Ultrasound image-based visual servoing of a surgical instrument through nonlinear model predictive control

Mickaël Sauvée, Philippe Poignet, Etienne Dombre

Mickaël Sauvée*
Philippe Poignet
Etienne Dombre

LIRMM (Laboratoire d'informatique, de Robotique
et de Microélectronique de Montpellier)
UMR 5506 CNRS Université de Montpellier 2
161 rue Ada, 34392 Montpellier Cedex 5, France
{sauvee, poignet, dombre}@lirmm.fr

Ultrasound Image-Based Visual Servoing of a Surgical Instrument Through Nonlinear Model Predictive Control

Abstract

Ultrasound image-guided interventions are widespread in surgery because of the non-invasive character of the procedures. However, hand/eye synchronization is relatively difficult for a surgeon. Ultrasound image-based visual servoing is one way to perform this kind of surgery. In this work, the control of instrument motion based on ultrasound images through nonlinear model predictive control is investigated. This new scheme ensures the convergence of the instrument to the desired position and also offers the possibility of satisfying constraints such as joint limits, actuator saturation and visibility preserving. This paper describes the proposed controller. The efficiency and the robustness of the proposed solution to control a six degree-of-freedom mechanical system is first illustrated by simulation. Experiments on a Mitsubishi PA10 robot highlight the efficiency of the vision control scheme to handle constraints of ultrasound image-based visual servoing.

KEY WORDS—human-centered and life-like robotics, medical robots and systems, mechanics, design and control, motion control, sensing and perception, computer vision, visual tracking

1. Introduction

Heart valves ensure the blood flows in only one direction (atria to ventricle, ventricle to artery). Valve surgery aims to repair or replace a cardiac valve in case of stenosis (narrowing) or regurgitation (leakage) (For figure see http://www.heart-valve-surgery.com/mitral-valve-regurgitation-symptoms-leaking.php). Originally, these interventions inside the heart were performed on the beating heart. As these operations were performed blindly, results were not guaranteed and depended largely on the surgeon's skill. Since the middle of the last century, cardiopulmonary bypass machines allowed the arresting of the heart while maintaining blood circulation and oxygenation. Valve surgery therefore became open heart surgery, providing direct visual access for the surgeon. However, this method may trigger physiologic complications, and many studies have highlighted the morbidity effects of the cardiopulmonary bypass machine (Picone et al. 1999; Diegeler et al. 2000; Wan et al. 2004) or cardioplegia (Rastan et al. 2005).

From blind and direct visual access intervention, the use of non-invasive imaging techniques can be envisaged. Performing a beating heart mitral valve repair based on ultrasound imaging has been investigated by Downing et al. (2002). The authors evaluate the feasibility of such an approach for suturing anterior and posterior mitral valve leaflets under image guidance; the results are promising.

Ultrasound (US) is very attractive to surgeons for examination or gesture guidance in cases of cardiac pathologies due to non-invasiveness, a sufficient frame rate to temporally discretize the cardiac cycle and a relatively low cost. US images are particularly helpful in percutaneous procedures or biopsies; the surgeon moves an instrument inside the patient and controls its position through a video monitor. However, the surgeon has to achieve hand/eye synchronization which is a complicated task, the difficulty being increased by the fact that he has to infer 3D structures from 2D US planes. The use of a robot with advanced US image-based visual servoing capabilities may therefore be an interesting solution.

The work presented here began as a French project funded by the CNRS (French National Centre for Scientific Research) named GABIE (Active Guidance based on Echographic Im-

* Previous address ECA Sinters, 5 rue Paul Mesplé, BP 10611, 31106 Toulouse Cedex 01, France

ages). The objective of the project was to provide tools for US image-guided robotic systems with surgical applications. More precisely, the project adressed mitral valve surgery under echographic images.

One of the most conventional types of mitral valve surgery consists of repairing broken cordages which work as springs and prevent the valve from opening during the ventricular diastol. The intervention consists of fixing a neo-cordage (Gore Tex fiber) at the mitral valve leaflet on one side and stapling the other side to the ventricle wall. This intervention is achieved by stopping the heart and performing a cardio-pulmonary bypass. Robotizing the intervention should avoid the bypass but introduces several difficulties such as US image analysis, specific instrument development, control design and definition of new robotized surgical techniques. Indeed, one scenario for the intervention could be to define a robotized task where motion of the surgeon's instrument tool tip is guided using information obtained by US image analysis with respect to the echographic plane to ensure the visibility of the instrument.

In the literature, US image-based robots may be charaterized as probe holders or surgical instrument holders. With probe holder robots, visual servoing is used to control the probe position with respect to anatomic structure. Abolmaesumi et al. (2001) describe a robot which manipulates a US probe for medical examinations of the carotid artery. The visual servo-control aims at automatically centering the region of interest. US image-based visual servoing is used by Batcha and Krupa (2006) to position the probe with respect to a desired image to assist the radiologist diagnostic in tumor detection. Simulation results are consistent with the proposed control law.

With instrument holder robots, the visual servoing is used to control the motion of the surgical instrument. Computer-guided pericardiocentesis has been realized based on US imaging (Chavanon et al. 2000). The Padyc robot (Schneider et al. 2000) moves the needle inside a pre-defined region computed on US images. Megali et al. (2001) evaluate the performance of a computer-assisted robotic US image-guided biopsy system in terms of position accuracy and execution time. Results show that accuracy is improved and execution time is decreased with the robotic system. Hong et al. (2004) present a robot specially designed to realize needle insertion for percutaneous cholecystectomy. Based on US imaging, the needle is guided to reach the gallbladder. In the case of transperineal biopsy of the prostate, Phee et al. (2005) also proposed the design of a US image-guided robot. In needle insertion intervention, control of the instrument motion is reduced to one direction (needle axis). Considering a more complex application such as grasping or cutting tissue, at least three degrees of freedom (dof) are needed (two translations and one rotation in the image plane). Thus, visual servoing is more complex. Vitrani et al. (2005) applied image-based visual servoing techniques (defined for classical camera; Hutchinson et al. 1996) to control three dof of an instrument. Experimental results in a box filled with water show good behavior of the proposed strategy and robustness to parameter estimation errors.

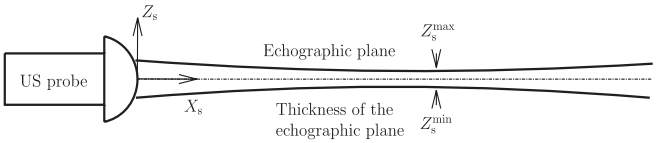


Fig. 1. Schematic view of US plane thickness.

Image guided for minimally invasive surgical atrial fibrillation ablation has been evaluated (Hastenteufel et al. 2006). The authors proposed to control the three dof of a special device (specific to atrial fibrillation ablation) with respect to information extracted from a US image. Stoll et al. (2006) describe how the control of the instrument is based on 3D US images. In this case, visual servoing becomes a position-based visual servoing since information extracted from the 3D US system is equivalent to the Cartesian position of the instrument.

In the above-mentioned examples, specific constraints of the application are not taken into account, particularly visibility constraint; the target must be kept inside the field of view of the probe and must intersect the echographic plane. In this paper, a new US image-based control scheme is proposed through nonlinear model predictive control (NMPC). This predictive approach can be seen simultaneously as an optimal controller and an online open loop motion planning that takes into account the future dynamics and constraints of the system, based on a prediction model of the process. The idea is to naturally handle constraints which can be applied to US Image-Based Visual Servoing (US-IBVS). This approach has already been exposed with a classical camera by (Sauvée et al. 2006). The effectiveness of this control strategy has been demonstrated to guarantee constraints of IBVS: actuator saturations, joint limits and visibility preserving.

This article aims to exploit the capacity of the nonlinear model predictive controller in the context of visual servoing with echographic image sequence to constrain the robot, keeping the instrument in the probe field of view. In-plane motion is measured and tracked. Out-of-plane motion compensation is guaranteed by imposing constraints on the instrument tool tip with respect to the echographic plane. Indeed, the echographic plane may be physically approximated by a thin plate with a thickness of a few millimeters (Figure 1). The instrument tool tip must therefore remain between the upper and lower limits of this plate, imposed by a constraint in the controller.

In Section 2, we provide a brief introduction to NMPC. Details of the NMPC implementation synthesized for US-IBVS are given in Section 3: we describe the models and cost function and introduce the physical limits. Simulations presented in Section 4 highlight the performance of this controller. In

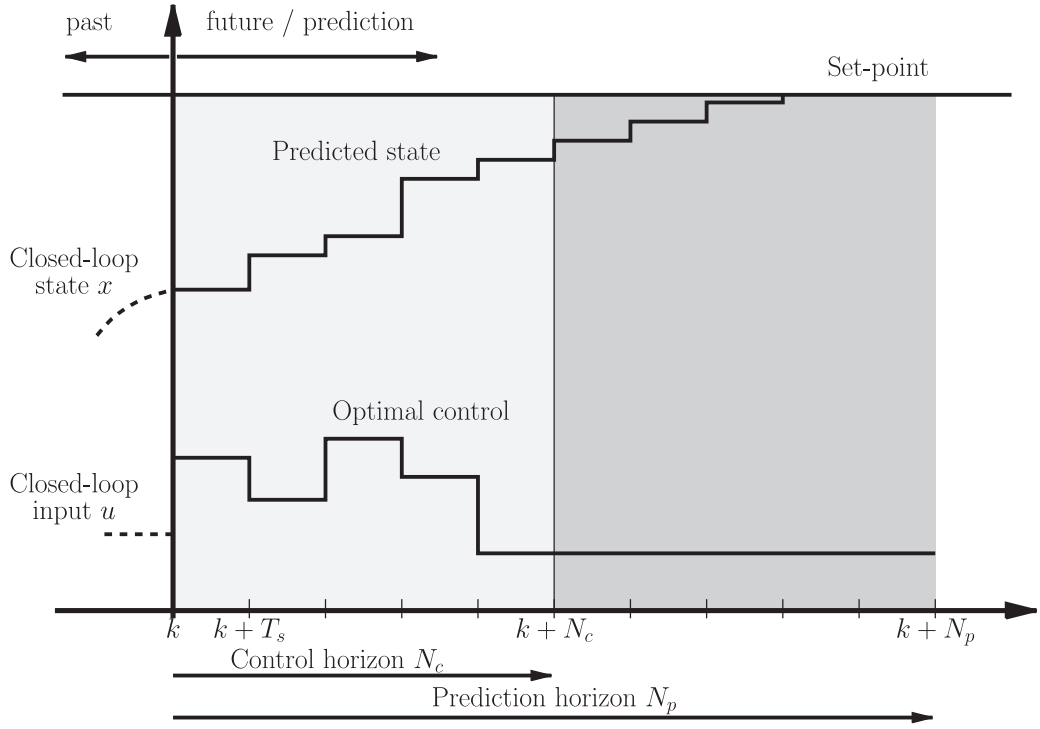


Fig. 2. Principle of model predictive control.

