Chapter 16 Points-Based Visual Servoing with Central Cameras

Hicham Hadj-Abdelkader, Youcef Mezouar, and Philippe Martinet

Abstract This chapter concerns hybrid visual servoing schemes from a set of points viewed by central camera. The main purpose is to decouple the velocity commands in order to obtain an adequate camera trajectory. The proposed schemes are model-free since they are based on the homography matrix between two views. The rotational motions are controlled using the estimated orientation between the current and the desired positions of the robot, while the translational motions are controlled using the combination between image points (onto the sphere or into the normalized plane) and 3D information extracted from the homography matrix. Real-time experimental results with a cartesian manipulator robot are presented and show clearly the decoupling properties of the proposed approaches.

16.1 Introduction

In vision-based control, the choice of the set of visual features to be used in the control scheme is still an open question, despite of the large quantity of results obtained in the last few years. The visual servoing schemes can be classified in three groups: position-based visual servoing (PBVS) [27], image-based visual servoing (IBVS) [7] and hybrid visual servoing [15]. In PBVS, the used information is defined in the 3D space which allow the control scheme to ensure nice decoupling properties between the degrees of freedom (DOF) (refer to [26]). Adequate 3D trajectories can thus be obtained such as a geodesic for the rotational motion and a straight line for the translational motion. However, this kind of control scheme is sensitive to measurement noises and the control may thus suffer from potential instabilities [3]. In IBVS the control is performed in the image space. Whatever the nature of the possible measures extracted from the image, the main question is how

LASMEA, University Blaise Pascal, Campus des Cezeaux, 63177 Aubiere, France, e-mail: {hadj,mezouar,martinet}@lasmea.univ-bpclermont.fr



Hicham Hadj-Abdelkader, Youcef Mezouar and Philippe Martinet

to combine them to obtain an adequate behavior of the system. In most works, the combination of different features is nothing but a simple stacking. If the error between the initial value of the features and the desired one is small, and if the task to realize constrains all the available DOF, that may be a good choice. However, as soon as the error is large, problems may appear such as reaching local minimum or task singularities [3]. Hybrid visual servoing is an alternative to the two previous control schemes. In this case, the visual features gather 2D and 3D information.

The way to design adequate visual features is directly linked to the modeling of their interaction with the robot motion, from which all control properties can be analyzed theoretically. If the interaction is too complex (i.e. highly nonlinear and coupled), the analysis becomes impossible and the behavior of the system is generally not satisfactory in difficult configurations where large displacements (especially rotational ones) have to be realized. To overcome these problems, it is possible to combine path-planning and visual servoing, since tracking planned trajectories allows the error to always remain small [20]. A second approach is to use the measures to build particular visual features that will ensure expected properties of the control scheme (refer for instance to [21, 14, 5, 13, 12, 4, 24]).

This chapter is concerned with homography-based visual servo control techniques with central catadioptric cameras. This framework, also called 2-1/2D visual servoing [15] in the case where the image features are points, exploits a combination of reconstructed Euclidean information and image features in the control design. The 3D information is extracted from an homography matrix relating two views of a reference plane. As a consequence, the 2-1/2D visual servoing scheme does not require any 3D model of the target. Unfortunately, in such approach when conventional cameras are used, the image of the target is not guaranteed to remain in the camera field of view. To overcome this deficiency, 2-1/2D visual servoing is first extended to the entire class of central cameras (including pinhole cameras, central catadioptric cameras and some fisheye cameras [6]). It will be shown that as when a conventional camera is employed, the resulting interaction matrix is blocktriangular with partial decoupling properties. Then two new control schemes will be proposed. The basic idea of the first one is to control the translational motions using a scaled 3D point directly obtained from the image points coordinates and the homography matrix. Compared to the conventional 2-1/2D visual servoing, it allows to obtain better camera trajectory since the translation is controlled in the 3D space while the interaction matrix remains block-triangular. Then, a hybrid scheme which allow us to fully decouple rotational motions from translational ones (i.e the resulting interaction matrix is square block-diagonal) will be proposed. For the three proposed control schemes, it will be also shown that the equilibrium point is globally stable even in the presence of errors in the norm of 3D points which appears in the interaction matrices.

16.2 Modeling

In this section, the unified cental projection model using the unitary sphere is briefly recalled. Then, Euclidean reconstruction from the generic homography matrix is addressed.

16.2.1 Generic Projection Model

Central imaging systems can be modeled using two consecutive projections: spherical projection succeeded by a perspective one. This geometric formulation called unified model has been proposed by Geyer and Daniilidis in [9] and has been intensively used by the vision and robotics community (structure from motion, calibration, visual servoing, etc).



Fig. 16.1 Unified central projection and two views geometry.

