2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

High Speed Visual Servoing with Ultrasonic Motors

Andrea Ranftl, Loïc Cuvillon, Jacques Gangloff, and Jos Vander Sloten

Abstract— **Visual servoing refers to the closed-loop position control of the robot end-effector [1] using visual feedback and should be distinguished from vision-based expert systems [2].**

In this work, we focus our interest on high performance visual servoing. Our goal is to maximize the bandwidth of the visual loop. This means usually, considering the average dynamics of robots, that the use of a high speed camera is necessary. To achieve this goal, the model of the visual loop must be known with a good accuracy. This includes the dynamics of the robot and the dynamics of the vision system.

The main objective of this work is to propose a new dynamic model for the vision sensor. The proposed model is validated by experiments. Ultrasonic motors are used in the experiments since they exhibit a very short response time and can be modeled by a simple transfer function, thus simplifying the decoupling between vision and actuation.

I. INTRODUCTION

A. State of the art in high-speed visual servoing

In [2] Corke et al. state that most of the visual servo systems are implemented on top of the existing positioning loop of the robot and thus not using a dedicated control strategy. To this end they initiate the first study of the influence of the dynamics of manipulator and vision system in the visual loop. The distinction between *visual kinematic control* and *visual dynamic control* is introduced. Visual kinematic control refers to the control of the robot path either in the 3D space [3] or in the image space [4] or even both [5]. The main goal of kinematic control is the robustness of the task and not the performance. Usually the dynamics of the manipulator and the effect of sampling are neglected yielding a continuous-time modeling of the visual loop.

With visual dynamic control the main goal is performance. The visual loop has to be fast and optimally tuned with a high order controller. So an accurate dynamic model of the loop is required. In [2] the robot is modeled with an integrator and the vision system is modeled with unit delays. Several control strategies are tested (P, PID, pole placement, and feed forward) as well as several control modes (position, velocity, and torque).

In [6] the vision system is modeled with a delay of two sampling periods. One is due to computation time of the control inputs the second delay models the transfer time of the image. The main contribution of this work is the development of a control strategy which takes the dynamics of the mechanical system into account. The controller is a GPC (Generalized Predictive Controller), which achieves a higher bandwidth than a classical PID controller for high-speed visual servoing.

In [7] Namiki et al. use a vision chip (a visual sensor that is able to perform highly parallel image processing at the pixel level). They achieve a sampling rate of 1000 Hz for the visual loop. They assume that the visual sensor can be modeled by a zero-order hold.

This study is generic. Like what is done at the LSIIT in Strasbourg, we plan to use it to improve the performance of robotized beating heart surgery. Nakamura et al. [8] were the first using high speed vision to track a beating heart. They demonstrated the feasibility of robotic tracking of fast heart motion using high speed vision ($\sim$1000Hz).

In [9] Ginhoux et al. propose a predictive control scheme for beating heart tracking using a 500Hz camera. The visual loop controller is tuned using a linearized dynamic model of the visual loop. In this work the image acquisition process is modeled by delays.

B. Presented approach

In this article we introduce a more accurate dynamic camera model for high-speed visual servoing systems. The theory is validated by experiments on two simple test-beds using ultrasonic actuators. With this model the heart-beat compensation strategy will be enhanced to obtain higher performance and stability. As in [9] and [10], the focus is on the control and we use standard techniques to do image-based visual servoing [1].

This article is organized as follows. In Section II the proposed camera model is explained and the theory is validated by experiments with the help of an ultrasonic device in Section III. The camera model is used in Section IV to enhance the performance of high-speed visual servoing systems using deadbeat and RST control. Conclusions and future perspectives are presented in Section V.

II. THE CAMERA MODEL

The goal is to obtain an accurate dynamic model of the camera. The camera is often modeled by a delay [6], [11]. This affects only the phase. In this work we want to go further in the modeling by also taking into account the gain introduced by the camera.

