Three-View Uncalibrated Visual Servoing

Azad Shademan and Martin Jägersand

Abstract— **A new uncalibrated visual servoing approach for motion control of 6-DOF manipulators is presented. Instead of image features, the elements of the scene-independent trifocal tensor of three views is used as features for visual servoing. These trifocal features depend only on the camera projection matrices and can be retrieved from point correspondences across the three views: initial, current, and desired. The Jacobian matrix that relates joint velocities and variations of the trifocal features is estimated online using least-squares. The visual servoing performance is evaluated for various easy and challenging motions, such as the translational motion along the view axis and the large rotation around the view axis. Simulations with a 6-DOF eye-in-hand manipulator show that the proposed uncalibrated approach rapidly converges in all cases.**

I. Introduction

Uncalibrated visual servoing studies vision-based motion control of robots without known intrinsic and extrinsic camera/robot calibration parameters, or object models [1], [2]. This is a demanding problem with increasing applications in unstructured environments, where no prior information can be assumed. Most existing uncalibrated image-based visual servoing methods consider the coordinates of interest points as visual features [2]–[5] with a proportional control law [6]. In this paper, the three-view projective geometry of the initial, current, and desired views is exploited for uncalibrated visual servoing of a 6-DOF manipulator. This geometry is encapsulated by the trifocal tensor, which is independent of the scene and depends only on the projective relations between the cameras [7].

The control law is important in the performance of a visual servo. Although pose-based [8] and hybrid [9] methods can provide global stability, they are not generally suitable for uncalibrated settings because they require a 3D quantity such as depth or pose. Conventional image-based methods provide local asymptotic stability [10], therefore are only safe to use when the desired state is close to the initial state. There are known degenerate cases, such as 180°-rotation around the z-axis, where the image-based control law makes the camera retreat from the features. There does not seem to be a general way to determine the basin of locality given an image-based control law. Recent work has studied and analyzed different image-based control laws [11], which highlights the inherent problems in purely image-based control laws (local asymptotic stability, local minima). The choice of visual features has an equally important role in the performance and is one of the open problems in visual servoing. To increase the stability region, image-moments have been proposed for planar [12] and non-planar objects [13]. More recently, Hadj-Abdelkader *et al.* have proposed features from a spherical projection model for a decoupled hybrid visual servoing [14]. These methods are either purely image-based or hybrid.

Two-view geometry has been extensively studied in visual servoing. Epipolar geometry can be used to estimate depth, which appears in the interaction matrix [15] for an improved stability [9]. Chesi *et al.* exploit the symmetry of epipolar geometry without point correspondences to control a holonomic mobile robot from a partially calibrated camera [16]. Mariottini *et al.* use epipolar geometry for visual servoing of non-holonomic mobile robots [17]. Becerra and Sagues develop a sliding mode control law using epipolar geometry for non-holonomic mobile robots [18]. Homography-based methods have also been studied in visual servoing. Benhimane and Malis develop a homography-based approach without reconstructing any 3D parameters [19]. Lopez-Nicolas *et al.* design a homography-based controller which considers the non-holonomic constraints [20]. These methods use the two-view geometry between the observed and desired views and ignore their relation with the *initial view*. In addition, epipolar geometry is not well-conditioned if the features are coplanar or the baseline is short. On the other hand, homography-based approaches require dominant planes [21].

Projective geometry of three views can also be used in vision-based motion control. The trifocal tensor relates three views in a similar manner that the fundamental matrix relates two views. The application of the trifocal tensor to visual servoing has been neglected in the visual servoing literature until very recently [21], [22]. Becerra and Sagues use a simplified trifocal tensor as measurement and estimate and track the pose of a non-holonomic mobile robot with Extended Kalman Filter (EKF) [21]. Lopez-Nicolas *et al.* also use the constrained camera motion on a mobile robot and linearize the input-output space for control [22]. This approach provides an analytic interaction matrix between the features (9 elements of the trifocal tensor) and robot velocities. To the best of our knowledge, the trifocal tensor has not yet been used to control a 6-DOF robot. This is likely due to the difficulty in linearizing the input-output space in the case of generalized 6-DOF camera motions.

The main contribution of this paper is to propose a 6-DOF uncalibrated visual servoing approach that uses the trifocal features. The advantage of uncalibrated methods over analytical methods is clear when the interaction matrix of a set of features cannot be derived easily. The trifocal features are a subset of the 27 elements of the $3 \times 3 \times 3$ trifocal

This work is supported by NSERC, CFI, iCORE, and the Alberta Advanced Education & Technology.

