Image Based Visual Servoing through Nonlinear Model Predictive Control

M. Sauvée  
ECA Sinters SA  
5 rue Paul Mesplé  
31106 Toulouse, France  
sauvee@lirmm.fr

P. Poignet, E. Dombre  
LIRMM, Robotic department  
161 rue Ada  
34392 Montpellier, France  
{poignet, dombre}@lirmm.fr  
Telephone: +33 4 67 41 85 61

E. Courtial  
LVR, Polytech’Orléans  
8 rue Léonard de Vinci  
45072 Orléans, France  
estelle.courtial@univ-orleans.fr

Abstract—Image based visual servoing (IBVS) is a vision sensor based control architecture. In classical approach, an image Jacobian matrix maps image space errors into errors in Cartesian space. Then a simple proportional control law can be applied guaranteeing local convergence to a desired set point. One of the main advantage of IBVS is its robustness w.r.t. camera and robot calibration errors and image measurement errors. Nevertheless, this scheme can not deal with nonlinear constraint such as joint limits and actuator saturation. Visual constraint is not ensured with classical IBVS. A new IBVS scheme based on Nonlinear Model Predictive Control (NMPC) is proposed considering the direct dynamic model of the robot, its joint and torque limits, the camera projection model and the visibility constraint. Simulations exhibit the efficiency and the robustness of the proposed solution to control a 6 degrees of freedom mechanical system.

I. INTRODUCTION

Visual servoing appears in the eighties through the use of image Jacobian [1]: the difference between reference and current image provides signal error that is mapped into the 3D Cartesian velocities of the end-effector through the interaction matrix. A proportional control law based on inverse image Jacobian allowed to perform complex task based on visual information (positioning, tracking, grasping). This architecture was called Image Based Visual Servoing (IBVS). In IBVS systems, the manipulator control input is computed through errors in image space. Considering an eye-to-hand configuration, the velocity screw \( \dot{r} = [v, \omega] \), composed of the linear velocities \( v = (T_x, T_y, T_z)^T \) and the angular velocities \( \omega = (\omega_x, \omega_y, \omega_z)^T \), of a point rigidly attached to the end-effector frame is related by the interaction matrix \( J_{im} \) to the motion of its image point whose coordinates in the image are given by \( p = (p_u, p_v)^T \):

\[
\dot{p} = J_{im}(p, f, z) \dot{r}
\]

where \( z \) is the depth of the considered point w.r.t. camera frame. \( J_{im} \) is defined as:

\[
J_{im} = \begin{bmatrix}
\frac{f^2}{z} & 0 & -\frac{p_u}{z} & -\frac{p_u p_v}{f} & \frac{f^2 + p_v^2}{f} & -p_v \\
0 & \frac{f^2}{z} & -\frac{p_v}{z} & -\frac{f^2 + p_v^2}{f} & \frac{p_u p_v}{f} & p_u
\end{bmatrix}
\]

in which \( f \) is the focal length of the camera in pixel. More details can be found in the tutorial proposed in [2].

It has been shown [3] that the minimum number of points to have the complete velocity screw of the object uniquely defined through the pseudoinverse \( J_{im}^+ \) is 4. So \( J_{im} \) is built by stacking matrices defined by (2) for each point of the rigid object.

Defining the robot Jacobian \( J_r \) that maps joint velocities \( \dot{q} \) to the velocity screw of the end-effector, (1) can be rewritten as:

\[
\dot{p} = J_{im} J_r \dot{q}
\]

Assume the mechanical arm as a perfect integrator, the simplest approach to IBVS is to design the control law using (3):

\[
u = K_p J_r^{-1} J_{im} \dot{p}
\]

where \( K_p \) is a diagonal gain matrix and \( u \) is the control input which is the joint velocity setpoint. Given a desired feature vector \( p^d \) in the image plane, thus IBVS controller vanishes error in image space and ensures local convergence.

IBVS main interest is its robustness w.r.t. camera and robot calibration errors as described by [4], and image measurement errors. However, IBVS may exhibit some drawbacks: visibility of the controlled features is not ensured, trajectory in operational and in joint space is not optimal, singularities of the image Jacobian exist for some motion. Error in image Jacobian estimation, which depends on the depth of the considered features, may cause convergence and stability problems. Approximation of image Jacobian at convergence is often chosen to simplify the determination problem of the interaction matrix. But this choice induces non optimal Cartesian trajectory for large motion between initial and desired position. Up to now, many solutions have been proposed in the literature to overcome these difficulties. Task function approach used by [5] makes possible the definition of secondary task. In [6] a cost function is added to the primary and priority vision control task to avoid robot joint limits and kinematic singularities. Addition of 3D Cartesian information is also explored. A feature vector augmented with depth information can be considered [7]. Another solution based on the use of Cartesian information was proposed by the authors of [8] named 2-1/2D visual servoing. It uses inputs express partly in the 3D Cartesian space and partly in the 2D image space. This method is based on the
decomposition of the homography matrix computed between current and desired images. In this approach the control of rotation is decoupled from the control of translation. That induces to avoid lost of visibility observed with classical IBVS especially during large rotation motion. Visibility preserving and optimal trajectory have been addressed by [9]. The use of path planning allows the authors to guarantee optimal trajectory and visibility constraint. Different feature preserving and optimal trajectory have been addressed by IBVS especially during large rotation motion. Visibility induces to avoid lost of visibility observed with classical current and desired images. In this approach the control of robot limitations and feature visibility, is presented. In the following, a brief description of NMPC theory is first given and then models used to evaluate this structure are presented. Simulations present the efficiency and the robustness of the proposed strategy. Conclusions and future work are detailed in the last section.