Section 5, experimental results are exhibited. The conclusions and future research are discussed in the final section.

2. 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 inputs. NMPC is an extension of MPC by considering a nonlinear system and constraints (see, for example, Allgöwer et al. 1999 for an introductory overview to NMPC theory).

A schematic representation of predictive control is given in Figure 2. At each iteration k, the future state is predicted from an estimated model of the process over the prediction horizon, and an optimal sequence of control is computed to force the predicted state to converge to the desired set-point. Only the first input of the optimal sequence is injected into the plant. The entire optimization is repeated after output measurement to take into account disturbances and model uncertainties.

The classical NMPC problem of computing an optimal control sequence $\mathbf{u}^{N_p}_{opt|k}$ can be formulated as:

$$\min_{\mathbf{u}_k^{N_p}} \mathcal{C}(\boldsymbol{\epsilon}_k, \mathbf{u}_k^{N_p}) \tag{1}$$

subject to:

$$\mathbf{x}_{i+1|k} = f(\mathbf{x}_{i|k}, \mathbf{u}_{i|k}),\ \mathbf{x}_{0|k} = \mathbf{x}_k \tag{2}$$

$$\boldsymbol{\epsilon}_{i|k} = \mathbf{x}^d - \mathbf{x}_{i|k},\ i \in [1, N_p] \tag{3}$$

$$\mathbf{x}_{i|k} \in \mathbb{X},\ i \in [1, N_p] \tag{4}$$

$$\mathbf{u}_{i|k} \in \mathbb{U},\ i \in [1, N_c],\ \forall i \geq N_c\ \mathbf{u}_{i|k} = \mathbf{u}_{N_c|k}. \tag{5}$$

The nonlinear differential equation (2) represents a model of the dynamics or the evolution of the system, where $\mathbf{x}_{i+1|k}$ represents the predicted state at time $i+1$ from time k. The initial prediction $\mathbf{x}_{0|k}$ is given by the state $\mathbf{x}_k$ measured at time k. $\mathbf{u}_k^{N_p}$ is a sequence $\mathbf{u}_{i|k},\ i \in [1, N_p]$, of input vectors to be minimized. $\boldsymbol{\epsilon}_k$ is a sequence $\boldsymbol{\epsilon}_{i|k},\ i \in [1, N_p]$, of prediction errors between the predicted state and the desired state $\mathbf{x}^d$. N_p is the prediction horizon. N_c is the control horizon, that is the number of degrees of freedom of the control input. $\mathbb{U}$ and $\mathbb{X}$ are the sets of feasible inputs and states, respectively.

The cost function $\mathcal{C}$ is inspired from LQG theory and is generally defined as a quadratic function of states and control inputs:

$$\mathcal{C} = \Phi(\boldsymbol{\epsilon}_{N_p|k}) + \sum_{i=1}^{N_p-1} L(\boldsymbol{\epsilon}_{i|k}, \mathbf{u}_{i|k}) \tag{6}$$

where L is a quadratic function of the states and inputs. Φ is a terminal constraint on the states at the end of the pre-

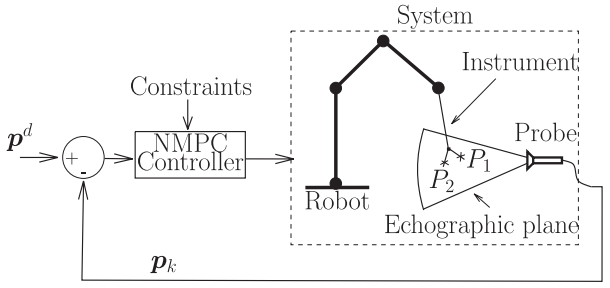


Fig. 3. Control scheme.

diction horizon, referred to as the state terminal constraint, which ensures asymptotic stability (Allgöwer et al. 1999). It should be noted that proof of closed-loop stability is still an open problem, especially when dealing with nonlinear system in the presence of constraints. More details on different approaches for proving stability can be found in Allgöwer et al. (1999).

From a practical point of view, one of the major difficulties is to find an efficient algorithm able to cope with the constrained optimization problem. We will present in the following sections a solution based on the Sequential Quadratic Programming (SQP) algorithm that has already shown its efficiency in such situations.

3. Application to US-IBVS

The proposed NMPC approach for US-IBVS is evaluated through the servoing of a long thin instrument composed of two jaws mounted on a six-dof manipulator. Figure 3 provides a schematic view of the application. In the following subsections, components of the implemented NMPC controller are detailed, including the modeling (robot model, projection model and geometrical configuration), the cost function and the physical limits that are the feasible sets of variables. We also include a subsection concerning model error adjustement that may be useful if modeling error.

3.1. Modeling

3.1.1. Robot Model

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

$$\mathbf{A}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) = \boldsymbol{\tau} \tag{7}$$

where $\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}$ are the vectors of joint positions, velocities and accelerations of the robot, $\mathbf{A}$ is the inertial matrix and the vector $\mathbf{h}$ represents the Coriolis, centrifugal, gravity and friction forces. The input $\boldsymbol{\tau}$ is the actuator torque vector. This is a nonlinear and coupled system. However, taking into account the low level velocity control of robotic systems (high bandwidth with respect to vision loop one), most researchers in visual servoing use a kinematic model (Hutchinson et al. 1996). Therefore, the discrete equation of the system, obtained by a first order discretization, is of the form

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \delta_t \mathbf{u}_k \tag{8}$$

where $\mathbf{q}_k$ represents the state vector, $\mathbf{u}_k$ is the input vector corresponding to the desired velocity and δ_t is the sampling period. This pure integrator will be considered as the model used for the open loop prediction in the NMPC scheme. To take into account the physical limits of the robot, input saturation and joint limits have been included (see Section 3.3).

3.1.2. Geometrical Configuration

To predict the position of the instrument in the image, the coordinates of the intersection points P_1 and P_2 (Figure 4) between tool tip and echographic plane must be expressed with respect to probe frame. The following frames are therefore defined to link the joint configuration of the robot to the coordinates of the two observed points in the image (Figure 4):

- the robot base frame R_b (reference frame),
- the end-effector frame R_e,
- the tool frame R_t,
- the probe frame R_s (the echographic plane relies on plane $X_s Y_s$).

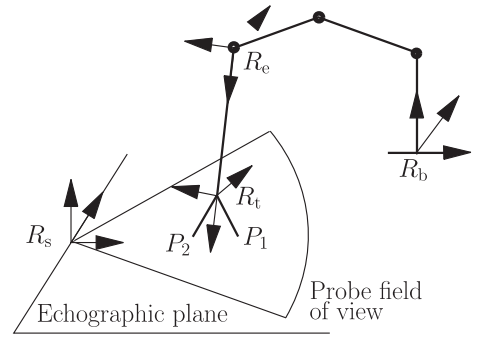
The direct geometric model (DGM) of the robot is used to define the rigid transformation from the robot base frame to the end-effector:

$${}^{b}\mathbf{T}_e = DGM(\mathbf{q}). \tag{9}$$

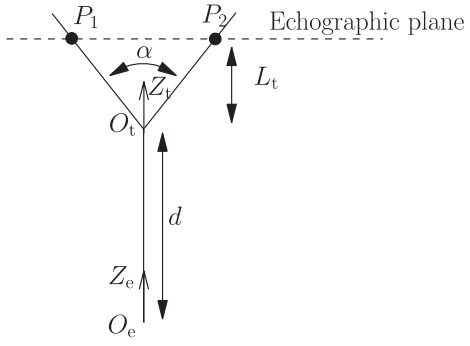
The instrument is directly attached to the end-effector. It is a forceps with two jaws as sketched in Figure 4. The tool frame R_t is located at point O_t. Using the rigid transformation between R_t and R_e (translation of distance d along the Z_e axis of R_e), we can define the position of point O_t with respect to R_b.

$${}^{b}\mathbf{I} = {}^{b}\mathbf{T}_e \, {}^{e}\mathbf{T}_t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{10}$$

The instrument intersects the echographic plane in two points: P_1 and P_2. Let L_t be the distance between the echographic plane and point O_t. L_t is measured at the beginning of



(a)



(b)

Fig. 4. (a) Frame description and (b) tool geometric parameters.

the experiment, and this initial value is subsequently used to predict the locations of P_1 and P_2 in frame R_s throughout the experiment. The instrument tool tip intersects the echographic plane for a given opening angle α ensuring that the points P_1 and P_2 are inside the segments O_tM_1 and O_tM_2, M_1 and M_2 being the extremities of the tool (Figure 4). The coordinates of the control points with respect to frame R_t are therefore defined by:

$$ {}^{t}\mathbf{P}_1 = \begin{pmatrix} -L_t \sin\left(\frac{\alpha}{2}\right) \\ 0 \\ L_t \end{pmatrix} \tag{11} $$

and

$$ {}^{t}\mathbf{P}_2 = \begin{pmatrix} L_t \sin\left(\frac{\alpha}{2}\right) \\ 0 \\ L_t \end{pmatrix}. \tag{12} $$

The coordinates of these two points are defined in R_s using the rigid transformation between R_b and R_s:

$$ {}^{s}\mathbf{P}_i = {}^{s}\mathbf{T}_b \, {}^{b}\mathbf{T}_e \, {}^{e}\mathbf{T}_t \, {}^{t}\mathbf{P}_i, \ i = 1, 2. \tag{13} $$

Two points should be noted here. First, the instrument is assumed to be perpendicular to the ultrasound image here. However, the proposed algorithm is based on the control of the error in the US images between the measured positions P_1 and P_2 and their desired positions as well as their depths with respect to the image plane. If the instrument is rotated along the line P_1P_2 (Figure 4), there will be no change in the definition of equations (11)–(12). If the ultrasound plane is not orthogonal to the direction of the instrument, equations (11)–(12) will have to be slightly modified or else additional constraints may have to be introduced. Second, we will also assume that the ultrasound probe does not move with respect to the robot base.

