



HAL
open science

Exploration autonome et construction de carte cognitive en environnement inconnu basées vision omnidirectionnelle uniquement

Romain Marie, Ouiddad Labbani-Igbida, El Mustapha Mouaddib

► To cite this version:

Romain Marie, Ouiddad Labbani-Igbida, El Mustapha Mouaddib. Exploration autonome et construction de carte cognitive en environnement inconnu basées vision omnidirectionnelle uniquement. Orasis, Congrès des jeunes chercheurs en vision par ordinateur, Jun 2013, Cluny, France. hal-00829404

HAL Id: hal-00829404

<https://hal.science/hal-00829404>

Submitted on 5 Jun 2013

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.

Exploration autonome et construction de carte cognitive en environnement inconnu basées vision omnidirectionnelle uniquement

Romain Marie

Ouiddad Labbani-Igbida

El Mustapha Mouaddib

Laboratoire Modélisation, Information & Systèmes
Université de Picardie Jules Verne, Amiens, France

prenom.nom@u-picardie.fr

Résumé

Dans ce papier, nous présentons une nouvelle méthode pour l'exploration autonome de robot et la construction de carte topologique en environnements inconnus à partir d'un unique capteur de vision catadioptrique. Nous proposons une approche incrémentale qui permet au robot d'extraire et de combiner plusieurs représentations spatiales construites à partir des images acquises : la détection de l'espace libre navigable, l'extraction des topologies locale et globale ainsi que la construction de signatures de lieu. Des résultats expérimentaux en environnement extérieur à large échelle démontrent l'efficacité de l'approche et ouvrent de nouvelles perspectives pour les méthodes de navigation autonome basées vision, qui sont encore aujourd'hui un problème ouvert en robotique.

Abstract

In this paper, we address the problem of exploration and cognitive map building in totally unknown environments for a mobile robot equipped with a sole catadioptric sensor. Multiple representations for spatial knowledge are built upon visual information only : free space detection, local space topology extraction, place signatures and topological links inferred from robot actions. We show how these spatial representations are nested together to incrementally build a spatial map of the explored robot space. The efficiency of the approach is evaluated in real world experiments, giving new solutions to the real-time simultaneous localization, navigation and topological map building.

1 Introduction

La localisation, navigation et cartographie simultanée basée vision ou apparence est un domaine très actif de la robotique mobile, du fait notamment de l'information très riche fournie par le capteur visuel et des étendues des applications dans des environnements complexes et incertains. Une littérature abondante existe sur le sujet et nous permet de distinguer deux grandes familles : les approches métriques nommées souvent *visual slam*, qui aboutissent à des cartes basées sur des caractéristiques géométriques et l'estimation de pose du robot [1, 2], et les méthodes basées

apparence [3, 4], dites topologiques, qui construisent une carte à un niveau plus structurel, en définissant des noeuds représentant des lieux, et des liens les reliant. Certains travaux utilisent la complémentarité de ces deux approches, en combinant l'information géométrique locale avec une information structurelle plus globale.

Les approches topologiques considèrent souvent une connaissance nécessaire a priori de l'environnement, et l'utilisent comme référence pour la localisation et navigation. Cette connaissance, dans la plupart des travaux, est acquise par un processus préalable d'apprentissage où le robot est guidé dans l'environnement. Le travail présenté dans cet article s'intéresse quant à lui à l'exploration et à la construction simultanées de carte topologique, à partir d'un unique capteur de vision catadioptrique, sans connaissance préalable du terrain. Pour cela, nous proposons une solution intégrant différentes représentations spatiales construites par le robot pendant un processus d'exploration autonome et uniquement à partir des images acquises :

- La topologie locale définie à partir du squelette de l'espace libre (lui-même extrait depuis l'image) ;
- Une signature visuelle construite parallèlement en sélectionnant dans l'image, l'information la plus pertinente.

Ces deux représentations sont fusionnées pour construire incrémentalement la carte topologique. Le squelette de l'espace libre permet de définir des trajectoires virtuelles qui guident l'exploration du robot, et de générer des liens reliant les différents lieux visités. Ces derniers sont indexés en utilisant les signatures construites caractérisant des noeuds dans le processus de construction de carte. De par leur invariance, elles permettent de localiser le robot et de gérer les problèmes de fermeture de boucle.

L'article est organisé comme suit. Après une brève présentation de l'état de l'art dans Section 2, on introduit le principe utilisé pour la détection de l'espace libre (Section 3). Le processus de construction de signature est présenté dans Section 4 alors que la Section 5 décrit comment les signatures et la topologie sont imbriquées pour construire une carte topologique. Pour valider l'approche, des résultats expérimentaux sont présentés en Section 6 avant de conclure

2 Etat de l'art

Les méthodes basées SLAM (Simultaneous Localization And Mapping) visuel (voir [5, 6] pour une revue) proposent des outils de sélection et de mise en correspondance de caractéristiques visuelles [7, 8, 9], menant à la construction de cartes géométriques précises de l'environnement.

