Asservissement visuel de droites par commande découplée et stéréovision

Fadi Alkhalil, Christophe Doignon, Pierre Graebling

▶ **To cite this version:**

Fadi Alkhalil, Christophe Doignon, Pierre Graebling. Asservissement visuel de droites par commande découplée et stéréovision. ORASIS - Congrès des jeunes chercheurs en vision par ordinateur, INRIA Grenoble Rhône-Alpes, Jun 2011, Praz-sur-Arly, France. inria-00595723

HAL Id: inria-00595723

https://inria.hal.science/inria-00595723

Submitted on 25 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asservissement visuel de droites par commande découplée et stéréovision

Fadi ALKHALIL Christophe DOIGNON Pierre GRAEBLING

Université de Strasbourg
Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT)
Equipe Automatique, Vision et Robotique
Pôle API - Bd Sébastien Brant, 67412 Illkirch
{fadi.alkhalil}@lsiit.u-strasbg.fr

Résumé

L'objectif général de ce travail est d'étudier le comportement et les propriétés de découplage d'une commande de robot manipulateur à l'aide d'une mesure par stéréovision. Pour ce faire, nous avons tenu compte des sensibilités différentes des degrés de liberté vis-à-vis des mesures visuelles et nous avons privilégié une représentation minimale des vecteurs d'état. Nous présentons dans cet article un asservissement 3-D basé sur une approche cinématique pour réaliser la commande du déplacement de caméras embarquées observant des droites dans la scène 3-D. Ces indices visuels ont été choisis car ils apparaissent dans de nombreuses applications relevant de la robotique médicale (aiguille de radiothérapie, axe d'un instrument de chirurgie mini-invasive, modélisation d'un bras d'un patient pour l'assistance visuelle à la physiothérapie,...).
En distinguant les cas où la cible correspond à une seule droite (4 ddl) ou à plusieurs droites (6 ddl en général), le torseur cinématique peut être décomposé afin de de réaliser un découplage complet des commandes. Ceci entraîne une réduction des dimensions des fonctions de tâche des sous-systèmes de commande et évite que la plus grande sensibilité aux erreurs de mesures de certaines composantes du torseur cinématique influe sur les autres. Comme cela a été initié dans [2] pour un système monoculaire, nous proposons ici, dans le cas de la stéréovision, une commande basée sur les coordonnées normalisées de Plücker tout en assurant aussi la stabilité globale. Faisant suite à des résultats de simulation à partir de données artificielles, la validation expérimentale a été réalisée (résultats préliminaires) à l'aide d'un robot manipulateur à 6 ddl et de deux caméras industrielles.

Mots Clef

Asservissement visuel, stéréo, commande découplée.

Abstract

The overall objective of this work is to study robotic-based stereo visual feedback with good control behaviour and decoupling properties. To this purpose, the key ideas (however not new) are both to look for a minimal data representation and to take care of degrees of freedom more sensitive to noise in the visual measurements so as to design efficient state vectors. A 3-D visual servoing based on kinematic approach is addressed here. Then, the components of the velocity screw are splitted in subvectors of low dimensions to get a better control of the on-board cameras motion.
In this article, straight lines are used as visual measurements since they are easy to segment from most images captured from man-made environments, while tracking cylinder-shaped surgical instrument in minimally invasive surgery or tracking arms for medical assistance in physiotherapy to get quantitative indicators for patient's assessment. To this end, new tasks functions based on the normalized Plücker representation (see [2]) and a stereo visual control scheme are presented. Considering one (4 DOF) or several straight lines (generally 6DOF) we design velocity control laws with full decoupling so as to improve the convergence behaviour of the servoing.
To assess this work, the control algorithms are firstly performed with simulations data. This allows to estimate the space parameters sensitivity and also to analyze the control behaviours. Finally, an experimental setup made up with two industrial cameras mounted on a 6-DOF robot plateform is used to validate the proposed approach.

Keywords

Visual servoing, stereovision, control with decoupling.

1 Introduction

Dans cet article, nous présentons une commande de robot par stéréovision embarquée. Les mesures visuelles sont constituées de droites qui servent à la définition des vecteurs d'état. La contribution principale de ce travail concerne les lois de commande cinématiques et les fonctions de tâches proposées, qui dans un cadre plus large serviront à réaliser des asservissements visuels sur des objets articulés.

1.1 Généralités et fonction de tâche

L'emploi d'un retour visuel dans le but d'effectuer une commande en boucle fermée de robot s'est largement répandu et concerne de nos jours tous les domaines de la robotique. Le retour permet d'effectuer une comparaison entre une situation désirée et la situation actuelle, issue des mesures visuelles. A cette fin, le formalisme de la fonction de tâche, présentée par Espiau *et al.* [7] il y a vingt ans, est un outil mathématique efficace pour formaliser le retour et établir l'influence mutuelle de plusieurs tâches. La mise en oeuvre d'un asservissement visuel de robot suivant cette approche permet de résoudre de manière rigoureuse le problème de la tâche hybride, combinant plusieurs objectifs à atteindre. Sous ce formalisme, une tâche visuelle peut alors être considérée comme principale et prioritaire. Les autres objectifs (visuels ou non) s'expriment sous la forme de tâches secondaires et évoluent dans le noyau du jacobien de la tâche principale sans effet sur celle-ci. Une hiérarchie des tâches peut alors en découler. Bien que ce formalisme dépasse largement le contexte de la commande référencée vision, de nombreuses techniques d'asservissement visuel y font référence. Elles diffèrent alors selon les mesures visuelles employées, l'espace de représentation des données, la définition des vecteurs d'état et le contexte applicatif. Ce dernier pourra imposer des conditions supplémentaires contraignant un peu plus le déplacement, comme par exemple en robotique mobile (robots non-holonomes et évitement d'obstacles) [23], la robotique domestique (redondance) [14], la chirurgie assistée (sécurité, invasion minimale) ([15]) la micro/meso-robotique [21] (forces de surfaces, d'adhésion, faible profondeur de champ,...).
Il en résulte que la définition de la fonction de tâche est un élément déterminant pour le comportement de la commande, pour filtrer certaines mesures visuelles ou pour construire des observateurs ou simplement pour étudier la convergence à travers une fonction de Lyapunov.

Dans la suite de l'introduction, les motivations et les travaux antérieurs sont exposés. Dans la section suivante (section 2), la modélisation des droites par la représentation de Plücker est abordée pour facilement appréhender dans la section 3 les fonctions de tâches et les lois de commandes. Les résultats sont analysés dans la section 4 et nous terminons par une conclusion accompagnée des perspectives envisagées (section 5).

