

MEMOIRE D'HABILITATION A DIRIGER DES RECHERCHES EN SCIENCES

Présenté à l'Université
d'Evry-Val d'Essonne

Sur la Réalité Augmentée pour la Téléopération

Par

Malik MALLEM
Maître de conférences

Soutenu publiquement le 19 Janvier 1999, devant le jury composé de

Messieurs :

Philippe COIFFET	Directeur de Recherches au CNRS	Rapporteur
Etienne COLLE	Professeur à l'Université d'Evry-Val d'Essonne	Rapporteur
Michel DHÔME	Directeur de Recherches au CNRS	Rapporteur
Francis ARTIGUE	Professeur à l'Université d'Evry-Val d'Essonne	Examineur
Florent CHAVAND	Professeur à l'Université d'Evry-Val d'Essonne	Examineur
Jean-Marie DETRICHE	Directeur Adjoint de CEREM/DPSA au CEA	Examineur
Claude LAURGEAU	Professeur à l'École Nationale Supérieure des Mines de Paris	Examineur

AVANT-PROPOS

Ce mémoire constitue une synthèse de nos travaux de recherche et développement, durant la dernière décennie, au Laboratoire de Robotique d'Evry puis au Centre d'Etudes Mécaniques d'Ile de France – groupe systèmes complexes dirigés par le Professeur Florent Chavand. Qu'il soit assuré de ma vive reconnaissance pour avoir dirigé et orienté nos travaux de recherche.

Je remercie vivement Messieurs : Philippe Coiffet Directeur de Recherches au CNRS, Etienne COLLE Professeur à l'Université d'Evry Val d'Essonne et Michel Dhôme Directeur de Recherches au CNRS, pour avoir accepté d'être rapporteurs de ce mémoire.

Je remercie également Messieurs : Francis Artigue Professeur à l'Université d'Evry Val d'Essonne, Jean-Marie Détriché Directeur Adjoint du département des procédés et systèmes avancés du Centre d'Etudes Nucléaires et Claude Lurgeau Professeur à l'Ecole Nationale Supérieure des Mines de Paris, pour avoir accepté d'être examinateurs de ce mémoire.

Mes remerciements et l'expression de ma sympathie à tous les chercheurs du laboratoire et notamment les docteurs et doctorants qui ont contribué de près à l'accomplissement des travaux présentés dans ce mémoire.

SOMMAIRE

0. DEFINITION DU SYSTEME MCIT	6
0.0Bref état de l'art.....	6
0.1 Le système MCIT.....	7
0.2 Architecture générale du système MCIT.....	7
0.3. Description fonctionnelle du système MCIT	9
0.4. Architecture informatique du système MCIT	10
0.5 Bilan.....	11
0.6 Bibliographie du chapitre 0.....	12
1. MODELISATION ET CALIBRATION DE CAPTEURS	13
1.0Bref état de l'art.....	13
1.1 La caméra.....	14
1.2.Dispositif de relevé 3D (DR3D).....	29
1.3 Choix du capteur de relevé 3D.....	37
1.4. Capteur mixte : Caméra + DR3D.....	40
1.5 Bilan.....	40
1.6 Annexe1	42
1.7 Bibliographie du chapitre 1.....	44
2. MODELISATION 3D	45
2.0 Bref état de l'art.....	45
2.1. Modélisation de prisme symétrique	45
2.2. Modélisation de cylindre.....	48
2.3. Modélisation d'objets à symétrie de révolution.....	54
2.4 Représentation informatique des données.....	55
2.5 Bilan.....	58
2.6 Annexe 2	60
2.7 Bibliographie du chapitre 2.....	61
3. RECONSTRUCTION 3D.....	62
3.0 Bref état de l'art.....	62
3.1. Reconstruction 3D basée sur les plans d'interprétation	63
3.2. Reconstruction polyédrique basée sur les arêtes	65
3.3 Reconstruction polyédrique fine	69
3.4. Reconstruction cylindrique	72
3.5 Expérimentation et résultats.....	74
3.6 Reconstruction basée sur les sommets	74
3.7 Bilan.....	76
3.8 Bibliographie du chapitre 3.....	77
4.APPARIEMENT D'UNE IMAGE 2D AVEC SON MODELE 3D.....	78
4.0 Bref état de l'art.....	78
4.1 Appariement 2D/3D polyédrique basée sur liaisons	79
4.2 Appariement 2D/3D polyédrique	79
4.3 Bilan.....	90
4.4 Bibliographie du chapitre 4.....	91

5. APPLICATIONS	92
5.0 Introduction.....	92
5.1. Animation de la représentation graphique de bras manipulateur superposée à l'image caméra	92
5.2. Planification de trajectoires en 2 1/2 D de bras manipulateur	93
5.3. Programmation CAO de tâches robotiques	98
5.4 Contrôle-commande de robot via internet.....	101
6. CONCLUSIONS ET PERSPECTIVES	103

INTRODUCTION

La Téléopération concerne en général la robotique en milieux difficiles d'accès à l'homme (nucléaire, spatial, sous-marin, ...). La complexité des tâches nécessite la présence d'un opérateur humain dans la boucle de perception-décision-commande. Celui-ci doit gérer un grand nombre d'informations: une coopération entre l'opérateur et le système est donc nécessaire. La vision de la scène, représentée par le robot et son environnement, est assurée par caméra(s). L'image vidéo peut être dégradée pour des raisons diverses (turbidité de l'eau, éblouissement de la caméra, présence d'obstacles, dégagement de fumée ou poussières, faible débit du canal de transmission). Il peut y avoir aussi d'importants retards de transmission d'informations.

En conséquence, l'opérateur a besoin d'assistance pour la perception de la scène distante et la télécommande de robot.

Afin de procurer à l'opérateur l'assistance nécessaire, notre orientation de recherche s'est portée sur la conception et la réalisation d'un système de Réalité Augmentée (RA), qui nous semble le plus approprié pour aider l'opérateur dans l'accomplissement de sa tâche.

En effet, un système de RA permet à l'opérateur de voir sur écran vidéo, en permanence, la scène réelle enrichie d'informations pertinentes pour la réalisation de la tâche. Ces informations peuvent être des indicateurs sur le déroulement de la tâche ou une représentation graphique de la scène superposée à l'image réelle pour la compléter voire la suppléer dans le cas où celle-ci est dégradée.

Les travaux que nous avons menés ont débuté par l'étude des différents retours d'informations, leur mode de présentation à l'opérateur, et leur impact sur celui-ci [MALLEM 1990].

Les contraintes liées à la RA, qui sont d'ordre temporelle et de précision nous ont amenés à entreprendre des recherches dans plusieurs thèmes.

En effet, ceux-ci concernent la modélisation et calibration de capteurs, la modélisation 3D, la reconstruction 3D, la reconnaissance d'objets polyédriques. Nous avons traité ces points dans l'ordre cité. Le respect de cet ordre est important car tout problème constitue un prérequis pour le ou les suivants. De la précision de la calibration des capteurs dépend celle de la modélisation 3D et de la reconstruction 3D et de ce fait celle de la reconnaissance d'objets.

La modélisation et calibration de capteurs consistent à calculer les transformations mathématiques entre leurs repères et celui de l'environnement afin de permettre, d'une part, la superposition d'images et, d'autre part, le relevé 3D.

La modélisation et reconstruction 3D consistent à construire et à mettre à jour la base de données 3D (BD3D) de l'environnement indispensable pour l'approche de reconnaissance d'objet que nous utilisons.

La reconnaissance d'objets consiste à partir des modèles 3D et du modèle 2D, résultat du traitement de l'image caméra, d'effectuer la mise en correspondance de ces deux modèles afin de reconnaître les objets modélisés et vus par la caméra.

Dans la suite de ce document, nous présentons, brièvement, pour chacun de ces problèmes, les méthodes que nous avons développées pour le résoudre et quelques résultats. Avant de présenter les différents chapitres traitant ces problèmes, nous décrivons brièvement le système de RA développé dans le laboratoire.

0. DEFINITION DU SYSTEME MCIT

Le système Multimédia de Contrôle et d'Interface en Téléopération (MCIT) est développé dans notre laboratoire afin d'apporter des assistances à la commande et au contrôle en téléopération ou télérobotique [SHERIDAN 1989]¹. Ce système est basé sur l'utilisation de capteurs et de base de données 3D. Avant de le décrire, nous présentons quelques systèmes comparables.

0.0Bref état de l'art

Le système de ETL ("ElectroTechnical Laboratory au Japon")

Ce système permet la supervision en vision synthétique pour les opérations d'assemblage et de désassemblage de pièces mécaniques [HASEGAWA 1991].

Ce système permet également de modéliser et de gérer une BD3D. La mise à jour est réalisée grâce aux informations fournies par une paire de caméras vidéo et un pointeur laser. La position 3D de l'objet à manipuler est obtenue par triangulation; l'image graphique obtenue est aussi superposée à l'image vidéo. Le système comprend :

- ❶ un module de modélisation interactive par la coopération caméra pointeur laser qui génère sur une base de données objets connus.
- ❷ un module de programmation, qui comprend le système de visualisation pour l'aide à la perception où le modèle synthétique est superposé à l'image caméra(monde réel).
- ❸ un module d'exécution qui englobe le manipulateur, le contrôleur de tâches ainsi que les transformateurs de coordonnées.

Ce système manipule des objets dont la géométrie est parfaitement connue dans de bonnes conditions de vision.

Le système du JPL ("Jet Propulsion Laboratory" aux USA)

C'est un système développé pour la téléopération spatiale. L'équipe du JPL, [BEJCZY 1990],[KIM 1993], propose une technique simple pour la restitution de l'information 3D par superposition d'aides synthétiques sur une image vidéo. L'objectif de ces chercheurs est de donner à l'opérateur une estimation de la distance et de l'orientation de la pince du robot par rapport à la cible. Pour suivre l'évolution de la tâche, une image synthétique est superposée à l'image du bras du robot. Une visualisation des forces et couples s'exerçant sur le bras est également présentée à l'opérateur.

Ce système permet de pallier les retards de transmission d'information en assurant un retour visuel prédictif sur l'évolution de la tâche.

Le système PYRAMIDE du CEN(Centre d'Etudes Nucléaires en France)

PYRAMIDE est un outil de modélisation graphique, développé au CEN, [BONNEAU 1993], pour l'exécution de tâches en téléopération lors d'une intervention sur accident en milieu hostile. C'est un logiciel qui permet de construire, mettre à jour et gérer une base de données géométriques d'un environnement structuré et de produire des modèles synthétiques.

La structure de PYRAMIDE comprend quatre modules principaux.

- ❶ le gestionnaire de la base de données, qui regroupe toutes les opérations de création, modification, sélection de primitives, destruction, ...

La modélisation est assurée par l'opérateur à l'aide d'une souris et sur une image caméra. C'est une modélisation interactive où l'opérateur manipule des entités connues selon trois dimensions (trois translations, trois rotations). Le système dispose d'une modélisation fine par stéréométrie également interactive sur deux images caméras.

- ❷ L'interface homme-machine, dont le rôle est de présenter les informations à l'opérateur et de gérer les outils de contrôle.

- ❸ le module de raisonnement géométrique, qui gère les traitements effectués sur les entités géométriques (triangulation, calibration, etc.).

- ❹ le module d'application, qui permet d'insérer PYRAMIDE dans un ensemble complet de téléopération (transmission de consignes, prise en compte d'informations embarquées, etc.).

Ce système permet une modélisation 3D en ligne de l'environnement et une superposition de l'image de synthèse sur une image vidéo statique.

¹ Les références bibliographiques du chapitre se trouvent dans le dernier paragraphe de celui-ci.

Une bibliographie plus complète se trouve dans les publications [COM 1994c]² et [COM 1995].

0.1 Le système MCIT

MCIT est basé sur l'approche de la réalité augmentée au même titre que les systèmes décrits précédemment. Ce système a pour but d'apporter à l'opérateur humain, en situation de téléopération ou de supervision de tâche, une assistance à la perception en vision indirecte sur écran vidéo de la scène et une assistance à la commande en désignant un objet directement sur l'écran vidéo. L'objet -pré modélisé - désigné est alors reconnu automatiquement par MCIT.

Ces assistances s'appuient sur deux architectures spécifiques, l'une matérielle et l'autre logicielle. L'architecture matérielle géographiquement répartie est constituée d'un ensemble de capteurs (télé-mètre et caméra), et de trois calculateurs en réseau. Cette architecture permet l'acquisition du monde réel par le biais de capteurs et la superposition à celui-ci du monde virtuel. La partie logicielle permet de créer, de mettre à jour ce monde virtuel et de le mettre en correspondance avec le monde réel.

Ce système offre à l'opérateur humain une interface graphique conviviale et lui permet d'effectuer une téléprogrammation de niveau tâche. Le monde réel sert d'interface entre l'opérateur humain et le monde virtuel représenté par une BD 3D. La désignation d'un objet, à manipuler par le robot, sur écran vidéo revient alors à le désigner dans le monde virtuel. L'exécution de la tâche peut être simulée en utilisant le monde virtuel avant d'être appliquée au monde réel. Après l'exécution de la tâche, le monde virtuel est superposé au monde réel. C'est l'approche de la réalité augmentée désigné ainsi car le monde réel est l'interface entre l'opérateur et le monde virtuel.

L'approche de la réalité virtuelle met en œuvre, généralement, un graphique d'un haut niveau de réalisme représentant le monde virtuel. L'opérateur humain effectue une simulation de la tâche dans le monde virtuel, avec immersion dans celui-ci[BURDEA 1993] . A l'issue de la simulation un fichier de consignes destiné au robot réel, est créé. Les modifications produites par l'exécution de celui-ci sont ensuite reportées dans le monde virtuel.

L'approche de la réalité augmentée contrairement à celle de la réalité virtuelle ne nécessite pas un graphique réaliste, une représentation fil de fer suffit. Car, les deux mondes étant superposés, il faut masquer le moins possible l'image vidéo. Par ailleurs, cette représentation fil de fer permet à l'opérateur de percevoir l'image vidéo même quand celle-ci est partiellement dégradée et donc de continuer la tâche. De plus, cette approche permet d'associer un objet réel à sa BD 3D.

Les architectures générale, fonctionnelle et matérielle du système MCIT sont présentées ci-après.

0.2 Architecture générale du système MCIT

La figure 0.1 représente l'architecture générale du système MCIT. Dans la station de contrôle commande, l'opérateur humain utilise un seul écran vidéo . Sur celui-ci est superposée une représentation graphique de la scène sur son image caméra. A partir de cet écran, l'opérateur humain peut percevoir la scène distante, et agir à distance sur celle-ci. En effet, il commande les capteurs pour réaliser le relevé 3D - nécessaire à la création de la base de données 3D- ou la reconstruction 3D, et le robot pour l'exécution de tâche.

Le site distant est constitué de capteurs extéroceptifs, dispositif de relevé 3D (DR3D) et caméra, et proprioceptifs qui renseignent en ligne sur la configuration du robot. Cette architecture permet l'acquisition du monde réel par le biais de capteurs et la superposition à celui-ci du monde virtuel. La partie logicielle permet de créer, de mettre à jour ce monde virtuel et de le mettre en correspondance avec le monde réel.

² Nos publications sont listées dans le chapitre 7 et les textes des plus significatives se trouvent dans le chapitre 8.

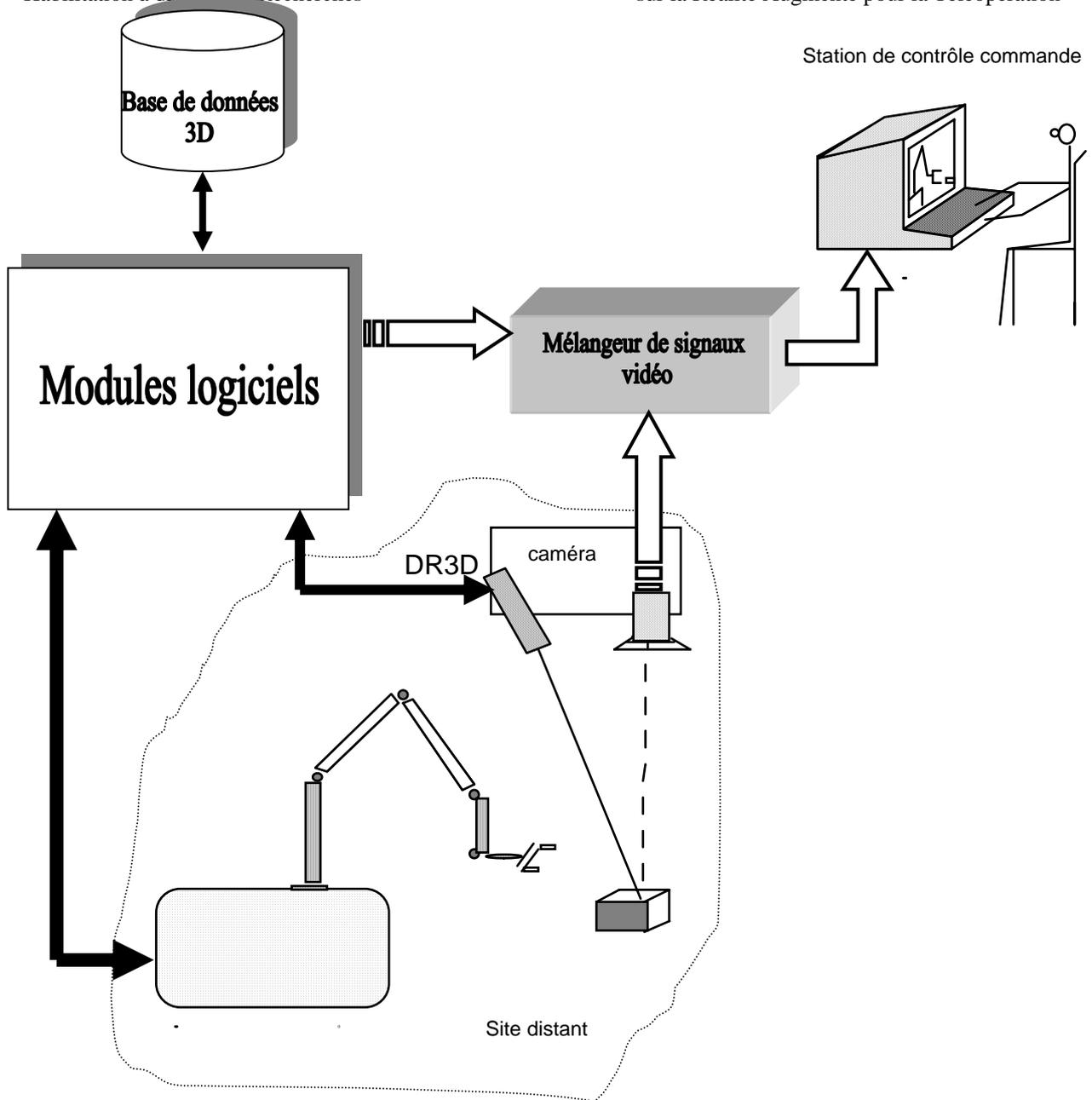


Figure 0.1: Architecture matérielle de MCIT

0.3. Description fonctionnelle du système MCIT

Le système MCIT comprend plusieurs modules (Figure 0.2):

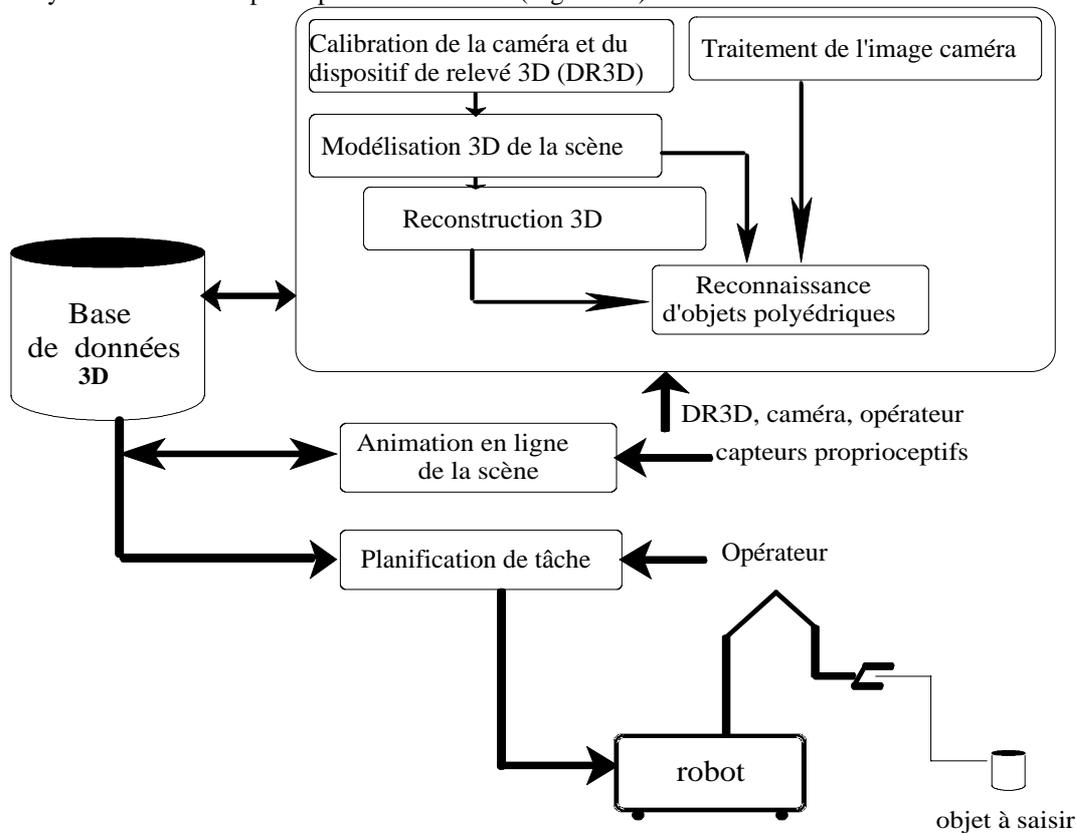


Figure 0.2 : Architecture fonctionnelle du système MCIT

Les principaux modules logiciels sont présentés brièvement.

Module de calibration de capteurs

Pour que la superposition d'images puisse être effectuée correctement, il faut connaître la transformation mathématique qui fait passer des coordonnées tridimensionnelles de l'objet aux coordonnées bidimensionnelles de son image et vice-versa. Cette transformation est obtenue par la calibration de la caméra.

Pour l'acquisition des données géométriques 3D de l'environnement, il faut également calibrer le dispositif de relevé 3D (DR3D).

Les procédures de calibration de la caméra et du dispositif de relevé 3D sont traitées dans le chapitre 1.

Module de modélisation 3D et de reconstruction 3D

Ce module consiste à construire et mettre à jour la base de données 3D de l'environnement par coopération avec l'opérateur humain.

Les procédures de modélisation 3D et de reconstruction 3D sont traitées dans les chapitres 2 & 3.

Module de reconnaissance d'objets

Ce module consiste à partir des modèles 3D fournis par le module de modélisation et du modèle 2D, résultat du traitement de l'image caméra, d'effectuer la mise en correspondance entre ces deux modèles afin de reconnaître les objets polyédriques modélisés et vus par la caméra. Ce module est décrit dans le chapitre 4.

Module d'animation

Ce module consiste à collecter les informations proprioceptives du robot et à l'animation, en ligne, de sa représentation synthétique, afin de permettre à l'opérateur de percevoir la scène même si l'image caméra est dégradée. Ce module est décrit dans le chapitre 5.

Module de planification de tâche

Ce module permet la génération automatique d'une trajectoire 2^{1/2} D d'un bras manipulateur opérant sur un plan de travail sur lequel sont entreposés des objets de faible hauteur. Il permet également de calculer les configurations de saisie par l'effecteur de l'objet désigné par l'opérateur humain sur écran vidéo. Ce module est décrit dans le chapitre 5.

0.4. Architecture informatique du système MCIT

La station de contrôle commande est constituée de 3 calculateurs reliés en réseau ethernet, dont les modules qu'ils supportent sont mentionnés dans la figure 0.3.

Cette architecture répartie est indispensable pour supporter l'ensemble des modules logiciels décrits précédemment. En effet, certains doivent s'exécuter en parallèle tels que les modules de perception de l'environnement et ceux de commande du robot.

Les différents logiciels organisés en client/serveur communiquent entre eux via internet, en utilisant les sockets selon le protocole TCP/IP.

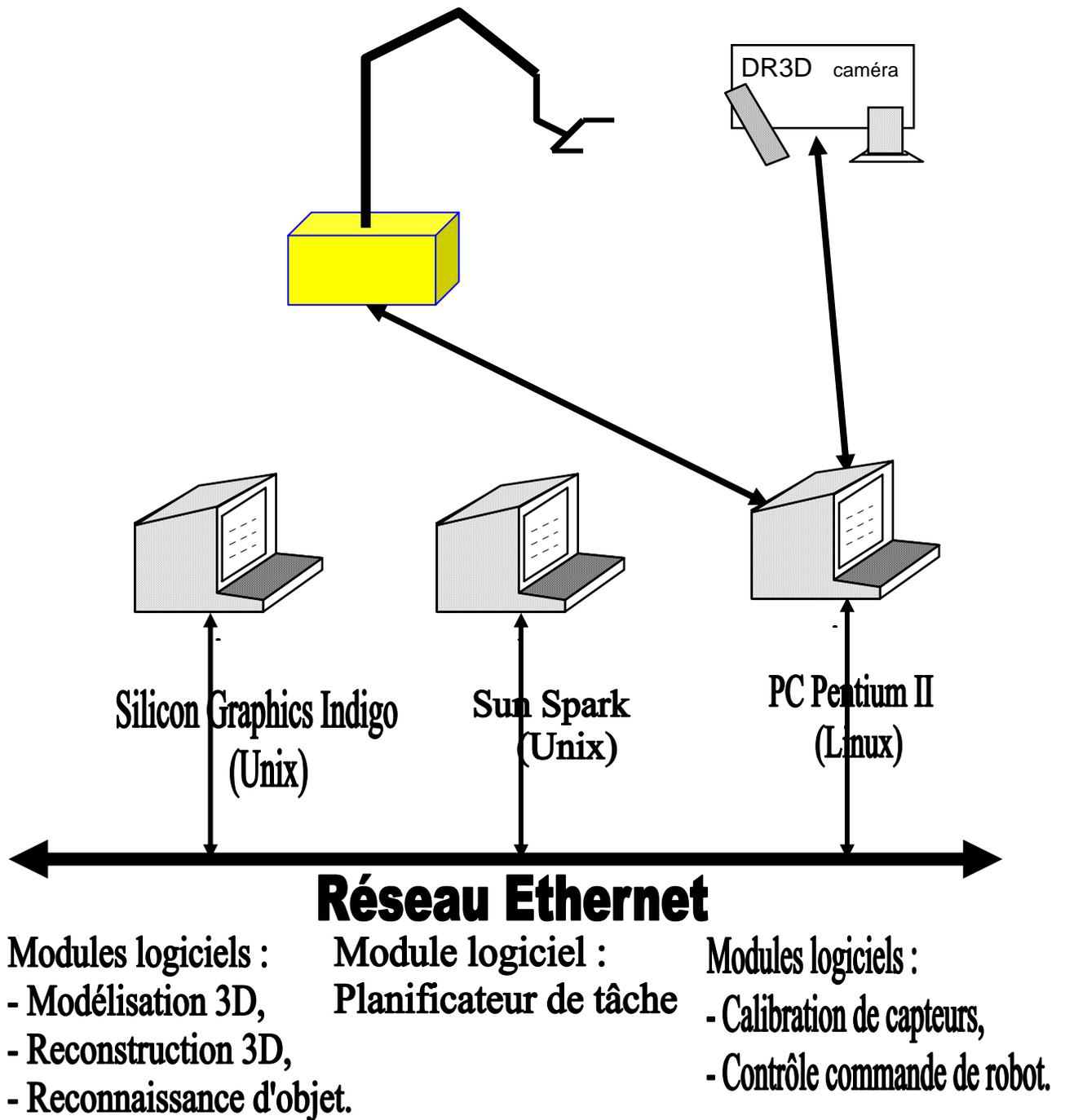


Figure 0.3 Architecture informatique du système MCIT

0.5 Bilan

La définition du système MCIT a fait l'objet de plusieurs publications, dont la 1^{ère} est parue dans "Symposium of Teleoperation and Control, Bristol, ENGLAND, Jul. 12-15, 1988"[COM 1988]³. La description d'une unité graphique intelligente, que nous avons réalisée et qui permet la superposition d'images, est publiée dans "MICROPROCESSORS AND MICROSYSTEMS, Butterworth-Heinemann Ltd, London Vol. 15 N°8, October 1991", [ART 1991]. Un brevet a également été déposé [BREVET 1991].

0.6 Bibliographie du chapitre 0

[**BEJCZY 1990**] BEJCZY A. K., KIM W. S. and VENEMA S. C. (1990): " The phantom robot: predictive displays for teleoperation with time delay". IEEE Intern. Conf. on Robotics and Automation, pp 546-551, May 13-18, 1990.

[**BONNEAU 1993**] P. BONNEAU and P. EVEN, " Man machine cooperation for 3D objects pose estimation "; IEEE/SMC'93, Systems, Man and Cybernetics, pp294 à 299, Le Touquet France October 17-20, 1993.

[**BURDEA 1993**]BURDEA G. et COIFFET P. « La réalité Virtuelle. »Hermès, France, 1993.

[**HASEGAWA 1991**] HASEGAWA T., SUEHIRO T., OGASAWARA T., MATSUI T., KITAGARI K. and TAGASE K. (1991): " A model-based tele-robot system with manipulation skills"; ISART Tokyo, Japon; pp 499-506; March 1991.

[**KIM 1993**] W. KIM, "Graphical Operator Interface for Space Telerobotics", in IEEE, Int. Conf. on Robotics and Automation (Atlanta, georgia), pp 761-768 May 2-6, 1993.

[**MALLEM 1990**]] MALLEM A.M.: " Aide à la perception en téléopération: superposition à une image caméra d'une image synthétique animée en 'temps réel' à partir d'informations capteurs". Thèse de Doctorat de l'université Paris XII, Avril 1990.

[**SHERIDAN 1989**] T.B. SHERIDAN, "Telerobotics", Automatica Vol. 25, n°4, pp 487-507, 1989.

³ Nos publications sont listées dans le chapitre 7 et les textes des plus significatives se trouvent dans le chapitre 8.

1. MODELISATION ET CALIBRATION DE CAPTEURS

La modélisation de capteurs à pour objectif de déterminer les relations mathématiques entre eux et l'environnement. Celles-ci permettent de résoudre les problèmes, de relevé 3D, de recalage 3D, et celui de la mise en coïncidence des points de vue réel et virtuel, tous liés à la RA.

La calibration des capteurs consiste à identifier les paramètres représentant les modèles définis dans la précédente étape.

Nous avons utilisé 3 types de capteurs :

- caméra mesurant un éclairage dans une direction donnée,
- télémètre laser à temps de vol monté sur une tourelle 2 ddl (site et azimut),
- capteur mixte constitué d'une caméra et d'un télémètre à faisceau laser visible.

Pour chacun d'eux, nous présentons les modèles utilisés et les méthodes de calibration mises en œuvre et les résultats obtenus. Auparavant, un bref état de l'art est présenté.

1.0 Bref état de l'art

Un modèle de capteur représente le formalisme mathématique permettant d'établir une relation analytique entre le repère associé au capteur et un autre repère de calibrage. Le modèle est alors constitué d'un ensemble de paramètres associés aux transformations mathématiques nécessaires au passage entre repère du capteur et celui du calibrage. Ces paramètres sont décrits ultérieurement.

Pour la caméra, on distingue trois modèles :

- le modèle du sténopé linéaire (sans distorsions),
- le modèle du sténopé non linéaire (avec distorsions),
- le modèle des deux plans.

Le modèle du sténopé est le plus largement utilisé, il est décrit dans §1.1.1. et l'identification de ses paramètres est explicitée dans §1.1.2. Deux types de paramètres sont considérés : ceux dits extrinsèques, qui représentent la position et l'orientation du repère de la caméra par rapport au repère de calibrage et ceux dits intrinsèques concernant le passage du repère de la caméra au repère d'affichage. Ce dernier type de paramètres représente notamment les distorsions de l'image dans le cas du modèle du sténopé avec distorsions.

Plusieurs chercheurs ont étudié ce modèle, et notamment le type de paramètres et comment les identifier.

Hung [Hung 1985] a utilisé le modèle du sténopé linéaire, a supposé connus les paramètres intrinsèques, et a proposé une méthode de calcul des paramètres externes en utilisant 4 points distincts coplanaires et non alignés.

Faugeras et Toscani ont [FAUGERAS 1987] ont développé le modèle du sténopé linéaire et ont proposé l'identification de ses paramètres en appliquant le filtrage de Kalman sans expliciter les formules de celui-ci. Celles ont été explicitées dans [AYACHE 1989].

Tsai [TSAI 1987] a appliqué le modèle du sténopé non linéaire. L'orientation de la caméra est exprimée en utilisant les angles de Bryant et l'identification de ceux-ci est basée sur une méthode d'optimisation avec contraintes. Tsai a montré, suite à une expérimentation, que la distorsion radiale est plus importante que la tangentielle. Il a proposé un modèle interne de la caméra permettant le calcul de la distorsion radiale dans le plan image. Une précision de 0.45 mm a été atteinte, sur la reconstruction de points, avec les conditions de calibration suivantes:

- distance "caméra/mire de calibration" : 1.00 m,
- focale de 25 mm,
- 60 points de calibration,
- résolution de la carte d'acquisition/visualisation non précisée.

Chaumette et Rives [CHAUMETTE 1989], ont appliqué le modèle du sténopé non linéaire. L'orientation de la caméra est exprimée en utilisant le vecteur rotation selon le formalisme de Rodrigues. Ils ont également développé un modèle interne de la caméra permettant le calcul de la distorsion radiale dans le plan d'affichage.

Martins et al [MARTINS 1981] ont proposé le modèle à deux plans qui permet de faire la correspondance entre le rayon optique dans l'espace objet et les coordonnées pixel de son image. Cette correspondance est obtenue par interpolation. Le rayon optique est défini par deux points appartenant chacun à un plan de référence. Une précision de 0.11mm a été atteinte, sur la reconstruction de points, avec les conditions de calibration suivantes:

- distance "caméra/mire de calibration" : 0.635 m
- focale de 35 mm
- deux plans de 36 points = 72 points de calibration
- résolution de la carte d'acquisition/visualisation non précisée.
- type d'interpolation: linéaire.

Lavest et al [LAVEST 1997] ont obtenu une précision subpixelique (de l'ordre du 1/100 de pixel) en améliorant la qualité de détection des amers, se trouvant dans une mire de calibration, dans plusieurs images.

Gremban et al [GREMBAN 1988] ont proposé le mixage des deux approches : ils appliquent le modèle sténopé pour le calcul de la projection perspective et celui des deux plans pour le calcul de la perspective inverse.

Le modèle des deux plans ne peut être utilisé que pour la projection inverse. Ce modèle ne représente pas une projection perspective car tous les rayons optiques ne convergent pas vers le même point. L'identification des paramètres de ce modèle est obtenue par interpolation. Une interpolation de type neuronale a été appliquée au CEMIF [BERTHOUSSE 1995].

Le modèle du sténopé est le plus utilisé. Nous avons développé des méthodes d'identification du sténopé linéaire [LOUKIL 1993]⁴, du sténopé non linéaire [TRIBOULET 1996].

Concernant la calibration de télémètre, la bibliographie est moins importante que celle de la calibration de la caméra.

Bolles et Kremers [BOLLES 1983] ont proposé une méthode de calibration d'une caméra et d'un plan laser. La caméra est calibrée et une triangulation sur un objet connu est appliquée afin d'estimer l'équation du plan laser.

Loukil [LOUKIL 1993], Chekhar [CHEKHAR 1994] et [TRIBOULET 1996] ont proposé des méthodes de calibration de télémètre 3D. Nous en présentons quelques unes dans le §1.2.2.

1.1 La caméra

1.1.1 Modèles de caméra

Le modèle géométrique de la caméra utilisé est celui du sténopé avec et sans distorsions. Dans le modèle du sténopé sans distorsions, un point 3D $P(X, Y, Z)$ et son image $p(u,v)$ sont considérés se trouver sur un rayon optique passant par le centre de la lentille et considéré non dévié (figure 1. 1).

On distingue 2 types de modèles, le modèle géométrique direct (MGD) et le modèle géométrique inverse (MGI). Le MGD est constitué par la relation $(u,v) = f(P)$, le MGI est représenté par : rayon visuel = $g(u,v)$.

Le MGD est constitué de :

- modèle interne (Mint), constitué des transformations permettant d'exprimer les coordonnées pixel, dans le plan image (R_i), du point p connu dans le plan rétinien (R_c').
- modèle externe (Mext), constitué des transformations permettant d'exprimer les coordonnées du point P , connu dans le repère du monde (R_o), dans le repère lié à la caméra (R_c).

⁴ Thèse de doctorat de Mr Abdelhamid Loukil, co encadrée avec Mr le Pr Florent Chavand

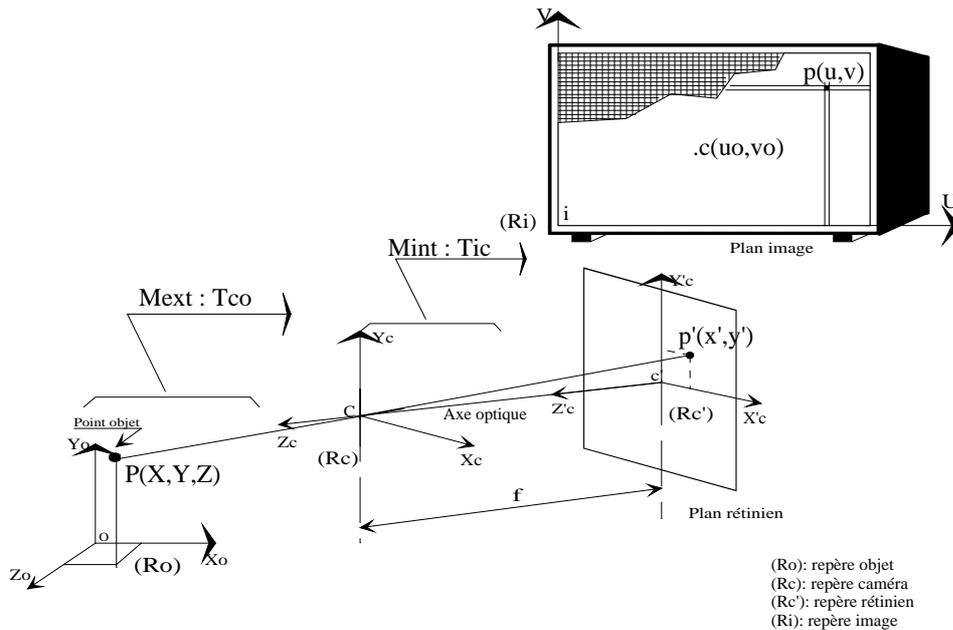


Figure 1.1. : Modèles géométriques de caméra

1.1.1.a Modèle interne de la caméra

Le modèle interne de la caméra, est représenté par les transformations permettant de passer du repère de la caméra au repère lié au plan rétinien, puis au plan image:

❶ projection dans le plan rétinien

C'est la transformation qui permet de passer du repère de la caméra (Rc) au repère rétinien (Rc') du plan rétinien. Un point P(Xc, Yc, Zc) dans (Rc), sa projection dans le plan rétinien donne p'(x', y') dans (Rc'). Les relations donnent:

$$Xc/Zc = x'/f \quad , \quad Yc/Zc = y'/f$$

❷ projection dans le plan image

C'est la transformation qui fait passer du repère (Rc') au repère (Ri). Cette transformation est une homothétie suivant les deux axes X'c et Y'c; puis un changement d'origine. On distingue 2 modèles internes :

- modèle linéaire ne tenant pas compte des distorsions :

$$\begin{cases} u = ki \cdot x' + uo = ku \cdot \frac{Xc}{Zc} + uo & ku = ki \cdot f \\ v = kj \cdot y' + vo = kv \cdot \frac{Xc}{Zc} + vo & kv = kj \cdot f \end{cases}$$

ce qui donne sous forme matricielle:

$$\begin{pmatrix} Zc \cdot u \\ Zc \cdot v \\ Zc \end{pmatrix} = \begin{pmatrix} ku & 0 & uo & 0 \\ 0 & kv & vo & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} Xc \\ Yc \\ Zc \\ 1 \end{pmatrix} = M \text{ int } (3 \times 4) \begin{pmatrix} Xc \\ Yc \\ Zc \\ 1 \end{pmatrix}$$

Mint(3x4) est le modèle interne, matrice de passage de (Rc) à (Ri) et les paramètres f, ki, kj, uo, vo s'appellent les paramètres internes de la caméra :

. uo, vo : représentent l'intersection entre l'axe optique et le plan image,

. ki, kj : les facteurs d'échelle exprimant l'homothétie entre plan caméra et plan image,

. f : distance focale de l'objectif.

- **modèle non linéaire tenant compte des distorsions :**

Lorsque l'objectif ou le capteur présente des défauts non négligeables, le modèle précédent n'est plus pertinent. Il faut affiner le modèle du sténopé en tenant compte de distorsions.

En présence d'une distorsion, l'image n'est pas exactement homologue à l'objet. Ce phénomène est causé par plusieurs défauts propres à l'objectif dont les plus importants sont des distorsions géométriques. On distingue deux types de distorsions, radiales et tangentielles.

Comme le montre la figure suivante, l'image d'un point M devait donner le point: ml avec un modèle linéaire, mt avec un modèle à distorsion tangentielle et mr avec un modèle à distorsion radiale. Pour la distorsion radiale la position du point mr est fonction du signe du facteur de distorsion (Kd) selon lequel il se trouvera à l'intérieur ou à l'extérieur du cercle.