Indeed, the camera is not a perfect sampler in the visual loop since usually the exposure time defined by the electronic shutter is equal to the sampling period and thus cannot be neglected. The effect of the exposure time on the transfer

This work was supported by the Research Found Belgium

A. Ranftl and J. Vander Sloten are with the Department of Mechanical Engineering, Katholieke Universiteit Leuven, Belgium, {Andrea.Ranftl,Jos.Vandersloten}@mech.kuleuven.be

L. Cuvillon and J. Gangloff are with the LSIIT (UMR CNRS 7005), Strasbourg I University, France, {loic,jacques}@eavr.u-strasbg.fr

function of the camera is of course a delay but also an attenuation when the object in the image is moving fast.

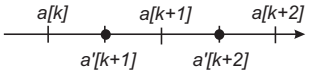

Fig. 1. Unidimensional case of the temporal sampling. The feature moves in one sampling period from position $a[k]$ to $a[k+1]$ and then from $a[k+1]$ to $a[k+2]$. The extracted feature positions are resp. $a'[k+1]$ and $a'[k+2]$ in the image.

Let us consider the unidimensional case of Fig. 1. Here we neglect the effect of spatial sampling due to the pixels. Let $a[k]$ be the position of a feature at time kT, T being the sampling period of the camera and also the duration of the exposition.

The image which is available at time $(k+1)T$ has an exposure beginning at time kT and ending at time $(k+1)T$. During this period, the feature moves from position $a[k]$ to position $a[k+1]$ yielding a trace in the image. If we suppose that the vision algorithm takes the center of mass of the blurred feature, then, the feature extracted from the $(k+1)$th image is located at position $a'[k+1]$ that is $a'[k+1] = \frac{a[k]+a[k+1]}{2}$.

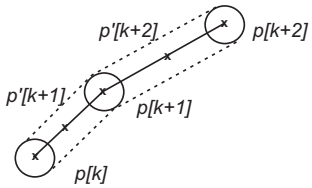

Fig. 2. Two-dimensional case of the temporal sampling. The real position of the feature is denoted by p and the extracted position in the image is denoted by p'.

Fig. 2 shows that it is possible to generalize the unidimensional case to 2D if we consider that the motion during the exposition time can be approximated by a line portion. The feature moves between two positions $p[k]$ and $p[k+1]$ during two time samples. In Fig. 2, the envelope of the moving feature is denoted by the dotted lines. The extracted feature at time $(k+1)T$ is located at the center of mass of the distorted moving feature, so at position $p'[k+1] = \frac{p[k]+p[k+1]}{2}$.

We see here that the camera acts like an averaging filter. So its discrete transfer function is:

$$S(z) = \frac{1+z^{-1}}{2} \tag{1}$$

In what follows, we will try to validate experimentally this model.

III. EXPERIMENTAL VALIDATIONS

A. Experimental setup

The experimental setup is described in Fig. 3. To validate the model in equation (1) we must identify the transfer function of the visual sensor. We use here a frequency approach. An optical marker moves in the field of view of the sensor making a sinusoidal wave with respect to time. The marker is attached to an arm actuated by an ultrasonic motor. We use an ultrasonic motor because of its high bandwidth. Indeed, the effects that we want to identify are high frequency effects. In other words, we would like to be fast enough to see the effect of the visual sampling, *i.e.* the blur induced by the moving marker.

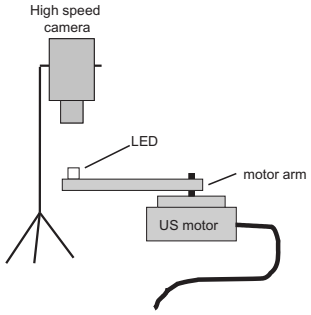

Fig. 3. Experimental setup for the identification experiment.