3.1.3. US Image Projection Model

The US image is modeled

$$ \begin{aligned} {}^{im}\mathbf{P}_i &= \begin{pmatrix} p_{u_i} \\ p_{v_i} \\ 1 \end{pmatrix} \\ &= \begin{bmatrix} S_x & 0 & 0 & u_s \\ 0 & S_y & 0 & v_s \\ 0 & 0 & 0 & 1 \end{bmatrix} {}^{s}\mathbf{P}_i, \ i = 1, 2. \end{aligned} \tag{14} $$

S_x and S_y are scale factors along the direction X_s and Y_s in meter by pixel and u_s and v_s represent the coordinates of the origin of the US rays. p_{u_i} and p_{v_i} are the pixel coordinates of point P_i, $i = 1, 2$. It should be noted that the estimation of the US probe parameters is achieved experimentally by performing motion of the instrument tool tip in the image plane. The Cartesian position of points P_1 and P_2 are then related to their image positions to compute the US probe parameters by way of a least squares method. Their values will be given in Section 5.

3.2. Cost Function

The function to be minimized is defined as a quadratic form of prediction errors $\boldsymbol{\epsilon}_{i|k}$ and the vector of control input $\mathbf{u}_{i|k}$:

$$ \mathcal{C} = \frac{1}{2} \sum_{i=1}^{N_p - 1} \boldsymbol{\epsilon}_{i|k}^T \mathbf{Q} \boldsymbol{\epsilon}_{i|k} + \mathbf{u}_{i|k}^T \mathbf{R} \mathbf{u}_{i|k} \tag{15} $$

where $\mathbf{Q}$ and $\mathbf{R}$ denote positive definite symmetric weighting matrices. $\boldsymbol{\epsilon}_{i|k}$ is equal to $\mathbf{p}^d - \mathbf{p}_{i|k} \cdot \mathbf{p}_{i|k}$ are the predictions over N_p of the pixel coordinates of points P_1 and P_2 in the image at the instant $k + i$. This prediction is obtained via equations

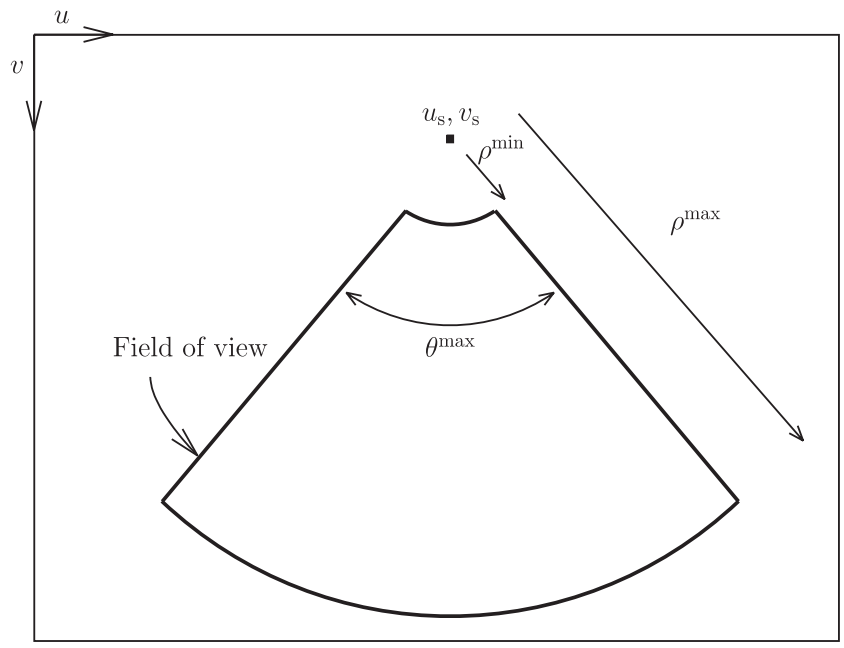


Fig. 5. Representation of US image with probe field of view.

(8)–(14). $\mathbf{p}^d$ is the desired pixel coordinate position in the image. The initial value $\boldsymbol{\epsilon}_{0|k}$ is given by $\boldsymbol{\epsilon}_{0|k} = \mathbf{p}^{\mathrm{d}} - \mathbf{p}_{0|k}$ where $\mathbf{p}_{0|k}$ are the pixel coordinates of points P_1 and P_2 measured in the image at instant k. $\mathbf{u}_{i|k}$ is the sequence of inputs to be optimized over the horizon. For computational reasons, the control horizon N_c (equation (5)) will be set equal to 1 in the following. This means that the control vector is kept constant over the horizon and that the control input has only one degree of freedom. This is a classical assumption when using a predictive approach.

3.3. Physical Limits

Four types of constraints have to be guaranteed to compute a physically valid solution of the US-IBVS. Two of them, related to the robot, are the input constraints

$$\dot{\mathbf{q}} \in \{\dot{\mathbf{q}}^{\min}; \dot{\mathbf{q}}^{\max}\} \tag{16}$$

and the joint boundaries

$$\mathbf{q} \in \{\mathbf{q}^{\min}; \mathbf{q}^{\max}\}. \tag{17}$$

The third, associated with the limits of the image, are referred to as visibility constraints and ensure that all the features will always be visible during motion. With a US probe, the field of view is a part of the disc as shown in Figure 5. This area can easily be defined in polar coordinates by:

$$\begin{aligned} p_u &= u_{\mathrm{s}} + \rho \sin(\theta) \\ p_v &= v_{\mathrm{s}} + \rho \cos(\theta). \end{aligned}$$

Conversely, for a point defined by coordinates (p_u, p_v), its polar coordinates are

$$\begin{aligned} \rho &= \sqrt{p_u^2 + p_v^2} \\ \theta &= \tan^{-1}\left(\frac{p_u - u_{\mathrm{s}}}{p_v - v_{\mathrm{s}}}\right). \end{aligned}$$

The visibility constraint for in-plane motion with respect to the probe field of view is expressed

$$\rho \in [\rho^{\min}; \rho^{\max}] \tag{18}$$

and

$$\theta \in \left[-\frac{\theta^{\max}}{2}; \frac{\theta^{\max}}{2}\right]. \tag{19}$$

To avoid lost of visibility due to being out of plane motion, an additional constraint must be imposed: the tool must intersect with the echographic plane. Since in practice the plane is a slice, which is very thin, the positions P_1 and P_2 along Z_{s} axis of frame R_{s} (Figure 1) must satisfy:

$$P_i^Z \in [Z_{\mathrm{s}}^{\min}, Z_{\mathrm{s}}^{\max}], \ i = 1, 2 \tag{20}$$

where $Z_{\mathrm{s}}^{\min}$ and $Z_{\mathrm{s}}^{\max}$ are the maximum displacement below and above the theoretical echographic plane, respectively.

3.4. Model Error Adjustment

Modeling error may be corrected by using a model error adjustment (Richalet 1993). In our case, the adjustment is simply

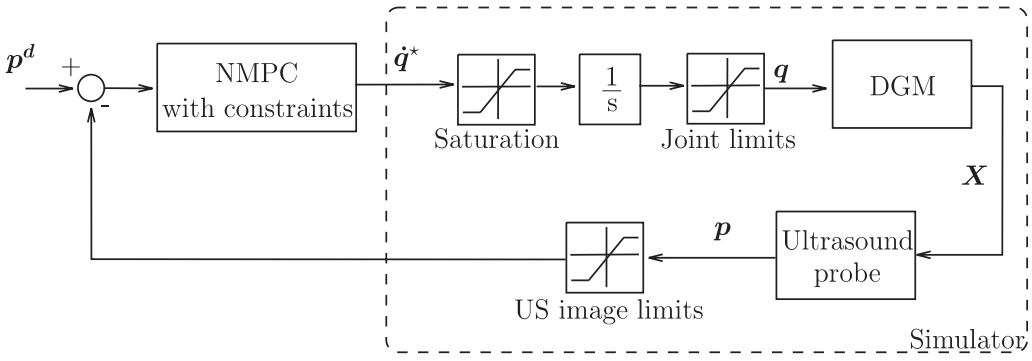


Fig. 6. Control scheme on a constrained linear model.

defined by the difference $\Delta\mathbf{P}_i$ between the predicted pixel coordinate output in the image and its measurement computed at each iteration k. Thus, the following correction has been introduced to the model:

$$^{\text{im}}\mathbf{P}_i^{\text{correct}} = {}^{\text{im}}\mathbf{P}_i + \Delta\mathbf{P}_i,\ i = 1, 2. \tag{21}$$

This compensation is based on a separated correction on P_1 and P_2, which neglects the angular discrepancies. However, this simple compensation provides acceptable simulated and experimental results, as shown in the following sections.