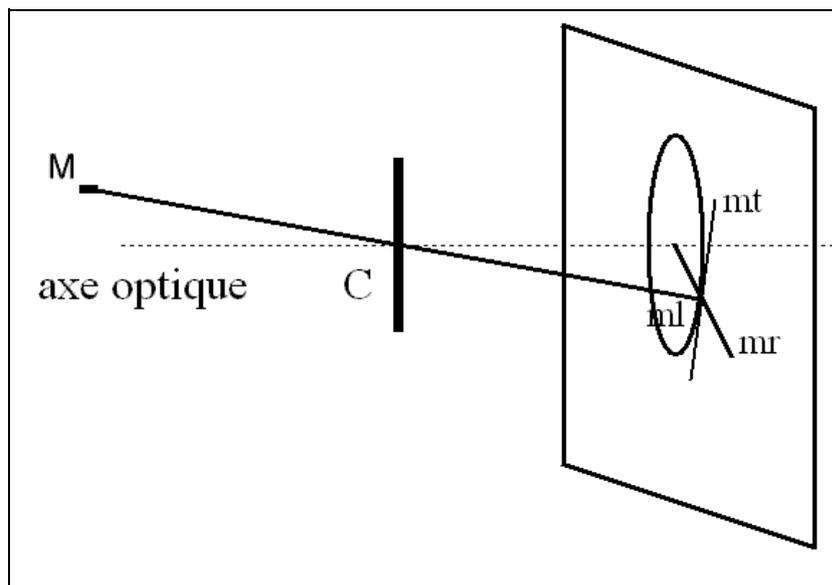


Figure 1.2 : Les distorsions radiales et tangentielles

La prise en compte de la distorsion tangentielle n'apporte pas beaucoup d'amélioration [TSAI87],[CHAUMETTE 1989], il suffit de considérer le coefficient de distorsion radiale. Le modèle interne est alors modifié, de la manière suivante :

$$\begin{cases} u = u_0 + k_u \cdot \frac{X_c}{Z_c} + k_u \cdot k_d \cdot \frac{X_c}{Z_c} \cdot \frac{(X_c^2 + Y_c^2)}{Z_c^2} & k_u = k_i \cdot f \\ v = v_0 + k_v \cdot \frac{Y_c}{Z_c} + k_v \cdot k_d \cdot \frac{Y_c}{Z_c} \cdot \frac{(X_c^2 + Y_c^2)}{Z_c^2} & k_v = k_j \cdot f \end{cases}$$

k_d est le facteur de distorsion radiale.

1.1.1.b Modèle externe de la caméra

C'est la transformation qui permet de passer du repère (R_0) au repère (R_c). Un point de la scène ne peut être connu que par rapport à un repère lié à celle-ci appelé repère de travail (R_0). Trois rotations possibles suivant les angles ($\theta_x, \theta_y, \theta_z$) et trois translations possibles suivant le vecteur $(t_a, t_b, t_c)^t$ permettent de définir la situation de (R_c) par rapport à (R_0). Un point P, de coordonnées X,Y,Z connu dans le repère R_0 , est représenté dans le repère R_c par :

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = M_{ext}(4 \times 4) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

où $R_{3 \times 3} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$ et $T_{3 \times 1} = \begin{pmatrix} t_a \\ t_b \\ t_c \end{pmatrix}$

$M_{ext}(4 \times 4)$ est le modèle externe, exprimant la situation (position et orientation) de la caméra par rapport au repère objet (R_o).

Les paramètres $(t_a, t_b, t_c, \theta_x, \theta_y, \theta_z)$, qui sont déterminés à partir de $M_{ext}(4 \times 4)$, sont appelés les paramètres externes de la caméra.

1.1.1.c Modèle global de la caméra : cas du sténopé linéaire

La matrice qui permet de passer directement du repère de travail (R_o) au repère (R_i) lié au plan d'affichage permet de déterminer le modèle global de la caméra en composant les modèles interne et externe :

$$\begin{pmatrix} Z_c.u \\ Z_c.v \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} k_i.f & 0 & u_o & 0 \\ 0 & k_j.f & v_o & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} k_i.f & 0 & u_o & 0 \\ 0 & k_j.f & v_o & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = M_{int} . M_{ext} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

qui s'écrit :

$$\begin{pmatrix} s.u \\ s.v \\ s \end{pmatrix} = C_{3 \times 4} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (1.1)$$

avec $s = Z_c$.

$$\text{et } C_{3 \times 4} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{pmatrix}$$

$C_{3 \times 4}$ représente le modèle global de la caméra, qui exprime la relation entre le repère objet (R_o) et le repère image (R_i).

Les paramètres c_{ij} sont appelés les paramètres globaux de la caméra.

1.1.1.d Modèle géométrique direct de la caméra

Ce modèle permet d'exprimer les coordonnées (u, v) du point p image du point P . La relation (1.1) est décrite par les équations :

$$\begin{cases} P^t \cdot c_1 + c_{14} - u(P^t \cdot c_3 + c_{34}) = 0 \\ P^t \cdot c_2 + c_{24} - v(P^t \cdot c_3 + c_{34}) = 0 \end{cases} \quad (1.2)$$

où :

(.) : produit scalaire,

P : vecteur de coordonnées X, Y, Z ,

$c1$: vecteur de coordonnées $(c11, c12, c13)$,

$c2$: vecteur de coordonnées $(c21, c22, c23)$,

$c3$: vecteur de coordonnées $(c31, c32, c33)$.

La relation (1.2) permet d'exprimer les coordonnées pixel (u, v) de l'image p du point P .

$$\begin{cases} u = \frac{P^t \cdot c_1 + c_{14}}{(P^t \cdot c_3 + c_{34})} \\ v = \frac{P^t \cdot c_2 + c_{24}}{(P^t \cdot c_3 + c_{34})} \end{cases} \quad (1.3)$$

Il s'agit du modèle géométrique direct de la caméra.

1.1.1.e Modèle géométrique inverse de la caméra

Ce modèle exprime l'équation du rayon optique passant par le point image p de coordonnées (u, v) .

La relation (1.2) donne deux équations indépendantes caractérisant le modèle inverse de la caméra qui peut se mettre sous la forme:

$$\begin{cases} P^t \cdot (c_1 - u \cdot c_3) + c_{14} - u c_{34} = 0 \\ P^t \cdot (c_2 - v \cdot c_3) + c_{24} - v c_{34} = 0 \end{cases} \quad (1.4)$$

qui peut s'écrire :

$$\begin{cases} P^t \cdot nu + au = 0 \\ P^t \cdot nv + av = 0 \end{cases} \quad (1.5)$$

$$\text{où } nu = \begin{pmatrix} c_{11} - u \cdot c_{31} \\ c_{21} - u \cdot c_{32} \\ c_{13} - u \cdot c_{33} \end{pmatrix}, \quad nv = \begin{pmatrix} c_{21} - v \cdot c_{31} \\ c_{22} - v \cdot c_{32} \\ c_{23} - v \cdot c_{33} \end{pmatrix}, \quad au = (c_{14} - u \cdot c_{34}), \quad av = (c_{24} - v \cdot c_{34})$$

La relation (1.5) représente deux équations de plan où nu et nv sont les normales à ceux-ci.

Le rayon optique, contenant le point P et son image p , est fourni par l'intersection de ces deux plans. Le vecteur directeur du rayon optique est $nuxnv$, (x : représente le produit vectoriel).

Celui-ci définit le modèle géométrique inverse de la caméra.

1.1.2 Calibration de caméra

La calibration de la caméra consiste à déterminer les paramètres du modèle de celle-ci. Le prérequis à ceci est l'acquisition d'un ensemble de points dans le repère objet et de leurs images.

Afin de calibrer la caméra, nous avons expérimenté successivement les modèles linéaires et non linéaire décrits dans le §1.1.1.

Pour le modèle linéaire, d'après la relation (1.2), un point donne deux équations. Afin d'identifier les paramètres globaux c_{ij} ($i=1..3$; $j=1..4$) de la caméra, il faut donc 6 points (non coplanaires, non optiquement alignés) et leurs images respectives. Pour avoir une meilleure précision lors de l'estimation de ces paramètres, par des méthodes d'optimisation mathématiques, plus de 6 points et leurs images sont nécessaires.

Nous appliquons alors, dans une première étape, la méthode des moindres carrés afin d'avoir une première estimation de ces paramètres. Dans une seconde phase, le filtrage de KALMAN est utilisé, en prenant comme estimé initial le résultat de la méthode des moindres carrés.

Pour le modèle non linéaire, nous appliquons la méthode d'optimisation de Levenberg-Marquart afin d'identifier les paramètres externes, où la rotation est représentée par le vecteur de rotation selon le formalisme de Rodrigues [CHAUMETTE 1989], et les paramètres internes incluant le facteur de distorsion radiale [TSAI 1987]. L'estimé initial est également fourni par la méthode des moindres carrés.

1.1.2.a Identification du modèle linéaire : méthode des moindres carrés

Réorganisons pour cela la relation (1.1) de la manière suivante:

$$\begin{pmatrix} X & Y & Z & 1 & 0 & 0 & 0 & 0 & -u.X & -u.Y & -u.Z \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 & -v.X & -v.Y & -v.Z \end{pmatrix} \cdot c = \begin{pmatrix} u \\ v \end{pmatrix} \quad (1.6)$$

où: $c=(c_{11},c_{12},c_{13},\dots, c_{33})^t$ est le vecteur des coefficients à identifier (une normalisation par c_{34} , qui doit être non nul, est appliquée).

Un échantillon de N points de référence mesurés :

$\tilde{P}_i = (X_i, Y_i, Z_i)^t$ ($i=1..N$) et leurs images $\tilde{p}_i = (u_i, v_i)^t$, donne $2N$ équations à 11 inconnues.

La relation (1.6) est linéaire en les coefficients c_{ij} . Pour un point P_i elle s'écrit : $\tilde{p}_i = H_i \cdot c$

où :

$$H_i = \begin{pmatrix} \tilde{P}_i^t & 1 & 0 & 0 & 0 & 0 & -u_i \cdot \tilde{P}_i^t \\ 0 & 0 & 0 & 0 & \tilde{P}_i^t & 1 & -v_i \cdot \tilde{P}_i^t \end{pmatrix}$$

Pour N points la relation (1.6), s'écrit : $\tilde{p} = H \cdot c$ (1.7)

où :

\tilde{p} : le vecteur ($2N$ lignes, 1 colonne) suivant:

$$\tilde{p} = (u_1, v_1, u_2, v_2, \dots, u_N, v_N)^t$$

et

H : la matrice ($2N$ lignes, 11 colonnes) suivante:

$$H = (H_1, H_2, \dots, H_N)^t$$

La résolution du système (1.7) est réalisée en appliquant la méthode de la pseudo-inverse :

$$\hat{c} = (H \cdot H^t)^{-1} \cdot H^t \cdot \tilde{p}$$

(H H^t) est singulière car sys 3s dimensionné (Moum d'eqn - que de variables / eqn, l'uné dep)

Le vecteur c est identifié en minimisant le critère suivant :

Critere = $Min\|(H \cdot \hat{c} - \tilde{p})\|^2$ où \tilde{p} est l'échantillon des N points images mesurés.

1.1.2.b Identification du modèle linéaire : filtrage de Kalman

La calibration de caméra par les moindres carrés ne prend pas en compte les bruits de mesures (erreur de saisie des coordonnées image et erreurs de mesure des coordonnées 3D de l'objet). Le filtrage *de Kalman* permet la modélisation du bruit de mesure.

L'équation (1.2) peut se mettre sous la forme:

$$g(m_i, c) = \begin{pmatrix} c_{11} \cdot X_i + c_{12} \cdot Y_i + c_{13} \cdot Z_i + c_{14} - c_{31} \cdot u \cdot X_i - c_{32} \cdot u \cdot Y_i - c_{33} \cdot u \cdot Z_i - u \\ c_{21} \cdot X_i + c_{22} \cdot Y_i + c_{23} \cdot Z_i + c_{24} - c_{31} \cdot v \cdot X_i - c_{32} \cdot v \cdot Y_i - c_{33} \cdot v \cdot Z_i - v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (1.8)$$

où :

$c = (c_{11}, c_{12}, c_{13}, \dots, c_{33})^t$: vecteur d'état du processus de Kalman, représentant les paramètres à estimer.

$m_i = (u_i, v_i, X_i, Y_i, Z_i)^t$: vecteur mesure exacte correspondant au point P_i ;

(X_i, Y_i, Z_i) sont les coordonnées 3D d'un point objet P_i , et (u, v) sont les coordonnées de son image en pixels.

Soit \tilde{m} le vecteur mesure approché, tel que:

$m_i - \tilde{m}_i = b_i$ représente un bruit des erreurs de mesure, supposées indépendantes entre elles, de moyenne nulle et de matrice de variance-covariance B_i connue :

$$E(b_i) = 0$$

$$E(b_i \cdot b_i^t) = B_i \neq 0$$

$$E(b_i \cdot b_j^t) = 0 \quad (i \neq j = 0, \dots, 4)$$

$E(.)$ dénote l'espérance de $(.)$.

B_i diagonale car bruit de mesure suppose, indépendants

$$B_i = \begin{pmatrix} \sigma_u^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_v^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_x^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_y^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_z^2 \end{pmatrix}$$

*$\sigma_u = \sigma_v = 1 \text{ pixel}$
 $\sigma_x = \sigma_y = \sigma_z = 0,1 \text{ mm}$
 $\sigma_x^2 = 10^{-8} \text{ m}$*

Soit c le vecteur à estimer et \hat{c}_{i-1} l'estimation au pas $i-1$;

$\hat{c}_{i-1} = c + s_{i-1}$ où s_{i-1} est un bruit indépendant des bruits de mesure b_i , associé aux paramètres du modèle et traduisant le degré de confiance de leur estimation, de moyenne nulle et de covariance S_{i-1} estimée par le filtre de Kalman.

on suppose que cette estimation est donnée avec une matrice de variance/covariance S_{i-1} symétrique définie positive :

$$S_{i-1} = E[(c_{i-1} - c) \cdot (c_{i-1} - c)^t]$$

$$S_{i-1} \neq 0 \text{ et } E(s_{i-1} \cdot b_i^t) = 0$$

$A \rightarrow N_0 = (S_{ij})_{12 \times 12} = 10^{10}$ (pas valeur car confiance faible ds les paramètres du modèle)

Disposant d'une estimation courante \hat{c}_{i-1} de c et la mesure courante \tilde{m}_i , l'équation (1.8) est linéarisée en

utilisant le développement de Taylor du 1^{er} ordre [AYACHE 1989] :

$$g(m_i, c) \approx g(\tilde{m}_i, \hat{c}_{i-1}) + \partial g / \partial m_i \cdot (m_i - \tilde{m}_i) + \partial g / \partial c \cdot (c - \hat{c}_{i-1}) \quad (1.9)$$

Notons par p_i, H_i et w_i les grandeurs suivantes:

$$p_i = \partial g / \partial c \cdot (\hat{c}_{i-1}) - g(\tilde{m}_i, \hat{c}_{i-1}) \text{ la nouvelle grandeur d'observation}$$

$$H_i = \partial g / \partial c$$

la transformation linéaire appliquée au vecteur c

$$w_i = \partial g / \partial m_i \cdot (m_i - \tilde{m}_i)$$

le vecteur bruit d'observation, dont la matrice de variance-covariance est : $W_i = (\partial g / \partial m) \cdot B_i \cdot (\partial g / \partial m)^t$

$$(2 \times 5) (5 \times 5) (5 \times 2) = (2 \times 2)$$

L'équation (1.9) peut se mettre sous la forme d'une équation d'observation linéaire :

$$p_i = H_i c + w_i \tag{1.10}$$

avec:

$$p_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix}$$

$$H_i = \frac{\partial g}{\partial c} = \begin{pmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_i X_i & -u_i Y_i & -u_i Z_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_i X_i & -v_i Y_i & -v_i Z_i \end{pmatrix}$$

$$w_i = \frac{\partial g}{\partial m_i} \cdot b_i$$

w_i est une loi gaussienne de moyenne nulle et de matrice de variances/covariances:

$$W_i = G_i \cdot B_i \cdot G_i^t$$

où:

$$G_i = \frac{\partial g}{\partial m_i} = \begin{pmatrix} -(c_{31} \cdot X_i + c_{32} \cdot Y_i + c_{33} \cdot Z_i + 1) & 0 & (c_{11} - u_i \cdot c_{31}) & (c_{12} - u_i \cdot c_{32}) & (c_{13} - u_i \cdot c_{33}) \\ 0 & -(c_{31} \cdot X_i + c_{32} \cdot Y_i + c_{33} \cdot Z_i + 1) & (c_{21} - v_i \cdot c_{31}) & (c_{22} - v_i \cdot c_{32}) & (c_{23} - v_i \cdot c_{33}) \end{pmatrix}$$

Le filtre de Kalman fournit un processus récursif qui calcule pour chaque mesure \tilde{m}_i , l'estimation optimale de c .

Les équations récursives du filtre de Kalman qui fournissent une nouvelle estimation (\hat{c}_i, S_i) de c et de la covariance de l'erreur à partir des précédentes valeurs (\hat{c}_{i-1}, S_{i-1}) sont:

12x11 $\hat{c}_i = \hat{c}_{i-1} + K_i (p_i - H_i \cdot \hat{c}_{i-1}) = 12 \times 11$ *Le gain de Kalman exprime une variation (Du et Dv)*

12x11 $K_i = S_{i-1} H_i^t (W_i + H_i S_{i-1} H_i^t)^{-1}$

12x12 $S_i = (I - K_i H_i) S_{i-1}$

où: I est la matrice unité de dimension (11x11).

Le critère à minimiser par le processus du filtrage de Kalman est:

$$Critere = Min [(c_o - c)^t \cdot S_0^{-1} \cdot (c_o - c) + \sum_{i=1}^N (p_i - H_i \cdot c)^t \cdot W_i^{-1} \cdot (p_i - H_i \cdot c)]$$

où : N est le nombre de points de mesure. *Erreur calcul* *Erreur mesure*

On remarque que le premier terme du critère calcule la distance entre c et son estimation initiale c_o pondérée par la matrice de covariance initiale S_o , tandis que les termes suivants calculent la somme des erreurs de mesure au carré, elles aussi pondérées par leurs covariances.

Algorithme:

- ↯ 1. Initialisation des matrices S_o, B_o .
- ↯ 2. Calculer l'estimé initial du vecteur d'état c par la méthode des moindres carrés avec un échantillon de 6 points de calibration.
- ↯ 3. Effectuer une mesure $(u_i, v_i, X_i, Y_i, Z_i)$.
- ↯ 4. Former la matrice G_i ; calculer W_i, p_i, H_i
- ↯ 5. Calculer le gain de Kalman : $K_i = S_{i-1} H_i^t \cdot (W_i + H_i S_{i-1} H_i^t)^{-1}$
Calculer la matrice de variances/covariance : $S_i = (I - K_i H_i) \cdot S_{i-1}$
Mettre à jour l'estimation du vecteur d'état: $\hat{c}_i = \hat{c}_{i-1} + K_i (p_i - H_i \cdot \hat{c}_{i-1})$
- ↯ 6. Aller en 3 ou arrêter le processus en fonction du critère à minimiser.

1.1.2.c Identification du modèle non linéaire : méthode de Levenberg-Marquart

La méthode d'optimisation de Levenberg-Marquart combine celle de descente du gradient, quand on est éloigné de la solution, et celle de Newton, quand on est près du minimum du critère à minimiser. Elle est appliquée afin de permettre l'identification des paramètres du modèle non linéaire de la caméra (§1.1.1a)

Le formalisme mathématique ainsi que l'algorithme de mise en œuvre de cette méthode est décrit dans l'annexe1.

Le vecteur de paramètres à estimer est constitué des paramètres intrinsèques incluant le facteur de distorsion radiale (§1.1.1a) et extrinsèques de la caméra. Le formalisme de Rodrigues est utilisé pour exprimer l'orientation de la caméra par rapport au repère de travail.

Le critère à minimiser est :

$$\text{Critere} = \text{Min} \left[\sum_{i=1}^N [(\tilde{u}_i - \hat{u}_i)^2 + (\tilde{v}_i - \hat{v}_i)^2] \right]$$

Où : N est le nombre de points,
(u_i, v_i) les coordonnées d'un point image.

1.1.3 Evaluation de la calibration de caméra

Afin de tester la validité de la méthode de calibration et de quantifier sa précision, nous avons défini une grandeur appelée "erreur de reconstruction".

Cette erreur de reconstruction peut être exprimée par rapport au repère objet (Ro) ou au repère d'affichage (Ri). Dans le cas où elle est calculée dans (Ro), c'est le MGI de la caméra (§ 1.1.1.e) qui est appliqué moyennant la connaissance de la distance caméra-mire de calibration. Dans le cas où elle est calculée dans (Ri), c'est le MGD de la caméra (§ 1.1.1.d) qui est appliqué. Cette quantité représente l'écart entre les coordonnées mesurées d'un point P de l'espace objet (resp. p de l'image), et les coordonnées calculées de ce même point.

La formulation mathématique de cette erreur est donnée ci dessous:

$$\Delta P = \left| \tilde{P} - \hat{P} \right| \quad (\text{resp. } \Delta p = \left| \tilde{p} - \hat{p} \right|)$$

1.1.3.a Evaluation de la calibration du modèle linéaire [LOUKIL 1993]

Cette évaluation est réalisée par une calibration manuelle qui nécessite la coopération de l'opérateur humain qui relève les coordonnées de points 3D dans l'espace (Ro) et leurs équivalents 2D, sur l'image 2D délivrée par la caméra, par rapport à (Ri).

Les conditions expérimentales sont spécifiées ci-dessous:

- une caméra CCD de focale 8.5 mm et de définition 756x581 photo éléments,
- la caméra est montée sur un banc de calibration (cf. figure 1.3),
- distance caméra-mire = 1,66 m,
- les points de calibration choisis ont leurs images dans la zone de l'écran où la distorsion géométrique est inférieure à 1 pixel,
- la définition de la carte graphique utilisée est de 512x256 pixels.

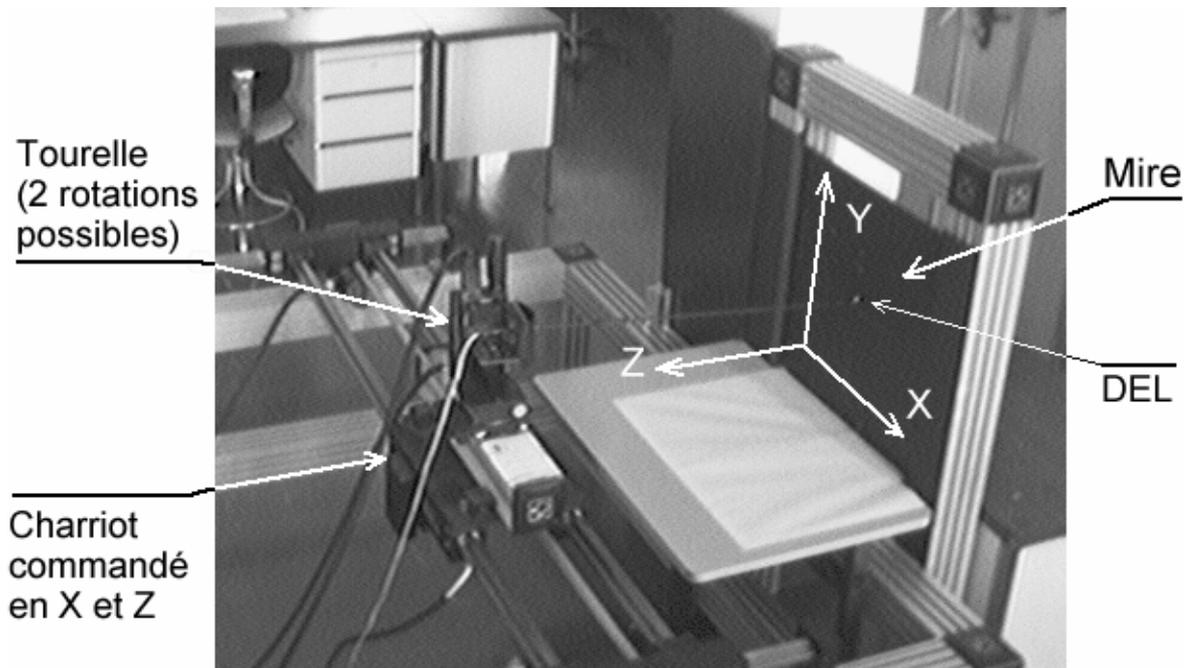
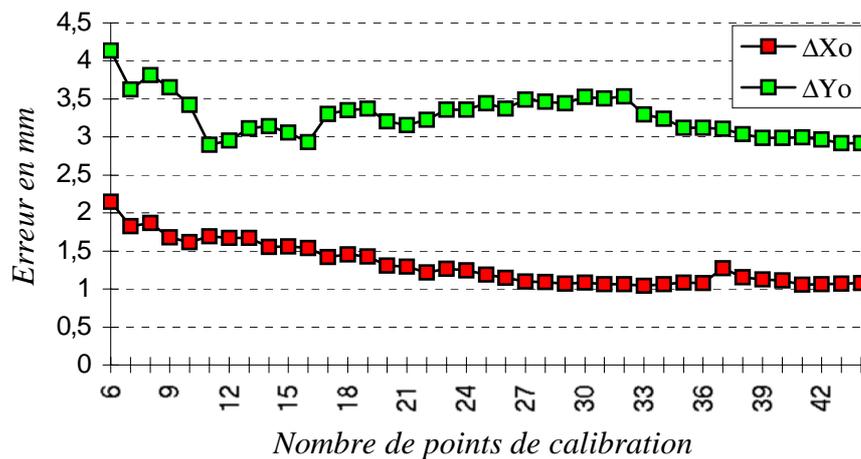
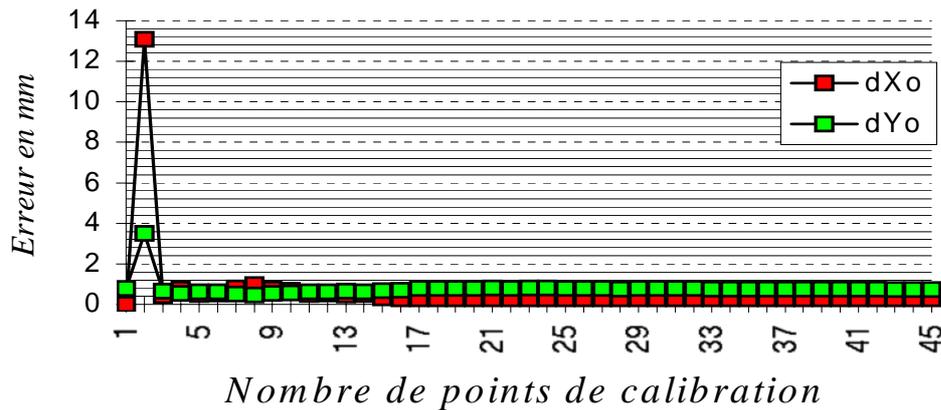


Figure 1.3 : banc de calibration

Les résultats de l'erreur de reconstruction, dans l'espace objet, sont représentées ci-dessous, respectivement en appliquant la méthode des moindres carrés puis celle du filtrage de Kalman, qui utilise comme estimé initiale le résultat de la méthode des moindres carrés.



Graphique 1.1 Variation de l'erreur de reconstruction en utilisant la méthode des moindres carrés(MMC)



Graphique 1.2 Variation de l'erreur de reconstruction en utilisant la méthode du filtrage de Kalman (FK)

Comme le montrent les graphiques ci-dessus, la calibration de caméra en utilisant le filtrage de Kalman donne des résultats nettement meilleurs que ceux obtenus en utilisant la méthode des moindres carrés car les bruits de mesure sont pris en compte.

- l'erreur de reconstruction sur l'axe (OXo): 0.04% pour le filtrage de Kalman et 0.15% pour les moindres carrés, soit 0.3 mm au lieu de 1 mm pour la distance caméra-mire utilisée.

- l'erreur de reconstruction sur l'axe (OYo): 0.08% pour le filtrage de Kalman et 0.35% pour les moindres carrés, soit 0.8 mm au lieu de 3.5 mm pour la distance caméra-mire utilisée.

Un calcul d'incertitude a montré qu'une erreur de pointage des points de calibration sur l'écran d'un pixel se traduit lorsque le modèle du capteur est parfaitement connu par une erreur de positionnement de 2 mm sur l'axe (OXo) et 4 mm sur l'axe (OYo), du repère Ro, pour une distance caméra/Mire de 1.66 m, soit en normant par rapport à Z: 0.12% en X et 0.24% en Y.

1.1.3.b Evaluation de l'apport du modèle non linéaire [LEJCZYK 1998]

Afin d'évaluer l'apport du modèle non linéaire de la caméra, nous avons mené deux séries d'expérimentation. La 1^{ère} concerne la calibration d'une caméra CCD et la 2^{ème} celle d'une caméra à tube. Afin de limiter les bruits de mesure nous avons réalisé ces 2 expérimentations en calibrant automatiquement la caméra.

La calibration automatique de la caméra est rendue possible par une pré calibration d'un mini robot à 4 degrés de liberté, constitué par la table XY du banc de calibration et de la tourelle à 2 degrés de liberté (site et azimut) sur lesquels est montée une tige au bout de laquelle est positionnée une diode électro luminescente (DEL) (figure 1.3bis). Ce mini robot permet la collecte des points de calibration.

Le processus de calibration est constitué des étapes suivantes :

- 1 - Calibration du mini robot en utilisant des points connus sur la mire de calibration,
- 2 - Pré calibration de la caméra avec 6 points fournis par les mouvements du mini robot calibré. Un point 3D est obtenu en utilisant le modèle géométrique direct du robot, son équivalent 2D est obtenu par traitement d'image (extraction du barycentre de la tâche lumineuse de la diode électro luminescente). Cette pré calibration permet d'obtenir le champ visuel de la caméra.
- 3 - Calibration de la caméra en utilisant une quarantaine de points répartis dans le champ visuel de la caméra et n'appartenant pas au même rayon visuel.

Les conditions expérimentales sont spécifiées ci-dessous:

- une caméra CCD de focale 25 mm et de définition 756x581 photo éléments,
- la caméra n'est pas embarquée sur le banc de calibration,
- distance caméra-plans de calibration : 1 - 2. m,
- les points de calibration sont répartis uniformément sur l'écran,
- la définition de la carte graphique utilisée est de 768x576 pixels.

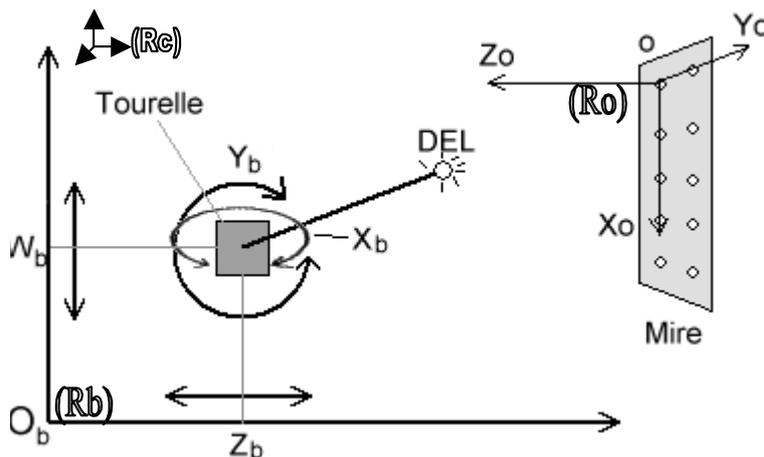


figure 1.3bis : Schéma général du dispositif de calibration (vue de dessus)

(Ro) est le repère de référence lié à la mire,

(Rb) est le repère lié au banc de calibration,

(Rc) est le repère lié à la caméra.

La calibration du mini robot consiste à déterminer la relation mathématique entre les 2 repères. Les points, représentés par des trous, 3D utilisés sont situés sur la mire du banc de calibration et sont connus au 1/10 de mm près.

Leurs équivalents 2D sont extraits automatiquement en calculant le barycentre de la tâche lumineuse formée par l'image de la diode électroluminescente, ceci permet d'obtenir une acquisition en sub-pixel. Une caméra est placée derrière la mire sur le trou concerné (figure 1.4) afin de permettre un positionnement précis de la diode électroluminescente au centre du trou.

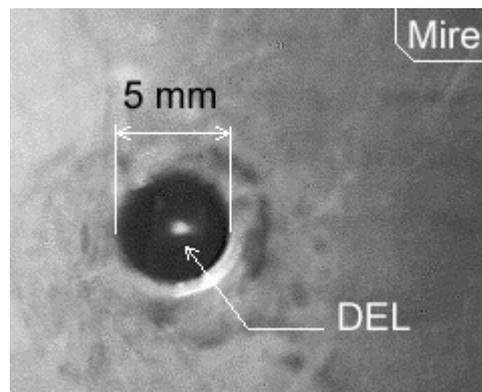


Figure 1.4 : Vue de la DEL au travers de l'un des trous de la mire

On obtient finalement des mesures beaucoup plus précises, comme le résume le tableau suivant :

Mesure effectuée	Précision (relevé manuel)	Précision (relevé automatique)
Relevé des coordonnées 3D d'un point dans l'espace par rapport à la mire	Quelques millimètres	Quelques dixièmes de millimètre
Relevé des coordonnées 2D d'un point à l'image	1 pixel	Sub pixel

Avant d'évaluer les apports du modèle non linéaire selon le type de caméra calibrée, nous avons évalué les améliorations en précision apportées par la calibration automatique.

. Comparaison entre calibration manuelle et calibration automatique

Afin de déterminer les apports en terme de précision de cette nouvelle méthode, une série de 10 calibrations manuelles et automatiques sont effectuées, en respectant les conditions expérimentales du § 1.1.3.b.

Les erreurs de reconstruction des points image pour les deux types de calibration sont résumés dans le tableau suivant :

		Moyenne relative de l'erreur	Ecart-type
Calibration manuelle	U	-0,001421	0,478
	V	0,005226	0,647
Calibration automatique	U	0,000713	0,322
	V	0,001679	0,363
Gain	U	50%	32%
	V	66%	44%

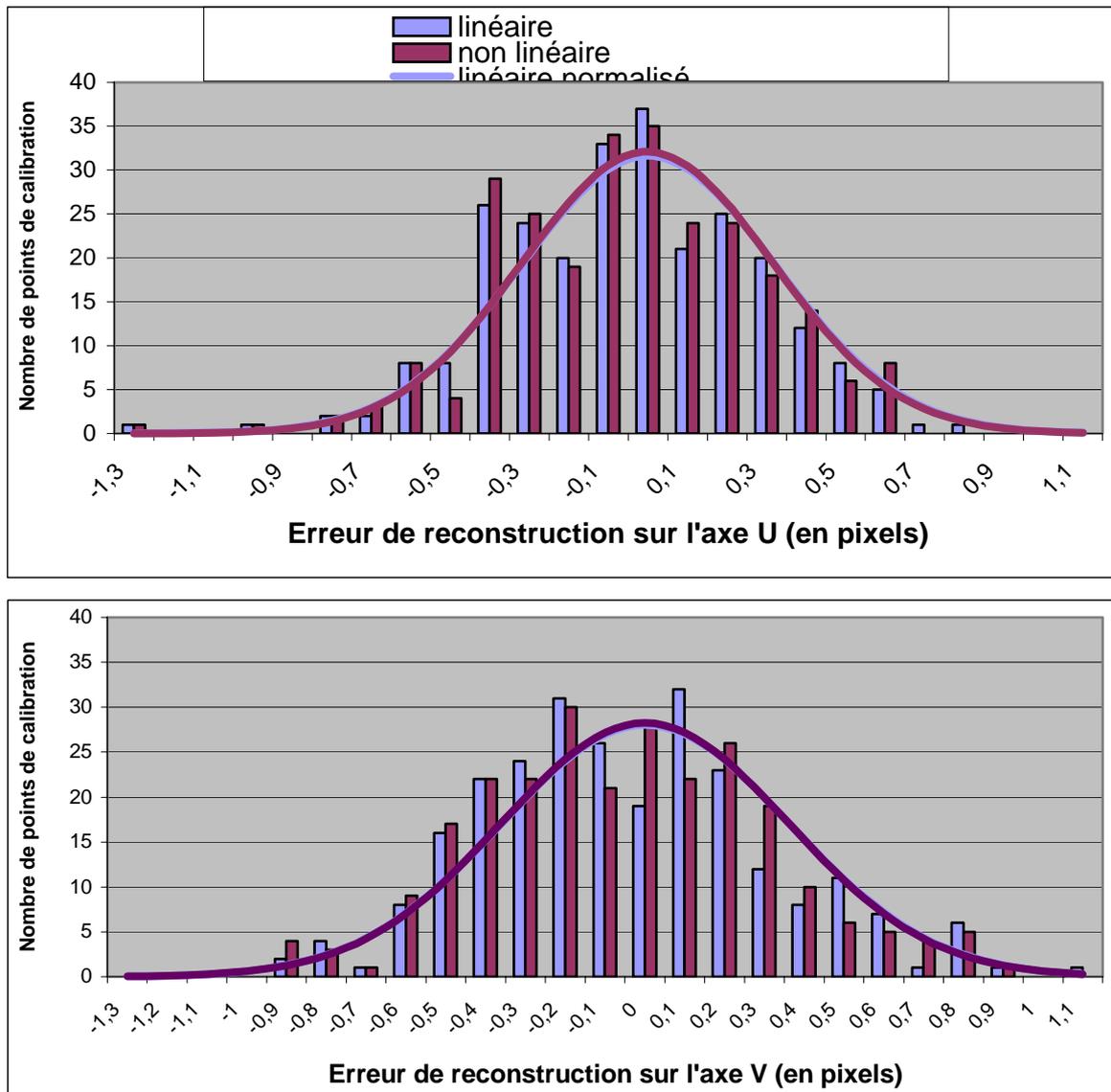
La moyenne de l'erreur est divisée par 2 voire 3. En effet, les points sont positionnés de façon beaucoup plus précise, grâce à l'utilisation du mini robot. Ce dernier est précis au dixième de millimètre près, alors que manuellement on atteint au mieux le millimètre.

Le gain sur l'écart-type est important car il est quasiment divisé par 2. Cela se traduira dans la pratique par un intervalle d'erreur divisé par 2 aussi.

. Utilité du modèle non linéaire

Afin de déterminer l'utilité du modèle non linéaire, nous avons utilisé les mêmes points de calibration que précédemment, mais appliqués aux 2 modèles de calibration (linéaire et non linéaire). Le test a été effectué avec 2 caméras différentes. L'une à capteur CCD, et l'autre à tube, en respectant les conditions expérimentales du § 1.1.3.b.

Résultats avec la caméra CCD



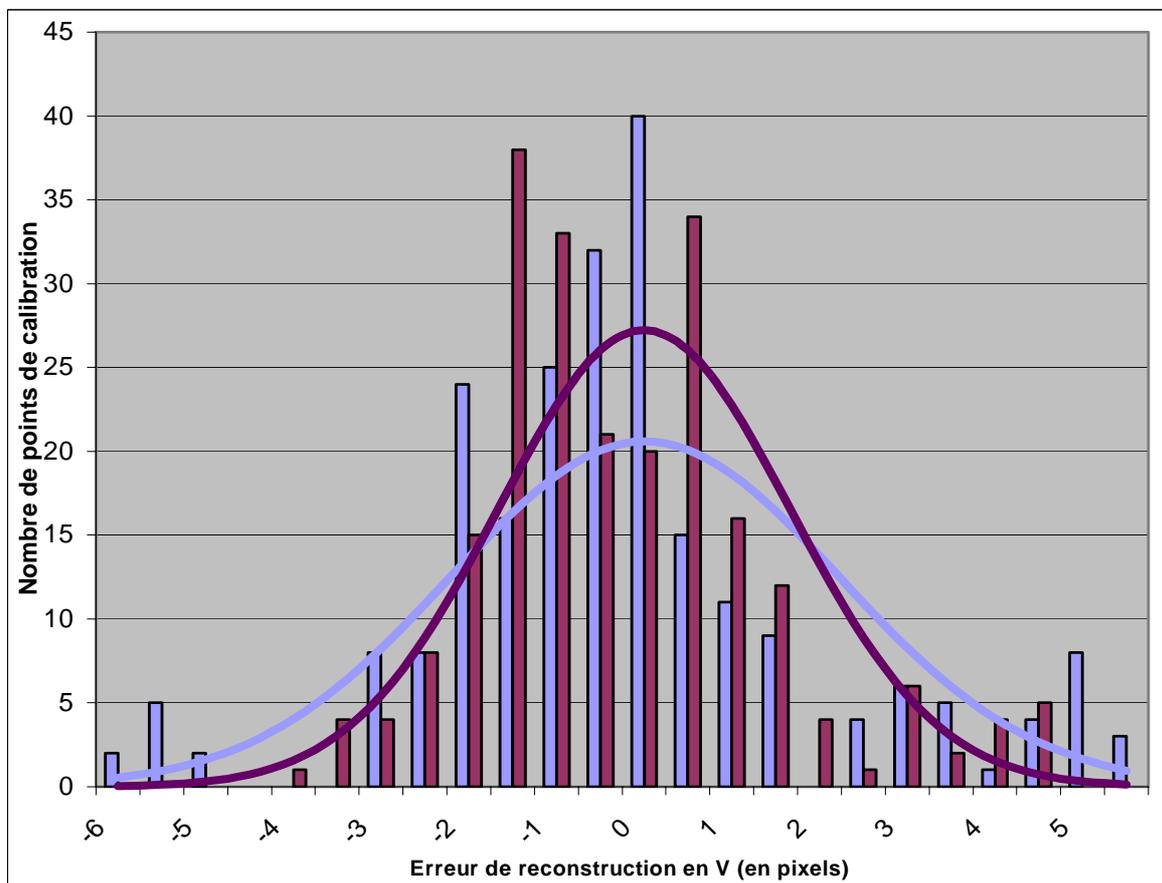
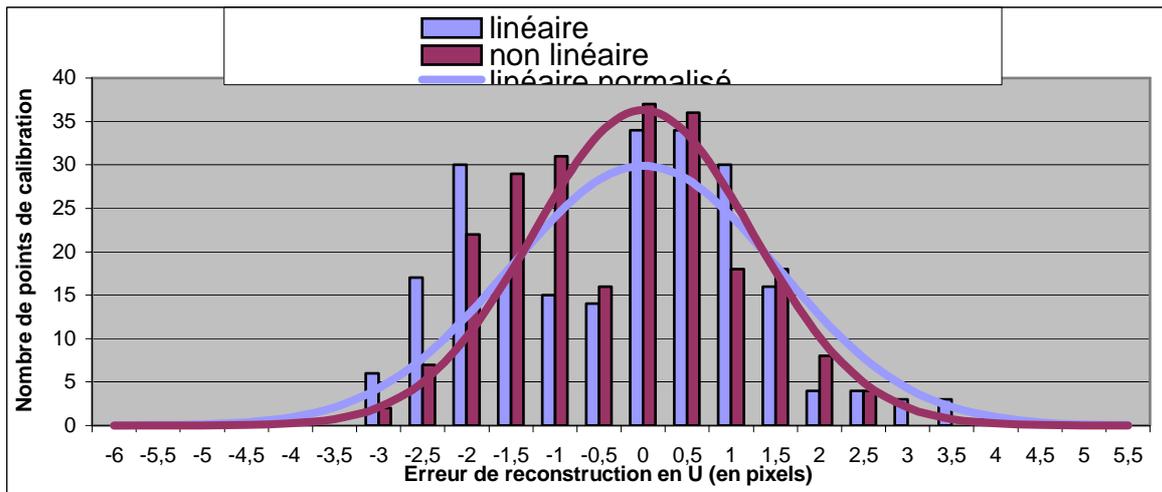
Graphique 1.3 : Répartition des erreurs de reconstruction sur les axes U et V pour une caméra CCD

En examinant les graphiques précédents, on remarque immédiatement que les résultats diffèrent très peu d'une méthode à l'autre, aussi bien pour l'axe U que V.

Cependant, les écarts sont bien centrés en 0 (les moyennes sont de l'ordre de un millièmme de pixel), ce qui indique qu'il n'existe pas de dérive ayant une composante statique.

