

Utilisation du modèle de projection sphérique pour l'asservissement visuel de sphères

Using a Spherical Projection Model for Visual Servoing wrt. Spheres

Roméo Tatsambon Fomena¹

François Chaumette²

¹ Université de Rennes 1

² INRIA

IRISA, Campus de Beaulieu, 35 042 Rennes-cedex
rtatsamb@irisa.fr

Résumé

Dans cet article, nous nous intéressons à l'utilisation du modèle de projection sphérique dans le cadre de la modélisation d'informations visuelles optimales pour l'asservissement visuel. Une nouvelle combinaison minimale de trois informations visuelles est proposée pour l'asservissement visuel de sphères à partir d'un système catadioptrique central quelconque. En utilisant cette combinaison et une loi de commande classique, la stabilité globale aux erreurs de modélisation et la stabilité locale aux erreurs d'étalonnage pour les caméras perspectives sont prouvées. Des expérimentations réalisées avec une caméra perspective confirment la validité des résultats théoriques proposés.

Mots Clefs

Asservissement visuel, vision par ordinateur, robotique.

Abstract

In this paper, we investigate the use of a spherical projection model to search for optimal visual features for visual servoing. A new minimal set of three visual features is proposed for visual servoing from spheres using any central catadioptric system. Using this set of features, a classical control method is proved to be globally stable even in the presence of modeling error and locally stable to calibration errors on perspective cameras. Using this type of cameras, experimental results are presented and validate the proposed theoretical results.

Keywords

Visual servoing, computer vision, robotics.

1 Introduction

Les techniques d'asservissement visuel consistent à utiliser les informations fournies par un capteur de vision afin de contrôler les mouvements d'un système dynamique [1], [2]. Avec un capteur de vision, la nature des informations visuelles potentielles est extrêmement riche. De cette richesse provient la difficulté majeure de l'asservissement visuel à savoir la sélection des informations qui fourniront un comportement optimal au système. Les qualités souhaitées sont nombreuses : stabilité locale voire globale, robustesse aux erreurs de mesure et aux erreurs de modélisation, absence de singularités et de minima locaux, trajectoire satisfaisante du robot, mais aussi des mesures dans l'image et enfin découplage et relation linéaire maximaux entre les informations visuelles et les degrés de liberté (ddl) commandés.

Plusieurs idées ont été proposées pour s'approcher d'un comportement optimal du système en utilisant des données 3D, hybrides (3D et 2D) ou uniquement 2D. Il est possible de contrôler le robot directement dans son espace cartésien en injectant des données 3D dans un schéma d'asservissement visuel [3]. Ces données sont obtenues à partir de la pose de l'objet cible dans le repère de la caméra. Ce schéma permet d'avoir de bonnes trajectoires dans l'espace 3D : lignes géodésiques pour l'orientation et lignes droites pour la position. Le fait qu'il n'y a pas de contrôle direct dans l'image peut cependant causer l'échec de l'asservissement dans le cas où les informations visuelles (ex-

traites de l'image perspective de la cible) utilisées pour l'estimation de la pose sortent du champ de vue. L'utilisation des coordonnées 3D de plusieurs points de la cible peut contribuer à garder la cible dans le champ de vue [4].

Une autre solution à ce problème de conservation de la visibilité est l'utilisation de lois de commandes hybrides. Cette solution a été exploitée dans l'asservissement 2 1/2D où un schéma de contrôle découplé, ne présentant aucune singularité dans l'espace de travail et tolérant une approximation grossière des paramètres d'étalonnage a été proposé [5].

Ce schéma reste tout de même plus sensible aux bruits de mesure (tout comme l'asservissement visuel 3D) que l'asservissement visuel 2D dans lequel les données issues de l'image sont directement intégrées dans la boucle de commande. Ce dernier type d'asservissement visuel est très intéressant à cause de sa robustesse aux erreurs sur les paramètres intrinsèques et extrinsèques de la caméra [6]. Il est évident qu'en utilisant uniquement les données 2D il n'y a aucun contrôle de la trajectoire 3D du système ; en plus il est possible de tomber sur des singularités ou de minima locaux qui limitent ainsi le domaine de convergence [7]. C'est l'une des raisons pour lesquelles beaucoup de travail a été fait (et reste encore à faire) pour améliorer le comportement du système en utilisant uniquement des données 2D.

Une information visuelle proportionnellement liée à la profondeur de l'objet observé a été établie dans [8] pour un bon comportement du système dans la direction de l'axe optique z . De manière similaire, pour obtenir un mouvement satisfaisant du robot dans l'espace cartésien, les mouvements de rotation et de translation de l'axe z peuvent être découplés des autres ddls [9]. Les coordonnées cylindriques des points peuvent aussi être exploitées pour obtenir un mouvement satisfaisant du système dans la direction z [10]. Une représentation générique et intuitive de l'image de la cible peut être obtenue en utilisant les moments 2D. Récemment, la théorie des moments invariants a été utilisée pour déterminer une combinaison particulière de moments 2D telle que l'interaction avec le système est quasiment linéaire et découplée lorsque des objets plans sont considérés [11].

Toutes les études ci-dessus sont basées sur le modèle classique de projection perspective qui n'est pas toujours adapté pour contrôler les systèmes. Pour une certaine classe de systèmes sous-actionnés, l'utilisation du modèle de projection sphérique est plus appropriée puisque la projection sphérique de l'image d'un point permet de retrouver la propriété de passivité qui est importante pour le contrôle de ces systèmes [12].

Beaucoup de travaux ont été réalisés dans le cadre de la modélisation. Cependant, très peu exploitent le modèle de projection sphérique. Cet article se propose d'utiliser ce modèle de projection pour modéliser des

informations visuelles optimales pour l'asservissement visuel de sphères. Une autre grande motivation qui milite en faveur de l'utilisation de ce modèle est sa simplicité comparée aux équations complexes générées par le modèle de projection omnidirectionnel, même à centre unique. Nous verrons qu'en utilisant le modèle de projection sphérique, la détermination d'informations visuelles optimales est assez simple et intuitive tandis que cela aurait été difficile d'obtenir les mêmes résultats à partir des projections perspective et omnidirectionnelle.