Consider the virtual unitary sphere centered in the origin of the mirror frame \mathcal{F}_m as shown in Fig.16.1 and the perspective camera centered in the origin of the camera frame \mathcal{F}_c . Without lost of generality, a simple translation of $-\xi$, along the Z axis of

the mirror frame, between \mathcal{F}_m and \mathcal{F}_c is considered. Let \mathcal{X} be a 3D point with coordinates $\mathbf{X} = [X Y Z]^\top$ in \mathcal{F}_m . The world point \mathcal{X} is projected in the image plane into the point of homogeneous coordinates $\mathbf{x}_i = [x_i y_i 1]^\top$. The image formation process can be split in three steps given in the following:

• first, the 3D world point X is mapped onto the unit sphere surface:

$$\mathbf{X}_{\mathbf{s}} = \frac{1}{\rho} \begin{bmatrix} \mathbf{X} \ \mathbf{Y} \ \mathbf{Z} \end{bmatrix}^{\mathsf{T}},\tag{16.1}$$

where $\rho = ||\mathbf{X}|| = \sqrt{X^2 + Y^2 + Z^2}$.

• then, the point X_s lying on the unitary sphere is perspectively projected on the normalized image plane $Z = 1 - \xi$ into a point of homogeneous coordinates:

$$\underline{\mathbf{x}} = \mathbf{f}(\mathbf{X}) = \left[\frac{X}{Z + \xi\rho} \frac{Y}{Z + \xi\rho} \ 1\right]^{\mathsf{T}}$$
(16.2)

(as it can be seen, the perspective projection model is obtained by setting $\xi = 0$);

finally, the 2D projective point <u>x</u> is mapped into the pixel image point with homogeneous coordinates <u>x</u>_i using the collineation matrix K:

$$\underline{\mathbf{x}}_{\mathbf{i}} = \mathbf{K}\underline{\mathbf{x}}$$

where the matrix \mathbf{K} contains the conventional camera intrinsic parameters coupled with mirror intrinsic parameters, and can be written as:

$$\mathbf{K} = \begin{bmatrix} f_u \ \alpha_{uv} \ u_0 \\ 0 \ f_v \ v_0 \\ 0 \ 0 \ 1 \end{bmatrix}.$$

The matrix **K** and the parameter ξ can be obtained after calibration using for instance the method proposed in [1]. The inverse projection from the image plane onto the unit sphere can be obtained by inverting the second and last steps. As a matter of fact, the point $\underline{\mathbf{x}}$ in the normalized image plane is obtained using the inverse mapping \mathbf{K}^{-1} :

$$\underline{\mathbf{x}} = [x \ y \ 1]^{\top} = \mathbf{K}^{-1} \underline{\mathbf{x}}_{\mathbf{i}}.$$
 (16.3)

The point onto the unit sphere is then obtained by inverting the nonlinear projection (16.2):

$$\mathbf{X}_{\mathbf{s}} = \mathbf{f}^{-1}(\underline{\mathbf{x}}) = \eta \left[x \ y \ 1 - \frac{\xi}{\eta} \right]^{\top}, \tag{16.4}$$

where

$$\eta = \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1}.$$

16.2.2 Scaled Euclidean Reconstruction

Several methods were proposed to obtain the Euclidean reconstruction from two views [8]. They are generally based on the estimation of the essential or homography matrices. The epipolar geometry of cameras obeying the unified model has been recently investigated [10, 23, 11]. For control purposes, the methods based on the essential matrix are not well suited since degenerate configurations such as pure rotation motion can induce unstable behavior of the control scheme. It is thus preferable to use methods based on the homography matrix.

It will be shown now how one can compute the Homographic relationship between two central views of points. Consider two positions \mathcal{F}_m and \mathcal{F}_m^{\star} of the central camera (see Fig.16.1). Those frames are related by the rotation matrix **R** and the translation vector **t**. Let (π) a 3D reference plane given in \mathcal{F}_m^{\star} by the vector $\pi^{\star \top} = [\mathbf{n}^{\star \top} - d^{\star}]$, where \mathbf{n}^{\star} is its unitary normal in \mathcal{F}_m^{\star} and d^{\star} is the distance from (π) to the origin of \mathcal{F}_m^{\star} .

Let \mathcal{X} be a 3D point with coordinates $\mathbf{X} = [X Y Z]^{\top}$ with respect to \mathcal{F}_m and with coordinates $\mathbf{X}^* = [X^* Y^* Z^*]^{\top}$ with respect to \mathcal{F}_m^* . Its projection in the unit sphere for the two camera positions is given by the coordinates $\mathbf{X}_s = \rho^{-1} \mathbf{X}$ and $\mathbf{X}_s^* = \rho^{*-1} \mathbf{X}^*$. The distance $d(\mathcal{X}, \pi)$ from the world point \mathcal{X} to the plane (π) is given by the scalar product $[\mathbf{X}^{\star^{\top}} \ 1] \cdot \pi^*$:

$$d(X,\pi) = \rho^* \mathbf{n}^{\star \top} \mathbf{X}_{\mathbf{s}}^* - d^*.$$
(16.5)

The relationship between the coordinates of X with respect to \mathcal{F}_m and \mathcal{F}_m^* can be written as a function of their spherical coordinates:

$$\rho \mathbf{X}_{\mathbf{s}} = \rho^* \mathbf{R} \mathbf{X}_{\mathbf{s}}^* + \mathbf{t}.$$
 (16.6)