En ce qui concerne l'écart type, on observe une diminution de 0,5% et 0,25% pour respectivement U et V. Cela indique que la caméra ne génère que très peu de distorsion. L'utilisation du modèle non linéaire ne s'impose donc pas pour ce type de caméra et pour cet objectif.

Résultats avec la caméra à tube



Graphique 1.4 : Répartition des erreurs de reconstruction sur les axes U et V pour une caméra CCD

Les graphiques permettent immédiatement de repérer une nette amélioration des résultats si on prend en compte les distorsions de cette caméra. En effet, le coefficient de distorsion K_d trouvé est de l'ordre de 0,5 alors qu'il était pratiquement nul pour la caméra de type CCD (de l'ordre de 10^{-3}).

Méthode		Moyenne relative de L'erreur	Ecart-type
Linéaire	U	0,004894	1,522342
	V	0,007631	2,218964
Non-linéaire	U	0,000123	1,252014
	V	0,000155	1,670155
Gain	U	97%	18%
	V	98%	25%

D'autre part, les résultats bruts ci-dessus indiquent que la moyenne est considérablement réduite : l'erreur résiduelle est due aux mesures et non plus à la méthode de calibration. L'écart a quand à lui sensiblement diminué.

Le modèle non – linéaire est donc très utile pour ce type de caméra, dans la mesure où un gain de précision est obtenu.

1.2. Dispositif de relevé 3D (DR3D)

Le dispositif de relevé 3D est constitué d'un télémètre laser monté sur un dispositif de rotation commandé en site et en azimut (figure 1.5).

Le support orientable permet d'orienter le laser suivant les angles θ_x et θ_y qui sont exprimés par rapport à un repère mécanique (Rm), Les mesures données par le télémètre sont exprimées par rapport à un repère qui lui est propre (Rv) (figure 1.6).

D'autre part les données contenues dans la base de données 3D sont exprimées par rapport au repère objet (Ro). Il est donc nécessaire de connaître les transformations permettant de passer d'un repère à un autre, ceci est présenté dans la définition des modèles du DR3D.

1.2.1 Modèles du DR3D

Le DR3D dispose également de paramètres intrinsèques et extrinsèques, nous présentons la formalisation de ceux-ci.

1.2.1.a. Modèle externe du DR3D : Tom

Le modèle externe est défini par la matrice de passage liant les deux repères : (Ro) (repère objet de travail) et (Rm) (repère mécanique du DR3D).

En utilisant le formalisme des matrices homogènes, la matrice *Tom* peut être décomposée en une matrice de rotation R_{3x3} et une translation T_{3x1} , représentant l'orientation et la position du repère (Rm)/(Rc)(figure 1.5).

$$Tom = \begin{pmatrix} R_{3x3} & T_{3x1} \\ 0 & 1 \end{pmatrix} \quad (1.12)$$

$$\text{où : } R_{3x3} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \text{ et } T_{3x1} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

Tom est déterminée par l'équation:

$$\begin{pmatrix} Xo \\ Yo \\ Zo \\ 1 \end{pmatrix} = Tom \cdot \begin{pmatrix} Xm \\ Ym \\ Zm \\ 1 \end{pmatrix}$$

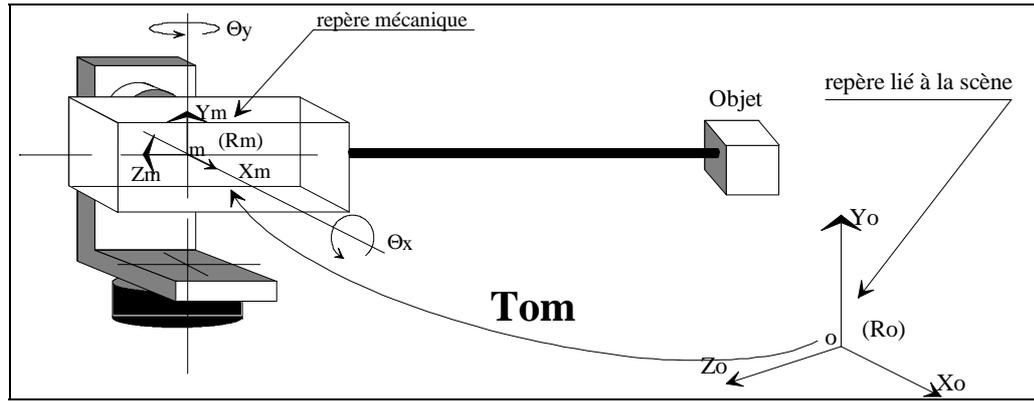


Figure 1.5 : Dispositif de relevé 3D et modèle externe

Tom est le modèle externe du DR3D et ses paramètres sont $t_x, t_y, t_z, (rij)_{3,3}$.

1.2.1.b. Modèle interne du DR3D : T_{mv}

C'est la transformation qui permet de passer de (R_m) à (R_v) (Figure 1.6). Connaissant la distance ρ donnée par le télémètre pour un point P de l'espace, on peut en déduire alors les coordonnées de ce point dans le repère mécanique par:

$$\begin{pmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{pmatrix} = T_{mv} \begin{pmatrix} 0 \\ 0 \\ -\rho \\ 1 \end{pmatrix} \quad (1.13)$$

En utilisant le formalisme des matrices homogènes, la matrice T_{mv} peut être décomposée en une matrice de rotation $R_{3 \times 3}$ d'angles γ_x et γ_y , représentant l'orientation du repère (R_v) , et une translation $T_{3 \times 1}$ de paramètres a, b , et c , représentant la position de celui-ci (figure 1.6) par rapport au repère (R_m) .

où:

$$T_{mv} = \begin{pmatrix} R_{\theta_x} \cdot R_{\theta_y} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & T_{a,b,c} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R_{\gamma_x} \cdot R_{\gamma_y} & 0 \\ 0 & 1 \end{pmatrix} = M_{int}$$

La relation (1.13) donne :

$$\begin{cases} X_m = -\rho \cdot (-C\theta_y \cdot C\gamma_x \cdot S\gamma_y + S\theta_x \cdot S\theta_y \cdot S\gamma_x - C\theta_x \cdot S\theta_y \cdot C\gamma_x \cdot C\gamma_y) + a \cdot C\theta_y + b \cdot S\theta_x \cdot S\theta_y - c \cdot C\theta_x \cdot S\theta_y \\ Y_m = -\rho \cdot (C\theta_x \cdot S\gamma_x + S\theta_x \cdot C\gamma_x \cdot C\gamma_y) + b \cdot C\theta_x + c \cdot S\theta_x \\ Z_m = -\rho \cdot (-S\theta_y \cdot C\gamma_x \cdot S\gamma_y - S\theta_x \cdot C\theta_y \cdot S\gamma_x + C\theta_x \cdot C\theta_y \cdot C\gamma_x \cdot C\gamma_y) + a \cdot S\theta_y - b \cdot S\theta_x \cdot C\theta_y - c \cdot C\theta_x \cdot C\theta_y \end{cases}$$

avec :

$$C\theta_x = \cos(\theta_x) \text{ et } S\theta_x = \sin(\theta_x),$$

T_{mv} est le modèle interne du DR3D, ses paramètres sont $a, b, c, \gamma_x, \gamma_y, \theta_x, \theta_y$.

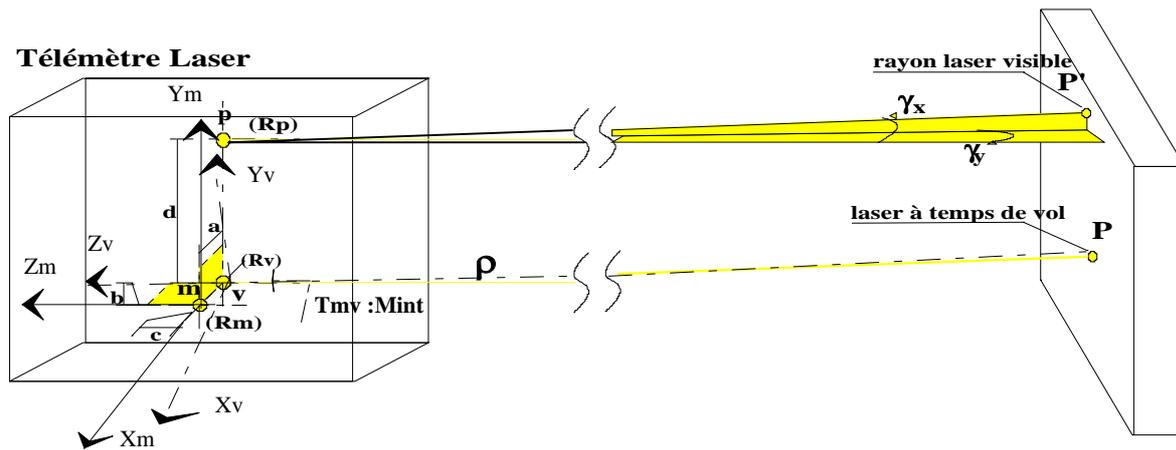


figure 1.6 : Repères liés au DR3D et modèle interne

(Rp) : repère du pointeur visible,
 (Rv) : repère du faisceau de mesure,
 d: entraxe pointeur visible - faisceau de mesure.

1.2.1.c. Modèle global du DR3D : T_{ov}

Le modèle global est donné par $T_{om}T_{mv}$, il est défini par les paramètres internes et externes du DR3D.

1.2.2 Calibration du DR3D

Le DR3D dispose également de paramètres intrinsèques et extrinsèques, nous présentons les méthodes d'identification de ceux-ci.

1.2.2.a. Calibration des paramètres internes

Le dispositif de relevé 3D est constitué d'un télémètre laser monté sur une tourelle orientable en site et en azimut. Le repère optique (Rv) est celui lié au Télémètre laser et, le repère mécanique (Rm) est celui du système rotationnel. L'origine de ces deux repères n'est pas confondu de part la technologie du support rotatif, et les axes de (Rv) et (Rm) ne sont pas forcément parallèles (Figure 1.6). Cela impose la détermination des deux paramètres de rotations (γ_x , γ_y) et des trois paramètres de translations (a, b, c) qui rentrent en jeu dans la transformation de (Rv) à (Rm) (figure 1.6).

Les trois paramètres de translation (a, b, c) ont été déduits directement des données constructeur et les deux paramètres de rotation ont été déterminés expérimentalement.

Une méthode géométrique est utilisée pour déterminer les paramètres γ_x , γ_y . Pour cela une surface plane parallèle au plan (Xm, Ym) du DR3D est déplacée suivant la direction (m, Zm). En mesurant les déplacements (dx et dy) du pointeur optique (impact du rayon laser visible sur cette surface) dans le repère (Rm) pour deux positions du plan mobile (figure 1.7), et connaissant le déplacement de ce plan (dl), on peut calculer les valeurs des angles γ_x et γ_y .

$$\text{tg}(\gamma_x) = dx/dl \quad \text{et} \quad \text{tg}(\gamma_y) = dy/dl$$

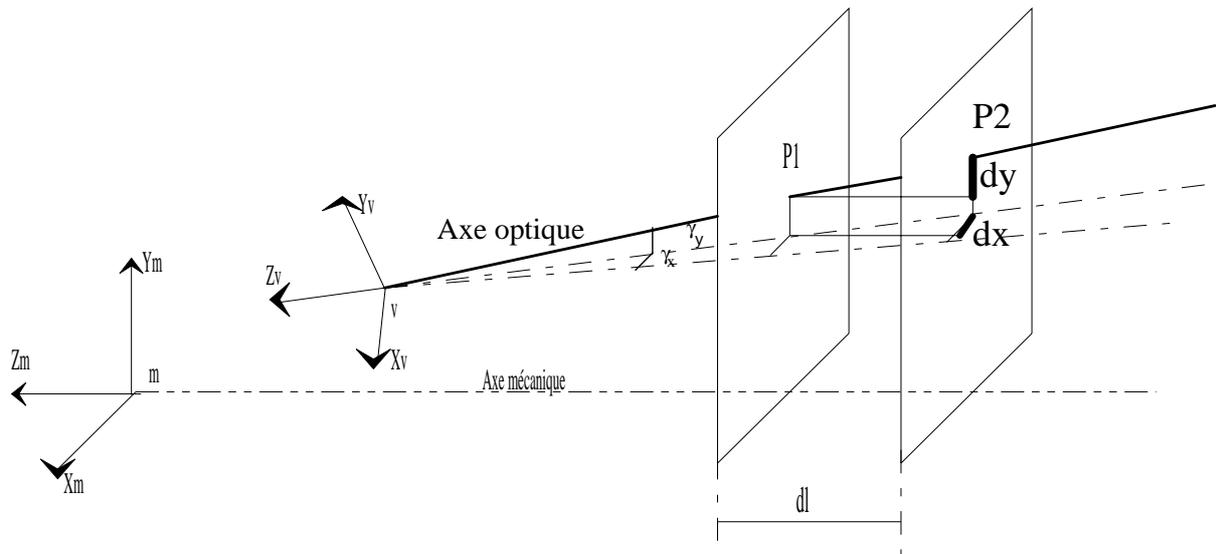


figure 1.7 : Détermination des paramètres $\underline{\gamma}_x, \underline{\gamma}_y$

1.2.2.b. Calibration des paramètres externes du DR3D : Méthode de propagation de contraintes

La calibration du modèle externe du DR3D consiste à estimer la matrice Tom connaissant un ensemble de points dont les coordonnées sont connues avec précision dans les deux repères (Ro) et (Rm). Cette estimation est réalisée par une méthode basée sur le formalisme de Lagrange, prenant en compte les contraintes de normalité, d'orthogonalité et d'orientation entre les paramètres à estimer. Ces contraintes ne sont pas toutes indépendantes.

. Mise en équation

La relation (1.12) peut s'écrire : $Tom = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} t1^t \\ t2^t \\ t3^t \\ t4^t \end{pmatrix}$

La connaissance des coordonnées de N points dans les deux repères (Ro) et (Rm) donne une équation matricielle de la forme:

$$\begin{pmatrix} X_1^0 & X_2^0 & - & - & X_N^0 \\ Y_1^0 & Y_2^0 & - & - & Y_N^0 \\ Z_1^0 & Z_2^0 & - & - & Y_N^0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_1^m & X_2^m & - & - & X_N^m \\ Y_1^m & Y_2^m & - & - & Y_N^m \\ Z_1^m & Z_2^m & - & - & Y_N^m \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (1.14)$$

. Résolution

La détermination de Tom revient à estimer les vecteurs $t_i, i=1..3$.

a. Estimation du vecteur t1 :

De l'équation (1.14), on peut former une nouvelle équation, fonction du seul vecteur t1 :

$$\begin{pmatrix} X_1^0 \\ X_2^0 \\ - \\ - \\ X_N^0 \end{pmatrix} = \begin{pmatrix} X_1^m & Y_1^m & Z_1^m & 1 \\ X_2^m & Y_2^m & Z_2^m & 1 \\ - & - & - & 1 \\ - & - & - & 1 \\ X_N^m & Y_N^m & Z_N^m & 1 \end{pmatrix} \cdot \begin{pmatrix} r_{11} \\ r_{12} \\ r_{13} \\ t_x \end{pmatrix}$$

Soit :

$$x1 = M1.t1$$

L'estimation du vecteur $t1$ revient donc à résoudre le système suivant:

$$\begin{cases} M1.t1 - x1 = 0 \\ r_{11}^2 + r_{12}^2 + r_{13}^2 = 1 \end{cases}$$

que l'on peut écrire sous la forme:

$$\begin{cases} A.r1 + B.x1 = 0 \\ \|r1\| = 1 \end{cases} \quad (1.15)$$

avec:

$$- r1^t = (r_{11}, r_{12}, r_{13})$$

$$- x1^t = (t_x, 1)$$

- $A(N \times 3)$ et $B(N \times 2)$ les matrices définies comme suit:

$$A = \begin{pmatrix} X_1^m & Y_1^m & Z_1^m \\ X_2^m & Y_2^m & Z_2^m \\ - & - & - \\ - & - & - \\ X_N^m & Y_N^m & Z_N^m \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 & -X_1^0 \\ 1 & -X_2^0 \\ 1 & - \\ 1 & - \\ 1 & -X_N^0 \end{pmatrix}$$

Le système (1.15) ainsi défini peut être résolu en termes de minimisation sous contraintes. Le critère à minimiser est donné par:

$$\text{Critere} = \text{Min} \left[\|A.r1 + B.x1\|^2 + \lambda (1 - \|r1\|^2) \right]$$

où λ est le multiplicateur de Lagrange.

Ce critère est à sa valeur minimale lorsque ses dérivées partielles par rapport à $r1$ et $x1$ sont nulles; ceci donne la solution:

$$\begin{cases} \frac{\partial \text{Critere}}{\partial r1} = A^t.A.r1 + A^t.B.x1 - \lambda r1 = 0 \\ \frac{\partial \text{Critere}}{\partial x1} = B^t.A.r1 + B^t.B.x1 = 0 \end{cases}$$

d'où :

$$x1 = -(B^t.B)^{-1}.B^t.A.r1$$

$$E.R1 = \lambda.r1$$

$$\text{avec: } E = A^t.A - A^t.B.(B^t.B)^{-1}.B^t.A \quad (E \text{ de dimension } 3 \times 3)$$

Critere à pour valeur λ si $r1$ est un vecteur propre unitaire de E (matrice symétrique positive) correspondant à la valeur propre λ . $r1$ est alors le vecteur propre unitaire de la plus petite valeur propre de E .

. Algorithme du calcul du vecteur $t1$:

1. calculer les valeurs propres de la matrice $E(3 \times 3)$
2. $r1$ est le vecteur propre de norme unité de la matrice E relatif à la plus petite valeur propre.
3. calculer $x1$ connaissant A , B et $r1$.
4. former le vecteur $t1$ à partir des deux vecteurs $r1$ et $x1$.

b. Estimation du vecteur $t2$:

La procédure est sensiblement identique à celle de l'estimation de $t1$; le seul changement réside dans l'ajout d'une équation supplémentaire pour vérifier l'orthogonalité des deux vecteurs $r1$ et $r2$ ($r1 = (r11, r12, r13)^t$, $r2 = (r21, r22, r23)^t$).

En ajoutant cette contrainte, l'équation (1.14) conduit à l'équation suivante:

$$\begin{pmatrix} Y_1^0 \\ Y_2^0 \\ - \\ - \\ Y_N^0 \\ 0 \end{pmatrix} = \begin{pmatrix} X_1^m & Y_1^m & Z_1^m & 1 \\ X_2^m & Y_2^m & Z_2^m & 1 \\ - & - & - & 1 \\ - & - & - & 1 \\ X_N^m & Y_N^m & Z_N^m & 1 \\ r11 & r12 & r13 & 0 \end{pmatrix} \cdot \begin{pmatrix} r21 \\ r22 \\ r23 \\ ty \end{pmatrix}$$

$$\Downarrow \qquad \qquad \qquad \Downarrow \qquad \qquad \qquad \Downarrow$$

$$y2 = \qquad \qquad \qquad M2 \qquad \qquad \qquad . \qquad t2$$

Cette équation peut se mettre sous la forme:

$$\begin{cases} A \cdot r2 + B \cdot x2 = 0 \\ \|r2\| = 1 \end{cases} \quad (1.16)$$

avec :

$$- r2 = (r21, r22, r23)^t$$

$$- x2 = (ty, 1)^t$$

- $A(N \times 3)$ et $B(N \times 2)$ les matrices définies comme suit :

$$A = \begin{pmatrix} X_1^m & Y_1^m & Z_1^m \\ X_2^m & Y_2^m & Z_2^m \\ - & - & - \\ - & - & - \\ X_N^m & Y_N^m & Z_N^m \\ r11 & r12 & r13 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 & -Y_1^0 \\ 1 & -Y_2^0 \\ 1 & - \\ 1 & -Y_N^0 \\ 0 & 0 \end{pmatrix}$$

Le critère à minimiser est alors:

$$\text{Critere} = \left[\|A \cdot r2 + B \cdot x2\|^2 + \lambda(1 - \|r2\|^2) \right]$$

Le vecteur t_2 est obtenu de la même manière que t_1 .

c. Estimation du vecteur t_3 :

Compte tenu des propriétés des matrices de rotation, on déduit le vecteur $r_3 = (r_{31}, r_{32}, r_{33})^t$ par le produit vectoriel, calculés dans les deux premières étapes, des deux vecteurs r_1 et r_2 donne la contrainte d'orientation :

$$r_3 = r_1 \times r_2 \quad \text{tels que :} \begin{cases} r_{31} = r_{12}.r_{23} - r_{13}.r_{22} \\ r_{32} = r_{13}.r_{21} - r_{11}.r_{23} \\ r_{33} = r_{11}.r_{22} - r_{12}.r_{21} \end{cases}$$

Connaissant le vecteur r_3 , nous pouvons déduire le paramètre t_z de l'équation matricielle suivante:

$$\begin{pmatrix} Z_1^0 \\ Z_2^0 \\ - \\ - \\ Z_N^0 \end{pmatrix} = \begin{pmatrix} X_1^m & Y_1^m & Z_1^m & 1 \\ X_2^m & Y_2^m & Z_2^m & 1 \\ - & - & - & 1 \\ - & - & - & 1 \\ X_N^m & Y_N^m & Z_N^m & 1 \end{pmatrix} \cdot \begin{pmatrix} r_{31} \\ r_{32} \\ r_{33} \\ t_z \end{pmatrix}$$

Etant donné que les coordonnées mesurées des points de calibration sont bruitées, nous pouvons faire un moyennage de t_z pour les N points de calibration.

$$t_z = \frac{1}{N} \sum_{k=1}^N \left(Z_k^0 - [X_k^m \cdot r_{31} + Y_k^m \cdot r_{32} + Z_k^m \cdot r_{33}] \right)$$

Note

L'estimation de la matrice Tom par cette méthode nécessite un nombre de points de calibration au moins égal à quatre ($N \geq 4$).

1.2.3 Evaluation de la calibration du DR3D

. Expérimentation et résultats

Afin de tester la validité du modèle externe estimé par notre méthode de calibration, nous avons calculé l'erreur de reconstruction pour chaque point qui a contribué à la calibration.

Cette erreur est définie comme étant l'écart entre les coordonnées des points dans le repère (R_0) et les coordonnées des mêmes points reconstruits à partir de la matrice Tom estimée et les mesures données par le DR3D (ρ, θ_x, θ_y) dans le repère (R_m).

L'erreur de reconstruction est définie par:

$$\bar{\Delta} = \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} \tilde{X}_0 \\ \tilde{Y}_0 \\ \tilde{Z}_0 \end{pmatrix} - \begin{pmatrix} \hat{X}_0 \\ \hat{Y}_0 \\ \hat{Z}_0 \end{pmatrix} \quad \text{avec :} \quad \begin{pmatrix} \hat{X}_0 \\ \hat{Y}_0 \\ \hat{Z}_0 \\ 1 \end{pmatrix} = \text{Tom} \cdot \begin{pmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{pmatrix}$$

où: X_m , Y_m , Z_m sont obtenues à partir de la mesure $(\rho, \theta_x, \theta_y)$ en utilisant le modèle interne du DR3D.
Les résultats que nous allons présenter ont été obtenus pour une position moyenne "mire/DR3D" égale à 1.4 m.

<i>N</i>	<i>Δx (mm)</i>	<i>Δy (mm)</i>	<i>Δz (mm)</i>
4	2.6	6.3	7.7
8	2.3	3.4	6.6
12	2.0	3.4	5.3
18	1.6	3.8	4.9

1.3 Choix du capteur de relevé 3D

Le choix du capteur le plus approprié pour le relevé 3D a été accompli suite à une étude portant sur l'évolution de l'incertitude sur les mesures. En effet, les erreurs limites sur la profondeur sont calculées pour les capteurs suivants :

- télémètre à temps de vol monté sur une tourelle à 2 degrés de liberté (télémètre 3D ou DR3D),
- ce même capteur couplé à une caméra vidéo (télémètre à triangulation),
- 2 caméras vidéo (stéréovision par triangulation).

1.3.1. Erreur limite du télémètre à temps de vol 3D (DR3D)

A partir de l'équation, donnant la profondeur, du modèle interne du DR3D (1.13b), l'incertitude relative à la mesure de profondeur (ρ), sur la distance (Z_m) objet - télémètre conduit à l'expression suivante:

$$\Delta Z_m = \Delta \rho + (-2\rho(\gamma_x + \gamma_y) + a)\Delta\theta$$

avec comme hypothèses :

- le télémètre est placé sur le zéro mécanique, tel que $\theta_x = \theta_y = 0$,
- $\gamma_x = -1.15^\circ$, $\gamma_y = -0.23^\circ$, $a = -0.5 \pm 0.001 \text{ mm}$,
- et $\Delta\theta_x = \Delta\theta_y = \Delta\gamma_x = \Delta\gamma_y = \Delta\theta = 0.01^\circ$ (données du constructeur du DR3D).

Compte tenu de ces valeurs numériques : $\Delta Z_m \approx \Delta \rho$.

L'incertitude sur la mesure ρ est prise égale à l'écart type sur la mesure $\sigma\rho (= 3 \text{ mm dans le meilleur cas et } 20 \text{ mm dans le pire cas})$ (cf. figure 1.9).

1.3.2. Erreur limite du télémètre à triangulation

Ce dispositif utilise le pointeur visible du télémètre laser et un réticule construit sur le système de visualisation. Le relevé 3D d'un point d'un objet est obtenu par triangulation. Celle-ci est obtenue par l'intersection entre la droite support du faisceau laser et celle du rayon visuel obtenue avec le réticule (cf. figure 1.8).

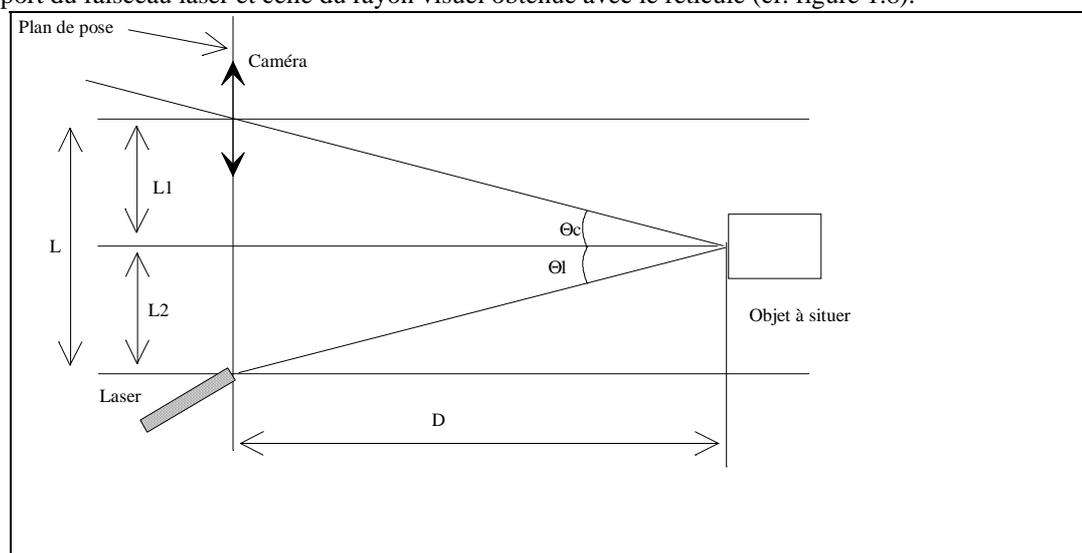


Figure 1.8 : Triangulation caméra - pointeur laser visible

La distance D et l'incertitude absolue sont exprimés par :

$$D = \frac{L}{\tan(\theta_C) + \tan(\theta_L)}$$

$$\Delta D = D \cdot \left(\frac{\Delta L}{L} + \frac{D}{L \cdot \cos^2(\theta_C)} \cdot \Delta\theta_C + \frac{D}{L \cdot \cos^2(\theta_L)} \cdot \Delta\theta_L \right)$$

avec:

D : la distance télémètre objet,

L : l'entre axe caméra - laser (connu),

θ_L : l'angle entre le faisceau laser et la perpendiculaire au plan de pose du télémètre,

θ_C : l'angle entre le rayon optique de la caméra et la perpendiculaire au plan de pose du télémètre.

L'incertitude absolue $\Delta\theta_C$ est obtenue à partir des données de la calibration de la caméra [LOUKIL 1993] : $\Delta\theta_C = 1.28 \cdot 10^{-3}$ rad.

L'incertitude absolue angulaire du télémètre est celle fournie par le constructeur :

$\Delta\theta_L = 1/100^\circ = 1.74 \cdot 10^{-4}$ rad.

L'incertitude absolue sur l'écart entre les 2 capteurs est : $\Delta L = 1$ mm.

La figure 1.9 montre l'évolution de cette erreur.

1.3.3. Erreur limite de la stéréovision

L'erreur limite de la stéréovision est calculée de la même façon que dans la partie précédente en remplaçant les paramètres du faisceau laser par ceux d'une caméra identique à la première: il s'agit de l'application de la triangulation en stéréovision.

$$\Delta D = D \cdot \left(\frac{\Delta L}{L} + \frac{D}{L \cdot \cos^2(\theta_{c1})} \cdot \Delta\theta_{c1} + \frac{D}{L \cdot \cos^2(\theta_{c2})} \cdot \Delta\theta_{c2} \right)$$

Les caractéristiques sont prises identiques pour les 2 caméras tels que :

$\Delta\theta_{c1} = \Delta\theta_{c2} = \Delta\theta_c = 0.01^\circ = 1.28 \cdot 10^{-3}$ rad et $\theta_{c1} = \theta_{c2} = \theta_c$.

L'équation précédente s'écrit alors :

$$\Delta D = D \cdot \left(\frac{\Delta L}{L} + \frac{2 \cdot D}{L \cdot \cos^2(\theta_c)} \cdot \Delta\theta_c \right)$$

et fournit aussi une forme parabolique.

La figure 1.9 montre l'évolution de l'incertitude de mesure en fonction de la profondeur dans les trois cas précédents :

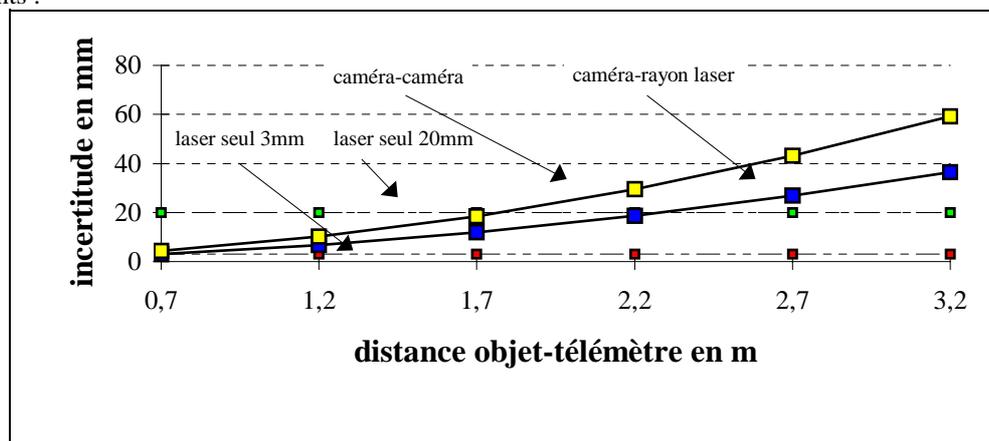


Figure 1.9: Comparaison des incertitudes pour les différents capteurs utilisés

Cette étude montre que :

- L'utilisation d'un télémètre laser à temps de vol monté sur une tourelle à 2 degrés de liberté procure une erreur de mesure quasiment constante dans la zone d'utilisation du télémètre et assimilable à celle donnée par le constructeur. L'inconvénient du télémètre 3D est la lenteur des mesures.
- L'utilisation de la coopération caméra - pointeur laser procure une erreur qui varie selon une loi parabolique.

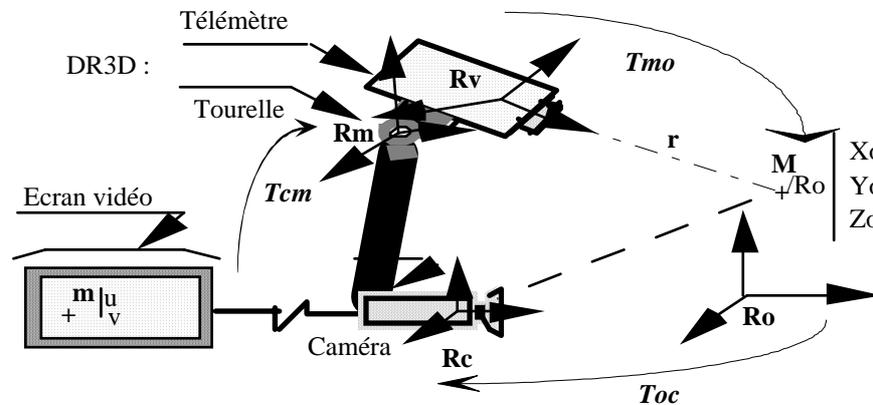
- L'utilisation de la coopération de deux caméras procure également une erreur qui varie aussi selon une loi parabolique mais plus importante dans ce cas.

Par ailleurs le télémètre laser à temps de vol présente une erreur fonction de l'angle d'inclinaison du faisceau par rapport à la surface (phénomène dû à la largeur du faisceau). Ceci nous a conduit à mettre en œuvre des procédés utilisant une coopération caméra - télémètre .

1.4. Capteur mixte : Caméra + DR3D

1.4.1 Description

Le dispositif de relevé 3D est constitué d'un télémètre laser monté sur un dispositif de rotation commandé en site et en azimut (DR3D) et de la caméra (cf. figure 1.10). Les repères du DR3D et de la caméra sont fixes l'un par rapport à l'autre.



R_m : Repère mécanique du DR3D

R_v : Repère du télémètre

R_c : Repère de la caméra

R_o : Repère objet

T_{mo} : modèle externe du DR3D

T_{oc} : modèle externe de la caméra

T_{cm} : matrice de passage entre repères de la caméra et du DR3D

Figure 1.10 : Capteur mixte caméra - DR3D

Ce capteur est utilisé pour effectuer un relevé 3D par triangulation. Dans ce cas, c'est le faisceau laser visible qui est utilisé.

1.4.2 Modèles

Les modèles de la caméra sont présentés au §1.1.1, et ceux du DR3D sont également présentés au §1.2.1. La matrice T_{cm} est constante, et est obtenue grâce aux modèles externes de la caméra et du DR3D.

1.4.3 Calibration

La calibration du capteur mixte revient à calibrer séparément la caméra (§1.1.2) et le DR3D (§1.2.2).

1.5 Bilan

Le modèle géométrique de la caméra utilisé est celui du sténopé avec et sans distorsions. Plusieurs méthodes d'optimisation mathématiques ont été utilisées. Pour le modèle linéaire, nous appliquons, dans une première étape, la méthode des moindres carrés afin d'avoir une première estimation de ces paramètres. Dans une seconde phase, le filtrage de KALMAN est utilisé, en prenant comme estimé initial le résultat de la méthode des moindres carrés. En présence de bruit de mesure, le filtrage de Kalman améliore la précision. Pour le modèle non linéaire, nous appliquons la méthode d'optimisation mathématique de Levenberg-Marquart en utilisant un estimé initial fourni par la méthode des moindres carrés.

La calibration de la caméra a été rendue automatique car elle permet de :

- minimiser les erreurs de saisie des points de calibration,
- d'obtenir une bonne précision dans l'estimation des paramètres de la caméra – l'erreur de reconstruction est divisée par deux - ,
- de s'affranchir de l'utilisation d'un objet étalon et
- d'effectuer un grand nombre de tests pour déterminer la stratégie quand au choix du modèle de caméra (linéaire ou non) à appliquer. Nous avons montré que le modèle non linéaire ne s'impose pas pour des caméras générant de faibles distorsions.

L'erreur de reconstruction est inférieure au pixel (de l'ordre du 1/10 de pixel) pour une caméra à faibles distorsions géométriques et du pixel pour une caméra à tube.

Par ailleurs, la calibration automatique, basée sur la pré calibration d'un mini robot peut être réalisée en ligne, si la caméra devient mobile, sans l'intervention de l'opérateur humain. Cette possibilité est très utile pour la réalisation de tâche en réalité augmentée.

Une procédure de calibration du dispositif de relevé 3D a été étudiée et testée, avec une application à la modélisation de volumes (chapitre 2). Le modèle géométrique a été séparé en un modèle interne et un modèle externe. Ce dernier est constitué par la matrice homogène qui situe le capteur par rapport au repère de référence. Les procédures de détermination de cette matrice permettent de tenir compte des contraintes de norme et d'orthogonalité de la matrice de rotation .

Le capteur mixte constitué du DR3D et de la caméra a également été calibré. Il est appliqué au relevé 3D par triangulation afin de générer des volumes englobants les objets de l'environnement du robot. Cette modélisation fait l'objet du chapitre suivant.

Ces travaux sur la calibration de capteurs ont été publiés dans plusieurs papiers dont , la revue Robotica 1992 [ART 1992], les conférences: International Symposium on Robotics Mechatronics and Manufacturing Systems'92 [COM 1992] , IEEE/SMC'93 [COM 1993b] , ORIA'94 [COM 1994c] et la revue "Automatique et Productique appliquée 8(4)/1995" [ART 1995].

1.6 Annexe1

Formalisme mathématique de la méthode de Levenberg-Marquart [PRESS 1992]:

Soient :

p : le vecteur des paramètres à estimer,

\hat{p}_k : le critère à la $K^{\text{ème}}$ itération,

$\nabla \hat{p}_k$: le gradient du critère d'erreur,

$H(\hat{p}_k)$: son hessien,

et Δp la Variation du vecteur de paramètres.

L'équation à résoudre est :

$$(H(\hat{p}_k) + \lambda_k \cdot I) \cdot \Delta p = -\nabla \hat{p}_k ; \quad (1.11)$$

$\lambda_k \geq 0$ et unidimensionnel ,

si $\lambda_k \rightarrow 0$: la méthode est équivalente à celle de Gauss – Newton,

si $\lambda_k \rightarrow \infty$: la méthode est équivalente à celle du gradient.

Calcul du gradient et du Hessien

Notons :

e : la fonction d'erreur

w : la fonction de poids

Si le critère d'erreur s'exprime sous la forme:

$$j(p) = \frac{1}{2} \sum_{i=1}^{nt} w_i [e(t_i, p)]^2$$

alors son gradient vérifie:

$$g(p) = \sum_{i=1}^{nt} w_i \frac{\partial e(t_i, p)}{\partial p} \cdot e(t_i, p)$$

et son Hessien s'écrit :

$$H(p) = \sum_{i=1}^{nt} w_i \frac{\partial e(t_i, p)}{\partial p} \cdot \frac{\partial e(t_i, p)}{\partial p^t} + \sum_{i=1}^{nt} w_i \frac{\partial^2 e(t_i, p)}{\partial p \partial p^t} \cdot e(t_i, p)$$

Si l'on néglige les termes du 2nd ordre, on obtient, donc, un hessien approché qui

est le produit du jacobien par le jacobien transposé :

$$H_a(p) = \sum_{i=1}^{nt} w_i \frac{\partial e(t_i, p)}{\partial p} \cdot \frac{\partial e(t_i, p)}{\partial p^t}$$

L'Algorithme, de mise en œuvre de la méthode de Levenberg-Marquart, est le suivant [PRESS 1992]:

1. Initialiser p , calculer le critère initial et initialiser λ à une valeur assez faible (10^{-3}).
2. Résoudre l'équation (1.11) et déterminer la correction Δp .
3. Avec le nouveau vecteur de paramètre, calculer le nouveau critère d'erreur
 - si celui-ci est inférieur au précédent critère calculé, alors
on conserve p , on divise λ par 10 et on recommence à l'étape 2.
 - sinon
on reprend le résultat précédent, on multiplie λ par 10 et on recommence à l'étape 2.

L'algorithme s'arrête quand le critère d'erreur est suffisamment petit, ou lorsque la norme de Δp devient trop faible (c.a.d: qu'à chaque cycle, le vecteur de paramètres varie peu).

Remarque :

Si l'estimation des paramètres est trop éloignée de la solution, la méthode de Levenberg-Marquard augmente le paramètre λ (pas de calcul) ce qui change de manière significative les valeurs des paramètres. En revanche, si l'estimation des paramètres se trouve au voisinage de la solution λ est diminué.

1.7 Bibliographie du chapitre 1

- [**AYACHE 1989**] AYACHE N.: " Vision stéréoscopique et perception multisensorielle: application à la robotique mobile"; InterEditions, 1989.
- [**BERTHOUBE 1995**] L. BERTHOUBE : "Vision en Robotique : de la reconstruction 3D à une approche comportementale", Thèse d'université, Evry Val d'Essonne, 1996.
- [**BOLLES 1983**] BOLLES R.C. , KREMERS J. and CAIN R.A. "Projector-Camera Range sensing of Three-Dimensional DATA", Machine Intelligence Research Applied to Industrial Automation, pp. 29-43, Jan. 1983.
- [**CHAUMETTE 1989**] F. CHAUMETTE et P. RIVES: " Modélisation et calibration d'une caméra", 7^{ème} Congrès AFCET, RFIA, pp. 527-531, 1989.
- [**CHEKHAR 1994**] Y. CHEKHAR, "Saisie et traitement d'images télémétriques -Application à la téléopération-", thèse de doctorat de l'université d'Evry Val d'Essonne; 13 DEC 1994.
- [**FAUGERAS 1987**] O. G. FAUGERAS, G. TOSCANI, "Camera Calibration for 3D Computer Vision", Proc. On Intern. Workshop on Industrial Application of Machine Vision and Machine Intelligence, pp. 240-247, Seiken Symposium, Tokyo, Feb. 1987.
- [**GREMBAN 1988**] GREMBAN K.D., THORP E.D. and KANADE T."Geometric camera calibration using systems of linear equations", Int. Conf. On Robotics & Automation, pp. 562-567, 1988.
- [**HUNG 1985**] HUNG Y. P., PEN-SHU YEH, HARWOOD D.: " Passive Ranging to known Planar Points Set", IEEE, Int. Conf. On Robotics and Automation, pp. 80-85, St Louis (US) , 1985.
- [**LAVEST 1997**]LAVEST J.M., et VIALA M. "Une mire d'étalonnage : à quelle précision", RFIA'98, 11^{ème} congrès de Reconnaissance des formes et d'intelligence artificielle, 20-22 janvier 1998, Clermont Ferrand, pp. III.35-III.43.
- [**LEJCZYK 1998**]S. LEJCZYK« Appariement d'image 2D avec son modèle 3D en vue de la reconnaissance d'objet. »Mémoire d'ingénieur, CEMIF-Systèmes Complexes, Université D' EVRY VAL D'ESSONNE, juin1998
- [**LOUKIL 1993**] A. LOUKIL, "Interface homme-machine da contrôle-commande an robotique téléopérée", thèse de doctorat de l'université d'Evry Val d'Essonne; 09 DEC 1993.
- [**MARTINS 1981**] MARTINS H. A., BIRK J. R. and KELLEY R. B. (1981): "Camera models based on data from two calibration planes". Computer graphics and image processing; Vol. 17, pp 173-180, 1981.
- [**PRESS 1992**] PRESS W. H., TEUKOLSKY S. A., VETTERLING W. T. and FLANNERY B. P. (1992): " Numerical Recipes in C: the art of scientific computing"; Cambridge University Press 1992.
- [**TRIBOULET 1996**] TRIBOULET J. "Système multicapteur pour la modélisation d'environnement", thèse de doctorat de l'université d'Evry Val d'Essonne; Janvier 1996.
- [**TSAI 1987**] TSAI R. Y. : " A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses". IEEE Journal of Robotics and Automation; Vol. RA-3, N° 4, pp 323-344, August 1987.

