

Tối ưu cập nhật bản đồ cục bộ trong môi trường động sử dụng phương pháp nhận dạng vật thể

Nguyễn Ngọc Tú*, Nguyễn Văn Thương, Trần Xuân Thịnh

Trung tâm Quang điện tử, Viện Ứng dụng Công nghệ

Ngày nhận bài 4/10/2021; ngày chuyển phản biện 7/10/2021; ngày nhận phản biện 3/11/2021; ngày chấp nhận đăng 8/11/2021

Tóm tắt:

Cập nhật bản đồ của môi trường thay đổi theo thời gian là một thách thức quan trọng trong quá trình hoạt động của các loại robot di động tự hành (Autonomous mobile robots). Trong bài báo này, các tác giả đề xuất một giải pháp mới để cập nhật bản đồ cục bộ của robot trong môi trường động theo thời gian thực. Phương pháp đề xuất dựa trên sự kết hợp giữa dữ liệu quét của LiDAR (Light detection and ranging) và thuật toán nhận dạng, nhằm mục đích phát hiện, định vị các vật thể mới xuất hiện cho phép robot cập nhật bản đồ một cách liên tục. Dữ liệu quét của LiDAR cùng một khu vực tại các thời điểm khác nhau được so sánh để phát hiện sự thay đổi; khi có đủ số lần thay đổi ở cùng một vị trí, bản đồ sẽ tự động cập nhật nhằm loại bỏ các đối tượng cũ hoặc thêm vào các đối tượng mới. Nhằm tối ưu việc cập nhật bản đồ, phương pháp nhận dạng và định vị vật thể cho các đối tượng di chuyển được sử dụng để quyết định đối tượng nào sẽ được thêm vào bản đồ. Kết quả thử nghiệm được hiển thị trong môi trường mô phỏng cho thấy khả năng cập nhật bản đồ trước những thay đổi môi trường trong thời gian dài.

Từ khóa: bản đồ cục bộ, môi trường động, phương pháp nhận dạng vật thể, robot tự hành.

Chỉ số phân loại: 2.2

Đặt vấn đề

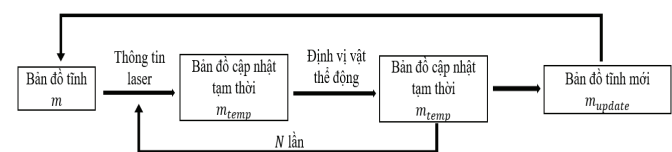
Để điều hướng tự động cho quá trình di chuyển, robot cần có thông tin của môi trường hoạt động dưới dạng mặt cắt của không gian theo hình chiếu bằng, được gọi là bản đồ 2D (bao gồm các điểm biểu diễn chướng ngại vật, vùng trống và vùng không xác định). Khi môi trường có sự thay đổi, đòi hỏi bản đồ phải được xây dựng lại, việc này tiêu tốn thời gian, gây ảnh hưởng lớn đến năng suất làm việc của robot. Vì vậy, một trong những mục tiêu của nghiên cứu robot tự hành là phát triển khả năng tự xây dựng và cập nhật bản đồ liên tục mỗi khi môi trường có sự thay đổi.

Các bản đồ được xây dựng bằng định vị và xây dựng bản đồ đồng thời (SLAM - Simultaneous localization and mapping) cung cấp cho robot trong việc tự định vị và lập kế hoạch đường đi, với điều kiện các vật thể nằm trong môi trường đó không thay đổi vị trí tương quan so với bản đồ được lập từ đầu [1]. Trong thực tế, hầu hết các môi trường hoạt động của robot trong nhà máy, nhà kho, văn phòng... luôn có thay đổi gây ra bởi sự tham gia giao thông của người, xe vận chuyển hàng, giá đỡ hàng... bởi quá trình hoạt động của hệ thống nội tại bên trong nó. Những thay đổi này có thể là do đối tượng có tính động như con người hoặc phương tiện di chuyển, hay các đối tượng có tính tĩnh như pallet, máy móc, thiết bị cố định [2-4]. Việc xem xét những thay đổi trong môi trường do các đối tượng có tính động hoặc tĩnh có thể cải thiện khả năng tự định vị của hệ thống robot [4] và cải thiện việc lập kế hoạch đường đi của các đội robot lớn [5] cũng như sự phối hợp của hệ thống nhiều robot. Hiện nay, một số nghiên cứu đã được đưa ra nhằm cập nhật bản đồ trong môi trường động [6]. D. Sun và cs (2016) [7] cải tiến thuật toán SLAM để có thể quét bản đồ khi robot vẫn đang

