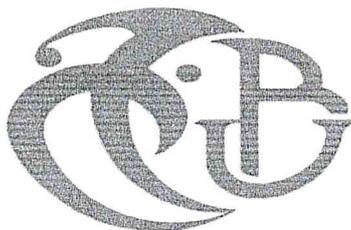


République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Saad Dahlab Blida

N° D'ordre :



Faculté des sciences

Département d'informatique

Mémoire Présenté par :

BESTAM Mohamed Abdessamed

BOUGHLALI Djalil

En vue d'obtenir le diplôme de master

Domaine : informatique

Filière : MI

Spécialité : Informatique

Option : Génie des systèmes informatique

Thème

« Détection de la pose de l'être humain en utilisant un capteur RGB-D »

Organisme d'accueil : Centre de Développement des Technologies Avancées

Soutenu le :

M.Président : *Ould-Khaoua*

M. Examineur *Mme. Chikhi*

M. Examineur *Mr. Derrar*

Mr. KAMECHE Abdallah Hicham Promoteur

Mme.KAHLOUCHE Souhila Encadrante

Promotion
2015 / 2016

AU NOM DE DIEU LE TRES HAUT, LE MISERICORDIEUX

« Les savants sont les héritiers des prophètes, or les prophètes n'ont laissé en guise d'héritage ni dinars, ni dirham mais la science »

REMERCIEMENTS

Nous adressons nos remerciements aux personnes qui nous ont aidés à réaliser ce mémoire.

En premier lieu **Monsieur Kameche**, En tant que directeur de mémoire, il nous guidés dans notre travail et nous a aidés à trouver des solutions pour avancer.

Nous remercions **Madame Kahlouche** qui nous a accueillis au niveau du Centre de Recherche et de Développement des Technologies avancées de Baba Hassen et qui nous a aidés pour la réalisation technique de notre travail.

Dédicace

Je dédicace ce mémoire à mes parents, à mon frère jumeau, mon frère aîné son épouse et mes petits neveux, ma sœur son époux, à ma petite nièce, et à Imene. Je les remercie pour leur soutien précieux.

Je tiens à associer à cette dédicace mon binôme,
tous mes proches et amis.

Merci pour vos encouragements.

DJALIL

Dédicace

Je dédicace ce mémoire à mes parents, mon oncle et sa femme, mes frères, ma sœur, mon beau frère et mes belles sœurs, mes neveux, ma fiancée et sa famille.

Je les remercie pour leur soutien précieux.

Je tiens à associer à cette dédicace mon binôme, tous mes proches et amis.

Merci pour vos encouragements.

Mohamed Abdessamed

RESUME :

Le travail présenté dans ce mémoire concerne la détection de l'être humain par un robot guide dans un milieu intérieur, et cela a l'aide d'un capteur RGB-D de type Kinect qui représente le capteur de vision du robot.

Ce dernier a été utilisé pour implémenter un nœud ROS.

Nous avons commencé par étudier la solution existante, puis concevoir et implémenter notre nouveau nœud ROS pour la détection de la pose de l'être humain.

Nous terminerons notre travail par des tests expérimentaux sur le robot B21r du CDTA

Mots-clés : capteur RGB-D, robot guide, Kinect, nœud ROS, détection de la pose

ABSTRACT

The work presented in this dissertation concerns the detection of the human being by a guide robot in an interior environment and this with the help of a RGBD-D collector type Kinect which represent the collector of the robot's vision.

The latter has been used in order to implement a ROS node.

We have started by studying the existing solution, then conserve and implement our new ROS node for the detection of the human being's pose.

We will finish our work by experimental tests on B21r robot of CDTA.

ملخص:

العمل المقدم في هذه المذكرة يتعلق بالتعرف على وضعية الإنسان في بيئة مغلقة عن طريق الروبوت، و هذا باستخدام جهاز الاستشعار "RGB-D" من نوع «Kinect»

الذي يمثل جهاز للرؤية بالنسبة للروبوت "B21r" وقد استخدم هذا الأخير لتنفيذ عقدة ROS.

بدأنا عملنا بدراسة الحل السابق ، ثم تصميم وتنفيذ عقدة "ROS" جديدة للكشف عن وضعية الإنسان. وسوف ننهي عملنا في الاختبارات التجريبية على الروبوت "B21r".

Table des Figures :

Figure 1:Le canard de Vaucanson, disparu, a été reconstitué partiellement.....	5
Figure 2:Euphonia, l'automate parlante de Faber.....	6
Figure 3:L'automate joueur d'échec de von Kempelen.....	6
Figure 4:Grey Walter et une tortue cybernétique sans carapace.....	7
Figure 5:Classification par fonctionnalité / domaine.....	9
Figure 6:PR2 by Willow Garage.....	10
Figure 7:Kawasaki FS-03N, robot de soudage.....	11
Figure 8: Un robot médical, Da Vinci, fabriqué par Intuitive Surgical.....	11
Figure 9:Un robot NAO (2014).....	14
Figure 10:Un robot ASIMO (2011).....	14
Figure 11: drones.....	15
Figure 12:ROV.....	16
Figure 13:ALISTAR3000.....	17
Figure 14:Le CURV-III, durant le sauvetage du Pisces III en 1973.....	17
Figure 15:Roomba 980.....	19
Figure 16:ROVIO.....	19
Figure 17: Robot-tondeuse Lizard S14N.....	20
Figure 18: Le Robot B21r.....	20
Figure 19: Classification des méthodes de détection d'objets.....	24
Figure 20: capture de mouvements avec marqueurs.....	33
Figure 121:capture de mouvements avec accéléromètres.....	33
Figure 22:capture de mouvements sans marqueurs.....	34
Figure 23:caméraKinect.....	35
Figure 24: les différents composants de la Kinect.....	36
Figure 25:Visualisation de la main (de face) avec rviz.....	40
Figure 26:Visualisation de la main (de dos) avec rviz.....	41
Figure 27:Architecture software du robot B21r sous ROS.....	42
Figure 138:Organigramme de détection de personne par le robot.....	42
Figure 29:Représentation des Angles d'Euler.....	44
Figure 30: Conversion Quaternion vers les angles d'Euler.....	44
Figure 31:la Pose avec les rotations.....	45

Figure 32: Organigramme pour la détection de la Pose.....	45
Figure 33: Récupération des coordonnées de l'épaule droite.....	46
Figure 34: Représentation du repère de la Kinect.....	47
Figure 35: Projections des coordonnées des épaules sur le plan (oxz).....	47
Figure 36: comparaison des coordonnées X.....	48
Figure 37: comparaison des coordonnées Z.....	48
Figure 38: Fonctionnement de ROS.....	54
Figure 39: Structure de la librairie OpenCV.....	57
Figure 40: Robot B21r.....	58
Figure 41: Calibration.....	59
Figure 42: Détection de la pose Face.....	60
Figure 43: Détection de la pose Droite.....	61
Figure 44: Détection de la pose Dos.....	62
Figure 45: Détection de la pose Droite.....	63
Figure 46: Organigramme pour la détection de la Pose	64

Liste des tableaux :

Tableau 14: Résultat de la détection de la pose avec l'algorithme d'un point d'articulation	46
Tableau 15: Résultats des calibrations de différentes distances.....	58
Tableau 16: Résultats primaires	64
Tableau 17: Résultats finaux	65

Introduction Générale

La recherche en robotique couvre un ensemble de sujets contribuant à comprendre les principes et à réaliser les fonctions qui permettent de donner à une machine, qui sera appelée robot, des capacités de perception, de décision, d'action et d'interaction adaptées à son environnement et aux tâches pour lesquelles elle est conçue.

La robotique doit sa progression considérable depuis quelques années, aussi bien sur le plan de la recherche que sur celui des applications et des usages, à la conjonction de trois facteurs technologiques :

- des capteurs (caméras, télémètres, centrales inertielles, capteurs d'effort, de contact, etc.) plus performants, miniaturisés et à coût de plus en plus réduit;
- des processeurs toujours plus rapides et de la mémoire de plus en plus disponible;
- des composants « mécatroniques » (moteurs, microcontrôleurs) et des matériaux permettant la construction de machines intégrées plus légères et plus précises.

Ces facteurs ont permis des avancées importantes de la recherche dans les domaines de la perception, de la représentation de l'espace, de la reconnaissance de scènes - objets, humains, situations - de la planification du mouvement et des actions ainsi que sur les architectures de contrôle, c'est-à-dire l'intégration de ces différentes fonctions.

Ces progrès sur les capacités opérationnelles ont favorisé à leur tour les recherches sur les capacités cognitives (prise de décision, autonomie, apprentissage, interaction avec les humains, détection et expression d'émotions).

De plus, le partage de données, de logiciels et le développement de plateformes communes entre différents centres de recherches ont accéléré l'identification des méthodes les plus performantes et les mises en œuvre expérimentales.

Un robot de service est un robot en interaction avec l'humain et/ou l'environnement. Il s'agit d'un système complexe évoluant avec d'autres systèmes complexes et dynamiques.

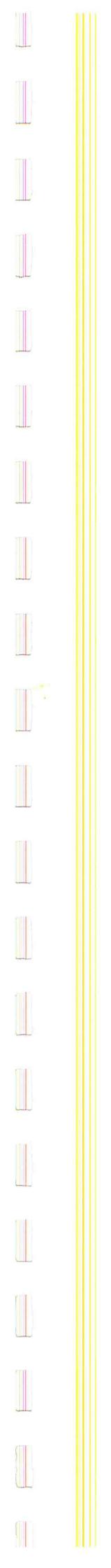
Dans le cadre du projet de recherche robot guide du CDTA, qui vise à rendre le robot B21r autonome et de le doter de fonctionnalités avancées lui permettant de jouer le rôle d'un robot guide dans un environnement public structuré type musée.

Notre contribution dans ce projet concerne la réalisation d'un système capable d'exploiter les données issues d'une caméra RGB-D embarquée sur le robot pour la reconnaissance de pose et par la suite l'interaction Homme/Robot.

Le présent mémoire est structuré en quatre grands chapitres :

- **Le chapitre 1** : Généralités sur la robotique où nous l'avons défini brièvement avec un historique approprié puis nous avons donné les types des robots selon leurs fonctionnalités en citant des exemples sur chaque type.
- **Le chapitre 2** : Nous l'avons entamé par une introduction sur l'estimation de la pose d'un être humain puis nous avons parlé sur les diverses détections telles que la détection des formes et nous avons expliqué les méthodes utilisées pour la détection de personnes ensuite nous avons cité les diverses techniques de capture de mouvement et enfin nous avons parlé sur les différentes méthodes de la détection de visage.
- **Le chapitre 3** : Consiste à décrire la conception et le développement de la solution proposée pour la détection de pose de l'être humain.
- **Le chapitre 4** : Où on expose les expériences et les résultats expérimentaux dans lequel nous présentons l'ensemble des tests effectués.

Enfin, nous clôturons ce mémoire par une conclusion générale dans laquelle on récapitule ce qui a été fait et mentionne des perspectives pour notre travail.



CHAPITRE 1 : GENERALITES

SUR LA ROBOTIQUE

1. INTRODUCTION

Sans remonter aux premiers concepts de machine remplaçant l'homme dès le XVIIe siècle, la robotique est née dans les années 1950, du croisement des besoins et des disponibilités de nouvelles technologies développées durant la seconde guerre mondiale : l'électronique, l'automatique, l'informatique... Les deux premières orientations de ces machines étaient de répondre aux besoins de l'industrie manufacturière et aux besoins de l'industrie en milieux hostiles à l'homme [1].

La robotique industrielle a répondu dans un premier temps au besoin de manipuler de manière répétitive et automatique des objets entre les machines de fabrication proprement dite. Ce robot devait donc avoir au minimum un organe de saisie des pièces pouvant suivre des trajectoires programmables. Le robot industriel est maintenant un produit majeur et répandu à plus d'un million d'exemplaires, utilisé dans de nombreuses étapes de fabrication (manipulation, peinture, soudure) ; il est complété maintenant dans le domaine manufacturier par des robots mobiles de transport ou d'assistance (AGV, assistance au montage) qui ont beaucoup de points communs avec les robots de service. La robotique est par ailleurs considérée comme un élément clé de la compétitivité des entreprises industrielles, au point de susciter des soutiens publics massifs.

La robotique en milieu hostile répond aux besoins de manipulation d'objets à distance du fait de leur dangerosité (chimie, explosif), de la nocivité du milieu pour l'homme (nucléaire) ou de la difficulté à s'y rendre (espace). Les premières réalisations ont été justifiées et financées par l'industrie nucléaire dans les années 1950. Ces robots sont des systèmes électromécaniques, télé opérés.

Les développements de l'électronique et de l'informatique ont permis d'introduire dans ces équipements l'assistance à l'opérateur sous diverses modalités (retour d'effort, mouvement automatique sur commande de haut niveau) afin d'améliorer la précision, la vitesse d'exécution et d'enrichir les domaines d'application (démontage, lutte contre le terrorisme, intervention post-accident). Depuis une vingtaine d'années, des moyens importants ont été mis en place pour étendre les applications robotiques sur les théâtres d'opérations militaires.

Après la robotique industrielle et la robotique d'intervention, une troisième orientation est apparue avec les progrès en miniaturisation, microélectronique et micromécanique et les nouvelles capacités des systèmes de traitement de l'information et de communication qui ont créé les conditions technologiques favorables au développement de robots mobiles autonomes

ou semi-autonomes pour réaliser des applications de service professionnel et de service personnel à des coûts et tailles raisonnables à moyen terme.

Les progrès en sciences cognitives et ceux réalisés dans la compréhension du vivant, ouvrent également la porte à de nouvelles avancées sur les capacités d'apprentissage et d'intelligence d'entités artificielles, matérielles et à leur introduction dans des environnements en forte interaction avec l'homme. Ces applications de la robotique de service sont très larges car elles peuvent couvrir des domaines très variés du monde professionnel (agriculture, médical, nettoyage...) et du monde grand public (jeux, éducation, tâches domestiques...) [2].

2. HISTORIQUE

On peut schématiquement distinguer trois principales ères en robotique : les automates, les robots n'étant pas dotés d'intelligence artificielle (I.A.), et ceux disposant d'une I.A.

2.1 Les automates (1ère génération)

Un automate, contrairement à un robot (même s'il ne dispose pas d'intelligence artificielle), obéit uniquement à un programme préétabli, que ce soit de manière mécanique ou électronique. De ce fait, il n'y a aucune adaptation possible entre l'automate et son environnement.

Dès l'Antiquité, on signale plusieurs automates, dont le pigeon volant d'Archytas de Tarente ou les fameuses scènes théâtrales animées de Héron d'Alexandrie.

Au XVIe siècle, Léonard de Vinci construisit le premier androïde capable de coordonner les mouvements de ses bras, de ses jambes et même de ses mâchoires. Au XVIIIe siècle – considéré comme l'âge d'or des automates - le célèbre canard de Jacques de Vaucanson (figure1), qui pouvait boire, se nourrir, caqueter, s'égoutter dans l'eau, digérer sa nourriture et même... déféquer, a ébloui par sa complexité les spectateurs de l'époque.

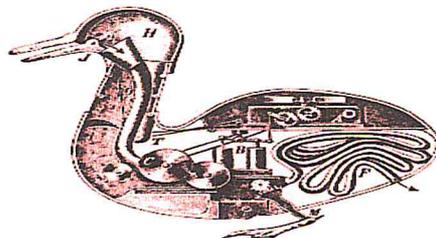


Figure 1: Le canard de Vaucanson, disparu, a été reconstitué partiellement [43]

À la même période, les horlogers Jaquet-Droz inventèrent une musicienne, un écrivain et un dessinateur réalisant vraiment les mouvements correspondant à la pratique de leur art.

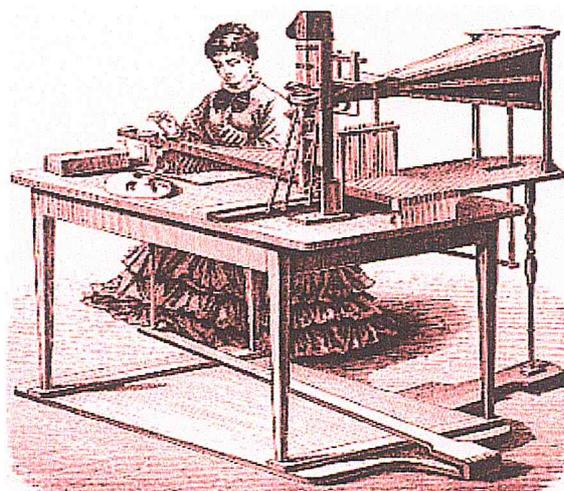


Figure 2: Euphonia, l'automate parlante de Faber[43]

Au XIXe siècle, l'automate parlante Euphonia d'Eugène Faber (figure 2) était supposée dialoguer avec les spectateurs et l'automate turc du baron von Kempelen jouait aux échecs – actionné peut-être par un humain caché dans le dispositif (figure 3).

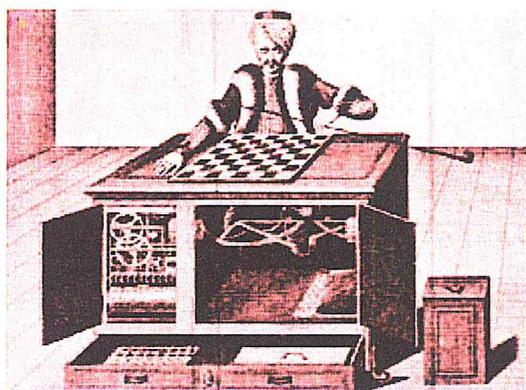


Figure 3: L'automate joueur d'échec de von Kempelen[43]

2.2 Les robots de seconde génération

Les robots de la seconde génération disposent d'organe(s) sensoriel(s), autrement dit des capteurs, pouvant influencer sur leurs comportements. Ils sont donc relativement adaptables à leur environnement.

Ce n'est qu'au tout début du XXe siècle que les robots firent leur apparition, suite aux travaux d'ingénieurs qui voulaient tester des hypothèses émises par des biologistes et des psychologues. Le chien électrique conçu par [Hammond et Miessneren, 1915] était attiré par une lumière, selon le phototropisme animal mis en évidence par [Loeb, 1918].

Les machines de [Russell,1913] et de [Stephens ,1929] [45], les tortues cybernétiques de [Grey Walter, 1950] (figure 4), le renard électronique de [Ducrocq, 1953] ou [l'homéostat d'Ashby,1952] [45] étaient, elles, dotées de capacités d'apprentissage directement issues des travaux des psychologues [Thorndike, 1911][46], [Hull ,1943] [46] et du physiologiste [Pavlov,1903][46] sur l'Homme et l'animal.

Ces réalisations sont des robots, car elles ne se comportent plus comme de simples automates dont les organes moteurs – leurs mécanismes – obéissent à un programme préétabli. À la différence des automates, ils ont des organes sensoriels – les capteurs – qui recueillent des informations de l'environnement qui vont, elles, influencer l'activité de leurs organes moteurs – les actionneurs.

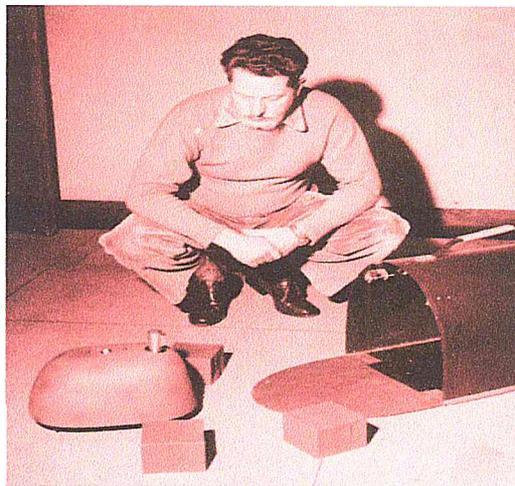


Figure 4:Grey Walter et une tortue cybernétique sans carapace [43]

2.3 Les robots dotés d'I.A. (3ème génération)

Pour être intégré dans cette famille, un robot doit parvenir à effectuer une tâche par lui-même, sans aucune aide extérieure.

En 1973, l'université de Waseda présente le tout premier humanoïde "intelligent" dénommé Wabot-1. [47]

Il est doué de la vision, peut manipuler des objets, effectue un semblant de marche et est même capable de débiter une conversation en japonais.

Hi-T-Hand d'Hitachi en 1974 [48] manipule des aiguilles à travers des trous grâce à la détection de force.

