

Automatic Tracking of an Organ Section with an Ultrasound Probe: Compensation of Respiratory Motion

Caroline Nadeau¹, Alexandre Krupa², and Jacques Gangloff³

¹ Université de Rennes I, IRISA, 35042 Rennes, France

² INRIA Rennes-Bretagne Atlantique, IRISA, 35042 Rennes, France

³ LSIT, UMR 7005 CNRS-Université de Strasbourg I, 67400 Illkirch, France

Abstract. In minimally invasive surgery or needle insertion procedures, the ultrasound imaging can easily and safely be used to visualize the target to reach. However the manual stabilization of the view of this target, which undergoes the physiological motions of the patient, can be a challenge for the surgeon. In this paper, we propose to perform this stabilization with a robotic arm equipped with a 2D ultrasound probe. The six degrees of freedom of the probe are controlled by an image-based approach, where we choose as visual feedback the image intensity. The accuracy of the control law is ensured by the consideration of the periodicity of the physiological motions in a predictive controller. Tracking tasks performed on a realistic abdominal phantom validate the proposed approach and its robustness to deformation is assessed on a gelatin-made deformable phantom.

Keywords: visual servoing, ultrasound, motion compensation

1 Introduction

Among the different medical imaging modalities, ultrasound (US) imaging is particularly attractive for providing intra-operatively real-time images. Indeed, this modality is safe and non-invasive for the patient and the US transducer, cheap and not cumbersome, can be easily used in an operating room. In particular, the US modality can provide to the surgeon a view of an organ or a tumor to reach throughout a minimally invasive surgery or a needle insertion procedure. In such applications, image-guided robotic systems could assist the surgeon insofar as they can follow an organ that undergoes physiological motions and thus automatically stabilize the US image of the target.

Some previous works dealing with US image-based robotic systems have focused on target tracking applications. In [1], the three in-plane motions of an US probe are controlled to automatically center in the US image the cross-section of an artery while an out-of-plane translation of the probe is manually performed. Five features extraction methods based on image similarity measures and contour segmentations are compared to track the artery center. In another work [2], the three translations of a XYZ stage robot equipped with two US probes and a HIFU transducer are controlled to follow a kidney stone while compensating for physiological motions during a lithotripsy procedure. In [3], an approach based on the speckle correlation observed in successive US

images is detailed. However, a region of fully developed speckle has to be segmented and a step of learning of the speckle decorrelation curves is required. Moreover in these different approaches, the visual error minimization is realized with a proportional control law which does not benefit from the model of the disturbance.

On the other hand, physiological motions compensation using predictive controller in the control loop of the robotic system is a promising approach to take advantage of the repetitiveness of these disturbances. It has been developed in previous works where the tracking of natural or artificial landmarks is performed using an endoscopic camera. This visual feedback gives then a measure of the organ motion which is combined with a model of the periodic or quasi-periodic disturbance to predict and anticipate this motion. Many contributions focus on cardiac motion compensation using endoscopic cameras. In [4], Nakamura *et al.* present a tele-operated system to assist surgeons during beating heart surgeries. The slave arm is synchronized with the heart beats thanks to the visual feedback of a high speed camera and provides to the surgeon a stabilized image of the heart. In the same context of beating heart surgery, a motion prediction scheme is developed in [5] to increase the robustness of the detection and the tracking of natural landmarks on the heart surface in a laparoscopic view. More generally, Bebek *et al.* [6] describe improvements in motion canceling by taking into account biological signals (ECG) in the predictive algorithm. Recently, a comparison between various predictive filtering methods has been proposed in [7] to predict the motion of the mitral valve annulus. Under the assumption of a major 1D translational motion of this annulus, the motion compensation of one degree of freedom (dof) based on an extended Kalman filter is validated in situation of high noise, time delay and heart rate variability.