L'objet que nous avons considéré dans cette étude est une sphère : l'image d'un tel objet ne présente aucune singularité quelle que soit la position du robot. Dans la section suivante, le modèle de projection sphérique est utilisé pour construire une nouvelle combinaison optimale de trois informations visuelles indépendantes pour l'asservissement visuel de sphères. La matrice d'interaction associée à ces trois informations visuelles présente un découplage maximal et une liaison linéaire entre les informations visuelles et les vitesses de translation du système. En section 3, nous montrons que cette combinaison est générale dans le sens où elle est mesurable sur tout système catadioptrique central. En utilisant la combinaison sélectionnée, une analyse théorique de la stabilité et de la robustesse d'une loi de commande classique relativement aux erreurs de modélisation et d'étalonnage pour les caméras perspectives est présentée en section 4. Enfin, en utilisant ce modèle classique de caméras, des résultats expérimentaux sont présentés en section 5.

2 Modélisation d'informations visuelles optimales

Dans cette section, le modèle de projection sphérique est utilisé pour concevoir un nouvel ensemble de trois informations visuelles indépendantes pour l'asservissement visuel de sphères. Dans un premier temps la projection sphérique de la sphère est présentée ; ensuite l'accent est mis sur les informations visuelles potentielles qui se dégagent de cette projection et enfin nous sélectionnons une combinaison optimale de trois informations visuelles.

2.1 Projection sphérique d'une sphère

Soit $S_{(O,R)}$ la sphère de rayon R et de centre O de coordonnées (X_O, Y_O, Z_O) dans le repère caméra. Soit $S_{p(C,1)}$ la sphère unité de centre le centre optique de la caméra C . La projection sphérique de $S_{(O,R)}$ sur $S_{p(C,1)}$ est un dôme. Ce dôme peut être caractérisé par le contour Γ de sa base. Ce contour est présenté sur la Figure 1(a). La forme analytique de Γ est déterminée ci-dessous.

L'équation de $S_{(O,R)}$ est donnée par

$$(X - X_O)^2 + (Y - Y_O)^2 + (Z - Z_O)^2 - R^2 = 0. \quad (1)$$

Soit $d = \sqrt{X^2 + Y^2 + Z^2}$. Un point M de $S(O,R)$ de coordonnées (X, Y, Z) est projeté sur $S_p(C,1)$ en un point M_S de coordonnées (X_S, Y_S, Z_S) telles que : $X_S = \frac{X}{d}$, $Y_S = \frac{Y}{d}$ et $Z_S = \frac{Z}{d}$. La multiplication de (1) par $\frac{1}{d^2}$ donne

$$\left(X_S - \frac{X_O}{d}\right)^2 + \left(Y_S - \frac{Y_O}{d}\right)^2 + \left(Z_S - \frac{Z_O}{d}\right)^2 - \frac{R^2}{d^2} = 0, \quad (2)$$

qui peut être réécrite comme une fonction polynôme en $\frac{1}{d}$

$$\frac{K_O^2}{d^2} - \frac{2}{d}(X_O X_S + Y_O Y_S + Z_O Z_S) + 1 = 0 \quad (3)$$

où $K_O^2 = X_O^2 + Y_O^2 + Z_O^2 - R^2$.

Le contour Γ est tel que le discriminant Δ de (3) est égal à 0 (une seule solution à (3)) :

$$(X_O X_S + Y_O Y_S + Z_O Z_S)^2 - K_O^2 = 0. \quad (4)$$

Puisque Γ est devant le centre de projection, de (4) on obtient l'équation du plan \mathcal{P} support de Γ (voir Figure. 1(a))

$$X_O X_S + Y_O Y_S + Z_O Z_S = K_O. \quad (5)$$

En résumé, Γ est l'intersection de la sphère unité et du plan \mathcal{P} :

$$\Gamma = \begin{cases} X_S^2 + Y_S^2 + Z_S^2 = 1 \\ X_O X_S + Y_O Y_S + Z_O Z_S = K_O. \end{cases} \quad (6)$$

Le contour Γ est donc un cercle. Il est plus simple d'extraire des informations visuelles optimales à partir de ce cercle qu'à partir de l'ellipse observée sur le plan image perspectif.

2.2 Informations visuelles potentielles

Soit $(C, \mathbf{u}, \mathbf{v}, \mathbf{n})$ un nouveau système de coordonnées où (\mathbf{u}, \mathbf{v}) est un système orthonormal sur \mathcal{P} et \mathbf{n} est le vecteur normal à \mathcal{P} (voir Figure 1(a)).

Le passage du repère $(C, \mathbf{x}, \mathbf{y}, \mathbf{z})$ au repère $(C, \mathbf{u}, \mathbf{v}, \mathbf{n})$ est une rotation

$$\mathbf{R} = \begin{pmatrix} \frac{Z_O}{\alpha_1} & 0 & -\frac{X_O}{\alpha_1} \\ -\frac{X_O Y_O}{\alpha_2} & \frac{X_O^2 + Z_O^2}{\alpha_2} & -\frac{Y_O Z_O}{\alpha_2} \\ \frac{X_O^{\alpha_2}}{d_O} & \frac{Y_O^{\alpha_2}}{d_O} & \frac{Z_O^{\alpha_2}}{d_O} \end{pmatrix}$$

$$\text{où } \begin{cases} d_O = \sqrt{X_O^2 + Y_O^2 + Z_O^2} \\ \alpha_1 = \sqrt{X_O^2 + Z_O^2} \\ \alpha_2 = \sqrt{X_O^2 + Z_O^2} \sqrt{d_O}. \end{cases}$$