De leur côté, les méthodes purement topologiques [3, 10] connaissent un essor important du fait des avancées récentes en reconnaissance d'images basée apparence. On peut ainsi caractériser et indexer un lieu visité par le robot en utilisant des signatures visuelles construites à partir de l'image. L'espace navigable du robot est représenté sous la forme d'un graphe où les nœuds correspondent à des lieux visités et les liens caractérisent les déplacements du robot entre les nœuds. Ces approches permettent une représentation compacte d'environnements très larges, sans utiliser une référence globale absolue, ni l'information géométrique. En revanche, elles sont généralement peu adaptées pour une navigation précise (et sécurisée). Dans [11], les auteurs proposent une carte persistante utilisant RatSLAM, un système SLAM bioinspiré, où le robot met en permanence à jour sa représentation du monde. Certaines approches hybrides étendent la notion de cartographie topologique en intégrant l'information métrique issue de l'odométrie du robot [12], ou bien une représentation géométrique permettant une mise à jour des nœuds [13].

La principale problématique dans les approches basées topologie consiste à identifier de manière fiable les différents lieux, tout en étant robuste au problème d'ambiguïté perceptive où des lieux distants peuvent présenter une apparence fortement similaire. Plusieurs solutions ont été proposées dans la littérature, basées sur des descripteurs locaux de l'image [1] ou globaux [4] (se référer à [14] pour une étude comparative de ces méthodes).

Ce papier partage aussi quelques idées avec le modèle spatial sémantique hiérarchique de Kuipers [15] et le contrôle bas niveau utilisant le Voronoï généralisé de Choset [16]. Ces travaux posent en effet quelques prérequis pour des algorithmes incrémentaux de localisation et de cartographie dans une approche topologique sans connaissance à priori de l'environnement.

3 Segmentation de l'espace libre

L'extraction de l'espace libre visible dans l'image est réalisée grâce à un algorithme de segmentation rapide présenté dans [17]. Il est basé sur la propagation d'un contour actif utilisant la texture du sol, et intègre notamment la gestion des faux obstacles isolés.

A noter que de par la nature des images omnidirectionnelles (voir Figure 1), l'initialisation du contour actif se trouve facilitée en définissant un contour initial englobant la projection du robot dans l'image. Bien évidemment, le robot n'a aucune connaissance à priori, que ce soit sur la taille ou la forme de l'espace libre.

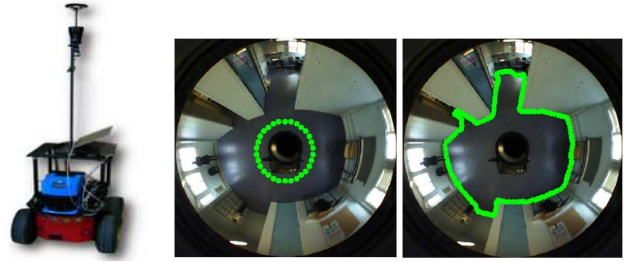


FIGURE 1 – (Gauche) : La plateforme expérimentale (robot Pioneer 3AT + capteur catadioptrique). (Droite) : Segmentation de l'espace libre en utilisant [17]. Le contour actif est initialisé autour de la projection du robot dans l'image (*centre*), et se propage jusqu'à atteindre les frontières de l'espace libre.

Les frontières de l'espace libre produisent une partition de l'image en deux domaines bien distincts (Figure 2) :

- 1- La région intérieure, correspondant à l'espace libre, est utilisée pour l'extraction de la topologie locale (et par extension globale), et sert de base à la commande du robot.
- 2- La partie extérieure sert à construire la signature, utilisée comme représentation du lieu visible dans l'image.

4 Construction de la signature de lieu

De façon similaire à [14], nous utilisons la théorie de l'invariance de Haar pour définir une signature visuelle calculée en temps réel par le robot, et qu'il utilise pour identifier le lieu où il se trouve. Tout en conservant l'idée originale, nous introduisons une nouvelle formulation :

- Nous proposons l'utilisation de descripteurs "région" à la place des descripteurs "contour" utilisés dans [14], qui, bien qu'ils soient performants pour la représentation de lieux en environnement intérieur, manquent de stabilité et de robustesse dans les environnements extérieurs non structurés.
- Le processus d'intégration n'est réalisé que sur le complément de l'espace libre dans l'image, ce qui permet d'éliminer la zone la moins porteuse d'information. Le sol est en effet souvent homogène et commun aux différents lieux (goudron dans un scénario urbain par exemple), et, de surcroît, occupe souvent une zone importante de l'image.
- L'étendue de l'espace libre est néanmoins implicitement prise en compte dans le processus de construction de la signature à travers une pondération utilisant les points frontières de l'espace libre.