4. Simulation

Simulations have been performed to highlight the performance and robustness of the proposed approach. Gathering all the equations of Sections 3.1–3.3, the implemented NMPC scheme for US-IBVS is then given by:

$$\mathcal{C} = \frac{1}{2}\sum_{i=1}^{N_p-1} \boldsymbol{\epsilon}_{i|k}^T \mathbf{Q}\boldsymbol{\epsilon}_{i|k} + \mathbf{u}_{i|k}^T \mathbf{R}\mathbf{u}_{i|k} \tag{22}$$

subject to

$$\mathbf{q}_{i|k+1} = \mathbf{q}_{i|k} + \delta_t \mathbf{u}_{i|k},\ \mathbf{q}_{0|k} = \mathbf{q}_k \tag{23}$$

$$\boldsymbol{\epsilon}_{i|k} = \mathbf{p}^d - \mathbf{p}_{i|k} = \mathbf{p}^d - g(\mathbf{q}_{i|k}) \tag{24}$$

$$\dot{\mathbf{q}} \in \{\dot{\mathbf{q}}^{\min}; \dot{\mathbf{q}}^{\max}\} \tag{25}$$

$$\mathbf{q} \in \{\mathbf{q}^{\min}; \mathbf{q}^{\max}\} \tag{26}$$

$$\rho(\mathbf{p}_{i|k}) \in [\rho^{\min}; \rho^{\max}] \tag{27}$$

and

$$\theta(\mathbf{p}_{i|k}) \in \left[-\frac{\theta^{\max}}{2}; \frac{\theta^{\max}}{2}\right] \tag{28}$$

$$P_i^Z \in \{Z_{\text{s}}^{\min}; Z_{\text{s}}^{\max}\},\ i = 1, 2 \tag{29}$$

where g represents geometric and projection models (equations (13)–(14)).

Simulations have been performed using an SQP optimization algorithm (*fmincon* function in Matlab 7 Optimization Toolbox). This classical package is able to cope with the constrained optimization problem as stated by equation (22) subject to equation (23) which corresponds to the box 'NMPC with constraints' in Figure 6. The sampling time δ_t is equal to 40 ms which is the sampling time obtained with a US acquisition system. The image size is set to 768 × 576 pixels and the coordinates of the origin of the US probe field of view is equal to (340, 40) pixels. The scale factors (S_x, S_y) are set to $(2500, 3000)\,\text{pixel}\,\text{m}^{-1}$. The image size and shape will also correspond to the visibility constraint given by equations (18) and (19), i.e. $\rho^{\min} = 40$ (pixels), $\rho^{\max} = 400$ (pixels), $\theta^{\max} = 80^\circ$. The constraint on the point position with respect to the echographic plane is defined by the interval $[-1\,\text{mm}, 1\,\text{mm}]$ along Z_{s} (Figure 1). These conditions are very similar to the experimental setup detailed in Section 5.1.

These simulations are the first stage in the evaluation and tuning of the algorithm. Section 4.1.1 is related to a practical question which arises when using the predictive approach, i.e. the choice of the prediction horizon, defined as a compromise between performance and computation time. The following sections discuss behavior performance during step responses. Behavior with respect to constraints (joint, input and visibility) has also been evaluated as well as robustness with respect to US projection model errors, noise measurement and in the presence of outliers.

4.1. Performances

4.1.1. Influence of N_p

Simulations have been performed to evaluate the influence of the prediction horizon length. Figure 7a displays the responses for a step along the pixel coordinate axis u for N_p equal to 2, 3 and 5. The system converges to the desired set-point, but the fastest response is obtained for $N_p = 2$. Practically, the optimal control obtained through minimization of the cost function

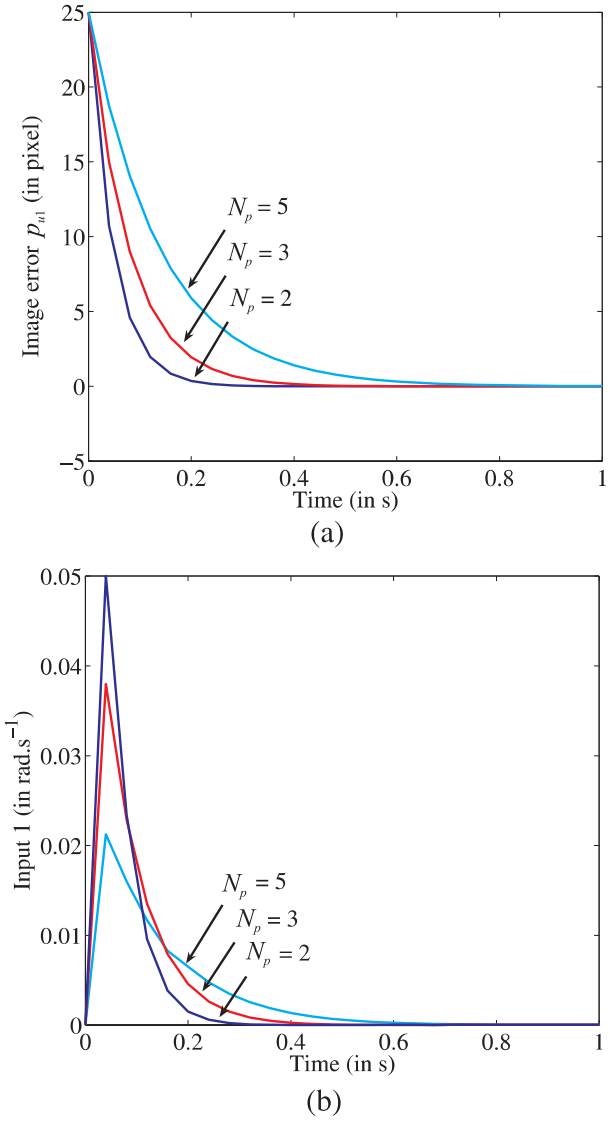


Fig. 7. Controller behavior with respect to N_p: (a) image coordinate error of point 1 and (b) optimal input of actuator 1.

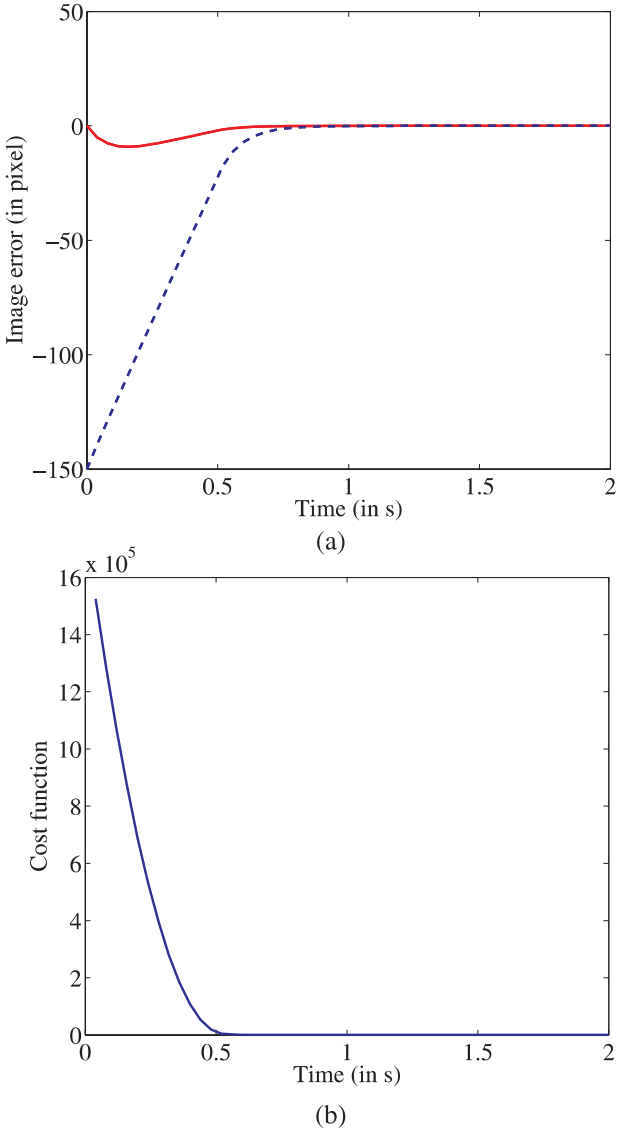


Fig. 8. Step response for a translation along v: (a) evolution of image error and (b) cost function.

is computed over the prediction horizon. With a short prediction horizon, the predicted state therefore has to reach the desired set-point more quickly than with a longer prediction horizon. The counterpart is that the controller with $N_p = 2$ is more demanding for the actuators than for $N_p = 5$ (Figure 7b). In the following simulations, a trade-off has been made between a fast response and actuator torques; therefore N_p is set at 3.

4.1.2. Step Response

Step responses along pixel coordinate axes u and v (translations in the image plane) and a rotation around the normal to the image plane have been tested. Since results obtained for the translation along u have already been presented in the previous section, only the results of the translation along v (Figure 8) and the rotation (Figure 9) are presented here. In each case, the cost function evolution is also plotted. The NMPC controller ensures convergence for each possible motion in the US image. As seen in Figures 8 and 9, the convergence for translation motion is faster than for rotational motion (0.6 s compared to 1.4 s).

4.2. Constraints Satisfaction

For sake of brevity, joint limit constraint satisfaction is not exposed in this paper; see Sauvée et al. (2006) for a presentation of this point.