Ultrasonic motors are a subgroup of piezoelectric and electrostrictive actuators [12], [13]. Their functionality is based on the quality of piezoelectric materials to change their shape if an electric field is applied. By applying an AC electric field at a defined frequency (usually higher than 20kHz, thus the name "ultrasonic"), it is possible to generate a rotating deformation on the surface of the piezoelectric material. This deformation drives the rotor of the motor by friction.

Ultrasonic motors show several advantages which make them a promising technique for positioning tasks and suite them to the presented research. These actuators have a very good torque/mass ratio. They have a very short response time (in the range of milliseconds) far shorter than electromagnetic motors. They do not exhibit complex dynamic modes like oscillations, delays, backlash or non minimum phase behavior. Usually they can be modeled by a first order system. This makes them very suitable for identification purposes. The main drawback of ultrasonic motors is the nonlinear relationship between the control signal and the velocity of the shaft: often there is a minimum speed for the motor yielding a dead-zone in the control.

In this work, we use a simple integrator to model the motor as shown in equation (2). The discrete transfer function $G(z)$ between the control signal u sent to a digital to analog converter and the angular position of the motor shaft x^m_{enc} measured by an encoder is the result of the linear to discrete transformation of an integrator combined with a zero-order hold for a sampling period T.

$$\begin{aligned} G(z) &= (1-z^{-1})\mathcal{Z}\left\{\frac{K_M}{s^2}\right\} \\ &= \frac{TK_M}{z-1} = TK_M\frac{z^{-1}}{1-z^{-1}} \end{aligned} \tag{2}$$

where K_M is a gain and $\mathcal{Z}$ represents the Z-transform.

The rotational ultrasonic piezo USR30 from Shinsei Cooperation Inc. (Japan) is used for the identification of the camera model. It is equipped with an encoder from Faul-

haber, which has a resolution of 500 pulses per revolution (see Fig. 4). The vision sensor is a high speed CCD camera

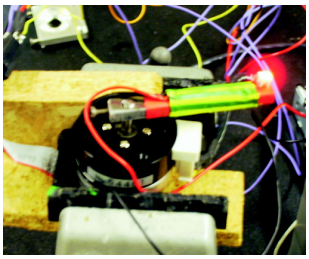

Fig. 4. Picture of the rotational ultrasonic motor and the optical marker.

DALSA^{TM}(Waterloo, ON, Canada), CAD6. It can acquire at a maximum of 955 frames per seconds with a resolution of 256 $\times$ 256 pixels. A dedicated real-time driver was developed for the frame grabber (Coreco-Imaging, PCDIG) that allows to synchronize accurately the acquisition process with the frame rate and so to avoid jitter that could yield frame drops. The driver allows also to control the frame rate by sending to the camera an external synchronization signal through the PCDIG board. The optical marker is a high intensity LED rigidly affixed to the motor arm which can be easily detected by a simple image processing algorithm. It uses the approach described above consisting of computing the center of mass of all illuminated pixels in an area of interest. To ensure a stable frequency, the controller computer runs with a real-time operating system called RTAI. This controller is in charge of image acquisition, image processing and control signal computation. The overall system is synchronized with the frequency of the image acquisition.

B. Experiments

To analyze the dynamics of a system it is in general necessary to observe the output of the system with respect to the input. Let $H(z)$ be the transfer function of the camera that we want to identify and $G(z)$, the transfer function of the piezo actuator. As shown in Fig. 5, the motor position is controlled with a simple proportional control loop using the feedback x^m_{enc}, the measurement provided by the encoder. This control loop allows to reject the effect of the dead-zone in the motor's drive. We use a sinusoidal input signal for the identification, so the reference x^d_{enc} of the position loop is a sinusoidal signal at different frequencies. Thanks to the position loop, x^m_{enc} is almost a perfect sinusoid. By recording simultaneously x^m_{enc} from the encoders and x^m_{cam} from the camera at different frequencies, it is possible to plot the Bode diagram of $H(z)$.