Authors are with the Department of Computing Science, University of Alberta, Edmonton, AB, T6G2E8, Canada. {azad,jag}@cs.ualberta.ca

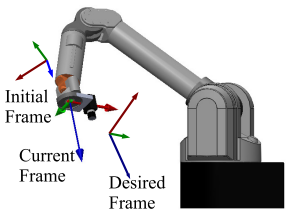

Fig. 1. The initial, current, and desired camera frames.

tensor. We estimate the interaction matrix, which relates joint velocities to the rate of change of trifocal features. This work does not fall into the conventional image-based, pose-based, or hybrid classifications of visual servoing based on the control law. There seems to be a missing visual servoing class, *projective visual servoing*, where the control loop is closed over projective measures (*e.g.,* controlling the epipoles [17], trifocal features [22], etc.). In essence, we control one such projective measure, which is found directly from images across three views, without explicitly recovering the camera pose or directly closing the loop in the image space.

This work is closely related to the recent works of Becerra and Sagues [18] and Lopez-Nicolas *et al.* [22] with distinct differences. First, we consider the 6-DOF motion of an eye-in-hand camera, but they consider the planar motion of the camera constrained to the non-holonomic motion model of a mobile robot. Second, they derive an analytic form of the Jacobian and use input-output linearization for control. In contrast, we estimate the Jacobian matrix directly from measurements of the trifocal features, because an analytic form is cumbersome to derive for the general motion. Further, we use a control law similar to the typical proportional control law in image-based visual servoing.

II. Background

A. The Geometry of Three Views

The initial, current, and desired camera frames for an eye-in-hand visual servoing system is shown in Fig. 1. A 3D point $\mathbf{X} \in \mathbb{R}^3$ in Euclidean space projects onto the image plane by a 3×4 projection matrix $\mathbf{P} = \mathbf{K}[\mathbf{R}| - \mathbf{R}\tilde{\mathbf{C}}]$, where $\mathbf{K}$ is the camera intrinsic matrix, $\mathbf{R}$ is the rotation of the camera frame w.r.t. the world frame, and $\tilde{\mathbf{C}}$ is the coordinate of the camera frame expressed in the world frame. The homogeneous coordinates of image point $\mathbf{x}$ can be found from $\mathbf{x} = \mathbf{PX}$. Let the initial, current, and desired projection matrices be $\mathbf{P}_i$, $\mathbf{P}_c$, and $\mathbf{P}_d$. A 3D point $\mathbf{X}$ projects to $\mathbf{x}_i = \mathbf{P}_i\mathbf{X}$, $\mathbf{x}_c = \mathbf{P}_c\mathbf{X}$, and $\mathbf{x}_d = \mathbf{P}_d\mathbf{X}$ across the three views. The three-view point correspondence is denoted by $\mathbf{x}_i \leftrightarrow \mathbf{x}_c \leftrightarrow \mathbf{x}_d$. In our case, the initial camera matrix $\mathbf{P}_i$ and the desired camera matrix $\mathbf{P}_d$ are constant. It is only the current camera matrix $\mathbf{P}_c$ and the image points in the current view that change as the robot configuration changes.

B. The Trifocal Tensor

The trifocal tensor encapsulates the geometry of three views in a similar manner that the fundamental matrix encapsulates the geometry of two views. They are independent of the scene and depend only on the camera projection matrices. Given one view and the fundamental matrix, image points can be transferred to the second view. Similarly, given two views and point (or line) correspondences across two-views, image points (or lines) can be transferred to a third view. The trifocal tensor is more general than combining the existing epipolar geometries between views (1,2) (1,3) and (2,3). For example, epipolar transfer fails for points lying on the trifocal plane[1] [7]. The trifocal tensor and its computation is described in detail by Hartley and Zisserman [7]. A brief description is presented here for the sake of completeness.

Consider the canonical representation $\mathbf{P}_i = [\mathbf{I}|\mathbf{0}]$, and $\mathbf{P}_c = [\mathbf{A}|\mathbf{a}_4]$, and $\mathbf{P}_d = [\mathbf{B}|\mathbf{b}_4]$, where $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices and vectors $\mathbf{a}_k$ and $\mathbf{b}_k$ are the k-th columns of $\mathbf{P}_c$ and $\mathbf{P}_d$, respectively. The set of three matrices $\{\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3\}$ constitute the trifocal tensor in matrix notation [7]:

$$\mathbf{T}_i = \mathbf{a}_k\mathbf{b}_4^\top - \mathbf{a}_4\mathbf{b}_k^\top .$$

The trifocal tensor is, in fact, a $3\times3\times3$ cube of cells with 27 elements. The equivalent camera projections are specified, up to a projective transformation, by only 18 parameters. This enforces 8 additional internal algebraic constraints on the elements of the trifocal tensor [7].

