

What can Prediction Bring to Image-Based Visual Servoing ?

Guillaume Allibert, Estelle Courtial

▶ **To cite this version:**

Guillaume Allibert, Estelle Courtial. What can Prediction Bring to Image-Based Visual Servoing ?. IEEE International Conference on Intelligent Robots and Systems, Oct 2009, Saint Louis, United States. hal-01994125

HAL Id: hal-01994125
https://hal.science/hal-01994125

Submitted on 25 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

What can Prediction Bring to Image-Based Visual Servoing ?

Guillaume Allibert and Estelle Courtial

***Abstract*— The purpose of this paper is to show what image prediction can bring to Image-Based Visual Servoing. The visual feature prediction is obtained thanks to the interaction matrix. Based on a Model Predictive Control strategy, the visual servoing task is formulated into an optimization problem. The error between the reference features and the predicted features is to be minimized over a receding prediction horizon. Numerous simulations highlight the interest of prediction, especially for difficult configurations such as large motion and rotation.**

I. INTRODUCTION

Since the eighties, visual servoing has been largely developed from a theoretical viewpoint but also from a practical viewpoint. Many applications of visual servoing, among others things, for aerial, submarine or medical robots have been reported in the literature [4],[8],[10]. The fundamental classification of visual servoing distinguishes three approaches: image-based control (2D), position-based control (3D) and a hybrid approach (2 1/2 D). Further details about visual servoing can be found in [5],[6]. Here, we focus our interest on Image-Based Visual Servoing (IBVS). In this context, the control task consists in determining the control input applied to the robotic system according to the error between the reference image and the current image from the camera. We consider a free-flying camera with six degree-of-freedom (6 dof). The relationship between the camera velocity screw τ and the time variation of the visual features $\dot{s}$ is given by:

$$\dot{s}(t) = L_s \tau(t) \tag{1}$$

where L_s is the interaction matrix related to s. The aim of visual servoing is to minimize the error $e(t)$ between the reference features s^*, assumed to be constant, and the measured features $s(t)$. It is defined by:

$$e(t) = s(t) - s^* \tag{2}$$

In order to satisfy an exponential decay of the error (2), mathematically expressed by:

$$\dot{e}(t) = -\lambda e(t) \quad \text{with } \lambda > 0 \tag{3}$$

and considering the open-loop model (1), we obtain the classical feedback control law written as:

$$\tau(t) = -\lambda \widehat{L_s}^{+} e(t) \tag{4}$$

where $\widehat{L_s}^{+}$ is the pseudo-inverse of the approximated interaction matrix.

G. Allibert and E. Courtial are with the Institut PRISME EA 4229, Polytech'Orléans, 8 rue Léonard de Vinci, 45072 Orléans Cedex France, guillaume.allibert@univ-orleans.fr, estelle.courtial@univ-orleans.fr

Another way to deal with IBVS task is to use advanced control laws such as optimal control [11], Linear Matrix Inequalities [3] and Model Predictive Control (MPC) [1], [2],[10],[13]. In [10], a predictive controller is used for motion compensation in target tracking applications. The prediction of the target motion is used to reject perturbation in order to cancel tracking errors. In [13], the predictive controller is used from ultrasound images for a medical application.

The strategy proposed in this paper exploits MPC for visual servoing tasks. The IBVS objective is formulated as solving on-line a nonlinear optimization problem expressed in the image plane [1],[2]. A nonlinear global model combining the robotic model and the camera model allows to predict the evolution of the visual features over a finite receding horizon. The error between the reference features and the predicted model features is to be minimized with respect to the robotic control inputs. This strategy, named Visual Predictive Control (VPC), offers two advantages. First, VPC can easily take into account constraints such as visibility constraints or state/input constraints. The constraint handling, especially visibility constraints, can be very useful to deal with obstacle avoidance. Secondly, the visual prediction can play a crucial role for difficult configurations.

The purpose of this paper is to show what image prediction brings to classical visual servoing. We propose an alternative approach based on the VPC strategy where the visual prediction is obtained no longer by the nonlinear global model but by the interaction matrix. In this case, no 3D data are required contrary to the global model [1]. On the other hand, no 3D constraint would directly be taken into account. The interest of the prediction is pointed out through many simulations describing difficult configurations such as large motion and rotation on a free-flying perspective camera.

The paper is organized as follows. In section II, the context of the study is stated and the principle of MPC is briefly recalled. Then, in section III, the proposed strategy and the interaction matrix is developed. In Section IV, simulation experiments highlight the efficiency of the proposed approach. The influence of the prediction horizon and the weighted matrix are addressed. Finally conclusions and future tasks are detailed in the last section.