II. NONLINEAR MODEL PREDICTIVE CONTROL

The NMPC is formulated as solving on-line a finite horizon open-loop optimal control problem. This optimization is subject to system dynamics and constraints involving states and input. NMPC is an extension of MPC considering nonlinear system and constraints (see [13] for an introduction to NMPC theory).

The classical problem can be formulated as:

$$\min_{u^N_p} J(x_k, u^N_p)$$  \hspace{1cm} (5)

subject to:

$$x_{i+1|k} = f(x_{i|k}, u_{i|k})$$, \hspace{1cm} $$x_{0|k} = x_k$$  \hspace{1cm} (6)

$$u_{i|k} \in U, \hspace{1cm} i \in [0, N_p - 1]$$  \hspace{1cm} (7)

$$x_{i|k} \in X, \hspace{1cm} i \in [0, N_p - 1]$$  \hspace{1cm} (8)

The nonlinear differential equation (6) describes the dynamics of the system, where $x_{i+1|k}$ represents the predicted state at time $i + 1$ from time $k$. The prediction is initiated with system states at time $k$. $u^N_p$ is a sequence of input vector. $N_p$ is the prediction horizon. $U$ and $X$ are respectively the sets of feasible inputs and states.

The cost function is inspired from LQG theory and is generally defined as a quadratic function of states and control input:

$$J = \Phi(x_{N_p|k}) + \sum_{i=1}^{N_p-1} L(x_{i|k}, u_{i|k})$$  \hspace{1cm} (9)

where $\Phi$ is a terminal constraint on the state at the end of the prediction horizon, called state terminal constraint, and $L$ is a quadratic function of the state and inputs.

Predictive control consists in computing control sequence $u^N_p$ over the prediction horizon $N_p$. In the following, the control input will be kept constant over the prediction horizon $N_p$, which corresponds to a control horizon equal to 1. Optimal control sequence minimizes the cost function $J$ w.r.t. the dynamic of the system (6) and the constraints (7) (8). Then only the first component of the input sequence is applied to the system.

III. APPLICATION TO IBVS

NMPC approach for IBVS is evaluated through the visual servoing of a 6 degrees of freedom (dof) manipulator in eye to hand configuration: the camera looks at a known object attached to the robot end-effector. Fig. 1 describes the control scheme.

A. Simulation setup

1) Robot model and feedback linearization:

The inverse dynamic model of the robot in joint space is written as:

$$A(q)\ddot{q} + h(q, \dot{q}) = \tau$$  \hspace{1cm} (10)

where $q, \dot{q}, \ddot{q}$ are the vectors of joint positions, velocities and accelerations of the robot, $A$ is the inertial matrix, and the vector $h$ stands for the Coriolis, centrifugal, gravity and friction forces. The input $\tau$ is the actuator torque vector. A linearizing and decoupling technique, based on the inverse dynamic model, is considered. Thus, in case of ideal identification of the dynamic model, the controlled system is reduced to a set of double integrators $\ddot{q} = \omega$ where $\omega$ is the new input. An internal state feedback is first applied through diagonal matrix gain $K_v$ on the velocity loop. In Fig.3, $\dot{A}$ and $\dot{h}$ stand respectively for the estimate of $A$ and $h$.

This set of double integrators plus the internal feedback will be considered as the model used for the open loop prediction in the NMPC scheme.
2) Camera model and geometrical configuration:

Classical pinhole model is used to describe the projection in the image plane (see Fig.2).