Dans le nouveau système de coordonnées $(C, \mathbf{u}, \mathbf{v}, \mathbf{n})$, (6) devient

$$\Gamma = \begin{cases} X'_S{}^2 + Y'_S{}^2 + Z'_S{}^2 = 1 \\ Z'_S = \sqrt{1 - \frac{R^2}{d_O^2}} \end{cases} \quad (7)$$

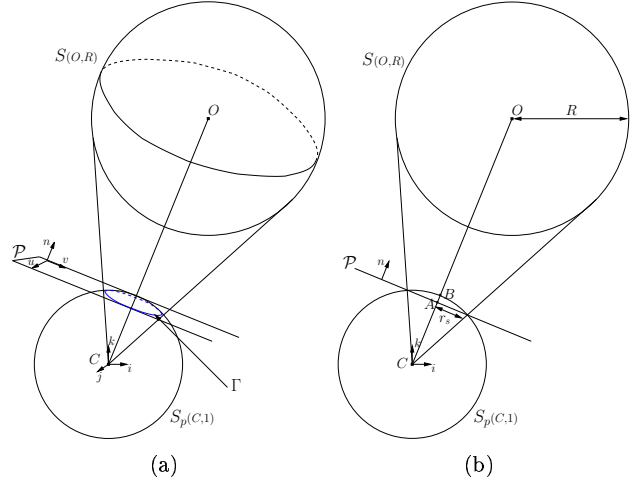


FIG. 1 – Projection sphérique d'une sphère. (a) Contour de la base du dôme. (b) Coupe perpendiculaire à \mathcal{P} .

où $(X'_S, Y'_S, Z'_S) = \mathbf{R}(X_S, Y_S, Z_S)$. De (7) on obtient facilement le rayon de Γ

$$r_s = R/d_O, \quad (8)$$

et le centre A de Γ dans le repère $(C, \mathbf{u}, \mathbf{v}, \mathbf{n})$

$$A \left(0, 0, \sqrt{1 - r_s^2}\right). \quad (9)$$

En utilisant \mathbf{R}^\top , on obtient les coordonnées de A dans le repère caméra

$$\begin{cases} X_A = X_O \sqrt{1 - r_s^2} / d_O \\ Y_A = Y_O \sqrt{1 - r_s^2} / d_O \\ Z_A = Z_O \sqrt{1 - r_s^2} / d_O. \end{cases} \quad (10)$$

En plus de A et r_s , le sommet B du dôme (voir Figure 1(b)) peut aussi fournir des informations visuelles potentielles. B correspond à la projection sphérique du centre O de $S(O,R)$. Les coordonnées de B dans le repère caméra sont données par

$$\begin{cases} X_B = X_O / d_O \\ Y_B = Y_O / d_O \\ Z_B = Z_O / d_O. \end{cases} \quad (11)$$

2.3 Informations visuelles sélectionnées

Dans cette partie nous présentons la matrice d'interaction associée à la combinaison optimale sélectionnée. Rappelons tout d'abord que la matrice d'interaction associée à un ensemble d'informations visuelles \mathbf{f} est définie par $\mathbf{f} = \mathbf{L}\mathbf{v}$ où $\mathbf{v} = (\mathbf{v}, \boldsymbol{\omega})$ est la vitesse instantanée de la caméra [13]; \mathbf{v} et $\boldsymbol{\omega}$ sont respectivement les vitesses de translation et de rotation de la caméra.

Trois paramètres suffisent pour caractériser la projection sphérique d'une sphère. C'est pourquoi nous devons sélectionner une combinaison de trois informations visuelles parmi $\{X_A, Y_A, Z_A, X_B, Y_B, Z_B, r_s\}$. Parmi toutes les combinaisons possibles, la combinaison $\mathbf{s} = (\frac{X_B}{r_s}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s}) = (\frac{X_O}{R}, \frac{Y_O}{R}, \frac{Z_O}{R})$ comparée aux autres est séduisante puisque sa matrice d'interaction est simple et découplée au maximum :

$$\mathbf{L}_s = \left[\begin{array}{c|c} -\frac{1}{R}\mathbf{I}_3 & [\mathbf{s}]_\times \end{array} \right], \quad (12)$$

où $[\mathbf{s}]_\times$ est la matrice du préproduit vectoriel associée à \mathbf{s} .

En plus de la propriété de découplage, \mathbf{L}_s présente la même dynamique ($\frac{1}{R}$) sur les composantes de translation du système. Puisque R est constant, il y a une relation linéaire entre les informations visuelles et les vitesses de translation du système. On peut aussi noter que cette matrice d'interaction présente la propriété de passivité qui est importante pour contrôler certains systèmes sous-actionnés [12]. Pour toutes ces raisons, nous proposons la combinaison $\mathbf{s} = (\frac{X_B}{r_s}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s})$ pour l'asservissement visuel de sphères.

Le seul paramètre 3D dans \mathbf{L}_s est la constante R . En pratique, \hat{R} (valeur estimée de R) est utilisée. À partir de l'analyse de stabilité aux erreurs de modélisations (de la cible) présentée en section 5, un domaine de validité de \hat{R} sera donné.

Pour l'instant nous montrons comment mesurer la combinaison \mathbf{s} choisie en utilisant un système catadioptrique central quelconque.

3 Application sur un système catadioptrique central quelconque

Dans cette partie, nous présentons tout d'abord la forme analytique de l'image de la sphère sur le plan image d'un système catadioptrique central quelconque. Ensuite nous montrons comment mesurer $\mathbf{s} = (\frac{X_B}{r_s}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s})$ à partir de cette forme analytique.

3.1 Équation de l'image catadioptrique de la sphère

Soient φ et ξ les paramètres du miroir d'une caméra catadioptrique centrale. L'image catadioptrique de la sphère est une ellipse. La formation de cette ellipse peut être décomposée en deux étapes (voir Figure 2) basées sur le modèle unifié de formation d'une image catadioptrique [14]. On rappelle que pour une caméra modélisée par une projection perspective classique, le modèle utilisé est applicable en posant simplement $\varphi = 1$ et $\xi = 0$.