1.2 Motivations

La classification des asservissements visuels (AV) est constituée essentiellement de trois catégories : les asservissements basés position (AV3D), basés image (AV2D) et hybrides (2 1/2 D). C'est à la première catégorie que cet article se réfère, pour laquelle la pose de la caméra par rapport à l'objet 3-D d'intérêt doit être estimée. Pour de plus amples détails, nous invitons le lecteur à se référer à [13, 5, 6]. L'avantage souvent mentionné de cette catégorie d'asservissement est d'offrir en théorie une stabilité globale, lorsque tous les paramètres sont parfaitement estimés (pose, paramètres intrinsèques, modèle de l'objet en vision monoculaire). Cependant, même si on peut étalonner très précisément le capteur visuel, des erreurs dans l'estimation de la pose auront un impact sur la précision du positionnement, voire sur la convergence même de l'asservissement.
Il est tout aussi connu que le principal inconvénient des AV3D est de ne pas assurer la visibilité des indices visuels puisque, pour cette catégorie d'AV, les fonctions de tâches sont liées linéairement à des vecteurs d'état composés de grandeurs 3-D. Par ailleurs, un des problèmes encore à l'étude est de parvenir à découpler des composantes de la commande (même si plusieurs solutions ont été apportées dans [5, 17] pour des indices visuels constitués de points géométriques), afin d'offrir le meilleur positionnement possible. De plus, le cas de la stéréovision est singulier car si ce même découplage peut être envisagé pour les AV3D par la stéréo [4], aucun modèle géométrique de l'objet d'intérêt est requis, au contraire, il peut être construit en ligne. Il est par conséquent pertinent d'établir le découplage entre les composantes qui n'ont pas la même sensibilité (d'un ordre de grandeur différent) aux incertitudes sur les mesures visuelles afin de rendre ces deux processus (asservissement et reconstruction) plus performants.

1.3 Etat de l'art sur les AV par stéréovision

La littérature concernant les asservissements par stéréovision (AVS) n'est pas aussi riche que pour les AV n'utilisant qu'une seule caméra. Au-delà du recouvrement important des méthodes qui s'y réfèrent, les raisons sont à rechercher dans l'important flux de données à acquérir et à traiter en temps réel, les problèmes techniques de synchronisation des acquisitions. Malgré ces difficultés pratiques, de nombreuses applications ne requièrent que l'emploi de capteurs monoculaires, en particulier lorsque l'environnement peut être contrôlé (fréquent dans l'industrie de la production). De plus, quand ce dernier est peu maitrisé, voire inconnu, plusieurs points de vues avec un seul capteur en mouvement permettent encore d'effectuer correctement un positionnement relatif, sous quelques hypothèses (SLAM, structure par le mouvement). Comme la profondeur peut être estimée avec un système stéréo sans faire appel à un modèle géométrique de l'objet d'intérêt (information qui, en monoculaire, intervient dans les premières composantes d'un jacobien image d'un AV2D ou dans le calcul de la pose pour un AV3D) et qu'un gain sur la précision d'un positionnement, sur un déplacement sans configuration singulières et sur l'interprétation des mesures visuelles peuvent être attendus avec un système de guidage par stéréovision, nous allons présenter quelques travaux antérieurs et clarifier ainsi le positionnement de ce travail.
Andersson [1] a sans doute réalisé un des premiers système de commande de robot par stéréovision. L'application est dédiée à un robot jouant au tennis de table. Le système de vision localise l'image d'une balle de ping-pong par une segmentation de la couleur qui se fait localement

grâce à l'apport d'un modèle dynamique de la trajectoire de la balle. Le système est calibré et utilise une loi de commande basée position. Graffunder *et al.* [8] ont effectué la simulation de la pose et du contrôle du déplacement d'une pince de robot à l'aide d'un système stéréo embarqué. La commande a été conçue pour être appliquée directement sur les articulations et les résultats de simulation ont validé la stabilité asymptotique de l'asservissement. Hollinghurst et Cipolla [12] ont réalisé un AVS utilisant des caméras déportées et un robot manipulateur à 5 DDL. Leur technique fait appel à un modèle affine de la projection perspective pour estimer les coordonnées cartésiennes des positions et orientations de l'organe terminal et de sa situation désirée par rapport à un objet polyédrique. La calibration est réalisée en observant quatre positions de référence sur le robot. Bien que le type d'objet d'intérêt à manipuler soit de géométrie simple et qu'une hypothèse de planarité est utilisée, la contribution de ces travaux est d'estimer automatiquement la pose et de réaliser la préhension de la cible. Une plateforme d'AVS a été réalisée par Hager *et al.* [10, 9]. Utilisant un système stéréovision déporté qui ne nécessite pas une calibration fine entre les capteurs et le robot, il permet le contrôle de ce dernier en utilisant des indices visuels, tels que des droites, des régions, des coins dans l'image trackés par le framework XVision. Pour son application, ce système permet le suivi visuel de l'organe terminal du robot et des objets d'intérêt et calcule une fonction d'erreur à minimiser. Il y a environ dix ans, Lamiroy *et al.* [16] ont introduit la contrainte épipolaire dans les AVS et plus récemment Cervera *et al.* [4] ont étudié trois approches de l'AVS par loi de commande cinématique pour des points (directement sur les images, basée sur la disparité, et par reconstruction et commande 3-D) appliquée à un système de vision embarqué sur un robot à 6 DDL. Dans ce travail, la deuxième et la troisième méthode font clairement la différence par rapport à la commande basée image car l'utilisation d'indices visuels 3-D permet d'établir des matrices jacobiennes linéaires et offre une plus grande précision (et aucune singularité) ainsi qu'un découplage.

2 Représentation géométrique et projection de droites

La modélisation géométrique d'une droite dans l'espace 3-D, sa projection dans le plan image d'une caméra perspective ainsi que sa reconstruction à partir d'une paire de vues sont rappelées dans cette section. La notation $\equiv$ employée par la suite s'entend pour une égalité à un facteur près.

2.1 Modélisation géométrique

Etant donné une droite L, ses coordonnées de Plücker sont constituées par un couple de vecteurs de dimension 3, $(\mathbf{v},\mathbf{w})$ non linéairement dépendants par $\mathbf{v}^\mathsf{T}\mathbf{w} = 0$. $\mathbf{v}$ est un vecteur renfermant la direction de la droite et $\mathbf{w}$ est un second vecteur dont la direction est perpendiculaire au plan d'interprétation ($\boldsymbol{\pi}^L$) qui contient la droite et l'origine du repère de référence. Ainsi, pour tout couple de points $\mathbf{M}$ et $\mathbf{N}$ sur L et représentés par leurs coordonnées homogènes, une matrice (4×4), $\mathtt{L}$, appelée matrice de Plücker, est construite selon la définition suivante :

$$\mathtt{L} \equiv \mathbf{M}\mathbf{N}^\mathsf{T} - \mathbf{N}\mathbf{M}^\mathsf{T} \equiv \begin{bmatrix} [\mathbf{w}]_\times & -\mathbf{v} \\ \mathbf{v}^\mathsf{T} & 0 \end{bmatrix} \qquad (1)$$

