IMAGE-BASED VISUAL SERVOING OF GOUGH-STEWARD PARALLEL MANIPULATORS USING LEGS OBSERVATION

Nicolas Andreff* Tej Dallej* Philippe Martinet*

* LASMEA-UMR 6602, CNRS-Université Blaise Pascal,
24 avenue des Landais, F-63177 Aubière Cedex,
firstname.lastname@lasmea.univ-bpclermont.fr,
http://www.lasmea.univ-bpclermont.fr/Control

Abstract: In this paper, a tight coupling between computer vision and parallel robotics is exhibited through the projective line geometry. Indeed, contrary to the usual methodology where the robot is modeled independently from the control law which will be implemented, we take into account, since the early modeling stage, that vision will be used for control. Hence, kinematic modeling and projective geometry are fused into a control-devoted projective kinematic model. Thus, a novel vision-based kinematic modeling of a Gough-Stewart manipulator is proposed through the image projection of its cylindrical legs. Using this model, a visual servoing scheme is presented, where the image projection of the non-rigidly linked legs are servoed, rather than the end-effector pose. Copyright ©2006 IFAC

Keywords: Vision-based control, Parallel mechanisms, Line geometry

1. INTRODUCTION

There exists a large amount of work on the control of parallel mechanisms1. In the focus of attention, Cartesian control is naturally achieved through the use of the inverse differential kinematic model which transforms Cartesian velocities into admissible joint velocities, taking into account the kinematic closures and avoiding to generate high internal forces (Dasgupta and Mruthyunjaya, 1998). It is noticeable that the inverse differential kinematic model of parallel mechanisms does not only depend on the joint configuration (as for serial mechanisms) but also on the end-effector pose.

Consequently, one needs to be able to estimate or measure the latter. As far as we know, all the effort has been put on the estimation of the end-effector pose through the forward kinematic model and the joint measurements. However, this yields much trouble, related to the fact that there is usually no closed-form solution to the forward kinematic problem for parallel mechanisms. Hence, one numerically inverts the inverse kinematic model, which is analytically defined for most of the parallel mechanisms. However, it is known (Merlet, 1990; Husty, 1996) that this nu-
merical inversion requires high order polynomial root determination, with several possible solutions (up to 40 real solutions (Dietmaier, 1998) for a Gough-Stewart platform (Gough and Whitehall, 1962; Stewart, 1965)). Much of the work is thus devoted to solving this problem accurately and in real-time (see for instance (Zhao and Peng, 2000)), or to designing parallel mechanisms with analytical forward kinematic model (Kim and Tsai, 2002; Gogu, 2004). Alternately, one of the promising paths lies in the use of the so-called metrological redundancy (Baron and Angelles, 1998), which simplifies the kinematic models by introducing additional sensors into the mechanism and thus yields easier control (Marquet, 2002).

Vision being an efficient way of estimating the end-effector pose, it is a good alternative to use it for Cartesian control of parallel mechanisms. It can be done in two appropriate manners. The first one consists in measuring the end-effector pose by vision, and then performing 3D pose visual servoing (Martinet et al., 1996) as in (Koreichi et al., 1998; Kino et al., 1999; Kallio et al., 2000) (for parallel robots with a reduced number of degrees of freedom). However, this approach consists solely in a simple adaptation of new classical control schemes, which, although efficient, are not very innovative. Moreover, the direct application of visual servoing techniques assumes implicitly that the robot inverse differential kinematic model is given and that it is calibrated. Therefore, modeling, identification and control have small interaction with each other. Indeed, the model is usually defined for control using proprioceptive sensing only and does not foresee the use of vision for control, then identification and control are defined later on with the constraints imposed by the model. This is useful for modularity but this might not be efficient in terms of accuracy as well as of experimental set-up time.

On the opposite, a second approach was introduced in (Andreff et al., 2005), where modeling, identification and control are performed knowing in advance that vision will be used at control time. To do so, the legs of the mechanism are servoed rather than its end-effector pose, which is useful in some practical cases. For instance, it is not wise to imagine observing the end-effector of a machining tool. On the opposite, it should not be a problem to observe the legs of the mechanism, even in such extreme cases.

In that method, the legs orientation were chosen as visual primitives and control was derived based on their reconstruction from the image. Although this reconstruction step consists only in computing the intersection of the two cylinder edges in the image, it might not be very accurate for intrinsically geometrical reasons. Indeed, if a leg is parallel to the image, its edges will appear as parallel lines in the image and their intersection will lie at the infinite. Thus, in a close case to this one, the reconstruction will not be highly accurate and will impair the control.

