

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

**MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE**

ECOLE NATIONALE POLYTECHNIQUE



Département d'Automatique

Master en Automatique

Thème :

**Construction 3D et Cartographies en utilisant un capteur
RGB-D**

Réalisé et présenté par:

Yahoui Anissa

Encadré par :

Pr .H.Chekireb

Soutenu le 24/06/2015

Juin 2015

Ecole Nationale Polytechnique, 10, AV. Hassen Badi, El-Harrach, Algérie

Remerciements

Louange à Allah le tout puissant qui m'a accordé le droit chemin, le savoir et l'opportunité de poursuivre nos études et la force pour réaliser ce modeste travail.

Au terme de notre cycle d'études, il me paraît opportun de nous acquitter d'un devoir noble, celui de remercier tous ceux qui ont contribué par leur assistance tant morale que physique à mon cursus universitaire et à la réalisation de ce mémoire.

Je tiens à exprimer mes vifs remerciements à mon encadreur :

Mr El-Hachemi Chekireb (ENP)

Pour son encadrement et ses conseils scientifiques

*Je remercie en particulier **Mm Chaib Khaoula** chercheur au (CDTA) pour son aide et sa bonne humeur .*

Je remercie les membres du jury d'avoir bien voulu étudier ce travail et participer à la commission d'examen afin de le juger.

Que tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste travail trouvent ici l'expression de notre sincère gratitude

Merci infiniment à tous.

ملخص :

هذه المهمة هي تحديد موقع نظام الروبوت عن طريق حساب موقفها في البيئة التي تعمل فيها. للقيام بذلك، عن طريق استخدام نظام الرؤية للحصول على صور RGB وعمق كل بكسل من الصورة عين كاميرا D-RGB، ومكس وظيفة والبناء ويحقق مشاهد رسم الخرائط D3 التي لاحظها الروبوت المتحرك.

الكلمات المفتاحية :

RGB-D الكاميرا ، البناء3D ، الموقع ، الروبوت المحمول، وظيفة مكس ، رسم الخرائط

Résumé :

Le travail consiste à localiser, un système robotisé en calculant sa position dans l'environnement dans lequel il évolue. Pour cela, est en utilisant système de vision qui permettant l'acquisition des images RGB et la profondeur de chaque pixel de l'image nommé caméra RGB-D , et des Mex function on réalise la construction 3D et la cartographie des scènes observées par un robot mobile .

Mots Clés :

Caméra RGB-D , Construction 3D , Localisation , Cartographie , robot mobile ,Mex function .

Abstract:

The job is to locate a robot system by calculating its position in the environment in which it operates. To do this, is by using vision system for acquiring RGB images and the depth of each pixel of the image appointed camera RGB-D, and Mex function is carried out the 3D structure and mapping of scenes observed by a mobile robot..

Keywords:

RGB-D camera, 3D Construction, Location, Mapping, Mobile robot, Mex function.

Sommaire

Sommaire

Liste des figures

INTRODUCTION GENERALE	1
Chapitre 1 : La Numérisation 3D	3
1. Introduction	3
2. Taxonomie de la Numérisation 3D	4
2.1 Numérisation 3D Passive	5
2. 2 Numérisation 3D active	6
3. Révolution technologique pour la Numérisation 3D :caméra de profondeur	8
4. Conclusion.....	10
Chapitre 2 : Microsoft Kinect en robotique	12
1. Introduction :	12
2. Description de la Kinect	13
3. Principe de fonctionnement de la Kinect	13
4. Acquisition des images par la Kinect	14
5. Calibrage de la Kinect	16
5.1 Paramètre intrinsèque de la Kinect	17
5.2 Paramètres extrinsèque de la Kinect	19
6. Conclusion.....	20
Chapitre 3 : Construction 3D et technique de plaquage de texture avec la Kinect	21
1. Introduction	21
2. Définition de la Construction 3D	22
3. Construction des nuages 3D et technique de plaquage de texture	22
3.1 Construction du nuage 3D	22
3.2 Plaquage de textures	26
4. Conclusion.....	28
CONCLUSION GENERALE	29
Bibliographie	

Table des figures

Fig 1 : Taxonomie de la Numérisation 3D	4
Fig 2 : Les approches passives et actives	8
Fig 3 : La Kinect sans cache.....	13
Fig 4 : (a) Le modèle IR émis par le projecteur, (b) l'image capturée par la caméra IR, (c) L'image enregistrée sous formes d'un tableau.....	14
Fig 5 : La technique de mesure de la profondeur.....	14
Fig 6 : L'acquisition des images par la Kinect	15
Fig 7 : Les images acquises par la Kinect (image RGB et de profondeur	15
Fig 8 : Principe du calibrage stéréoscopique.....	16
Fig 9 : Projection perspective.....	17
Fig 10 : Distance Base ligne.....	19
Fig 11 : Organigramme de la construction du nuage 3D	22
Fig 12 : Image de profondeur.....	23
Fig 13 : Le Nuage 3D N_{IR} en rouge (le repère de la caméra IR)	25
Fig 14 : Le Nuage 3D N_{RGB} en bleu c'est dans Le repère caméra RGB (N_{IR} rouge)	26
Fig 15 : Image réelle (RGB), construction 3D avec placage de texture.....	27
Fig 16 : Organigramme de la construction du nuage 3D avec placage de textures	27

INTRODUCTION GENERALE

Depuis plusieurs années, le développement des robots complètement autonomes, est un axe de recherche particulièrement actif dans le domaine de la robotique mobile.

Dans cette optique, plusieurs travaux ont déjà montré des résultats convaincants sur la faisabilité d'une plateforme robotique autonome en utilisant principalement des capteurs laser haut de gamme, dont la commercialisation à grande échelle n'est pas encore envisageable. Il est nécessaire de développer des méthodes alternatives basées sur des capteurs bas coût : les systèmes de vision sont une bonne option, car ils fournissent une information très riche sur l'environnement, qui peut être utilisée pour la navigation.

Une étape fondamentale pour la navigation autonome d'un robot est la localisation. Cette action anodine est effectuée en permanence par les êtres vivants pour se déplacer dans l'espace. Différents mécanismes sensoriels sont utilisés, en particulier la vision, mais aussi l'ouïe, l'odorat, le champ magnétique terrestre pour certains oiseaux migrateurs, un système sonar pour les chauves-souris ou les dauphins, ou encore un compas solaire pour certaines fourmis. Ces systèmes de localisation reposent sur des mécanismes naturels complexes, étudiés en particulier par les biologistes. Ces mécanismes ne sont cependant pas encore compris parfaitement. En robotique mobile, l'idée principale est de reproduire ces mécanismes, en utilisant des capteurs spécifiques, afin d'effectuer une tâche de manière autonome.

La localisation et la construction d'une carte 3D par un robot mobile est une étape indispensable pour n'importe quelle tâche de navigation autonome: en effet avant de planifier un déplacement, il est préférable de connaître sa position par rapport aux obstacles, de savoir si le chemin est navigable, ou simplement de connaître la position du robot pour naviguer vers la destination désirée. L'utilisation de la stéréovision (deux caméras) a été pendant plusieurs années le meilleur moyen de navigation visuelle car elle est capable de reconstruire des scènes 3D à partir des images 2D. Cependant, ces algorithmes de reconstruction nécessitent des algorithmes robustes d'extraction, de mise en correspondance et de triangulation des points clefs. C'est une tâche très difficile et coûteuse en temps de calcul. La caméra RGB-D (Red, Green, Blue and Depth) elle est un nouveau système de vision qui permet l'acquisition des images RGB et la profondeur de chaque pixel de l'image et donc facilite énormément la construction 3D et même le plaquage des textures. Microsoft a conçu une caméra RGB-D à faible coût à base du capteur Kinect (Kinect Sensor).

Problématique et Objectif :

Pour se localiser, un système robotisé doit être capable de calculer sa position et de percevoir l'environnement dans lequel il évolue. Pour cela, il est nécessaire que le robot dispose de nombreux capteurs mesurant aussi bien son état interne que l'environnement dans lequel il évolue. On cherche donc un nouveau système de vision qui permet l'acquisition des images RGB et la profondeur de chaque pixel de l'image et facilite la construction 3D et le plaquage des textures. Ce type de matériel, nommé caméra RGB-D, existe depuis plusieurs années mais elle est excessivement chère pour la plupart des applications robotiques. Cependant un nouveau type de matériel permettant d'atteindre cet objectif a fait son apparition récemment. La Kinect de Microsoft/PrimeSense destinée au monde du jeu vidéo, à la console de jeux vidéo Xbox 360, qui permet de contrôler des jeux vidéo sans utiliser le pad traditionnel, ou la manette . La Kinect représente une véritable révolution dans le domaine de la robotique. Pour une somme modique (99 Euros), elle permet de disposer d'un capteur fournissant à la fois des données 3D et des images.