C. Estimation of the Trifocal Tensor

The trifocal tensor can be estimated from point or line correspondences across three views. Each point correspondence provides 4 independent equations. One may use the normalized linear solution, however, this does not consider the internal constraints and may lead to a geometrically invalid tensor. A more elaborate solution uses the linear solution as the initial condition of geometric minimization. The geometric minimization algorithm provides a geometrically valid tensor. However, these methods are vulnerable to outliers as they consider all correspondences as inliers. Robust estimation can be used to handle outliers. In particular, RANSAC [23] can be used on the random samples of 6 point correspondences across the three views, from which the camera projections are found up to a projectivity [24]. The trifocal tensor is then retrieved from these camera projections. This ensures a geometrically valid solution, while avoiding outliers. For the purpose of this paper, we have avoided outliers in controlled laboratory settings and estimate the trifocal tensor from the 6 point algorithm [24]. This allows us to focus on the development of the uncalibrated control law and Jacobian estimation (see Section III). Generalizations to the robust trifocal estimation would be straightforward.

[1]The trifocal plane is the plane passing through the three camera centers.

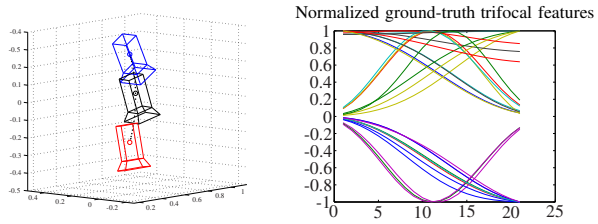

Fig. 2. (a) A sample camera trajectory from the initial (top, blue) to the desired (bottom, red) pose. An arbitrary camera on the trajectory is also shown (middle, black). (b) The evolution of the elements of the trifocal tensor along this trajectory. The elements are normalized with respect to their maximum absolute value to fit in the same graph.

III. Visual Servoing with Trifocal Tensor

A. Problem Formulation

The visual servoing problem is usually set up having known initial and desired features and measuring current features. The goal of visual servoing is to drive a robot to a desired configuration by regulating a task function to zero. The task function is defined on some relevant features. In our case, these features are defined according to the projective geometry of three views. Let $\mathbf{F} : \mathbb{R}^6 \rightarrow \mathbb{R}^M$ denote the sensory-motor function from configuration $\mathbf{q} \in \mathbb{R}^6$ of a robot with 6 joints (Fig. 1), to the feature vector $\mathbf{s} \in \mathbb{R}^M$ with M features. For visual servoing, measurements with $M \geq 6$ features is needed.

The trifocal tensor has 27 elements, which can be found from point correspondences across the three views. A subset of the elements of the trifocal tensor goes into vector $\mathbf{s}$. Fig. 2 shows a sample camera trajectory with the evolution of the trifocal features from the initial to desired pose.

The time derivative of the sensory-motor function leads to the $M \times 6$ sensory-motor Jacobian[2], $\mathbf{J}(\mathbf{q})$:

$$\frac{\partial \mathbf{s}}{\partial t} = \frac{\partial \mathbf{F}(\mathbf{q})}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial t}, \tag{1}$$

$$\dot{\mathbf{s}} = \mathbf{J}(\mathbf{q})\dot{\mathbf{q}}. \tag{2}$$

With an estimate $\widehat{\mathbf{J}}(\mathbf{q})$ for $\mathbf{J}(\mathbf{q})$, the discrete-time form of (2) becomes

$$\Delta \mathbf{s} \simeq \widehat{\mathbf{J}}(\mathbf{q})\Delta \mathbf{q}. \tag{3}$$

The sensory-motor Jacobian is an integral part of the visual servoing control law. A weighted Jacobian may be written as follows [11]:

$$\widehat{\mathbf{J}}_\beta = \beta \widehat{\mathbf{J}}_\mathbf{d} + (1-\beta)\widehat{\mathbf{J}}(\mathbf{q}), \tag{4}$$

where $\widehat{\mathbf{J}}_\mathbf{d} = \widehat{\mathbf{J}}(\mathbf{q_d})$ is the estimated Jacobian at the desired state. The control law in uncalibrated visual servoing is defined entirely in the feature space. To reach a visual goal $\mathbf{s_d}$, the general Jacobian $\widehat{\mathbf{J}}_\beta$ can be used in the following control law:

$$\dot{\mathbf{q}} = -\lambda \widehat{\mathbf{J}}_\beta^\dagger (\mathbf{s} - \mathbf{s_d}), \tag{5}$$

where $\lambda < 1$ is a positive constant to make joint velocity small, and $\widehat{\mathbf{J}}_\beta^\dagger$ is the Moore-Penrose pseudoinverse of $\widehat{\mathbf{J}}_\beta$.