In this paper, we propose to control the six dof of a robotic arm holding a 2D US probe in order to follow a desired abdominal cross section while compensating for the respiratory motion. The considered applications are for instance the assistance for diagnoses or hepatic tumor biopsies where the liver and the tumors mainly undergo the respiratory motion [8]. Other clinic applications, such as prostate cancer brachytherapy have been identified in [3] that could benefit from such robotic image stabilization. To deal with the low quality of the US images, we propose to directly use the intensity of the image as visual feature, which has been successfully applied in camera-based visual servoing [10]. The structure of our paper is as follows: the second section details the principle of the intensity-based approach and the computation of the interaction matrix that links the intensity features variation to the motion of the 2D probe. The section 3 presents the proportional and predictive control laws implemented and compared with simulation results. The results of the tracking task performed on an abdominal and on a deformable phantom are presented in section 4 and included in a video⁴.

2 US Image-based Approach

The US images of the human body present a low quality due to the noise, called speckle, generated by the propagation of the US waves in the soft tissues. Because of this speckle, the processing of the US images is often more complex than with other

⁴ http://www.irisa.fr/lagadic/demo/demo-us-servoing-intensity/intensity_us_servoing.html

imaging modalities, making the extraction of natural landmarks particularly complicated. We propose therefore to consider as visual features \mathbf{s} the intensity values of the pixels of a region of interest (ROI) of the US image:

$$\mathbf{s} = \{I(1, 1), \dots, I(u, v), \dots, I(M, N)\}, \quad (1)$$

where M and N are respectively the width and the height of the ROI and where $I(u, v)$ represents the intensity of the pixel of coordinates (u, v) in the US image.

Given $\mathcal{R}_p(\mathbf{x}_p, \mathbf{y}_p, \mathbf{z}_p)$ the frame attached to the US probe where $(\mathbf{x}_p, \mathbf{y}_p)$ defines the image plane and \mathbf{z}_p corresponds to the elevation axis, the coordinates $\mathbf{x} = (x, y, z)$ of the image features in this frame are such as:

$$(x, y, z) = (s_x(u - u_0), s_y(v - v_0), 0),$$

with (s_x, s_y) the image pixel size and (u_0, v_0) the pixel coordinates of the image center, and $z = 0$ since the considered visual features belong to the US image plane.

The computation of the interaction matrix \mathbf{L}_s that links the variation of these image features to the motion of the probe is based on the constancy of the US wave reflexion by a physical 3D point. Given such a 3D point at the position \mathbf{x} at the time t , which moves to the new position $\mathbf{x} + d\mathbf{x}$ at the time $t + dt$ and considering that the US reflexion is converted in an intensity value in a B-mode US image, the US reflexion conservation yields to the following intensity conservation equation:

$$I(\mathbf{x} + d\mathbf{x}, t + dt) - I(\mathbf{x}, t) = 0. \quad (2)$$

We can expand this equation in the form of a first order Taylor series:

$$\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial z} dz + \frac{\partial I}{\partial t} dt = 0. \quad (3)$$

Then, the time variation of each pixel intensity $I(u, v)$ can be expressed as a function of the corresponding 3D point motion :

$$\dot{I} = -\nabla I \dot{\mathbf{x}}, \quad (4)$$

with $\nabla I = (\nabla I_x \ \nabla I_y \ \nabla I_z)$ the 3D image gradient, which is computed from the current probe image and at least two out-of-plane additional images. In practice, with a conventional 2D US probe mounted on a robotic arm, a small back and forth translational motion along the elevation direction is applied to acquire these additional images.

According to the kinematics fundamental relationship, the velocity of the 3D point $\dot{\mathbf{x}}$ is linked to the probe velocity \mathbf{v}_c through the interaction matrix $L_{\mathbf{x}}$:

$$\dot{\mathbf{x}} = L_{\mathbf{x}} \mathbf{v}_c, \quad L_{\mathbf{x}} = \begin{bmatrix} -1 & 0 & 0 & 0 & -z & y \\ 0 & -1 & 0 & z & 0 & -x \\ 0 & 0 & -1 & -y & x & 0 \end{bmatrix}. \quad (5)$$

The interaction matrix L_I of size 1×6 associated to each visual feature is then defined as $L_I = -\nabla I L_{\mathbf{x}}$ and the complete interaction matrix \mathbf{L}_s , defined as $\dot{\mathbf{s}} = \mathbf{L}_s \mathbf{v}_c$, is built by stacking the $M \times N$ matrices L_I associated to each pixel.