II. VISUAL PREDICTION IN IBVS BASED ON OPTIMIZATION PROBLEM

A. Context of the study

Let us consider a point visual feature denoted s. For a 3D point with coordinates $P = (X,Y,Z)^T$ in the camera frame, which is projected in the image plane as a 2D point in

normalized coordinates such that $s=(u,v)^T$ with $u=X/Z$ and $v=Y/Z$, it has been shown that the interaction matrix, related to s, is given by:

$$L_{s_1}=\begin{pmatrix} -\frac{1}{Z} & 0 & \frac{u}{Z} & uv & -(1+u^2) & v \\ 0 & -\frac{1}{Z} & \frac{v}{Z} & 1+v^2 & -uv & -u \end{pmatrix} \quad (5)$$

The value Z is the depth of the point relative to the camera frame. To avoid the estimation of this parameter at each iteration, the depth computed or measured at the reference position Z^* is generally used in the control scheme (4). Consequently, the interaction matrix (5) depends only on the current time measure of the visual feature s.
To control a 6 dof free-flying camera, a minimum of three points is necessary. To a value of s corresponds a single camera pose w.r.t. the target if four points are considered.
The global interaction matrix is then obtained by stacking the L_{s_i} for $i\in[1,...,4]$:

$$L_s=\begin{pmatrix} L_{s_1} \\ L_{s_2} \\ L_{s_3} \\ L_{s_4} \end{pmatrix} \quad (6)$$

In this work, we propose to use (1), (5) and (6) to predict, at the current time, the evolution of the visual features over a finite prediction horizon.

B. Nonlinear Model Predictive Control

Nonlinear Model Predictive Control, also named Receding Horizon Control (RHC), has been largely developed for the control of constrained nonlinear processes [12]. The control problem (setpoint or trajectory tracking) is formulated as solving on-line a nonlinear optimization problem. Based on the process model, the controller predicts the behavior of the system over a prediction horizon N_p. The difference between the reference trajectory and the predicted model behavior defines the cost function J to be minimized with respect to a control sequence $\tilde{u}$. Due to disturbances and model mismatches, this procedure is repeated at each sampling instant. Only the first component of the optimal control sequence is really applied to the process. At the next sampling instant, when the measurements are updated, the finite horizon moves one step forward and the procedure starts again.
The control structure considered is the well-known Internal Model Control (IMC) structure (see Fig. 1).

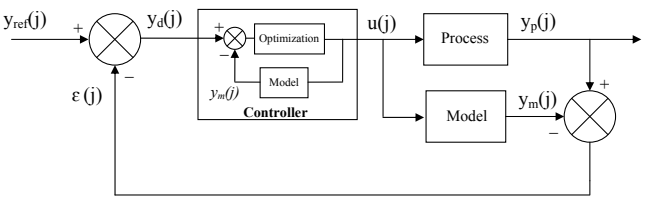


Fig. 1. Internal Model Control Structure

According to this control scheme, we can write:

$$\begin{aligned} y_d(j) &= y_{ref}(j)-\varepsilon(j) \\ y_d(j) &= y_{ref}(j)-(y_p(j)-y_m(j)) \\ y_d(j)-y_m(j) &= y_{ref}(j)-y_p(j) \end{aligned} \quad (7)$$

The tracking of the reference trajectory y_{ref} by the process output y_p is equivalent to the tracking of the desired trajectory y_d by the model output y_m.
The cost function J is defined as a quadratic function of states and control inputs. Due to the IMC structure, J can be written in discrete-time as:

$$J(x,u)=\sum_{j=k+1}^{k+Np}[y_d(j)-y_m(j)]^T Q\,[y_d(j)-y_m(j)] \quad (8)$$

The mathematical formulation of NMPC strategy is then given by:

$$\min_{\tilde{u}}\ J(x,u) \quad (9)$$

subject to:

