



HAL
open science

Éléments d'état de l'art sur la cartographie sémantique et son applicabilité en environnement industriel

Abdessalem Achour, Hiba Al-Assaad, Madeleine El-Zaher, Yohan Dupuis

► To cite this version:

Abdessalem Achour, Hiba Al-Assaad, Madeleine El-Zaher, Yohan Dupuis. Éléments d'état de l'art sur la cartographie sémantique et son applicabilité en environnement industriel. 20èmes Rencontres des Jeunes Chercheurs en Intelligence Artificielle, Jun 2022, Saint-Etienne, France. hal-03765383

HAL Id: hal-03765383

<https://hal.science/hal-03765383>

Submitted on 31 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Éléments d'état de l'art sur la cartographie sémantique et son applicabilité en environnement industriel

A. Achour^{1,3}, H. Al-Assaad¹, M. El Zaher¹, Y. Dupuis²

¹ CESI campus de Toulouse, LINEACT CESI

² CESI, LINEACT CESI

³ ENSAM, 75013 Paris, France

Résumé

La sécurité dans la collaboration homme-robot est l'un des principaux défis de l'industrie 5.0. Les robots doivent à présent s'adapter à l'humain. Ils évaluent et améliorent le système de navigation sur la base de technologies de cartographie sémantique qui permet de conduire un raisonnement clair. Ces cartes permettent aux robots de comprendre l'environnement qui les entoure pour assurer la sécurité des humains. L'objectif de cet article de synthèse est de présenter les différentes méthodes de cartographie sémantique mono-robot et l'évolution de ces systèmes, existant dans la littérature, et basés sur des techniques d'IA. Dans cette perspective, nous mettrons en avant les méthodes récentes et applicables dans le domaine industriel.

Mots-clés

Cartographie sémantique mono-robot, Fusion multi-sources, Collaboration homme-robot, Environnement industriel.

Abstract

Safety in human-robot collaboration is one of the main challenges of Industry 5.0. Now, robots must adapt to humans. They evaluate and improve the navigation system based on semantic mapping technologies that enable clear reasoning. These maps allow robots to understand the environment around them to ensure human safety. The objective of this review article is to present the different single-robot semantic mapping methods and the evolution of these systems, existing in the literature, and based on AI techniques. In this perspective, we will highlight the recent methods applicable in the industrial domain.

Keywords

Single-robot semantic mapping, Multi-source fusion, Human-Robot collaboration, Industrial environment.

1 Introduction

Dans la littérature, il existe de nombreux travaux qui s'intéressent au problème de la cartographie d'intérieur afin de proposer de nouvelles approches capables de générer une meilleure représentation de l'environnement. La majorité

des approches proposées sont basées sur des données de perceptions des robots mobiles. Elles permettent de générer deux grandes catégories de cartes : les cartes métriques et topologiques. Ces cartes sont principalement utilisées pour permettre au robot de se localiser dans son environnement et de planifier sa trajectoire. Les cartes métriques fournissent une représentation géométrique des objets de l'environnement dans un cadre de référence global. Un problème majeur de ces cartes est l'incertitude des mouvements des robots due à l'accumulation des erreurs de perception et de localisation. Les cartes topologiques représentent l'environnement sous la forme d'un graphe, où les nœuds représentent des lieux et les arêtes représentent les relations entre ces lieux. D'après [1], l'avantage principal des cartes topologiques est qu'elles font abstraction des problèmes d'incertitude dans le mouvement du robot : comme il navigue localement entre les nœuds, il n'y a pas une accumulation globale d'erreurs. Par ailleurs, ce manque des informations métriques peut être considéré comme l'inconvénient principal de ces cartes. Afin de surmonter les limites de ces deux cartes, des cartes hybrides combinant les deux ont été proposées. Ces cartes permettent d'utiliser la carte métrique pour une localisation précise et un graphe topologique global pour se déplacer d'un endroit à un autre [2]. Mais, malgré les avantages de ces différentes cartes, elles sont toutes considérées comme des représentations de bas niveau de l'environnement.

Depuis quelques années, afin de faire face aux nouveaux besoins dans le domaine de la robotique mobile, un nouveau concept, appelé cartographie sémantique, est proposé. Ce dernier permet d'ajouter des informations sémantiques à la carte afin d'avoir des robots plus efficaces, avec une diversité d'actions planifiées et capables de réagir à des événements inattendus [3]. La carte sémantique représente l'environnement avec des concepts de haut niveau et ajoute une signification sémantique aux éléments cartographiés. *Nüchter et al.* [4] la définissent comme une carte qui contient, en plus des informations spatiales, une association des éléments physiques cartographiés avec des entités de classes connues. Une base d'informations supplémentaires sur ces entités, structurée et indépendante de la carte, est également nécessaire pour permettre au robot de raisonner. Plusieurs

travaux sur la cartographie sémantique ont été réalisés, mais la plupart d'entre eux se concentrent sur la représentation d'environnements statiques. En effet, l'intérêt de la majorité de ces travaux est de déployer une carte sémantique dans le système de navigation des robots de service afin que les utilisateurs puissent les faire fonctionner à l'aide de commandes de haut niveau, par exemple en utilisant des commandes vocales ou gestuelles [5].