L'objectif de ce travail est localiser un robot mobile en utilisant une Kinect puis et avec la technique du plaquage on construit une cartographie de la scène observé.

Organisation du manuscrit

Ce manuscrit est divisé en trois principaux chapitres :

1. Dans le premier Chapitre, une classification et un état de l'art sur les différentes méthodes de reconstruction 3D est présenté.
2. Le deuxième Chapitre lui est consacré à notre capteur RGB-D (Kinect) intitulé : Microsoft Kinect et la robotique, ou on décrit la Kinect, on explique son principe de fonctionnement et on détail sa calibration.
3. Dans le troisième Chapitre on détaillera les deux techniques étudiés à savoir : la reconstruction 3D d'une scène à base du capteur RGB-D qui permet au robot la navigation dans l'environnement observé par le capteur , et le plaquage de textures qui se fait par un algorithme Matlab qui permet de cartographier la scène

CHAPITRE 1

La Numérisation 3D

1. Introduction

2. Taxonomie de la Numérisation 3D

2.1 Numérisation 3D Passive

2.2 Numérisation 3D active

3. Révolution technologique pour la Numérisation 3D : caméra de profondeur

4. Conclusion

1. Introduction

Du moment que dans notre étude on s'intéresse à la construction 3D ,on consacra ce chapitre à l'étude des techniques de numérisation pour ce faire Une taxonomie globale de la numérisation 3D est d'abord mise en œuvre. Ensuite, une étude de l'état de l'art de la numérisation 3D passive et active est réalisée.

2. Taxonomie de la Numérisation 3D

La numérisation 3D distingue les techniques qui nécessitent un contact physique entre le système d'acquisition et l'objet à numériser et des techniques capables de récupérer l'information 3D sans contact. La figure présente une taxonomie des techniques de numérisation 3D avec et sans contact [Curless 2000]. Les techniques avec contact sondent le sujet grâce à un contact physique et fournissent une bonne précision. Elles se partagent entre des techniques qui détruisent l'objet pour le numériser en le découpant et des techniques non-destructives à l'exemple du bras articulé de palpation et de la MMT : Machine à Mesurer Tridimensionnelle. Le bras de palpation est un capteur positionné au bout d'un bras que l'on peut déplacer dans toutes les directions. On vient placer le capteur sur l'objet à numériser pour prendre les coordonnées x, y, z dans l'espace de ce point. La reconstruction de l'objet s'obtient par une multiplication des prises de points. Cette technologie est destinée à reconstituer des objets techniques de géométrie simple.

La numérisation 3D par palpation est lente puisque la mesure s'effectue point par point. La technologie MMT utilise une tête de mesure, une table sur laquelle la pièce à mesurer est immobilisée, trois liaisons glissières permettant de positionner la tête de mesure en tout point de l'espace et des règles graduées optiques ou électriques permettant de connaître la position de chacune des glissières. Les techniques de numérisation avec contact constituent des outils de métrologie utilisés essentiellement en industrie pour leur haute précision qui est de l'ordre du micron.

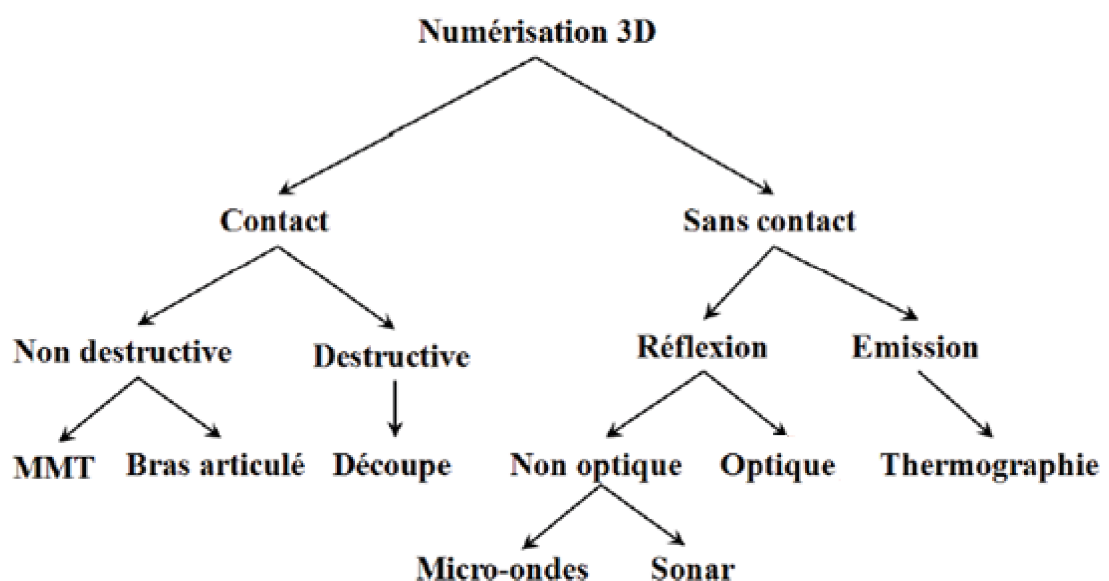


Fig.1 : Taxonomie de la Numérisation 3D

La famille de techniques de numérisation 3D sans contact se base sur le principe de l'émission ou de la réflexion d'ondes. La thermographie récupère l'information 3D en utilisant le principe de l'émission des ondes thermiques en évaluant les performances thermiques de l'objet. Elle est utilisée par exemple pour une localisation précise 3D des sources de déperdition de chaleur dans un bâtiment. Le principe de la réflexion permet aussi de récupérer l'information 3D en analysant le profil d'ondes optiques ou non-optiques réfléchies par la surface de l'objet. Les ondes non-optiques peuvent être des ondes sonar ou des micro-ondes.

La numérisation 3D optique peut être assurée par des méthodes passives ou actives. Les méthodes passives récupèrent l'information 3D en utilisant l'apparence observée de la scène et des objets dans des images sans assistance par une source de lumière. Par contre, les méthodes actives nécessitent la projection d'une lumière sur la surface de l'objet pour le numériser. La figure 2.2 présente une classification non-exhaustive des techniques optiques. Dans les deux sections suivantes, nous décrivons les différentes approches passives et actives.

2.1 Numérisation 3D Passive

La numérisation passive peut être mono-caméra ou multi-caméras, mono-image ou multi-images. Une seule caméra avec plusieurs images successives permet d'estimer la forme 3D d'un objet par une étude du mouvement de l'objet (shape from motion), par focalisation ou dé-focalisation (shape from focus/defocus). Les deux approches « shape from motion » et « shape from focus/defocus » sont ainsi mono-caméra et multi-images. La profondeur peut aussi être retrouvée par une analyse de l'ombrage (shape from shading) qui utilise une seule caméra et une seule image. Autrement, la numérisation passive peut employer deux ou plusieurs caméras pour retrouver la dimension perdue. Dans ce cas, la reconstruction 3D peut se faire par stéréovision, par photogrammétrie ou par une analyse des silhouettes visuelles de l'objet créées par plusieurs acquisitions 2D prises de différents points de vue (shape from silhouettes). Ces approches sont donc multi-caméras et mono-image. Le tableau 2.1 propose les différentes caractéristiques des approches passives. Le critère Nbre de caméras indique le nombre de caméras utilisées. Le critère Nbre images/caméra constitue le nombre d'images utilisées par chaque caméra du système. Le critère Image de profondeur permet de distinguer les approches qui fournissent une image de profondeur comme résultat ou non. Le critère Directe caractérise les approches qui fournissent directement une information de profondeur

ou non. Finalement, le critère Orientation de la surface permet de discerner les approches qui fournissent des cartes d'orientation de la surface comme information de profondeur

	Nbre de caméras	Nbre images/caméra	Image de profondeur	Directe	Orientation de la surface
Approche passive					
Stéréovision	2	1	X	X	
Photogrammétrie	2	1	X	X	
Shape from motion	1	2	X	X	
Shape from focus/defocus	1	2			
Shape from shading	1	1	X		X
Shape from silhouettes	N>2	1	X	X	
Shape from texture	1	1			X

Tableau.1 : Les Approches Passives

2.12. 2 Numérisation 3D active

Le principe de la numérisation active consiste à émettre un rayonnement et détecter sa réflexion afin de sonder un objet ou un environnement. Différents types de source de rayonnement sont utilisés : faisceau laser, lumière blanche, ultrason, rayon X ou lumière infrarouge. L'information 3D se calcule soit par l'estimation du temps de vol des rayons lumineux heurtant la surface de l'objet, par le principe de la triangulation optique ou encore par le principe de l'interférométrie ou de moiré. Temps de vol La technique de temps de vol utilise un télémètre laser pour mesurer le temps nécessaire au trajet aller-retour de l'impulsion d'un faisceau laser réfléchi [Schuon et al. 2009]. Connaissant la vitesse de la lumière c , le temps de retour t détermine la distance parcourue par la lumière est égale à $(c.t) / 2$. Évidemment, la précision de la numérisation 3D dépend de celle de la mesure du temps de retour. Cette technique est rapide mais elle fournit un résultat très bruité ce qui nécessite une étape de post-traitement comme le proposent [Cui et al. 2010] qui utilisent une caméra TOF Swissranger SR4000 et font recours à la super-résolution pour débruiter l'information 3D reconstruite.