Object point \( M_i (i \in \{1, \ldots, 4\} ) \), defined by its Cartesian coordinates \( wX \) in frame \( R_w \) (attached to object plane), is related to its image point \( m_i \), defined by its image coordinates \( (p_{u_i}, p_{v_i})^T \) through:

\[
s_i \begin{pmatrix} p_{u_i} \\ p_{v_i} \\ 1 \end{pmatrix} = K P^c T_w \begin{pmatrix} wX_i \\ wY_i \\ wZ_i \end{pmatrix}
\]  

(11)

The pinhole model is composed of:

- \( s_i \), a scale factor induced by perspective projection,
- \( K \), the \((3 \times 3)\) intrinsic parameter matrix, composed of optical center coordinates \( (p_{u_0}, p_{v_0}) \) and focal length \( f \),
- \( P \), a \((3 \times 4)\) projection matrix,
- \( c T_w \), the rigid transformation matrix between object frame and camera frame \( R_c \).

Let’s assume that the transformation between the camera frame and robot base frame is known. In order to focus on the control evaluation, the object has a known geometry and is described by 4 points (so output \( p_k \) is 8-elements vector obtained by stacking image coordinates of each point).

This object is attached to the robot end-effector, with known transformation between end-effector frame \( R_e \) and object frame \( R_w \). Using the direct geometric model, Cartesian coordinates of the end-effector \( \theta X_e \), expressed in robot base frame \( R_0 \), can be computed from joint positions.

Error in image space from current to desired position is defined by:

\[
e_k = p^d - p_k
\]  

where \( p^d \) is the desired feature vector (position of the 4 points in image coordinates).

B. NMPC design

1) Cost function:

The function to be minimized is defined as a quadratic form of measurement errors \( \epsilon_k \) and vector of input control sequence \( u_k^{N_F} \):

\[
J = \frac{1}{2} \sum_{i=1}^{N_F-1} \epsilon_i^T Q \epsilon_i + u_i^T R u_i
\]  

(13)

where \( Q \) and \( R \) denote positive definite, symmetric weighting matrices.

2) Constraint definitions:

Three types of constraints have to be guaranteed for computing physically valid solution of the NMPC. Two of them are related to the robot are respectively the torque constraints

\[
\tau \in \{\tau^{\min}, \tau^{\max}\}
\]  

(14)

and the joint boundaries

\[
q \in \{q^{\min}, q^{\max}\}
\]  

(15)

The third one is associated to the limits of the image called the visibility constraint, ensuring that all the features are always visible:

\[
(p_{u_i|k}, p_{v_i|k}) \in [p_{u_i}^{\min}, p_{u_i}^{\max}, p_{v_i}^{\min}, p_{v_i}^{\max}]
\]  

(16)

with \( p_{u_i}^{\min} = p_{u_i}^{\min} = 0 \) and \( p_{v_i}^{\max} = p_{v_i}^{\max} = 512 \).

IV. SIMULATIONS

Through simulations, two aspects have been evaluated:

1) performances of the proposed approach: convergence and behavior w.r.t. constraints (joint, torque, visibility),
2) robustness w.r.t. robot and camera model error as well as noise measurement.

The simulations have been performed under Matlab 7 on PC pentium IV 2.4 Ghz, using SQP optimization algorithm. The sample time \( \delta_t \) is equal to 5 ms which corresponds to the sampling time obtained with a high speed camera (DALSTAR 1M75) that will be used in the experimental setup. The prediction horizon is equal to \( 10 \times \delta_t \) and the control horizon is equal to 1.

A proportional integral (PI) IBVS controller and the NMPC approach have been compared (see Fig.3). The PI gains have been tuned to reach the best trade-off between speed and stability.

The simulations are performed using the direct dynamic model of 6 dof robot which is linearized through its inverse dynamic model (Fig.3) as explained in section III-A.1. The model used for computing the NMPC scheme is the double set of integrators obtained with the exact feedback linearization. Joint torques are limited to 10 N.m. Object frame is
chosen equal to end-effector frame. The lens focal length of the camera is 10 mm and optical center is located at $(p_{u_0}, p_{v_0}) = (256, 256)$ pixel. Image size corresponds to visibility constraints given by Eq. (16).

A. Performances

First, step response on each Cartesian axis have been tested. Displacement of 0.03 m and rotation of 0.3 rad have been imposed between initial and desired position. In this paper, only results along the X direction are detailed. Performances along the other directions are very similar. Step responses along X direction are compared for both controllers (PI and NMPC), in image space (Fig.4) and in Cartesian space (Fig.5). NMPC ensures convergence as in classical IBVS, but with a much faster response time. Fig. 6 shows torques of the actuators ($\tau_2$, $\tau_4$ respectively for the joints 2 and 4) involved in motion along X direction. Saturation constraint of the control input is quickly reached with the NMPC that is able to deal with these constraints ensuring faster motions with small overshoot while the PI controller gains have been tuned to obtain a trade-off between speed and stability.

Fig.7 illustrates the behavior, when joint motions are bounded. Indeed, during a step along X direction of 0.4 m, the boundary of joint 2 has been reached. In Fig. 7, the continuous line represents the image point coordinate error for each feature obtained with the NMPC approach, which computes a feasible solution taking into account the joint boundaries. Convergence in the image space is guaranteed. The dot line corresponds to IBVS case that leads to a maximum image point coordinate steady state error of 20 pixels.