Récemment, un nouveau contexte industriel, l'industrie 5.0, où les humains et les robots partagent le même espace de travail [6], se met en place. En effet, les barrières physiques entre les humains et les robots sont totalement ou partiellement supprimées. Ainsi, en plus de la collaboration directe, où l'humain et le robot travaillent simultanément sur une tâche commune, un nouveau mode de collaboration apparaît, où l'humain et le robot travaillent à proximité l'un de l'autre pour effectuer des tâches différentes. Il est donc question d'améliorer les capacités des robots pour les rendre plus adaptés à cette nouvelle façon de coopération homme-robot. En effet, il est nécessaire de disposer d'une carte sémantique à jour des éléments statiques et dynamiques de l'environnement [3]. Par exemple, cette carte peut être utilisée par les robots mobiles pour éviter les collisions homme-robot / robot-robot ou pour effectuer des tâches plus complexes, pour analyser et optimiser le flux des déplacements des robots et des opérateurs, et pour fournir un inventaire en cas d'accident ou d'incendie. Alors, notre objectif est de proposer un état de l'art préliminaire des méthodes de cartographie sémantique pour étudier leur applicabilité dans un environnement industriel.

Dans ce travail, nous présentons, dans une première section, le processus de cartographie sémantique. Nous étudions les différentes méthodes d'acquisition de données géométriques et de données sémantiques proposées, ainsi que les méthodes de représentation de ces données. Ensuite, dans une deuxième section, nous étudions les caractéristiques et les besoins de l'environnement industriel. Cette étude vise à identifier les méthodes de cartographie utilisées dans le domaine domestique, qui seront applicables dans le domaine industriel.

2 Cartographie sémantique

La cartographie sémantique est le processus qui consiste à générer une carte sémantique à partir de données provenant de sources différentes, potentiellement hétérogènes. Ces données sont traitées pour générer une représentation géométrique et obtenir des informations sémantiques de l'environnement (Fig. 1). Ensuite, cette représentation et les informations sémantiques obtenues sont organisées dans une structure pour obtenir une carte sémantique. Il existe des approches de cartographie sémantique mono-robot et des approches multi-robots. Les approches étudiées dans cette section sont uniquement celles basées sur un seul robot qui peut être assisté par un humain.

2.1 Acquisition et traitement de données

Dans cette section, les trois types d'acquisition de données utilisés dans le processus de cartographie sémantique sont

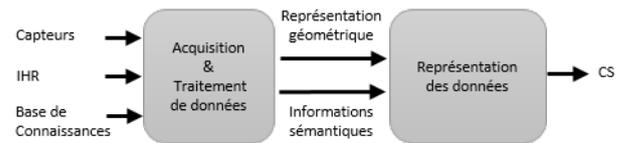


FIGURE 1 – Le processus de cartographie sémantique (CS = Carte Sémantique)

présentés : par les capteurs, par l'interaction homme-robot (IHR) et par le raisonnement.

2.1.1 Acquisition des données par les capteurs

Les robots mobiles sont dotés de différents capteurs pour réaliser la cartographie. Les données acquises par ces capteurs sont utilisées pour la représentation géométrique de l'environnement et l'extraction d'informations sémantiques, permettant de relier les entités physiques, tels que les lieux et les objets, à une base de données contenant des informations supplémentaires sur l'environnement. En général, la cartographie géométrique est basée sur les méthodes classiques SLAM (Simultaneous Localization And Mapping), qui consistent à estimer simultanément l'état d'un robot et à construire un modèle de l'environnement perçu par ses capteurs [7]. Il convient de noter que ce travail ne couvre pas les systèmes qui utilisent des informations sémantiques, telles que des informations sur les objets dans l'environnement, au profit du SLAM. Ils s'intéressent toutefois aux approches qui utilisent les systèmes SLAM classiques avec d'autres techniques telles que la segmentation et la reconnaissance d'objets pour créer une carte sémantique de l'environnement.

Qi et al. [8] proposent une approche pour créer une carte d'occupation sémantique d'un environnement domestique qui permet aux utilisateurs finaux de robots de spécifier les tâches de navigation d'une manière plus pratique en utilisant le langage naturel. Ils utilisent la méthode SLAM hors ligne basée sur des capteurs sonars, proposée dans [9], pour la création d'une grille d'occupation. Ensuite, grâce à l'odométrie et à une caméra stéréo, ils utilisent une méthode basée sur la détection d'objets et la triangulation pour y ajouter les espaces topologiques des objets détectés avec leurs étiquettes. Dans [8], les capteurs sonars ont été préférés aux capteurs laser, plus précis, car ces derniers sont peu coûteux, légers et nécessitent peu de ressources de calcul. Ce travail compare la carte obtenue à une autre créée en utilisant des capteurs laser. Le résultat final montre que la carte obtenue par le laser est plus fine, mais elle manque certains obstacles puisque le capteur scanne un plan 2D. Par contre, la carte obtenue par les capteurs sonars représente parfaitement ces obstacles car ils scannent en 3D.

Li et al. proposent dans [10] une méthode de reconstruction 3D et de segmentation de l'environnement à partir d'un flux vidéo destinée aux robots opérant dans des environnements intérieurs et extérieurs. En effet, une méthode SLAM monoculaire est utilisée car elle permet de représenter un environnement semi-dense sans avoir besoin d'informations de profondeur. De plus, elle permet de surmonter les limites

