Vision-based control for rigid body stabilization

Rita Cunha, Carlos Silvestre, João Hespanha, and A. Pedro Aguiar

Abstract— **This paper addresses the problem of stabilizing to a desired equilibrium point an eye-in-hand system, which consists of a single camera mounted on a rigid body free to move in $\mathrm{SE}(3)$. It is assumed that there is a collection of landmarks fixed in the environment and that the image coordinates of those landmarks are provided to the system by an on-board CCD camera. The proposed method addresses not only the problem of stabilization but also that of maintaining feature visibility along the system's trajectory. The resulting solution consists of a feedback control law based on the image coordinates and reconstructed depth information that guarantees *i)* almost global asymptotic stability of the desired equilibrium point; *ii)* positive invariance of a conveniently defined subset of $\mathrm{SE}(3)$, to enforce feature visibility; and *iii)* exponential stability of an error vector directly defined in the image plane.**

I. Introduction

Computer vision has long been thought of as an extremely flexible means of sensing the environment and acquiring valuable information for feedback control. Over the last decade, awareness of this potential has brought about a widespread interest in the field of *vision-based* control, also known as *visual-servoing*. Vision-based control can be used to perform a variety of tasks such as positioning a manipulator's end-effector with respect to an object to be grasped [1] or landing an UAV over a predefined target [2].

Visual-servoing is traditionally classified as either *position-based* or *image-based* [3]. Despite their proved merits, each of these methods exhibits a number of pitfalls, leaving behind unsolved problems, as reported in [1], [4]. The main problem in position-based methods derives from the fact that the feedback law is designed in the configuration space, without taking into the account the mapping to the image plane. Thus, the resulting camera trajectories are likely to cause the loss of features in the field of view (FOV), which precludes the reconstruction of the 3-D pose and consequently leads to system failure. Likewise, the drawbacks of the classical image-based approach should not be overlooked. In particular, asymptotic stability of the system is only guaranteed locally [5] and, more importantly, an analytical characterization of the region of attraction for the desired equilibrium point is yet to be established. Furthermore, the image-based approach does not entirely solve the problem of image feature loss; for example, to produce a straight line trajectory for the features in the image plane, the camera may have to move behind the observed object, causing the features to be occluded.

In view of the above considerations, one of the main questions in vision-based control, which continues to challenge researchers, is the FOV problem. As described in [1], the FOV problem presents two challenges: the features should not leave the image boundaries and they should also not become occluded by the object on which they are marked. There are several recent approaches that address the problem explicitly (see [6]–[9] and references therein). These include methods that partition the system's degrees of freedom into position and rotation, allowing for the definition of decoupled control laws [6] and hybrid methods that switch between different controllers combining, for example, position-based and image-based strategies [8], [9].

In this paper, building on results presented in [10], we propose an alternative solution that guarantees almost globally asymptotically stability (GAS) of a target configuration defined in $\mathrm{SE}(3)$. More importantly, the current approach ensures the positive invariance of a subset of $\mathrm{SE}(3)$, specifically defined to enforce feature visibility throughout the closed-loop trajectories of the camera. In contrast to most strategies, which only consider the problem of keeping the features inside the camera's FOV, the proposed method also takes into account the second type of feature loss, which is due to self-occlusions. To this end, the invariant set is defined so that the camera not only points towards the features, but also remains in front of them.

We reinforce the likelihood of maintaining feature visibility by showing that a relevant error directly defined in the image plane is exponentially stable inside the predefined invariant set. The image error is given by the difference between the current and desired images of a predefined point, which if chosen judiciously can provide a valuable measure for feature visibility. Note that a result of almost GAS for the desired configuration in $\mathrm{SE}(3)$ is not a sufficient condition for stability in the image plane. Consequently, although the current approach is position-based in the sense that depth information must be recovered from the image measurements and a priori knowledge of the features' geometry is required, it can also be interpreted as image-based, with the advantages relative to other solutions of taking into account self-occlusions and providing a formal characterization for the region of attraction.

The paper is organized as follows. Section II introduces

This work was partially supported by Fundação para a Ciência e a Tecnologia (ISR/IST pluriannual funding) through the POS_Conhecimento Program that includes FEDER funds and by the project PTDC/EEA-ACR/72853/2006 HELICIM. Hespanha's research was supported by the NSF Grant ECS-0242798.

R. Cunha, C. Silvestre, and A. P. Aguiar are with the Department of Electrical Engineering and Computer Science, and Institute for Systems and Robotics, Instituto Superior Técnico, 1046-001 Lisboa, Portugal. {rita,cjs,pedro}@isr.ist.utl.pt

J. Hespanha is with Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560, USA. hespanha@ece.ucsb.edu

the vision-based control problem. Section III describes the construction of an almost globally asymptotically stabilizing controller for the system at hand. An exact expression for the region of attraction is derived in Section III-A and the positive invariance of a set of FOV-related configurations together with the exponentially stability of the image error are formally established in Section III-B. Simulation results that illustrate the behavior of the control system are presented in Section IV. For the sake of brevity, most of the proofs and technical results are either omitted or only outlined in the paper, and the reader is referred to [11] for a comprehensive presentation of this material.

II. Problem Formulation

Consider a fully-actuated rigid-body, attached to a coordinate frame $\{B\}$ and let $(\mathbf{p}, R) = ({}^{B}\mathbf{p}_{I},\ {}^{B}_{I}R) \in \mathrm{SE}(3)$ denote the configuration of an inertial coordinate frame $\{I\}$ with respect to $\{B\}$, such that

$$\dot{\mathbf{p}} = -\mathbf{v} - S(\boldsymbol{\omega})\mathbf{p} \tag{1a}$$

$$\dot{R} = -S(\boldsymbol{\omega})R, \tag{1b}$$

where $\mathbf{v}$ and $\boldsymbol{\omega} \in \mathbb{R}^3$ are the body-fixed linear and angular velocities, respectively and S denotes the map from $\mathbb{R}^3$ to the space of skew-symmetric matrices $\mathrm{so}(3) = \{M \in \mathbb{R}^{3\times 3} : M^T = -M\}$ defined so that $S(\mathbf{a})\mathbf{b} = \mathbf{a} \times \mathbf{b}$, where $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ and $\times$ is the vector cross product.

