UNIVERSITÉ DE MONTRÉAL

ASSERVISSEMENT D'UN BRAS ROBOTIQUE D'ASSISTANCE À L'AIDE D'UN SYSTÈME DE STÉRÉO VISION ARTIFICIELLE ET D'UN SUIVEUR DE REGARD

YANN-SEING LAW-KAM CIO

DÉPARTEMENT DE MÉCANIQUE

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE MÉCANIQUE)

AOÛT 2017

© Yann-Seing Law-Kam Cio, 2017.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ASSERVISSEMENT D'UN BRAS ROBOTIQUE D'ASSISTANCE À L'AIDE D'UN SYSTÈME DE STÉRÉO VISION ARTIFICIELLE ET D'UN SUIVEUR DE REGARD

présenté par : LAW-KAM CIO, Yann-Seing

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. BALAZINSKI Marek, Doctorat ès sciences, président

M. ACHICHE Sofiane, Ph. D., membre et directeur de recherche

M. RAISON Maxime, Ph. D., membre et codirecteur de recherche

M. MOHEBBI Abolfazl, Ph. D., membre

DÉDICACE

Je dédie le travail de ce mémoire à ma mère, mon père et mon frère.

REMERCIEMENTS

Je voudrais, tout d'abord, remercier mes directeurs pour toutes les heures consacrées au projet de maîtrise afin de rendre cette expérience des plus plaisantes. Un gros merci de m'avoir fait confiance, de m'avoir bien guidé et merci pour votre disponibilité, et ce, tout au long de ma maîtrise.

J'aimerais remercier mon directeur de maîtrise, le professeur Sofiane Achiche de m'avoir fait découvrir les plaisirs de l'enseignement par l'entremise de charges de travaux dirigés et laboratoires. J'aimerais aussi souligner la générosité du professeur Sofiane Achiche qui m'a appuyé au-delà du projet de recherche ainsi que son support plaisant et agréable qu'il apporte au laboratoire.

Mon codirecteur, le professeur Maxime Raison titulaire de la Chaire de recherche en génie de la réadaptation pédiatrique m'a mis sur la voie pour faire une maîtrise recherche dans le domaine des technologies d'assistance. En m'ayant pris pour un projet d'initiation à la recherche, le professeur Maxime Raison m'a permis de découvrir les joies de la recherche et j'en suis très reconnaissant.

J'aimerais souligner l'accueil chaleureux et l'assistance technique de toute l'équipe de Kinova Robotics. Un gros merci pour ma superviseure industrielle Martine Blouin, chargée de projet en contrôle robotique dans la section recherche et développement, pour m'avoir conseillé et aidé tout au long du projet de maîtrise.

Je voudrais aussi mentionner ma reconnaissance envers les organismes du CRSNG, FQRNT et de Kinova Robotics pour leur soutien financier par l'entremise d'une bourse en milieu pratique BMP innovation. Je remercie aussi Polytechnique Montréal pour leur financement ainsi que leurs locaux de travail.

Je voudrais remercier ma mère, mon père ainsi que mon frère pour tout ce qu'ils ont fait afin que je puisse me rendre où j'en suis aujourd'hui. J'aimerais aussi remercier Sui-Meing Wong, grand ami de la famille, pour son support moral et ses bons conseils avant ou après les entraînements.

L'atmosphère de travail a été des plus agréables grâce à mes collègues du Laboratoire de Conception de Systèmes Intelligents et Mécatroniques (CoSIM) du Prof. Sofiane Achiche. Une

mention particulière à mon collègue Cédric Leblond Ménard est de mise pour son aide technique et sa motivation pour le projet de recherche. Je voudrais aussi remercier mes amis de m'avoir appuyé dans les moments plus difficiles.

Je voudrais remercier le professeur Marek Balazinski d'avoir accepté d'être le président du jury de ce mémoire. Je voudrais aussi remercier Dr. Abolfazl Mohebbi d'agir en tant que membre du jury.

RÉSUMÉ

L'utilisation récente de bras robotiques sériels dans le but d'assister des personnes ayant des problèmes de motricités sévères des membres supérieurs soulève une nouvelle problématique au niveau de l'interaction humain-machine (IHM). En effet, jusqu'à maintenant le « joystick » est utilisé pour contrôler un bras robotiques d'assistance (BRA). Pour les utilisateurs ayant des problèmes de motricité sévères des membres supérieurs, ce type de contrôle n'est pas une option adéquate. Ce mémoire présente une autre option afin de pallier cette problématique.

La solution présentée est composée de deux composantes principales. La première est une caméra de stéréo vision utilisée afin d'informer le BRA des objets présents dans son espace de travail. Il est important qu'un BRA soit conscient de ce qui est présent dans son espace de travail puisqu'il doit être en mesure d'éviter les objets non voulus lorsqu'il parcourt un trajet afin d'atteindre l'objet d'intérêt pour l'utilisateur.

La deuxième composante est l'IHM qui est dans ce travail représentée par un suiveur de regard à bas coût. Effectivement, le suiveur de regard a été choisi puisque, généralement, les yeux d'un patient ayant des problèmes sévères de motricités au niveau des membres supérieurs restent toujours fonctionnels. Le suiveur de regard est généralement utilisé avec un écran pour des applications en 2D ce qui n'est pas intuitif pour l'utilisateur puisque celui-ci doit constamment regarder une reproduction 2D de la scène sur un écran. En d'autres mots, il faut rendre le suiveur de regard viable dans un environnement 3D sans l'utilisation d'un écran, ce qui a été fait dans ce mémoire.

Un système de stéréo vision, un suiveur de regard ainsi qu'un BRA sont les composantes principales du système présenté qui se nomme PoGARA qui est une abréviation pour **P**oint **o**f **G**aze **A**ssistive **R**obotic **A**rm. En utilisant PoGARA, l'utilisateur a été capable d'atteindre et de prendre un objet pour 80% des essais avec un temps moyen de 13.7 secondes sans obstacles, 15.3 secondes avec un obstacle et 16.3 secondes avec deux obstacles.

ABSTRACT

The recent increased interest in the use of serial robots to assist individuals with severe upper limb disability brought-up an important issue which is the design of the right human computer interaction (HCI). Indeed, so far, the control of assistive robotic arms (ARA) is often done using a joystick. For the users who have a severe upper limb disability, this type of control is not a suitable option. In this master's thesis, a novel solution is presented to overcome this issue.

The developed solution is composed of two main components. The first one is a stereo vision system which is used to inform the ARA of the content of its workspace. It is important for the ARA to be aware of what is present in its workspace since it needs to avoid the unwanted objects while it is on its way to grasp the object of interest.

The second component is the actual HCI, where an eye tracker is used. Indeed, the eye tracker was chosen since the eyes, often, remain functional even for patients with severe upper limb disability. However, usually, low-cost, commercially available eye trackers are mainly designed for 2D applications with a screen which is not intuitive for the user since he needs to constantly watch a reproduction of the scene on a 2D screen instead of the 3D scene itself. In other words, the eye tracker needs to be made viable for usage in a 3D environment without the use of a screen. This was achieved in this master thesis work.

A stereo vision system, an eye tracker as well as an ARA are the main components of the developed system named PoGARA which is short for **P**oint **o**f **G**aze **A**ssistive **R**obotic **A**rm. Using PoGARA, during the tests, the user was able to reach and grasp an object for 80% of the trials with an average time of 13.7 seconds without obstacles, 15.3 seconds with one obstacles and 16.3 seconds with two obstacles.

TABLE DES MATIÈRES

DÉDICACE III

REMERCIEMENTS IV

RÉSUMÉ VI

ABSTRACT VII

TABLE DES MATIÈRES VIII

LISTE DES TABLEAUX X

LISTE DES FIGURES XI

CHAPITRE 1 INTRODUCTION 1

1.1 Mise en contexte 1

CHAPITRE 2 REVUE DE LA LITTÉRATURE 3

2.1 Robot d'assistance 3

2.2 Vision par ordinateur 3

2.2.1 Capteurs visuels 3

2.2.2 Planification de trajectoire et évitement d'obstacle 7

2.2.3 Interface utilisateur 19

CHAPITRE 3 OBJECTIFS, QUESTIONS DE RECHERCHE ET HYPOTHÈSES 21

3.1 Présentation de l'objectif principal et des objectifs spécifiques 21

3.2 Question de recherche 21

3.3 Hypothèses 22

CHAPITRE 4 MÉTHODOLOGIE 23

4.1 Cahier des charges 23

4.1.1 Besoins essentiels 23

4.1.2 Besoins facultatifs 23

4.1.3 Contraintes 23

4.2 Robot d'assistance 24

4.3 Vision par ordinateur 24

4.3.1 Capteur visuel 24

4.3.2 Planification de trajectoire et évitement d'obstacle 25

4.3.3 Interface humain-machine 25

CHAPITRE 5 ARTICLE 1: ASSISTIVE ROBOTIC ARM CONTROL USING ARTIFICIAL STEREOVISION AND EYE-TRACKING 27

5.1 Introduction 28

5.2 Literature review on vision for robotic control 29

5.3 System design description 32

5.3.1 Visual sensors 32

5.3.2 Path Planning and Obstacle Avoidance 33

5.3.3 User Interface 33

5.3.4 System Integration 34

5.4 Results & Discussion 39

5.5 Limitations and future work 43

5.6 Conclusion 44

5.7 References 45

CHAPITRE 6 DISCUSSION GÉNÉRALE 50

CHAPITRE 7 CONCLUSION ET RECOMMENDATIONS 52

RÉFÉRENCES 53

LISTE DES TABLEAUX

Tableau 2-1 Comparaison des différents bras robotiques d'assistance 3

Tableau 4-1 Tableau récapitulatif des différents capteurs visuels potentiels.....24

Table 5-1 Results of the tests with the object of interest alone shown in Figure 5-7 40

Table 5-2 Results of the tests with the object of interest and one obstacle shown in Figure 5-8 41

Table 5-3 Results of the tests with the object of interest and two obstacles shown in Figure 5-9 42

LISTE DES FIGURES

Figure 2-1 Exemple : Kinect 2(gauche) (Figure tirée de [9])et ses caractéristiques (droite) 4

Figure 2-2 Exemple : Banana Pi Camera Module(gauche)(Figure tiré de [13]) et ses caractéristiques (droite) 5

Figure 2-3 Exemple : Mic+600/D/TC de Microsonic(gauche) (Figure tirée de[15]) et ses caractéristiques(droite) 6

Figure 2-4 Exemple : Tara(droite) (Figure tirée de [17]) et ses caractéristiques (gauche) 7

Figure 2-5 Les trois versions du "Bug algorithm". Version 1(haut). Version 2(milieu). Version 3(bas). (Figure tirée de [19]) 9

Figure 2-6 Présentation du principe de l'algorithme APF 10

Figure 2-7 Situations menant à des minimums locaux. Obstacles symétriques (haut). Obstacles en forme de U (bas) (Figure tirée de [19]) 11

Figure 2-8 Étape 1(haut): détection d'obstacle. Étape 2(milieu et bas): Transformation des données en histogramme polaire. (Figure tirée de [19]) 12

Figure 2-9 Présentations du principe de la méthode "Follow the gap" 13

Figure 2-10 Présentation du principe de NHNA. (Figure tirée de [19]) 14

Figure 2-11 La couche réactive ("DH algorithm") s'adapte à des obstacles en mouvements. (Figure tirée de [19]) 15

Figure 2-12 Image initiale avec les pistes (gauche). Image présentant le chemin emprunté avec une ligne pointillée en gras (droite). (Figure tirée de [19]) 16

Figure 2-13 L'image 2D a été divisée en carrés. Ceux en gris sont libres et ceux en noir sont des obstacles. Le chemin est trouvé avec un algorithme de recherche. (Figure tirée de [21]) 17

Figure 2-14 Présentation du principe en "multi query" (gauche) et en "single query"(droite) (Figure tirée de [18] 18

Figure 4-1 Photo du bras robotique d'assistance JACO de Kinova Robotics (Figure tirée de [4]) 24

Figure 5-1 Tara stereovision camera (left). Features of Tara (right) 33

Figure 5-2 Presentation of PoGARA system. JACO is the ARA. Eye Tribe is the eye tracker and Tara is the stereovision camera 34

Figure 5-3 Image from the right (top-left) and left (bottom-left) camera and the disparity map (right) 35

Figure 5-4 Image of the scene. The blue cube is the position of the Tara. The green cube with the white line is the gaze of the user. The white lines with the red cube is the JACO arm 36

Figure 5-5 Method based on vector field histogram for collision detection 38

Figure 5-6 The path that the JACO's end effector would take 39

Figure 5-7 Position of the object of interest (left). Object of interest grasped by the JACO arm (right).....40

Figure 5-8 Position of the object of interest and the obstacle (left). Object of interest grasped by the JACO arm (right).....41

Figure 5-9 Position of the object of interest and the two obstacles (left). Object of interest grasp by the JACO arm (right).....42

CHAPITRE 1 INTRODUCTION

1.1 Mise en contexte

Depuis le premier bras robotique d'assistance (BRA), cette technologie a attiré l'intérêt autant en entreprise qu'en recherche. Les BRAs les plus connus sont le bras Manus prédécesseur du iARM (Exact dynamics, Didam, Pays-Bas) [1], WAM Arm (Barrett Technology, LCC, Newton, États-Unis) [2], DLR LWR-III (DLR Robotics and Mechatronics center, Weichselbaum, Allemagne) [3] and JACO and MICO (Kinova Robotics, Boisbriand, Canada) [4]. Néanmoins, l'un des principaux défis des bras d'assistance est leur grande contrainte de polyvalence afin d'aider le plus de personnes atteintes de problèmes variés de motricité des membres supérieurs; allant de la dystrophie musculaire à la tétraplégie. Le bras d'assistance devrait être capable de s'adapter à l'utilisateur afin qu'il puisse les aider à recouvrer leur autonomie et de leur qualité de vie. Afin de surmonter ce défi, une interface utilisateur doit être choisie et conçue en considérant les limites physiques de l'utilisateur afin de lui permettre de plus facilement contrôler le bras d'assistance. À ce jour, la technologie la plus utilisée pour contrôler les bras robotiques d'assistance est le « joystick ». Ceci peut être adéquat pour des utilisateurs ayant des problèmes de motricité telle que les premiers stades de la dystrophie musculaire. Par contre, pour les problèmes de motricité plus sévères comme le stade avancé de la dystrophie musculaire ou la spasticité involontaire, le « joystick » n'est pas une technologie adaptée.

Ce mémoire propose une solution alternative pour les utilisateurs avec des problèmes moteurs sévères en modifiant le contrôle du bras JACO. En effet, la nouveauté est l'utilisation du suiveur de regard comme technologie d'IHM dans le contexte du contrôle par vision d'un BRA.

Ce mémoire présente un système composé du BRA JACO, d'un système de stéréo vision, du planificateur de trajectoire A* avec évitement d'obstacles intégré et d'un suiveur de regard.

Tout d'abord, une brève revue de la littérature sur les différentes composantes du système de vision est présentée au Chapitre 2. Le Chapitre 3 présente les objectifs, la question de recherche et l'hypothèse du projet. En se basant sur les Chapitre 2 et Chapitre 3, le Chapitre 4 présente le cahier des charges du système à concevoir ainsi que la méthodologie suivie. L'article scientifique, soumis à une revue avec comité de lecture, est présenté dans le Chapitre 5. Le Chapitre 6 contient la

discussion générale ainsi que les résultats et les limitations du projet. Enfin, pour clore ce mémoire de maîtrise un résumé suivi d’une conclusion sont présentés au Chapitre 7.