By multiplying and dividing the translation vector by the distance d^* and according to (16.5), the expression (16.6) can be rewritten as:

$$\rho \mathbf{X}_{\mathbf{s}} = \rho^* \mathbf{H} \mathbf{X}_{\mathbf{s}}^* + \alpha \mathbf{t}, \tag{16.7}$$

with $\mathbf{H} = \mathbf{R} + \frac{\mathbf{t}}{d^{\star}} \mathbf{n}^{\star \top}$ and $\alpha = -\frac{d(X, \pi)}{d^{\star}}$. **H** is the Euclidean homography matrix written as a function of the camera displacement and of the plane coordinates with respect to \mathcal{F}_m^{\star} . It has the same form as in the conventional perspective case (it can be decomposed into a rotation matrix and a rank 1 matrix). If the world point X belongs to the reference plane (π) (i.e. $\alpha = 0$) then (16.7) becomes:

$X_s \propto H X_s^{\star}$.

The homography matrix **H** related to the plane (π) can be estimated up to a scale factor by solving the linear equation $\mathbf{X}_{s} \otimes \mathbf{H}\mathbf{X}_{s}^{\star} = 0$ (where \otimes denotes the crossproduct) using, at least, four couples of coordinates (\mathbf{X}_{sk} ; \mathbf{X}_{sk}^{\star}) (where $k = 1 \cdots n$ with $n \ge 4$), corresponding to the spherical projection of world points X_{k} belonging to (π). If only three points belonging to (π) are available then at least five supplementary points are necessary to estimate the homography matrix by using for example the linear algorithm proposed in [15].

From the estimated homography matrix, the camera motion parameters (that is the rotation **R** and the scaled translation $\mathbf{t}_{d^*} = \frac{1}{d^*}\mathbf{t}$) and the structure of the observed scene (for example the vector \mathbf{n}^*) can thus be determined (refer to [8, 28]). It can also be shown that the ratio $\sigma = \frac{\rho}{\rho^*}$ can be computed as:

$$\sigma = \frac{\rho}{\rho^{\star}} = \det(\mathbf{H}) \frac{\mathbf{n}^{\star \top} \mathbf{X}_{\mathbf{s}}^{\star}}{\mathbf{n}^{\star \top} \mathbf{R}^{\top} \mathbf{X}_{\mathbf{s}}}.$$
 (16.8)

In the sequel, the rotation parameters and the ratio σ , extracted from the estimated homography are used to define the task function for the proposed hybrid visual servoing schemes.

16.3 Visual Servoing

16.3.1 Task Function and Interaction Matrices

As usual when designing a visual servoing scheme, the visual feature vector \mathbf{s} is often expressed as function of the 3D representation of the observed object such as a set of 3D points. In order to control the movements of the robotic system from visual features, one defines a task function to be regulated to $\mathbf{0}$ as [22] :

$$\mathbf{e} = \mathbf{L}^+(\mathbf{s} - \mathbf{s}^\star),$$

where \cdot^+ denote the pseudo-inverse and **L** is the interaction matrix which links the variation of **s** to the camera velocities. If the observed object is motionless, one gets:

$$\dot{\mathbf{s}} = \mathbf{L}\boldsymbol{\tau},$$

where τ is a 6D vector denoting the velocity screw of the central camera. The vector τ contains the instantaneous linear velocity **v** and the instantaneous angular velocity ω of the sensor frame expressed in the same frame. In the sequel, the sensor frame is chosen as the mirror frame \mathcal{F}_m .

A simple control law can be designed by imposing an exponential decay of the task function e toward 0:

$$\dot{\mathbf{e}} = -\lambda \mathbf{e},$$

where λ is a proportional gain. The corresponding control law is:

$$\tau = -\lambda \mathbf{L}^+ (\mathbf{s} - \mathbf{s}^\star). \tag{16.9}$$

16 Points-Based Visual Servoing with Central Cameras

In order to compute the control law (16.9), the interaction matrix **L** or its pseudoinverse (its inverse if **L** is square) should be provided. In practice, an approximation $\widehat{\mathbf{L}}$ of the interaction matrix is used. If the task function **e** is correctly computed, the global asymptotic stability of the system can be obtained if the necessary and sufficient condition $\mathbf{L}\widehat{\mathbf{L}}^+ > 0$ is satisfied.

When the visual features are related to the projection of 3D points, the vector **s** is function of the 3D coordinates $\mathbf{X} = [X \ Y \ Z]^{\top}$ of the 3D point X. In that case, the interaction matrix related to **s** can be written as:

$$\mathbf{L} = \frac{\partial \mathbf{s}}{\partial \mathbf{X}} \mathbf{L}_{\mathbf{X}},$$

 $J_s = \frac{\partial s}{\partial X}$ is the Jacobian matrix linking the variations of s and X, and L_X is the interaction matrix related to the 3D point X:

$$\dot{\mathbf{X}} = \mathbf{L}_{\mathbf{X}}\tau = \begin{pmatrix} -\mathbf{I}_3 & [\mathbf{X}]_{\times} \end{pmatrix}\tau, \tag{16.10}$$

where $[\mathbf{a}]_{\times}$ is the anti-symmetric matrix of the vector \mathbf{a} .