Consider also a target configuration $(\mathbf{p}^*, R^*) = ({}^{D}\mathbf{p}_{I},\ {}^{D}_{I}R) \in \mathrm{SE}(3)$, defined as the configuration of $\{I\}$ with respect to the desired body frame $\{D\}$, which is assumed to be fixed in the workspace, i.e. $\dot{\mathbf{p}}^* = 0$ and $\dot{R}^* = 0$. Introducing the error variables

$$\mathbf{e} = \mathbf{p} - \mathbf{p}^* \in \mathbb{R}^3, \quad R_e = R^{*T}R \in \mathrm{SO}(3), \tag{2}$$

we can write the respective state equations as

$$\dot{\mathbf{e}} = -\mathbf{v} - S(\boldsymbol{\omega})(\mathbf{e} + \mathbf{p}^*) \tag{3a}$$

$$\dot{R}_e = -S(R^{*T}\boldsymbol{\omega})R_e. \tag{3b}$$

As illustrated in Fig. 1, it is assumed that there is a collection of n feature points placed at fixed positions in the environment and that the image coordinates $\mathbf{y}_j$ and $\mathbf{y}_j^* \in \mathbb{R}^2$, $j \in \{1, 2, \ldots, n\}$, acquired at the current and desired configurations $(\mathbf{p}, R)$ and $(\mathbf{p}^*, R^*)$, respectively, are both available for feedback.

The feature points, whose position coordinates in $\{I\}$ are denoted by $\mathbf{x}_j \in \mathbb{R}^3$, $j \in \{1, 2, \ldots, n\}$, are required to satisfy the following assumptions:

Assumption 1: The n points are not all coplanar.

Assumption 2: The origin of $\{I\}$ is placed so that the feature inertial coordinates $\mathbf{x}_j$ verify $\mathbf{x}_j^T\mathbf{n}_\pi \leq 0$ for some vector $\mathbf{n}_\pi \in \mathbb{R}^3$, $j \in \{1, 2, \ldots, n\}$, meaning that all feature points are "below" a plane Π orthogonal to $\mathbf{n}_\pi$ and that crosses the origin of $\{I\}$.

Regarding the geometric condition introduced in Assumption 1, it can be shown that it implies a set of algebraic conditions as stated in the following lemma.

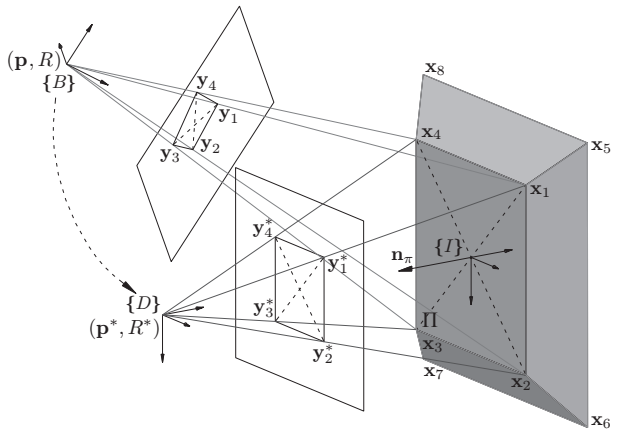


Fig. 1. Setup for the vision-based control problem.

Lemma 2.1: If Assumption 1 is verified then the matrix $X = [\mathbf{x}_1 \cdots \mathbf{x}_n] \in \mathbb{R}^{3\times n}$ is such that *i)* $XX^T > 0$ and *ii)* there is a vector $\mathbf{a} \in \mathbb{R}^n$ such that $X\mathbf{a} = 0$ and $\mathbf{1}^T\mathbf{a} \neq 0$, where $\mathbf{1} = [1 \cdots 1]^T \in \mathbb{R}^n$.

Proof: See [11]. ■

Though only 2-D image measurements are available, the feedback law proposed in this paper will be based on

$$Q = [\mathbf{q}_1\ \ldots\ \mathbf{q}_n] \in \mathbb{R}^{3\times n}, \tag{4}$$

where

$$\mathbf{q}_j = R\mathbf{x}_j + \mathbf{p}, \quad j \in \{1, 2, \ldots, n\}, \tag{5}$$

are the 3-D position coordinates of the feature points expressed in $\{B\}$. Similarly, we define the desired matrix $Q^* = [\mathbf{q}_1^*\ \ldots\ \mathbf{q}_n^*]^T \in \mathbb{R}^{3\times n}$, where $\mathbf{q}_j^* = R^*\mathbf{x}_j + \mathbf{p}^*$.

Using the perspective camera model, the image $\mathbf{y}_j$ of the point $\mathbf{q}_j$ can be written as

$$\mathbf{y}_j = \lambda_j A\mathbf{q}_j \tag{6}$$

where $A \in \mathbb{R}^{2\times 3}$ is the camera calibration matrix assumed to be known and λ_j is an unknown scalar encoding depth information and given by $\lambda_j = (\mathbf{e}_3^T\mathbf{q}_j)^{-1}$, $\mathbf{e}_3 = [0\ 0\ 1]^T$. Therefore, to reconstruct the position coordinates $\mathbf{q}_j$, the depth variables λ_j need to be recovered from the image measurements $\mathbf{y}_j$.