CHAPITRE 2 REVUE DE LA LITTÉRATURE

2.1 Robot d'assistance

Le système à développer dans le cadre de ce travail de maîtrise a pour but d'assister les personnes avec des troubles moteurs au niveau des membres supérieurs. De ce fait le BRA devrait être facile à installer sur un fauteuil roulant ou sur une table. De plus, le BRA doit respecter les mesures de sécurité pour les bras d'assistance car il est en permanente et étroite collaboration avec l'utilisateur. Les bras industriels tels que le FANUC[5] ou le KUKA[6] ne sont pas adéquats pour usage dans le cadre de ce projet de maîtrise puisqu'ils sont trop volumineux et lourds. De plus, leur design n'a pas été basé autour de la problématique important qu'est la sécurité des utilisateurs. Sachant cela, les technologies des BRAs sont plus adaptées à la situation. En effet, les BRAs sont plus petits et légers. De plus, ils sont déjà utilisés, avec un joystick, pour assister des utilisateurs à effectuer des tâches quotidiennes. Par conséquent, la sécurité des utilisateurs est la préoccupation la plus importante dans leur design. Le Tableau 2-1 présente une comparaison des caractéristiques, importantes pour ce projet, des BRAs les plus populaires [1-4, 7, 8].

Tableau 2-1 Comparaison des différents bras robotiques d'assistance

Caractéristiques\Nom	iARM	JACO	MICO	WAM Arm	DLR LWR-III
Portée (mm)	900	900	700	1000	936
Poids (kg)	9	5.2	5.4	27	14
Consommation d'énergie(W)	<24	25	25	45	100

2.2 Vision par ordinateur

2.2.1 Capteurs visuels

Cette section présentera un résumé des différents capteurs visuels existants sur le marché. Le résumé de chaque capteur couvrira dix caractéristiques principales soit le développement, la précision, la vitesse, la robustesse, la calibration, la taille, le poids, la consommation d'énergie, la possibilité de fonctionner à l'extérieur et le prix. Ces caractéristiques ont été choisies après avoir

cherchées et analysées les feuilles de spécifications des différents capteurs visuels. Idéalement, le capteur choisi devrait respecter le plus possible les conditions suivantes:

1. Développement : le traitement des données brutes est minimal
2. Précision : positionner un point 3D en minimisant l'erreur.
3. Vitesse : les images filtrées devraient donner un vidéo en temps réel.
4. Robustesse : être capable de répéter une action donnée dans une situation donnée.
5. Calibration : le prétraitement nécessaire pour l'utilisation du capteur.
6. Taille : ces dimensions en largeur, longueur et hauteur du capteur.
7. Poids : le poids du capteur.
8. Consommation d'énergie : le nombre de Watt nécessaire ou la source externe nécessaire.
9. Fonctionne à l'extérieur : le capteur visuel devrait fonctionner même s'il y a des variations de lumière particulièrement la lumière du soleil.
10. Prix : il faut avoir le meilleur rapport qualité-prix.

2.2.1.1 Capteurs infrarouges et caméra

Caractéristiques	Valeurs
Dimensions(mm)	249X66X67
Masse(g)	1400
Prix ($ CAD)	135
Consommation d'énergie (W)	32
Résolution camera(pixels)	1920X1080 à 30FPS
Portée (m)	0,5-4,5

Figure 2-1 Exemple : Kinect 2(gauche) (Figure tirée de [9])et ses caractéristiques (droite)

Le capteur le plus utilisé afin d'avoir la meilleure précision à un prix abordable est la Kinect (Figure 2-1) qui utilise un projecteur infrarouge, un capteur infrarouge et une caméra couleur. Ce type de capteur excelle au niveau de la précision, la robustesse, le prix ainsi que le développement. Par contre, la consommation d'énergie, la taille et le poids sont très importants. De plus, les capteurs infrarouges peuvent être saturés par les rayons du soleil, donc, ce genre de capteur ne peut pas fonctionner à l'extérieur.[9-12]

2.2.1.2 Caméra simple

Caractéristiques	Valeurs
Dimensions(mm)	36X32X10
Masse(g)	5
Prix ($ CAD)	20
Consommation d'énergie (W)	0.219
Résolution camera(pixels)	640X480 à 30 FPS
Portée (m)	0.2-1.8

Figure 2-2 Exemple : Banana Pi Camera Module(gauche)(Figure tiré de [13]) et ses caractéristiques (droite)

Les meilleurs atouts d'une simple caméra RGB sont au niveau de son prix abordable, sa taille raisonnable, son poids relativement bas, sa faible consommation d'énergie et la possibilité de son utilisation à l'extérieur. Par contre, la contrainte principale de ce type de capteurs est qu'il donne en sortie des données dans un plan 2D. Cela signifie donc que la précision et la robustesse du système dépendent grandement de la qualité du traitement d'image. En effet, avec un capteur 2D, il est encore possible d'évaluer la profondeur d'un objet par une transformation homothétique. En d'autres mots, si l'objet paraît plus gros alors il est plus proche de la caméra et vice-versa. Une autre façon d'utiliser la transformation homothétique est d'utiliser une base de données où il serait

possible d'obtenir les dimensions de différents objets à une certaine distance. Avec ces informations, il est, alors, possible d'évaluer la profondeur de l'objet avec une simple règle de trois [14]. Un exemple de ce type de caméra est présenté à la Figure 2-2.

2.2.1.3 Capteurs ultrasons

Caractéristiques	Valeurs
Dimensions(mm)	65X65X105
Masse(g)	230
Prix ($ CAD)	287
Consommation d'énergie (W)	1.8-6
Résolution (mm)	0.025-2.4
Portée (m)	0.6-8

Figure 2-3 Exemple : Mic+600/D/TC de Microsonic(gauche) (Figure tirée de[15]) et ses caractéristiques(droite)

Même si ce genre de capteur est discret, léger et offre une précision et une consommation d'énergie acceptables. Leurs prix et la consommation d'énergie augmentent assez rapidement avec la précision. De plus, le traitement des données brutes nécessaire reste assez important, surtout si l'on considère que face à une surface lisse en angle, les ondes sonores vont plutôt se disperser au lieu de converger vers le capteur sonore. Ce phénomène peut être observé lorsque les ondes sonores rencontrent beaucoup d'objets, ce qui est le cas dans un environnement intérieur. [16]. La Figure 2-3 présente un capteur sonore ayant une bonne précision, mais dont le coût et la consommation d'énergie sont relativement importants.

2.2.1.4 Stéréovision

Caractéristiques	Valeurs
Dimensions (mm)	100X30X35
Masse(g)	80,5
Prix ($ CAD)	336
Consommation d'énergie (W)	2,27
Résolution camera(pixels)	1504X480 à 60 FPS
Portée (m)	0,5-3

Figure 2-4 Exemple : Tara(droite) (Figure tirée de [17]) et ses caractéristiques (gauche)

La stéréovision est un type de capteur qui dépend beaucoup du développement. En effet, la précision, la vitesse ainsi que la robustesse dépendent tous du traitement des données brutes de la stéréovision. Généralement, il faut faire des compris entre la précision ainsi que la robustesse et la vitesse. Pour avoir un système plus robuste et précis, il faut augmenter la quantité de traitement de données ce qui diminue la vitesse du système. La taille et le poids des capteurs sont généralement inférieurs à 200x50x50mm et 500 g respectivement, comme il possible de remarquer à la Figure 2-4 qui montre un exemple de stéréo vision, Tara. De plus, la consommation d'énergie est faible et, avec un traitement d'image, le capteur peut fonctionner à l'extérieur. Les deux inconvénients les plus importants sont le prix qui peut être relativement élevé et la calibration du système.[16]

2.2.2 Planification de trajectoire et évitement d'obstacle

La planification de trajectoire et l'évitement d'obstacle est une étape importante dans le développement d'un système de vision pour le contrôle d'un bras robotique. La revue de la littérature des algorithmes en 2D et 3D pour robot mobile permettra de trouver la meilleure méthode

à adapter pour le projet. Considérant que la plupart du temps les scènes obtenues par le système de vision auront beaucoup d'objets, l'algorithme doit être en mesure d'atteindre un objet sans toucher au restant des objets. Dans la section suivante, différents algorithmes fréquemment utilisés dans le domaine seront présentés :[18-20]

« Bug algorithms »: le principe de cette méthode est de diminuer la distance entre le point de départ et celui d'arrivée. Plusieurs versions de cet algorithme existent. La principale différence entre chaque version est la manière utilisée pour contourner un obstacle (Figure 2-5). La première version suit le périmètre complet de l'obstacle afin de trouver le côté le plus proche du point d'arrivée. À partir de ce côté, il s'éloigne de l'obstacle en prenant la direction menant vers le point d'arrivée. La deuxième version, quant à elle, calcule une direction initiale à partir des points de départ et d'arrivée. Lorsqu'un obstacle est rencontré, on suit le périmètre de l'objet jusqu'à ce que l'on retrouve la direction initiale. Face à un obstacle, la troisième version contourne l'objet en suivant les points qui se rapprochent du point d'arrivée.

Avantages	**Désavantages**
• Ne contiens pas de minimum local	• Évite un obstacle toujours de la même façon.
• Converge vers le but.	• N'emprunte pas le chemin optimal.
	• Risque d'éloigner le robot du but.

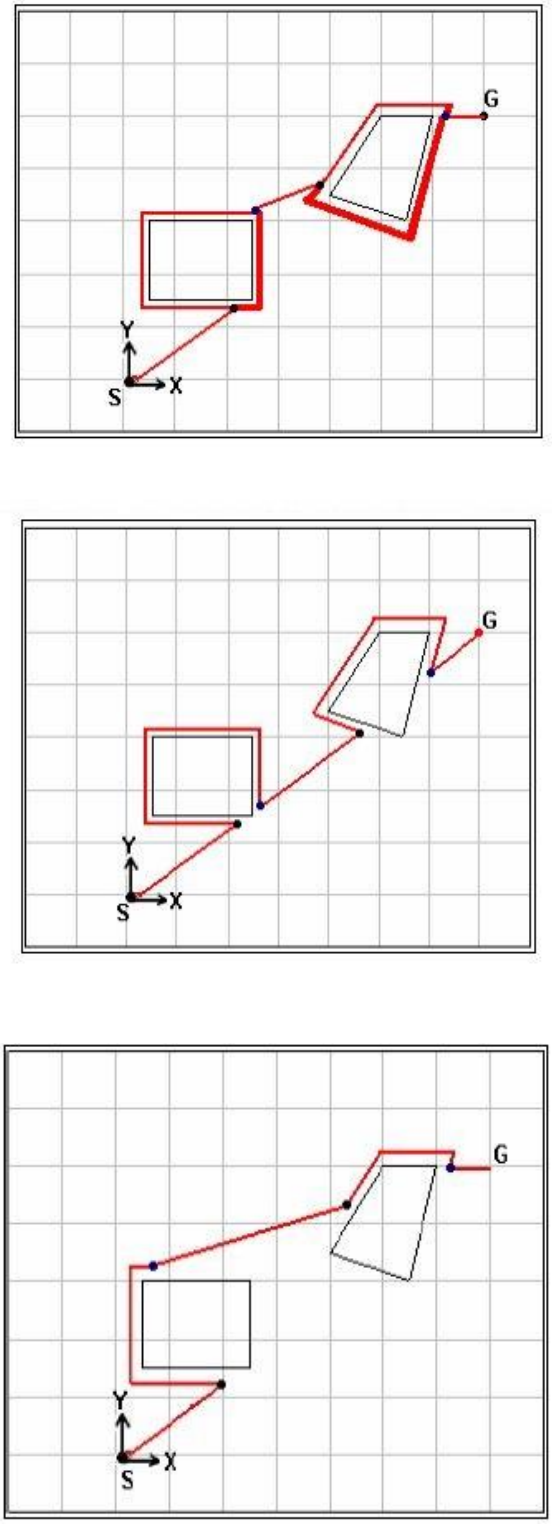


Figure 2-5 Les trois versions du "Bug algorithm". Version 1(haut). Version 2(milieu). Version 3(bas). (Figure tirée de [19])

« Artificial Potential field algorithm » (APF): cette méthode commence par assigner des charges virtuelles positives au robot et aux obstacles de la scène ainsi qu'une charge virtuelle négative au point d'arrivé. Par la suite, il est possible de calculer la somme des forces d'attraction et de répulsion appliquées sur l'effecteur. À partir de cela, le gradient des forces potentielles dirige l'effecteur au point voulu. Le principe général est présenté à la Figure 2-6 et les situations qui présentent des minimums locaux sont présentées à la Figure 2-7.

Avantages	Désavantages
• Rapide (chemin optimal) • Pas besoin de connaître les obstacles avec précision	• Présence de minimum local • Pas de convergence lorsqu'il atteint un minimum local. • Manque de précision au niveau des calculs et les contraintes ne sont pas prises en considération

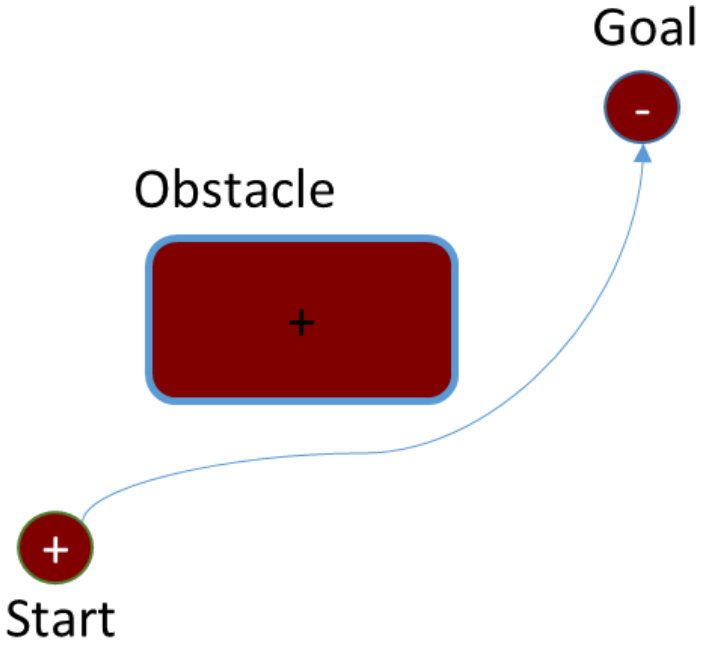


Figure 2-6 Présentation du principe de l'algorithme APF

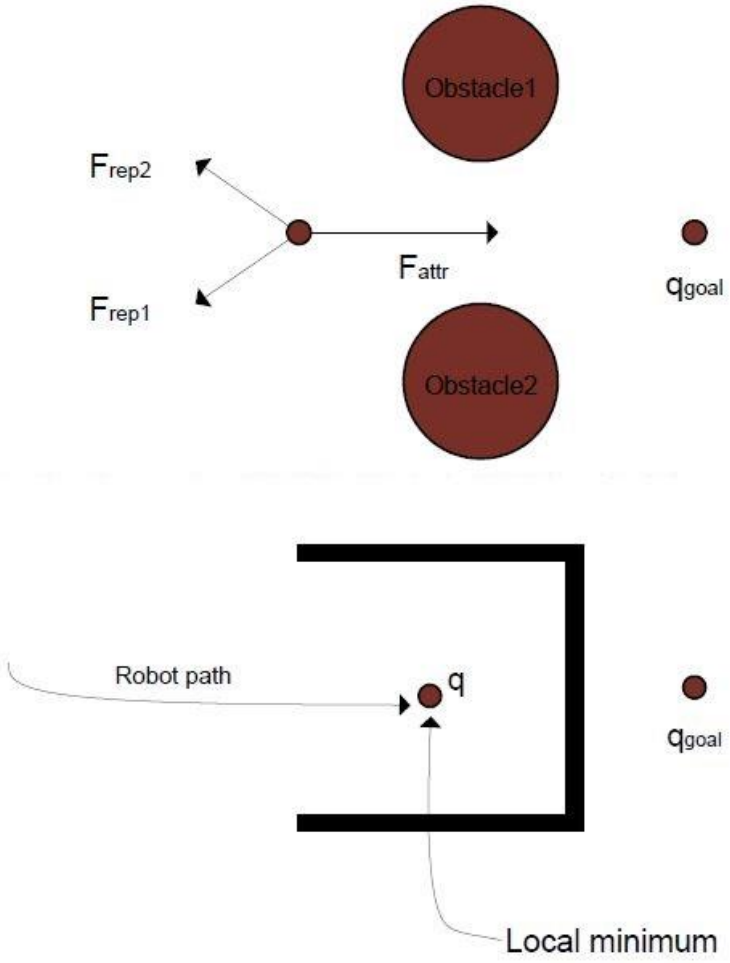


Figure 2-7 Situations menant à des minimums locaux. Obstacles symétriques (haut). Obstacles en forme de U (bas) (Figure tirée de [19])

« Vector Field Histogram »: cette méthode s'inspire d'un radar sonore. En effet, la détection d'objet est faite tout autour de la position initiale. À partir de cette détection, les données recueilles sont représentées par un histogramme polaire de la densité d'obstacle en fonction des directions possibles. Les grandes étapes de la méthode sont présentées à la Figure 2-8.