hoạt động nhằm tạo một bản đồ mới. X. Hu và cs (2018) [8] sử dụng kết hợp khả năng định vị của robot và thông tin môi trường từ thiết bị LiDAR [9] để cập nhật bản đồ. Tuy nhiên, các tác giả này chỉ sử dụng dữ liệu từ LiDAR để cập nhật bản đồ, dẫn đến bản đồ có thể sẽ cập nhật cả những thay đổi tạm thời do vật chuyển động gây ra. Để khắc phục hạn chế này, nhóm tác giả đề xuất một giải pháp mới là ngoài thông tin từ LiDAR, tích hợp thêm một camera 3D và thuật toán nhận dạng vật thể [10] để loại bỏ chính xác những thông tin tạm thời do các vật thể động gây ra, giúp tối ưu hóa việc cập nhật bản đồ.

Phương pháp nghiên cứu

Bản đồ ban đầu của môi trường được xây dựng bằng thuật toán SLAM tiêu chuẩn, chỉ chứa các phần tĩnh của môi trường (tức là các bức tường và các công trình lắp đặt cố định), bản đồ này được gọi là bản đồ tĩnh. Những thay đổi của môi trường được ghi nhận bởi cảm biến LiDAR. Một camera 3D và phần mềm nhận dạng đối tượng được sử dụng để có thể quan sát và định vị vật thể dựa trên ảnh độ sâu. Các thông tin từ LiDAR và camera 3D được kết hợp sử dụng để phát hiện những thay đổi trong môi trường so với bản đồ tĩnh. Sau đó, robot bắt đầu xây dựng bản đồ tạm thời và sẽ được cập nhật vào bản đồ hiện tại khi đạt đủ yêu cầu. Hình 1 mô tả sơ đồ của thuật toán được đề xuất.



Hình 1. Sơ đồ của phương pháp nghiên cứu đề xuất.

*Tác giả liên hệ: Email: ngoctu@cfoc.vn

Optimising local map updates in a dynamic environment using an object detection method

Ngoc Tu Nguyen*, Van Thuong Nguyen, Xuan Thinh Tran

Center for Optoelectronics, National Center for Technological Progress

Received 4 October 2021; accepted 8 November 2021

Abstract:

Keeping maps up to date in a time-varying environment is a big challenge in the development of autonomous mobile robots. In this paper, the authors propose a new solution to update the local map of the robot in a dynamic environment in real time. This proposed method works based on a combination of lidar scan data and object recognition and positioning algorithms to allow the robot to update the map continuously. Laser scan data for the same part of the environment at different times are compared to perform change detection. When there are enough changes in a location, the map will be edited to remove old features or add new features. To optimise updates for moving objects, authors use object recognition and positioning to decide which objects should be added to the map. The results are displayed in a simulated environment, showing the ability to maintain an up-to-date map despite long-term environmental changes.

Keywords: autonomous robot, dynamic environment, local map, object detection.