In view of the above, the primary control objective can be defined as that of designing an output-feedback controller that drives $(\mathbf{p}, R)$ to $(\mathbf{p}^*, R^*)$. Of course that in the case of vision-based control systems, a simple convergence result is not sufficient to avoid failure, the FOV problem needs to be explicitly addressed. Thus, we consider the secondary goal of keeping the features inside the camera's FOV along the closed-loop system's trajectories.

As discussed in the introduction, feature loss can occur for two reasons: the features may either leave the camera's FOV or become occluded by the object on which they are marked (see for example the camera configuration $\{B_3\}$ shown in Fig. 2). The likelihood of the first type of feature loss can be greatly reduced by ensuring the asymptotic stability of a relevant error directly defined in the image plane. We

propose that such error be given by the difference between the images of the current and desired position vectors $\mathbf{p}$ and $\mathbf{p}^*$, respectively. Then, the image error vector can be written as

$$\bar{\mathbf{y}}_e = \bar{\mathbf{y}} - \bar{\mathbf{y}}^* \in \mathbb{R}^2, \tag{7}$$

where $\bar{\mathbf{y}} = (\mathbf{e}_3^T\mathbf{p})^{-1}A\mathbf{p}$ and $\bar{\mathbf{y}}^* = (\mathbf{e}_3^T\mathbf{p}^*)^{-1}A\mathbf{p}^*$ are the image coordinates of $\mathbf{p}$ and $\mathbf{p}^*$, respectively. Since $\mathbf{p}$ is the position of the inertial frame $\{I\}$ expressed in the body frame $\{B\}$, it is important to place $\{I\}$ close to the feature points, so that $\bar{\mathbf{y}}_e$ can provide an adequate measure for feature visibility. Also note that even if $\bar{\mathbf{y}}_e = 0$ is asymptotically stable, $\bar{\mathbf{y}}$ may become invalid if $\mathbf{e}_3^T\mathbf{p}$ crosses the origin. Therefore, we will also consider

$$\mathbf{e}_3^T\mathbf{p} > 0, \tag{8}$$

as a necessary condition for keeping feature validity.

To address the second type of feature loss, we introduce the condition

$$\mathbf{n}_\pi^{T\,I}\mathbf{p} = -\mathbf{n}_\pi^T R^T\mathbf{p} > 0, \tag{9}$$

which guarantees that the camera is placed "above" the plane Π. Note that the limit case $\mathbf{n}_\pi^{T\,I}\mathbf{p} = 0$ yields line-segment images for features that belong to Π, such as the image obtained from $\{B_2\}$ in Fig. 2. Thus, this condition is excluded from the valid set.

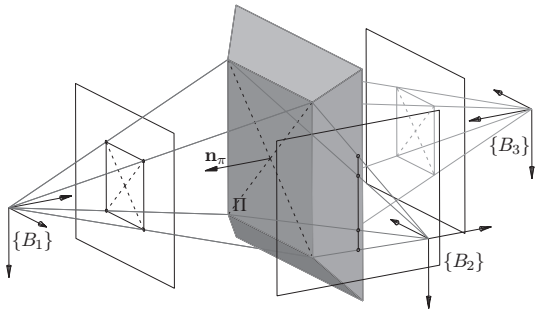


Fig. 2. Valid ($\{B_1\}$) and invalid ($\{B_2\}$ and $\{B_3\}$) configurations.

In summary, the problem addressed in this paper can be stated as follows:

Problem 1: Consider the rigid-body kinematic model described in error coordinates by (3). Design a controller for $\mathbf{v}$ and $\boldsymbol{\omega}$, based on Q and Q^*, such that

- *i)* $(\mathbf{e}, R_e) = (0, I_3)$ is almost globally asymptotically stable, i.e. its region of attraction $\mathcal{R}_A$ coincides with $\mathrm{SE}(3)$ except for a set of zero measure;
- *ii)* a predefined set $\mathcal{J} \subset \mathcal{R}_A$ on which (8) and (9) hold is guaranteed to be positively invariant;
- *iii)* the image error $\bar{\mathbf{y}}_e$ converges to zero exponentially fast.

For the sake of completeness, we recall the definition of positively invariant set [12] and almost Global Asymptotic Stability [13] for a system of the form $\dot{x} = f(x)$. A set M is said to be positively invariant, if for every $x(0) \in M$, $x(t)$ remains in M for all $t \geq 0$. The equilibrium point $x = x^*$ is said to be an almost GAS if it is stable and, except for a set of zero measure, all initial conditions converge asymptotically to it.

III. A Control Law for Visual-Servoing

In the following, we propose a solution to the problem of vision-based control that builds upon the results presented in [10]. To this end, we introduce the angle-axis representation for rotations, according to which $R_e = \mathrm{rot}(\theta, \mathbf{n}) = I_3 + \sin\theta S(\mathbf{n}) + (1 - \cos\theta)S(\mathbf{n})^2$ represents a rotation of angle $\theta \in \mathbb{R}$ about the unitary axis $\mathbf{n} \in \mathbb{S}^2$, and define the function $\mathrm{sign} : \mathbb{R} \mapsto \{1, -1\}$ such that $\mathrm{sign}(x) = 1$ if $x \geq 0$ and $\mathrm{sign}(x) = -1$ if $x < 0$. It is also convenient to present the following proposition, which has been adapted from [10].

Lemma 3.1: Consider the feedback law for $\boldsymbol{\omega}$ given by

$$\boldsymbol{\omega} = k_\omega R^* S^{-1}(R_e NN^T - NN^T R_e^T), \tag{10}$$

where $S^{-1} : \mathrm{so}(3) \mapsto \mathbb{R}^3$ corresponds to the inverse of the skew map S and $N \in \mathbb{R}^{3\times m}$, $m \geq 2$ is such that its two largest singular values verify $\sigma_1 > \sigma_2 > 0$. Then, the interconnection of (3b) and (10) has an almost GAS equilibrium point at $R_e = I_3$ with region of attraction $\mathrm{SO}(3) \setminus \{R_e : \mathrm{tr}(I_3 - R_e) = 4\}$. Moreover, almost every initial condition $\mathbf{n}(0)$ for the axis of rotation $\mathbf{n}(t)$ converges asymptotically to $\mathrm{sign}(\mathbf{n}(0)^T\mathbf{n}_1)\mathbf{n}_1$, where $\mathbf{n}_1$ is a unitary eigenvector of NN^T associated with σ_1^2.

