

KHẢO SÁT TÍCH HỢP THÔNG TIN NGỮ NGHĨA VÀO LẬP KẾ HOẠCH CHO RÔ-BỐT TỰ HÀNH

INTEGRATING THE SEMANTIC INFORMATION INTO PLANNING FOR AUTONOMOUS ROBOT: A SURVEY

Nguyễn Văn Hùng^{1,*}, Trương Xuân Tùng²,
Nghiêm Hoàng Nam², Nguyễn Văn Tuấn¹

DOI: <http://doi.org/10.57001/huinh5804.2025.057>

TÓM TẮT

Lập kế hoạch là khả năng thiết yếu, đóng vai trò then chốt trong các hệ thống rô-bốt. Sử dụng thông tin ngữ nghĩa để tăng cường hiệu quả và khả năng ở cả lập kế hoạch bậc thấp cũng như lập kế hoạch bậc cao là lĩnh vực mới do kế thừa kết quả của lĩnh vực phân đoạn ngữ nghĩa và phát hiện đối tượng. Nó có tiềm năng trang bị cho rô-bốt các hành vi thông minh hơn. Cho nên, mặc dù là lĩnh vực mới nhưng nó cũng nhận được sự quan tâm lớn của cộng đồng và bước đầu cũng đạt được một số kết quả có thể áp dụng vào thực tế. Tuy vậy chưa có một thống kê, phân loại và tổng hợp về các công bố đã xuất bản. Mục tiêu của bài báo là lấp đầy khoảng trống đó bằng khảo sát các công bố liên quan, từ đó rút ra nguyên lý chung bằng cách nào và làm thế nào thông tin ngữ nghĩa có thể tăng cường cho quá trình lập kế hoạch. Các nguyên lý chung này có thể làm cơ sở định hướng cho các nghiên cứu trong tương lai liên quan tới lĩnh vực này.

Từ khóa: *Lập kế hoạch tự động, lập kế hoạch nhiệm vụ, lập kế hoạch lộ trình, TAMP, lập kế hoạch ngữ nghĩa, lập kế hoạch dựa trên học máy.*

ABSTRACT

Planning is an essential capability and plays a pivotal role in robotic systems. Utilizing semantic information to enhance efficiency and capability in both low-level and high-level planning is a new field that builds upon advancements in semantic segmentation and object detection. It holds the potential to equip robots with more intelligent behaviors. Therefore, despite being a nascent field, it has garnered significant attention from the community and has already achieved some practical results. However, there is yet to be a comprehensive statistical analysis, classification, and synthesis of published works in this area. The goal of this paper is to fill that gap by surveying related publications and deriving general principles on how semantic information can enhance the planning process. These general principles can serve as a foundation to guide future research in this domain.

Keywords: *Automated planning, task planning, path planning, TAMP, semantic planning, learning-based planning.*

¹Viện Tự động hóa Kỹ thuật Quân sự, Viện Khoa học và Công nghệ Quân sự

²Học viện Kỹ thuật Quân sự

*Email: nvhung.v2k@gmail.com

Ngày nhận bài: 09/01/2025

Ngày nhận bài sửa sau phản biện: 28/02/2025

Ngày chấp nhận đăng: 28/3/2025

1. GIỚI THIỆU

Lập kế hoạch (planning) trong lĩnh vực rô-bốt là quá trình xác định chuỗi hành động, lộ trình, hoặc chuyển động để rô-bốt đạt được một mục tiêu nhất định, đồng thời thỏa mãn các ràng buộc liên quan đến môi trường, động học, động lực học và hiệu suất. Kết quả của quá trình lập kế hoạch là quỹ đạo $\tau(t)$ hoặc thao tác $q(s)$ làm đầu vào cho bộ điều khiển (minh họa bởi hình 1). Quá trình lập kế hoạch thông thường là sự kết hợp của Lập kế hoạch nhiệm vụ (task planning) và Lập kế hoạch chuyển động (motion planning). Lập kế hoạch nhiệm vụ thao tác với phần biểu tượng mức cao (high-level symbol), còn lập kế hoạch chuyển động xử lý chuyển động và ràng buộc ở mức hình học.

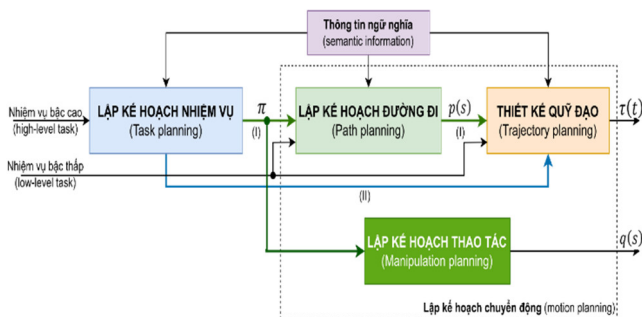
Lập kế hoạch nhiệm vụ (hay còn gọi là automated planning) [1] là quá trình tạo ra một chuỗi các hành vi π mà rô-bốt cần thực hiện để đạt được một mục tiêu cụ thể (từ trạng thái ban đầu s_0 tới trạng thái đích s_g), trước khi chuyển sang các giai đoạn chi tiết hơn như lập kế hoạch đường đi hoặc thiết kế quỹ đạo.

Còn lập kế hoạch chuyển động [2, 3] là quá trình tìm ra quỹ đạo hoặc thao tác cho một tác nhân (cánh tay

rô-bốt, thiết bị bay không người lái hoặc xe tự hành) hành động (hoặc tác động) từ vị trí ban đầu đến vị trí mục tiêu, đồng thời tránh các va chạm với chướng ngại vật trong môi trường. Nó gồm Lập kế hoạch đường đi (path planning), Thiết kế quỹ đạo (trajectory planning) và Lập kế hoạch thao tác (manipulation planning).

Lập kế hoạch đường đi (path planning) [4-9] là quá trình xác định một đường đi khả thi $p(s)$ để rô-bốt di chuyển từ điểm bắt đầu đến điểm đích trong môi trường hoạt động, đồng thời tránh va chạm với các chướng ngại vật và tuân thủ các ràng buộc (hạn chế) khác.

Thiết kế quỹ đạo (trajectory planning) [10-15] là quá trình tạo một chuỗi chuyển động $\tau(t)$ từ trạng thái bắt đầu tới trạng thái mục tiêu, đảm bảo tính khả thi động học và động lực học, đồng thời tối ưu chỉ tiêu chất lượng nào đó. Khác với lập kế hoạch đường đi, tập trung vào việc xác định đường đi hình học không va chạm, lập kế hoạch quỹ đạo còn xem xét yếu tố thời gian và các ràng buộc vật lý trong chuyển động.