[2]This Jacobian is also called the interaction matrix. We use Jacobian and interaction matrix interchangeably in this paper.

B. Online Jacobian Estimation

It is natural to assume that the robot keeps record of the sensory-motor information while it operates in the environment. This information may be stored in the memory for uncalibrated Jacobian estimation. Let the robot memory include P sensory-motor data pairs The uncalibrated sensory-motor Jacobian at a new sensory-motor query point $d_c = (\mathbf{s}_c, \mathbf{q}_c)$ can be estimated by solving the following optimization problem [5]:

$$\widehat{\mathbf{J}}(\mathbf{q})\Big|_{\mathbf{q}=\mathbf{q}_c} = \arg\min_{\mathbf{J}} \sum_{k:\, \mathbf{q}_k \in N(\mathbf{q}_c)} \rho(\Delta \mathbf{s}_k - \mathbf{J}\Delta \mathbf{q}_k), \tag{6}$$

where $N(\mathbf{q}_c) = \{\mathbf{q}_p : \|\mathbf{q}_c - \mathbf{q}_p\| < r\,, p = 1, \cdots, P\}$ is a neighborhood of $\mathbf{q}_c$ in the joint space, $\Delta \mathbf{s}_k = \mathbf{s}_c - \mathbf{s}_k$, $\Delta \mathbf{q}_k = \mathbf{q}_c - \mathbf{q}_k$, and $\rho(\cdot)$ is a cost function such as the L_2-norm [4] or robust M-estimator [5]. This method fits the best hyperplane to the sensory-motor data around $\mathbf{q}_c$ and has been validated for coordinates of interest points as visual features in both eye-to-hand [4] and eye-in-hand [5] configurations.

C. Reference Jacobian

In order to evaluate the online Jacobian estimate in (6), we need a reference Jacobian to compare against. Since the analytic form of the Jacobian is not available, we will estimate the value of the reference Jacobian by small orthogonal exploratory motions [25]. Orthogonal exploratory motions can be used to numerically estimate a Jacobian with reasonable accuracy. Consider a small observable displacement δ of a joint. The orthogonal motions can be measured from 6 observed feature displacements $\Delta \mathbf{s}^{(1)}, \cdots, \Delta \mathbf{s}^{(6)}$:

$$\begin{aligned} \Delta \mathbf{s}^{(1)} &\simeq \widehat{\mathbf{J}} \begin{bmatrix} \delta & 0 & \dots & 0 \end{bmatrix}^\top \\ &\vdots \quad\quad \vdots \\ \Delta \mathbf{s}^{(6)} &\simeq \widehat{\mathbf{J}} \begin{bmatrix} 0 & \dots & 0 & \delta \end{bmatrix}^\top, \end{aligned} \tag{7}$$

The Jacobian can then be directly recovered:

$$\mathbf{J_R} \simeq \frac{1}{\delta} \begin{bmatrix} \Delta \mathbf{s}^{(1)} & \Delta \mathbf{s}^{(2)} & \dots & \Delta \mathbf{s}^{(6)} \end{bmatrix} \tag{8}$$

In simulations, we can choose a very small value for δ and precisely measure $\Delta \mathbf{s}^{(i)}, i = 1 \cdots 6$. The main difference between the reference Jacobian in (8) and the online estimates in (6) is that there is no guarantee that the $\Delta \mathbf{s}_k$'s in (6) are measured after orthogonal motions.

The orthogonal motions is not practical for online estimation of the Jacobian in real robotic systems, because of extra undesired motions. However, it is a meaningful reference to evaluate the accuracy of the Jacobian estimation and we use it to evaluate the accuracy of the online Jacobian estimates. Jacobian estimation error is measured by the Frobenius norm of the online estimate, $\widehat{\mathbf{J}}$, to the reference Jacobian, $\mathbf{J_R}$:

$$\nu = \|\mathbf{J_R} - \widehat{\mathbf{J}}\|_F = \left|\left| \mathrm{diag}\left((\mathbf{J_R} - \widehat{\mathbf{J}})^\top (\mathbf{J_R} - \widehat{\mathbf{J}}) \right) \right|\right|_2, \tag{9}$$

where $||\cdot||_F$ is the Frobenius norm and $||\cdot||_2$ is the L_2-norm.

IV. Experiments

In this section, we experimentally evaluate the performance of the new features with the proposed control law in (5). Note that this control law uses a new Jacobian and is specified in the space of the trifocal features. To evaluate this new control law, we follow the recommendations in Good Experimental Methodology in robotics prepared by EURON [26]. Specifically, we limit this work to evaluations through controlled simulations. The results for small (local) motions include pure translation, translation along z-axis, and an arbitrary motion including rotations. We also present experimental results for the degenerate case of rotational motion around the z-axis, which is one of the "hard" configurations in visual servoing [6], [10], [11].