des méthodes basées sur des capteurs à petit échelle, tels que les caméras stéréo et les caméras RGB-D, qui ne fournissent une représentation fiable de l'environnement que dans leur plage de fonctionnement limitée et perdent en efficacité lors du passage d'un environnement à un autre. Dans ce travail, un modèle CNN (Convolutional Neural Network) pré-entraîné est utilisé pour l'extraction d'informations sémantiques en parallèle à la reconstruction 3D. *Niko et al.* [11] ont également proposé une approche qui utilise une caméra RGB-D pour cartographier un environnement 3D et reconstruire à la volée des modèles d'objets détectés. En effet, ORB-SLAM2 est utilisé pour la cartographie et la localisation de la caméra sur chaque image RGB-D, en parallèle les objets sont détectés sur les images RGB et le nuage de points associé est segmenté en 3D. Les travaux [10, 11] utilisent des techniques d'apprentissage profond basées sur l'image pour l'acquisition d'informations sémantiques grâce aux grands progrès réalisés par ces derniers dans la détection d'objets et la segmentation basée sur les images 2D. Dans les travaux de [12], les auteurs proposent également une méthode de détection d'objets basée sur une caméra RGB-D de type Kinect pour effectuer une navigation sémantique. L'approche proposée fusionne deux techniques de détection d'objets : la détection des contours à partir de l'image de profondeur et la détermination de la similarité basée sur les descripteurs à partir des images RGB. Par ailleurs, *Zender et al.* [13] proposent un système qui crée une représentation multi-couches d'un environnement d'intérieur. Ils utilisent un algorithme EKF-SLAM basé sur le laser pour créer les couches de bas niveau de la représentation et l'algorithme de reconnaissance d'objets RFCH (Receptive Field Cooccurrence Histograms) basé sur la caméra RGB pour établir un lien avec la couche de haut niveau.

En conclusion, le choix des capteurs ne dépend pas seulement du type de carte que l'on souhaite obtenir, notamment une carte 2D/3D, un nuage de points ou une carte en couleurs. Mais, il dépend également de niveau de la description de l'environnement que l'on souhaite obtenir. De plus, les capteurs les plus populaires pour la cartographie géométrique sont les lasers 2D/3D, et récemment les caméras RGB-D sont de plus en plus utilisées [14]. Ces caméras sont relativement peu coûteuses et peuvent fournir simultanément une image couleur et une carte de profondeur précise et de haute résolution sur une portée de 5 à 7 mètres, caractérisant la distance des objets perçus dans l'image. D'autre part, les capteurs visuels sont fréquemment choisis pour l'acquisition d'informations sémantiques, car ils permettent une représentation de l'environnement proche de celle perçue par l'œil humain. Ils permettent également de collecter des données de bas niveau, telles que des points, et des données de haut niveau, comme les catégories d'objets [2].

2.1.2 Acquisition des données sémantiques par l'interaction homme-robot

Selon [5], l'extraction automatique d'informations sémantiques à partir de capteurs est limitée et n'est pas assez robuste. En effet, la reconnaissance d'objets ou de lieux

est une tâche complexe pour les robots. Pour ces raisons, certaines techniques intègrent l'humain dans le processus d'identification d'objets ou de lieux pour la création d'une carte augmentée [3, 14].

Dans le travail de [5], un utilisateur guide le robot dans une visite de l'environnement pour sa cartographie. L'algorithme RBPF-SLAM basé sur des données laser et des données d'odométrie est utilisé pour générer en temps réel une carte d'occupation. En même temps, l'utilisateur dispose d'une application vocale basée sur IFLYTEK sur son téléphone portable pour labelliser les lieux sur cette carte. Afin d'éviter les erreurs dans la reconnaissance des commandes vocales données par l'opérateur, une étape de confirmation des informations est appliquée. Dans la même optique [13], le robot possède des connaissances préalables sur les concepts spatiaux, et le rôle de l'utilisateur est de l'assister dans le processus de labellisation des lieux. En effet, tout en marchant avec le robot, l'utilisateur exprime ce qu'il considère comme pertinent, par exemple : « C'est le couloir » ou « C'est la station de recharge ».

Par ailleurs, les auteurs de l'article [15] proposent une approche où le robot est guidé par l'utilisateur dans une visite de l'environnement opérationnel. Il perçoit son environnement et détecte l'objet pointé par un laser. Ensuite, il le segmente dans l'image et estime sa position et son orientation. Enfin, il reçoit sa description de l'utilisateur via un module de reconnaissance vocale et l'associe à sa position pour construire la représentation de l'objet à ajouter à la carte. Dans [16], *Bastianelli et al.* ont amélioré cette approche en implémentant un système qui permet d'acquérir de nouveaux objets dans la représentation grâce à une interaction continue et en ligne avec l'utilisateur après la création initiale de la carte sémantique. Les connaissances sont alors acquises selon les besoins et ajoutées progressivement à la représentation de l'environnement par le robot. Par exemple, si l'utilisateur donne une commande vocale indiquant un emplacement inconnu pour le robot, un processus d'acquisition intégré par un réseau de Petri est lancé. Pendant ce processus, l'utilisateur peut guider le système avec des commandes vocales telles que « Tournez à droite », « Suivez-moi » ou « Allez à la cuisine ». Lorsque le robot se trouve devant un nouvel emplacement ou objet à mémoriser, l'utilisateur peut pointer l'objet et indiquer au robot sa référence, par exemple « C'est la porte de secours ».

Pronobis et Jensfelt [17] proposent un algorithme qui combine des informations sur la présence d'objets avec des propriétés sémantiques des lieux telles que la taille et l'apparence pour classer les pièces. Pendant le processus de cartographie, l'utilisateur peut introduire des informations supplémentaires sur les objets présents dans la pièce à l'aide d'une interface sur le PC utilisé par le robot. Si l'utilisateur fournit une information sur l'existence d'un objet, le robot la traite comme une autre source d'information. Les travaux de *Crespo et al.* [18] ont mis en œuvre des dialogues en langage naturel entre les utilisateurs et le robot par une interface vocale et le clavier pour ajouter les catégories d'objets et leurs relations sémantiques à la carte. Par exemple, le robot peut interroger l'utilisateur sur les utilisations possibles

d'un objet ou sur les possibilités d'interaction avec les objets de l'environnement.