Hình 1. Quy trình lập kế hoạch điều khiển rô-bốt gồm Lập kế hoạch nhiệm vụ và Lập kế hoạch chuyển động

Lập kế hoạch thao tác (manipulation planning) tập trung vào việc lập một chuỗi các thao tác $q(s)$ cho rô-bốt để tương tác (cầm, nắm, gắp, mang) với các đối tượng nhằm thay đổi trạng thái của các vật thể trong môi trường.

Ngày nay, nhiệm vụ (yêu cầu, sứ mệnh) mà rô-bốt cần thực hiện ngày càng phức tạp, mang tính trừu tượng và thường được mô tả bằng ngôn ngữ tự nhiên thông qua ngữ nghĩa, ví dụ “nấu ăn”, “mang đồ ăn tới giường bệnh”,... Để hoàn thành được một nhiệm vụ như vậy thường phải kết hợp nhiều hành vi bậc thấp như di chuyển, gắp, thả. Hay nói cách khác là nhiệm vụ bậc cao được thực thi bằng một chuỗi các hành vi bậc thấp.

Cách tiếp cận trực quan với quá trình lập kế hoạch là tách riêng biệt lập kế hoạch nhiệm vụ và lập kế hoạch chuyển động. Khi lập kế hoạch nhiệm vụ coi như các hành vi bậc thấp là hộp đen xác định. Kết quả của lập kế hoạch nhiệm vụ là đầu vào cho lập kế hoạch chuyển động. Mặc

dù vậy trên thực tế, lập kế hoạch nhiệm vụ và lập kế hoạch chuyển động không độc lập với nhau, nên việc tách rời chúng sẽ làm quá trình lập kế hoạch thiếu hiệu quả [16]. Nên chúng thường được thực hiện đồng thời, vấn đề này gọi là TAMP (viết tắt của cụm từ Task And Motion Planning) [16-19].

Với sự ra đời, bùng nổ và đạt được những bước tiến vượt bậc, trí tuệ nhân tạo (AI) đã len lỏi vào hầu hết các lĩnh vực, trong đó có lĩnh vực rô-bốt. Nhờ áp dụng trí tuệ nhân tạo, các phương pháp phân đoạn ngữ nghĩa [20-24] và phát hiện đối tượng [25, 26] đã đạt được hiệu suất và độ chính xác cao. Từ đó cho phép sử dụng thông tin ngữ nghĩa để hỗ trợ, nâng cao hiệu quả và khả năng của quá trình lập kế hoạch [27]. Các nghiên cứu đã công bố cho thấy rằng thông tin ngữ nghĩa được sử dụng vào trong cả lập kế hoạch nhiệm vụ và lập kế hoạch chuyển động (minh họa bởi hình 1).

Ngoài ra, một vấn đề là hầu hết nghiên cứu hiện tại đều dựa vào độ chính xác của mô hình mô tả mục tiêu hoặc kiến thức có trước về môi trường để lập kế hoạch. Tuy nhiên, trong thực tế rất khó để có được điều này, vì kiến thức và dữ liệu liên quan đến môi trường nhiều khi bị hạn chế hoặc không có sẵn. Sử dụng thông tin ngữ nghĩa cho phép vượt qua rào cản này bằng áp dụng học máy (learning-based) vào quá trình lập kế hoạch [28]. Hơn nữa, với sự ra đời và áp dụng các phương pháp học tăng cường RL (viết tắt cụm từ Reinforcement Learning) là một giải pháp thay thế cho việc tách biệt các giai đoạn lập bản đồ và lập kế hoạch rời rạc.

Cấu trúc nội dung bài báo gồm hai phần chính:

- Thông tin ngữ nghĩa trong lập kế hoạch chuyển động: khảo sát và rút ra cách thức chung sử dụng thông tin ngữ nghĩa trực tiếp vào hỗ trợ và nâng cao quá trình lập kế hoạch chuyển động. Mỗi loại cách thức đi kèm một vài công bố điển hình.

- Lập kế hoạch ngữ nghĩa: trình bày nguyên lý chung sử dụng thông tin ngữ nghĩa để tăng cường hiệu quả và khả năng lập kế hoạch giúp cho rô-bốt có hành vi, tương tác thông minh hơn. Tiếp theo là khảo sát ứng dụng thực tế của nguyên lý đó trong một số bài toán lớn như tìm kiếm, kiểm tra đối tượng và khám phá môi trường.

2. THÔNG TIN NGỮ NGHĨA TRONG LẬP KẾ HOẠCH CHUYỂN ĐỘNG

Các thuật toán lập kế hoạch chuyển động [2] chỉ cần tới thông tin hình học thuần túy, tức là thông tin hình dạng, kích thước và vị trí tương đối của các đối tượng. Loại bản đồ mô tả môi trường mà rô-bốt hoạt động kiểu này

được gọi là bản đồ hình học (metric map). Nền ý tưởng đầu tiên để sử dụng thông tin ngữ nghĩa vào hỗ trợ lập kế hoạch bậc thấp là dùng nó vào tăng cường bản đồ hình học như hỗ trợ tạo bản đồ, thay đổi cấu trúc,... Thông thường thông tin ngữ nghĩa được mã hóa vào bản đồ bằng tạo thêm một hoặc nhiều lớp.

Cách tiếp theo là thay đổi trực tiếp vào cấu trúc cơ sở của bộ tạo quỹ đạo bằng thêm các phần xử lý thông tin ngữ nghĩa; Hoặc tác động vào tăng cường các thuật toán đã có như sử dụng thông tin ngữ nghĩa để dẫn quá trình lấy mẫu ưu tiên vùng có nhiều thông tin cần cho quá trình lên kế hoạch, sử dụng thông tin ngữ nghĩa để điều chỉnh heuristic, hàm tối ưu để điều hướng tới vùng có xu hướng tới đích nhanh hơn.

Deeken và cộng sự [29] sử dụng hệ thống thông tin địa lý (GIS) với các lớp hình học và ngữ nghĩa được sử dụng để xây dựng các bản đồ chi phí (costmap) được khai thác bởi gói ROS move_base (http://wiki.ros.org/move_base) để thực hiện nhiệm vụ lấy và giao hàng. Đây là nỗ lực thành công đầu tiên chứng minh rằng thông tin ngữ nghĩa có thể được sử dụng trong các mô-đun điều hướng.

Sadat và cộng sự [30] sử dụng mạng nơ-ron để tạo thêm các lớp trung gian kiểu xác suất chiếm chỗ (occupancy probability) có thông tin ngữ nghĩa để mô tả trạng thái hiện tại và dự đoán trạng thái tương lai của các đối tượng trong môi trường. Từ đó quỹ đạo được tạo bằng tối ưu hàm phạt trên cơ sở thông tin từ bản đồ này.

