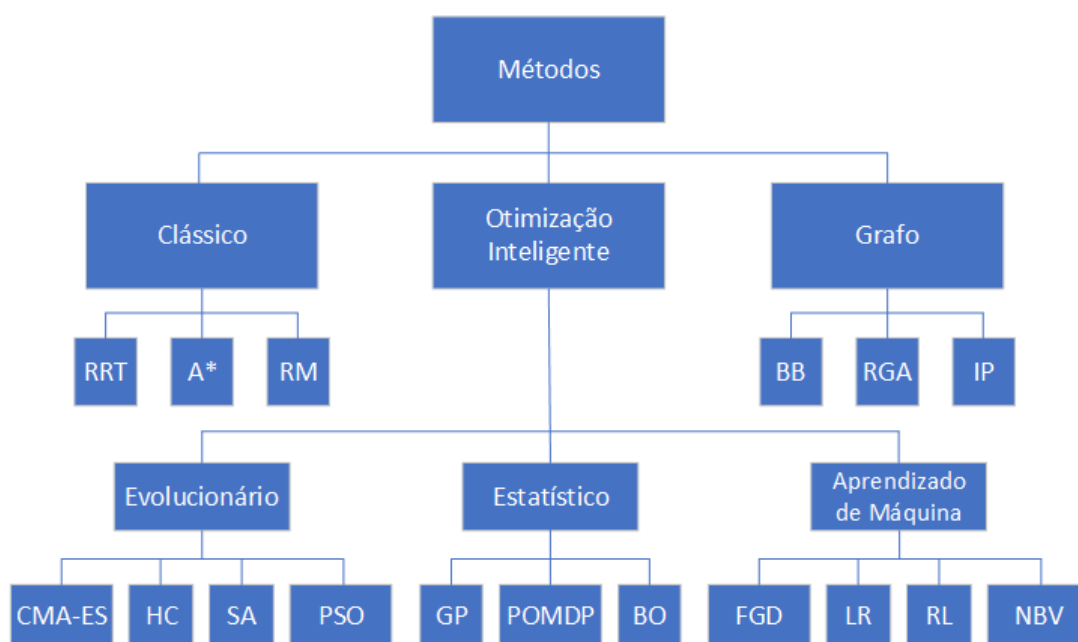
É importante destacar que esse resumo pode orientar a aplicação dos métodos, relacionando-os com os problemas estudados, que são IPP, EA e PR.

Conforme mostrado na Figura 3.1 foi desenvolvida a taxonomia dos métodos analisados para IPP, AE e PR no ambiente de monitoramento para a revisão da literatura publicados por (SANTOS; VIVALDINI, 2022b). Os métodos foram divididos em categorias Clássicas, baseados em Grafos e de Otimização Inteligente (Evolucionária, Baseada em Estatística e Aprendizado de Máquina).

| Ano  | Autor                | Metodos    | Aplicação  | Problema |
|------|----------------------|------------|--|----------|
| 2014 | Marchant and Ramos   | BO-GP      | Monitoramento de ambientes                           | IPP      |
| 2015 | Rossi                | HC         | Monitoramento para localização de vazamento de gás   | EA       |
| 2016 | Lim, Hsu and Lee     | RAId       | Região de desastre para localização de sobreviventes | IPP      |
| 2016 | Meng et al.          | RRT        | Monitoramento de ambientes                           | PR       |
| 2017 | Hitz et al.          | CMA-ES     | Monitoramento cyanobacteria toxica                   | IPP      |
| 2019 | Francis and Ramos    | RL and FGD | Exploração   | EA       |
| 2019 | Vivaldini et al.     | BO-RRT     | Classificação de árvores doentes                     | PR       |
| 2020 | Wang et al.          | NBV        | Exploração   | EA       |
| 2021 | Hari et al.          | TSP        | Persistent monitoring missions                       | PR       |
| 2022 | Blanchard and Sapsis | BO-POMDP   | Monitoramento para detecção de anomalia              | IPP e EA |

Tabela 3.1: Os métodos usados para IPP, AE e PR em monitoramento de ambientes. Adaptado (SANTOS; VIVALDINI, 2022b)

FIGURA 3.1 – A Taxonomia dos métodos analisados para IPP, EA e RP em Monitoramento de Ambientes. Adaptado (SANTOS; VIVALDINI, 2022b)



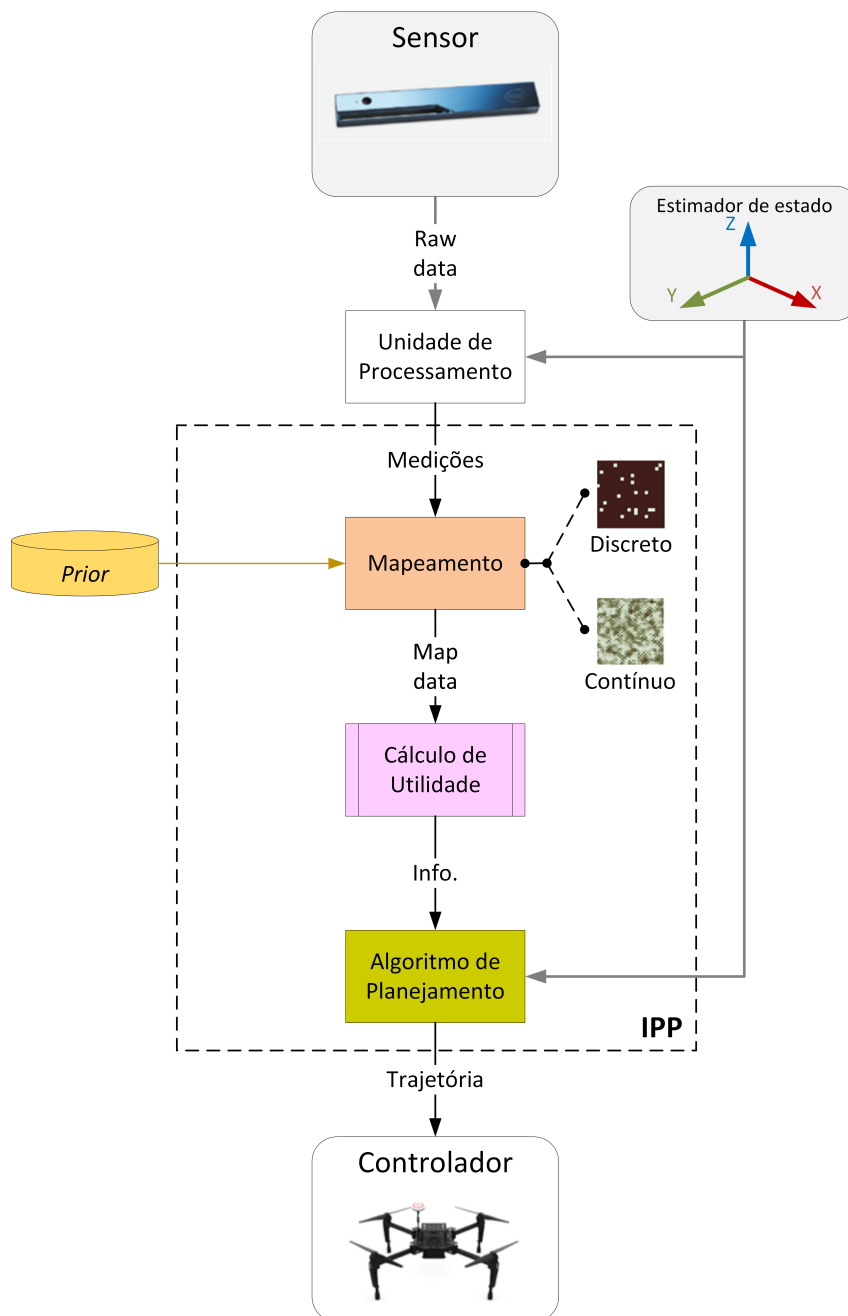
## 3.7 Arquiteturas de IPP, PR e Exploração Autônoma para UAVs

As arquiteturas apresentadas a seguir estão relacionadas aos problemas de IPP, Planejamento de Rota e Exploração Autônoma aplicadas em UAVs. O intuito deste tópico é apresentar a forma como estão sendo elaboradas as arquiteturas e o funcionamento geral dos métodos em cada módulo.

A arquitetura da Figura 3.2 foi proposta por Popović et al. (2020) para o planejamento de trajetória informativa. Segundo os autores, nesta arquitetura são mostrados os principais elementos do IPP destacando as etapas de **Mapeamento**, **Cálculo de Utilidade** e **Algoritmo de Planejamento**. Importante destacar que o módulo Cálculo de Utilidade também é conhecido como Ganho de Informação.

De um modo geral, o funcionamento da arquitetura é dado da seguinte forma: o mapa do ambiente é construído usando medições extraídas de um fluxo de dados de um sensor. Em um determinado momento, o conhecimento do mapa é usado pela unidade de planejamento para encontrar as trajetórias mais úteis para a coleta de dados.

Vale ressaltar que o módulo **Cálculo de Utilidade** proposto pelos autores não considera a estratégia de planejamento geral do orçamento restante (conhecido como *budget*). Neste caso, em uma situação em que a bateria estivesse com pouca carga, não seria possível realizar qualquer manobra para finalizar a missão, por exemplo, retornar o robô ao ponto inicial ou pousar o UAV com segurança. Palazzolo e Stachniss (2018), Hollinger e Sukhatme (2014) mostram como esses aspectos podem ser facilmente tratados por meio de um planejamento com objetivos em função do tempo. Tais funções de utilidade também permitiriam definir objetivos precisos com variação no tempo, por exemplo, encaminhando o robô para primeiramente explorar um ambiente de modo uniforme antes de se concentrar em áreas específicas de interesse. Outro ponto a se considerar sobre o módulo Cálculo de Utilidade é que os autores propuseram um método Míope para realizar o Planejamento de Trajetória Informativa, ou seja, escolhem a melhor ação baseado na maior recompensa imediata e não consideram o planejamento futuro das recompensas podendo aumentar a chance de se obter soluções que estejam em ótimos locais ou até mesmo podem efetuar cálculos repetitivos.

FIGURA 3.2 – *Framework* para Planejamento de Trajetória Informativa

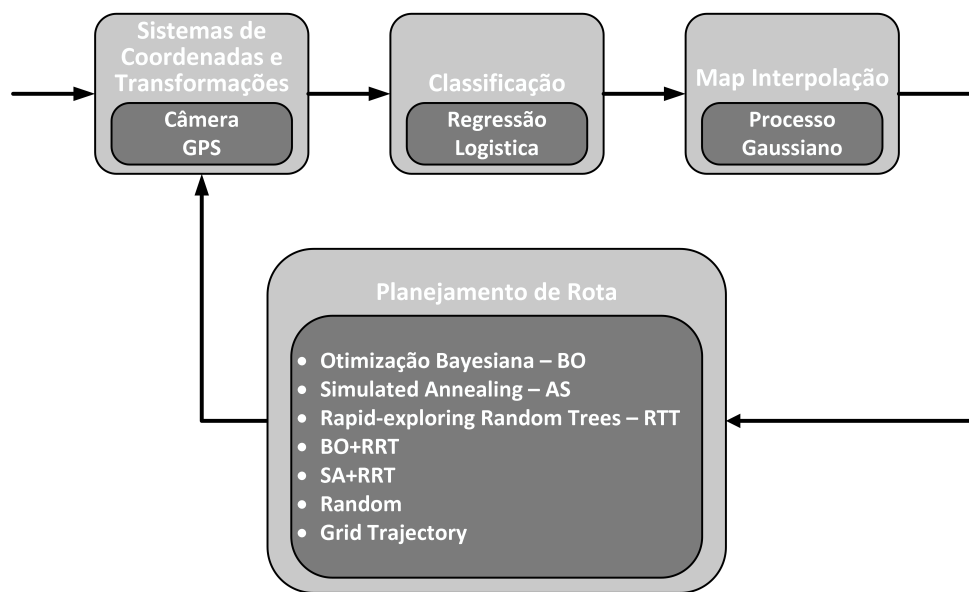
Fonte: Adaptado de Popović et al. (2020).

Já a arquitetura proposta por Vivaldini et al. (2018), Figura 3.3, possui os módulos de **Sistemas de Coordenadas e Transformações**, de **Classificação**, de **Interpolação do Mapa** e de **Planejamento de Rota**.

O módulo de **Sistemas de Coordenadas e Transformações** faz uso da câmera e do GPS. O módulo de **Classificação** trata do problema de classificação do solo e dos eucaliptos saudáveis

e doentes. O módulo de **Interpolação do Mapa** usa o *Processo Gaussiano* e o módulo de **Planejamento de Rota** faz uso dos algoritmos de planejamento de trajetória que são propostos, desenvolvidos e posteriormente testados. Assim essa arquitetura combina percepção, representação do ambiente e planejamento de rota. Vale ressaltar que um ponto a se considerar, seria mudanças na altitude durante o voo, onde é possível gerar mapas de resolução múltipla, pois com resoluções mais baixas é possível produzir um campo de visão maior para levantamentos iniciais mais rápidos. Entretanto, para resoluções mais altas pode-se produzir representações mais detalhadas para melhor precisão na busca pelo alvo da missão.

**FIGURA 3.3 – Framework para Planejamento de Rota para Classificação Ativa**

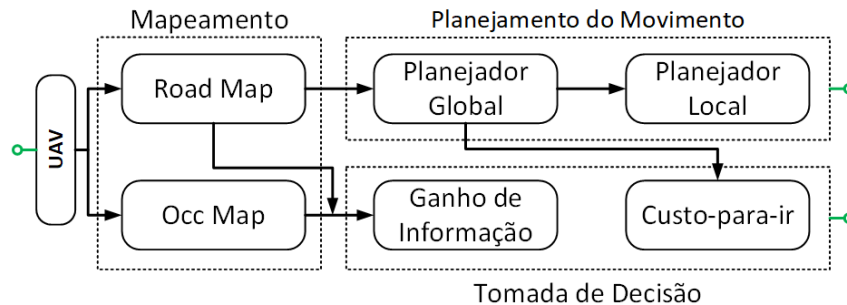


Fonte: Adaptado de Vivaldini et al. (2018)

A arquitetura desenvolvida por Wang et al. (2020), Figura 3.4, possui as etapas de **Mapeamento** na qual o *3D Occupancy Grid Mapping* fornece a construção do modelo do ambiente durante a exploração. O módulo de construção do **Road Map** tem a tarefa de construir um mapa topológico de forma incremental junto com o processo de exploração. Para determinar um NBV, tanto o ganho de informação quanto o custo inicial são levados em consideração. Assim, o caminho obtido no mapa é considerado na arquitetura um planejador global. No planejador local foi usado o algoritmo de *Potencial Fields* para fechar a lacuna entre o plano global e o controle do veículo para ajudar a coletar mais informações ao longo da rota de navegação. Apesar de ser considerado o ganho de informação e o custo inicial para determinar o NBV, tem-se nesses casos o que alguns autores da literatura consideram como algoritmos míopes, pois definem a melhor ação como aquela com maior recompensa imediata. Ao deixar de levar em

conta o futuro de recompensas potenciais, algoritmos míopes muitas vezes produzem resultados abaixo do comportamento ideal (MORERE; MARCHANT; RAMOS, 2017).

**FIGURA 3.4 – Framework para Exploração Autônoma**



Fonte: Adaptado de Wang et al. (2020)

### 3.8 Considerações Finais

De acordo com os trabalhos analisados da literatura percebe-se que os problemas IPP, Planejamento de Rota e Exploração Autônoma apresentados, tratam o planejamento de trajetória em ambientes desconhecidos como um problema de tomada de decisão autônoma do robô aéreo. Isso é feito para descobrir qual a rota a ser percorrida a fim de coletar as informações referentes ao ambiente.

Entretanto, percebe-se algumas diferenças no enfoque dado. O Planejamento de trajetória Informativa em ambientes desconhecidos tem como objetivo principal a coleta de informações com interesse em possíveis alvos. O planejamento de Rota em ambientes desconhecidos, em alguns casos, é semelhante ao IPP. Já a Exploração Autônoma possui o enfoque na construção do mapa, ou seja, produzir uma representação consistente de um ambiente desconhecido.

Vale destacar que os três problemas consideram que a trajetória que o UAV deve realizar em relação a missão deve ser feita de forma a percorrer o máximo de área possível. Isso é feito de modo a maximizar os ganhos de informações coletadas sobre um ambiente desconhecido ao passo em que se minimiza as incertezas do modelo. Para isso, são usadas diferentes restrições, tais como, tempo máximo de monitoramento, bateria, distância, restrições relativas a colisões, suavização de curvas dentre outras.

Em relação as abordagens apresentadas, embora sejam diferentes, os métodos usados para solução dos problemas IPP, Planejamento de Rota e a Exploração Autônoma são similares.

Foi possível perceber que BO-GP vêm sendo usado para coletar dados e para a estrutura do mapa do ambiente. O BO-POMDP vêm sendo usado para tomada de decisão sequencial sob incertezas em contexto de IPP obtendo resultados significativos.

As arquiteturas foram apresentadas com destaque para os módulos de cada uma e o funcionamento dos mesmos em relação ao métodos utilizados em cada etapa.

No presente trabalho, a ideia é desenvolver uma abordagem BO-POMDP em problemas de Planejamento de Trajetória Explorativa e Informativa. Pretende-se usar a arquitetura, proposta em Popović et al. (2020), no entanto a abordagem proposta é Não-Miope, pois podem calcular com eficiência a melhor ação, verificando vários passos à frente obtendo, assim, maior chance de soluções ótimas globais. Outro ponto a se considerar é que diferentes dos autores Morere, Marchant e Ramos (2017), Blanchard e Sapsis (2022) neste trabalho será considerado para realizar o planejamento da trajetória em diferentes altitudes, pois em situações onde onde é necessário voar em altitudes mais elevadas pode-se obter uma visão mais ampla sobre o ambiente, enquanto que em outras situações, voar em altitudes mais baixas pode ser necessário para coletar dados de forma mais precisa e eficaz. Essa consideração proporcionaria uma compreensão mais abrangente e precisa do ambiente. Para mais detalhes da proposta serão abordados no Capítulo 4.



# Capítulo 4

## PROPOSTA

---

---

Neste capítulo é apresentada a abordagem proposta intitulada de Planejamento de Trajetória Explorativa e Informativa (*Explorative Informative Path Planning* - EIPP). Os módulos que compõem a arquitetura e os respectivos métodos que compõem a abordagem BO-POMDP (um sistema para a tomada de decisão sequencial sob incertezas) são detalhados. Também será apresentada a plataforma adotada para a realização dos testes e as métricas que serão usadas para avaliação da abordagem BO-POMDP. Por fim, serão abordadas as considerações finais à respeito da proposta e o que se espera com a execução da mesma.

### **4.1 Planejamento de Trajetória Explorativa e Informativa - EIPP**

O foco deste trabalho está na proposta de uma estratégia eficiente de coleta de dados para realização de trajetórias contínuas usando um robô aéreo em ambientes desconhecidos. Para isso, é necessário que o UAV visite uma determinada área, explore o ambiente e tome uma decisão de qual trajetória seguir para maximizar a área de busca visando a uma região de interesse/alvo e minimizando as incertezas sobre o ambiente. Portanto, para que o UAV obtenha as informações sobre o ambiente e planeje as trajetórias a serem realizadas, será feito a exploração no espaço contínuo 3D para realizar uma tomada de decisão que envolva o planejamento de trajetória explorativa e informativa.

Diante disso, será apresentada a arquitetura modular para o Planejamento de Trajetória Explorativa e Informativa usando a abordagem BO-POMDP para uma tomada de decisão sequencial sob incertezas. Essa arquitetura proposta EIPP foi baseada no framework para IPP de (POPOVIĆ et al., 2020) (Ver Capítulo 3 Figura 3.2).

É importante destacar que uma das diferenças desse framework para o presente trabalho é que a abordagem proposta por (POPOVIĆ et al., 2020) é considerada míope, definem a melhor ação como aquela de maior recompensa imediata e não consideram o planejamento futuro das recompensas podendo aumentar a chance de se obter soluções que estejam em ótimos locais. Em contraste, neste trabalho adota-se a abordagem não-míope, a qual possui um planejamento mais estratégico, pois os planejadores baseados em POMDP podem calcular com eficiência a melhor ação, verificando vários passos à frente obtendo, assim, maior chance de soluções ótimas globais.

O abordagem BO-POMDP não-míope é baseada no trabalho de Morere, Marchant e Ramos (2017), um sistema para tomada de decisão sequencial sob incertezas. No entanto, um diferencial para esta abordagem é o planejamento da trajetória em diferentes altitudes para a tomada de decisão. Em situações onde é necessário voar em altitudes mais elevadas pode-se obter uma visão mais ampla sobre o ambiente, enquanto que em outras situações, voar em altitudes mais baixas pode ser necessário para coletar dados de forma mais precisa e eficaz .

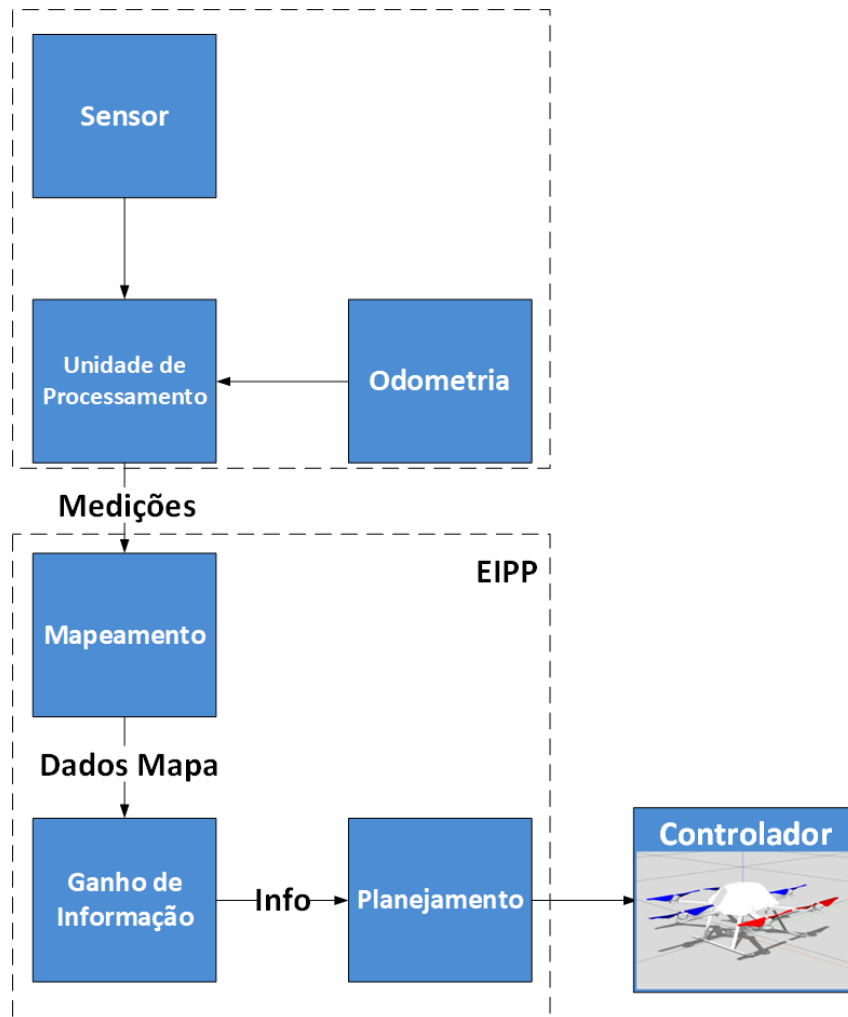
Entretanto, realizar um planejamento de trajetória em 3D pode ser computacionalmente custosos para abordagens BO-POMDP. Assim, é proposto, nesta . Isto é, após estabelecer a taxa de exploração/exploração os pontos mais promissores são selecionados para a inicialização do algoritmo. Essa estratégia pode acelerar a convergência do algoritmo.

Já em relação ao framework, o funcionamento do mesmo acontece da seguinte forma: o mapa do ambiente é construído usando medições extraídas de um fluxo de dados de um sensor. Em um determinado momento, o conhecimento do mapa é usado pela unidade de planejamento para encontrar as trajetórias mais úteis para a coleta de dados. O framework em questão possui os módulos de **Mapeamento**, **Ganho de Informação** e **Planejamento**.

Na Figura 4.1, a entrada de dados para o framework EIPP é obtida através de dados obtidos dos sensores, e processadas no módulo **Mapeamento** para construção do mapa. A partir dos dados obtidos pelo **Mapeamento**, o módulo **Ganho de Informação** gera informações que serão utilizadas no módulo “Planejamento” ou sistema de decisão. O módulo “Planejamento” realiza uma ação no ambiente definindo a trajetória que o UAV irá seguir. Ao observar o ambiente, o módulo BO-POMDP atualiza a crença e usa uma “Seleção de Política” para definir o próximo estado. Este próximo estado será utilizado como entrada para o módulo “Planejamento”, realizando assim uma nova ação no ambiente. O processo é repetido até que a missão seja concluída. Na fase EIPP existe o Processo Gaussiano que irá ser usado no modelo BO. Segundo Krause, Singh e Guestrin (2008), esta técnica tem sido amplamente utilizada no contexto do IPP em diferentes aplicações envolvendo UAVs para coleta de dados em ambientes desconhecidos. Essa

estrutura permite o uso de diferentes funções do kernel para expressar relacionamentos de dados dentro do ambiente. Neste trabalho será usado o Matérn 3/2 kernel (JADIDI, 2017; POPOVIĆ et al., 2020).

