Proceedings of the American Control Conference
Chicago, Illinois • June 2000

Robust Controller Design for Active Vision Systems[1]

Mario Sznaier Tamer Inanc Octavia Camps
Department of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802
email: {msznaier,tinanc}@gandalf.ee.psu.edu, camps@whale.ee.psu.edu

Abstract

Recent hardware developments have rendered controlled active vision a viable option for a broad range of problems, spanning applications as diverse as Intelligent Vehicle Highway Systems, MEMS microassembly and assisting individuals with disabilities. However, realizing this potential requires having a framework for synthesizing robust active vision systems, capable of moving beyond carefully controlled environments. In this paper we show how recently developed robust identification and control synthesis techniques can be brought to bear on the problem. These results are experimentally validated using a Bisight robotic head.

1 Introduction

Visual servoing systems appeared as far back as late 1970's [8]. An excellent survey of the state–of–the art as of 1996 can be found in [9]. Earlier systems dealt with stability issues, at the expense of performance, by experimentally detuning the controller. Latter approaches combine PID controllers with some prediction to explicitly address the time delay required to find the target in the image[4, 2]. However, these predictors can tolerate only small amounts of model uncertainty. Moreover, the combination PID controller/predictor must be tuned using a potentially lengthy trial and error process. Performance can be improved by using a two–degrees of freedom controller [3], but this approach does not improve robustness.

Alternatively, the use of LQG controllers has been proposed [11, 7, 6]. Although LQG controllers have the potential to minimize the effect of noise [6], it has been experimentally corroborated [11] (using a slightly defocused camera) that the resulting systems are potentially fragile. This difficulty can be overcome, to a certain extent, by using a self–tuning controller [10]. However, this approach does not provide a–priori stability guarantees, nor does it allow for trading–off robustness against performance.

The issue of rendering the closed–loop system insensitive to calibration errors has been recently addressed in [5] using time–varying proportional feedback. While this approach has been successfully used for robot motion control [5], the gain matrices must be empirically tuned to achieve good performance. In addition, implementing the control action requires computing the image Jacobian.

It has been long recognized that dynamic control effects coupled with the presence of uncertainty are the factors limiting performance in visual closed–loop systems. Indeed, the motivation behind the approaches mentioned above is to obtain a compromise between these factors leading to acceptable performance. However, only very recently there have been attempts to explicitly address some of these tradeoffs in a systematic way [14].

In this paper we show how recently developed robust identification and control synthesis techniques can be used to synthesize robust active vision systems capable of delivering good performance for a range of conditions. These techniques eliminate the need for tuning controller parameters a posteriori, by trial and error, and for calibrating the system.

2 Preliminaries

2.1 Notation

$\mathcal{L}_\infty$ denotes the Lebesgue space of complex valued matrix functions essentially bounded on the unit circle, with the norm: $\|G(z)\|_\infty \doteq ess\sup_{|z|=1} \overline{\sigma}(G(z))$ where $\overline{\sigma}$ denotes the largest singular value. $\mathcal{H}_\infty$ denotes the subspace of functions in $\mathcal{L}_\infty$ with a bounded analytic continuation inside the unit disk, with the norm $\|G(z)\|_\infty \doteq ess\sup_{|z|<1} \overline{\sigma}(G(z))$. Finally, $\mathcal{H}_{\infty,\rho}$ denotes the Banach space of transfer matrices in $\mathcal{H}_\infty$ which have analytic continuation inside the disk of radius $\rho > 1$, with the norm $\|G(z)\|_{\infty,\rho} \doteq \sup_{|z|<\rho} \overline{\sigma}(G(z))$.

[1] This work was partially supported by NSF grant IIS–9911161

2.2 Statement of the Problem

The control–related issues involved in active vision can be illustrated by considering the problem of smooth tracking of a non–cooperative target, illustrated in the block diagram shown in Figure 1. Here the goal is to internally stabilize the plant and to track the reference signal y_{ref}, using as measurements images possibly corrupted by noise.