3 Control Strategy

3.1 Classical Control Law

An image-based visual servoing control scheme consists in minimizing the error between a current set of visual features \mathbf{s} and a desired one \mathbf{s}^* . In the case of a number p of visual features exceeding the number m of dof of the controlled system, a combination matrix \mathbf{C} of size $m \times p$ and full rank m is introduced to define the task function to minimize $\mathbf{e}(t) = \mathbf{C}(\mathbf{s}(t) - \mathbf{s}^*)$.

In an eye-in-hand system, the instantaneous velocity applied to the imaging sensor \mathbf{v}_c is computed from the task function and the interaction matrix \mathbf{L}_s . More particularly, the classical visual servoing control law described in [9] is based on an exponential decrease of the task function ($\dot{\mathbf{e}} = -\lambda \mathbf{e}, \lambda > 0$) and presents the best behavior with $\mathbf{C} = \widehat{\mathbf{L}}_s^+$, where $\widehat{\mathbf{L}}_s^+$ is the pseudo-inverse of the matrix estimate $\widehat{\mathbf{L}}_s$, which yields to:

$$\mathbf{v}_c = -\lambda \left(\widehat{\mathbf{L}}_s^+ \widehat{\mathbf{L}}_s \right)^{-1} \left(\widehat{\mathbf{L}}_s^+ (\mathbf{s}(t) - \mathbf{s}^*) \right) = -\lambda \widehat{\mathbf{L}}_s^+ (\mathbf{s}(t) - \mathbf{s}^*) \quad (6)$$

3.2 Predictive Controller

In the previous formulation, the controller does not take advantage of some knowledge on the disturbance model and more particularly of its periodicity. We propose then to implement in the control loop a predictive controller, based on the principle of the unconstrained Generalized Predictive Controller (GPC) [11], in order to anticipate the effect of the disturbance. We consider a system with multiple inputs \mathbf{u} and outputs \mathbf{y} . The former are the six velocity components applied to the probe and the latter correspond to the image features observed in the current image \mathbf{s} . In practice, in order to avoid an excessive number of visual components and to ensure the controllability of the system, we introduce the combination matrix in the system loop to consider an output vector of dimension 6: $\mathbf{y} = \mathbf{C}\mathbf{s} = \widehat{\mathbf{L}}_s^+ \mathbf{s}$. The system is described with an ARIMAX model:

$$\mathbf{A}(z^{-1})\mathbf{y}(k) = \mathbf{B}(z^{-1})\mathbf{u}(k-1) + \frac{\mathbf{P}(z^{-1})}{\Delta(z^{-1})}\mathbf{b}(k), \quad (7)$$

where the polynomials \mathbf{A} and \mathbf{B} represent the theoretical model of the robotic system, approximated by a pure integrator. The noise is modeled by a white noise term $\mathbf{b}(k)$ which is colored by the polynomial matrix \mathbf{P} and made non stationary by the operator $\Delta(z^{-1})$. In the classical GPC, this operator is taken as an integrator:

$$\Delta(z^{-1}) = 1 - z^{-1}.$$

In the case of periodic disturbances with a known fixed period, Gangloff et al [12] proposed to consider a repetitive noise model by modifying this operator as follows:

$$\Delta_R(z^{-1}) = 1 - \alpha z^{-T}, \quad 0 < \alpha \leq 1$$

where α is a forgetting factor making the noise model repetitive with a period T corresponding to the number of sampling periods in one period of the disturbance signal.

To ensure that this repetitive noise model only affects the disturbance rejection and not the response to the visual reference, the new controller called R-GPC is a combination of two GPC (see Fig. 1). The controller GPC_1 is applied to the theoretical model of the system with no noise term which allows the robot to anticipate the future references $\mathbf{r} = \mathbf{C}\mathbf{s}^*$. The second controller GPC_2 includes the repetitive noise model previously defined to cancel the disturbance effect on the output vector.