Vers la fin des années 70 Hans Morava présente les premiers robots capables d'évoluer à l'extérieur, de façon autonome. [48]

3. DEFINITION

Dans ce qui suit l'origine du mot et aussi quelques définitions du robot : [4]

3.1 Origine des termes

Le mot Robot a été utilisé pour la première fois en 1921 par Karel Capek dans sa pièce R.U.R. (Rossums Universal Robots).

Il provient du tchèque "robota" qui signifie corvée, travail obligatoire.

Le terme robotique a été employé pour la première fois par Asimov en 1941.

3.2 Définition d'un robot

Le Petit Larousse définit un robot comme étant un appareil automatique capable de manipuler des objets, ou d'exécuter des opérations selon un programme fixe ou modifiable. En fait, l'image que chacun se fait d'un robot est généralement vague, souvent un robot est défini comme un manipulateur automatique à cycles programmables.

Pour « mériter » le nom de robot, un système doit posséder une certaine flexibilité, caractérisée par les propriétés suivantes :

-La versatilité : Un robot doit avoir la capacité de pouvoir exécuter une variété de tâches, ou la même tâche de différente manière ;

-L'auto-adaptativité : Un robot doit pouvoir s'adapter à un environnement changeant au cours de l'exécution de ses tâches.

L'Association Française de Normalisation (A.F.N.O.R.) définit un robot comme étant un système mécanique de type manipulateur commandé en position, reprogrammable, polyvalent (i.e., à usages multiples), à plusieurs degrés de liberté, capable de manipuler des matériaux, des pièces, des outils et des dispositifs spécialisés, au cours de mouvements variables et programmés pour l'exécution d'une variété de tâches. Il a souvent l'apparence d'un, ou plusieurs, bras se terminant par un poignet. Son unité de commande utilise, notamment, un dispositif de mémoire et éventuellement de perception et d'adaptation à l'environnement et aux circonstances. Ces machines polyvalentes sont généralement étudiées pour effectuer la même fonction de façon cyclique et peuvent être adaptées à d'autres fonctions sans modification permanente du matériel. [5]

4. TYPES DES ROBOTS

Dans la figure 5 nous proposons une classification des robots en fonction de leurs fonctionnalités :

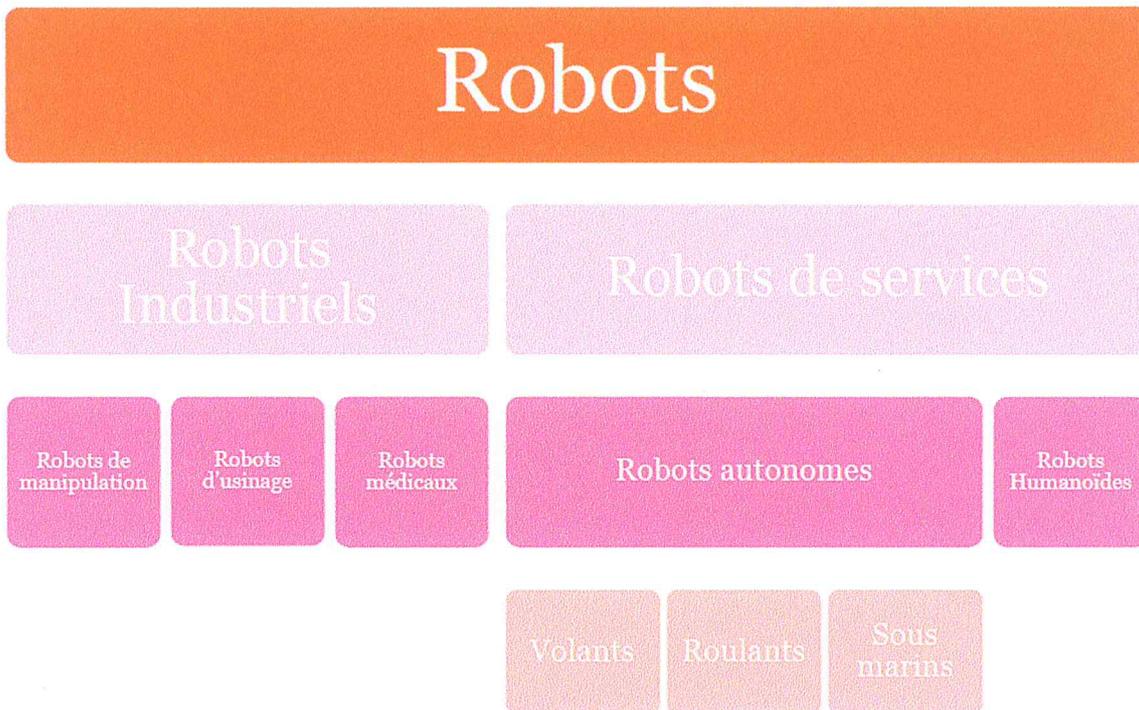


Figure 5: Classification par fonctionnalité / domaine [6]

4.1 ROBOTS INDUSTRIELS

Par *robot industriel* (RI) on entend une machine automatique reprogrammable et polyvalente, conçue pour déplacer des matériaux, des pièces, des outils, ou des dispositifs spécialisés au travers de déplacements variables et programmables pour accomplir diverses tâches.

Cette définition s'apparente à la norme proposée par l'AFNOR (Norme enregistrée NF E 61400 en août 1983 et acceptée entre temps au niveau international ISO. [6]

Il est clair que cette définition reste encore trop vague pour permettre de classer systématiquement une machine comme étant, ou n'étant pas, un robot industriel.

Les premiers robots industriels sont apparus sur le marché aux Etats Unis en 1962. Utilisés tout d'abord uniquement dans l'industrie automobile, ils ont peu à peu conquis d'autres secteurs de l'économie tels que la métallurgie, la microtechnique, l'agroalimentaire, la chimie-pharmacie, voire le bâtiment. [6]

Discrédités dans les années 80 par les excès technologiques et une fréquente inadéquation aux besoins, les robots industriels reviennent en force. Devenus moins chers et plus efficaces, ils offrent des solutions rentables aux entreprises à la recherche de compétitivité.

En 2008, on peut estimer que le nombre de robots industriels installés dans le monde a largement dépassé 1 million.

Environ 350'000 RI sont installés au Japon. Les deux dernières années, on a pu observer une croissance annuelle de près de 20% suivant les régions. La densité de robots pour 10'000 ouvriers varie entre 50 et 150 pour les pays industrialisés (Europe et USA) alors qu'elle atteint 330 au Japon et environ 150 en Corée. [7]

Les robots industriels se divisent en 3 types :

4.1.1 Robots de manipulation

Pour dire qu'un robot est de manipulation il doit manipuler des objets dans le but de les déplacer d'une position à une autre comme le robot PR2 (figure 6)

Par exemple :

- L'Assemblage de pièces
- L'Emballage de pièces (éléments)
- Le déplacement

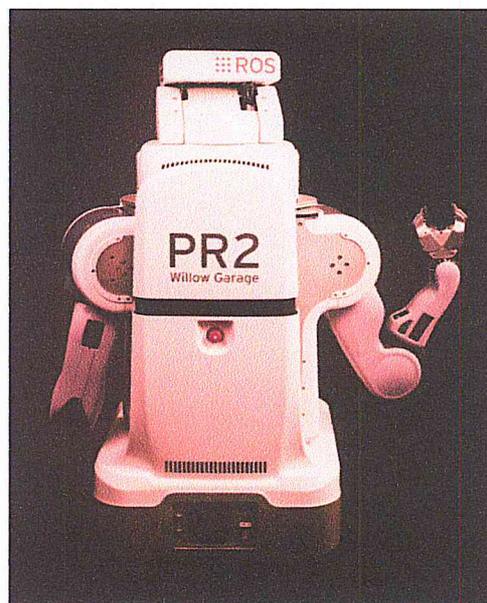


Figure 6:PR2 par Willow Garage [50]

4.1.2 Robots d'usinage

Leur rôle est d'usiner des pièces en manipulant des outils d'usinage comme le robot FS-03N (figure 7).

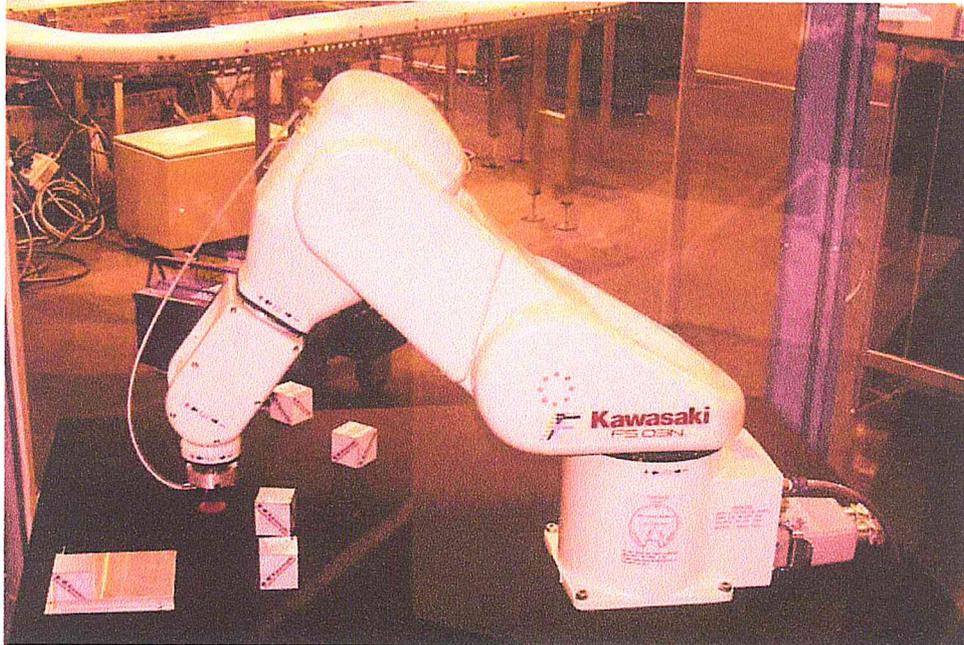


Figure 7:Kawasaki FS-03N, robot de soudage [49]

4.1.3 Robots médicaux

Ceux sont destinés à des applications médicales en rapport avec l'homme par exemple :

- Réadaptation motrice.
- Chirurgie par la manipulation d'outils. (Figure 8)

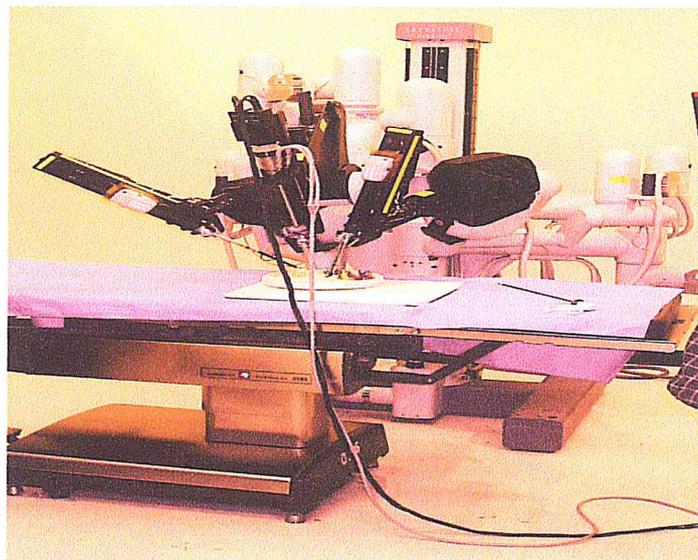


Figure 8: Un robot médical, Da Vinci, fabriqué par Intuitive Surgical [51]

4.2 ROBOTS DE SERVICES

La robotique de service est un champ d'application de la robotique de plus en plus médiatisé.

L'action des entreprises du secteur au monde est de plus en plus forte pour promouvoir des applications et des robots plus ou moins technologiquement perfectionnés, de l'aspirateur personnel au robot d'assistance militaire.

Afin de fournir à la filière et aux pouvoirs publics une vision claire de la réalité de cette industrie et de ses marchés, ainsi que de leur potentiel à moyen terme dans le monde.

Elle s'est développée en premier lieu dans le domaine nucléaire dès les années 1950 avec des mobiles téléopérés.

Aujourd'hui, les développements techniques matériels, algorithmiques et en traitement du signal rendent envisageable la commercialisation de robots en grande quantité, pour servir des marchés professionnels ou domestiques en environnement beaucoup plus ouvert que la robotique industrielle déjà bien implantée dans les grandes entreprises industrielles.

La robotique de service peut être définie comme le pan de la robotique qui assiste l'homme dans ses activités professionnelles comme dans sa vie courante, contribuant à améliorer ses conditions de travail, sa sécurité, son bien-être, etc. [8]

Selon les typologies utilisées, on distinguera :

4.2.1 Les robots humanoïdes

Un robot « humanoïde » est un robot dont l'apparence générale rappelle celle d'un corps humain. Habituellement, les robots « humanoïdes » possèdent un torse avec une tête, deux bras et deux jambes. En revanche, certains prototypes ne représentent qu'une partie du corps. Par exemple, certains robots « humanoïdes » n'ont que la tête (visage avec des yeux et une bouche). Pour simplifier, un humanoïde est un robot qui ressemble physiquement à un humain. [9]

Par exemple :

ASIMO est un robot humanoïde développé par Honda.(figure10) Ce nom est l'abréviation de «Advanced Step in InnovativeMobility ». [52]

La conception de ce robot est tournée vers un but utile. En effet, ce robot devrait pouvoir venir en aide aux personnes à mobilité réduite, âgées ou, encore malades. Il est conçu, aussi, pour effectuer des tâches dangereuses pour les humains. Par exemple, « ASIMO » pourrait aller dans des milieux hostiles à l'homme.

L'une des principales nouveautés technologiques de ce robot est le fait qu'il est capable de modifier sa trajectoire tout en marchant, à l'inverse des robots d'anciennes générations qui étaient contraints de stopper leur mouvement avant d'effectuer une rotation sur place.

De même, il lui est possible de détecter les mouvements des objets, ainsi que leur trajectoire. Il est aussi capable de monter et descendre des escaliers, reconnaître des visages, comprendre la parole humaine, étudier son environnement et garder son équilibre sur des surfaces mobiles.

NAO est un robot humanoïde autonome, de taille moyenne (environ 57 cm), (voir figure 9) développé par la société française «AldebaranRobotics». [53]

Ce robot possède une multitude de capteurs et d'actionneurs pour interagir avec son environnement. Pour apercevoir son environnement il dispose de :

- deux caméras vidéo
- quatre micros
- deux capteurs ultrasons
- des capteurs capacitifs (sensitifs)
- des capteurs de pressions (les pieds)

Ce robot est un "robot de compagnie".

En effet, NAO dispose de nombreuses fonctions, telles que jouer aux échecs, jouer au football. Cependant ces fonctions sont à programmer, avant utilisation du robot. Il a aussi une capacité très prononcée à "aimer", en effet, de très nombreuses expressions peuvent être lues sur le visage de « Nao ».

Néanmoins, il est, surtout, destiné à rendre la vie plus facile à l'Homme :

- Il pourra, grâce à sa synthèse vocale, lire n'importe quel contenu texte sur Internet, faire office de professeur de gym ou de cuisine. Il pourra raconter des blagues aux enfants, ou rappeler à un malade qu'il est l'heure de prendre ses médicaments.

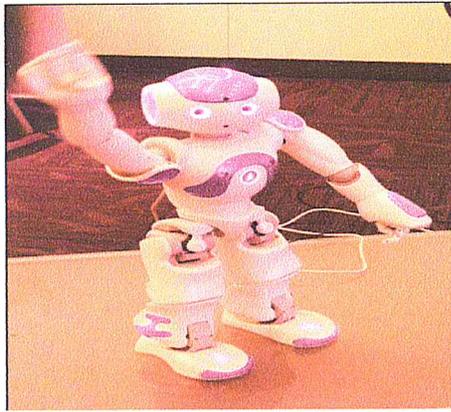


Figure 9: Un robot NAO (2014).[53]

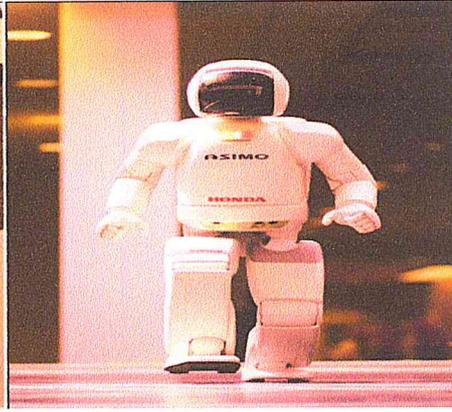


Figure 10: Un robot ASIMO (2011)[52]

4.2.2 Les robots autonomes

Le concept de robot mobile autonome est apparu vers la fin des années soixante, de deux sources totalement différentes : tout d'abord des recherches menées au Stanford Research Institute sur les possibilités d'équiper des machines de capacités de déduction et de réaction logique à des événements extérieurs. On a ainsi construit Shakey [54], machine à roues reliée à un ordinateur et équipée d'une caméra lui permettant d'acquérir des images de son environnement. Elle évolue dans un univers de cubes et de pyramides de tailles et de couleurs différentes. Shakey a pour mission de prendre un objet et de le porter ailleurs, quelque soit sa position ; chaque mission dure près de cinquante minutes. [10]

D'autre part, l'industrie nucléaire a besoin de machines permettant d'agir à distance dans des environnements encombrés et inaccessibles à l'homme. L'entreprise américaine Général Electric développe alors un quadrupède pour essayer de résoudre ce problème, tandis que les projets Luna [55] et Mars Rover [56] s'échafaudent dans le but d'explorer des planètes sans que l'homme ne prenne part au voyage.

Pendant plusieurs années, laboratoires, industriels, informaticiens et mécaniciens vont continuer leurs travaux en parallèle. On accède ainsi côté industriel à la télé-opération et à une partie de la robotique classique, tandis que du côté informatique, on assiste à de grands progrès dans le domaine de l'intelligence artificielle. Ainsi, vers la fin des années soixante-dix, trois pôles géographiques principaux se distinguent (France, Japon, Etats-Unis). La synthèse de tous les travaux réalisés jusqu'alors donne enfin naissance aux robots mobiles autonomes (du robot domestique au robot militaire).

L'industrie de production, les sociétés d'exploitation minière, les expéditions de recherche sous-marine... Les domaines d'utilisation de robots autonomes sont très variés, allant de la production en chaîne dans une usine de voitures à l'exploration d'autres planètes, comme

c'est le cas avec Mars Pathfinder [57], robot mobile autonome d'exploration de la planète Mars.

C'est dans cet environnement de plus en plus automatisé que se fait sentir le besoin d'outils capables, non seulement d'effectuer des tâches répétitives ou encore impossibles à l'homme (porter des charges lourdes, découpage ultra précis, ...), mais aussi de manifester une certaine autonomie de déplacement dans des milieux hostiles à l'homme. On en voit désormais les applications sur des chantiers tels que le désamiantage d'immeubles, la décontamination radioactive, les expériences en milieu dangereux... Aussi a-t-on besoin de robots mobiles autonomes capables de se déplacer d'un point à un autre sur une simple demande de l'utilisateur, qui n'a ainsi plus besoin d'être un expert en pilotage.

Le meilleur exemple de planification complexe de la trajectoire est fourni par les drones, ces robots volants destinés aussi bien à l'espionnage militaire qu'aux études botaniques nécessitant des prises de vues aériennes.

Les types des robots autonomes :

a- les robots volants les drones

"Un drone ou UAV (Unmanned Aerial Vehicle) est un aéronef inhabité, autonome ou télé-piloté. Ses missions peuvent être variées, pour des usages civils ou militaires"

A l'origine, le mot drone désigne un bourdon en anglais, référence faite au bruit caractéristique émit par beaucoup de ces appareils. [11]

Leur taille varie de celle d'un insecte aux maintenant célèbres Prédateur américains, drones militaires faisant près de 15 mètres d'envergure. (Figure 11)

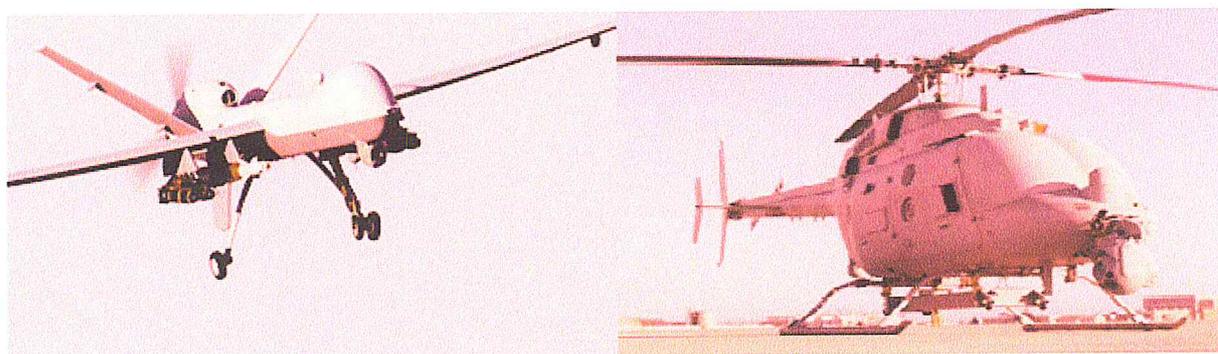


Figure 11: drones

Les limites des drones : Si les drones présentent des avantages certains, des reproches peuvent leur être faits sans difficultés. En effet, par leur taille parfois très petite ou leurs capacités opérationnelles les populations peuvent se sentir espionnées ou même en danger par l'usage des drones.

b- les robots sous-marins

Les robots Sous-marins sont des véhicules qui sont capables de fonctionner sous l'eau sans un occupant humain. Ils peuvent être contrôlés par un opérateur humain directement ou être autonome et fonctionner indépendamment de commandes humaines. [12]

Quelques robots sous marins :

-**ROV** (figure 12) est un petit sous-marin contrôlé à distance (généralement filoguidé).