$$\begin{cases} x(j+1)=f(x(j),u(j)) \\ y_m(j)=h(x(j)) \end{cases} \quad (10)$$

$x\in\mathbb{R}^n$, $u\in\mathbb{R}^m$, $y\in\mathbb{R}^p$ are respectively the states, the inputs and the outputs.
The first nonlinear equation (10) describes the dynamics of the system where $x(j+1)$ represents the predicted state at time $j+1$ from the current time k. The predicted states are initialized with the system states at time k which guarantees the implicit feedback of the IMC structure when modeling errors and disturbances are equal to zero. The error $\varepsilon(j),\ j\in[k+1,k+N_p]$ is assumed to be constant over the prediction horizon and equal to the measurement $\varepsilon(k)$. $\tilde{u}=\{u_k,u_{k+1},...,u_{k+N_c},...,u_{k+N_p-1}\}$ is the optimal control sequence. From $u(k+N_c+1)$ to $u(k+N_p-1)$, the control input is constant and equal to $u(k+N_c)$ where N_c is the control horizon. The weighted matrix Q is a symmetric definite positive matrix.
One of the main advantages of NMPC is the capability to explicitly handle constraints on the states ($x(j)\in\mathbb{X}\subset\mathbb{R}^n,\ j\in[k+1,k+N_p]$, $\mathbb{X}$ is the set of admissible states) and on the control inputs ($u(j)\in\mathbb{U}\subset\mathbb{R}^m,\ j\in[k,k+N_p-1]$, $\mathbb{U}$ is the set of admissible inputs). These constraints are easily added to the problem (9).

III. 2D Visual Servoing through NMPC

The IBVS task is formulated into the minimization of an image error over a prediction horizon.
Due to the IMC structure combined with visual servoing task, the equations (7) become:

$$\begin{aligned} \varepsilon(j) &= s(j)-s_m(j) \\ s_d(j) &= s^*-\varepsilon(j) \end{aligned} \quad (11)$$

where s^*, s and s_m are respectively the reference features, the current features and the predicted model features. The latter are obtained thanks to equation (1), discretized with a first order approximation (T_e is the sampling period):

$$s_m(j+1)=s_m(j)+T_e\,\widehat{L_s}(j)\tau(j) \quad (12)$$

In [9], this dynamic equation is solved to reconstruct the image data in case of occlusion.
Consequently, the 2D visual predictive controller can be written as:

$$\min_{\tilde{\tau}} \quad \{J(s,\tau) = \sum_{j=k+1}^{k+Np} [s_d(j) - s_m(j)]^T \; Q \; [s_d(j) - s_m(j)]\} \tag{13}$$

with:

- $\tilde{\tau} = \{\tau_k, \tau_{k+1}, ..., \tau_{k+N_c}, ..., \tau_{k+N_p-1}\}$: the optimal sequence of the camera velocity screw;
- Q : a symmetric definite positive matrix;
- N_p, N_c : the prediction and control horizons.

The optimization problem (13) is solved at each sampling time T_e. Only the first component of the optimal control sequence, $\tau(k)$, is really applied. Numerous optimization routines are available in software libraries to solve this kind of problem: conjugate gradient methods, quasi-Newton's methods, etc. In our case, a large-scale algorithm is used and more precisely, the function *fminunc* from Matlab optimization toolbox.
To sum up, the different steps of the procedure are:
Step 1: Image acquisition and initializations;
Test 1: $\|s(t) - s^*\| \leq \beta$ (for small $\beta > 0$)
Yes: End ;
No: Step 2;
Step 2: Model initialization with measurements;
Step 3: Solution of the optimization problem (13);
Step 4: Application of the optimal control $\tau(k)$;
Step 5: New image acquisition and go to **Test 1**;

The setting parameters of this approach are the prediction horizon (N_p), the control horizon (N_c) and the weighted matrix (Q):

- the prediction horizon is chosen in order to satisfy a compromise between scheme stability (long horizon) and numerical feasibility in term of computational time required (short horizon);
- the control input is usually kept constant over the prediction horizon which corresponds to a control horizon equal to 1;
- the matrix Q is often the identity matrix or can be a time-varying matrix for stabilization task.

In the next section, the role of these parameters is discussed for different configurations.

IV. SIMULATION EXPERIMENTS

Different simulations illustrate what the image prediction can bring to 2D visual servoing. The results are compared with the classical approach introduced in section I (cf. Equations (1), (2), (3)).
For all simulations, the sampling time T_e is equal to 40 ms. This choice allows to consider a real time application with an usual camera (25 frames per second).
The control task consists in positioning a perspective free-flying camera with respect to the perspective projection of four points. These four points form a square of 20 cm in length in Cartesian space. The reference image is obtained when the target pose expressed in the camera frame (R_C) is equal to $P_{T/C} = (0,0,0.5,0,0,0)^T$ (see Fig. 2). The first three components correspond to the translation expressed in meters and the last three components are the roll, the pitch and the yaw angles expressed in radians. The coordinates of the four points in the reference image are: $s^* = (u_{d1}, v_{d1}, u_{d2}, v_{d2}, u_{d3}, v_{d3}, u_{d4}, v_{d4})^T = (-0.2, 0.2, 0.2, 0.2, 0.2, -0.2, -0.2, -0.2)^T$ (see Fig. 3).