In loose terms, the proposed solution can be described as comprising two sequential steps:

- *i)* aligning the position vector $\mathbf{p}$ with the axis defined by $\mathbf{p}^*$ using solely rotational motion (the inertial position ${}^I\mathbf{p} = -R^T\mathbf{p}$ remains unchanged);
- *ii)* ensuring the convergence of $(\mathbf{p}, R)$ to $(\mathbf{p}^*, R^*)$ using a controller that also guarantees that $\mathbf{n}(t) \to \mathrm{sign}(\mathbf{n}(0)^T\mathbf{n}_1)\mathbf{n}_1$ as $t \to \infty$, where $\mathbf{n}_1 = {}^I\mathbf{p}^*/\|{}^I\mathbf{p}^*\|$.

This strategy will allow for the definition of a positively invariant set $\mathcal{J} \in \mathrm{SE}(3)$ on which both (8) and (9) hold, and therefore depends on the desired configuration $(\mathbf{p}^*, R^*)$ and feature plane Π. The set $\mathcal{J}$ is defined as

$$\mathcal{J} = \mathcal{J}_1 \cap \mathcal{J}_2 \setminus \mathcal{N}_A, \tag{11}$$

where $\mathcal{N}_A$ is a zero measure set that will be explicitely defined shortly and

$$\mathcal{J}_1 = \{(\mathbf{e}, R_e) : \mathbf{e}_3^T(\mathbf{e} + \mathbf{p}^*) > 0\}, \tag{12}$$

$$\mathcal{J}_2 = \{(\mathbf{e}, R_e) : {}^I\mathbf{p} = -R_e^T R^{*T}(\mathbf{e} + \mathbf{p}^*) \in \mathcal{C}\}, \tag{13}$$

with the set $\mathcal{C} \subset \mathbb{R}^3$ given by

$$\mathcal{C} = \Big\{ {}^I\mathbf{p} : {}^{I}\mathbf{p}^{*T\,I}\mathbf{p} > \cos\alpha_\pi \, \|{}^I\mathbf{p}^*\|\|{}^I\mathbf{p}\|, \quad \cos\alpha_\pi = \tfrac{\|S(\mathbf{n}_\pi){}^I\mathbf{p}^*\|}{\|{}^I\mathbf{p}^*\|} \Big\}. \tag{14}$$

As illustrated in Fig. 3, $\mathcal{C}$ defines an unbounded cone in the space of inertial positions, which results from revolving the vector $-S(\mathbf{n}_\pi)^2{}^I\mathbf{p}^*$ (i.e. the projection of ${}^I\mathbf{p}^*$ onto to the feature plane) around ${}^I\mathbf{p}^*$. It is easy to see that $\mathcal{C}$ is placed "above" the plane Π, and therefore (9) holds inside $\mathcal{J}_2$. Also note that the opening angle of the cone has a maximum of $\alpha_\pi = \pi/2$ when ${}^I\mathbf{p}^*$ is perpendicular to the plane (in this case $\mathcal{C}$ coincides with the half-space above the plane) and decreases to zero as ${}^I\mathbf{p}^*$ approaches the plane.

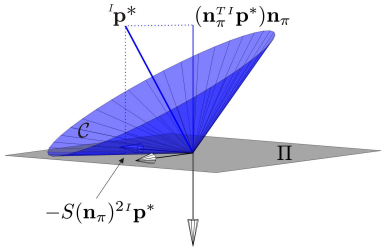


Fig. 3. Set $\mathcal{C} \subset \mathbb{R}^3$ for ${}^I\mathbf{p}^* = [3\ 3\ 8]^T$ and $\mathbf{n}_\pi = \mathbf{e}_3$.

Having described the goals and properties of the proposed solution, a feedback law that yields this result can be constructed as follows:

1) Pick $N \in \mathbb{R}^{3\times m}$, $m \geq 2$ such that its two largest singular values verify $\sigma_1 > \sigma_2 > 0$ and ${}^I\mathbf{p}^*$ is an eigenvector of NN^T associated with σ_1^2.

2) Define the controller

$$\mathbf{v} = \begin{cases} 0 & \text{if } \mathbf{e}+\mathbf{p}^* \notin \mathcal{C}_\gamma, \quad (15a) \\ k_v\mathbf{e} + S(\mathbf{e}+\mathbf{p}^*)\boldsymbol{\omega} & \text{otherwise}, \quad (15b) \end{cases}$$

$$\boldsymbol{\omega} = \begin{cases} -k_0 S(\mathbf{e})\mathbf{p}^* & \text{if } \mathbf{e}+\mathbf{p}^* \notin \mathcal{C}_\gamma, (16a) \\ k_\omega R^* S^{-1}(R_e NN^T - NN^T R_e^T) & \text{otherwise}, \quad (16b) \end{cases}$$

where the set $\mathcal{C}_\gamma$ is given by

$$\mathcal{C}_\gamma = \{\mathbf{p} : \mathbf{p}^T\mathbf{p}^* \geq \cos\gamma \|\mathbf{p}\|\|\mathbf{p}^*\|\}, \quad (17)$$

and γ, k_v, k_0, and k_ω are positive scalars.

The following lemma provides an expression for (15)-(16) in terms of Q and Q^*.