**FIGURA 4.1 – Arquitetura Proposta para o Planejamento de Trajetória Explorativa e Informativa**



Fonte: elaborada pela autora

### 4.1.1 Processo Gaussiano - GP

Para o funcionamento do Processo Gaussiano as Equações 2.5 e 2.6 do Capítulo 2 são aplicadas para inferir a distribuição do ambiente nos pontos de consulta de  $X_*$ . O cálculo é realizado repetidamente, para mapeamento e planejamento online, à medida que as medições são acumuladas no conjunto de dados de treinamento  $(x_i, y_i) | i = 1, \dots, n$  durante uma missão. Isso corresponde à maneira como os GPs são aplicadas para resolver problemas de regressão convencionais. Na configuração de monitoramento de ambiente, no entanto, os hiperparâmetros

aprendidos e um mapa de um ambiente anterior correlacionado  $p(f|X)$  são aproveitados para unir recursivamente os novos dados do sensor à medida que são obtidos.

Usando essa abordagem descrita acima baseada no trabalho de (POPOVIĆ et al., 2020) a Equação 2.6 é aplicada apenas uma vez, no início do procedimento, para capturar as correlações espaciais dentro da distribuição.

Além disso, o conjunto de dados de treinamento não aumenta de tamanho, permitindo tempo constante nas atualizações no que diz respeito ao número de medições do ambiente. Para isso, os valores nos pontos de treinamento de entrada  $X$  precisam ser inicializados. Para ambientes desconhecidos, podem simplesmente ser definidos para uma média uniforme constante.

Dada uma média uniforme e as correlações espaciais capturadas pela Equação 2.6, o mapa  $p(f|X) \sim GP(\mu^-, P^-)$  é usado como um prior para incorporar novas medições de sensor.

Seja  $z = [z_1 \dots z_m]$  novas medições independentes recebidas nos pontos  $[x_1 \dots x_m]^T \subset X$  no ambiente. Cada medição é modelado assumindo um ruído gaussiano como  $p(z_i|f_i, x_i) \sim N(\mu_{s,i}, \sigma_{s,i})$  com  $\sigma^{s,i}$  sendo a variância que pode ser constante ou não.

Para unir as medições em  $z$  com o mapa anterior  $p(f|X)$  é utilizado o estimador máximo a posteriori. Formulado da seguinte forma:

$$\operatorname{argmax} p(f|z, X) \quad (4.1)$$

As equações de atualização do Filtro de Kalman (FK) são aplicadas diretamente para calcular a densidade posterior  $p(f|z, X) \propto p(z|f, X) \times p(f|X) \sim GP(\mu^+, P^+)$

$$\mu^+ = \mu^- + Kv \quad (4.2)$$

$$P^+ = P^- - KHP^- \quad (4.3)$$

onde  $K = P^- H^T S^{-1}$  é o ganho de Kalman e  $v = z - H\mu^-$  e  $S = HP^- H^T + R$  são os termos de medição de covariância.  $R$  é uma matriz diagonal  $m \times m$  de variâncias  $\sigma_{s,i}$  dependentes da altitude associado a cada medição ruidosa  $z_i$ , e  $H$  é uma matriz  $m \times m$  dado um modelo de medição linear que seleciona intrinsecamente parte do estado latente  $[f_1 \dots f_m]$  observado através de  $z$ .

As atualizações do mapa nas Equações 4.2 e 4.3 são repetidas sequencialmente, sempre que novos dados são registrados, e podem ser realizadas em tempo constante. Observe que, como todos os modelos são lineares neste caso, a atualização KF acima produz a solução posterior ótima. Finalmente, um benefício importante desta abordagem é que ela é independente do tipo de sensor utilizado, pois permite unir dados heterogêneos em uma única representação do mapa de campo ambiental.

### 4.1.2 Formulação BO-POMDP

A ideia de usar BO-POMDP (*Bayesian Optimization with Partially Observable Markov Decision Process*) é resolver o problema de encontrar a melhor sequência de trajetórias ao longo do caminho enquanto coleta as informações. O valor de uma trajetória depende da quantidade de informações coletadas e do quanto essas informações reduzem a incerteza e a crença global no fenômeno monitorado. Este valor é fornecido pela função de recompensa que será a responsável pela avaliação do ganho de informação. No caso de métodos míopes, em que o planejador olha apenas um passo à frente, a métrica é calculada para uma única trajetória (HITZ et al., 2017; POPOVIĆ et al., 2020; BLANCHARD; SAPSIS, 2022). No entanto, no caso não míope, o valor de uma trajetória deve assumir todas as amostras futuras esperadas. Idealmente, o valor de uma trajetória deveria ser calculado com todas as amostras futuras ao longo de um horizonte infinito. Na prática, no entanto, o horizonte (ou olhar para frente) precisa ser finito por questões computacionais.

Diante disso, a formulação do POMDP a seguir é baseada nos trabalhos (MARCHANT; RAMOS, 2014; MORERE; MARCHANT; RAMOS, 2017). Usar um POMDP ao invés de um processo de decisão de Markov simples nos permite codificar o fato do robô executar ações apenas com base em observações ruidosas da função objetivo  $f$ . O estado do sistema  $\{f, p\}$  é completamente descrito por  $f$  e pelo pose do robô  $p$ .

Vale destacar que em situações reais nem sempre é dado o pose  $p$  do robô. No entanto, o interesse não está em aprender as transições do sistema, a função de transição  $T$  funciona de forma que o robô possa simular sequências de ações. Da mesma forma, o robô também recebe uma função de recompensa aproximada  $R(\{b(f), p\}, a)$  com base em sua crença atual  $b(f)$  em vez de  $f$ . A função de recompensa aproximada é baseada na Equação 2.14, fazendo o *trade-off* entre *exploration* e *exploitation*.

Assim, segue a definição da formulação POMDP:

- $S$ : O estado é uma tupla  $\{f, p\}$ , onde  $f$  é a função objetivo e  $p$  é o pose o robô.

- $A$ : As ações são definidas pelo parâmetro  $\theta$ , cada par de parâmetro de ação e pose definem totalmente uma trajetória  $T(\theta, p)$ .
- $T$ : A função de transição modela a probabilidade  $T(\{f, p\}, \theta, \{f', p'\})$  resultando no estado  $\{f', p'\}$  dada uma trajetória  $T(\theta, p)$  que foi tomada no estado  $\{f, p\}$ . Assume-se que a função de transição é determinística, independente da função objetivo  $f$  porque nenhuma transição afeta  $f$ , nem  $f$  afeta as transições.
- $R$ : A função de recompensa calcula a soma das recompensas obtidas ao longo da trajetória  $T(\theta, p)$ , e é definida como na Equação 4.4, onde o  $custo(T(o, p))$  é a aplicação específica do custo ao mover-se  $T(\theta, p)$ .
- $\Omega$ : As observações são avaliações ruidosas de  $f$  ao longo das trajetórias.
- $O$ : Em cada local, a função de observação  $O$  é definida exclusivamente pela pose do robô.

$$\bar{R}(\{b(f), p\}, \theta) = R(\theta, p, b(f)) * custo(T(\Theta, p)) \quad (4.4)$$

Vale destacar que o pose do robô é  $p = (x, y, z)$  para que as trajetórias selecionadas contêm mudanças de altitude para realizar a exploração no espaço contínuo 3D gerando trajetórias que maximizem o ganho de informação com a localização de possíveis alvos/regiões de interesse.

A seguir é apresentada a função aquisição  $h(x)$  que será usada neste trabalho. A Equação 4.5 foi proposta a primeira vez por Vivaldini et al. (2018) para o problema de IPP, onde uma distribuição Gaussiana com média  $\mu$ , amplitude  $\sigma_v$  e desvio padrão  $\sigma_l$  são empregados. Vale ressaltar que no contexto do trabalho os autores lidaram com a minimização da incerteza em relação a exploração do ambiente durante o processo de otimização, o sinal negativo foi usado para inverter a distribuição gaussiana. No entanto, no presente trabalho a função de aquisição será usada para a maximização no processo de otimização. Neste sentido, será usado a amplitude  $\sigma_v$  positiva.

$$h(x) = \sigma_v^2 * \exp\left(\frac{1}{2} \left(\frac{\mu(x) - 0,5}{\sigma_l(x)}\right)^2\right) \quad (4.5)$$

### 4.1.3 Função de Recompensa

A função de recompensa baseada na Equação 4.4 em conjunto com a função de aquisição serão responsáveis pela avaliação das trajetórias que serão selecionadas. Essa avaliação visa o ganho de informação referente a trajetória.

Reescrevendo, a nova função de Recompensa baseada na Equação 4.5 é dada da seguinte forma:

$$R(\{b(f), p\}, \theta) = \sigma_v^2 * \exp\left(\frac{1}{2} \left(\frac{\mu(b(f)) - 0,5}{(\sigma_l(b(f)))}\right)^2\right) * \text{ganho}/\text{custo}(T(\Theta, p)) \quad (4.6)$$

onde o custo de executar a trajetória é dado pela avaliação da distância entre os pontos avaliados. Assim, tem-se  $\frac{\text{ganho}}{\text{custo}}$ , o ganho é calculado pelas Equações 4.2 e 4.3 e a distância é calculado pela distância euclidiana entre os pontos.

Resolver um POMDP é equivalente a encontrar a política ótima  $\pi^*$ . A política ótima é definida como aquela que maximiza a soma infinita esperada de recompensas  $R$  a partir do estado de crença  $b_o$ . Segue formalmente na Equação 4.7.

$$\pi^* = \arg \max_{\theta \in A} \sum_{t=0}^n \gamma^t R(\theta, p, b(f)) \quad (4.7)$$

### 4.1.4 Algoritmo proposto para BO-POMDP

Nos últimos anos foram propostos diferentes métodos para resolver o POMDPs, como a *Value iteration* (KURNIAWATI; YADAV, 2016), *Branch-and-Bound* (BROCK; TRINKLE; RAMOS, 2009) e *Monte-Carlo Tree Search* (SILVER; VENESS, 2010; MARCHANT; RAMOS, 2014; MORERE; MARCHANT; RAMOS, 2017). Esses métodos podem ser usados para encontrar a política ideal  $\pi^*$ , concedendo a soma máxima de recompensas com desconto como definido na Equação 4.7. Na prática, precisamos aproximar a soma infinita por uma finita, portanto, restringindo  $\pi^*$  a ser uma política ótima em um horizonte de  $n$  ações. Neste trabalho é considerado  $n = 3$  já que aumentar o horizonte de ações aumenta o tempo computacional viável para resolver o problema.

O algoritmo proposto neste trabalho é baseado em Busca em Profundidade (*Depth-First Search - DFS*). O algoritmo começa em um nó raiz e explora cada um dos seus ramos (no caso do presente trabalho, até a camada 3), antes de retroceder(*backtracking*) calculando a função de

recompensa para avaliar.

---

**Algoritmo 4.1: Algoritmo BO-POMDP**


---

```

1 Função Grid_Search (no_Inicial, grid_Map)
2   Enquanto (exploração não for finalizada) Faça
3     nivel_Arvore_0 = 0; nivel_Arvore_1 = 1;
4     nivel_Arvore_2 = 2; nivel_Arvore_3 = 3;
5     lista_Pontos_Nivel_1 = Encontra_Pontos(no_Inicial, grid_Map);
6     Para_Todo (no_a em lista_Pontos_Nivel_1) Faça
7       Atualiza_Grid (grid_Map);
8       lista_Pontos_Nivel_2 = Encontra_Pontos (no_a, grid_Map);
9       Para_Todo (no_b da lista_Pontos_Nivel_2 que ainda não foi expandido) Faça
10        Atualiza_Grid (grid_Map);
11        lista_Pontos_Nivel_3 = Encontra_Pontos(no_b, grid_Map);
12      Fim Para_Todo
13    Fim Para_Todo
14    maior_Recompensa_Acumulada = Maior_Recompensa(lista_Pontos_Nivel_3);
15    Para_Todo (nivel de nivel_Arvore_1 até nivel_Arvore_3) Faça
16      Se (nivel == 1) Então
17        Executa_Trajectoria (maior_Recompensa_Acumulada.coordenada_nivel_1);
18      Fim Se
19      Se (nivel == 2) Então
20        Executa_Trajectoria (maior_Recompensa_Acumulada.coordenada_nivel_2);
21      Fim Se
22      Se (nivel == 3) Então
23        Executa_Trajectoria (maior_Recompensa_Acumulada.coordenada_nivel_3);
24      Fim Se
25    Fim Para_Todo
26  Fim Enquanto
27 Fim Grid_Serach

```

---



---

**Algoritmo 4.2: Encontra\_Pontos**


---

```

1 Função Encontra_Pontos (no_raiz, grid_Map)
2   lista_Pontos = Gera_Pontos(no_raiz);
3   Para_Todo (ponto de lista_Pontos) Faça
4     Avaliar_Ponto (ponto); //equação 4.6
5   Fim Para_Todo
6   lista_Pontos_Oderanada = Ordena_Por_Valor_Recompensa(lista_Pontos);
7   Numero_Nos = lista_Pontos_Oderanada.tamanho * taxa_Exploração_Exploração
8   Para_todo (indice_lista de 0 até Numero_Nos-1) Faça
9     lista_Nos_Returno.adiciona(lista_Pontos_Oderanada[indice_lista]);
10  Fim ParaTodos
11  Retorna lista_Nos_Returno;
12 Fim Encontra_Pontos

```

---



O Algoritmo de busca em árvore para resolver BO-POMDP é usado para fazer a busca de novos nós dentro do espaço de ação do UAV para que ele possa realizar o monitoramento do ambiente. O Algoritmo é iniciado com um laço que vai possibilitar que o algoritmo seja repetido até que os pontos sejam avaliados.

Na linha 1, o algoritmo prossegue com a função GridSearch recebendo um nó inicial (nó no qual o UAV se encontra no momento) e um gridMap (que é o mapa do momento pela perspectiva do UAV). Em seguida, as linhas 3 e 4 são definidas variáveis que representam os 4 níveis da árvore de busca tem seus valores atribuídos. Essa árvore de busca guarda, em cada um dos seus nós, informações sobre o ambiente que serão úteis quando ocorrer a definição de qual trajetória o UAV irá percorrer. O nível 0 é o nível onde o UAV se encontra no momento do início do processo. Os outros níveis, representam os níveis da árvore de busca que, na medida que o processamento para tomada de decisão da trajetória prossegue, tem seus nós expandidos para o próximo nível até que se alcance uma profundidade igual a 3.

Em seguida, na linha 5, a função EncontraPontos é chamada e recebe o nó inicial, o gridMap atual (gridMap correspondente à posição do UAV no momento) e o nível da árvore em que o nó inicial se encontra (nível 0). Essa função retornará uma lista de pontos que juntos formam os nós do nível 0 (mais detalhes sobre essa função serão abordados mais à frente neste tópico). Essa lista de pontos retornada pela função EncontraPontos é armazenada na listaPontosNivel1, que guarda os pontos (nós) que formam o nível 1 da árvore de busca.

Com esses nós, seria possível já escolher o nó com maior valor de recompensa e enviar o UAV para aquela coordenada. Entretanto, deseja-se evitar que o algoritmo tenha um comportamento guloso e, sendo assim, os nós são armazenados para que mais níveis da árvore sejam explorados corroborando para que a tomada de decisão tenha a tendência de ser mais precisa ao final do processo.

Em seguida, na linha 6, um laço é iniciado para que seja possível percorrer toda a lista de pontos do nível 1, expandindo, posteriormente, cada um deles para o próximo nível. No entanto, antes que cada um dos pontos seja expandido, é necessário realizar a atualização do gridMap. Essa atualização (linha 7) se dá a partir do nó que se quer expandir (nó pai), simulando, portanto, a presença do UAV naquela coordenada e, assim, gerando os pontos do próximo nível (nível 2).

Na linha 8 a função Encontra Pontos é novamente chamada com o nó atual (nó pai), o gridMap deste nó e o nível em que o nó pai se encontra (nível 1 da árvore). A função EncontraPontos retorna uma lista de nós que forma parte dos nós do nível 2 da árvore de busca. Os nós do nível 1 vão sendo expandidos passo a passo até formarem os nós do nível 2 da árvore de busca.

Como os nós armazenam, juntamente com o valor da própria recompensa, os valores das recompensas dos seus nós antecessores na árvore de busca, apenas com os nós do nível 3 é possível saber qual a melhor trajetória a ser seguida. Assim, a trajetória a ser percorrida pelo UAV, será aquela cuja soma de todas as recompensas obtidas pelo caminho seja a maior.

Desse modo, na linha 14, a função `MaiorRecompensa` recebe a lista de nós do nível 3 e retorna o nó do nível 3 que corresponde ao nó cuja recompensa acumulada é a maior. (considerando os 3 [três] níveis antecessores da árvore em questão).

Com esse nó escolhido, um laço é realizado para possibilitar a execução da trajetória pelo UAV partindo do nível 0, passando pelo nível 1, pelo nível 2 até chegar ao nível 3.

O Algoritmo `EncontraPontos` é usado para gerar as listas de pontos que formarão os nós de cada um dos níveis da árvore de busca. O algoritmo se inicia com a função recebendo o nó raiz( nó base para a geração dos demais nós daquela iteração, o `gridMap` do nó raiz e o nível da árvore em que a expansão se encontra. Uma função que gera os pontos é chamada e os pontos gerados são armazenados em uma lista.

Um laço é então criado na linha 3 para percorrer a lista e avaliar (de acordo com o valor de recompensa) cada um dos pontos gerados. A lista com os pontos avaliadas é ordenada e, com base na taxa de `ExploraçãoExplotação` definida anteriormente, um subconjunto dos nós com maior valor de recompensa é separada para ser enviada para a função `EncontraPontos` do algoritmo `GridSearch`.

Essa estratégia foi adotada, pois com a Taxa de `ExploraçãoExplotação` é possível gerar uma quantidade maior de pontos (Explorar o espaço de busca) e, posteriormente, com o subconjunto de pontos que serão retornados pela função definido, é possível usá-los para explorar determinadas regiões do espaço de busca.

#### 4.1.4.1 Exemplo do Funcionamento do Algoritmo

Na Figura 4.2 tem-se um exemplo de busca em profundidade em árvore com nível 3. O funcionamento do algoritmo proposto no presente trabalho dá-se da seguinte forma. No exemplo da Figura, o algoritmo inicia-se no nó X (camada 0) tendo 3 pontos possíveis para a escolha. Essa escolha da-se calculando a função de recompensa e como neste trabalho tem-se o objetivo de maximizar o ganho de informação, então, a soma da recompensa de maior valor é selecionada.

No caso de métodos Míope, a recompensa de maior valor imediato é selecionada, neste caso, o valor de 20 (rosa) na camada 1. A busca continua na camada 2 e a recompensa de maior valor imediato é 50 (rosa) e na camada 3 o valor é 80 (rosa), então, considerando o

algoritmo Míope tem-se as sequencias dos nós dados da seguinte forma:  $X \rightarrow 20 \rightarrow 50 \rightarrow 80$ . É interessante notar que para o método Míope, neste exemplo, as quantidades dos pontos pontos que serão avaliados na camadas são sempre três.

Já no caso do presente trabalho que é considerado o método Não-Míope, a partir do nó  $X$  é calculada a função de recompensa para todos os pontos da camada 1. O algoritmo prossegue para a camada 2 e amplia as possibilidades de escolhas, ou seja, para cada ponto são selecionados mais três pontos totalizando 9 pontos e na camada 3 são avaliados 27 pontos, ou seja, tem-se um crescimento exponencial de possibilidades a cada camada. Na última camada é feito um backtracking para verificar qual é a maior soma das recompensas e assim, neste exemplo, têm a sequencia destacada em verde  $X \rightarrow 15 \rightarrow 69 \rightarrow 140$ .

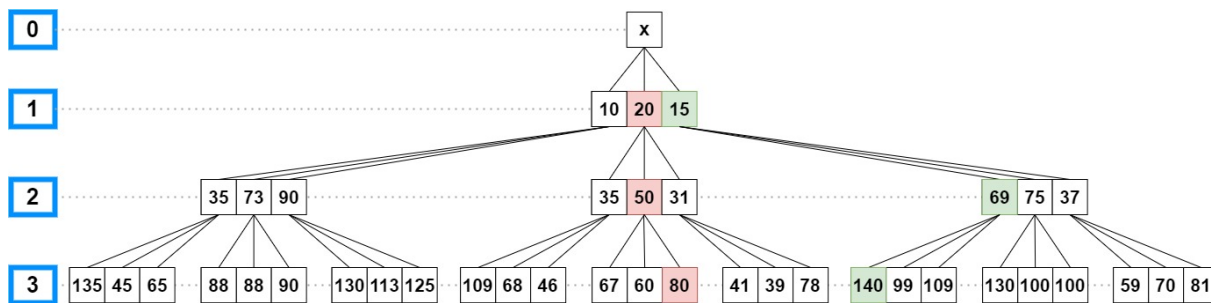
Note que os pontos selecionados pelo método Míope foram diferentes do método Não-Míope. O primeiro escolhe o caminho que, à primeira vista, é mais econômico, já o segundo possui um planejamento mais estratégico e explora mais o espaço de busca e pode ter maior chance de soluções ótimas globais, no entanto, o tempo computacional pode ser maior também.

Na Figura 4.3 tem-se um exemplo com 6 camadas na árvore para explorar o espaço de busca. Considere duas situações neste exemplo:

A primeira diz respeito a escolha anterior da Figura 4.2 com as sequencia da trajetória destacadas em verde, ou seja, o UAV já executou a trajetória selecionada  $X \rightarrow 15 \rightarrow 69 \rightarrow 140$ . A partir da camada 3 o nó com valor (140) é o nó raiz inicial, assim o algoritmo recomeça e todo o procedimento é repetido e a nova sequencia selecionada por intermédio do calculo da soma da função de recompensa é a seguinte  $140 \rightarrow 150 \rightarrow 240 \rightarrow 280$  destaque em laranja.

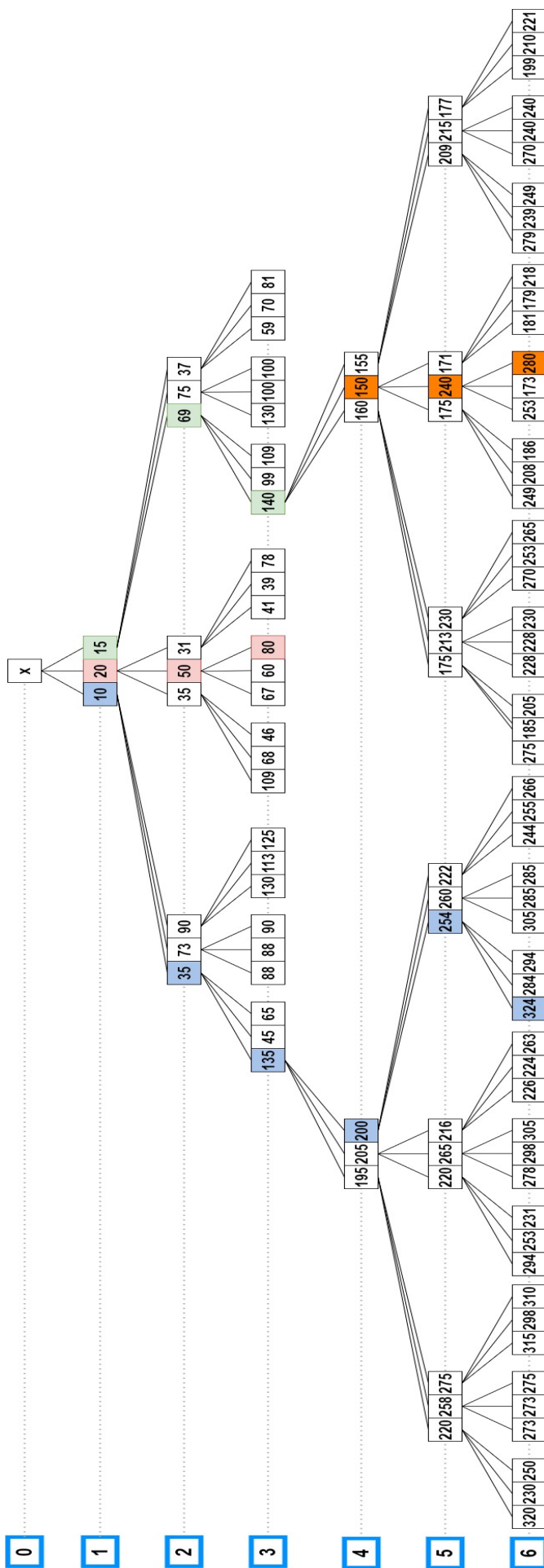
A segunda situação diz respeito a busca em profundidade na árvore com nível 6, ou seja, inicializando o algoritmo no nó  $X$  o procedimento é repetidos para todos os  $729$  pontos,  $pt^n$ , que neste exemplo representa a quantidade de pontos elevado pelo numero da camada  $3^6 = 729$ . A sequênciã em azul é a selecionada  $X \rightarrow 10 \rightarrow 35 \rightarrow 135 \rightarrow 200 \rightarrow 254 \rightarrow 32$ . Note que a sequênciã selecionada foi diferente da situação anterior. Isso pode ocorrer pelo fato de aumentarmos as possibilidades de escolhas, e assim, a tomada de decisão considera longas sequências de ações, aumentando o tamanho do espaço de busca e tornando o algoritmo computacionalmente custoso e tempo inviável, logo, o *trade-off* entre exploração e tempo computacional precisam ser levados em consideração.