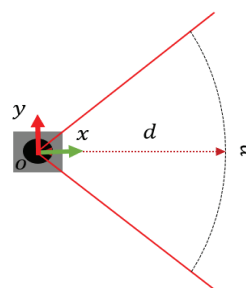
Classification number: 2.2

Cập nhật bản đồ tạm thời dựa trên thông tin cảm biến LiDAR

Trước tiên, bản đồ lưới và tín hiệu quét từ laser sẽ được biểu diễn dưới dạng ma trận. Để thuận tiện cho việc theo dõi, thuật toán quy ước như sau:

- Tại các vị trí có tồn tại vật thể, giá trị điểm ảnh trên bản đồ bằng 1.
- Tại các vị trí trống, giá trị điểm ảnh trên bản đồ bằng 0.
- Tọa độ của một điểm trong môi trường so với cảm biến LiDAR ký hiệu là $p(x, y)$.
- Tọa độ của $p(x, y)$ so với bản đồ ký hiệu là $P(X, Y)$.

Trong phạm vi nghiên cứu này, chúng tôi chỉ đề cập đến cập nhật bản đồ cục bộ, do đó chỉ các thông tin quét nằm trong phạm vi nhỏ sẽ được thu thập. Cụ thể, góc quét α sẽ tương ứng với góc rộng (vùng nhìn thấy) của camera và khoảng cách quét d sẽ trong phạm vi mà camera vẫn có thể nhận dạng được đối tượng (hình 2).

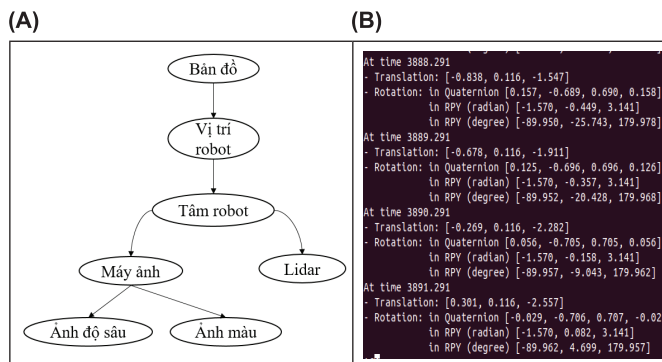


Hình 2. Mô hình quét laser.

Trước tiên, tín hiệu quét tại góc θ (với $\theta \leq \alpha$) từ laser sẽ được thu thập. Nếu xuất hiện vật cản nằm trong khoảng giới hạn d , tín hiệu sẽ trả về khoảng cách d_θ từ vật tới cảm biến. Từ d_θ và θ , tọa độ của điểm được quét $p_{\theta,d_\theta}(x, y)$ so với cảm biến được tính như sau:

$$\begin{aligned} x &= d_\theta \times \cos(\theta) \\ y &= d_\theta \times \sin(\theta) \end{aligned} \quad (1)$$

Sau khi xác định được tọa độ điểm quét so với cảm biến, ta cần chuyển đổi để thu được tọa độ điểm quét $p_{\theta,d_\theta}(X, Y)$ so với bản đồ để có thể so sánh với các điểm tương ứng trên bản đồ, nhờ vào thông tin chuyển vị. Hình 3 minh họa mối liên kết giữa cảm biến LiDAR và bản đồ. Các vector tịnh tiến T và quay R được cập nhật liên tục để đảm bảo độ chính xác khi chuyển đổi tọa độ điểm quét.



Hình 3. Mối liên hệ giữa bản đồ, thông tin từ LiDAR và máy ảnh (A) và các vector tịnh tiến, vector quay chuyển đổi thông tin giữa bản đồ và LiDAR được cập nhật liên tục (B).

Tiếp theo, ta tiến hành kiểm tra giá trị tại điểm $p_{\theta,d_\theta}(X, Y)$ trên bản đồ. Tại đây, nếu $p_{\theta,d_\theta}(X, Y)=1$, trong bản đồ đã tồn tại sẵn điểm tại vị trí này, bản đồ sẽ bỏ qua vị trí này và không cập nhật thêm. Ngược lại, trong trường hợp $p_{\theta,d_\theta}(X, Y)=0$, tức là không tồn tại vật cản trước đó, giá trị điểm ảnh tại vị trí này sẽ được cập nhật lên 1 và đưa vào bản đồ tạm thời (phương trình 2).