Avantages	**Désavantages**
• Calculs précis	• Consomme plus de ressources

- Pas besoin de connaître l'information au préalable.
- Générer une grille 2D ainsi qu'un histogramme polaire 1D demande beaucoup de temps.
- Pas de convergence dans le cas d'un obstacle en forme de U ainsi que lorsqu'il y a deux objets symétriques.

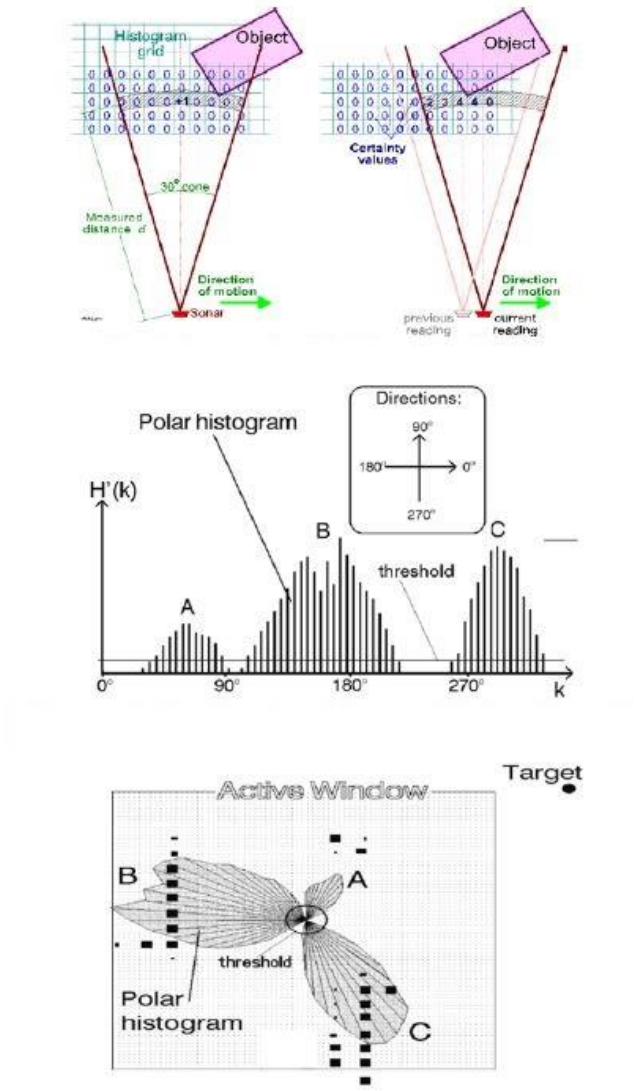


Figure 2-8 Étape 1(haut): détection d'obstacle. Étape 2(milieu et bas): Transformation des données en histogramme polaire. (Figure tirée de [19])

« Follow the gap method »: cette méthode est décrite en trois étapes principales. La première étape consiste à faire une liste de tous les interstices entre les obstacles et de trouver le plus grand d'entre eux. Par la suite, il faut trouver l'angle d'approche permettant de passer à travers le centre du plus

grand interstice trouvé. Finalement, il faut trouver l'angle du robot à la suite de son passage à travers l'interstice. Le principe général est présenté dans la Figure 2-9.

Avantages

- Choisit le chemin le plus court et permet l'évitement d'objets symétriques.
- Prend moins de temps étant donné que la prise de décision est locale.
- Choisit le chemin le plus court.

Désavantages

- Ne converge pas pour les obstacles en forme de U.

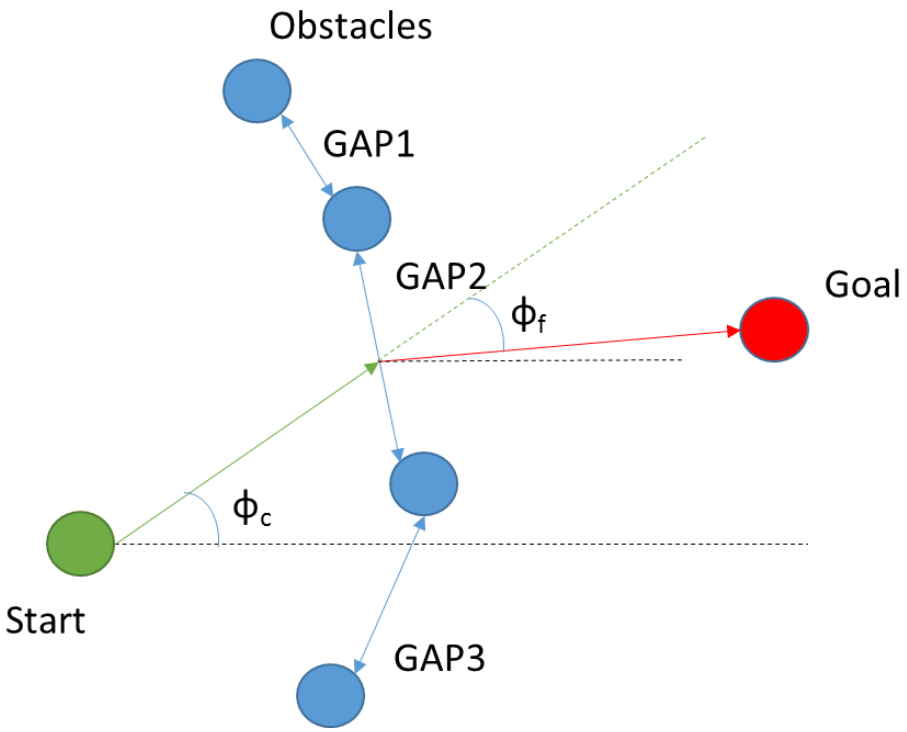


Figure 2-9 Présentations du principe de la méthode "Follow the gap"

« New Hybrid navigation algorithm (NHNA) »: les algorithmes hybrides utilisent deux couches afin de trouver une trajectoire adéquate. La première couche est la couche délibérative. Cette couche planifie la trajectoire avec les informations de départ. Pour cette couche, l'algorithme A* est utilisé. La deuxième couche est la couche réactive. Cette couche permet de gérer les nouvelles informations que le robot découvre lorsqu'il exécute la trajectoire donnée par la couche

délibérative. Le « Bug algorithm » est utilisé pour la couche réactive (Figure 2-11). Un exemple de cet algorithme est présenté à la Figure 2-10.

Avantages

- Génère le chemin optimal, par contre, il n'y a pas de limite en ce qui concerne l'altération du chemin de référence.

Désavantages

- Besoin d'informations apriori.
- Demande un important temps de calcul.
- Dans certaines conditions le robot s'arrête devant un obstacle.

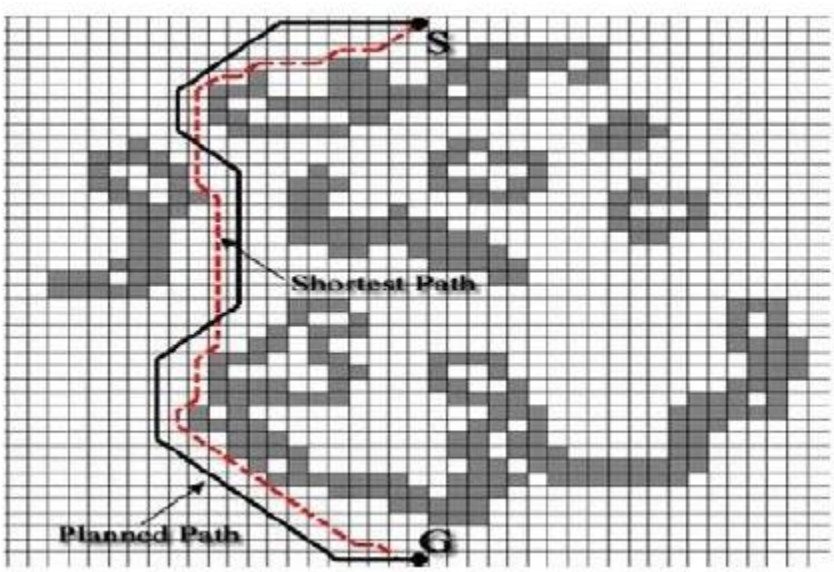


Figure 2-10 Présentation du principe de NHNA. (Figure tirée de [19])

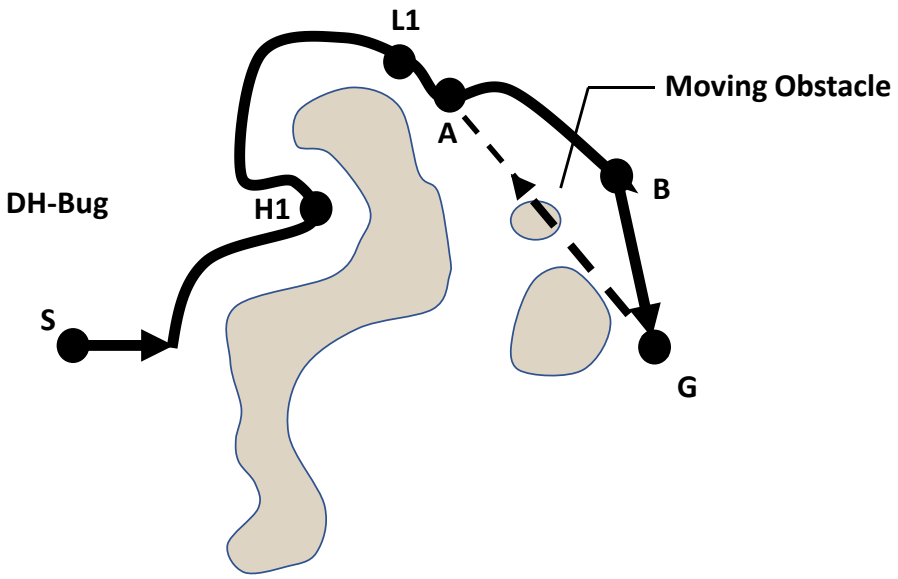


Figure 2-11 La couche réactive ("DH algorithm") s'adapte à des obstacles en mouvements. (Figure tirée de [19])

« Zybrid navigation algorithm with roaming trails (HNA) »: cette méthode est basée sur la méthode NHNA. La différence se trouve au niveau de la couche réactive. Effectivement, la couche délibérative reste la même (l'algorithme A*) tandis que la couche réactive est remplacée par l'algorithme des potentiels artificiels (AFP). L'illustration de la Figure 2-12 montre comment l'algorithme fonctionne.

Avantages

- Converge à part en présence d'obstacles en forme de U
- Meilleure limitation de déviations du chemin de référence.

Désavantages

- Besoin d'informations apriori.
- Demande un important temps de calcul.

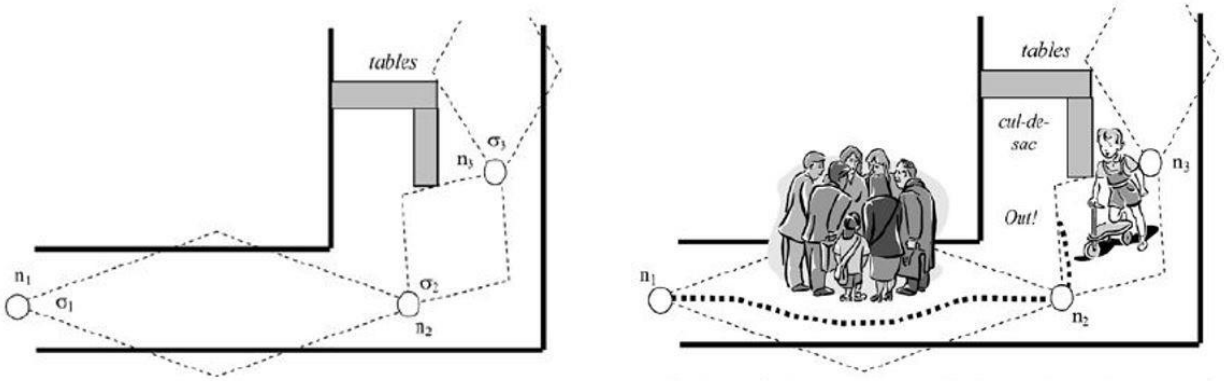


Figure 2-12 Image initiale avec les pistes (gauche). Image présentant le chemin emprunté avec une ligne pointillée en gras (droite). (Figure tirée de [19])

« **Combinatorial roadmaps and Search algorithms** »: cette méthode est composée de deux étapes principales. La première étape est la décomposition de l'image en nœuds. Ces nœuds sont représentés par des formes géométriques simples (carrés, triangles, etc.). Chaque nœud peut avoir deux états possibles, soit, occupé ou libre. Par conséquent, avec le tableau de nœuds, il est possible de trouver plusieurs trajectoires différentes pouvant amener le robot du point de départ au point d'arrivée. La deuxième étape sert à sélectionner la meilleure trajectoire à l'aide d'un algorithme de recherche. Par exemple, on peut utiliser le « octree » afin de décomposer l'image 3D en cube d'une certaine résolution. Par la suite, il est possible d'utiliser l'algorithme de recherche A* dans le but de trouver la trajectoire la plus courte entre le point de départ et celui d'arrivée. (Exemple : Octree + A*, Octree + Dijkstra's algorithm). Un exemple 2D de ce type d'algorithme est présenté à la Figure 2-13.

Avantages	**Désavantages**
• Généralement, la méthode choisit le chemin optimal ou optimal proche	• Demande un important temps de calcul.