Comme leurs homologues aériens, ces drones permettent une acquisition rapide et sécurisée d'informations globales ou précises, physicochimiques et visuelles (sous forme numérique notamment), assez rapidement, à distance de l'opérateur et parfois "en masse".

Certains ont une fonction de plateforme pouvant être équipée à la demande de préleveurs ou de divers capteurs. De nombreux prospectivistes estiment que leur usage civil, encore émergent, est une source potentielle de progrès scientifique et opérationnel.



Figure 8:ROV [58]

-**Alistar 3000** : C'est un engin sous-marin autonome est capable d'effectuer différentes sortes de tâches d'inspection sur les champs Offshore, à des profondeurs jusqu'à 3 000 m, sans intervention humaine : il n'a pas besoin d'être téléguidé comme un ROV.

Pesant 2 100 kg, il mesure 5 m de long, et peut emporter 150 kg de charge utile (Figure 13).



Figure 13:ALISTAR3000 [59]

- **CURV-III** : C'est un robot sous-marin de la Marine américaine. Il constitue la quatrième génération du CURV (Cable-controlled Undersea Recovery Vehicle), c'est un prototype de sous-marin téléguidé développé par le Space and Naval Warfare Systems Center San Diego au début des années 1960. (Figure14)

Il a effectué l'opération de sauvetage sous-marine la plus profonde de l'histoire, lorsqu'il sauve à 480 mètres de profondeur l'équipage du sous-marin canadien *Pisces III* en 1973.

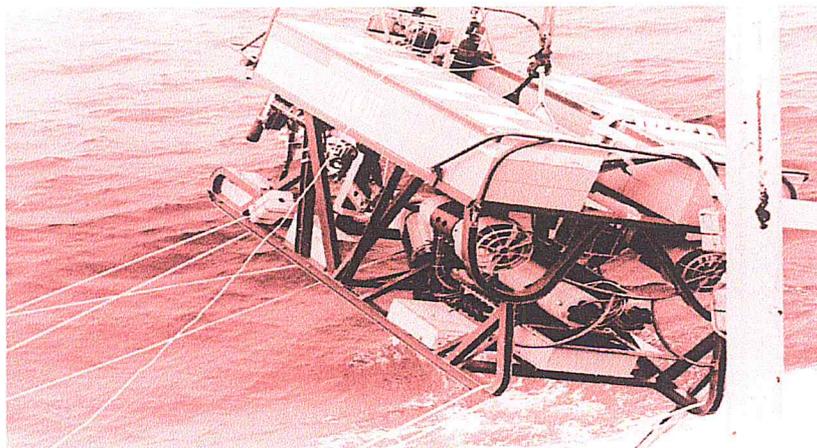


Figure 14:Le CURV-III, durant le sauvetage du Pisces III en 1973[60]

c- les robots roulants

La mobilité par roues est la structure mécanique la plus communément appliquée.

Cette technique assure selon l'agencement et les dimensions des roues un déplacement dans toutes les directions avec une accélération et une vitesse importante. Le franchissement d'obstacles ou l'escalade de marches d'escalier est possible dans une certaine mesure. Toutes les configurations (nombre, agencement, fonction) des roues sont appliquées.

La très grande majorité de robots mobiles développés jusqu'à ce jour sont équipés de roues plutôt que de jambes.

Les robots à roues sont plus populaires parmi le marché des consommateurs pour plusieurs raisons :

- dissipent moins d'énergie
- Ils sont plus bien contrôlés que les autres types de robots.
- trois roues sont suffisantes pour l'équilibre statique et dynamique.

Les robots à roues sont utilisés lorsque le roboticien veut focaliser son énergie et son temps sur un autre aspect que le déplacement ou lorsque le projet de recherche porte sur autre chose que :

- la reconnaissance de forme
- la saisie d'un objet
- l'interaction sociale

Par contre, Les roues comportent aussi un certain nombre d'inconvénients face aux mécanismes à jambes ; Entre autres, elles ne sont utilisables que sur des terrains relativement plats et dont la surface est dure.

Dans le cas de surfaces moins stables, les robots à roues ne peuvent pas naviguer bien sur des obstacles, tels que les terrains rocheux, de fortes baisses, ou les zones à faible friction.

Les roues peuvent glisser ou s'enliser, ce qui pose des difficultés de contrôle importantes.

Sur des terrains accidentés, les roues doivent être plus grandes que les obstacles qu'elles risquent de rencontrer. [13]

Les types :

- **Robots à 2 roues :**

Le Roomba (figure 15) est un robot aspirateur fabriqué par la société i Robot il a la forme d'un disque de 34 cm de diamètre et 9 cm de haut. Un large détecteur de choc ainsi qu'un capteur infrarouge sont montés sur la partie avant. Une poignée et les boutons de contrôle sont logés sur la face supérieure de l'appareil. La face inférieure comporte deux roues principales à suspension et le mécanisme d'aspiration composé de deux brosses contrarotatives. La partie arrière accueille le bac à poussière amovible. [14]



Figure 15:Roomba 980[14]

- **Robots à (3 roues):**

Le Rovio (figure 16) est un robot de surveillance mobile contrôlable depuis n'importe où dans le monde pour peu que vous ayez une connexion internet.

Cette webcam wi-fi dispose de 3 roues omnidirectionnelles, d'un bras articulé permettant de contrôler l'orientation de la caméra, et une fois les batteries presque vides, il retourne se recharger automatiquement sur sa base.

Il peut faire office de gardien également car il dispose d'un détecteur de mouvement et prend une photo qu'il l'envoie par mail.

Il dispose également d'une lampe qui permet de voir, ou du moins s'orienter dans le noir. Enfin, il dispose d'un micro qui permet de contacter vos proches lorsque vous êtes en déplacement. [15]



Figure 16:ROVIO [15]

- **Robot à 4 roues:**

Le Robot-tondeuse Lizard S14N (figure 17) Fonctionnant sans câble de délimitation de terrain, la robot-tondeuse Lizard S14N intègre une technologie faisant intervenir 6 capteurs aptes à reconnaître toutes sortes de surfaces.

A chaque fois qu'il ne détecte plus d'herbe ou qu'il est face à un obstacle, cet appareil change de direction grâce à ses 4 roues motrices. Ce modèle se déplace aisément et rapidement.

Pouvant intervenir sur 400m², cette tondeuse robotique est adaptée aux petits jardins.

Il est doté d'une lame longue de 25 cm garantissant une hauteur de coupe de 34 à 40 mm. Avec seulement 3 heures de charge dans un endroit sec et abrité, ce modèle est prêt à l'utilisation. Il fonctionne grâce à une batterie Lithium. [61]

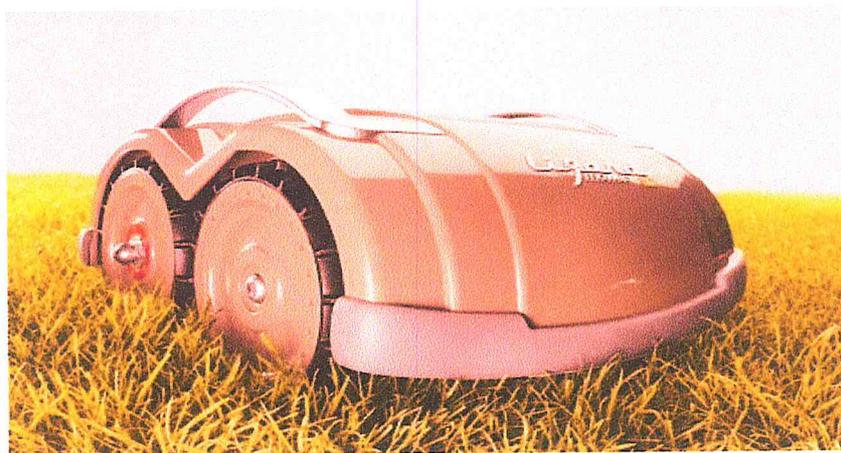


Figure 17: Robot-tondeuse Lizard S14N[61]

Le robot B21r par i-Robot : (figure 18)



Figure 18: Le Robot B21r

5. CONCLUSION

Ce pan de la robotique est l'objet de la présente étude, et mérite qu'on s'y attarde dans la suite du texte.

À ce titre, la robotique ludique dans toutes ses composantes fait partie du champ considéré dans le cadre de cette étude.

Dans le prochain chapitre nous allons parler de la détection des formes et ses différentes méthodes, ensuite nous allons entamer la détection de la pose et les diverses solutions existantes puis nous allons évoquer la caméra Kinect de Windows et son fonctionnement.

Enfin nous allons aborder la détection du visage et les différentes méthodes utilisées.

CHAPITRE 2 :
LA DETECTION DE POSE DE
L'ETRE HUMAIN

1. INTRODUCTION

L'estimation de la pose d'une personne et son suivi dans une séquence vidéo consiste à déterminer les coordonnées spatiales de tout ou partie des membres du corps pour chacune des images.

C'est un objectif primordial en vision par ordinateur et les applications en sont nombreuses.

Elles concernent :

- les interfaces homme/machine
- la communication au travers d'avatars
- l'analyse technique des gestes artistiques et sportifs
- la surveillance des lieux.

La paramétrisation de la pose humaine, même simplifiée comprends plusieurs dizaines de paramètres ; les gestes parfois rapides, les vêtements amples, les occultations ou encore les changements dans la scène tels que les mouvements dans l'arrière plan ou les variations de luminosité, contribuent à faire de l'estimation ou du suivi de la pose un défi scientifique.

Dans notre travail nous nous intéressons en particulier à la détection de l'être humain pour un robot guide intelligent et interactif, capable d'exécuter de façon autonome des tâches complexes comme renseigner les usagers (affichage, son, ...), ou les conduire à une destination demandée. [17]

La détection de l'être humain est la détection de tout ce qui puisse nous renseigner sur la personne notamment : la détection de sa silhouette, la détection de sa position, la détection de son visage.

Nous allons présenter dans la partie qui suit les différentes méthodes de la détection de formes.

2. DETECTION DES FORMES

La détection de formes se divise en deux parties :

2.1 Détection des objets

A partir des travaux présentés dans [62], [63], [64], nous présentons une classification des méthodes proposées jusqu'à présent pour la détection d'objets en mouvement.

En fonction de la modélisation de l'arrière-plan, nous distinguons deux grandes familles de méthodes : Détection sans modélisation de l'arrière-plan et détection avec modélisation de l'arrière-plan.

La figure 19 illustre une classification possible des méthodes proposées jusqu'à présent pour la détection d'objets en mouvement. [18]

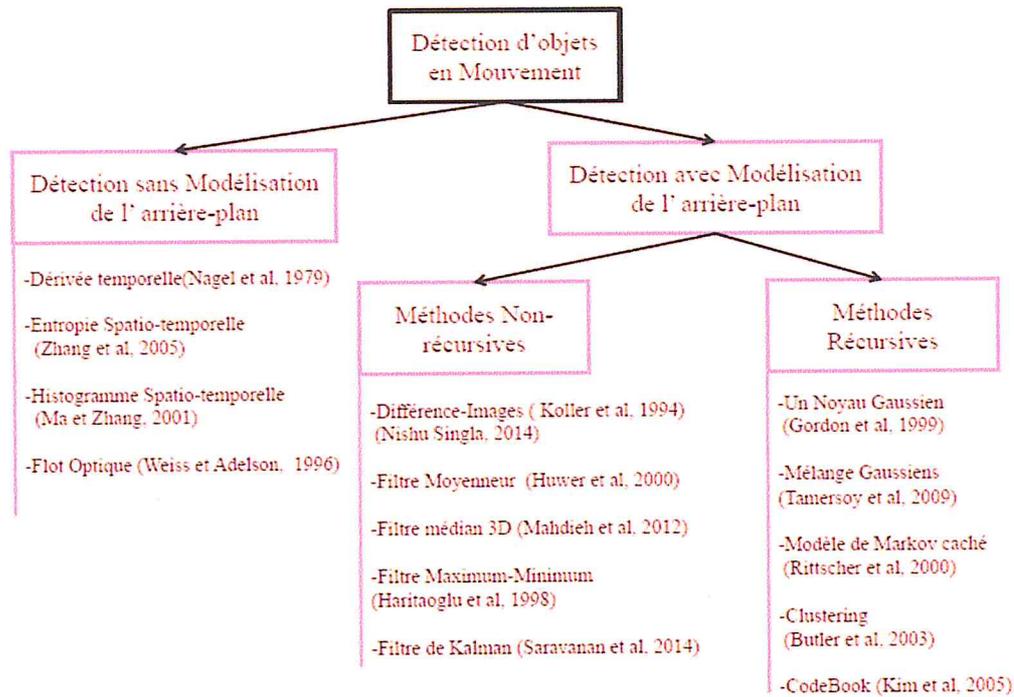


Figure 19: Classification des méthodes de détection d'objets [18]

2.1.1 Détection sans modélisation de l'arrière-plan :

Ce type de détection est basé sur des mesures simples pour estimer la variation temporelle de la couleur entre deux images consécutives comme : l'histogramme spatio-temporelle et l'entropie spatio-temporelle.

Ces méthodes sont souvent classées comme des estimateurs de mouvement que des détecteurs de mouvement.

Dans cette section nous présentons en détail les méthodes les plus utilisées dans la détection de mouvement sans modélisation d'arrière-plan.

a- Dérivée temporelle :

Cette méthode aussi connue comme la différence inter-image [Radke 2005].

Elle permet de détecter les régions en mouvement et de mesurer le changement de la luminosité inter-pixel entre deux images (Frames) consécutives d'une séquence de vidéo.

Malgré la simplicité et l'efficacité de cette méthode dans des applications et des conditions très spécifiques, elle reste insuffisante dans des conditions réelles d'application, par exemple en présence : de mouvement lent ou saccadé. Cette méthode est la base de toutes les autres méthodes proposées dans la littérature pour la soustraction de fond. [18]

b-Flot optique

En physique, le flot optique (en anglais, Optical Flow) est le mouvement apparent des objets, surface et contours d'une scène visuelle causée par le mouvement relatif entre l'observateur (œil humain ou camera) et la scène observée [Radke 2005].

En traitement et analyse des vidéos, le flot optique est défini comme un champ de vecteurs à 2D inter-pixel pour passer d'une image dans une séquence vidéo à une autre au fil du temps. L'estimation de mouvement en 2D à travers le mouvement apparent est une tâche très complexe parce que le mouvement apparent est très sensible par rapport aux bruits et au changement de la luminosité. Plusieurs méthodes ont été proposées dans la littérature pour l'estimation du flot optique inter-pixel allant des méthodes différentielles ; vers des méthodes basées sur la corrélation. [18]

La méthode différentielle proposée par Lucas et Kanade [Baker 2011] est la plus utilisée par rapport aux autres méthodes existantes et cela pour sa robustesse et sa rapidité.

2.1.2 Détection avec modélisation de l'arrière-plan

La soustraction de fond (en anglais, Background subtraction) est une classe des techniques d'extraction d'objets en mouvement utilisées dans plusieurs applications comme la surveillance.

La validité de ces techniques reste conditionnée par l'arrière-plan utilisé, le développement des arrière-plans valide est une tâche très complexe :

- Premièrement, l'arrière-plan doit être robuste au changement de la luminosité et aux mouvements lents.
- Deuxièmement, il faut éviter la détection des objets mobiles dans le fond et les ombres projetées par les objets mobiles.

Un bon modèle de fond devrait rapidement répondre aux changements en arrière-plan et s'adapter automatiquement aux changements survenant dans le fond.

Dans la littérature, plusieurs méthodes ont été proposées pour la modélisation de l'arrière-plan.

Ces méthodes peuvent être décomposées en deux grandes familles : les méthodes récursives et les méthodes non-récursives. [18]

Les méthodes récursives, ne nécessitant pas une trace de l'historique de la luminosité de chaque pixel.

Les méthodes non-récursives, nécessitent un buffer pour garder une trace de l'historique de la luminosité de chaque pixel.

Dans cette partie, nous présentons en détail quelques méthodes, les plus recommandées jusqu'à présent pour la modélisation de l'arrière-plan.

a-Modélisation récursive

Simple gaussien

La méthode de modélisation de l'arrière-plan par un simple gaussien permet de modéliser la distribution temporelle de l'intensité de chaque pixel par une simple distribution gaussienne.

Pour chaque pixel dans l'image courante d'une séquence vidéo, le modèle utilise les valeurs des pixels précédents comme des échantillons pour créer une fonction de distribution de probabilité caractérisée par une moyenne et un écart type.

Si l'intensité des pixels précédents de chaque image qui compose la séquence ont été stockés pour calculer la moyenne et l'écart type, alors la mémoire nécessaire pour cette méthode est très importante ce qui limite l'application de cette méthode en temps réel.

Un simple seuillage est utilisé par la suite pour classier les pixels avant-plan ou arrière-plan, les pixels classifiés comme avant-plan représentant le masque d'objets en mouvement. La méthode de soustraction entre l'arrière-plan estimé et l'image courante pour extraire l'avant-plan.[18]

La modélisation de l'arrière-plan par un simple gaussien est robuste si et seulement si la dynamique de la scène est inférieure par rapport au changement introduit par les objets en mouvements.

Cette méthode échoue quand un mouvement rapide aura lieu dans l'arrière-plan, ce qui limite la robustesse de cette méthode dans des conditions réelles d'application.

Mélanges gaussiens

La modélisation de l'arrière-plan (image de référence) par un simple gaussien est insuffisante surtout lorsqu'un changement rapide et important aura lieu dans l'arrière-plan. Dans ce cas une distribution multimodale est nécessaire pour décrire la dynamique de la scène.

Face à ce problème, les auteurs de [Stauffer 2000], proposent une modélisation de l'arrière-plan par un mélange gaussien (en anglais, Mixture of Gaussian).

Cette modélisation qui permet la représentation de l'image en arrière-plan avec une distribution multimodale.

Dans un mélange gaussien, la distribution temporelle au fil du temps de l'intensité de chaque pixel est modélisée.

Le mélange gaussien est caractérisé par une moyenne, une matrice de covariance et une probabilité a priori [18]

Malgré la recommandation de cette méthode par plusieurs chercheurs pour la détection de mouvement, les résultats d'application de cette méthode restent insuffisants dans des conditions réelles d'application en particulier en présence de changements d'éclairage, mouvement dans le fond de la scène et mouvement non-rigide [Zang 2006], [Bouwman 2008].

CodeBook

"CodeBook" ou en français, Livre de code, est une méthode proposée par [Kim 2005] pour la modélisation de l'arrière-plan.

Cette méthode est basée sur une technique de quantification, où l'historique d'observations de chaque pixel dans une séquence vidéo est utilisé pour construire un codebook pour chaque pixel.

Ce codebook est caractérisé par un ou plusieurs codewords.

Le nombre de codewords par pixel dépend de la variation du fond.

Initialement, le codebook de chaque pixel est vide et ne contient pas des codeword's. Quand un échantillon de pixel est observé alors cet échantillon sera classé comme un codeword dans le codebook, et les valeurs seront utilisées pour estimer la mesure qui représente le codeword.

Si un codeword existe déjà dans le codebook alors l'échantillon de ce nouveau pixel est comparé avec chaque codeword dans le codebook utilisant la mesure d'altération de la couleur.

Cette méthode a été prouvée dans le même travail comme robuste par rapport au mouvement rapide et au changement d'éclairage.