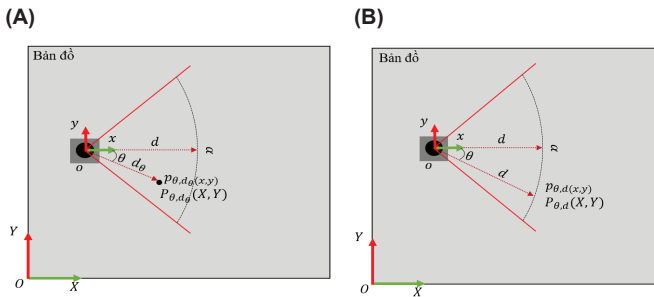
$$if \begin{cases} P_{\theta,d_\theta}(X, Y) = 1, & do \text{ nothing} \\ P_{\theta,d_\theta}(X, Y) = 0, & set \ P_{\theta,d_\theta}(X, Y) = 1. \end{cases} \quad (2)$$

Trong trường hợp tín hiệu trả về của laser cho thấy không xuất hiện vật cản trong khoảng cách d , giá trị pixel tại vị trí $p_{\theta,d_\theta}(X, Y)$ trên bản đồ sẽ được kiểm tra, tương ứng với điểm $p_{\theta,d_\theta}(X, Y)$ so với cảm biến LiDAR. Nếu $p_{\theta,d_\theta}(X, Y)=1$, trên bản đồ tính có vật thể tại vị trí này trước đó. Điều này có nghĩa là vật thể đã từng tồn tại ở vị trí đó được di dời sang vị trí khác, do vậy laser không còn quét

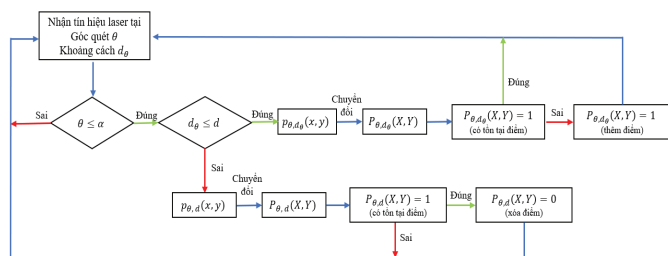
được tín hiệu từ vật thể này nữa. Lúc này, giá trị $p_{\theta,d\theta}(X, Y)$ sẽ được gán về 0. Nếu $p_{\theta,d\theta}(X, Y) = 0$, bản đồ không có gì thay đổi tại vị trí này (phương trình 3).

$$if \begin{cases} P_{\theta,d}(X, Y) = 1, & set P_{\theta,d\theta}(X, Y) = 0 \\ P_{\theta,d}(X, Y) = 0, & do nothing \end{cases} \quad (3)$$

Minh họa các điểm quét tại góc quét θ trong trường hợp có vật cản trong khoảng giới hạn d (A) và không có vật cản trong khoảng giới hạn d (B).



Hình 4. Minh họa điểm quét tại góc θ trong trường hợp có vật cản trong khoảng giới hạn d (A) và không có vật cản trong khoảng giới hạn d (B).



Hình 5. Sơ đồ phương pháp cập nhật bản đồ tạm thời dựa trên thông tin cảm biến LIDAR.