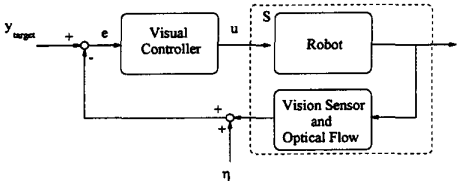


Figure 1: Block diagram of a visual tracking system.

It can be shown [11] that a simplified model of the plant is given by the following state–space model:

$$\begin{aligned} \underline{x}(k+1) &= A_f \underline{x}(k) + B_f \underline{u}(k) + E_f \underline{v}(k) \\ \underline{y}(k) &= \underline{x}(k) + \underline{\eta}(k) \end{aligned} \tag{1}$$

where the state $\underline{x} \doteq (x, y)^T \in R^2$ is the position of the feature in the image plane, $\underline{v} \doteq (v_x, v_y)^T \in R^2$ is the unknown (but bounded) target velocity, $\underline{u} \doteq (\theta_x, \theta_y)^T \in R^2$ is the control input (pan and tilt motion of the camera), $\underline{y}$ is the measured position of the feature corrupted by the noise $\underline{\eta}$, and where the matrices are given by:

$$A_f = I_2, E_f = \frac{\tau f}{Z_s(k)} I_2, B_f = \tau \begin{bmatrix} \frac{x(k)y(k)}{f} & -(f + \frac{x^2(k)}{f}) \\ (f + \frac{y^2(k)}{f}) & -\frac{x(k)y(k)}{f} \end{bmatrix} \tag{2}$$

where f is the focal distance of the camera, Z_s its distance to the object, and τ the sampling interval. LTI controllers can be synthesized by linearizing equation (1) around the present position [11, 3, 7, 10]. However, experimental results show that the performance of the resulting system can be far worse than expected based on simulations using the model (1). This is largely due to the fact that this simple model does take into account neither the time–delay required by the image processing algorithms to locate the object in each frame, nor potentially destabilizing modelling errors such as variations in the optical parameters and in the depth Z_s, and unmodelled head dynamics [15].

3 Plant Modelling

3.1 Control Oriented Identification

As pointed out above, the simple model (1) is often inadequate to synthesize controllers that achieve acceptable performance. A more detailed model could be obtained by incorporating additional factors such as a model of the dynamics of the robotic head. However, the resulting model still depends on several parameters, including the optical parameters of the cameras and the mass of the pan/tilt unit. Thus, even if a more detailed model were to be used, an identification step is still required to obtain the values of these parameters.

We avoid these difficulties by using recently developed non–parametric robust identification techniques. A good tutorial to the field and relevant references can be found in [13], Chapter 10.

The identification technique starts from experimental data and some mild *a priori* assumptions on the plant to generate a nominal model as well as bounds on the worst case identification error suitable to be used by robust control synthesis methods. The experimental data used in this paper consists of the first N_t samples, corrupted by additive noise, of the time response of the system $y(k) = h(k) + \eta_t(k)$, $k = 0, \ldots, N_t - 1$ corresponding to a known input $u(k)$. The *a priori* assumptions are: 1) The system to be identified belongs to the class $\mathcal{H}_\infty(\rho, K) \stackrel{\text{def}}{=} \{H \in \mathcal{H}_{\infty,\rho} : \|H\|_{\infty,\rho} \leq K\}$; and 2) $\sup_k |\eta_t(k)| \leq \epsilon_t$. With these assumptions it can be shown [12] that the problem of identifying a model that interpolates the data points within the experimental errors reduces to a LMI feasibility problem.