A. System Description

We consider the uncalibrated eye-in-hand visual servoing set up with a 6-DOF PUMA 560. The camera is mounted on the end-effector of the manipulator. We emphasize that our proposed method is general to arbitrary camera/robot configurations and the camera does not have to be on the end-effector. No assumptions on camera calibration or camera/robot calibration have been made. Simulations are implemented in MATLAB using the Robotics Toolbox [27] and the Epipolar Geometry Toolbox [28]. We use 6 points in a general configuration to evaluate the performance. Generalization to more point correspondences is straightforward by adopting the RANSAC robust estimation of the trifocal tensor [7]. For a valid trifocal tensor estimation, these 6 points should not be collinear in any of the views. This is somewhat limiting in translational motion experiments. We have chosen the initial and desired states to avoid collinearity.

B. Visual Servoing - Translation Along x-axis, y-axis, and z-axis

The first experiment considers a small translation of $[5.4, 5.4, -5.4]cm$ along the three axes. The control parameter is chosen $\beta = 0.5$ for this experiment. Fig. 3 (top-left) shows the position of the initial and desired camera. The robot is not shown in this figure. Other values of β result in a similar trajectory in this case, because the initial and desired states are very close. Fig. 3 (top-right) shows the evolution of the projections in the image. Note that at the desired image some of the features are collinear. Fig. 3 (bottom-left) shows the evolution of the trifocal features during servoing. Instead of showing all of the features, we normalize the features and show their mean-square-error (MSE). Apart from a discrepancy at the start, the features rapidly converge to zero. The source of this discrepancy is most likely due to the conditioning of the trifocal tensor estimation. Fig. 3 (bottom-right) shows the evolution of joint values during servoing. Joints 1, 4, 5, and 6 have large motions and cancel out each other to result in a linear end-effector motion.

C. Visual Servoing - Translation Along z-axis

Visual servoing along the z-axis is usually more challenging than along the xy-plane. This is due to poor motion

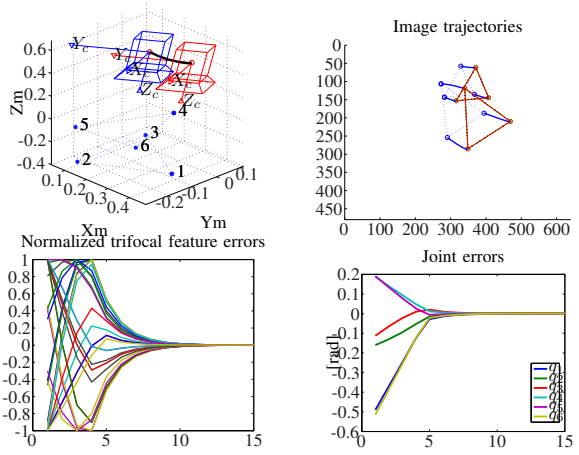

Fig. 3. Translation Along x-axis, y-axis, and z-axis. (Top-Left) Initial/Desired cameras and the camera trajectory. (Top-Right) Initial/Desired image coordinates and image trajectory. (Bottom-Left) The normalized trifocal feature errors. (Bottom-Right) Joint values in [rad].

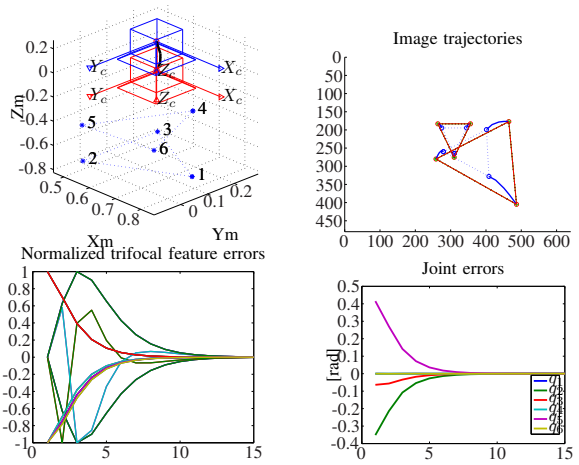

Fig. 4. Translation Along z-axis. (Top-Left) Initial/Desired cameras and the camera trajectory. (Top-Right) Initial/Desired image coordinates and image trajectory. (Bottom-Left) The normalized trifocal feature errors. (Bottom-Right) Joint values in [rad].

resolvability when the camera moves towards the points [29]. Fig. 4 (top-left) shows the camera trajectory for a $17.9cm$ translation along the z-axis. The initial and desired cameras are parallel to the plane passing through three of the six landmarks. The camera trajectory is almost linear with control parameter β is 0.5 (other values have similar performance). Fig. 4 (top-right) shows the image trajectories. Note that 4 non-coplanar points have very similar initial and desired values, but the trifocal features capture the geometry adequately. Fig. 4 (bottom-left) shows a rapid convergence of the trifocal features, and Fig. 4 (bottom-right) shows the convergence of the joint values.