If one considers *n* visual features related to the same 3D point *X*, the global interaction matrix **L** for the features vector $\mathbf{s} = [\mathbf{s_1} \ \mathbf{s_2} \ \cdots \ \mathbf{s_n}]^\top$ can be written:

$$\mathbf{L} = \left(\mathbf{J}_{\mathbf{s}_1}^{\top} \mathbf{J}_{\mathbf{s}_2}^{\top} \cdots \mathbf{J}_{\mathbf{s}_n}^{\top} \right)^{\top} \mathbf{L}_{\mathbf{X}}.$$

16.3.2 Interaction Matrix for 2D Point

Consider a 3D point X with coordinates $\mathbf{X} = [X Y Z]^{\top}$ with respect to the mirror frame \mathcal{F}_m . Its central projection on the normalized image plane is obtained using (16.1) and it is given by the point of homogeneous coordinates $\mathbf{x} = [x \ y \ 1]^{\top}$. If the visual feature is chosen as $\mathbf{s} = [x \ y]^{\top}$, the interaction matrix \mathbf{L} is:

$$\mathbf{L} = \mathbf{J}_{\mathbf{S}} \mathbf{L}_{\mathbf{X}}$$

where

$$\mathbf{J}_{\mathbf{s}} = \frac{1}{\rho(Z + \xi\rho)^2} \begin{pmatrix} \rho Z + \xi(Y^2 + Z^2) & -\xi XY & -X(\rho + \xi Z) \\ -\xi XY & \rho Z + \xi(X^2 + Z^2) & -Y(\rho + \xi Z) \end{pmatrix}$$

After few developments, the analytical expression of the interaction matrix L can be written as:

$$\mathbf{L} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}, \tag{16.11}$$

where

$$\mathbf{A} = \rho^{-1} \begin{pmatrix} -\frac{\gamma + \xi (x^2 + y^2)}{1 + \xi \gamma} + \xi x^2 & \xi xy & \gamma x \\ \xi xy & -\frac{\gamma + \xi (x^2 + y^2)}{1 + \xi \gamma} + \xi y^2 & \gamma y \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} xy & -\gamma \frac{\gamma + \xi \left(x^2 + y^2\right)}{1 + \xi \gamma} + y^2 & y \\ \gamma \frac{\gamma + \xi \left(x^2 + y^2\right)}{1 + \xi \gamma} - x^2 & -xy & -x \end{pmatrix}$$

with $\gamma = \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}$.

16.3.3 Decoupled Visual Servoing

In visual servoing scheme, the control properties are directly linked to the interaction between the designed features and the camera (or the robot) motion. The behavior of the camera depends on the coupling between the features and the camera velocities. For example, the interaction matrix in (16.11) related to the image coordinates of 2D points is highly nonlinear and coupled. Thereof, large displacements of the camera became difficult to realize.

Several approaches have been proposed to overcome these problems. Most of them ensure a good decoupling properties by combining 2D and 3D information when defining the input of the control law. The related control schemes are called hybrid visual servoing. In this work, three model free decoupled control schemes are proposed. Let us first define the observation vector as:

$$\mathbf{s} = \left[\begin{array}{cc} \mathbf{\tilde{s}}^\top & \theta \mathbf{u}^\top \end{array} \right]^\top.$$

The vector $\tilde{\mathbf{s}}$ is chosen to be variant to the translational motions of the camera and can be variant or invariant to the rotational motions, whereas the vector $\theta \mathbf{u}$, representing the rotational information between the current and the desired positions of the camera, is invariant to the translational motions. Consequently, the global interaction matrix \mathbf{L} related to the features vector \mathbf{s} is a block-triangular matrix:

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_{\mathbf{\tilde{s}}_{\mathbf{v}}} & \mathbf{L}_{\mathbf{\tilde{s}}_{\omega}} \\ \mathbf{0}_{\mathbf{3}} & \mathbf{L}_{\omega} \end{pmatrix}.$$

Note that when \tilde{s} is invariant to rotational motions, L becomes a block-diagonal matrix.

16.3.3.1 Interaction Matrix L_{ω}

The rotation matrix between the current and the desired positions of the central camera, can be obtained from the estimated homography matrix **H**. Several representations of the rotation are possible. The representation θ **u** (where θ is the rotation angle and **u** is a unit vector along the rotation axis) is chosen since it provides the largest possible domain for the rotation angle. The corresponding interaction matrix can be obtained from the time derivative of θ **u** since it can be expressed with respect

316 and 16 Points-Based Visual Servoing with Central Cameras

to the central camera velocity screw τ :

$$\frac{d(\theta \mathbf{u})}{dt} = \begin{pmatrix} \mathbf{0}_3 & \mathbf{L}_\omega \end{pmatrix} \tau,$$

where \mathbf{L}_{ω} is given by [17]:

$$\mathbf{L}_{\omega} = \mathbf{I}_{3} - \frac{\theta}{2} [\mathbf{u}]_{\times} + \left(1 - \frac{\operatorname{sinc}(\theta)}{\operatorname{sinc}^{2}(\frac{\theta}{2})}\right) [\mathbf{u}]_{\times}^{2}.$$

Note also that theoretically in this case $\mathbf{L}_{\omega}^{-1} \theta \mathbf{u} = \theta \mathbf{u}$. This nice property can advantageously be exploited to compute the control vector.