In practice, this case is rapidly encountered. Indeed, since tracking lines in the image might be hard in the presence of camera distortion, one would chose a rather long focal lens (6 mm is here a long focal). Then, to observe all the legs in the image, one would place the camera at some distance from the manipulator (say, 1 m away from a desktop Gough-Stewart manipulator). In such an easy to set-up case, the cylinder edges appear nearly parallel in the image and control becomes unstable.

Consequently, following the original idea of visually servoing the legs and the common habit in visual servoing to derive control laws as close as possible to the image space, we propose in this paper to servo the leg edges rather than the leg orientation.

The outline of the paper is as follows. Section 2 recalls the model used for lines, then uses it to express the Gough-Stewart platform kinematics in the (static) camera frame and finally, recalls some useful geometric properties associated to the fact that most parallel mechanisms have cylindrical legs. Section 3 introduces, under the cylindrical legs assumption, the novel control law, expressed in the image and using the apparent edges of the legs as visual primitives for control. Section 4 presents the first ever experimental results of visual servoing using legs observation in its two variants: the already published leg orientation-based control and the novel edge-based control.

2. MODELING

2.1 Line modeling

A line \( L \) in space, expressed in the camera frame, is defined by its binormalized Plücker coordinates (Andreff et al., 2002):

\[
L \equiv (\hat{u}, \hat{n}, \hat{n})
\]  

(1)

where \( \hat{u} \) is the unit vector giving the spatial orientation of the line, \( \hat{n} \) is the unit vector defining the so-called interpretation plane of line \( L \) and \( \hat{n} \) is a non-negative scalar. The latter are defined by \( \hat{n} \hat{n} = \hat{P} \times \hat{u} \) where \( \hat{P} \) is any point on the line. Notice that, using this notation, the well-known Plücker coordinates (Plücker, 1865) are the couple \( (\hat{u}, \hat{n}) \).

The projection of such a line in the image plane, expressed in the camera frame, has for characteristic equation:
\[ c_n^T c p = 0 \]  
where \( c p \) are the coordinates in the camera frame of any point in the image plane, lying on the line.

With the intrinsic parameters \( K \), one can easily obtain the conversion from the line equation in the camera frame \( c n \) to the same in pixel coordinates \( p n \) and vice-versa:

\[ p n = \frac{K^{-T} c n}{\|K^{-T} c n\|} \]  
\[ c n = \frac{K^T p n}{\|K^T p n\|} \]

Notice that for numerical reasons, one should use normalized pixel coordinates. Namely, let us define the pixel frame by its origin located at the image center (i.e. the intersection of the image diagonals) and such that the pixel coordinates vary approximately between -1 and +1, by dividing them by, say, the image horizontal dimension in pixels.

### 2.2 Vision-based kinematics of an hexapod

Consider the hexapod in Figure 1. It has 6 cylindrical legs of varying length \( q_i, i \in 1..6 \), attached to the base by spherical joints located in points \( A_i \) and to the moving platform (end-effector) by spherical joints located in points \( B_i \).

Rather than using the standard scalar inverse kinematic model of such an hexapod given by

\[ \forall i \in 1..6, \quad q_i^2 = A_i B_i^T A_i B_i \]  
expressing that \( q_i \) is the length of vector \( A_i B_i \), it is preferable for the subsequent derivation to use the vector form, introduced as the vision-based kinematics of the hexapod expressed in the camera frame in (Andreff et al., 2005):

\[ q_i^c \hat{u}_i = \tau_i \times \hat{u}_i + \tau_i e - A_i \]  
where \( \hat{u}_i \) is the spatial orientation of the \( i \)th leg. From the inverse kinematic model, one easily obtains the differential inverse kinematic model:

\[ \dot{q}_i^c = J_i^c \dot{\tau}_c \]  
\[ J_i^c = \begin{bmatrix} \hat{u}_i \times (A_i x \hat{u}_i)^T \\ \vdots \\ \hat{u}_i \times (A_6 x \hat{u}_6)^T \end{bmatrix} \]

where \( \dot{\tau}_c \) is the Cartesian velocity of the camera frame, considered as attached to the base frame and moving with respect to a fixed end-effector, and \( \hat{u}_i, i = 1..6 \) are the unit vectors giving the pointing direction of each leg in the camera frame. Under the assumption that the legs are cylinders, those vectors can be easily detected as the intersection of the two cylinder edges in the image plane.

### 2.3 Cylindrical leg observation

Let us remark now that each cylinder edge is a line in space, with binormalized Plücker expressed in the camera frame \( (\hat{u}_i, \hat{v}_i, \hat{w}_i) \). Moreover, the attachment point \( A_i \) is lying on the cylinder axis at distance \( R \) from the edge. Consequently, a cylinder edge is entirely defined by the following constraints, expressed here in the camera frame:

\[ c A_i^T \hat{u}_i = -R \]  
\[ c \hat{v}_i^T p_i = 1 \]  
\[ c \hat{w}_i^T p_i = 0 \]

### 3. EDGE-BASED VISUAL SERVOING

In the following subsection, the \( i \) subscript denoting the leg number was removed for clarity sake.