Les approches actives fournissent généralement une meilleure précision que les approches passives. Cependant, une numérisation active d'un visage par exemple est souvent intrusive et nécessite une collaboration explicite du sujet. L'utilisation de la projection de lumière est aussi utilisée pour renforcer les approches passives de stéréovision ou de shape from X. En effet, une solution pour réussir la stéréovision pour les objets mal-texturés consiste à projeter une lumière codée pour texturer l'objet pour garantir la convergence du processus d'appariement. [Zhang & Yau 2008] utilisent un système étalonné formé d'un vidéoprojecteur et de deux caméras. Ils projettent trois patrons sinusoïdaux, calculent les valeurs de phases locales pour chaque pixel. Ensuite, ils proposent d'assurer un appariement stéréo sur les cartes 2D de phase gauche et droite au lieu de l'assurer sur les images elles-mêmes. La densité du modèle 3D calculé par la stéréovision active dépend de la richesse du patron texturant et de l'aptitude de la lumière codée à cerner le maximum de détails de la forme. L'utilisation de mesures de similarités spatio-temporelles a été aussi proposée dans [Zhang et al. 2004] pour des meilleurs résultats.

Approche active	Nbre de caméras	Nbre images/caméra	Image de profondeur	Directe	Orientation de la surface
Triangulation laser	1	1	X	X	
Time-Of-Flight	1	1	X	X	
Lumière structurée	1	$N \geq 3$	X	X	
Interférométrie	1	$N \geq 2$	X	X	
Moiré	1	1	X		
Stéréovision active	2	$N \geq 1$	X	X	
Active shape from defocus	1	1	X		
Stéréophotométrie	1	$N \geq 2$			X

Tableau.2 : Les Approches Actives

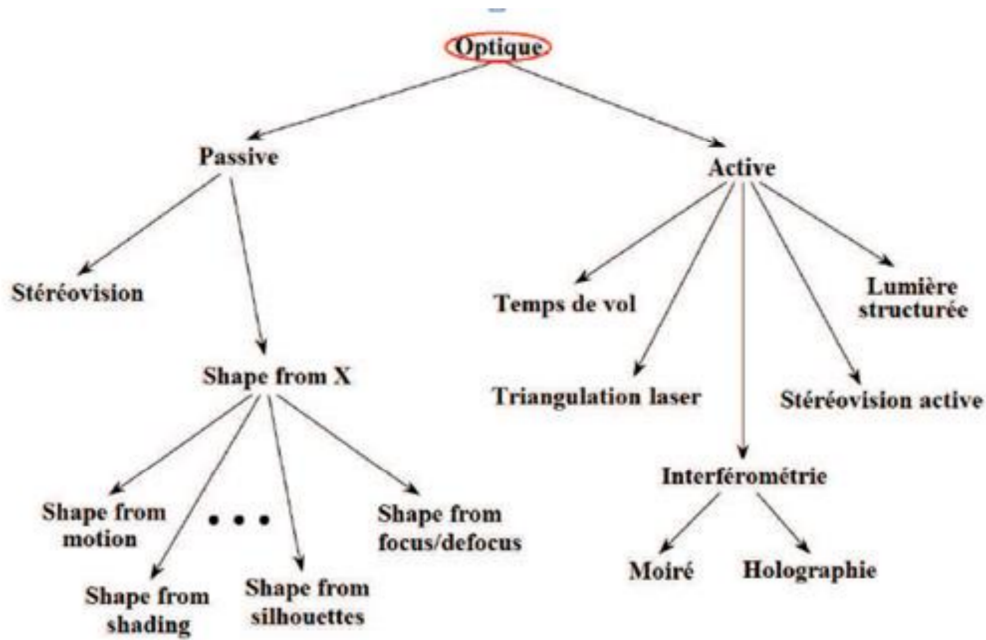


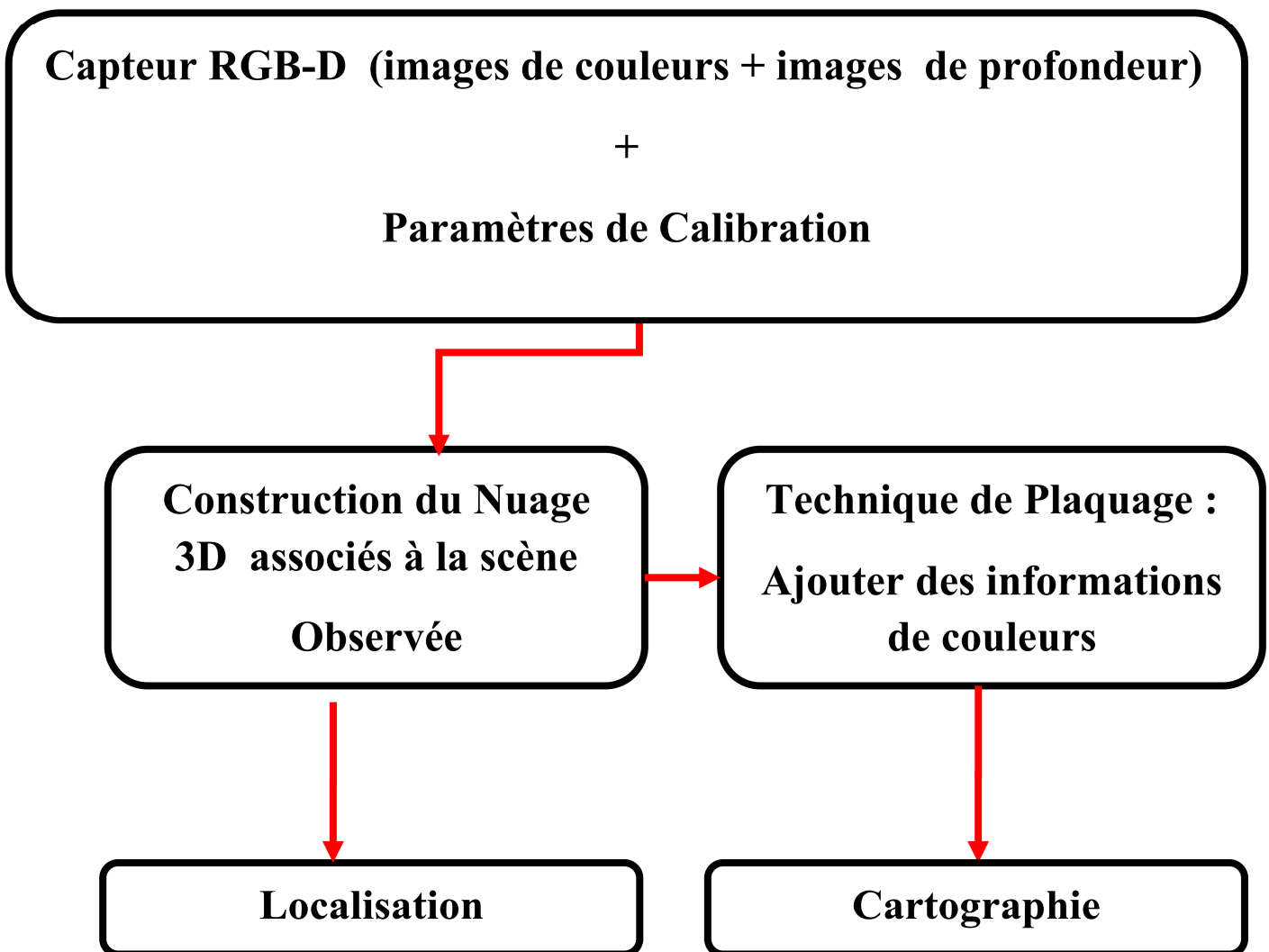
Fig.2 : Les approches passives et actives

3. Révolution technologique pour la Numérisation 3D : caméra de profondeur

La vision est une des solutions aux problèmes de Numérisation grâce à l'intégration d'un système de mesures 3D à une caméra. Ces dernières années un nouveau type de caméra capable d'acquérir des images couleurs et des images de profondeur est apparu qui a engendré une révolution dans le domaine de la vision 3D. Ce type de caméra, nommé caméra RGB-D (Red, Green, Blue-Depth), existe depuis plusieurs années, mais avec un coût prohibitif pour la plupart des applications robotiques. Cependant, à la fin de 2010 un nouveau type de caméra RGB-D à faible coût appelé la Kinect a fait son apparition. Vu les avantages de la Kinect que ce soit dans la technique de détection de profondeur et le coût faible; la société Vision par ordinateur (Microsoft/Prime Sense) a rapidement découvert que cette technologie pourrait servir à d'autres fins comme les jeux vidéo et la robotique, car cette nouvelle catégorie de caméra est dédiée en premier lieu à la détection des mouvements ainsi que les profondeurs des points qui constituent l'environnement ce qui facilite énormément la construction 3D et même le plaquage des textures et donc assurer les deux étapes indispensable pour la navigation autonome d'un robot mobile à savoir la localisation de ce dernier dans la carte générée.