Lemma 3.2: Under Assumptions 1 and 2, the controller defined in (15)-(16) can be rewritten as

$$\mathbf{v} = \begin{cases} 0 & \text{if } Q\mathbf{a} \notin \mathcal{C}_\gamma, \\ k_v(Q-Q^*)\mathbf{a} + S(Q\mathbf{a})\boldsymbol{\omega} & \text{otherwise}, \end{cases} \quad (18)$$

$$\boldsymbol{\omega} = \begin{cases} -k_0 S(Q\mathbf{a})Q^*\mathbf{a} & \text{if } Q\mathbf{a} \notin \mathcal{C}_\gamma, \\ k_\omega S^{-1}(QPP^TQ^* - Q^*PP^TQ^T) & \text{otherwise}, \end{cases} \quad (19)$$

where $P = (I_n - \mathbf{a}\mathbf{1}^T)X^T(XX^T)^{-1}N$.

Proof: See [11]. ■

Remark 3.1: It is straightforward to incorporate an hysteresis in the switching rule for (15)-(16) so as to obtain a chattering-free commutation in the presence of small disturbances.

A. Stability Analysis

In this section, we analyze the stability of the closed-loop system and show that claim *i)* of Problem 1 is verified by the proposed controller, as stated in the following result.

Theorem 3.3: Let Σ denote the closed-loop system that results from the feedback interconnection of (3) and (15)-(16). The point $(\mathbf{e}, R_e) = (0, I_3)$ is an almost GAS equilibrium point of Σ and the corresponding region of attraction is given by $\mathcal{R}_A = \mathrm{SE}(3) \setminus \mathcal{N}_A$, where

$$\mathcal{N}_A = \{(\mathbf{e}, R_e) : \mathrm{tr}(I_3 - R_\gamma(\mathbf{e}+\mathbf{p}^*)R_e) = 4 \text{ or } \mathbf{e} = a\mathbf{p}^*, a < -1\}, \quad (20)$$

and the function $R_\gamma : \mathbb{R}^3 \mapsto \mathrm{SO}(3)$ is given by

$$R_\gamma(\mathbf{p}) = \begin{cases} I_3 & \text{if } \mathbf{p} \in \mathcal{C}_\gamma \text{ or } S(\mathbf{p})\mathbf{p}^* = 0, \quad (21a) \\ \mathrm{rot}\left(\mathrm{acos}(\frac{\mathbf{p}^T\mathbf{p}^*}{\|\mathbf{p}\|\|\mathbf{p}^*\|}) - \gamma, R^{*T}\frac{S(\mathbf{p})\mathbf{p}^*}{\|S(\mathbf{p})\mathbf{p}^*\|}\right) & \text{otherwise}. \quad (21b) \end{cases}$$

To prove Theorem 3.3, we follow a constructive approach that begins by focusing on the position system and then proceeds to analyze the overall closed-loop system. Direct substitution of (15)-(16) in (1a) yields an autonomous system for $\mathbf{p}$, which can be written as

$$\dot{\mathbf{p}} = \begin{cases} -k_0 S(\mathbf{p})^2\mathbf{p}^* & \text{if } \mathbf{p} \notin \mathcal{C}_\gamma, \quad (22a) \\ -k_v(\mathbf{p}-\mathbf{p}^*) & \text{otherwise}. \quad (22b) \end{cases}$$

As illustrated in Fig. 4, when $\mathbf{p}(0) \neq b\mathbf{p}^*$, $b < 0$, $\mathbf{p}(t)$ moves towards $\|\mathbf{p}(0)\|\frac{\mathbf{p}^*}{\|\mathbf{p}^*\|}$ through the shortest arc of circumference that results from connecting $\mathbf{p}(0)$ to $\|\mathbf{p}(0)\|\frac{\mathbf{p}^*}{\|\mathbf{p}^*\|}$ until it reaches the cone $\mathcal{C}_\gamma$. From then on, the motion of $\mathbf{p}(t)$ is governed by (22b) and so it converges to $\mathbf{p}^*$ describing a straight line trajectory. An important result can be derived from this analysis. It is easy to see that once $\mathbf{p}$ is inside the cone $\mathcal{C}_\gamma$ it will not leave that set and consequently there will be at most one switching (reached in finite time).

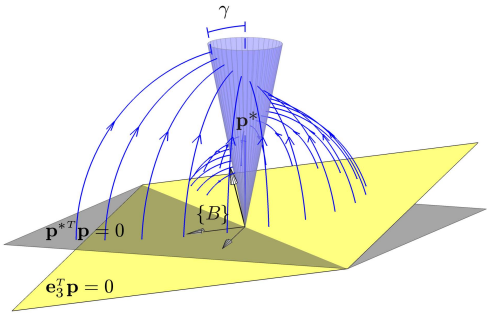


Fig. 4. Phase portrait of the position system (22).

This shows that the time evolution of the system is divided into two stages, determined by the control laws (15a)-(16a) and (15b)-(16b), respectively.

Proof: [Theorem 3.3] Considering the Lyapunov function $V_1 = \frac{1}{2}\mathbf{e}^T\mathbf{e}$ and substituting (15)-(16) in $\dot{V}_1 = \mathbf{e}^T(-\mathbf{v} - S(\boldsymbol{\omega})\mathbf{p})$ yields

$$\dot{V}_1 = \begin{cases} k_0\mathbf{p}^{*T}S(\mathbf{p})^2\mathbf{p}^* \leq 0 & \text{if } \mathbf{p}(t) \notin \mathcal{C}_\gamma, \\ -k_v\mathbf{e}^T\mathbf{e} < 0 & \text{otherwise}. \end{cases} \quad (23)$$

Since the switching condition $\mathbf{p}(t) \notin \mathcal{C}_\gamma$ is reached before $S(\mathbf{p})\mathbf{p}^* = 0$, we have that $\dot{V}_1 < 0$ (provided that $\mathbf{p}(0) \neq a\mathbf{p}^*$, $a < 0$), and therefore $\mathbf{e} = 0$ is an almost GAS equilibrium point of the autonomous position error system.