2. MODELISATION 3D

Dans cette partie, des méthodes de modélisation de volumes englobants d'objets polyédriques et cylindriques, présentant une symétrie, sont présentées en utilisant les capteurs décrits dans le chapitre précédent. Cette modélisation est réalisée interactivement avec la coopération de l'utilisateur qui indique la classe d'objet à modéliser et télécommande les capteurs afin de réaliser le relevé 3D nécessaire au calcul des dimensions et l'attitude spatiale de l'objet.

Trois classes d'objets sont considérées : celle des prismes symétriques (parallélépipède, pyramides tronquées, pyramides,...), celle des cylindres (cylindres, cônes tronqués, cônes) et celle des objets à symétrie de révolution, à axe droit, n'appartenant pas aux deux catégories précédentes et à génératrice non forcément rectiligne.

L'idée originale pour toutes ces méthodes, consiste à utiliser le capteur mixte constitué par la caméra et le DR3D. L'image caméra sert à positionner l'objet dans une pyramide infinie ayant pour sommet le centre optique de la caméra. Le DR3D fournit l'information de profondeur qui permet de situer l'objet dans la pyramide.

Les méthodes développées pour chacune des classes des primitives sont décrites, les résultats obtenus après mise en œuvre sont également présentés. Auparavant, un bref état de l'art est présenté.

2.0 Bref état de l'art

La modélisation géométrique 3D d'objets nécessite l'utilisation d'un dispositif de relevé 3D. Celui-ci peut être constitué par une caméra mobile, de deux voire trois caméras, de télémètre déposé sur un support rotatif ou à triangulation.

Les travaux concernant la vision monoculaire dynamique pour la modélisation d'objet de révolution [DHOME 1990], [SAYD 1996].

Les travaux concernant la vision stéréoscopique voire trinoculaire sont nombreux [FAUGERAS 1993], [AYACHE 1991].

Ceux concernant la modélisation d'images de profondeur le sont moins :

Trassoudaine [TRASSOUDAIN 1993] a proposé des méthodes de relevé 3D télémétrique par balayage de routes. Des méthodes de relevé 3D d'objets en utilisant un télémètre à temps de vol monté sur un dispositif rotatif ont été développé au CEMIF [COLLE 1994].

Badcock et Jarvis [BADCOCK 1994] ont proposé une méthode permettant la modélisation de polyèdres en utilisant un télémètre à triangulation. Les données fournies, par plusieurs points de vue de l'objet à modéliser, sont fusionnées afin d'obtenir les sommets et les arêtes de l'objet.

Neveu et al [NEVEU 1995], Solina et al [SOLINA 1994] ont présenté une modélisation par des superquadriques de la surface d'objets de forme courbe et présentant une symétrie.

Kumar et al [KUMAR 1995] ont travaillé sur la modélisation par des hyperquadriques de la surface d'objets, de forme courbe non forcément symétrique, en utilisant des données télémétriques.

Pour notre part, nous avons développé des méthodes de modélisation de volumes englobants d'objets polyédriques et cylindriques présentant une symétrie, [N'ZI 1995]⁵, [FABRE 1996], en utilisant les capteurs décrits dans le chapitre précédent. Nous utilisons le couple caméra télémètre, car ce capteur mixte présente un bon compromis distance/précision comme l'a montré l'étude du § 1.3.

2.1. Modélisation de prisme symétrique

Pour un prisme symétrique, le relevé 3D de deux faces suffit pour le modéliser. Nous utilisons une face extrême dite principale et une des faces latérales dite secondaire. Les faces choisies sont celles pour lesquelles les mesures télémétriques sont les meilleures et qui sont vues simultanément par la caméra et par le télémètre.

L'image caméra et le modèle géométrique inverse de celle-ci permettent de déterminer les plans tangents (plans d'interprétation) aux contours de chacune des deux faces -un plan d'interprétation est un plan qui passe par une arête de l'objet et le centre optique de la caméra - . Ces plans permettent de définir une pyramide infinie ayant pour sommet le centre optique de la caméra . Le télémètre permet de déterminer l'équation du plan porteur de la face. L'intersection entre la pyramide et le plan porteur de la face permet de définir le contour de celle-ci.

Le volume englobant est ensuite généré en translatant la face principale selon un vecteur ayant une extrémité commune avec la face secondaire.

⁵ Thèse de doctorat de Mr Edie Camille N'ZI co encadrée avec Mr le Pr Florent Chavand

La figure 2.1 représente les plans d'interprétation P_i ($i=1..4$) qui sont déterminés grâce aux images des arêtes AB, BC, CD et DA, correspondantes à une face d'un prisme vu simultanément par la caméra et par le DR3D.

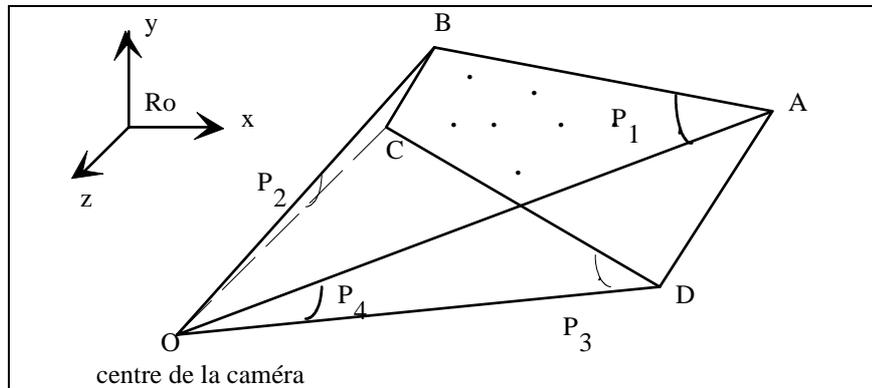


Figure 2.1 : Plans d'interprétation (P_1, P_2, P_3, P_4)

Le DR3D permet de relever un ensemble de points 3D appartenant à la surface à modéliser. La modélisation précédente est réitérée sur la surface secondaire (figure 2.2), qui doit également être vue par la caméra et le DR3D simultanément. Des contraintes type point commun/arête commune (figure 2.2) avec la face préalablement modélisée sont prises en compte pour le calcul de la nouvelle face. Le volume englobant l'objet est calculé par la translation de la face principale selon un vecteur défini par la face secondaire.

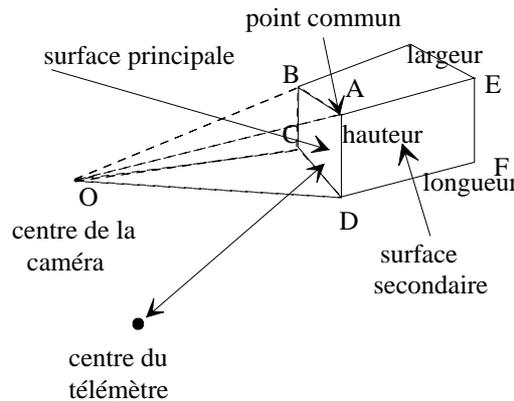


Figure 2.2: Modélisation 3D de polyèdre

Pour une face qui n'est pas parallèle à l'axe des z, le coefficient directeur suivant z est toujours différent de zéro, donc une normalisation est possible.

l'équation du plan devient alors:

$$ax + by + z + d = 0$$

Pour n points, on obtient n équations qui peuvent être mises sous la forme matricielle:

$A \cdot X = B$ avec:

$$A = \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_i & y_i & 1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & 1 \end{bmatrix} \quad X = \begin{bmatrix} a \\ b \\ d \end{bmatrix} \quad B = \begin{bmatrix} -z_1 \\ \vdots \\ -z_i \\ \vdots \\ -z_n \end{bmatrix}$$

Expérimentation et Résultats

Deux expériences sont menées, la 1^{ère} dans laquelle les faces principale et secondaire (fig. 2.2) sont modélisées de manière indépendantes, la 2^{ème} dans laquelle les liens entre les deux faces sont pris en compte.

Les expériences sont basées sur l'utilisation d'un parallélépipède de dimensions : Largeur = 9cm, hauteur=14cm, longueur=9.4cm. L'objet a été modélisé dans différentes configurations dépendant de la distance capteur-objet. Pour chaque configuration, la moyenne de trois modélisations a été prise en compte.

Les résultats de la 1^{ère} expérience (figure 2.3) montrent que les erreurs sont inférieures en profondeur à 5.5‰, en hauteur à 3.5‰, en longueur à 4‰ et en largeur à 7.5‰ (Figure 2.3).

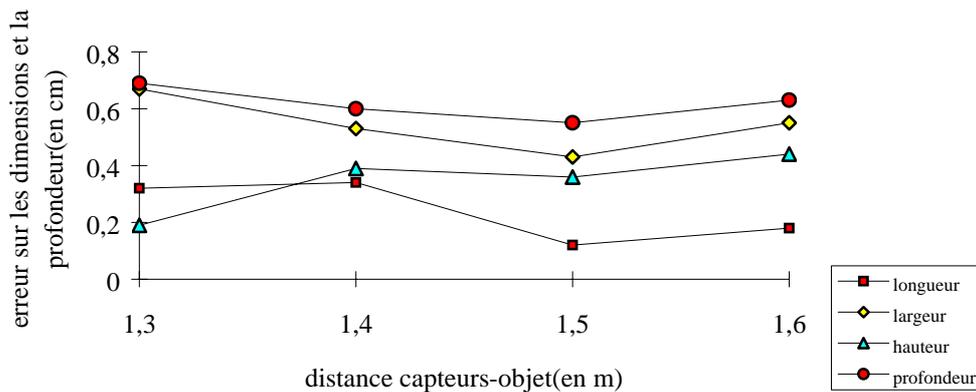


Figure 2.3 : Erreurs de la modélisation contrainte (faces liées) de parallélépipède

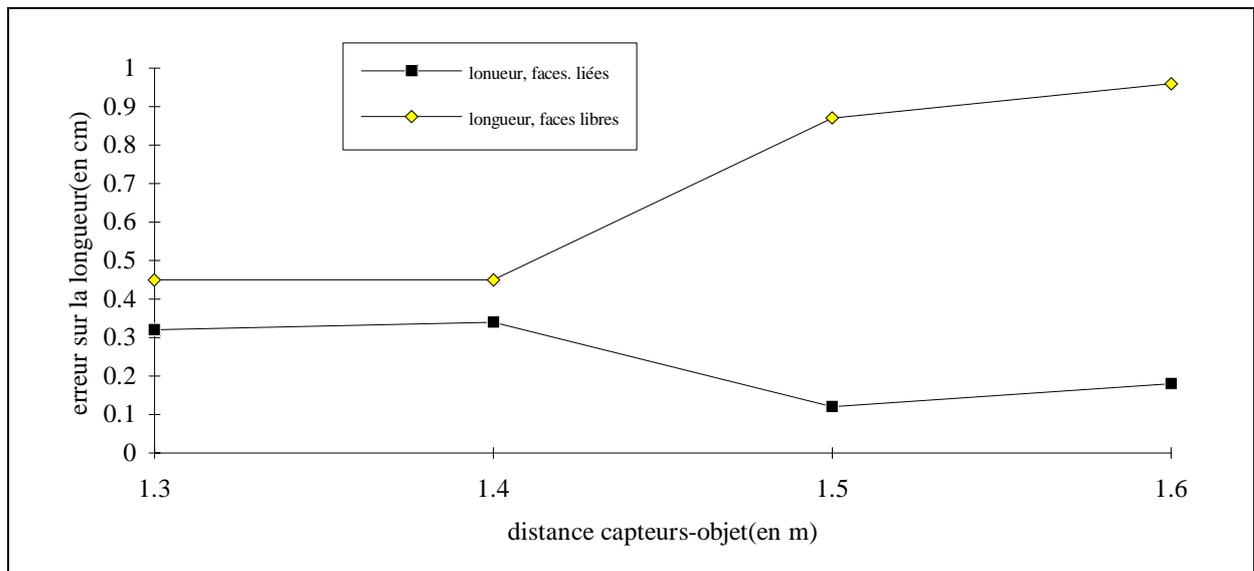


Figure 2.4 : Résultats de modélisations 3D contrainte (faces liées) et non contrainte (faces non liées)

Nous constatons que les résultats sur la longueur sont meilleurs quand la face secondaire est modélisée en tenant compte de la contrainte de liaison (figure 2.4). L'inclinaison de la face à modéliser a une influence non négligeable sur l'information de profondeur donnée par le télémètre quand la distance télémètre objet augmente. Pour cela il est conseillé de commencer la modélisation par la surface principale.

La modélisation de tout objet prismatique symétrique peut s'appuyer sur cette méthode.

2.2. Modélisation de cylindre

Dans cette partie sont présentées différentes méthodes employées pour la détermination des caractéristiques d'une primitive cylindrique observée à l'aide d'une caméra et du DR3D. Pour chaque méthode nous déterminons la position de l'axe de la primitive cylindrique et son rayon.

La définition des différents plans d'interprétation est illustrée (Figure 2.5). La caméra permet de définir l'équation des deux plans d'interprétation P_{c1} et P_{c2} relatifs aux génératrices du cylindre.

En tenant compte de l'attitude du cylindre (vertical, horizontal ou incliné) nous effectuons avec le DR3D deux balayages à deux niveaux différents. A l'aide des points, obtenus après balayage, on détermine l'équation d'un plan frontal P_{tl3} tangent au cylindre. Compte tenu du diamètre du faisceau laser, le profil du cylindre relevé ne correspond pas au profil exact (Figure 2.6). Cependant on peut détecter les extrêmes. Chaque balayage permet d'obtenir les coordonnées des points latéraux et de déterminer les plans P_{tl1} et P_{tl2} .

Pour la définition du plan P_{tl3} , on considère deux droites tangentes au cylindre. Chaque balayage permet de définir une droite. Pour chaque droite, on ne considère qu'un nombre réduit de points (Figure 2.6).

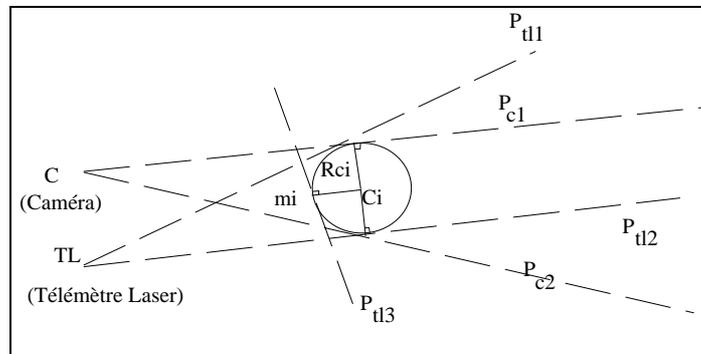


Figure 2.5 : Définition des différents plans d'interprétation

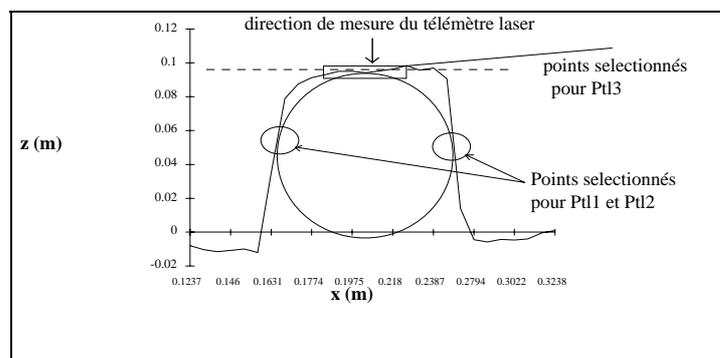


Figure 2.6 : Profil d'un cylindre

Les différents plans d'interprétation étant définis, nous décrivons trois méthodes de modélisation de cylindre ainsi que les résultats des expériences menées.

2.2.1 Coopération caméra/DR3D: plans tangents

Dans cette approche les plans P_{c1} , P_{c2} et P_{tl3} , qui constituent trois plans distincts tangents au cylindre, sont utilisés afin de calculer deux points C_1 et C_2 se trouvant sur l'axe du cylindre et qui sont équidistants à ces trois plans. Les deux altitudes de C_1 et C_2 sont connues et correspondent à celles des balayages télémétriques. Pour chaque balayage, les projections orthogonales de C_j ($j=1,2$) sur les plans P_{c1} , P_{c2} et P_{tl3} sont calculées.

Ces plans sont définis par les équations suivantes:

$$(P_i : a_i x + b_i y + c_i z + d_i = 0, i=1..3)$$

Ces P_i correspondent au plan $Pc1$, $Pc2$ et $Ptl3$

Un premier groupe de trois relations permet d'exprimer que C_j est à la distance R_{c_j} de ces trois plans:

$$\left| \frac{a_i x_{c_j} + b_i y_{c_j} + c_i z_{c_j} + d_i}{\sqrt{a_i^2 + b_i^2 + c_i^2}} \right| = R_{c_j} \text{ avec } i = 1, 2, 3 \text{ et } j=1, 2.$$

les inconnues sont ici $x_{c_j}, z_{c_j}, R_{c_j}$ une des coordonnées peut être fixée a priori: y_{c_j} par exemple (dans le cas d'un balayage horizontal); dans le cas d'un balayage vertical le paramètre fixé est x_{c_j} .

La valeur absolue est levée en calculant l'orientation du vecteur normal au plan P_i . Le signe de la projection de C_j sur $Ptl3$ est l'opposé du signe de la projection du point TL (connu), centre du télémètre, sur $Ptl3$. Le signe de la projection de C_j sur $Pc1$ ou $Pc2$ est le même que le signe de la projection d'un point mi , point donné par le balayage, sur $Pc1$ ou $Pc2$ (figure 2.5).

En posant:

$$x = \begin{pmatrix} x_{c1} \\ z_{c1} \\ R_{c1} \end{pmatrix} \quad A = \begin{pmatrix} a_1 & c_1 & -\sqrt{a_1^2 + b_1^2 + c_1^2} \\ a_2 & c_2 & -\sqrt{a_2^2 + b_2^2 + c_2^2} \\ a_3 & c_3 & -\sqrt{a_3^2 + b_3^2 + c_3^2} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} -b_1 y_{c1} - d_1 \\ -b_2 y_{c1} - d_2 \\ -b_3 y_{c1} - d_3 \end{pmatrix}$$

Le système peut alors s'écrire:

$$AX = B$$

Que l'on résoud alors par une méthode type moindres carrés.

$$X = (A^t A)^{-1} A^t B$$

La solution exprimant X permet de définir les inconnues x_{c1}, z_{c1}, R_{c1} .

On procède de la même manière avec le deuxième point C_2 et on détermine alors les inconnues x_{c2}, z_{c2}, R_{c2} avec ici y_{c2} fixé.

Ainsi les coordonnées des points (x_{c1}, y_{c1}, z_{c1}) et (x_{c2}, y_{c2}, z_{c2}) permettent de déterminer l'équation de l'axe du cylindre.

La hauteur du cylindre est déterminée en utilisant les plans $Pc3$, $Pc4$ et $Ptl3$.

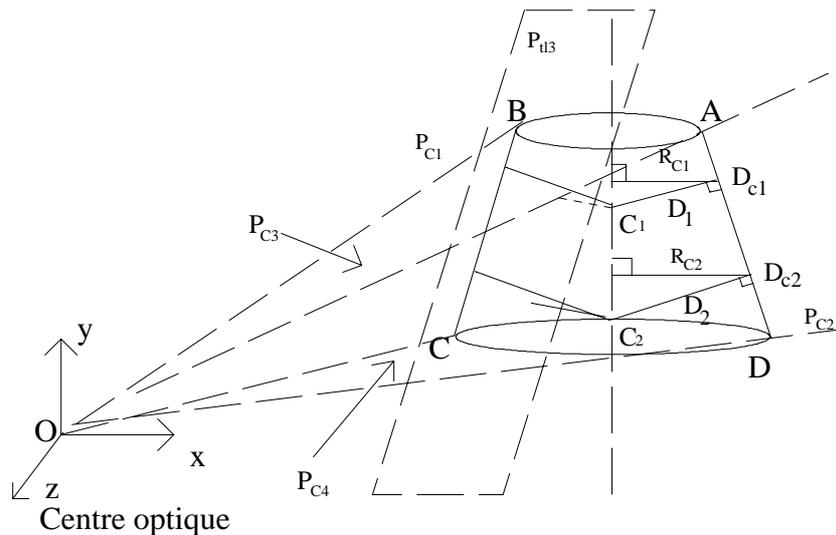


Figure 2.7: Méthode des trois plans tangents sur deux altitudes différentes

Pour cette approche nous utilisons deux plans latéraux tangents au cylindre et définis par la caméra et un plan facial tangent au cylindre et défini par le télémètre.

L'expérience a été réalisée avec un cylindre droit dont les caractéristiques sont les suivantes:

Diamètre : 10.2 cm , hauteur : 16.6 cm

Cette approche permet d'obtenir des erreurs inférieures, sur le rayon à 2%, la hauteur à 8‰ et la profondeur à 5.5‰ sur une distance objet-captteur de 1m à 1.6m (**Figure 2.8**). Le temps de modélisation est inférieur à 20 secondes.

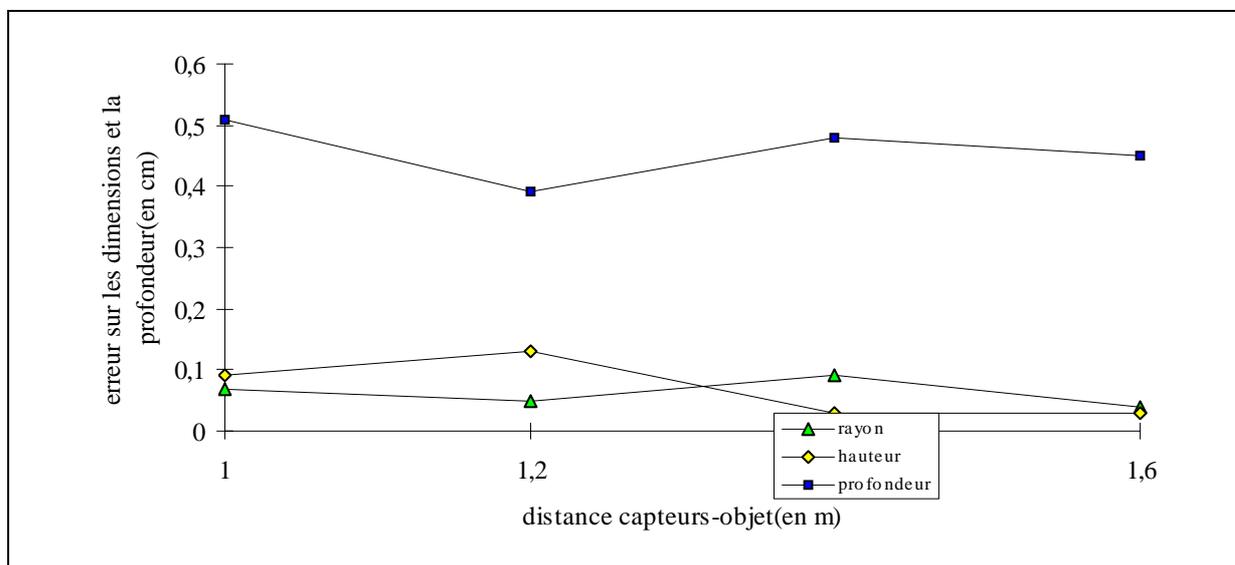


Figure 2.8: Résultats de la méthode des plans tangents (caméra+télémètre)

Ce principe peut aussi être appliqué en utilisant uniquement les données télémétriques: deux plans d'interprétation et un plan tangent issus de deux balayages.

2.2.2. Télémètre seul: plans tangents

Cette approche s'apparente à la précédente, la résolution est identique. Cependant, la détermination de l'axe du cylindre s'appuie désormais sur des plans tangents Pt1, Pt2 et Pt3 donnés par le DR3D (figure 2.5).

Cette approche est identique à la précédente. La détermination de l'axe du cylindre s'appuie sur trois plans qui lui sont tangents. Ces plans sont obtenus à l'aide des mesures données par le télémètre.

Les erreurs sur le rayon et sur la hauteur sont du même ordre de grandeur que pour la méthode précédente (**Figure 2.9**). Par contre les erreurs sur la profondeur sont plus grandes, inférieures à 6.5‰. Le balayage nécessite un temps plus important (2 à 4 minutes) ce qui augmente considérablement le temps pour la modélisation.

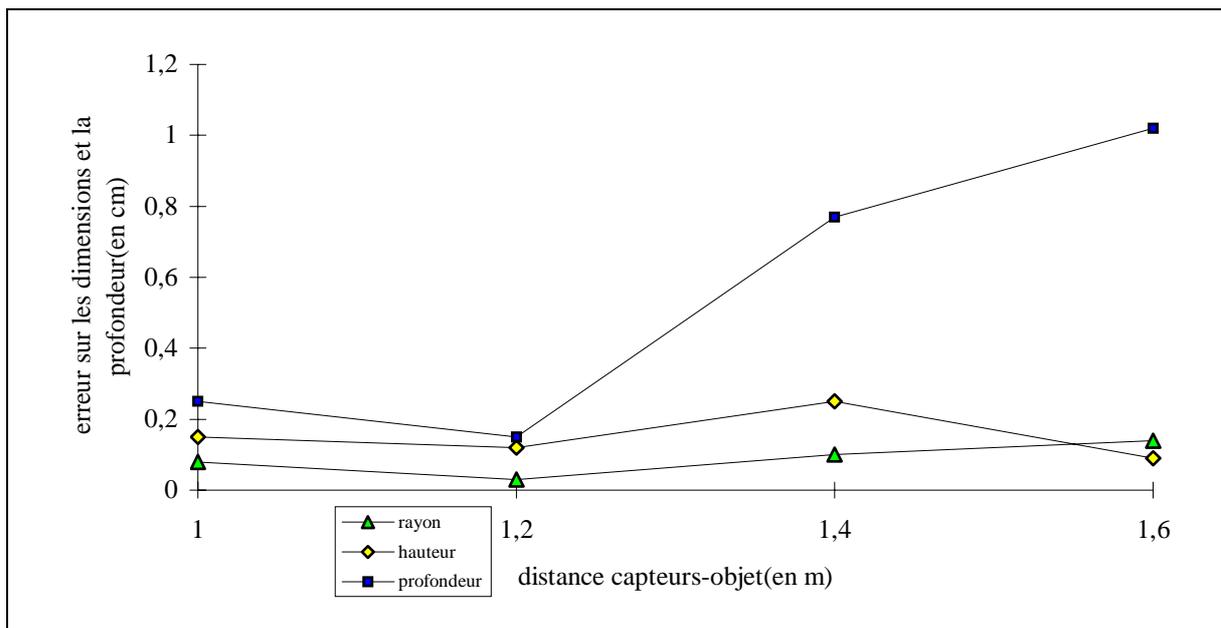


Figure 2.9 : Résultats de la méthode des plans tangents (télémètre)

2.2.3 Coopération caméra télémètre: plans bissecteurs

Les plans d'interprétation, P_{c1} et P_{c2} , fournis par la caméra permettent de définir l'équation du plan bissecteur P_{mc} .

Les plans d'interprétation donnés par le DR3D P_{it1} et P_{it2} permettent de définir le plan bissecteur P_{mit} . La détermination de l'équation de l'axe du cylindre est alors obtenue par l'intersection des deux plans bissecteurs P_{mc} et P_{mit} (Figure 2.10).

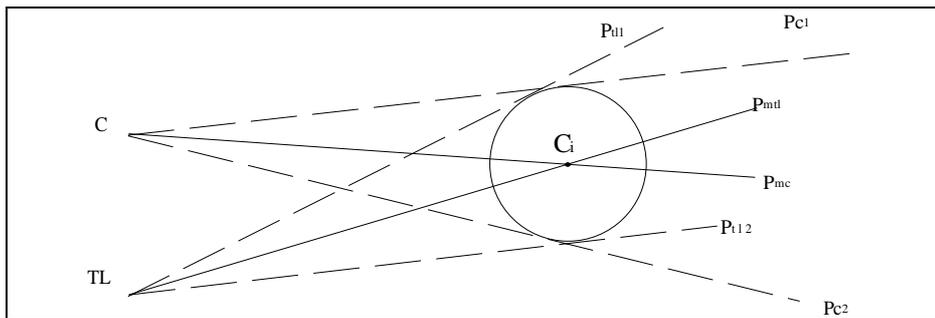


Figure 2.10: Plans bissecteurs de la caméra et du télémètre

Les plans, P_{it1} et P_{it2} étant obtenus par balayage télémétrique, le temps de modélisation est de l'ordre de 2 à 3 mn.

Les erreurs sur le rayon et sur la hauteur sont du même ordre que les deux précédentes méthodes (Figure 2.11). Les erreurs sur la profondeur sont inférieures à 1.5%. Cependant l'erreur en profondeur augmente considérablement quand la distance capteur-objet augmente.

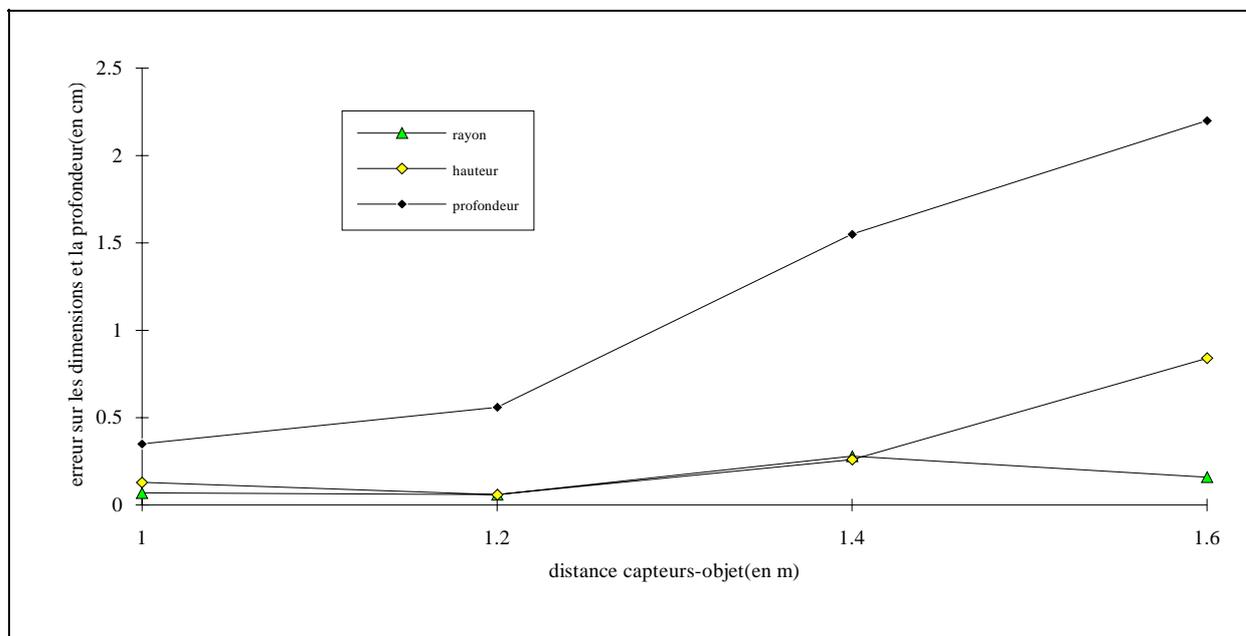


Figure 2.11 : Modélisation 3D par plans bissecteurs (caméra/télémètre)

La **Figure 2.12** représente une récapitulation des erreurs en profondeur observées sur les trois méthodes.

La méthode des plans tangents (caméra + télémètre) est la plus rapide en temps de modélisation (moins de 20s). C'est aussi la plus précise pour des distances de 1.4m à 1.6m. C'est donc celle qui a été retenue.

Pour la méthode basée sur les plans bissecteurs, l'erreur en profondeur augmente rapidement avec la distance capteur-objet. En effet, plus l'objet est éloigné, plus l'angle entre les deux plans bissecteurs est faible. Une erreur sur la détermination d'un plan se traduit par une erreur importante sur la position de l'intersection de ces plans.

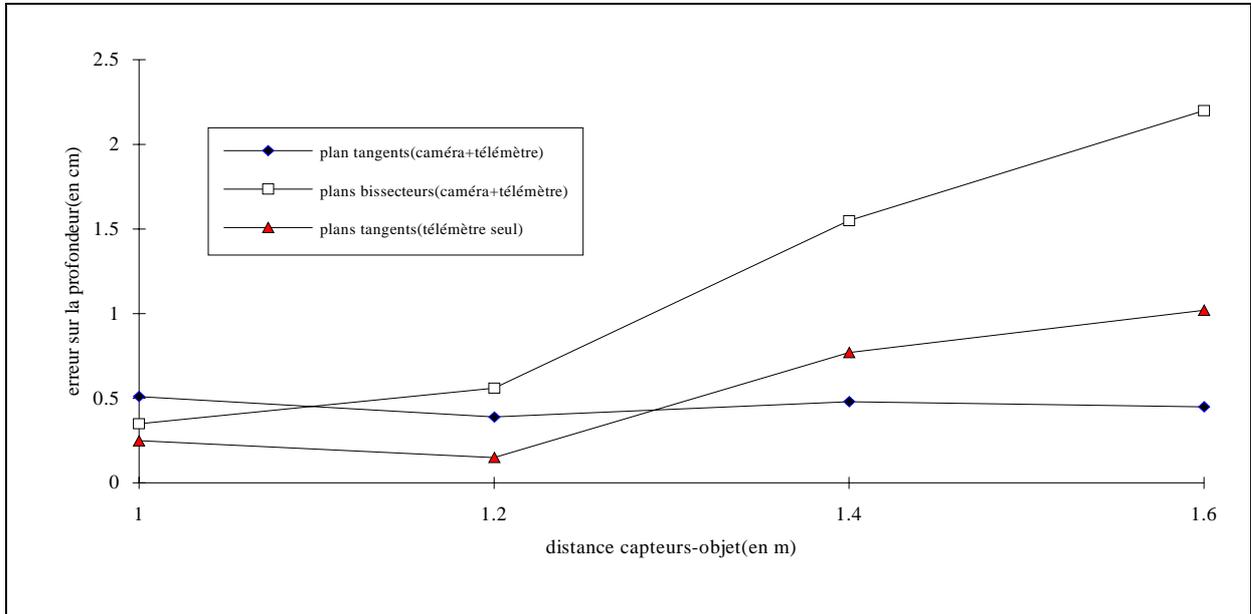


Figure 2.12: Erreur sur la profondeur des différentes méthodes de modélisation 3D

2.3. Modélisation d'objets à symétrie de révolution

La modélisation d'objets à symétrie de révolution est basée sur l'utilisation des courbes B-Spline. Elle est applicable à des objets d'axe droit. Le principe de modélisation consiste à représenter une partie du contour du bord de l'objet appelé contour de base (cf. figure 2.13), par une courbe B-spline, et à effectuer un balayage rotoïde de celui-ci autour de l'axe afin de générer le corps de révolution.

Nous commençons par décrire les données nécessaires à cette modélisation avant d'expliquer la méthode mathématique.

2.3.1 Données de modélisation

Des points caractéristiques (croix sur la figure 2.13) du contour de base de l'objet sont désignés sur l'image caméra de celui-ci. L'axe est connu, il peut être calculé par la méthode décrite dans le § 2.2.1.

2.3.2. Méthode de modélisation

La courbe B-spline est calculée pour le contour de base. La détermination des coordonnées 3D des points du contour de base est possible car ceux-ci sont situés sur la bordure visible de l'objet, et qu'ils sont coplanaires avec l'axe. Aux points obtenus sont alors appliquées des rotations autour de l'axe.

Ci-dessous un exemple de résultat obtenu :

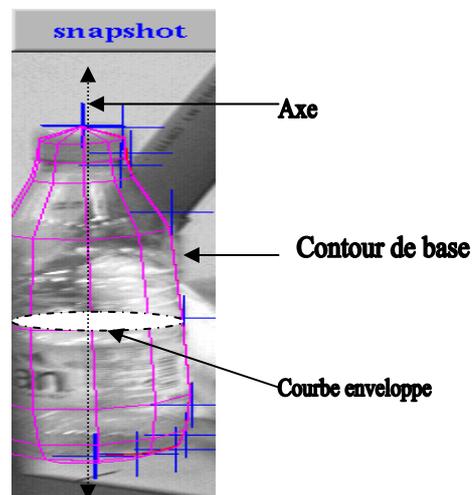


Figure 2. 13 Exemple de modélisation d'une bouteille

Nous présentons la méthode de génération de la courbe représentant la contour de base, ainsi que la méthode de construction du volume englobant l'objet de révolution :

Génération des courbes :

La génération des courbes se fait par la méthode des B-Spline, leur formalisme mathématique est décrit dans l'annexe2.

Génération du volume englobant

Afin de générer le volume englobant l'objet de révolution, il est nécessaire de calculer les coordonnées 3D du contour de base et ceux des courbes enveloppes.

Détermination des coordonnées 3D des points du contour de base :

On sait que les points du contour de base sont situés sur la bordure visible de l'objet, et cet objet est à symétrie de révolution, donc ils sont inclus dans un plan P , qui contient d'une part l'axe de révolution \vec{A} , et d'autre part un vecteur \vec{V} parallèle au plan image de la caméra. Ce 2nd vecteur est défini par :

$$\vec{V} = \vec{I} \times \vec{A}$$

où \vec{I} est la normale au plan image de la caméra, et \times le produit vectoriel.

Ainsi, on connaît le vecteur normal au plan P ($\vec{N} = \vec{A} \times \vec{V}$), on connaît un point de ce plan (une des extrémités du segment formant l'axe), on peut donc calculer les coordonnées 3D des points du contour de base.

Soient

x et y les coordonnées d'un point sur l'image vidéo, correspondant à un point X du contour de base,

\vec{N} : le vecteur normal au plan contenant le contour de base, et \vec{P} un vecteur contenant les coordonnées d'un point appartenant à ce plan.

Le point X appartenant à ce plan, d'où (en notations algébriques) :

$$\vec{N} \cdot \vec{X} = \vec{N} \cdot \vec{P}$$

D'autre part, soient C et D deux points du rayon optique (C : centre optique de la caméra) passant par X , alors :

$$X = C + \lambda \vec{CD} \quad \text{ou } \lambda \text{ est une inconnue et } \vec{CD} \text{ le vecteur directeur du rayon optique.}$$

D'où :

$$\lambda = \frac{\vec{N} \cdot \vec{P} - \vec{N} \cdot \vec{C}}{\vec{N} \cdot \vec{CD}}$$

On obtient alors X .

Détermination des courbes enveloppes:

Pour calculer l'image X' d'un point X , du contour de base, par la rotation d'angle θ autour d'un axe dirigé par un vecteur \vec{A} , on calcule la matrice de passage P du repère de travail à un repère orthonormé dont \vec{A} est l'axe des abscisses, on a alors :

$$X' = \text{rot}_{\vec{A}, \theta}(X) = P^{-1} \cdot R \cdot P \cdot X$$

où R est la rotation d'angle θ autour de \vec{A} .

Cette méthode présente l'avantage de ne nécessiter aucun appariement. En revanche, l'acquisition des points caractéristiques de l'axe de rotation est contraignante. Cette approche de modélisation des corps de révolution est une 1^{ère} ébauche d'un travail qui doit se poursuivre afin de permettre une modélisation plus précise et plus rapide.

2.4 Représentation informatique des données

Le modèle informatique de la scène constitue l'élément indispensable à tout système d'aide à la perception et à la commande. En effet, il permet de traduire les caractéristiques de la scène sous une forme exploitable par le calculateur.

L'environnement de téléopération est constitué de deux types d'objets: les objets connus (robots, outils, pièces à manipuler, ...) et les objets non connus (obstacles,...).

Concernant les objets connus, leur géométrie est bien définie a priori. L'opérateur peut intégrer hors ligne le modèle de ces objets dans la base de données géométriques 3D (BD3D).

Dans le cas où la position de l'objet de ce type n'est pas connue a priori, ou bien s'il est déplacé, l'opérateur utilise les capteurs d'environnement (caméra et/ou DR3D) pour mettre à jour cette base de données.

Pour les objets non connus une coopération entre l'opérateur humain et le système est établie pour construire le modèle informatique de la scène.

Nous présentons brièvement les principales méthodes de modélisation informatique d'objets solides et celle que nous avons retenue. Cette partie est achevée par une brève description de la structure interne des données conçue pour le système MCIT.

2.4.1 Modèles d'objets 3D

Nous définissons le modèle informatique comme une représentation de l'objet comportant des informations de classe (pyramide, cylindre, corps de révolution), de géométrie (coordonnées 3D, équations,...), de topologie(liens entre les éléments géométriques), ...

L'acquisition des données sensorielles permet de prélever l'information du monde réel, par l'intermédiaire de divers capteurs. De ces données brutes (informations de profondeur, de luminance, ...) doit être extraite une description symbolique, de plus ou moins haut niveau, utilisable (point 3D, droite, plan, cylindre,...). Cette opération de description des données perceptuelles est appelée modélisation. Ces données peuvent être organisées sous différentes formes appelées modèles[REQUICHA 1980] (cf. figure 2.14).

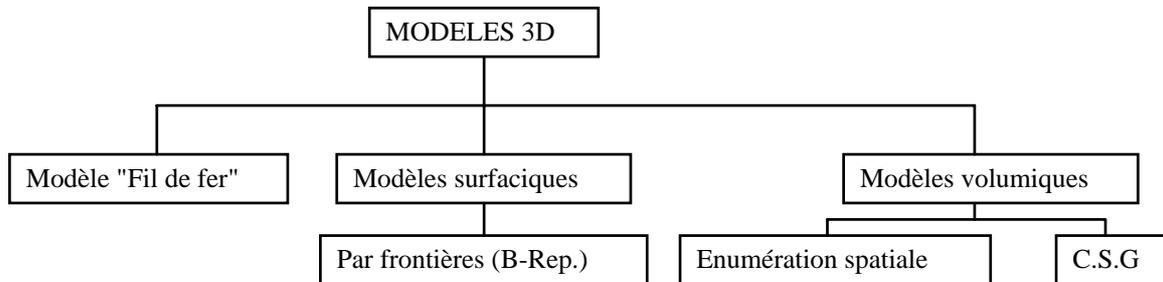


Figure 2.14: Taxinomie des modèles 3D

a. Modèle "fil de fer"

Le modèle "fil de fer" ("wire frame") retient de l'objet les coordonnées des sommets et les arêtes les joignant. La figure 2.15 montre les ambiguïtés qui peuvent découler de son utilisation, puisque ne connaissant que les arêtes et les sommets, plusieurs interprétations d'un même modèle peuvent être faites. D'autre part, les calculs géométriques (dimensions, volumes) sont difficiles, voire impossibles car les informations concernant les surfaces des objets sont inexistantes.