In practice, estimated camera parameters are used. The estimated rotation parameter $\widehat{\theta u}$ can be written as a nonlinear function of the real ones $\psi(\theta u)$. Since $\widehat{\mathbf{L}_{\omega}^{-1}} \ \widehat{\theta u} = \widehat{\theta u}$, the closed-loop equation of the rotation control is:

$$\frac{\theta \mathbf{u}}{dt} = -\lambda \mathbf{L}_{\omega} \psi(\theta \mathbf{u}).$$

The asymptotic stability of this system has been studied for conventional camera $(\xi = 0)$ since in this case the function ψ has a simple analytical form [16]. However, the stability analysis remains an open problem when $\xi \neq 0$ since the nonlinear function ψ is much more complex in this case.

16.3.3.2 2-1/2D Visual Servoing

2-1/2D visual servoing has been first proposed by Malis and Chaumette in case of conventional camera ($\xi = 0$). In this section, the original scheme is extended to the entire class of central cameras. In order to control the translational motion, let us define \tilde{s} as:

$$\mathbf{\tilde{s}} = \begin{bmatrix} \mathbf{\tilde{s}_1}^\top & \mathbf{\tilde{s}_2} \end{bmatrix}^\top,$$

where $\tilde{\mathbf{s}}_1 = [x \ y]^\top$ and $\tilde{\mathbf{s}}_2 = \log(\rho)$ are respectively the coordinates of an image point and the logarithm of the norm of its corresponding 3D point. The error between the current value $\log(\rho)$ and the desired value $\log(\rho^*)$ can be estimated using (16.8) since $\tilde{\mathbf{s}}_2 - \tilde{\mathbf{s}}_2^* = \log(\sigma)$.

The corresponding interaction matrix $L_{\tilde{s}}$ can be written as:

$$\mathbf{L}_{\tilde{\mathbf{s}}} = \left(\left. \mathbf{J}_{\tilde{\mathbf{s}}_1}^{\top} \right. \left. \mathbf{J}_{\tilde{\mathbf{s}}_2}^{\top} \right)^{\top} \mathbf{L}_X,$$

where the Jacobian matrix $J_{\tilde{s}_1}$ is given by (16.11), and $J_{\tilde{s}_2}$ can be easily computed:

$$\mathbf{J}_{\mathbf{\tilde{s}}_2} = \rho^{-2} \mathbf{X}^\top.$$

 $\mathbf{L}_{\tilde{\mathbf{s}}} = \begin{pmatrix} \mathbf{L}_{\tilde{\mathbf{s}}_{\mathbf{v}}} & \mathbf{L}_{\tilde{\mathbf{s}}_{\omega}} \end{pmatrix}$ can be obtained by stacking the interaction matrix in (16.11) and:

Hicham Hadj-Abdelkader, Youcef Mezouar, and Philippe Martinet

$$\mathbf{L}_{\tilde{\mathbf{s}}_{2}} = \mathbf{J}_{\tilde{\mathbf{s}}_{2}} \mathbf{L}_{\mathbf{X}} = \frac{1}{\sigma \rho^{\star}} \Big(-\Phi_{X} - \Phi_{Y} \Phi_{\frac{\xi^{2}(x^{2} + y^{2}) - 1}{1 + \gamma \xi}} 0 0 0 \Big),$$
(16.12)

with $\Phi = \frac{Z+\xi\rho}{\rho} = \frac{1+\gamma\xi}{\gamma+\xi(x^2+\gamma^2)}$. Note that the parameter ρ^* can be estimated only once during an off-line learning stage. If the system is supposed correctly calibrated and that measurements are noiseless, then the control law is asymptotically stable for any positive value ρ^* . However, the robustness with respect to calibration and measurement errors still remains an open problem.

16.3.3.3 Norm-ratio-based Visual Servoing

As it can be seen in (16.12), the ratio between ρ and ρ^* is invariant to rotational motion. In the sequel, this property will be exploited in a new control scheme allowing us to decouple translational motions from the rotational ones. At this end, let us now define \tilde{s} as:

$$\mathbf{\tilde{s}} = \left[\log(\rho_1) \log(\rho_2) \log(\rho_3)\right]^{\top}.$$

The interaction matrix $J_{\tilde{s}}$ corresponding to \tilde{s} is obtained by stacking the interaction matrices given by (16.12) for each point. In this case, the global interaction matrix L is a block-diagonal matrix:

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_{\mathbf{\tilde{s}}_{\mathbf{v}}} & \mathbf{0}_{\mathbf{3}} \\ \mathbf{0}_{\mathbf{3}} & \mathbf{L}_{\omega} \end{pmatrix}.$$

As above-mentioned, the translational and rotational controls are fully decoupled. If the system is correctly calibrated and the measurements are noiseless, the system is stable since $\frac{\widehat{\rho^{\star}_{i}}}{\rho_{i}^{\star}}$ is positive.