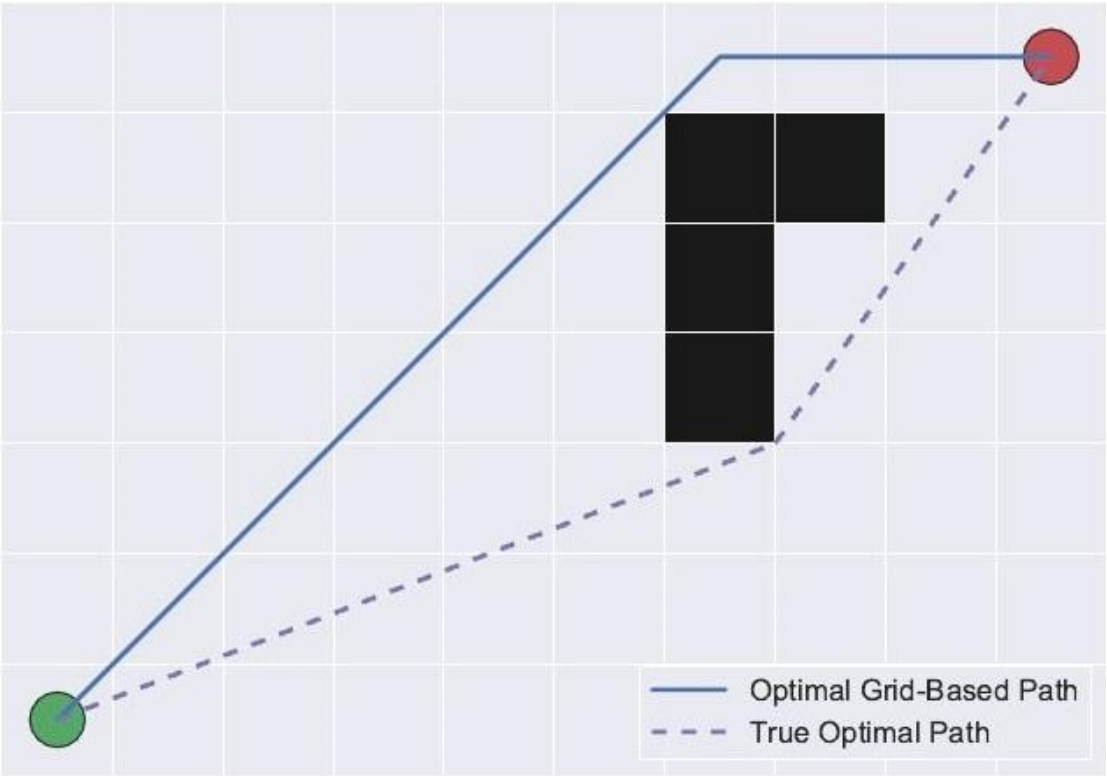


Figure 2-13 L'image 2D a été divisée en carrés. Ceux en gris sont libres et ceux en noir sont des obstacles. Le chemin est trouvé avec un algorithme de recherche. (Figure tirée de [21])

Sample-based planning »: ce type de planificateur de trajectoire peut se subdiviser en deux soit la requête unique (single query) et les requêtes multiples (multi query) (Figure 2-14). Pour la requête unique, la trajectoire est construite à partir de ramifications omnidirectionnelles dont la racine est le point de départ et/ou le point d'arrivée. Même si les ramifications sont omnidirectionnelles, il y a une direction qui est favorisée. Pour les requêtes multiples, plusieurs racines sont choisies. Ces racines sont des points ne faisant pas partie des obstacles. Chaque racine a des ramifications omnidirectionnelles. À un certain point, les ramifications de différentes racines fusionneront et formeront une trajectoire sans collision que le robot pourra prendre. (Exemple : RRT, PRM)

Avantages

- Rapide si une solution existe
- Si une solution existe, la méthode la trouvera

Désavantages

- Dépendamment du choix des points, il peut y avoir de la redondance.
- Si une solution n'existe pas, la méthode n'a pas de mécanisme d'arrêt.
- Dépendamment de l'algorithme choisi, le chemin trouvé ne sera pas optimal.
- Certaines méthodes dépendent de la condition initiale.

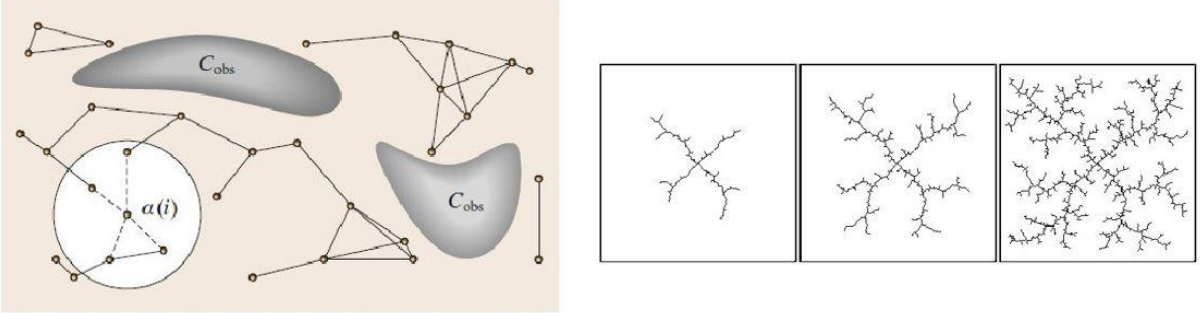


Figure 2-14 Présentation du principe en "multi query" (gauche) et en "single query"(droite) (Figure tirée de [18]

« AI based »: Généralement, un planificateur de trajectoire basé sur l'intelligence artificielle est un problème d'optimisation qui est résolu par un apprentissage à travers une base de données de bons et de mauvais exemples. (Exemple : Neural Network, Evolutionary Algorithm)

Avantages

- Pas besoin d'une approche analytique, le réseau apprend par lui-même.
- Une fois l'apprentissage fait, l'algorithme est rapide.

Désavantages

- L'apprentissage est lent.
- Les poids acquis de l'entrainement sont adéquats que pour certains cas connus.

- Demande beaucoup de ressources dans certains cas complexes.

2.2.3 Interface utilisateur

Comme mentionné ci-haut, dans ce travail on se veut de réduire l'utilisation du « joystick » pour contrôler le bras JACO. En effet, en considérant que les utilisateurs du bras JACO sont des personnes ayant des difficultés motrices, l'utilisation du « joystick » ralentit l'exécution d'une tâche. Par conséquent, il faut trouver un moyen de traduire l'intention de l'utilisateur afin que ce dernier, en harmonie avec un système de vision, puisse mieux contrôler le bras JACO. Trois options ont été retenues : un écran tactile, un dispositif vocal et un suiveur de regard.

2.2.3.1 Écran tactile

Pour l'écran tactile, l'approche utilisée serait d'afficher la scène captée par la stéréovision. Ensuite, l'utilisateur peut choisir l'un des objets de la scène à attraper. Il serait aussi possible d'inclure une liste d'action de base que l'on pourrait faire suite à la préhension de l'objet, par exemple, déplacer l'objet.

Avantages	Désavantages
• Vu de toute la scène, donc plus d'informations connues	• Affichages peut ralentir le programme • Besoins de plusieurs fonctionnalités d'affichage (par exemple : « zoom »)

2.2.3.2 Dispositif vocal

Pour le dispositif vocal, l'idée serait de donner des indications vocales afin de prendre des objets et d'effectuer des tâches. Pour se faire, il faut un algorithme afin de comprendre le sens d'une phrase dans un langage donné. Ensuite, il faut un autre algorithme afin d'associer les objets de la scène à des mots.[22-31]

Avantages	Désavantages
• Pas besoin d'affichage ni de l'utilisation des mains. • Interfaçage plus facile et intuitif pour l'utilisateur	• Besoin de beaucoup de traitement visuel et auditif • Besoin d'une grande base de données • Besoin d'un bon capteur de voix afin de différencier le bruit d'arrière-plan de l'utilisateur. • Limitation du système face à un objet ou une commande inconnu

2.2.3.3 Suiveur de regard

Pour le suiveur de regard, l'utilisateur regarderait l'objet afin que le bras puisse le prendre. Par la suite, il serait possible de regarder l'endroit où déposer l'objet pris. Pour utiliser cette approche, il faut, dans un premier temps, localiser les yeux d'une personne. Dans un deuxième temps, il faut trouver la direction du regard.[32-39]

Avantages	Désavantages
• Plus intuitif : l'objet qu'on regarde est celui qu'on veut généralement • Existe des méthodes pouvant aller dehors, mais le traitement est plus important	• Capteur infrarouge est la meilleure méthode, mais ne peut pas aller dehors. • Méthode de confirmation : une fois que l'on regarde l'objet voulu, il faut indiquer au bras d'aller le chercher • Une fois l'objet prit, difficile de dire ce qui doit être fait. • La détection des yeux peut être difficile.

CHAPITRE 3 OBJECTIFS, QUESTIONS DE RECHERCHE ET HYPOTHÈSES

3.1 Présentation de l'objectif principal et des objectifs spécifiques

L'objectif principal est de réduire le temps d'exécution et d'augmenter le taux de réussite de l'action d'approche et de préhension d'un objet à l'aide d'un système de vision artificielle et un système IHM.

Pour atteindre cet objectif principal, il faut atteindre les objectifs spécifiques suivants :

1. SO1 : Intégrer du capteur visuel au BRA
   1.1. Faire le filtrage et le traitement des données brutes acquises.
   1.2. Faire la traduction des données et établir la communication entre le capteur visuel et le bras JACO.
2. SO2 : Intégrer une méthode de planification de trajectoire et d'évitement d'obstacles au BRA
   2.1. S'assurer que l'entièreté du bras JACO ne rentre pas en contact avec des obstacles.
   2.2. Définir l'angle de préhension de l'objet par l'effecteur.
3. SO3 : Intégration de la méthode d'IHM choisie au BRA
   3.1. Comprendre et réagir aux intentions de l'utilisateur du bras JACO.
   3.2. Faire la traduction des données et établir la communication entre le système d'IHM et le capteur visuel ainsi que le bras JACO.

3.2 Question de recherche

Est-ce que l'utilisation d'un système de contrôle par vision artificielle et suivi du regard de l'utilisateur permettra d'améliorer les performances d'utilisation du bras JACO pour exécuter des tâches quotidiennes ?

3.3 Hypothèses

Hypothèse # 1 : Intégrer un système de vision artificielle permet de réduire le temps d'exécution de la tâche d'approche et de préhension d'un objet.

Hypothèse # 2 : Intégrer un système de vision artificielle permet d'augmenter le taux de réussite de la tâche d'approche et de préhension d'un objet.

CHAPITRE 4 MÉTHODOLOGIE

4.1 Cahier des charges

Le but de cette section est de souligner le cahier des charges de l'entièreté du projet afin de connaître les besoins essentiels, les besoins facultatifs et les contraintes du projet

4.1.1 Besoins essentiels

Les besoins essentiels sont les caractéristiques nécessaires pour faire fonctionner le système et pour effectuer la phase d'approche et de préhension d'un objet. Pour se faire, il faut être en mesure d'obtenir les caractéristiques suivantes :

- Obtenir des informations de la scène
- Être capable d'éviter des obstacles
- Être apte à approcher et attraper un objet
- Être capable de sélectionner un objet

4.1.2 Besoins facultatifs

Les besoins facultatifs sont des caractéristiques permettant d'améliorer ou d'apporter de nouvelles fonctionnalités. Cependant, ces caractéristiques ne sont pas nécessaires pour combler les fonctionnalités de base du système. Voici quelques exemples de ce genre de besoins :

- Être capable de fonctionner à l'extérieur
- Identifier les objets
- Établir les différentes actions que l'on peut faire avec un objet

4.1.3 Contraintes

La contrainte principale est de minimiser le coût du système de vision et du suiveur de regard que l'on ajoute au BRA. Tandis que la contrainte secondaire est de rendre l'intégration du système de vision et du suiveur de regard au fauteuil roulant discret et convivial pour l'utilisateur.

4.2 Robot d'assistance

Plusieurs bras d'assistance peuvent être sélectionnés pour ce projet. Le BRA JACO (Figure 4-1) a été choisi puisqu'il a la plus grande portée (90 cm) pour un poids léger (5.2 kg) et une consommation d'énergie typique faible (25W). De plus, il a été possible d'avoir accès à l'utilisation de ce BRA à la Polytechnique de Montréal.

Figure 4-1 Photo du bras robotique d'assistance JACO de Kinova Robotics (Figure tirée de [4])

4.3 Vision par ordinateur

4.3.1 Capteur visuel

Tableau 4-1 Tableau récapitulatif des différents capteurs visuels potentiels

Caractéristiques\Type de capteur	Infrarouge et combinaison	Caméra	Ultrason	Stéréovision
Développement	Faible	Faible	Élevé	Moyen
Précision	Élevé	Faible	Faible	Moyen
Vitesse	Moyen	Élevé	Moyen	Moyen
Robustesse	Élevé	Faible	Faible	Moyen
Calibration	Faible	Faible	Moyen	Élevé
Taille	Élevé	Faible	Faible	Moyen
Poids	Élevé	Moyen	Faible	Moyen
Consommation d'énergie	Élevé	Faible	Élevé	Moyen
Fonctionne à l'extérieur	Non	Oui	Oui	Oui
Prix	Bon marché	Bon marché	Chère	Chère

Le choix du capteur visuel est la stéréo vision puisque cette technologie offre les meilleurs compromis tout en subvenant aux besoins du projet (Tableau 4-1). En effet, les deux principaux défis sont le développement et la calibration. Ceux-ci peuvent bénéficier de la bibliothèque OpenCV[40] qui contient déjà des algorithmes de calibrations pour la stéréo vision et pour créer des nuages de points 3D de la scène. Pour le prix, la stéréo vision nommée Tara de e-con Systems [17] peut être acheter pour 249$ US qui est un prix raisonnable considérant que la Kinect pour Xbox One [9] offre le meilleur ratio qualité prix pour un coût de 99,99$ US.

4.3.2 Planification de trajectoire et évitement d'obstacle

La solution choisie est la « combinatorial roadmaps and search algorithms » qui est la décomposition d'un nuage 3D en nœuds suivie d'un algorithme de recherche afin de trouver la trajectoire idéale. Ces types d'algorithmes ont prouvés qu'ils sont en mesure de trouver le chemin optimal ou proche permettant d'atteindre une destination donnée. Par contre, leur faiblesse est la rapidité de calcul pour trouver ce chemin puisqu'il considère l'entièreté des informations du capteur visuel. Présentement, le système prend une photo de la scène, l'analyse et envoie le chemin au bras JACO. Sachant cela, la rapidité n'est pas aussi importante que l'optimisation du chemin pour l'instant. Par conséquent, la décomposition en cube de la scène et l'algorithme de recherche A* sont utilisés. De plus, un détecteur de collision du bras entier a été rajouté dans l'implémentation de l'algorithme A*.

4.3.3 Interface humain-machine

La solution choisie est le suiveur de regard. Considérant que l'utilisateur a des difficultés motrices des membres supérieurs, l'écran tactile ne serait pas une solution idéale puisqu'il ne pourrait pas bénéficier des fonctionnalités demandant des mouvements de précision comme le « zoom ». Pour l'utilisation d'un dispositif vocale, la quantité de traitement de signal vocal et d'image ainsi que l'intégration des deux technologies seraient très importantes particulièrement si l'on considère que l'environnement est inconnu. En effet, dans la vie de tous les jours, plusieurs objets sont utilisés, par conséquent, il faudrait une base de données énorme. En plus, l'utilisateur pourrait souffrir de trouble de la parole due au problème de motricité (par exemple, la dystrophie musculaire). Par conséquent, il ne reste que le suiveur de regard. Généralement, les muscles des yeux sont parmi les derniers muscles affectés par un problème de motricités des membres supérieurs[41]. Le suiveur

de regard est une solution intuitive et le traitement de données n’est pas aussi compliqué que le traitement relié à l’intégration du dispositif vocale. De plus, la reconnaissance d’objet et l’association de mots ne sont pas nécessaires. Le suiveur de regard choisi est le Eye tribe[42] pour le prix raisonnable de 100$ US. Le désavantage de ce suiveur de regard est le fait qu’il soit conçu pour une analyse en 2D. Cela veut dire qu’il y a une quantité importante de développement nécessaire pour le faire fonctionner en 3D avec une précision acceptable.

CHAPITRE 5 ARTICLE 1: ASSISTIVE ROBOTIC ARM CONTROL USING ARTIFICIAL STEREOVISION AND EYE-TRACKING

L'article présenté par Yann-Seing Law-Kam Cio, Maxime Raison, Cédric Leblond-Ménard et Sofiane Achiche dans ce chapitre a été soumis le 17 juillet 2017 dans le journal *Machine Vision and applications*.

Abstract

Assistive robotic arms have become more and more popular to help users with upper limb disabilities to achieve autonomy in their daily tasks, such as eating, drinking, and grasping objects in general. Usually, these robotic arms are controlled with an adapted joystick. Joysticks are user friendly when it comes to a general approach to an object. However, they are not as intuitive when having to accurately approach an object, such as the case when obstacles are present. This is because they only give access to three degrees of freedom (DOF) at a time to control an arm with six DOF. This can be problematic for users with disabilities. Further, for users with severe disabilities, joysticks are difficult to use. Consequently, some studies have already proposed a vision-based control system, sometimes combined with an adapted user interface technology such as a tablet. The combined use of artificial stereovision and eye tracking seems to be a promising solution, as the user's vision is usually dissociated from their upper limb disability. But to our knowledge, this integration of technologies has never been proposed yet. Consequently, the objective of this study is to develop the control of an assistive robotic arm using the low-cost combination of artificial stereovision and eye tracking. Using the developed control system, the robotic arm successfully reached and grasped an object for 80% of the trials with an average time of 13.7 seconds without obstacles, 15.3 seconds with one obstacle and 16.3 seconds with two obstacles. Furthermore, the cost of the eye tracker and the stereovision remains below 400$ which represents around 1% of the price of the JACO arm.