Bien que rarement utilisée, cette méthode de modélisation a l'intérêt de permettre une création et une visualisation rapide de la scène, car elle permet une modification aisée des points et des arêtes. La place mémoire utilisée est également très faible.

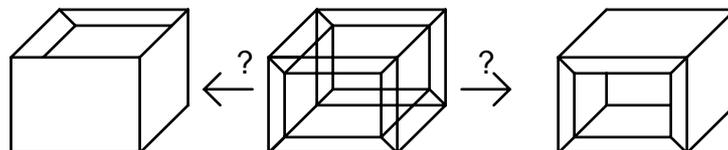


Figure 2.15: Ambiguïté du modèle "fil de fer"

Le terme "fil de fer" est, également utilisé lorsqu'une représentation volumique/surfacique laisse apparaître les arêtes.

b. Modèle surfacique

Ce modèle représente l'objet par ses bords. Le modèle le plus utilisé est la B-REP ("Boundary REPresentation").

Les surfaces représentées peuvent être : planes, surfaces de révolution, surfaces super/hyper quadriques.

La scène est construite interactivement par assemblage de primitives géométriques de base qui sont: point, segment, polygone et polyèdre. Les liaisons entre ces primitives sont définies par des relations topologiques d'incidence, de contiguïté et d'inclusion (cf. figure 2.16).

La principale méthode utilisée pour l'élaboration de polyèdre consiste à balayer un contour de base plan ou courbe le long d'une trajectoire linéaire, circulaire ou quelconque permettant de modéliser le solide correspondant.

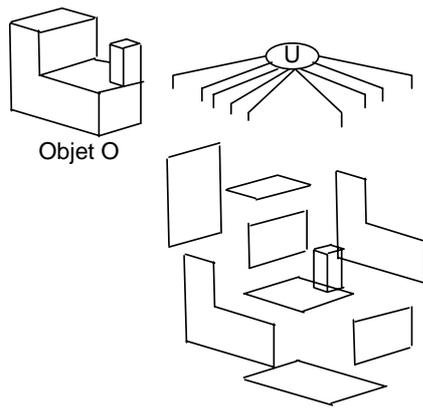


Figure 2.16: La B-REP ("Boundary REPresentation").

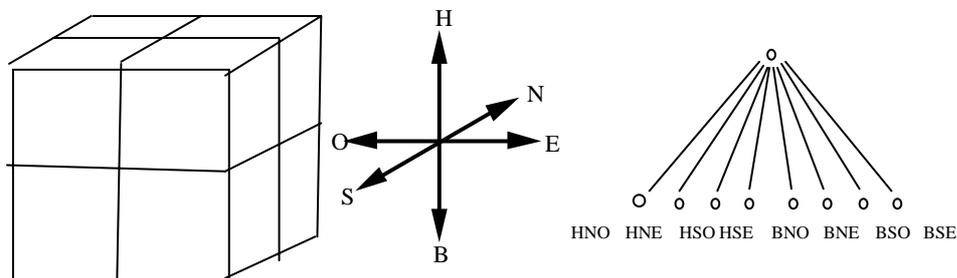
c. Modèle volumique

Dans ce type de modélisation, l'objet est défini par le volume qu'il occupe. On distingue deux types de modèles: l'énumération spatiale et la représentation C.S.G. ("Constructive Solid Geometry").

c.1 Enumération spatiale

Sous ce terme nous plaçons toutes les représentations fondées sur une discrétisation de l'espace en éléments de volumes identiques ou non.

L'espace 3D est discrétisé de manière récursive en cubes élémentaires par exemple. Si un cube n'est pas homogène, il est découpé en huit cubes plus petits (octree) et ainsi de suite. On obtient une description arborescente dont les feuilles sont des "voxels" (cubes homogènes). La figure 2.17 illustre le principe.



Les noeuds sont ordonnés du Haut vers le Bas, du Nord vers le Sud et de l'Ouest vers l'Est

Figure 2.17: Enumération spatiale

L'inconvénient majeur de cette méthode est la place mémoire importante occupée qui est proportionnelle à la finesse de la représentation (taille du voxel). En revanche, l'application de cette méthode à la modélisation d'environnement pour la navigation de robots mobiles est plus plausible.

c.2 Représentation CSG ("Constructive Solid Geometry")

La démarche de cette méthode s'apparente à celle de l'artisan qui usine des pièces élémentaires avant de les assembler pour en obtenir une plus complexe. Elle est décrite par un arbre de construction où les feuilles sont des solides de base (parallélépipède, cylindres,...) et les nœuds sont des opérateurs tels que: union, intersection, différence,... (cf. figure 2.18).

La représentation CSG est très adaptée à la représentation d'objets mécaniques, donc de robots. Elle permet d'avoir une interface conviviale mais nécessite une conversion à un modèle B-REP pour la visualisation. Par

ailleurs, cette représentation n'est pas unique, l'objet O peut être représenté par 2 arbres différents en fonction des volumes élémentaires choisis et des opérateurs ensemblistes qui leurs sont appliqués(cf. figure 2.18).

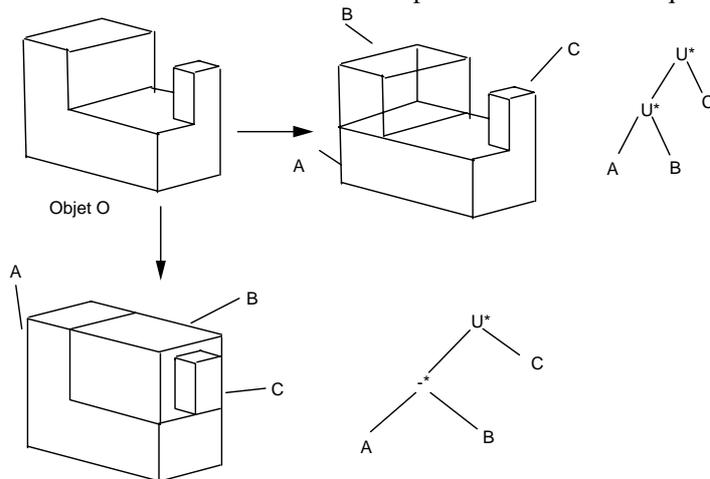


Figure 2.18: Représentation CSG ("Constructive Solid Geometry")

La liste des modèles présentée n'est pas exhaustive. Les plus utilisées sont la CSG et la B-REP. Plusieurs méthodes peuvent être associées les unes aux autres: par exemple la CSG et la B-REP. La représentation CSG permet plus facilement de créer une représentation de la scène que la B-REP. Cette simplicité d'utilisation se fait au détriment de la rapidité d'animation de l'image. En effet, les données canoniques simples ayant servi à la représentation de la scène ne sont pas utilisées directement par le matériel graphique. Une conversion de ces données en points et segments est nécessaire. De plus, le calcul des surfaces très utiles pour les applications de robotique, est difficile avec la CSG, puisque les formules mathématiques de calcul de surface nécessitent la connaissance de points de frontières. Il est certain que la disponibilité des deux méthodes est souhaitable. La CSG permettrait la saisie des données pour la création d'une représentation de la scène et la B-REP l'animation et la visualisation. La solution retenue est celle utilisant la B-REP, elle est plus générale car elle permet de modéliser un grand nombre d'objets et plus proche des systèmes graphiques d'où une animation rapide de l'image synthétique.

2.4.2 La structure interne de données de MCIT

La représentation interne des données est une étape primordiale car tous les algorithmes en dépendent. Celle-ci est construite interactivement par assemblage de primitives géométriques de base (segment, cercle, polygone, ...). Les liaisons entre ces primitives sont des relations topologiques d'incidence, de contiguïté et de contour (modèle B-Rep).

La structure interne des données, du système MCIT, est représentée par un arbre n-aire, les nœuds de cet arbre constituent les informations topologiques qui définissent les relations entre les objets de l'environnement et les relations au sein même d'un objet. Un objet est représenté par les listes de ses faces, arêtes et sommets. L'enregistrement arête est muni d'informations topologiques complètes permettant un accès en temps constant à tous les éléments géométriques voisins[PREPARATA 1985]. Les feuilles de cet arbre constituent les données géométriques cinématiques ou graphiques des objets.

L'arborescence et la hiérarchie de la structure de données et la représentation par frontières permettent de décrire l'environnement avec des niveaux de complexité différents suivant les besoins des applications. L'organisation adéquate des données topologiques permet un accès en temps constant aux données géométriques et cinématiques et optimise ainsi les temps de calcul. Notamment pour la visualisation et pour les méthodes d'appariement 2D/3D décrites dans le chapitre 4.

2.5 Bilan

Les méthodes de modélisation 3D de volumes englobants présentées sont basées sur l'utilisation des capteurs décrits dans le précédent chapitre. Cette modélisation permet d'introduire des objets non préalablement modélisés dans la base de données. Cet ajout en ligne d'objet permet à la tâche effectuée par le robot de se poursuivre. Ces méthodes ont permis également d'évaluer la précision du relevé 3D basé sur l'utilisation de différentes combinaisons de capteurs.

L'erreur sur la profondeur est inférieure au centimètre, pour les distance DR3D-objet inférieure à 1.5m, et les erreurs sur les dimensions de l'objet sont inférieures à 5mm.

Par ailleurs, la connaissance de la géométrie des objets permet de remettre à jour celle-ci si ceux-ci changent de situation . Ceci fait l'objet du chapitre suivant.

Ces travaux ont été publiés dans " Revue d'Automatique et de Productique Appliquées, vol. 8 N°4/1995, HERMES, Paris", [ART 1995].

2.6 Annexe 2

Formalisme mathématique des B-splines :

La courbe B-spline selon $n+1$ points de contrôle P_i est donnée par [FOLEY 1990] :

$$Spline(t) = \sum_{i=0}^n \vec{P}_i N_{i,k}(t)$$

où $N_{i,k}(t)$ est le polynôme de Riesenfeld défini récursivement par :

$$N_{i,1}(t) = \begin{cases} 1 & \text{si } T[i] \leq t \leq T[i+1] \\ 0 & \text{sin on} \end{cases}$$

$$N_{i,j}(t) = \frac{t - T[i]}{T[i+j-1] - T[i]} N_{i,j-1}(t) + \frac{T[i+j] - t}{T[i+j] - T[i+1]} N_{i+1,j-1}(t)$$

Avec comme convention pour le calcul de la formule ci-dessus $0/0 = 0$ et où T est défini par :

$$T[i] = \begin{cases} 0 & \text{si } i < j \\ i - j + 1 & \text{si } j \leq i \leq d \\ d - k + 1 & \text{si } i \leq n \end{cases}$$

2.7 Bibliographie du chapitre 2

- [**AYACHE 1991**] N. AYACHE, F. LUSTMAN, "Trinocular Stereo Vision for Robotics", IEEE Trans. on PAMI,13(1), pp.73-85, January 1991.
- [**BADCOCK 1994**] J.M BADCOCK, R.A. JARVIS, "Wire-frame modeling of polyhedral objects from rangefinder data", Robotica, Vol 12, pp. 65-75, 1994.
- [**COLLE 1994**] E. COLLE , Y. CHEKHAR, "Processus de construction d'images 3D", RAPA, Ed. HERMES, Vol. 7(2), 1994.
- [**DHOME 1990**] DHOME M., LAPRESTE J.T., RIVES G.,and RICHTIN M. "Spatial Localisation of Modelled Objects of Revolution in Monocular Perspective Vision", European Conference on Computer Vision, volume 24, 1990, pp. 475-488.
- [**FABRE 1996**] FABRE R. "Modélisation et Reconstruction 3D d'objets de révolution", stage d'ingénieur, CEMIF-Systèmes Complexes, Université D' EVRY VAL D'ESSONNE, septembre 1996.
- [**FAUGERAS 1993**] O. G. FAUGERAS, "Three Dimensionnal Computer Vision: A Geometric Viewpoint", MIT Press, Boston 1993.
- [**FOLEY 1990**]FOLEY D.F., VAN DAM A. , FEINER S.K. , HUGHES J.F. "Computer Graphics: Principle and Practice" 2nd Edition, Addison-Wesley Publishing Compagny, 1990.
- [**KUMAR 1995**] KUMAR S. , HAN S., GOLDFOF D. and BOWYER K."On Recovering Hyperquadrics from Range Data", IEEE Transaction on PAMI, Vol. 17(11), pp. 1079-1083, 1995.
- [**N'ZI 1995**] N'ZI E.C. "Modélisation et Reconstruction 3D d'Environnement : Application à la Téléopération", Thèse de doctorat de l'université d'Evry Val d'Essonne, Evry, 7/12/1995.
- [**NEVEU 1995**] NEVEU M., FAUDOT D., DERDOURI B. "Superquadriques B-déformables pour la reconstruction 3D", Technique et Sciences Informatiques, Vol. 14(10), pp. 1291-1314, 1995.
- [**PREPARATA 1985**]PREPARATA F.P. and SHAMOS M.I. "Computational Geometry: an Introduction", Springer Verlag 1985.
- [**REQUICHA 1980**]REQUICHA A.A.G."Representations for Rigid Solids: Theory, Methods and Systems", ACM Computing Surveys, vol.12, N°4, Dec. 1980, pp. 437-464.
- [**SAYD 1996**]P. SAYD "Modélisation de cylindres généralisées en vision artificielle", Thèse de doctorat de l'université Blaise Pascal , Clermont Ferrand, 18/12/1996.
- [**SOLINA 1994**], F. SOLINA, A. LEONARDIS, A. MARCEL, "A Direct Part-Level Segmentation of Range Images Using Volumetric Models", IEEE, Robotics and Automation, San Diego, pp. 2254-2259.California, May 8-13, 1994.
- [**TRASSOUDAIN 1993**] Laurent TRASSOUDAIN, "Solution multisensorielle temps reel pour la detection d'obstacles sur route", thèse de doctorat de l'université Blaise Pascal de Clermont-Ferrand, 16 FEV93.

3. RECONSTRUCTION 3D

L'approche de reconstruction 3D utilisée est basée sur la détermination des transformations géométriques rigides qui permettent d'obtenir la nouvelle situation de l'objet dans l'espace. L'obtention de celle-ci permet d'effectuer un recalage de l'objet, c'est à dire la mise en correspondance de ses représentations réelle et virtuelle.

On distingue deux types de recalages qui dépendent de l'ampleur du mouvement :

- recalage à grande amplitude dans le cas de grand mouvement,
- recalage à faible amplitude ou fin pour petit mouvement.

Ces deux types de recalages sont applicables successivement à un objet. Deux cas de figure peuvent se présenter:

❶ Le cas de la création de l'objet, cette création peut se faire de deux manières:

- les dimensions de l'objet sont fournis par l'opérateur, dans ce cas en général, l'écart entre la représentation graphique obtenue de l'objet et sa représentation réelle peut être important.

- les coordonnées de l'objet sont fournies par les capteurs, dans ce cas l'écart entre la représentation graphique obtenue de l'objet et sa représentation réelle est généralement faible. Dans ce cas nous n'appliquons qu'un recalage fin.

❷ Le cas de la modification de l'environnement. L'objet étant modélisé et ses coordonnées 3D intégrées dans la base de données. Il s'agit de retrouver sa nouvelle situation en appliquant le recalage.

Dans le cas d'une aide à la commande (exemple d'une réalisation de tâche), il faut permettre au robot de manipuler correctement les objets présents dans son environnement; pour cela il est indispensable de bien situer ces objets par rapport à un repère de référence. Egalement dans le cas d'une assistance à la perception de la scène en vision indirecte, on superpose au monde réel un monde virtuel grâce à la base de données géométriques 3D des objets de l'environnement. Donc si la situation du monde réel change, les mondes réel et virtuel ne sont plus superposés; il faudra alors procéder à un recalage.

Nous présentons un bref état de l'art contenant les principales références bibliographiques aux travaux antérieurs aux nôtres que nous décrivons par la suite.

3.0 Bref état de l'art

Deux approches sont utilisées pour la reconstruction 3D polyédrique:

- Une approche numérique basée sur l'utilisation d'un nombre important de données tout en cherchant à optimiser un critère de mise en correspondance du modèle de l'objet avec son image. Les méthodes correspondantes sont citées dans [HORAUD 1993].
- Une approche analytique basée sur la résolution d'équations permettant d'obtenir la nouvelle situation de l'objet géométriquement [DHOME 1989].

Ganapathy [GANAPATHY 1984] considère simplement les neuf composantes de la matrice de rotation de manière indépendante. Le système étant linéaire, il est facile à résoudre. Cependant, aucune contrainte d'orthogonalité n'est appliquée à la matrice de rotation, de plus cette méthode est extrêmement sensible au bruit.

Lowe [LOWE 1987] a proposé une méthode itérative fournissant une solution unique en manipulant un triplet de segments de droite. Celle-ci est obtenue en optimisant, par la méthode de Newton, d'un critère d'erreur quadratique. Celui-ci est constitué par la somme des carrés des distances entre chaque projection d'un segment et de son correspondant dans l'image. L'optimisation pose des problèmes de minima locaux et d'initialisation des paramètres à estimer.

Yuan [YUAN 1989] a séparé la rotation de la translation et a proposé de déterminer la rotation en tenant compte des contraintes d'orthogonalité de la matrice de rotation. La solution commune à ces contraintes est obtenue en utilisant la méthode de Newton. L'auteur a noté que des minima locaux peuvent être trouvés avec cette méthode. Le minimum global peut être atteint si on fournit une initialisation proche de la solution.

Dhome [DHOME 1989] a proposé une méthode géométrique basée sur l'utilisation d'un triplet de segments de droite. Cette méthode utilise des repères intermédiaires permettant de simplifier le calcul de la rotation. Des règles sont utilisées pour réduire le nombre de solutions.

Horaud [HORAUD 1993] a proposé une méthode qui minimise un critère basé sur l'utilisation du formalisme N.R. V décrit dans le § 3.1. La rotation est représentée par les quaternions unitaires. L'optimisation

utilise l'algorithme de la région de confiance. Il s'avère que l'algorithme est très rapide quand il converge, mais qu'il faut l'initialiser avec de bonnes conditions initiales pour obtenir le résultat attendu. Cet algorithme semble avoir une bonne robustesse au bruit.

Pour notre part, nous avons utilisé une méthode existante et développé deux autres toutes basées sur la connaissance a priori du modèle 3D l'objet et de son image.

La 1^{ère} est géométrique qui consiste à faire un choix judicieux de repères afin de simplifier les calculs de la transformation rigide. Elle basée sur la méthode de Dhome [DHOME 1989], cependant la résolution du système est réalisée différemment (§ 3.2). Cette méthode peut être utilisée uniquement pour les objets polyédriques.

La 2^{ème} est numérique est basée sur la dérivation du modèle de la caméra. Cette méthode est valable uniquement pour les faibles mouvements de l'objet (§ 3.3).

La 3^{ème} est géométrique, elle est valable pour des objets cylindriques à axe droit.

Ces trois méthodes ont fait l'objet d'une partie du travail de Mr E.C. N'ZI, [N'ZI 1995].

La 4^{ème} est également numérique, elle est basée sur la connaissance des distances entre les sommets d'un objet. Un processus itératif, basé sur la minimisation d'un critère, est appliqué [MOREAU 1995]. De plus, la 3^{ème} méthode numérique peut être utilisée pour des objets de forme quelconque (§ 3.6).

Avant de décrire ces méthodes, nous appelons le principe de la reconstruction 3D basée sur les plans d'interprétation.

3.1. Reconstruction 3D basée sur les plans d'interprétation

L'objectif de la reconstruction polyédrique 3D est de retrouver la rotation et la translation subies par un objet dans l'espace (Figure 3.1). La situation initiale de l'objet représentée par un quadrilatère est exprimée dans (R_0) . Il s'agit de calculer R et T , c'est à dire la situation de $(R_m)/(R_0)$ et donc la nouvelle situation de l'objet. (R_c) est le repère lié à la caméra.

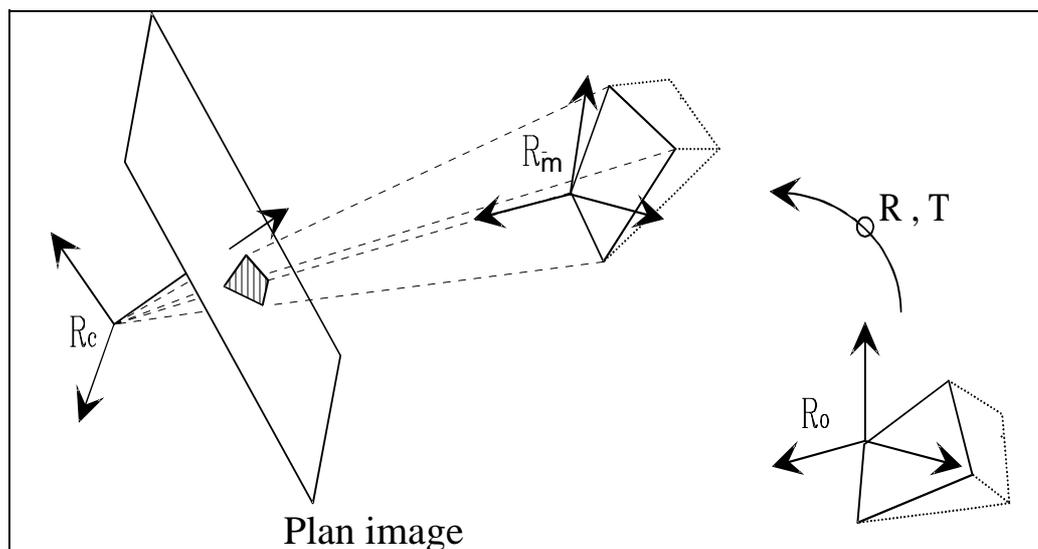


Figure 3.1 : Reconstruction 3D

Toutes les méthodes de localisation tridimensionnelle reposent sur des bases mathématiques communes exploitant des indices visuels 2D provenant de la caméra. Parmi tous les indices visuels, les droites sont facilement extractibles, ce sont de plus des indices robustes. C'est la raison pour laquelle nous les utilisons.

La caméra étant considérée calibrée, on connaît donc la correspondance entre les points dans l'espace et leurs projections sur le plan rétinien. Le vecteur 3D, \vec{V}' (Figure 3.2) donne le vecteur \vec{v}' par projection perspective de centre C sur le plan rétinien. A l'aide des coordonnées 2D, on définit le plan d'interprétation qui passe par le vecteur \vec{v}' et le centre optique C. Ce plan contient également le vecteur \vec{V} de l'espace, \vec{N} étant la normale à ce plan.

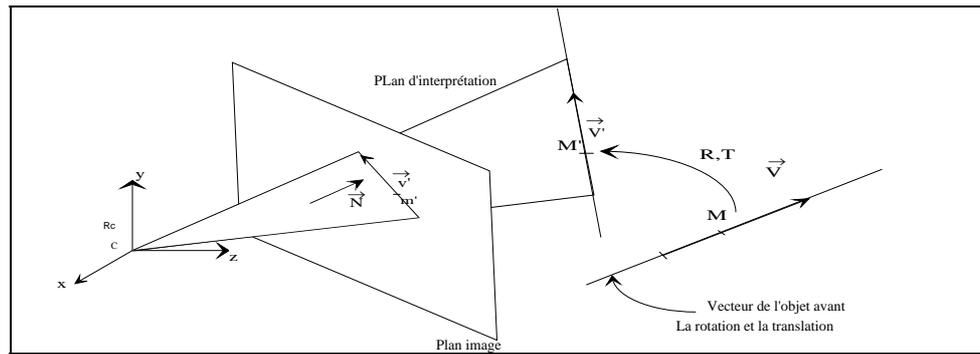


Figure 3.2 : Définition de plan d'interprétation

La localisation fait intervenir trois repères principaux (Figure 3.1)

- ❶ (R_c) repère lié à la caméra.
- ❷ (R_m) repère lié à l'objet représentant la position courante.
- ❸ (R_o) repère de travail, représente (R_m) à l'instant initial.

Localiser l'objet évoluant dans son environnement revient à déterminer la position de (R_m) par rapport à (R_o). Les repères (R_c) et (R_o) sont connus l'un par rapport à l'autre. On supposera par la suite pour simplifier les expressions, mais sans rien enlever à la généralité du problème, que (R_o) est confondu avec (R_c). L'estimation de la transformation amenant (R_m) à (R_o) revient à déterminer une rotation R et une translation T . On peut alors connaître les coordonnées de l'objet dans le repère (R_o) donc dans (R_c). On a ainsi pour tout point M et vecteur \vec{V} de l'objet exprimés dans (R_m), leurs expressions respectives $M_{/R_c}$ et $\vec{V}_{/R_c}$ dans (R_c) :

$$M_{/R_c} = T + RM_{/R_m} \quad (3.1)$$

$$\vec{V}_{/R_c} = R\vec{V}_{/R_m} \quad (3.2)$$

A partir de ces relations, on peut poser les équations de base de la localisation. Pour un vecteur 3D, \vec{V}' , auquel on associe son projeté \vec{v}' et la normale \vec{N} au plan d'interprétation on a :

$$\vec{N} \cdot \vec{V}' = 0 \quad (3.3)$$

$$\vec{N} \cdot \vec{CM}' = 0 \quad (3.4)$$

Les vecteurs présents dans ces deux relations sont exprimés dans une même base. Comme les expressions sont données dans (R_c), repère de centre C ; en tenant compte des équations (3.1) et (3.2) on obtient:

$$\vec{N} \cdot (R\vec{V}) = 0 \quad (3.5)$$

$$\vec{N} \cdot (T + RM) = 0 \quad (3.6)$$

\vec{V} est un vecteur reliant deux points de l'objet et exprimé dans (R_m) (connu),
 M est un point de l'objet exprimé dans (R_m) (connu).

Le problème est donc le suivant: étant donné un ensemble de points 3D décrits dans un référentiel 3D "Objet" et leurs projections 2D décrites dans un référentiel "caméra". Si on connaît les paramètres de la caméra, il faut déterminer la transformation rigide (rotation et translation) entre le référentiel objet à un instant donné et le référentiel caméra.

Les solutions proposées peuvent se regrouper en deux catégories:

- ❶ Une solution numérique, Le système d'équation à résoudre est non linéaire. Ce type d'approche permet de tenir compte des données redondantes obtenues quand le nombre de points de correspondance est grand. Ces méthodes numériques sont basées sur la résolution des équations (3.5) et (3.6) en minimisant un critère d'erreur

qui est, généralement, la somme des carrés des distances entre chaque projection d'un point (ou d'une arête) et de son correspondant(e) dans l'image.

L'inconvénient majeur de cette approche est la lenteur de convergence des algorithmes d'optimisation. Une bibliographie traitant de cette approche peut être trouvée dans [ART 1997], [COM 1997a].

② Une solution analytique, qui est celle que nous allons utiliser et qui est présentée dans le paragraphe suivant.

3.2. Reconstruction polyédrique basée sur les arêtes

La détermination directe de l'orientation de $(Rm)/(Ro)$ est complexe car elle fait intervenir trois angles. Afin de simplifier les calculs nous avons utilisé l'approche présentée dans [DHOMÉ 1989] qui consiste à introduire des repères intermédiaires permettant de ne faire intervenir que deux angles.

3.2.1. Les différents repères

Les différents repères mis en jeu sont (Figure 3.1):

- ① (Ro) = repère de travail (repère objet).
- ② (Rc) = repère de la caméra.
- ③ (Ri) = repère de l'image.

On définit les matrices de passage suivantes:

→ $C_{3 \times 4}$ = modèle global de calibration, matrice de passage du repère (Ro) au repère image (Ri) .

④ (Rm) = repère lié à l'objet et qui subit donc le même mouvement que l'objet. Dans la position initiale de l'objet, le repère (Rm) est le même que le repère (Ro) .

⑤ $(Rm1)$ = repère défini à partir des arêtes 3D de l'objet courant.

⑥ $(Rc1)$ = repère défini à partir des indices visuels 2D de l'objet courant.

On définit les matrices de passage suivantes:

→ $Ro\ c1$ = matrice de passage de (Ro) à $(Rc1)$.

→ $Rm\ m1$ = matrice de passage de (Rm) à $(Rm1)$.

→ R, T = matrices de rotation et de translation qui permettent de passer de (Ro) à (Rm) .

→ $R\ c1\ m1$ = matrice de passage du repère $(Rc1)$ à $(Rm1)$.

La rotation R cherchée est donnée par: $R = Ro\ c1\ Rc1\ m1\ Rm\ m1$

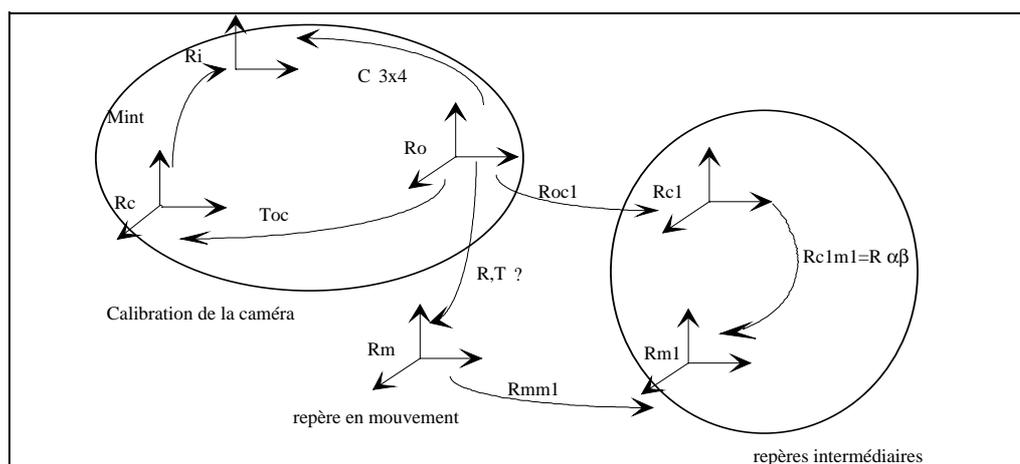


Figure 3.3 : Les différents repères de la reconstruction 3D basée sur les arêtes

3.2.2. Méthode

Soient trois arêtes $L1, L2, L3$ du modèle d'un objet définis dans un système d'axes (Rm) lié au modèle, et trois segments $l1, l2, l3$ détectés dans une image et définis dans le système d'axes (Rc) lié à la caméra.

L'objectif est de déterminer la rotation R et la translation T à appliquer à Li pour i variant de 1 à 3 afin que leurs images coïncident avec celles de li .

Le principe est de résoudre les deux équations (3.5) et (3.6). La rotation contient trois variables (les trois angles de Bryant), donc en principe avec trois vecteurs 3D différents et leurs images, on doit pouvoir trouver R . Une fois la rotation obtenue, on déduit la translation avec également trois vecteurs images distincts et trois points appartenant aux segments 3D. Dans le cas de la translation, les trois segments considérés ne doivent pas être concourants, sinon les trois équations de plans d'interprétation sont liées et représentent seulement deux équations indépendantes. On montrera que cette condition ne concerne pas le calcul de la rotation.

La méthode que nous utilisons s'appuie sur trois droites qui peuvent être construites à partir de 3, 4 ou 6 points comme le montre la (Figure 3.4). Nous utilisons 3 points formant trois segments.

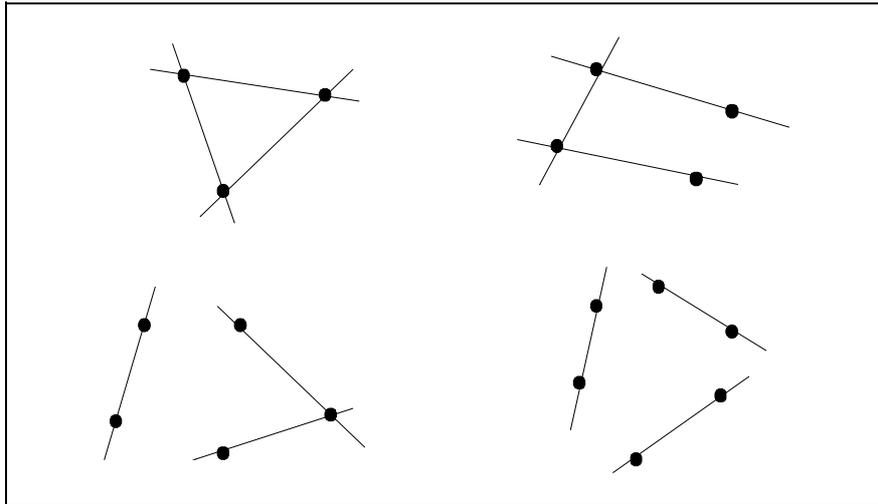


Figure 3.4: Exemple de configuration de trois droites (3, 4, 5 ou 6 points)

Afin de déterminer R , des repères intermédiaires ont été définis. La matrice de rotation R , devient une composition de passage entre ces repères (Figure 3.3).

Les deux repères $Rc1$ et $Rm1$ sont choisis de telle façon qu'ils aient chacun un axe dans le même plan qui est le plan d'interprétation de l'une des arêtes. Ainsi le passage de $Rc1$ à $Rm1$ se fait seulement par deux rotations.

3.2.3. Détermination de la rotation et de la translation

Dans une première étape nous allons décrire, comment ces transformations sont obtenues pour le calcul de la rotation R . Et, dans une seconde étape, nous allons décrire le calcul de la translation T qui est fonction de R .

La détermination des matrices de rotation est basée sur la connaissance de trois segments $L1, L2, L3$ de la position courante de l'objet. Chaque segment est défini par la connaissance de ces deux points extrêmes. Soient $l1, l2, l3$ l'image caméra de ces segments. Soit \vec{V}_i ($i=1$ à 3) le vecteur directeur associé au segment Li ($i=1$ à 3).

Nous présentons, ci-après, les méthodes d'obtention des matrices composant la rotation R .

3.2.3.1. Détermination de $Roc1$, matrice de passage de (Ro) à $(Rc1)$

La détermination du repère $Roc1$ est basée sur l'utilisation de l'image d'une arête de l'objet et du plan d'interprétation correspondant. L'image de l'arête $L1$ est utilisée.

La matrice globale de calibration ($C3 \times 4$) permet de relier les coordonnées (u, v) d'un point aux coordonnées 3D de ce point dans le repère (Ro) (§ 1.1.1.e).

$$\begin{cases} nu_{ij} \cdot P + au_{ij} = 0 \\ nv_{ij} \cdot P + av_{ij} = 0 \end{cases} \quad (3.7)$$

où nu_{ij} et nv_{ij} sont les deux normales associées aux deux plans donnant le rayon visuel passant par le j ème ($j=1$ à 2) point extrême de l'image li du segment Li ($i=1$ à 3). au_{ij} et av_{ij} sont les constantes associées.

$nu_{ij} \times nv_{ij}$ est le vecteur directeur de ce rayon visuel.

L'équation (3.7) représente deux équations de plans dans le repère (Ro) . Le produit vectoriel de la normale des deux plans donne le vecteur directeur du rayon visuel du point.

La désignation du premier point extrême de LI permet de définir le rayon visuel de ce point à l'aide de l'équation (3.7). Ce rayon visuel définit l'axe \vec{x} du repère (RcI) . La désignation du deuxième point extrême de LI donne un second rayon visuel. Le produit vectoriel des vecteurs directeurs de ces deux rayons visuels permet de définir la normale $\vec{N1}$ au plan d'interprétation contenant LI .

$$\vec{N1} = (\vec{nu}_{11} \times \vec{nv}_{11}) \times (\vec{nu}_{12} \times \vec{nv}_{12})$$

Cette normale définit l'axe \vec{y} du repère (RcI) . On définit l'axe \vec{z} de (RcI) par le produit vectoriel de \vec{x} et \vec{y} . La connaissance de $(\vec{x}, \vec{y}, \vec{z})$ permet de définir la matrice de passage de (Ro) à (RcI) .

$$RocI = \left((\vec{nu}_{11} \times \vec{nv}_{11})^*, (\vec{N1})^*, ((\vec{nu}_{11} \times \vec{nv}_{11}) \times \vec{N1})^* \right),$$

(*) vecteur normé

On peut conclure que pour la détermination de $RocI$, il suffit d'utiliser l'image de LI .

3.2.3.2. Rmm1, matrice de passage de (Rm) à (RmI)

On définit (RmI) tel que son axe \vec{x} soit porté par LI , son axe \vec{z} est donné par le produit vectoriel de $\vec{V1}$ et $\vec{V2}$. L'axe \vec{y} est donné par le produit vectoriel de \vec{z} et \vec{x} . La connaissance de $(\vec{x}, \vec{y}, \vec{z})$ permet de définir la matrice de passage de (Rm) à (RmI) .

$$RmmI = \left((\vec{V1})^*, ((\vec{V1} \times \vec{V2}) \times \vec{V1})^*, (\vec{V1} \times \vec{V2})^* \right),$$

(*) vecteur normé

On peut également remarquer que $L1, L2$ exprimés dans (Rm) ont suffi pour déterminer $RmmI$.

3.2.3.3. Rc1m1, matrice de passage de (RcI) à (RmI)

L'axe \vec{x} du repère (RmI) est porté par LI , donc se trouve dans le plan d'interprétation de LI et l'axe \vec{x} du repère (RcI) également dans ce même plan.

L'axe \vec{y} du repère (RcI) est perpendiculaire à ce plan (**Figure 3.5**). Donc la matrice de passage de (RcI) à (RmI) ne dépend plus que de deux rotations, une rotation (\vec{y}, β) pour que les deux axes des \vec{x} soient parallèles et ensuite une rotation (\vec{x}, α) pour que les deux axes \vec{z} soient colinéaires. Une fois ces deux correspondances faites, les deux axes des \vec{y} deviennent automatiquement parallèles. Il reste alors à calculer α et β .

Lorsque $Rc1mI$ est déterminée, on calcule la rotation R cherchée (matrice de passage de (Ro) à (Rm)) par :

$$R = RocI \cdot Rc1mI \cdot RmIm$$

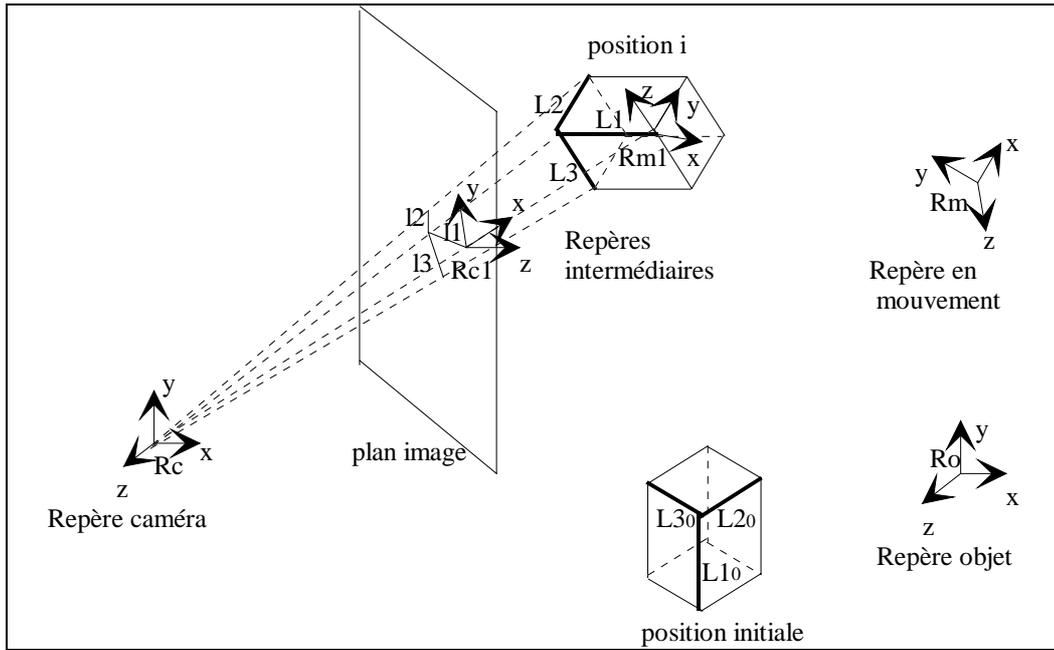


Figure 3.5 : Passage de (R_{c1}) à (R_{m1})

3.2.3.4. Détermination de R

Pour une arête 3D quelconque L_i , auquel on associe son projeté l_i et la normale \vec{N}_i au plan d'interprétation, on a :

$$\vec{N}_i(R_{c1}) \cdot \vec{V}_i(R_{c1}) = 0 \quad (i=1 \text{ à } 3)$$

Cependant on connaît les \vec{N}_i par rapport à (R_o) et les \vec{L}_i par rapport à (R_m) , donc nous allons appliquer les différentes matrices de passage, l'équation précédente devient :

$$[R_{c1o} \vec{N}_i(R_o)] \cdot [R_{c1m1} R_{m1m} \vec{V}_i(R_m)] = 0 \tag{3.8}$$

Remarque :

$$R_{m1m} \vec{V}_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} \quad \text{avec } (a \neq 0) \text{ par définition de } R_{m1m}$$

$$\text{la matrice } R_{c1m1} = R \beta \cdot R \alpha = \begin{pmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Ainsi, on peut remarquer que le vecteur $(R_{m1m} \vec{V}_1)$ est invariant en fonction de α , donc l'équation donnée par $L1$ ne sera pas utilisée pour calculer le couple (α, β) . Ce qui impose une contrainte sur les deux autres arêtes considérés à savoir $L2$ et $L3$ doivent être distinctes, non parallèles à $L1$.

Le développement de l'équation (3.8) aboutit à la résolution d'un polynôme de degré 8 dont la solution est donnée par des méthodes numériques. Nous avons utilisé la méthode de calcul des valeurs propres à partir de la

matrice "Compagnon". On obtient alors les angles α et β qui caractérisent la matrice de passage $RcImI$. La rotation R subie par l'objet est ensuite calculée par la formule :

$$R = RocI RcImI RmIm$$

. Détermination de la translation T

Le calcul du vecteur T est effectué directement dans le repère (Ro) qui est le repère de travail. Cependant on va exprimer les équations dans (Rc) avant de passer à (Ro) . Pour tout point M dans (Rm) , son expression dans (Rc) est donnée par :

$$M_{/Rc} = T + RcmM_{/Rm} \quad (3.9)$$

Comme le vecteur \overrightarrow{CM} se trouve dans un plan d'interprétation :

$$\vec{N}_{/Rc} \cdot \overrightarrow{CM}_{/Rc} = 0 \quad (3.10)$$

Tous les vecteurs présents dans ces expressions sont donnés dans (Rc) , repère de centre C . En remplaçant $M_{/Rc}$ donné par (3.9), l'équation (3.10) devient :

$$\vec{N}_{/Rc} \cdot (T + RcmM_{/Rm}) = 0 \quad (3.11)$$

Par contre si on veut exprimer tous les vecteurs dans (Ro) , étant donné que le point C n'est pas l'origine de (Ro) , il faut tenir compte des coordonnées de C par rapport à (Ro) . Le modèle externe de calibration de la caméra Toc donne les coordonnées du point C par rapport à (Ro) . D'autre part ce point reste fixe une fois la calibration faite et ne subit pas la rotation, l'équation (3.11) devient :

$$\vec{N}_{/Ro} \cdot (T + RomM - C_{/Ro}) = 0$$

où :

T est la translation de (Ro) à (Rm) .