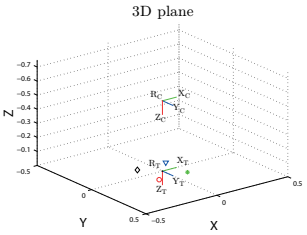


Fig. 2. 3D desired posture

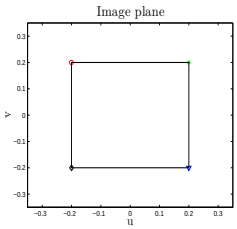


Fig. 3. 2D reference image

To avoid the estimation of the depth parameter $Z(t)$ at each iteration, its value Z^* measured at the reference position is used ($Z^* = 0.5m$). The discrete-time formulation of the interaction matrix (5) is used in the predictive approach.
The VPC requires to set three parameters: the prediction horizon N_p, the control horizon N_c and the weighted matrix Q.

- The control horizon is kept constant and equal to 1 ($N_c = 1$). Only one control is calculated over the prediction horizon.
- The weighted matrix is either the identity matrix $Q = I_{8\times 8}$, constant over the prediction horizon, or a time-varying matrix $Q = Q(t)$, weighting the error at each sampling instant more and more over the prediction horizon. It gives importance to the control objective at the end of the prediction horizon. The time variation of Q is given by: $Q(j) = 2 * Q(j-1)$ with $Q(1) = I_{8\times 8}$.
- The choice of the prediction horizon is discussed below.

Even if VPC can handle constraints, no constraint will be considered here in order to compare the proposed approach with the classical visual servoing. In both cases, the control inputs are normalized if needed. This is accomplished by computing a scaling factor based on the knowledge of the velocity bounds (25 cm/s for the translation speed and 0.25 rad/s for the rotation speed). The scaling factor is similar to multiply all the control inputs by a gain coefficient in order to satisfy the bounds.

A. Pure rotation around the optical axis of the camera

The required camera motion is a pure rotation of $\frac{\pi}{2}$ radians around the optical axis Z. We compare the classical IBVS with three cases:

- $N_p = 1$ (see Fig. 5);
- $N_p = 10$ with $Q = I$ and $Q = Q(t)$ (see Fig. 6, 7);
- $N_p = 20$ with $Q = I$ and $Q = Q(t)$ (see Fig. 8, 9);

For $N_p = 1$, the results are similar to the classical IBVS (see Fig. 4). The visual features try to follow the shortest path in

the image plane which is the straight line. Consequently, the camera moves and rotates respectively along and around the optical axis. The only difference is that the control law of Fig. 4 has an exponential decrease.

Remark: if the depth Z is calculated at each sampling time (Z=Z(t)), the trajectories in the image plane are pure straight lines.

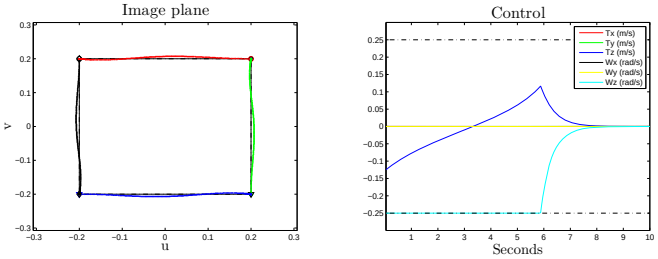


Fig. 4. Classical IBVS

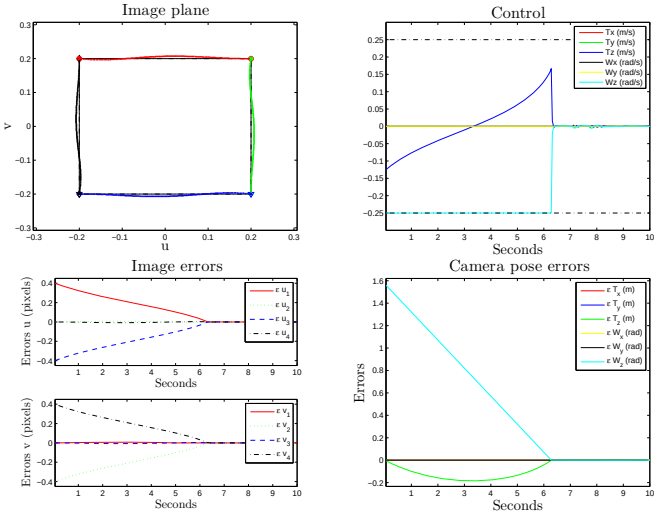


Fig. 5. $N_p = 1$ and $Q = I$

When the prediction horizon increases ($N_p = 10$ or $N_p = 20$), the optimization routine gives a solution which decreases the translation motion around the optical axis whatever the choice of Q. Indeed, the only control which minimizes the cost function over the prediction horizon, consists in achieving only the rotation. Consequently,the image prediction allows to decouple the control in order to effectuate only the desired motion: the pure rotation around the optical axis.