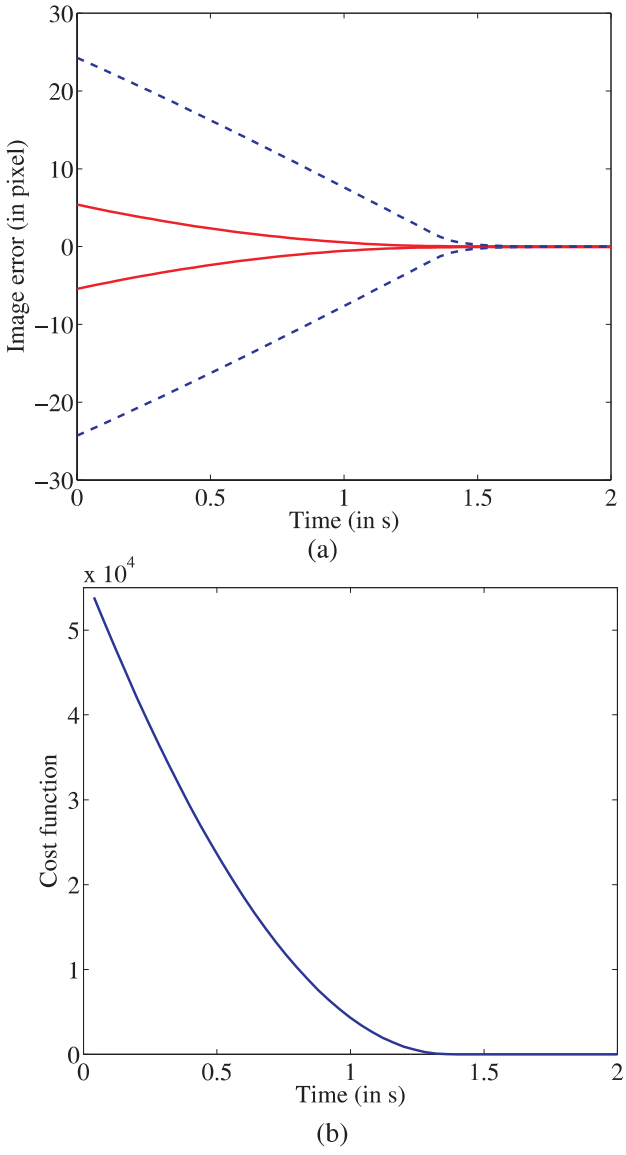


Fig. 9. Step response for a rotation about the normal to the image plane: (a) evolution of image error and (b) cost function.

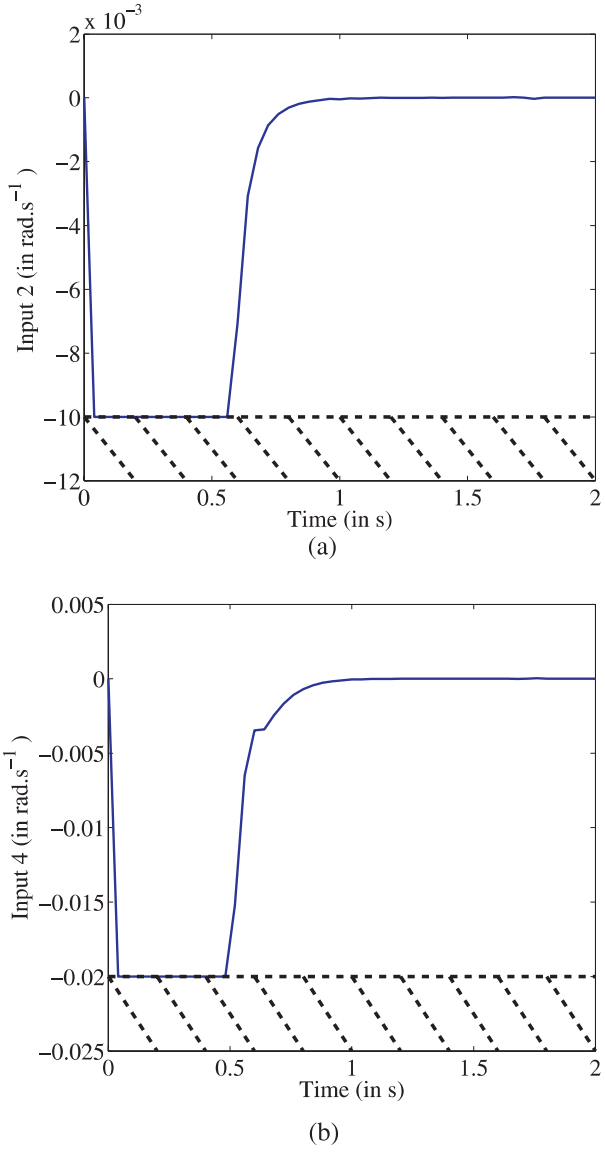


Fig. 10. Input velocities with saturation constraints for a complex motion: (a) input 2 and (b) input 4.

To highlight constraint satisfaction on actuator saturation, a motion of 100 pixels has been simulated along pixel coordinate axis v. Figure 10 shows input velocities of actuators involved ($\dot{q}_2$, $\dot{q}_6$ for joints 2 and 6, respectively) in this motion. Saturation constraint of the control input is quickly reached, ensuring fast motions, and is naturally handled by the controller.

The visibility constraint is satisfied when both points remain in the field of view of the probe and intersect the US plane. To emphasize the respect of the field of view, simulations of a desired motion with two different fields of view (normal and reduced) have been achieved. In case of a reduced field of view, the desired position is close to the limit which induces a change in the controller behavior to ensure that the image instrument trajectory respects the visibility constraint (Figure 11).

The height of points P_1 and P_2 with respect to frame R_s is limited to ± 1 mm around 0. The evolution of the coordinates along the Z_s axis is plotted in Figure 12 for a complex motion composed of two translations and one rotation. The limit is reached for both points and, in this case, a small rotation along the Y axis of frame R_t can be observed (${}^sP_1^Z = 1\,\text{mm}$ and ${}^sP_2^Z = -1\,\text{mm}$).

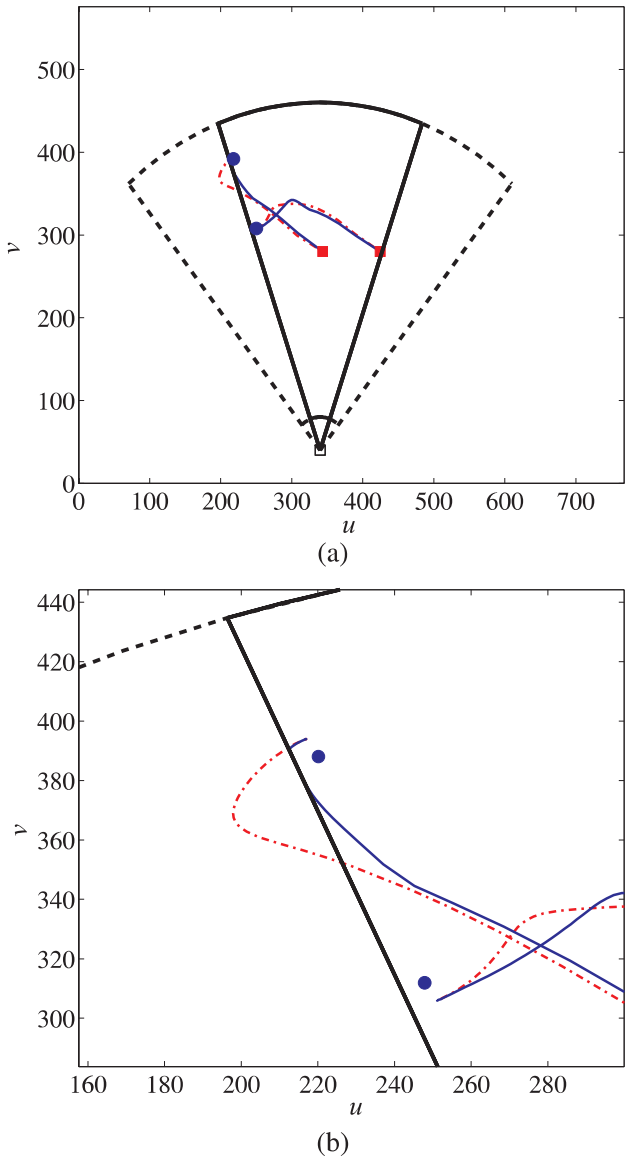


Fig. 11. Visibility constraints: (a) image trajectories of the two points for the same initial and desired positions but with different fields of view (dashed line: normal, continuous line: reduced) and (b) zoom on in the region of the desired position.

4.3. Robustness

The robustness of the proposed control scheme has been evaluated in the presence of (1) outlier and noise measurement and (2) modeling error.

Feature extraction is sensitive to outlier detection and noise measurement. Outliers can occur for different reasons such as change in image intensity or occlusions. Control based on visual servoing must be robust to outliers. Then outliers have been randomly introduced in the simulation on the output of the projection model by adding given value on measured data. Behavior of the proposed scheme is shown on Figure 13. Convergence is still observed in the image (Figure 13a and b) and in Cartesian space (Figure 13c) as long as the measurement remains visible in the image.