Donc on va exploiter ce nouveau type de capteur pour rendre le robot autonome .

Ceci en mesurant aussi bien son état n'interne que l'environnement dans lequel il évolue. ce choix de capteur est bien évidemment compatible a notre application robotique envisagé a savoir la localisation à travers le calcule de position du système robotique ainsi que cartographier l'environnement pour une éventuelle planification et navigation dans un environnement inconnu.



Organigramme résumant Notre étude

4.Conclusion :

La Numérisation 3D est d'intérêt majeurs dans divers applications , ou on a exposé les différentes méthodes existante ceci dit elle a connu une grande révolution avec l'apparition des caméras de profondeur surtout dans le domaine de la robotique ou la robotique mobile voit une grande évolution dans l'optique de développer des robots autonomes capable de naviguer et cartographier l'environnement observée dans les deux prochains chapitre nous détailleront cette question .

CHAPITRE 2

CHAPITRE 2

Microsoft Kinect en robotique

1. Introduction

2. Description de la Kinect

3. Principe de fonctionnement de la Kinect

4. Acquisition des images par la Kinect

5. Calibrage de la Kinect

5.1 Paramètre intrinsèque de la Kinect

5.2 Paramètres extrinsèque de la Kinect

6. Conclusion

1. Introduction

Les problèmes de la vision 3D sont résolus avec l'arrivée des caméras de profondeur RGB-D (D-IMAGE, Z-Cam, la Kinect de Microsoft), La Kinect représente une véritable révolution en particulier dans le domaine de la robotique. Pour une somme modique, elle permet de disposer d'un capteur fournissant à la fois des données 3D et des images. La caméra RGB-D (Red, Green, Blue and Depth) est un nouveau système de vision qui permet l'acquisition des images RGB et la profondeur de chaque pixel de l'image et donc facilite énormément la construction 3D et même le plaquage des textures. Microsoft a conçu une caméra RGB-D à faible coût à base du capteur Kinect (Kinect Sensor). Pour cela, il est nécessaire que le robot dispose de nombreux capteurs mesurant aussi bien son état interne que l'environnement dans lequel il évolue. Le choix des capteurs dépend bien évidemment de l'application envisagée.

Ce type de matériel, nommé caméra RGB-D, existe depuis plusieurs années mais elle est excessivement chère pour la plupart des applications robotiques. Cependant un nouveau type de matériel permettant d'atteindre cet objectif a fait son apparition récemment. La Kinect de Microsoft/PrimeSense destinée au monde du jeu vidéo, et aux applications robotiques en particulier.

2. Description de la Kinect

Le capteur Kinect est une barre horizontale reliée à une petite base avec un pivot motorisé, conçu pour être placé au-dessus ou en dessous de l'affichage vidéo. Le dispositif comporte une caméra RGB, un capteur de profondeur constitué par un projecteur laser et une caméra infrarouge (IR) et un micro phone.

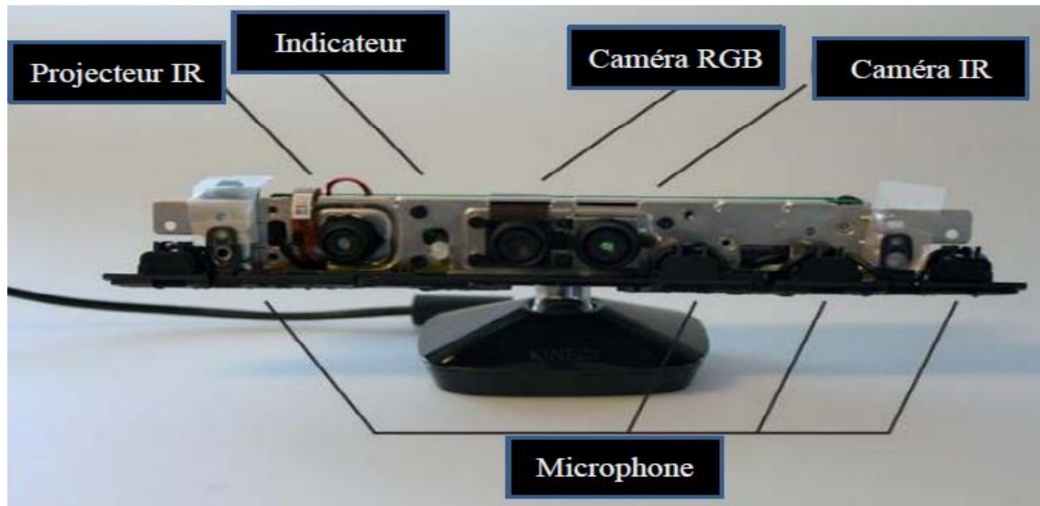


Fig. 3 : La Kinect sans cache

Le capteur de profondeur se compose d'un projecteur laser infrarouge combiné à une caméra IR. C'est un capteur du type CMOS (Complementary Metal Oxide Semiconductor) monochrome, qui capture des données vidéo en 3D.

3. Principe de fonctionnement de la Kinect

Le principe de fonctionnement de la Kinect consiste à projeter un flux laser sur la scène, qui appartient à la gamme infrarouge (figure (a)), et à capturer simultanément une image IR avec un capteur CMOS monochrome équipé d'un filtre IR (figure(b)). En décodant la déformation des flux émis, le processeur de la Kinect utilise la position relative des points dans le modèle pour calculer la profondeur, qui provient de chaque pixel dans l'image figure

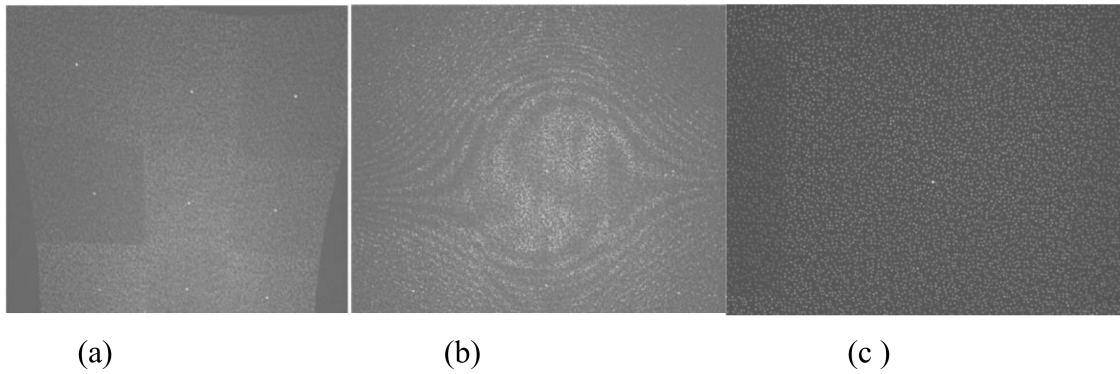


Fig.4 : (a) Le modèle IR émis par le projecteur, (b) l'image capturée par la caméra IR, (c) L'image enregistrée sous formes d'un tableau

Ainsi les valeurs de profondeur réelle sont les estimations de la distance entre l'objet et le plan de la caméra IR figure ci-dessous. De plus, le capteur de profondeur est considéré comme un dispositif qui peut retourner les coordonnées cartésiennes (x, y, z) des objets 3D.

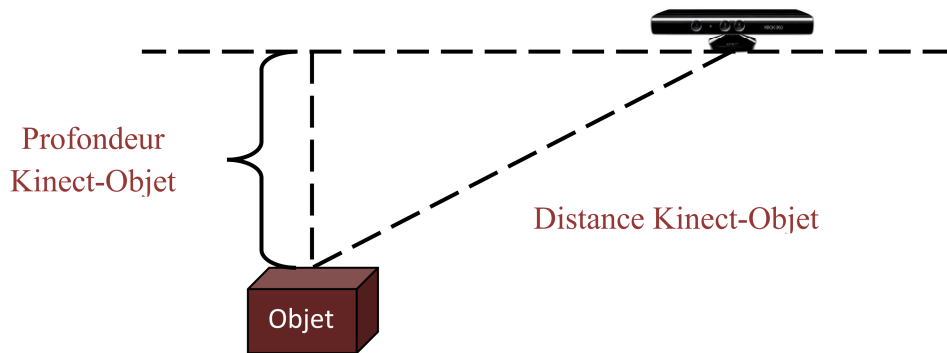


Fig.5 : La technique de mesure de la profondeur.