16.3.3.4 Scaled 3D Point-based Visual Servoing

Visual servoing scheme based on 3D points benefits of nice decoupling properties [19] [2]. Recently, Tatsambon *et al.* show in [25] that similar decoupling properties than the ones obtained with 3D points can be obtained using visual features related to the spherical projection of a sphere: the 3D coordinates of the center of the sphere computed up to a scale (the inverse of the sphere radius). However, even if such an approach is theoretically attractive, it is limited by a major practical issue since spherical object has to be observed.

Consider a 3D point X with coordinates $\mathbf{X} = [X Y Z]^{\top}$ with respect to the frame \mathcal{F}_m . The corresponding point onto the unit sphere is \mathbf{X}_s and $\mathbf{X} = \rho \mathbf{X}_s$.

Let us now choose \tilde{s} as:

$$\tilde{\mathbf{s}} = \sigma \, \mathbf{X}_{\mathbf{s}} = \frac{1}{\rho^{\star}} \, \mathbf{X},\tag{16.13}$$

318

where ρ^* is the 2-norm of X with respect to the desired position \mathcal{F}^* of the camera. The feature vector \tilde{s} is thus defined as a vector containing the 3D point coordinates up to a constant scale factor. Its corresponding interaction matrix can be obtained directly from (16.10):

$$\mathbf{L}_{\tilde{\mathbf{s}}} = \frac{1}{\rho^{\star}} \mathbf{L}_{\mathbf{X}} = \left(-\frac{1}{\rho^{\star}} \mathbf{I}_{\mathbf{3}} \ [\tilde{\mathbf{s}}]_{\times} \right).$$

As it is shown in the expression of $L_{\tilde{s}}$, the only unknown parameter is ρ^* which appears as a gain on the translational velocities. A nonzero positive value attributed to ρ^* will thus ensure the global asymptotic stability of the control law. The ratio between the real value of ρ^* and the estimated one ρ^* will act as an over-gain in the translational velocities.

Note that a similar approach using conventional camera has been proposed by Malis and Chaumette in [16] in order to enhance the stability domain. However, an adaptive control law has to be used in order that the reference point remains in the camera field of view during the servoing task. This is not a crucial issue in our case since our approach can be used with a large field of view.

16.4 Results

The proposed hybrid visual servoing schemes have been validated with a series of experiments. They were carried out on a 6 DOF manipulator robot in eye-in-hand configuration. A fisheye camera is mounted on the end-effector of the robot (see Fig. 16.2). The estimated camera calibration parameters are $\xi = 1.634$, $f_u = 695$, $f_v = 694.9$, $\alpha_{uv} = 0$, $u_0 = 400.4$ and $v_0 = 304.4$. In order to simplify the features extraction and tracking, the target is composed of a set of white marks printed into a black background. These marks are tracked and their centers of gravity are extracted using the VISP library [18]. The experiments are detailed in the sequel by denoting with:

- A the 2D point-based control law,
- **B** the hybrid scheme presented in Section 16.3.3.2,
- C the hybrid scheme presented in Section 16.3.3.3,
- **D** the hybrid scheme presented in Section 16.3.3.4.

Experiment 1. A large generic displacement is considered. It is composed of a translation $\mathbf{t} = [80\ 80\ -40]$ cm and of a rotation $\theta \mathbf{u} = [0\ 50\ 140]$ deg. The behaviors of the proposed control schemes are compared with conventional IBVS. Since a very large rotation about the *Z*-axis (around 140 deg) is considered, the control **A** fails, the robot reaching quickly its joint limits. The rotation about the *Z*-axis is thus reduced to 40 deg for the control **A**.

Fig. 16.3 shows the results obtained using the control **A**. The interaction matrix depends on 3D parameters, points coordinates and calibration parameters. If one



Fig. 16.2 Experimental setup: Eye-in-Hand configuration.

supposes that the camera-robot system is correctly calibrated and that the measurements are noiseless, the 3D parameters should be accurately estimated to guarantee a quasi-exponential decreasing of the task function **e** (leading to straight line trajectories of the points in the image plane). In this experiment, the 3D parameters ρ_i (which appears in the interaction matrix (16.11)) are set to a constant values $\widehat{\rho_i} = \widehat{\rho^*}_i$ (where $\widehat{\rho_i^*}$ denotes the estimated value of ρ_i at the desired configuration). Consequently, the points trajectories are no more straight lines until around the 300th iteration (where ρ_i became very close to ρ_i^*). After the 300th iteration, one can observe that the errors are decreasing exponentially and the image trajectories became roughly straight.

The results obtained with the hybrid control **B**, **C** and **D** are shown in Fig. 16.4, Fig. 16.5 and Fig. 16.6 respectively. The parameter ρ^* is set to $\widehat{\rho^*} = 2\rho^*$ in those cases.