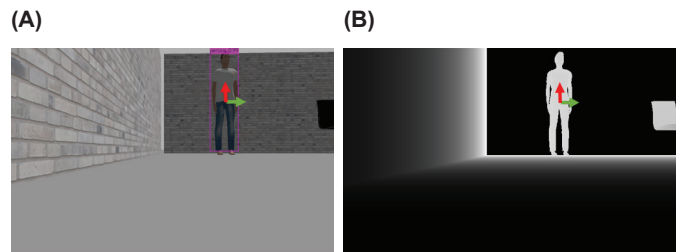
Nhận dạng, định vị vật thể động - tĩnh

Trong quá trình di chuyển, đối với môi trường xuất hiện các đối tượng chuyển động liên tục, robot sẽ gặp phải các vật cản. Các vật cản này có thể xuất hiện tạm thời rồi biến mất được gọi là vật cản động (như xe kéo, con người) hoặc xuất hiện và duy trì trong thời gian dài được gọi là vật cản tĩnh (như thùng hàng, máy móc, thiết bị mới). Các vật cản tĩnh sẽ ảnh hưởng đến quỹ đạo di chuyển của robot trong thời gian lâu dài, trong khi các vật cản động chỉ ảnh hưởng tại thời điểm tức thời. Do đó, robot cần phân biệt được đâu là vật cản động, đâu là vật cản tĩnh lâu dài để có thể đưa ra quyết định cập nhật vào bản đồ.

Để có thể nhận dạng và định vị được vật thể trong thời gian thực, cần có một mô hình nhận dạng kết hợp được hai yếu tố nhanh và chính xác. Hiện nay, các mô hình AI (trí tuệ nhân tạo) đang được phát triển mạnh và ngày càng đáp ứng được yêu cầu về tốc độ và độ chính xác, được sử dụng rộng rãi trong nhiều ứng dụng như hệ thống tự lái xe, video giám sát, tầm nhìn robot trong công nghiệp. Trong nghiên cứu này, chúng tôi đã chọn mô hình YOLO (You only look one) để áp dụng vào bài toán nhận dạng vật thể bởi độ chính xác cũng như tốc độ xử lý của mô hình này.

YOLO là một mô hình mạng tích chập neuron có khả năng nhận dạng vật thể với ưu điểm nổi trội là nhanh (100 fps khi sử dụng với GPU GTX Titan X), độ chính xác cao (lên tới 93%) [11]. Đầu vào của mô hình là một ảnh. Mô hình sẽ nhận dạng ảnh đó có đối tượng nào hay không, sau đó xác định tọa độ của đối tượng trong ảnh. Hiện tại, phiên bản YOLOv4 [11] đầy đủ cho phép nhận diện hơn 80 loại đối tượng khác nhau. Nếu cần nhận diện một đối tượng mới không được hỗ trợ bởi YOLO, chúng ta cần dạy máy (training) lại. Trong phạm vi nghiên cứu này, chúng tôi sử dụng mô hình YOLOv4 với mạng đã được học từ trước và tích hợp vào phần mềm thư viện ROS (Robot operating system).

Tuy nhiên, đầu ra của mô hình YOLO là tọa độ của đối tượng so với trục tọa độ trong ảnh. Để có thể chuyển đổi giữa tọa độ của đối tượng trong ảnh ra tọa độ của đối tượng trong bản đồ, chúng tôi kết hợp sử dụng ảnh độ sâu thu được từ camera để xác định khoảng cách từ đối tượng đến camera (hình 6), sau đó chuyển đổi ra tọa độ bản đồ. Thông tin tọa độ sẽ được sử dụng để kết hợp với bản đồ tạm thời thu được từ nội dung trước để cập nhật thành bản đồ mới hoàn thiện. Nội dung chi tiết về kết hợp thông tin để cập nhật bản đồ sẽ được trình bày ở phần sau.

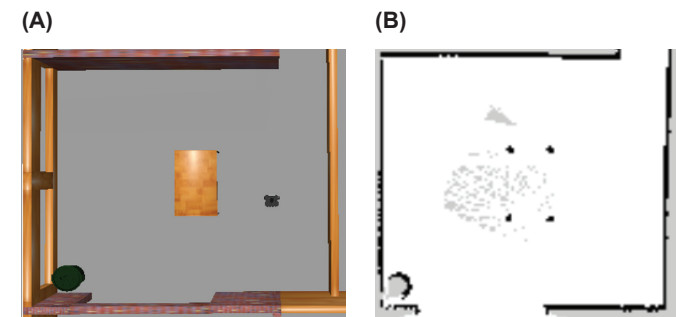


Hình 6. Thông tin của ảnh nhận dạng đối tượng (A) sẽ được chuyển sang ảnh độ sâu (B) để xác định tọa độ của vật so với robot.