Commençons par rappeler le formalisme des invariants intégraux de Haar,

$$H(\mathbf{x}) = \frac{1}{|G|} \int_G f(g\mathbf{x})dg; |G| = \int_G dg; g \in G \quad (1)$$

L'invariant H d'une configuration \mathbf{x} (ici l'image acquise à un instant donné), est défini par l'application d'une fonction noyau f (à support compact) intégrée sur l'ensemble

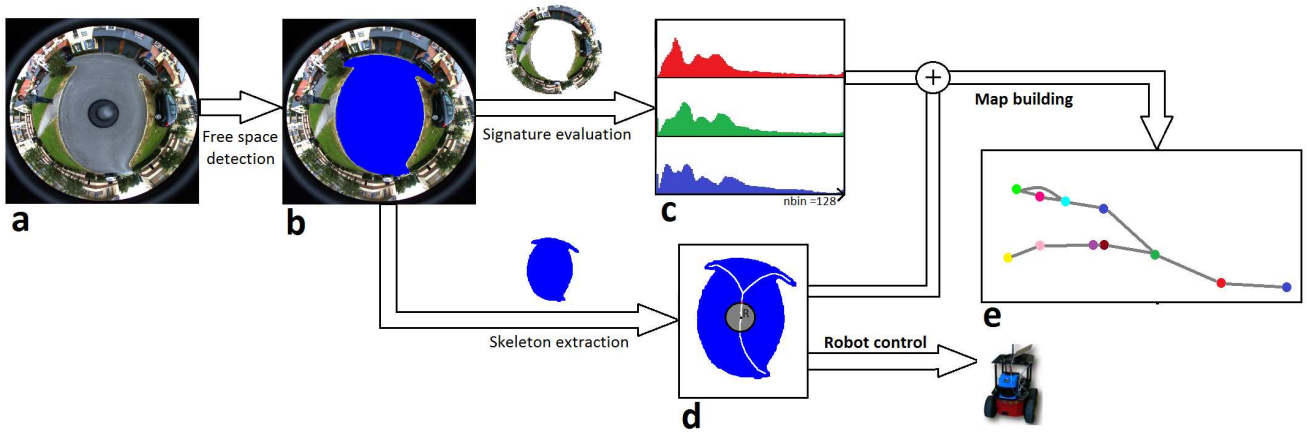


FIGURE 2 – Etapes de la construction de carte topologique. (a) Image omnidirectionnelle, (b) Extraction de l’espace libre, (c) Signature invariante construite depuis le complément de l’espace libre, (d) Squelette de l’espace libre, et (e) Fusion squelette / signature pour l’obtention de la carte topologique.

des transformations g d’un groupe G sur l’espace des configurations. L’invariance de H est directement liée à G , qui constitue dans ce travail l’ensemble des déplacements possibles du robot (translations et rotations), dans un monde supposé plan.

Une configuration \mathbf{x} est obtenue en dépliant l’image omnidirectionnelle sur le plan $\Pi(\theta, \varphi)$ calculé à partir de la projection inverse déduite du modèle de la sphère équivalente [18]. La région correspondante à l’espace libre est ensuite soustraite du domaine de \mathbf{x} afin de calculer l’invariant uniquement sur son complément $\Omega_{\overline{FS}} = \Pi_{\theta, \varphi} \setminus \Omega_{FS}$ (Fig. 3).

L’invariant de Haar s’écrit donc :

$$\begin{aligned}
 H(\mathbf{x}) &= \frac{1}{|G_{\theta, \varphi}|} \int_{\Omega_{\overline{FS}}} f(\mathbf{x}(\theta, \varphi)) \sin \theta d\theta d\varphi \\
 &= \frac{1}{|G_{\varphi}|} \int_{\varphi=0}^{2\pi} \left[\frac{1}{|G_{\theta}|} \int_{\theta=\pi/2}^{\theta_0(\varphi)} f(\mathbf{x}(\theta, \varphi)) \sin \theta d\theta \right] d\varphi \quad (2)
 \end{aligned}$$

avec φ variant dans $[0; 2\pi[$ (champ de vue de l’image omnidirectionnelle), et θ dans $[\pi/2; \theta_0(\varphi)]$ où $\theta_0(\varphi)$ est la limite supérieure de θ qui dépend de l’étendue de l’espace libre dans la direction correspondante à φ ($\theta_0(\varphi) \rightarrow \pi$ si $\Omega_{\overline{FS}}$ couvre l’image entière). $|G_{\varphi}| = \int_0^{2\pi} d\varphi$ et $|G_{\theta}| = \int_{\pi/2}^{\theta_0(\varphi)} \sin \theta d\theta$.

Enfin, l’invariant de Haar nécessite de définir une fonction noyau f . Bien que théoriquement on puisse utiliser n’importe quelle fonction intégrable, le choix de f peut s’avérer critique dans la pratique. Dans [14], un noyau basé sur une différence de gaussiennes est proposé et donne des résultats convaincants en environnement intérieur. Néanmoins, en extérieur, les scènes manquent souvent d’éléments présentant des contours structurés, ce qui réduit le taux de reconnaissance de la méthode.