To analyze the rotation system, assume that the initial condition is given by $(\mathbf{e}(t_0), R_e(t_0))$ and consider the following two cases:

Case 1: $\mathbf{p}(t_0) = \mathbf{e}(t_0) + \mathbf{p}^* \in \mathcal{C}_\gamma$. Since the system is started directly inside the domain of application of the second controller (15b)-(16b), it follows immediately from Lemma 3.1 and (16b) that $R_e = I_3$ is an asymptotically

stable equilibrium point and that $R_e(t) \to I_3$ as $t \to \infty$, provided that $\mathrm{tr}(I_3 - R_e(t_0)) \neq 4$.

Case 2: $\mathbf{p}(t_0) \notin \mathcal{C}_\gamma$. Let $t_1 > t_0$ denote the time instant at which the switching occurs. Then, $R_e(t_1)$ provides the initial condition for the second stage, which according to the arguments of Case 1 will converge to I_3, provided that $\mathrm{tr}(I_3 - R_e(t_1)) \neq 4$. Straightforward derivations, which are omitted for the sake of brevity, show that $R_e(t_1)$ can be written as $R_e(t_1) = R_\gamma(\mathbf{p}(t_0))R_e(t_0)$, where R_γ is given by (21b). ■

B. Additional Stability and Convergence Properties

In this section, we show that claims *ii)* and *iii)* of Problem 1 are also verified by the proposed controller. By analyzing the convergence behavior of the closed-loop system, we will be able to prove that the set $\mathcal{J}$ defined in (11) is positively invariant. Next, we will show that the closed-loop system for the image error vector $\bar{\mathbf{y}}_e$ defined in (7) is asymptotically stable inside $\mathcal{J}$.

Theorem 3.4: Consider the switched system Σ that results from the feedback interconnection of (3) and (15)-(16) and the set $\mathcal{J}$ defined in (11) as a function of the desired configuration $(\mathbf{p}^*, R^*) \in \mathrm{SE}(3)$ and feature plane Π. The set $\mathcal{J}$ is positively invariant with respect to Σ if, given an initial condition $(\mathbf{e}(0), R_e(0)) \in \mathcal{J}$, the design variable γ is such that $0 < \gamma < \frac{1}{2}(\alpha_\pi - \alpha_1)$, where $\alpha_\pi \in (0, \pi/2]$ is the angle between ${}^I\mathbf{p}^*$ and Π and $\alpha_1 \in [0, \alpha_\pi)$ the angle between ${}^I\mathbf{p}^*$ and ${}^I\mathbf{p}(0)$.

Proof: [Outline] Since the positive invariance of a collection of sets implies that their intersection is also positively invariant and $\mathcal{J} = \mathcal{R}_A \cap \mathcal{J}_1 \cap \mathcal{J}_2$, we consider each set separately. The positive invariance of $\mathcal{R}_A$ follows from the fact that it is the region of attraction of an asymptotically stable equilibrium point. By simply observing the phase portrait in Figure 4, it is straightforward to verify that if $\mathbf{e}_3^T\mathbf{p}(0) > 0$ and $\mathbf{e}_3^T\mathbf{p}^* > 0$ then

$$\min_{\mathbf{p}(t)} \mathbf{e}_3^T\mathbf{p}(t) \geq \min\left\{\mathbf{e}_3^T\mathbf{p}(0), \tfrac{\|\mathbf{p}(0)\|}{\|\mathbf{p}^*\|}\mathbf{e}_3^T\mathbf{p}^*, \mathbf{e}_3^T\mathbf{p}^*\right\} > 0, \quad (24)$$

and so $\mathcal{J}_1$ is positively invariant.

Considering now the set $\mathcal{J}_2$, we recall that it can be identified with the set $\mathcal{C}$ defined in (14). It is easy to observe that the invariance condition will not be violated during the first stage, since ${}^I\dot{\mathbf{p}}(t) = R(t)^T\mathbf{v}(t) = 0$ while $\mathbf{p}(t) \notin \mathcal{C}_\gamma$. For the second stage ($\mathbf{p}(t) \in \mathcal{C}_\gamma$), we consider a particular case and show that the intersection set $\mathcal{J}_2 \cap \{(\mathbf{e}, R_e) : \mathbf{e} = 0\}$ is positively invariant. This result can be easily extended for $\mathbf{e} \neq 0$ as shown in [11]. Recalling that $\mathbf{e} = 0$ is an equilibrium point of the position error system, ${}^I\mathbf{p}$ can be written as ${}^I\mathbf{p}(t) = R_e(t)^T{}^I\mathbf{p}^*$ for all $t > 0$ and the function $W : \mathrm{SO}(3) \mapsto \mathbb{R}$ given by $W(R_e) = {}^I\mathbf{p}^{*T}(I_3 - R_e^T){}^I\mathbf{p}^*$ takes the form $W({}^I\mathbf{p}) = {}^I\mathbf{p}^{*T}({}^I\mathbf{p}^* - {}^I\mathbf{p})$. Intersecting the level curves of W with the cone $\mathcal{C}$, it is straightforward to observe that the nonincreasing monotonicity of W guarantees the positive invariance of $\mathcal{J}_2 \cap \{(\mathbf{e}, R_e) : \mathbf{e} = 0\}$. To show that $\dot{W} \leq 0$, recall that ${}^I\mathbf{p}^*$ is an eigenvector of NN^T associated with σ_1^2. Then, $\dot{W}$ can be written as $\dot{W} = -k_\omega {}^I\mathbf{p}^{*T}(\sigma_1^2 I_3 - R_e^T NN^T R_e^T){}^I\mathbf{p}^* \leq 0$. ■