et est indépendante de la position des deux points sur la droite, puisque ($\mathbf{v} \equiv \mathbf{M} - \mathbf{N}$ et $\mathbf{w} \equiv \mathbf{M} \times \mathbf{N}$). C'est une représentation pratique puisqu'il est très simple de manipuler les transformations géométriques. En effet, à une transformation représentée génériquement par une matrice (4×4) $\mathtt{H}$, la matrice de Plücker transformée $\mathtt{L}'$ s'exprime par $\mathtt{L}' \equiv \mathtt{H}\,\mathtt{L}\,\mathtt{H}^\mathsf{T}$ et les coordonnées de Plücker correspondants $\mathbf{v}', \mathbf{w}', \mathbf{v}, \mathbf{w}$ sont linéairement reliées [3]. La matrice duale de Plücker, $\mathtt{L}^\star$, correspond à l'intersection de deux plans $\boldsymbol{\pi}_1$ et $\boldsymbol{\pi}_2$ et représente la même droite 3-D puisque point et plan sont des grandeurs duales dans l'espace 3-D. Elle a comme expression :

$$\mathtt{L}^\star \equiv \boldsymbol{\pi}_1\,(\boldsymbol{\pi}_2)^\mathsf{T} \;-\; \boldsymbol{\pi}_2\,(\boldsymbol{\pi}_1)^\mathsf{T} \equiv \begin{bmatrix} [\mathbf{v}]_\times & -\mathbf{w} \\ \mathbf{w}^\mathsf{T} & 0 \end{bmatrix} \qquad (2)$$

Les matrices $\mathtt{L}$ et $\mathtt{L}^\star$ sont anti-symétriques, homogènes, singulières et de rang 2. En calculant le polynôme caractéristique de $\mathtt{L}$, on peut facilement montrer que les valeurs propres son égales à $\{\pm i\,\mu,\ 0,\ 0\}$ où $\mu = \sqrt{\mathbf{v}^\mathsf{T}\mathbf{v} \;+\; \mathbf{w}^\mathsf{T}\mathbf{w}}$ et μ peut être arbitrairement fixé à une valeur quelconque non nulle pour normaliser $\mathtt{L}$. En particulier, quand $\mathtt{L}$ est normalisée pour que le vecteur $\mathbf{v}$ soit unitaire dans un espace Euclidien, alors la norme du vecteur $\mathbf{w}$ correspond à la distance orthogonale d séparant la droite de l'origine du repère de référence.

2.2 Projection perspective d'une droite

La projection perspective de L est la droite image $\mathbf{l}$ définie par $[\mathbf{l}]_\times \equiv \mathtt{P}\,\mathtt{L}\,\mathtt{P}^\mathsf{T}$, où $\mathtt{P}$ est la matrice caméra (de dimensions (3×4)) et $[\mathbf{l}]_\times$ est la matrice anti-symétrique associée au vecteur $\mathbf{l}$.

En considérant la projection inverse (Fig. 1-a), le plan d'interprétation est alors défini par $\boldsymbol{\pi}^L \equiv (\mathtt{P})^\mathsf{T}\,\mathbf{l}$. L'estimation des coordonnées de Plücker peut être obtenue par l'emploi d'une paire de vues (indicées par r et l) en utilisant l'intersection des plans d'interprétation des projections perspectives inverses des droites image $(\mathbf{l}^r, \mathbf{l}^l)$. On a alors $\boldsymbol{\pi}^r \equiv (\mathtt{P}^r)^\mathsf{T}\,\mathbf{l}^r$ et $\boldsymbol{\pi}^l \equiv (\mathtt{P}^l)^\mathsf{T}\,\mathbf{l}^l$.

Au regard de la définition de $\mathtt{L}^\star$ (2), ceci permet de retrouver les coordonnées de Plücker pourvu que les matrices caméras $\mathtt{P}^r$ et $\mathtt{P}^l$ soient exprimées dans un même repère. De cette manière, les vecteurs $\mathbf{v}$ et $\mathbf{w}$ sont identifiés (au signe près de l'un des deux car $(\mathbf{v}, \mathbf{w})$ et $(\mathbf{v}, -\mathbf{w})$ ne correspondent pas à la même matrice !). En notant et décomposant les matrices caméra tel que $\mathtt{P}^r = [\mathtt{M}^r \;\; \boldsymbol{\sigma}^r]$ et $\mathtt{P}^l = [\mathtt{M}^l \;\; \boldsymbol{\sigma}^l]$ où $\mathtt{M}$ est une matrice (3×3) et où $\boldsymbol{\sigma}$ est la dernière colonne de $\mathtt{P}$, on a :

$$\mathbf{v} \;\equiv\; \mathtt{M}^{r\mathsf{T}}\mathbf{l}^r \;\times\; \mathtt{M}^{l\mathsf{T}}\mathbf{l}^l \qquad (3)$$

$$\mathbf{w} \;\equiv\; \mathtt{M}^{l\mathsf{T}}\mathbf{l}^l\mathbf{l}^{r\mathsf{T}}\mathtt{M}^r\boldsymbol{\sigma}^r \;-\; \mathtt{M}^{r\mathsf{T}}\mathbf{l}^r\mathbf{l}^{l\mathsf{T}}\mathtt{M}^l\boldsymbol{\sigma}^l\,. \qquad (4)$$

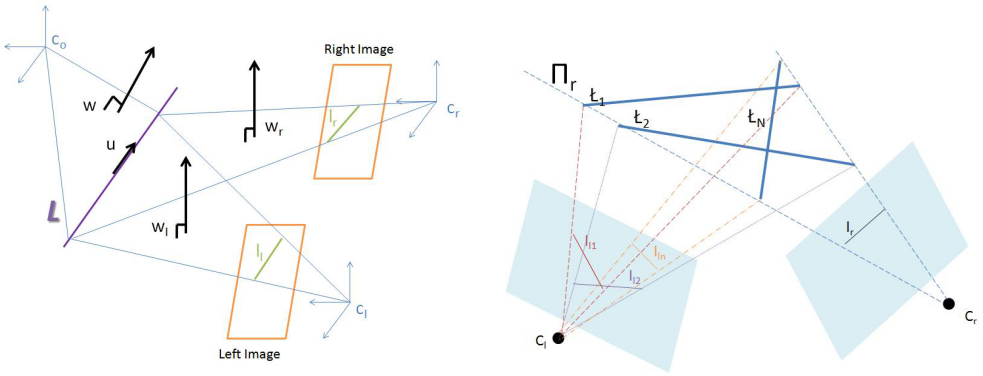

FIGURE 1 – (a) La projection perspective d'une droite 3-D L et ses coordonnées de Plücker $(\mathbf{v}, \mathbf{w})$. (b) L'orientation de la droite 3-D dans le plan d'interprétation n'est pas unique ($L_1, L_2, ...L_N$ sont aussi solutions) à partir d'une seule projection ($\mathbf{l}_r$). L'unicité de l'orientation est obtenue avec une seconde vue ($\mathbf{l}_l$) qui permet aussi d'estimer la distance orthogonale.