En conclusion, les méthodes d'acquisition d'informations par l'interaction homme-robot ajoutent une variété d'informations sémantiques, notamment des données sur les objets de la scène (la catégorie de l'objet, l'espace qu'il occupe et son occurrence dans une pièce), des données sur les lieux (la catégorie de la pièce et ces caractéristiques), et des données sur les relations sémantiques des objets (par exemple l'utilité d'un objet). De plus, il existe de nombreux travaux pour l'acquisition d'informations sémantiques dans des scènes statiques, mais seuls quelques travaux abordent le problème de l'acquisition de données dans une scène dynamique [16].

2.1.3 Acquisition des données sémantiques par le raisonnement

La troisième méthode d'acquisition de données est celle par raisonnement. Elle consiste à raisonner, en utilisant les données déjà obtenues par les autres méthodes et les connaissances fournies par une base de connaissances de sens commun, afin d'obtenir de nouvelles informations sur l'environnement. Les connaissances de sens commun sont celles que tous les humains possèdent et qu'ils ont acquises depuis leur naissance sans même en avoir conscience [19]. Cette base est généralement représentée comme une structure composée de concepts décrivant l'environnement liés par des lois.

Par exemple, *Galindo et al.* [20] proposent une structure hiérarchique d'informations conceptuelles modélisée par le système NeoClassic. Le niveau le plus bas de cette structure est composé de symboles liés à des éléments physiques détectés dans l'environnement. La structure conceptuelle est composée de lois qui relient ces symboles à des concepts et relient les concepts entre eux pour permettre un raisonnement sur les symboles de catégorie connue. Par exemple, si un élément physique détecté est lié au symbole « *four* », que ce dernier est lié au concept « *four* », et que le concept « *four* » est lié au concept « *cuisine* » par un lien « *contient* », alors le lieu actuel du robot peut être déduit comme étant la cuisine. D'autre part, dans [13], les connaissances de sens commun sur un environnement intérieur d'un bureau sont modélisées dans une ontologie. L'ontologie constitue un modèle de données représentant un ensemble de concepts dans un domaine, ainsi que les relations entre ces concepts. Sur la base de cette description, le système utilise un logiciel de raisonnement pour déduire des catégories plus spécifiques pour les zones topologiques connues. En effet, si une zone est classée comme une pièce et qu'elle contient un canapé selon les informations acquises, donc grâce aux connaissances conceptuelles données dans l'ontologie, cette zone peut être classée comme une instance de salon. De même, une ontologie est utilisée dans la carte sémantique de [21] pour créer une stratégie de planification des tâches pour les robots de service. Dans cette ontologie, le concept « *objets* » est séparé en « *objets statiques* » et « *objets dynamiques* ». Ensuite, les objets statiques sont liés aux objets dynamiques par des relations pro-

babilistes qui permettent d'obtenir par raisonnement les positions approximatives des objets dynamiques. En outre, les auteurs de [17] proposent une structure conceptuelle probabiliste où les lois reliant les concepts peuvent être prédéfinies, acquises ou déduites et peuvent être probabilistes ou pas. Les catégories et les propriétés des pièces sont estimées ensemble à l'aide d'un modèle de graphe de chaîne probabiliste. Les propriétés de la pièce telles que les objets, l'apparence, la surface et la forme sont extraites des capteurs, puis utilisées avec les relations pièces-objets fournies par la base de connaissances « *Open Mind Indoor Common Sense* » pour déduire la catégorie de la pièce. Les auteurs de [18] proposent un modèle relationnel qui permet de stocker facilement des informations dans des tables gérées par un moteur de raisonnement. La conception de ce modèle définit implicitement les relations entre les entités. Il n'est donc pas nécessaire de définir des lois entre elles comme pour d'autres structures. En utilisant ce modèle, si le robot identifie des objets dans la même pièce qui ont une utilité commune, il est possible de déduire que ces objets sont situés dans une pièce où le robot pourrait trouver d'autres objets ayant des utilités connexes. Par conséquent, de nouvelles catégories de pièces sont créées de manière autonome. L'identification des objets et des pièces permet également de détecter d'éventuelles incohérences.

Alors, dans le processus de cartographie sémantique, un système de raisonnement peut être mis en œuvre pour intégrer des informations supplémentaires dans la représentation de l'environnement. Ce système comprend une base de données contenant des connaissances de sens commun sur l'environnement d'application et un moteur de raisonnement pour exploiter ces données. La base de données est composée d'informations prédéfinies avant le processus de cartographie et d'informations acquises par les capteurs du robot. En utilisant ces informations structurées et le moteur de raisonnement, de nouvelles informations sont déduites pour enrichir la représentation.

2.2 Représentation de la carte sémantique

Afin d'exploiter les données sémantiques extraites par les méthodes présentées précédemment, il est nécessaire de les organiser dans une structure adaptée à l'application.

La méthode la plus simple est d'attacher directement les informations sémantiques aux éléments physiques associés dans la carte géométrique. Par exemple, *Sunderhauf et al.* [11] utilisent ORB-SLAM2 pour créer un nuage de points de l'environnement, et les points associés à chaque objet détecté sont modélisés par une couleur différente. Dans la même optique, une carte d'occupation sémantique est créée en ajoutant des MBRs (Minimum Bounding Rectangles), qui représentent les espaces occupés par les objets, à leurs zones topologiques dans la carte [8]. Chaque nouvelle catégorie d'objet est représentée par une couleur différente. *Zhao et al.* [5] utilisent également le module Gmapping de ROS pour créer une carte d'occupation de l'environnement. Au cours de ce processus, lorsque le robot reçoit une information sémantique, il la combine avec sa position et ajoute un nouveau nœud sémantique à la carte pour la représenter.