La translation a trois paramètres $(Tabc)$, donc 3 vecteurs suffisent pour les déterminer. On peut prendre plus de 3 vecteurs et utiliser une résolution du type moindres carrés. La méthode utilisant trois arêtes est la plus simple et est celle qui a été implantée. Elle consiste à déterminer trois arêtes afin de trouver trois équations indépendantes dont la résolution analytique permettra d'estimer les trois inconnues a , b et c .

3.3 Reconstruction polyédrique fine

Notre objectif dans ce recalage est d'affiner le recalage à grande amplitude à l'aide d'une méthode de reconstruction appliquée aux petites rotations, pour réduire les erreurs. Nous utilisons le vecteur instantané de rotation pour exprimer la rotation.

Nous commençons par rappeler l'expression du vecteur de rotation. Ensuite, nous décrivons la méthode de calcul de R et T pour le recalage fin.

3.3.1. Vecteur instantané de rotation

Une rotation quelconque peut être représentée par trois angles θ_x , θ_y , θ_z dénommés angles de Bryant. Chacun de ces angles quantifie la rotation autour de l'un des axes d'un repère donné. Une rotation et un repère étant donnés, ces angles ne dépendent que de l'ordre dans lequel s'effectuent les rotations.

L'avantage de ce formalisme est qu'il représente directement la commande de l'articulation du robot (on effectue une rotation de 20° suivant x , puis une rotation de 34° suivant y ... signifie l'envoi direct d'une commande au moteur d'axe x puis à celui d'axe y ...). L'inconvénient est que le produit de ces trois rotations n'est pas

commutatif. L'ordre dans lequel s'effectuent ces rotations conditionne la valeur des angles. De ce fait, la rotation inverse ne s'exprime pas de façon simple. La matrice rotation s'exprime de la façon suivante :

$$R(\theta_x, \theta_y, \theta_z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{pmatrix} \begin{pmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{pmatrix} \begin{pmatrix} \cos(\theta_z) & -\sin(\theta_z) & 0 \\ \sin(\theta_z) & \cos(\theta_z) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Une autre façon d'exprimer une rotation est d'utiliser le vecteur instantané de rotation (Figure 3.6). Une rotation en effet se caractérise par deux entités:

- ❶ une direction autour de laquelle la rotation s'effectue,
- ❷ un angle qui quantifie la rotation dans un plan orthogonal à sa direction. Ainsi, on peut réunir ces deux entités dans un vecteur \vec{R} dont la direction est celle de la rotation et dont la norme est l'angle de cette rotation : $\vec{R} = \theta \cdot \vec{r}$ où \vec{r} est un vecteur unitaire.

L'expression de la matrice de rotation associée au vecteur \vec{R} est la suivante :

$$R(\vec{r}, \theta) = I_{3 \times 3} + \sin(\theta)X(\vec{r}) + (1 - \cos(\theta))X(\vec{r})^2$$

où : $X(\vec{r}) = \begin{bmatrix} 0 & -rz & ry \\ rz & 0 & -rx \\ -ry & rx & 0 \end{bmatrix}$ est la matrice du produit vectoriel, avec $\vec{r} = \begin{bmatrix} rx \\ ry \\ rz \end{bmatrix}$

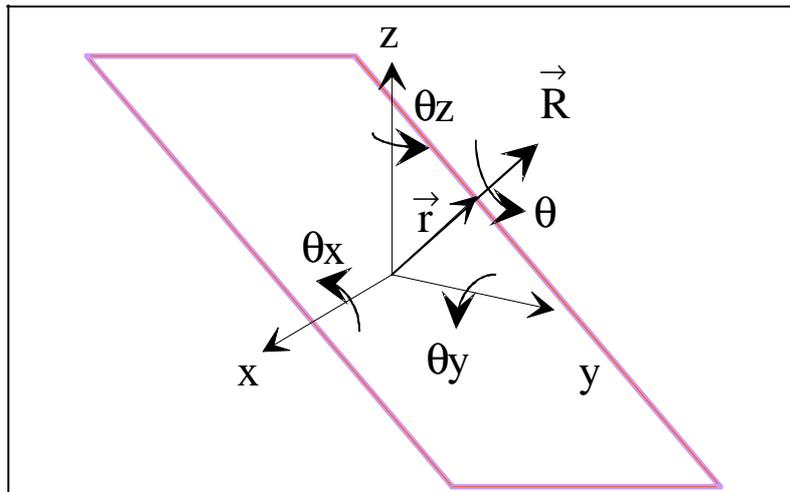


Figure 3.6 : Définition du vecteur instantané de rotation

Cette représentation possède l'avantage de s'appuyer sur la géométrie classique. C'est un outil mathématique très commode grâce à l'interprétation physique que l'on peut en faire : le mouvement d'un solide se décompose en une translation de son centre de gravité et en une rotation autour d'un axe de vecteur directeur \vec{r} .

3.3.2. Détermination de la rotation et de la translation

On considère un objet solide polyédrique dont la position initiale est connue et une image de celui-ci à l'issue d'un petit mouvement. Suite à ces hypothèses, en ne considérant qu'une seule arête de l'objet, on a :

$$\vec{V}' = \vec{V} + \vec{R} \times \vec{V} \quad \text{ou} \quad \vec{V}' = \vec{V} + d\vec{V}, \quad \text{avec } d\vec{V} = \vec{R} \times \vec{V}$$

\vec{V} : un vecteur 3D d'une arête de l'objet dans la position initiale, \vec{V}' le même vecteur après faible mouvement, \vec{R} le vecteur de rotation élémentaire ayant pour direction l'axe de la rotation, pour module son angle θ . L'écriture de cette équation revient à assimiler l'arc issu de la rotation à la tangente de l'angle (Figure 3.7).

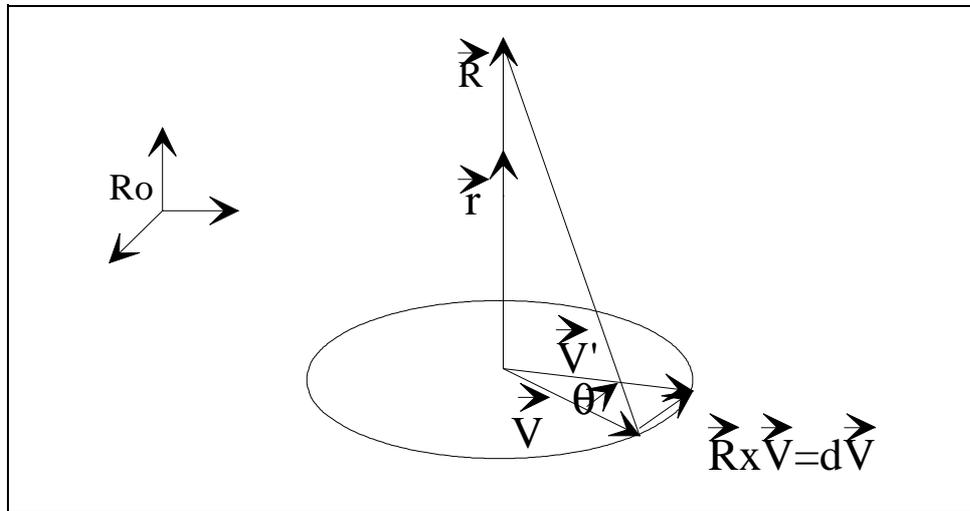


Figure 3.7 : Approximation des petits angles

Pour un point P, ces expressions deviennent:

$$dP = T + \vec{R} \times P \tag{3.12}$$

Comme cela a été vu au chapitre 1 (§ 1.1.1.e) un point P et son image (u,v) dans la caméra sont liés par les équations :

$$\begin{cases} n_u \cdot P + a_u = 0 \\ n_v \cdot P + a_v = 0 \end{cases}$$

qui donnent après dérivation (cas de petits déplacement : en rotation et translation):

$$\begin{cases} dn_u \cdot P + n_u \cdot dP + da_u = 0 \\ dn_v \cdot P + n_v \cdot dP + da_v = 0 \end{cases}$$

dP est donné par l'équation (3.12) et en utilisant une propriété du produit mixte de trois vecteurs i, j, k quelconques: $i \cdot (j \times k) = j \cdot (k \times i)$ ces deux équations sont écrites de la manière suivante :

$$\begin{cases} n_u \cdot T + (P \times n_u) \cdot R + dn_u \cdot P + da_u = 0 \\ n_v \cdot T + (P \times n_v) \cdot R + dn_v \cdot P + da_v = 0 \end{cases} \tag{3.13}$$

L'équation (3.13) peut se mettre sous la forme:

$$AX = B \quad \text{où } X = (a, b, c, Rx, Ry, Rz)^T$$

où $R = \begin{pmatrix} Rx \\ Ry \\ Rz \end{pmatrix}$ et $T = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ sont les inconnus

La résolution de ce système linéaire permet de déterminer R et T:

Un point donne deux équations. Pour obtenir les 6 inconnues un minimum de trois points sont nécessaires. Ces trois points doivent avoir des images distinctes et non alignées. On utilisera la redondance d'informations en prenant $n(n>3)$ points. Le système peut alors être résolu par une méthode de type moindres carrés.

Ce recalage, dit de faible mouvement, permet ainsi de réduire les erreurs obtenues par la méthode de recalage pour les grands mouvements (rotation de: 0 à 2π et translation quelconque). Il peut également être réitéré.

3.4. Reconstruction cylindrique

Trois méthodes de reconstruction de cylindre, à axe droit, ont été développées. Celles-ci supposent que le cylindre est pré modélisé, son rayon et sa hauteur connus(calculés par l'une des méthodes du §2.2).

La 1^{ère} méthode se base sur la connaissance du rayon et du modèle inverse de la caméra pour déterminer l'axe du cylindre.

La 2^{nde} permet d'obtenir la direction de l'axe du cylindre par l'intersection des plans d'interprétation contenant les génératrices du cylindre.

La 3^{ème} est basée sur l'utilisation des plans d'interprétation et exploite la même idée que la méthode décrite dans le §2.2.1.

3.4.1. Première méthode

Soit une primitive cylindrique dont les faces supérieure et inférieure sont perpendiculaires à l'axe (Figure 3.8).

Soient $A1$ et $A2$ les centres des faces supérieure et inférieure, en $A1$ et $A2$ le rayon est respectivement $R1$ et $R2$. Soient li ($i=1,2$) les images des segments (diamètre des faces) Li ($i=1,2$), $a1$ et $a2$ sont respectivement les images des points $A1$ et $A2$.

Soit $D1$ le rayon visuel tangent au cercle de rayon $R1$. Si $d1$ est la distance de $a1$ à $D1$ alors nous avons :

$$\vec{CA1} = \vec{Ca1} \cdot \frac{R1}{d1} \quad \text{où } C \text{ est le centre de la caméra.}$$

C est connu par la calibration de la caméra et $R1$ est connu, car c'est un recalage dont les caractéristiques de la primitive sont connues(modélisée par une méthode du §2.2), $a1$ et $d1$ sont calculés à l'aide des indices visuels 2D, on détermine ainsi $A1$. Par le même principe on obtient $A2$ ce qui permet d'en déduire la situation de l'axe de la primitive cylindrique.

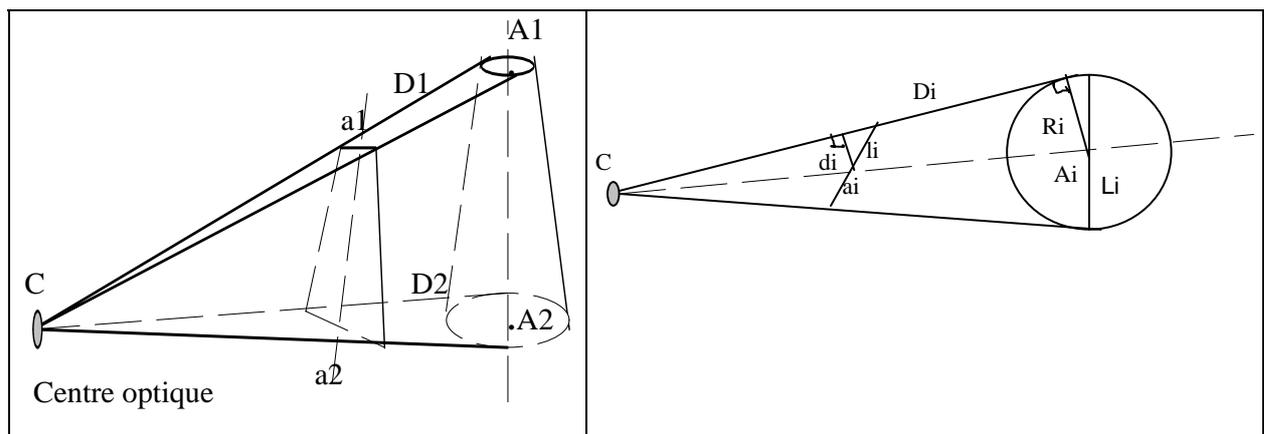


Figure 3.8: Recalage d'une primitive cylindrique

3.4.2. Deuxième méthode

Soit un cylindre, la projection de ses deux génératrices latérales L_1 et L_2 dans le plan image fournissent deux droites l_1 et l_2 . Ces deux droites l_1 et l_2 sont sécantes dans le plan image si les deux génératrices (L_1 et L_2) ne sont pas parallèles au plan image. Leur intersection constitue le point de fuite (p) ou foyer d'expansion. On montre que si C est le centre de la caméra alors la direction Cp est celle de l'axe du cylindre (Figure 3.9).

Soient \vec{v}_1 et \vec{v}_2 les vecteurs directeurs de L_1 et L_2 , P_1 et P_2 les plans d'interprétation passant par L_1 et L_2 et leurs normales associées \vec{n}_1 et \vec{n}_2 . L'intersection entre P_1 et P_2 est définie par la droite Cp . D'après ce que nous venons d'énoncer on peut écrire:

$$\textcircled{1} \vec{v}_1 \perp \vec{n}_1 \text{ et } \vec{v}_2 \perp \vec{n}_2$$

$$\textcircled{2} L_1 // L_2 \Rightarrow \vec{v}_1 = \vec{v}_2 = \vec{v} \Rightarrow \vec{v} \perp \vec{n}_1 \text{ et } \vec{v} \perp \vec{n}_2$$

$$\textcircled{3} \vec{n}_1 \perp \vec{Cp} \text{ et } \vec{n}_2 \perp \vec{Cp} \Rightarrow \vec{v} = \vec{Cp} \quad (\vec{v} = \text{vecteur directeur de l'axe})$$

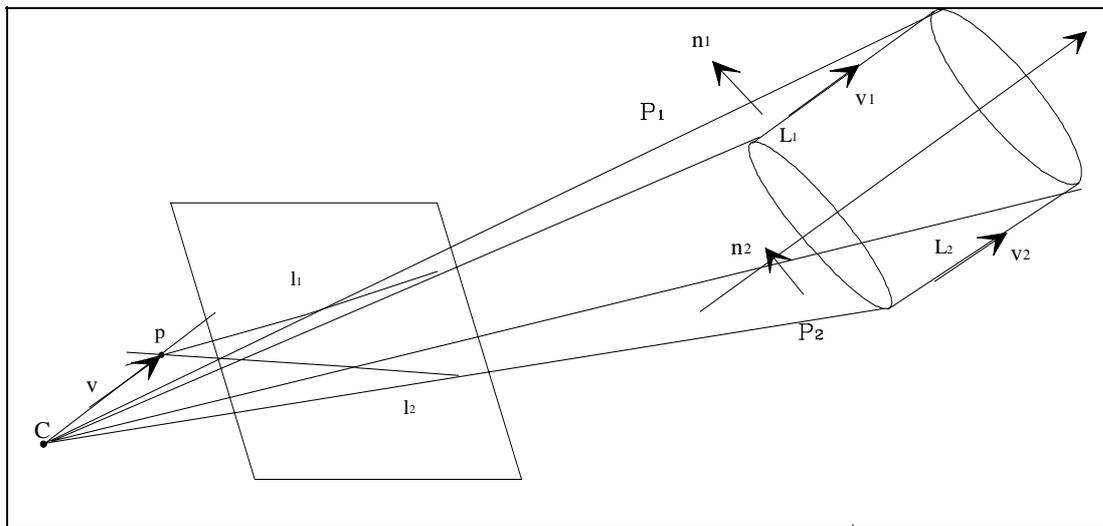


Figure 3.9: Détermination de l'axe du cylindre en utilisant la caméra seule

Cette méthode ne peut s'appliquer qu'à un cylindre et ne renseigne que sur la direction de son axe.

3.4.3. Troisième méthode

Soient $Pc1$, $Pc2$, $Pc3$ et $Pc4$ les plans d'interprétations qui délimitent le cylindre. Comme précédemment, $A1$ et $A2$ sont les centres des faces supérieure et inférieure (Figure 3.10). En $A1$ et $A2$, le rayon est R (rayon du cylindre). Ce rayon est également la distance de A_i ($i=1$ à 2) aux plans d'interprétations $Pc1$ ou $Pc2$. Soient (X_i, Y_i, Z_i) les coordonnées de A_i ($i=1, 2$), la projection de A_i sur un plan Pc_i ($i=1$ ou 2) défini par (Pc_i) : $a_i X_i + b_i Y_i + c_i Z_i + d = 0$ est donnée par:

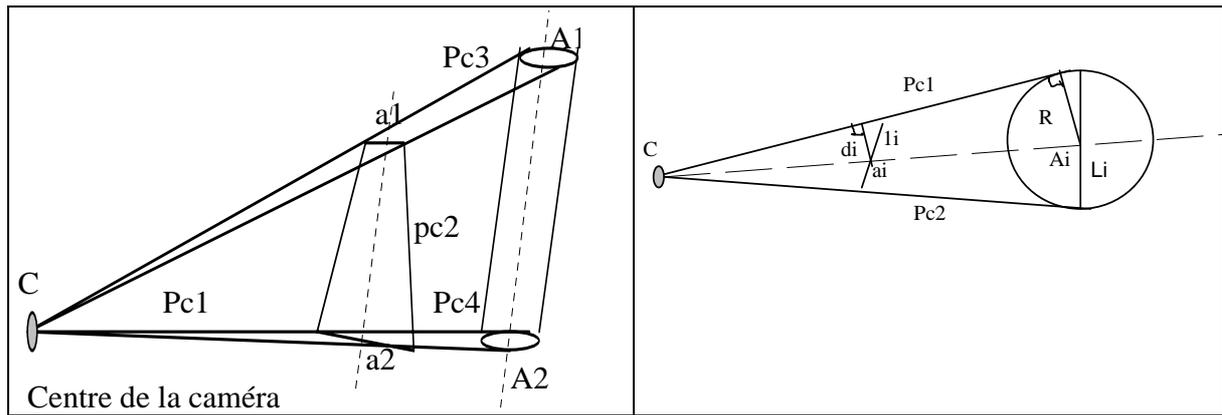


Figure 3.10: Recalage, cas particulier du cylindre

$$\frac{|a_i X_i + b_i Y_i + c_i Z_i + d_i|}{\sqrt{a_i^2 + b_i^2 + c_i^2}} = R$$

Un choix judicieux du vecteur normal au plan P_{ci} , permet d'enlever les valeurs absolues; l'expression étant toujours positive. Il vient:

$$\frac{a_i X_i + b_i Y_i + c_i Z_i + d_i}{\sqrt{a_i^2 + b_i^2 + c_i^2}} = R$$

D'autre part, $A1$ et $A2$ appartiennent respectivement aux plans P_{c3} et P_{c4} , plans d'interprétation passant par les faces supérieure et inférieure; on arrive à un système de trois équations à trois inconnues ce qui permet de déterminer les points $A1$ et $A2$.

3.5 Expérimentation et résultats

Les algorithmes de recalage ont été testés sur des objets réels. Les indices 2D sont donnés directement par désignation sur l'image caméra de la nouvelle position de l'objet. Pour le recalage à grand mouvement, l'erreur sur la rotation est inférieure à 0.6° et sur la translation inférieure à 4cm. La précision de la translation étant liée directement à celle de la rotation. Dans le cas de faible mouvement, l'algorithme de recalage fin donne de bons résultats, l'erreur sur la rotation est inférieure à 0.3° et sur la translation inférieure à 5mm. Ces résultats sont présentés dans la publication [ART 1997].

Dans le cas des objets cylindriques, l'erreur sur la profondeur peut être importante, 2 à 4cm si la distance caméra-objet est supérieure à 1.6m. En deçà de cette distance, l'erreur est inférieure à 1cm. L'erreur en X et Y reste par contre toujours faible (inférieure à 5mm).

3.6 Reconstruction basée sur les sommets

L'objectif de cette méthode est de réduire le nombre d'inconnues du système, en le formalisant différemment, sans utiliser les plans d'interprétation fournis par la calibration de la caméra. Elle est basée sur la connaissance des distances respectives entre certains sommets visibles de l'objet à reconstruire. La correspondance entre les points 3D et 2D est faite manuellement par l'opérateur.

L'idée est de positionner directement dans l'espace 3D les points qui définissent les sommets de l'objet à reconstruire. Ces points sont situés sur une droite connue (le rayon optique associé à ce point) et le principe de la méthode consiste à définir la distance entre le point recherché et un point arbitraire du rayon optique.

Notons qu'à partir des points reconstruits, la transformation rigide qui permet de passer du repère du monde au repère local de l'objet est calculée, la vérification de la reconstruction peut alors être réalisée.

3.6.1 Mise en équation du problème

Etant donné un objet, on note P_i ($i \in E$), l'ensemble de ses sommets. A chaque sommet P_i , on associe un rayon optique de vecteur directeur \vec{V}_i (cf. figure 3.11), passant par le point P_i et un point C_i , choisis arbitrairement.

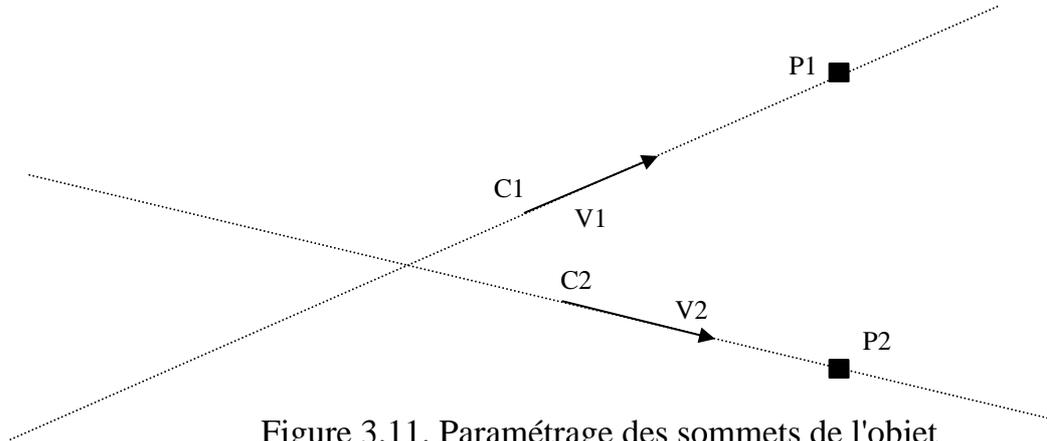


Figure 3.11. Paramétrage des sommets de l'objet

Formulation mathématique :

$$\forall i \in E, P_i = C_i + \rho_i \vec{V}_i \quad (\rho_i \text{ réel non nul})$$

$$\text{soit : } |P_i - P_j|^2 = d_{ij}^2$$

où : d_{ij} est la distance entre 2 sommets visibles de l'objet,

d'où :

$$|C_j - C_i + \rho_j \vec{V}_j - \rho_i \vec{V}_i|^2 = d_{ij}^2$$

$$\Leftrightarrow \rho_i^2 + \rho_j^2 + \alpha_{i,j} \rho_i \rho_j + \beta_{i,j} \rho_i + \gamma_{i,j} \rho_j + \delta_{i,j} = 0$$

où $\alpha_{i,j}$, $\beta_{i,j}$, $\gamma_{i,j}$, $\delta_{i,j}$ sont des coefficients dépendant des C_i et \vec{V}_i

3.6.2 Résolution des équations

La méthode d'optimisation mathématique de Levenberg-Marquardt (cf. § 1.1.2.c) est appliquée pour résoudre le système non linéaire, qui est constitué des quatre (3 points suffisent, nous en utilisons 4 pour une meilleure convergence de l'optimisation) équations relatives aux sommets à reconstruire et des six contraintes sur les distances entre eux, en minimisant le critère suivant :

$$\min_{\rho} C = \sum_{i,j} \left(|C_j - C_i + \rho_j \vec{V}_j - \rho_i \vec{V}_i|^2 - d_{ij}^2 \right)^2$$

En utilisant le modèle du sténopé pour la caméra, les points C_i et C_j sont confondus avec le point C , centre optique de la caméra. Le système à résoudre devient plus simple. Par ailleurs, cette méthode peut être utilisée avec un modèle de caméra à deux plans.

Cette méthode a été testée en utilisant des objets prismatique et cylindrique. Elle a été comparée à la méthode de reconstruction basée sur la connaissance des arêtes. Les résultats, qui se trouvent dans la publication [COM 1997a], montrent que cette méthode basée sur les sommets est plus précise. En revanche, elle est très sensible aux erreurs de mesure.

3.7 Bilan

Les algorithmes de recalage ont été testés .Les indices 2D sont donnés directement par désignation sur l'image caméra de la nouvelle position de l'objet. L'algorithme de reconstruction des objets polyédriques est en deux parties. La première partie concerne le recalage à grand mouvement et est basé sur la méthode de DHOME. La deuxième partie est un algorithme de recalage fin qui utilise une redondance d'informations et qui optimise les résultats du premier recalage. Les méthodes de recalage des primitives cylindriques et reconstruction 3D basée sur les sommets ont également été présentées.

Concernant le recalage à grand mouvement , Dhome travaille avec les modèles interne et externe de la caméra séparément. Il s'appuie sur les paramètres internes et après chaque mouvement de l'objet recalcule le modèle externe (matrice de passage de (Rm) à (Rc)) qui permettrait la superposition mondes réel et virtuel. Or le fait de séparer le modèle externe et interne peut présenter un inconvénient (problème de stabilité) lorsque la méthode de calibration donne directement les paramètres globaux.

L'originalité de notre approche consiste à travailler avec le modèle global cela nous libère des contraintes d'instabilité des modèles interne et externe pris séparément. Nous avons par ailleurs réduit le nombre de repères pour faciliter la compréhension de la méthode. La mise en équation a été effectuée en utilisant un logiciel de calcul formel, dont la syntaxe est compatible avec celle du langage C. Ceci a permis d'intégrer les équations générées par ce logiciel dans les programmes écrits en C.

L'équation obtenue est un polynôme de degré 8. Afin de la résoudre nous avons appliqué une méthode numérique basée sur le calcul de valeurs propres de la matrice "compagnon" construite à partir du polynôme. Dhome a utilisé la méthode itérative de Bairstow.

Afin d'améliorer les résultats obtenues par cette première résolution valable dans le cas général, nous avons appliqué une seconde résolution valable pour les faibles mouvements, ce qui permet de linéariser le problème et d'introduire des informations redondantes.

Les algorithmes de recalage ont été testés sur des objets réels. Les indices 2D sont donnés directement par désignation sur l'image caméra de la nouvelle position de l'objet. Pour le recalage à grand mouvement, l'erreur sur la rotation est inférieure à 0.6° et sur la translation inférieure à 4cm. La précision de la translation étant liée directement à celle de la rotation. Dans le cas de faible mouvement, l'algorithme de recalage fin donne de bons résultats, l'erreur sur la rotation est inférieure à 0.3° et sur la translation inférieure à 5mm.

Le recalage basé sur les sommets utilise un formalisme nouveau qui est plus général car il peut s'appliquer à des objets non polyédriques. L'erreur de reconstruction obtenue est inférieure au centimètre pour une distance caméra-objet inférieure à 1.5m.

Les travaux, décrits dans les paragraphes 3.2 et 3.3, sont publiés dans la revue ROBOTICA [ART97] et dans les conférences : IFIP'96 [COM 1996a], IEEE/IMACS SPRANN'94 [COM 1994a].

Les travaux, décrits dans le paragraphes 3.6, sont publiés dans " IFAC'97, SYROCO, Nantes, France, Sep. 3-5 th, 1997, [COM 1997a].

3.8 Bibliographie du chapitre 3

[DHOME 1989] M. DHOME, M. RICHTIN, J.T. LAPRESTE, G. RIVES, " Determination of the Attitude of 3D Objects from a Single Perspective View", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.11, n°12 pp 1265-1278,December 1989.

[GANAPATHY 1984] GANAPATHY S. (1984): "Decomposition of transformation matrices for robot vision". Proceedings Int. Conf. on Robotics and Automation; Atlanta; pp 130-139; March 13-15,1984.

[HORAUD 1987] R. HORAUD, "New Methods for Matching 3D Objects with Single Perspective Views", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol PAMI-9, n°3,pp 401-412, May 1987

[HORAUD 1993] R. HORAUD, O. MONGA, "Vision par ordinateur, outils fondamentaux", Hermes, Paris, France, 1993

[LOWE 1987] D. LOWE, "Three dimensionnal Object Recognition from Single Two-dimensional Images", Artificial Intelligence, 31, pp 355-395, 1987.

[MOREAU 1995] G. MOREAU "Etude de Méthodes de Reconstruction 3D en Vision Artificielle" rapport de DEA CMPI, Université de Paris XII-Val de Marne, Juin 1995.

[N'ZI 1995] N'ZI E.C. "Modélisation et Reconstruction 3D d'Environnement : Application à la Téléopération", Thèse de doctorat de l'université d'Evry Val d'Essonne, Evry, 7/12/1995.

[YUAN 1989] J.S.C. YUAN, " A General Photogrammetric Method for Determining Object Position and Orientation", IEEE Trans. on Robotics and Automation, 5(2), pp 129-142, April 1989.

4.Appariement d'une image 2D avec son modèle 3D

4.0 Bref état de l'art

Le problème de la mise en correspondance est l'un des problèmes les plus difficiles de la vision par ordinateur. Ce problème peut être traité en 2D/2D, ce qui est le cas de la vision stéréoscopique ou de l'analyse de séquence d'images, ou en 2D/3D basé sur l'utilisation de modèles 3D. Nous nous intéressons à cette dernière catégorie d'appariement. Nous présentons les principales approches de résolution de ce problème.

Grimson et al [GRIMSON 1984], ont proposé une méthode de recherche arborescente. L'arbre est constitué de tous les appariements possibles. Un nœud de l'arbre contient une hypothèse de correspondance entre une primitive 2D et une primitive 3D. Le problème est de rechercher dans l'arbre la suite d'hypothèses qui permettent une mise en correspondance correcte.

Bolles et Cain [BOLLES 1982], procèdent en extrayant en premier dans une image une primitive caractéristique et discriminante telle que coin, trou,...("link focus"). L'appariement est poursuivie avec les primitives 2D voisines à la primitive caractéristique. Cette technique a également été utilisée par Bolles et Horaud [HORAUD 1984], [BOLLES 1986] pour la reconnaissance d'objet 3D à partir de données de profondeur.

Bolles et Fischler [BOLLES 1981], ont proposé une méthode permettant l'extraction aléatoire d'un ensemble de primitives 2D qui sont appariées avec des primitives 3D. L'alignement obtenu (transformation rigide) est généralisé à toute l'image. Le processus est réitéré jusqu'à la convergence de l'algorithme.

Ayache et Faugeras [AYACHE 1986], ont proposé l'utilisation d'un triplet de segments ayant un sommet commun. Cette technique permet de réduire considérablement le nombre d'appariement possible.

Lowe [LOWE 1987] a proposé d'améliorer le traitement d'image, il a établi des critères de regroupement des indices 2D afin d'obtenir des primitives consistantes qui facilitent l'appariement. Par ailleurs, un degré de probabilité est attribué à une hypothèse d'appariement afin de réduire l'espace des solutions et enfin une reconstruction 3D est effectuée afin de vérifier les hypothèses d'appariement.

Lamdan et Wolfson [LAM DAN 1988] ont introduit une méthode géométrique. Celle-ci consiste en la création hors ligne d'une table de hachage et à une reconnaissance en ligne. La table de hachage contient les différents aspects des objets du modèle 3D exprimés dans des bases différentes (une base peut être représentée par 2, 3 points ou arêtes). Cette expression est basée sur des invariants ou des quasi invariants géométriques[BINFORD 1993], [GROS 1995]. La reconnaissance en ligne est obtenu par mise en correspondance de points 2D, exprimés à leur tour dans différentes bases, avec le contenu de la table de hachage.

Jain and Hoffman [JAIN 1988] ont développé une méthode utilisant le raisonnement par évidence pour reconnaître un objet 3D dans une image de profondeur. Pour cela, ils utilisent un système à règles de production. La prémisses d'une règle est composée de conditions sur les différentes primitives de l'objet et la partie action attribue un degré de probabilité sur l'appartenance de l'objet à la base de données.

Plusieurs chercheurs [IKEUCHI 1987], [HANSEN 1989], [PLATINGA 1990], [STEWMAN 1988], représentent un objet 3D par son graphe d'aspects qui donne une description topologique du modèle 3D où chaque aspect représente des faces du modèle visibles simultanément par un observateur. Les avantages et inconvénients de l'utilisation du graphe d'aspects sont décrits dans [BOWYER 1992].

Pampagnin et Devy [PAMPAGNIN 1988] ont introduit le graphe de compatibilité qui regroupe dans ses nœuds toutes les hypothèses d'appariement entre une chaîne de segments 2D et une face d'un aspect du modèle 3D et dont les arcs représentent les compatibilités entre hypothèses d'appariement.

Nous avons développé deux méthodes d'appariement d'objets polyédriques basées sur la connaissance de leur modèle 3D et de leur image caméra 2D:

La 1^{ère} est particulière car elle suppose une connaissance géométrique à priori. Elle utilise la technique de la recherche arborescente et le raisonnement par l'évidence. L'originalité de cette méthode est qu'elle prend en compte les contraintes physiques(liaisons) entre les objets.

La 2^{nde} est plus générale. Elle utilise les méthodes du graphe d'appariement et celle du hachage géométrique. La méthode du graphe d'appariement permet d'apparier des faces du modèle 3D avec des chaînes de segments dans une image. La méthode du hachage géométrique permet de mettre en correspondance des sommets du modèle 3D avec des sommets de segments dans une image. L'originalité réside dans la l'utilisation de l'une de ces méthodes, selon la qualité du traitement de l'image, afin de permettre une reconnaissance de l'objet même si le traitement de l'image ne fournit par suffisamment de chaînes cohérentes de segments.

Nous décrivons dans les paragraphes suivants chacune de ces méthodes.

4.1 Appariement 2D/3D polyédrique basée sur liaisons

La 1^{ère} est basée sur l'utilisation d'un modèle 3D tenant compte des liaisons entre objets polyédriques. Nous nous sommes restreints à la liaison appui plan, car c'est la plus couramment rencontrée, mettant en jeu 3 degrés de liberté entre 2 objets. Les liaisons permettent, si elles ne sont pas rompues dans la scène réelle, de restreindre l'appariement dans une zone privilégiée de l'espace.

La méthode consiste à reconnaître un objet polyédrique parmi plusieurs posés sur un support plan. L'équation de celui-ci est connue (fournie par la calibration de la caméra). Les segments extraits de l'image réelle sont projetés dans ce plan de pose. Les segments obtenus sont appariés avec les arêtes du modèle 3D en utilisant leur longueur et l'angle formé entre leur support et la perpendiculaire au plan image comme critères d'appariement. L'étape de vérification se trouve également simplifiée car elle met en œuvre une reconstruction à 3 degrés de liberté au lieu de 6 [COM 1996b].

4.2 Appariement 2D/3D polyédrique

Cette méthode résout le problème de l'appariement des objets polyédriques en s'appuyant sur trois techniques : l'organisation perceptive des indices visuels pour décrire un modèle 2D de l'image, le graphe d'aspects pour décrire la topologie du modèle 3D et un graphe d'appariement ou une table de hachage pour la construction des hypothèses d'appariement. Ensuite, la transformation rigide entre le modèle 3D et le repère de travail est calculée dans la phase de vérification durant laquelle sont appliquées les méthodes de reconstruction 3D décrites dans le chapitre 3. Les invariants géométriques, basés sur les rapports des surfaces du modèle 3D et de l'image, sont également appliqués pour éliminer les hypothèses invalides. Des tests partiels sur la précision du système de reconnaissance sont effectués.

La figure 4.1 représente la structure générale du système de reconnaissance automatique d'objets polyédriques :

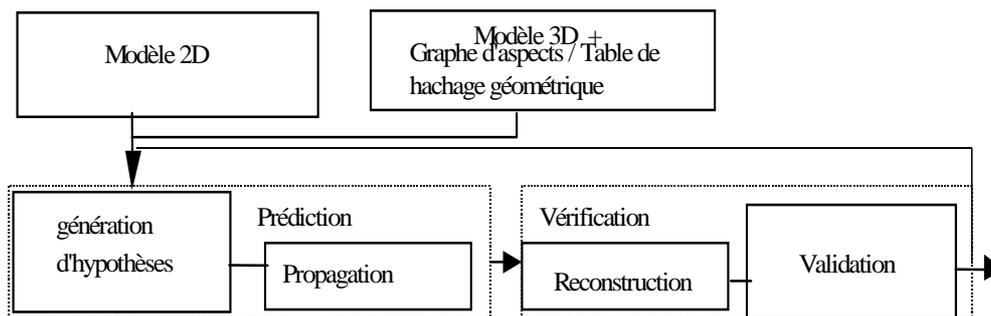


Figure 4.1: Structure générale du système de reconnaissance d'objet.

Nous décrivons brièvement les étapes principales du système de reconnaissance automatique, qui sont :

- le traitement de l'image, à l'issue duquel un modèle 2D de la scène est établi,
- la construction automatique des graphes d'aspects à partir du modèle 3D,
- l'appariement entre modèle 2D et le modèle 3D, qui permet la reconnaissance d'objet.

4.2.1. Traitement et Modélisation de l'image 2D

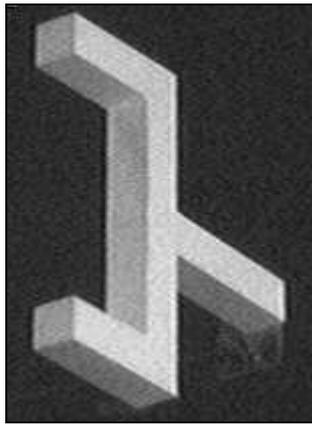
Le traitement d'image appliqué consiste à extraire des primitives 2D (sommet, segment, chaîne de segments) formant le modèle 2D.

Ce traitement est constitué du filtrage du gradient, de l'amincissement et du suivi de contour par hystérésis, de l'extraction des segments et des chaînes de segments. L'extraction de segments est réalisée à partir des points de contours en appliquant la Transformé de Hough (HT).

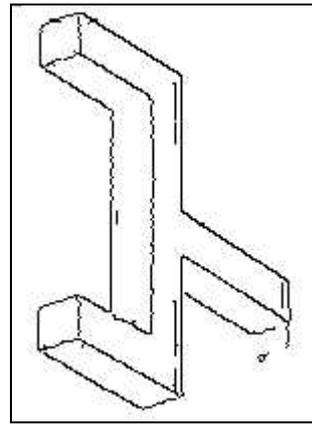
L'objectif est de développer un traitement d'image minimal, pour fournir les éléments nécessaires à l'appariement. L'extraction du contour, se fait alors, par une application successive des filtres classiques :

- le filtre de Sobel dans deux axes, pour obtenir l'amplitude et l'angle du gradient,
- le filtre des maxima locaux, pour amincir le contour,
- le seuillage par hystérésis, pour obtenir un suivi simple du contour avec un bruit minimal.

La Figure 4.2 montre les points de contour obtenus (image binaire), à partir de l'image d'entrée(a).



(a) image de l'objet Pince



(b) le résultat de l'extraction des points de contour

Figure 4.2 : Traitement de l'image de l'objet pince

La transformée de Hough (TH) est appliquée pour la segmentation de l'image. Ce choix vient du fait que la TH est robuste vis à vis du bruit. L'application de la TH classique, sans connaissance de l'angle de contour, est lente car il faut représenter chaque point du contour par une sinusoïde complète dans la matrice de Hough (MH). Chaque point (x, y) du contour peut correspondre à la cellule (a_i, d_i) dans la MH (a_i : angle de contour par rapport à la verticale, d_i : distance du centre de l'image à la droite support du point i), tel que :

$$d_i = x \cdot \cos a_i - y \cdot \sin a_i$$

Nous avons appliqué l'angle de contour, calculé à partir de l'angle du gradient, dans l'équation ci-dessus pour avoir une seule correspondance dans la MH. La segmentation est devenue nettement plus rapide, mais l'erreur sur l'angle de contour a entraîné des segments incomplets. Pour pallier ceci, nous avons développé deux méthodes :

La 1^{ère} consiste à explorer les cellules voisines du segment à extraire. L'inconvénient de cette méthode est la difficulté de fixer un critère d'arrêt.

La 2^{ème} méthode consiste à définir une tolérance sur l'angle du gradient et à explorer uniquement la cellule correspondante à celui-ci. Chaque point du contour est alors représenté par une portion de sinusoïde dans la MH.

La figure 4.3 illustre le fonctionnement de la 2^{ème} méthode, dont l'algorithme itératif est constitué des étapes suivantes :

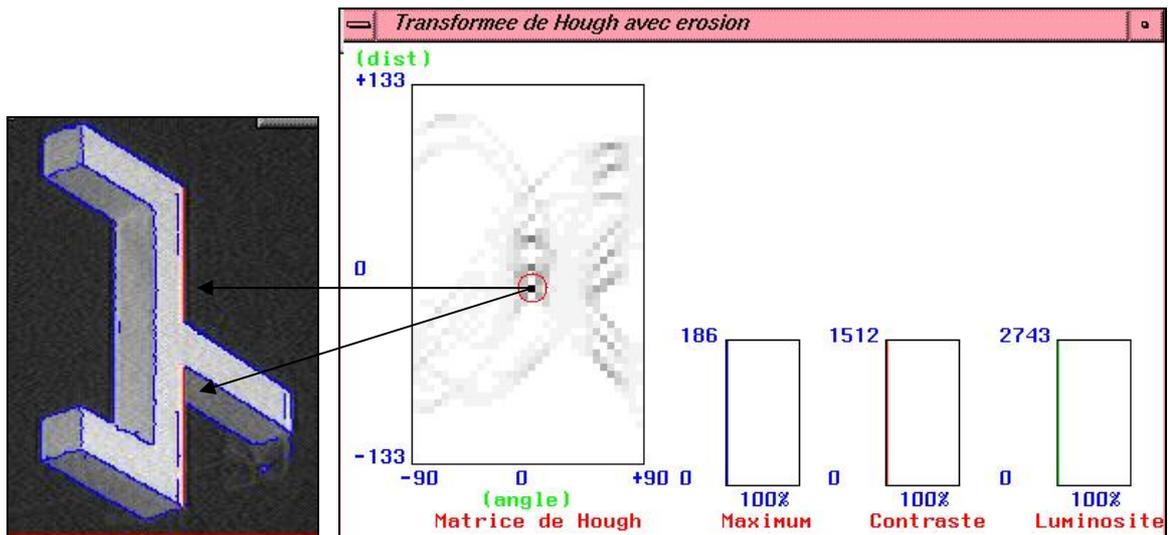
- 1- calcul du vote (nombre de points) pour toutes les cellules de la MH,
- 2- détermination de la cellule à vote maximal : elle représente le segment courant,
- 3- retrait de la MH de tous les points appartenant au segment courant,
- 4- test du critère d'arrêt, s'il n'est pas satisfait retourner à 1.