Kết quả và bàn luận

Phương pháp, công cụ mô phỏng

Để đánh giá kết quả của phương pháp này, chúng tôi tiến hành mô phỏng trong chương trình Gazebo với robot Turtlebot (hình 7A). LiDAR chúng tôi sử dụng mô phỏng là LiDAR LDS-01, với góc quét 360°, giới hạn khoảng cách quét từ 0,12 đến 3,5 m. Camera độ sâu được sử dụng trong nghiên cứu này là 3D Camera Intel RealSense R200.



Hình 7. Môi trường ban đầu (A) và bản đồ tĩnh (B) được tạo từ môi trường ban đầu.

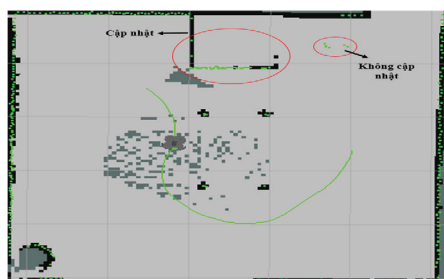
Ban đầu, một bản đồ hoàn chỉnh về môi trường với các vật thể tĩnh được xây dựng bằng Hector SLAM (hình 7B). Các thông số xây dựng bản đồ được cài đặt như sau: Độ phân giải bản đồ 0,05 m, chu kỳ cập nhật bản đồ là 2 s. Sau đó, một mô hình người và một mô hình khối lập phương (lần lượt đại diện cho vật thể động và tĩnh) được đặt vào trong môi trường tĩnh ban đầu (hình 7A). Robot mô phỏng Turtlebot sẽ tiến hành thực hiện nhiệm vụ đi tới 2 điểm cố định trên bản đồ, trong đó quỹ đạo đường đi của robot có đi qua các vật thể đã đặt thêm vào. Giá trị các thông số được cài đặt trong thí nghiệm này như sau: Khoảng cách giới hạn $d=3$ m, góc giới hạn $\alpha=60^\circ$, kích thước cửa số $M \times M=5 \times 5$, điều kiện số lần thay đổi để cập nhật tại một vị trí $n=3$.

Kết quả mô phỏng

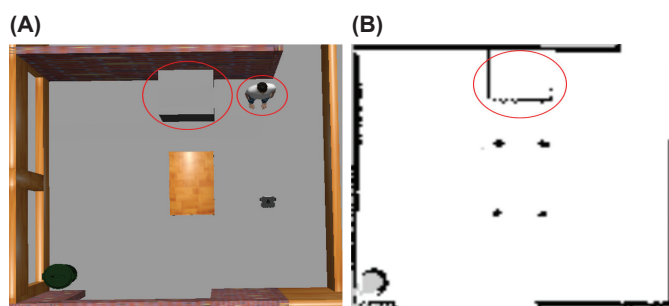
Như được thể hiện ở hình 8, robot đã phát hiện và nhận dạng được con người xuất hiện trong môi trường. Dựa vào kết quả nhận dạng và định vị này, robot sẽ xử lý dữ liệu để cập nhật vào bản đồ. Hình 9 cho thấy, mặc dù robot nhận được tín hiệu laser từ cả 2 vật thể, nhưng chỉ đối tượng tĩnh là mô hình khối được cập nhật vào bản đồ, trong khi đối tượng con người bị loại bỏ. Kết quả cuối cùng sau 3 lần di chuyển qua các vật thể, bản đồ ban đầu đã được cập nhật (hình 10).



Hình 8. Kết quả nhận dạng và định vị con người.