3 Asservissement visuel stéréo

3.1 Analyse du mouvement de droites 3-D

L'analyse cinématique de droites 3-D a été étudiée par plusieurs auteurs, dans l'objectif d'effectuer un suivi visuel sur cet indice apparaissant dans les scènes urbaines ou d'intérieur. Les approches proposées ont permis d'établir les relations mathématiques entre les variations des droites apparentes à partir d'une ou plusieurs vues et le déplacement 3-D correspondant [18]. Rives *et al.* [20] et Navab [19] se sont intéressés à la modélisation par la représentation de Plücker et ont établi les relations cinématiques du champ de mouvement (*motion field* - pour Rives) et les relations géométriques du déplacement (*structure from motion* - à l'aide de trois vues) pour Navab. Plus récemment Andreff [2] a défini la représentation de Plücker binormalisée dans un contexte Euclidien ($\mathbf{w}_n = \mathbf{w}/||\mathbf{w}||, \mathbf{v}, d = ||\mathbf{w}||$) puis développé des propriétés qui en découlent, en particulier l'orientation propre de la droite dans le plan d'interprétation (qui ne modifie pas son apparence dans l'image, voir la Fig. 1-b) mais contribue à l'obtention d'un alignement correct. La localisation d'une droite est ainsi contraint par trois éléments bien identifiés : son image (qui fournit la direction de $\mathbf{w}$), l'alignement dans le plan d'interprétation (qui avec la condition $\mathbf{v}^\mathsf{T}\mathbf{w}$ définit la direction de $\mathbf{v}$) et sa distance orthogonale d (parfois appelée profondeur - $d = ||\mathbf{w}||$). Ceci a permis à Andreff de proposer, pour la vision monoculaire, plusieurs schémas de commandes cinématiques pour lesquels la matrice d'interaction exhibe un découplage partiel entre la vitesse de translation $\boldsymbol{\nu}$ et la vitesse de rotation $\boldsymbol{\Omega}$ de la caméra durant son déplacement, selon les relations suivantes :

$$\begin{pmatrix} \dot{\mathbf{w}}_n \\ \dot{\mathbf{v}} \\ \dot{d} \end{pmatrix} = \begin{pmatrix} -\frac{1}{d}(\mathbf{v} \times \mathbf{w}_n)\, \mathbf{w}_n^\mathsf{T} & -[\mathbf{w}_n]_\times \\ \mathbf{0}_{3\times 3} & -[\mathbf{v}]_\times \\ (\mathbf{v} \times \mathbf{w}_n)^\mathsf{T} & \mathbf{0}^\mathsf{T} \end{pmatrix} \boldsymbol{\tau}^c \tag{5}$$

$\boldsymbol{\tau}^{c\mathsf{T}} = (\boldsymbol{\nu}^{c\mathsf{T}}, \boldsymbol{\Omega}^{c\mathsf{T}})$ est le torseur cinématique de la caméra exprimé dans le repère de la caméra (R_c), et à son origine. Nous nous appuirons par la suite sur cette equation (equation 5) pour décrire la méthode que nous proposons.

3.2 Fonction de tâche $\mathbf{e}_\nu$ pour la commande en position (*profondeur)*

En considérant la dernière ligne de (5) $\dot{d} = \boldsymbol{\nu}^{c\mathsf{T}}(\mathbf{v} \times \mathbf{w}_n)$, on observe que tant que la profondeur d n'est pas asservie complètement, il existe une contribution (composante) de la vitesse de translation $\boldsymbol{\nu}^{c\mathsf{T}}$ selon la direction $(\mathbf{v} \times \mathbf{w}_n)$. Par conséquent, et sous réserve de convergence, il est possible d'imposer que la commande $\boldsymbol{\nu}^c$ se fasse uniquement selon cette direction, ce qui revient à aligner $\boldsymbol{\nu}^c$ avec $(\mathbf{v} \times \mathbf{w}_n)$ et dans ce cas $\mathbf{v}^\mathsf{T}\boldsymbol{\nu}^c = 0$ et $\mathbf{w}_n^\mathsf{T}\boldsymbol{\nu}^c = 0$.

Proposition : Si $d^\star$ représente la distance orthogonale (ou profondeur) désirée, alors la fonction de tâche $\mathbf{e}_\nu(t)$ définie par

$$\mathbf{e}_\nu(t) = \mathbf{C}_\nu(t)\,(d(t) - d^\star) \tag{6}$$

où $\mathbf{C}_\nu(t) = \mathbf{v}(\mathbf{t}) \times \mathbf{w}_n(t)$ est la matrice de combinaison, fournit une loi de commande cinématique globalement et asymptotiquement stable, $\boldsymbol{\nu}^c = -\lambda_\nu \mathbf{e}_\nu$, pourvu que le scalaire (gain) λ_ν soit réel positif.

Preuve : le choix proposé ci-dessus pour la commande $\boldsymbol{\nu}^c$ simplifie également la première ligne de (5), c'est-à-dire que l'on a aussi $\dot{\mathbf{w}}_n = \boldsymbol{\Omega} \times \mathbf{w}_n$. La cinématique "naturelle" de la droite est donc contraint, car à une variation de la direction du plan d'interprétation ($\dot{\mathbf{w}}_n$) aurait dû normalement correspondre une contribution sur la vitesse de translation ; celle-ci est d'autant plus faible que d est grand et est donc annulée ici par la commande.
Pour établir la stabilité de cet asservissement, une fonction basée sur la norme de $\mathbf{e}_\nu$ est choisie comme fonction de Lyapunov $\mathcal{L}(t) = 1/2||\mathbf{e}_\nu(t)||^2$. En supposant que l'objet

d'intérêt est statique (en profondeur), $\dot{d}^\star = 0$), la dérivée temporelle première de la fonction de tâche vaut

$$\begin{aligned}
\dot{\mathbf{e}}_\nu &= \mathbf{C}_\nu\, \dot{d} + (\dot{\mathbf{v}} \times \mathbf{w}_n + \mathbf{v} \times \dot{\mathbf{w}}_n)(d - d^\star) \\
&= \mathbf{C}_\nu\, \dot{d} + ((\boldsymbol{\Omega} \times \mathbf{v}) \times \mathbf{w}_n + \mathbf{v} \times (\boldsymbol{\Omega} \times \mathbf{w}_n))(d - d^\star) \\
&= \mathbf{C}_\nu\, \dot{d} + ((\mathbf{w}_n^\mathsf{T}\boldsymbol{\Omega})\mathbf{v} - (\mathbf{v}^\mathsf{T}\boldsymbol{\Omega})\mathbf{w}_n)(d - d^\star) \\
&= \mathbf{C}_\nu\, \dot{d} + \boldsymbol{\Omega} \times (\mathbf{v} \times \mathbf{w}_n)(d - d^\star) \\
&= \mathbf{C}_\nu\, \dot{d} + \boldsymbol{\Omega} \times \mathbf{e}_\nu
\end{aligned}$$