**FIGURA 4.2 – Exemplo do funcionamento do algoritmo BO-POMDP - Árvore com 3 níveis**



Fonte: elaborada pela autora

**FIGURA 4.3 – Exemplo do funcionamento do algoritmo BO-POMDP - Árvore com 6 níveis**



#### 4.1.4.2 Seleção de Pontos

É de suma importância a definição dos pontos que serão inicializados na abordagem proposta BO-POMDP, pois se iniciarmos com muitos pontos, teríamos um número grande de possibilidades para a definição das trajetórias o que aumenta a exploração, porém teríamos um tempo computacional inviável para aplicações que envolvem monitoramento de ambientes desconhecidos com UAV, devido a restrições de tempo e também de bateria. Neste tipo de situação, selecionar os pontos considerando a taxa de exploração/exploração pode acelerar a convergência do algoritmo Não-Míope proposto neste trabalho e alcançar soluções em tempo computacional viável para a aplicação.

## 4.2 Ambiente de Simulação

O ROS (*Robot Operating System*) e o GAZEBO foram usados como protocolo de comunicação e plataforma de simulação 3D, respectivamente. Esta ferramenta tem sido amplamente adotada para o teste e avaliação do software do robô em várias missões de planejamento de trajetória informativa, exploração autônoma, entre outras aplicações usando UAV (WANG et al., 2020; POPOVIĆ et al., 2020; Zhang et al., 2015; MEYER et al., 2012; KANCIR; DIGUET; SEVAUX, 2019; BáčA et al., 2020).

O ROS surgiu em 2007, com objetivo de integrar com um *framework* as ideias de desenvolvimento dos robôs, permitindo assim, a reusabilidade de código. É um *meta-operating system robotic* que fornece abstração de *hardware*, controle de dispositivo de baixo nível, implementação de funcionalidade comumente usada, troca de mensagens entre processos e gerenciamento de pacotes (ROS.ORG, 2014). Tornou-se o protocolo de comunicação padrão em aplicações de robótica, devido à sua transparência e eficiência durante a transferência de dados entre diferentes módulos de *software* ou máquinas.

Além disso, o ROS é um *framework open-source* para desenvolvimento de software para robôs que contém diversas ferramentas, bibliotecas e convenções. Ele visa simplificar a tarefa de criar comportamentos complexos e robustos para robôs.

O Gazebo é uma iniciativa que começou a ser desenvolvida por Dr. Andrew Howard e Nate Koenig em 2002 na Universidade do Sul da Califórnia. O Gazebo surgiu da necessidade de simular robôs em ambiente externo sob diversas condições com um alto grau de fidelidade aos ambientes reais. Em 2009, o projeto do Gazebo foi integrado ao projeto do ROS e, desde então, o Gazebo vem sendo utilizado como ferramenta de simulação robótica na comunidade do ROS.

Hoje em dia ele é mantido pela *Open Source Robotics Foundation* (FOUNDATION, 2014) .

O Gazebo como simulador oferece a capacidade de simular populações de robôs em diversos ambientes com acurácia e eficiência permitindo, assim, o teste de algoritmos, o design de robôs, teste de regressão e treinar sistemas de inteligência artificial em cenários realísticos (KOENIG; HOWARD, 2004; AGUERO et al., 2015).

### 4.3 Unmanned Aerial Vehicle - UAV

Neste trabalho foi usado RotorS<sup>1</sup> simulador que é considerado um rotorcraft do tipo UAV e também chamado de *Micro Aerial Vehicle* (MAV) devido ao seu peso.

A estrutura de simulação RotorS foi desenvolvida para reduzir o tempo de teste de campo e separar problemas para teste, facilitando a depuração e, finalmente, reduzindo travamentos de MAVs reais. Além disso, todos os componentes encontrados em um MAV real são simulados pelos plugins do Gazebo e pelos mecanismos de física.

Com relação especificamente ao MAV, ele possui um número fixo de rotores, que podem ser colocados em locais especificados pelo usuário, e alguns sensores conectados a ele. Cada rotor tem dinâmica motora e é responsável pelos efeitos aerodinâmicos mais dominantes de acordo com (MARTIN; SALAUN, 2010). Podem ser acoplados diversos sensores, incluindo uma Unidade de Medição Inercial (IMU), um sensor de odometria genérico e um sensor inercial visual composto por uma câmera estéreo e uma IMU, além de sensores desenvolvidos pelo usuário. Para simular condições realistas, foi implementado modelos de ruído para os sensores aplicados. Todos os dados de simulação podem ser registrados usando rosbags, tendo acesso direto aos dados de simulação. Conseqüentemente, dados reais estão disponíveis para cada sensor, o que facilita a depuração e avaliação de novos algoritmos (FURRER et al., 2016).

Assim, para resolver tarefas de nível superior, como planejamento de trajetória, os MAVs simulados podem ser usados sem qualquer modificação adicional nos modelos. Além da simulação de MAVs, o framework também inclui um controlador de posição e um estimador de estado, que funcionam com os modelos fornecidos. Na Figura 4.4 tem-se um ambiente construído no Gazebo com o MAV.

---

<sup>1</sup>[https://github.com/ethz-asl/rotors\\_simulator/wiki](https://github.com/ethz-asl/rotors_simulator/wiki)

**FIGURA 4.4 – Simulador RotorS com o ambiente construído a partir de modelos padrões do Gazebo e um Firefly hex-rotor helicopter. (FURRER et al., 2016)**



Fonte: Retirado do artigo RotorS - A Modular Gazebo MAV Simulator Framework

## 4.4 Considerações Finais

Neste capítulo foi apresentada a proposta do trabalho intitulada de Planejamento de Trajetória Explorativa e Informativa. Foi descrito o funcionamento de cada módulo apresentando os métodos adotados em cada etapa. Também foi destacada a contribuição deste trabalho que está na formulação matemática da abordagem BO-POMDP para tomada de decisão sequencial sob incertezas em que são geradas trajetórias contínuas no espaço 3D para realizar o EIPP.

Também foi apresentada a plataforma adotada para a realização dos testes e a estratégia de usar um UAV para visitar e explorar uma área, analisando-a, em tempo real, para que seja possível definir a rota a ser seguida.

Por fim, espera-se que, com esse trabalho, seja possível, por intermédio dos experimentos, mostrar o desenvolvimento da arquitetura proposta e o desempenho da formulação BO-POMDP para uma tomada de decisão sequencial sob incertezas. Nesse processo tem-se a estratégias de explorar o ambiente no espaço contínuo 3D para gerar trajetórias que maximizem o ganho de informação com a localização de possíveis alvos.



# Capítulo 5

## EXPERIMENTOS E ANÁLISES DOS RESULTADOS

---

---

Para analisar o desempenho da proposta do trabalho, foram realizados experimentos envolvendo a abordagem não-Míope (proposta neste trabalho), a abordagem Míope e a abordagem Randômica. Os experimentos foram realizados em 3 (três) cenários. Durante a execução dos experimentos foram considerados os seguintes aspectos:

- Raiz do Erro médio quadrático em relação às previsões e os valores reais obtidos;
- A incerteza associada ao ambiente durante o decorrer do planejamento da trajetória;
- A distância percorrida pelo UAV;
- Tempo necessário para o UAV explorar o ambiente com o uso da estratégia de Planejamento de Trajetória Informativa e Explorativa;

A seguir, mais detalhes sobre os resultados obtidos a partir das execuções dos experimentos.

### 5.1 Cenários

Foram definidos 3 (três) cenários para as execuções dos experimentos. Esses cenários foram incluídos no simulador Gazebo de modo a possibilitar um ambiente virtual onde o UAV executasse o planejamento da trajetória explorativa informativa.

Os cenários possuíam, cada um deles, as dimensões de  $30m \times 30m \times 30m$  (largura x comprimento x altura). Em todas as execuções a posição inicial do UAV foi especificada como sendo de  $7.5m \times 7.5m \times 8.66m$ . Essa definição foi realizada para que as abordagens possuíssem as mesmas condições iniciais. Outro ponto a se destacar é que a velocidade e a aceleração máximas de referência foram de  $5m/s$  e  $2m/s^2$  respectivamente.

Em relação ao cenário 1, este possui apenas 1 região de interesse e foi incluído para possibilitar a avaliação das abordagens em situações com menos regiões de interesse conforme Figura 5.1a.

Já o cenário 2 contém 3 regiões de interesse que estão dispostas nas extremidades da base do ambiente conforme pode ser visto na Figura 5.1b.

O cenário 3 é o que possui mais regiões de interesse, com um total de 6. Essas regiões estão distribuídas pelo ambiente permitindo um planejamento que visite áreas distintas do cenário conforme pode ser visto na Figura 5.1c.

A representação do ambiente é baseada no Processo Gaussiano com área de  $30m \times 30m$  e resolução de  $0.75m$ . Isso faz com que se tenha uma matriz de  $40 \times 40$  com 1600 grids. Assumindo que a área é desconhecida à priori, o mapa é inicializado uniformemente com uma média (*mean*) informada à priori de 50%.

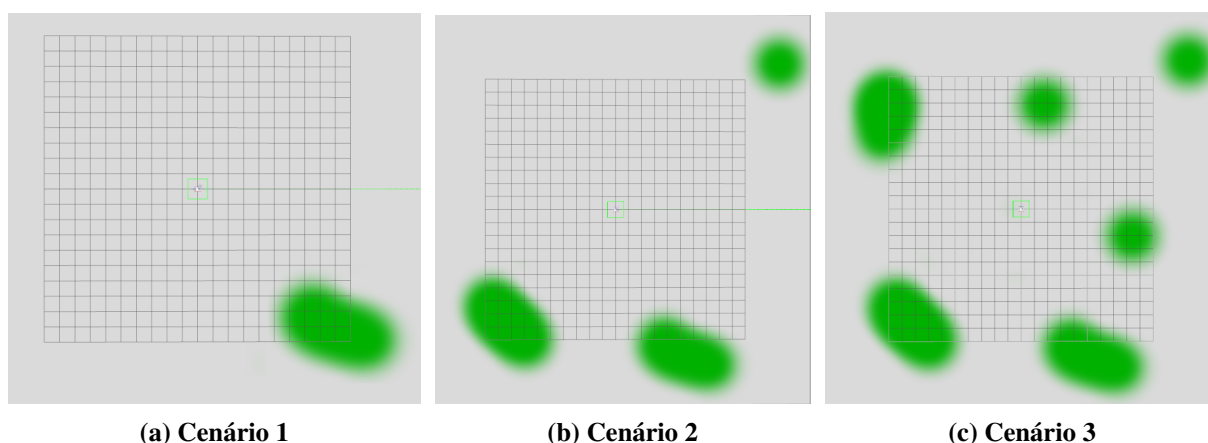


FIGURA 5.1 – Cenários usados no simulador Gazebo

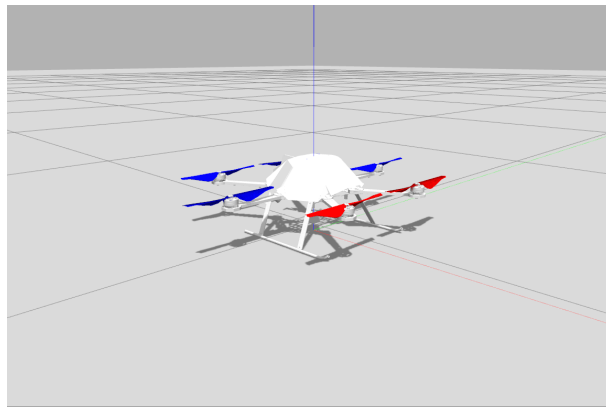
## 5.2 Abordagens

A abordagem proposta neste trabalho para o Planejamento de Trajetória Informativa e Explorativa é a abordagem BO-POMDP. Essa abordagem é considerada não-míope, pois pode ser usada para realizar a tomada de decisão com mais eficiência. Essa tomada de decisão se dá considerando mais passos à frente o que aumenta a chance de que sejam encontradas solução melhores ou até mesmo o ótimo global. Para verificar o desempenho do método proposto, serão realizados testes considerando outras abordagens como a míope e a randômica. A abordagem míope realiza a tomada de decisão a partir da escolha da melhor ação seguindo o critério de maior recompensa imediata. A abordagem randômica realiza a tomada de decisão a partir da definição de trajetórias de forma aleatória. As abordagens que serão testadas neste trabalho são:

- BO-POMDP com taxa de exploração/exploração de 15% usando a Equação 4.6;
- BO-POMDP com taxa de exploração/exploração de 30% usando a Equação 4.6;
- Míope baseado no trabalho de (POPOVIĆ et al., 2020) Equações 4.2 e 4.3;
- Randômico.

## 5.3 Experimentos

Os experimentos foram realizados em um computador Intel Core i7-7700 3.60GHz com 8 núcleos e 16GB de memória RAM. O modelo UAV é um AscTec Firefly equipado com uma câmera voltada para baixo, que possui uma resolução de imagem de 360 pixels  $\times$  480 pixels e um FoV ( $35.4^\circ$ ,  $47.2^\circ$ ) nas direções  $x$  e  $y$ , respectivamente, como mostra a Figura 5.2.



**FIGURA 5.2 – UAV no ambiente gazebo**

Foram realizados os experimentos relacionados a:

- Avaliações do erro usando a métrica RMSE (Raiz do Erro Médio Quadrático);
- Componente de avaliação de incerteza sobre o ambiente  $\text{Tr}(P)$  (*Trace of P*);
- Cálculo das distâncias percorridas pelo UAV;
- Cálculo do tempo para realização da exploração do ambiente;

A métricas RMSE e o componente de avaliação de incerteza  $\text{Tr}(P)$  foram usadas no trabalho para medir o desempenho dos algoritmos nos testes realizados. O RMSE mede o erro ao quadrado entre as previsões e os valores reais, enquanto que o  $\text{Tr}(P)$  avalia o trace da matriz de covariância do GP, ou seja, avalia a incerteza associada ao ambiente. Além disso a distância

percorrida e o tempo foram usados para determinar a eficiência das abordagens, uma vez que se busca realizar um planejamento de trajetória explorativa e informativa do ambiente visando percorrer a menor distância possível em um menor tempo de execução. Vale destacar que todos os algoritmos foram executados 30 vezes com 10 execuções realizadas em cada um dos cenários. Para calcular as métricas foram considerados valores médios das execuções. Cada um deles serão detalhados a seguir.

### 5.3.1 RMSE

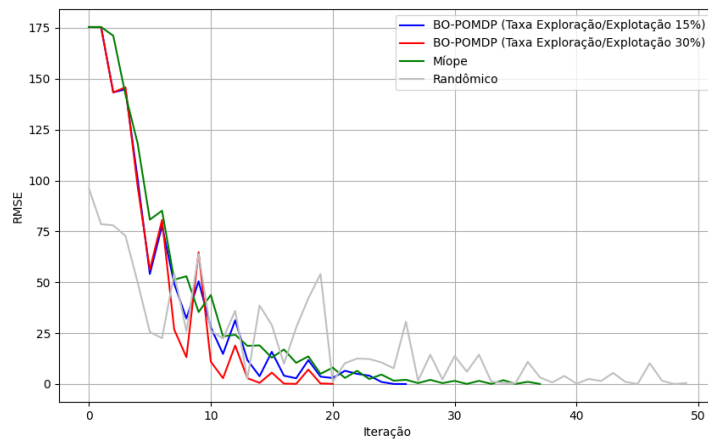
As execuções dos testes retornaram a média do RMSE de cada um dos 3 cenários. Os valores de RMSE obtidos foram equivalentes havendo pequenas variações como pode ser visto na Figura 5.3.

Pode-se perceber que o algoritmo BO-MOMDP que é não-miope, com taxa de exploração / exploração de 30% obteve menor erro em relação aos demais, o que possibilita inferir que o planejamento de trajetória explorativa e informativa foi mais assertivo obtendo um planejamento que maximiza os ganhos de informação em uma dada região de interesse. Vale destacar que o algoritmo em questão convergiu em menos iterações quando comparado com os demais, corroborando com a ideia de que algoritmos Não-Míopes podem calcular com eficiência a melhor ação verificando vários passos à frente e obtendo, assim, melhores soluções. Já os métodos míopes realizam um planejamento onde a melhor ação é definida seguindo o critério de recompensa imediata e não considerando o planejamento futuro o que aumenta o número de iterações e também aumenta as chances de se obter soluções que estejam em ótimos locais.

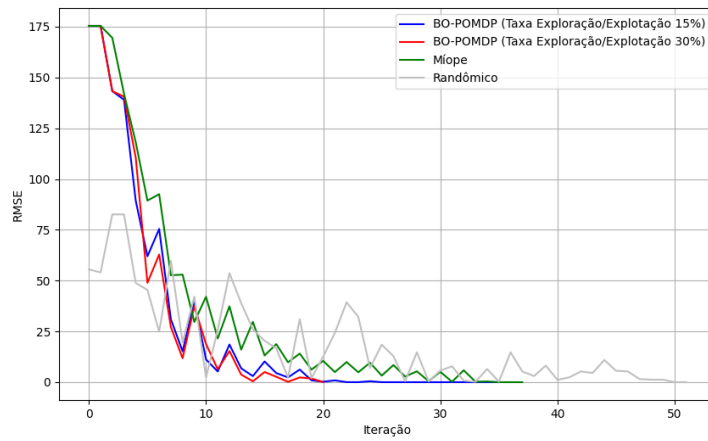
### 5.3.2 Trace of P - Tr(P)

Com a análise da métrica  $Tr(P)$  é demonstrado que a incerteza sobre o ambiente diminui no decorrer das iterações dos algoritmos. O algoritmo BO-POMDP não-Míope com taxa de exploração/exploração de 30% obteve melhores resultados em comparação aos algoritmos míope e randômico. Esse fato, corrobora a hipótese do trabalho de que identificar os pontos mais promissores em vários passos a frente e estabelecer uma taxa adequada de exploração/exploração pode acelerar a convergência do algoritmo, representando um elemento de relevância crucial dessa abordagem.

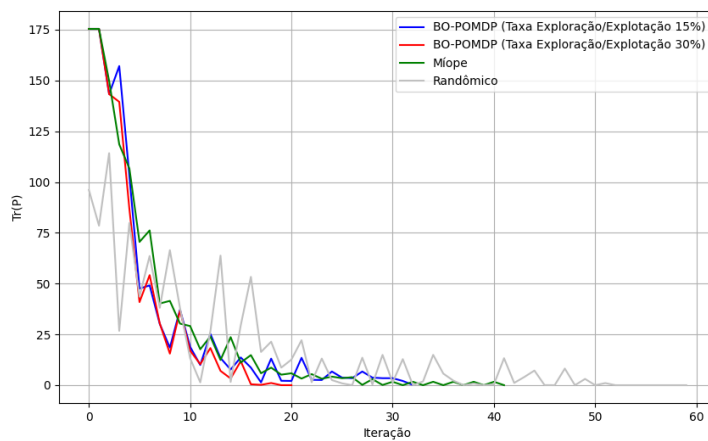
Nos 3 cenários (Ver Figura 5.4), o algoritmo proposto, tanto no modelo com taxa de exploração / exploração de 15% como no modelo com taxa de exploração / exploração de 30% obtiveram resultados de  $Tr(P)$  que indicaram menores incertezas sobre o ambiente do que as



(a) Cenário 1

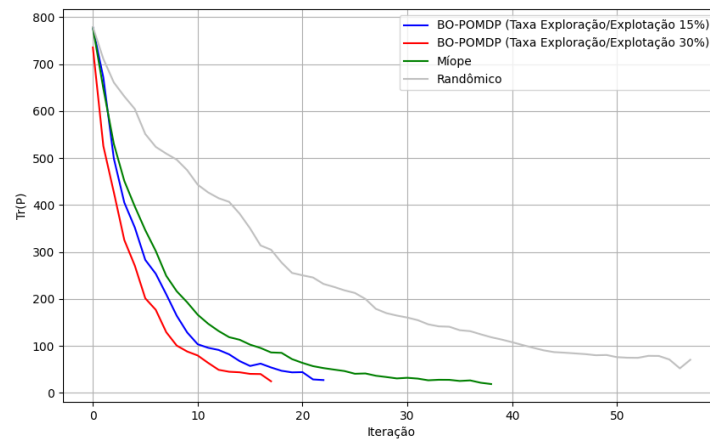


(b) Cenário 2

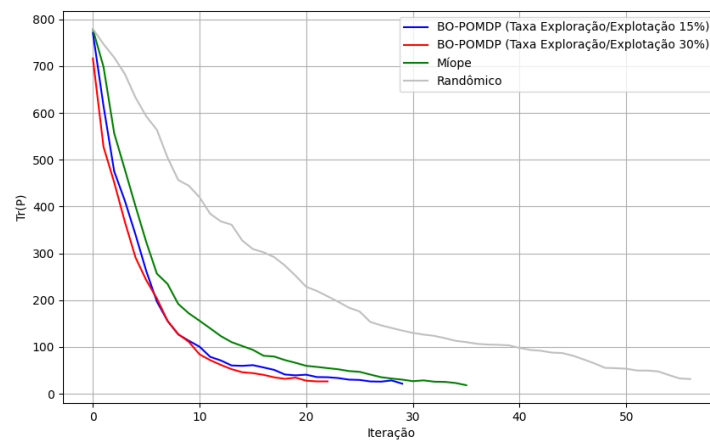


(c) Cenário 3

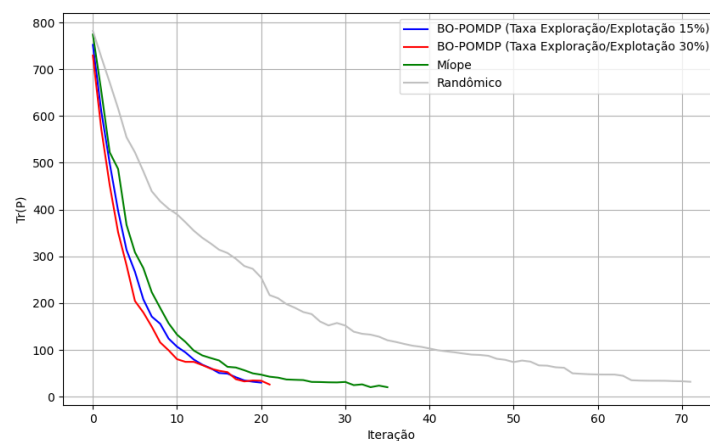
**FIGURA 5.3 – Testes realizados nos 3 cenários para RMSE**



(a) Cenário 1



(b) Cenário 2



(c) Cenário 3

**FIGURA 5.4 – Testes realizados nos 3 cenários para  $Tr(P)$**

abordagens míope e randômico. No cenário 1 (uma região de interesse) a abordagem BO-POMDP 30% convergiu em menos interações na média entre as execuções e, assim, obteve menor incerteza sobre o ambiente. Já no cenário 3 (6 regiões de interesse) a convergência dos algoritmos BO-POMDP 15% e POMDP 30% foram equivalentes inclusive no número de iterações médias necessárias para finalização da exploração do ambiente. Como esperado, o algoritmo Randômico obteve, nos 3 cenários, uma maior incerteza sobre o ambiente o que provoca, conseqüentemente, uma maior quantidade de iterações necessárias para realização da exploração.

### 5.3.3 Distância

As distâncias calculadas a partir das trajetórias percorridas pelo UAV nas execuções foram comparadas para as 4 abordagens (BO-POMDP 15%, BO-POMDP 30%, Míope e Randômica).