La première étape est la projection sphérique de $S_{(O,R)}$ dans le repère virtuel centré en V . Ce résultat a été présenté au point 2.1. Puisque $S_{(O,R)}$ est décrit dans

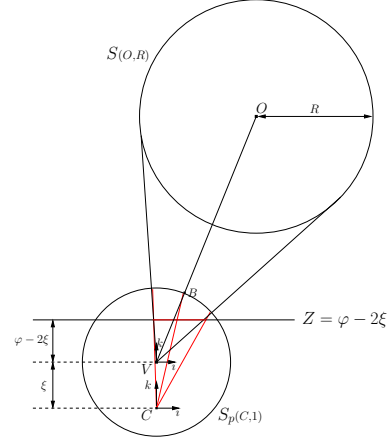


FIG. 2 – Image catadioptrique de la sphère.

le repère virtuel centré en V , on obtient

$$\Gamma = \begin{cases} X_S^{V^2} + Y_S^{V^2} + Z_S^{V^2} = 1 \\ X_O X_S^V + Y_O Y_S^V + Z_O Z_S^V = \sqrt{X_O^2 + Y_O^2 + Z_O^2 - R^2}. \end{cases} \quad (13)$$

Pour la seconde étape, le contour Γ est exprimé dans le repère caméra et projeté sur le plan image catadioptrique. Γ est donc l'intersection de la sphère

$$X_S^2 + Y_S^2 + (Z_S - \xi)^2 = 1 \quad (14)$$

et du plan

$$X_O X_S + Y_O Y_S + Z_O Z_S = K_O + \xi Z_O. \quad (15)$$

Les équations de projection sur le plan image catadioptrique ne sont autres que

$$\begin{cases} x_o = \frac{X_S}{Z_S} \\ y_o = \frac{Y_S}{Z_S}. \end{cases} \quad (16)$$

En injectant (16) dans (15) on a

$$\frac{1}{Z_S} = \frac{X_O x_o + Y_O y_o + Z_O}{K_O + \xi Z_O} \quad (17)$$

et (16) dans (14) donne

$$x_o^2 + y_o^2 + 1 - 2\frac{\xi}{Z_S} + \frac{\xi^2 - 1}{Z_S^2} = 0. \quad (18)$$

Enfin, en injectant (17) dans (18) on obtient l'équation de l'ellipse

$$k_0 x_o^2 + k_1 y_o^2 + 2k_2 x_o y_o + 2k_3 x_o + 2k_4 y_o + k_5 = 0 \quad (19)$$

avec

$$\begin{cases} k_0 = (K_O + \xi Z_O)^2 + (\xi^2 - 1) X_O^2 \\ k_1 = (K_O + \xi Z_O)^2 + (\xi^2 - 1) Y_O^2 \\ k_2 = (\xi^2 - 1) X_O Y_O \\ k_3 = X_O ((\xi^2 - 1) Z_O - \xi (K_O + \xi Z_O)) \\ k_4 = Y_O ((\xi^2 - 1) Z_O - \xi (K_O + \xi Z_O)) \\ k_5 = (K_O + \xi Z_O)^2 + (\xi^2 - 1) Z_O^2 - 2\xi Z_O (K_O + \xi Z_O). \end{cases}$$

3.2 Mesure des informations visuelles

Maintenant nous montrons comment calculer \mathbf{s} en utilisant les moments $\mu = (x_g, y_g, n_{20}, n_{11}, n_{02})$ de l'ellipse observée sur le plan image catadioptrique.

Partant de (8) et (11), on rappelle tout d'abord que : $\frac{X_B}{r_s} = \frac{X_O}{R}$, $\frac{Y_B}{r_s} = \frac{Y_O}{R}$ et $\frac{Z_B}{r_s} = \frac{Z_O}{R}$.

En utilisant (19), les moments de l'ellipse sur le plan image catadioptrique peuvent être exprimés en utilisant les paramètres 3D :

$$\begin{cases} x_g = X_O H_1 / H_2 \\ y_g = Y_O H_1 / H_2 \\ 4n_{20} = (H_2 - (\xi^2 - 1) X_O^2) R^2 / H_2^2 \\ 4n_{11} = -X_O Y_O (\xi^2 - 1) R^2 / H_2^2 \\ 4n_{02} = (H_2 - (\xi^2 - 1) Y_O^2) R^2 / H_2^2 \end{cases} \quad (20)$$

$$\text{avec } \begin{cases} H_1 = Z_O + \xi K_O \\ H_2 = H_1^2 + (\xi^2 - 1) R^2. \end{cases}$$

Après quelques calculs, on obtient en utilisant (20)

$$\begin{cases} \frac{X_B}{r_s} = x_g \frac{h_2}{\sqrt{h_2 + (1 - \xi^2)}} \\ \frac{Y_B}{r_s} = y_g \frac{h_2}{\sqrt{h_2 + (1 - \xi^2)}} \end{cases} \quad (21)$$

où $h_2 = 1/f(\mu)$ avec $f(\mu) = \frac{4n_{20}y_g^2 + 4n_{02}x_g^2 - 8n_{11}x_g y_g}{x_g^2 + y_g^2}$.

Il est possible de démontrer que $f(\mu)$ est continue même lorsque $x_g = y_g = 0$ auquel cas $f(\mu) = 4n_{20}$.