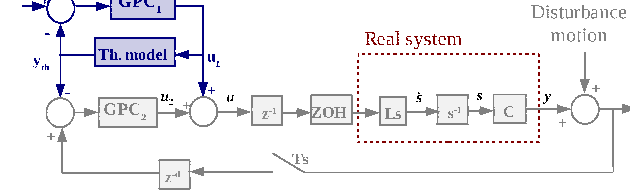


Fig. 1. Bloc diagram of the R-GPC including a control delay of one sampling period and a delay of d sampling periods due to the measure.

The cost function minimized by the R-GPC approach is therefore composed of the cost functions of GPC_1 and GPC_2 as detailed in [12]:

$$J(\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2, k) = \sum_{j=N_1}^{N_2} \|\hat{\mathbf{y}}_{th}(k+j) - \mathbf{r}(k+j)\|^2 + \sum_{j=N_1}^{N_2} \|\hat{\boldsymbol{\varepsilon}}(k+j)\|^2 + \lambda \sum_{j=1}^{N_u} \|\delta \mathbf{u}_1(k+j-1)\|^2 + \mu \sum_{j=1}^{N_u} \|\delta \mathbf{u}_2(k+j-1)\|^2 \quad (8)$$

where N_1, N_2 are the bounds of the prediction horizon and N_u is the length of the control horizon. λ and μ weight the control energies respectively for reference tracking and disturbance rejection. With this structure, the control signal $\mathbf{u}(t)$ applied to the real system is composed of the two components $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$. The former corresponds to the input of the theoretical system model without disturbance that generates the theoretical output \mathbf{y}_{th} , the latter allows the rejection of the error $\boldsymbol{\varepsilon}(t)$ due to noises and disturbance.

4 Results and Discussion

4.1 Simulation Results with a Human Liver

To validate our approach, we use a software simulator that we have developed to reconstruct and display a dense volume from a set of parallel images. In addition to this display functionality, the simulator allows us to move the reconstructed volume wrt a fixed Cartesian frame and to control a 2D virtual probe which generates an US image by cubic interpolation process. For the simulation experiments, an US complete volume of a human liver is loaded in the simulator. This volume is created from a set of 218 parallel images of resolution 308×278 , the voxel size is $0.5 \times 0.5 \times 0.5 \text{mm}^3$.

In abdominal surgery, the hepatic motions are mainly induced by the respiration [8]. The cranio-caudal translation is the most important motion with an established range of 10 to 26mm and additional translations of about 8mm in antero-posterior and lateral directions are observed. In line with these data, a periodic translational motion is applied to the liver volume with amplitudes of 16mm, 5.5mm and 7mm along the x , y and z axes.

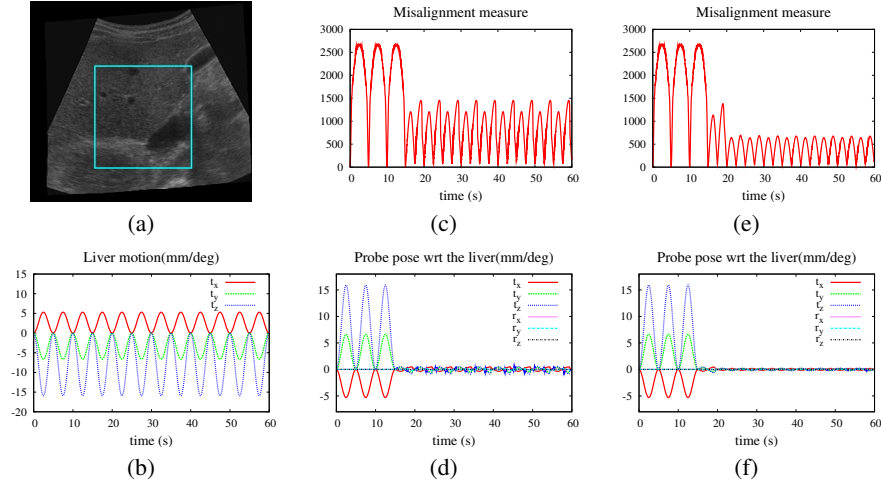


Fig. 2. Results of the tracking of a liver cross-section (a) that undergoes a periodic translational motion (b). For the same task, the results obtained with the proportional control law (c,d) and with the predictive one (e,f) are compared in term of visual error (c,e) and pose error (d,f).