Hình 9. Bản đồ cập nhật tạm thời trong khi robot đang di chuyển. Thông tin laser trả về bao gồm cả thông tin của vật thể tĩnh và động nhưng bản đồ chỉ cập nhật thông tin của vật thể tĩnh.



Hình 10. Kết quả cập nhật bản đồ. (A) Môi trường thay đổi với vật thể tĩnh và động (trong khoanh tròn đỏ); (B) Bản đồ cập nhật chỉ bao gồm các thông tin của vật thể tĩnh, trong khi vật thể động không được cập nhật vào bản đồ.

Kết luận

Trong bài báo này, nhóm tác giả đã giới thiệu một phương pháp mới nhằm tối ưu hóa cập nhật bản đồ cục bộ cho robot trong môi trường có tính biến động cao như nhà máy, kho, xưởng. Phương pháp này sử dụng thông tin quét từ cảm biến LiDAR để phát hiện những đối tượng trong phạm vi nhất định xung quanh robot, kết hợp với thuật toán nhận dạng vật thể để nhận dạng và định vị các vật thể động. Sau đó, thông tin quét từ cảm biến và thông tin về vật thể sẽ được kết hợp để cập nhật vào bản đồ. Kết quả thí nghiệm trên môi trường mô phỏng cho thấy robot có khả năng cập nhật bản đồ chính xác, tối ưu khi có thể loại bỏ được những thông tin của vật thể động không cần thiết, giúp robot có thể hoạt động trong thời gian dài trong môi trường động mà không cần xây dựng lại bản đồ. Trong các nội dung nghiên cứu tiếp theo, chúng tôi sẽ tiếp tục thử nghiệm khả năng cập nhật bản đồ trong môi trường thực tế để có thể cải thiện cũng như tối ưu hiệu quả của phương pháp đề xuất, hướng tới tích hợp giải pháp vào các sản phẩm robot tự hành sẵn có trong công nghiệp.

LỜI CẢM ƠN

Nghiên cứu này được thực hiện thông qua đề tài cấp bộ “Nghiên cứu phát triển robot tự hành điều hướng thông minh sử dụng công nghệ LiDAR”. Nhóm tác giả xin trân trọng cảm ơn.

TÀI LIỆU THAM KHẢO

- [1] F. Abrate, et al. (2010), *Map Updating in Dynamic Environments*, ISR/ROBOTIK, pp.1-8.
- [2] N. Mitsou, C. Tzafestas (2007), *Temporal Occupancy Grid for Mobile Robot Dynamic Environment Mapping*, Control Automation, pp.1-8.
- [3] A.N. Walcott (2011), *Long-Term Robot Mapping in Dynamic Environments*, Thesis, Cambridge, USA.
- [4] D. Meyer-Delius, et al. (2010), “Temporary maps for robust localization in semi-static environments”, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.5750-5755.
- [5] A. Kleiner, et al. (2011), “Armo: Adaptive road map optimization for large robot teams”, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.3276-3282.
- [6] A.A. Panchpor, et al. (2018), “A survey of methods for mobile robot localization and mapping in dynamic indoor environments”, *Conference on Signal Processing and Communication Engineering Systems (SPACES)*, pp.138-144.
- [7] D. Sun, et al. (2016), “Towards effective localization in dynamic environments”, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.4517-4523.
- [8] X. Hu, et al. (2018), “Long-term localization of mobile robots in dynamic changing environments”, *IEEE Chinese Automation Congress*, pp.384-389.
- [9] <https://en.wikipedia.org/wiki/Lidar>.
- [10] Joseph Redmon, et al. (2016), *You Only Look Once: Unified, Real-Time Object Detection*, CVPR, pp.779-788.
- [11] Alexey Bochkovskiy, et al. (2020), “YOLOv4: Optimal speed and accuracy of object detection”, *arXiv: 2004.10934*, DOI: 10.48550/arXiv:2004.10934.