Dans le cas des systèmes paracatadioptriques (où $\xi = 1$, qui correspond au couplage d'un miroir parabolique et d'une caméra orthographique), on a :

$$\frac{Z_B}{r_s} = \frac{h_2 - \left(\frac{X_B^2}{r_s^2} + \frac{Y_B^2}{r_s^2} - 1 \right)}{2\sqrt{h_2}} \quad (22)$$

et pour tous les autres systèmes catadioptriques ($\xi \neq 1$)

$$\frac{Z_B}{r_s} = \frac{h_1 - \xi \sqrt{h_1^2 + (1 - \xi^2) \left(\frac{X_B^2}{r_s^2} + \frac{Y_B^2}{r_s^2} - 1 \right)}}{(1 - \xi^2)} \quad (23)$$

où $h_1 = \sqrt{h_2 + (1 - \xi^2)}$. On peut noter que pour la projection perspective classique, on a simplement

$$\begin{cases} \frac{X_B}{r_s} = x_g \frac{h_2}{\sqrt{h_2 + 1}} \\ \frac{Y_B}{r_s} = y_g \frac{h_2}{\sqrt{h_2 + 1}} \\ \frac{Z_B}{r_s} = \sqrt{h_2 + 1}. \end{cases} \quad (24)$$

Pour conclure, la combinaison proposée $\mathbf{s} = \left(\frac{X_B}{r_s}, \frac{Y_B}{r_s}, \frac{Z_B}{r_s} \right)$ est générale puisqu'elle est mesurable sur tout système catadioptrique central.

4 Analyse de stabilité

La stabilité du système est analysée séparément : une condition nécessaire et suffisante est donnée pour la stabilité globale aux erreurs de modélisation et pour

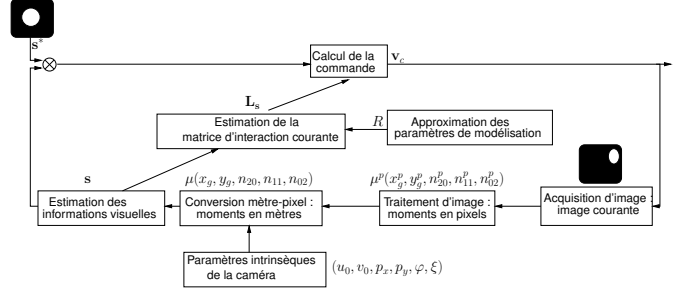


FIG. 3 – Boucle fermée.

la stabilité locale aux erreurs d'étalonnage sur des caméras perspectives.

Considérons l'asservissement visuel de sphères avec la combinaison \mathbf{s} choisie. La loi de commande utilisée est

$$\mathbf{v}_c = -\lambda \widehat{\mathbf{L}}_{\mathbf{s}}^+ (\mathbf{s} - \mathbf{s}^*) \quad (25)$$

où \mathbf{v}_c est la vitesse de la caméra envoyée au contrôleur bas-niveau du robot, λ est un gain positif et $\widehat{\mathbf{L}}_{\mathbf{s}}^+$ est la pseudo-inverse d'une approximation de la matrice d'interaction associée à \mathbf{s} .

La Figure 3 décrit la boucle fermée de l'asservissement visuel. Pour un système catadioptrique avec φ et ξ comme paramètres du miroir, la conversion mètre-pixel est donnée par [14] :

$$\begin{cases} x_g = \frac{x_g^p - u_0}{p_x(\varphi - \xi)} \\ y_g = \frac{y_g^p - v_0}{p_y(\varphi - \xi)} \\ n_{20} = \frac{n_{20}^p}{p_x^2(\varphi - \xi)^2} \\ n_{11} = \frac{n_{11}^p}{p_x p_y(\varphi - \xi)^2} \\ n_{02} = \frac{n_{02}^p}{p_y^2(\varphi - \xi)^2} \end{cases} \quad (26)$$

où u_0, v_0, p_x and p_y sont les paramètres intrinsèques de la caméra (pour rappel la configuration $\varphi = 1$ et $\xi = 0$ correspond à une caméra perspective classique).

De la Figure 3, il ressort trois sources potentielles d'erreur dans la boucle fermée : les erreurs de modélisation causées par l'approximation de R , les erreurs d'étalonnage affectant la conversion mètre-pixel et les erreurs de traitement d'images.

4.1 Stabilité aux erreurs de modélisation

En supposant que nous ne sommes pas en présence des deux derniers types d'erreurs, l'équation de la boucle fermée peut être écrite comme :

$$\dot{\mathbf{s}} = -\lambda \mathbf{L}_{\mathbf{s}} \widehat{\mathbf{L}}_{\mathbf{s}}^+ (\mathbf{s} - \mathbf{s}^*) \quad (27)$$

avec, en utilisant (12)

$$\widehat{\mathbf{L}}_{\mathbf{s}}^+ = \begin{bmatrix} -\frac{\hat{R} r_s^2}{r_s^2 + \hat{R}^2} \left(\hat{R}^2 \mathbf{s} \mathbf{s}^\top + \mathbf{I}_3 \right) \\ -\frac{\hat{R}^2 r_s^2}{r_s^2 + \hat{R}^2} [\mathbf{s}]_\times \end{bmatrix}.$$

Une condition suffisante pour la stabilité asymptotique globale aux erreurs de modélisation est $\mathbf{L}_s \widehat{\mathbf{L}}_s^+ > 0$. Les valeurs propres de $\mathbf{L}_s \widehat{\mathbf{L}}_s^+ > 0$ peuvent être calculées. Elles sont données par $\frac{\hat{R}}{R}$ et $\frac{\hat{R}(r_s^2 + R\hat{R})}{R(r_s^2 + \hat{R}^2)}$ (qui est une valeur propre double). Ainsi, on a :

$$\begin{aligned} \mathbf{L}_s \widehat{\mathbf{L}}_s^+ > 0 &\iff \frac{\hat{R}(r_s^2 + R\hat{R})}{R(r_s^2 + \hat{R}^2)} > 0 \text{ et } \frac{\hat{R}}{R} > 0 \\ &\iff \hat{R} > 0. \end{aligned}$$

Cette condition est également nécessaire. En effet, lorsque $\hat{R} \leq 0$, on a $\mathbf{L}_s \widehat{\mathbf{L}}_s^+ \leq 0$, donc le système diverge. Le domaine de robustesse aux erreurs de modélisation est donc extrêmement large : $\hat{R} \in]0, +\infty[$. D'un point de vue pratique, une estimation très approximative de R est donc suffisante.