Finally, the last simulation highlights the efficiency of the NMPC approach for tackling the visibility constraint problem. Indeed during image based visual servoing, possible motion can make features of interest going out of the image which results in infeasible motion. Motion in image space have then to be limited using (16). This constraint avoids to loose object feature during motion. To evaluate the ability of the proposed scheme to deal with this constraint, motion along X direction and a rotation of yaw angle is simulated.
Fig. 7. Errors of feature coordinates in image plane with joint boundaries.

Fig. 8. Feature coordinates in image plane with visibility limits.

Fig. 8 shows the convergence of each feature coordinates in the image plane. The continuous lines are the coordinates of the features controlled by the NMPC approach and the dot lines represent the coordinates obtained with the classical IBVS. The NMPC approach ensures to keep all the features in the image during the motion, whereas some of them are going out of the image with classical IBVS, which will be impossible to control during experiments.

During all the simulations, the mean computation time at each sampling period for solving the optimization problem is about 0.4 s in Matlab environment which does not compromise future real time implementation.

B. Robustness

Evaluation of robustness has been performed successively in presence of i) robot dynamic modelling error and ii) camera calibration error and noise measurement.

Robot dynamic modelling error. Errors on the inverse dynamic model parameters used for the feedback linearization (Fig. 3) have been introduced whereas the model used in the NMPC is still a set of double integrators. Simulations have been computed by adding more or less 30% of error on the parameters. To ensure stability, a terminal constraint on the state $x_T^N, p_x^N$, has been added to the cost function (13). Fig. 9 illustrates that convergence is still guaranteed.

Camera calibration error and measurement noise. Estimation of intrinsic parameters (Eq. 11) is generally subject to calibration error ($f = f + \Delta f$ and $(\tilde{p}_{u0}, \tilde{p}_{v0}) = (p_{u0}, p_{v0}) + (\Delta p_{u0}, \Delta p_{v0})$). Camera calibration error induces errors on feature coordinates in image plane during the prediction phase of the NMPC approach. To compensate this error, an homography matrix between measured and computed image point coordinates is used to adjust the model. Thus at each step, after measuring the image point coordinate related to the current joint position, image point coordinates are computed using the projection model. Then, the homography matrix $H_{\text{error}}$ which can be defined to express the image transformation between the measured image to the estimated image (17) is evaluated. The linear system $Lh_{\text{error}} = g$ is then solved where vector $h_{\text{error}}$ contains the coefficients of matrix $H_{\text{error}}$, matrix $L$ and vector $g$ are obtained by stacking for each point component $L_i$ and $g_i$ (18). As we consider only four points, a constraint must be added on matrix $H_{\text{error}}$ to solve the linear problem. Usually, the homography matrix is chosen in the projective group by setting $H_{\text{error}}(3,3) = 1$.

\[
\delta \begin{pmatrix}
\tilde{p}_u \\
\tilde{p}_v \\
1
\end{pmatrix} = H_{\text{error}} \begin{pmatrix}
p_u \\
p_v \\
1
\end{pmatrix} \quad (17)
\]

\[
L_i = \begin{pmatrix}
\tilde{p}_{u_i} & \tilde{p}_{v_i} & 1 & 0 & \ldots \\
0 & 0 & p_{u_i} & p_{v_i} & \ldots \\
0 & 0 & \tilde{p}_{u_i} & \tilde{p}_{v_i} & 1 & -p_{u_i} & p_{v_i} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & p_{u_i} & p_{v_i}
\end{pmatrix} \quad (18)
\]

\[
g_i = \begin{pmatrix}
p_{u_i} \\
p_{v_i}
\end{pmatrix} \quad (19)
\]

Errors of 20% have been used on intrinsic parameters. Simulations have also been performed with additional white noise of 5 pixels amplitude on the measure output. Fig.10 shows the convergence of the feature coordinates for a step along X direction.

V. Conclusion

In this paper, a new Nonlinear Model Predictive Control scheme applied to IBVS has been proposed. Using cost function based on errors in image plane, convergence and stability of robot motion have been obtained through nonlinear constraint optimization. Constraints due to mechanical system (joint position limits and actuator saturation) and sensor characteristics (visibility constraint) can naturally be taken into account in the predictive control scheme. Simulations
have shown the efficiency of the NMPC approach to overcome classical IBVS difficulties. This new scheme exhibits a good robustness to modelling errors by adding a terminal constraint to ensure the stability and shows good properties in case of camera calibration errors and measurement noise. On going work concerns experimental evaluation to evaluate the proposed approach on robotic platform.

REFERENCES