$$\begin{aligned}
\dot{\mathcal{L}}(t) &= \mathbf{e}_\nu^\mathsf{T}\dot{\mathbf{e}}_\nu = \mathbf{C}_\nu^\mathsf{T}\mathbf{C}_\nu\, \dot{d}\,(d - d^\star) + \overbrace{\mathbf{e}_\nu^\mathsf{T}(\boldsymbol{\Omega} \times \mathbf{e}_\nu)}^{=0} \\
&= \dot{d}\,(d - d^\star) \quad (\text{car } ||\mathbf{C}_\nu|| = 1) \\
&= \boldsymbol{\nu}^{c\mathsf{T}}\, \mathbf{C}_\nu\, (d - d^\star) \;=\; -\lambda_\nu\, \mathbf{e}_\nu^\mathsf{T}\, \mathbf{C}_\nu\, (d - d^\star) \\
&= -\lambda_\nu\, (d(t) - d^\star)^2
\end{aligned}$$

est négative pour tout $t \in \mathbb{R}^+$ si $\lambda_\nu \in \mathbb{R}^{\star +}$.

Remarques :

1. Avec cette stratégie pour établir la commande $\boldsymbol{\nu}^c$, on peut noter que si la norme de la fonction de tâche n'est pas influencée par la commande en orientation, la direction du vecteur $\mathbf{e}_\nu$ peut en dépendre.
2. Il est possible de choisir une matrice (ou vecteur) de combinaison constante, comme $\mathbf{C}_\nu^\star$ (pour les coordonnées désirées de Plücker $\mathbf{v} = \mathbf{v}^\star$ et $\mathbf{w}_n = \mathbf{w}_n^\star$) par exemple. Cependant, la mise à jour de la matrice de combinaison permet de tenir compte de l'évolution éventuelle de l'orientation de la droite et d'asservir la profondeur toujours selon la direction orthogonale, ce qui amène à l'obtention d'une trajectoire minimale.

3.3 Fonction de tâche $\mathbf{e}_\omega$ pour la commande en rotation

Dans cette partie, nous abordons l'asservissement en orientation comme un problème d'alignement grâce aux propriétés issues de la représentation par les coordonnés binormalisées de Plücker. L'asservissement en orientation consiste alors à amener le référentiel actuel du système stéréo (R_s) vers le référentiel désiré $(R_{s^\star})$. De par le choix de la commande en position, il est possible de commander l'orientation à partir des changements de direction des deux vecteurs $\mathbf{v}$ and $\mathbf{w}_n$ car $\boldsymbol{\Omega} \times \mathbf{w}_n = \dot{\mathbf{w}}_n$ et $\boldsymbol{\Omega} \times \mathbf{v} = \dot{\mathbf{v}}$. Une tâche d'alignement de droites peut alors être résolue simplement sous la forme d'une commande purement en orientation. Pour se faire, la matrice de rotation ${}^{s^\star}\mathbf{R}_s$ est calculée à partir de deux paires de coordonnées de Plücker. Tout d'abord, la matrice de rotation ${}^s\mathbf{R}_L$ correspondant à l'orientation du référentiel sur la droite (R_L) par rapport au référentiel courant du système de vision stéréo (R_s) est calculée à partir des estimations de $\mathbf{v}(t)$, $\mathbf{w}_n(t)$ et $\mathbf{v}(t) \times \mathbf{w_n}(t)$ qui forment ses colonnes. Puis la matrice de rotation ${}^{s^\star}\mathbf{R}_L$ issue de l'orientation de (R_L) par rapport au référentiel désiré du système de vision $(R_{s^\star})$ est calculée à partir des estimations de $\mathbf{v}^\star$, $\mathbf{w}_n^\star$ et de $\mathbf{v}^\star \times \mathbf{w}_n^\star$. Finalement, avec le produit ${}^{s^\star}\mathbf{R}_s = {}^L\mathbf{R}_{s^\star}^\mathsf{T}\, {}^L\mathbf{R}_s$, la représentation angle/axe $\theta\ \mathbf{u}$ fournit directement l'orientation relative, et la loi de commande déjà présentée dans [17] pour des points peut s'appliquer ici aussi. La commande en orientation suivante $\boldsymbol{\Omega} = -\lambda_\omega\, \mathbf{e}_\omega = -\lambda_\omega\, \theta\, \mathbf{u}$ est globalement et asymptotiquement stable (λ_ω positif).

3.4 Asservissement à partir de N droites

Dans les deux parties précédentes, nous avons considéré l'asservissement visuel stéréo à partir d'une seule droite en estimant les 4 ddl qui la définissent dans l'espace 3-D, puis en proposant un système commandable en présentant deux lois de commandes cinématiques. Bien que le torseur complet ait été calculé, on peut montrer que les composantes de la vitesse de translation, $\boldsymbol{\nu}^c = (\nu_x^c, \nu_y^c, \nu_z^c)^\mathsf{T}$, sont linéairement liées (en exprimant par exemple ce torseur dans le référentiel (R_L) de la droite).
En abordant ici le cas d'un objet d'intérêt de la scène composé de plusieurs droites rigidement liées, les coordonnées de Plücker ne sont pas libres, mais il est possible de les combiner pour établir une commande du mouvement de l'ensemble. Nous allons, dans les deux paragraphes suivants, décrire les deux schémas de commande que nous proposons dans ce cas.

Commande en position pour N droites. Faisons l'hypothèse dans ce paragraphe que la matrice de rotation ${}^{s^\star}\mathbf{R}_s$ entre (R_s) et $(R_{s^\star})$ ait été préalablement estimée et soit connue. Nous nous intéressons dans un premier temps à déterminer une solution pour le vecteur de position ${}^{s^\star}\mathbf{t}_s$ entre les origines des deux référentiels (courant et désiré) du système de vision stéréo embarqué. Etant donné que la matrice de passage entre les coordonnées de Plücker actuelles $(\mathbf{w}(t), \mathbf{v}(t))$ et celles du référentiel désiré $(\mathbf{w}(\infty), \mathbf{v}(\infty))$ s'exprime par

$$\begin{pmatrix} \mathbf{w}_j(\infty) \\ \mathbf{v}_j(\infty) \end{pmatrix} = \begin{pmatrix} {}^{s^\star}\mathbf{R}_s & [{}^{s^\star}\mathbf{t}_s]_\times\, {}^{s^\star}\mathbf{R}_s \\ \mathbf{0} & {}^{s^\star}\mathbf{R}_s \end{pmatrix} \begin{pmatrix} \mathbf{w}_j(t) \\ \mathbf{v}_j(t) \end{pmatrix} \tag{7}$$