3.2 Identification results

To identify the transfer function from the command input u to the pan/tilt unit to the displacement y of the target in the image (in pixels) the system was excited with a step input of amplitude 110 encoder units (roughly an angular displacement of 2.5°) and the position of a target (originally located at the center of the image) was measured. By repeatedly measuring the location of the centroid of the target in the absence of input, the experimental noise measurement was determined to be $\epsilon_t = 4$pixels[1]. Finally, by measuring the time–constants of the pan and tilt unit, ρ was estimated to be around 2. The identification algorithm proposed in [12] using $N_t = 26$ samples, followed by a model reduction step, yielded the following discrete model:

$$G_{nom} = \frac{0.016z^5 + 0.09z^4 + 0.19z^3 + 0.08z^2 + 0.045z + 0.08}{z^5 + 0.39z^4 + 0.84z^3 + 0.38z^2 + 0.29z + 0.08} \tag{3}$$

Figure 2 compares the step response of this model against the experimental data points (normalized by the input). Here 'o' denotes an experimental data point used in the identification, while '$*$' denotes additional experimental data, plotted for validation purposes. It

[1] This experimental error is mainly due to fluctuating conditions such as ambient light.

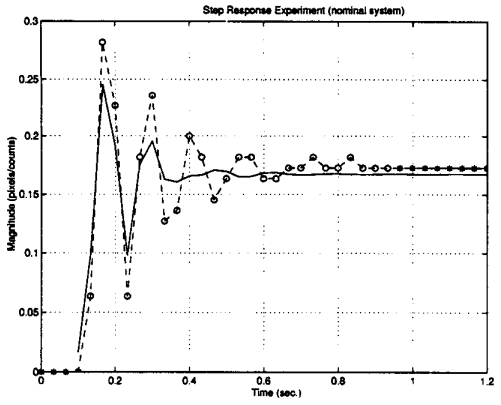


Figure 2: G_{nom} step response versus experimental data

is worth noticing the existence of a time delay T_d of approximately 67msecs.

4 Controller Design

Once a nominal model has been identified, standard control techniques can be used to synthesize suitable controllers. Note however that when using these techniques, the time delay T_d must be taken into account in order to guarantee stability [15].

Figure (3)a compares simulation and experimental results obtained combining the model G_{nom} with a PID controller, tuned to achieve a compromise between settling time and overshoot. These results match (within the experimental error), provided that a time–delay of two sampling periods is added to the model. However, the controller is tuned to the nominal model G_{nom}, leading to potentially *fragile* closed loop systems. This effect is illustrated in Figure (3b), showing the result of an experiment where the focal length f of the system is increased and then reduced to its previous value. As shown there, while the system exhibits good performance for the nominal value of f, it becomes unstable as f is changed.

The loss of stability can be easily explained from the simple model (1) by noting that (for small values of the target displacement) f enters the model as a gain that multiplies the control action. This is corroborated by the plots in Figure 4(a) showing a magnitude plot of the identified models for different values of f[2].

[2]Stability can be guaranteed by designing a PID controller for the maximum value of f. However, this leads to very slow closed–loop systems for the medium and minimum values of f due to the small value of the gain that is required.