Achat và cộng sự [31] đưa thông tin ngữ nghĩa vào bản đồ chi phí nhiều lớp (multi-layer costmap), trong đó mỗi lớp được gán chi phí khác nhau dựa trên tính khả thi khi di chuyển hoặc khả năng quan sát. Ngoài ra, tác giả mở rộng các thuật toán lập kế hoạch truyền thống như A* bằng cách đưa thêm trọng số thông tin ngữ nghĩa vào hàm heuristic. Cuối cùng, đường đi đã lên kế hoạch được xử lý tiếp bằng một thuật toán tối ưu hóa nhằm rút ngắn và tăng cường.

Thông tin ngữ nghĩa được sử dụng hỗ trợ tạo bản đồ cấu trúc nhiều lớp, tức là tổ chức thông tin dạng phân cấp: cấp độ cao (Semantic Layer) mô tả ngữ nghĩa như loại phòng, đối tượng (bàn, ghế, tường); cấp độ trung gian (Spatial Layer) mô tả không gian, như các khu vực hay phòng cụ thể; cấp độ thấp (Geometric Layer) chứa thông tin hình học, ví dụ S-Graphs+ [32]. Dựa trên loại bản đồ này, Kremer và cộng sự đề xuất thuật toán lập kế hoạch phân cấp (Hierarchical planning) S-nav [33] để thực hiện lập kế hoạch nhiều mức, các lớp ngữ nghĩa mức cao cho phép lập kế hoạch nhanh, kết quả của lớp trên sẽ được sử dụng để dẫn quá trình lập kế hoạch ở mức thấp

hơn. Quá trình lập kế hoạch kết thúc sau khi lập kế hoạch xong ở lớp hình học.

Ryll và cộng sự [34] đề xuất bộ lập kế hoạch ngữ nghĩa (semantic planner) dẫn đường cho thiết bị bay không người lái trang bị RGB camera trong môi trường đô thị chưa biết (unknown urban environment) nhằm mục đích mở rộng đường chân trời bởi hạn chế về thông tin cảm biến. Bản đồ xác suất (probabilistic map) được tạo từ ảnh đã được phân đoạn ngữ nghĩa bằng cách chiếu ảnh đó lên bản đồ với ưu tiên đường đi vì cho rằng quá trình bay ở không gian bên trên đường sẽ an toàn hơn. Quá trình ưu tiên thực hiện bằng gán trọng số dương cho đường và trọng số âm cho các đối tượng khác như thảm cỏ, vỉa hè,... từ bản đồ tiến hành tạo graph, quỹ đạo toàn cục có được sử dụng graph search.

Một cách tiếp cận tiếp theo là đưa thông tin ngữ nghĩa vào thay đổi hàm tối ưu. Bartolomei và cộng sự [35] sử dụng kiến trúc tạo quỹ đạo hai giai đoạn để tạo quỹ đạo cho thiết bị bay không người lái (UAV). Giai đoạn thứ nhất đưa thông tin ngữ nghĩa vào như là một hạng tử của hàm tối ưu của Kinodynamic A* [36] sao cho xu hướng tạo quỹ đạo tới vùng có nhiều thông tin hơn. Giai đoạn thứ hai là tối ưu B-Spline [37] dựa trên kết quả của lớp thứ nhất giữ cho các điểm mốc (landmark) ở trong trường nhìn. Bartolomei và cộng sự [38] trên cơ sở phát triển nghiên cứu [35] để có thể thực thi trong môi trường thực tế bằng cách nhận diện động vùng có nhiều thông tin hơn sử dụng học tăng cường (RL - reinforcement learning).

3. LẬP KẾ HOẠCH NGỮ NGHĨA

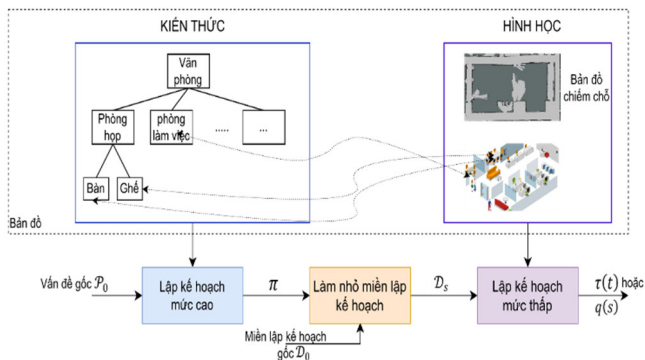
3.1. Nguyên lý chung

Thông tin ngữ nghĩa được sử dụng để tăng cường hiệu quả và khả năng của quá trình lập kế hoạch. Ý tưởng chung khi đưa thông tin ngữ nghĩa vào hỗ trợ và tăng cường hiệu quả quá trình lập kế hoạch là sử dụng thông tin ngữ nghĩa để làm nhỏ đi miền lập kế hoạch gốc $\mathcal{D}_0$ thành miền $\mathcal{D}_s \subseteq \mathcal{D}_0$. Như vậy, thay vì lập kế hoạch mức thấp trong miền $\mathcal{D}_0$ thì sẽ được thực hiện ở miền $\mathcal{D}_s$ có kích thước nhỏ hơn.

Có thể phân loại phương pháp lập kế hoạch ngữ nghĩa thành hai loại ứng với hai cấu trúc lưu trữ kiến thức trong bản đồ ngữ nghĩa. Phương pháp thứ nhất là kiến thức ngữ nghĩa được lưu trữ tách riêng với thông tin hình học như T-Box [39], ontology [40-42],... Mỗi quan hệ giữa hai phần sẽ được duy trì (minh họa bằng hình 2, mối quan hệ giữa kiến thức và hình học được biểu thị bằng đường nét đứt).

Lập kế hoạch mức cao tạo ra một chuỗi hành vi π từ 'kiến thức' và vấn đề lập kế hoạch gốc $\mathcal{P}_0$. Quá trình làm

nhỏ miền lập kế hoạch là loại bỏ các thông tin không liên quan đến quá trình lập kế hoạch dựa vào chuỗi hành vi π . Kết quả là được miền lập kế hoạch mới $\mathcal{D}_s$.