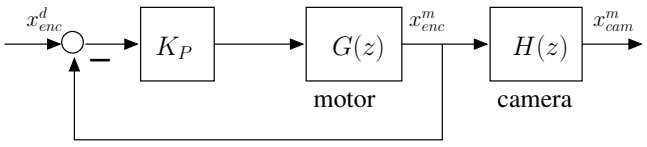

Fig. 5. Closed-loop identification.

C. Results and Discussion

Data series are measured for two different camera frequencies. At 100 images/second, the ultrasonic motor is controlled to follow a sinusoidal motion with a frequency between 1Hz and 20 Hz. At 200 images/s, the sinusoidal excitation is between 2 Hz and 40 Hz.

On the upper graph in Fig. 6 the gain magnitude is plotted in dB. The two solid lines represent the theoretical averaging filter with camera sampling frequencies of 100 Hz and 200Hz. The dotted lines represent the measured values (for each measurement, a $\times$ is plotted). The curve is then interpolated. The same holds for the lower graphic which represents the phase of the theoretical *vs.* the experimental system for the two camera frame rates. Of course these Bode plots are only drawn up to the Nyquist frequency $\frac{\pi}{T}$.

For both gain and phase, the measured results show the same behavior than the theoretical frequency response of the averaging filter. It should be mentioned that the phase measurements have been compensated for internal delays in the system. These include image processing, image transfer as well as the Direct Memory Access (DMA) transfer time from the frame grabber's local memory to the host computer's memory. Altogether this costs around 1 ms delay. For 200Hz the desired plot of the theoretical averaging filter and the measurement results match in a very satisfactory way. For 100Hz this matches for lower frequencies very nicely, while for higher frequencies measurement errors appeared. But still the results show the same behavior as the averaging filter.

These plots show also that the effect of the motion blur on the gain is significant for a standard tuning. Indeed, there is a practical law that specifies the sampling frequency of the control loop with respect to the bandwidth of the plant, *i.e.*, the sampling frequency should be at least 5 times greater than the highest frequency of the system. So at a sampling frequency of 100 Hz, the corresponding maximal frequency component of the plant would be 20 Hz. According to the diagram, a 20 Hz (125.66 rad/s) frequency would at least undertake a 25dB attenuation which is of course significant. This shows that the effect of the motion has to be taken into account when the visual servo loop deals with the highest dynamic modes of the system.

With this experiments it is shown that the averaging filter can model the dynamic behavior of a vision sensor. In the following Section IV this result is used to develop a high performing control strategy for a high-speed visual servoing loop using ultrasonic actuation. The necessity of using the camera model in this plant is shown.

IV. DEVELOPMENT OF A CONTROL STRATEGY

A. Experimental set-up

Since we have now a more accurate model of the camera, we would like to validate this model by using it to optimally tune a high-speed visual loop using ultrasonic actuation.

We use the same high speed DALSA camera as in Section II for the high-speed visual feedback. For the actuation, we use the rotational Shinsei USR30 that was used