**Tabela 5.1: Distâncias obtidas em cada execução dos testes no Cenário 1**

| Cenário       | Teste    | BO-PMDP                   | BO-PMDP                   | Míope          | Randômico      |
|---------------|----------|---------------------------|---------------------------|----------------|----------------|
|               |          | Exploração/Explotação 15% | Exploração/Explotação 30% |                |                |
| Cenário 1     | Teste 01 | 233.399                   | 277.865                   | 479.608        | 919.801        |
|               | Teste 02 | 213.326                   | <b>138.486</b>            | 460.556        | <b>900.942</b> |
|               | Teste 03 | 241.968                   | 268.681                   | 540.006        | 905.692        |
|               | Teste 04 | 383.166                   | 233.468                   | 542.449        | 1032.664       |
|               | Teste 05 | 403.728                   | 288.479                   | 534.411        | 969.330        |
|               | Teste 06 | 379.803                   | 309.783                   | <b>333.385</b> | 908.427        |
|               | Teste 07 | <b>195.818</b>            | 279.048                   | 447.942        | 913.765        |
|               | Teste 08 | 366.921                   | 289.426                   | 434.229        | 973.709        |
|               | Teste 09 | 413.153                   | 219.579                   | 346.421        | 1034.071       |
|               | Teste 10 | 339.516                   | 260.600                   | 386.312        | 930.682        |
| Média         |          | 317.080                   | <b>256.541</b>            | 450.532        | 948.908        |
| Desvio Padrão |          | 85.726                    | 49.325                    | 77.025         | 51.122         |

**Tabela 5.2: Distâncias obtidas em cada execução dos testes no Cenário 2**

| Cenário       | Teste    | BO-PMDP                   | BO-PMDP                   | Míope          | Randômico      |
|---------------|----------|---------------------------|---------------------------|----------------|----------------|
|               |          | Exploração/Explotação 15% | Exploração/Explotação 30% |                |                |
| Cenário 2     | Teste 01 | 442.451                   | 326.358                   | 448.346        | 940.306        |
|               | Teste 02 | 351.365                   | 336.680                   | 505.834        | 921.248        |
|               | Teste 03 | 376.145                   | 266.932                   | <b>329.005</b> | 957.095        |
|               | Teste 04 | 454.567                   | 385.940                   | 446.471        | 1068.707       |
|               | Teste 05 | 311.501                   | 323.343                   | 398.831        | 1042.884       |
|               | Teste 06 | 272.981                   | 344.076                   | 489.900        | 983.384        |
|               | Teste 07 | 352.634                   | 263.068                   | 480.513        | 1003.611       |
|               | Teste 08 | <b>191.096</b>            | <b>180.767</b>            | 476.483        | 1037.129       |
|               | Teste 09 | 197.063                   | 193.239                   | 351.426        | 949.640        |
|               | Teste 10 | 202.689                   | 260.614                   | 484.451        | <b>894.611</b> |
| Média         |          | 315.249                   | <b>288.102</b>            | 441.126        | 979.862        |
| Desvio Padrão |          | 97.669                    | 66.767                    | 61.220         | 57.175         |

**Tabela 5.3: Distâncias obtidas em cada execução dos testes no Cenário 3**

| Cenário       | Teste    | BO-PMDP                   |                           | Míope          | Randômico      |
|---------------|----------|---------------------------|---------------------------|----------------|----------------|
|               |          | Exploração/Exploração 15% | Exploração/Exploração 30% |                |                |
| Cenário 3     | Teste 01 | 364.728                   | 343.351                   | 414.708        | 1220.365       |
|               | Teste 02 | 235.410                   | 330.978                   | 469.408        | 960.350        |
|               | Teste 03 | 316.884                   | 285.408                   | 316.229        | 941.018        |
|               | Teste 04 | 257.690                   | 251.649                   | 343.388        | 912.065        |
|               | Teste 05 | 303.441                   | 235.277                   | 508.691        | 944.058        |
|               | Teste 06 | <b>146.742</b>            | 248.415                   | 486.897        | <b>507.990</b> |
|               | Teste 07 | 278.939                   | 253.020                   | 421.015        | 969.239        |
|               | Teste 08 | 338.448                   | <b>137.774</b>            | 391.936        | 861.385        |
|               | Teste 09 | 605.249                   | 146.791                   | <b>218.817</b> | 940.917        |
|               | Teste 10 | 362.396                   | 317.775                   | 307.470        | 892.486        |
| Média         |          | 320.993                   | <b>255.044</b>            | 387.856        | 914.987        |
| Desvio Padrão |          | 119.500                   | 70.192                    | 91.313         | 173.040        |

Na Tabela 5.1 que apresenta os dados do cenário 1 demonstra que a abordagem BO-POMDP 30% realizou a exploração do ambiente com uma menor distância percorrida na média das execuções. Esse fato corrobora com a ideia de que verificar vários passos à frente possibilita uma tomada de decisão que leva à execução de trajetórias mais precisas e que resultam em menores distâncias percorridas. É possível perceber também que nos cenários com mais pontos de interesse (Tabelas 5.2 e 5.3) há uma maior variação das distâncias percorridas em cada um dos testes executados. Isso demonstra que uma quantidade maior de regiões de interesse aumenta as possibilidades de rotas que o UAV pode seguir levando a diferentes trajetórias realizadas em cada um dos testes e, conseqüentemente, levando à uma maior variedade de distâncias percorridas.

**Tabela 5.4: Distâncias - Abordagens BO-POMDP e Abordagem Míope**

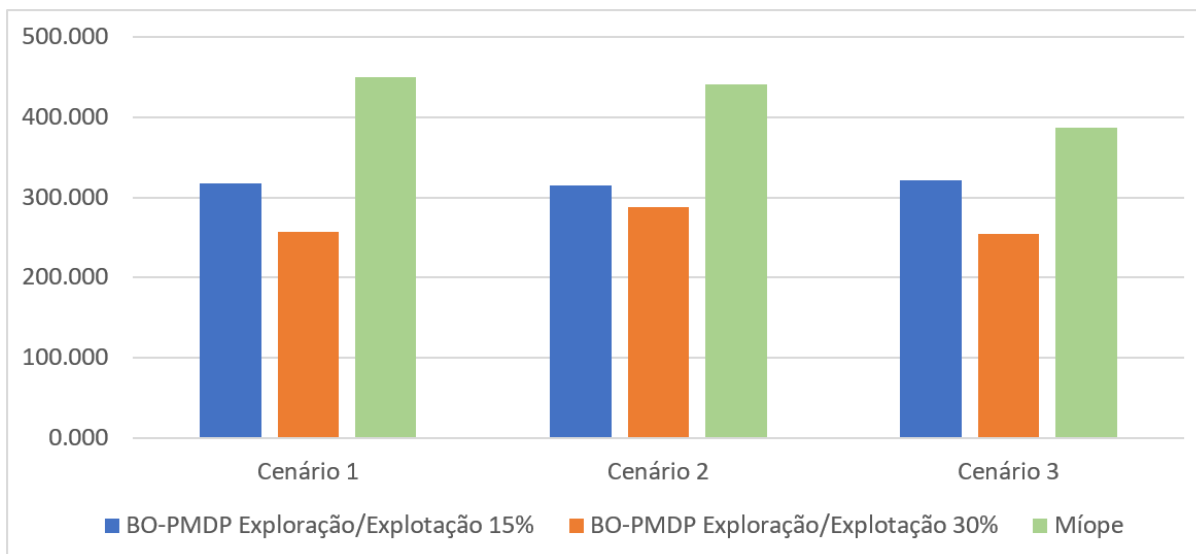
| Cenário   | BO-PMDP                   |                           | Míope   |
|-----------|---------------------------|---------------------------|---------|
|           | Exploração/Exploração 15% | Exploração/Exploração 30% |         |
| Cenário 1 | 317.080                   | <b>256.541</b>            | 450.532 |
| Cenário 2 | 315.249                   | <b>288.102</b>            | 441.126 |
| Cenário 3 | 320.993                   | <b>255.044</b>            | 387.856 |

**Tabela 5.5: Distâncias - Abordagens BO-POMDP e Abordagem Randômica**

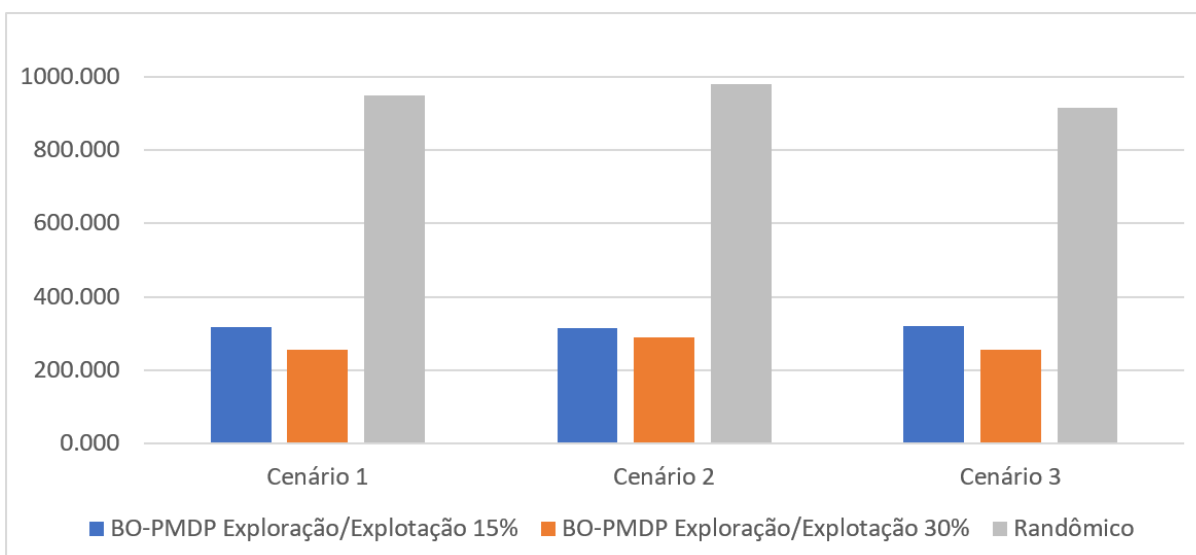
| Cenário   | BO-PMDP                   |                           | Randômico |
|-----------|---------------------------|---------------------------|-----------|
|           | Exploração/Exploração 15% | Exploração/Exploração 30% |           |
| Cenário 1 | 317.080                   | <b>256.541</b>            | 948.908   |
| Cenário 2 | 315.249                   | <b>288.102</b>            | 979.862   |
| Cenário 3 | 320.993                   | <b>255.044</b>            | 914.987   |

Nas Tabelas 5.4 e 5.5 e nas Figuras 5.5 e 5.6 é possível comparar as medidas das execuções dos algoritmos nos 3 cenários. O algoritmo BO-POMDP 30% obteve as menores médias





**FIGURA 5.5 – Gráfico da Abordagem das distâncias obtidas pela abordagem BO-POMDP e pela abordagem Míope**



**FIGURA 5.6 – Gráfico das distâncias obtidas pela abordagem BO-POMDP e pela abordagem Randômica**

de distância percorrida. O algoritmo BO-POMDP 30%, mesmo com uma taxa menor de exploração/exploração, obteve menores distâncias na média, quando comparado com o método míope e o método randômico.

### 5.3.4 Tempo

Os tempos obtidos a partir das trajetórias percorridas pelo UAV nas execuções foram comparados para as 4 abordagens (BO-POMDP 15%, BO-POMDP 30%, Míope e Randômica).

Tabela 5.6: Tempo Cenário 1

| Cenário             | Teste    | BO-PMDP                   |                           | Míope          | Randômico      |
|---------------------|----------|---------------------------|---------------------------|----------------|----------------|
|                     |          | Exploração/Exploração 15% | Exploração/Exploração 30% |                |                |
| Cenário 1           | Teste 01 | 146.760                   | 446.940                   | 280.230        | 420.200        |
|                     | Teste 02 | 200.190                   | <b>220.200</b>            | 233.700        | <b>413.580</b> |
|                     | Teste 03 | 153.440                   | 320.140                   | 273.490        | 420.170        |
|                     | Teste 04 | 287.000                   | 240.150                   | 347.090        | 420.480        |
|                     | Teste 05 | 326.770                   | 327.010                   | 293.710        | 420.230        |
|                     | Teste 06 | 246.910                   | 366.930                   | 200.070        | 420.200        |
|                     | Teste 07 | <b>140.220</b>            | 386.910                   | 280.150        | 413.590        |
|                     | Teste 08 | 233.660                   | 373.550                   | 246.900        | 420.240        |
|                     | Teste 09 | 246.850                   | 313.470                   | <b>193.500</b> | 420.200        |
|                     | Teste 10 | 227.010                   | 373.640                   | 280.410        | 420.140        |
| Média (em segundos) |          | <b>220.881</b>            | 336.894                   | 262.925        | 418.903        |
| Média (em minutos)  |          | <b>3.681</b>              | 5.615                     | 4.382          | 6.982          |
| Desvio Padrão       |          | 61.566                    | 68.404                    | 45.858         | 2.804          |

Tabela 5.7: Tempo Cenário 2

| Cenário             | Teste    | BO-PMDP                   |                           | Míope          | Randômico      |
|---------------------|----------|---------------------------|---------------------------|----------------|----------------|
|                     |          | Exploração/Exploração 15% | Exploração/Exploração 30% |                |                |
| Cenário 2           | Teste 01 | 293.440                   | 440.330                   | 280.030        | 420.260        |
|                     | Teste 02 | 406.900                   | 473.780                   | 300.100        | 420.260        |
|                     | Teste 03 | 273.580                   | 320.110                   | <b>206.770</b> | <b>420.150</b> |
|                     | Teste 04 | 347.000                   | 460.310                   | 246.880        | 420.240        |
|                     | Teste 05 | 293.450                   | 393.580                   | 273.630        | 420.240        |
|                     | Teste 06 | 193.480                   | 453.580                   | 266.860        | 420.190        |
|                     | Teste 07 | 226.800                   | 300.380                   | 253.810        | 420.180        |
|                     | Teste 08 | <b>133.410</b>            | 300.290                   | 253.460        | 420.190        |
|                     | Teste 09 | 200.150                   | <b>293.470</b>            | 226.900        | 420.360        |
|                     | Teste 10 | 193.390                   | 313.710                   | 280.240        | 420.160        |
| Média (em segundos) |          | <b>256.160</b>            | 374.954                   | 258.868        | 420.223        |
| Média (em minutos)  |          | <b>4.269</b>              | 6.249                     | 4.314          | 7.004          |
| Desvio Padrão       |          | 82.323                    | 76.296                    | 27.530         | 0.063          |

Tabela 5.8: Tempo Cenário 3

| Cenário             | Teste    | BO-PMDP                   |                           | Míope          | Randômico      |
|---------------------|----------|---------------------------|---------------------------|----------------|----------------|
|                     |          | Exploração/Exploração 15% | Exploração/Exploração 30% |                |                |
| Cenário 3           | Teste 01 | 260.210                   | 460.240                   | 213.520        | 520.220        |
|                     | Teste 02 | 193.350                   | 400.200                   | 320.110        | 413.650        |
|                     | Teste 03 | 200.070                   | 333.710                   | 206.820        | 420.180        |
|                     | Teste 04 | 173.420                   | 240.300                   | 240.280        | 420.180        |
|                     | Teste 05 | 226.980                   | 320.330                   | 266.870        | 420.240        |
|                     | Teste 06 | <b>120.050</b>            | 260.190                   | 253.460        | <b>226.770</b> |
|                     | Teste 07 | 180.180                   | 393.510                   | 240.070        | 420.360        |
|                     | Teste 08 | 240.170                   | 180.080                   | 240.130        | 420.290        |
|                     | Teste 09 | 400.100                   | <b>160.250</b>            | 200.270        | 413.510        |
|                     | Teste 10 | 293.540                   | 400.440                   | <b>180.190</b> | 420.410        |
| Média (em segundos) |          | <b>228.807</b>            | 314.925                   | 236.172        | 409.581        |
| Média (em minutos)  |          | <b>3.813</b>              | 5.249                     | 3.936          | 6.826          |
| Desvio Padrão       |          | 77.455                    | 101.615                   | 39.626         | 71.786         |

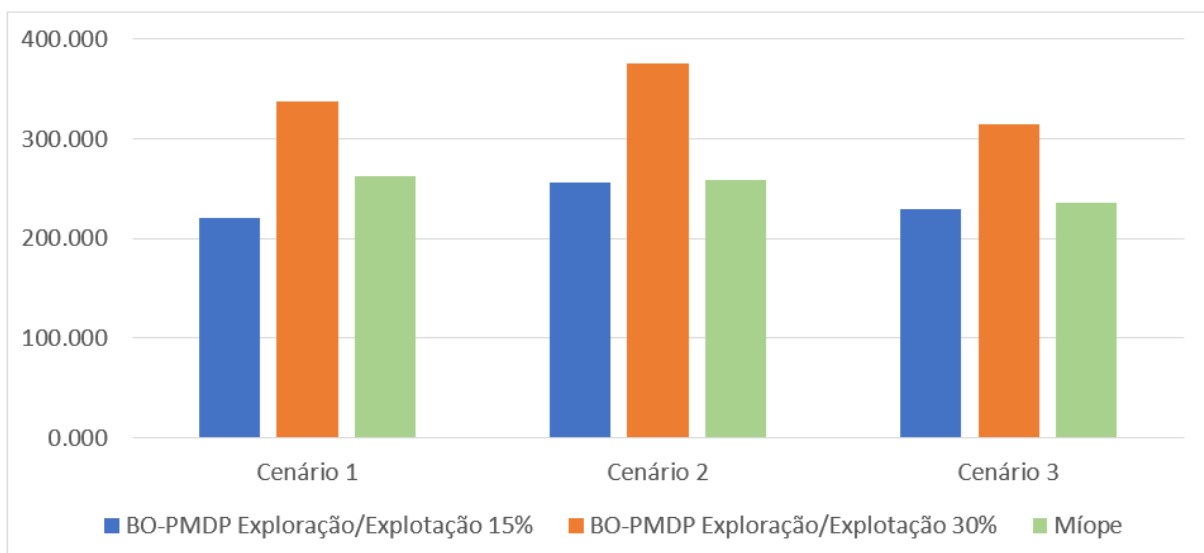
**Tabela 5.9: Tempos - Abordagens BO-POMDP e Abordagem Míope**

| Cenário   | BO-PMDP                   |                           | Míope   |
|-----------|---------------------------|---------------------------|---------|
|           | Exploração/Exploração 15% | Exploração/Exploração 30% |         |
| Cenário 1 | <b>220.881</b>            | 336.894                   | 262.925 |
| Cenário 2 | <b>256.160</b>            | 374.954                   | 258.868 |
| Cenário 3 | <b>228.807</b>            | 314.925                   | 236.172 |

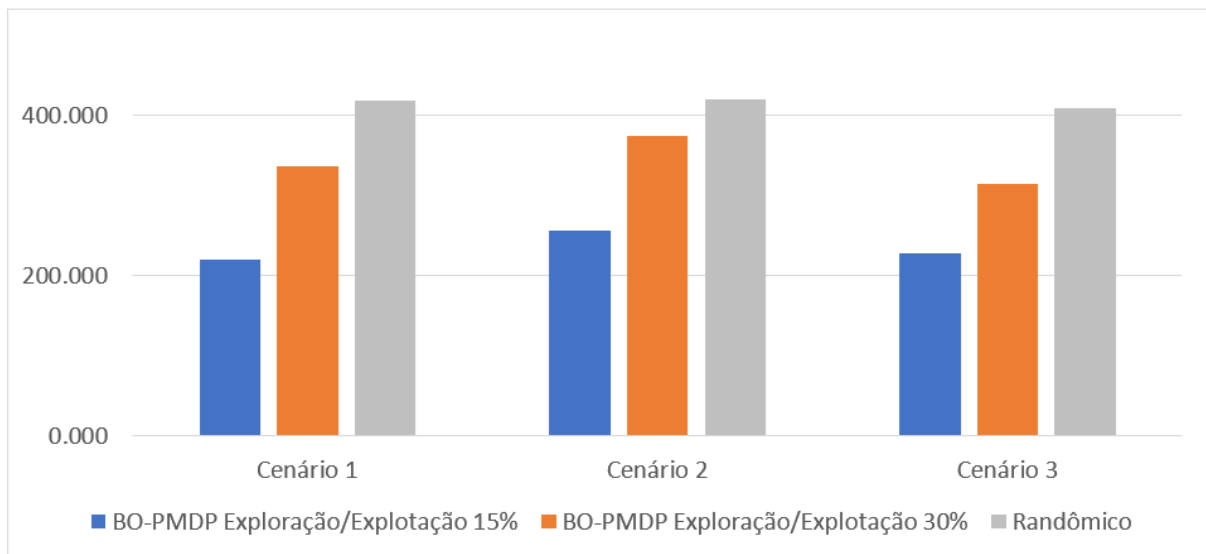
**Tabela 5.10: Tempos - Abordagens BO-POMDP e Abordagem Randômica**

| Cenário   | BO-PMDP                   |                           | Randômico |
|-----------|---------------------------|---------------------------|-----------|
|           | Exploração/Exploração 15% | Exploração/Exploração 30% |           |
| Cenário 1 | <b>220.881</b>            | 336.894                   | 418.903   |
| Cenário 2 | <b>256.160</b>            | 374.954                   | 420.223   |
| Cenário 3 | <b>228.807</b>            | 314.925                   | 409.581   |

Nas Tabelas 5.6, 5.7, 5.8 é possível perceber que a abordagem BO-POMDP 15% obteve menor tempo de execução na média. Mesmo com essa abordagem considerando vários passos à frente, ainda assim, ela obteve um menor tempo de execução na média do que abordagem míope e do que a abordagem randômica. Isso demonstra que um planejamento de trajetória explorativa informativa que leva em consideração a obtenção de recompensas futuras resulta em uma definição mais assertiva dos pontos que o UAV deve percorrer.

**FIGURA 5.7 – Gráfico dos tempos obtidos pela abordagem BO-POMDP e pela abordagem Míope**

Nas Tabelas 5.9 e 5.10 e nas Figuras 5.7 e 5.8 é possível comparar os tempos das execuções dos algoritmos nos 3 cenários. O algoritmo BO-POMDP 15% obteve as menores médias de tempo. Nos entanto, os valores são próximos aos valores do míope o que torna possível afirmar que foram tempos equivalentes. Isso indica que essa abordagem, realiza o planejamento de



**FIGURA 5.8** – Gráfico dos tempos obtidos pela abordagem BO-POMDP e pela abordagem Randômica

**Tabela 5.11: Sumarização das Médias das Distâncias e das Médias dos Tempos**

| Abordagens                        | Distância | Tempo   |
|-----------------------------------|-----------|---------|
| BO-PMDP Exploração/Exploração 15% | 317.080   | 220.881 |
| BO-PMDP Exploração/Exploração 30% | 256.541   | 336.894 |
| Míope                             | 450.532   | 262.925 |
| Randômico                         | 948.908   | 418.903 |

trajetória explorativa informativa em tempo semelhante percorrendo uma menor distância.

Nas Tabelas 5.11, 5.12 e 5.13 tem-se a demonstração de sumarizações das distâncias médias e dos tempos médios obtidos por cada um dos algoritmos avaliados. Nela, é possível perceber que o algoritmo BO-POMDP 15% realizou a exploração dos ambientes em uma menor distância média com tempo médio equivalente ao do algoritmo míope, demonstrando a eficácia da abordagem proposta. Também é possível perceber que o algoritmo BO-POMDP 30% obteve maior tempo médio de execução quando comparado com o algoritmo míope. No entanto, como a definição do sua trajetória leva em consideração recompensas futuras, mesmo com um tempo maior de execução, esse algoritmo proporcionou a realização da exploração do ambiente em uma menor distância, como pode ser visto na Tabela 5.4.