4. Acquisition des images par la Kinect

nous présentons une solutions soft qui peuvent être utilisées pour lire et afficher les images acquises par la Kinect. Voir figure .

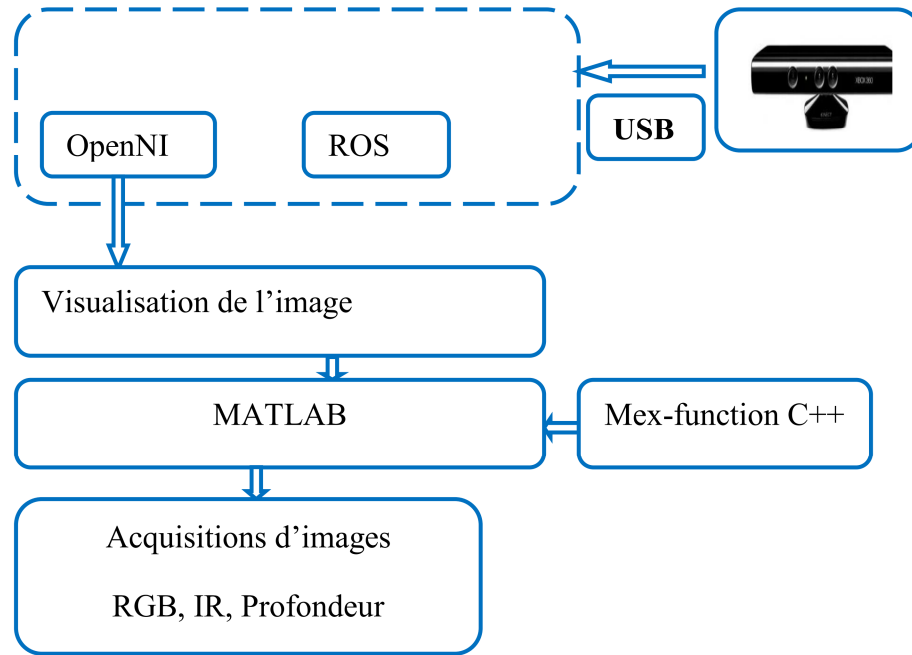


Fig.6 :L'acquisition des images par la Kinect

Pour l'acquisition des images de la Kinect nous avons utilisé OpenNI comme open source car il est le plus simple pour visualiser les images, nous avons utilisé MATLAB comme logiciel de simulation à l'aide des Mex-fonction qui sont compilés en C++ pour acquérir les images IR, RGB et l'image de profondeur qui sont présentés dans la figure ci-dessous :

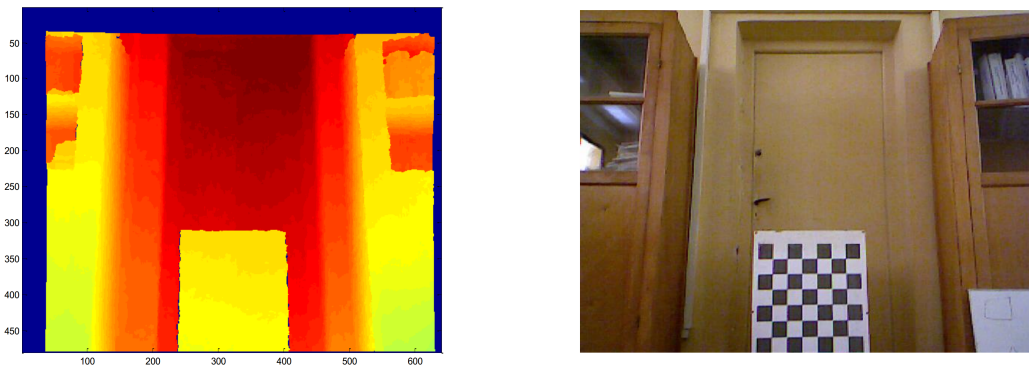


Fig.7 : Les images acquises par la Kinect (image RGB et de profondeur).

5. Calibrage de la Kinect

Pour que la Kinect soit prête à l'utilisation, elle doit être étalonnée (calibrée). Ceci se fait en deux phases : la première consiste à calibrer les deux caméras toutes seules (RGB et IR) pour obtenir les paramètres intrinsèques et la deuxième phase consiste à déterminer les paramètres extrinsèques, en

considérant les deux caméras comme un banc stéréoscopique dans lequel nous nous intéressons à la distance entre les deux caméras..

Calibrage des deux caméras IR et RGB :

Le modèle sténopé modélise une caméra idéale (Projection Perspective Simple). Le système optique induit des distorsions géométriques qui affectent la projection des points. Ainsi, un point projeté dans une image ne correspond pas au modèle sténopé. La distorsion est d'autant plus élevée que le champ de vue de la caméra est grand. Cependant, il est possible de modéliser la distorsion en enrichissant le modèle sténopé par des termes supplémentaires (le modèle devient non linéaire).

Le calibrage permet de déterminer les paramètres intrinsèques et extrinsèques d'une ou plusieurs caméras en utilisant une mire dont le modèle est connu à l'avance. Pour réaliser le calibrage de deux caméras stéréoscopiques, les paramètres à estimer sont :

- **Les paramètres intrinsèques de chaque caméra**
- **Les paramètres extrinsèques** (la matrice de rotation translation R et T de la caméra IR par rapport à la caméra RGB).
- **Les coefficients de distorsion.**

Les paramètres de la
caméra gauche
(caméra RGB)

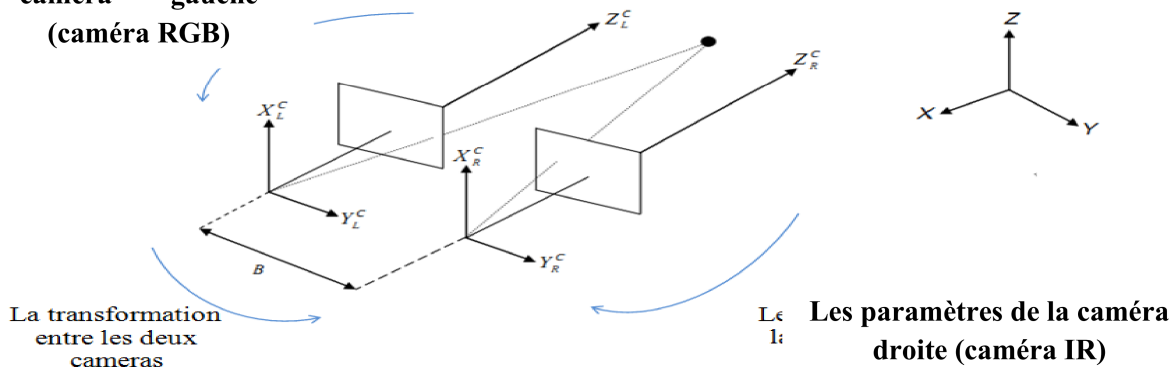


Fig.8 : Principe du calibrage stéréoscopique.

Pour déterminer les paramètres du modèle de la caméra, nous la placerons devant une mire (objet étalon) de calibrage. C'est un ensemble de points dont les coordonnées sont parfaitement connues dans un repère de la mire qui est différent du repère caméra. Chaque point de la mire se projette dans l'image, puis nous mesurons ses coordonnées dans le repère image. Le point Q est défini par ses coordonnées (x_w, y_w, z_w) dans le repère univers. Il est également précisé par ses coordonnées (x_c, y_c, z_c) dans le repère caméra.

5.1 Paramètres intrinsèques

Le changement de repère qui lie les coordonnées du point Q dans le repère image (figure I-23) avec le repère lié à la caméra est défini par les paramètres intrinsèques de la caméra.