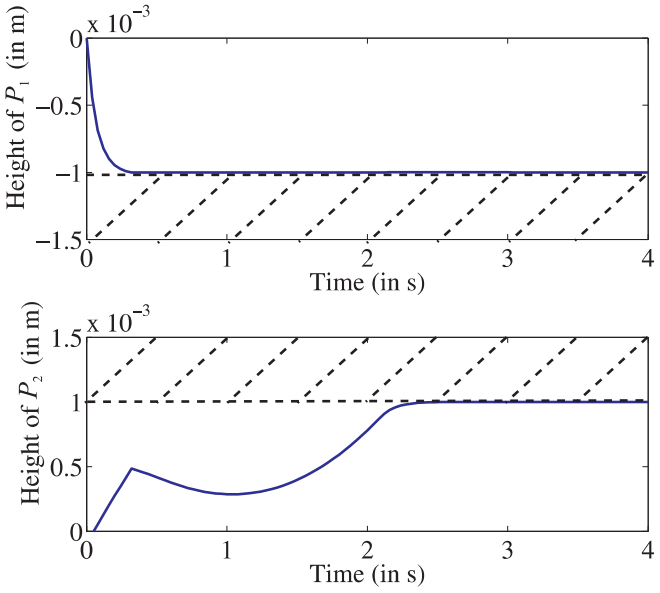


Fig. 12. Visibility constraints: the height of points P_1 and P_2 is kept within the desired intervals.

Simulations under white noise with 5 pixels amplitude have also been performed. Convergence and constraint satisfaction are still observed (Figure 14).

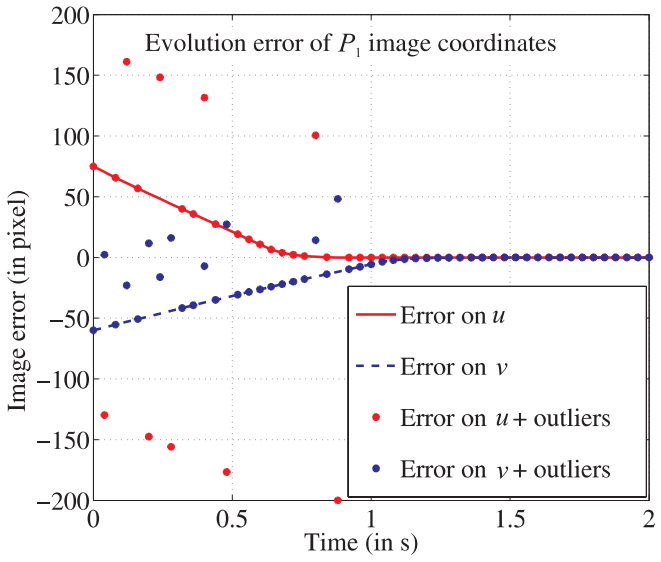
Finally, an error of 30% has been added in simulation on the projection model (equation (14)). Simulations performed with the model error adjustment exhibit a good behavior. Convergence to the desired position is observed (Figure 15a), even if the model gives a position outside the field of view (Figure 15b).

5. Experiments

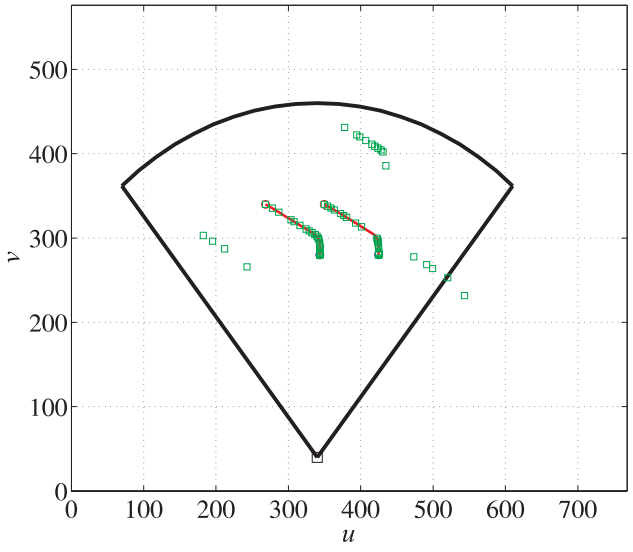
5.1. Experimental Setup

5.1.1. Robotic Platform

The proposed control algorithm has been implemented on a Mitsubishi PA10 robot. The experimental setup is depicted in Figure 16. This is a lightweight robot with seven revolute joints. However, only joints from numbers 2–7 have been considered for this experimental evaluation in order that control issues due to redundancy can be neglected. Velocity has been limited to 1% of its maximum value on each axis (see Table 1) to cope with the harsh exciting conditions of the step response. The instrument as described in Figure 4 is attached to the robot



(a)



(b)

Fig. 13. Robustness to outlier detection: (a) image error and (b) image trajectory.

Table 1. Joint limits (rad)and actuator saturation (rad s^{-1}).

1	2	3	4	5	6
$\pm\pi/2$	$\pm\pi$	± 2.5	$\pm 3\pi/2$	$\pm\pi$	$\pm 2\pi$
0.01	0.01	0.02	0.02	0.063	0.063

end-effector. A standard Pentium II PC running under QNX real-time operating system ensures communication via Arcnet

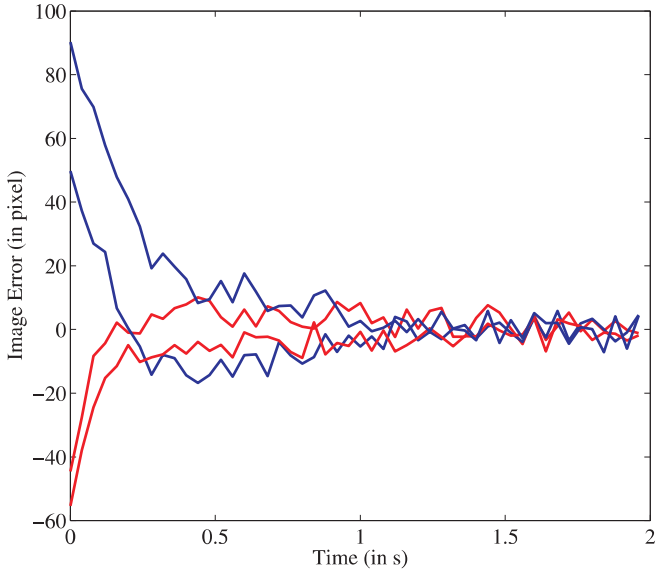


Fig. 14. Robustness with respect to noise measurement.

Table 2. Estimated parameters of the US probe.

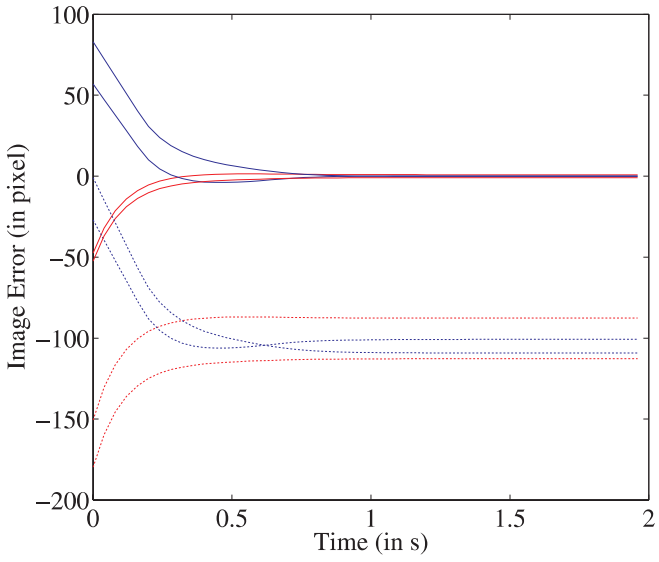
S_x (in pixel m^{-1})	2450
S_y (in pixel m^{-1})	2900
u_s (in pixel)	340
v_s (in pixel)	150

LAN with a low-level control unit at a sampling rate of 10 ms. In the setup, the computer receives the desired joint velocities generated by the NMPC controller through an Ethernet link and sends them to the robot controller. Due to the large sample rate of the vision control loop with respect to the robot control one, a zero-order hold has been implemented.

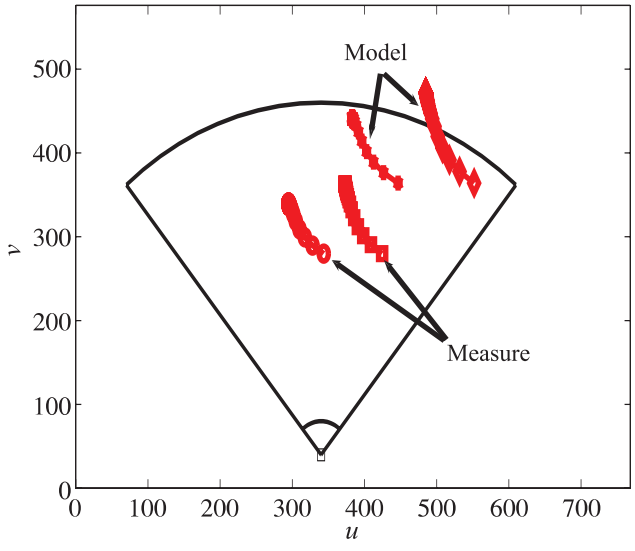
5.1.2. Image Acquisition and Analysis

We used an Acuson Cypress® from Siemens with a transthoracic probe. Display output is used to acquire images through a Matrox Meteor II acquisition board at a frame rate of 25 Hz. The estimation of the US probe parameters has been achieved experimentally by performing motion of the instrument tool tip in the image plane. The Cartesian position of points P_1 and P_2 are then related to image point positions to compute the US probe parameters through a least squares method. Estimated parameters for the US projection model are listed in Table 2.