Hình 2. Phương pháp tăng cường hiệu quả lập kế hoạch bằng giảm nhỏ miền lập kế hoạch sử dụng thông tin ngữ nghĩa dựa trên bản đồ lưu trữ “kiến thức” riêng rẽ với thông tin hình học

Phương pháp thứ hai là dựa trên cấu trúc phân tầng của bản đồ [32, 43-49], tức là bản đồ được lưu dưới dạng nhiều lớp, giữa các lớp có mối liên hệ với nhau (mô tả bằng hình 3). Trong đó ‘kiến thức’ được lưu trữ ở các lớp bên trên và có tính trừu tượng, còn thông tin ‘hình học’ thì được lưu ở các lớp bên dưới. Quá trình lập kế hoạch được thực hiện theo từng lớp, kết quả của lớp trên sẽ là đầu vào cho lớp dưới, miền lập kế hoạch sẽ được giảm dần qua từng lần lập kế hoạch $\mathcal{D}_0 \rightarrow \mathcal{D}_1 \rightarrow \dots$ cho tới khi đạt được quỹ đạo $\tau(t)$ hoặc hành vi $q(s)$.

Kỹ thuật làm nhỏ miền lập kế hoạch có thể thực hiện bằng lập kế hoạch từng lớp, các phương pháp TAMP [17] hoặc sử dụng học máy như mô hình ngôn ngữ lớn (LLM-enabled planner), học sâu, học sâu tăng cường,... [50-55].

Ngoài việc tăng cường hiệu quả khi lập kế hoạch, thông tin ngữ nghĩa còn được sử dụng để tăng cường khả năng lập kế hoạch:

• **Hỗ trợ hoàn thành nhiệm vụ chưa tường minh:**

Thông thường vấn đề lập kế hoạch $\mathcal{P}_0$ được truyền đạt rõ ràng, tức là yêu cầu đạt được và các nhiệm vụ con cần thực hiện để đạt được mục đích đã tường minh, ví dụ yêu cầu “cất cho tôi quyển sách này vào bàn làm việc ở phòng ngủ”. Tuy nhiên trong một số trường hợp yêu cầu lại không được tường minh như vậy, như ví dụ trên nhưng với yêu cầu là “cất cho tôi quyển sách vào bàn làm việc”, kiến thức từ thông tin ngữ nghĩa cho phép rô-bốt biết rằng bàn làm việc ở trong phòng ngủ. Thông tin ngữ nghĩa giúp rô-bốt có khả năng hoàn thành nhiệm vụ bằng ngôn ngữ tự nhiên chưa tường minh [56-61].

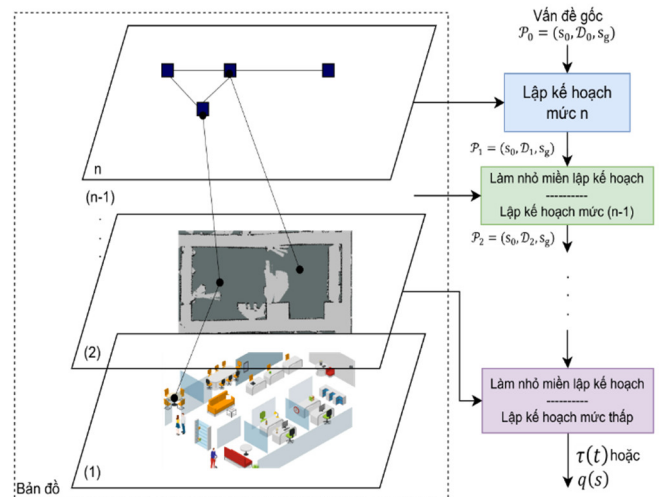
• **Suy ra sự tồn tại của đối tượng:**

Trong thực tế rất hay xảy ra trường hợp đối tượng tồn tại ở môi trường thực tế

nhưng trên bản đồ lại không có do chỉ quan sát một phần hoặc đối tượng chưa được nhận dạng để thêm vào bản đồ. Mà rô-bốt lại nhận được lệnh tương tác với đối tượng đó. Ví dụ: rô-bốt cần thực hiện yêu cầu “lấy cho tôi quả táo”, nếu quả táo không tồn tại trên bản đồ thì lệnh không được thực thi. Nhưng với kiến thức từ bản đồ cung cấp cho rô-bốt khả năng suy đoán nhiều khả năng quả táo nằm trong tủ lạnh, cho nên yêu cầu gốc chuyển sang lập kế hoạch điều hướng tới tủ lạnh, sau đó thực hiện các hành vi tiếp theo để hoàn thành yêu cầu gốc [62-64].

• **Hỗ trợ trong môi trường động:**

Trong môi trường có các đối tượng di chuyển (người, xe cộ), thông tin ngữ nghĩa giúp dự đoán hành vi của các đối tượng này, ví dụ: “người thường đi theo lối đi bộ” hoặc “xe ô tô di chuyển trên đường” [44, 65, 66].



Hình 3. Phương pháp tăng cường hiệu quả lập kế hoạch bằng lập kế hoạch phân tầng. Miền lập kế hoạch sẽ thu nhỏ dần qua từng lớp lập kế hoạch

Các kỹ thuật tăng cường hiệu quả và khả năng của quá trình lập kế hoạch sử dụng thông tin ngữ nghĩa thường không tồn tại riêng rẽ mà được kết hợp nhằm giải quyết các bài toán cụ thể như tìm đối tượng (object search), kiểm tra (inspection), khám phá (exploration), lập kế hoạch cho môi trường động,...

Ngoài ra, thông tin ngữ nghĩa còn cho phép lập kế hoạch bằng mô hình ngôn ngữ lớn (LLM, viết tắt của Large Language Model) [54, 67] và mô hình ngôn ngữ thị giác (VLM, viết tắt của Visual Language Model, là sự kết hợp của xử lý ngôn ngữ tự nhiên NLP và thị giác máy tính CV) [68, 69].

3.2. Lập kế hoạch một số bài toán điển hình sử dụng thông tin ngữ nghĩa

Bài toán tìm đối tượng (object search) là xác định vị trí của một hoặc nhiều đối tượng trong môi trường sao cho

tối ưu hóa thời gian, năng lượng hoặc tài nguyên sử dụng trong quá trình tìm kiếm. Anderson và cộng sự [70] phân loại bài toán tìm đối tượng thành ba loại: PointGoal (một điểm có tọa độ trong không gian, ví dụ (1,2,4)), ObjectGoal (là một đối tượng trong không gian, ví dụ “tủ lạnh”, “giường”,...) và AreaGoal (là một vùng trong không gian, như “bếp”, “nhà kho”,...). ObjectGoal và AreaGoal được cung cấp dưới dạng ngôn ngữ tự nhiên hoặc đa phương tiện (ảnh, video).

Bài toán kiểm tra đối tượng (inspection) là đánh giá trạng thái hoặc đặc điểm của một đối tượng đã biết vị trí. Như vậy để kiểm tra đối tượng thì trước tiên phải giải quyết bài toán tìm đối tượng. Khám phá (exploration) là khả năng hoạt động trong môi trường không biết trước (không có bản đồ trước) hoặc biết một phần, đồng thời tạo bản đồ có thể sử dụng lại. Một chiến lược khám phá tốt là có thể tạo ra bản đồ tốt trong lượng thời gian hợp lý [71-75].