Pour palier ce problème, nous proposons un noyau $f_{\theta, \varphi}$ qui combine une composante monomiale photométrique et une composante relationnelle basée sur les distances :

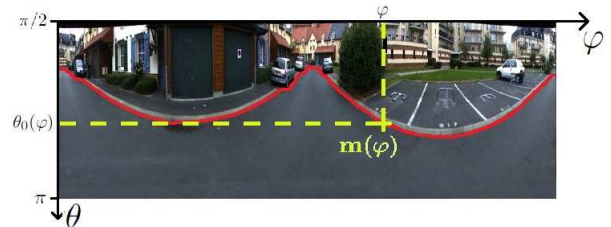


FIGURE 3 – Représentation du plan $\Pi(\theta, \varphi) = \Omega_{\overline{FS}} \cup \Omega_{FS}$. En rouge, les frontières de l’espace libre séparent l’espace navigable Ω_{FS} (en dessous) et son complément $\Omega_{\overline{FS}}$ (au dessus). Elles constituent le lieu de projection des points $\mathbf{m}(\varphi) = \mathbf{x}(\theta_0(\varphi), \varphi)$, utilisés dans le calcul de $f_r(\mathbf{x}(\theta, \varphi))$.

- Le noyau monomial f_m code la structure photométrique locale des pixels et est défini par :

$$f_m(\mathbf{x}(\theta, \varphi)) = \sum_j \sqrt{\prod \mathbf{x}^j(\theta', \varphi')} \quad (3)$$

où $\theta' = \theta + kd\theta$ et $\varphi' = \varphi + ld\varphi$ avec k, l, j entiers, et \sum_j l’ordre du noyau (on utilise ici un noyau d’ordre 2 généré par $(k=0, l=1)$ et $(k=2, l=0)$, ce qui donne : $f_m(\mathbf{x}(\theta, \varphi)) = \sqrt{\mathbf{x}(\theta, \varphi + d\varphi) \cdot \mathbf{x}(\theta + 2d\theta, \varphi)}$.

- Le noyau relationnel f_r ajoute quant à lui une information de distance entre le pixel et le robot, basée sur les contours de l’espace libre :

$$f_r(\mathbf{x}(\theta, \varphi)) = \frac{d_2(\mathbf{m}(\varphi), \mathbf{m}(\varphi - d\varphi)) + d_2(\mathbf{m}(\varphi), \mathbf{m}(\varphi + d\varphi))}{2} \quad (4)$$

où d_2 correspond à la distance Euclidienne et $\mathbf{m}(\varphi)$ sont les coordonnées 2D métriques des points frontières de l’espace libre $\mathbf{x}(\theta_0, \varphi)$. Ces points (d’altitude zéro) sont reconstruits à partir des équations de reprojection du capteur catadioptrique (calibré).

Le noyau final $f_{\theta, \varphi} = f_r(\mathbf{x}(\theta, \varphi)) \cdot f_m(\mathbf{x}(\theta, \varphi))$ est appliqué séparément sur chacun des 3 canaux RGB de l’image

couleur, afin d'augmenter la discriminance de la signature obtenue. Pour chaque canal $c = r, g, b$, on obtient donc un invariant exprimé dans le domaine discret par :

$$H_c(\mathbf{x}) = \frac{1}{|G_\varphi|} \sum_{\varphi} \left[\frac{1}{|G_\theta|} \sum_{\theta} f_{\theta, \varphi}(\mathbf{x}(\theta, \varphi)) \sin \theta d\theta \right] d\varphi \quad (5)$$

Pour améliorer la discriminance de la signature, on remplace l'intégration sur φ ($\int_{\varphi} [\dots] d\varphi / |G_\varphi|$) par un histogramme normalisé (128 bins). Les différentes étapes nécessaires à l'obtention de la signature de lieu sont résumées dans Algo. (1).

Algorithm 1 Construction de la signature.

Entree : Image omni, Frontières de l'espace libre $\partial\Omega$ (Fig. 2-(b))
Sortie : Signature de Haar $H(\mathbf{x})$

```

Calcul de  $\Pi_{\theta, \varphi}$  et  $\mathbf{x}$  avec la résolution désirée ;
Projection des limites de l'espace libre  $\partial\Omega$  dans  $\Pi_{\theta, \varphi}$  ;
Pour chaque  $c$  dans  $r, g, b$ 
  Pour chaque  $\varphi$ 
     $sum_c = 0$  (initialisation)
    Pour chaque  $\theta$ 
      Si  $\mathbf{x}(\theta, \varphi) \in \Omega_{\overline{FS}}$  (complément de l'espace libre)
         $sum_c = sum_c + f_{\theta, \varphi}(\mathbf{x}(\theta, \varphi)) \cdot \sin \theta d\theta$ ; (Eq. 5)
      finSi
    finPour
    Ajout de  $sum_c$  dans l'histogramme  $Hist_c$ ;
  finPour
Normalisation de  $Hist_c$ ;
finPour
return  $H(\mathbf{x}) = \{Hist_c\}_{c=r,g,b}$  (Fig.2-(c)).