The instrument position is extracted from US images through a blob analysis. First, the image is binarized with respect to a fixed threshold. Opening and closing operations are

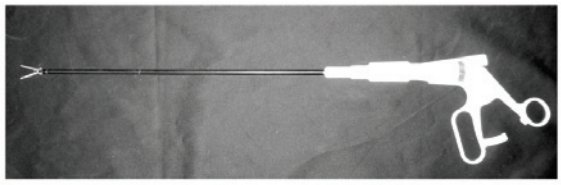


(a)

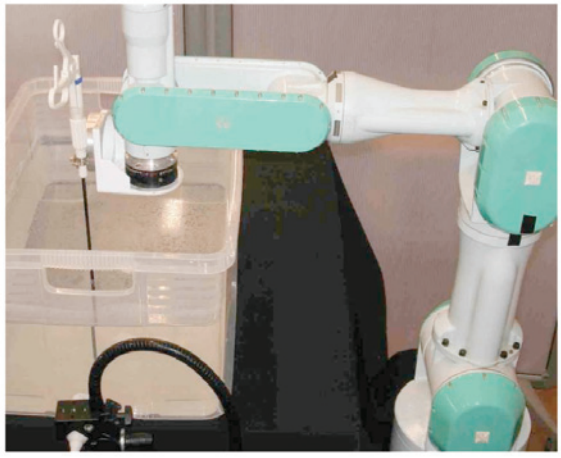


(b)

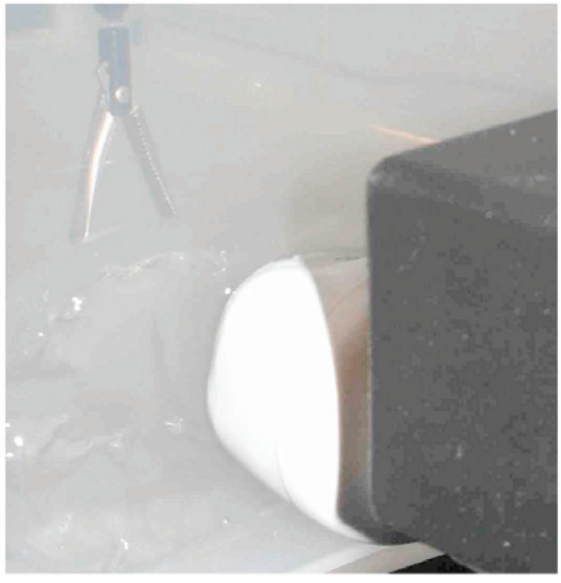
Fig. 15. Robustness with respect to projection model errors: (a) image error (measured: solid line, modeled: dash-dot line) and (b) image trajectories.



(a)



(b)



(c)

Fig. 16. Experimental setup: (a) minimally invasive instrument with tool tip composed of two jaws, (b) overview of the robot which holds the tool in a box full of water and (c) detail of the setup with the US probe in contact with the box.

then performed to remove small particles and fill holes that can appear. Finally, detection of connected pixels (blob) is realized. The selection of the good blobs is based on area analysis (image of the instrument is represented by two blobs of minimal area). Thus, image analysis provides the coordinates of the image instrument points by computing blob centers of gravity. All these operations are performed directly using the acquisition board capabilities and the MIL library.

5.1.3. Software Architecture

The vision control loop has been implemented on a Pentium IV bi-Xeon PC running under Windows 2000. The architec-

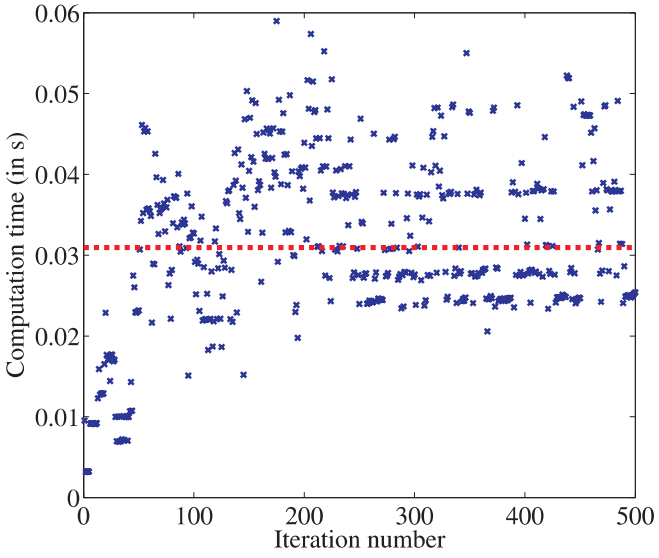
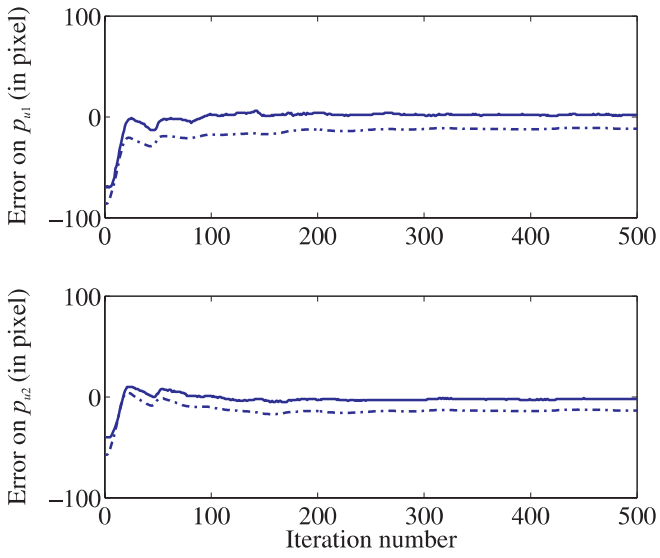


Fig. 17. Evolution of the optimization thread computation time for a complex motion (mean time is represented by the dashed line).

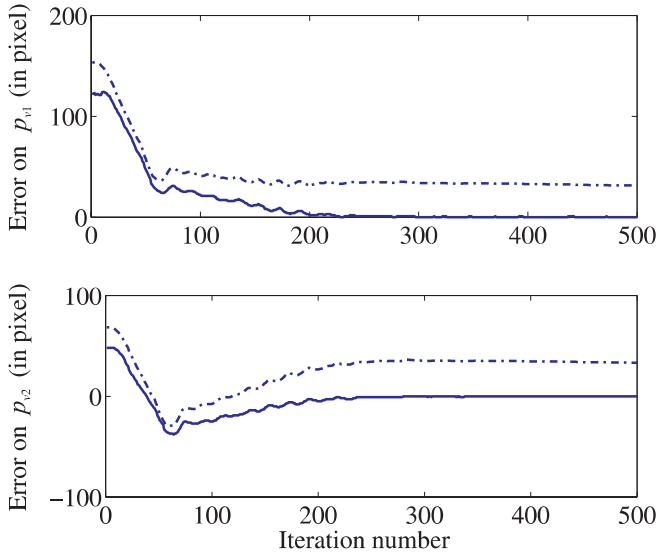
ture is based on a multithread application and the software has been divided into three threads, which allows the exploitation of the hardware capabilities of the computer. A first thread is dedicated to image processing to detect the two image features at a low level on the acquisition board, as described previously.

In a second thread, the optimization procedure using CFSQP (a C library of the Feasible Sequential Quadratic Programming algorithm) computes at each iteration a feasible solution with respect to the constraints. Indeed, this optimization package is a portable standard C source code that is based on SQP programming (as for simulation) but modified to provide feasible iterates that satisfy the constraints. As mentioned before, it has been shown to be an efficient way of solving constrained optimization problems. Details of CFSQP can be read of in Lawrence et al. (1997).

Finally, a third thread handles communication with the PC supporting the robot controller; this thread receives the joint positions and sends the optimal control input. The optimization thread is clearly the most expensive in terms of computation time. Experiments have shown a mean computation time for minimizing the criterion to be close to 30 ms. Computation time is plotted in Figure 17, exhibiting some data over 30 ms. Moreover, the vision loop is limited by the image acquisition frame rate (40 ms) and the time used to perform the blob analysis. Thus, a sampling time of 50 ms has been chosen during open-loop prediction.



(a)



(b)

Fig. 18. Error in the image plane for a complex motion: (a) coordinates along pixel coordinate axis u and (b) coordinates along pixel coordinate axis v (measured: solid line and model prediction: dash dot line).

5.2. Results

5.2.1. Performance

Evaluation of the proposed controller has been achieved through step responses for each image motion (two translations and one rotation of the instrument). In all cases, convergence is satisfied. For brevity, only convergence for a complex

Fig. 19. Image trajectory for a complex motion between initial (circle) and desired (square) positions.

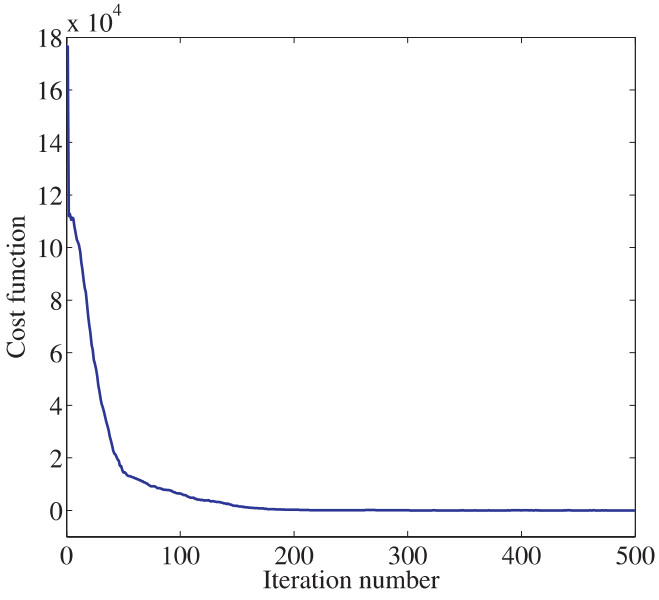


Fig. 20. Evolution of the cost function for a complex motion.

motion is plotted in Figure 18. The desired position is reached even when an erroneous model is used. The corresponding image trajectory is shown in Figure 19. The controller forces the system to make the error vanish by translation motion and then corrects the instrument orientation to ensure image convergence.