#### 3.1 Interaction matrix

The interaction matrix \( J^T \) relating the Cartesian velocity \( \tau_c \) to the time derivative of the cylinder edges \( \dot{\hat{u}}_i \), expressed in the pixel frame:

\[ c \hat{u}_i^T \dot{\hat{u}}_i = N^T \tau_c \]  

can be decomposed into the product of three matrices:

\[ N^T = p J \dot{\hat{u}} J \alpha M^T \]  

From right to left, the first one is the interaction matrix associated to the leg orientation, it thus relates the time derivative of a leg orientation to \( \tau_c \). The second transforms leg orientation velocities into leg edge velocities both expressed in the camera frame. Finally, the third one is associated to the camera-to-pixel change of frame. Below, the expression of the leg orientation interaction matrix is first recalled then the last two matrices are derived.

#### 3.1.1. Leg orientation interaction matrix

The control proposed in (Andreff et al., 2005) servoed the geodesic error between the current and desired legs orientation \( \hat{u} \times \dot{\hat{u}} \) and thus introduced the interaction matrix associated to a leg orientation:

\[ \ddot{\hat{u}} = M \tau_c \]  
\[ M = -\frac{1}{q} (I_3 - \hat{u} \hat{u}^T) \begin{bmatrix} I_3 & -\hat{u}^T \\ \hat{u} & 0 \end{bmatrix} \]

#### 3.1.2. Edge velocity in the camera frame

Let us first derive the time derivative of a cylinder edge, expressed in the camera frame, and under the kinematic constraint that the cylinder is attached to the base by a spherical or universal joint located in \( A \). To do so, let us differentiate (9)-(11):
\[ c_n^i \hat{J}^t A = 0 \]  
\[ \hat{J}^T c_n \hat{n}^i = 0 \]  
\[ \hat{J}^T c_n \hat{u} + c_n^i \hat{J}^T c_n \hat{u} = 0 \]  

From (17) and the fact that \( (\mathbf{u}, \mathbf{n}, \mathbf{u} \times \mathbf{n}) \) form an orthonormal basis (Andreff et al., 2002), one can state:
\[ \hat{c}_n^j = \alpha \hat{c}_n^j \mathbf{u} + \beta \mathbf{u} \times \hat{c}_n^j \]  
Inserting this expression into (16) and (18) yields
\[ \alpha = -c_n^i \hat{J}^T \mathbf{u}, \quad \beta = \frac{c_n^i \hat{J}^T (\mathbf{u} \times \hat{c}_n^j)}{c_n^i \hat{J}^T (\mathbf{u} \times \hat{c}_n^j)} c_n^i \hat{J}^T \mathbf{u} \]

Consequently, one obtains the relationship between the time derivative of a leg edge, expressed in the camera frame, and the time derivative of the leg orientation
\[ \hat{c}_n^j = h \mathbf{J}_n \hat{c}_n^j \quad (= h \mathbf{J}_n M^T c_n \tau_c) \]  
\[ h \mathbf{J}_n = -\left( I_3 - \frac{(\mathbf{u} \times \hat{c}_n^j) \hat{J}^T}{c_n^i \hat{J}^T (\mathbf{u} \times \hat{c}_n^j)} \right) \mathbf{u} \hat{J}^T \]

### 3.1.3. Image line velocity in pixel coordinates

Let us now derive the Jacobian associated to the change of frame, where the time derivative of an image line is expressed, from the camera frame to the pixel frame. Note that this paragraph holds for any image line, not only for cylinder edges.

Rewriting (3) as
\[ \hat{p}_n = \mu(c_n^i) K^{-T} \hat{n} \]

we can differentiate the latter with time:
\[ \hat{p}_n = \frac{d\mu(c_n^i)}{dt} K^{-T} \hat{n} + \mu(c_n^i) K^{-T} \hat{c}_n \]

Taking into account again that \( \hat{p}_n \) is a unit vector (17), one gets
\[ \left( \frac{d\mu(c_n^i)}{dt} K^{-T} \hat{n} + \mu(c_n^i) \right) \hat{p}_n K^{-T} \hat{c}_n = 0 \]

Using (3) again, this simplifies into
\[ \frac{d\mu(c_n^i)}{dt} = -\mu(c_n^i) \hat{p}_n K^{-T} \hat{c}_n \hat{n} \]