D'autres chercheurs [Ilyas 2009] prouvent l'insuffisance de cette méthode dans certaines conditions réelles d'application comme : la présence de mouvement lent dans le fond, la présence d'objets avec des couleurs semblables, etc.

Dans le même cadre de la modélisation réursive d'arrière-plan pour la détection de mouvement, plusieurs méthodes ont été proposées.

Cet axe reste très actif à l'heure actuelle et en développement remarquable ces dernières années. [18]

Plusieurs autres méthodes pour la modélisation de l'arrière-plan ont été proposées comme la modélisation par un modèle de Markov cachée [Rittscher 2000], les méthodes prédictives [Koller 1994].

b-Modélisation non-réursive

Les méthodes non-réursives de la modélisation de l'arrière-plan se basent essentiellement sur l'historique de l'intensité des pixels afin de construire un modèle robuste de l'arrière-plan. Dans cette orientation il y a eu beaucoup d'investissement et plusieurs méthodes ont été proposées.

La forme la plus basique d'un modèle d'arrière-plan est l'utilisation des images de références [François 1999] ou de l'image précédente [Radke2005].

Ces méthodes restent très limitées dans plusieurs conditions comme : le mouvement rapide, le changement d'éclairage, la stabilité et la non- stabilité d'objets, etc.

Face à ces problèmes d'autres chercheurs comme dans [François 1999], proposent la modélisation de l'arrière-plan utilisant un filtre moyenné où la valeur de chaque pixel de l'arrière-plan courant est égale à la moyenne de toutes les valeurs des pixels gardés dans l'historique. Dans ce même contexte d'autres chercheurs [Cucchiara 2003] proposent l'utilisation d'un filtre de type médian pour la mise à jour de l'arrière-plan. Le problème de ces méthodes est que les informations des pixels proviennent non seulement de l'arrière-plan précédent mais aussi de l'avant plan précédent.

Un filtre de Kalman est aussi utilisé par [Weng 2006], où l'arrière-plan courant est calculé par la prédiction de la valeur des pixels en utilisant l'historique.

Le grand souci de ces méthodes c'est le temps de calcul et le changement d'éclairage.

D'autres chercheurs [Haritaoglu 1998], proposent la modélisation de chaque pixel par une variable aléatoire dans un espace de caractéristiques spécifiques puis l'associe à une fonction de distribution de probabilité. L'estimation de cette distribution de probabilité se fait par un estimateur à noyau.

Cette méthode est considérée comme une forme générale de la mixture gaussienne non-paramétrique. Sa robustesse est strictement conditionnée par la dimension des données et l'espace des caractéristiques utilisées. La modélisation non récursive de l'arrière-plan pour la détection de mouvement a fait l'objet de plusieurs recherches. Ces méthodes nécessitent une grande mémoire et une capacité de calcul importante. [18]

2.2 Détection de l'humain

La **détection de personnes** est un domaine de la vision par ordinateur consistant à détecter un humain dans une image numérique où l'on cherche à détecter la présence et la localisation précise, dans une image, d'une ou plusieurs personnes, en général dans une posture proche de celle de la station debout ou de la marche.

On parle également de détection de piéton, en raison de l'importance des applications en vidéosurveillance et pour les systèmes de vision embarqués dans des véhicules.

Étudiée à partir de la fin des années 1990, la détection de personne s'est révélée être un sujet assez difficile, en raison de la grande variété d'apparences des personnes, de l'articulation du corps humain et des phénomènes d'occultations.

Il est facile pour un humain de reconnaître un autre humain autour de lui ou sur une image mais c'est un problème très complexe pour un système automatisé.

Pourtant, beaucoup de systèmes ont besoin d'avoir des informations sur la présence ou l'absence de personnes dans leur environnement.

Les applications potentielles d'un système de détection automatique de la présence humaine sont nombreuses.

Nous pouvons par exemple citer les systèmes de transport intelligent, la robotique, la surveillance, la domotique intelligente, l'indexation d'images ou de vidéos,

D'une manière générale, le principe de la détection de personnes dans une image ou une vidéo est le même. [20]

Cependant, pour une image, le détecteur ne dispose d'aucune information a priori et doit donc parcourir l'image entière avec une fenêtre de détection.

Pour une vidéo, il est possible de simplifier le problème en se focalisant par exemple sur les zones de l'image où il y a eu un mouvement. Il est également possible de raisonner en fonction de la position des personnes présentes dans les images précédentes ou même de caractériser les personnes par leur mouvement.

Plusieurs méthodes sont utilisées pour détecter une silhouette parmi ces méthodes nous citons :

2.2.1 Par détecteurs

Différentes technologies existantes sur le marché : [21]

Détecteur à infrarouge (IR) : Ils détectent le mouvement du corps humain par la mesure du rayonnement infrarouge émis par le corps humain.

Ils fournissent une indication de changement d'occupation d'un lieu : absence ou présence.

Ils ne permettent pas de connaître le taux d'occupation d'un local ou le nombre d'occupants.

Détecteurs ultrasoniques (US) : Les détecteurs US sont de type émetteur/récepteur et fonctionnent sur le principe de l'effet Doppler.

Le détecteur est capable de mesurer l'écart de fréquence et de générer ainsi un signal de présence.

Les détecteurs US ont une portée limitée mais peuvent détecter des mouvements mineurs et ce même autour de certains obstacles.

Détecteurs à double technologie : Les détecteurs de présence à infrarouges risquent de ne pas détecter les mouvements légers.

Par contre des détecteurs à ultrasons peuvent être trop sensibles et risquent de déclencher l'allumage de l'éclairage lors du passage "d'une mouche".

Pour éviter cet inconvénient tout en gardant une sensibilité importante, certains détecteurs, appelés "détecteurs à double technologie" combinent ultrasons et infrarouge. Cette combinaison permet d'augmenter la fiabilité des détecteurs et élimine les détections indésirables.

2.2.2 Par Caméras

Pour détecter un être humain 3 étapes sont nécessaires :

-**Soustraction de l'arrière-plan** : filtrer l'image en séparant l'avant-plan de l'arrière-plan avec une distribution gaussienne.

-**Suivi** : Après avoir filtré l'image en séparant de l'avant-plan de l'arrière plan, nous regroupons en composantes connectées les pixels de l'avant-plan.

À chaque instant nous disposons donc de la liste des composantes connectées présentes et de la liste des objets suivis. [22]

-**Classification** :

- un classifieur de la partie supérieure du corps humain, vu de face ou de dos.
- un classifieur de la partie supérieure du corps humain, vu de gauche.
- et finalement un classifieur de la partie supérieure du corps humain, vu de droite.

Finalement, les résultats des trois classifieurs pour construire un indice de confiance de l'appartenance à la classe humain de l'objet suivi. Cet indice évoluera dans le temps selon les résultats de détection à chaque instant.

3. DETECTION DE LA POSE

L'estimation de la pose de l'être humain revient à déterminer la configuration des parties du corps de l'être humain à partir de l'image (les épaules, les genoux, les hanches, les chevilles, les coudes, les mains et la tête). [23]

La Détection de la pose humaine est un sujet très intéressant et difficile en vision par ordinateur en raison de sa large gamme d'application. La détection de la pose peut être considérée comme un sous-champ de reconnaissance de geste, tant que la pose est déjà un "geste statique".

Il existe un large domaine d'application de l'estimation de la pose de l'être humain notamment :

- les jeux vidéo et la réalité virtuelle.
- la télésurveillance pour détecter les activités anormales.
- la robotique pour le contrôle des robots.
- Le domaine sportif pour analyser les performances des athlètes sportifs.
- Le domaine médical pour les diagnostics orthopédiques des patients.

-Le trafic routier pour contrôler les airbags, détection de sommeil du conducteur, détection des piétons.

Il existe plusieurs techniques de détection de pose de l'être humain, la prochaine partie illustre ces différentes techniques.

4. SOLUTIONS EXISTANTES

La capture de mouvements ou captation de mouvements (motion capture ou mocap) est une technique permettant de capter les mouvements d'un élément réel afin de les renvoyer dans un univers virtuel : les mouvements sont enregistrés ou restitués en temps réel vers d'autres systèmes.[25]

On distingue deux grandes familles de détection :

4.1 Techniques invasives (avec marqueurs)

Les techniques de capture de mouvements envahissantes sont habituellement plus précises.

Cependant, elles nécessitent que le sujet porte des senseurs, marqueurs et/ou accéléromètre. . [25]

Toutes les données sont recueillies en temps réel et une animation est générée. Le but ultime habituellement visé par les méthodes envahissantes est de recréer le plus fidèlement les mouvements d'un sujet. Par contre, elles sont généralement très dispendieuses et le port d'équipements spécialisés peut encombrer les mouvements du sujet. Finalement, les animations générées, malgré qu'elles soient très précises, nécessitent habituellement des retouches d'artistes lorsqu'elles sont utilisées dans des applications de multimédia ou dans le cinéma, car le sujet ne produit pas nécessairement le mouvement espéré.

Il existe deux grands types de captures de mouvements avec marqueurs :

- Le premier type utilise des points de références colorés afin de situer les points-clés dans l'espace (voir figure 20). Cette technique utilise des caméras RGB à haute résolution afin de suivre les points-clés. Dans cette technique, le sujet sera également habillé de façon à augmenter le contraste entre lui et les points suivis, afin de mettre ceux-ci en évidence. Par exemple, trois points blancs peuvent être fixés à chaque joint d'un sujet habillé en noir. Plusieurs caméras observeront le sujet et permettront le suivi de la position et la rotation de chaque joint.

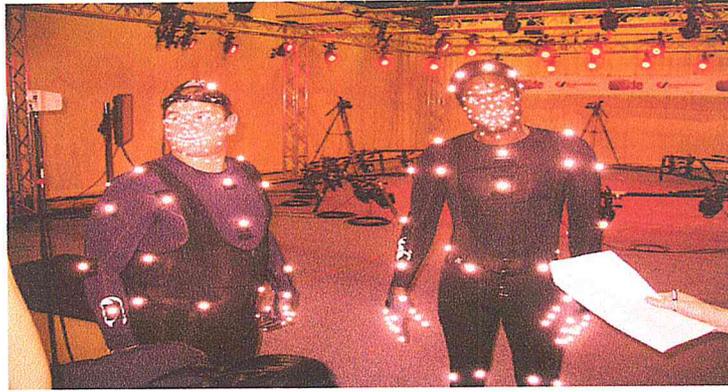


Figure 20: capture de mouvements avec marqueurs. [25]

- Le deuxième type de capture de mouvement avec marqueur utilise des accéléromètres, gyroscopes et un système de positionnement pour chaque joint du sujet. Les senseurs envoient directement leur information de positions et de rotations à un poste de travail.

Cette technique a pour avantage de ne pas nécessiter que les joints du sujet soient dans le champ de vision d'une caméra. (Voir figure 21)



Figure 11: capture de mouvements avec accéléromètres. [25]

4.2 Capture sans marqueurs

Les techniques de capture de mouvement sans marqueur (habituellement appelées *Markerless Motion Capture*) sont souvent basées sur des caméras. Aussi, depuis la venue de la caméra *3DKinect de Microsoft*, plusieurs solutions sont basées sur les caméras de profondeur.

Une technique largement utilisée avec des caméras RGB normales demande une calibration initiale, permettant de détecter l'arrière-plan. Ensuite, cet arrière-plan peut être soustrait des images nouvellement capturées afin d'isoler le sujet capturé.

Cette technique a été jumelée à la capture sonore de chaque caméra utilisée pour la synchronisation des images, le suivi de trait et la densité de détails de l'arrière-plan afin de stabiliser les résultats obtenus

Le squelette du sujet capturé est ensuite extrait des images capturées par la caméra

RGB et on peut en extraire les positions de chaque joint dans l'espace. Par contre, puisque les caméras RGB ne fournissent aucune information de profondeur sur l'image capturée, cette technique souffre des mêmes limites de précisions quant à la position des joints dans l'axe de profondeur.

La technique précurseur au suivi avec des caméras 3D utilise un modèle articulé du corps humain afin de préciser la capture de mouvement.

Cette technique offre une façon facile d'extraire le squelette des images de profondeur fournies par la caméra 3D.

Une approche utilisant cette technique consiste à générer une représentation voxel du modèle articulé à partir des images, qui a l'avantage d'être adaptable selon le niveau de précision requis grâce à la nature évolutive des voxels.

Par contre, la limite de la précision de la capture est toujours déterminée par la résolution en profondeur de la caméra utilisée. (Voire figure 22)

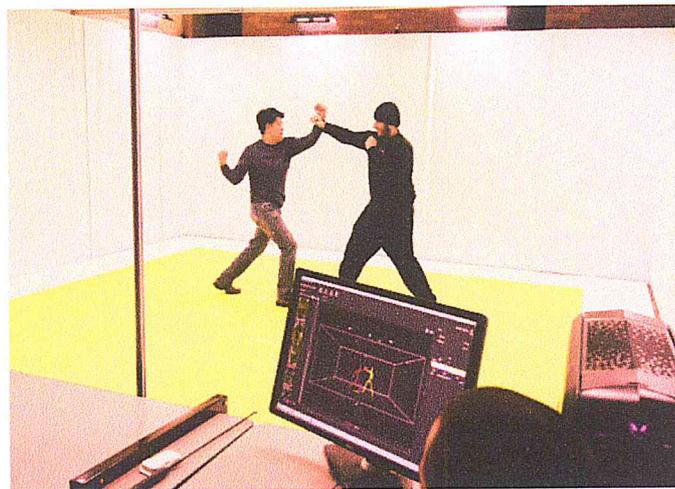


Figure 22: capture de mouvements sans marqueurs. [25]

5. CAMERA KINECT

La Kinect [26] est une caméra 3D créée initialement par Microsoft pour l'industrie des jeux vidéo. En effet, la Kinect permet à l'utilisateur de contrôler le jeu sans utiliser de manette mais uniquement grâce aux mouvements du corps. (voir figure 23)



Figure 23: caméra Kinect [26].

Devant l'étendue de possibilités offertes par la Kinect, des logiciels de développement « open-source » ont été créés pour manipuler la Kinect par ordinateur et créer des applications utilisant la Kinect.

5.1 Kinect en chiffre :

La Kinect est sortie le 4 novembre 2010 en Amérique du Nord et le 10 novembre 2010 en Europe. Cette nouvelle technologie est au protée de tous utilisateurs en raison de son prix très bas (150€).

Plus de 8 millions d'exemplaires de la Kinect ont été vendus partout dans le monde 2 mois après sa sortie, dont un million en seulement 10 jours. Le 11 mars 2011, la Kinect entre au livre Guinness des records comme étant « l'accessoire high-tech le plus vendu dans un court laps de temps » avec 10 millions d'unités vendues, soit une moyenne officielle de 133 333 Kinects vendues chaque jour à travers le monde.

Aujourd'hui, plus de 100 millions d'exemplaires ont été vendus.

5.2 Caractéristiques de la Kinect :

Dans ce qui suit nous allons détailler tous les composants de la Kinect ainsi son fonctionnement :

Capteurs :

- Caméra couleur et capteur de profondeur
- Micro à reconnaissance vocale

Motorisation :

- Ensemble motorisé pour suivre les déplacements
- Inclinaison: ± 27 degrés

Champ de vision :

- Champ de vision horizontal : 57 degrés
- Champ de vision vertical : 43 degrés
- Portée du capteur : 0,8 m – 4,0 m

Flux de données :

- Infrarouge 640x480 à 30 images par seconde
- RVB 640x480 à 30 images par seconde
- Audio 16 bits à 16 kHz

Système de reconnaissance physique :

- Jusqu'à 6 personnes et 2 joueurs actifs (4 joueurs actifs avec le SDK 1.0)
- 20 articulations par squelette

5.3 Fonctionnement de la Kinect

Comme le montre la figure 24, la Kinect est composée de plusieurs capteurs audio, d'une caméra RVB (Rouge, Vert, Bleu), d'un projecteur infrarouge, d'une caméra infrarouge, et d'un pied motorisé.

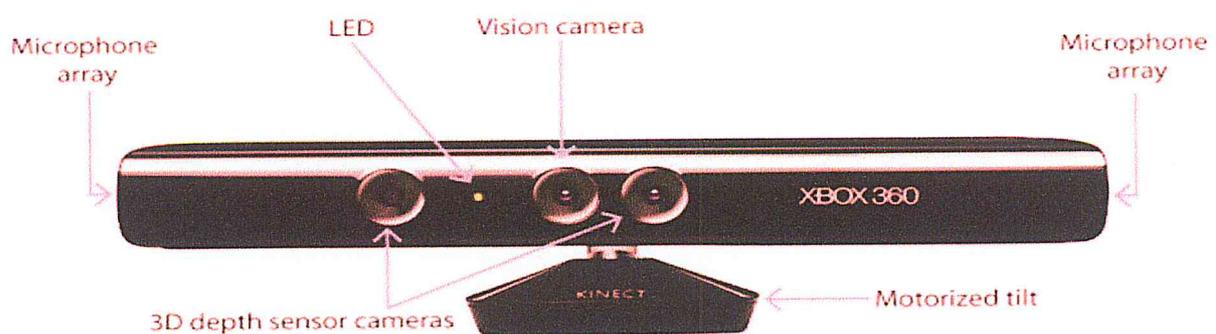


Figure 24: les différents composants de la Kinect

Grâce à tous ces composants la Kinect fournit trois flux différents qui sont le flux audio, le flux vidéo, et le flux de profondeur.

6. LA DETECTION DU VISAGE :

La détection de visage dans l'image est un traitement indispensable.

Le traitement consiste à rechercher dans une image la position des visages et de les extraire sous la forme d'un ensemble d'images dans le but de faciliter leur traitement ultérieur.

Un visage est considéré correctement détecté si la taille d'image extraite ne dépasse pas 20% de la taille réelle de la région faciale, et qu'elle contient essentiellement les yeux, le nez et la bouche [27]- [28].

Quelques appareils photos numériques récents emploient la détection de visages pour la mise au point automatique.

Elle est également recherchée dans le domaine des économies d'énergie ; les télévisions et les ordinateurs peuvent économiser l'énergie en réduisant l'éclat.

Le système peut identifier la direction de visage de l'utilisateur de TV. Quand l'utilisateur ne regarde pas l'écran, l'éclat de TV est abaissé, et quand le visage revient à l'écran, l'éclat est augmenté [29].

Les premières difficultés rencontrées par les méthodes de détections de visages sont les variations de pose (vue de profil ou de face), d'expression, de rotation du visage, d'âge et d'illumination.

6.1 Méthodes de détection de visage

Une classification des méthodes de localisation faciale a été proposée par Yang et al [28].

Les méthodes sont divisées en quatre catégories. Ces catégories peuvent se chevaucher si

Un algorithme peut appartenir à deux ou plusieurs catégories. Cette classification peut être faite comme suit :

-Knowledge-based methods :

Ces méthodes se basent sur la connaissance des différents éléments qui constituent un visage et des relations qui existent entre eux. Ainsi, les positions relatives de différents éléments clés tels que la bouche, le nez et les yeux sont mesurées pour servir ensuite à la classification 'visage' 'nonvisage' chez Chiang et al. [30]. Le problème dans ce type de méthode est qu'il est difficile de bien définir de manière unique un visage.

Si la définition est trop détaillée, certains visages seront ratés tandis que si la description est trop générale, le taux de faux positifs montera en flèche.

-Feature invariant approaches :

Ces approches utilisent les éléments invariants aux variations d'illumination, d'orientation ou d'expression tels que la texture ou la signature de couleur de la peau pour la détection.

- Template matching methods :

Des modèles caractéristiques d'un visage entier ou de sous-partie de visage (bouche, oeil, nez) sont créés.

La localisation se fait ensuite sur la base de la corrélation de ces modèles avec les candidats [31].

-Appearance-based methods :

Ces méthodes utilisent le même principe que présenté au point précédent mais se basent sur des modèles appris à partir d'un ensemble d'essai.

Ces méthodes présentent l'avantage de s'exécuter très rapidement mais demandent un long temps d'entraînement.

Les méthodes appartenant à cette catégorie ont montré de bons résultats par rapport aux trois autres types de méthodes [32].

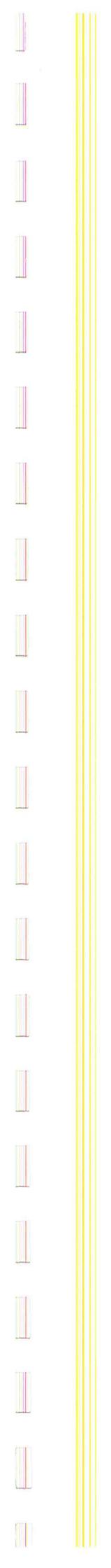
On peut citer parmi celles-ci, la méthode basée sur les réseaux de neurones de Rowley et al. [33], la méthode de Schneiderman et Kanade [34] basée sur un classifieur de Bayes naïf ainsi que le fameux algorithme de Viola et Jones [35] fonctionnant en temps réel.