The results of the simulation are displayed in Fig. 2. During three breathing periods no compensation is applied, then at $t = 15s$, the predictive controller R-GPC is launched and compared with the classical control law (6). From the first period the probe follows the liver motion and the misalignment measure defined as $\mathcal{L}(\mathbf{s}) = (\mathbf{s} - \mathbf{s}^*)^\top (\mathbf{s} - \mathbf{s}^*)$ is significantly reduced with both approaches (see Fig. 2(c) and (e)). With the R-GPC, the periodic disturbance is predicted and the accuracy of the tracking is improved.

4.2 Experimental Results with the Predictive Controller

The experiments have been performed with a robotic arm equipped with a convex 2-5 MHz US transducer on an abdominal phantom and a deformable one (see Fig. 3).

For the considered medical applications, we combine the visual control with a force control insofar as the US probe relies on the patient skin. We therefore add to the visual task a force constraint which guarantees a constant force of 1N applied to the phantom. The force control is used to servo the translational motion along the y -axis of the probe frame while the five remaining dof are controlled by the visual servoing control scheme. The Fig. 4 shows the results of one tracking experiment realized with the predictive controller where a sinusoidal perturbation is applied to the abdominal phantom



Fig. 3. Experimental setup with the abdominal (left) and gelatin-made (right) phantoms.

with a period of 6s, mainly along the x and z axes in translation and around the y -axis in rotation. The visual error is reduced from the second period, which guarantees the stabilization of the organ cross section in the US image during the disturbance cycles.

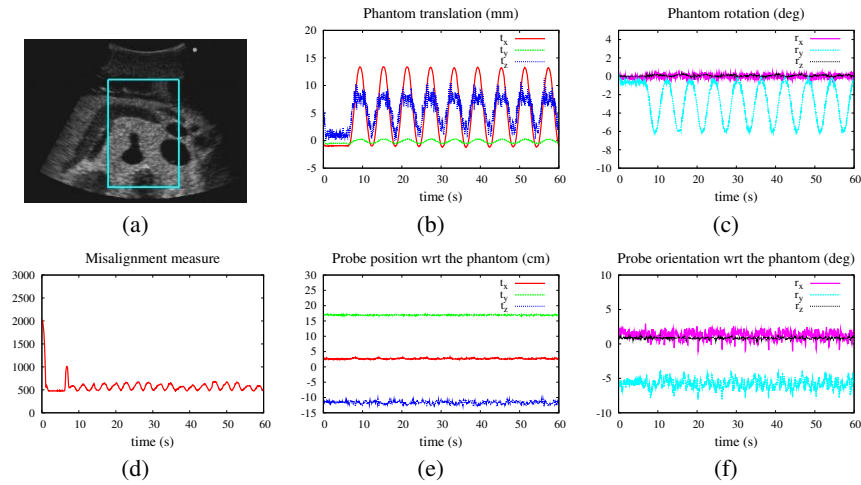


Fig. 4. Tracking of an abdominal cross section (a) while the phantom undergoes a periodic external motion applied at $t = 6$ s in translation (b) and rotation (c). The low value of the visual error (d) and the constant position (e) and orientation (f) of the probe wrt the phantom validate the task.

The robustness of the control law when dealing with deformable objects is assessed using a gelatin-made phantom filled with cereal fibers to create US speckle and olives modeling hepatic tumors (see Fig. 5). At $t = 5$ s, a periodic translational motion of about 2cm is applied to this phantom which is also pressed against a wall to create local deformations. In this particular case, the relative pose of the probe wrt the phantom can not be measured and the success of the task is visually validated by the display of the view of the US probe and the corresponding image difference with the desired image (see Fig. 5(b) and (c)) at $t = 37$ s, which corresponds to the maximum error in the misalignment measure (see Fig. 5(d)) observed during the tracking task.