It can be first observed that the three control laws allow to achieve the large rotation about the Z-axis (i.e. 140 deg) and that, as expected, a rough estimation of the parameter ρ^* does not affect the system stability. It can be also observed that the decoupling properties have been significantly improved with respect to the 2D points visual servoing. Finally, let us note that, in Fig. 16.4(b) the trajectory of the point used to define the 2-1/2D task function should be a straight line. This is clearly not observed since once again ρ is not correctly estimated.

The control **C** allows to fully decouple translational and rotational motions. However, the computation of the 3D features $\frac{\rho_i}{\rho_i^*}$ increase the sensitivity of the control scheme to measurement noise as it can be observed in Fig. 16.5 (see between 200th and 300th iterations). The control law **D** provides very nice decoupling properties (refer to Fig. 16.6). In this case, translational velocities are directly related to the visual features (used to control the translational DOF) through a constant diagonal

16 Points-Based Visual Servoing with Central Cameras



Fig. 16.3 A, 2D points-based visual servoing: (a) initial image; (b) desired image and image-points trajectories; (c) translational velocities in m/s; (d) rotational velocities in deg/s; and (e) error vector components.

matrix. Furthermore, it can be observed that this control scheme is less sensitive to noise measurement than the previous one.

Experiment 2. In this set of experiments, the three hybrid schemes are compared when only a translational motion $\mathbf{t} = [80\ 80\ -40]$ cm has to be realized. The results are shown in Fig. 16.7. It can be observed that the behavior of the three control schemes is similar. These results confirm also that the control scheme based on the features $\frac{\rho_i}{\rho_i^*}$ seems to be the most sensitive to measurement noises. One can also



Fig. 16.4 B, 2-1/2D visual servoing: (a) initial image; (b) desired image and image-points trajectories; (c) translational velocities in m/s; (d) rotational velocities in deg/s; and (e) error vector components.

observe a nonzero rotational velocities at the beginning due to the measurement noises and calibration errors.

Experiment 3. In this set of experiments, only a rotational motion about the *Z*-axis is considered. The control laws **B**, **C** and **D** are first tested with a huge rotation of 140 deg. In this case, only the control law **C** allows to reach the desired configuration. When using the control laws **B** and **D** the robot reached its joint limits due to the coupling between rotational and translational motions. In the results shown in Fig. 16.8, the rotation about the *Z*-axis is reduced to 90 deg for the control laws

16 Points-Based Visual Servoing with Central Cameras



Fig. 16.5 C, norm-ratio-based visual servoing: (a) initial image; (b) desired image and imagepoints trajectories; (c) translational velocities in m/s; (d) rotational velocities in deg/s; and (e) error vector components.

B and **D**. Finally, the full decoupling between translational and rotational motions provided by the control scheme **C** can be clearly observed.

16.5 Conclusion

In this chapter, it has been shown how a generic projection model can be exploited to design vision-based control laws valid for all cameras obeying the unique



Fig. 16.6 D, scaled 3D point-based visual servoing: (a) initial image; (b) desired image and imagepoints trajectories; (c) translational velocities in m/s; (d) rotational velocities in deg/s; and (e) error vector components.

viewpoint constraint. First, the problem of estimating Homographic relationship between two spherical views related to a reference plane has been addressed. Then, three homography-based control schemes have been presented. The task functions are defined to allow as much as possible nice decoupling properties of the control laws. In all cases, the rotational control is achieved using the orientation error extracted from the estimated homography matrix. In the first control scheme, the visual features used to control the translational motions are chosen as the combination of the 2D coordinates of an image point and the ratio of the norms of the corresponding

16 Points-Based Visual Servoing with Central Cameras



Fig. 16.7 A comparison between the hybrid visual servoing schemes under a pure translation displacement: (a) control scheme in Section 16.3.3.2; (b) control scheme in Section 16.3.3.3; and (c) control scheme in Section 16.3.3.4.

3D point at the current and desired configurations (which can be computed from the homography matrix). In a second control scheme, a scaled 3D point, computed from the corresponding image point and the homography matrix, is exploited to control efficiently the translations. It allows to obtain properties similar to 3D point-based visual servoing while being model free. The last control law allows to fully decouple translational and rotational motions (the interaction matrix is block-diagonal) by employing three ratios of the norms related to three 3D points. From a practical point of view, large camera motions can be achieved since the developed control laws are partially or fully decoupled and valid for a large class of wide field of view cameras. Experimental results have confirmed this last point. The stability analysis under modeling errors of the proposed control laws still remain an important theoretical point to be addressed in future works.