7. CONCLUSION

Dans ce chapitre, nous avons vu les techniques de détection de la pose de l'être humain avec marqueurs ainsi que le principe de la détection de l'être humain par un capteur

RGB-D ou Kinect, qui revient en fait à détecter sa silhouette, cependant cette détection ne donne pas une idée sur l'orientation de son corps, cette information peut être récupérée par les techniques de détection du visage, pour cela nous avons présenté les techniques permettant la détection du visage de l'être humain.

Dans le chapitre suivant nous allons présenter l'aspect développement de notre solution.



CHAPITRE 3 :
DEVELOPPEMENT DE LA
METHODE DE DETECTION DE
POSE

1. INTRODUCTION

Dans ce chapitre nous allons présenter l'architecture software sous ROS (Robot Operating System) du robot B21r, cette architecture est appelée à s'enrichir et à intégrer de nouvelles fonctionnalités à chaque fois que l'on a besoins, dans le but d'augmenter le niveau d'intelligence et d'autonomie du robot.

Nous allons situer notre contribution dans la globalité du projet robot guide. Nous commençons par une description de la solution existante ensuite nous allons détailler la solution que nous avons proposé et implémenté

Cette solution sera utilisée dans une tâche de suivi de personne pour être par la suite implémentée dans une application de navigation sociale.

2. PROBLEMATIQUE

Malgré que les travaux de Shotton aient résolu une très grande partie du problème de la pose de l'être humain, l'orientation du corps par rapport à la Kinect reste ambiguë, les figures suivantes résumant ce problème.

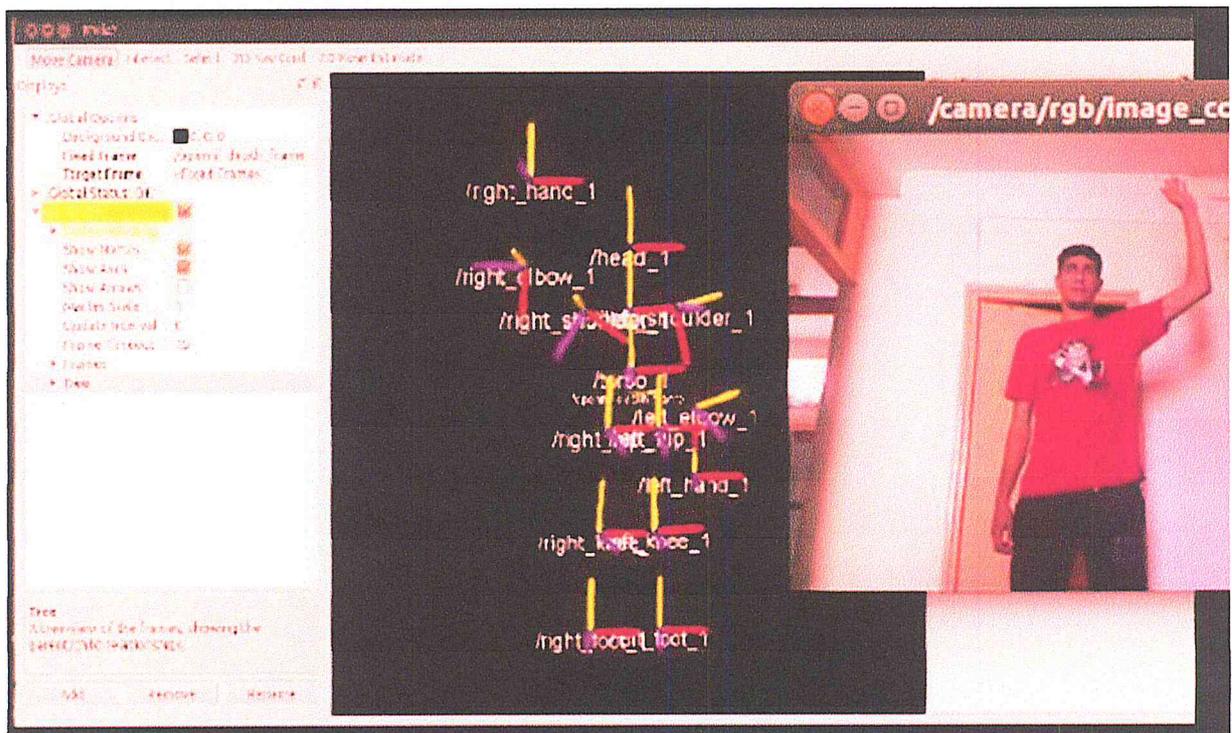


Figure 25: Visualisation de la main (de face) avec rviz

La figure ci-dessus montre une personne de face levant la main gauche, pour le squelette vu par la Kinect c'est la main droite qui est levée.

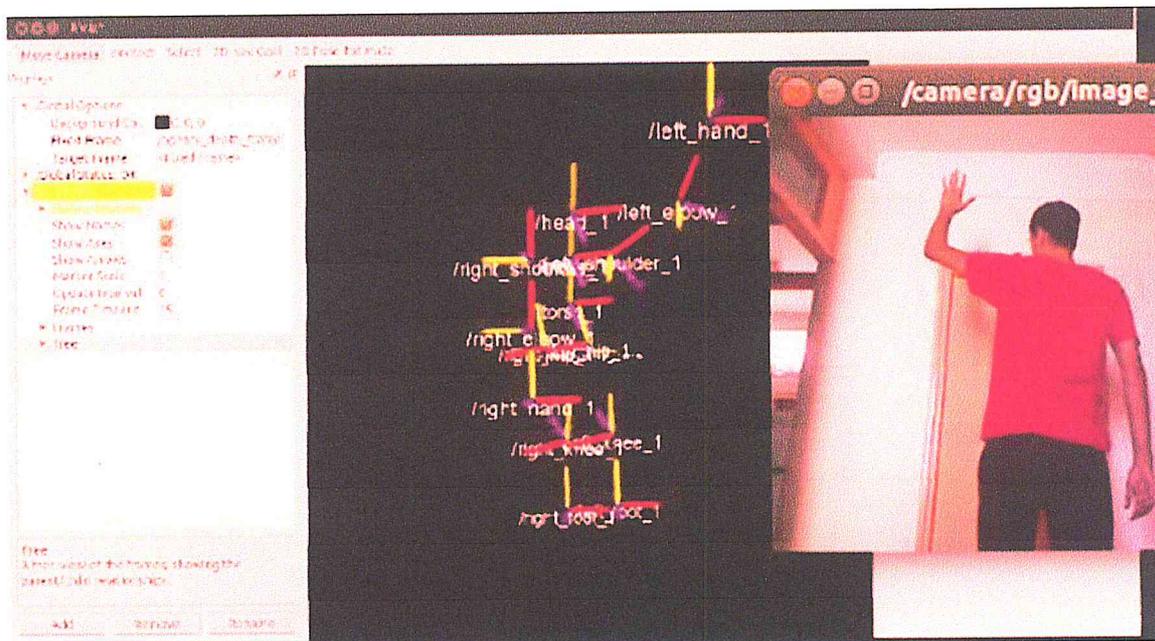


Figure 26: Visualisation de la main (de dos) avec rviz.

La même personne levant la même main (gauche) se met de dos, et le squelette vu par la Kinect affiche aussi la main gauche levée.

Nous pouvons voir par cette expérience que ce que nous affiche la Kinect (gauche ou droite) c'est par rapport à elle-même, non pas par rapport à la personne détectée.

Nous essayons d'apporter une solution à ce problème par la détection du visage pouvoir déterminer si la pose de la personne détectée est de face ou de dos et pouvoir déterminer par conséquence la pose gauche et droite.

3. ARCHITECTURE SOFTWARE DU ROBOT B21r

L'architecture software du robot B21r sous ROS est composée des nœuds suivants (voir figure27) :

-Le nœud **Hip_sk** publie dans le topic **TORSO**, par contre le nœud **Face_detect** publie dans le topic **HEAD**.

Les résultats publiés par les nœuds précédents sont utilisés par le nœud **Orientation** pour le calcul de la pose de l'être humain, le résultat de la pose est publié dans le topic **POSE**, qui sera utilisée par le nœud **Humantracker** pour permettre à l'être humain d'interagir avec le robot à travers sa pose, c'est cette partie de l'architecture qui nous intéresse dans notre travail.

-Le nœud **Rflex** : ce nœud gère le bas niveau du robot, il récupère les valeurs de vitesses linéaires et angulaires calculées par le nœud **Humantracker**, à partir du topic **Cmd_Velocity** pour commander le robot.

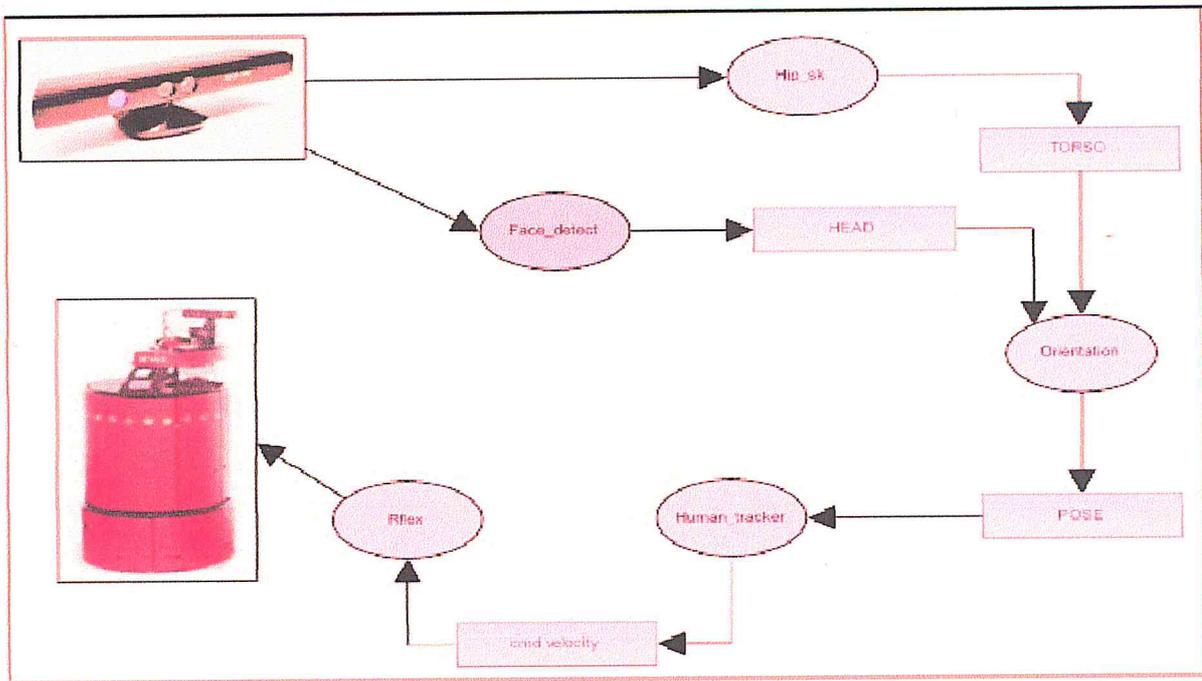


Figure 27: Architecture software du robot B21r sous ROS

4. DETECTION DE PERSONNE PAR UN ROBOT GUIDE

La figure 28 montre les différentes étapes pour détecter un être humain dans un scénario de navigation sociale.

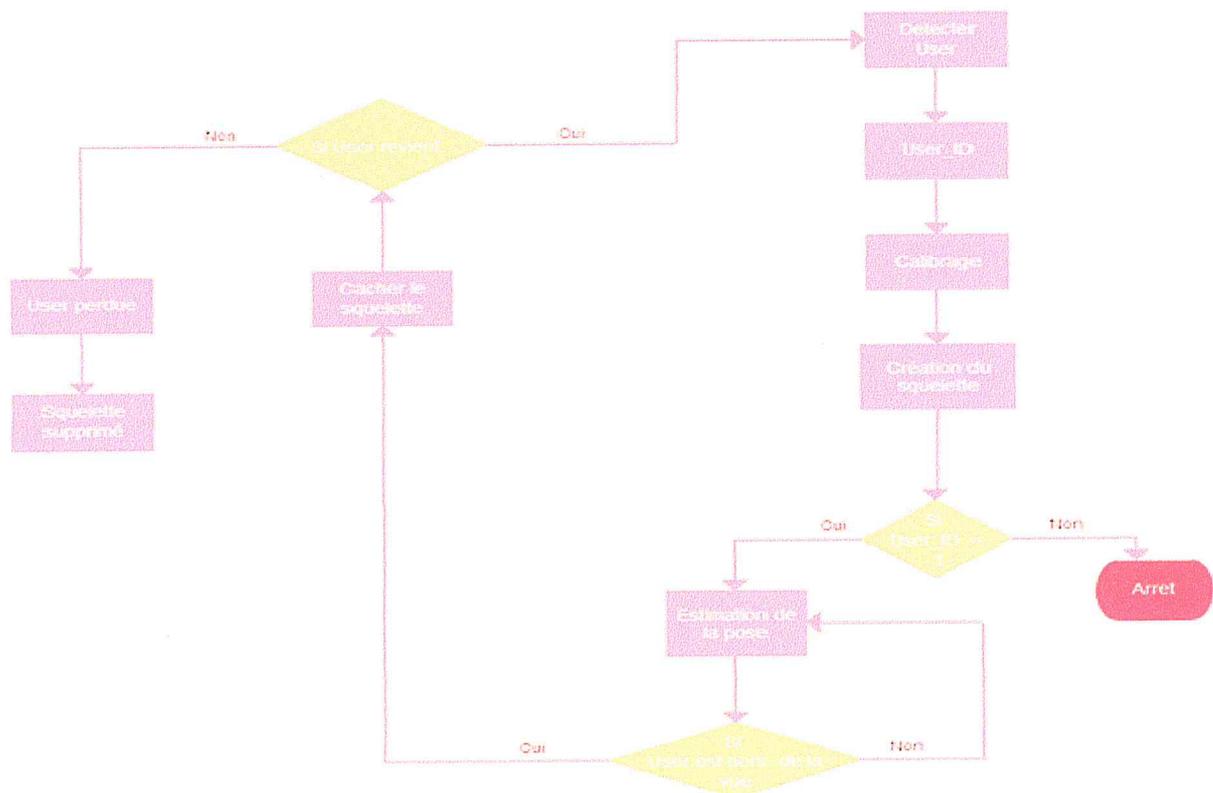


Figure 28: Organigramme de détection de personne par le robot

4.1 Détection : Initialement la personne à suivre doit se tenir debout en face de la Kinect en position Psi cette position est nécessaire pour que le Framework OpenNI procède au calibrage et à la segmentation des utilisateurs dans la scène. La distance idéale pour avoir une bonne détection est de 1.5 m.

A chaque utilisateur détecté, est affecté un identificateur (ID), ce dernier est libéré si l'utilisateur disparaît du champ visuel plus de 10 secondes.

4.2 Calibration : Cette étape a pour objectif, la reconnaissance et la sauvegarde des données de la première personne détectée dans le champ visuel de la Kinect, en localisant ses repères anatomiquement et en déterminant les longueurs approximatives de ses membres (bras et jambe, ..).

4.3 Création du squelette : Une fois un nouvel utilisateur est détecté, un nouveau squelette est créé et ajouté à la scène.

4.4 Estimation de la Pose : Cette étape représente l'objet de notre travail, qui a pour but de détecter 4 positions différentes de l'utilisateur qui sont : face, dos, gauche et droite, l'algorithme sera détaillé plus loin.

4.5 Mise à jour du squelette : Une fois que l'utilisateur quitte le champ de vision de la Kinect plus de 10 seconds, son squelette est supprimé de la scène et son ID est libéré.

Notre contribution dans ce travail concerne le nœud Orientation. En effet le travail que nous avons réalisé, est une amélioration d'un nœud existant.

5. PRINCIPE DE L'ALGORITHME D'UN POINT D'ARTICULATION

Le Principe de l'algorithme est d'étudier les rotations du Torse sur l'axe vertical (l'axe Y dans notre cas) pour pouvoir déterminer la pose de l'être humain détecté.

Les coordonnées de rotation standard pour le torse sont calculées en forme de matrice (Matrice de Rotation), avec le nœud `Hip_sk`.

Pour simplifier le calcul matriciel on a choisi de faire une conversion vers les angles d'Euler.

5.1 Description d'Angles d'Euler

Dans un système de représentation d'angle pour les données de mouvement 3D, les angles d'Euler et la transformation du système de coordonnées sont largement utilisés, pour obtenir un système de coordonnées à partir d'un autre système de coordonnées par trois rotations.

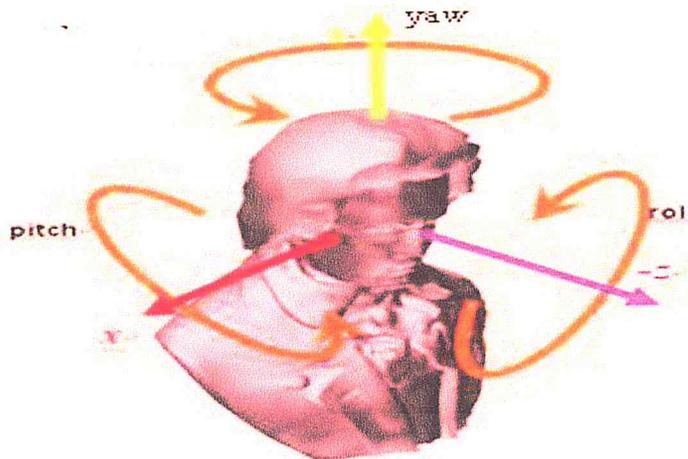


Figure 29: Représentation des Angles d'Euler

Les Angles d'Euler sont des concepts cinématique.

Ils sont utilisés pour identifier de manière unique la position de la rotation rigide du corps d'un groupe de trois paramètres angulaires indépendants, indiquée par l'angle de lacet (axe y), l'angle de tangage (axe x) et l'angle de roulis (axe z), comme le montre la figure 29

Les angles d'Euler sont un système de rotation autour d'un certain axe. Cependant, il ya une certaine règle d'ordre à respecter, la séquence de rotation d'angle d'Euler peut être Z - X - Y, X - Y - Z, X - Z - Y, etc.

Le système à suivre pour avoir les coordonnées quaternion est le suivant :

$$qw = \sqrt{(1 + m[0] + m[4] + m[8]) / 2}$$

$$qx = (m[7] - m[5]) / (4 * qw)$$

$$qy = (m[2] - m[6]) / (4 * qw)$$

$$qz = (m[3] - m[1]) / (4 * qw)$$

Après calculer les coordonnées quaternion on les a convertis vers les angles d'Euler car ces derniers sont faciles à implémenter, pour cela on a utilisé les équations suivantes :

```
double sqw = qw*qw;
double sqx = qx*qx;
double sqy = qy*qy;
double sqz = qz*qz;

double euler_y = (atan2(2.0 * (qx*qy + qz*qw), (sqx - sqy - sqz + sqw)) * (180.0f/pi));
double euler_x = (atan2(2.0 * (qy*qz + qx*qw), (-sqx - sqy + sqz + sqw)) * (180.0f/pi));
double euler_z = (asin(-2.0 * (qx*qz - qy*qw)) * (180.0f/pi));
```

Figure 30: Conversion Quaternion vers les angles d'Euler

5.2 Calcul de la pose

On peut déduire la pose de l'être humain détecté selon la variation d'Euler_Y qui a été déjà calculé, voir figure 31.