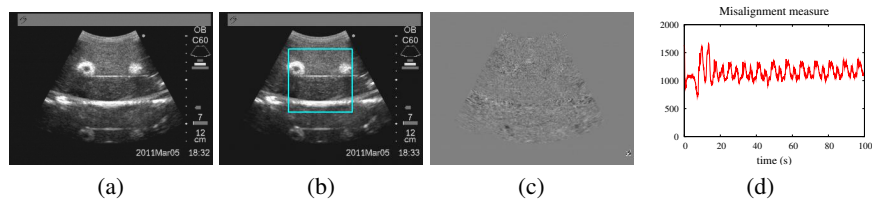


Fig. 5. Tracking results with the deformable phantom of an US slice containing two tumors.

The tracking results validate the intensity-based approach to deal with the stabilization of a 2D US image by compensating periodic disturbances with a predictive controller. The limitation due to the low quality of the US images is overcome by the choice of the image intensity as visual feature and the six dof of the probe are controlled thanks to the modeling of the interaction matrix associated to these features.

Acknowledgment. The authors acknowledge the support of the ANR project US-Comp of the French National Research Agency.

References

1. Abolmaesumi, P., Salcudean, S., Zhu, W., Sirouspour, M., DiMaio, S.: Image-guided control of a robot for medical ultrasound. *IEEE Trans. on Rob.* 18, 11–23 (2002)
2. Lee, D., Koizumi, N., Ota, K., Yoshizawa, S., Ito, A., Kaneko, Y., Matsumoto, Y., Mitsuishi, M.: Ultrasound-based visual servoing system for lithotripsy. In: *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.877–882 (2007)
3. Krupa, A., Fichtinger, G., Hager, G.D.: Real time motion stabilization with B-mode ultrasound using image speckle information and visual servoing. *Int. J. of Rob. Res.* 28, 1334–1354 (2009)
4. Nakamura, Y., Kishi, K., Kawakami, H.: Heartbeat synchronization for robotic cardiac surgery. In: *IEEE Int. Conf. on Robotics and Automation ICRA01*, pp. 2014–2019 (2001)
5. Ortmaier, T., Groger, M., Boehm, D., Falk, V., Hirzinger, G.: Motion estimation in beating heart surgery. *IEEE Trans. on Biomedical Engineering* 52, 1729–1740 (2005)
6. Bebek, O., Cavusoglu, M.C.: Intelligent control algorithms for robotic-assisted beating heart surgery. *IEEE Trans. on Rob.* 23, 468–480 (2007)
7. Yuen, S. G., Kettler, D. T., Novotny, P. M., Plowes, R. D., Howe, R. D.: Robotic motion compensation for beating heart intracardiac surgery. *Int. J. of Rob. Res.* 28, 1355–1372 (2009)
8. Clifford, M. A., Banovac, F., Levy, E., Cleary, K.: Assessment of hepatic motion secondary to respiration for computer assisted interventions. *Computer Aided Surgery* 7, 291–299 (2002)
9. Espiau, B., Chaumette, F., Rives P.: A new approach to visual servoing in robotics. *IEEE Trans. on Rob.* 8, 313–326 (1992)
10. Collewet, C., Marchand, E.: Photometric visual servoing. *IEEE Trans. on Rob.* 27 (2011)
11. Clarke, D. W., Mohtadi, C., Tuffs, P. S.: Generalized predictive control - Part I: The basic algorithm. *Automatica* 23, 137–160 (1987)
12. Gangloff, J., Ginhoux, R., De Mathelin, M., Soler, L., Marescaux, J.: Model predictive control for compensation of cyclic organ motions in teleoperated laparoscopic surgery. *IEEE Trans. on Control System Technology* 14, 235–246 (2006)