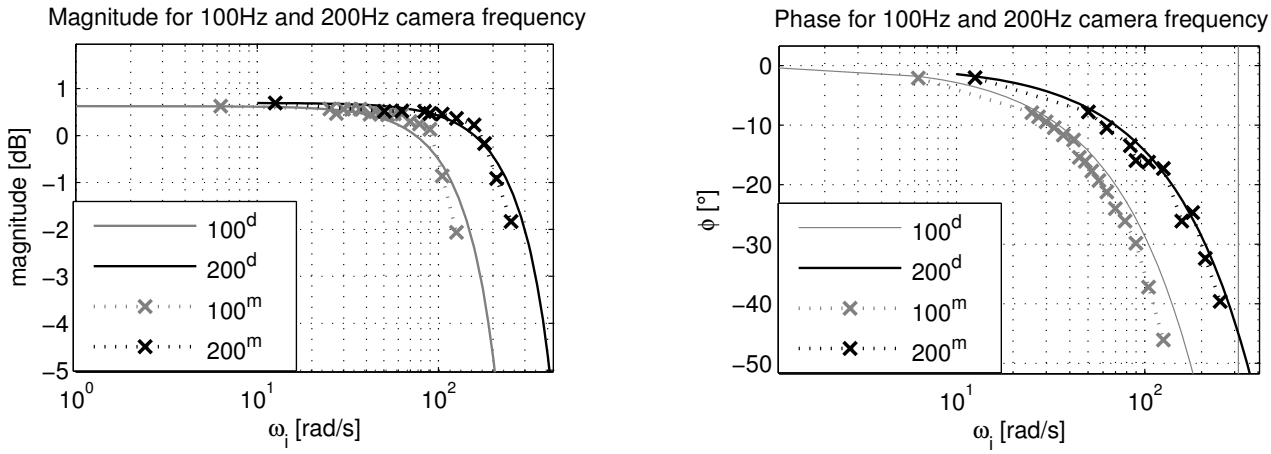

Fig. 6. Resulting Bode diagram for the experiments with the rotational ultrasonic motor. The solid lines are the ideal lines of the averaging filter for the two camera frequencies (100^d corresponds to the desired values for a camera frequency of 100Hz) and the dotted lines are the experimental values (100^m corresponds to the measured values for a camera frequency of 100Hz).

for identification and also a linear ultrasonic motor, the M-661 linear piezo motor stage from Physik Instrumente GmbH&Co.KG [14] with a travel range of 20mm (see Fig. 7). Though it shows a nicer dynamic behavior than the rotational ultrasonic actuator it was not possible to use it for the identification experiment due to the lack of an encoder. The ultrasonic motors are controlled by the PC running RTAI

Fig. 7. Piezo linear motor.

and the visual servo loop is synchronized with the image acquisition. So, in this case, the feedback value is not the encoder value but the position of the LED in the image. It is an image-based visual servoing since the reference, given as a position of the LED in the image, is compared with the measured pixel position. The interaction matrix, which is very simple in this case, is identified in a preliminary procedure.

B. Control Strategy

Our goal is to show that taking into account the model of the camera in the tuning of the controller yields better performance for the visual loop. We plan to use this approach in the future for a system that can track the beating heart in beating heart surgery. Of course, the dynamics of the manipulator have also to be taken into account. As already mentioned, the actuator dynamics are in this case very simple because we use ultrasonic actuators. So we use the model given by equation (2). In the next two sections, we try two control strategies for the high speed visual loop: deadbeat control and RST control. Both need an accurate model of the plant to be efficient. So we can validate in this way the proposed model.

1) Deadbeat Control: A deadbeat controller has a finite settling time. This means that, within a finite number of samples, the control signal $U[k]$ stays constant, the error $E[k]$ reaches zero and the output reaches the steady-state. Furthermore, the feedback system exhibits the smallest response time which is possible. The control scheme of the

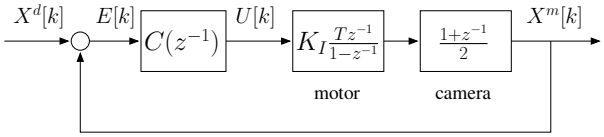

Fig. 8. Control scheme of a deadbeat controller.

deadbeat control loop is presented in Fig. 8. The gain K_I is identified in a preliminary phase. It accounts for the ratio between the velocity control signal $U[k]$ sent to the power amplifier and the position $X^m[k]$ measured in the image. K_I is the simplest form of the well-known interaction matrix in 2D visual servoing. The design of $C(z^{-1})$ is based on pole-zero compensation of the system consisting of motor dynamics and camera model.