4.2 Stabilité aux erreurs d'étalonnage

En considérant uniquement des erreurs d'étalonnage, (27) devient

$$\dot{\mathbf{e}} = -\lambda \mathbf{L}_s \widehat{\mathbf{L}}_s^+ \mathbf{E}(\mathbf{s}) \mathbf{e} \quad (28)$$

avec $\mathbf{e} = \mathbf{s} - \mathbf{s}^*$, \mathbf{E} tel que $\hat{\mathbf{s}} = \mathbf{E}(\mathbf{s}) \mathbf{s}$ et

$$\widehat{\mathbf{L}}_s^+ = \begin{bmatrix} -\frac{R\hat{r}_s^2}{\hat{r}_s^2 + R^2} (R^2 \hat{\mathbf{s}} \hat{\mathbf{s}}^\top + \mathbf{I}_3) \\ -\frac{R^2 \hat{r}_s^2}{\hat{r}_s^2 + R^2} [\hat{\mathbf{s}}]_\times \end{bmatrix}.$$

La non-linéarité de (28) rend complexe l'analyse de la stabilité globale aux erreurs d'étalonnage. Cette analyse peut être simplifiée en considérant la stabilité locale (c'est-à-dire pour $\mathbf{s} = \mathbf{s}^*$) ce qui permet de linéariser \mathbf{E} en \mathbf{s}^* . Le terme \mathbf{E} peut être linéarisé en utilisant la dérivée temporelle de \mathbf{s} . En effet, partant de (24) (c'est-à-dire que l'on considère ici une caméra perspective), on obtient après quelques développements

$$\begin{cases} \left(\frac{\dot{X}_B}{r_s} \right) = -\frac{u_0(p_x^2 + p_y^2)}{p_x(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*})} \frac{Z_B}{r_s} \\ \left(\frac{\dot{Y}_B}{r_s} \right) = -\frac{v_0(p_x^2 + p_y^2)}{p_y(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*})} \frac{Z_B}{r_s} \\ \left(\frac{\dot{Z}_B}{r_s} \right) = \frac{p_x p_x + p_y p_y}{p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}} \frac{Z_B}{r_s}, \end{cases}$$

d'où on déduit

$$\mathbf{E}^* = \begin{bmatrix} 1 & 0 & \alpha_{u_0} \\ 0 & 1 & \alpha_{v_0} \\ 0 & 0 & 1 + \alpha_{p_x p_y} \end{bmatrix}$$

$$\text{avec } \begin{cases} \alpha_{u_0} = -\frac{\Delta u_0(p_x^2 + p_y^2)}{p_x(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*})} \\ \alpha_{v_0} = -\frac{\Delta v_0(p_x^2 + p_y^2)}{p_y(p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*})} \\ \alpha_{p_x p_y} = \frac{p_x \Delta p_x + p_y \Delta p_y}{p_x^2 + p_y^2 + 4n_{20}^{p*} + 4n_{02}^{p*}} \end{cases}$$

$$\text{où } \begin{cases} \Delta u_0 = \widehat{u}_0 - u_0 \\ \Delta v_0 = \widehat{v}_0 - v_0 \\ \Delta p_x = \widehat{p}_x - p_x \\ \Delta p_y = \widehat{p}_y - p_y. \end{cases}$$

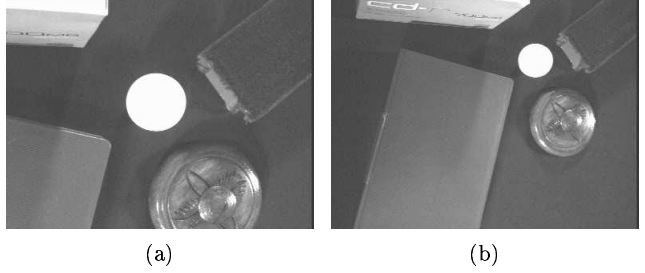


FIG. 4 – (a) Image désirée. (b) Image initiale.

On obtient alors :

$$\begin{aligned} \mathbf{L}_{s^*} \widehat{\mathbf{L}}_{s^*}^+ \mathbf{E}^* &= \frac{\hat{r}_s^{*2}}{\hat{r}_s^{*2} + R^2} \left(1 + R^2 \mathbf{s}^{*\top} \mathbf{E}^* \mathbf{s}^* \right) \mathbf{E}^* \\ &+ \left(\frac{R^2 \hat{r}_s^{*2}}{\hat{r}_s^{*2} + R^2} \left(\mathbf{E}^* \mathbf{s}^* \mathbf{s}^{*\top} \mathbf{E}^{*\top} - \mathbf{E}^* \mathbf{s}^* \mathbf{s}^{*\top} \right) \right) \mathbf{E}^*. \end{aligned} \quad (29)$$

Les valeurs propres γ_i de $\mathbf{L}_{s^*} \widehat{\mathbf{L}}_{s^*}^+ \mathbf{E}^*$ peuvent être obtenues après des calculs complexes. Elles sont données par

$$\begin{cases} \gamma_1 = \frac{\hat{r}_s^{*2} r_s^{*2} + R^2(1 + \alpha_{p_x p_y})}{\hat{r}_s^{*2} + R^2} \\ \gamma_2 = a + b^{\frac{1}{2}} \\ \gamma_3 = a - b^{\frac{1}{2}} \end{cases} \quad (30)$$

où a et b sont des valeurs réelles qui sont trop complexes pour être écrites ici. À partir de (30) il est possible de montrer qu'une condition suffisante pour assurer $Re(\gamma_i) > 0$ est $\Delta p_x > -p_x$ et $\Delta p_y > -p_y$; c'est-à-dire $\widehat{p}_x > 0$ et $\widehat{p}_y > 0$ ce qui est, une fois de plus, un excellent résultat.