D. Visual Servoing - Arbitrary General Motion

The results for an arbitrary small general motion are presented in Fig. 5. The desired camera frame is rotated by $[-35.2°, -14.5°, -12.8°]$ and the translated by

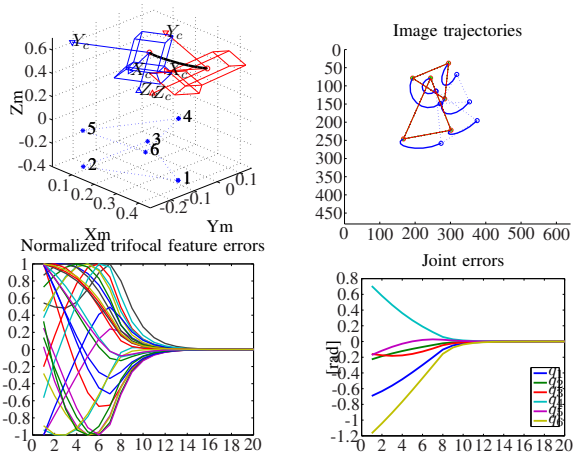

Fig. 5. Small general motion. (Top-Left) Initial/Desired cameras and the camera trajectory. (Top-Right) Initial/Desired image coordinates and image trajectory. (Bottom-Left) The normalized trifocal feature error. (Bottom-Right) Joint values in [rad].

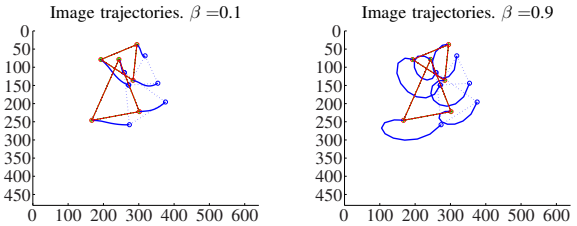

Fig. 6. Comparison of the control parameter β in (4) for the experiment of Section IV-D. For $\beta = 1$ the control law uses the constant desired Jacobian $\widehat{\mathbf{J}}_{\mathbf{d}}$ and for $\beta = 0$ the control law chooses the current estimate at each iteration. (Left) $\beta = 0.1$, and (Right) $\beta = 0.9$.

$[8.6, 8.6, -8.6]cm$. In this experiment, the control parameter is chosen to be $\beta = 0.5$. Other values of β also work. With a larger β the robot moves a longer trajectory and the image trajectory is more circular. Fig. 6 shows the image trajectory for $\beta = 0.1$ and $\beta = 0.9$.

E. Visual Servoing - 85°-Rotation Around and Translation Along z-axis

This experiment includes a $85°$-rotation around the z-axis and $25cm$ motion along the z-axis. During this experiment, we noticed that some of the elements of the trifocal tensor are constantly 0. This is because of the special type of camera motion in this experiment. Specifically,

$$\mathcal{T}_1^{23} = \mathcal{T}_1^{33} = \mathcal{T}_2^{13} = \mathcal{T}_2^{33} = 0.$$

In this case, we use the remaining 23 elements of the tensor in the trifocal feature vector $\mathbf{s}$.

Fig. 7 summarizes this experiment. It can be seen that camera translation is not entirely linear in the Euclidean space, however, the image trajectories are rotational, which is more desired than a linear image trajectory. Since we have made no attempt to decouple the translation from rotation, the non-linear camera trajectory is expected. Nonetheless, the rapid convergence suggests that the trifocal tensor is a suitable feature for uncalibrated visual servoing.

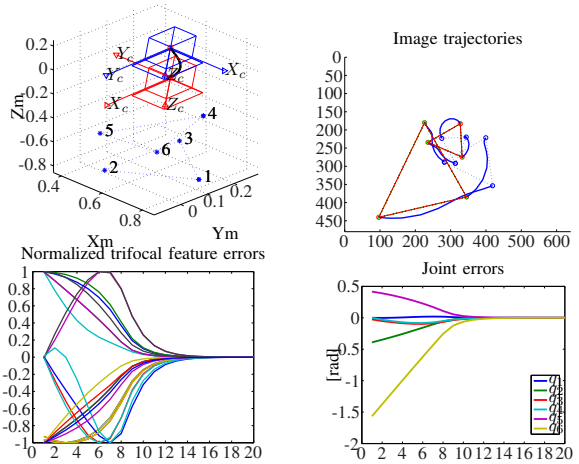

Fig. 7. $85°$-rotation around the z-axis and $25cm$ motion along the z-axis. (Top-Left) Initial/Desired cameras and the camera trajectory. (Top-Right) Initial/Desired image coordinates and image trajectory. (Bottom-Left) The normalized trifocal features errors. (Bottom-Right) Joint values in [rad].