To conclude this section, we analyze the stability of the closed-loop system for the image coordinates $\bar{\mathbf{y}}$. Recalling that $\bar{\mathbf{y}} = (\mathbf{e}_3^T\mathbf{p})^{-1}A\mathbf{p}$, consider applying the coordinate transformation

$$\begin{bmatrix}\bar{\mathbf{y}} \\ p_z\end{bmatrix} = \begin{bmatrix}(\mathbf{e}_3^T\mathbf{p})^{-1}A\mathbf{p} \\ \mathbf{e}_3^T\mathbf{p}\end{bmatrix} \quad (25)$$

to the closed-loop system for $\mathbf{p}$ given in (22). Simple algebra shows that the resulting system can be written as

$$\dot{\bar{\mathbf{y}}} = -k_y(\bar{\mathbf{y}} - \bar{\mathbf{y}}^*),\ k_y = \begin{cases} k_0\|\mathbf{p}\|^2\frac{p_z^*}{p_z} & \text{if } \mathbf{p} \notin \mathcal{C}_\gamma, \\ k_v\frac{p_z^*}{p_z} & \text{otherwise,}\end{cases} \quad (26a)$$

and

$$\dot{p}_z = \begin{cases} -k_0\mathbf{e}_3^T S(\mathbf{p})^2\mathbf{p}^* & \text{if } \mathbf{p} \notin \mathcal{C}_\gamma, \\ -k_v(p_z - p_z^*) & \text{otherwise,}\end{cases} \quad (26b)$$

respectively, where $p_z^* = \mathbf{e}_3^T\mathbf{p}^*$, and $\bar{\mathbf{y}}^* = p_z^{*-1}A\mathbf{p}^*$. By noting that $\mathbf{p}$ can be expressed as a function of $\bar{\mathbf{y}}$ and p_z, more specifically $\mathbf{p} = p_z A_1^{-1}\begin{bmatrix}\bar{\mathbf{y}}^T & 1\end{bmatrix}^T$, where $A_1 = \begin{bmatrix}A^T & \mathbf{e}_3\end{bmatrix}^T$, we can conclude that (26) is an autonomous system and derive the following theorem.

Theorem 3.5: If $p_z^* > 0$, the system given by (26) has an asymptotically stable equilibrium point at $(\bar{\mathbf{y}}, p_z) = (\bar{\mathbf{y}}^*, p_z^*)$ and, inside the positively invariant set $\{(\bar{\mathbf{y}}, p_z) : p_z > 0\}$, $\bar{\mathbf{y}}$ converges exponentially fast to $\bar{\mathbf{y}}^*$ and $\|\bar{\mathbf{y}} - \bar{\mathbf{y}}^*\|$ is monotonically decreasing.

Proof: See [11]. ■

IV. Simulation Results

The simulation results presented in this section attest to the stability and convergence properties of the proposed vision-based controller. To implement the feedback law (15)-(16) given the target configuration $(\mathbf{p}^*, R^*)$, we need to select both the matrix $N \in \mathbb{R}^{3\times m}$ and the set of feature points in the form of matrix $X \in \mathbb{R}^{3\times n}$, with $m \geq 2$ and $n \geq 4$.

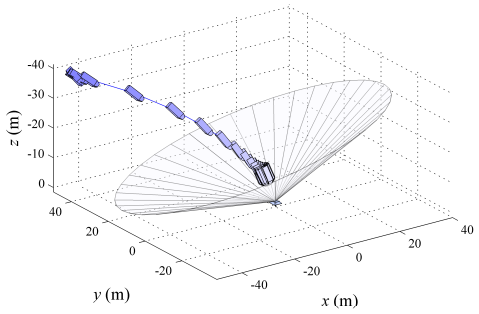
The matrix N is required to be such that the two largest eigenvalues verify $\sigma_1 > \sigma_2 > 0$ and $NN^T{}^I\mathbf{p}^* = \sigma_1^2{}^I\mathbf{p}^*$. Defining the unitary vector $\mathbf{n}_1 = {}^I\mathbf{p}^*/\|{}^I\mathbf{p}^*\|$ and assuming that $\mathbf{n}_1$ and $\mathbf{e}_1 = [1\ 0\ 0]^T$ are not collinear, a possible choice for N is given by

$$N = US,\ U = \begin{bmatrix}\mathbf{n}_1 & \frac{S(\mathbf{n}_1)\mathbf{e}_1}{\|S(\mathbf{n}_1)\mathbf{e}_1\|} & \frac{S(\mathbf{n}_1)^2\mathbf{e}_1}{\|S(\mathbf{n}_1)^2\mathbf{e}_1\|}\end{bmatrix},\ S = \begin{bmatrix}\sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0\end{bmatrix}. \quad (27)$$

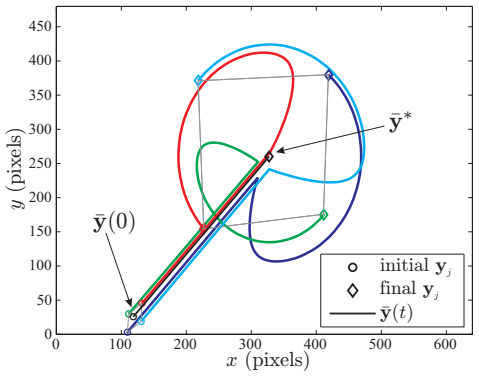
Regarding X, it may seem that it can be formed by virtually any set of feature points satisfying Assumptions 1 and 2. However, since the visual-servoing problem is concerned with keeping feature visibility and the proposed solution only guarantees the positive invariance of $\mathcal{J}$, the matrix X should carefully chosen. To meet the assumptions and ensure that the positive invariance of $\mathcal{J}$ does not lose its significance, we consider a set of $n = 8$ feature points such that $X = \begin{bmatrix}X_1 & X_2\end{bmatrix}$, where $X_1 = \begin{bmatrix}a_1 & a_1 & -a_1 & -a_1 \\ b_1 & -b_1 & -b_1 & b_1 \\ 0 & 0 & 0 & 0\end{bmatrix}$, $X_2 = \begin{bmatrix}a_2 & a_2 & -a_2 & -a_2 \\ b_2 & -b_2 & -b_2 & b_2 \\ -c & -c & -c & -c\end{bmatrix}$, and $0 < a_1 < a_2$, $0 < b_1 < b_2$, and $c > 0$. Note that the choice of feature configurations is not limited to the one just proposed.