Inserting this result in (24) yields
\[ \hat{p}_n = \left(-K^{-T} \hat{n} \mu(c_n^i) \hat{p}_n \right) + \mu(c_n^i) I_3 ) K^{-T} \hat{n} \]

which simplifies into
\[ \hat{p}_n = \mu(c_n^i) (I_3 - \hat{p}_n \hat{p}_n T) K^{-T} \hat{n} \]

Introducing (4) in (23) proves that
\[ \mu(c_n^i) = ||K^T p_n|| \]

from which we finally obtain the relationship between the time derivative of a line, expressed in the image frame, and the same expressed in the camera frame.

\[ \hat{p}_n = \frac{p}{J_c} c_n \]

where \( \frac{p}{J_c} h J_a M^T c_n \tau_c \)

\[ p_{J_c} = ||K^T p_n|| (I_3 - \hat{p}_n \hat{p}_n T) K^{-T} \]

### 3.2 Control

Since we want to drive the unit vectors associated to the leg edges to their desired values, we choose to servo the geodesic errors
\[ \mathbf{e}_{1,2} = \hat{p}_n \hat{p}_n^T, \quad j = 1, 2, i = 1, 6 \]

whose time derivatives are
\[ \dot{e}_{1,2} = L^T e_{1,2} \]

\[ L^T = -[p_n^i]_{ik} N_{i,j}^T \]

where \( i = 1.6 \) denotes the legs and \( j = 1.2 \) the edges.

Now, the standard visual servoing method applies: we stack each individual errors in a single over-constrained vector \( \mathbf{e} \) and each associated individual interaction matrices \( L^T \) into a compound one \( L^T \) and impose a first-order convergence to \( \mathbf{e} \). This yields the following pseudo-control vector \( \tau_c \)
\[ \tau_c = -\lambda L^T \mathbf{e} \]

which is fed to the actuators through the vision-based differential inverse kinematic model (8) to deliver the final control signal
\[ \mathbf{q} = -\lambda^2 J_{inv} \hat{L}_T^T \mathbf{e} \]

where the \( \hat{\mathbf{L}} \) means estimated.

Notice that this control makes use of the detected edges in the image, the joint values, the intrinsic parameters and the attachment points of the legs onto the base expressed in the camera frame. However, notice that neither the attachment points of the legs onto the mobile platform nor the radius of the legs are used here explicitly, which reduces the number of kinematic parameters to be calibrated.

### 4. RESULTS

The experimental robot has an analog joint position-controller that we interfaced with Linux-RTAI. Joint velocity control is emulated through this position-controller with an approximate 20ms sampling period. Frame grabbing, line tracking and numerical computation are performed using ViSP, an open C++ library for visual servoing (Marchand et al., 2005).

It also has to be noticed that the mechanism presents high frictional disturbances that have not yet been compensated for since friction seems to
Fig. 2. Composition of the desired (black) and initial (white) configurations.

Fig. 3. Behavior of orientation-based control: Evolution of the controlled leg orientation errors (left) and of the non-controlled edge errors (right) with respect to time.

Fig. 4. Behavior of edge-based control: Evolution of the non-controlled leg orientation errors (left) and of the controlled leg edge errors (right) with respect to time.

Fig. 5. Comparison of the behavior of orientation-based (solid line) and edge-based control (dotted line): evolution of the norm of the legs orientation error (left) and of the edges error (right) with respect to time.

5. CONCLUSION

We extended previous results concerning (PBVS-like) leg orientation-based control of a Gough-Stewart to an (IBVS-like) edge-based visual servoing scheme. It benefits from the advantages of the orientation-based visual servoing of the Gough-Stewart legs: reduced calibration parameters set, low dependence on the joint values and ability to servo the robot even though the end-effector is not visible. However, we improved the practical robustness (although it still has to be proven theoretically) by servoing the legs in the image: almost all the calibration parameters (intrinsic parameters of the camera and base points) and numerical errors remain located in the interaction matrix.
To do so, we took advantage of a common use of line geometry in kinematics, vision and visual servoing. This allows for an optimal modeling of Gough-Stewart parallel robots, provided that vision is used at control time. This modeling was established under the hypothesis that the camera is calibrated, but this result might be extendable to the use of an uncalibrated camera. Nevertheless, this extension is not necessary since the control is done in the very projective space associated to image lines, while the reconstructed or calibrated Euclidean terms only appear in the interaction matrix where extreme accuracy is not required.

However, self-occlusions of the mechanism with respect to a single camera are still a matter of study, although the observation of edges should simplify the problem since the two edges of a given leg are seldom hidden simultaneously. A way to overcome the occlusion problem is to turn oneself to multi-camera perception systems.

ACKNOWLEDGMENT

This study was jointly funded by CPER Auvergne 2003-2005 program and by the CNRS-ROBEA program through the MP2 project. The authors warmly thank Eric Marchand for providing them with an early version of ViSP.

REFERENCES