Les cartes obtenues par ces méthodes permettent une interprétation directe des informations sémantiques en position. Par conséquent, elles peuvent être facilement déployées dans les systèmes de navigation des robots mobiles pour effectuer des tâches de navigation sémantique. Cependant, ces représentations simples ne permettent pas au robot d'effectuer des tâches complexes telles que le raisonnement de mission. Par conséquent, d'autres représentations plus intéressantes ont été proposées, à savoir des représentations hiérarchiques, où les informations spécifiques sont placées dans les niveaux inférieurs et les informations abstraites dans les niveaux supérieurs.

Pronobis et al. [17] proposent une représentation qui donne au robot la capacité de déduire les catégories sémantiques des pièces en utilisant les propriétés de l'espace, notamment l'apparence, la surface, la forme et les objets. Ils représentent la connaissance spatiale par une structure hiérarchique à 4 couches : la couche sensorielle, la couche topologique, la couche des catégories et la couche conceptuelle. Le niveau le plus bas de la structure contient les données de perception des capteurs et le 4^{ème} niveau contient les connaissances conceptuelles abstraites. La couche sensorielle contient une carte métrique de l'environnement. La couche topologique contient un graphe topologique avec des nœuds représentant des lieux et des arêtes codant le chemin vers un autre nœud. La couche des catégories contient les objets reconnus et les propriétés de l'espace obtenues par des modèles de classification géométriques et visuels. Enfin, la couche conceptuelle contient une ontologie statique de connaissances de sens commun et des relations reliant les concepts de cette ontologie aux connaissances de bas niveau perçues dans les trois autres couches. Cette structure permet non seulement au robot de classer les pièces, mais aussi de prédire l'existence d'objets, les propriétés de l'espace et les espaces inexplorés. Par ailleurs, il existe d'autres travaux qui séparent les connaissances spatiales des connaissances conceptuelles pour obtenir une représentation plus organisée de l'environnement. Par exemple, dans [20], un modèle de représentation avec deux structures hiérarchiques est proposé pour permettre au robot d'exploiter les informations sémantiques recueillies par ces capteurs dans les tâches de navigation. La première structure est une hiérarchie qui représente les connaissances spatiales par trois niveaux d'abstraction. Le premier niveau, le plus bas, contient des cartes d'occupation locale, des lieux et des images, stockés par le robot. Le deuxième niveau contient un graphe topologique de l'environnement. Et le troisième niveau contient un nœud abstrait qui représente l'ensemble de l'environnement. La deuxième structure est une hiérarchie qui représente les connaissances conceptuelles par quatre niveaux. Elle est codée manuellement en amont avec le langage NeoClassic. Le niveau le plus haut contient un nœud appelé « *Thing* » à partir duquel sortent deux branches « *Pièces* » et « *Objets* ». Le troisième niveau contient les catégories d'objets et de pièces, et le niveau le plus bas contient les instances d'objets et de lieux reconnus par le robot. Les deux hiérarchies sont liées par le processus d'ancrage qui consiste à lier les enti-

tés reconnues à différents niveaux de la hiérarchie spatiale aux symboles qui les désignent dans la hiérarchie conceptuelle. Ces liens permettent au robot d'exploiter les informations sémantiques pour déterminer les erreurs de localisation en raisonnant sur les emplacements attendus des objets et pour effectuer des tâches de navigation sémantique. Également, dans [16], la connaissance du robot est divisée en deux parties : la connaissance du monde, qui représente la connaissance spécifique d'un certain environnement, et la connaissance du domaine, qui est la connaissance générale du domaine d'application. La représentation de la connaissance du monde contient les éléments suivants : une grille d'occupation, des cartes cellulaires pour représenter les emplacements locaux, un graphe topologique de l'environnement global et des instances d'éléments physiques reconnus avec leurs catégories et propriétés. D'autre part, la connaissance du domaine est composée comme dans [20] par les concepts impliqués dans l'environnement ainsi que leurs propriétés et relations. Cependant, pour ce travail, contrairement à [20, 17], l'objectif de l'information sémantique ajoutée par la connaissance du domaine n'est pas de classer automatiquement les pièces ou d'effectuer une navigation sémantique, mais plutôt d'obtenir et d'associer à chaque objet perçu par le robot ces propriétés spatiales et fonctionnelles afin de pouvoir raisonner sur les tâches liées à ces objets.

En conclusion [14], les représentations qui consistent à visualiser les informations sémantiques par différentes couleurs sur la carte permettent de montrer les résultats de l'acquisition des connaissances sémantiques, mais elles ne sont pas faciles à implémenter dans les tâches des robots. Par contre, les représentations par structure hiérarchique ne permettent pas de visualiser ces connaissances sur une carte, mais elles permettent de structurer et d'organiser les informations par niveaux d'abstraction pour rendre le robot capable de raisonner.

3 Synthèse

Cette section présente une synthèse des méthodes de cartographie sémantique proposées dans ce travail et étudie leur applicabilité dans un environnement industriel d'intérieur. En effet, ces environnements présentent des caractéristiques et des contraintes spécifiques à prendre en compte pour leur cartographie sémantique. Ils sont généralement de grande taille, dynamiques et à structure changeante [22]. Ils possèdent leurs propres concepts et informations sémantiques. De plus, dans le contexte de l'industrie 5.0, où l'homme et le robot partagent le même espace de travail, le renforcement de la sécurité de la collaboration homme-robot est un besoin majeur de cet environnement.