As shown in Fig. 1, the feature points correspond to the vertices of a polyhedron that results from chopping the top off a pyramid and the origin of $\{I\}$ coincides with the centroid of the polyhedron's upper face. By aligning the plane Π with this upper face, the positive invariance of $\mathcal{J}$, or equivalently of $\mathcal{C}$, will guarantee that the inertial position ${}^{I}\mathbf{p}(t)$ remains above the features, while converging to ${}^{I}\mathbf{p}^{*} = -R^{*T}\mathbf{p}^{*}$. This choice of feature geometry also simplifies the process of recovering the depth variables from images coordinates, since the 3-D reconstruction algorithm for planar scenes can be directly applied to the pairs (Y_1, X_1) and (Y_2, X_2) [2].

The simulation results that follow were obtained with the polyhedron parameters set to $a_1 = b_1 = 0.56$, $a_2 = b_2 = 1.4$, and $c = 0.28$ and a target position and orientation given by $\mathbf{p}^{*} = [0.1\ 0.25\ 10]^{T}$ and $R^{*} = \mathrm{rot}(-0.4, [0.77\ 0.63\ -0.1]^{T})$, respectively. The corresponding set $\mathcal{C}$, which results from choosing $\mathbf{n}_{\pi} = [0\ 0\ -1]^{T}$, is shown in Fig. 5(a).



(a) 3-D trajectory



(b) 2-D trajectory in the image plane

Fig. 5. System trajectories.

Figure 5 illustrates the convergence behavior that can be achieved with the proposed switching controller. The first and second stages of the trajectory can be easily identified. While the first controller is being applied, the body rotates around itself (see Fig. 5(a)), yielding a translational motion of the image coordinates $\mathbf{y}_j$ in the image plane (see Fig. 5(b)). Regarding the second controller, which takes the body to its desired configuration, the resulting 3-D trajectory involves both rotational and translational motions, which can be identified as a rotation and zooming in of the feature points in the image plane. Figure 6 shows that the position error $\mathbf{e}$ and the angle of rotation θ converge exponentially fast to zero.

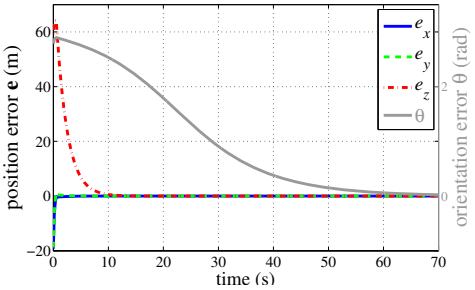


Fig. 6. Time evolution of $\mathbf{e} = [e_x\ e_y\ e_z]^{T}$ and θ.

V. Conclusions

The paper presented a vision-based solution to the problem of stabilization on $\mathrm{SE}(3)$. Based on the image coordinates of a set of feature points and reconstructed depth information, a switching controller was defined to ensure that the features remain visible while the system converges to an almost GAS target configuration. Exponential stability of an error vector directly defined in the image plane was also established. Simulation results were presented, which support the adequacy of the proposed method.

References

[1] N. J. Cowan, J. D. Weingarten, and D. E. Koditschek, "Visual servoing via navigation functions," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 4, pp. 521–533, Aug. 2002.

[2] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry, *An Invitation to 3-D Vision From Images to Geometric Models*, ser. Interdisciplinary Applied Mathematics. Springer, 2004, vol. 26.

[3] S. A. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[4] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The Confluence of Vision and Control*, ser. LNCIS, D. Kriegman, G. . Hager, and A. Morse, Eds. Springer-Verlag, 1998, vol. 237, pp. 66–78.

[5] R. Kelly, R. Carelli, O. Nasisi, B. Kuchen, and F. Reyes, "Stable visual servoing of camera-in-hand robotic systems," *IEEE/ASME Transactions on Mechatronics*, vol. 5, no. 1, pp. 39–48, Mar. 2000.

[6] E. Malis and F. Chaumette, "Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 2, pp. 176–186, Apr. 2002.

[7] B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice, "Position based visual servoing: keeping the object in the field of vision," in *IEEE Conference on Robotics and Automation*, 2002, pp. 1624–1625.

[8] G. Chesi, K. Hashimoto, D. Prattichizzo, and A. Vicino, "Keeping features in the field of view in eye-in-hand visual servoing: A switching approach," *IEEE Transactions on Robotics*, vol. 20, no. 5, pp. 908–913, Oct. 2004.

[9] N. R. Gans and S. A. Hutchinson, "Stable visual servoing through hybrid switched-system control," *IEEE Transactions on Robotics*, vol. 23, no. 3, pp. 530–540, Jun. 2007.

[10] R. Cunha, C. Silvestre, and J. Hespanha, "Output-feedback control for point stabilization on SE(3)," in *IEEE Conference on Decision and Control*, San Diego, CA, Dec. 2006.

[11] R. Cunha, "Advanced motion control for autonomous air vehicles," Ph.D. dissertation, Instituto Superior Técnico, Lisbon, 2007.

[12] H. Khalil, *Nonlinear Systems, Third Edition*. New Jersey: Prentice Hall, 2000.

[13] D. Angeli, "An almost global notion of input-to-state stability," *IEEE Transactions on Automatic Control*, vol. 49, no. 6, pp. 866–874, Jun. 2004.