5 Résultats expérimentaux

Les expérimentations ont été effectuées avec une caméra perspective conventionnelle embarquée sur un système robotique à six ddl. Nous rappelons que ce type de caméra est un type de système catadioptrique avec $\xi = 0$ et $\varphi = 1$ comme paramètres du miroir. Ces expérimentations consistent en une tâche de centrage. La cible est une balle de ping-pong de rayon 2 cm. L'utilisation d'un objet aussi simple permet de calculer facilement les moments de l'ellipse à la cadence vidéo sans problème de traitement d'images. La valeur désirée \mathbf{s}^* est fixée à $(0, 0, \tau^*)$ où τ^* permet de fixer la profondeur entre la caméra et la balle. La Figure 4 montre les images désirée et initiale utilisées pour chaque expérimentation. Le même gain $\lambda = 0.5$ a été utilisé.

5.1 Cas idéal

Pour valider les informations visuelles sélectionnées, nous considérons tout d'abord le cas idéal où $\hat{R} = R$. En effet, lorsque $\hat{R} = R$ on a un comportement idéal

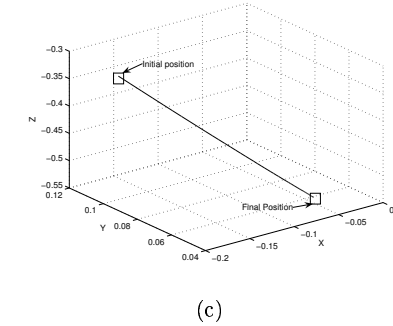
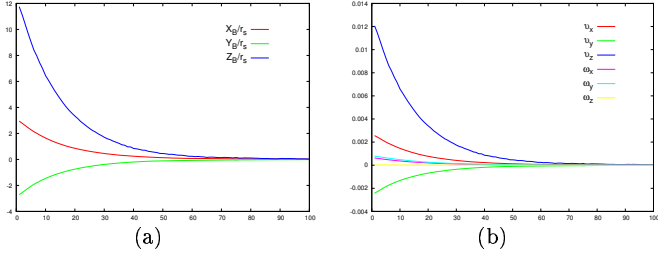


FIG. 5 – Cas idéal. (a) $\mathbf{s} - \mathbf{s}^*$. (b) Vitesses de la caméra (m/s et dg/s). (c) Trajectoire de la caméra dans l'espace cartésien.

du système puisque $\mathbf{L}_s \widehat{\mathbf{L}}_s^+ = \mathbf{I}_3$. Comme prévu, une décroissance exponentielle pure de l'erreur sur les informations visuelles peut être observée sur la Figure 5(a). Les vitesses de la caméra sont quant à elles affichées sur la Figure 5(b). Enfin, la trajectoire de la caméra dans l'espace cartésien est, comme prévu, une ligne droite pure (voir Figure 5(c)).

5.2 Erreurs de modélisation

La stabilité aux erreurs de modélisation a été prouvée dans cet article. Cette preuve a été validée par deux expérimentations. Les résultats dans les cas $\hat{R} = 10R$ et $\hat{R} = 0.1R$ sont présentés respectivement sur les Figures 6 et 7. On peut noter que dans les deux cas, le système converge encore.

La Figure 6(b) montre une grande vitesse sur les composantes en translation du système tandis que la Figure 7(b) montre une faible vitesse sur ces mêmes composantes du système. En effet, partant de (25) (où $\widehat{\mathbf{L}}_s^+$ est définie dans (27)), on obtient aisément

$$\begin{cases} \mathbf{v} = \lambda \frac{\hat{R} r_s^2}{r_s^2 + \hat{R}^2} \left(\hat{R}^2 \mathbf{s} \mathbf{s}^\top + \mathbf{I}_3 \right) (\mathbf{s} - \mathbf{s}^*) \\ \boldsymbol{\omega} = \lambda \frac{\hat{R}^2 r_s^2}{r_s^2 + \hat{R}^2} [\mathbf{s}]_\times (\mathbf{s} - \mathbf{s}^*). \end{cases} \quad (31)$$

où \mathbf{v} et $\boldsymbol{\omega}$ sont respectivement les composantes en translation et en rotation de la vitesse de la caméra. Lorsque \hat{R} tend vers $+\infty$, (31) tend vers

$$\begin{cases} \mathbf{v} = \infty \\ \boldsymbol{\omega} = \lambda r_s^2 [\mathbf{s}]_\times (\mathbf{s} - \mathbf{s}^*) \end{cases}$$

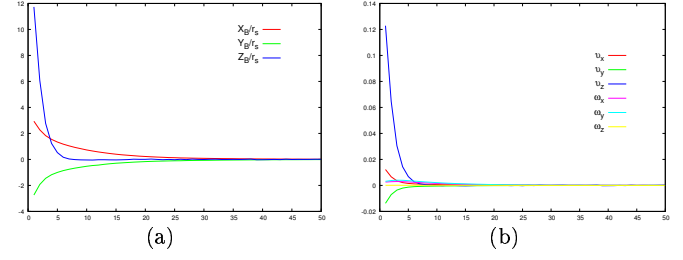


FIG. 6 – Erreurs de modélisation $\hat{R} = 10R$. (a) $\mathbf{s} - \mathbf{s}^*$. (b) Vitesses de la caméra (m/s et dg/s).