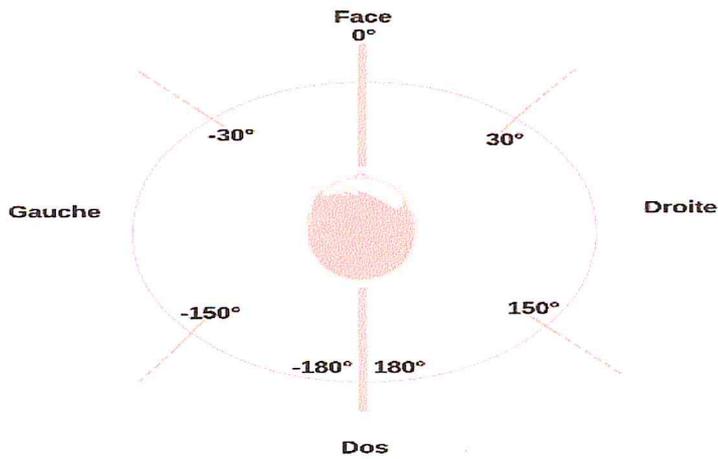


Figure 31: la Pose avec les rotations

5.3 L'organigramme

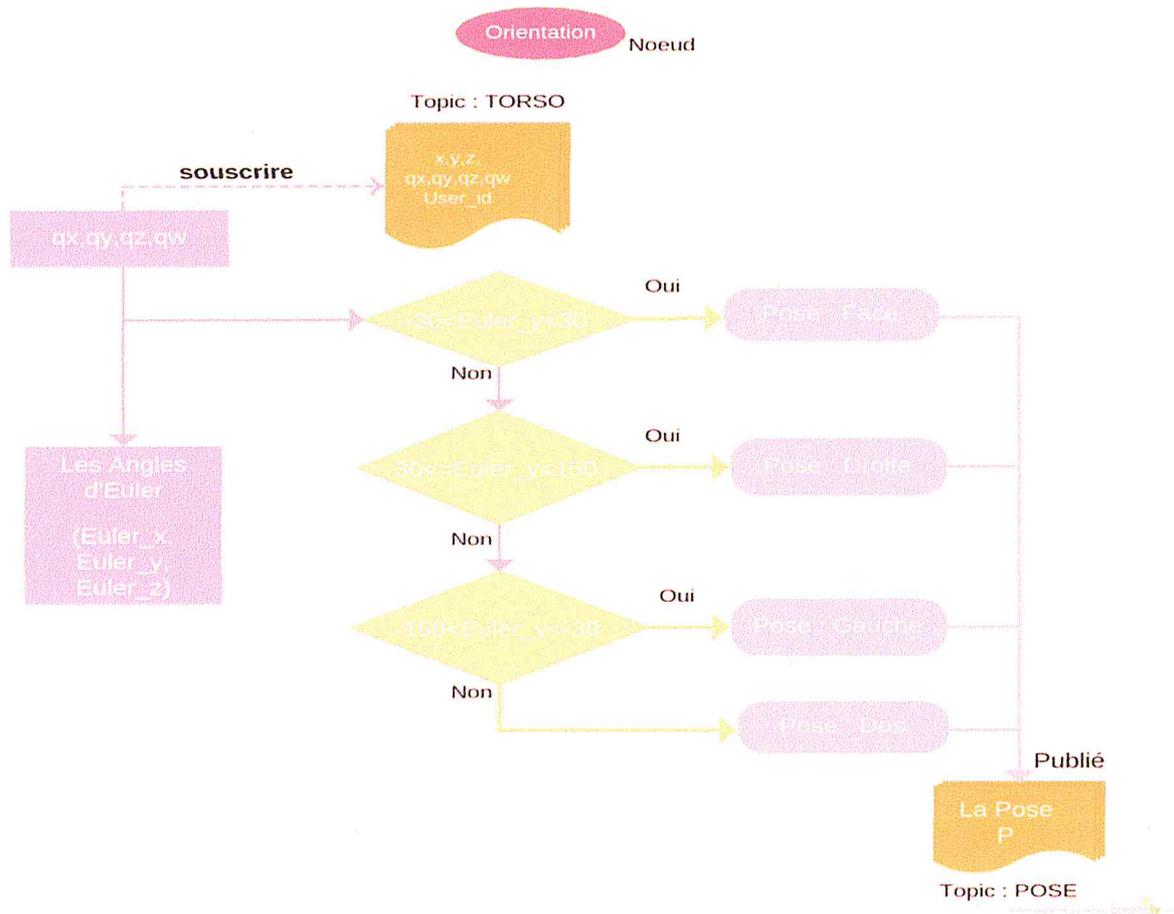


Figure 32: Organigramme pour la détection de la Pose

5.4 Résultats de l'algorithme

Après avoir testé le programme les résultats sont résumés dans le tableau suivant :

Pose	Face	Dos	Droite	Gauche
Test_Face	78%	15%	5%	2%
Test_Dos	10%	84%	2%	4%
Test_Droite	1%	3%	86%	10%
Test_Gauche	5%	3%	12%	80%

Tableau 1:Résultat de la détection de la pose avec l'algorithme d'un point d'articulation

Nous constatons d'après le tableau ci-dessus que la fiabilité des résultats est assez acceptable pour la détection des poses.

Notre contribution donc se situe à ce niveau dans le but d'améliorer la solution existante.

6. ALGORITHME PROPOSE

Comme nous l'avons dit précédemment notre solution sera une amélioration de l'algorithme expliqué au dessus, donc on va présenter un organigramme général du programme et puis on va le détailler.

Les entrées de notre algorithme sont les positions (X, Y, Z) des épaules publiées par le nœud `Hip_sk`, ainsi que la largeur et la hauteur du visage fourni par le nœud `Face_detect` tout en gardant les rotations du Torse sur l'axe vertical.(figure33)

```
g_UserGenerator.GetSkeletonCap().GetSkeletonJointPosition(user, XN_SKEL_RIGHT_SHOULDER, joint_position);
double x1 = joint_position.position.X;
double y1 = joint_position.position.Y;
double z1 = joint_position.position.Z;

position.x1 = x1;
position.y1 = y1;
position.z1 = z1;
```

Figure 33:Récupération des coordonnées de l'épaule droite

Les coordonnées sont exprimées dans le repère de la Kinect qui est représenté dans la figure 34

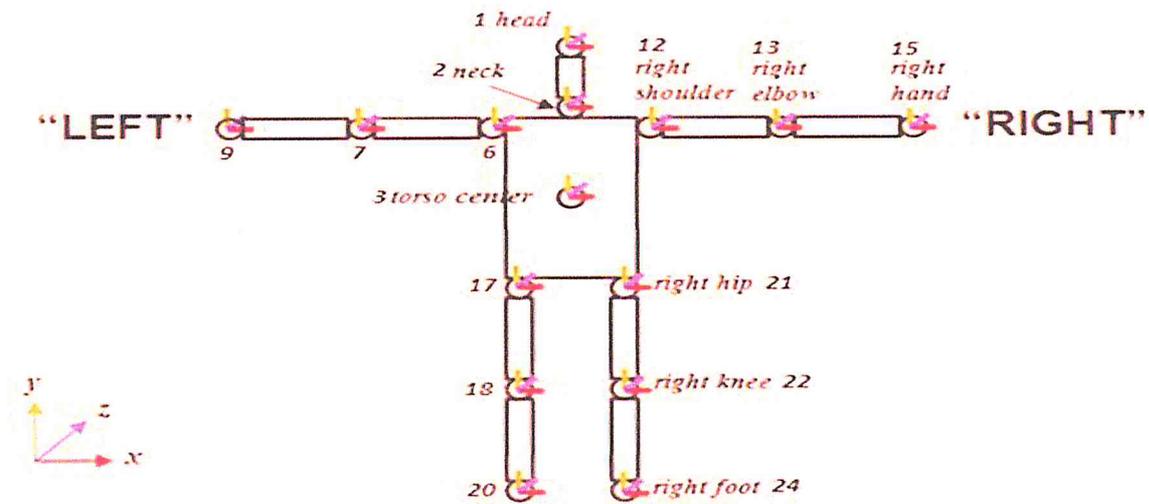


Figure 34: Représentation du repère de la Kinect.

Le principe de l'algorithme est basé sur les trois articulations définissant l'épaule gauche, l'épaule droite, le torse, et la détection du visage.

6.1 Calcul de l'orientation du corps

Lorsque l'être humain est en face à la Kinect, ses épaules sont parallèle à l'axe des x, et donc la distance dx entre les projections des deux articulations (épaule gauche et épaule droite) est maximale et elle est égale à la distance entre ses épaules par contre sa projection sur l'axe des z donne une valeur nulle comme la montre la figure 35.

Cette propriété permet de détecter si la personne se trouve parallèle à la Kinect (face ou dos) ou alors elle est de profil.

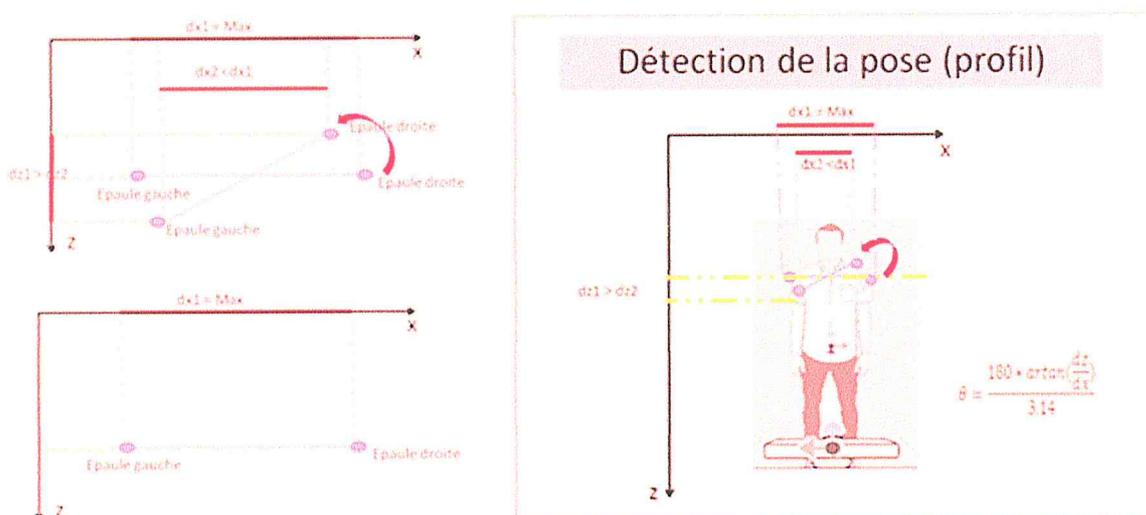


Figure 35: Projections des coordonnées des épaules sur le plan (oxz).

Une fois la pose de l'être humain est détectée face ou dos, une hypothèse est posée par l'algorithme pour différencier entre les deux poses, cette hypothèse se base sur la comparaison entre les coordonnées X des deux points épaule gauche et épaule droite voir figure 36.

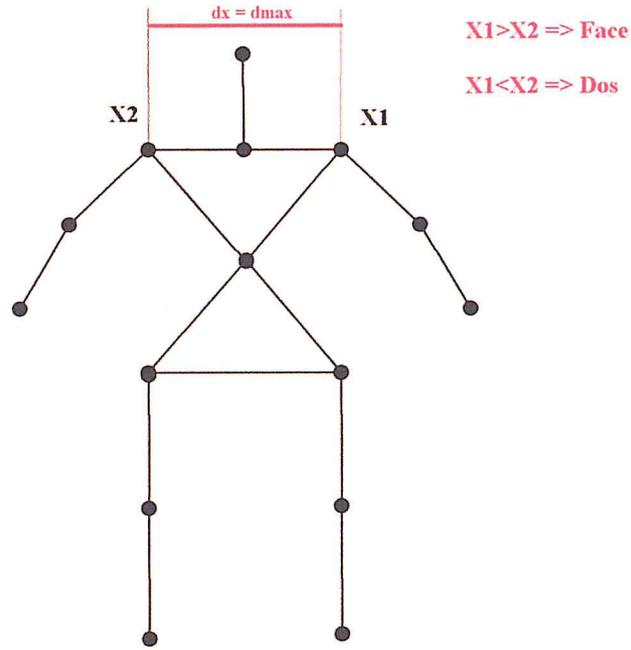


Figure 36: comparaison des coordonnées X

Et pour la différenciation entre l'orientation gauche et droite, les projections des articulations des épaules sur l'axe des Z sont utilisées pour avoir une idée sur l'orientation de la personne voir figure 37.

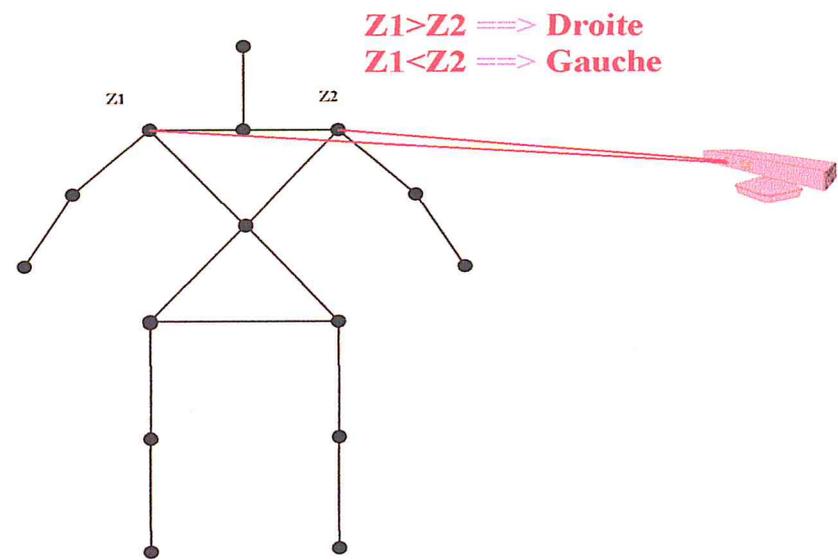
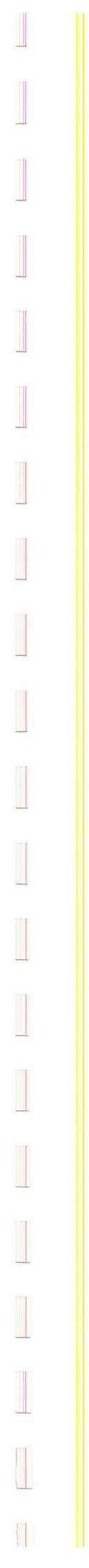


Figure 37: comparaison des coordonnées Z

7. CONCLUSION

Dans ce chapitre, nous avons présenté l'architecture globale du robot B21r ainsi que la solution existante en expliquant les différentes étapes de cette solution et puis on a cité notre contribution dans cette solution où on a présenté les différentes solutions apportées pour enrichir et améliorer cette solution en commençant par la présentation et l'explication de l'algorithme de détection de pose humaine avec les rotations des points, après cela on a présenté la détection des visages en 3D en utilisant la bibliothèque OpenCV.

Dans le chapitre suivant nous allons parler sur les outils utilisés et les différents résultats après avoir implémenté et tester les solutions.



CHAPITRE 4 :
IMPLEMENTATION ET
REALISATION

1. INTRODUCTION

Dans ce chapitre nous allons définir les différents outils utilisés dans notre travail et discuter les résultats obtenus après avoir implémenté les différentes solutions proposées sur le robot B21r.

2. ROS

Comme son nom l'indique, ROS (Robot Operating System) est un système d'exploitation pour robots. De même que les systèmes d'exploitation pour PC, serveurs ou appareils autonomes, ROS est un système d'exploitation complet pour la robotique de service.

ROS est un méta système d'exploitation, quelque chose entre le système d'exploitation et le middleware.[36]

Il fournit des services proches d'un système d'exploitation (abstraction du matériel, gestion de la concurrence, des processus...) mais aussi des fonctionnalités de haut niveau (appels asynchrones, appels synchrones, base de données centralisée de données, système de paramétrage du robot...).

2.1 Langage de programmation ROS

ROS n'est pas dépendant d'un langage. A ce jour, trois bibliothèques principales ont été définies pour ROS qui permettent de programmer respectivement ROS en Python, en Lisp ou en C++. En plus de ces trois bibliothèques, deux bibliothèques expérimentales sont proposées, qui permettent de programmer ROS en Java.

2.2 Les Systèmes de fichier ROS

Les ressources de ROS sont organisées dans une structure hiérarchique sur disque.

Deux concepts importants se détachent : [37]

- **Le package** : C'est l'unité principale d'organisation logicielle de ROS.

Un package est un répertoire qui contient les nœuds (nous verrons ci-dessous ce qu'est un nœud), les bibliothèques externes, des données, des fichiers de configuration et un fichier de configuration xml nommé manifest.xml.

- **La stack** : Une stack est une collection de packages.

Elle propose une agrégation de fonctionnalités telles que la navigation, la localisation...

Une stack est un répertoire qui contient les répertoires des packages ainsi qu'un fichier de configuration nommé stack.xml.

En plus de ces deux notions très importantes, on relève également la notion de distribution.

- **Une Distribution**, comme dans Linux, est un ensemble de stacks versionnées.

Les dernières distributions en date de ROS :

- 22 Juillet 2014 - Indigo Igloo
- 4 Septembre 2013 – Hydro Medusa
- 31 Décembre 2012 – GroovyGalapagos
- v 23 April 2012 – Fuerte
- v 30 Aout 2011 – Electric Emys
- v 2 Mars 2011 – Diamondback
- v 3 Aout 2010 – C Turtle
- v 1 Mars 2010 – Box Turtle
- v 22 Janvier 2010 – ROS 1.0

2.3 Les notions de base de ROS

Le principe de base d'un OS robotique est de faire fonctionner en parallèle un grand nombre d'exécutables qui doivent pouvoir échanger de l'information de manière synchrone ou asynchrone. Par exemple, un OS robotique doit interroger à une fréquence définie les capteurs du robot (capteur de distance à ultrasons ou infrarouge, capteur de pression, capteur de température, gyroscope, accéléromètre, caméras, microphones...), récupérer ces informations, les traiter (faire ce que l'on appelle la fusion de données), les passer à des algorithmes de traitement (traitement de la parole, vision artificielle, localisation et cartographie simultanée,...) et enfin contrôler les moteurs en retour.

Tout ce processus s'effectue en continu et en parallèle. D'autre part, l'OS robotique doit assurer la gestion de la concurrence afin d'assurer l'accès efficace aux ressources du robot.

[38]

Nous décrivons ci-dessous les concepts regroupés dans ROS sous le nom de « ROS Computation Graph » et qui permettent d'atteindre ces objectifs.

Il s'agit des concepts utilisés par le système en cours de fonctionnement tandis que le « ROS File System » décrit dans le paragraphe précédent correspond aux concepts statiques. (Voir figure 38)

- **Les Nœuds** : ROS répond à tout cette problématique grâce à des notions de base simples. La première notion est la notion de nœud. Dans ROS, un nœud est une instance d'un exécutable. Un nœud peut correspondre à un capteur, un moteur, un algorithme de traitement, de surveillance... Chaque nœud qui se lance se déclare au 25Master. On retrouve ici l'architecture microkernel où chaque ressource est un nœud indépendant.

- **Le Master** : Le Master est un service de déclaration et d'enregistrement des nœuds qui permet ainsi à des nœuds de se connaître et d'échanger de l'information.

Le Master comprend une sous-partie très utilisée qui est le Paramètre Server. Celui-ci, également implémenté sous forme de XMLRPC, comme son nom l'indique est une sorte de base de données centralisée dans laquelle les nœuds peuvent stocker de l'information et ainsi partager des paramètres globaux.

- **Les Topics** : L'échange de l'information s'effectue soit de manière asynchrone via un topic ou de manière synchrone via un service.

Un topic est un système de transport de l'information basé sur le système de l'abonnement / publication (subscribe / publish).

Un ou plusieurs nœuds pourront publier de l'information sur un topic et un ou plusieurs nœuds pourront lire l'information sur ce topic.

- **Les Messages** : Un message est une structure de donnée composite. Un message est composé d'une combinaison de types primitifs (chaines de caractères, booléens, entiers, flottants...) et de message (le message est une structure récursive).

Par exemple un nœud représentant un servomoteur du robot, publiera certainement son état sur un topic (selon ce que vous aurez programmé) avec un message contenant par exemple un entier représentant la position du moteur, un flottant représentant sa température, un autre

flottant représentant sa vitesse, La description des messages est stockée dans «Nom_package/msg/monMessageType.msg.» Ce fichier décrit la structure des messages.

▪ **Les Services** : Le topic est un mode de communication asynchrone permettant une communication many-to-many. Le service en revanche répond à une autre nécessité, celle d'une communication synchrone entre deux nœuds. Cette notion se rapproche de la notion d'appel de procédure distante (remote procedure call). La description des services est stockée dans nom_package/srv/monServiceType.srv. Ce fichier décrit les structures de données des requêtes et des réponses.