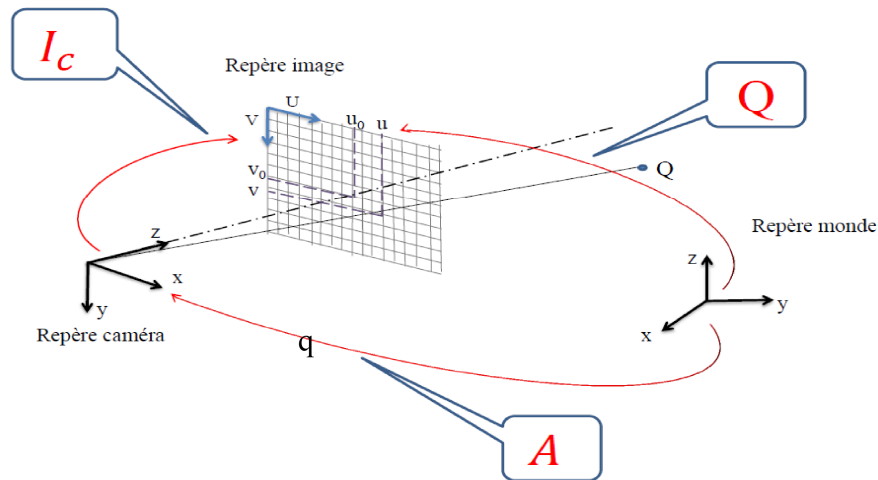


Fig.9 : Projection perspective

Dans le cas où le système de coordonnées lié à la caméra est représenté par le repère $R_c (X_c, Y_c, Z_c)$. En effet, l'origine de ce repère se trouve en C alors que le plan X_c-Y_c est parallèle au plan image, et l'axe Z_c est confondu avec l'axe optique. Soient $(x, y$ et $z)$ les coordonnées d'un point P dans le repère R_c ; dans ce même repère, les coordonnées $(x', y'$ et $z')$ de sa projection psur le plan image sont exprimées comme suit :

$$\begin{cases} x' = f \frac{x}{z} \\ y' = f \frac{y}{z} \\ z' = f \end{cases}$$

En outre ce même point p est mesuré en pixels par les coordonnées (u, v) dans un repère bidimensionnel $U-V$ associé à l'image. La relation entre (u, v) et (x', y') fait intervenir les coordonnées en pixels de C dans le repère image (u_0, v_0) , le facteur d'échelle horizontal k_u et celui vertical k_v sont tous deux exprimés en pixels/mm car les pixels d'une caméra sont rarement carrés. La transformation du repère caméra au repère image s'écrit (pour le point p) :

$$\begin{aligned} u &= -k_u x' + u_0 \\ v &= +k_v y' + v_0 \end{aligned}$$

En remplaçant (x', y') par leurs expressions, nous obtenons :

$$u = -k_u f \frac{x}{z} + u_0$$

$$v = k_v f \frac{y}{z} + v_0$$

En posant $d = z$ on obtient

$$du = -k_u f x + du_0$$

$$dv = k_v f y + dv_0$$

D'où l'écriture matricielle :

$$\begin{pmatrix} du \\ dv \\ d \end{pmatrix} = \begin{pmatrix} -k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

En posant $\alpha_u = -k_u f$ et $\alpha_v = k_v f$ le système prend la forme :

$$\begin{pmatrix} du \\ dv \\ d \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{pmatrix} du \\ dv \\ d \end{pmatrix} = Ic \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ et } Ic = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Donc pour une caméra donnée (c), la transformation perspective entre l'espace projectif et le plan projectif et une application linéaire exprimée par la matrice Ic qui est fonction des paramètres intrinsèques qui sont les facteurs d'échelles de chaque caméra et les paramètres extrinsèques qui sont la matrice de rotation translation et de la caméra IR par rapport à la caméra RGB.

5.2 Paramètres extrinsèques

Les paramètres extrinsèques permettent de définir la transformation géométrique entre le repère caméra et la scène , ils sont représentés par la matrice A .

$$A = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Erreur ! Il n'y a pas de texte répondant à ce style dans ce document.-1)

L'outil Toolbox calib-gui qui nous permet d'obtenir la matrice de rotation et de translation entre les deux repères des caméras ainsi que les positions de la mire par rapport à la caméra pendant le calibrage .

Les résultats de calibrage obtenus sont représentés dans le tableau I-1, ils montrent que la matrice de rotation entre les deux caméras est égale à l'identité. La première composante du vecteur de translation (selon x) est égale à 26.67705 mm , dont on sait a priori (mesure prise manuellement) qu'elle est égale à 26 mm , donc nous pouvons dire que notre estimation des paramètres est juste avec une petite erreur de 0.67705 mm , voir la figure.

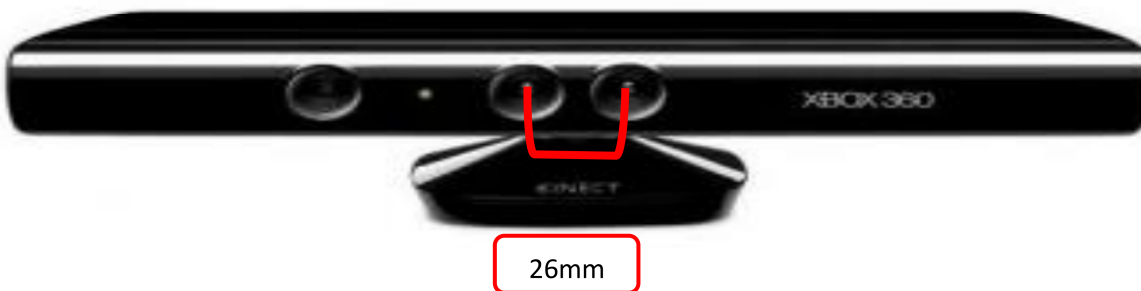


Fig.10 : Distance Base ligne

Les Résultats de calibrage stéréoscopique de la Kinect réalisé au niveau du Laboratoire Robotique et Productique de l'UER Automatique de l'EMP (KHaoula, 2014)

Caméra droite IR	Le focal	$\alpha_u = 526.97494$ $\alpha_v = 528.35283$
	Le point central	$u_0 = 324.17104$ $v_0 = 254.74485$
Caméra gauche RGB	Le focal	$\alpha_u = 582.52835$ $\alpha_v = 584.54913$
	Le point central	$u_0 = 322.26189$ $v_0 = 240.73137$
Transformation entre les deux caméras	Vecteur de rotation	$\alpha = -0.00070$ $\beta = -0.00273$ $\gamma = -0.00060$
	Vecteur de translation	$t_x = 26.67705$ $t_y = 2.82631$ $t_z = 0.30783$

7. Conclusion

Ce type de matériel, nommé caméra RGB-D, représente une véritable révolution dans le domaine de la robotique, elle permet l'acquisition des images RGB et la profondeur de chaque pixel de l'image et facilite la construction 3D et le plaquage des textures mesurant aussi bien son état interne que l'environnement dans lequel il évolue permettant d'atteindre l'objectif d'autonomie à travers la tâche navigation

CHAPITRE 3

Construction 3D et technique de plaquage de texture avec la Kinect

- 1. Introduction**
- 2. Définition de la Construction 3D**
- 3. Construction des nuages 3D et technique de plaquage de texture**
 - 3.1 Construction du nuage 3D**
 - 3.2 Plaquage de textures**
- 4. Conclusion**

1. Introduction

Donner la possibilité aux robots mobiles d'être autonomes, c'est lui donner la capacité de se localiser dans de nouveaux environnements en outre le doter de l'habilité de générer des cartes précises 3D reflétant la scène qu'il observe , ainsi pouvoir mesurer les distances dans l'espace 3D et d'ajouter des informations de couleurs (textures) à la carte est d'un grand intérêt.

La construction d'une carte locale de l'environnement est un problème dual du problème de localisation. Ici, il s'agit de déterminer une représentation de l'environnement obtenue sous la forme d'une carte géométrique ou d'une carte d'occupation.

En effet, le passage d'une cartes 2D à des cartes 3D nécessite des algorithmes de reconstruction assurant l'extraction des nuages des points , le dosage des couleurs de ces dernier et le plaquage des textures .

2. Définition de la Construction 3D

La reconstruction 3D à partir d'images désigne la technique qui permet d'obtenir une représentation en trois dimensions d'une scène à partir d'un ensemble d'images prises sous différents points de vue. D'une manière plus générale, le problème appelé reconstruction 3D est le suivant : nous disposons d'une ou plusieurs représentations en 2D d'un objet et nous souhaitons déterminer les coordonnées des éléments visibles sur ces représentations dans un repère de l'espace réel 3D.

Dans la robotique mobile, la navigation par l'approche réactive est généralement utilisée dans le cas où l'environnement est inconnu. Dans ce cas le robot doit réagir tout seul sans planification de trajectoire, mais l'inconvénient de ces approches par rapport aux autres est que le chemin parcouru par le robot n'est pas optimal. De ce fait, nous souhaitons utiliser le premier parcours du robot pour reconstruire une carte.

3. Construction des nuages 3D et technique de plaquage de texture

3.1 Construction du nuage 3D :

L'organigramme présenté ci-dessous définit les étapes de la technique pour la reconstruction 3D étudié dans le mémoire (Khaoula, 2014)

Dans ce qui suit nous allons expliquer chaque étape en détail.