One can see that an error remains at the end of the motion around two pixels on coordinates p_u of the two points, whereas the convergence is reached as shown in the cost function evolution (Figure 20). It has to be linked to the authorized displacement with respect to the US plane. In this experiment,

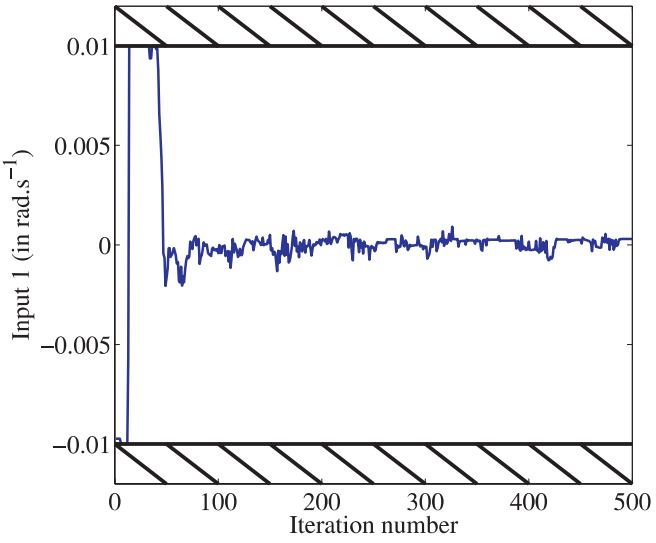


Fig. 21. Input saturation on joint 1.

maximum displacement with respect to this plane is 1.1 mm which can induce the observed error in the image.

5.2.2. *Constraint Satisfaction*

The behavior of the proposed control scheme to satisfy the constraints is illustrated through the respect of the saturation and the intersection constraints. The saturation of the robot actuators is observed and naturally satisfied by the NMPC controller. In Figure 21, behavior of joint 1 is presented in the presence of the saturations listed in Table 1.

The height of the two points is naturally constrained within the interval $\{Z_s^{\min}, Z_s^{\max}\}$ where $Z_s^{\max} = -Z_s^{\min} = 0.0011$ m (Figure 22). Nevertheless, an overshoot is observed. This behavior may be the result of a combination of two effects. On one side, the dynamics of the system are not taken into account in the prediction model, in particular friction force effects. On the other side, inequality constraints are 'soft' constraints in the CFSQP algorithm, meaning that it accepts small constraint violation during the optimal input search. However, this problem may be practically solved by imposing smaller software joint limits. Without this constraint, however, the out of plane motion becomes much larger.

6. Conclusion

In this paper, a Nonlinear Model Predictive Control scheme applied to US-IBVS has been proposed. Using cost function based on errors in the image plane, convergent and stable robot motion has been obtained through nonlinear constraint optimization. The proposed visual servoing scheme ensures the

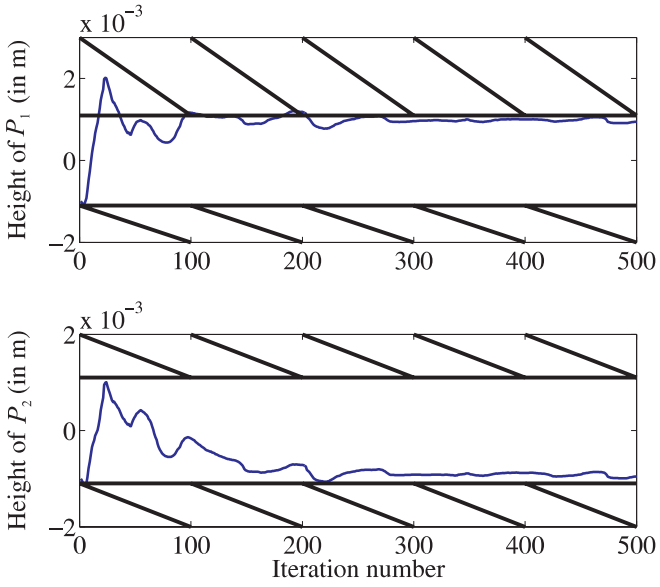


Fig. 22. Constraint satisfaction on z coordinates of points P_1 and P_2 with respect to probe frame R_s.

convergence of the system to the desired set-point. It also exhibits robustness with respect to sensor calibration, outlier detection and noise measurement. The major contribution of this approach is the constraint handling. Inequality constraints are naturally taken into account, which allows us to specify joint limits, actuator saturation and sensor working range.

Experimental validation has been performed on a robotic platform. The results prove the efficiency of the proposed strategy to handle the above-mentioned constraints. Future work concerns enhancement of the robustness of the proposed approach. Evolution of the proposed algorithm may be able to take into account displacement of the US probe, as in real intervention where it is needed to move the probe to observe different structures. The minimization of the computation time will also be considered to increase the dynamics of the system. *In vivo* experiments on animals will allow the evaluation of the effectiveness of the control strategy.

Acknowledgements

This work was supported by the Centre National de la Recherche Scientifique (CNRS) in the framework of the ROBEA project.

References

Abolmaesumi, P., Salcudean, S., Zhu, W., Sirouspour, M., and DiMaio, S. (2001). Imageguided control of a robot for medical ultrasound. *IEEE Transactions on Robotics and Automation,* **18**(1): 11–23.

Allgöwer, F., Badgwell, T. A., Qin, S. J., Rawlings, J. B., and Wright, S. J. (1999). Nonlinear predictive control and moving horizon estimation – an introductory overwiew. In *Advances in Control: Highlights of ECC'99*, pp. 391–449.

Batcha, W. and Krupa, A. (2006). Towards ultrasound image-based visual servoing. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'06)*, pp. 4112–4117.

Chavanon, O., Carrat, L., Pasqualini, C., Dubois, E., Blin, D., and Troccaz, J. (2000). Computer-guided pericardiocentesis: experimental results and clinical perspectives. *Herz,* **25**(8): 761–768.

Diegeler, A., Hirsch, R., Schneider, F., Schilling, L., Falk, V., Rauch, T., and Mohr, F. (2000). Neuromonitoring and neurocognitive outcome in off-pump versus conventional coronary bypass operation. *Annals of Thoracic Surgery,* **69**: 1162–1166.

Downing, S. D., Herzog, W. A., McLaughlin, J. S., and Gilbert, T. P. (2002). Beating-heart mitral valve surgery: preliminary model and methodology. *Journal of Thoracic and Cardiovascular Surgery,* **123**(6): 1141–1146.

Hastenteufel, M., Yang, S., Christoph, C., Vetter, M., Meinzer, H.-P., and Wolf, I. (2006). Image-based guidance for minimally invasive surgical fibrillation ablation. *International Journal of Medical Robotics and Computer Assisted Surgery,* **2**: 60–69.

Hong, J., Dohi, T., Hashizume, M., Konishi, K., and Hata, N. (2004). An ultrasound-driven needle insertion robot for percutaneous cholecystostomy. *Physics in Medicine and Biology,* **49**(3): 441–455.

Hutchinson, S., Hager, G. D., and Corke, P. (1996). A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation,* **12**(5): 651–670.

Lawrence, G., Zhou, J. L., and Tits, A. L. (1997). *User's guide for CFSQP (Version 2.5).*

Megali, G., Tonet, O., Stefanini, C., Boccadoro, M., Papaspyropoulos, V., Angelini, L., and Dario, P. (2001). A computer-assisted robotic ultrasound-guided biopsy system for videoassisted surgery. In *Proceedings of Medical Imaging Computing and Computer-Assisted Intervention (MICCAI'01).*

Phee, L., Xiao, D., Yuen, J., Chan, C. F., Ho, H., Thng, C. H., Cheng, C., and Ng, W. S. (2005). Ultrasound guided robotic system for transperineal biopsy of the prostate. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'05)*, pp. 1327–1332.

Picone, A. L., Lutz, C. J., Finck, C., Carney, D., Gatto, L. A., Paskanik, A., Searles, B., Snyder, K., and Nieman, G. (1999). Multiple sequential insults cause post-pump syndrome. *Annals of Thoracic Surgery,* **67**: 978–985.

Rastan, A. J., Bittner, H. B., Gummert, J. F., Walther, T., Schewick, C. V., Girdauskas, E., and Mohr, F. W. (2005). On-pump beating heart versus off-pump coronary artery

bypass surgery – evidence of pump-induced myocardial injury. *European Journal of Cardio-Thoracic Surgery,* **27**: 1057–1064.

Richalet, J. (1993). *Pratique de la commande prédictive*, Chapter 4, pp. 193–199. Hermès.

Sauvée, M., Poignet, P., and Dombre, E. (2006). Towards nonlinear model predictive control for image based visual servoing. In *Proceedings of IFAC Workshop on Nonlinear Model Predictive Control for Fast System (NMPCFS'06)*, pp. 147–152.

Schneider, O., Troccaz, J., Chavanon, O., and Blin, D. (2000). PADyC: a synergistic robot for cardiac puncturing. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'00)*, Volume 3, San Fransisco, USA, pp. 2883–2888.

Stoll, J., Novotny, P., Howe, R., and Dupont, P. (2006). Real-time 3d ultrasound-based servoing of a surgical instrument. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'06)*, pp. 613–618.

Vitrani, M.-A., Morel, G., and Ortmaier, T. (2005). Automatic guidance of a surgical instrument with ultrasound based visual servoing. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'05)*, pp. 508–513.

Wan, I. Y. P., Arifi, A. A., Wan, S., Yip, J. H. Y., Sihoe, A. D. L., Thung, K. H., Wong, E. M. C., and Yim, A. P. C. (2004). Beating heart revascularization with or without cardiopulmonary bypass: evaluation of inflammatory response in a prospective randomiezd study. *Journal of Thoracic and Cardiovascular Surgery,* **127**(6): 1624–1631.