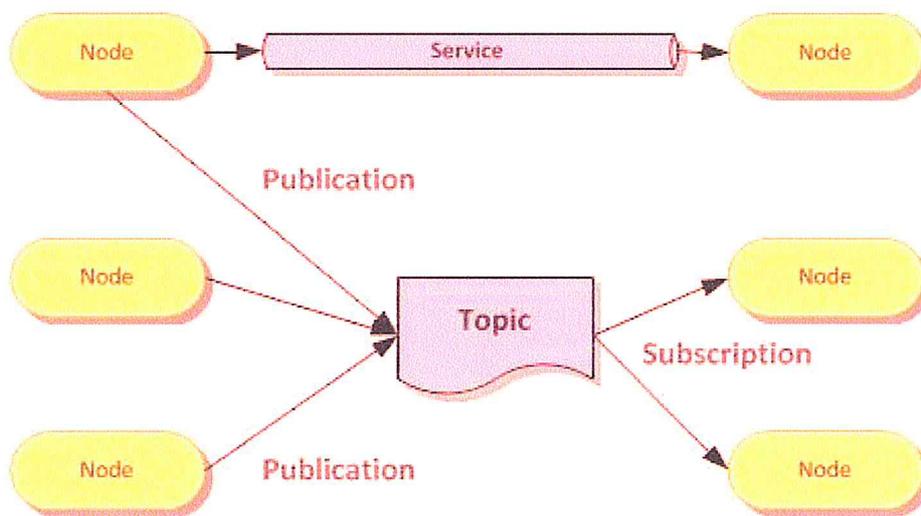


Figure 38: Fonctionnement de ROS [38]

2.4 Outils dans ROS

Comme nous l'avons dit plus haut, ROS est une collection d'outils et d'algorithmes.

Certains sont très utilisés lors de la programmation, de la simulation ou de l'exécution des comportements des robots. Citons quelques outils ou algorithmes que le programmeur de

ROS retrouvera souvent :

- **Stage** : un simulateur 2D
- **Gazebo** : un simulateur 3D
- **Rviz** : un système de visualisation 3D (contrairement à Gazebo, il n'inclut pas de moteur physique)

- **Le package tf** qui permet de manipuler des coordonnées et des transformations (faire de la cinématique inverse, manipuler des matrices).
- **Opencv** : traitement d'images
- **PointCloudLibrary** : reconstruction d'environnement 3D à partir de mesures d'un laser.

2.5 ROS et la Kinect

ROS a lancé un driver qui permet l'accès aux différentes composantes de la Kinect.

Ce driver a été créé par OPENNI, la grand Liberia des capteurs de vision comme les caméras et les capteurs RGB-D.

Pour cela nous avons choisi la Kinect de Microsoft grâce à sa librairie qui est open-source et ce type de capteur RGB-D ne coute rien que les autres type de capteur RGB-D (exemple :la Kinect Asus), aussi elle est facile à manipuler sous ROS et que ce type de capteur est très utilisé dans le domaine de la modélisation 3D ou de la détection et suivi de l'être humain.

3. LES LANGAGES UTILISES

On a utilisé deux langages dans notre travail :

3.1 Le langage C++

C'est un langage dit « de bas niveau », dérivé du C, même que le code C peut le plus souvent être directement utilisé avec un compilateur C++ (quelques précaution doivent toutefois être prises) ; il peut être procédural (un programme sans objets ni classes) ; ou bien orienté-objet (comme Java), avec une syntaxe similaire à ce dernier. [39]

3.2 Le langage PYTHON

Python est un langage de script de haut niveau, structuré et open source.

Il est multi-paradigme et multi-usage.

Développé à l'origine par Guido Van Rossum en 1993, il est, comme la plupart des applications et outils open source, maintenu par une équipe de développeurs un peu partout dans le monde.

Conçu pour être orienté objet, il n'en dispose pas moins d'outils permettant de se livrer à la programmation fonctionnelle ou impérative, c'est d'ailleurs une des raisons qui lui vaut son appellation de « langage agile ».

Parmi les autres raisons, citons la rapidité de développement (qualité propre aux langages interprétés), la grande quantité de modules fournis dans la distribution de base ainsi que le nombre d'interfaces disponibles avec des bibliothèques écrites en C, C++ ou Fortran.

Il est également apprécié pour la clarté de sa syntaxe, ce qui l'oppose au langage Perl. [40]

4. OpenNi (Open Natural Interaction)

La Librairie OPENNI déposé une application sous Linux qui permet de détecter les personnes sur la scène (elle détecte 4 personnes, et permet le suivi de 2 personnes).

Cette application permet de visualiser l'utilisateur comme un squelette, sauf qu'il ne donne pas les configurations des points intéressants comme la tête et les deux mains.

Le système de base et les algorithmes de cette application sont très importants.[41]

Pour notre projet, nous allons améliorer cette application pour l'utiliser dans le cas de notre robot.

5. OpenCV

C'est une librairie open source de traitement et analyse d'images et vidéos avec des interfaces pour les principaux langages de programmation C, C++, Java, C#, Python ...

- Elle est optimisée pour les applications temps réelles.
- Elle fournit une API bas et haut niveau · fournit une interface pour le langage de programmation parallèle IPP.
- Elle est utilisée aussi bien dans les laboratoires de recherche que dans l'industrie.

Fonctions :

- manipulation d'images (chargement, sauvegarde, copie, conversion...) .
- manipulation et acquisition de vidéos.
- manipulations de matrices et algèbre linéaire .
- structure de données utilitaires variées (listes, files, ensembles, graphes...).
- traitement d'images (filtrage, détections de discontinuités, morphologie mathématique...).
- analyse d'images (composantes connexes, ajustement de primitives, transformée de distance...).
- vision (calibration de caméra, stereovision, recherche d'association...).

- reconnaissance de forme (modèles de markov, ACP...).
- interface graphique (affichage d'images, de vidéos, gestion des évènements...).

Modules

- CV - Main OpenCV functions.
- CVAUX - Auxiliary (experimental) OpenCV functions.
- CXCORE - Data structures and linear algebra support.
- HighGUI - GUI functions. [42]

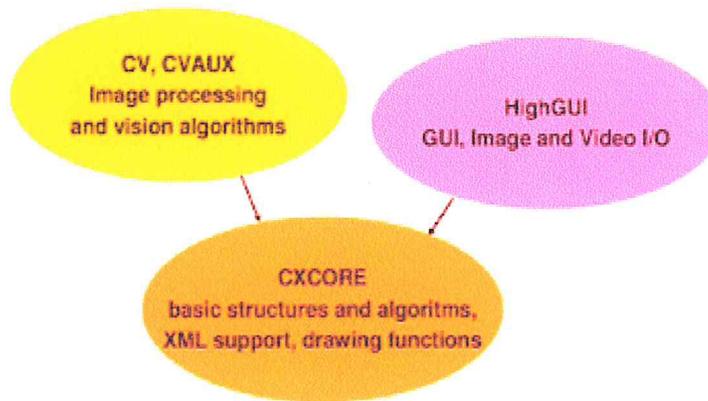


Figure 39: Structure de la librairie OpenCV[41]

6. LE ROBOT B21r

Le robot mobile B21r est une plate forme expérimentale construite par la société i Robot pouvant se déplacer sur un terrain non accidenté ayant comme type de traction, la traction synchrone.(voire figure 40)

Il dispose de quatre roues décentrées orientables tournant selon deux axes : une rotation selon l'axe y pour engendrer la translation, et une rotation selon l'axe verticale au sol pour engendrer une rotation sur lui même.

Ce robot mobile est muni de deux ceintures de capteurs à ultrasons, une ceinture de capteurs infrarouges, un laser, des capteurs tactiles placés le long de ses parois, et une caméra CCD N/B. [16]



Figure 40: Robot B21r

7. SIMULATIONS ET RESULTATS

Dans ce qui suit nous allons exposer les résultats de notre travail de la calibration jusqu'à l'affichage de notre algorithme.

7.1 La calibration

Distance entre la kinect et l'humain (mètre)	Calibration	Temps de réponse (seconde)
0.25m	Echouée	/
0,5m	Echouée	/
0,75m	Réussie	6 s
1m	Réussie	4 s
1,5m	Réussie	1 s
2m	Réussie	3 s
2,5m	Réussie	4,5 s
3m	Réussie	6 s
3,5m	Réussie	7 s
4m	Echouée	/
4,5m	Echouée	/
5m	Echouée	/

Tableau 2: résultats des calibrations de différentes distances

Après faire plusieurs tests de calibration depuis différentes distances (entre 0.25m et 5m) nous remarquons que la calibration se fait entre [0.75m et 3.5m] et en dehors de ce domaine la calibration devient difficile à faire et échouée la plupart du temps.

On remarque aussi que la distance optimale pour faire la calibration est de 1,5m et ça prend en moyenne une seconde.

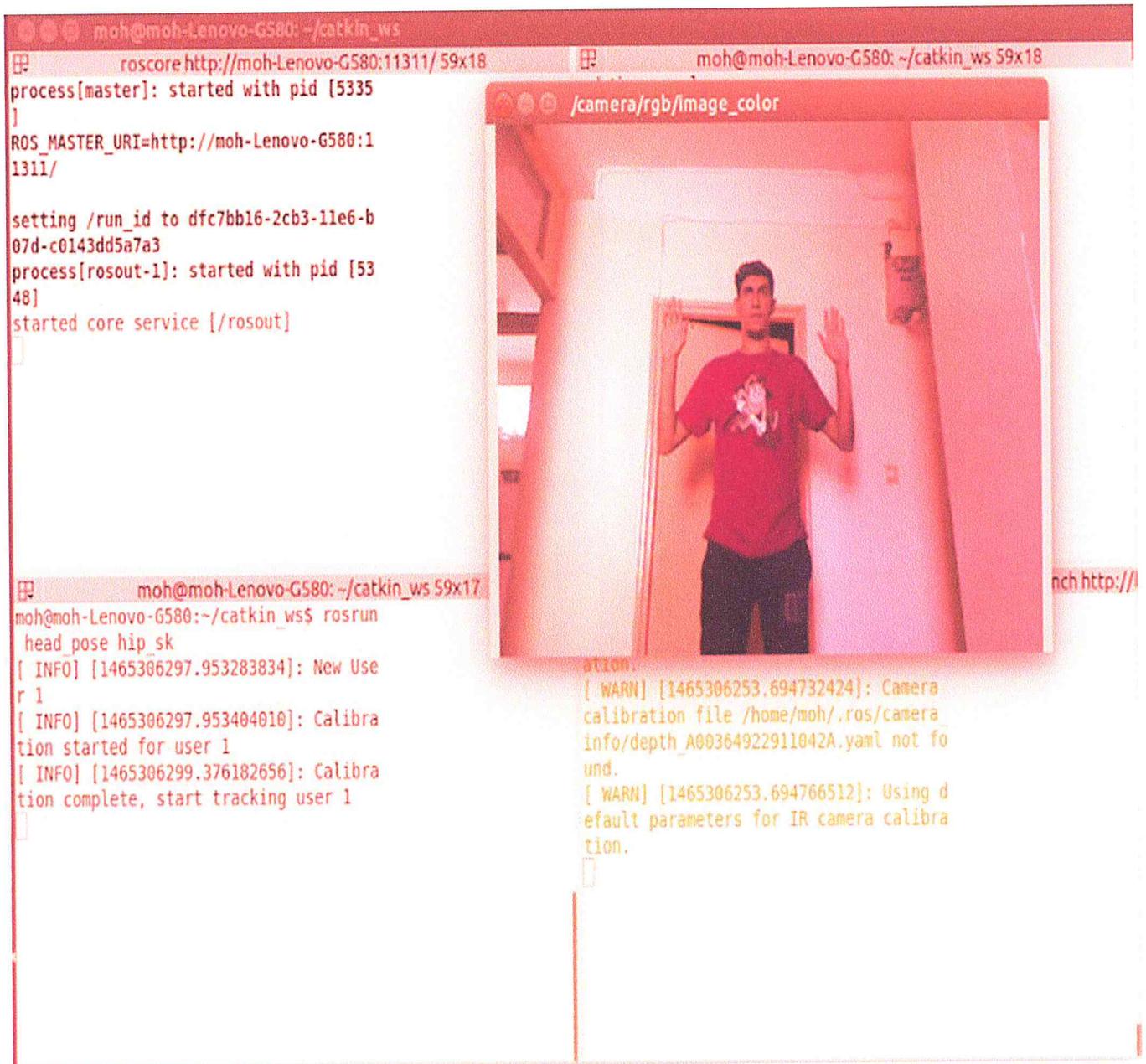


Figure 41: Calibration

7.2 La détection de la pose

Dans les figures suivantes nous allons montrer les résultats obtenus lors de l'exécution des nœuds de notre solution

7.2.1 La détection de la pose Face

Dans la figure 42 nous montrons les résultats de la pose calculés par notre solution quand l'être humain est de face.

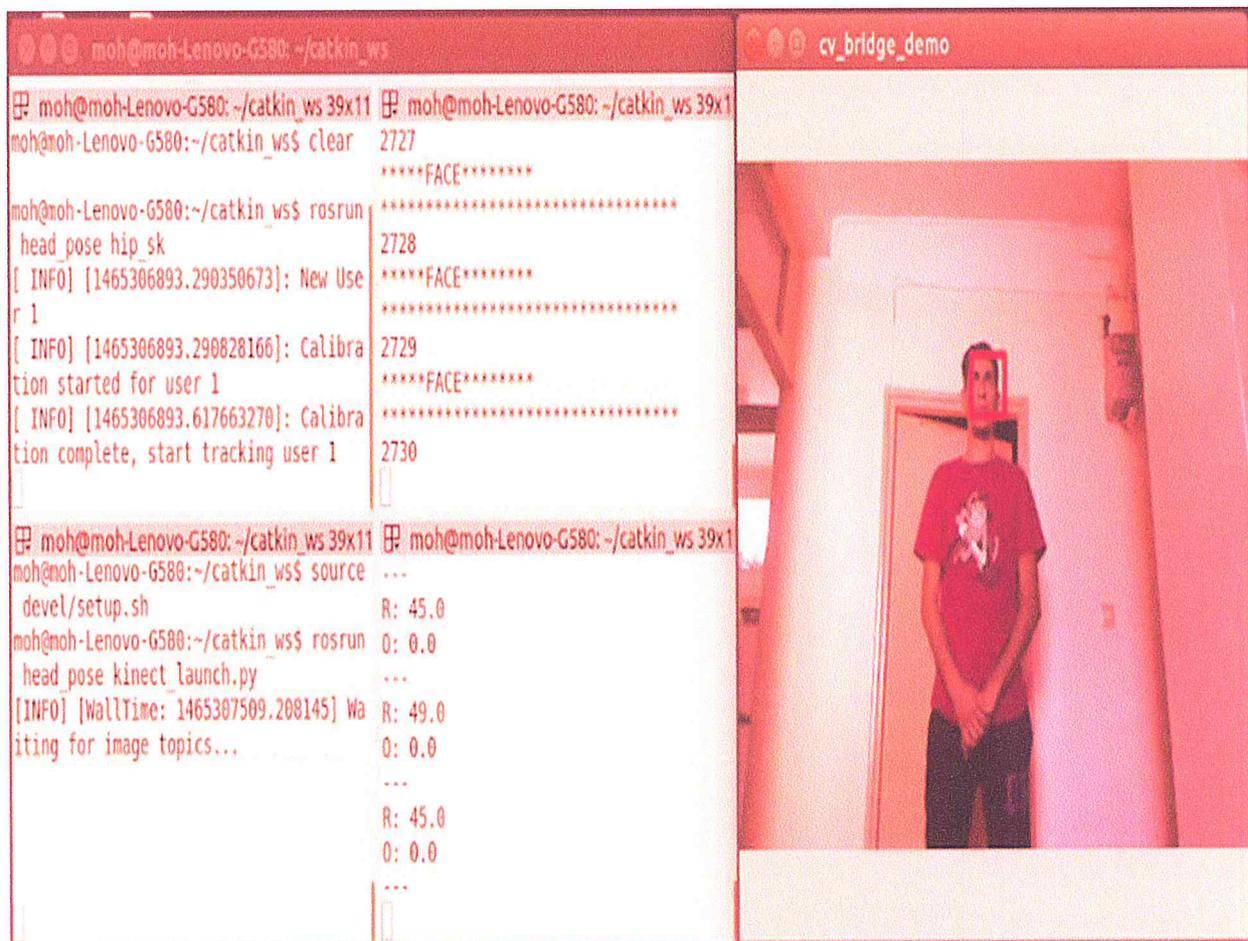


Figure 42: Détection de la pose Face

7.2.2 La détection de la pose Droite

La figure 43 nous montre les résultats affichés lors l'être humain est de droite.

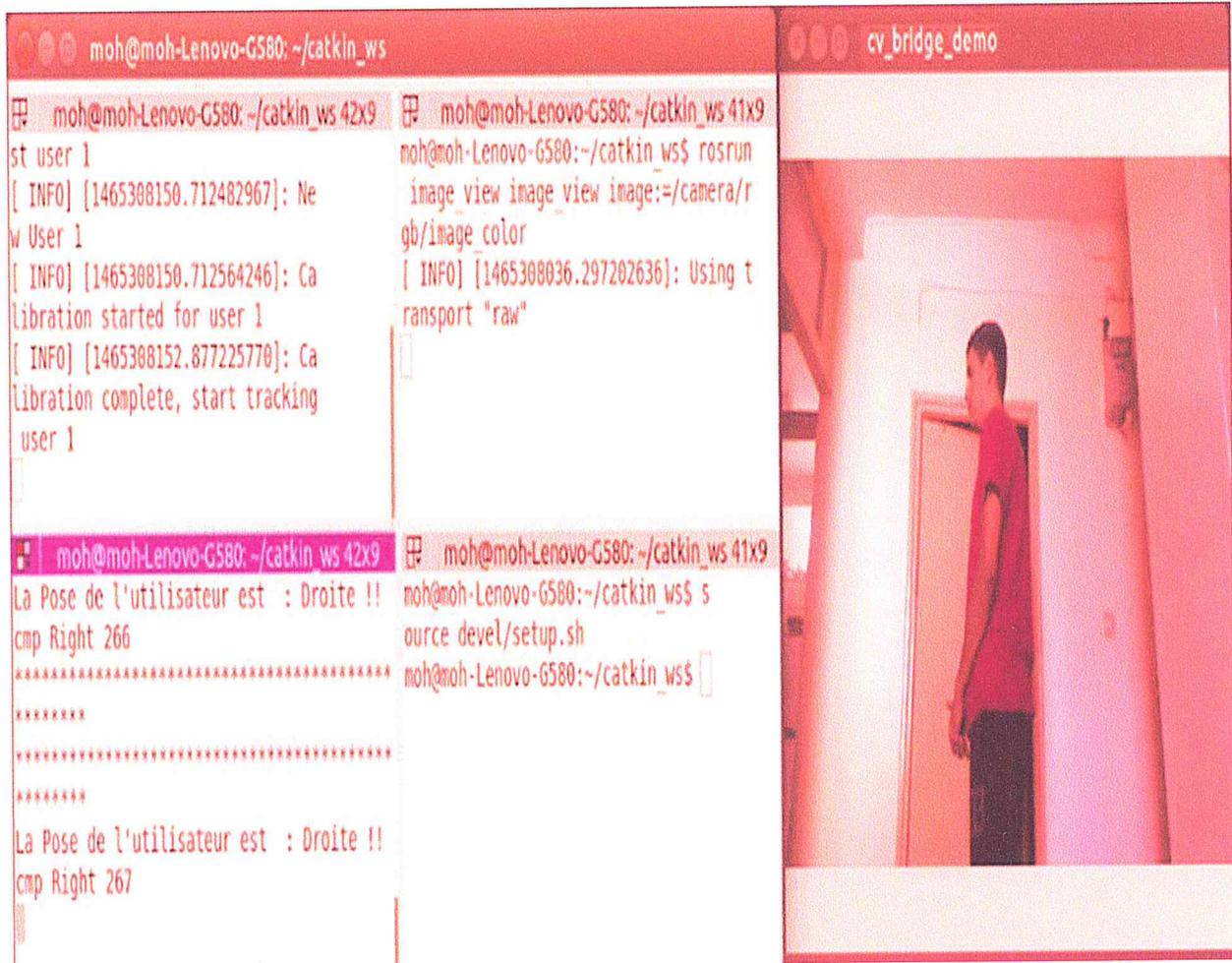


Figure 43: Détection de la pose Droite

7.2.3 La détection de la pose Dos

La figure 44 nous montre les résultats affichés lors l'être humain est de dos.