5.1 Introduction

Recently, assistive robotic arms (ARA) caught the interest of clinics, industry, and research. The most common assistive arms are the Manus arm which is the predecessor of the iARM (Exact dynamics, Didam, The Netherlands) [1], the WAM arm (Barrett Technology, LCC, Newton, USA) [2], the DLR LWR-III arm (DLR Robotics and Mechatronics center, ,Weichselbaum,Germany), [3] and the JACO and MICO arms (Kinova Robotics, Boisbriand,Canada) [4]. The benefits to the patients of using ARAs have already been proven [43-47]. However, one of the main challenges with the ARAs is that they must be versatile enough to support the most amount of people with various upper limb disabilities, including muscular dystrophies, cerebral palsy, and tetraplegia, and help them recover most of their autonomy and improve their quality of life.

To overcome this challenge, an adequate human-computer interaction (HCI) interface must be used to control the ARA. Currently, the most common technology used to control the ARA is a joystick, which might be suitable for individuals who have less severe upper limb disabilities such as early stage muscular dystrophies. However, it is not adapted for individuals with severe upper limb disabilities such as late stage muscular dystrophies or severe spasticity. Indeed, most ARAs have more degrees of freedom (DOF) than the number of control modes of the joysticks, e.g. 6 DOF vs. 3 modes of control.

Two main issues come with the use of a joystick. The first one is that the ARA has no information about its workspace. By providing such information to the ARA, it becomes aware of the possible positions it can reach and the constraints that it must respect. To fix this issue a visual sensor can be used to map the workspace in 3D.

The second issue is the number of control modes of the joystick. By controlling a 6 DOF robotic arm with 3 modes of control, the user can easily be confused, and therefore yield a slower execution of the task at hand. Reducing the number of control modes by changing the HCI to a more adequate and intuitive one should speed up the execution of the task.

The work presented in this paper offers an alternative solution for better user support by modifying the control interface of an ARA using a low-cost combination of stereovision and eye tracking. The low-cost aspect of the solution is crucial, as one does not want to increase the overall cost of the ARA significantly. By minimizing the overall cost of the solution, the user will be more inclined

to adopt the solution. It is worth noting that in this study we use the JACO robotic arm from Kinova Robotics as an ARA.

5.2 Literature review on vision for robotic control

Vision systems for robotic control have been thoroughly researched in various industrial fields, since the composition of these systems is usually complex. The following sections will describe the main systems with the goal to define the main requirements needed for the design of our integrated system.

The AROMA system is presented in [48, 49]. It is composed of an infrared (IR) depth-color sensor combined with voice recognition to control an ARA. To reach and grasp an object, the authors used a direction-based command system combined with a visual servoing path planner. In [50], an IR depth-color sensor is used as well. The authors of this paper focused on image processing and object recognition. Inspired by mechanical probing, we developed a method to segment the images. Then, using a database-based neural network, we were able to identify a specific object on the scene. Once the user chooses an object through a computer interface, a decision tree is used to decide how to grasp this specific object. In [51], an eye-in-hand ARA is presented. By fixing a stereo camera on the end effector of the ARA, it allows the ARA to reach and grasp the selected object. Next, word-based voice control and hand-based gesture control with an IR depth-color sensor is presented in [52, 53]. By recognizing the object on the table and associating it with a word, the user can select the object by saying the word. Then the user can command the ARA with a hand gesture. In [54], stereovision is explored, the objects are recognized using a neural network and a virtual 3D scene is created using a 3D reconstruction of the identified objects. With this model, it is possible to find a collision-free path. For the HCI part, the author uses a chin joystick to control the system. In [55], the author uses a camera on the shoulder of the ARA to make a graphical display for the user. Another camera is within the end effector of the ARA to reach and grasp an object using visual servoing. As for the HCI, the author is comparing the use of either a joystick or a touch screen combined with a graphical display to pinpoint the object. The author also compares a graphical display in which the user can move the camera and another one where the user cannot move the camera.

The work reported in [56] focuses mainly on path planning and obstacle avoidance. The authors use a probabilistic roadmap to find a collision-free path. Furthermore, the path planning is calculated on a chip, which allows the ARA to find a path in less than 1 millisecond. To obtain the whole scene, four IR depth-color sensors are used which allow the complete reconstruction of the objects. In [57], the authors look at different HCI to control the ARA. The first method presents a list of selections for the user to choose from. As for the HCI, a noise sensor reads the sound generated by the user to go over the selections. The second method allows the user to use a joystick to select an item on the 2D image presented on a graphic interface. The third method allows the user to pick a 3D point in the scene where the ARA can reach. Once the ARA has grasped the target object, the user can activate another mode that allows him to place the object somewhere in the scene. In every mode, the ARA follows a collision-free path that is generated in the robot operating system (ROS) with the information given by an IR depth-color sensor.

Finally, a system with multiple options of HCI, depending on the severity of the upper body disability, is presented in [58]. The first option is a touchscreen, which is ideal for people who are still able to use their fingers. The second option is a trackball combined with a jelly switch, ideal for people that cannot use their fingers but still can use their hand. With the palm of their hand the users can use the trackball to select an object on the graphical display and the jelly switch to stop or make the robot move. The third option is a combination of a speech recognition device and a jelly switch. The speech recognition device allows the user to use voice commands to control the robot. Finally, the head tracker switches on to control the cursor of the graphical display. A stereovision camera fixed on top of the end effector of the ARA feeds the images to the graphical display.

From these state-of-the-art systems, one can conclude that to build a vision-based system to control ARA three basic components need to be considered: 1) selecting the visual sensors and their location, 2) developing a solution to detect the objects of interest for the user and 3) deciding upon a user interface technology. These three components are added to the path planning and obstacle avoidance algorithm that will ensure the autonomy of the ARA.

For this paper, the developed vision-based ARA system, the Point of Gaze Assistive Robotic Arm (PoGARA), is composed of the JACO arm, a stereovision camera, an A* obstacle avoidance algorithm and a low-cost eye tracker. It is important to note that the literature shows that even for

people with severe muscular dystrophies, their vision remains functional [41]. Hence, the development of the system presented here focuses on taking advantage of the availability of the user's vision. There is little work in the literature that focuses on the use of the visual-motor abilities of patients to their benefit through the use of an eye tracker in 3D as a user interface technology. This is most probably due to the high cost of high-end accurate eye trackers such as the Face-Lab 5 [14] and the low accuracy of low-cost ones such as the Eye Tribe (The Eye Tribe, Copenhagen, Denmark) [42]. For instance, in the work reported in [59], the authors use an active stereovision camera to capture the scene and to follow the gaze of the user. First, the stereovision system finds the position of the user's face to find the direction of the gaze. Then, the stereovision camera moves along the direction of the gaze to detect the object of interest. Finally, an ARA reaches and grasps the object of interest. With this system, however, it is not possible to capture the scene and to track the gaze of the user at the same time. In other words, if there is a change in the scene or in the direction of the user's gaze, there will be a significant delay before the system will consider these changes. In [60], the study presents a simulation of a vision-based control system combined with an eye tracker to control a prosthetic arm. Here, a stereovision camera tracks the gaze of the user and a depth sensor is used to capture the scene. The depth sensor is fixed on the user's shoulder, restraining the user from moving his shoulder. The author proposes tongue control to select the object of interest. Through simulation tests, the author associated four actions to four tongue positions to indicate the target object to be grasped and the grasp position of the palm. Since the simulation was in the depth plane of the monitor, only the information of one eye was taken into account. In the simulation environment, such approximation is acceptable. However, in the real environment it will not be possible to use only the gaze of one eye as well as associate an object to a tongue position. The research carried out in [61] presents a combination of two depth sensors and an eye tracker to help the user control an upper body exoskeleton. The eye tracking is used with a 2D screen. This means that the user has to look at a screen to get an object instead of looking directly at the object in the real 3D scene. The author uses two Kinect cameras to get the position of the user's hand and mouth as well as the workspace. The first Kinect camera captures the workspace from a top view. The second Kinect camera is positioned in front of the user to follow the user's face. To identify objects and the user's hand, the author uses colored markers. This set-up, however, is not suited for an unstructured environment and for a wheelchair mounted with an ARA. This is because of the number of Kinect cameras, which are not very energy efficient, would

quickly deplete the battery. Additionally, Kinect cameras are not designed to work outside. Furthermore, the user needs to look at a 2D screen to select an object, which is very unnatural. Finally, the work presented in [62] presents a teleoperation system composed of a six-DOF force feedback joystick, a gaze tacker, a monitor and a Kinect camera. The six-DOF force feedback joystick is used to make the user feel any kind of force or constraint applied to the ARA. The use of the gaze tracker to pinpoint the target position is combined with the force feedback joystick to find the final position and orientation of the ARA. Such a system is not suited for our aimed application, since the joystick is still a big part of the system. In addition, the use of force feedback can be very imprecise with users suffering from muscular weakness or dystrophies. Finally, as mentioned above, the user has the extra burden to manipulate his environment by looking at a 2D screen.

Hence, to the best of our knowledge, there is no system that integrates a low-cost eye tracker combined with stereovision for the control of an ARA for patients with upper body disabilities. Therefore, the objective of this study is to control an ARA with a low-cost eye tracker and stereovision. This paper details the steps followed for the development of such a system.

Section 3 presents the system design based on the main basic components: the visual sensors, the path planning and obstacle avoidance algorithms and the user interface. Once the choice of each component is known, the integration of the whole system will be presented. Next section 4, presents the results and discussion followed by the limitations and future work presented in section 5. Finally, section 6 ends the paper with a conclusion.

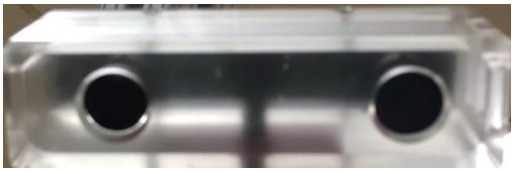
5.3 System design description

This section describes the different parts of the integrated system.

5.3.1 Visual sensors

The chosen visual sensor is the stereovision camera system. The two main challenges of using this technology are the data processing and the calibration, both of which can be aided by the software OpenCV [40]. This software has built-in algorithms to calibrate the cameras and map 3D scatter points of the scene. As for the price, the Tara stereovision camera (e-con Systems, St-Louis, USA) [17] can be bought for 249$ (May 03, 2017) which is a comparable price considering that the Kinect camera for Xbox One [9] offers an excellent quality-price ratio for a cost of 99,99$ (May

03, 2017). The Tara stereovision camera and its features of interest for this study are presented in Figure 5-1.



Features	Values
Dimensions (mm)	100X30X35
Mass(g)	80,5
Price ($ US)	336
Energy consumption (W)	2,27
Camera resolution(pixels)	1504X480 à 60 FPS
Scope (m)	0,5-3

Figure 5-1 Tara stereovision camera (left). Features of Tara (right)

5.3.2 Path Planning and Obstacle Avoidance

Path planning and obstacle avoidance are an important step in the development of a vision system for the control of an ARA. Considering that most of the time, the scene captured by the vision system will include more than just the object of interest, the algorithm must be able to find the path to get the object of interest from the objects cluster.

The chosen solution consists of combinatorial roadmaps and search algorithms [18-20]. These types of algorithms have proven that they can find the optimal, or near optimal, path to reach a point. Usually, the downside of these path planners is the speed since they consider all the information of the sensor to make their analysis. Currently, the developed system takes a picture of the scene then analyzes the scene to finally send the path to the JACO arm. Knowing this, the speed of the system is not as important as the optimization of the path. For this reason, a combination of a cube decomposition of the scene and the search algorithm A* is used. An implementation of this search algorithm is presented in [21]. Furthermore, a collision detector for the whole arm is added to the implementation of the A* algorithm.

5.3.3 User Interface

The challenge of the user interface technology is to find a way to make the ARA understand the intention of the user, and use this information in harmony with the vision system to control the JACO arm.

Considering that the user has an upper body disability, the tactile screen would not be an ideal solution since precision functionalities like zooming in and out on the screen would be hard for the user to execute. As for vocally controlled devices, the amount of treatment for a vocal-visual integration is enormous and difficult since we work in an unstructured environment [22-31]. Indeed, in day-to-day life, we come across a multitude of objects, thus, a huge database is necessary. Furthermore, the user could be suffering from a speech disorder, which can be caused by muscular dystrophy. This leaves us with the eye tracker option.

Generally, the muscles of the eyes are among the last ones that are affected by an upper-body disability. The eye tracker is an intuitive solution and the data treatment is not as complicated as it is for vocal devices [32-39]. Also, there is no need for object~~s~~ recognition or word association.

In order to keep the price of the integrated system as low as possible, we selected a low-cost eye tracker, the Eye Tribe (100$). The downside of this eye tracker is that it is mainly designed for 2D analysis and on-screen projections of gaze. This means that an important algorithmic improvement is needed to make it work in 3D with an acceptable precision.

5.3.4 System Integration

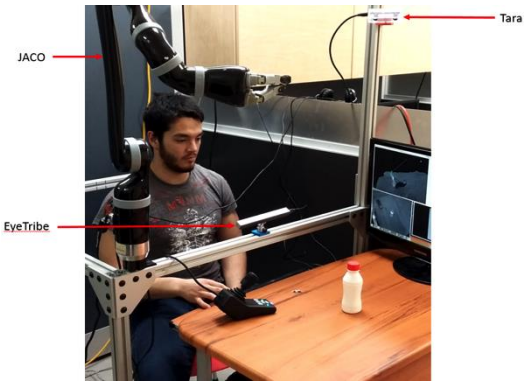


Figure 5-2 Presentation of PoGARA system. JACO is the ARA. Eye Tribe is the eye tracker and Tara is the stereovision camera

In the next subsections, we will describe the steps necessary for the set-up of the PoGARA shown in Figure 5-2. Even if some steps seem trivial, they remain important for the reproducibility of the system by other researchers.

5.3.4.1 Calibration

There are two main components that need to be calibrated for this vision system: the stereovision camera and the eye tracker.

Using OpenCV, it is possible to calibrate the stereovision camera, Tara. Indeed, with a chessboard, the calibration is done by finding the corners of every square of the boards for each camera. Then, we match those corners and find the parameters of the mathematical model used by the stereovision camera.

As for the Eye Tribe, the first calibration is done with its specific software, Eye Tribe UI, and using a screen. The screen is at a certain known distance from the Eye Tribe facing the user. With the same set-up, the second calibration is done to find the field of view of the user's gaze. To do this, the user has to fix the middle of the Eye Tribe device at a given distance between him and the Eye Tribe.

5.3.4.2 3D Scene Acquisition and Treatment

The first step is the acquisition of the 3D scene using the stereovision camera, Tara. With OpenCV library, it is possible to find a disparity map, which shows the difference between the left and right cameras as it is shown in Figure 5-3. Then with this disparity map, it is possible to yield the 3D map of the real distance in the frame of the stereovision camera as shown in Figure 5-4.