Giải quyết bài toán thực tế như tìm hồng hộc, sửa chữa, chữa cháy, tìm kiếm cứu hộ cứu nạn,... là quá trình kết hợp các phương pháp giải quyết các bài toán cơ bản trên.

Cách thủ công để giải quyết bài toán tìm đối tượng là định nghĩa trước các tuyến đường (path), điểm quan sát (viewpoint) hoặc đặt các thẻ có thể nhận dạng [76, 77], chẳng hạn như Apriltags [78] hoặc mã QR. Cách làm này đơn giản nhưng chỉ giải quyết được bài toán đơn lẻ, cụ thể, môi trường tĩnh và hoàn toàn xác định.

Với yêu cầu khả năng tự hành thì rô-bốt phải tự động chọn mục tiêu, điều hướng và nhận dạng đối tượng. Devendra Singh và cộng sự [79] sử dụng chính sách dựa trên học máy để chọn đích dựa trên các thông tin ngữ nghĩa có trước và loại đối tượng đích. Muhammad Fadhil và cộng sự đề xuất SEEK [80] để kết hợp kiến thức ngữ nghĩa trước đó với các quan sát của rô-bốt để tìm kiếm và điều hướng đến các đối tượng mục tiêu hiệu quả hơn. Cụ thể, SEEK sử dụng kiến thức ngữ nghĩa quan hệ để hướng dẫn rô-bốt tìm kiếm các khu vực có triển vọng cao hơn trong môi trường.

Khoanh vùng tìm kiếm để nhanh chóng tìm ra đối tượng dựa trên thông tin ngữ nghĩa là hướng nghiên cứu được phát triển và ứng dụng [81-83]. Tung Dang và cộng sự [84] đề xuất framework cho thiết bị bay không người lái tìm kiếm cứu hộ tai nạn hầm mỏ sử dụng thông tin ngữ nghĩa liên quan tới nạn nhân như điện thoại, công cụ khai thác mỏ, ba lô,... Thông tin ngữ nghĩa được sử dụng bởi mô hình ngôn ngữ lớn (LLM) như một phương pháp hướng dẫn để tìm thiên vị theo hướng mà đối tượng mục tiêu có thể được đặt [52, 85-87].

Các phương pháp học máy rất cần lượng dữ liệu huấn luyện lớn, cho nên việc tự động hóa quá trình thu thập dữ liệu là rất quan trọng trong việc giải phóng sức lao động của con người. Julius và cộng sự [88] dùng thông tin ngữ nghĩa để tự động hóa quá trình thiết bị bay không người lái thu thập ảnh địa hình tập trung vào phần có nhiều thông tin nhất đã được dán nhãn.

Với bài toán khám phá, thông tin ngữ nghĩa được sử dụng để đẩy nhanh quá trình khám phá các môi trường chưa biết (unknown environment) hoặc được biết một phần (partially-unknown environment). Việc kết hợp kiến thức ngữ nghĩa với các phương pháp tìm kiếm cổ điển như frontier đã được chứng minh là một chiến lược khám phá đặc biệt hiệu quả. Yu và cộng sự [89] dùng học máy kết hợp với phương pháp frontier để tạo “frontier semantic policy” dựa trên thông tin ngữ nghĩa từ bản đồ ngữ nghĩa.

Muhammad Fadhil và cộng sự đề xuất framework tên là SB2G (viết tắt cụm từ Semantic Belief Behavior Graph) [90] tích hợp thông tin ngữ nghĩa vào hệ thống lập kế hoạch và điều khiển của rô-bốt, cho phép xác định và tương tác với nhiều đối tượng khác nhau mà không cần kiến thức về môi trường trước đó.

Thông tin ngữ nghĩa cho phép dự đoán chuyển động, tư thế của người và đối tượng trong môi trường, do đó có thể đoán biết trạng thái tương lai gần của môi trường [44, 65, 66, 91-96]. Điều này rất hữu ích với lập kế hoạch trong môi trường động.

Ngoài ra, điều khiển dựa vào hành vi (behavior-based control) [97] là một phương pháp điều khiển rô-bốt dựa trên các hành vi, được thiết kế để phản hồi trực tiếp với môi trường, thay vì dựa vào mô hình hóa toàn diện hoặc lập kế hoạch chi tiết. Wolf và cộng sự [98] dẫn đường rô-bốt trong môi trường off-road sử dụng điều khiển hành vi dựa trên thông tin ngữ nghĩa. Trong đó kiến thức được lưu trữ bằng ontology.

4. KẾT LUẬN

Quá trình lập kế hoạch sử dụng thông tin ngữ nghĩa bước đầu kết quả còn nhiều hạn chế về độ chính xác, độ phủ, thời gian thực,... Nhưng nó mở ra vô vàn cách tiếp cận vì thông tin ngữ nghĩa vô cùng đa dạng về số lượng, chủng loại và đặc biệt là nó cho phép áp dụng công nghệ trí tuệ nhân tạo. Có thể nói rằng việc sử dụng thông tin ngữ nghĩa vào quá trình lập kế hoạch là một bước ngoặt trên con đường hướng tới trang bị cho rô-bốt khả năng và hành vi tiệm cận trí thông minh sinh học. Các nguyên lý chung trình bày trong bài khảo sát này giúp định hướng các nghiên cứu tiếp theo.

Ngoài việc tham gia trực tiếp hỗ trợ, nâng cao hiệu quả và khả năng của quá trình lập kế hoạch, thông tin ngữ nghĩa còn hỗ trợ gián tiếp thông qua cải thiện chất lượng bản đồ và nâng cao độ chính xác định vị [99-103]. Đây cũng là hướng nghiên cứu vô cùng sôi động và đạt được những kết quả vượt trội.