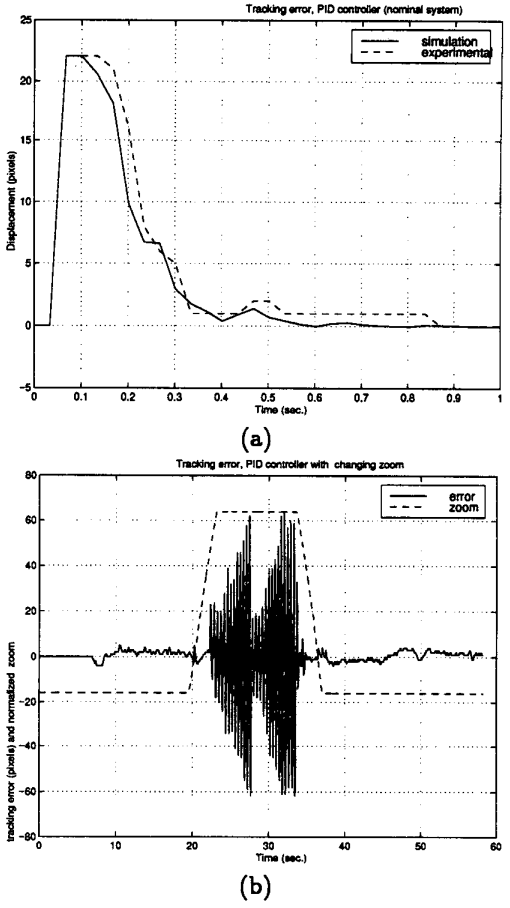
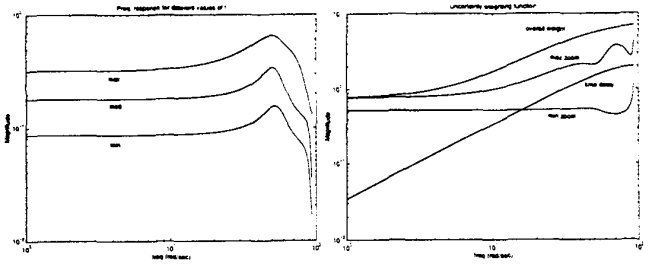


Figure 3: Tracking error for a PID controller: (a) nominal (b) changing f.

In this paper we will model the variation in the plant dynamics due to changes in f as multiplicative dynamic uncertainty, i.e:

$$G_f(z) = G_{nom}(z)(1 + \Delta(z)W_u(z)) \qquad (4)$$

where G_{nom} denotes the nominal transfer function (3), $W_u(z)$ is a fixed weighting function and $\Delta(z) \in \mathcal{BH}_\infty$ represents dynamic model uncertainty. $W_u(z)$ should be selected so that the family (4) covers all possible plants. Equivalently, the magnitude bode plot of $W_u(z)$ should cover the plots of $\frac{G_f(z)-G_{nom}(z)}{G_{nom}(z)}$ for all possible plants. The uncertainties corresponding to the minimum and maximum values of the zoom f are shown in Figure 4(b), where we also show the plot corresponding to uncertainty in the time–delay of up to one sampling period $T_d = 0.033$ seconds (the total time delay fluctuates between two and three sampling periods depending on the time required to process the image). Based on these



(a) (b)

Figure 4: (a) Open loop magnitude plots for different f's (b) Uncertainty weight selection

plots we selected $W_u(z) = \frac{4.7z-3.9}{z-0.08}$.

Robust performance (i.e. guaranteed performance for all plants in the family G_f) can be addressed by using the main loop theorem to recast the problem into an equivalent robust stability problem with an additional fictitious perturbation block Δ_p, as shown in Figure 5. Here the weighting function W_p is used to shape the frequency response of the transfer function $T_{ey_{ref}}$ and thus impose performance specifications on the tracking error $e(t)$.

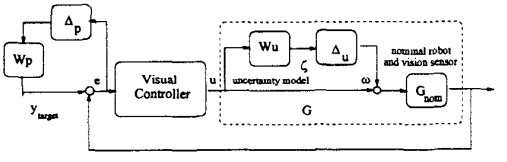


Figure 5: Setup for robust performance synthesis.

The selection of W_p entails a trade–off among different performance specifications, including good regulation, peak control action and settling time. In this case a good compromise is given by $W_p(z) = \frac{2.3z+0.56}{z-0.72}$. This function has a low pass characteristic, penalizing large tracking errors and leading to a closed loop function with a bandwidth on the order of $10Hz$ (equivalently, settling times on the order of 0.5 sec.) Finally, in order to guarantee perfect tracking of step displacements, the controller was augmented with an integrator.