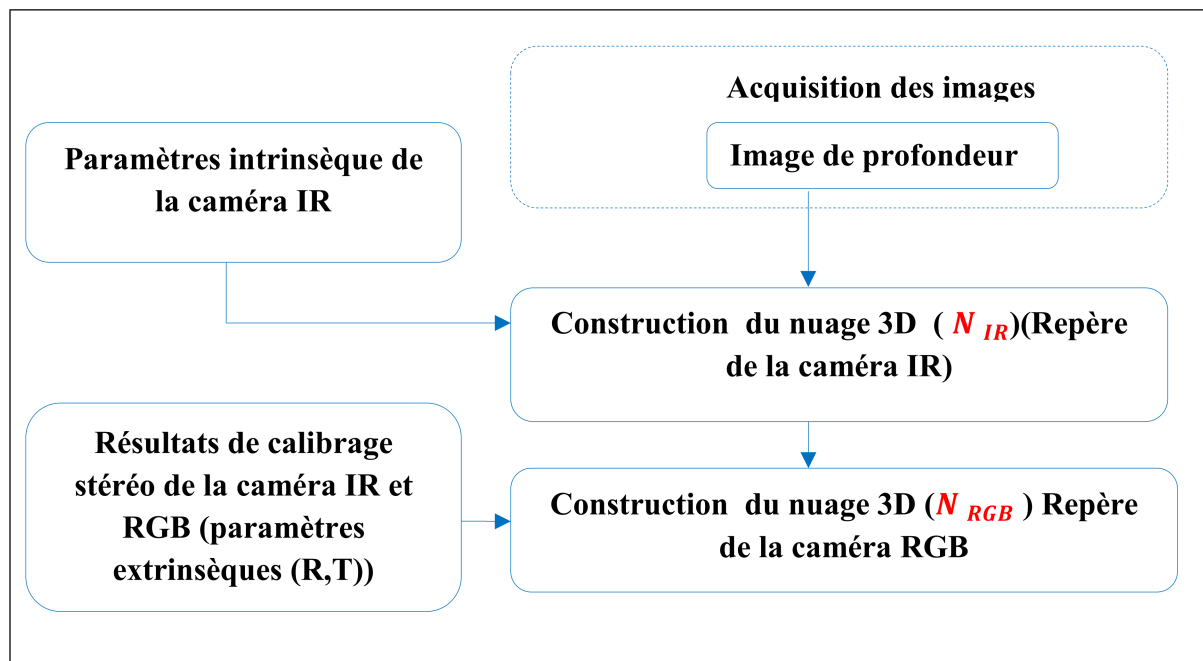


Fig.11 : Organigramme de la construction du nuage 3D

Afin de construire le nuage 3D nous passons par plusieurs étapes : Nous avons considéré que le calibrage des deux caméras (IR, RGB) est fait, donc nous avons comme résultats de calibrage (les paramètres intrinsèques et extrinsèques).

1. acquérir l'image de profondeur (Depth) par la Kinect voir figure :

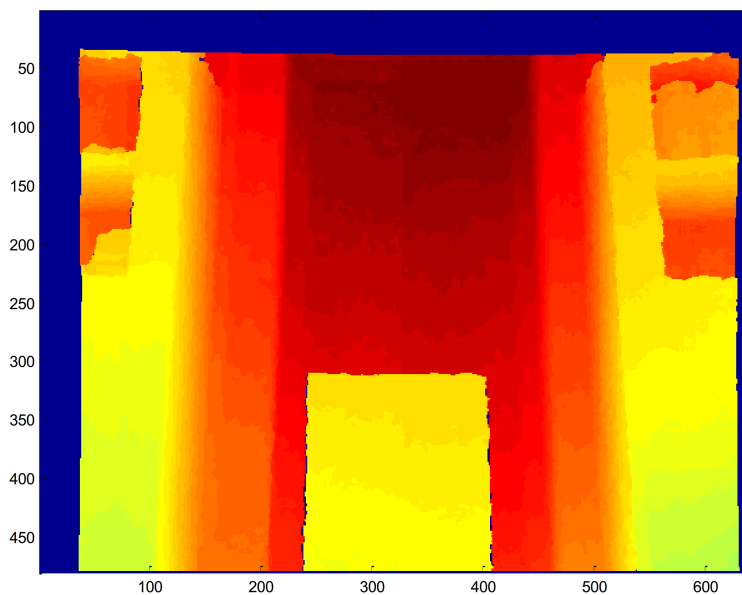


Fig.12 :Image de profondeur

2. Utiliser l'image de profondeur et les résultats du calibrage de la caméra IR (les paramètres intrinsèques $I_{C_{IR}}$) Pour construire le nuage3D (N_{IR}) dans le repère de la caméra IR, présenté dans la figure .

$$N_{IR} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Nous avons:

$$\begin{bmatrix} SU \\ SV \\ S \end{bmatrix} = I_{C_{IR}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Avec :

$$I_{CIR} = \begin{bmatrix} \alpha_u & 0 & U_0 \\ 0 & \alpha_v & V_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Donc :

$$N_{IR} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = I_{CIR}^{-1} \begin{bmatrix} SU \\ SV \\ S \end{bmatrix}$$

Nous obtenons la valeur de profondeur pour chaque pixel à partir de l'image de profondeur

Donc : $S = d = Z$

$$I_{CIR}^{-1} = \begin{bmatrix} \frac{1}{\alpha_u} & 0 & -\frac{1}{U_0} \\ 0 & \frac{1}{\alpha_v} & -\frac{1}{V_0} \\ 0 & 0 & 1 \end{bmatrix}$$

Après développement :

$$N_{IR} = \begin{cases} X = \frac{S}{\alpha_u} * (U - U_0) \\ Y = \frac{S}{\alpha_v} * (V - V_0) \\ Z = S = d \end{cases}$$

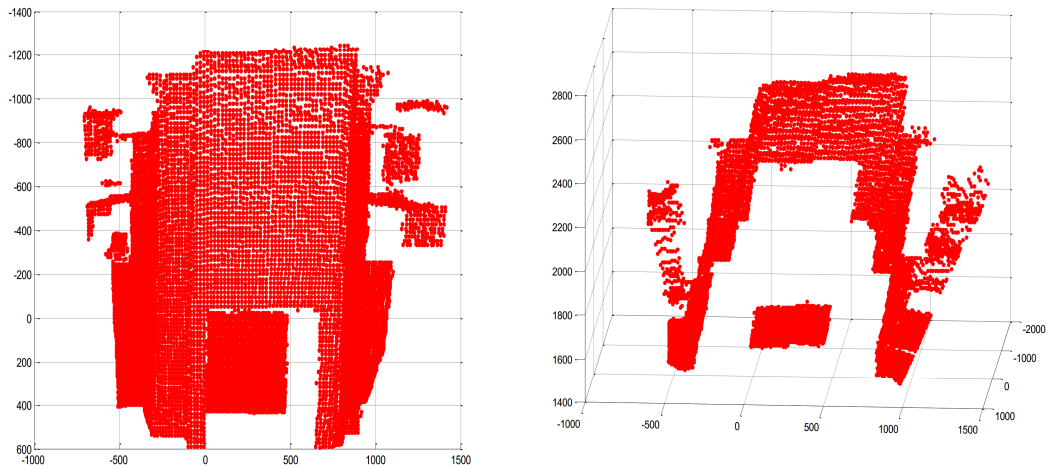


Fig.13 : Le Nuage 3D N_{IR} en rouge (le repère de la caméra IR)

3. Ensuite pour construire le nuage 3D (N_{RGB}) dans le repère de la caméra RGB on utilise les résultats du calibrage stéréo entre les de caméra IR et RGB les paramètres extrinsèques T (vecteur de translation) et R (matrice de rotation)

$$N1 = N_{RGB} = R_{EX} * N_{IR} + T_{EX}$$

$$N1 = \begin{bmatrix} X1 \\ Y1 \\ Z1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

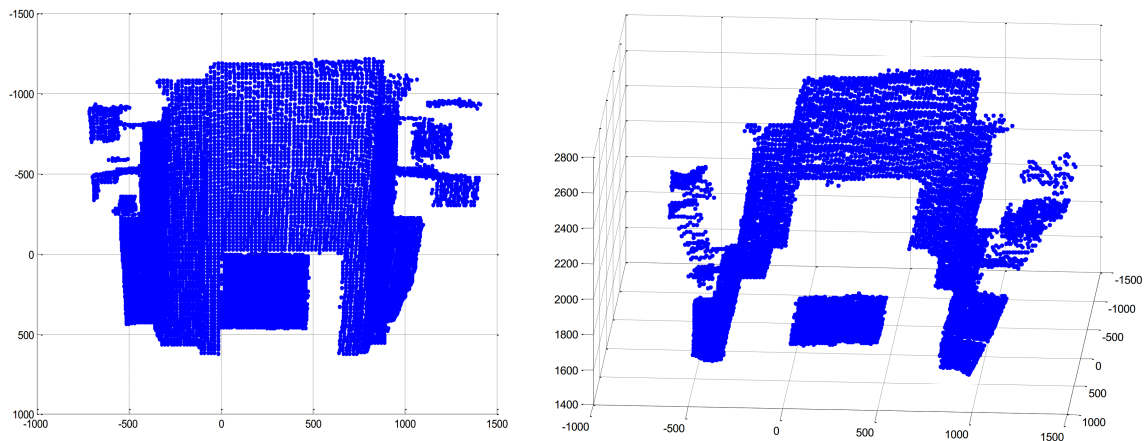


Fig.14 : Le Nuage 3D N_{RGB} en bleu c'est dans Le repère caméra RGB (N_{IR} rouge)