### 5.3.5 Testes de Wilcoxon

Para tornar possível a análise da significância estatística das diferentes abordagens executadas nesse trabalho, foi aplicada o teste de Wilcoxon. Nesse teste, comparações são realizadas

**Tabela 5.12: Sumarização das Médias das Distâncias e das Médias dos Tempos**

| Abordagens                        | Distância | Tempo   |
|-----------------------------------|-----------|---------|
| BO-PMDP Exploração/Exploração 15% | 315.249   | 256.160 |
| BO-PMDP Exploração/Exploração 30% | 288.102   | 374.954 |
| Míope                             | 441.126   | 258.868 |
| Randômico                         | 979.862   | 420.223 |

**Tabela 5.13: Sumarização das Médias das Distâncias e das Médias dos Tempos**

| Abordagens                        | Distância | Tempo   |
|-----------------------------------|-----------|---------|
| BO-PMDP Exploração/Exploração 15% | 320.993   | 228.807 |
| BO-PMDP Exploração/Exploração 30% | 255.044   | 314.925 |
| Míope                             | 387.856   | 236.172 |
| Randômico                         | 914.987   | 409.581 |

entre os dados obtidos a partir das execuções das abordagens. Os resultados obtidos indicam que uma hipótese nula  $H_0$  previamente estabelecida é rejeitada ou não. Nesse teste, um valor  $p$  é obtido de cada experimento. O valor  $p$  é comparado a um nível de significância pré-definido (para o presente trabalho o valor foi de 0,1). Se o valor  $p$  for menor ou igual ao nível de significância, rejeita-se a hipótese nula. Isso sugere que os resultados são estatisticamente significativos e não podem ser atribuídos ao acaso. Se o valor  $p$  for maior que o nível de significância a hipótese é não-rejeitada. Para o presente trabalho as hipóteses foram:

- $H_0$  em relação a BO-POMDP e Abordagem míope:
  - “Não há diferença significativa entre a abordagem BO-POMDP e a abordagem míope”
- $H_0$  em relação a BO-POMDP e Abordagem Randômica:
  - “Não há diferença significativa entre o método BO-POMDP e a abordagem randômica”

**Tabela 5.14: Teste de Wilcoxon - Cenário 1 - Distâncias**

| Cenário 1    |                     | R- | R+ | $H_0$ ( $\alpha = 0,1$ ) | Valor $p$ |
|--------------|---------------------|----|----|--------------------------|-----------|
| BO-POMDP 15% | BO-POMDP 30%        | 45 | 10 | Rejeitada                | 0,08398   |
| BO-POMDP 15% | Abordagem Míope     | 4  | 51 | Rejeitada                | 0,01367   |
| BO-POMDP 15% | Abordagem Randômica | 0  | 55 | Rejeitada                | 0,001953  |
| BO-POMDP 30% | Abordagem Míope     | 0  | 55 | Rejeitada                | 0,001953  |
| BO-POMDP 30% | Abordagem Randômica | 0  | 55 | Rejeitada                | 0,001954  |

**Tabela 5.15: Teste de Wilcoxon - Cenário 2 - Distâncias**

| Cenário 2    |                     | R- | R+ | H0 ( $\alpha = 0,1$ ) | Valor p  |
|--------------|---------------------|----|----|-----------------------|----------|
| BO-POMDP 15% | BO-POMDP 30%        | 40 | 15 | Não-Rejeitada         | 0,2324   |
| BO-POMDP 15% | Abordagem Míope     | 5  | 50 | Rejeitada             | 0,01953  |
| BO-POMDP 15% | Abordagem Randômica | 0  | 55 | Rejeitada             | 0,001953 |
| BO-POMDP 30% | Abordagem Míope     | 0  | 55 | Rejeitada             | 0,001953 |
| BO-POMDP 30% | Abordagem Randômica | 0  | 55 | Rejeitada             | 0,001953 |

**Tabela 5.16: Teste de Wilcoxon - Cenário 3 - Distâncias**

| Cenário 3    |                     | R- | R+ | H0 ( $\alpha = 0,1$ ) | Valor p  |
|--------------|---------------------|----|----|-----------------------|----------|
| BO-POMDP 15% | BO-POMDP 30%        | 40 | 15 | Não-Rejeitada         | 0,2324   |
| BO-POMDP 15% | Abordagem Míope     | 15 | 40 | Não-Rejeitada         | 0,2324   |
| BO-POMDP 15% | Abordagem Randômica | 0  | 55 | Rejeitada             | 0,001953 |
| BO-POMDP 30% | Abordagem Míope     | 1  | 54 | Rejeitada             | 0,003906 |
| BO-POMDP 30% | Abordagem Randômica | 0  | 55 | Rejeitada             | 0,001953 |

Nas Tabelas 5.14, 5.15 e 5.16 dos 3 cenários a hipótese  $H_0$  foi rejeitada em todas as comparações com a abordagem míope e com a abordagem randômica, com apenas uma exceção. Isso indica que, em relação às distâncias percorridas pelo método BO-POMDP proposto, foram obtidas distâncias que foram suficientemente grandes para serem estatisticamente significantes. Isso indica que o método BO-POMDP percorreu uma distância menor do que as outras abordagens durante o processo de exploração e de obtenção das informações sobre o ambiente. A exceção se deu na comparação entre o método BO-POMDP com 15% de taxa de exploração/exploração, indicando que não há diferença significativa das distâncias percorridas quando esse método é comparado com o método míope.

Considerando as Tabelas 5.17, 5.18 e 5.19 pode-se fazer as seguintes inferências:

Na comparação entre as abordagens BO-POMDP 15% e Míope, a hipótese  $H_0$  foi não-rejeitada nos 3 cenários indicando que o tempo necessário para a execução da trajetória do UAV não foi suficientemente grande para ser estatisticamente significativo.

Na comparação entre a abordagens BO-POMDP 15% e Randômica a hipótese  $H_0$  foi re-

**Tabela 5.17: Teste de Wilcoxon - Cenário 1 - Tempo**

| Cenário 1    |                     | R- | R+ | H0 ( $\alpha = 0,1$ ) | Valor p  |
|--------------|---------------------|----|----|-----------------------|----------|
| BO-POMDP 15% | BO-POMDP 30%        | 3  | 52 | Rejeitada             | 0.009766 |
| BO-POMDP 15% | Abordagem Míope     | 11 | 44 | Não-Rejeitada         | 0.1055   |
| BO-POMDP 15% | Abordagem Randômica | 0  | 55 | Rejeitada             | 0.001953 |
| BO-POMDP 30% | Abordagem Míope     | 48 | 7  | Rejeitada             | 0.03711  |
| BO-POMDP 30% | Abordagem Randômica | 2  | 53 | Rejeitada             | 0.005859 |

**Tabela 5.18: Teste de Wilcoxon - Cenário 2 - Tempo**

| Cenário 2    |                     | R- | R+ | H0 ( $\alpha = 0,1$ ) | Valor p  |
|--------------|---------------------|----|----|-----------------------|----------|
| BO-POMDP 15% | BO-POMDP 30%        | 0  | 55 | Rejeitada             | 0.001953 |
| BO-POMDP 15% | Abordagem Míope     | 25 | 30 | Não-Rejeitada         | 0.8457   |
| BO-POMDP 15% | Abordagem Randômica | 0  | 55 | Rejeitada             | 0.001953 |
| BO-POMDP 30% | Abordagem Míope     | 55 | 0  | Rejeitada             | 0.001953 |
| BO-POMDP 30% | Abordagem Randômica | 13 | 42 | Não-Rejeitada         | 0.1602   |

**Tabela 5.19: Teste de Wilcoxon - Cenário 3 - Tempo**

| Cenário 3    |                     | R- | R+ | H0 ( $\alpha = 0,1$ ) | Valor p  |
|--------------|---------------------|----|----|-----------------------|----------|
| BO-POMDP 15% | BO-POMDP 30%        | 11 | 44 | Não-Rejeitada         | 0.1055   |
| BO-POMDP 15% | Abordagem Míope     | 22 | 33 | Não-Rejeitada         | 0,625    |
| BO-POMDP 15% | Abordagem Randômica | 0  | 55 | Rejeitada             | 0,001953 |
| BO-POMDP 30% | Abordagem Míope     | 47 | 8  | Rejeitada             | 0,04883  |
| BO-POMDP 30% | Abordagem Randômica | 3  | 52 | Rejeitada             | 0,009766 |

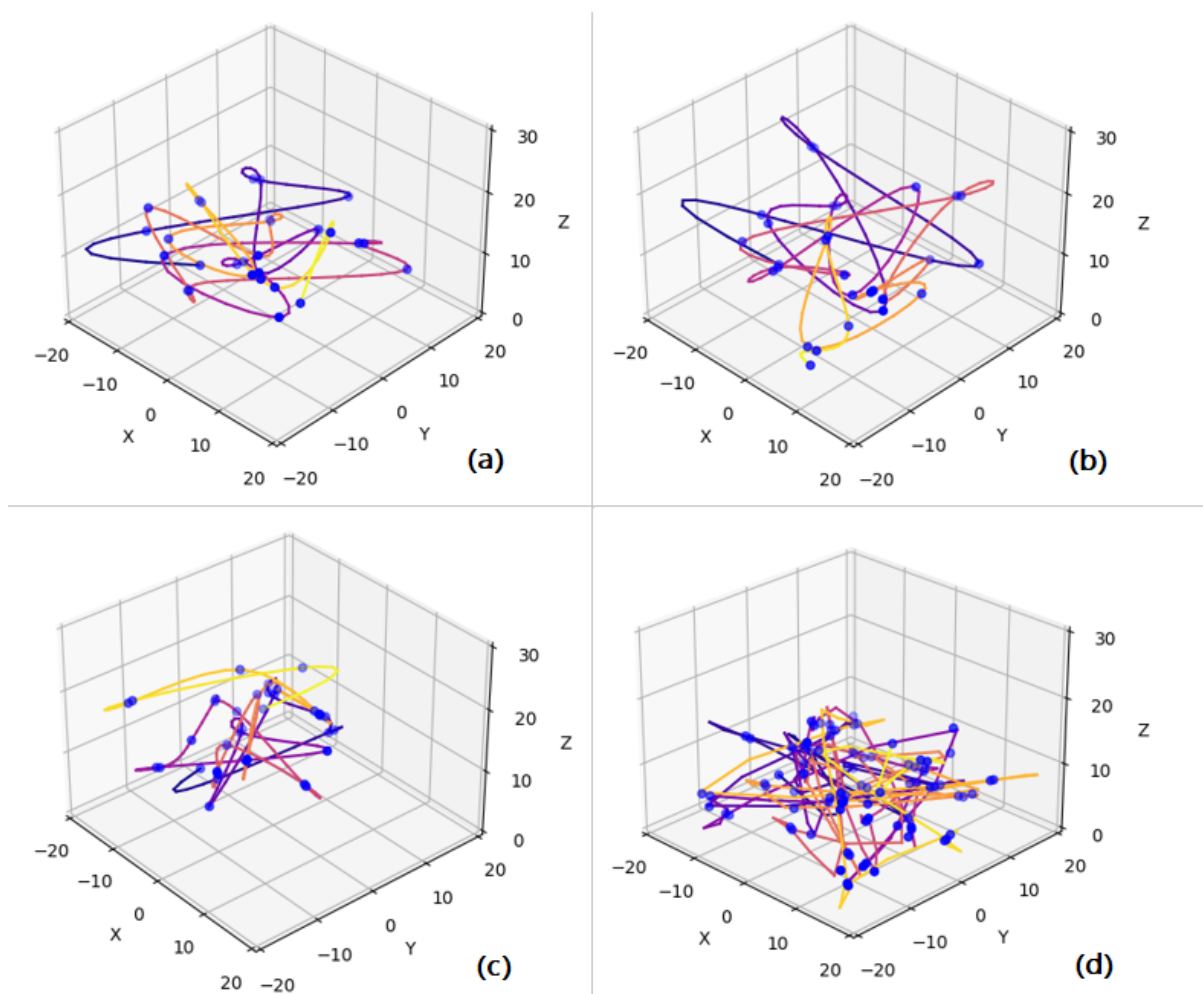
jeitada nos 3 cenários, indicando que o tempo necessário para a execução da trajetória do UAV foi suficientemente grande para ser estatisticamente significativa. Nesse caso, a abordagem BO-POMDP 15% realizou um planejamento de trajetória explorativa e informativa que resultou em melhores escolhas de trajetória possibilitando a execução em um menor tempo quando comparado ao método randômico.

Na comparação entre a abordagens BO-POMDP 30% e Míope a hipótese H0 foi rejeitada nos 3 cenários, indicando que o tempo necessário para a execução da trajetória do UAV foi suficientemente grande para ser estatisticamente significativa. Como esperado, o fato da abordagem BO-POMDP 30% explorar mais o espaço de busca de soluções visando recompensas futuras, leva a um maior tempo de execução. Entretanto, como pode ser visto nas Tabelas 5.11, 5.12 e 5.13, a distância percorrida foi menor, indicando a possibilidade do uso dessa abordagem em situações em que seja necessária um planejamento explorativo e informativa mais assertivo e proporcionando um consumo menor de recursos e ocasionando a redução da possibilidade de colisões durante o trajeto. Nesse caso, a abordagem BO-POMDP 15% realizou um planejamento de trajetória explorativa e informativa que resultou em melhores escolhas de trajetória, possibilitando a execução em um menor tempo quando comparado ao método randômico.

Na comparação entre a abordagens BO-POMDP 30% e randômica, a hipótese H0 foi rejeitada nos cenários 1 e 3, indicando que, nesses casos, o tempo necessário para a execução da trajetória do UAV foi suficientemente grande para ser estatisticamente significativa. No cenário 2 a hipótese H0 foi não-rejeitada, indicando que, para este cenário, os tempos de execução das abordagens não foram suficientemente grandes para ser estatisticamente significantes.

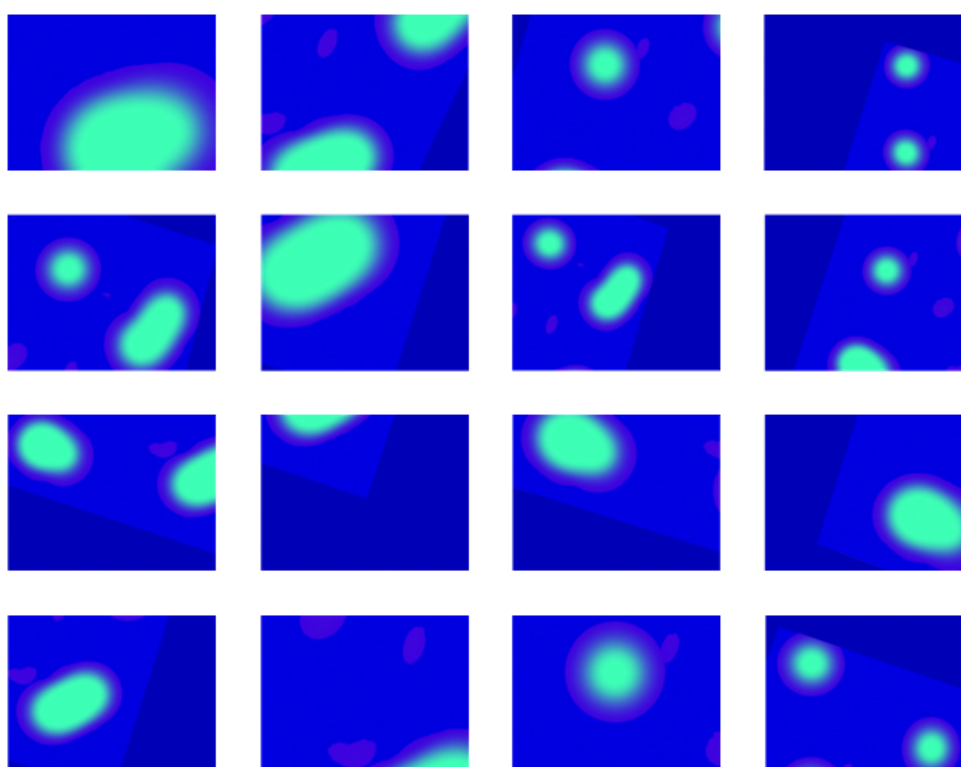
### 5.3.6 Exemplo de Execução

No exemplo abaixo, Figura 5.9, tem-se a realização do Planejamento das Trajetórias Explorativas e Informativas usando as 4 abordagens: abordagem proposta BO-POMDP com taxa de exploração/exploração em 15%, abordagem proposta BO-POMDP com taxa de exploração/exploração em 30%, abordagem míope e abordagem randômica. Na Figura 5.10 tem-se as imagens capturadas pelo UAV durante a execução de um planejamento de trajetória explorativa informativa. Os pontos de destaque em verde indicam as regiões de interesse que o UAV deve priorizar.



**FIGURA 5.9** – Em a) tem-se a realização do EIPP usando o algoritmo BO-POMDP com taxa de exploração/exploração em 15%, em b) tem-se a realização do EIPP usando o algoritmo BO-POMDP com taxa de exploração/exploração em 30%, em c) tem-se a execução do algoritmo Míope e em d) tem-se a execução do algoritmo randômico.





**FIGURA 5.10** – Imagens capturadas pelo UAV durante a execução de um planejamento de trajetória explorativa informativa

# Capítulo 6

## CONCLUSÃO

---

---

Neste Capítulo apresentam-se as conclusões deste trabalho, as contribuições, limitações e trabalhos futuros.

### 6.1 Conclusão

Neste trabalho, apresentou-se a abordagem não-míope para o Planejamento de Trajetória Explorativa e Informativa (*Explorative Informative Path Planning - EIPP*) em ambientes desconhecidos usando UAV, a fim de obter um planejamento que maximize o ganho de informação com a localização de possíveis alvos ou regiões de interesse e, conseqüentemente, minimize as incertezas sobre o ambiente. Para essa abordagem, foi desenvolvido um sistema para a tomada de decisão sequencial sob incertezas em ambiente 3D baseado na abordagem BO-POMDP.

Testes foram realizados em 3 cenários com diferentes regiões de interesse em um ambiente com dimensão de 30m x 30m x 30 m. As métricas RMSE e  $Tr(P)$  foram usados para avaliar a abordagem BO-POMDP. Com a métrica  $Tr(P)$  foi demonstrados que a incerteza sobre o ambiente diminui no decorrer das iterações dos algoritmos. Além disso, o algoritmo BO-POMDP não-Míope obteve melhores resultados quando comparado às abordagens míope e randômica, no que diz respeito à distância necessária percorrida pelo UAV. Esse fato se deu em um tempo de trajetória equivalente à abordagem míope, demonstrado que a busca por soluções melhores, a partir do planejamento que leve em consideração recompensas futuras, leva a escolha de soluções mais assertivas e, conseqüentemente, a uma exploração mais consistente e precisa do espaço de busca.

Já com a métrica RMSE é demonstrado que o algoritmo BO-POMDP não-Míope obteve o menor erro em relação aos demais, o que permite inferir que o planejamento de trajetória

explorativo e informativo foi assertivo para obter um planejamento que maximiza os ganhos de informação em uma dada região de interesse. Vale destacar também que o algoritmo em questão convergiu em menos iterações comparado com os demais.

Vale ressaltar que o algoritmo Míope no Cenário 2, obteve a maior diferença em relação ao erro comparado a cada um dos cenários e precisou de mais iterações para convergir, quando comparado com o algoritmo Não-Míope. Já os métodos míopes, realizam um planejamento que define a melhor ação como aquela de maior recompensa imediata e não consideram o planejamento futuro das recompensas. Esse fato, apesar de proporcionar uma definição de trajetória rápida, acarreta no aumento da chance de se obter soluções que estejam em ótimos locais.

Assim, cumpriu-se o objetivo do trabalho que foi desenvolver uma estratégia para EIPP com o uso da abordagem BO-POMDP a fim de se obter um planejamento que maximize os ganhos de informação em uma dada região de interesse e consequentemente minimize as incertezas sobre o ambiente considerando um espaço contínuo 3D.

## 6.2 Contribuições

Um das principais contribuições deste trabalho reside na incorporação de trajetórias contínuas em um espaço tridimensional (3D), com o objetivo de maximizar o ganho de informação ao localizar possíveis alvos ou regiões de interesse. Para alcançar esse objetivo, foi elaborado um sistema de tomada de decisão sequencial sob incertezas, com base na abordagem BO-POMDP para o problema de Planejamento de Trajetória Explorativa Informativa - EIPP.

Além disso, este trabalho usa a abordagem BO-POMDP em um contexto que se diferencia ao considerar a variação de altitude no planejamento das trajetórias, o que amplia a compreensão abrangente do ambiente e a flexibilidade de exploração.

A estratégia relacionada a taxa de exploração/exploração para seleção dos pontos a serem avaliados (pontos mais promissores) na abordagem BO-POMDP proposto acelerou a convergência do algoritmo. Já que o problema EIPP possui um *trade-off* relacionado ao tempo necessário para monitorar exaustivamente uma grande área para obter o máximo de informações sobre o meio ambiente e buscar uma solução ótimo global.

Outras contribuições estão relacionadas à investigação do comportamento de abordagens não-míopes em comparação as abordagens consideradas míopes no contexto de Planejamento de Trajetória Explorativa e Informativa.

Além disso, houve contribuições na área de robótica no que refere-se a definição dos pro-

blemas de Planejamento de Trajetória Informativa, Exploração Autônoma e Planejamento de Rotas usando UAVs em monitoramento de ambientes desconhecidos. Os conceitos e diferenças entre esses problemas, bem como a taxonomia dos métodos adotados para resolvê-los foram apresentados.

## 6.3 Limitações e Trabalhos Futuros

A partir deste trabalho surgem novas oportunidades de pesquisa, tanto no sentido de melhorar a abordagem proposta, quanto na elaboração de novas ideias de como usar a abordagem BO-POMDP no contexto de Planejamento de Trajetória Explorativa e Informativa.

A abordagem BO-POMDP possui uma limitação relacionada ao horizonte de planejamento fixado em profundidade 3. Assim, investigar a complexidade do algoritmo modificando esse horizonte de planejamento e verificar como isso influencia no Planejamento de Trajetória Informativa. Outra ideia é realizar busca de métodos que usam a Computação Paralela para a execução do algoritmo que poderiam trazer ganhos significativos, uma vez que o problema EIPP possui um trade-off relacionado ao tempo computacional e buscar uma solução ótima global.

Como trabalho futuro, a investigação de novas funções de recompensa usando diferentes funções de aquisições, como por exemplo, Expected Improvement - EI e Probability of Improvement - PI podem resultar em melhorias significativas.

Por fim, considerar um sistema multi-agente pode estender a abordagem BO-POMDP para MEIPP (Multi-agent Explorative and Informative Path Planning ) para monitoramento de ambientes desconhecidos e complexos. Com isso, novos desafios de pesquisa surgiriam para alcançar um comportamento coordenado das UAVs, bem como melhorar a eficiência computacional do procedimento de otimização e reformulação da função de recompensa 4.6.

## REFERÊNCIAS

---

---

- AGUERO, C. et al. Inside the virtual robotics challenge: Simulating real-time robotic disaster response. *Automation Science and Engineering, IEEE Transactions on*, v. 12, n. 2, p. 494–506, April 2015. ISSN 1545-5955.
- BARBEDO, J. G. A. A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses. *Drones*, Multidisciplinary Digital Publishing Institute, v. 3, n. 2, p. 40, 2019.
- BERNARDINI, S.; FOX, M.; LONG, D. Planning the behaviour of low-cost quadcopters for surveillance missions. In: *Proc. of Int. Conf. on Automated Planning and Scheduling*. Portsmouth, USA: [s.n.], 2014.
- BINNEY, J.; KRAUSE, A.; SUKHATME, G. S. Informative path planning for an autonomous underwater vehicle. In: IEEE. *2010 IEEE International Conference on Robotics and Automation*. [S.l.], 2010. p. 4791–4796.
- BINNEY, J.; SUKHATME, G. S. Branch and bound for informative path planning. In: IEEE. *2012 IEEE International Conference on Robotics and Automation*. [S.l.], 2012. p. 2147–2154.
- BISWAS, S.; ANAVATTI, S. G.; GARRATT, M. A. A particle swarm optimization based path planning method for autonomous systems in unknown terrain. In: IEEE. *2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*. [S.l.], 2019. p. 57–63.
- BLANCHARD, A.; SAPSIS, T. Informative path planning for anomaly detection in environment exploration and monitoring. *Ocean Engineering*, Elsevier, v. 243, p. 110242, 2022.
- Bourgault, F.; Makarenko, A. A.; Williams, S. B.; Grocholsky, B.; Durrant-Whyte, H. F. Information based adaptive robotic exploration. v. 1, p. 540–545 vol.1, 2002.
- BRAEKERS, K.; RAMAEKERS, K.; NIEUWENHUYSE, I. V. The vehicle routing problem: State of the art classification and review. *Computers & Industrial Engineering*, Elsevier, v. 99, p. 300–313, 2016.
- BROCHU, E.; CORA, V.; FREITAS, N. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *CoRR*, abs/1012.2599, 12 2010.

- BROCK, O.; TRINKLE, J.; RAMOS, F. Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces. In: \_\_\_\_\_. *Robotics: Science and Systems IV*. [S.l.: s.n.], 2009. p. 65–72.
- BáčA, T.; PETRLÍK, M.; VRBA, M.; SPURNÝ, V.; PěNIČKA, R.; HERT, D.; SASKA, M. The mrs uav system: Pushing the frontiers of reproducible research, real-world deployment, and education with autonomous unmanned aerial vehicles. In: . [S.l.: s.n.], 2020.
- CAPRAZ, S.; AZYIKMIS, H.; OZSOY, A. An optimized gpu-accelerated route planning of multi-uav systems using simulated annealing. *International Journal of Machine Learning and Computing*, v. 10, n. 3, 2020.
- CARRILLO, H.; DAMES, P.; KUMAR, V.; CASTELLANOS, J. A. Autonomous robotic exploration using a utility function based on rényi’s general theory of entropy. *Autonomous Robots*, Springer, v. 42, n. 2, p. 235–256, 2018.
- Cox, D. D.; John, S. A statistical method for global optimization. In: *[Proceedings] 1992 IEEE International Conference on Systems, Man, and Cybernetics*. [S.l.: s.n.], 1992. p. 1241–1246 vol.2.
- EKSIOGLU, B.; VURAL, A. V.; REISMAN, A. The vehicle routing problem: A taxonomic review. *Computers & Industrial Engineering*, Elsevier, v. 57, n. 4, p. 1472–1483, 2009.
- ERGEZER, H.; LEBLEBICIOĞLU, K. Online path planning for unmanned aerial vehicles to maximize instantaneous information. *International Journal of Advanced Robotic Systems*, SAGE Publications Sage UK: London, England, v. 18, n. 3, p. 17298814211010379, 2021.
- FAA. *Unmanned Aircraft Systems*. [S.l.]: Federal Aviation Administration, 2016.
- FEDOROV, V. V. *Theory of optimal experiments*. [S.l.]: Elsevier, 2013.
- FERNÁNDEZ, I. M. R.; DENNISTON, C. E.; CARON, D. A.; SUKHATME, G. S. Informative path planning to estimate quantiles for environmental analysis. *IEEE Robotics and Automation Letters*, IEEE, v. 7, n. 4, p. 10280–10287, 2022.
- FOUNDATION, O. S. R. *Gazebo: Robot simulation made easy*. 2014. <http://gazebosim.org/media>. Acessado: 01-11-2020.
- FRANCIS, G.; OTT, L.; MARCHANT, R.; RAMOS, F. Occupancy map building through bayesian exploration. *The International Journal of Robotics Research*, SAGE Publications Sage UK: London, England, v. 38, n. 7, p. 769–792, 2019.
- FRANCIS, G.; OTT, L.; RAMOS, F. Functional path optimisation for exploration in continuous occupancy maps. In: *Robotics Research*. [S.l.]: Springer, 2020. p. 859–875.
- FURRER, F.; BURRI, M.; ACHELNIK, M.; SIEGWART, R. Rotors—a modular gazebo mav simulator framework. *Robot Operating System (ROS) The Complete Reference (Volume 1)*, Springer, p. 595–625, 2016.
- GALCERAN, E.; CARRERAS, M. A survey on coverage path planning for robotics. *Robotics and Autonomous systems*, Elsevier, v. 61, n. 12, p. 1258–1276, 2013.
- HANSEN, N. *The CMA Evolution Strategy: A Tutorial*. 2023.