Using μ–synthesis with first order scales we obtained the following controller guaranteeing robust performance:

$$K_\mu(z) = \frac{1.28z^{10}+3.43z^9+3.62z^8+2.87z^7+2.72z^6}{z^{10}+1.91z^9+1.32z^8+1.47z^7+1.99z^6} \frac{+2.09z^5+1.24z^4+0.623z^3+0.151z^2-4.4\cdot10^{-3}z}{+1.33z^5+0.78z^4+0.59z^3+0.29z^2+0.06z-0.002}$$

and the overall controller (including the integrator) is given by $K(z) = K_\mu(z) \times \frac{z}{z-1}$.

5 Experimental Validation

The step responses of the closed loop system obtained with the controller $K(z)$ (simulation and experimental) for two different values of f are shown in Figure 6. Performance in the nominal case is similar (in terms of overshoot and settling time) to that achieved with the PID controller[3]. On the other hand, changes in f result in somewhat degraded performance, but the system remains stable, as opposed to the PID case.

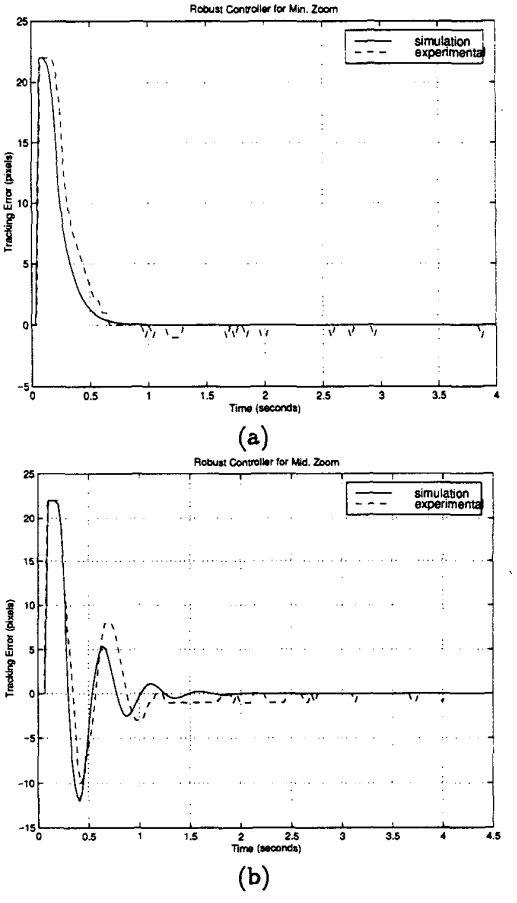


Figure 6: Tracking error for a robust controller: (a) nominal (b)off–nominal

Finally, Figure 7 shows the result when an uncooperative target (a radio controlled truck) was tracked, while the focal length f was increased and then returned to its original value. As shown there, the closed loop system remains stable and achieves tracking with zero steady state error.

[3]Similar results were also obtained with an optimal LQG controller. These results are omitted for space reasons.

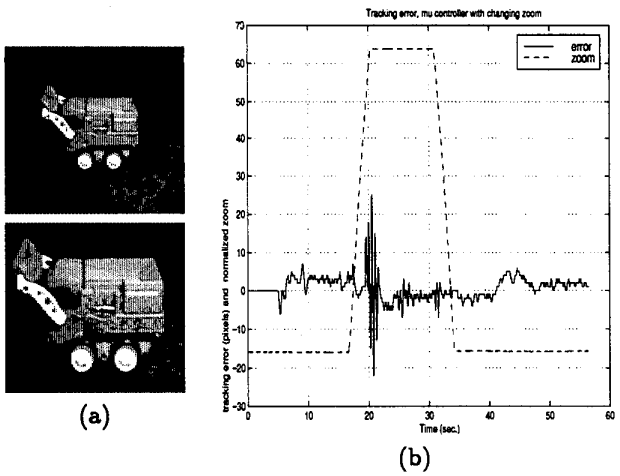


Figure 7: (a) Target images for two different values of f. (b) Tracking error using the μ controller.