TÀI LIỆU THAM KHẢO

- [1]. M. Ghallab, D. Nau, P. Traverso, *Automated planning and acting*. Cambridge University Press, 2016.
- [2]. Steven M. LaValle, *Planning algorithms*. Cambridge University Press, 2006.
- [3]. W. Chen, W. Chi, S. Ji, et al., "A survey of autonomous robots and multi-robot navigation: Perception, planning and collaboration," *Biomimetic Intelligence and Robotics*, 5(2), 100203, 2025.
- [4]. E.W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, 1(1), 269–271, 1959.
- [5]. P.E. Hart, N.J. Nilsson, B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics*, 4(2), 100–107, 1968.
- [6]. S. LaValle, *Rapidly-exploring random trees: A new tool for path planning*. Research Report 9811, 1998.
- [7]. J.J. Kuffner, S.M. LaValle, "RRT-connect: An efficient approach to single-query path planning," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation, Symposia Proceedings* (Cat. No. 00CH37065), IEEE, 995–1001, 2000.
- [8]. S. Karaman, E. Frazzoli, *Incremental sampling-based algorithms for optimal motion planning*. MIT Press, 2011.
- [9]. J.D. Gammell, S.S. Srinivasa, T.D. Barfoot, "Informed RRT*: Optimal sampling-based path planning focused via direct sampling of an admissible ellipsoidal heuristic," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2997–3004, 2014.
- [10]. C. Richter, A. Bry, N. Roy, "Polynomial Trajectory Planning for Aggressive Quadrotor Flight in Dense Indoor Environments," in *Springer Tracts in Advanced Robotics*, 649–666, 2016.
- [11]. H. Oleynikova, M. Burri, Z. Taylor, et al., "Continuous-Time Trajectory Optimization for Online UAV Replanning," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016.
- [12]. B. Lopez, *Low-latency trajectory planning for high-speed navigation in unknown environments*. PhD Thesis, Massachusetts Institute of Technology, 2016.
- [13]. M. Hehn, R. D'Andrea, "Real-Time Trajectory Generation for Quadcopters," *IEEE Transactions on Robotics*, 31(4), 877–892, 2015.
- [14]. M. van Nieuwstadt, R.M. Murray, "Real Time Trajectory Generation for Differentially Flat Systems," *IFAC Proceedings Volumes*, 29(1), 2301–2306, 1996.
- [15]. J. Tordesillas, B.T. Lopez, J. Carter, et al., "Real-Time Planning with Multi-Fidelity Models for Agile Flights in Unknown Environments," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019.
- [16]. J. Ferrer-Mestres, G. Francès, H. Geffner, *Combined Task and Motion Planning as Classical AI Planning*. 2017. <https://arxiv.org/abs/1706.06927>.
- [17]. C.R. Garrett, R. Chitnis, R. Holladay, et al., "Integrated task and motion planning," *Annual review of control, robotics, and autonomous systems*, 4(1), 265–293, 2021.
- [18]. M. Mansouri, F. Pecora, P. Schüller, "Combining Task and Motion Planning: Challenges and Guidelines," *Frontiers in Robotics and AI*, 8, 2021.
- [19]. L.P. Kaelbling, T. Lozano-Pérez, "Hierarchical task and motion planning in the now," in *2011 IEEE International Conference on Robotics and Automation*, 1470–1477, 2011.
- [20]. A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, et al., *A review on deep learning techniques applied to semantic segmentation*. 2017. arXiv preprint arXiv:170406857.
- [21]. Q. Hu, B. Yang, L. Xie, et al., "RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [22]. A. Milioto, I. Vizzo, J. Behley, et al., "RangeNet ++: Fast and Accurate LiDAR Semantic Segmentation," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4213–4220, 2019.
- [23]. E. Romera, J.M. Álvarez, L.M. Bergasa, et al., "ERFNet: Efficient Residual Factorized ConvNet for Real-Time Semantic Segmentation," *IEEE Transactions on Intelligent Transportation Systems*, 19(1), 263–272, 2018.
- [24]. L.C. Chen, G. Papandreou, I. Kokkinos, et al., "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40, 834–848, 2018.
- [25]. J. Redmon, S. Divvala, R. Girshick, et al., *You Only Look Once: Unified, Real-Time Object Detection*. 2016. <https://arxiv.org/abs/1506.02640>.
- [26]. L. Schmid, O. Andersson, A. Sulser, et al., "Dynablox: Real-Time Detection of Diverse Dynamic Objects in Complex Environments," *IEEE Robotics and Automation Letters*, 8(10), 6259–6266, 2023.
- [27]. J. Crespo, J.C. Castillo, O.M. Mozos, et al., "Semantic Information for Robot Navigation: A Survey," *Applied Sciences*, 10(2), 2020.
- [28]. H.X. Pham, H.M. La, D. Feil-Seifer, et al., *Autonomous UAV Navigation Using Reinforcement Learning*. 2018. <https://arxiv.org/abs/1801.05086>.
- [29]. H. Deeken, S. Puetz, T. Wiemann, et al., "Integrating Semantic Information in Navigational Planning," in *ISR/Robotik 2014, 41st International Symposium on Robotics*, 1–8, 2014.
- [30]. A. Sadat, S. Casas, M. Ren, et al., *Perceive, Predict, and Plan: Safe Motion Planning Through Interpretable Semantic Representations*. 2020. <https://arxiv.org/abs/2008.05930>.
- [31]. S. Achat, J. Marzat, and J. Moras, "Path Planning Incorporating Semantic Information for Autonomous Robot Navigation," in *19th*

International Conference on Informatics in Control, Automation and Robotics (ICINCO) 2022, Lisbonne, Portugal, SCITEPRESS - Science and Technology Publications, 285–295, 2022.

[32]. H. Bavlé, J.L. Sanchez-Lopez, M. Shaheer, et al., "S-graphs+: Real-time localization and mapping leveraging hierarchical representations," *IEEE Robotics and Automation Letters*, 8(8), 4927–4934, 2023.

[33]. P. Kremer, H. Bavlé, J.L. Sanchez-Lopez, et al., *S-Nav: Semantic-Geometric Planning for Mobile Robots*. 2023. <https://arxiv.org/abs/2307.01613>.

[34]. M. Ryll, J. Ware, J. Carter, et al., "Semantic Trajectory Planning for Long-Distant Unmanned Aerial Vehicle Navigation in Urban Environments," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1551–1558, 2020.

[35]. L. Bartolomei, L. Teixeira, M. Chli, "Perception-aware Path Planning for UAVs using Semantic Segmentation," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 5808–5815, 2020.

[36]. B. Zhou, F. Gao, L. Wang, et al., "Robust and efficient quadrotor trajectory generation for fast autonomous flight," *IEEE Robotics and Automation Letters*, 4(4), 3529–3536, 2019.

[37]. G. Rousseau, C.S. Maniu, S. Tebbani, et al., "Minimum-time B-spline trajectories with corridor constraints. Application to cinematographic quadrotor flight plans," *Control Engineering Practice*, 89, 190–203, 2019.

[38]. L. Bartolomei, L. Teixeira, M. Chli, "Semantic-aware Active Perception for UAVs using Deep Reinforcement Learning," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3101–3108, 2021.

[39]. C. Galindo, J.-A. Fernández-Madriral, J. González, et al., "Robot task planning using semantic maps," *Robotics and Autonomous Systems*, 56(11), 955–966, 2008.

[40]. T.R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?," *International Journal of Human-Computer Studies*, 43(5–6), 907–928, 1995.