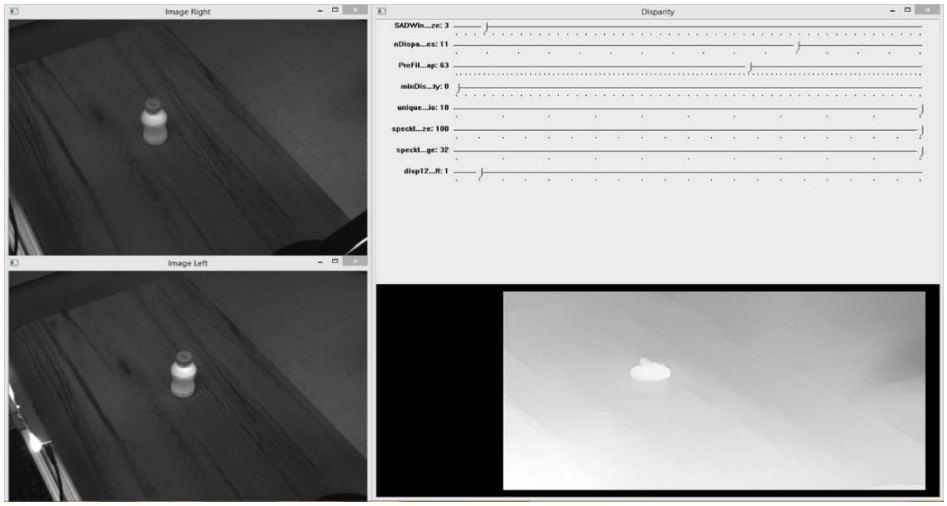


Figure 5-3 Image from the right (top-left) and left (bottom-left) camera and the disparity map (right)

The second step is to divide the 3D scene into nodes, where a node is a voxel of a certain size.

To do this, the library Octomap [63, 64] was used to divide the 3D scene into a 3D grid of nodes. The nodes can have two states: free or occupied.

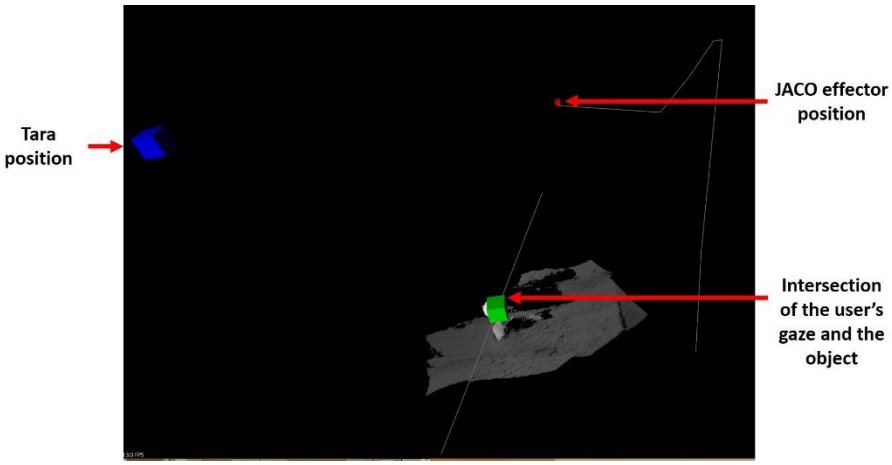


Figure 5-4 Image of the scene. The blue cube is the position of the Tara. The green cube with the white line is the gaze of the user. The white lines with the red cube is the JACO arm

5.3.4.3 Object selection and path planning.

The first step is to find the eye position and the direction of the gaze of the user with the Eye Tribe. Knowing the real distance between the user and the Eye Tribe and the [x,y] position of the left and right eye given by the Eye Tribe, the real 3D eye position of the user is found.

As for the 3D target point, it is possible to find it with the [x,y] coordinates on the calibration screen given by the Eye Tribe and the 3D calibration position of the screen. By changing the 3D eye position and the 3D target point into the same frame of reference, it is possible to find the gaze of the user. The frame of reference here is that of the stereovision camera since this frame contains the real 3D coordinates of the scene. Once the object is chosen, a path planning is needed. Using the 3D grid of nodes representing the scene, the search algorithm A* is used to find the optimal path to reach the object.

The A* algorithm makes sure that the end effector of the robotic arm is traveling in a collision-free path but it does not in any way guarantee that the rest of the arm will not collide with obstacles. For this reason, numerical inverse kinematics based on forward kinematics is used to obtain the Cartesian position of the whole arm. With this information and the 3D grid of nodes representing

the scene, a collision detector can be implemented. The implementation of the A* algorithm comes from [21]. Based on that implementation, the collision detector was inserted in the function U.Top(). The U.Top() function outputs the next node based on the obstacles, the current node and the final node. The implementation of the collision detector and the U.Top() function are presented in the form of the following pseudo-code:

```
function Collision_detector (thetas, obstacle)
    XYZpositions = DIRKIN (thetas)
    XYZsplits = split(XYZpositions)
    While (i<XYZsplits.size()) Do
        While(j<obstacle.size()) Do
            Dist = Euclediandist(XYZsplits(i), obstacle(j));
            Radius = radiusvalue(jointnumber, disteffgoal);
            If (Dist < Radius) Then
            return false;
    return true;
end
```

```
function U.Top()
  While done! = true and U.empty() != true Do
    out = U.back();
    U.pop_back();
    If (out.bptr != nullptr) Then
        While (i < rotangle.size()) Do
            EulerXYZ = orientation (out.bptr->orientation ,rotangles[i]);
            Out.thetas = INVKIN (out.bptr->thetas, out.pose);
            If(rotangles.size() > 1)
                Done = Collision_detector (out.thetas, obstacles)
                If (done == true) Then
                Break;
            else
                Done = true;
            end
  return out;
end
```

The functions **DIRKIN** and **INVKIN** are the direct and inverse kinematics, respectively. The output of the direct kinematics is the Cartesian position of each joint of the JACO arm. As for the inverse kinematics, the inputs are the initial condition and the goal. The initial condition is the previous angular position of each joint and the goal is the wanted pose (i.e. position and orientation) of JACO's end effector.

The **Split** function provides the intermediate points between two consecutive joints of the JACO arm. The distance between two consecutive joints is divided by sections of 5 cm to make sure that most of the arm is considered when checking for collisions. The output of the function is the 3D positions of all the intermediate points of the joints as well as the joint positions.

The **radiusvalue** function calculates the minimum distance needed between the arm and an obstacle based on the segment of the arm (*jointnumber*) and the distance between the goal and the end effector (*disteffgoal*).

The **orientation** function calculates the orientation of the end effector based on the XYZ Euler angle convention. The next orientation is based on the previous orientation. The input argument *rotangles* is a vector of rotation angles that can be applied to the previous orientation. If a collision is detected, then a rotation angle is applied to the previous orientation. The values of the *rotangles* vector depend on the obstacles around the object of interest. To choose the value of the vector, a method inspired by the vector field histogram, presented in [19], is used. Figure 5-5 gives an example of this method. 0° is the initial orientation. The range of angles [340°, 20°] and [60°, 120°] would be the forbidden angles since there are obstacles in the way. Knowing this, it is possible to see that there are more forbidden angles in the range [0°,180°] than in the range [180°, 360°], which means the final orientation should be within the range of [180°,340°]. It is then possible to deduce that the *rotangles* vector would be composed of [0°, -5°, -10°, -15°, -20°, -25°, -30°, -35°, -40°, -45°].

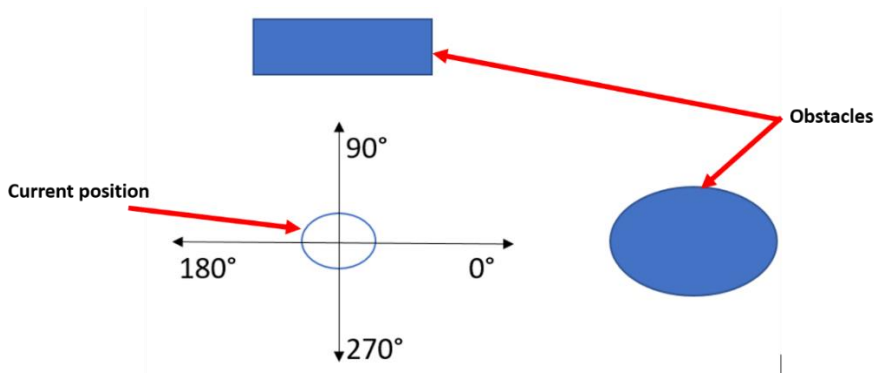


Figure 5-5 Method based on vector field histogram for collision detection

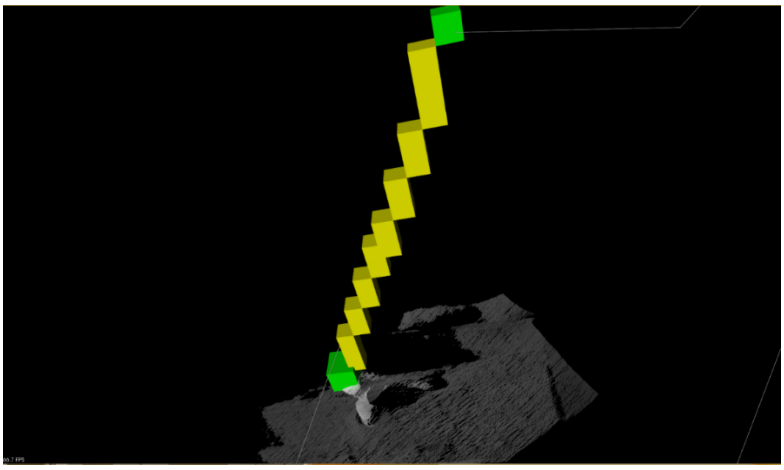


Figure 5-6 The path that the JACO's end effector would take

Once the path is found (Figure 5-6), a series of 6 angular positions, one per joint, is sent to the JACO arm to execute the trajectory. It is important to note that most of the trajectory is made with the gripper in its closed position. Once the end effector is within 4 nodes to the object of interest, the gripper opens. When the trajectory is fully completed, the gripper grasps the object. We choose 4 nodes to consider the length of the fingers of the end effector in the open hand position.

5.4 Results & Discussion

In this section, the experimental tests and the results will be presented and discussed. As mentioned before, the objective is to approach and grasp an object in a scene fast enough to replace the use of a joystick by users of the JACO arm. Knowing this, three experimental tests were carried out: grasping the object of interest alone, with one obstacle and with two obstacles.

For each of these tests, five trials were carried out with the object of interest in the same position to evaluate the repeatability. The data extracted from these tests are the time to find the path of the end effector (in seconds), the time to execute the path (in seconds) and the success of the trial.

The success of the trial can take three possible values. The value “1” means that the arm didn’t touch any of the obstacles on the way to grasp the object. The value “0.5” means that the arm clearly touches an obstacle or the object of interest but still grasps the object of interest. The value “0” means that either the arm did not grasp the object of interest or that it hit an obstacle hard enough to make it fall.

It is important to note that all the objects are empty plastic bottles. This means that the reach-and-grasp task is harder to execute since it is easier to make a light object fall than a heavier one.

Figure 5-7 shows the set-up of the tests consisting of the object of interest alone. The results of these tests are presented in Table 5-1.

Figure 5-7 Position of the object of interest (left). Object of interest grasped by the JACO arm (right)

Table 5-1 Results of the tests with the object of interest alone shown in Figure 5-7

Trials	Path planning time (s)	Execution time(s)	Success	Comments
1	0.38	12.7	1.0	
2	0.46	13.9	1.0	
3	0.42	14.4	0.0	Finger tips hit the object of interest
4	0.38	13.7	1.0	
5	0.33	14.0	1.0	
Average	0.39	13.7	-	
Standard deviation	0.05	0.6	-	

Based on the results of Table 5-1, the average time to find a suitable path is 0.39 seconds. This means that once the user has chosen an object, he will not have to wait a long time before the JACO arm starts moving. As for the execution, the average time is 13.7 seconds.

As far as the success of the trials, it is possible to see that 4 out of 5 trials have been collision-free and successful. For the third trial, while the effector was open the fingertips hit the object of interest, making it fall and consequently preventing it from being grasped by the end effector. This

is due to the fact that the obstacle avoidance algorithm didn't take into consideration the open hand scenario when calculating the trajectory.

Figure 5-8 shows the set-up of the tests consisting of the object of interest with one obstacle. The results of these tests are presented in Table 5-2.

Figure 5-8 Position of the object of interest and the obstacle (left). Object of interest grasped by the JACO arm (right)

Table 5-2 Results of the tests with the object of interest and one obstacle shown in Figure 5-8

Trials	Path planning time (s)	Execution time(s)	Success	
1	0.40	15.4	1.0	
2	0.34	14.7	1.0	
3	0.39	16.1	0.0	Finger tips hit the object of interest
4	0.36	14.3	1.0	
5	0.47	15.8	1.0	
Average	0.39	15.3	-	
Standard deviation	0.05	0.8	-	

For the situation where there is an obstacle with the object of interest on the table, we decided to place the obstacle in the way of the robotic arm. Indeed, if we look at the photo on the right of Figure 5-7 and the photo on the right of Figure 5-8, we can see that the wrist of the arm would have collided with the obstacle if no obstacle avoidance algorithm had been included. It is important to note that the distance between the obstacle and the object of interest is around 20 centimeters.

It is interesting to see that the mean and standard deviation of the path planning time for this situation is the same as the previous situation. This means that the addition of one obstacle to this current situation does not slow down the path planner. However, it is possible to see the impact of

the added obstacle in the execution time. Indeed, the average time of 15.3 seconds makes this situation 1.6 seconds slower than the previous one. As for the success of the trials, in most of them the JACO dodged the obstacle and grasped the object of interest perfectly. However, as in the previous situation, in one of the trials the fingertips bumped the object of interest hard enough to make it fall, preventing the arm from grasping it.

Figure 5-9 shows the set-up of the tests consisting of the object of interest with two obstacles. The results of these tests are presented in Table 5-3.

Figure 5-9 Position of the object of interest and the two obstacles (left). Object of interest grasp by the JACO arm (right)

Table 5-3 Results of the tests with the object of interest and two obstacles shown in Figure 5-9

Trials	Path planning time (s)	Execution time(s)	Success	Comments
1	0.59	17.0	1.0	
2	0.46	14.7	1.0	
3	0.45	14.9	1.0	
4	0.52	20.7	0.0	Made an obstacle fall down
5	0.48	14.4	1.0	
Average	0.50	16.3	-	
Standard deviation	0.06	2.7	-	

For the situation where the object of interest is surrounded by two obstacles, we kept the first obstacle in the same position as the previous test. The second obstacle was placed in such a way as to make the arm pass in between the two obstacles before reaching and grasping the object of interest. This makes the grasping task more complex.

The effect of adding another obstacle is mostly noticed in the path planning time and the success of the trials. The average path planning time increased by 0.11 seconds. As for the success of the

trials, we were able to grasp the object of interest in all but one of the trials. In the second trial, the red and white bottle fell down because the wrist of the JACO bumped into it. As for the average execution time, it is 1 second slower than the previous situation. This result was predictable since there were more obstacles to avoid in this situation. However, it remains an acceptable increase.

Overall, the objective of this study was achieved as we were able to use a low-cost eye tracker and stereovision to control an ARA. Indeed, the overall cost of the control system is below $400, which means that the added cost is around 1% of the price of the JACO arm. Moreover, our results show that in 4 out of 5 trials (80%) for each situation, the JACO arm was able to reach and grasp the object of interest with an average execution time of 13.7 seconds without obstacles, 15.3 seconds with one obstacle and 16.3 seconds with two obstacles. However, the failed trials have allowed us to pinpoint an issue with the obstacle avoidance. Indeed, the current obstacle avoidance algorithm does not take into account the fingers when the hand of the ARA is open. This was the reason for 2 out of 3 failed trials. For the last failed trial, the end effector tried to reorient itself at a certain collision free direction. However, the maneuver leading to that collision free direction did not take into account the second obstacle.

It is worth noting that the image processing currently used is a threshold in XYZ to make a bounding box of the scene based on the reach of the JACO arm.

5.5 Limitations and future work

The main limitations of the current version of the PoGARA system remain the implementation of the image processing and the robustness against IR.

The image processing currently used combined with the filter presented by [50] could improve the current version of the PoGARA. Ideally, the filtered 3D point should have all the objects on the table in a bounding box and rectify the table level to be the same at any depth. These modifications would benefit the user as well as the path planning and collision avoidance. The targeting of the object by the user could be made significantly better by creating a bounding box around the object. As for the path planning and collision avoidance, the orientation and approach could be improved by doing the image processing the same way as in [50].