For the system in Fig. 8 we designed the controller to track a first-order reference (a ramp). This has the effect to add an integrator in the controller so that nonlinearities like a dead-zone can be rejected. The obtained controller is given by equation (3).

$$U[k] = -\frac{1}{4}U[k-1] + \frac{3}{4}U[k-2] + \frac{1}{2K}(5E[k] - 3E[k-1]). \quad (3)$$

To assess the need for an accurate model of the camera, we designed another deadbeat controller (equation (4)) without using the averaging filter (the model of the camera was

simply replaced by an unitary gain).

$$U[k] = U[k-1] + \frac{1}{K}(2E[k] - E[k-1]) \quad (4)$$

2) RST Control: The RST-controller gains its names from the three polynomials which are shown in Fig. 9. R is placed in the feedback loop, S is directly acting on the plant input, and T acts as a prefilter. These are all polynomials in z^{-1}. As the above described deadbeat controller, the design is based on pole-zero compensation of the system poles and zeros, which are stable. The main difference with the deadbeat controller is the possibility to influence the response time by defining a target closed-loop transfer function. So it is possible to specify a shape for the response which results in a smoother control signal. Here we use a second order target transfer function with natural frequency ω_n and damping ξ. The higher the product $\omega_n T$ is set, the slower the feedback

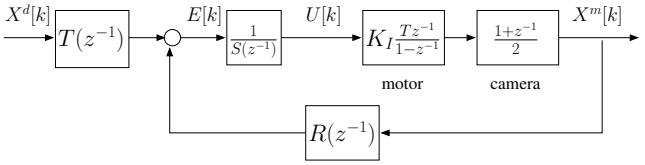

Fig. 9. RST controller. Thereby R, S, and T are polynomials in z^{-1}.

system reacts on changes of the reference input. Furthermore, the damping ratio ξ is set to a fixed value of $\frac{1}{2}\sqrt{2}$.

In addition it should be mentioned that the deadbeat controller is a special case of the RST-controller which has the shortest possible response time, thus usually yielding more powerful control signals.

C. Results

1) Rotational ultrasonic motor: In these experiments, we use the deadbeat controller with integrator (3).

The drawback of the integrator is to yield a limit cycle: the output is oscillating around the desired value without stabilizing. To avoid this effect, we added a threshold on the integrator: we integrate the error only if its absolute value raises over a given threshold.

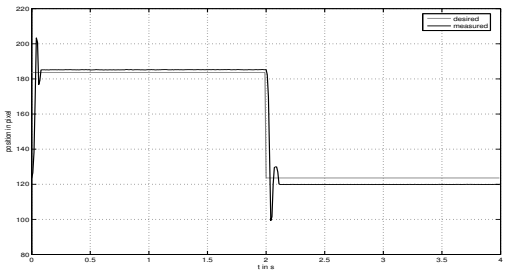

Fig. 10. Deadbeat control of the rotational piezo at 100Hz.

Fig. 10 shows the step response of the visual loop at 100Hz with the deadbeat controller. The output stabilizes in less than 0.1s. There is a little steady-state error due to the threshold on the integrator as explained above. Fig. 11 shows the step response of the visual loop at 100Hz with the deadbeat controller of equation (4), *i.e.* using a simplified camera model. Since the response is totally unstable, this clearly demonstrate the need for an accurate model.

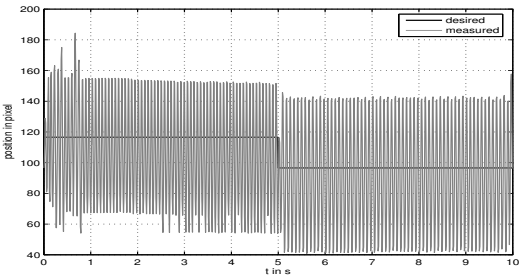

Fig. 11. Deadbeat control of the rotational piezo at 100Hz without camera model.