avec $j = 1, ..., N$ pour N droites. De l'équation (7), un système linéaire est formé $\mathbf{H}\, {}^{s^\star}\mathbf{t}_s = \mathbf{Q}$ où ${}^{s^\star}\mathbf{t}_s$ est le vecteur inconnu et où la matrice $(3N \times 3)$ $\mathbf{H}$ et le vecteur $(3N \times 1)$ $\mathbf{Q}$ ont les expressions suivantes :

$$\begin{aligned}
\mathbf{H} &= -\begin{pmatrix} [\mathbf{v}_1(\infty)]_\times & [\mathbf{v}_2(\infty)]_\times & \cdots & [\mathbf{v}_N(\infty)]_\times \end{pmatrix}^\mathsf{T}, \\
\mathbf{Q} &= \begin{pmatrix} \mathbf{q}_1^\mathsf{T} & \mathbf{q}_2^\mathsf{T} & \cdots & \mathbf{q}_N^\mathsf{T} \end{pmatrix}^\mathsf{T}, \\
\mathbf{q}_j &= {}^{s^\star}\mathbf{R}_s\, \mathbf{w}_j(t) \,-\, \mathbf{w}_j(\infty)\,.
\end{aligned}$$

Le calcul du produit matriciel $\mathbf{H}^\mathsf{T}\mathbf{H}$ révèle les cas de dégénérescence puisque

$$\begin{aligned}
\mathbf{H}^\mathsf{T}\mathbf{H} &= -[\mathbf{v}_1(\infty)]_\times^2 - [\mathbf{v}_2(\infty)]_\times^2 - ... - [\mathbf{v}_N(\infty)]_\times^2 \quad (8) \\
&= N\,\mathbf{I} \,-\, \textstyle\sum_{j=1}^{j=N} \mathbf{v}_j(\infty)\, \mathbf{v}_j(\infty)^\mathsf{T} \quad (9)
\end{aligned}$$

est de rang 3 si $N \geq 2$ et s'il existe un vecteur $\mathbf{v}_k$ de l'ensemble tel que $\mathbf{v}_k \times \mathbf{v}_j \neq 0$ (au moins deux droites

ne sont pas parallèles). En particulier, pour $N = 2$, $\det(\mathbf{H}^\mathsf{T}\mathbf{H}) = 2||\mathbf{v}_1 \times \mathbf{v}_2||^2$.

Une solution peut être obtenue directement par la méthode des moindres carrés. Par ailleurs, les composantes du système linéaire précédent nécessitent l'estimation au préalable des coordonnées de Plücker (dans $\mathbf{H}$ et $\mathbf{Q}$), qui, elles-mêmes sont estimées à partir des mesures visuelles d'une paire de vues. Ainsi, par rapport à l'approche classique de l'asservissement visuel basé position et monoculaire, ici nous ne disposons pas de modèles (supposés exacts) de la géométrie de l'objet d'intérêt mais nous disposons avec les coordonnées de Plücker de primitives visuels 3-D estimées. Bien qu'une approche basée primitive tente d'asservir toutes les primitives, dans le cas de plusieurs droites un problème de commandabilité apparaît avec la redondance qu'elles renferment [11].
Finalement, la fonction de tâche emploie directement le vecteur de position relative $\mathbf{e}_\nu = {}^{s^\star}\mathbf{t}_s$, et, en imposant un comportement exponentiel décroissant de $\mathbf{e}_\nu$, la commande de la vitesse de translation est $\boldsymbol{\nu}^c = -\lambda_\nu \, {}^{s^\star}\mathbf{R}_s^\mathsf{T} \, \mathbf{e}_\nu$.

Commande en orientation pour N droites. Avant de décrire notre solution, nous rappelons un important lemme que l'on doit à Umeyama [22] et qui fournit une solution par les moindres carrés pour l'estimation des composantes de la rotation entre deux ensembles de points (rigides) à recaler dans un espace de dimension m.
Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices $(m \times n)$, $\mathbf{R}$ une matrice $(m \times m)$ de rotation et UDV^T la décomposition en valeurs singulières de $\mathbf{AB}^\mathsf{T}$ ($UU^\mathsf{T} = VV^\mathsf{T} = \mathbf{I}$, D =diag(d_i), $d_1 \geq d_2 \geq \cdots \geq d_m \geq 0$), alors le minimum de la quantité $||\mathbf{A} - \mathbf{RB}||^2$ en fonction de $\mathbf{R}$ est

$$\min_R ||\mathbf{A} - \mathbf{RB}||^2 = ||\mathbf{A}||^2 + ||\mathbf{B}||^2 - 2\, tr(DS) \quad (10)$$

avec

$$S = \begin{cases} \mathbf{I} & \text{si } \det(U)\det(V) = 1, \\ \text{diag}(1, \cdots, 1, -1) & \text{si } \det(U)\det(V) = -1 \end{cases} \quad (11)$$

Preuves et détails peuvent être trouvés dans [22]. Bien que les points soient les entitées géométriques utilisées dans [22], en considérant la dernière ligne de (7), les deux matrices $\mathbf{A}$ et $\mathbf{B}$ peuvent être construites aussi pour des droites car la matrice de rotation fait appel aux vecteurs directeurs $\mathbf{v}_j$ uniquement. $\mathbf{A} = (\mathbf{v}_1(\infty)^\mathsf{T}, \cdots, \mathbf{v}_N(\infty)^\mathsf{T})^\mathsf{T}$ et $\mathbf{B} = (\mathbf{v}_1(t)^\mathsf{T}, \cdots, \mathbf{v}_N(t)^\mathsf{T})^\mathsf{T}$ amenant à une solution pour la rotation ${}^{s^\star}\mathbf{R}_s$. Au-delà de la possibilité qu'offre le lemme ci-dessus pour trouver une solution linéaire unique pour la matrice de rotation, le minimum du critère (10) fournit l'erreur résiduelle, qui permettra par la suite d'incorporer une estimation robuste initiée par cette solution et tenir compte ainsi des mesures abérrantes, des erreurs de correspondance ou de l'absence de certaines données de mesures (gestion des occultations) des droites. Finalement, la fonction de tâche et la commande en position sont identiques à celles déjà décrites pour une seule droite.

4 Résultats

4.1 Simulations

Pour la simulation, le robot est représenté par un intégrateur pur, ce qui permet de simuler le déplacement des caméras par des relations cinématiques. Pour les résultats présentés ici, nous supposons que le système est parfaitement calibré, car notre objectif est de tester la convergence des lois de commande issues des fonctions de tâches ($\mathbf{e}_\nu$ et $\mathbf{e}_\omega$) proposées dans la section précédente. Les courbes sur la figure 2 montrent le comportement de l'AV stéréo à travers les vitesses de rotation (Fig. 2-a,c) et de translation (Fig. 2-b,d) dans le cas d'une droite (Fig. 2-a,b) ou de deux droites (Fig. 2-c,d). On observe une décroissance de toutes les composantes du torseur cinématique.