In Fig. 9, the visual feature trajectories are almost circles. The translation along the optical axis is near to zero. Only the rotation is done.

In Fig. 10, the evolution of Δ_Z ($\Delta_Z = \max(Z(t)_{T/C} - Z^*)$ for $t \in [0, 10s]$) is represented versus the horizon prediction for $Q = I$ and $Q = Q(t)$. We can remark that $Q = Q(t)$ enhances the control decoupling.

In the classical IBVS approach, a π radians rotation around the optical axis is known to lead to a failure of the control law [7]. As IBVS strategy chooses the shortest path in the

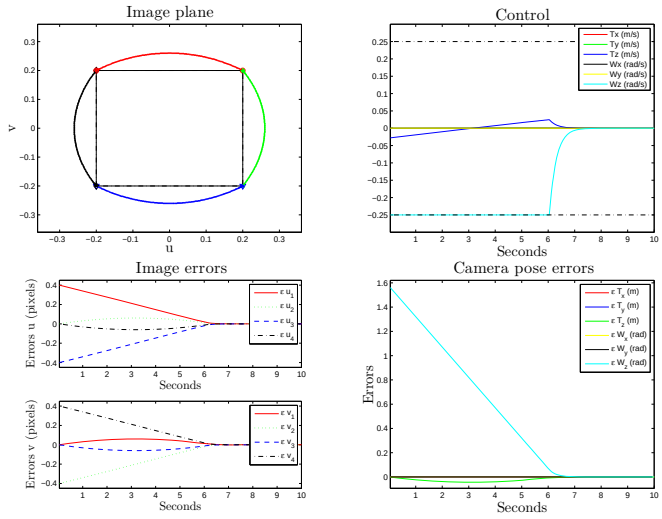


Fig. 6. $N_p = 10$ and $Q = I$

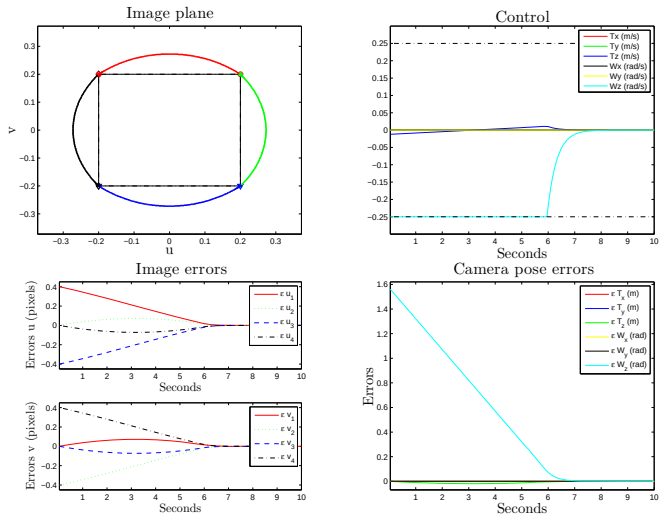


Fig. 7. $N_p = 10$ and $Q = Q(t)$

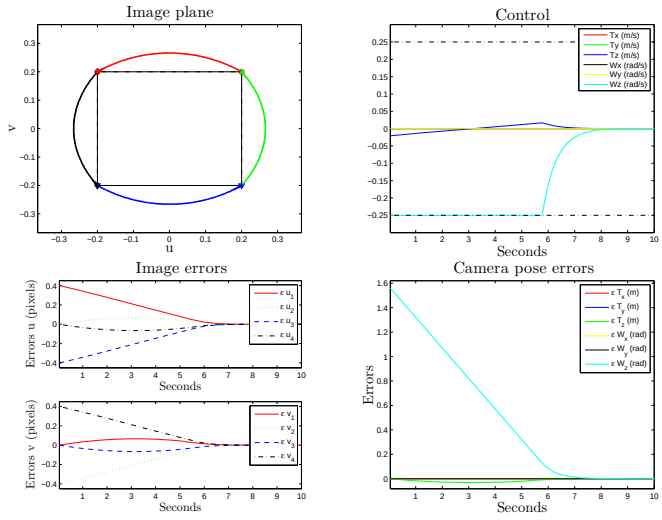


Fig. 8. $N_p = 20$ and $Q = I$

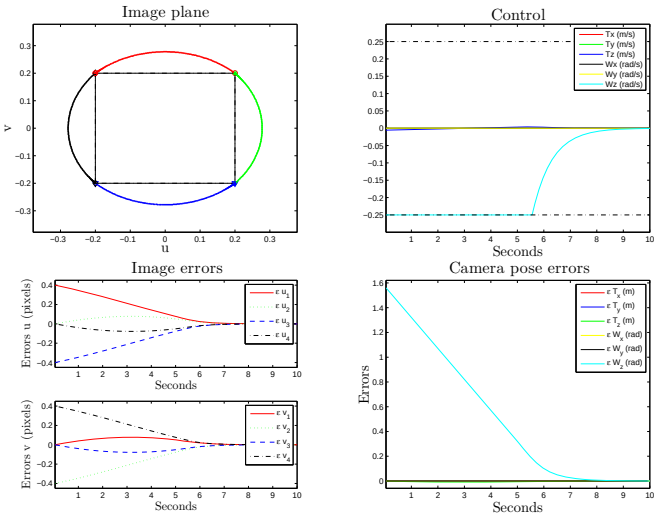


Fig. 9. $N_p = 20$ and $Q = Q(t)$

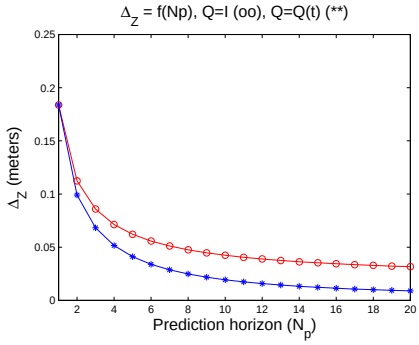