Comme le montre le tableau 1, la plupart des approches dans la littérature [13, 16, 5, 10, 18, 8, 21], sont proposées pour cartographier des environnements domestiques. Ceci est évident puisque ces dernières années, plusieurs travaux ont porté sur la facilitation de l'intégration des robots de service dans les environnements domestiques et sur l'amélioration de leur fonctionnement en intégrant la navigation

Réf	Environnements	Approche	Sources d'informations sémantiques	La sémantique ajouté	
				Sur les objets	Sur les lieux
[16]	- Env. domestique - Env. de bureaux - Dynamique - Petite échelle	Approche incrémentale par interaction homme-robot	IHR	Catégories, positions, tailles et caractéristiques	Catégories, positions du robot
[18]	- Env. domestique	-	Raisonnement	Catégories, instances utilités, caractéristiques, relations avec les lieux	Catégories et instances
[10]	- Env. d'intérieur et d'extérieur - Moyennement dense	LSD-SLAM monoculaire, Segmentation 2D avec CNN	Capteurs	Catégories	-
[17]	- Env. de bureaux - Large et non structuré	EKF-SLAM, Reconnaissance et Raisonnement d'instances, Classification des propriétés	Capteurs - IHR - Raisonnement	Catégories, relations avec d'autres concepts de l'environnement	Apparence, surface et forme
[8]	- Env. domestique - Petite échelle	SLAM basé sur des capteurs sonars [9], Détection et triangulation d'objets	Capteurs	Catégories	Catégories
[11]	- Env. de bureaux - Espacé	RGBD-SLAM (ORB-SLAM2), Détection d'objets, Segmentation 3D avec CNN	Capteurs	Catégories et instances	-
[21]	- Env. domestique - Dynamique - Non structuré	Monocular SLAM, Détection d'objets	Capteurs - Raisonnement	Catégories, relations avec d'autres concepts de l'environnement	Catégories
[13]	- Env. domestique - Petite échelle	SLAM à base de vision, Reconnaissance d'objets	Capteurs - IHR - Raisonnement	Catégories et instances	Catégories et instances
[5]	- Env. domestique - Petite échelle	RBPF-SLAM basé sur les données laser et odométrie, Reconnaissance vocale	IHR	-	Catégories

TABLE 1 – Résumé des différentes approches de cartographie sémantique citées dans cet article (Env. = Environnement, IHR = Interaction Homme-Robot).

sémantique [3]. D'autres travaux [17, 11], se sont intéressés aux environnements de bureaux ayant les mêmes caractéristiques qu'un environnement domestique. Contrairement aux environnements industriels, ces derniers sont généralement de petite taille et leur structure ne varie pas aussi rapidement que celle d'un environnement industriel.

Les méthodes SLAM sont choisies en fonction du type de l'environnement à cartographier, par exemple, SLAM semi-dense est utilisé dans [10] pour cartographier un environnement moyennement dense. D'autre part, EKF-SLAM est utilisée dans [17] pour cartographier un grand environnement de bureaux non structuré, et RBPF-SLAM est utilisée dans [5] pour cartographier un environnement domestique de petite taille. Les méthodes SLAM basées sur des capteurs laser sont considérées comme les plus robustes dans l'environnement industriel [22], car elles peuvent fournir un positionnement précis dans un environnement changeant et en présence d'obstacles dynamiques. Alors, les approches [13, 17, 16, 5] utilisant ces capteurs conviennent le mieux à cet environnement. En particulier, l'approche de *Pronobis et al.* [17] qui traite un environnement large et non structuré. En ce qui concerne les informations sémantiques collectées, elles dépendent principalement de l'application à réaliser. Si l'objectif est de permettre la navigation sémantique comme dans [8, 5], une carte d'occupation sémantique, où les catégories d'objets et les lieux topologiques sont associés à leurs positions dans la grille d'occupation, est suffisante. Dans les travaux récents, ces informations sont collectées par des techniques de détection/reconnaissance d'objets ou des techniques de segmentation basées sur l'apprentissage profond à partir d'images RGB ou RGB-D grâce à leurs avancées. Par contre, pour permettre au robot de raisonner, par exemple de déterminer des lieux à partir d'objets [17] ou d'analyser une situation pour prendre une décision [18], la représentation de la carte doit être plus riche en informations sémantiques, par exemple la descrip-

tion des objets et leur utilité. Ces données sont généralement fournies par une base de connaissances commune ou bien introduites par l'interaction homme robot.

Toutes les méthodes d'acquisition de données sémantiques mentionnées dans cet article peuvent être adaptées à l'environnement industriel. Les méthodes d'acquisition basées sur l'interaction homme-robot [13, 17, 5, 16] peuvent être transposées telles quelles, puisque les informations sont fournies par des humains. Pour les méthodes basées sur des capteurs [17, 10, 11, 8, 21], il est nécessaire de collecter des données de l'environnement industriel pour entraîner les modèles de détection/reconnaissance d'objets ou les modèles de segmentation 2D/3D. Par ailleurs, pour l'acquisition basée sur le raisonnement, il est nécessaire de créer de nouvelles structures de connaissances adaptées à l'environnement industriel. En effet, les ontologies de connaissances conceptuelles proposées [13, 17, 21] sont composées de concepts et de liaisons liées aux environnements domestiques, par exemples les catégories de pièces (cuisine, salon, etc..), les catégories d'objets (meuble) et les liaisons (le four se trouve dans la cuisine). Ces ontologies ne permettent pas le raisonnement dans des environnements industriels. Donc, il faut créer d'autres ontologies avec de nouveaux concepts et relations. Cependant, le modèle relationnel sémantique proposé dans [18] peut être adapté directement à l'environnement industriel. Ce dernier comprend la représentation physique et conceptuelle des objets et des lieux, les utilités des objets et les relations sémantiques liant les objets et les lieux. Son avantage est que les relations entre les concepts sont incluses implicitement par la conception du modèle. Il suffit donc de modifier les informations sur l'environnement domestique par des informations sur l'environnement industriel.