- HANSEN, N.; MÜLLER, S. D.; KOUMOUTSAKOS, P. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es). *Evolutionary computation*, MIT Press, v. 11, n. 1, p. 1–18, 2003.
- HARI, S. K. K.; RATHINAM, S.; DARBHA, S.; KALYANAM, K.; MANYAM, S. G.; CASBEER, D. Optimal uav route planning for persistent monitoring missions. *IEEE Transactions on Robotics*, v. 37, n. 2, p. 550–566, 2021.
- HITZ, G.; GALCERAN, E.; GARNEAU, M.-È.; POMERLEAU, F.; SIEGWART, R. Adaptive continuous-space informative path planning for online environmental monitoring. *Journal of Field Robotics*, Wiley Online Library, v. 34, n. 8, p. 1427–1449, 2017.
- HOLLINGER, G. A.; SUKHATME, G. S. Sampling-based robotic information gathering algorithms. *The International Journal of Robotics Research*, SAGE Publications Sage UK: London, England, v. 33, n. 9, p. 1271–1287, 2014.
- JADIDI, M. G. *Gaussian processes for information-theoretic robotic mapping and exploration*. Tese (Doutorado), 2017.
- JONES, D. A taxonomy of global optimization methods based on response surfaces. *J. of Global Optimization*, v. 21, p. 345–383, 12 2001.
- KANCIR, P.; DIGUET, J.; SEVAUX, M. Development of tools for multi vehicles simulation with robot operating system and ardupilot. In: . [S.l.: s.n.], 2019.
- KANISTRAS, K.; MARTINS, G.; RUTHERFORD, M. J.; VALAVANIS, K. P. A survey of unmanned aerial vehicles (uavs) for traffic monitoring. In: IEEE. *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*. [S.l.], 2013. p. 221–234.
- KARAKAYA, M. Uav route planning for maximum target coverage. *Computer Science and Engineering: An International Journal (CSEIJ)*, v. 4, 2014.
- KIM, S. J.; LIM, G. J.; CHO, J.; CÔTÉ, M. J. Drone-aided healthcare services for patients with chronic diseases in rural areas. *Journal of Intelligent and Robotic Systems*, v. 88, p. 163–180, 2017.
- KOENIG, N.; HOWARD, A. Design and use paradigms for gazebo, an open-source multi-robot simulator. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*. Sendai, Japan: [s.n.], 2004. p. 2149–2154.
- KRAUSE, A.; SINGH, A.; GUESTRIN, C. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, v. 9, n. Feb, p. 235–284, 2008.
- KRELL, E.; SHETA, A.; BALASUBRAMANIAN, A. P. R.; KING, S. A. Collision-free autonomous robot navigation in unknown environments utilizing pso for path planning. *Journal of Artificial Intelligence and Soft Computing Research*, Sciendo, v. 9, n. 4, p. 267–282, 2019.
- KURNIAWATI, H.; YADAV, V. An online pomdp solver for uncertainty planning in dynamic environment. In: *Robotics Research*. [S.l.]: Springer, 2016. p. 611–629.

- KWON, W.; PARK, J. H.; LEE, M.; HER, J.; KIM, S.-H.; SEO, J.-W. Robust autonomous navigation of unmanned aerial vehicles (uavs) for warehouses' inventory application. *IEEE Robotics and Automation Letters*, IEEE, v. 5, n. 1, p. 243–249, 2019.
- LAVALLE, M.; KUFFNER, S. J. J. Rapidly-exploring random trees: Progress and prospects. In: *Proc. Workshop on the Algorithmic Foundations of Robotics*. San Francisco: [s.n.], 2000.
- LI, H.; SAVKIN, A. V.; VUCETIC, B. Autonomous area exploration and mapping in underground mine environments by unmanned aerial vehicles. *Robotica*, Cambridge University Press, v. 38, n. 3, p. 442–456, 2020.
- Li, T.; Wang, C.; Max Q.-H., M.; Silva, C. W. d. Coverage sampling planner for uav-enabled environmental exploration and field mapping. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. [S.l.: s.n.], 2019. p. 2509–2516.
- LI, W.; TAN, M.; WANG, L.; WANG, Q. A cubic spline method combining improved particle swarm optimization for robot path planning in dynamic uncertain environment. *International Journal of Advanced Robotic Systems*, SAGE Publications Sage UK: London, England, v. 17, n. 1, p. 1729881419891661, 2020.
- LIM, Z. W.; HSU, D.; LEE, W. S. Adaptive informative path planning in metric spaces. *The International Journal of Robotics Research*, SAGE Publications Sage UK: London, England, v. 35, n. 5, p. 585–598, 2016.
- MAJEED, A.; LEE, S. A fast global flight path planning algorithm based on space circumscription and sparse visibility graph for unmanned aerial vehicle. *Electronics*, Multidisciplinary Digital Publishing Institute, v. 7, n. 12, p. 375, 2018.
- MALANDRINO, F.; CHIASSERINI, C.-F.; CASETTI, C.; CHIARAVIGLIO, L.; SENACHERIBBE, A. Planning uav activities for efficient user coverage in disaster areas. *Ad Hoc Networks*, Elsevier, v. 89, p. 177–185, 2019.
- MARCHANT, R.; RAMOS, F. Bayesian optimisation for informative continuous path planning. In: IEEE. *2014 IEEE International Conference on Robotics and Automation (ICRA)*. [S.l.], 2014. p. 6136–6143.
- MARCHANT, R.; RAMOS, F.; SANNER, S. Sequential bayesian optimisation for spatial-temporal monitoring. *Uncertainty in Artificial Intelligence - Proceedings of the 30th Conference, UAI 2014*, p. 553–562, 01 2014.
- MARTIN, P.; SALAUN, E. The true role of accelerometer feedback in quadrotor control. In: IEEE. *2010 IEEE international conference on robotics and automation*. [S.l.], 2010. p. 1623–1629.
- MEDEIRO, F. L. L.; SILVA, J. D. S. da. A dijkstra algorithm for fixed-wing uav motion planning based on terrain elevation. *Advances in Artificial Intelligence - Lecture Notes in Computer Science*, v. 6404, p. 213–22, 2010.
- MENG, H.; XIN, G. Uav route planning based on the genetic simulated annealing algorithm. In: *International Conference on Mechatronics and Automation*. Xi'an, China: [s.n.], 2010.



- MENG, L.; QING, S.; QINJUN, Z.; YONGLIANG, Z. Route planning for unmanned aerial vehicle based on rolling rrt in unknown environment. In: IEEE. *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCCIC)*. [S.l.], 2016. p. 1–4.
- MEYER, J.; SENDOBRY, A.; KOHLBRECHER, S.; KLINGAUF, U.; STRYK, O. V. Comprehensive simulation of quadrotor uavs using ros and gazebo. In: . [S.l.: s.n.], 2012. v. 7628, p. 400–411. ISBN 978-3-642-34326-1.
- MORERE, P.; FRANCIS, G.; BLAU, T.; RAMOS, F. Reinforcement learning with probabilistically complete exploration. *arXiv preprint arXiv:2001.06940*, 2020.
- MORERE, P.; MARCHANT, R.; RAMOS, F. Sequential bayesian optimization as a pomdp for environment monitoring with uavs. In: IEEE. *2017 IEEE International Conference on Robotics and Automation (ICRA)*. [S.l.], 2017. p. 6381–6388.
- OLEYNIKOVA, H.; TAYLOR, Z.; MILLANE, A.; SIEGWART, R.; NIETO, J. A complete system for vision-based micro-aerial vehicle mapping, planning, and flight in cluttered environments. *arXiv preprint arXiv:1812.03892*, 2018.
- PALAZZOLO, E.; STACHNISS, C. Effective exploration for mavs based on the expected information gain. *Drones*, Multidisciplinary Digital Publishing Institute, v. 2, n. 1, p. 9, 2018.
- Park, H.; Lee, B. H.; Morrison, J. R. Persistent uav security presence service: Architecture and prototype implementation. In: *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*. [S.l.: s.n.], 2017. p. 1800–1807.
- POPOVIĆ, M.; VIDAL-CALLEJA, T.; CHUNG, J. J.; NIETO, J.; SIEGWART, R. Informative path planning and mapping for active sensing under localization uncertainty. *arXiv preprint arXiv:1902.09660*, 2019.
- POPOVIĆ, M.; VIDAL-CALLEJA, T.; HITZ, G.; JEN, C. J.; INKYU, S.; SIEGWART, R.; NIETO, J. An informative path planning framework for uav-based terrain monitoring. *Autonomous Robots*, Springer Nature BV, v. 44, n. 6, p. 889–911, 2020.
- Popović, M.; Hitz, G.; Nieto, J.; Sa, I.; Siegwart, R.; Galceran, E. Online informative path planning for active classification using uavs. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. [S.l.: s.n.], 2016. p. 5753–5758.
- POPOVIĆ, M.; VIDAL-CALLEJA, T.; HITZ, G.; SA, I.; SIEGWART, R.; NIETTO, J. Multiresolution mapping and informative path planning for uav-based terrain monitoring. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. [S.l.: s.n.], 2017. p. 1382–1388.
- ROS.ORG. *Robot Operating System - ROS*. 2014. <https://www.ros.org/about-ros/>. Acessado: 01-11-2020.
- ROSSI, M.; BRUNELLI, D. Autonomous gas detection and mapping with unmanned aerial vehicles. *IEEE Transactions on Instrumentation and Measurement*, IEEE, v. 65, n. 4, p. 765–775, 2015.

- SANTOS, M. A. A.; VIVALDINI, K. C. T. A modeling strategy using bayesian optimization with pomdp for exploration and informative path planning of uavs in monitoring of forest. In: *2022 International Conference on Computational Science and Computational Intelligence (CSCI)*. [S.l.: s.n.], 2022. p. 1–6.
- SANTOS, M. A. A.; VIVALDINI, K. C. T. A review of the informative path planning, autonomous exploration and route planning using uav in environment monitoring. In: *2022 International Conference on Computational Science and Computational Intelligence (CSCI)*. [S.l.: s.n.], 2022. p. 1–6.
- SCHOPFERER, S.; BENDERS, S. Minimum-risk path planning for long-range and low-altitude flights of autonomous unmanned aircraft. In: *AIAA Scitech 2020 Forum*. [S.l.: s.n.], 2020. p. 0137.
- SILVER, D.; VENESS, J. Monte-carlo planning in large pomdps. In: . [S.l.: s.n.], 2010. p. 2164–2172.
- SIM, R.; ROY, N. Global a-optimal robot exploration in slam. In: *IEEE. Proceedings of the 2005 IEEE international conference on robotics and automation*. [S.l.], 2005. p. 661–666.
- SINGH, A.; KRAUSE, A.; GUESTRIN, C.; KAISER, W. J. Efficient informative sensing using multiple robots. *Journal of Artificial Intelligence Research*, v. 34, p. 707–755, 2009.
- SMALLWOOD, R. D.; SONDIK, E. J. The optimal control of partially observable markov processes over a finite horizon. *Operations research*, INFORMS, v. 21, n. 5, p. 1071–1088, 1973.
- SRIVASTAVA, K.; BHUTORIA, A. J.; SHARMA, J. K.; SINHA, A.; PANDEY, P. C. Uavs technology for the development of gui based application for precision agriculture and environmental research. *Remote Sensing Applications: Society and Environment*, Elsevier, v. 16, p. 100258, 2019.
- STACHNISS, C. *Robotic mapping and exploration*. [S.l.]: Springer, 2009.
- TAN, U.; RABASTE, O.; ADNET, C.; OVARLEZ, J.-P. On the eclipsing phenomenon with phase codes. In: . [S.l.: s.n.], 2019. p. 1–5.
- THIBBOTUWAWA, A.; BOCEWICZ, G.; NIELSEN, P.; BANASZAK, Z. Unmanned aerial vehicle routing problems: a literature review. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 10, n. 13, p. 4504, 2020.
- TURKER, T.; SAHINGOZ, O. K.; YILMAZ, G. 2d path planning for uavs in radar threatening environment using simulated annealing algorithm. In: *International Conference on Unmanned Aircraft Systems*. Denver, CO, USA: [s.n.], 2015.
- VASUDEVAN, S.; RAMOS, F.; NETTLETON, E.; DURRANT-WHYTE, H. Gaussian process modeling of large-scale terrain. *Journal of Field Robotics*, Wiley Online Library, v. 26, n. 10, p. 812–840, 2009.
- VIVALDINI, K. C.; MARTINELLI, T. H.; GUIZILINI, V. C.; SOUZA, J. R.; OLIVEIRA, M. D.; RAMOS, F. T.; WOLF, D. F. Uav route planning for active disease classification. *Autonomous Robots*, Springer, v. 43, n. 5, p. 1137–1153, 2018.

VIVALDINI, K. C. T.; V., G.; OLIVEIRA, M. D. C.; MARTINELLI, T. H.; WOLF, D. F.; RAMOS, F. Route planning for active classification with uavs. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. [S.l.: s.n.], 2016. p. 2563–2568.

WANG, C.; MA, H.; CHENC, W.; LIU, L.; MENG, M. Q.-H. Efficient autonomous exploration with incrementally built topological map in 3d environments. *IEEE Transactions on Instrumentation and Measurement*, IEEE, 2020.

WEINSTEIN, A.; SCHUMACHER, C. Uav scheduling via the vehicle routing problem with time windows. p. 17, 01 2007.

WILLIAMS, C. K.; RASMUSSEN, C. E. *Gaussian processes for machine learning*. [S.l.]: MIT press Cambridge, MA, 2006.

YAMAUCHI, B. A frontier-based approach for autonomous exploration. In: IEEE. *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. Towards New Computational Principles for Robotics and Automation*. [S.l.], 1997. p. 146–151.

YANG, K.; GAN, S. K.; SUKKARIEH, A. Gaussian process-based rrt planner for the exploration of an unknown and cluttered environment with an uav. *Advanced Robotics*, v. 27, 2013.

YANG, K.; GAN, S. K.; SUKKARIEH, S. A gaussian process-based rrt planner for the exploration of an unknown and cluttered environment with a uav. *Advanced Robotics*, Taylor & Francis, v. 27, n. 6, p. 431–443, 2013.

YUAN, S. et al. Marine environmental monitoring with unmanned vehicle platforms: Present applications and future prospects. *Science of The Total Environment*, Elsevier, v. 858, p. 159741, 2023.

Zhang, M.; Lan, M.; Lin, J.; Wang, S.; Liu, K.; Lin, F.; Chen, B. M. A high fidelity simulator for a quadrotor uav using ros and gazebo. In: *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*. [S.l.: s.n.], 2015. p. 002846–002851.

ZHU, H.-Y.; NIU, Y.-F.; SHEN, L.-C.; ZHANG, G.-Z. State of the art and trends of autonomous control of uav systems. *Journal of National University of Defense Technology*, China National Publishing Industry Trading Corporation, No. 504 Anhuali . . . , v. 32, n. 3, p. 115–120, 2010.

# Apêndice A

## I

---

---

Seguem as principais publicações relacionadas a esta tese.

- SANTOS, Marcela Ap Aniceto; VIVALDINI, Kelen Cristiane Teixeira. A Review of the Informative Path Planning, Autonomous Exploration and Route Planning Using UAV in Environment Monitoring. In: 2022 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2022. p. 445-450.
- SANTOS, Marcela Ap Aniceto; VIVALDINI, Kelen Cristiane Teixeira. A Modeling Strategy Using Bayesian Optimization with POMDP for Exploration and Informative Path Planning of UAVs in Monitoring of Forest. In: 2022 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2022. p. 451-456.

Outras Publicações:

- Rocha, L., Santos, M. A., Araújo, I., Vivaldini, K. (2021, June). A uav global planner to improve path planning in unstructured environments. In 2021 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 688-697). IEEE.

Submissões de outros artigos:

- Title: Explorative and Informative Path Planning in Monitoring Environments with UAV: A BO-POMDP Approach

# A Review of the Informative Path Planning, Autonomous Exploration and Route Planning using UAV in Environment Monitoring

Marcela Ap. Aniceto dos Santos  
Department of Computer Science.  
Federal University of São Carlos)  
São Carlos, Brazil  
marcelaaniceto@estudante.ufscar.br  
Full/Regular Research Paper - CSCI-RTCS

Kelen Cristiane Teixeira Vivaldini  
Department of Computer Science.  
Federal University of São Carlos)  
São Carlos, Brazil  
vivaldini@ufscar.br

**Abstract**—Unmanned Aerial Vehicles (UAVs) have been used in several applications for monitoring environments and mapping. To carry out the mapping of these environments, the UAV needs to decide which path to follow to collect as much information about the environment to maximize the search area. In the literature, these issues are being addressed within the area of Informative Path Planning (IPP), Route Planning (PR) and Autonomous Exploration of environments. As a way to clarify these problems and their objectives in robotics, this article aims to present a comprehensive review on these areas highlighting their approaches. For this, a comprehensive review of the main existing methods to solve them was carried out and this study serves as a starting point and a guide for everyone interested in exploring the monitoring area for data acquisition in unknown environments.

**Index Terms**—UAV, environment monitoring, informative path planning, route planning, autonomous exploration.

## I. INTRODUCTION

UAVs have resources related to flexibility, security, ease of operation and low cost of ownership. These characteristics have facilitated the adoption of these vehicles in several areas, being used successfully in applications of traffic monitoring [1], monitoring of diseases in eucalyptus [2], monitoring in agriculture [3], [4] search and rescue scenarios [5], industrial inspection [6], environmental disaster areas [7], [8] surveillance missions, among other applications [9]–[12].

The monitoring area for data acquisition in unknown environments has been the study of many researches in the field of robotics, in these applications, it is necessary that UAVs visit a certain area, analyze the environment and make a decision on which path it should following to maximize the search area in order to collect information and/or knowledge about the environment minimizing uncertainties or the location of certain patterns.

In the literature, we found that these problems have been addressed in three areas of robotics research: Informative Path Planning, Route Planning and Autonomous Environment Exploration. These researches, despite being different, solve the same problem by giving different approaches and focus.

The Informative Path Planning problem is autonomous decision-making to define which route the UAV should follow to collect information about the environment. This way, paths need to be planned to maximize the information gathered about an unknown environment while satisfying the given budget constraint [13].

In the IPP approach, the researches on monitoring unknown environments using UAV have adopted some methods, such as: Bayesian Optimization (BO) using Gaussian Processes (GP), to collect a set of information sequentially considering a set of specific constraints for a given problem. [14] and [15] consider the BO-POMDP (Partially Observable Markov Decision Process) formulation to perform sequential decision-making under uncertainty [14]–[16]. Other approaches are Covariance Matrix Evolution Strategy - CMA-ES, Interior Point - IP, Simulated Annealing - SA [13], [17], [18] among other methods has been adopted.

Route Planning using UAV can be considered a variant of the classic problem in the literature, which is the Vehicle Routing Problem (VRP) [19]. For decision-making regarding finding optimal routes in unknown environments, UAV Route Planning aims to monitor the environment in order to maximize the visited area, increasing knowledge about the environment or minimizing uncertainties [2]. Some methods used for route planning are BO using GP [20] [2], Rapidly-exploring Random Tree - RRT [21], SA [22], among others.

In Autonomous Exploration, the problem of path planning in unknown environments is to produce a consistent representation of the environment. Autonomous Exploration also involves decision-making, selecting the trajectories that a robot should follow to minimize the overall uncertainty in the model and maximize the gain of information about the environment. Essentially, exploration is a path optimization procedure for finding trajectories that efficiently learn the environment [23] [24].

Some of the adopted methods are algorithms based on Occupancy Grid Mapping: Octomap with Road Map [25] [26] and Graph Simultaneous Localization and Mapping - SLAM

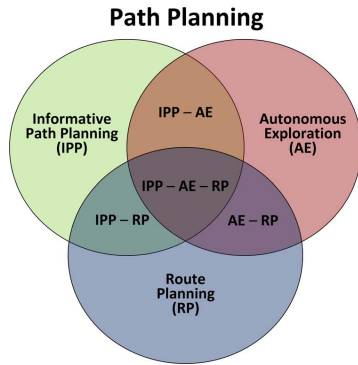


Fig. 1. The Informative Path Planning, Autonomous Exploration and Route Planning Problem's.

[27] to perform the environment mapping.

Algorithms based on Occupancy Continuous Mapping are also considered, such as Hilbert maps together with BO and Stochastic Gradient Descent [28] [29] that also perform the environment mapping. However, the representation of the occupation states is continuous.

It can be noticed that the IPP, RP and AE problems have in common the search for decision-making involving UAVs to obtain optimal routes in unknown environments. However, they consider different restrictions and different approaches as a solution.

Therefore, as a way to analyze these approaches and provide a review for robotics, this article contributes to the understanding of Informative Path Planning, Autonomous Exploration and Route Planning using UAVs in unknown environments. The concepts and differences between these problems, as well as the main methods adopted to solve them will be presented.

The article is divided as follows. In Section II the description of the problems: Informative Path Planning, Autonomous Exploration and Route Planning. In Section III, an overview of the literature highlighting the main methods to solve the problems. And finally, in Section IV, the conclusion of this comprehensive review.

## II. PROBLEM DESCRIPTIONS

This section presents the description of the Informative Path Planning, Autonomous Exploration and Route Planning problems based on the definitions given by the work carried out in the area of robotics research.

As shown in Figure 1, the Informative Path Planning, Route Planning and Autonomous Exploration problems are highlighted. The intersection between them form subcategories such as: IPP - AE [12], IPP - RP [2], AE - RP [25] and IPP - AE - RP. These subcategories are the union of the IPP, AE and RP problem's.

### A. Informative Path Planning - IPP

Based on the works in the literature [14], [17], [30]–[32], IPP is characterized by the lack of knowledge about the environment a priori. In this case, it is necessary to obtain a map of the environment or get a graph.

According to [30], if the scientist manually specifies the robot's exact trajectory while collecting sensor measurements, the problem will be relatively simple. Thus, the robot autonomously decides which path to follow during the collection of measurements, based on a probabilistic model of a dataset to be studied is known as IPP. Another definition is that the task of choosing trajectories to maximize information gain is known as informative path planning and is a fundamental monitoring concept [15]. In the same sense as the previous work [32] [32] is defined the IPP.

Another definition of IPP is designing the route of a vehicle, which must follow in such a way that a certain goal is maximized and a goal is achieved. The IPP maximizes the information collected from targets in a region of interest (ROI) [9].

So, IPP is an NP-hard optimization problem and has a trade-off between map completeness and practical efficiency. This trade-off is related to the time to exhaustively monitor a large area to get as much information about the environment and search for an optimal global solution [18], [31], [32].

To carry out IPP, it is necessary to calculate the route to obtain the maximum of gain information about the environment, not being concerned at first with the movement that the UAV will make to carry out the trajectory, but with which route will be taken to collect the information about the environment. This must be done by monitoring as much area as possible and thus maximizing the information gains collected from the environment.

### B. Autonomous Exploration - AE

In the same way the IPP, the AE is characterized by a lack of knowledge about the environment.

Thus, according to the works in the literature, they treat the problem of trajectory planning in unknown environments as an AE that aims to produce a consistent representation of the environment. AE also involves making complex decisions, selecting the trajectories that a robot will follow to maximize the information collected about the environment and also minimize the overall uncertainty of model the [24], [33].

Other authors emphasize that AE is a major precondition for building a map of an unknown environment with a robot to provide the data of interest. However, achieving this task efficiently in large-scale or high-dimensional environments is still challenging. [25], [26].