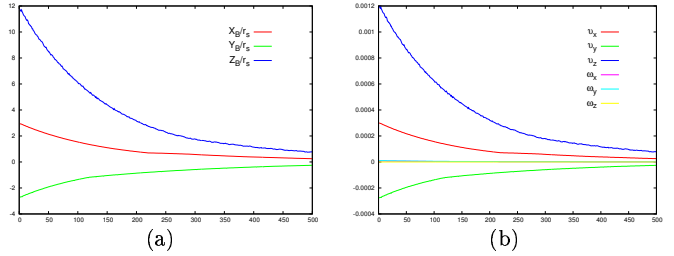


FIG. 7 – Erreurs de modélisation $\hat{R} = 0.1R$. (a) $\mathbf{s} - \mathbf{s}^*$. (b) Vitesses de la caméra (m/s et dg/s).

ce qui explique la convergence rapide observée sur la Figure 6 (50 itérations) lorsque $\hat{R} = 10R$. Lorsque \hat{R} tend vers 0, partant de (31) on a \mathbf{v} et $\boldsymbol{\omega}$ qui tendent vers 0. Cela explique la convergence lente observée sur la Figure 7 (500 itérations) lorsque $\hat{R} = 0.1R$. En pratique, le comportement du système peut être amélioré en utilisant un gain λ élevé (dans le cas d'une sous-approximation de \hat{R}) et en saturant \mathbf{v}_c au besoin (dans le cas d'une sur-approximation de \hat{R}).

5.3 Erreurs d'étalonnage

Enfin, nous validons les résultats obtenus lors de l'analyse de stabilité aux erreurs d'étalonnage. La condition suffisante $\hat{p}_x > 0$ & $\hat{p}_y > 0$ est validée en introduisant des erreurs dans les paramètres intrinsèques de la caméra : 35% p_x , -47% p_y , -25% u_0 et 57% v_0 . Les résultats obtenus sont donnés sur la Figure 8. Une fois de plus, le système converge (même si seule la stabilité locale a été considérée). La décroissance rapide de l'erreur sur $\frac{YB}{r_s}$ comparée aux autres erreurs est simplement due à la grande marge d'erreur introduite sur p_y . Bien évidemment nous avons observé que si $\hat{p}_x < 0$ ou $\hat{p}_y > 0$, le système diverge, ce qui n'est pas surprenant.

6 Conclusions et perspectives

Dans cet article, le modèle de projection sphérique a été utilisé dans le cadre de la modélisation d'informations visuelles optimales pour l'asservissement visuel. Une nouvelle combinaison de trois informations

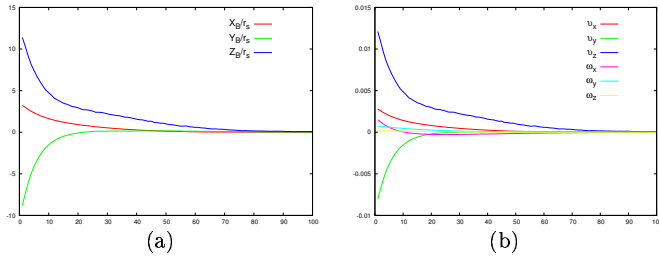


FIG. 8 – Erreurs d'étalonnage. (a) $s - s^*$. (b) Vitesses de la caméra (m/s et dg/s).

visuelles indépendantes a été proposée pour l'asservissement visuel de sphères. Cette combinaison peut être mesurée sur l'image de la sphère obtenue à partir d'un système catadioptrique central quelconque. La matrice d'interaction associée à cette combinaison est linéaire sur les composantes de translation et présente un découplage maximal. En utilisant cette combinaison, la stabilité globale aux erreurs de modélisation et la stabilité locale aux erreurs d'étalonnage sur une caméra perspective d'une loi de commande classique ont été prouvées. Ces résultats théoriques ont été validés expérimentalement sur un système robotique équipé d'une caméra perspective.

Les travaux futurs seront consacrés à la recherche des mêmes résultats pour d'autres primitives géométriques telles que les droites par exemple.

Références

[1] K. Hashimoto, Ed., *Visual servoing : Real time control of robot manipulators based on visual sensory feedback*, ser. World Scientific series in Robotics and Automated systems. Singapore : World Scientific Press, 1993, vol. 7.

[2] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Trans. on Robotics and Automation*, vol. 12, no. 3, pp. 651–670, Oct. 1996.

[3] W. Wilson, C. Hulls, and G. Bell, "Relative end-effector control using cartesian position-based visual servoing," *IEEE Trans. on Robotics and Automation*, vol. 12, no. 5, pp. 684–696, Oct. 1996.

[4] P. Martinet, J. Gallice, and D. Khadraoui, "Vision based control law using 3d visual features," in *World Automation Congress, WAC, Robotics and manufacturing systems*, vol. 3, Montpellier, France, Mai 1996, pp. 497–502.

[5] E. Malis, F. Chaumette, and S. Boudet, "2 1/2 d visual vervoing," *IEEE Trans. on Robotics and Automation*, vol. 15, no. 2, pp. 238–250, Apr. 1999.

[6] B. Espiau, "Effect of camera calibration errors on visual servoing in robotics," in *3rd Int. Sym-*

posium on Experimental Robotics, Kyoto, Japan, October 1993.

[7] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The Confluence of Vision and Control*, D. Kriegman, G. Hager, and A. S. Morse, Eds. LNCIS Series, No 237, Springer-Verlag, 1998, pp. 66–78.

[8] R. Mahony, P. Corke, and F. Chaumette, "Choice of image features for depth-axis control in image-based visual servo control," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'02*, vol. 1, Lausanne, Switzerland, Oct. 2002, pp. 390–395.

[9] P. Corke and S. Hutchinson, "A new partitioned approach to image-based visual visual servo control," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 4, pp. 507–515, Aug. 2001.

[10] M. Iwatsuki and N. Okiyama, "A new formulation for visual servoing based on cylindrical coordinate system," *IEEE Trans. on Robotics*, vol. 21, no. 2, pp. 266–273, Apr. 2005.

[11] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. on Robotics*, vol. 21, no. 6, pp. 1116–1127, Dec. 2005.

[12] T. Hamel and R. Mahony, "Visual servoing of an under-actuated dynamic rigid-body system : an image-based approach," *IEEE Trans. on Robotics and Automation*, vol. 18, no. 2, pp. 187–198, Apr. 2002.

[13] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. on Robotics and Automation*, vol. 8, no. 3, pp. 313–326, June 1992.

[14] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," *European Conference on Computer Vision*, vol. 29, pp. 159–179, May 2000.