References

[1] Barreto J, Araujo H (2002) Geometric properties of central catadioptric line images. In: 7th European Conference on Computer Vision, Copenhagen, Den-



Fig. 16.8 A comparison between the hybrid visual servoing schemes under a pure rotational displacement: (a) control scheme in Section 16.3.3.2; (b) control scheme in Section 16.3.3.3; and (c) control scheme in Section 16.3.3.4.

mark, pp 237-251

- [2] Cervera E, Pobil APD, Berry F, Martinet P (2003) Improving image-based visual servoing with three-dimensional features. International Journal of Robotics Research 22(10-11):821–840
- [3] Chaumette F (1998) Potential problems of stability and convergence in imagebased and position-based visual servoing. The Confluence of Vision and Control, D Kriegman, G Hager, A Morse (eds), LNCIS Series, Springer Verlag 237:66–78
- [4] Chaumette F (2004) Image moments: A general and useful set of features for visual servoing. IEEE Transaction on Robotics and Automation 20(4):713– 723
- [5] Corke PI, Hutchinson SA (2001) A new partitioned approach to imagebased visual servo control. IEEE Transaction on Robotics and Automation 17(4):507–515
- [6] Courbon J, Mezouar Y, Eck L, Martinet P (2007) A generic fisheye camera model for robotic applications. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, California, USA, pp 1683–1688
- [7] Espiau B, Chaumette F, Rives P (1992) A new approach to visual servoing in robotics. IEEE Transaction on Robotics and Automation 8(3):313–326

- [8] Faugeras O, Lustman F (1988) Motion and structure from motion in a piecewise planar environment. International Journal of Pattern Recognition and Artificial Intelligence 2(3):485–508
- [9] Geyer C, Daniilidis K (2000) A unifying theory for central panoramic systems and practical implications. In: European Conference on Computer Vision, Dublin, Ireland, pp 159–179
- [10] Geyer C, Daniilidis K (2003) Mirrors in motion: Epipolar geometry and motion estimation. In: International Conference on Computer Vision, Nice, France, pp 766–773
- [11] Hadj-Abdelkader H, Mezouar Y, Andreff N, Martinet P (2005) 2 1/2 d visual servoing with central catadioptric cameras. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, Canada, pp 2342–2347
- [12] Hamel T, Mahony R (2002) Visual servoing of an under-actuated dynamic rigid body system: an image-based approach. IEEE Transaction on Robotics and Automation 18(2):187–198
- [13] Iwatsuki M, Okiyama N (2002) A new formulation of visual servoing based on cylindrical coordinates system with shiftable origin. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland, pp 266–273
- [14] Lee JS, Suh I, You BJ, Oh SR (1999) A novel visual servoing approach involving disturbance observer. In: IEEE International Conference on Robotics and Automation, Detroit, Michigan, pp 269–274
- [15] Malis E, Chaumette F (2000) 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
- [16] Malis E, Chaumette F (2002) Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods. IEEE Transaction on Robotics and Automation 18(2):176–186
- [17] Malis E, Chaumette F, Boudet S (1999) 2 1/2 d visual servoing. IEEE Transaction on Robotics and Automation 15(2):238–250
- [18] Marchand E, Spindler F, Chaumette F (2005) Visp for visual servoing: a generic software platform with a wide class of robot control skills. IEEE Robotics and Automation Magazine 12(4):40–52
- [19] Martinet P, Daucher N, Gallice J, Dhome M (1997) Robot control using 3d monocular pose estimation. In: Workshop on New Trends in Image Based Robot Servoing, IEEE/RSJ International Conference on Intelligent Robots and Systems, Grenoble, France, pp 1–12
- [20] Mezouar Y, Chaumette F (2002) Path planning for robust image-based control. IEEE Transaction on Robotics and Automation 18(4):534–549
- [21] Rives P, Azinheira J (2004) Linear structures following by an airship using vanishing points and horizon line in a visual servoing scheme. In: IEEE International Conference on Robotics and Automation, New Orleans, Louisiana, pp 255–260
- [22] Samson C, Espiau B, Borgne ML (1991, ISBN: 0198538057) Robot Control: The Task Function Approach. Oxford University Press

¹⁶ Points-Based Visual Servoing with Central Cameras

- [23] Svoboda T, Pajdla T, Hlavac V (1998) Motion estimation using central panoramic cameras. In: IEEE Conference on Intelligent Vehicles, Stuttgart, Germany, pp 335–340
- [24] Tahri O, Chaumette F, Mezouar Y (2008) New decoupled visual servoing scheme based on invariants projection onto a sphere. In: IEEE International Conference on Robotics and Automation, Pasadena, CA, pp 3238–3243
- [25] Tatsambon Fomena R, Chaumette F (2007) Visual servoing from spheres using a spherical projection model. In: IEEE International Conference on Robotics and Automation, Roma, Italia, pp 2080–2085
- [26] Thuilot B, Martinet P, Cordesses L, Gallice J (2002) Position based visual servoing : keeping the object in the field of vision. In: IEEE International Conference on Robotics and Automation, Washington DC, USA, pp 1624–1629
- [27] Wilson W, Hulls CW, Bell G (1996) Relative end-effector control using cartesian position-based visual servoing. IEEE Transaction on Robotics and Automation 12(5):684–696
- [28] Zhang Z, Hanson AR (1995) Scaled euclidean 3d reconstruction based on externally uncalibrated cameras. In: IEEE Symposium on Computer Vision, Coral Gables, FL, pp 37–42

328