According to [24] Autonomous exploration can be seen as active learning that aims to minimize uncertainty and produce high-fidelity maps [34], [35], where exploration requires the simultaneous solution of mapping, path planning and location.

So, the AE is a high-level task dedicated to building the model of an unknown environment in which the vehicle frequently makes decisions to select the trajectories that a

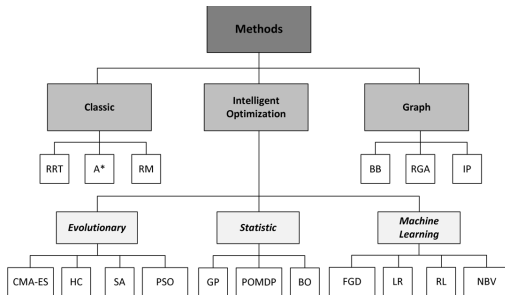


Fig. 2. The Taxonomy of analyzed methods for IPP, AE and RP in Monitoring Environment.

robot will follow to maximize the information collected about the environment and also minimize the uncertainty of model [25].

### C. Route Planning - RP

When classifying routing literature, it can be segregated based on the problem type with an emphasis on the VRP, which has given the major research contributions in the domain of vehicle routing [36], [37] and is used as an input for all the routing problems in general [19]. So, when considering de UAV Route Planning monitoring the environment in order to maximize the visited area, increasing knowledge about the environment or minimizing uncertainties there is a similar problem with IPP.

In RP in unknown environments, the target is known, but the position of these targets is unknown. So an area must be tracked without knowing the location of the targets and therefore, it is necessary for the UAV to make a decision to identify these targets and plan its route, regardless of human support, this approach is known as route planning for an active classification. Thus, for decision making it is necessary to verify the information and extract the relevant data for the optimization of an efficient route planning according to each application [2].

Route planning is an important part of the unmanned aerial vehicle mission planning system [21]. There are many uncertain factors in the task environment of UAV, or the UAV is flying in a completely unknown environment [38]. At this time, the off-line route planning method is no longer applicable. The online route planning method can be used to generate a feasible route based on real-time detection of environmental information [39], [40].

## III. LITERATURE REVIEW

This section presents a brief section on the challenges of planning missions for efficient data acquisition. A state-of-the-art overview highlighting the main methods developed to solve Informative Path Planning, Exploration and Route Planning problems.

As shown in the figure 2 was done the taxonomy of analyzed methods for IPP, AE and RP in the monitoring environment for the literature review. The methods were divided into classical, graph-based and intelligent optimization categories (Evolutionary, Statistic-based and Machine Learning).

### A. Challenges

The monitoring area for efficient data acquisition in unknown environments has been the study of several researches in the field of robotics. According to the literature, it can be seen that the problems of Informative Path Planning, Route Planning and Autonomous Exploration address the same problem, but with different approaches.

So, there are still some challenges of how to plan missions to obtain efficient data in complex and unknown environments. To solve these problems in UAV path planning, it is necessary to make optimal decisions for various mission-critical operations performed by UAVs. These decisions require a map or graph of the mission environment so that UAVs are aware of their locations or close to their target/objective [41]. Keeping the focus on the aforementioned points, this reviews several UAV informative planning techniques used in recent years. The objective of the techniques is not only to find an ideal and shortest path but also to provide the environment map, as the environment is unknown in the cases of the informative path planning, autonomous exploration, and route planning problems discussed in this work.

### B. Methods used for Informative Path Planning

In the search for solutions for IPP there are several methods proposed in the literature. Graph-based algorithms can be used to find trajectories to gain information about the environment [30], [42], [43]. Mixed Integer Linear Programming problems are generally solved using a linear-programming based branch-and-bound algorithm.

The key to Branch and Bound - BB [30] methods is finding an easily computable upper bound for the objective function. If the upper bound is loose (is often much higher than the actual objective), then a few branches of the search tree will be pruned. The goal is to find a function which is as tight as possible, while still being a valid upper bound.

The Branch and Bound was used by [30] to solve the IPP. The algorithm uses the objective function monotonicity to provide an acceleration dependent on the objective function versus the Brute Force Search. Results that suggest that by maximizing the reduction of variance in a Gaussian process model, the acceleration of the algorithm is significant were presented. To validate the algorithm, the execution time in different scenarios was analyzed.

Interior Point - IP is a certain class of algorithms that solve linear and nonlinear optimization problems [44]. The IP was used to approximate gradient-based optimization [13].

It can be noticed that in situations where there is a large dataset, graph-based methods are limited in terms of problem-solving. In these situations, these methods have exponential growth, which makes it difficult to find a solution.

Recursive Greedy Algorithm - RGA [43] can also be used, however, they tend to converge to local optimal.

Another methods are Statistics-based BO is a global optimization technique that possesses major advantages when used to find the maximum of partially observed objective functions that are costly to evaluate, lack gradient information, and can only be inferred indirectly from noisy observations [45].

BO is robust to this setting because it builds a statistical model over the objective. More specifically, it places a prior over the space of functions and combines it with noisy samples to produce an incremental prediction for the unknown function. The prior usually takes the form of a Gaussian Process (GP) [46]. The key component for the effectiveness of BO is the use of an Acquisition Function (AF) that guides the search for the optimum by selecting the locations where samples are gathered based on the posterior in each iteration.

In this way, Path-planning algorithms for environment exploration come in two flavors. Approaches in which the UAV decides on its next move one step at a time are referred to as myopic [15], [47].

Myopic algorithms are suitable for most situations but lack a mechanism for anticipation, which may be problematic in cases where path-planning decisions may have negative long-term consequences.

The main tool for this is the partially observable Markov decision process, which assigns a reward to each admissible sequence of actions. Nonmyopic approaches are computationally complex and incredibly expensive, which is why myopic approaches are often preferred [12].

One approach that we can highlight is the one developed by [13], as they are adaptive and non-adaptive strategies using CMA-ES. Non-adaptive approaches explore an environment using a pre-determined sequence of actions to execute the route [48]. Adaptive approaches allow routes to change as information is collected, making them suitable for planning based on specific interests [13], [17], [31], [32].

The CMA-ES is a generic global optimization routine based on the concepts of evolutionary algorithms which has been successfully applied to high-dimensional, nonlinear, non-convex problems in the continuous domain. As an evolutionary strategy, the CMA-ES operates by iteratively sampling candidate solutions according to a multivariate Gaussian distribution in the search space

As shown in Table I some methods are used for IPP in environment monitoring and application.

### C. Methods used for Autonomous Exploration

The BO method [24] is also was used in an approach for AE and for building maps. This method finds optimal continuous paths rather than discrete detection locations that satisfy UAV security and motion constraints. By balancing the reward function and the risk associated with each path, the optimizer minimizes the number of function evaluations that are computationally costly.

Another method can be used is the Functional Gradient Descent (FGD) to efficiently optimize the exploratory paths

TABLE I  
THE METHODS USED FOR IPP IN ENVIRONMENT MONITORING.

| Year | Author               | Methods  | Application                     |
|------|----------------------|----------|---------------------------------|
| 2014 | Marchant and Ramos   | BO-GP    | Environment monitoring          |
| 2016 | Lim, Hsu and Lee     | RAId     | Disaster region                 |
| 2017 | Hitz et al.          | CMA-ES   | Monitoring toxic cyanobacteria  |
| 2022 | Blanchard and Sapsis | BO-POMDP | Anomaly detection in monitoring |

TABLE II  
THE METHODS USED FOR AE IN ENVIRONMENT MONITORING.

| Year | Author            | Methods    | Application                             |
|------|-------------------|------------|---|
| 2015 | Rossi             | HC         | Monitoring for gas leakage localization |
| 2019 | Francis and Ramos | RL and FGD | Exploration                             |
| 2020 | Wang et al.       | NBV        | Exploration                             |

in continuous occupancy maps [23]. Stochastic FGD was adopted to overcome the limitations of standard FGD methods in order to ensure convergence. This process allows for the optimization of the entire path, resulting in continuous smooth paths that maximize the overall map quality, keeping the robot safe from collisions. In addition, the results were compared with exploration methods such as RRT planner [49] and Frontier [50].

A method based on evolutionary algorithm [51] is an approach for monitoring environment in a gas leak location. The profile of gas concentration measurements was modeled using a 2D Gaussian distribution model and the search was performed applying an exploration strategy based on the Hill-Climbing. The gas source location strategy optimizes the speed of the aerial robot while minimizing the monitoring system's energy consumption.

It is interesting to observe in the work of [51], that despite the authors considering the proposed algorithm as an exploratory strategy that aims to build a map of the environment for locating gas leaks. It could be presented as IPP since the information collected from the environment as its main mission is the detection of gas in an unknown environment and for this to happen autonomously, a map is built online so that the UAV can locate itself in the environment.

Another approach to efficient UAV exploration was proposed by [25], where a Road Map - RM was built incrementally along with the exploration process that explicitly displays the topological structure of the 3D environment. By simplifying the environment, a road map can efficiently provide the information gain and cost for a candidate region to be explored, which are two quantities for Next-Best-View (NBV) evaluation, suggesting thus the efficiency for determining the NBV. In addition, a local planner was made based on the Potential Fields method that drives the robot to the information-rich area during the navigation process.

As shown in Table III some methods used for Autonomous Exploration in environment monitoring



TABLE III  
THE METHODS USED FOR RP IN ENVIRONMENT MONITORING.

| Year | Author           | Methods | Application                      |
|------|------------------|---------|----------------------------------|
| 2016 | Meng et al.      | RRT     | Environment monitoring           |
| 2019 | Vivaldini et al. | BO-RRT  | Classification of diseased trees |
| 2021 | Hari et al.      | TSP     | Persistent monitoring missions   |

#### D. Methods used for Route Planning

The methods classics can be used for RP like Rapidly-exploring Random Tree (RRT) algorithm is an efficient path planning method, which can quickly find the feasible solution in complex environment [21].

The RRT with GP-generated occupancy maps was used to explore unknown environments [52] developed an algorithm that combines Rolling. The method consists of a route planner that collects information about a search area, focusing on parts with greater uncertainty and following a disorganized geometry.

Another algorithm based on the RRT for the UAV route planning problem in an unknown environment [21]. According to the current information from the environment, the local route planning is carried out at the same time, the new information from the environment is detected and the next stage of the route is generated.

A route planning methodology with active classification for UAVs, in order to increase knowledge of the visited areas and minimize uncertainties in the classification of diseased trees [20]. Five different route planning algorithms were evaluated, continuous BO, discrete BO, Random points, pre-established trajectory and RRT. The main advantage of continuous BO is the combination of RP (building a route between the origin and objective points) and active classification (which allows you to choose objectives based on how they affect the uncertainty of the environment).

As an extension of [20], [2] developed a route planning framework for active classification using UAVs to maximize the information collected within a given distance, limited by flight range. The authors proposed BO+RRT, where BO chooses the destination points and the RRT suggests the trajectory between these points. The Logistic Regression classifier was used to classify the sick/healthy/soil trees and a GP was used to interpolate this information, producing a navigation map.

It is interesting to note that the work of [20] and [2] could be presented as IPP since, for decision making which route to take to collect information from the environment, it was done to cover as much area as possible to maximize the gains information collected about an environment.

As shown in Table III some methods used for Route Planning in environment monitoring.

#### IV. CONCLUSION

This article analyzed the state of the art in the monitoring area for efficient data acquisition in unknown environments, addressing its main challenges. Informative Path Planning,

Route Planning and Autonomous Exploration were theoretically clarified and the main approaches in this area were presented.

According to the works analyzed in the literature, it can be seen that the IPP, PR, and AE problems presented, treat Informative Planning as an autonomous decision-making problem of the aerial robot to know which route to take to collect information about the unknown environment.

However, there are some differences in the approach given, IPP and RP in unknown environments have as main objectives the collection of information, such as temperature, and humidity, among others, using an onboard sensor. AE, on the other hand, focuses on the construction of the map, that is, producing a consistent representation of an unknown environment.

It can be noticed that the three problems consider that the trajectory that the UAV must take in order to the mission it should be done in such a way as to cover as much area as possible to maximize the gains of information collected about an unknown environment and minimize model uncertainties. For this, different restrictions are used, such as maximum monitoring time, battery, distance, collision avoidance, curve smoothing, among others.

Regarding the approaches presented, although they are presented in different ways, the methods used to solve the IPP, PR, and AE problems are solved in similar ways with similar methods.

#### ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

#### REFERENCES

- [1] K. Kanistras, G. Martins, M. J. Rutherford, and K. P. Valavanis, "A survey of unmanned aerial vehicles (uavs) for traffic monitoring," in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2013, pp. 221–234.
- [2] K. C. Vivaldini, T. H. Martinelli, V. C. Guizilini, J. R. Souza, M. D. Oliveira, F. T. Ramos, and D. F. Wolf, "Uav route planning for active disease classification," *Autonomous Robots*, vol. 43, no. 5, pp. 1137–1153, 2018.
- [3] J. G. A. Barbedo, "A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses," *Drones*, vol. 3, no. 2, p. 40, 2019.
- [4] K. Srivastava, A. J. Bhutoria, J. K. Sharma, A. Sinha, and P. C. Pandey, "Uavs technology for the development of gui based application for precision agriculture and environmental research," *Remote Sensing Applications: Society and Environment*, vol. 16, p. 100258, 2019.
- [5] H. Oleynikova, Z. Taylor, A. Millane, R. Siegart, and J. Nieto, "A complete system for vision-based micro-aerial vehicle mapping, planning, and flight in cluttered environments," *arXiv preprint arXiv:1812.03892*, 2018.
- [6] W. Kwon, J. H. Park, M. Lee, J. Her, S.-H. Kim, and J.-W. Seo, "Robust autonomous navigation of unmanned aerial vehicles (uavs) for warehouses' inventory application," *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 243–249, 2019.
- [7] F. Malandrino, C.-F. Chiasserini, C. Casetti, L. Chiaraviglio, and A. Senacheribbe, "Planning uav activities for efficient user coverage in disaster areas," *Ad Hoc Networks*, vol. 89, pp. 177–185, 2019.
- [8] S. Schopferer and S. Benders, "Minimum-risk path planning for long-range and low-altitude flights of autonomous unmanned aircraft," in *AIAA Scitech 2020 Forum*, 2020, p. 0137.

- [9] H. Ergezer and K. Leblebicioğlu, "Online path planning for unmanned aerial vehicles to maximize instantaneous information," *International Journal of Advanced Robotic Systems*, vol. 18, no. 3, p. 17298814211010379, 2021.
- [10] T. Elmokadem and A. V. Savkin, "A hybrid approach for autonomous collision-free uav navigation in 3d partially unknown dynamic environments," *Drones*, vol. 5, no. 3, p. 57, 2021.
- [11] X. Zhu, X. Zhu, R. Yan, and R. Peng, "Optimal routing, aborting and hitting strategies of uavs executing hitting the targets considering the defense range of targets," *Reliability Engineering & System Safety*, p. 107811, 2021.
- [12] S. K. K. Hari, S. Rathinam, S. Darbha, K. Kalyanam, S. G. Manyam, and D. Casbeer, "Optimal uav route planning for persistent monitoring missions," *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 550–566, 2021.
- [13] P. Marija, T. Vidal-Calleja, G. Hitz, C. J. Jen, S. Inkyu, R. Siegwart, and J. Nieto, "An informative path planning framework for uav-based terrain monitoring," *Autonomous Robots*, vol. 44, no. 6, pp. 889–911, 2020.
- [14] R. Marchant and F. Ramos, "Bayesian optimisation for informative continuous path planning," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 6136–6143.
- [15] P. Morere, R. Marchant, and F. Ramos, "Sequential bayesian optimization as a pomdp for environment monitoring with uavs," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 6381–6388.
- [16] —, "Continuous state-action-observation pomdps for trajectory planning with bayesian optimisation," 10 2018, pp. 8779–8786.
- [17] G. Hitz, E. Galceran, M.-E. Garneau, F. Pomerleau, and R. Siegwart, "Adaptive continuous-space informative path planning for online environmental monitoring," *Journal of Field Robotics*, vol. 34, no. 8, pp. 1427–1449, 2017.
- [18] M. Popović, T. Vidal-Calleja, G. Hitz, I. Sa, R. Siegwart, and J. Nieto, "Multiresolution mapping and informative path planning for uav-based terrain monitoring," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 1382–1388.
- [19] A. Thibbotuwawa, G. Bocewicz, P. Nielsen, and Z. Banaszak, "Unmanned aerial vehicle routing problems: a literature review," *Applied Sciences*, vol. 10, no. 13, p. 4504, 2020.
- [20] K. C. T. Vivaldini, V. Guizilini, M. D. C. Oliveira, T. H. Martinelli, F. Ramos, and D. F. Wolf, "Route planning for active classification with uavs," in *2016 - IEEE Int. Conf. on Robotics and Automation (ICRA)*, Stockholm, Sweden, 2016.
- [21] L. Meng, S. Qing, Z. Qijun, and Z. Yongliang, "Route planning for unmanned aerial vehicle based on rolling rrt in unknown environment," in *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICIC)*. IEEE, 2016, pp. 1–4.
- [22] S. Capraz, H. Azyikmis, and A. Ozsoy, "An optimized gpu-accelerated route planning of multi-uav systems using simulated annealing," *International Journal of Machine Learning and Computing*, vol. 10, no. 3, 2020.
- [23] G. Francis, L. Ott, and F. Ramos, "Functional path optimisation for exploration in continuous occupancy maps," in *Robotics Research*. Springer, 2020, pp. 859–875.
- [24] G. Francis, L. Ott, R. Marchant, and F. Ramos, "Occupancy map building through bayesian exploration," *The International Journal of Robotics Research*, vol. 38, no. 7, pp. 769–792, 2019.
- [25] C. Wang, H. Ma, W. Chenc, L. Liu, and M. Q.-H. Meng, "Efficient autonomous exploration with incrementally built topological map in 3d environments," *IEEE Transactions on Instrumentation and Measurement*, 2020.
- [26] H. Li, A. V. Savkin, and B. Vucetic, "Autonomous area exploration and mapping in underground mine environments by unmanned aerial vehicles," *Robotica*, vol. 38, no. 3, p. 442–456, 2020.
- [27] H. Carrillo, P. Dames, V. Kumar, and J. A. Castellanos, "Autonomous robotic exploration using occupancy grid maps and graph slam based on shannon and rényi entropy," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 487–494.
- [28] F. Ramos and L. Ott, "Hilbert maps: scalable continuous occupancy mapping with stochastic gradient descent," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1717–1730, 2016.
- [29] R. Senanayake and F. Ramos, "Bayesian hilbert maps for dynamic continuous occupancy mapping," in *Conference on Robot Learning*, 2017, pp. 458–471.
- [30] J. Binney and G. S. Sukhatme, "Branch and bound for informative path planning," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 2147–2154.
- [31] Z. W. Lim, D. Hsu, and W. S. Lee, "Adaptive informative path planning in metric spaces," *The International Journal of Robotics Research*, vol. 35, no. 5, pp. 585–598, 2016.
- [32] M. Popovic, T. Vidal-Calleja, J. J. Chung, J. Nieto, and R. Siegwart, "Informative path planning and mapping for active sensing under localization uncertainty," *arXiv preprint arXiv:1902.09660*, 2019.
- [33] P. Morere, G. Francis, T. Blau, and F. Ramos, "Reinforcement learning with probabilistically complete exploration," *arXiv preprint arXiv:2001.06940*, 2020.
- [34] A. A. Makarenko, S. B. Williams, F. Bourgault, and H. F. Durrant-Whyte, "An experiment in integrated exploration," in *IEEE/RSJ international conference on intelligent robots and systems*, vol. 1. IEEE, 2002, pp. 534–539.
- [35] C. Stachniss, *Robotic mapping and exploration*. Springer, 2009, vol. 55.
- [36] B. Eksioglu, A. V. Vural, and A. Reisman, "The vehicle routing problem: A taxonomic review," *Computers & Industrial Engineering*, vol. 57, no. 4, pp. 1472–1483, 2009.
- [37] K. Braekers, K. Ramaekers, and I. Van Nieuwenhuysse, "The vehicle routing problem: State of the art classification and review," *Computers & Industrial Engineering*, vol. 99, pp. 300–313, 2016.
- [38] H.-Y. Zhu, Y.-F. Niu, L.-C. Shen, and G.-Z. Zhang, "State of the art and trends of autonomous control of uav systems," *Journal of National University of Defense Technology*, vol. 32, no. 3, pp. 115–120, 2010.
- [39] S. Lai, K. Wang, H. Qin, J. Q. Cui, and B. M. Chen, "A robust online path planning approach in cluttered environments for micro rotorcraft drones," *Control Theory and Technology*, vol. 14, no. 1, pp. 83–96, 2016.
- [40] M. Lavalle and S. J. J. Kuffner, "Rapidly-exploring random trees: Progress and prospects," in *Proc. Workshop on the Algorithmic Foundations of Robotics*, San Francisco, 2000.
- [41] S. Aggarwal and N. Kumar, "Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges," *Computer Communications*, 2019.
- [42] A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, pp. 707–755, 2009.
- [43] J. Binney, A. Krause, and G. S. Sukhatme, "Informative path planning for an autonomous underwater vehicle," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 4791–4796.
- [44] I. Dikin, "Iterative solution of problems of linear and quadratic programming," in *Doklady Akademii Nauk*, vol. 174, no. 4. Russian Academy of Sciences, 1967, pp. 747–748.
- [45] E. Brochu, V. M. Cora, and N. De Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," *arXiv preprint arXiv:1012.2599*, 2010.
- [46] M. Seeger, "Gaussian processes for machine learning," *International journal of neural systems*, vol. 14, no. 02, pp. 69–106, 2004.
- [47] R. Marchant, F. Ramos, and S. Sanner, "Sequential bayesian optimisation for spatial-temporal monitoring," *Uncertainty in Artificial Intelligence - Proceedings of the 30th Conference, UAI 2014*, pp. 553–562, 01 2014.
- [48] E. Galceran and M. Carreras, "A survey on coverage path planning for robotics," *Robotics and Autonomous systems*, vol. 61, no. 12, pp. 1258–1276, 2013.
- [49] K. Yang, S. Keat Gan, and S. Sukkariéh, "A gaussian process-based rrt planner for the exploration of an unknown and cluttered environment with a uav," *Advanced Robotics*, vol. 27, no. 6, pp. 431–443, 2013.
- [50] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97: Towards New Computational Principles for Robotics and Automation*. IEEE, 1997, pp. 146–151.
- [51] M. Rossi and D. Brunelli, "Autonomous gas detection and mapping with unmanned aerial vehicles," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 4, pp. 765–775, 2015.
- [52] K. Yang, S. K. Gan, and A. Sukkariéh, "Gaussian process-based rrt planner for the exploration of an unknown and cluttered environment with an uav," *Advanced Robotics*, vol. 27, 2013.

## A Modeling Strategy using Bayesian Optimization with POMDP for Exploration and Informative Path Planning of UAVs in Monitoring of Forest

Marcela Ap. Aniceto dos Santos  
*Department of Computer Science.*  
*Federal University of São Carlos*  
*São Carlos, Brazil*  
 marcelaaniceto@estudante.ufscar.br

Kelen Cristiane Teixeira Vivaldini  
*Department of Computer Science.*  
*Federal University of São Carlos*  
*São Carlos, Brazil*  
 vivaldini@ufscar.br

**Abstract**—Unmanned Aerial Vehicles (UAVs) have been used for several applications in monitoring complex and unknown environments. The challenge is to plan missions for the UAV in situations where the vehicle needs to visit and explore an area and analyze it in real-time to define the route to be followed. This visit occurs with the search area's maximization from the definition of the trajectories, making it possible to collect information to acquire knowledge about the environment and provide a map. This type of problem is known as Informative Path Planning (IPP) and Autonomous Exploration (AE). In this context, Bayesian Optimization (BO) has been adopted. Moreover, there is a need to define the planning for decision-making based on information about the environment considering specific restrictions. Partially Observable Markov Decision Processes (POMDP) can be used to define the planner responsible for the decision-making. Therefore, considering these methods and the strategy of exploring the environment in a continuous 3D space, this paper proposes a development of a modeling strategy that is exploration and informative path planning using Sequential Bayesian Optimization with POMDP in forest monitoring with Canopy gap. A sequential decision-maker was developed under uncertainty based on the Sequential Bayesian Optimization approach with Partially Observable Markov Decision Process (BO-POMDP).

**Index Terms**—UAV, informative path planning, autonomous exploration, bayesian optimization, POMDP.

### 1. Introduction

Currently, the increased autonomy in the decision-making process in the Unmanned Aerial Vehicles (UAVs) have been providing growth in their use in several areas of exploring and monitoring, such as traffic monitoring [1], monitoring of diseases in eucalyptus [2], agriculture [3] [4], search and rescue scenarios [5], industrial inspection [6], areas of environmental disaster [7], water quality monitoring [8], surveillance missions [9], among other applications.



(a) Deforestation



(b) Canopy Gap

Figure 1: Collection of images from National Geographic Brazil and Instituto Socio Ambiente.