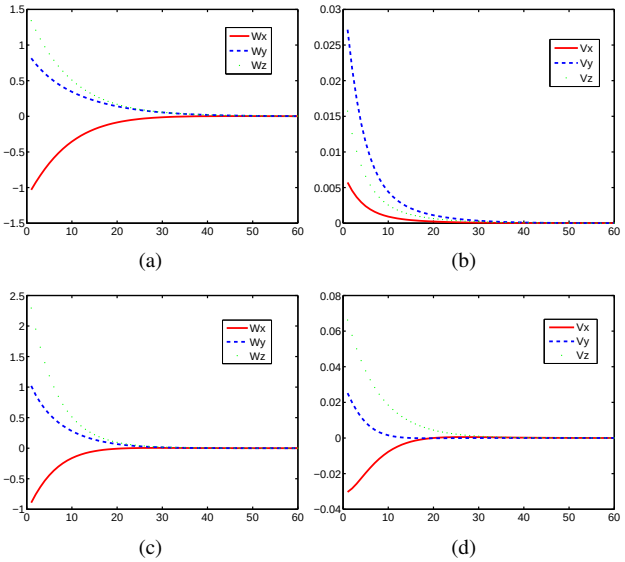

FIGURE 2 – Simulations des AV stéréo à partir d'une droite (a,b) ou de deux droites coplanaires (c,d). Vitesses de rotation (deg/s) (a,c) et de translation (m/s) (b,d).

4.2 Validation expérimentale

Pour mener les expériences, nous avons utilisé un robot manipulateur Adept Viper S650 à 6 DDL. Le contrôleur du robot de Cerebellum Automation permet une commande en position et en vitesse, et l'interface avec le PC de vision et un PC de supervision/configuration est assuré par une liaison firewire au niveau matériel et par les outils de développement et bibliothèques CIDE.
Le système de vision est constitué de deux caméras USB 2.0 SUMIX SMX-150M de résolution maximum de (1280×1024) et utilisées ici à la resolution de (600×400) pour acquérir 40 images/s. Il est fixé rigidement à l'organe terminal du robot (voir Fig. 3) et dans une configuration où les caméras sont parallèles. Deux séries d'expériences ont été effectuées, la première pour un AV à partir d'une seule droite, la seconde pour commander les 6 DDL du robot avec deux droites coplanaires. La tâche robotique

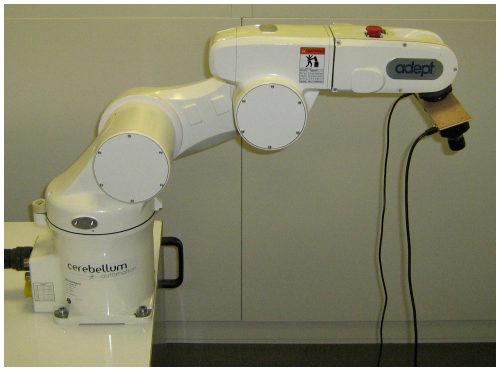

FIGURE 3 – Le système de vision stéréo embarqué sur un robot manipulateur.

consiste alors à contrôler le déplacement l'organe terminal à une position et à une orientation désirée. Des gains de 0.05 ont été appliqués pour λ_ν et pour λ_ω. Pour ces premiers résultats, nous avons supposé que les droites restent visibles par les deux caméras embarquées.

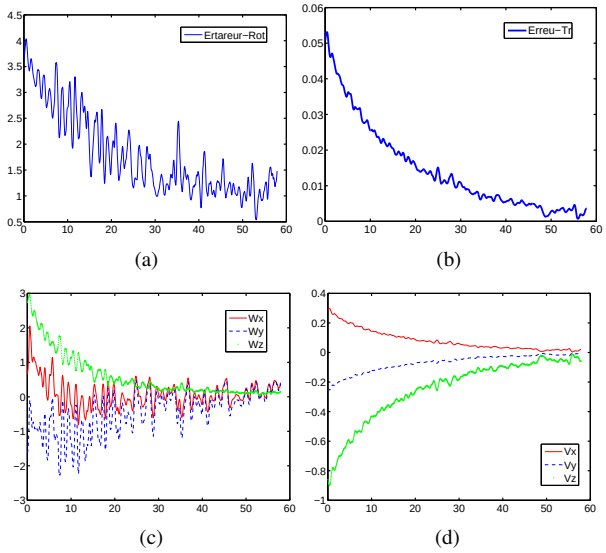

FIGURE 4 – AVS sur une seule droite. (a) Erreur en orientation (deg). (b) Erreur en position (m). (c) Vitesses de rotation (deg/s). (d) Vitesses de translation (m/s).

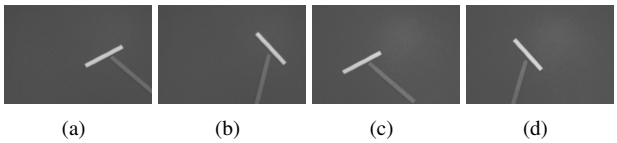

FIGURE 5 – Images initiales et finales. (a-b) : pour la caméra de gauche, (c-d) pour la caméra de droite.

Pour la première série, comme l'objet d'intérêt consiste en une seule droite, seuls 4 ddl sont contraints par la

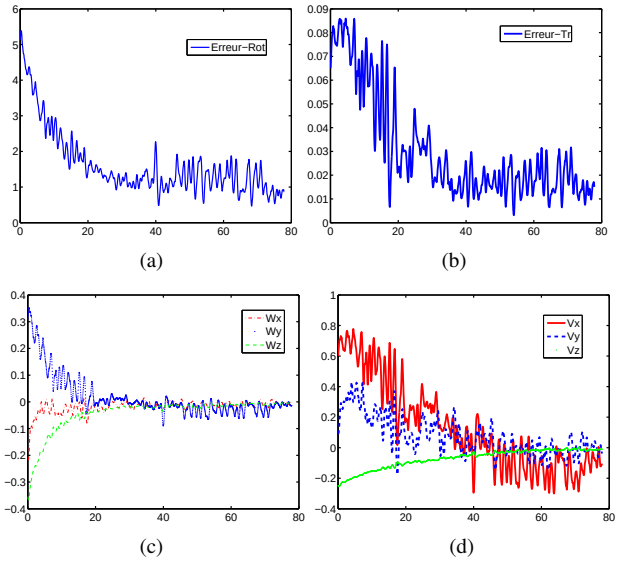

FIGURE 6 – AVS sur deux droites. (a) Erreur en orientation (deg). (b) Erreur en position (m). (c) Vitesses de rotation (deg/s). (d) Vitesses de translation (m/s).

vision (1 ddl de position, 3 ddl d'orientation). Néanmoins, la commande en vitesse met à contribution l'ensemble des articulations du robot (par l'utilisation du jacobien inverse) bien qu'elle s'exprime dans l'espace opérationnel de l'organe terminal.