```

On utilise la mesure basée sur l'intersection d'histogrammes (Eq. (6)) pour évaluer la similarité entre deux signatures. D'après [19], cette mesure possède de bonnes propriétés pour la classification d'images couleur.

$$d_{\cap}(H, K) = \frac{1}{3} \sum_{c=r,g,b} \frac{\sum_i \min(h_{c,i}; k_{c,i})}{\sum_i k_{c,i}} \quad (6)$$

avec $H = \{h_i\}_{i=1, \dots, n}$ et $K = \{k_i\}_{i=1, \dots, n}$ les histogrammes des deux signatures comparées, n le nombre de bins de chaque histogramme. Etant donné que les signatures sont normalisées, on a $\sum_i k_{c,i} = 1$. A noter que deux signatures totalement identiques vont présenter une similarité de 1.0, et que plus la différence augmentera, plus le score obtenu sera faible.

5 Exploration autonome et construction de carte topologique

A l'état initial, le robot commence l'exploration de l'espace libre. Il se déplace de manière autonome en suivant les chemins disponibles extraits automatiquement¹ depuis le squelette de l'espace libre. La stratégie d'exploration est basée sur la sélection d'une des branches du squelette extrait (Figure 2-(d)). En priorité, le robot suivra la branche

¹. Le processus d'extraction du squelette est actuellement en cours de soumission. On le considère donc acquis dans ce travail.

la plus proche de sa trajectoire actuelle. Lorsqu'il atteint une intersection, la branche la plus à gauche dans l'image est sélectionnée par défaut (l'image étant inversée, le robot se déplacera donc vers la droite). Les branches non visitées sont enregistrées en tant que points d'intérêt topologiques, mais, à l'heure actuelle, aucune stratégie n'est définie pour revenir les explorer plus tard (c'est l'une des orientations de nos travaux en cours).

Dans le même temps, une signature correspondant à la position initiale est construite $H_{t=0}$, et un premier noeud topologique n_0 est ajouté à la carte. Au cours de l'exploration, de nouvelles signatures sont créées et permettent au robot de se localiser et de définir de nouveaux lieux (voir Algo. 2). Un lieu est défini autour d'un noeud. Pour changer de lieu, la différence de la signature courante avec la signature du noeud doit être inférieure à un seuil τ_1 ($\tau_1 = 0.9$ dans ce cas d'étude). Dans le cas où la mesure de similarité est inférieure à ce seuil, le robot élargit la recherche à l'ensemble des noeuds de la carte et trouve l'appariement le plus proche. Si la similarité est suffisamment marquée (plus grande qu'un second seuil $\tau_2 = 0.8$), le robot considère qu'il a atteint un lieu déjà visité (fermeture de boucle). Dans le cas contraire, il crée un nouveau noeud qui intègre la nouvelle signature et qui sera associé à un lieu pour l'instant inconnu. Cette stratégie à deux seuils offre une certaine robustesse au processus de construction de la carte : avec le premier seuil, on détecte un potentiel déplacement dans la carte topologique, tandis qu'avec le second, on décide si le nouveau noeud atteint existe déjà dans la carte (fermeture de boucle) ou bien s'il s'agit d'un nouveau lieu. Dans tous les cas, à chaque fois que le robot change de position (de noeud), un lien reliant le dernier noeud visité et celui qui vient d'être atteint est créé.

Algorithm 2 Construction et mise à jour de la carte

Entree : Signature H_t du noeud courant n_i , Signature courante H_{t+1} , Map \mathcal{M}_t
Sortie : Nouvelle map \mathcal{M}_{t+1} (mise à jour) avec noeud courant

```

CurrentNode  $\leftarrow n_i$ ;
Si  $d_{\cap}(H_t, H_{t+1}) < \tau_1$ 
  Trouver  $n^* = \arg \max_{n_k \in \mathcal{M}(t)} d_{\cap}(H_k, H_{t+1})$ ; ( $H^*$  sign. de  $n^*$ )
Si  $d_{\cap}(H^*, H_{t+1}) \geq \tau_2$ 
  Ajouter connexion  $n_i \mapsto n^*$ ;
  CurrentNode  $\leftarrow n^*$ ;
Sinon
  Créer noeud  $n$  défini par la signature  $H_{t+1}$ ;
  Ajouter connexion  $n_i \mapsto n$ ;
  CurrentNode  $\leftarrow n$ ;
finSi
 $\mathcal{M}_{t+1} = \mathcal{M}_t \cup \{CurrentNode\}$ ;
Sinon
 $\mathcal{M}_{t+1} = \mathcal{M}_t$ ;
finSi
return  $\mathcal{M}_{t+1}, CurrentNode$ 

```

L'algorithme global d'exploration et de construction et mise à jour de carte topologique est résumé dans Algo. 3. L'ensemble du processus, de l'acquisition de l'image jus-

qu'à la commande des roues est réalisé de manière totalement autonome par le robot. Ce dernier s'arrête lorsqu'un temps maximal d'exploration (prédéfini) est atteint, ou si la carte n'évolue plus pendant une période de temps fixée (exploration supposée complète).