Concernant la représentation sémantique des objets dynamiques, il existe peu de travaux qui s'intéressent à ce problème. Les approches [13, 17, 5, 10, 11, 18, 8] s'intéressent

uniquement à la cartographie sémantique initiale de l'environnement, et par conséquent ne permettent pas la représentation d'objets dynamiques. L'approche [16], prévoit après cette cartographie initiale, un système de mise à jour en ligne de la carte sémantique avec les nouveaux objets introduits dans l'environnement. Le processus d'acquisition de nouvelles données sémantiques est déclenché par une commande vocale pendant le fonctionnement du robot, mais l'opérateur doit l'assister tout au long de ce processus. Cette approche permet de mettre à jour la carte, mais pas en temps réel et nécessite la présence permanente d'un opérateur pour y procéder. Elle n'est pas adaptée à l'environnement industriel car l'espace de travail dans ce dernier est très dynamique et changeant. Il est partagé entre les robots et les humains, et des objets sont introduits et d'autres sont retirés à chaque instant. Donc, la représentation des objets dynamiques doit être mise à jour en temps réel pour assurer la sécurité dans les collaborations homme-robot et robot-robot. Par ailleurs, dans [21], le système de planification de tâches basé sur une carte sémantique, traite également les objets dynamiques. Ce dernier permet de déterminer les positions approximatives de ces objets à travers des liens sémantiques probabilistes qui les relient aux objets statiques, mais sans pour autant les représenter ou les mettre à jour sur la carte.

En conclusion, certaines des approches de cartographie sémantique mono-robot peuvent être adaptées pour représenter de grands environnements industriels. En effet, afin de cartographier les éléments statiques de cet environnement, il est nécessaire d'opter pour des approches basées sur des algorithmes SLAM adaptés aux environnements denses et de grande taille. En outre, les méthodes basées sur le laser sont considérées comme les plus robustes pour l'environnement industriel car elles permettent un positionnement précis dans un environnement changeant et en présence d'obstacles. Pour l'acquisition de données sémantiques, des techniques de segmentation, de reconnaissance d'objets et d'IHR peuvent être utilisées. De plus, il est nécessaire de disposer d'une base de connaissances conceptuelles adaptée à cet environnement et à la mission du robot afin de lui fournir d'autres informations sémantiques. En ce qui concerne la représentation des éléments dynamiques, il existe seulement quelques méthodes qui considèrent l'environnement dynamique et la mise à jour de la carte. Les systèmes proposés sont basés sur l'IHR et ne sont pas capables de mettre à jour les positions de ces éléments en temps réel. Cependant, cette condition est nécessaire dans les environnements industriels, où la structure change rapidement, les objets sont introduits ou retirés à tout moment, et les opérateurs sont à proximité des robots. Par conséquent, les approches présentées sont capables de créer une carte sémantique initiale de l'environnement avec une représentation des objets statiques et une représentation initiale des objets dynamiques. Mais, elles ne sont pas en mesure de mettre à jour la carte avec leurs nouvelles positions et les nouveaux éléments introduits en temps réel.

Pour répondre à ce besoin, les extensions possibles des différentes méthodes citées dans cet article sont les méthodes

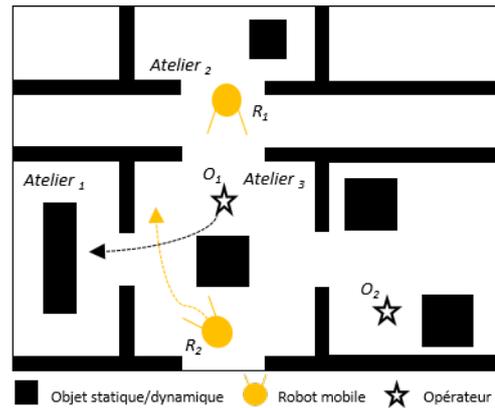


FIGURE 2 – Cas d'utilisation de la cartographie sémantique multi-robots dans un environnement industriel

de cartographie sémantique multi-robots. La Figure 2 présente un exemple de situation rencontrée dans un environnement industriel. Les trajectoires à suivre par le robot R_2 et l'opérateur O_1 illustrées par des flèches, sont concordantes. Dans ce cas, R_2 ne peut pas voir O_1 à cause de l'obstacle au milieu de l'Atelier₃. Il est donc nécessaire de partager des informations entre R_1 et R_2 pour rendre O_1 visible à R_2 . Ainsi, R_2 peut anticiper un accident avec O_1 par un raisonnement et une perception augmentée et partagée avec R_1 . Il y a peu de travaux qui se concentrent sur les problématiques de la cartographie multi-robots [23, 24, 25]. En effet, selon [24], la majorité des travaux existants s'intéressent soit à la cartographie sémantique mono-robot, soit à la cartographie géométrique collaborative [26].

4 Conclusion

Cet article propose un état de l'art des méthodes de cartographie sémantique mono-robot. Dans la littérature, les méthodes SLAM classiques sont généralement utilisées pour générer la représentation géométrique de l'environnement. D'autre part, plusieurs méthodes sont utilisées pour l'acquisition d'informations sémantiques, notamment des méthodes basées sur les capteurs, sur les IHR et sur le raisonnement. Cette étude montre que la plupart des travaux de cartographie sémantique se concentrent sur les environnements domestiques afin d'y faciliter l'intégration des robots de service. Le présent travail étudie donc la capacité de transposition de ces méthodes dans des environnements industriels de grande taille, dynamiques et structurellement changeants, où les robots et les humains partagent le même espace de travail.

En conclusion, il se révèle que les méthodes mono-robot sont efficaces pour cartographier un environnement statique, et peuvent donc être utilisées pour la cartographie sémantique initiale d'un environnement industriel. Cependant, elles ne seront pas en mesure d'établir une cartographie en temps réel de cet environnement dynamique. Une piste possible pour adresser cette problématique est la fusion de cartes sémantiques individuelles de différents robots pour créer une carte sémantique partagée, représentant

en temps réel les objets statiques et dynamiques de l'environnement.