Trois critères d'arrêt ont été défini empiriquement. Ce sont des critères dynamiques, qui évoluent après chaque itération. Ils représentent des taux exprimés par rapport à l'état initial de la MH.

Le 1^{er} critère, appelé maximum, représente le vote maximal dans une itération donnée.

Le 2nd critère, appelé contraste, la somme des différences avec la moyenne des votes des différentes cellules restantes à explorer.

Le 3^{ème} critère, appelé luminosité, la somme des différences au carré avec la moyenne des votes des différentes cellules restantes à explorer.

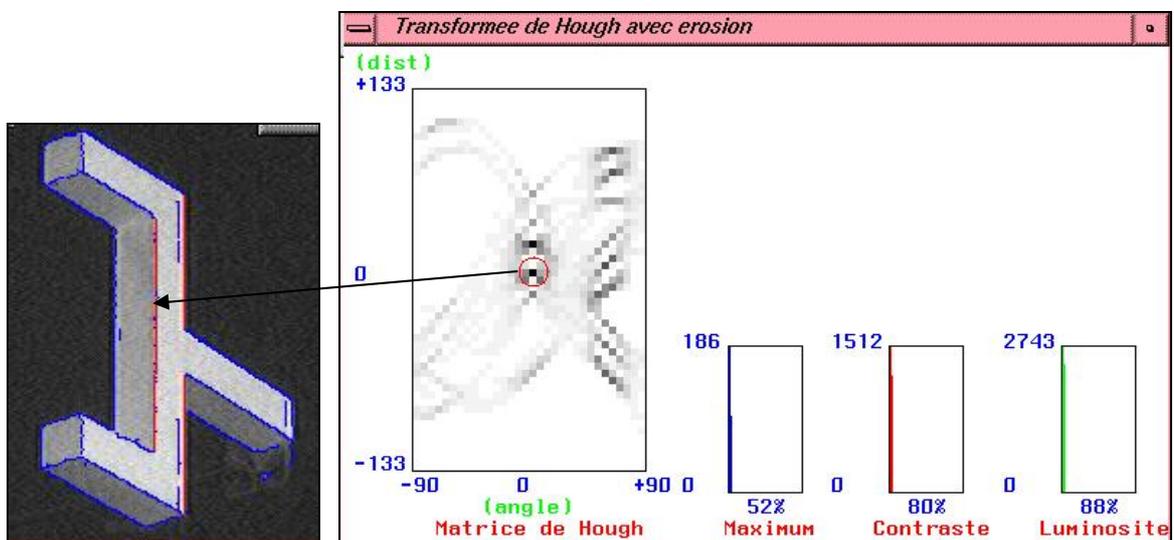


(a) première itération :

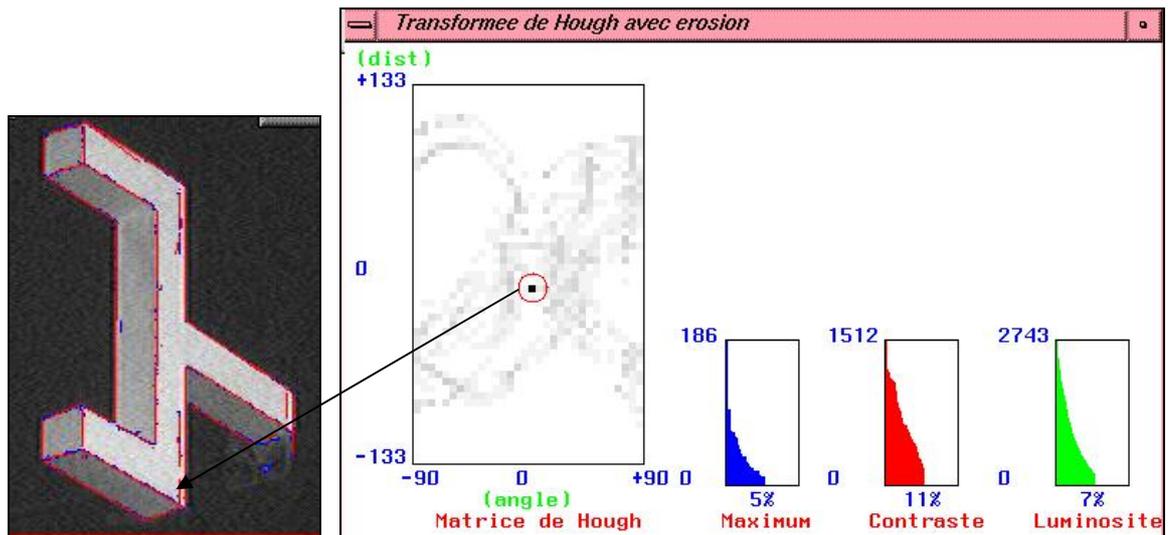
- le vote maximal(cellule de niveau de gris maximum) dans la MH donne le segment le plus long.
- le retrait de la MH de tous les points appartenant au segment courant .

Notes :

- Un point est représenté par ses coordonnées polaires dans la matrice de Hough (l'angle de contour par rapport à la verticale et la distance de la droite support du point au centre de l'image).
- Dans les rectangles représentant les critères(maximum,contraste,luminosité), l'axe des abscisses représente le nombre de segments traités, l'axe des ordonnées représente le nombre de points par segment(taux qui correspond à la longueur du segment en cours de traitement par rapport à celle du 1^{er} segment traité).



(b) seconde itération : le vote maximal s'est déplacé dans la MH, maximum, contraste et luminosité ont diminué.



(c) dernière itération : maximum, contraste et luminosité sont proche de 10% (limite fixée empiriquement pour l'arrêt du programme)

Figure 4.3 : La segmentation par la transformée de Hough améliorée

Pour la fusion de segments, nous utilisons les formules développées par Lowe [LOWE 1987] pour regrouper les segments de droite selon leurs proximité, parallélisme. A partir des informations de proximité entre les segments 2D, les points d'intersection (sommets 2D) sont calculés.

L'étape suivante consiste à parcourir la liste de segments qui se trouvent en proximité pour établir des chaînes qui contiendront certaines descriptions : chaîne ouverte ou fermée, nombre de segments,... Les sommets, segments et chaînes de segments sont organisés dans une structure de données appelée *modèle 2D*.

L'opérateur peut intervenir pour changer ou tester ce modèle (figure 4.4).

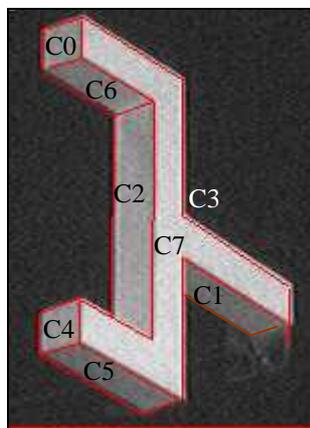


Figure 4.4 : Le modèle 2D de l'image de l'objet Pince

4.2.2. Construction des graphes d'aspects pour polyédre

Le graphe d'aspects donne une description topologique du modèle 3D où chaque aspect représente des éléments géométriques (sommets, arête, face), du modèle 3D, visibles simultanément par un observateur (caméra).

Un graphe d'aspects doit être généré pour chaque objet que l'on veut intégrer dans la base de données des modèles 3D. Deux aspects d'un objet sont différents si l'une des primitives (sommet ou arête) au moins disparaît ou apparaît lorsqu'on passe de l'un à l'autre.

La notion de graphe d'aspects correspond à un ensemble de vues d'un objet topologiquement identiques. Pour chaque aspect de l'objet, il existe un ensemble de points de vues de l'espace pour lesquels l'image de l'objet observé change géométriquement tout en ayant la même topologie. Cet ensemble de points de vue forme une région de l'espace appelé cellule. Lorsque l'œil de l'observateur se déplace d'une cellule à une autre, il voit la topologie de l'image de l'objet changer : des primitives disparaissent et d'autres deviennent visibles. Ce changement est appelé *événement visuel*.

Un graphe d'aspects est la représentation sous forme de graphe de la parcellisation de l'espace en cellules. Il a pour nœuds les aspects et pour arcs les événements visuels.

Deux méthodes de calcul du graphe d'aspects ont été développées. La 1^{ère} est basée sur l'utilisation de la sphère de Gauss et la 2^{ème} sur la parcellisation de l'espace par des plans pour la constitution des différentes cellules.

Nous décrivons brièvement chacune de ces méthodes.

4.2.2.1. Méthode exhaustive

Cette méthode a été développée pour servir à l'appariement basée sur le hachage géométrique (§4.2.3.2.), utilisant les sommets du modèle 3D. Un aspect est alors représenté par un ensemble de sommets visibles simultanément.

On considère une sphère de Gauss centrée sur l'objet que l'on discrétise selon les longitude et latitude (figure 4.5) pour obtenir une série de points repartis sur la surface de cette sphère. Un point P représente ainsi un point de vue adopté pour projeter le modèle.

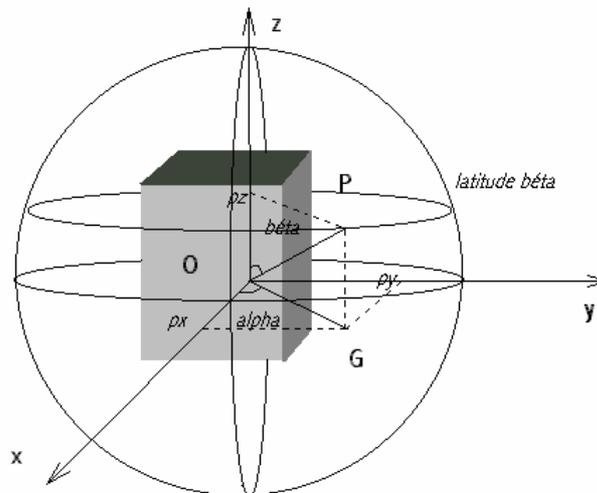
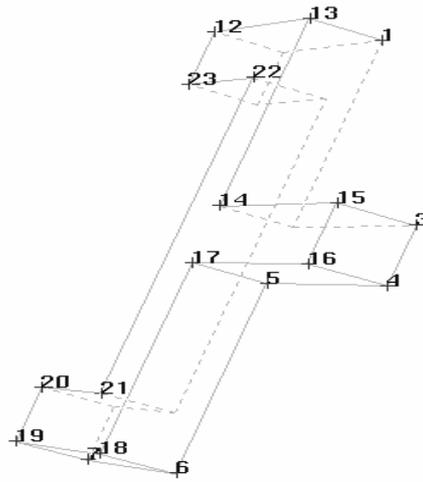


Figure 4.5 : sphère englobant un modèle de type 'boîte'

Pour obtenir un aspect concret du modèle les points cachés sont éliminés lors de la projection (Figure 4.6).



La Figure 4.6 : un aspect de l'objet pince
(tous les segments contenant au moins un point caché sont en pointillé)

Un ensemble d'aspects est ainsi défini, dans lequel il suffit alors d'éliminer la redondance. Deux aspects sont reliés dans le graphe, si et seulement si, les points correspondants sur la sphère sont voisins.

Cette méthode a le mérite d'être rapide et de fournir un graphe correspondant uniquement à ce que l'observateur peut voir. De plus elle s'applique à n'importe quel type de polyèdre immédiatement avec la même complexité algorithmique. Cependant le graphe d'aspects est généralement incomplet et sa précision de celle de la discrétisation et du rayon de la sphère.

4.2.2.2. Méthode analytique

Cette méthode a été développée pour servir à l'appariement basé sur un graphe (§4.2.3.1.). Elle fournit le graphe complet d'un objet polyédrique. Le problème a été abordé en deux étapes de complexité croissante ; il s'agissait tout d'abord de traiter les polyèdres convexes, puis de s'étendre au cas général en incluant les polyèdres concaves. Afin d'obtenir l'ensemble des aspects d'un polyèdre, un graphe d'incidence est créé.

Le graphe d'incidence (GI) est la représentation d'une parcellisation de l'espace. Pour le construire, la liste des plans support des faces de l'objet est calculée. L'intersection entre trois plans, supports de faces de l'objet, constitue le noyau du graphe d'incidence. A chaque itération, un nouveau plan de la liste subdivisera les plans antérieurs. En exploitant toute la liste des plans, le graphe d'incidence est entièrement construit. Chaque cellule du graphe représente une parcelle de l'espace, contenant tous les points de vue à partir desquels on voit toujours le même aspect du polyèdre. Un codage est attribué à la cellule. Celui-ci est représenté par n bits où n est le nombre de faces du polyèdre. L'état d'un bit définit la visibilité d'une face. Une face n'est considérée visible, d'un point de vue donné, que si ce celui-ci est situé du bon côté du plan support de la face, c'est à dire du côté où se trouve la normale à la face (la normale à une face est orientée vers l'extérieur de la matière).

Les feuilles du GI représentent les différents aspects, les nœuds constituent les codages intermédiaires et les arcs les événements visuels représentant le passage d'un aspect à un autre.

La parcellisation de l'espace est complétée pour le cas concave en ajoutant celle obtenue par des plans auxiliaires. Ceux-ci sont issues des événements visuels de type segment/point.

Plantiga & Dyer [PLATINGA 1990] classifient les événements visuels que l'on va rencontrer d'une manière très accessible : événements point/segment ou segment/segment/segment. Sur la figure 4.7 nous constatons comment un simple plan (ici celui en diagonale) peut séparer ce qui est une cellule simple dans un graphe d'incidence en deux parties : l'une où une face est entièrement visible, l'autre où elle est partiellement cachée.

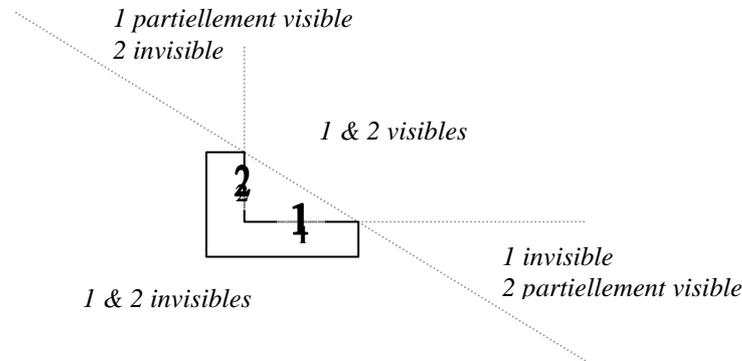


figure 4.7 : exemple d'un plan limite d'un événement visuel

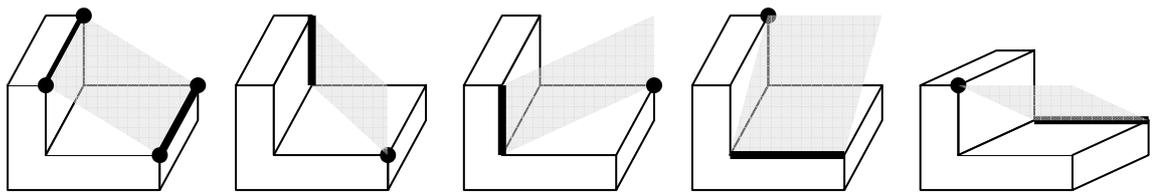


figure 4.8 : les plans intervenant dans les événements segment/point d'un polyèdre concave simple

On parle d'événement visuel segment/point lorsque l'on passe d'un côté à l'autre d'un tel plan. Sur la figure 4.8 les points et les segments impliqués sont en gras.

Cet algorithme est issu des travaux de [STEWMAN 1988] qui utilise celui de EDELSBRUNNER 1986] pour la construction du graphe d'incidence. L'idée générale provient de leur algorithme mais de nombreuses modifications y ont été apportées afin de limiter le nombre de plans auxiliaires pris en compte.

La construction du graphe d'aspects pour un polyèdre convexe fonctionne parfaitement et elle est relativement rapide, les modèles de tests convexes sont traités en quelques secondes. Les modèles concaves pour les tests sont traités en quelques secondes pour les modèles simples (figure 4.11) et en plusieurs minutes pour les modèles complexes.

Pour les polyèdres concaves il faudrait incorporer le traitement des événements du type segment/segment/segment afin de couvrir l'ensemble des aspects..



Figure 4.9 : quelques aspects obtenus de l'objet Pince

Le graphe d'aspects sert à vérifier si une hypothèse est valide du point de vue de la visibilité de ses éléments. La construction du graphe d'aspects de chaque modèle 3D est réalisée hors ligne. L'utilisation du graphe d'aspects, en ligne, permet de simplifier et d'accélérer la reconnaissance d'un polyèdre dans une image.

4.2.3. Mise en correspondance 2D/3D

L'appariement est constitué d'une étape de prédiction et d'une autre de vérification.

La prédiction confronte un aspect(vue) de l'objet, extrait d'un graphe de l'ensemble des aspects de l'objet, à l'image de celui-ci.

L'étape de vérification teste la véracité de cette hypothèse en utilisant les méthodes de recalage 3D. En effet, celles-ci permettent le calcul des transformations rigides (rotation et translation) subies par l'objet en connaissant le contour de l'objet dans l'image et son modèle 3D. Ces transformations rigides, ainsi que le modèle de la caméra, sont appliquées au modèle 3D de l'objet. La projection obtenue est ensuite comparée au contour de l'objet. Deux méthodes de mise en correspondance 2D/3D ont été développées, l'une basée sur le graphe d'appariement, l'autre basée sur le hachage géométrique. L'appariement effectué est applicable aux objets polyédriques.

4.2.3.1. Méthode du graphe d'appariement

La méthode du *graphe d'appariement* (Gapp) permet d'incorporer les contraintes topologiques des objets, grâce au graphe d'aspects. Cela mène à une réduction importante du nombre d'hypothèses possibles sur l'appariement entre le modèle 2D et le modèle 3D. Les nœuds du GApp représentent les hypothèses d'appariement entre une chaîne de segments 2D et une face d'un aspect du modèle 3D. Les nœuds sont directement issues de la topologie des faces 3D et des chaînes de segments 2D. Un arc relie deux nœuds si les 3 critères suivants sont satisfaits:

- Critère de visibilité : les deux faces 3D apparaissent dans l'un des aspects du modèle 3D (la visibilité des deux faces est possible).
- Critère de connexité : les deux chaînes ont la même relation de connexité que les deux faces (sommet en commun, segment en commun ou rien).
- Critère d'invariance : le test des invariants géométriques, basés sur des rapports de surfaces entre faces et chaînes fermées de segments, est satisfait[ZHU 1995].

Nous cherchons, ensuite, des sous graphes complets du Gapp, appelés *cliques maximales*. Une clique maximale doit, aussi, vérifier les critères de visibilité et de connexité :

- toutes les faces 3D de la clique doivent apparaître dans un aspect,
- tous les appariements, entre les segments 2D et les arêtes 3D de la clique, doivent être cohérents,
- l'erreur issue de l'application des invariants géométriques ne dépasse pas un seuil fixé.

Les cliques maximales forment les hypothèses plausibles sur l'appariement. Elles sont munies d'un score, basé sur le nombre d'appariement sommet3D/point2D obtenu.

Le Gapp est appliqué sur le modèle 2D et chacun des modèles 3D sélectionnés. Les hypothèses issues de tous les Gapps sont ordonnées selon leur score. Une méthode de prédiction/vérification classique est appliquée, aux hypothèses une par une. Si une hypothèse réussit, ses éléments sont retirés du modèle 2D de l'image. Nous continuons cette opération jusqu'à l'exploration de tout le modèle 2D ou de toutes les hypothèses. La transformation rigide, entre le modèle 3D et le repère de l'affichage, est calculée dans la phase de vérification, pour chaque hypothèse réussie. Enfin, tous les modèles reconnus sont superposés sur l'image.

Cette méthode est tributaire de la qualité de l'image caméra. Si celle-ci est dégradée, le résultat du traitement de l'image s'avère insuffisant - difficulté d'extraction des chaînes de segments - pour appliquer la méthode du graphe d'appariement. D'où l'utilisation de la méthode du hachage géométrique basée sur les sommets des objets.

4.2.3.2. Méthode du hachage géométrique

De la même manière que la précédente méthode, un ensemble d'aspects d'un objet est calculé hors ligne et l'appariement est effectué en ligne. Les aspects sont calculés par la méthode exhaustive (§4.2.2.1.).

Dans notre implémentation, un aspect est représenté uniquement par les sommets visibles de l'objet du point de vue considéré. Une base de transformation dans le plan de l'aspect est définie, elle peut être formée de

deux ou de trois points de l'aspect. Un nombre défini de ces bases, dans lesquelles nous projetons les aspects, est utilisé. Après avoir projeté les points dans chaque base, nous les stockons dans une structure appelée *table de hachage*. Celle-ci est un tableau à deux dimensions, de taille fixe, dont chaque entrée est une coordonnée (les deux dimensions correspondent donc à l'abscisse et à l'ordonnée d'un point). Les dimensions de la table de hachage étant fixes, nous sommes obligés d'approximer les coordonnées des points projetés. Un élément de la table de hachage est donc un point exprimé dans un repère absolu normalisé. Il sera modélisé par une liste dont chaque élément est un triplet : le modèle auquel il appartient, un aspect de celui-ci et une base dans laquelle l'aspect est exprimé.

L'appariement en ligne est effectué en confrontant une série de points, extraite de la scène réelle, avec le contenu de la table de hachage. Ces points, qui ont une valeur expérimentale, sont pour la plupart des sommets de l'objet à reconnaître. Ils sont ensuite exprimés dans différentes bases, ces bases étant construites à l'aide des points reconnus de l'image. En insérant ces points dans la table de hachage, nous obtenons une série de correspondances (triplets : modèle/aspect/base) avec un ou plusieurs aspects préalablement enregistrés lors de la phase hors ligne de préparation de la table de hachage. Pour chaque correspondance obtenue, nous ajoutons un vote pour un triplet modèle/aspect/base, dans une structure réservée à cet effet. Nous définissons ensuite un nombre de votes minimal à obtenir, afin qu'un triplet modèle/aspect/base soit considéré comme une hypothèse plausible. C'est ici que les propriétés d'invariances interviennent. En effet, quelque soit la position ou la taille de l'objet, il sera reconnu. Car, un ou plusieurs aspects, qui lui sont topologiquement semblables, ont été préalablement introduits dans la table de hachage. Il suffira alors d'appliquer une série de transformations affines sur un aspect reconnu, afin de réaliser l'appariement avec l'objet représenté dans la scène. Les hypothèses données par la table de hachage, sont toutes testées par une procédure de prédiction/vérification.

Des tests expérimentaux ont montré que les bases à deux points sont plus performantes que les bases à trois. La figure 4.10b montre que les bases à deux points rendent la table de hachage mieux équilibrée. Le niveau de gris d'une cellule représente la taille de sa liste. Le noir représente une liste vide, tandis que le blanc est une liste à taille maximale.

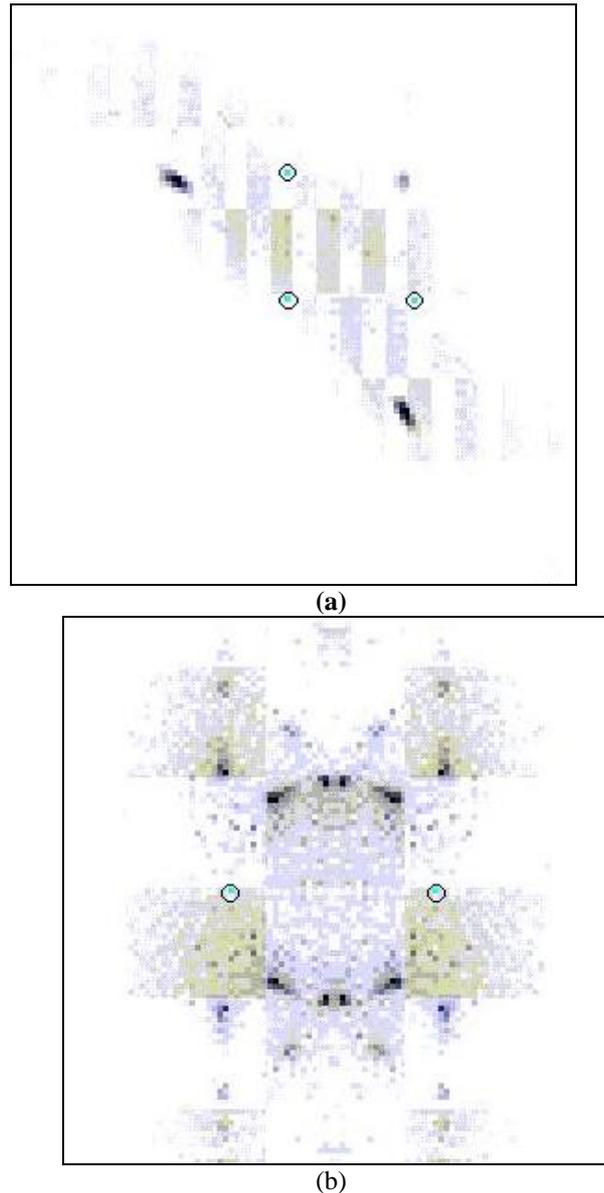


Figure 4.10: La table de hachage (a) a été réalisée à l'aide de bases à trois points, la (b) à l'aide de bases à deux points.

Grâce à la reconnaissance 2D/3D présentée ci-dessus, le suivi d'objet dans l'image caméra est rendu possible.

4.2.4 Evaluation

Afin de valider cette méthode de reconnaissance d'objet polyédrique, nous utilisons l'objet pince (figure 4.2), monté sur le robot à 4ddl décrit dans le §1.1.3.b., auquel le protocole expérimental suivant a été appliqué :

- Calibration du robot à 4 ddl décrit dans le §1.1.3.b. (on note ses coordonnées par rapport à la mire, ainsi que les coordonnées des moteurs du banc correspondantes, et ce, dans 3 positions différentes. On détermine alors les 6 paramètres de la tige.
- Calibration automatique de la caméra.
- Calibration de la pince montée sur le robot à 4 ddl selon la même méthode que celle appliquée au robot.
- Saisie d'une série d'images de la pince avec la caméra calibrée, en notant pour chacune d'entre elles les positions des moteurs du robot.

- Reconnaissance automatique de l'objet, et enregistrement des coordonnées 3D trouvées.
- Evaluation de l'erreur global du système.

Un ensemble de modèles 3D fut sélectionné de la base pour effectuer la reconnaissance (Figure 4.11). La méthode de graphes a été appliquée pour l'appariement. Elle donne un nombre total de 423 hypothèses. C'est l'hypothèse 2 sur le modèle de l'objet Pince, ayant un score 80 qui a réussi.

Nous donnons une description de cette hypothèse :

Chaîne /Face : C7F1 C5F2 C0F9 C6F10 C4F13

Segment /Arête: S3A0 S12A7 S11A8 S7A11 S23A12 S22A13 S0A14 S16A15 S15A16 S20A17 S21A18 S17A19 S13A20 S1A21 S10A22 S8A23 S4A24 S24A25 S18A31 S14A32 S2A33 S9A35

Point / Sommet: P13S0 P12S1 P9S7 P6S8 P14S9 P8S11 P11S12 P0S13 P18S14 P4S16 P17S17 P1S18 P3S19 P5S20 P16S21 P15S22 P7S23

Le modèle reconnu, vérifié par projection, est superposé sur son image Figure 4.12.

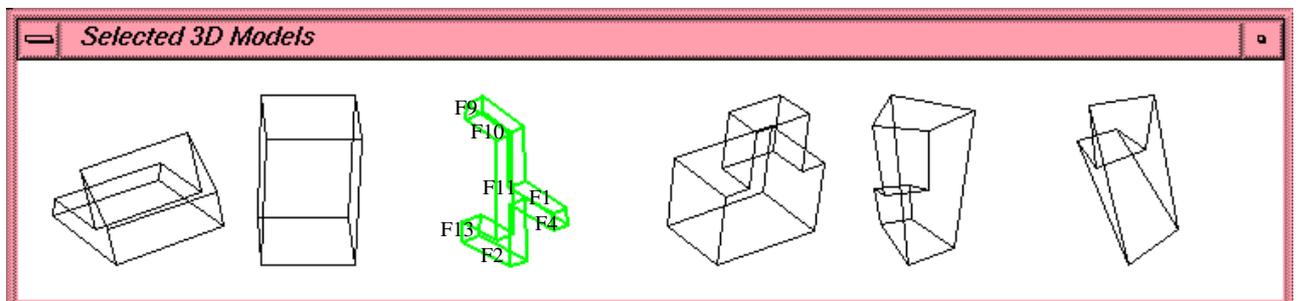


Figure 4.11 : les modèles 3D sélectionnés de la base, le modèle en double épaisseur a été reconnu et ses faces visibles sont marquées

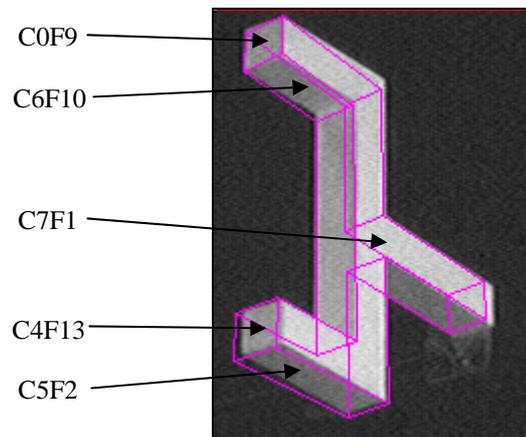


Figure 4.12 : le modèle de l'objet Pince est superposé sur son image avec erreur moyenne de 2.6 pixels, les hypothèses chaîne/face sont marquées

Les erreurs de reconstruction dans le plan image et dans le repère objet ont été calculées :

- l'erreur moyenne est de 2,6 pixels dans le plan image,
- les erreurs moyennes et écarts type dans le repère objet sont présentées dans le tableau ci-dessous :

	X	Y	Z
Moyenne de l'erreur(mm)	4,03	-2,75	-7,90
Ecart-type(mm)	1,67	1,04	6,50

Il est intéressant d'examiner les erreurs relevées par sommet de l'objet. Par exemple, le graphique suivant (figure 4.13) reprend la moyenne de l'erreur, exprimées en mètre, en X pour chacun des sommets :

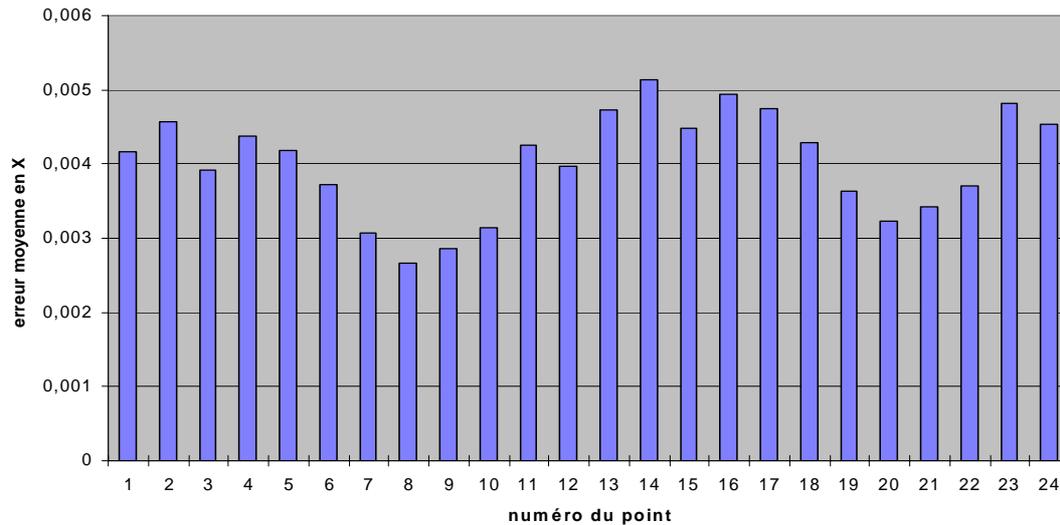


Figure 4.13 Erreur moyenne en X pour chaque sommet

On se rend compte que l'erreur pour certains sommets (8-9 et 20-21) est presque 2 fois moins élevée en moyenne que pour d'autres (2-4, 13-17 et 23-24). En fait la première série de points faisait partie des chaînes utilisées lors de l'appariement, et a donc été prise en compte dans le processus de minimisation de l'erreur d'appariement. Ce n'est pas le cas des sommets de la deuxième série.

Il paraît judicieux de sélectionner les sommets visibles apportant la meilleure précision dans le processus d'optimisation.

4.3 Bilan

Dans ce chapitre, nous avons étudié et développé les chaînons manquants pour l'aboutissement à un système de reconnaissance automatique d'objet polyédrique (traitement d'image + appariement 2D/3D).

Pour le traitement d'image, nous avons amélioré l'application de la transformée de Hough pour l'extraction de segments.

Concernant le difficile problème d'appariement 2D/3D, nous avons développé une méthode particulière à hypothèses simplificatrices dans un premier temps. Ensuite, nous avons utilisé et proposé des améliorations de méthodes plus générales basées sur le graphe d'appariement et sur le hachage géométrique.

Nous avons pris en compte, lors de l'élaboration des hypothèses dans le graphe d'appariement, les invariants géométriques basés sur les rapports entre les surfaces du modèle et celle de son image.

Pour la construction de la table de hachage, nous appliquons des quasi-invariants obtenus par des bases à deux points.

Le système est actuellement en cours d'évaluation. Des résultats partiels sont obtenus, certes pas avec la précision attendue. Néanmoins ils sont encourageants et nous pensons pouvoir les améliorer.

La méthode basée sur les liaisons est publiée dans le congrès IEEE CESA'SMC [COM 96b].

Le travail portant sur le §4.2 est en cours d'évaluation, il fait l'objet de la thèse de Doctorat de Mr Mudar Shaheen. Une publication est parue dans : "International Symposium on Artificial Intelligence, Robotics, and

Intellectual Human Activity Support for Nuclear Applications, 8.1-8.11, Nov. 19th-21st, 1997, Riken, Wako-shi, Saitama, Japan."[COM 97b].

4.4 Bibliographie du chapitre 4

- [AYACHE 1986] Ayache N. and Faugeras O.D., "HYPER: A new approach for the recognition and positioning of two-dimensional objects", IEEE Trans. PAMI, 8(1), January 1986, pp.44-54.
- [BINFORD 1993] BINFORD T.O. and LEVITT T.S. "Quasi Invariants : Theory and exploitation" In Proceeding of DARPA, Image understanding Workshop, 1993, pp. 819-829.
- [BOLLES 1981] Bolles R.C. and Fischler M.A. "A RANSAC-based approach to model fitting and its application to finding cylinders in range data" Intern. joint conference on Artificial Intelligence, pp. 637-643, Vancouver, Canada, 1981.
- [BOLLES 1982] Bolles R.C. and Cain R.A. "Recognizing and locating partially visible objects : the local feature -focus method" Intern. Journal of Robotics Research , 1(3), 1982, pp. 57-82.
- [BOLLES 1986] Bolles R.C. and Horaud R. "3DPO : A three Dimensional part orientation System " , Intern. . Journal of Robotics Research , 5(3), 1986, pp. 3-26.
- [BOWYER 1992] Bowyer K.W., "Workshop panel report - why aspect graphs are not (yet) practical for computer vision", CVGIP: Image Understanding, vol. 55, No. 2, March 1992, pp. 212-218.
- [EDELBRUNNER 1986] Edelsbrunner H., O'Rourke J. and Seidel R., "Constructing arrangements of lines and hyperplanes with applications", SIAM Journal on Computing, 15(2), May 1986, pp. 341-363.
- [GROS 1995]GROS P. "Matching and Clustering : Two steps towards objects modelling in Computer Vision". The International Journal of Robotics Research, 14(5), October 1995.
- [HANSEN 1989] Hansen C and Henderson T.C. "CAGD-based Computer Vision", IEEE Transactions on PAMI, 11(11), November 1989, pp. 1181-1193.
- [HORAUD 1984] Horaud R. and Bolles R.C., "3DPO's strategy for matching three-dimensional objects in range data", IEEE Conf. on Robotics & Automation, Atlanta, March 13-15 1984, pp. 78-85.
- [IKEUCHI 1987] Ikeuchi K, "Generating an interpretation tree from a CAD model for 3D-Object recognition in bin-picking tasks". International Journal of Computer Vision 1(2), 1987, pp. 145-165.
- [IKEUCHI 1988] Ikeuchi K and Kanade T., "Applying sensor models to automatic generation of 3D Object recognition program ", 2nd International Conference on Computer Vision, Tampa(FL), November 1988, pp. 228-237.
- [GRIMSON 1984] Grimson W.E.L. and Lozano-Perez T., "Model-based recognition and localization from sparse range or tactile data", International Journal of Robotics Research, 5(3), Fall 1984, pp. 3-34.
- [JAIN 1988] Jain A.K. and Hoffman R., "Evidence-based recognition of 3-D objects", IEEE Trans. on PAMI, 10(6), November 1988, pp. 783-802.
- [LAMDMAN 1988] Lamdan Y. and Wolfson H.J., "Geometric hashing: A general and efficient model-based recognition scheme", Second IEEE Int'l Conf. Computer Vision, Tampa, FL, December 1988, pp. 238-249.
- [LOWE 1987] Lowe D.G., "Three dimensional object recognition from single two-dimensional images", Artificial Intelligence 31(1987), Elsevier Science Publishers B.V.(North-Holland), pp. 355-395.
- [PAMPAGNIN 1988] Pampagnin L.H. and Devy M., "Identification d'un objet 3D par mise en correspondance entre une seule image et un modèle", Rapport interne, LAAS-France, 1988.
- [PLATINGA 1990] Plantinga H. and Dyer C., "Visibility, occlusion and the aspect graph", Int. Journal of Computer Vision, 5(2), 1990, pp. 137-160.
- [STEWMAN 1988] Stewman J.H. and Bowyer K.W., "Creating the perspective projection aspect graph of polyhedral objects", in IEEE International Conference on Computer Vision, 1988, pp. 494-500.
- [ZHU 1995] Zhu y. , Seneviratne L.D. and Earles S.W.E. "A new structure of invariants for 3D point sets from single view" IEEE international conference on Robotics and Automation, Nagaya(Japan), May 21-27, 1995, pp. 1726-1731.

5. APPLICATIONS

5.0 Introduction

Afin de valider les différents modules de MCIT, plusieurs applications ont été mises en œuvre.

La 1^{ère} avait pour objectif l'assistance de l'opérateur à la perception de l'évolution du robot opérant dans un milieu perçu par une caméra et où la visibilité est dégradée. Cette assistance s'est concrétisée par la superposition d'une représentation graphique du robot sur son image caméra. Cette superposition en ligne est rendue possible par la connaissance du modèle 3D du robot, de son modèle géométrique direct et des informations capteurs renseignant à tout instant sur la configuration du robot. Le robot MAT2 (bras manipulateur 6R, construit par la société LA CALHENE) a été pris comme exemple.

La 2^{ème} application consistait en la simulation des tâches du robot MAT2, dont la mise en œuvre devenait laborieuse. Cette nécessité de simulation nous a amené à développer un module de planification de trajectoires de robot mobile dans un premier temps puis pour un bras manipulateur (MAT2) en 2D dans un second temps. Ce module est décrit dans le §5.2.

La 3^{ème} application a eu pour objectif, dans le cadre de collaborations avec la société Aleph Technologies (Alma actuellement) et le service de télé robotique du CEN, l'ajout au logiciel industriel ACT de Aleph Technologies de deux modules. Le premier concerne la calibration de la caméra et le second la modélisation des objets en tenant compte des liaisons physiques entre eux.

L'application actuelle consiste en l'utilisation des modules de MCIT via internet.

Nous décrivons dans ce qui suit chacune de ces applications.

5.1. Animation de la représentation graphique de bras manipulateur superposée à l'image caméra.

Cette application consiste à animer la représentation graphique du robot en la superposant à l'image caméra. Le robot étant équipé de capteurs articulaires (capteurs proprioceptifs), ceux ci mesurent le vecteur des variables articulaires du robot \vec{q} . La position et l'orientation de chaque segment, de la représentation graphique, du robot sont calculées en prenant en compte le vecteur \vec{q} dans le calcul du modèle géométrique direct du robot. Si $Toi(\vec{q})$ est la matrice homogène de passage du repère (Ro) de travail au repère (Ri) lié à l'articulation i , nous avons:

$$\begin{pmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{pmatrix}_{/(Ro)} = Toi(\vec{q}) \cdot \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}_{/(Ri)}$$

où X_i , Y_i , et Z_i sont les coordonnées dans (Ri) d'un point du bras attaché à ce repère.

La visualisation de la représentation graphique du robot sur l'écran se fait alors en utilisant l'équation suivante:

$$\begin{pmatrix} s.u \\ s.v \\ s \end{pmatrix} = C \cdot Toi(\vec{q}) \cdot \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}_{/(Ri)}$$

C étant le modèle global de la caméra défini pendant la phase de la calibration de celle ci (cf. §1.1.2.).

Ce travail a été publié dans : "ROBOTICA, International Journal of Information, Education and Research in Robotics and Artificial Intelligence, Edited by J. Rose, Cambridge University Press, England vol. 10, 1992", [ART 1992]

5.2. Planification de trajectoires en 2 1/2 D de bras manipulateur

La planification de trajectoires est une aide précieuse à la téléopération principalement en robotique d'intervention. Elle permet une meilleure efficacité des déplacements du robot par calcul d'une trajectoire optimale (économies de temps et d'énergie). De plus, elle simplifie grandement la programmation des tâches (interface homme-machine).

Il s'agissait de mettre au point et de développer une méthode permettant la planification automatique de trajectoires pour un robot mobile en 2D et pour un bras manipulateur en 2 1/2D. Le robot évolue dans un environnement encombré par des obstacles (bi-dimensionnels pour le robot mobile et tri-dimensionnels pour le bras manipulateur). Partant d'une position initiale, il devra rejoindre une position finale sans entrer en collision avec les obstacles de la scène. Les paramètres géométriques du robot sont pris en compte afin de générer des trajectoires réalistes. L'analyse du sujet est volontairement décomposée en trois phases: étude du cas du point mobile, étude du robot mobile puis étude du bras manipulateur.