Algorithm 3 Vue sommaire de l'algorithme.

Faire

- Acquisition de l'image omnidirectionnelle ;
- 1. Détection de l'espace libre dans l'image (Algorithme [17]) ;
- 2. Calcul de la signature de lieu (Algorithme 1) ;
- 3. Création / Mise à jour de la carte topologique (Algorithme 2) ;
- 4. Déplacement du robot le long du squelette de l'espace libre

Tant que *NotEnd*

6 Validation et expérimentations en conditions réelles

Toutes les expériences sont menées sur une plateforme Pioneer 3-AT (voir Fig. 1) équipée d'un capteur de vision omnidirectionnelle paracatadioptrique couleur. L'ensemble des traitements est réalisé en ligne directement sur le robot par un processeur Intel I7-2620m, qui est capable de traiter, pour chaque image (de résolution 800×600), l'ensemble des opérations (calcul de l'espace libre, extraction du squelette, construction de la signature, mise à jour de la carte et commande du robot), en moins de 200ms (sans aucune parallélisation des calculs).

Plusieurs expériences ont été menées (en intérieur comme en extérieur) pour tester la capacité de nos algorithmes à explorer et à cartographier de manière autonome différents types d'environnements inconnus. Les scénarios proposés vont d'un simple couloir en intérieur jusqu'à plusieurs centaines de mètres dans différentes situations en extérieur. Dans ce papier, nous présentons l'étude d'un scénario en extérieur sur un peu plus de 350m présentant des lieux différents avec ambiguïté perceptive importante (voir fig.4).

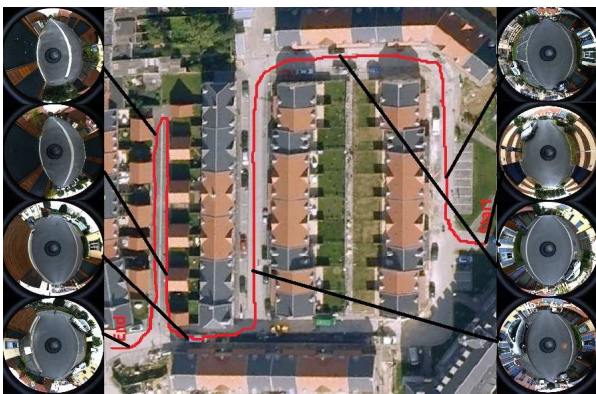


FIGURE 4 – Vue Google Earth de l'environnement de test, avec un ensemble de vues présentées à gauche et à droite. En rouge, La trajectoire que le robot a suivie sans aucune intervention humaine.

La figure 5 présente la carte topologique construite par le robot lors de l'exploration autonome. Malgré les fluctuations qui surviennent dans la segmentation des espaces

libres instantanés, l'extraction du squelette (construite pour être robuste au bruit), permet de guider le robot de manière sûre tout au long de l'exploration. De plus, le squelette porte les propriétés géométriques et topologiques de l'environnement. Le robot empruntant deux fois le même chemin suivra donc une trajectoire très similaire (sous réserve que la topologie de l'environnement ne change pas trop). Cette propriété ajoute une stabilité et une robustesse supplémentaire à la construction de signatures et donc à la localisation du robot dans la carte topologique.

Sur un parcours de plus de 350m, un total de 50 noeuds avec des liens cohérents composent la carte (voir Fig. 5). La fermeture de boucle (passage par des noeuds déjà visités) à la fin de la séquence est correctement détectée, et la topologie globale de la carte est conservée. Ces résultats sont obtenus avec les valeurs de seuil $\tau_1 = 0.9$ et $\tau_2 = 0.8$. Comme attendu, on peut observer que la répartition des noeuds sur la carte n'est pas homogène, ce qui est spécifique des représentations topologiques : tant que la perception visuelle ne change pas (ou peu), il n'est pas nécessaire d'ajouter de nouveaux noeuds. Dans sa représentation topologique, le robot reste au même lieu. Ainsi, il peut se déplacer de plus de 20m avant de considérer que les informations visuelles ont suffisamment changé (il crée alors un nouveau noeud). A l'inverse, dans certaines situations (carrefour par exemple), les signatures risquent de changer très rapidement (par exemple lorsqu'un objet caché par un obstacle apparaît d'un coup dans l'image). Il peut ainsi y avoir plusieurs noeuds construits sur une distance réduite (ce phénomène est bien visible Fig. 5 sur la partie en bas à gauche, où la disposition des bâtiments fait que la vision du robot change très rapidement).



FIGURE 5 – Carte topologique construite par le robot, superposée à la vue Google Earth. En rouge, la position géographique des signatures ayant amenées à la création d'un nouveau noeud. En noir, les liens existant entre les noeuds.