For the robustness against sunlight, an ultraviolet-IR cut off filter needs to be added to Tara's lens. In addition, the exposure of Tara is currently not activated. This way an exposure to sunlight will not blind the system.

Additionally, the PoGARA should move toward capturing a video instead of an image to get closer to a “real time” vision system. Another factor to be considered is the user’s comfort. The user shouldn’t notice a significant lag between the command he gives and the result he gets. The components of the system that would be affected by the real-time implementation are the path planning and obstacle avoidance modules. For this reason, a dynamic path planner needs to be implemented. A recent path planner called HD* has been used for drones [21] . It could be adapted to the PoGARA to improve the path planning speed and to smooth the trajectories.

5.6 Conclusion

A high-level control system was presented to control a robotic arm to assist people with upper limb disabilities. By using a stereovision camera and an eye tracker combined with a path planner of the full robotic arm, we were able to control the assistive robotic arm JACO to reach and grasp an object with and without obstacles 80% of the time. The average times needed to accomplish a task were 13.7 seconds without obstacles, 15.3 seconds with one obstacle and 16.3 seconds with two obstacles. We deem these times to be satisfactory. Although the presented system has some important weaknesses in terms of image processing, the results still show that it has the potential to improve the quality of life and the autonomy of people with severe upper body disabilities by exploiting their vision to better function in their environment. The total cost of the system remains below $400, which represents around 1% of the price of the JACO arm. This low-cost solution proves that the efficiency of assistive robotic arms can easily be increased without necessarily increasing the global cost in a significant manner.

5.7 References

1. Exact Dynamics. http://www.exactdynamics.nl/site/. Accessed 2017
2. Barrett Technology LCC: WAM Arm. http://www.barrett.com/products-arm.htm. Accessed 2017
3. DLR Robotics and Mechatronics Center: Light-Weight Robotics. http://www.dlr.de/rmc/rm/en/desktopdefault.aspx/tabid-3803/6175_read-8963/. Accessed 2017
4. Kinova Robotics: Assistive robotics. http://www.kinovarobotics.com/assistive-robotics/products/robot-arms/. Accessed 2017
5. Routhier, F., Archambault, P., Cyr, M.-C., Maheu, V., Lemay, M., Gélinas, I.: BENEFITS OF JACO ROBOTIC ARM ON INDEPENDENT LIVING AND SOCIAL PARTICIPATION: AN EXPLORATORY STUDY.
6. Maheu, V., Archambault, P.S., Frappier, J., Routhier, F.: Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities. In: Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on 2011, pp. 1-5. IEEE
7. Chung, C.-S., Wang, H., Cooper, R.A.: Functional assessment and performance evaluation for assistive robotic manipulators: Literature review. The journal of spinal cord medicine **36**(4), 273-289 (2013).
8. Groothuis, S.S., Stramigioli, S., Carloni, R.: Lending a helping hand: Toward novel assistive robotic arms. IEEE robotics & automation magazine **20**(1), 20-29 (2013).
9. Campeau-Lecours, A., Maheu, V., Lepage, S., Lamontagne, H., Latour, S., Paquet, L., Hardie, N.: JACO Assistive Robotic Device: Empowering People With Disabilities Through Innovative Algorithms. In: Rehabilitation Engineering and Assistive Technology Society of North America (RESNA) annual conference 2016
10. Ka, H., Ding, D., Cooper, R.A.: ARoMA-V2: Assistive Robotic Manipulation Assistance with Computer Vision and Voice Recognition. In: The 9th Conference on Rehabilitation Engineering and Assistive Technology Society of Korea 2015. RESKO
11. Ka, H., Ding, D., Cooper, R.A.: Three Dimentional Computer Vision-Based Alternative Control Method for Assistive Robotic Manipulator. Symbiosis **1**(1) (2016).

12. Bousquet-Jette, C., Achiche, S., Beaini, D., Cio, Y.L.-K., Leblond-Ménard, C., Raison, M.: Fast scene analysis using vision and artificial intelligence for object prehension by an assistive robot. Engineering Applications of Artificial Intelligence **63**, 33-44 (2017).
13. Kim, D.-J., Lovelett, R., Behal, A.: Eye-in-hand stereo visual servoing of an assistive robot arm in unstructured environments. In: Robotics and Automation, 2009. ICRA'09. IEEE International Conference on 2009, pp. 2326-2331. IEEE
14. Jiang, H., Wachs, J.P., Duerstock, B.S.: Integrated vision-based robotic arm interface for operators with upper limb mobility impairments. In: Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on 2013, pp. 1-6. IEEE
15. Jiang, H., Zhang, T., Wachs, J.P., Duerstock, B.S.: Enhanced control of a wheelchair-mounted robotic manipulator using 3-D vision and multimodal interaction. Computer Vision and Image Understanding **149**, 21-31 (2016).
16. Ristic-Durrant, D., Grigorescu, S.M., Graser, A., Cojbasic, Z., Nikolic, V.: Robust stereo-vision based 3d object reconstruction for the assistive robot friend. Advances in Electrical and Computer Engineering **11**(4), 15-22 (2011).
17. Tsui, K., Yanco, H., Kontak, D., Beliveau, L.: Development and evaluation of a flexible interface for a wheelchair mounted robotic arm. In: Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction 2008, pp. 105-112. ACM
18. Murray, S., Floyd-Jones, W., Qi, Y., Sorin, D., Konidaris, G.: Robot motion planning on a chip. In: Robotics: Science and Systems 2016
19. Quintero, C.P., Ramirez, O., Jägersand, M.: Vibi: Assistive vision-based interface for robot manipulation. In: Robotics and Automation (ICRA), 2015 IEEE International Conference on 2015, pp. 4458-4463. IEEE
20. Kim, D.-J., Lovelett, R., Behal, A.: An empirical study with simulated adl tasks using a vision-guided assistive robot arm. In: Rehabilitation Robotics, 2009. ICORR 2009. IEEE International Conference on 2009, pp. 504-509. IEEE
21. Tanriverdi, V., Jacob, R.J.: Interacting with eye movements in virtual environments. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems 2000, pp. 265-272. ACM

22. Leroux, M., Raison, M., Adadja, T., Achiche, S.: Combination of eyetracking and computer vision for robotics control. In: Technologies for Practical Robot Applications (TePRA), 2015 IEEE International Conference on 2015, pp. 1-6. IEEE
23. THEEYETRIBE: Eye Tracking 101. https://s3.eu-central-1.amazonaws.com/theeyetribe.com/theeyetribe.com/dev/general/index.html. Accessed 2017
24. Atienza, R., Zelinsky, A.: Intuitive interface through active 3D gaze tracking. In: Active Media Technology, 2005.(AMT 2005). Proceedings of the 2005 International Conference on 2005, pp. 16-21. IEEE
25. Buckley, M., Vaidyanathan, R., Mayol-Cuevas, W.: Sensor suites for assistive arm prosthetics. In: Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on 2011, pp. 1-6. IEEE
26. Klauer, C., Schauer, T., Reichenfelser, W., Karner, J., Zwicker, S., Gandolla, M., Ambrosini, E., Ferrante, S., Hack, M., Jedlitschka, A.: Feedback control of arm movements using Neuro-Muscular Electrical Stimulation (NMES) combined with a lockable, passive exoskeleton for gravity compensation. Biosignal Processing and Computational Methods to Enhance Sensory Motor Neuroprosthetics, 44 (2016).
27. Webb, J.D.: Gaze control for remote robotics. Colorado School of Mines (2016)
28. OpenCV. http://opencv.org/. Accessed 2017
29. e-con Systems Your Product Development Partner: Tara - USB 3.0 Stereo Vision Camera. https://www.e-consystems.com/3D-USB-stereo-camera.asp. Accessed 2017
30. Microsoft: Kinect hardware. https://developer.microsoft.com/en-us/windows/kinect/hardware. Accessed 2017
31. Pluzhnikov, S.: Motion Planning and Control of Robot Manipulators. Institutt for teknisk kybernetikk (2012)
32. Zohaib, M., Pasha, M., Riaz, R., Javaid, N., Ilahi, M., Khan, R.: Control strategies for mobile robot with obstacle avoidance. arXiv preprint arXiv:1306.1144 (2013).
33. Yang, L., Qi, J., Song, D., Xiao, J., Han, J., Xia, Y.: Survey of Robot 3D Path Planning Algorithms. Journal of Control Science and Engineering **2016**, 5 (2016).
34. Solomon, M.D.: Development of a Real-Time Hierarchical 3D Path Planning Algorithm for Unmanned Aerial Vehicles. (2016)

35. Cutler, R., Davis, L.: Look who's talking: Speaker detection using video and audio correlation. In: Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on 2000, pp. 1589-1592. IEEE

36. Huang, J., Potamianos, G., Connell, J., Neti, C.: Audio-visual speech recognition using an infrared headset. Speech Communication **44**(1), 83-96 (2004).

37. Johnson, M., Lapkin, S., Long, V., Sanchez, P., Suominen, H., Basilakis, J., Dawson, L.: A systematic review of speech recognition technology in health care. BMC medical informatics and decision making **14**(1), 94 (2014).

38. Deshpande, S., Shriram, R.: Real time text detection and recognition on hand held objects to assist blind people. In: Automatic Control and Dynamic Optimization Techniques (ICACDOT), International Conference on 2016, pp. 1020-1024. IEEE

39. Tomas, W., Anders, B.: Semantic and Verbatim Word Spotting using Deep Neural Networks. In: International Conference on Frontiers in Handwriting Recognition (ICFHR), October 23-26, 2016, Shenzhen, China. 2016

40. Aksoy, E.E., Ovchinnikova, E., Orhan, A., Yang, Y., Asfour, T.: Unsupervised Linking of Visual Features to Textual Descriptions in Long Manipulation Activities. IEEE Robotics and Automation Letters **2**(3), 1397-1404 (2017).

41. Fan, J., Chen, T., Lu, S.: Unsupervised Feature Learning for Land-Use Scene Recognition. IEEE Transactions on Geoscience and Remote Sensing (2017).

42. Franco, A., Maltoni, D., Papi, S.: Grocery product detection and recognition. Expert Systems with Applications (2017).

43. Wang, Y., Cen, Y., Zhao, R., Cen, Y., Hu, S., Voronin, V., Wang, H.: Separable vocabulary and feature fusion for image retrieval based on sparse representation. Neurocomputing **236**, 14-22 (2017).

44. Zhang, W., Yan, J., Shi, W., Feng, T., Deng, D.: Refining deep convolutional features for improving fine-grained image recognition. EURASIP Journal on Image and Video Processing **2017**(1), 27 (2017).

45. Sibert, L.E., Jacob, R.J.: Evaluation of eye gaze interaction. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems 2000, pp. 281-288. ACM

46. Jacob, R., Karn, K.S.: Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. Mind **2**(3), 4 (2003).

47. Al-Rahayfeh, A., Faezipour, M.: Eye tracking and head movement detection: A state-of-art survey. IEEE journal of translational engineering in health and medicine **1**, 2100212-2100212 (2013).

48. Chennamma, H., Yuan, X.: A survey on eye-gaze tracking techniques. arXiv preprint arXiv:1312.6410 (2013).

49. Sharma, A., Abrol, P.: Eye gaze techniques for human computer interaction: A research survey. International Journal of Computer Applications **71**(9) (2013).

50. Bazrafkan, S., Kar, A., Costache, C.: Eye Gaze for Consumer Electronics: Controlling and commanding intelligent systems. IEEE Consumer Electronics Magazine **4**(4), 65-71 (2015).

51. Chandra, S., Sharma, G., Malhotra, S., Jha, D., Mittal, A.P.: Eye tracking based human computer interaction: Applications and their uses. In: Man and Machine Interfacing (MAMI), 2015 International Conference on 2015, pp. 1-5. IEEE

52. Fatima, R., Usmani, A., Zaheer, Z.: Eye movement based human computer interaction. In: Recent Advances in Information Technology (RAIT), 2016 3rd International Conference on 2016, pp. 489-494. IEEE

53. Octomap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees. https://octomap.github.io/. Accessed 2017

54. Hornung, A., Wurm, K.M., Bennewitz, M., Stachniss, C., Burgard, W.: OctoMap: An efficient probabilistic 3D mapping framework based on octrees. Autonomous Robots **34**(3), 189-206 (2013).

CHAPITRE 6 DISCUSSION GÉNÉRALE

Le système PoGARA présenté dans ce mémoire a été capable de mettre en valeur le potentiel d'un suiveur de regard en tant que technologie d'IHM pour les personnes ayant des problèmes de motricités sévères des membres supérieurs. Bien que le système PoGARA permet d'aller chercher un objet en présence d'obstacle, il y a certaines limitations importantes qui restent à résoudre. Les plus importantes sont la surexposition à la lumière, la direction de préhension d'un objet et le traitement du nuage 3D.

Premièrement, la surexposition à la lumière de la stéréo vision par une source comme le soleil rend le système complètement inutilisable. En effet, si la stéréo vision est « aveuglé », c'est-à-dire qu'elle n'est pas en mesure de capter des informations de la scène, alors il sera impossible pour l'utilisateur d'indiquer l'objet d'intérêt et le BRA ne sera pas en mesure de connaître son environnement de travail. Pour pallier à ce problème, il faut activer le contrôle de l'exposition de la stéréo vision ainsi que rajouter un filtre réduisant les rayons infrarouges et ultraviolets aux lentilles. Pour le suiveur de regard, le Eye tribe utilise des rayons infrarouges afin de suivre nos yeux. Par conséquent, en présence des rayons du soleil, le capteur infrarouge de l'Eye tribe se sature et le dispositif n'est plus capable de suivre nos yeux. Donc, l'utilisateur ne peut plus contrôler le BRA avec les yeux. Pour contourner ce problème, il faudrait utiliser un suiveur de regard par caméra ou stéréo caméra avec des filtres coupant les rayons infrarouges et ultraviolets. De cette manière, il sera possible de contrôler le BRA même avec une surexposition de lumières.

Deuxièmement, il n'y a pas de façon pour choisir la direction de préhension d'un objet. En ce moment, le système PoGARA ne peut pas prendre d'objet par le haut. En effet, PoGARA peut prendre un objet par le côté uniquement. Quand la scène a beaucoup d'objets rapprochés les uns les autres, une approche par le haut serait plus adéquate, car le poignet du BRA a moins de chance d'accrocher des obstacles. Par conséquent, l'arbre de décision présenté dans [50] est une solution simple et efficace pour prendre cette décision.

Troisièmement, un traitement du nuage 3D plus approfondi aiderait beaucoup pour la préhension de l'objet. En ce moment, le suivi de regard permet de choisir un nœud et son centre sera la position finale de l'effecteur du BRA. Par conséquent, si le haut d'un objet est choisi, le BRA prendra le haut de l'objet, ce qui n'est pas idéal. En effet, la préhension d'un objet est à son meilleur lorsqu'il est pris à son centre de masse [50]. Le traitement du nuage 3D présenté dans [50] devrait être

appliqué au nuage 3D obtenu par le système PoGARA. L'auteur de l'article présente un traitement de nuage 3D permettant d'identifier et de regrouper tous les voxels d'un même objet et de trouver son centre de masse. Néanmoins, l'auteur utilise une Kinect V2 qui donne des nuages 3D de meilleure qualité que la stéréo vision, par conséquent, il n'est pas possible d'utiliser exactement la même méthode; il faudra l'adapter au système PoGARA.

En se référant au cahier des charges présenté précédemment au Chapitre 4, il est possible de remarquer que le système PoGARA remplit les besoins essentiels. En effet, le système est capable d'obtenir un nuage de point 3D de la scène, de sélectionner un objet et d'attraper l'objet d'intérêt sans faire tomber des obstacles. Par contre, le système n'est pas en mesure de subvenir aux besoins facultatifs. Le système PoGARA ne peut pas fonctionner à l'extérieur; il est aussi incapable d'identifier des objets et de leur associer des actions possibles. En ce qui concerne les contraintes du projet, le coût ajouté au BRA est seulement de 349$ US. Par conséquent, la contrainte principale du projet, qui était de minimiser les coûts ajoutés, est respectée. Par contre, la contrainte secondaire concernant l'intégration discrète du système PoGARA sur un fauteuil roulant serait facilement atteinte car le suiveur du regard et la caméra choisis sont de petite dimension.