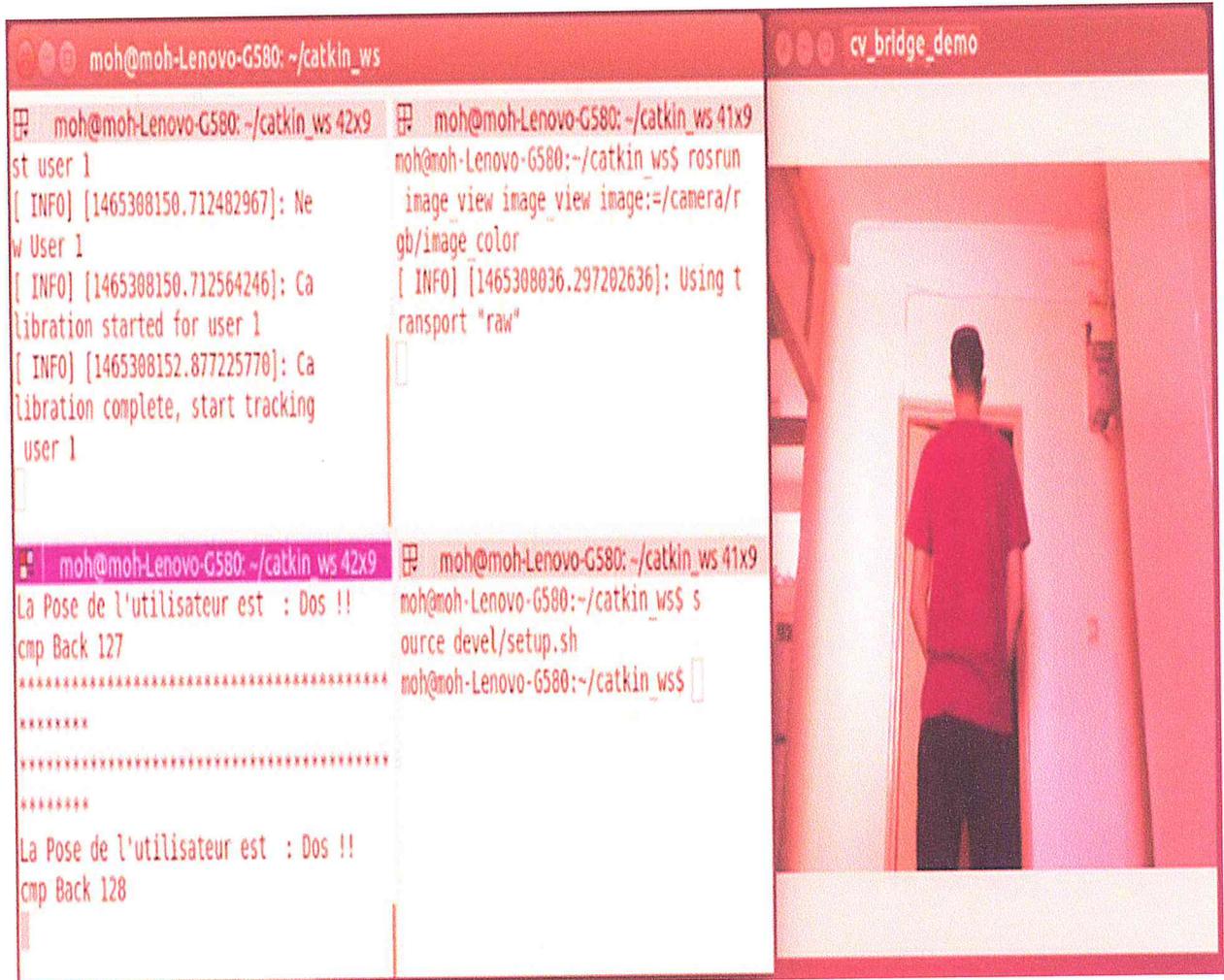


Figure 44: Détection de la pose Dos

7.2.4 La détection de la pose Gauche

La figure 45 nous montre les résultats affichés lors l'être humain est de Gauche.

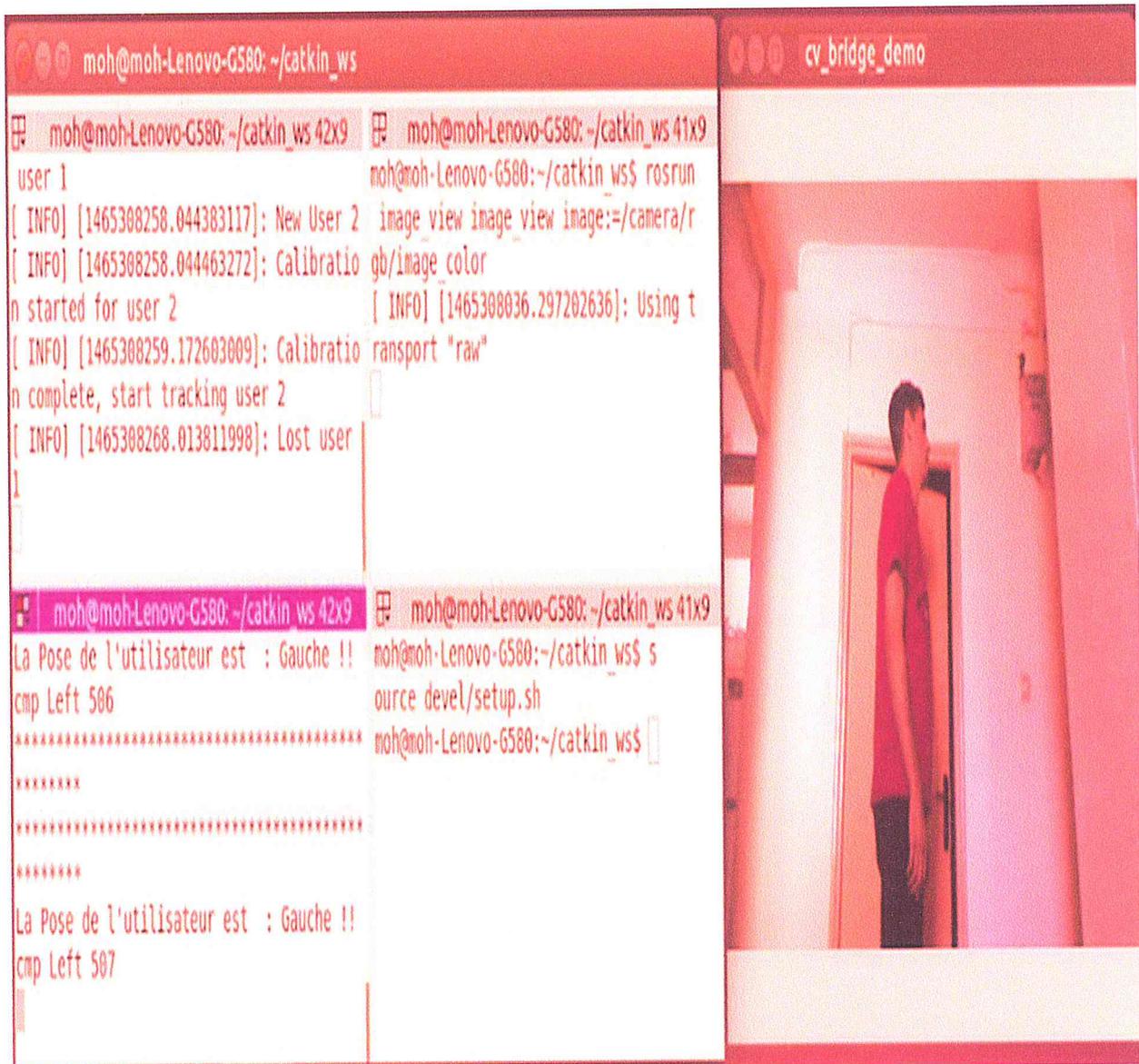


Figure 45: Détection de la pose Droite

Temps d'exécution :

Le temps d'exécution est calculé à l'aide de la fonction "ros::Time",

Cette fonction nous donne le temps réel durant l'exécution de programme, donc il suffit d'appeler la fonction au début et à la fin du programme pour calculer le temps d'exécution voire le code suivant :

```

ros::Timebegin = ros::Time::now(); // au début du traitement.
ros::Time end= ros::Time::now(); // à la fin du traitement.
float temps = end - begin;
cout<<"Le temps d'exécution :"<<temps<<endl;//afficher le temps Remarque : Le temps
d'exécution est calculé en second.
temp_d'exécution = 0.0005s.

```

7.3 Les Résultats

Les tableaux ci-dessous représentent les résultats obtenus :

a. Résultats obtenus avant l'ajout de la détection faciale (tableau 3)

Pose	Face	Dos	Droite	Gauche
Test_Face	84%	8%	5%	3%
Test_Dos	4%	90%	2%	4%
Test_Droite	2%	3%	91%	4%
Test_Gauche	4%	2%	5%	89%

Tableau 3: résultats primaires

D'après le tableau précédent on remarque que la fiabilité s'est améliorée mais elle reste toujours insuffisante par conséquent on a décidé d'ajouter la détection faciale pour améliorer le test face et le test dos et on a ajouté aussi les angles d'Euler pour améliorer le test gauche et le test droit. L'algorithme est résumé dans la figure 46

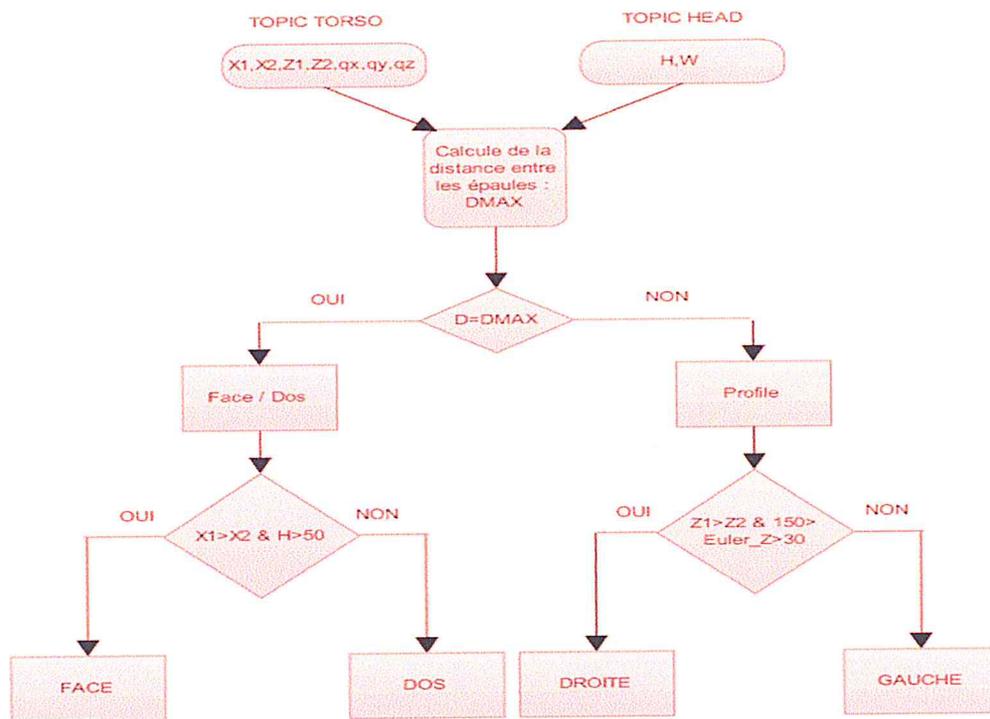


Figure 46 : Organigramme pour la détection de la Pose

b. Résultats obtenus après l'ajout de la détection faciale (tableau 3)

Pose	Face	Dos	Droite	Gauche
Test_Face	96%	3%	1%	0%
Test_Dos	3%	95%	0%	2%
Test_Droite	2%	1%	92%	5%
Test_Gauche	2%	2%	3%	93%

Tableau 4: résultats finaux

Après l'exécution de notre solution on remarque que les résultats se sont améliorés de manière significative grâce à l'utilisation de la combinaison entre la détection faciale et les coordonnées X des épaules (les positions Face et Dos) ainsi que la combinaison entre les coordonnées Z des épaules et les angles d'Euler (les positions Gauche et Droite).

8. CONCLUSION

Dans ce chapitre nous avons présenté les différents outils utilisés tel que le système d'exploitation ROS les langages de programmation C++ et python ainsi que les bibliothèques utilisées tel que OpenCV et OpenNI

Et enfin nous avons discuté les résultats obtenus après l'implémentation de différents algorithmes.

Conclusion Générale :

Le travail présenté dans ce mémoire traite le domaine de la détection de l'être humain par un robot de service muni d'une caméra RGB-D de type Kinect.

Nous avons à cet effet réalisé une solution qui améliore les résultats obtenus par l'exécution de l'algorithme « un point d'articulation ».

Notre solution permet la détection de visage en temps réel en lui rajoutant également l'aspect suivi (tracking).

Cette contribution qui nous a permis d'améliorer les résultats d'autres solutions est basée sur le calcul des coordonnées des articulations ainsi que les rotations lors de la détection de la pose.

Ce projet nous a été très bénéfique, il nous a permis d'une part de découvrir le domaine de la robotique et de l'environnement de développement des projets liées à ce domaine, et d'autre part, de mettre en pratique les connaissances acquises tout le long de notre cursus académique.

Bibliographie

- [1]: World robotics, par l'International Federation of Robotics, éditions 2006 à 2011
- [2]: International Assessment of Research and development in Robotics, rapport pour la National Science Foundation (NSF) USA, par George Bekey et al. 2006.
- [3]: Le développement industriel futur de la robotique personnelle et de service en France,
Date de parution : 12 avril 2012 Couverture : Nathalie Palous, Brigitte Baroin Édition :
Nicole Merle-Lamoot, Gilles Pannetier
- [4] : Introduction à la robotique par Laetitia Matignon Université de Caen, France
- [5] : ROBOTIQUE, ISTIA, Université Angers Jean-Louis Boimond
- [6] : Bases de la Robotique, Dr M. Bouri, Septembre 2015
- [7] : Définition, historique, domaines, poids économique, caractéristiques R.Clavel, 2009
- [8] : Le développement industriel futur de la robotique personnelle et de service en France,
Nicole Merle-Lamoot, Gilles Pannetier, 12 avril 2012
- [10]: Journal of Robotics and Autonomous Systems, n° 36, pp. 103-119, Elsevier Science Publishers, 2001
- [12]: Les Sous-marins Autonomes, Pierre Van Rie, Février 2010
- [13] : A.Chohra, « Planification et control de la navigation des Véhicules Autonomes Intelligents (VIA) en environnements dynamiques », Thèse de doctorat d'état en électronique 1998.
- [17] : Philippe Noriega. Modèle du corps humain pour le suivi de gestes en monoculaire. Modélisation et simulation. Université Pierre et Marie Curie - Paris VI, 2007.

- [18]: Thanarat Horprasert, David Harwood et Larry S Davis.
A statistical approach for real-time robust background subtraction and shadow detection.
In IEEE ICCV, volume 99, pages 256_261, 1999.
- [19]: Nizar Fakhfakh. Détection et localisation tridimensionnelle par stéréovision d'objets en mouvement dans des environnements complexes : application aux passages à niveau. Interface homme-machine [cs.HC]. Ecole Centrale de Lille, 2011.
- [20]: L.Hammoudi, Application de techniques d'apprentissage pour la détection et la reconnaissance d'individus », Thèse de doctorat d'état en informatique 2011.
- [21]: Antoine Belconde. Modélisation de la détection de présence humaine. Sciences de l'ingénieur [physics]. Université d'Orléans, 2010
- [22]: Yannick Benezeth, Bruno Emile, Hélène Laurent, Christophe Rosenberger. Détection de la présence humaine et caractérisation de l'activité. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2009, Dijon, France. 2009
- [23]: G. Dumont. Animation de scènes tridimensionnelles : la mécanique des solides comme modèle de synthèse du mouvement. PhD thesis, Université de Rennes 1, 1990.
- [24]: Maeva Le Goic, Etude du contrôle postural chez l'homme, Université René Descartes - Paris V, 2013.
- [25]: B. Hasler, N.and Rosenhahn, T. Thormahlen, M. Wand, J. Gall, et H.-P. Seidel,
“ Marker less Motion Capture with unsynchronized moving cameras ”, dans Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on 20-25 June 2009, June 2009, pp. 224– 231
- [26]: Maxime DEVANNE. Modélisation 3D d'un corps humain à partir de caméras 3D Kinect. Laboratoire d'Informatique Fondamental de Lille.2011

- [27] : Boudjellal. S. Détection et identification de personne par méthode biométrique. Mémoire de Magister : Electronique-Téledétection. Tizi-ouzou : Université Mouloud Mammeri, 95p.
- [28]: Ming-Hsuan Yang, David J. Kriegman et Narendra Ahuja. Detecting faces image: A survey. Dans IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002, vol.24 (1), 34-58.
- [30]: Cheng-Chin, Wen-Kai Tai, Mau-Tsuen Yang, Yi-Timg Huang, and chi-Janng Huang. A novel method for detecting lips, eyes and faces in real time. Real-Time Imaging, 9(4) : 277-287, 2003.
- [31] : M.Van Wambeke, Reconnaissance et suivi de visages et implémentation en robotique temps-réel. Master Ingénieur Civil en Génie Biomédical. Belgique. Université catholique de Lovain. 2010, 93p.
- [32]: Wenlong Zheng and Suchendra M. Bhandarkar. Face detection and haking using a boosted adaptative particle filter. Journal of visual communication and Imog Representation, 20(1) : 9-27, 2009.
- [33]: H. A. Rowley, S. Baluja, and T. Kanade. Neural Network-based face detection. IEEE Transachen on Pattern Analysis and Machine Intelligence, 20(1): 23-38, 1998.
- [34]: H. Schneidermand and T. Kanade. Probabilistic modeling of local appearance and spacial relationships for object recognition. Computer Vision and Pattern Recognition, IEEE Computer Society Conférence on 0:45, 1998.
- [35]: Paul Viola and Michael Jones. Robust real-time object detection. In Second international work shop on statistical and computation atheories of vision, Vancouver, Canada, July 13 2001.
- [39] : Emmanuel Courcelle, Introduction au C++et à la programmation objet, Formation permanente du C.N.R.S. - Janvier 2013

[40] : S. Vialette IGM-LabInfo, introduction a Python, Université Paris-Est, 21 janvier 2008

[42] : Walid HIZEM ; Capteur Intelligent pour la Reconnaissance du visage ; UNIVERSITE PIERRE ET MARIE CURIE - PARIS 6 ; Thèse ; Paris ; 2009.

[43] : Agnès Guillot, La robotique de A à Z, 20/01/2003

[44] : Jean-Arcady Meyer • Agnès Guillot, Vers une robotique animale, Octobre 2002.

[45]: Julian Warner, From Writing To Computers, 1994

[46]: Jean Piaget, Cheminements dans l'œuvre scientifique, 1999

[48]: Robot Grippers Par Gareth J. Monkman, Stefan Hesse, Ralf Steinmann, Henrik Schunk
page 446

[62]: Wang, L., Hu, W. et Tan, T. (2003). Recent developments in human motion analysis. Pattern Recognition, 36:585-601.

[63]: Hu, W., Tan, T., Wang, L. et Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. IEEE Trans. on Systems, Man, and Cybernetics, 34(3):334-352

[64]: Piccardi, M. (2004). Background subtraction techniques: A review. In Proc. IEEE Conf. on Systems, Man, and Cybernetics, volume IV, pages 3099-3104, La Hague, Pays-Bas.

Webographie

[9] : http://www.futura-sciences.com/fr/doc/t/technologie/d/hrp-2-le-robot-humanoide-franco-japonais_688/c3/221/p2/

[11] : <http://polyaeronefsautonomespilotes.jimdo.com/>

[14] : <http://www.irobot.com/>

[15] : <http://www.rovio.com/>

[16] : <http://www.cdta.dz/>

[29] : Commercial Security Devices[en ligne] (page consultée le 01/10/2013).

<<http://www.commercialSecuritydevises.com/fr/détection-de-visage.html>>

[36] : www.ros.org

[37] : <http://www.willowgarage.com/>

[38] : <https://code.ros.org/gf/>

[41] : <http://openni.org/>

[47] : <http://www.humanoid.waseda.ac.jp/>

[49] : <https://robotics.kawasaki.com>

[50] : <https://www.willowgarage.com/>

[51] : <http://www.intuitivesurgical.com/>

[52] : <http://asimo.honda.com/>

[53] : <http://www.ald.softbankrobotics.com/>

[54] : <https://www.sri.com/work/timeline-innovation/timeline.php?timeline=computing-digital#!&innovation=shakey-the-robot>

[55] : <http://www.lunaticadventures.com/>

[56] : <http://mars.nasa.gov/>

[57] : http://www.nasa.gov/mission_pages/mars-pathfinder/

[58] : <http://www.openrov.com/>

[59] : <http://www.ecagroup.com/en/solutions/alistar-3000-auv-autonomous-underwater-vehicle>

[60] : <http://www.navalunderseamuseum.org/curv-iii/>

[61] : <http://www.ambrogiorobot.com/>