F. Visual Servoing - Large Rotation Around and Translation along z-axis

One of the most challenging image-based visual servoing configurations is the $180°$ rotation around the z-axis [6], [10]. This is due to the nature of the image-based control law which makes the camera to retreat from the object instead of rotation around the view axis. It is important to evaluate a visual servo for large z-axis rotations, close to $180°$, for example a $170°$ rotation [6], [10], [11].

We consider a translation of $50cm$ and a $170°$-rotation to provide a common ground to compare this method against other approaches [11]. We use the same trifocal feature vector of the previous experiment. The proposed trifocal features successfully handle this case as illustrated in Fig. 8. Note that the desired and initial camera frames are rotated at $170°$ in Fig. 8 (top-left). The image trajectories show a spiral motion in Fig. 8 (top-right), which is the desired case. The results in Fig. 8 are obtained with control parameter $\beta = 0.05$. For $\beta = 0$, which corresponds to using the current Jacobian estimate in (5), control also converges, but with a slight abrupt motion at the start of the control loop. Choosing $\beta = 1$, which corresponds to using the constant value of the desired Jacobian in (5), was not successful. This result is expected as such a large motion is not local and the values of the Jacobian matrices, at the initial and desired states, are significantly different.

V. Conclusions

We presented a new uncalibrated visual servoing approach based on the trifocal tensor. The trifocal tensor encapsulates the geometry of the initial, desired, and current views and can be estimated from point correspondences up to a projectivity. We use the elements of the trifocal tensor as features and construct a task function from the desired and current estimation of the trifocal tensor. Such control laws may be considered as *projective visual servoing*, where the control

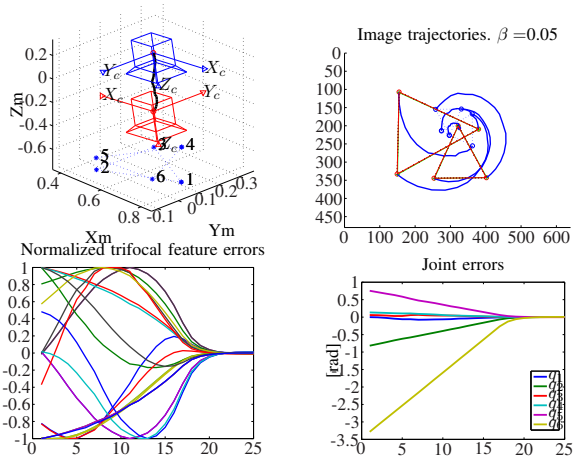

Fig. 8. 170°-rotation around and $50cm$ translation along the z-axis. (Top-Left) Initial/Desired cameras and the camera trajectory. (Top-Right) Initial/Desired image coordinates and image trajectory. (Bottom-Left) The mean-square-error of the normalized trifocal features. (Bottom-Right) Joint values in [rad].

law uses projective measures instead of image-based, pose-based, or hybrid. We also presented an online estimation of the sensory-motor Jacobian. Unlike conventional image-based methods, where the Jacobian relates the joint velocities to image measurements (coordinates, moments, etc.), the proposed Jacobian directly relates the joint velocities to the rate of change in the elements of the trifocal tensor. Experimental results show that the proposed method is very promising and handles well "hard" configurations (such as large rotation around the view axis).

The degenerate cases of image-based visual servoing [10] do not necessarily translate to degenerate cases here as shown by the large-rotation experiment (Section IV-F). Therefore, this approach is different than the integration of online calibration of the camera intrinsic parameters and using the image Jacobian. In our approach, the feedback loop is closed on the elements of the trifocal tensor, which are found from further computation of the visually-tracked interest points. The degenerate cases for the estimation of the trifocal tensor could result in the failure of the visual servo and will be studied in the future.

REFERENCES

[1] F. Chaumette and S. Hutchinson, "Visual servo control. part II: Advanced approaches [tutorial]," *IEEE Robot. Automat. Mag.*, vol. 14, no. 1, pp. 109–118, Mar. 2007.

[2] M. Jägersand, O. Fuentes, and R. Nelson, "Experimental evaluation of uncalibrated visual servoing for precision manipulation," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 4, April 1997, pp. 2874–2880.