[41]. M. Tenorth, L. Kunze, D. Jain, et al., "KNOWROB-MAP - knowledge-linked semantic object maps," in *2010 10th IEEE-RAS International Conference on Humanoid Robots*, 430–435, 2010.

[42]. M.A. Mohammad, I. Kaloskamps, Y. Hicks, et al., "Ontology-based Framework for Risk Assessment in Road Scenes Using Videos," *Procedia Computer Science*, 60, 1532–1541, 2015.

[43]. I. Armeni, Z.Y. He, J. Gwak, et al., *3D Scene Graph: A Structure for Unified Semantics, 3D Space, and Camera*. 2019. <https://github.com/StanfordVL/3DSceneGraph>.

[44]. A. Rosinol, A. Gupta, M. Abate, et al., *3D Dynamic Scene Graphs: Actionable Spatial Perception with Places, Objects, and Humans*. 2020. <https://arxiv.org/abs/2002.06289>.

[45]. S.C. Wu, J. Wald, K. Tateno, et al., *SceneGraphFusion: Incremental 3D Scene Graph Prediction from RGB-D Sequences*. 2021. <https://github.com/ShunChengWu/SceneGraphFusion>.

[46]. A. Rosinol, A. Violette, M. Abate, et al., *Kimera: from SLAM to Spatial Perception with 3D Dynamic Scene Graphs*. 2021. <https://arxiv.org/abs/2101.06894>.

[47]. N. Hughes, Y. Chang, L. Carlone, "Hydra: A Real-time Spatial Perception System for 3D Scene Graph Construction and Optimization," *Robotics: Science and Systems (RSS)*, 2022.

[48]. Y. Chang, L. Ballotta, L. Carlone, "D-Lite: Navigation-Oriented Compression of 3D Scene Graphs for Multi-Robot Collaboration," *IEEE Robotics and Automation Letters*, 8(11), 7527–7534, 2023.

[49]. S. Looper, J. Rodriguez-Puigvert, R. Siegwart, et al., "3D VSG: Long-term Semantic Scene Change Prediction through 3D Variable Scene Graphs," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2023.

[50]. S. Vasudevan, S. Gächter, V. Nguyen, et al., "Cognitive maps for mobile robots - an object based approach," *Robotics and Autonomous Systems*, 55(5), 359–371, 2007.

[51]. D. Schleich, T. Klamt, S. Behnke, "Value Iteration Networks on Multiple Levels of Abstraction," *Robotics: Science and Systems XV, Robotics: Science and Systems Foundation*, 2019.

[52]. D. Shah, M. Equi, B. Osinski, et al., *Navigation with Large Language Models: Semantic Guesswork as a Heuristic for Planning*. 2023. <https://arxiv.org/abs/2310.10103>.

[53]. K. Rana, J. Haviland, S. Garg, et al., "SayPlan: Grounding Large Language Models using 3D Scene Graphs for Scalable Task Planning," in *7th Annual Conference on Robot Learning*, 2023.

[54]. I. Singh, V. Blukis, A. Mousavian, et al., *ProgPrompt: Generating Situated Robot Task Plans using Large Language Models*. 2022. <https://arxiv.org/abs/2209.11302>.

[55]. Y. Xie, C. Yu, T. Zhu, et al., *Translating Natural Language to Planning Goals with Large-Language Models*. 2023. <https://arxiv.org/abs/2302.05128>.

[56]. Zachary Ravichandran, V. Murali, M. Tzes, et al., *SPINE: Online Semantic Planning for Missions with Incomplete Natural Language Specifications in Unstructured Environments*. 2024. <https://arxiv.org/abs/2410.03035>.

[57]. Y. Chen, R. Gandhi, Y. Zhang, et al., *Nl2tl: Transforming natural languages to temporal logics using large language models*. 2023. arXiv preprint arXiv:230507766.

[58]. Z. Dai, A. Asgharivaskasi, T. Duong, et al., "Optimal scene graph planning with large language model guidance," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE. 14062–14069, 2024.

[59]. J.X. Liu, Z. Yang, I. Idrees, et al., *Lang2tl: Translating natural language commands to temporal robot task specification*. 2023. arXiv preprint arXiv:230211649.

[60]. J.X. Liu, Z. Yang, I. Idrees, et al., "Grounding complex natural language commands for temporal tasks in unseen environments," in *Conference on Robot Learning*, PMLR, 1084–1110, 2023.