Les méthodes de résolution du problème de planification de trajectoires pour un robot mobile sont principalement: la méthode du *Grphe des Visibilités*, applicable essentiellement à un robot mobile en translation; les méthodes de *décomposition cellulaires*, concernant le point mobile ou le robot mobile assimilé en tant que tel; les méthodes de *propagation de coût*, concernant également le point mobile dans un environnement décrit grossièrement; la méthode des *potentiels fictifs*, concernant le robot mobile⁶. Cette dernière est fortement limitée par l'existence de zones de blocage (minima locaux).

La méthode développée combine la planification, au niveau *global*, utilisant le *Grphe des Visibilités*, et l'algorithme A* pour la détermination du chemin optimal, avec une procédure de validation *locale* du chemin trouvé à l'aide de *potentiels fictifs*.

Cette méthode hybride combinant les deux niveaux de recherche global et local permet de résoudre une grande partie des problèmes de planification pour un robot mobile. L'étude a été menée en utilisant un modèle particulier de plate-forme à deux roues motrices et une roue folle. Néanmoins, son extension à d'autres types de plates-formes ne devrait pas poser de problème, à condition de modifier l'étude géométrique des déplacements du robot.

Cette méthode mixte est étendue au cas 2 1/2D et permet de résoudre avec une assez bonne efficacité (rapidité d'exécution), sans grande modification de principe, une grande partie des problèmes de déplacement d'un bras manipulateur. La représentation graphique est effectuée en 3 dimensions mais le raisonnement permettant le déplacement du robot se basera sur la modélisation d'un environnement en 2 dimensions.

Le planificateur de tâche du système MCIT met en œuvre 2 sous modules [COM 93a]:

- un sous module de génération de trajectoire en 2D,
- un sous module de génération de configurations de saisie d'un objet cible par l'effecteur d'un robot dans un environnement 2 1/2D contenant des objets de faible hauteur posés sur un plan de travail.

Le sous module de génération de trajectoire en 2D applique une approche hybride mettant en œuvre une méthode globale: le graphe de visibilité (GV), et une méthode locale, le champs de potentiel (CP).

Le GV est utilisé pour calculer toutes les trajectoires possibles reliant la position de départ à la position d'arrivée du robot. Une trajectoire est obtenue en reliant le point de départ au point d'arrivée en passant par les sommets S_i des obstacles O_i (cf. fig.5.1).

⁶ Latombe J.C. « Robot Motion Planning. » Kluwer Academic Publishers, USA, 1991.

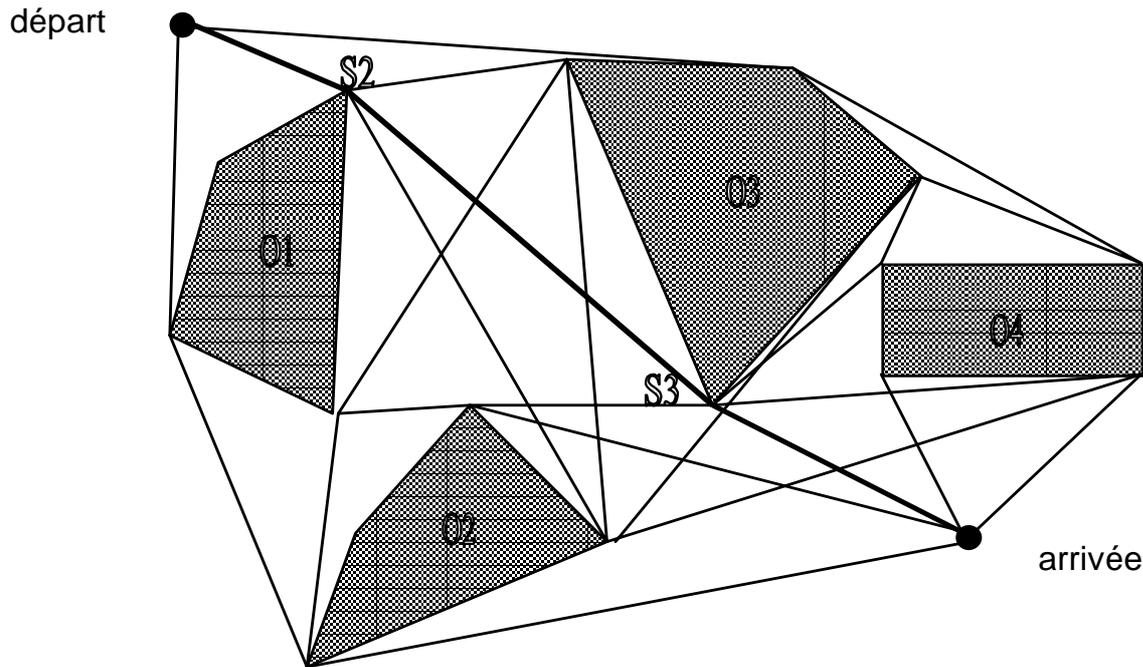


Figure 5.1: Graphe de visibilité

Le sommet S_i de l'obstacle O_i est lié au sommet S_{i+1} si :

- S_{i+1} est un sommet d'un obstacle autre que O_i , ou le but à atteindre, tel que la droite $S_i S_{i+1}$ n'intersecte pas O_i ,
- le trajet $S_i S_{i+1}$ est libre d'obstacles.

L'algorithme GV a été développé en tenant compte de l'encombrement du robot par dilatation des obstacles (en fonction des dimensions du robot).

A partir du GV, toutes les trajectoires possibles sont obtenues. La trajectoire qui minimise la distance euclidienne entre les points de départ et d'arrivée est obtenue grâce à l'algorithme A^* .

La trajectoire à suivre par le robot étant définie par la méthode globale, afin de guider le robot entre deux sommets successifs S_i et S_{i+1} et pour fixer son orientation, un champ de potentiel artificiel est créé par ceux-ci.

Le robot se situant entre deux sommets S_i et S_{i+1} , l'obstacle O_i crée un champ de potentiel répulsif qui a tendance à éloigner le robot de celui-ci, l'obstacle O_{i+1} quant à lui crée un potentiel attractif. La consigne appliquée au robot pour aller du point de départ au point d'arrivée est calculée en fonction de la résultante du champ de potentiels en chaque point de la trajectoire. Les représentations et calculs des forces sont illustrés par la figure 5.2 et les formules suivantes :

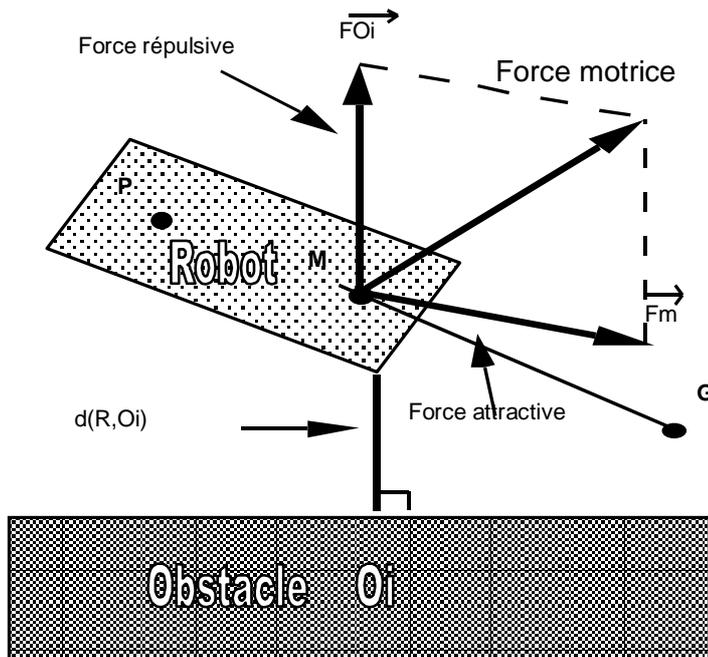


Figure 5.2 : Champ de potentiels

$$\|\vec{F}_{O_i}\| = C_o \times 1/d(R, O_i)^2$$

$$\|\vec{F}_m\| = C_g = \text{constante}$$

(Pour les buts intermédiaires G_i)

$$\|\vec{F}_m\| = C_g \times 1/d(R, G)^2$$

(Pour le but final G)

Le sous module de génération de configuration de saisie permet de calculer la position et l'orientation de l'effecteur pour la meilleure prise de l'objet. L'algorithme implanté est utilisé pour des objets convexes ou des objets concaves. L'orientation de l'effecteur pour la saisie est calculée de telle manière que l'une des mâchoires de l'effecteur soit en contact avec une face de l'objet à saisir, l'autre le sera sur la face ou l'arête opposée (cf. fig. 5.3).

La tâche qui peut être planifiée est une tâche de saisie avec désignation de l'objet sur l'écran vidéo ce qui constitue une commande de niveau objet.

Les opérations qui sont effectuées pour cette commande niveau objet sont les suivantes pour une saisie d'objet:

- ↳ désigner l'objet à saisir sur l'écran vidéo,
- ↳ rechercher toutes les configurations de saisie et toutes les configurations d'approche (sous module générateur de configurations de saisie),
- ↳ rechercher les trajectoires à suivre pour déplacer l'effecteur de la position initiale jusqu'aux points d'approches (sous module générateur de trajectoires),
- ↳ choisir la trajectoire dont le chemin est le plus court,
- ↳ envoyer les consignes niveau effecteur au robot.

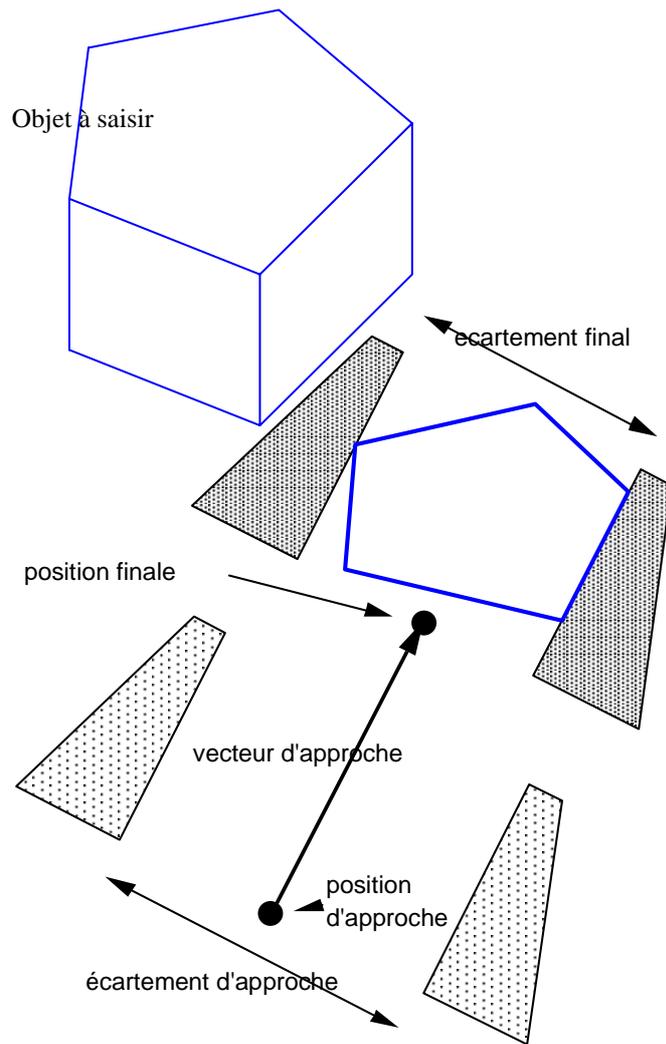


Figure 5.3: saisie d'un objet

La figure 5.4 montre un exemple de simulation, de la tâche de saisie-dépose, effectuée avec une représentation graphique du robot MAT2.

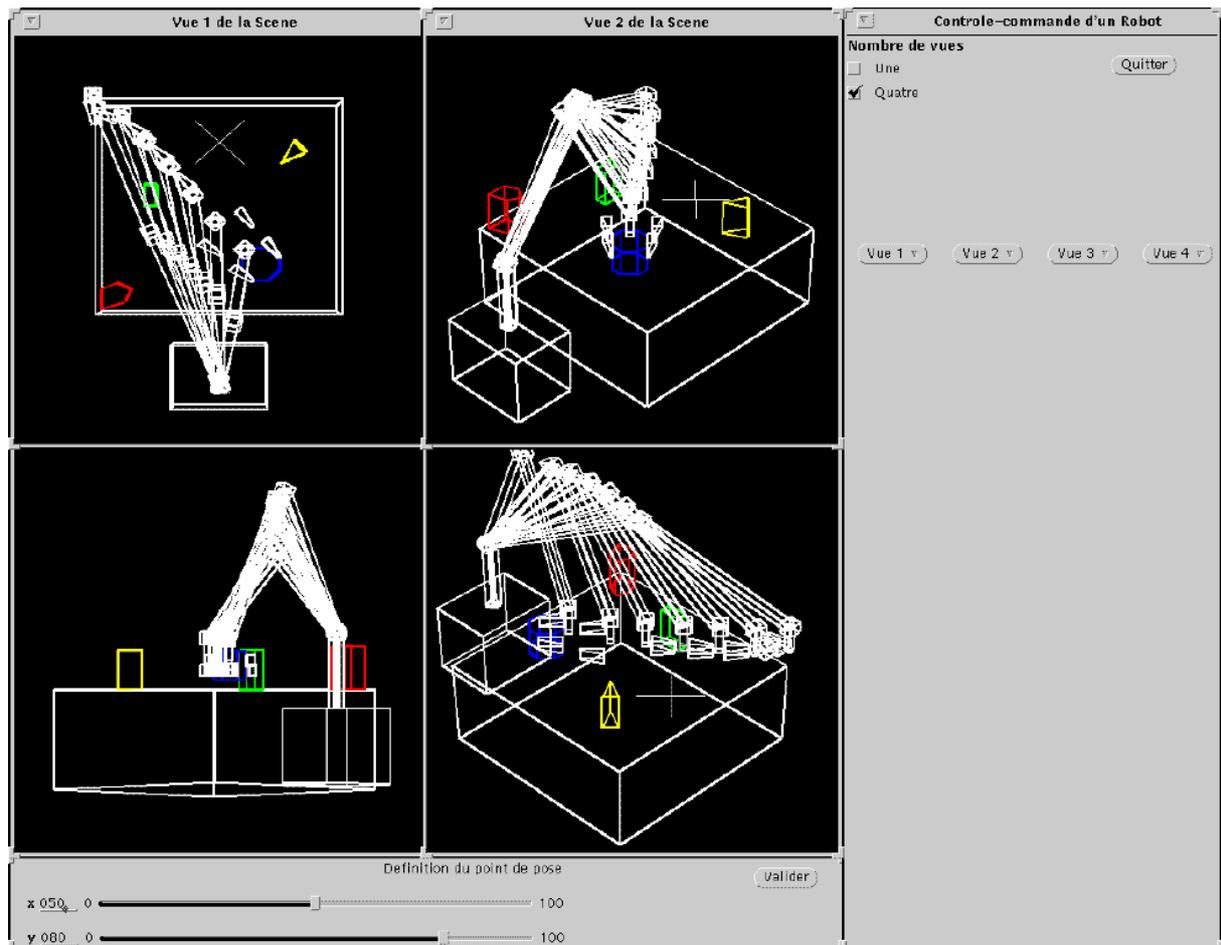


Figure 5.4 : Simulation de tâche de saisie dépose d'objet avec le MAT2

Ce travail a été publié dans le congrès IEEE, SEE, COMPEURO93, "International Conference on Computer in Design, Manufacturing and Production, , Evry, -France-, May 24-27 1993 ", [COM 1993a]

5.3. Programmation CAO de tâches robotiques

L'objectif de la programmation graphique est d'effectuer un contrôle au niveau tâche d'un site robotique téléopéré. Il s'agit de faciliter le travail de l'opérateur en rendant la programmation de la tâche simple et en effectuant les tests de faisabilité et de sécurité de celle-ci.

Afin de réaliser cet objectif, nous utilisons la notion de liaison entre les objets de l'environnement. Le concept de liaison permet une description minimale des tâches et permet donc de les programmer en spécifiant un nombre minimum de paramètres. Une liaison peut être réelle (liaison appui-plan entre un livre posé sur une table) ou virtuelle (repère d'une caméra virtuelle). De plus, par les relations qu'il apporte entre les objets, le concept liaison enrichit la représentation des connaissances, notamment la réalité du contact entre les objets. La richesse du concept liaison permettra à terme de faciliter la reconstruction 3D d'un ensemble d'objets ainsi qu'une commande minimale.

En effet, les liaisons entre les objets étant enregistrées, les paramètres utiles au calcul des trajectoires associées à la tâche sont disponibles. Des algorithmes de génération de trajectoire avec contraintes peuvent alors être utilisés tels la commande minimale à partir de la jacobienne⁷, ou le calcul de trajectoire dans le cas de manipulation au contact. Ces précédents travaux font appel au formalisme des liaisons sans pour autant utiliser une base de données liaisons, les paramètres des liaisons étant à la charge de l'opérateur. Notre contribution consiste donc en l'introduction d'une base de données liaison.

Les paramètres physiques statiques (frottement, pesanteur) sont introduits. Ceci permet, avec la notion de liaison, d'effectuer des tests de faisabilité d'une action, par exemple en introduisant des butées virtuelles aux liaisons. Par contre, l'inertie et les forces de contact ne sont pas introduites. En conséquence, il n'est pas possible dans l'état actuel de faire une simulation dynamique et donc de créer par exemple un retour d'effort utilisé pour le contrôle manuel de niveau effecteur ainsi que pour la téléprogrammation qui se base sur une simulation de ce contrôle manuel.

Le concept objet informatique permet, quand à lui, de mettre en œuvre la notion de liaison. Une organisation arborescente et hiérarchique des connaissances du monde est obtenue par ce concept ainsi qu'un accès naturel. Il permet une meilleure modularité et une facilité d'extension du système.

Le logiciel de base utilisé est le modèleur et le générateur ACT de l'entreprise Aleph Technologies. Une base de données liaisons est ajoutée. Les actions du robot sont définies par des créations et des destructions de liaisons.

Une base de données procédés est également utilisée représentant les différentes tâches. Un procédé peut être décrit de manière simple et systématique comme une succession d'actions élémentaires. Il permet de simplifier la programmation tout en assurant la sécurité des opérations.

. PLANIFICATION DE TACHE DE SAISIE - DEPOSE BASEE SUR LES LIASONS

Dans la figure 5.5, Quatre repères et trois matrices homogènes sont utilisées pour exprimer la liaison entre Objet1 et Objet2. R1 et R2 sont les repères respectifs de ces 2 objets. Rplan1 est un repère attaché à la face de dessus de Objet1, et Rplan2 est un repère attaché à la face de dessous de Objet2. T1, T2 et T3 sont des matrices homogènes 3D, dont la composition exprime la liaison entre Objet1 et Objet2. T2 est la seule transformation variable. La liaison entre les 2 objets est de type appui plan (3 degrés de liberté). L'avantage principal des liaisons est qu'elles permettent la connaissance à priori du nombre de degrés de liberté entre objets. Aussi, seuls ceux - la sont sujets à modification lors de la programmation de tâches.

⁷ Dombre E. et Khalil W. « Modélisation et commande des robots. », Hermès, Paris, 1988.

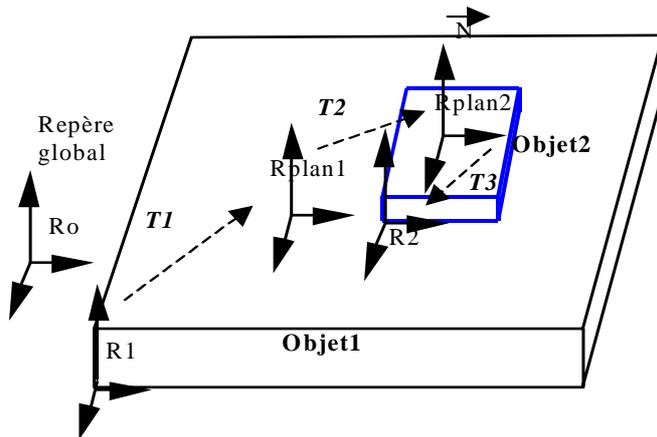


Fig.5.5 Modélisation d'objets en utilisant le concept de liaison

Lors de la phase de programmation de tâche, en utilisant la connaissance sur les liaisons entre les objets de l'environnement, le système assure un contrôle efficace de l'exécution.

Dans une tâche de saisie - dépose, les actions suivantes sont réalisées :

- A1 - Approche de l'objet à saisir(Objet2) : l'effecteur du robot est positionné selon la normale N au plan2.
- A2 - Saisie de Objet2.
- A3 - Dégageement de Objet2.
- A4 - Transport de Objet2 à sa destination.
- A5 - Dépose de Objet2 à sa destination(supposée être sur Objet1)
- A6 - Libération de Objet2.
- A7 - Déplacement de l'effecteur à sa position de repos.

La table ci-dessous illustre les modifications subies par les liaisons entre l'objet père(table) et l'objet fils(livre) lors de la tâche de saisie - dépose :

ACTION	LIAISON	MOUVEMENT DU ROBOT	UTILISATION DE EFFECTEUR
A1	Non affectée	Oui	Non
A2	Effecteur - Objet2 créée	Non	Oui
A3	Objet1-Objet2 détruite	Oui	Non
A4	Non affectée	Oui	Non
A5	Objet1-Objet2 créée	Oui	Non
A6	Effecteur - Objet2 détruite	Non	Oui
A7	Non affectée	Oui	Non

Table3: Modification de liaisons dans la tâche de saisie - dépose

Ce travail a été publié dans le symposium : " Second Intern. Symposium on Methods and Models in Automation and robotics. 30/8-2/9/1995, Miedzyzdroje, Poland.", [COM 1995].

Cette programmation de tâche de saisie - dépose a été intégrée dans le logiciel industriel de CAO - Robotique ACT, ainsi que la calibration de la caméra. Les figures ci-dessous illustre ces intégrations :

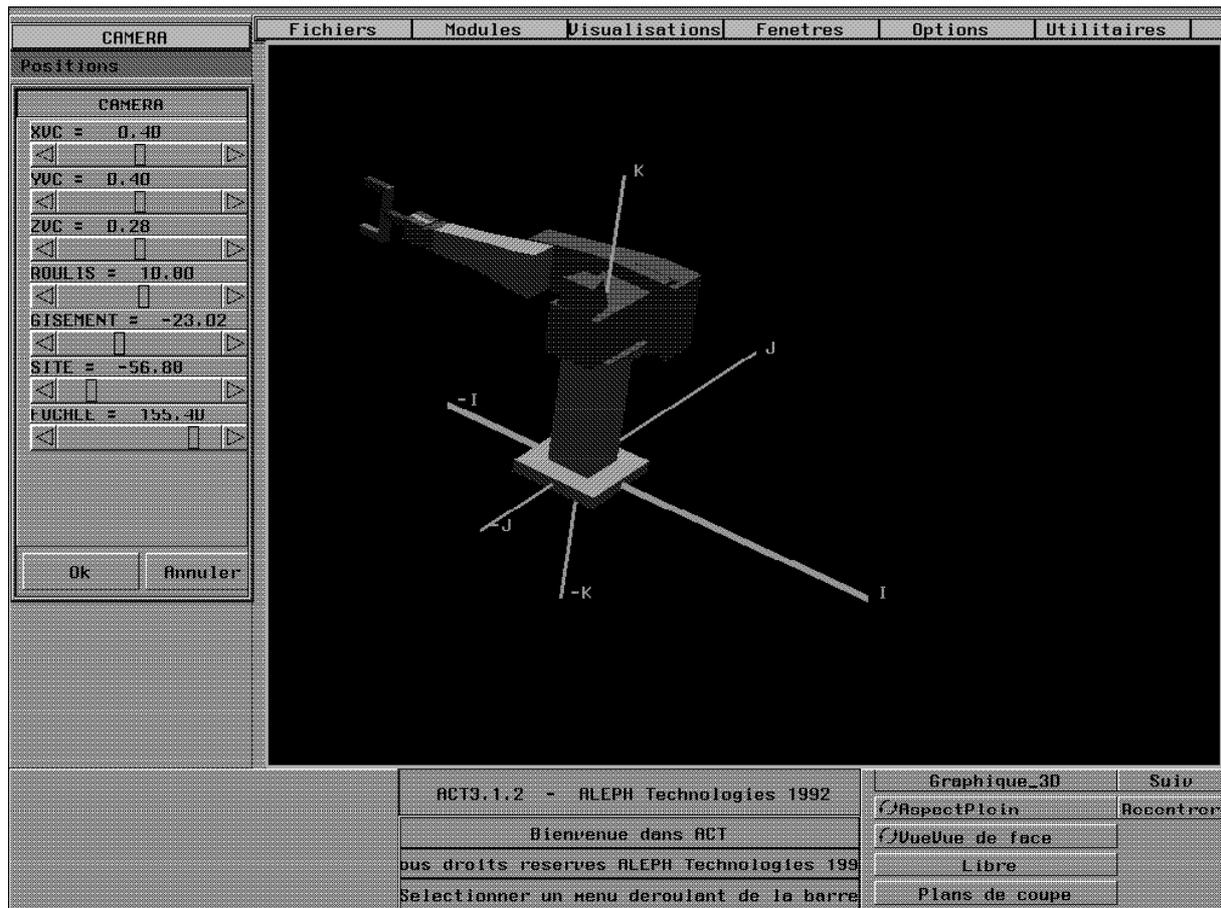


Figure 5.6 : intégration de la caméra dans le logiciel ACT
 [ENC-ING1997b]

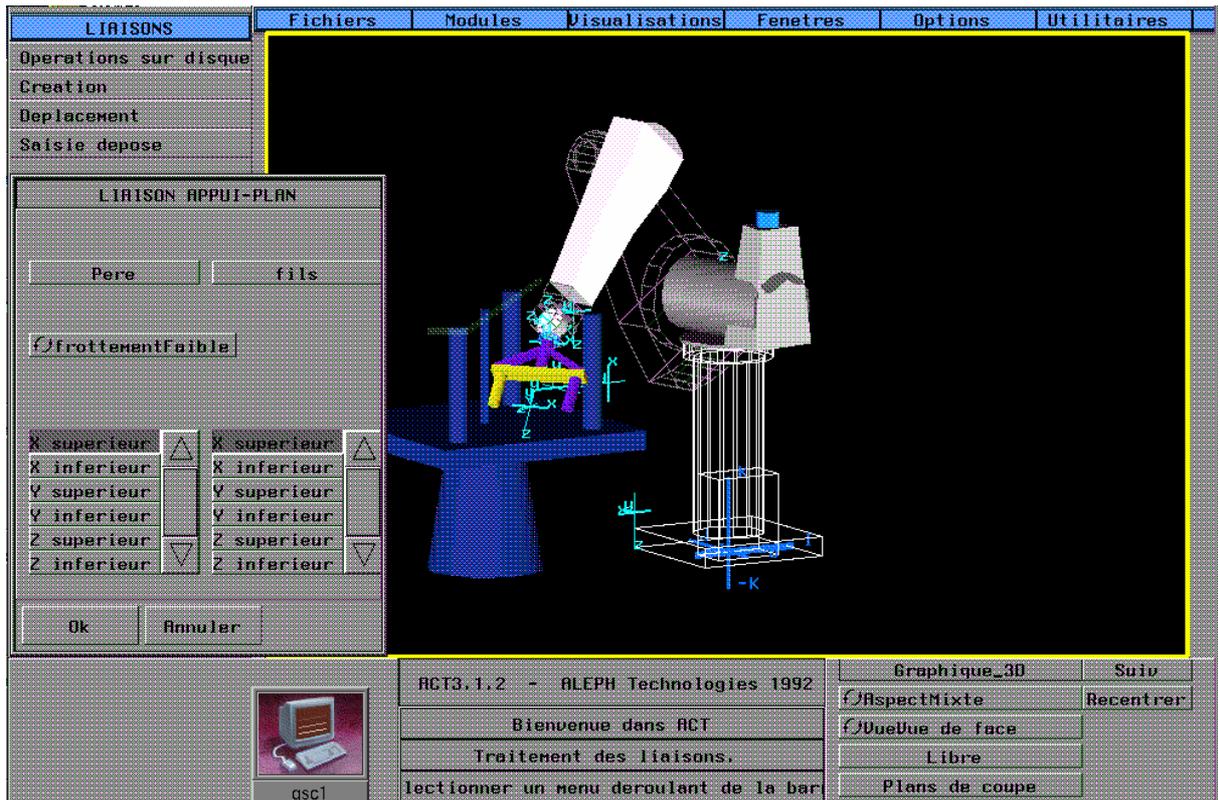


Figure 5.7 : Intégration des liaisons dans le logiciel de CAO ACT

5.4 Contrôle-commande de robot via internet

Ce contrôle commande via internet est possible par l'accès au site Web du Cemif Système Complexes dont le serveur a pour adresse : "gsc26.cemif.univ-evry.fr" par tout ordinateur doté du navigateur Netscape/internet explorer. Cet accès permet une visualisation en quasi temps réel de l'image d'un robot du Cemif et la télécommande graphique du robot. En effet, le robot et son environnement sont représentés par leur modèle 3D. Les modèles géométriques direct et inverse du robot sont également pris en compte. La désignation sur écran de l'un des composants de la scène équivaut à une désignation dans le modèle 3D.

Deux programmes ont été développés, l'un correspond au serveur et l'autre au client écrit en JAVA. Le serveur a pour rôle de transmettre au client le retour vidéo et d'appliquer au robot les consignes fournies par le

client. Concernant le transfert des informations vidéo, seule la première image est comprimée et envoyée dans son intégralité, ensuite sont transmises les différences entre une image et la suivante. Le serveur est exécuté donc sur la machine locale. Le client est exécuté sur la machine distante et permet d'une part une supervision du robot et d'autre part une commande de niveau objet de chacun des composants du robot.

La télécommande du robot est bien entendu réservée à un seul client à la fois, il s'agit du premier connecté.

La figure 5.8 montre l'accès au site du CEMIF via le navigateur Netscape et l'interface développée.



Figure 5.8 : Contrôle commande de robot via internet

6. Conclusions et Perspectives

Nous avons, pendant cette dernière décennie, étudié et mis en œuvre des méthodes et des outils indispensables à un système de réalité augmentée pour la Téléopération.

Les contraintes liées à la réalité augmentée, qui sont d'ordre temporelle et de précision nous ont amené à entreprendre des recherches dans plusieurs thèmes.

En effet, ceux-ci concernent les modélisation et calibration de capteurs, la modélisation 3D, la reconstruction 3D, la reconnaissance d'objets polyédriques. Nous avons traité ces points dans l'ordre cité. Le respect de cet ordre est important car tout problème constitue un prérequis pour le ou les suivants. De la précision de la calibration des capteurs dépend celle de la modélisation 3D et de la reconstruction 3D et de ce fait celle de la reconnaissance d'objets.

Pour la calibration de capteurs (caméra et télémètre), nous avons étudié et mis en œuvre des méthodes permettant d'améliorer la précision. L'erreur de reconstruction est inférieure au pixel (de l'ordre du 1/10 de pixel) pour une caméra à faibles distorsions géométriques et de l'ordre de 1 – 2 pixels pour une caméra à tube.

Par ailleurs, la calibration automatique, basée sur la pré calibration d'un mini robot peut être réalisée en ligne, si la caméra devient mobile, sans l'intervention de l'opérateur. Cette possibilité est très utile pour la réalisation de tâches en réalité augmentée.

Une procédure de calibration du dispositif de relevé 3D a été étudiée et testée, avec une application à la modélisation de volumes. Le modèle géométrique a été séparé en un modèle interne et un modèle externe. Ce dernier est constitué par la matrice homogène qui situe le capteur par rapport au repère de référence. Les procédures de détermination de cette matrice permettent de tenir compte des contraintes de norme et d'orthogonalité de la matrice de rotation.

La modélisation 3D, basée sur l'utilisation du capteur mixte (caméra+télémètre), a concerné la construction de volumes englobants. L'objectif de celle-ci est de compléter la base de données 3D de l'environnement du robot et, ainsi, de permettre à la tâche de celui-ci de ne pas s'interrompre. Par ailleurs, cette modélisation a permis également d'évaluer la précision du relevé 3D basé sur l'utilisation de différentes combinaisons de capteurs. L'erreur sur la profondeur est inférieure au centimètre, pour les distances DR3D-objet inférieures à 1.5m, et les erreurs sur les dimensions de l'objet sont inférieures à 5mm.

La reconstruction 3D est basée sur l'utilisation d'un modèle 3D et d'une image caméra 2D. Plusieurs méthodes basées sur des indices visuels différents ont été développées et comparées. Elles constituent un outil incontournable pour la reconnaissance d'objet.

Les méthodes de recalage ont été testées sur des objets réels. Les indices 2D sont donnés directement par désignation sur l'image caméra de la nouvelle position de l'objet. Pour le recalage à grand mouvement, l'erreur sur la rotation est inférieure à 0.6° et sur la translation inférieure à 4cm. La précision de la translation étant liée directement à celle de la rotation. Dans le cas de faible mouvement, l'algorithme de recalage fin donne de bons résultats, l'erreur sur la rotation est inférieure à 0.3° et sur la translation inférieure à 5mm.

Le recalage basé sur les sommets utilise un formalisme nouveau qui est plus général car il peut s'appliquer à des objets non polyédriques. L'erreur de reconstruction obtenue est inférieure au centimètre pour une distance caméra-objet inférieure à 1.5m.

Pour la reconnaissance automatique d'objet polyédrique, un traitement d'image a été développé ainsi que deux méthodes d'appariement utilisant l'une les faces et l'autre les sommets de l'objet à reconnaître.

Nous avons développé une méthode particulière à hypothèses simplificatrices dans un premier temps. Ensuite, nous avons utilisé et proposé des améliorations de méthodes plus générales basées sur le graphe d'appariement et sur le hachage géométrique.

Nous avons pris en compte, lors de l'élaboration des hypothèses dans le graphe d'appariement, des informations topologiques et des invariants géométriques entre le modèle 3D de l'objet et de son image.

Pour la construction de la table de hachage, nous appliquons des quasi-invariants obtenus par des bases à deux points.

Comme perspectives :

A court terme, nous poursuivons l'évaluation des méthodes de reconnaissance automatique d'objet afin d'améliorer leur précision et permettre un suivi d'objet en ligne. L'utilisation des invariants entre une image et son modèle est l'une des voies à poursuivre.

La méthode de reconstruction 3D, basée sur les sommets, qui introduit des déformations de modèle peut être utilisée, dans le cadre du projet TRV(Téléopération et Réalité Virtuelle) du CNRS, pour la détection des écarts entre les modèles et la réalité.

A moyen terme, nous envisageons de développer la télé programmation de tâches de téléopération via internet. Plusieurs problèmes se posent dont celui du retard de transmission d'informations entre les sites maître et esclave et celui de la précision de la mise en correspondance des mondes réel et virtuel.

Afin d'éviter le problème du retard de transmission d'informations entre les sites maître et esclave, nous allons mettre a contribution les outils décrits dans ce mémoire. L'objectif est de minimiser le débit d'informations à transférer. Cette minimisation est rendue possible par d'importants traitements, dans les sites maître et esclave, et par une sélection judicieuse des informations pertinentes à transférer.

Le site maître sera doté de traitements tels que la reconnaissance en ligne de l'environnement proche du robot, et la prédiction de la trajectoire du robot. La reconnaissance de l'environnement du robot permet d'effectuer une mise en conformité précise des modèles virtuel et réel. Ceci permet d'effectuer une commande de haut niveau du robot. Celle-ci sera soumise dans un premier temps au module de prédiction de trajectoires puis ensuite envoyée au robot réel. Un retour d'effort virtuel devra compléter le module de prédiction de trajectoires.

Le site esclave est équipé aussi d'un calculateur permettant de transmettre, en temps réel au site maître, des informations capteurs minimales renseignant sur l'état du robot et de son environnement et d'effectuer la commande du robot en prenant en compte les consignes issues du site maître. Toutes ces informations capteurs servent à la mise en conformité des modèles réel et virtuel.

7. LISTE DES PUBLICATIONS

PUBLICATIONS DANS REVUES (à comité de lecture)

[ART 1997] E.C. N'ZI, M. MALLEM, F. CHAVAND "Interactive building and updating of a 3D database for teleoperation" Robotica, International Journal of Information, Education and Research in Robotics and Artificial Intelligence, Edited by J. Rose, Cambridge University Press, England vol. 15, part 5, pp. 493-510, oct.1997.

[ART 1995] E.C. N'ZI, J. TRIBOULET, M. MALLEM, F. CHAVAND "Mise à jour d'une BD 3D en Téléopération : Modélisation d'objets non connus" Revue d'Automatique et de Productique Appliquées, vol. 8 N°4/1995, pp. 565-583, HERMES, Paris.

[ART 1992] M. MALLEM, F. CHAVAND and E. COLLE
"Computer-assisted visual perception in teleoperated robotics"
ROBOTICA, International Journal of Information, Education and Research in Robotics and Artificial Intelligence, Edited by J. Rose, Cambridge University Press, England vol. 10, pp. 93-103, 1992.

[ART 1991] M. MALLEM, E. COLLE and F. CHAVAND
"A VME-bus tool for the creation and animation of a synthetic picture, mixed with a camera image",
MICROPROCESSORS AND MICROSYSTEMS, Butterworth-Heinemann Ltd, London Vol. 15 N°8, pp. 427-433, October 1991.

[ART 1990] M. MALLEM, F. CHAVAND
"The perception function in teleoperation"
Revue d'Automatique et de Productique Appliquées, vol. 3 N°3-4/1990, pp. 71-77, HERMES, Paris

COMMUNICATIONS DANS CONGRES - COLLOQUES (à comité de lecture et actes)

[COM 1998] F. CHAVAND, M. MALLEM, "Enhanced Reality Versus Virtual Reality for Teleoperated Robotics and Teleoperation", CESA'98, IEEE-SMC, pp. 626-631 Hamamet, Tunisie, April 1-4th, 1998

[COM 1997a] G. MOREAU, M. MALLEM, F. CHAVAND, C. N'ZI "Two 3D Recovering Methods for Robot Control" IFAC'97, SYROCO ,pp. 531-537, Nantes, France, Sep. 3-5th, 1997

[COM 1997b] F. CHAVAND, M. SHAHEEN, M. MALLEM, "« Matching between a 2D-Image and its 3D-model : Application to Updating 3D Information of Working Spaces »International Symposium on Artificial Intelligence, Robotics, and Intellectual Human Activity Support for Nuclear Applications, 8.1-8.11, Nov. 19th-21st, 1997, Riken, Wako-shi, Saitama, Japan.

[COM 1996a] E.C. N'ZI, M. MALLEM and F. CHAVAND_ "Method for Updating the Environment 3D Geometrical Database using single camera view", IFIP WG 7.6 Working Conference, Noisy le Grand, France, May 28-30 1996.

[COM 1996b] M. MALLEM, M. SHAHEEN, X. DOURILLE and F. CHAVAND "Method of Matching Object 3D model and its image using Geometric Constraints" CESA'96, IEEE-SMC, Lille, France, July 9-12, 1996.

[COM 1995] F. CHAVAND and M. MALLEM "Design of a Man-Machine Interface in Teleoperation : a Real-Virtual approach", Second Intern. Symposium on Methods and Models in Automation and robotics. 30/8-2/9/1995, pp. 539-548, Miedzyzdroje, Poland.

[COM 1994a] F. CHAVAND, M. MALLEM, E.C. N'ZI, J. TRIBOULET "Methods for Updating the Environment Geometric 3D Database in Telerobotics", IEEE/IMACS SPRANN'94, April 25-27 1994, pp. 628-631, Lille, France.

[COM 1994b] M. BENREGUIEG, H. MAAREF, M. MALLEM, C. BARRET "Réalisation d'un suivi de couloir par un robot mobile à l'aide de la logique floue", Intern. conference on Electrical and Electronics Engineering, Oct. 1-3, 1994, pp. 170-176, Oran, Algérie.

[COM 1994c] M. MALLEM, F; CHAVAND, E.C. N'ZI "Multimedia Control Interface in Teleoperation based on Real-Virtual Approach" ORIA'94, December 5-6 1994, pp. 307-318, Marseille, France.

[COM 1993a] M. MALLEM, S. ROUGEAUX, H. MELLANGER, F. CHAVAND
"A Trajectory Generation Module for 2D and 2D 1/2 Environment"
IEEE, SEE, COMPEURO93, International Conference on Computer in Design, Manufacturing and Production, pp. 478-485, Evry, -France-, May 24-27 1993

[COM 1993b] M. MALLEM, E. COLLE and F. CHAVAND
"A Multimedia Control Interface in Teleoperation"
IEEE, SMC'93 Conference: Systems Engineering in the service of humans , vol 8 pp126-131, Touquet -FRANCE- Oct.17-20 1993

[COM 1992] F. CHAVAND, M.MALLEM, E.COLLE, A.LOUKIL
"System for the remote monitoring and control of robotized devices"
International Symposium on Robotics Mechatronics and Manufacturing Systems`92 KOBE, JAPAN Sep. 16-20 1992, pp 1077-1082

[COM 1991a] M. MALLEM, F. CHAVAND
"The perception function in teleoperation"
Second France Israël Symposium, Advances in Non-Manufacturing robotics,
pp. 7.19-7.27, INSTN, CEN-SACLAY, France, April 9-10, 1991

[COM 1991b] M. MALLEM, F. CHAVAND
"Assistances à la perception apportées à un opérateur en robotique téléopérée", European Control Conference (ECC91), pp. 1448-1453, Grenoble -France-, July 2-5 1991

[COM 1991c] F. CHAVAND, M.MALLEM, E.COLLE, A.LOUKIL
"Système d'aide au contrôle-commande à distance de dispositifs robotisés",
4° Symposium international sur l'offshore, la robotique et l'intelligence
artificielle, Marseille, FRANCE, pp. 199-205, 11-12 Decembre 1991

[COM 1990] M. MALLEM, F. CHAVAND and J.P. GAILLARD
"Visual Enhancements for Teleoperation"
Ninth European Annual Conference on HUMAN DECISION MAKING AND MANUAL CONTROL, pp. 1-10, Varese, Italy, September 10-12, 1990

[COM 1989] M. MALLEM et F. CHAVAND
"Aides à la perception en téléopération: image synthétique superposée à l'image vidéo"
Colloque d'échanges de technologie entre milieux critiques:
la téléopération, Marseille, FRANCE, 11-12 Decembre 1989

[COM 1988] F. CHAVAND, E. COLLE, J.P. GAILLARD, M. MALLEM and J.P. STROMBONI
"Visual assistance to the operator in teleoperation "
Symposium of Teleoperation and Control, p. 237-248, Bristol, ENGLAND, Jul. 12-15, 1988

VALORISATION

[BREVET 1991] F. CHAVAND, E. COLLE, M. MALLEM
Procédé et dispositif d'aide à la commande d'engins robotisés lors d'un processus de téléopération effectué en environnement hostile. Brevet N° 91 11160 (dépôt:10 /9/1991, accordé: 17/3/95).