Fig. 10. $\Delta_Z = f(N_p)$ for $\frac{\pi}{2}$ rotation around the optical axis

image plane which is the straight line, the camera performs an infinite retreat corresponding to a singularity. In contrast, the predictive controller with $N_p = 20$ and $Q = Q(t)$ generates a decoupling control: the π rotation around the optical axis is effectuated almost without translation (see Fig. 11).

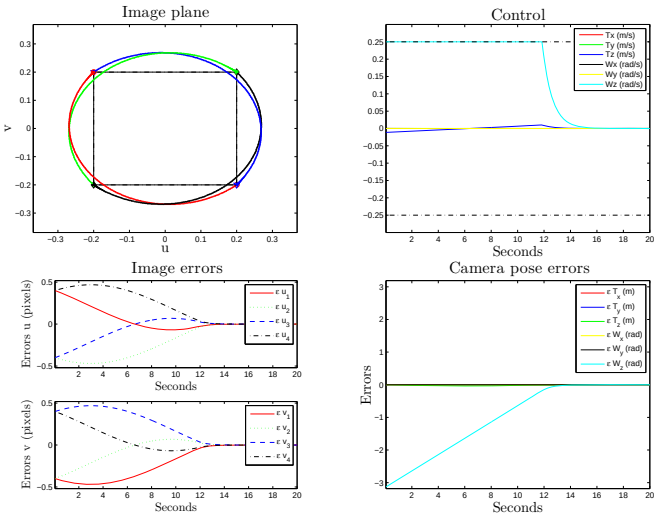


Fig. 11. π rotation: $N_p = 20$ and $Q = Q(t)$

B. *Large displacements*

The initial target pose expressed in the camera frame is given by: $P_{T/C} = (0.08, -0.13, 0.66, 0.46, 0.25, -1.1)^T$.

For $N_p = 1$ (see Fig. 12) and whatever the choice of Q, the obtained results are once again approximately the same as the ones obtained with the classical IBVS. The evolution of the target position (resp. orientation) expressed in the camera frame shows the coupling of the dof involving useless movement.