[3] J. A. Piepmeier, G. V. McMurray, and H. Lipkin, "Uncalibrated dynamic visual servoing," *IEEE Trans. Robot. Automat.*, vol. 20, no. 1, pp. 143–147, Feb. 2004.

[4] A. M. Farahmand, A. Shademan, and M. Jägersand, "Global visual-motor estimation for uncalibrated visual servoing," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2007, pp. 1969–1974.

[5] A. Shademan, A.-M. Farahmand, and M. Jägersand, "Robust jacobian estimation for uncalibrated visual servoing," in *Proc. IEEE Int. Conf. Robot. Automat.*, Anchorage, AK, May 2010, pp. 5564–5569.

[6] F. Chaumette and S. Hutchinson, "Visual servo control. part I: Basic approaches," *IEEE Robot. Automat. Mag.*, vol. 13, no. 4, pp. 82–90, Dec. 2006.

[7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, UK: Cambridge University Press, 2003.

[8] W. J. Wilson, C. C. W. Hulls, and G. S. Bell, "Relative end-effector control using cartesian position based visual servoing," *IEEE Trans. Robot. Automat.*, vol. 12, no. 5, pp. 684–696, October 1996.

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

[10] 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. Morse, Eds. LNCS Series, No 237, Springer-Verlag, 1998, pp. 66–78.

[11] M. Marey and F. Chaumette, "Analysis of classical and new visual servoing control laws," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 19–23, 2008, pp. 3244–3249.

[12] F. Chaumette, "Image moments: a general and useful set of features for visual servoing," *IEEE Trans. Robot.*, vol. 20, no. 4, pp. 713–723, Aug. 2004.

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

[14] H. Hadj-Abdelkader, Y. Mezouar, and P. Martinet, "Decoupled visual servoing based on the spherical projection of a set of points," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 12–17, 2009, pp. 1110–1115.

[15] E. Malis and F. Chaumette, "2 1/2 d visual servoing with respect to unknown objects through a new estimation scheme of camera displacement," *Int. J. Comput. Vision*, vol. 37, no. 1, pp. 79–97, 2000.

[16] G. Chesi, D. Prattichizzo, and A. Vicino, "A visual servoing algorithm based on epipolar geometry," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2001, pp. 737–742.

[17] G. Mariottini, G. Oriolo, and D. Prattichizzo, "Image-based visual servoing for nonholonomic mobile robots using epipolar geometry," *IEEE Trans. Robot.*, vol. 23, no. 1, pp. 87–100, Feb. 2007.

[18] H. M. Becerra and C. Sagues, "A sliding mode control law for epipolar visual servoing of differential-drive robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 3058–3063.

[19] S. Benhimane and E. Malis, "Homography-based 2d visual servoing," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 15–19, 2006, pp. 2397–2402.

[20] G. López-Nicolás, C. Sagüés, J. J. Guerrero, D. Kragic, and P. Jensfelt, "Nonholonomic epipolar visual servoing," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2006, pp. 2378–2384.

[21] H. M. Becerra and C. Sagues, "Pose-estimation-based visual servoing for differential-drive robots using the 1d trifocal tensor," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 5942–5947.

[22] G. López-Nicolás, J. J. Guerrero, and C. Sagüés, "Visual control through the trifocal tensor for nonholonomic robots," *Robot. Auton. Syst.*, vol. 58, pp. 216–226, 2010.

[23] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[24] R. Hartley and N. Dano, "Reconstruction from six-point sequences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2000, pp. 480–486.

[25] H. Sutanto, R. Sharma, and V. Varma, "The role of exploratory movement in visual servoing without calibration," *Robotics and Autonomous Systems*, vol. 23, pp. 153–169, 1998.

[26] F. Bonsignorio, J. Hallam, and A. P. del Pobil. (March, 2010) GEM guidlines. EURON Special Interest Group on Good Experimental Methodology. Retrieved. [Online]. Available: http://www.heronrobots.com/EuronGEMSig/Downloads/GemSigGuidelinesBeta.pdf

[27] P. Corke, "A robotics toolbox for MATLAB," *IEEE Robot. Automat. Mag.*, vol. 3, no. 1, pp. 24–32, Mar. 1996.

[28] G. Mariottini and D. Prattichizzo, "EGT: a toolbox for multiple view geometry and visual servoing," *IEEE Robot. Automat. Mag.*, vol. 3, no. 12, December 2005.

[29] B. J. Nelson and P. K. Khosla, "Vision resolvability for visually servoed manipulation," *Journal of Robotic Systems*, vol. 13, no. 2, pp. 75–93, 1996.