We were able to tune the RST controller so that the limit cycle vanishes for $\omega_n T = 10$. Indeed, with the RST approach, it is possible to slow down the closed loop up to a point where there is no more oscillation. The resulting step response for a visual servoing running at 500Hz is given in Fig. 12. There are some high-frequency oscillations due to the control dead-zone but the response reaches its steady-state after approximately 0.1s. The deadbeat control is not

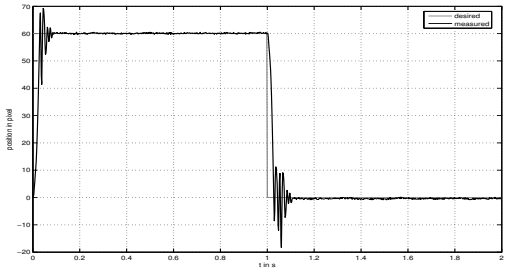

Fig. 12. RST-control of the rotational piezo at 500Hz with additional integrator and using a natural period $\omega_n T = 10$.

possible at 500Hz. Indeed, the response time of the deadbeat controller decreases with the sampling period T. So, due to saturation issues, the 500Hz deadbeat controller is unstable with the rotational motor.

By using the simplified camera model with the RST controller, we obtained similar results than with deadbeat control *i.e.* a totally unstable behavior.

2) Linear ultrasonic motor: In Fig. 13 to 14 we present step responses of visual loops using a linear piezo actuator. This actuator does not have a dead-zone in the control signal: it is almost linear. So we were able to successfully use deadbeat and RST controllers without integrator with almost no steady-state error.

The deadbeat controller as well as the RST-controller are tested at a sampling frequency of 500Hz. As for the rotational piezo, the actuator must follow two steps which are predefined in the image.

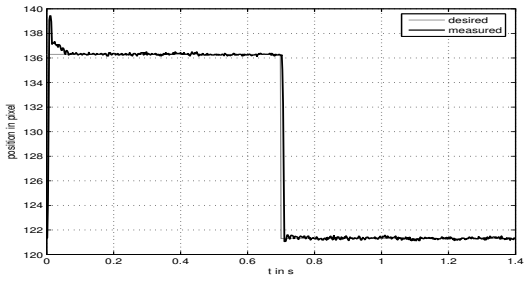

Fig. 13. Deadbeat control of the linear piezo at 500Hz.

In Fig. 13, during the first step, the trajectory of the translator shows an overshoot while during the second step, the response is almost perfectly flat. We can observe this in all the plots of the linear motor. We suspect that the model of this motor is not symmetrical: a different gain in the two directions could yield such a behavior.

In Fig. 13 the maximum settling time is less than 0.1s and if we consider only the second step, it is less than 0.01s ! This gives an idea on the performance that can be achieved by such a visual servo loop.

Fig. 14 presents similar results for an RST-controller at 500Hz. The natural period is set to $\omega_n T = 10$ and there is no integrator. These plots are very similar to the responses obtained with deadbeat control.

When using the simplified model of the camera for tuning the deadbeat or RST controller, we obtained again a totally unstable behavior of the visual loop. This proves that the proposed model is useful to achieve the best performance.

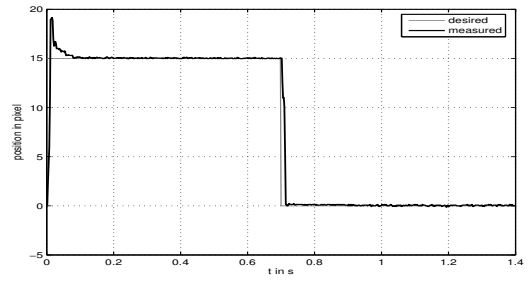

Fig. 14. RST-control of the linear piezo at 500Hz and a natural period $\omega_n T = 10$.

V. CONCLUSION

This work shows that it is possible to obtain very high performances with a visual servo loop if all the dynamic effects are taken into account, including the model of the camera.