Afin de valider la formulation des signatures comme représentation pertinente et empreinte de lieu, nous la com-

parons à deux variantes du processus de construction de signature. On s'intéresse aux effets sur la carte construite lorsque *i*) le noyau relationnel est ignoré (seule l'information photométrique est prise en compte dans Eq. (3)-(5)), et *ii*) lorsque toute l'image est prise en compte au lieu de restreindre le calcul d'invariant au complément de l'espace libre.

Les graphiques Fig. 6 présentent pour chaque signature le noeud auquel elle a été associée (c-à-d, la position du robot dans la carte topologique à chaque image acquise). Commençons par comparer la signature proposée dans ce papier (Fig. 6-(a)) avec celle utilisant uniquement le noyau monomial sur $\Omega_{\overline{FS}}$ (Fig. 6-(b)). L'aspect global des résultats est similaire, mais deux différences majeures sont observables. Pour le même environnement, 82 noeuds sont nécessaires pour construire la carte topologique en utilisant le seul noyau monomial, contre seulement 50 avec la nouvelle formulation combinant noyau monomial et noyau relationnel. Cela s'explique par le fait que, pour les objets proches du robot (voitures, murs, ...), un faible déplacement de ce dernier engendre des variations perceptives importantes de ces objets. C'est exactement pour compenser ce phénomène que le noyau relationnel a été introduit.

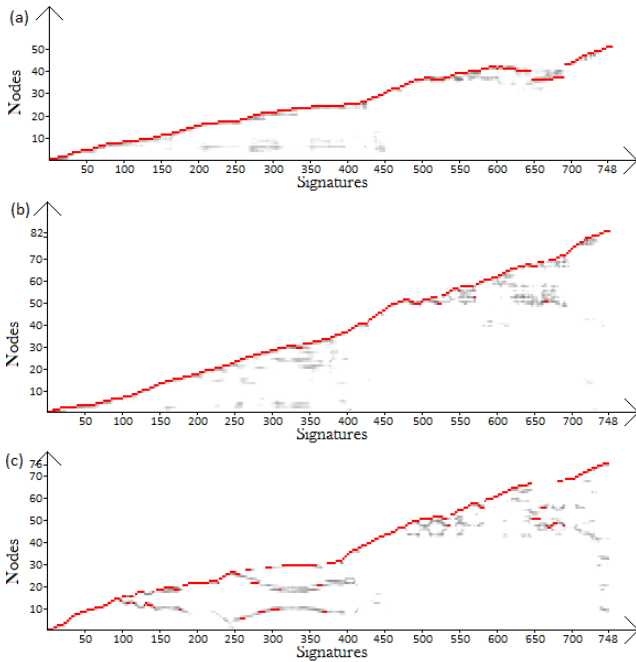


FIGURE 6 – Graphe de correspondances entre noeuds et signatures. En rouge la position déduite de chaque signature pour : (a) L'approche proposée combinant noyaux monomial et relationnel, et intégrant sur $\Omega_{\overline{FS}}$; (b) noyau monomial seul sur $\Omega_{\overline{FS}}$; et (c) noyau monomial sur l'ensemble de l'image ($\Omega_{FS} \cup \Omega_{\overline{FS}}$). En gris, on présente également les noeuds dont la ressemblance est inférieure au noeud choisi, mais tout de même suffisamment élevée pour être notifiée (plus le pixel est sombre, plus la similarité est proche de celle du noeud choisi).

La deuxième différence est l'incapacité du noyau monomial à localiser la fermeture de boucle dans certaines si-

tuations (voir fin de parcours Fig. 5). Dans cet exemple, la ligne continue de séparation des voies navigables délimite naturellement les frontières de l'espace libre dans l'image, le robot emprunte alors des chemins différents à l'aller et au retour sur ce parcours. Les variations de perception visuelle sur chacun des chemins conduisent à des signatures (basées noyau monomial) différentes. Ainsi, des nouveaux noeuds sont créés dans la carte et ne permettant pas l'identification des lieux visités. Le noyau relationnel en revanche permet de compenser cette différence, et notre méthode identifie correctement que le robot est déjà passé par là.

Les résultats pour les signatures calculées par intégration sur l'ensemble de l'image sont donnés Fig. 6-(c). Les performances chutent de manière significative. Les liens entre les noeuds ne sont pas fidèles à la réalité. Non seulement la fermeture de boucle n'est pas détectée, mais en plus de fausses fermetures de boucle le sont. Même en faisant varier les valeurs des seuils, il n'a pas été possible d'atteindre un résultat de cartographie cohérent avec la réalité terrain sur l'ensemble du parcours. Avec un capteur positionné à 1.35m au dessus du sol, l'espace libre Ω_{FS} occupe environ 40% de l'image. De ce fait, de petites variations dans la texture ou la taille de l'espace libre domine les changements qui peuvent survenir dans la scène (apparition d'un nouvel obstacle par exemple), et donc induit de larges variations et instabilités dans la signature.