Références

- [1] F. Wang, C. Zhang, F. Tang, H. Jiang, Y. Wu, and Y. Liu, "Lightweight object-level topological semantic mapping and long-term global localization based on graph matching," *ArXiv*, vol. abs/2201.05977, p. 9, 2022.
- [2] Q. Liu, R. Li, H. Hu, and D. Gu, "Extracting semantic information from visual data : A survey," *Robotics*, vol. 5, no. 1, p. 8, 2016.
- [3] J. Crespo Herrero, J. C. Castillo Montoya, Ó. Martínez Mozos, and R. I. Barber Castaño, "Semantic information for robot navigation : a survey," *Applied Sciences*, vol. 10, no. 2, p. 497, 2020.
- [4] A. Nüchter and J. Hertzberg, "Towards semantic maps for mobile robots," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 915–926, 2008.
- [5] C. Zhao, W. Mei, and W. Pan, "Building a grid-semantic map for the navigation of service robots through human–robot interaction," *Digital Communications and Networks*, vol. 1, no. 4, pp. 253–266, 2015.
- [6] D. TIHAY and N. Perrin, "Robotique collaborative : perception et attentes des industriels," *Hygiène et Sécurité du Travail, INRS*, no. 250, pp. 50–57, 2018.
- [7] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping : Toward the robust-perception age," *IEEE Transactions on robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [8] X. Qi, W. Wang, M. Yuan, Y. Wang, M. Li, L. Xue, and Y. Sun, "Building semantic grid maps for domestic robot navigation," *International Journal of Advanced Robotic Systems*, vol. 17, no. 1, p. 172988141990006, 2020.
- [9] K. Lee, S.-J. Lee, M. Kölsch, and W. K. Chung, "Enhanced maximum likelihood grid map with reprocessing incorrect sonar measurements," *Autonomous Robots*, vol. 35, no. 2, pp. 123–141, 2013.
- [10] X. Li and R. Belaroussi, "Semi-dense 3d semantic mapping from monocular slam," *arXiv preprint arXiv :1611.04144*, 2016.
- [11] N. Sünderhauf, T. T. Pham, Y. Latif, M. Milford, and I. Reid, "Meaningful maps with object-oriented semantic mapping," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5079–5085, IEEE, 2017.
- [12] C. Astua, R. Barber, J. Crespo, and A. Jardon, "Object detection techniques applied on mobile robot semantic navigation," *Sensors*, vol. 14, no. 4, pp. 6734–6757, 2014.
- [13] H. Zender, O. M. Mozos, P. Jensfelt, G.-J. Kruijff, and W. Burgard, "Conceptual spatial representations for indoor mobile robots," *Robotics and Autonomous Systems*, vol. 56, no. 6, pp. 493–502, 2008.
- [14] X. Han, S. Li, X. Wang, and W. Zhou, "Semantic mapping for mobile robots in indoor scenes : A survey," *Information*, vol. 12, no. 2, p. 92, 2021.
- [15] G. Randelli, T. M. Bonanni, L. Iocchi, and D. Nardi, "Knowledge acquisition through human–robot multimodal interaction," *Intelligent Service Robotics*, vol. 6, no. 1, pp. 19–31, 2013.
- [16] E. Bastianelli, D. D. Bloisi, R. Capobianco, F. Cossu, G. Gemignani, L. Iocchi, and D. Nardi, "On-line semantic mapping," in *2013 16th International Conference on Advanced Robotics (ICAR)*, pp. 1–6, IEEE, 2013.
- [17] A. Pronobis and P. Jensfelt, "Large-scale semantic mapping and reasoning with heterogeneous modalities," in *2012 IEEE international conference on robotics and automation*, pp. 3515–3522, IEEE, 2012.
- [18] J. Crespo, R. Barber, and O. Mozos, "Relational model for robotic semantic navigation in indoor environments," *Journal of Intelligent & Robotic Systems*, vol. 86, no. 3, pp. 617–639, 2017.
- [19] K. Darlington, "Common sense knowledge, crucial for the success of ai systems," *OpenMind BBVA*, 2020.
- [20] C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J.-A. Fernandez-Madrigal, and J. González, "Multi-hierarchical semantic maps for mobile robotics," in *2005 IEEE/RSJ international conference on intelligent robots and systems*, pp. 2278–2283, IEEE, 2005.
- [21] Z. Wang and G. Tian, "Hybrid offline and online task planning for service robot using object-level semantic map and probabilistic inference," *Information Sciences*, vol. 593, pp. 78–98, 2022.
- [22] S. Macenski and I. Jambrecic, "Slam toolbox : Slam for the dynamic world," *Journal of Open Source Software*, vol. 6, no. 61, p. 2783, 2021.
- [23] G. S. Martins, J. F. Ferreira, D. Portugal, and M. S. Couceiro, "Modsem : modular framework for distributed semantic mapping," *Poster Papers*, p. 12, 2019.
- [24] Y. Yue, C. Zhao, R. Li, C. Yang, J. Zhang, M. Wen, Y. Wang, and D. Wang, "A hierarchical framework for collaborative probabilistic semantic mapping," in *2020 IEEE international conference on robotics and automation (ICRA)*, pp. 9659–9665, IEEE, 2020.
- [25] V. Tchuiev and V. Indelman, "Distributed consistent multi-robot semantic localization and mapping," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4649–4656, 2020.
- [26] R. Dubois, *Méthodes de partage d'informations visuelles et inertielles pour la localisation et la cartographie simultanées décentralisées multi-robots*. PhD thesis, École centrale de Nantes, 2021.