- [61]. B. Liu, Y. Jiang, X. Zhang, et al., *Llm+ p: Empowering large language models with optimal planning proficiency*. 2023. arXiv preprint arXiv:230411477.
- [62]. J. Kolb, K.M. Feigh, "Inferring Belief States in Partially-Observable Human-Robot Teams," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 7115–7122, 2024.
- [63]. Y. Chen, H. Kurniawati, "Pomdp planning for object search in partially unknown environment," *Advances in Neural Information Processing Systems*, 36, 53146–53157, 2023.
- [64]. M.M. de Almeida, R. Moghe, M. Akella, "Real-Time Minimum Snap Trajectory Generation for Quadcopters: Algorithm Speed-up Through Machine Learning," in *2019 International Conference on Robotics and Automation (ICRA)*, 683–689, 2019.
- [65]. J. Tordesillas, J.P. How, "PANTHER: Perception-Aware Trajectory Planner in Dynamic Environments," *IEEE Access*, 10, 22662–22677, 2022.
- [66]. J. Tordesillas, J.P. How, "Deep-panther: Learning-based perception-aware trajectory planner in dynamic environments," *IEEE Robotics and Automation Letters*, 8(3), 1399–1406, 2023.
- [67]. K. Yadav, R. Ramrakhya, S.K. Ramakrishnan, et al., *Habitat-Matterport 3D Semantics Dataset*. 2023. <https://arxiv.org/abs/2210.05633>.
- [68]. A. Majumdar, G. Aggarwal, B. Devnani, et al., *ZSON: Zero-Shot Object-Goal Navigation using Multimodal Goal Embeddings*. 2023. <https://arxiv.org/abs/2206.12403>.
- [69]. S.Y. Gadre, M. Wortsman, G. Ilharco, et al., *CoWs on Pasture: Baselines and Benchmarks for Language-Driven Zero-Shot Object Navigation*. 2022. <https://arxiv.org/abs/2203.10421>.
- [70]. P. Anderson, A. Chang, D.S. Chaplot, et al., *On Evaluation of Embodied Navigation Agents*. 2018. <https://arxiv.org/abs/1807.06757>.
- [71]. 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"*, 146–151, 1997.
- [72]. A. Bircher, K. Alexis, M. Burri, et al., "Structural inspection path planning via iterative viewpoint resampling with application to aerial robotics," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 6423–6430, 2015.
- [73]. T. Gervet, S. Chintala, D. Batra, et al., "Navigating to objects in the real world," *Science Robotics*, 8(79), eadf6991, 2023.
- [74]. D.S. Chaplot, D. Gandhi, S. Gupta, et al., *Learning to explore using active neural slam*. 2020. arXiv preprint arXiv:200405155.
- [75]. J.A. Placed, J. Strader, H. Carrillo, et al., "A survey on active simultaneous localization and mapping: State of the art and new frontiers," *IEEE Transactions on Robotics*, 39(3), 1686–1705, 2023.
- [76]. M. Kalaitzakis, B. Cain, S. Carroll, et al., "Fiducial markers for pose estimation: Overview, applications and experimental comparison of the artag, apriltag, aruco and stag markers," *Journal of Intelligent & Robotic Systems*, 101, 1–26, 2021.
- [77]. P. Kellnhofer, A. Recasens, S. Stent, et al., "Gaze360: Physically unconstrained gaze estimation in the wild," in *Proceedings of the IEEE/CVF international conference on computer vision*, 6912–6921, 2019.
- [78]. J. Wang, E. Olson, "AprilTag 2: Efficient and robust fiducial detection," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4193–4198, 2016.
- [79]. D.S. Chaplot, D. Gandhi, A. Gupta, et al., *Object Goal Navigation using Goal-Oriented Semantic Exploration*. 2020. <https://arxiv.org/abs/2007.00643>.
- [80]. M.F. Ginting, S.K. Kim, D.D. Fan, et al., *SEEK: Semantic Reasoning for Object Goal Navigation in Real World Inspection Tasks*. 2024. <https://arxiv.org/abs/2405.09822>.
- [81]. A.A.B. Abdelnabi and G. Rabadi, "Human Detection from Unmanned Aerial Vehicles' Images for Search and Rescue Missions: A State-of-the-Art Review," *IEEE Access*, 2024.
- [82]. N. Zhang, F. Nex, G. Vosselman, et al., "Training a disaster victim detection network for UAV search and rescue using harmonious composite images," *Remote Sensing*, 14(13), 2977, 2022.
- [83]. M. Lyu, Y. Zhao, C. Huang, et al., "Unmanned aerial vehicles for search and rescue: A survey," *Remote Sensing*, 15(13), 3266, 2023.
- [84]. T. Dang, F. Mascari, S. Khattak, et al., "Autonomous Search for Underground Mine Rescue Using Aerial Robots," in *2020 IEEE Aerospace Conference*, 1–8, 2020.
- [85]. S. Huang, L. Dong, W. Wang, et al., "Language is not all you need: Aligning perception with language models," *Advances in Neural Information Processing Systems*, 36, 72096–72109, 2023.
- [86]. W. Huang, F. Xia, D. Shah, et al., "Grounded decoding: Guiding text generation with grounded models for embodied agents," *Advances in Neural Information Processing Systems*, 36, 59636–59661, 2023.
- [87]. D. Shah, B. Osiński, S. Levine, et al., "Lm-nav: Robotic navigation with large pre-trained models of language, vision, and action," in *Conference on robot learning*, PMLR, 492–504, 2023.
- [88]. J. Rückin, L. Jin, F. Magistri, et al., *Informative Path Planning for Active Learning in Aerial Semantic Mapping*. 2022. <https://arxiv.org/abs/2203.01652>.
- [89]. B. Yu, H. Kasaei, M. Cao, "Frontier Semantic Exploration for Visual Target Navigation," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 4099–4105, 2023.
- [90]. M.F. Ginting, D.D. Fan, S.K. Kim, et al., *Semantic Belief Behavior Graph: Enabling Autonomous Robot Inspection in Unknown Environments*. 2024. <https://arxiv.org/abs/2401.17191>.
- [91]. X. Jin, H. Xiao, X. Shen, et al., "Predicting scene parsing and motion dynamics in the future," *Advances in neural information processing systems*, 30, 2017.
- [92]. Y. Wu, R. Gao, J. Park, et al., "Future video synthesis with object motion prediction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5539–5548, 2020.

[93]. R. Asghar, W. Liu, L. Rummelhard, et al., *Flow-guided Motion Prediction with Semantics and Dynamic Occupancy Grid Maps*. 2024. arXiv preprint arXiv:240715675.

[94]. E. Corona, A. Pumarola, G. Alenya, et al., "Context-aware human motion prediction," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 6992–7001, 2020.

[95]. G. Xu, J. Tao, W. Li, et al., "Learning semantic latent directions for accurate and controllable human motion prediction," in *European Conference on Computer Vision*, Springer, 56–73, 2024.

[96]. M. Liu, S. Tang, Y. Li, et al., "Forecasting Human-Object Interaction: Joint Prediction of Motor Attention and Actions in First Person Video," in *Computer Vision – ECCV 2020: 16th European Conference, Proceedings, Part I*, Berlin, Heidelberg, Springer-Verlag, 704–721, 2020.

[97]. M. Proetzsch, T. Luksch, K. Berns, "Development of complex robotic systems using the behavior-based control architecture iB2C," *Robotics and Autonomous Systems*, 58(1), 46–67, 2010.

[98]. P. Wolf, T. Ropertz, P. Feldmann, et al., "Combining Ontologies and Behavior-based Control for Aware Navigation in Challenging Off-road Environments," in *ICINCO 2019 - 16th International Conference on Informatics in Control, Automation and Robotics*, (2), 135–146, 2019.

[99]. R.F. Salas-Moreno, R.A. Newcombe, H. Strasdat, et al., "SLAM++: Simultaneous Localisation and Mapping at the Level of Objects," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 1352–1359, 2013.

[100]. S.W. Chen, G.V. Nardari, E.S. Lee, et al., "Sloam: Semantic lidar odometry and mapping for forest inventory," *IEEE Robotics and Automation Letters*, 5(2), 612–619, 2020.

[101]. X. Liu, G.V. Nardari, F. Cladera, et al., "Large-Scale Autonomous Flight With Real-Time Semantic SLAM Under Dense Forest Canopy," *IEEE Robotics and Automation Letters*, 7(2), 5512–5519, 2022.

[102]. T. Shan, B. Englot, "LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4758–4765, 2018.

[103]. Y. Tao, X. Liu, I. Spasojevic, et al., "3D Active Metric-Semantic SLAM," *IEEE Robotics and Automation Letters*, 9(3), 2989–2996, 2024.

AUTHORS INFORMATION

**Nguyen Van Hung¹, Truong Xuan Tung² Nghiem Hoang Nam²,
Nguyen Van Tuan¹**

¹Control, Automation in Production and Improvement of Technology
Institute, Academy of Military Science and Technology, Vietnam

²Le Quy Don Technical University, Vietnam