Les objectifs du projet présenté au Chapitre 3 ont été atteints de manière générale. Avec les résultats obtenus, il est possible de dire que l'objectif général ainsi que les hypothèses du projet ont été atteints puisque le taux de réussite des essais est de 80% et que le temps moyen pour exécuter l'action est de 13.7 secondes sans obstacles, 15.3 secondes avec un obstacle et 16.3 secondes avec deux obstacles. En ce qui concerne les objectifs plus spécifiques, l'intégration du système en entier a été réussie. En effet, le système est en mesure d'obtenir des informations de la scène, de comprendre et traduire l'intention de l'utilisateur et d'aller chercher l'objet d'intérêt. Le seul objectif spécifique qui n'a pas été complètement atteint est le traitement de la scène comme il a été mentionné plus tôt.

CHAPITRE 7 CONCLUSION ET RECOMMENDATIONS

En somme, ce mémoire présente un nouveau système d'assistance, PoGARA, qui combine une stéréo vision artificielle et un suiveur de regard pour le contrôle d'un BRA. Ce système a pu démontrer le potentiel du suiveur de regard dans les technologies d'assistances dans le but d'améliorer la qualité de vie et l'autonomie des patients atteints d'un problème de motricité sévère des membres supérieurs. En effet, pour l'action d'approche et de préhension d'un objet, le système PoGARA a un taux de réussite de 80 % et un temps d'exécution de 13.7 secondes sans obstacles, 15.3 secondes avec un obstacle et 16.3 secondes avec deux obstacles. Sachant que les tâches quotidiennes commencent généralement par une action d'approche et de préhension d'un objet, ces résultats sont prometteurs pour l'amélioration de la technologie des BRAs en termes de temps d'exécution et de taux de réussite. Néanmoins, le système présente des limitations importantes pour l'adoption de la technologie dans la vie de tous les jours. Pour commencer, la lumière du soleil aveugle la stéréo vision empêchant l'acquisition de la scène 3D. Pour remédier à ce problème, il a été suggéré d'installer des filtres coupant les rayons infrarouges et ultraviolets. De plus, la lumière du soleil sature le suiveur de regard rendant la détection des yeux de l'utilisateur impossible. La solution à ce problème est de changer le suiveur de regard par un autre qui n'utilise pas les rayons infrarouges pour trouver les yeux de l'utilisateur. Une autre limitation importante est que le système prend les objets uniquement par le côté. Par conséquent, s'il y a beaucoup d'objets dans la scène, il ne sera pas possible d'aller en chercher un sans faire tomber les autres. Pour pallier cette problématique, il est recommandé de choisir l'orientation de l'effecteur de l'ARA en se basant sur le nombre d'obstacles et leurs distances par rapport à l'objet d'intérêt.

Finalement, les futurs travaux du système PoGARA devraient être sur le traitement du nuage 3D principalement. Effectivement, segmenter le nuage 3D de la scène afin de rassembler les voxels d'un même objet améliorait la planification de trajectoire et l'évitement d'obstacle ainsi que la préhension de l'objet. En rassemblant les voxels d'un même objet, il sera possible de faire un meilleur filtrage du bruit, par conséquent, le planificateur de trajectoire ne prendra pas un voxel de bruit comme un obstacle. De plus, la préhension d'objet sera plus facile puisque le centre de masse de l'objet pourra être calculé et le BRA aura une meilleure prise de l'objet d'intérêt.

RÉFÉRENCES

[1] Exact Dynamics. (2017). Available: http://www.exactdynamics.nl/site/

[2] Barrett Technology LCC. (2017). *WAM Arm*. Available: http://www.barrett.com/products-arm.htm

[3] DLR Robotics and Mechatronics Center. (2017). *Light-Weight Robotics*. Available: http://www.dlr.de/rmc/rm/en/desktopdefault.aspx/tabid-3803/6175_read-8963/

[4] Kinova Robotics. (2017). *Assistive robotics*. Available: http://www.kinovarobotics.com/assistive-robotics/products/robot-arms/

[5] FANUC corporation. Available: http://www.fanuc.co.jp/eindex.html

[6] K. AG. Available: https://www.kuka.com/

[7] Assistive innovations. (2017). *Robotic arms*. Available: https://assistive-innovations.com/en/robotic-arms/

[8] G. Hirzinger *et al.*, "DLR's torque-controlled light weight robot III-are we reaching the technological limits now?," in *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, 2002, vol. 2, pp. 1710-1716: IEEE.

[9] Microsoft. (2017). *Kinect hardware*. Available: https://developer.microsoft.com/en-us/windows/kinect/hardware

[10] DUO 3D. (2017). *DUO MLX*. Available: https://duo3d.com/product/duo-minilx-lv1#tab=specs

[11] I.D.E.A. (2017). *PrimeSense 3D sensor*. Available: http://www.2-bit-toys.com/category/gear/

[12] H.-H. Wu and A. Bainbridge-Smith, "Advantages of using a Kinect Camera in various applications," *University of Canterbury*, 2011.

[13] Robotshop. (2017). *Cameras & Vision Sensors*. Available: http://www.robotshop.com/en/cameras-vision-sensors.html?p=1

[14] M. Leroux, M. Raison, T. Adadja, and S. Achiche, "Combination of eyetracking and computer vision for robotics control," in *Technologies for Practical Robot Applications (TePRA), 2015 IEEE International Conference on*, 2015, pp. 1-6: IEEE.

[15] Automation24. (2017). *Ultrasonic sensor microsonic mic+600/D/TC*. Available: http://www.automation24.com/position-sensors/ultrasonic-sensor-microsonic-mic600/d/tc-i7-187-0.htm

[16] R. Jafri, S. A. Ali, H. R. Arabnia, and S. Fatima, "Computer vision-based object recognition for the visually impaired in an indoors environment: a survey," *The Visual Computer,* vol. 30, no. 11, pp. 1197-1222, 2014.

[17] e-con Systems Your Product Development Partner. (2017). *Tara - USB 3.0 Stereo Vision Camera*. Available: https://www.e-consystems.com/3D-USB-stereo-camera.asp

[18] S. Pluzhnikov, "Motion Planning and Control of Robot Manipulators," Institutt for teknisk kybernetikk, 2012.

[19] M. Zohaib, M. Pasha, R. Riaz, N. Javaid, M. Ilahi, and R. Khan, "Control strategies for mobile robot with obstacle avoidance," *arXiv preprint arXiv:1306.1144,* 2013.

[20] L. Yang, J. Qi, D. Song, J. Xiao, J. Han, and Y. Xia, "Survey of Robot 3D Path Planning Algorithms," *Journal of Control Science and Engineering,* vol. 2016, p. 5, 2016.

[21] M. D. Solomon, "Development of a Real-Time Hierarchical 3D Path Planning Algorithm for Unmanned Aerial Vehicles," 2016.

[22] R. Cutler and L. Davis, "Look who's talking: Speaker detection using video and audio correlation," in *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*, 2000, vol. 3, pp. 1589-1592: IEEE.

[23] J. Huang, G. Potamianos, J. Connell, and C. Neti, "Audio-visual speech recognition using an infrared headset," *Speech Communication,* vol. 44, no. 1, pp. 83-96, 2004.

[24] M. Johnson *et al.*, "A systematic review of speech recognition technology in health care," *BMC medical informatics and decision making,* vol. 14, no. 1, p. 94, 2014.

[25] S. Deshpande and R. Shriram, "Real time text detection and recognition on hand held objects to assist blind people," in *Automatic Control and Dynamic Optimization Techniques (ICACDOT), International Conference on*, 2016, pp. 1020-1024: IEEE.

[26] W. Tomas and B. Anders, "Semantic and Verbatim Word Spotting using Deep Neural Networks," in *International Conference on Frontiers in Handwriting Recognition (ICFHR), October 23-26, 2016, Shenzhen, China.*, 2016.

[27] E. E. Aksoy, E. Ovchinnikova, A. Orhan, Y. Yang, and T. Asfour, "Unsupervised Linking of Visual Features to Textual Descriptions in Long Manipulation Activities," *IEEE Robotics and Automation Letters,* vol. 2, no. 3, pp. 1397-1404, 2017.

[28] J. Fan, T. Chen, and S. Lu, "Unsupervised Feature Learning for Land-Use Scene Recognition," *IEEE Transactions on Geoscience and Remote Sensing,* 2017.

[29] A. Franco, D. Maltoni, and S. Papi, "Grocery product detection and recognition," *Expert Systems with Applications,* 2017.

[30] Y. Wang *et al.*, "Separable vocabulary and feature fusion for image retrieval based on sparse representation," *Neurocomputing,* vol. 236, pp. 14-22, 2017.

[31] W. Zhang, J. Yan, W. Shi, T. Feng, and D. Deng, "Refining deep convolutional features for improving fine-grained image recognition," *EURASIP Journal on Image and Video Processing,* vol. 2017, no. 1, p. 27, 2017.

[32] L. E. Sibert and R. J. Jacob, "Evaluation of eye gaze interaction," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2000, pp. 281-288: ACM.

[33] R. Jacob and K. S. Karn, "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises," *Mind,* vol. 2, no. 3, p. 4, 2003.

[34] A. Al-Rahayfeh and M. Faezipour, "Eye tracking and head movement detection: A state-of-art survey," *IEEE journal of translational engineering in health and medicine,* vol. 1, pp. 2100212-2100212, 2013.

[35] H. Chennamma and X. Yuan, "A survey on eye-gaze tracking techniques," *arXiv preprint arXiv:1312.6410,* 2013.

[36] A. Sharma and P. Abrol, "Eye gaze techniques for human computer interaction: A research survey," *International Journal of Computer Applications,* vol. 71, no. 9, 2013.

[37] S. Bazrafkan, A. Kar, and C. Costache, "Eye Gaze for Consumer Electronics: Controlling and commanding intelligent systems," *IEEE Consumer Electronics Magazine,* vol. 4, no. 4, pp. 65-71, 2015.

[38] S. Chandra, G. Sharma, S. Malhotra, D. Jha, and A. P. Mittal, "Eye tracking based human computer interaction: Applications and their uses," in *Man and Machine Interfacing (MAMI), 2015 International Conference on*, 2015, pp. 1-5: IEEE.

[39] R. Fatima, A. Usmani, and Z. Zaheer, "Eye movement based human computer interaction," in *Recent Advances in Information Technology (RAIT), 2016 3rd International Conference on*, 2016, pp. 489-494: IEEE.

[40] OpenCV. (2017). Available: http://opencv.org/

[41] V. Tanriverdi and R. J. Jacob, "Interacting with eye movements in virtual environments," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2000, pp. 265-272: ACM.

[42] THEEYETRIBE. (2017). *Eye Tracking 101*. Available: https://s3.eu-central-1.amazonaws.com/theeyetribe.com/theeyetribe.com/dev/general/index.html

[43] F. Routhier, P. Archambault, M.-C. Cyr, V. Maheu, M. Lemay, and I. Gélinas, "BENEFITS OF JACO ROBOTIC ARM ON INDEPENDENT LIVING AND SOCIAL PARTICIPATION: AN EXPLORATORY STUDY."

[44] V. Maheu, P. S. Archambault, J. Frappier, and F. Routhier, "Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities," in *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on*, 2011, pp. 1-5: IEEE.

[45] C.-S. Chung, H. Wang, and R. A. Cooper, "Functional assessment and performance evaluation for assistive robotic manipulators: Literature review," *The journal of spinal cord medicine,* vol. 36, no. 4, pp. 273-289, 2013.

[46] S. S. Groothuis, S. Stramigioli, and R. Carloni, "Lending a helping hand: Toward novel assistive robotic arms," *IEEE robotics & automation magazine*, vol. 20, no. 1, pp. 20-29, 2013.

[47] A. Campeau-Lecours *et al.*, "JACO Assistive Robotic Device: Empowering People With Disabilities Through Innovative Algorithms," in *Rehabilitation Engineering and Assistive Technology Society of North America (RESNA) annual conference*, 2016.

[48] H. Ka, D. Ding, and R. A. Cooper, "ARoMA-V2: Assistive Robotic Manipulation Assistance with Computer Vision and Voice Recognition," in *The 9th Conference on Rehabilitation Engineering and Assistive Technology Society of Korea*, 2015: RESKO.

[49] H. Ka, D. Ding, and R. A. Cooper, "Three Dimentional Computer Vision-Based Alternative Control Method for Assistive Robotic Manipulator," *Symbiosis*, vol. 1, no. 1, 2016.

[50] C. Bousquet-Jette, S. Achiche, D. Beaini, Y. L.-K. Cio, C. Leblond-Ménard, and M. Raison, "Fast scene analysis using vision and artificial intelligence for object prehension by an assistive robot," *Engineering Applications of Artificial Intelligence*, vol. 63, pp. 33-44, 2017.

[51] D.-J. Kim, R. Lovelett, and A. Behal, "Eye-in-hand stereo visual servoing of an assistive robot arm in unstructured environments," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, 2009, pp. 2326-2331: IEEE.

[52] H. Jiang, J. P. Wachs, and B. S. Duerstock, "Integrated vision-based robotic arm interface for operators with upper limb mobility impairments," in *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*, 2013, pp. 1-6: IEEE.

[53] H. Jiang, T. Zhang, J. P. Wachs, and B. S. Duerstock, "Enhanced control of a wheelchair-mounted robotic manipulator using 3-D vision and multimodal interaction," *Computer Vision and Image Understanding*, vol. 149, pp. 21-31, 2016.

[54] D. Ristic-Durrant, S. M. Grigorescu, A. Graser, Z. Cojbasic, and V. Nikolic, "Robust stereo-vision based 3d object reconstruction for the assistive robot friend," *Advances in Electrical and Computer Engineering*, vol. 11, no. 4, pp. 15-22, 2011.

[55] K. Tsui, H. Yanco, D. Kontak, and L. Beliveau, "Development and evaluation of a flexible interface for a wheelchair mounted robotic arm," in *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, 2008, pp. 105-112: ACM.

[56] S. Murray, W. Floyd-Jones, Y. Qi, D. Sorin, and G. Konidaris, "Robot motion planning on a chip," in *Robotics: Science and Systems*, 2016.

[57] C. P. Quintero, O. Ramirez, and M. Jägersand, "Vibi: Assistive vision-based interface for robot manipulation," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, 2015, pp. 4458-4463: IEEE.

[58] D.-J. Kim, R. Lovelett, and A. Behal, "An empirical study with simulated adl tasks using a vision-guided assistive robot arm," in *Rehabilitation Robotics, 2009. ICORR 2009. IEEE International Conference on*, 2009, pp. 504-509: IEEE.

[59] R. Atienza and A. Zelinsky, "Intuitive interface through active 3D gaze tracking," in *Active Media Technology, 2005.(AMT 2005). Proceedings of the 2005 International Conference on*, 2005, pp. 16-21: IEEE.

[60] M. Buckley, R. Vaidyanathan, and W. Mayol-Cuevas, "Sensor suites for assistive arm prosthetics," in *Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on*, 2011, pp. 1-6: IEEE.

[61] C. Klauer *et al.*, "Feedback control of arm movements using Neuro-Muscular Electrical Stimulation (NMES) combined with a lockable, passive exoskeleton for gravity compensation," *Biosignal Processing and Computational Methods to Enhance Sensory Motor Neuroprosthetics*, p. 44, 2016.

[62] J. D. Webb, "Gaze control for remote robotics," Colorado School of Mines, 2016.

[63] Octomap. (2017). *An Efficient Probabilistic 3D Mapping Framework Based on Octrees*. Available: https://octomap.github.io/

[64] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189-206, 2013.