As can be observed, the environments monitoring has been subject to a lot of research in the robotics area.

One of these subject is found in forests with canopy gaps is the biological invasion carried out by exotic plants. This type of invasion is threatening the conservation of biodiversity by causing the displacement of native species and, consequently, influencing the modification of the functioning of [10] ecosystems. In 1 has some examples in forests with canopy gap.

The formation of canopy gaps can be caused by the fall of branches, which generate less opening in the canopy and destruction of the understory; by the death of standing trees, whose crowns are gradually fragmented; and by uprooting the trees, which, in addition to causing destruction in the understory, cause soil disturbance in the root region [11]. Other possible causes for gap formation can be found in forests with deforestation and environmental disasters.

For these applications, it is necessary for UAVs to visit a certain area, analyze the environment and make a decision on which route it should take to maximize the search area for the collection of information and/or knowledge about the environment minimizing uncertainties or the location of certain targets or patterns. Depending on the type of monitoring, the UAV can be equipped with appropriate

sensors to measure/capture the target/patterns and provide information/observation necessary for the application.

In the literature, these problems are being addressed in 2 different ways in the robotic area, found in Informative Path Planning (IPP) and Autonomous Exploration (AE) of Environments research. Thus, despite different research, the same problem is solved using different approaches.

IPP is an autonomous decision-making problem to know which route the UAV must take to collect information related to the environment [12] [13]. Thus, the path planning must be done to cover as much area as possible to maximize the gains information collected about an unknown environment [14] [15]. In the IPP approach, some researches have adopted to solve the problem the Bayesian Optimization (BO) using Gaussian Process [16] [13].

In Autonomous Exploration, the path planning problem in environment monitoring aims to produce a consistent representation of the environment. Autonomous Exploration also involves decision-making, selecting the paths that a robot should follow to minimize the general uncertainty of the model and maximize the gain of information about the environment [17] [18]. Some of the methods adopted are algorithms based occupancy Continuous Mapping have also been used, such as algorithms Bayesian and Stochastic Gradient Descent [19] [20].

We can observe that both problems to carry out path planning in environments monitoring, the UAV needs to visit a certain area, analyze the environment and make a decision on which route it must take to maximize the search area aiming at the collection of information of a target.

In this context, planning under uncertainty is critical to robotics. So, the Partially Observable Markov Decision Process (POMDP) is a mathematical framework for such planning problems. It is powerful due to its careful quantification of the non-deterministic effects of actions and partial observability of the states [21].

Therefore, the focus of this paper is the development of a modeling strategy that is an exploration and informative path planning using Sequential Bayesian Optimization with POMDP in forest monitoring with canopy gap. For this, continuous trajectories that maximize the gain of information with the location of possible targets will be considered. In this approach, a sequential decision maker under uncertainty will be developed based on the BO-POMDP approach.

The contributions of this work are:

- 1 Development of the modeling strategy using Sequential Bayesian Optimization with POMDP for a sequential decision maker under uncertainty in order to carry out Exploration and Informative Path Planning in 3D space.
- 2 The strategy of exploring the environment as it focuses on specific areas of interest in the search for possible targets.

The rest of the paper is organized as follows. In section 2 related works, in section 3 a description of the Gaussian Process, Bayesian Optimisation, and then present the classic

POMDP formulation. In section 4, the proposal of this work is presented, and in section 5, the conclusion is presented.

## 2. Related Work

Several proposed methods are found in the literature in the search for solutions for Informative Path Planning (IPP). Graph-based algorithms can find trajectories to gain information about the environment. [22] [23] [12].

The Branch and Bound method was used by [12] to solve the IPP. The algorithm uses the objective function monotonicity to provide an acceleration dependent on the objective function versus the Brute Force Search. The results suggest that by maximizing the reduction of variance in a Gaussian process model, the acceleration of the algorithm is significant were presented. To validate the algorithm, the execution time in different scenarios was analyzed.

It can be noticed that in situations where there is a large data set, graph-based methods are limited in terms of problem solving. In these situations, the methods have exponential growth, which makes the search for a solution difficult.

Greedy methods can also be used, but they tend to converge on great locations [23].

One approach to Informative Path Planning was developed by [16], applied to environment monitoring using BO. The approach uses a layered algorithm based on the BO method, where the first layer optimizes the environmental phenomenon to find areas of interest (high pollution and high temperature). This layer uses an incremental spatial-temporal model of the phenomenon, given by a previous Gaussian Process [24], [25], which takes into account the uncertainty and predicted values propagated in time. The second layer is used to find the best set of parameters that determine a continuous path the robot takes. Algorithm validation was performed in a large-scale environment to monitor the ozone concentration in the US and in a mobile robot that monitors the dynamics of light changes.

Another approach using BO-POMDP was proposed by [13]. The improvement of this approach lies in the reformulation of the POMDP reward function to balance exploitation (high gradient areas) and exploration (high uncertainty areas). This means that encouraging the robot to exploit (high-gradient areas) is critical to achieving high-quality monitoring, as reconstructed maps generally have greater errors in complex and variable areas.

It was possible to notice that BO-GP has been used for data acquisition and to structure the environment map. BO-POMDP has been used for sequential decision-making under uncertainties in the context of IPP, obtaining significant results. However, the mathematical model of BO-POMDP was not defined for 3D environments in Exploration and IPP problems.

## 3. Background

We start with a description of Gaussian Process, Bayesian Optimisation, and then present the classic POMDP

formulation.

### 3.1. Gaussian process

A Gaussian process is a Bayesian technique for non-parametric regression and classification [24]. It is non-parametric because it does not maintain an explicit model of the underlying phenomenon but learns the transformation between input and output directly from training samples (observations). It is Bayesian in the sense that it places prior distributions (hypotheses) on observed variables and updates these distributions to produce a posterior distribution as new data become available. Formally a GP is defined:

**Definição 1.** A Gaussian Process is a collection of random variables where any finite number has a joint Gaussian distribution.

A GP is completely defined by the mean function  $m(x)$  and the covariance function  $K(x, x')$  of a real process  $f(x)$ , like the Equation 1.

$$m(x) = [f(x)] K(x, x') = [(f(x) - m(x))(f(x') - m(x')))] \quad (1)$$

Thus, a Gaussian Process can be represented by the Equation 2.

$$f(x) \sim GP(m(x), K(x, x')) \quad (2)$$

The function  $K$ , known as Kernel,  $x$  is the input of a dimensional space  $D$  such that  $x \in R^D$ .

In addition, an additive noise model can be assumed  $y_i = f(x_i) + \epsilon$  for noisy observations  $y$  de  $f$ , where  $\epsilon \sim iid N(0, \sigma_n^2)$  it is an independent gaussian noise.

Thus, given a  $N$  training set of input  $X = \{x_i\}_{i=1}^N$  and outputs  $Y = \{y_i\}_{i=1}^N$  both with  $i = 1 \dots N$ , we can calculate the predictive distribution of  $f$  at an unknown query location  $x^*$  by calculating the later  $p(f(x^*)|y, X, x^*)$ .

For a GP, this predictive distribution is Gaussian,  $f(x^*)N(\bar{f}^*, cov(f^*))$ , where:

$$\bar{f}^* = K(x^*, X)K_{-1}^X(y - m(X))$$

$$cov(f^*) = (x^*, x^*) - K(x^*, X)K_{-1}^X K(X, x^*) \quad (3)$$

$K(A, B)$  is the covariance matrix whose element  $(i, j)$  is calculated to be  $k_{i,j} = k(x_i, x_j)$ ,  $x_i \in A$  and  $x_j \in B$ ;

$K_X = K(X, X) + \sigma_n^2 I$  is the covariance matrix between observations, with  $I$  being the identity matrix.

It can be noticed that the kernel  $K(\cdot)$  determines the generalization properties of the GP model and is chosen to describe the characteristics of the environment. Different Kernels can be used. In the context of IPP, some are more applicable, such as: exponential square and functions Matérn [13] [15]. Equation 4 presents the square exponential kernel.

$$K_{SE}(x, x') = \sigma^2 e^{-\frac{(x-x')^2}{2l^2}} \quad (4)$$

### 3.2. Bayesian Optimisation

The BO algorithm uses a priori Gaussian Process and evidence (acquisition function) to define a posterior distribution along with the function space. The Bayesian model allows a sophisticated way for the a priori information (initial parameters of the covariance function of a GP and the same) to describe the attributes of the objective function, such as the smoothness or the positions most likely to be chosen (the maximum of the function), even when the objective function itself is not known [26].

Bayesian Optimisation is a technique aiming to find the optimum  $\hat{x} \in R^D$  of an objective function  $f : R^D \rightarrow R$  by gathering noisy observations from it. Then,

$$\hat{x} = \arg \max_x f(x). \quad (5)$$

Noisy observations are assumed to result from an additive Gaussian noise on the function evaluation. The  $i$ th observation is defined as  $y_i = f(x_i) + \epsilon$ , where  $\epsilon \sim iid N(0, \delta_n^2)$  is the noise associated with each independent observation.

Most implementations of BO use Gaussian processes to model the objective function  $f$ . The GP model is updated with the data couple  $(x_i, y_i)$  every time a noisy observation  $y_i$  is made at location  $x_i$ . The search of where to get an observation next  $h(x)$ . At each iteration in the BO algorithm, the location at which to evaluate  $f$  is determined by finding  $\arg \max_x h(x)$ , therefore reporting the maximization problem from  $f$  to  $h$ ,  $h$  is much easier to optimize with traditional techniques such as [27] and, contrary to  $f$ , is cheap to evaluate. An implementation of BO is detailed in algorithm 1.

---

#### Algorithm 1 Bayesian Optimisation

---

- 1: Let  $x_t$  be the sampling point at iteration  $t$ .
  - 2: Let  $D = (x_1, y_1), \dots, (x_n, y_n)$  be the data.
  - 3:  $h$  be an acquisition function
  - 4: **for**  $t = 1, 2, 3, \dots$  **do**
  - 5:     Find  $x_t = \arg \max_x h(x)$
  - 6:     Sample the objective function
  - 7:     Augment the data  $D$
  - 8:     Recompute the GP model with  $D$
  - 9: **end for**
- 

### 3.3. Partially Observable Markov Decision Processes

Partially observable Markov decision processes (POMDP) are a framework for decision-making under uncertainty [28]. Unlike in Markov decision processes, the agent cannot directly observe its current state. It must instead act relying on a belief of the underlying state, built from observations. It is, therefore, important for the agent to maintain a probability distribution over the set of possible states [13].

A POMDP is fully defined by the tuple  $\langle S, A, T, R, \Omega, O, \gamma \rangle$ , with:

- $S$ : Set of states  $\{s_1, s_2 \dots s_n\}$ .
- $A$ : Set of actions  $\{a_1, a_2 \dots a_m\}$ .
- $T : S \times A \times S \rightarrow [0, 1]$  is a transition function interpreted as the probability to transition to state  $s'$  when executing action  $a$  in state  $s$ , i.e.  $T(s, a, s') = p(s'|s, a)$ .
- $R : S \times A \rightarrow R$ : is a reward function defining the reward of executing action  $a$  in state  $s$ , i.e.  $R(s, a)$ .
- $\Omega$ : Set of observations  $\{\omega_1, \omega_2 \dots \omega_n\}$ .
- $O : S \times A \times \Omega \rightarrow [0, 1]$ : is an observation function that represents the probability of observing  $o$  when action  $a$  was executed and led to state  $s$ , i.e.  $O(o, a, s) = p(o|a, s)$ .
- $\gamma \in [0, 1]$  is the discount factor.

Solving a POMDP is equivalent to finding the optimal  $\pi^* : \Omega \rightarrow A$ . The optimal policy is defined as the one maximizing the expected infinite sum of discounted rewards  $r_t$  starting from belief state  $b_0$ .

$$\pi^* = \arg \max_{\pi} E \left[ \sum_{t=0}^{\infty} \gamma^t r_t^{\pi} | b_0 \right] \quad (6)$$

where  $r_t^{\pi}$  is the reward given for following the  $\pi$  policy at the time  $t$ . In this context, the focus is on planning the transition and reward functions.

#### 4. A Modeling Strategy for Explorative and Informative Path Planning Approach

The focus of this work is on the proposal of an efficient data acquisition strategy to perform continuous trajectories using an aerial robot in the Monitoring of forests. For this, it is necessary for the UAV to visit a certain area, explore the environment and make a decision on which route to take to maximize the search area in order to gather information. Therefore, in order for the UAV to have information about the environment, it will be explored in continuous 3D space to carry out informative path planning.

Therefore, the strategy is to explore the environment in order to find the canopy gaps and, after identifying them, carry out the planning with possible changes in altitudes. In this way, clearings will be explored and informative path planning will be carried out with the aim of identifying invasive plants. An example of an invasive plant is the *Hedychium Coronarium*.

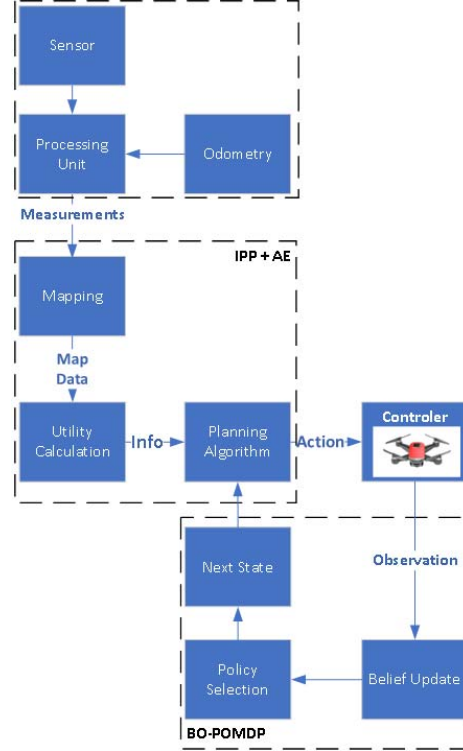


Figure 2: Proposed Architecture Diagram for Exploration and Informative Path Planning

In the architecture shown in Figure 2, was developed based on [15], the environment map is built using measurements extracted from a sensor data stream. After the measurements are performed in the AE phase, the IPP phase is started, where the mapping takes place and where the "Utility Calculation" module generates information that will be used in the "Planning Algorithm" module. The "Planning Algorithm" module performs an action in the environment defining the path that the UAV will take. By observing the environment, the BO-POMDP module updates the belief and uses a "Policy Selection" to define the next state. This next state will be used as input to the "Planning Algorithm" module, thus performing a new action in the environment. The process is repeated until the mission is completed.

In the IPP phase there is the Gaussian Process that will be used in the BO model. According to [15], this technique has been widely used in the context of IPP in different applications involving UAVs for data collection in unknown environments. This structure allows the use of different kernel functions to express data relationships within the [29] environment. The idea will be use path signatures in the

kernel definition for BO in the path space [30]. Furthermore, [24] highlights that this framework is capable of scaling large datasets through approximations.

#### 4.1. Formulation BO-POMDP

The main idea of using BO-POMDPs (Bayesian Optimization with Partially Observable Markov Decision Process) is to solve the problem of finding the best sequence of trajectories along the path while collecting information about the environment. The value of a trajectory depends on the amount of information collected and how much this information conditions uncertainty and a global phenomenon in the monitored.

This value is captured by a metric that balances exploration and exploitation. In the case of myopic methods, where the planner looks only one step ahead, the metric is calculated for a single trajectory. However, in the non-myopic case, the value of a trajectory must assume all expected future samples. Ideally, the value of a trajectory should be calculated with all future samples along an infinite horizon. In practice, however, the horizon (or looking ahead) needs to be finite for tractability reasons.

The proposed POMDP is based on the works [16] [13]. The difference is in the development of an integrated approach with the BO-POMDP mathematical formulation for a sequential decision maker under uncertainty to carry out Exploration and Informative Path Planning considering 3D space.

Using a POMDP instead of a simple Markov decision process allows to code the fact that the UAV performs actions only based on noisy observations of the objective function  $f$ . The system state  $\{f, p\}$  is completely described by  $f$  and the UAV pose  $p$ , in some situations the pose can be changed by points of interest.

In a real situation, the pose  $p$  of the UAV is not always given. However, the interest is not in learning how the system transitions, the  $T$  transition function works in such a way that the UAV can simulate sequences of actions. Likewise, the UAV also receives an approximate reward function  $R(\{b(f), p\}, a)$  based on its current function  $b(f)$  instead of  $f$ . The approximate reward function is based on Equation 6, making the trade-off between exploration and exploitation.

- $S$ : the state is a tuple  $\{f, p\}$ , where  $f$  is the objective function and  $p$  is the pose UAV.
- $A$ : The actions are defined by the parameter  $\Theta$ , each pair of action and pose parameter fully define a  $T(\Theta, p)$  trajectory.
- $T$ : The transition function models the probability  $T(\{f, p\}, \Theta, \{f', p'\})$  resulting in the state  $\{f', p'\}$  given a trajectory  $T(\Theta, p)$  which was taken in state  $\{f, p\}$ . The transition function is assumed to be deterministic, independent of the objective function. So it can be rewritten.  $T(p, \Theta, p') = \gamma(T(\Theta, p)|_{u=1} - p')$ , where  $T(\Theta, p)|_{u=1}$  is the resulting pose after executing the full path where  $\gamma$  is the dirac function.

- $R$ : The reward function calculates the sum of rewards obtained along the trajectory  $T(\Theta, p)$ , and is defined as Equation 7, where the  $cost(T(o, p))$  is the specific application of cost when moving  $T(\Theta, p)$ .
- $\Omega$ : Observations are noisy evaluations of  $f$  along trajectories.
- $O$ : At each location, the observation function  $O$  is uniquely defined by the pose of the robot.  $O$  depends on the distribution  $p(O|f(x))$  from noisy observations of the objective function  $f$ . In a discrete case,  $O$  is collected along the trajectories, resulting in a set of observations  $\{o_i\}$  (Eq. 8).

$$\bar{R}(\{b(f), p\}, \Theta) = r(\Theta, p, b(f)) + custo(T(\Theta, p)) \quad (7)$$

$$O(\{o_i\}, \Theta, \{f, p\}) = \prod_{x_i \in T(\Theta, p)} p(o_i|f(x_i)) \quad (8)$$

Our formulation will include in the set  $S$  robot state  $\{f, p\}$ , where  $f$  is the objective function and  $p$  is the pose of the robot, altitude changes to perform the continuous 3D space exploration generate trajectories that maximize the gain of information with the location of possible targets.

Finally, in recent years, different methods have been proposed to solve POMDPs, such as Value iteration [31], Branch-and-Bound [32] and Monte-Carlo Tree Search [13], [33]. These methods can be used to find the optimal  $\pi^*$  policy, granting the maximum sum of discounted rewards as defined in Equation 6. In practice, we need to approximate the infinite sum by a finite, thus restricting  $\pi^*$  to be an optimal policy in a horizon of  $n$  actions.

## 5. Conclusion

In this paper, we present a modeling strategy using Sequential Bayesian Optimization with POMDP for the Exploration and Informative Path Planning of UAVs in the monitoring of forest with canopy gaps in 3D space. The strategy is to explore the environment and make a decision on which route to take to maximize the search area in order to gather information. It is hoped, that the formulation BO-POMDP for Explorative and Informative Path Planning can be developed. As a future work, we intend to analyze the performance of the proposed architecture to carry out the Exploration and Informative Path Planning, experiments will be adopting different scenarios (simulated and real environments).

## Acknowledgment

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The authors acknowledge CAPES for the financial support.

## References

- [1] K. Kanistras, G. Martins, M. J. Rutherford, and K. P. Valavanis, "A survey of unmanned aerial vehicles (uavs) for traffic monitoring," in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2013, pp. 221–234.
- [2] K. C. Vivaldini, T. H. Martinelli, V. C. Guizilini, J. R. Souza, M. D. Oliveira, F. T. Ramos, and D. F. Wolf, "Uav route planning for active disease classification," *Autonomous Robots*, vol. 43, no. 5, pp. 1137–1153, 2018.
- [3] J. G. A. Barbedo, "A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses," *Drones*, vol. 3, no. 2, p. 40, 2019.
- [4] K. Srivastava, A. J. Bhutoria, J. K. Sharma, A. Sinha, and P. C. Pandey, "Uavs technology for the development of gui based application for precision agriculture and environmental research," *Remote Sensing Applications: Society and Environment*, vol. 16, p. 100258, 2019.
- [5] H. Oleynikova, Z. Taylor, A. Millane, R. Siegwart, and J. Nieto, "A complete system for vision-based micro-aerial vehicle mapping, planning, and flight in cluttered environments," *arXiv preprint arXiv:1812.03892*, 2018.
- [6] W. Kwon, J. H. Park, M. Lee, J. Her, S.-H. Kim, and J.-W. Seo, "Robust autonomous navigation of unmanned aerial vehicles (uavs) for warehouses' inventory application," *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 243–249, 2019.
- [7] F. Malandrino, C.-F. Chiasserini, C. Casetti, L. Chiaraviglio, and A. Senacheribbe, "Planning uav activities for efficient user coverage in disaster areas," *Ad Hoc Networks*, vol. 89, pp. 177–185, 2019.
- [8] T. Li, C. Wang, M. Max Q.-H., and C. W. d. Silva, "Coverage sampling planner for uav-enabled environmental exploration and field mapping," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 2509–2516.
- [9] S. Schopferer and S. Benders, "Minimum-risk path planning for long-range and low-altitude flights of autonomous unmanned aircraft," in *AIAA Scitech 2020 Forum*, 2020, p. 0137.
- [10] M. Enserink, "Biological invaders sweep in," 1999.
- [11] R. A. Lima, "Estrutura e regeneração de clareiras em florestas pluviais tropicais," *Brazilian Journal of Botany*, vol. 28, pp. 651–670, 2005.
- [12] J. Binney and G. S. Sukhatme, "Branch and bound for informative path planning," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 2147–2154.
- [13] P. Morere, R. Marchant, and F. Ramos, "Sequential bayesian optimization as a pomdp for environment monitoring with uavs," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 6381–6388.
- [14] M. Popovic, T. Vidal-Calleja, J. J. Chung, J. Nieto, and R. Siegwart, "Informative path planning and mapping for active sensing under localization uncertainty," *arXiv preprint arXiv:1902.09660*, 2019.
- [15] P. Marija, T. Vidal-Calleja, G. Hitz, C. J. Jen, S. Inkyu, R. Siegwart, and J. Nieto, "An informative path planning framework for uav-based terrain monitoring," *Autonomous Robots*, vol. 44, no. 6, pp. 889–911, 2020.
- [16] R. Marchant and F. Ramos, "Bayesian optimisation for informative continuous path planning," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 6136–6143.
- [17] G. Francis, L. Ott, and F. Ramos, "Functional path optimisation for exploration in continuous occupancy maps," in *Robotics Research*. Springer, 2020, pp. 859–875.
- [18] G. Francis, L. Ott, R. Marchant, and F. Ramos, "Occupancy map building through bayesian exploration," *The International Journal of Robotics Research*, vol. 38, no. 7, pp. 769–792, 2019.
- [19] F. Ramos and L. Ott, "Hilbert maps: scalable continuous occupancy mapping with stochastic gradient descent," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1717–1730, 2016.
- [20] R. Senanayake and F. Ramos, "Bayesian hilbert maps for dynamic continuous occupancy mapping," in *Conference on Robot Learning*, 2017, pp. 458–471.
- [21] H. Kurniawati, "Partially observable markov decision processes (pomdps) and robotics," *arXiv preprint arXiv:2107.07599*, 2021.
- [22] A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, pp. 707–755, 2009.
- [23] J. Binney, A. Krause, and G. S. Sukhatme, "Informative path planning for an autonomous underwater vehicle," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 4791–4796.
- [24] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.
- [25] S. Vasudevan, F. Ramos, E. Nettleton, and H. Durrant-Whyte, "Gaussian process modeling of large-scale terrain," *Journal of Field Robotics*, vol. 26, no. 10, pp. 812–840, 2009.
- [26] E. Brochu, V. Cora, and N. Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," *CoRR*, vol. abs/1012.2599, 12 2010.
- [27] D. R. Jones, C. D. Pertunnen, and B. E. Stuckman, "Lipschitzian optimization without the lipschitz constant," *Journal of optimization Theory and Applications*, vol. 79, no. 1, pp. 157–181, 1993.
- [28] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable markov processes over a finite horizon," *Operations research*, vol. 21, no. 5, pp. 1071–1088, 1973.
- [29] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, no. Feb, pp. 235–284, 2008.
- [30] P. Kidger, P. Bonnier, I. Perez Arribas, C. Salvi, and T. Lyons, "Deep signature transforms," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [31] H. Kurniawati and V. Yadav, "An online pomdp solver for uncertainty planning in dynamic environment," in *Robotics Research*. Springer, 2016, pp. 611–629.
- [32] O. Brock, J. Trinkle, and F. Ramos, *SARSOP: Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces*, 2009, pp. 65–72.
- [33] D. Silver and J. Veness, "Monte-carlo planning in large pomdps," 01 2010, pp. 2164–2172.