7 Conclusion

Dans ce papier, nous avons présenté une nouvelle approche pour l'exploration autonome et la construction de carte topologique en environnement initialement inconnu. Une carte cognitive est construite et enrichie incrémentalement pendant l'exploration. L'approche implique une intégration temps réel de différentes représentations spatiales conduisant à une exploration sûre en évitant les obstacles. Basée sur un algorithme efficace de segmentation de l'espace libre dans l'image, elle (*i*) offre une solution pour l'exploration autonome en déplaçant le robot sur le squelette de l'environnement local; et (*ii*) propose une procédure d'indexation des lieux explorés par des signatures invariantes construites en sélectionnant uniquement l'information pertinente dans l'image. Le squelette constitue le support des graphes topologiques de la carte cognitive et les signatures sont des empreintes des lieux (noeuds) sur la carte.

A travers l'étude d'un scénario en conditions réelles, nous avons validé à la fois l'exploration autonome et la construction de carte topologique dans un large environnement extérieur. Nous travaillons maintenant sur une stratégie d'exploration à plus haut niveau qui utilise les branches non explorées du squelette pour définir des objectifs précis et influencer directement sur la prise de décision lors de la détection de jonctions topologiques complexes.

Remerciements : Ce travail est financé par l'ANR (projet R-Discover), la Direction Générale de l'Armement (DGA) et le Fond Européen de Développement Régional en Picardie (FE-

DER).

Références

- [1] S. Se, D. Lowe, and J. Little, "Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks," *IJRR*, vol. 21, pp. 735–758, 2002.
- [2] R. Sim, P. Elinas, and J. J. Little, "A study of the rao-blackwellised particle filter for efficient and accurate vision-based slam," *IJCV*, vol. 74, no. 3, pp. 303–318, 2007.
- [3] C. Valgren, A. J. Lilienthal, and T. Duckett, "Incremental topological mapping using omnidirectional vision," in *IROS*, 2006, pp. 3441–3447.
- [4] A. C. Murillo and J. Kosecka, "Experiments in place recognition using gist panoramas," in *IEEE Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras, held with ICCV*, 2009, pp. 2196–2203.
- [5] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping (slam) : Part i," *Robotics Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.
- [6] N. Muhammad, D. Fofi, and S. Ainouz, "Current state of the art of vision based slam," in *SPIE Image Processing : Machine Vision Applications II*, 2009, pp. 72 510F–72 510F–12.
- [7] E. Eade and T. Drummond, "Scalable monocular slam," in *CVPR*, 2006, pp. 469–476.
- [8] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "Monoslam : Real-time single camera slam," *PAMI Transactions*, vol. 29, no. 6, pp. 1052–1067, 2007.
- [9] K. Konolige and M. Agrawal, "Frameslam : From bundle adjustment to real-time visual mapping," *Transactions on Robotics*, vol. 24, no. 5, pp. 1066–1077, 2008.
- [10] M. Cummins and P. Newman, "Appearance-only slam at large scale with fab-map 2.0," *IJRR*, vol. 30, no. 9, pp. 1100–1123, 2011.
- [11] M. Milford and G. Wyeth, "Persistent navigation and mapping using a biologically inspired slam system," *International Journal of Robotics Research*, vol. 29, no. 9, pp. 1131–1153, 2010.
- [12] A. Angeli, S. Doncieux, J.-A. Meyer, and D. Filliat, "Visual topological slam and global localization," in *IEEE ICRA*, 2009, pp. 2029–2034.
- [13] F. Dayoub, G. Cielniak, and T. Duckett, "Long-term experiments with an adaptive spherical view representation for navigation in changing environments," *Robotics and Autonomous Systems*, vol. 59, no. 5, pp. 285–295, 2011.
- [14] O. Labbani-Igbida, C. Charron, and M. Mouaddib, "Haar invariant signatures and spatial recognition using omnidirectional visual information only," *Autonomous Robots Journal*, vol. 30, no. 3, pp. 333–349, 2011.
- [15] B. Kuipers, J. Modayil, P. Beeson, M. MacMahon, and F. Savelli, "Local metrical and global topological maps in the hybrid spatial semantic hierarchy," in *ICRA*, 2004, pp. 4845–4851.
- [16] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (slam) : Toward exact localization without explicit localization," *Transactions on Robotics and Automation*, vol. 17, no. 2, pp. 125–137, 2001.
- [17] P. Merveilleux, O. Labbani-Igbida, and M. Mouaddib, "Robust free space segmentation using active contours and monocular omnidirectional vision," in *ICIP*, 2011, pp. 2877–2880.
- [18] C. Geyer and K. Daniilidis, "Catadioptric projective geometry," *IJCV*, vol. 45, no. 3, pp. 223–243, 2001.
- [19] A. Barla, F. Odone, and A. Verri, "Histogram intersection kernel for image classification," in *ICIP*, 2003, pp. 513–516.