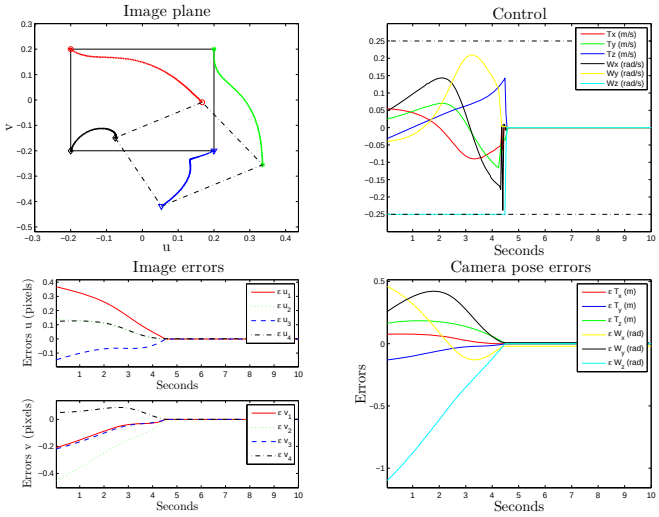


Fig. 12. $N_p = 1$ and $Q = I$

For $N_p = 5$ and $Q = I$ (see Fig. 13), the results are clearly better. The feature trajectories in the image plane are more straightforward and the useless displacements in the Cartesian space are minimized.

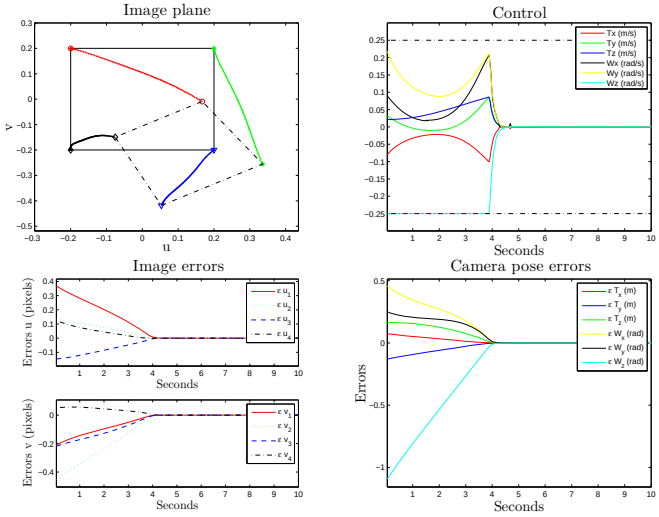


Fig. 13. $N_p = 5$ and $Q = I$

At least, for $N_p = 10$ and $Q = Q(t)$ (see Fig. 14), with approximatively the same feature trajectories in the image plane (the trajectories seem to be smoother), we can remark that the controls obtained are decoupled. This can be verified by the camera position and orientation.

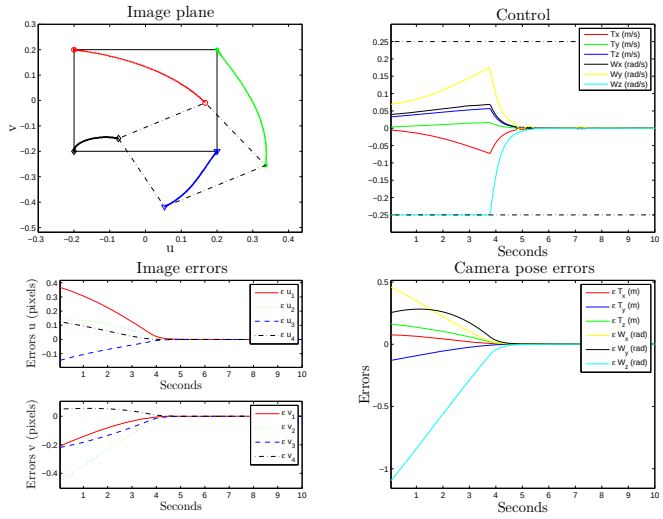


Fig. 14. $N_p = 10$ and $Q = Q(t)$

C. Robustness

The robustness w.r.t modeling errors (25% on the intrinsic camera parameters in the model) and disturbances (white noise added to the output s) is tested (see Fig. 15). Due to the IMC structure, the VPC converges to the desired image in spite of a large displacement to achieve.