3.2 Plaquage de textures

Maintenant qu'on a le nuage de point RGB dans le repère camera RGB et pour récupérer les couleurs associé au Nuage on doit passer par les pixels

1. On calcule la valeur du pixel ($U1, V1$) pour chaque point de $N1$ à l'aide des résultats de calibrage de la caméra RGB ($I_{C_{RGB}}$)

$$\begin{bmatrix} SU1 \\ SV1 \\ S \end{bmatrix} = \begin{bmatrix} \alpha_u' & 0 & U_0' \\ 0 & \alpha_v' & V_0' \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X1 \\ Y1 \\ Z1 \end{bmatrix}$$

$$S = Z1$$

Après développement :

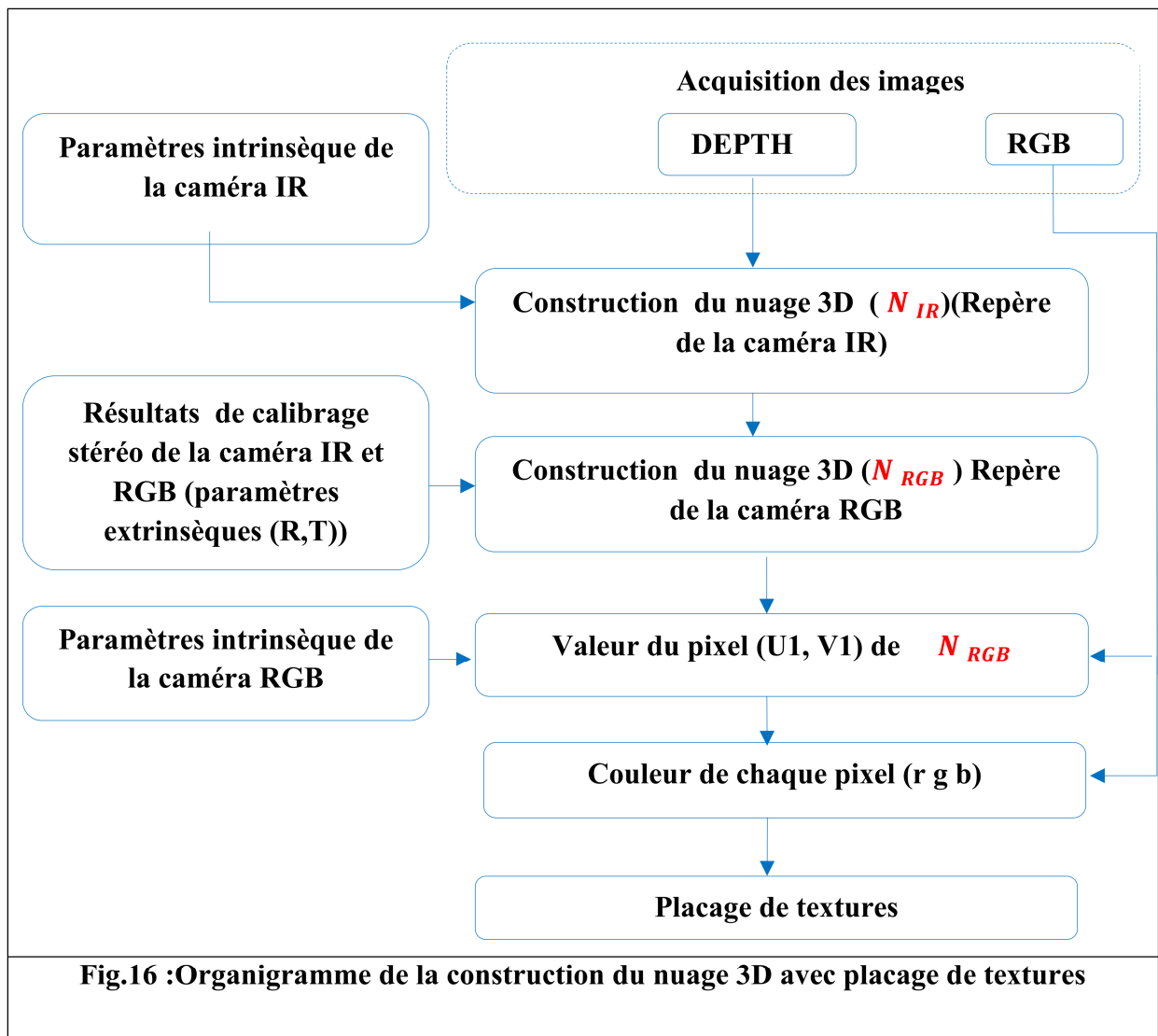
$$U1 = \frac{(\alpha_u' * X1 + U_0')}{S}$$

$$V1 = \frac{(\alpha_v' * Y1 + V_0')}{S}$$

2. après on calcule la couleur de chaque pixel (r, g, b) ou on les détermine à partir de l'image RGB pour faire le plaquage des textures avec la fonction scatter3 sur matlab



Fig.15 :Image réelle (RGB), construction 3D avec placage de texture



4. Conclusion

Notre système de vision (Kinect) nous a permis d'acquérir des images RGB, IR et la profondeur de chaque pixel. Ce qui facilite la construction 3D et le plaquage des textures a cet effet le robot devient capable de générer des cartes précises 3D, mesurer les distances dans l'espace 3D, ajouter des informations de couleurs/textures à la carte construite et se localiser dans cette carte qui étapes indispensables pour n'importe quelle tâche de navigation autonome, En effet, Notre approche nous permet le passage de cartes 2D à des cartes 3D à l'aide du capteur RGB-D et les Mex function .

CONCLUSION GENERALE

Le travail présenté consiste à étudier et résoudre la problématique de la localisation et la cartographie 3D pour un robot mobile, en utilisant une caméra de profondeur (Kinect).

D'abord, nous avons présenté un état de l'art sur la numérisation 3D ou on expose les capteurs visuels passifs et leurs limitations pour introduire notre capteur Kinect. Lors de cette étude, nous avons montré la description, le principe de fonctionnement, la partie calibrage de cette caméra. Notre attention s'est portée sur l'avantage d'une telle caméra pour obtenir directement des informations différentes sur l'environnement qui seront utilisées pour la reconstruction 3D du robot en temps réel.

Après cela, nous avons présenté la construction du nuage 3D à partir d'une image acquise par la Kinect puis par la suite et avec technique de plaquage de texture nous avons cartographié la scène observée.

Bibliographie

- A. LE JEUNE, S. P. (2012). Utilisation de la Kinect. *Linux Magazine France* , 16-29.
- Abdullah Hojaij, J. Z. (2014). A Two Phase RGB-D Visual Servoing Controller. *International Conference on Intelligent Robots and Systems*. Chicago, IL, USA: IEEE.
- Céline Teuliere, E. M. (2012). Direct 3D servoing using dense depth maps. *Conf. on Intelligent Robots and Systems, IROS'12*. Vilamoura, Portugal: IEEE.
- Céline Teuliere, E. M. (2010). Using multiple hypothesis in model-based tracking. *Conf. on Robotics and Automation, ICRA'10*,. Anchorage, Alaska, United.
- Chaumette, F. (2002). Asservissement visuel. Dans KHALIL, *La commande des robots* . Hermès. .
- hamza, d. (2015). *Conception et réalisation d'un driver pour la plateforme robotique B21r à l'aide de l'outil Ros*. algerie: USTHB.
- Lapayre, B. B.-C. (2012). *Projet Kinect*. l'Université de Franche-Comté.
- Martinet, P. e. (1999). *ASSERVISSEMENT VISUEL*. Universit e Blaise Pascal Clermont-Ferrand.
- Martinez, A., & Farnandez, E. (2013). *Learning ROS For Robotics Programing*. Birmingham B3 2PB, UK: Packt Publishing.
- oudah. (2012). *LE CONTRÔLE VISUEL APPLIQUÉ DANS LA ROBOTIQUE MOBILE*. ALGERIE: Thèse préparée au sein du Laboratoire de Commande des Processus de l'ENP .
- Tamtsia, A. Y. (2013). *Nouvelles contributions `a l'application des moments en asservissement visuel*. france: Universite Blaise Pascal - Clermont-Ferrand II.
- Wenhao Fu, H. H.-A. (2013). Visual servoing based mobile robot navigation able. *IEEE* .