Les figures 4-a,b montrent l'obtention de la convergence asymptotique vers zéro et l'allure exponentielle décroissante des composantes de la vitesse de translation (Fig. 4-d). Ceci est moins évident pour la vitesse de rotation (Fig. 4-c) où la composante Ω_z évolue peu. De plus, nous n'assurons pas de découplage entre les composantes de la vitesse de rotation. Le bruit est peu important dans ce test où le traitement d'images rapide permet de maintenir un débit élevé pour l'acquisition (40 images/s/caméra). Finalement on constate, en fin d'asservissement, que l'erreur statique est faible (environ 3 mm en position, 1.5 deg en orientation).

Dans la seconde série, l'acquisition des images est fixée à un débit de 20 images/s/caméra. Les 6 DDL contraints par la vision permet de commander indépendamment les 3 DDL de position de l'organe terminal (et les 3 DDL de l'orientation, comme pour la première série). Les figures (Fig. 6-a,b) montrent la convergence des fonctions de tâche ($\mathbf{e}_\nu$ et $\mathbf{e}_\omega$) vers zéro. On observe que l'influence du bruit est plus important sur l'asservissement de position (Fig. 6-d) que dans la première série, ce qui n'est pas le cas de l'asservissement en orientation (Fig. 6-c) qui reste très stable. Il s'ensuit que l'erreur statique est d'environ 13 mm en position mais reste à environ 1 deg en orientation.

5 Conclusion et perspectives

Nous avons présenté dans cet article une méthode d'asservissement de robot manipulateur par stéréovision embarquée. L'approche proposée, basée sur la position 3-D et les coordonnées de Plücker binormalisées de droites, a permis d'obtenir un découplage des commandes des vitesses de translation et de rotation. Un des résultats théorique est d'assurer la stabilité globale, avec une ou plusieurs droites. Les premiers résultats expérimentaux obtenus, avec des gains faibles, montrent un comportement tout à fait satisfaisant de la convergence de l'AV, malgré le fait que le système n'est que grossièrement calibré.

Par la suite, un effort sera porté sur la validation expérimentale pour améliorer la vitesse de convergence afin de considérer un objet d'intérêt en mouvement. Outre le problème de visibilité des indices visuels évoqué au regard d'une estimation robuste, nous planifions d'étendre la méthode d'AV stéréo pour qu'elle incorpore le suivi d'un objet articulé, simultanément au positionnement du robot.

Références

[1] R. Andersson. Dynamic sensing in ping-pong playing robot. *IEEE Transactions on Robotics and Automation*, 5(6) :728–739, 1989.

[2] N. Andreff, B. Espiau, and R. Horaud. Visual servoing from lines. In *IEEE International Conference on Robotics and Automation*, pages 2070–2075, San Francisco, USA, 2000.

[3] A. Bartoli and P. Sturm. The 3d line motion matrix and alignment of line reconstructions. In *Proceedings of the IEEE Int'l Conf. on CVPR*, pages 287–292, Hawaii, USA, 2001.

[4] E. Cervera, A.P. del Pobil, F. Berry, and P. Martinet. Improving image-based visual servoing with three-dimensional features. *Int'l Journal of Robotics Research*, 22(10-11) :821–839, 2003.

[5] F. Chaumette and S. Hutchinson. Visual servo control part 1 : Basic approaches. *IEEE Robotics and Automation Magazine*, 13(4) :82–90, Décembre 2006.

[6] F. Chaumette and S. Hutchinson. Visual servo control part ii : Advanced approaches. *IEEE Robotics and Automation Magazine*, pages 109–118, Mars 2007.

[7] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation*, 8(3) :313–326, Juin 1992.

[8] A. Graffunder, Z. Ren, and I. Hartmann. Robot control using stereo eyes in hand technique. In *IFAC Int'l. Symposium on Robot control*, volume 2, pages 729–734, 1991.

[9] G. D. Hager. A modular system for robust positioning using feedback from stereo vision. *IEEE Trans. on Robotics and Automation*, 13(4) :582–595, 1997.

[10] G.D. Hager, W.-C. Chang, and A.S. Morse. Robot hand-eye coordination based on stereo vision. *IEEE Control System Magazine*, 15(1) :30–39, Février 1995.

[11] K. Hashimoto, T. Ebine, and H. Kimura. Visual servoing with hand-eye manipulator - optimal control approach. *IEEE Trans. on Robotics and Automation*, 12(5) :766–774, 1996.

[12] N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand-eye coordination. *Image and Vision Computing*, 12(3) :187–192, 1994.

[13] S. Hutchinson, G. D. Hager, and P.I. Corke. A tutorial on visual servo control. *IEEE Trans. on Robotics and Automation*, 12(5) :651–670, 1996.

[14] D. Kragic and H.I. Christensen. A framework for visual servoing. In *Proceedings of the Int'l Conf. on Computer Vision Systems*, pages 345–354. Springer-Verlag, 2003.

[15] A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel, L. Soler, J. Leroy, and J. Marescaux. Autonomous retrieval and 3d positioning of surgical instruments in robotized laparoscopic surgery. *IEEE Trans. on Robotics and Automation*, 19(5) :842–853, Octobre 2003.

[16] B. Lamiroy, B. Espiau, N. Andreff, and R. Horaud. Controlling robots with two cameras : How to do it properly. In *IEEE Int'l Conf. on Robotics and Automation*, pages 2100–2105, San Francisco, CA, USA, Avril 2000.

[17] E. Malis and F. Chaumette. 2 1/2 d visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. *International Journal of Computer Vision*, 37(1) :79–97, 2000.

[18] A. Mitiche. *Computational Analysis of Visual Motion*. Plenum Press, New York, 1994.

[19] N. Navab and O.D. Faugeras. The critical set of lines for camera displacement estimation : A mixed euclidean-projective and constructive approach. *Int'l Journal of Computer Vision*, 23(1) :17–44, Mai 1997.

[20] P. Rives and B. Espiau. Estimation récursive de primitives 3d au moyen d'une caméra mobile. *INRIA Research report 351*, 1987.

[21] B. Tamadazte, E. Marchand, S. Dembélé, and N. Le Fort-Piat. Cad model-based tracking and 3d visual-based control for mems microassembly. *Int'l. Journal of Robotics Research*, 29(11) :1416–1434, 2010.

[22] S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(4) :376–380, Avril 1991.

[23] H. Zhang and J. Ostrowski. Visual motion planning for mobile robots. *IEEE Trans. on Robotics and Automation*, 18(2) :199–208, 2002.