We propose a dynamic model of the camera that accounts for the sampling action of this sensor in the visual loop.

This model is validated by experiments. The experimental Bode plot of the camera fits pretty well the theoretical expectations. Furthermore, we show that a visual servo loop tuned using the accurate model of the camera behaves far better than the same loop tuned using a simplified model.

In this work, we use piezo actuators because of their high dynamic performances. We had to tackle with some drawbacks of these actuators, like nonlinearities in the control signal. In future work, we plan to use these techniques to improve existing work on robotized beating heart surgery.

VI. ACKNOWLEDGMENTS

REFERENCES

[1] S. A. Hutchinson, G. D. Hager, and P. I. Corke, "A Tutorial on Visual Servo Control," *IEEE Trans. Robot. Automat.*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[2] P. I. Corke and M. C. Good, "Dynamic effects in visual closed-loop systems," *IEEE Trans. Robot. Automat.*, vol. 12, no. 5, pp. 671–683, Oct. 1996.

[3] W. J. Wilson, C. C. W. Hulls, and G. Bell, "Relative end-effector control using cartesian position based visual servoing," *IEEE Trans. Robot. Automat.*, vol. 12, no. 5, pp. 684–696, Oct. 1996.

[4] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," *IEEE Trans. Robot. Automat.*, vol. 8, no. 3, pp. 313–326, June 1992.

[5] E. Malis and S. Benhimane, "A unified approach to visual tracking and servoing," *Robotics and Autonomous Systems*, vol. 52, no. 1, pp. 39–52, 2005.

[6] J. A. Gangloff and M. F. de Mathelin, "High-speed visual servoing of a 6-d.o.f. manipulator using multivariable predictive control," *Advanced Robotics*, vol. 17, no. 10, pp. 993–1021, 2003.

[7] A. Namiki, Y. Nakabo, I. Ishii, and M. Ishikawa, "1-ms sensory-motor fusion system," *IEEE/ASME Trans. Mechatron.*, vol. 5, no. 3, pp. 244–252, Sept. 2000.

[8] Y. Nakamura, K. Kishi, and H. Kawakami, "Heartbeat synchronization for robotic cardiac surgery," in *Proc. 2001 IEEE Int. Conf. Robot. & Auomat.*, May 2001, pp. 2014–2019.

[9] R. Ginhoux, J. Gangloff, M. de Mathelin, L. Soler, M. M. A. Sanchez, and J. Marescaux, "Active filtering of physiological motion in robotized surgery using predictive control," *IEEE Transactions on Robotics*, vol. 21, no. 1, pp. 67–79, Feb. 2005.

[10] L. Cuvillon, J. Gangloff, M. de Mathelin, and A. Forgione, "Towards robotized beating heart TECABG: Assessment of the heart dynamics using high-speed vision," in *Proceedings of the 8th conference of Medical Image Computing and Computer-Assisted Intervention*, ser. Lecture Notes in Computer Science. Springer, Oct. 2005, pp. 551–558.

[11] N. P. Papanikolopoulos, P. K. Khosla, and T. Kanade, "Visual tracking of a moving target by a camera mounted on a robot:a combination of control and vision," *IEEE Trans. Robot. Automat.*, vol. 9, no. 1, pp. 14–35, 1993.

[12] K. Uchino, "Piezoelectric ultrasonic motors: overview," *Smart Materials and Structures*, vol. 7, no. 3, pp. 273–285, June 1998.

[13] C. Niezrecki, D. Brei, S. Balakrishnan, and A. Moskalik, "Piezoelectric Actuation: State of the Art," *The Shock and Vibration Digest*, vol. 33, no. 4, pp. 269–280, July 2001.

[14] *Technical Documentation PI MP56E User Manual M-661 / M-662 PILine Linear Piezo Motor Stages*, 1st ed., Physik Instrumente (PI) Gmbh & Co. KG, Karlsruhe, Germany, Oct. 2003.