6 Conclusions

Recent hardware developments have opened up the possibility of applying active vision techniques to a broad range of real–world problems, such as Intelligent Vehicle Highway Systems, robotic–assisted surgery, 3D reconstruction, inspection, vision assisted grasping, MEMS microassembly and automated spacecraft docking. A salient feature common to all these applications is that using a feedback structure incorporating the visual information in the loop (as opposed to open loop control) offers the possibility of achieving acceptable performance even in the presence of process modelling errors and measurement noise, stemming for instance from poorly calibrated cameras, blurring or only partially determined feature correspondences between images.

However, as noted in [1], synthesizing practical active vision systems capable of moving beyond carefully controlled environments requires controllers that can accommodate (perhaps substantial) uncertainty, stemming for instance from uncertain time delays, unmodelled dynamics and changing optical parameters.

In this paper we addressed these problems by using recently developed non–parametric robust identification techniques to model the system as a family of plants, combined with μ–synthesis techniques to obtain a controller that guarantees acceptable performance for all members of this family. A salient feature of the method is that it requires very few assumptions on the system (for instance the order of the model does not need to be fixed *a-priori*), and that the controller synthesis process does not entail *a posteriori* experimental tuning.

The methodology was experimentally illustrated using a setup with a BiSight pan/tilt stereo head with motorized lenses. It was shown that the responses of the closed loop system obtained with the resulting controller using experimental data and simulations are in close agreement. Furthermore while the responses of the PID and μ controllers are similar for the nominal plant, the latter controller also achieves acceptable performance for off nominal values of the parameters, while the former fails to stabilize the plant.

References

[1] In *IEEE Conf. Decision and Control*. 1998. Invited Session: Active Vision: A New Challenge for Control Theory, O. Camps and M. Sznaier Organizers.

[2] C. Brown, D. Coombs, and J. Soong. Real time smooth pursuit tracking. In A. Blake and A Yuille, editors, *Artificial Intelligence*, pages 123–136. MIT Press, 1993.

[3] P. I. Corke and M. C. Good. Dynamic effects in visual closed–loop systems. *IEEE Transactions on Robotics and Automation*, 12(5):671–683, 1996.

[4] N. J. Ferrier and J. J. Clark. The Harvard binocular head. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(1):69–84, 1993.

[5] G. Hager. A modular system for robust positioning using feedabck from stereo vision. *IEEE Transactions on Robotics and Automation*, 13(4):582–595, 1997.

[6] K. Hashimoto, T. Ebine, and H. Kimura. Visual servoing with hand–eye manipulator–optimal control approach. *IEEE Transactions on Robotics and Automation*, 12(5):766–774, 1996.

[7] K. Hashimoto and H. Kimura. LQ optimal and nonlinear approaches to visual servoing. In K. Hashimoto, editor, *Visual Servoing*, volume 7, pages 165–198. World Scientific Series in Robotics and Automated Systems, World Scientific, Singapore, 1993.

[8] J. Hill and W. T. Park. Real time control of a robot with a mobile camera. In 9^{th} *ISIR*, pages 233–246, Washington D.C., March 1979.

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

[10] N. P. Papanikolopoulos. Integrating computer vision and control for vision assisted robotics tasks. In *1995 American Control Conference*, pages 904–908, Seattle, WA, June 1995.

[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 Transactions on Robotics and Automation*, 9(1):14–35, 1993.

[12] P. A. Parrilo, M. Sznaier, and R. S. Sanchez Pena. Mixed time/frequency domain based robust identification. *Automatica*, 34(11):1375–1389, 1998.

[13] R. S. Sanchez Pena and M. Sznaier. *Robust Systems Theory and Applications*. John Wiley, New Jersey, 1998.

[14] H. Rotstein and E. Rivlin. Optimal servoing for active foveated vision. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 177–182, June 1996.

[15] M. Sznaier and O. I. Camps. Control issues in active vision: Open questions and some answers. In *IEEE Conf. Decision and Control*, pages 3238–3244, 1998.