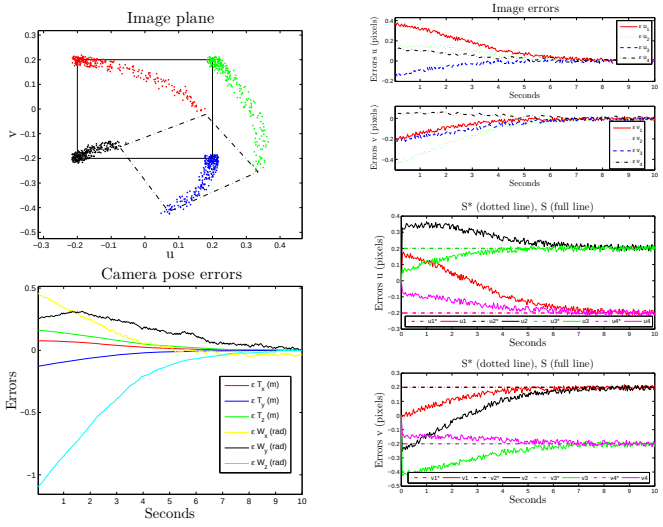


Fig. 15. Robustness simulation: $N_p = 10$ and $Q = Q(t)$

For all simulations presented in this section, the computational time needed to solve the optimization problem is about 30ms. It can be largely reduced in using C/C++.

V. CONCLUSIONS

In this paper, an alternative predictive control approach for 2D visual servoing has been proposed. The control objective is formulated into an optimization problem in the image plane. The difference between the reference features and the predicted features is to be minimized over the predicted horizon in regard to the camera velocity screw inputs. The predicted features are computed thanks to the interaction matrix. This allows to avoid 3D data necessary in the case of a global nonlinear model [1]. We have shown, through simulations, what image prediction can bring to IBVS: a decoupled control. This decoupling control can be of great interest for large displacements or rotations around the optical axis. Due to the IMC structure, the VPC approach is also robust in regard to modeling errors and disturbances.

The visual servoing task formulated into an optimization problem is suited to deal with constraint handling. For instance, visibility constraints which ensure that the visual features stay in the camera field of view, can easily be added to the optimization problem.

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge François Chaumette for his original idea and for the fruitful discussions.

REFERENCES

[1] G. Allibert, E. Courtial, Y. Touré, *Visual Predictive Control for Manipulators with Catadioptric Camera*, IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 510-515, Pasadena, USA, May 2008.

[2] G. Allibert, E. Courtial, Y. Touré, *Real-time visual predictive controller for image-based trajectory tracking of mobile robot*, 17th IFAC World Congress, pp. 11244-11249 Seoul, Korea, July 2008.

[3] P. Danès, D. Bellot, *Towards an LMI approach to multicriteria visual servoing in robotics*, European Journal of Control, 12(1):86-110, 2006.

[4] O. Bourquardez, R. Mahony, N. Guenard, F. Chaumette,T. Hamel, L. Eck, *Image-based visual servo control of the translation kinematics of a quadrotor aerial vehicle*, IEEE Trans. on Robotics, to appear, 2009.

[5] F. Chaumette, S. Hutchinson, *Visual Servo Control, Part I: Basic Approaches*, IEEE Robotics and Automation Magazine, Vol. 14, pp 82-90, December 2006.

[6] F. Chaumette, S. Hutchinson, *Visual Servo Control, Part II: Advanced Approaches*, IEEE Robotics and Automation Magazine, Vol. 14, pp 109-118, March 2007.

[7] F. Chaumette, *Potential problems of stability and convergence in image-based and position-based visual servoing*, The Confluence of Vision and Control, Lecture Note in Control and Informations Systems, Vol 237, pp. 66-78, Springer-Verlag, 1998.

[8] M. Dunbabin, P. Corke, G. Buskey, *Low-cost vision-based AUV guidance system for reef navigation*, IEEE Int. Conf. on Robotics and Automation (ICRA), pp.7-12, New Orleans, USA, April 2004.

[9] D. Folio, V. Cadenat, *Dealing with visual features loss during a vision-based task for a mobile robot*, International Journal of Optomechatronics, 2(3):185-204, July 2008.

[10] J. Gangloff, M. De Mathelin, *Hight speed visual servoing of a 6 dof manipulator using MIMO predictive control*, IEEE Int. Conf. on Robotics and Automation, (ICRA), San Francisco, USA, April 2000.

[11] K. Hashimoto, H. Kimura, *LQ optimal and nonlinear approaches to visual servoing*, in Visual Servoing, K. Hashimoto, Ed. (Robotics and Automated Systems). Singapore:World Scientific, 1993, vol. 7, pp. 165-198.

[12] D.Q. Mayne, J.B. Rawlings, C.V. Rao, P.O.M. Scokaert, *Constrained model predictive control: Stability and optimality*, Automatica, Volume 36, Number 6, pp. 789-814(26), June 2000.

[13] M. Sauvée, P. Poignet, E. Dombre, E. Courtial, *Image Based Visual Servoing through Nonlinear Model Predictive Control*, 45th IEEE CDC, San Diego, USA, December 2006.