January 1997

Extending Linear System Models to Characterize the Performance Bounds of a Fixating Active Vision System

Ulf M. Cahn von Seelen
Sensar, Inc.

Ruzena Bajcsy
University of Pennsylvania

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-97-22.

Extending Linear System Models to Characterize the Performance Bounds of a Fixating Active Vision System

Abstract

If active vision systems are to be used reliably in practical applications, it is crucial to understand their limits and failure modes. In the work presented here, we derive, theoretically and experimentally, bounds on the performance of an active vision system in a fixation task. In particular, we characterize the tracking limits that are imposed by the finite field of view. Two classes of target motion are examined: sinusoidal motions, representative for targets moving with high turning rates, and constant-velocity motions, exemplary for slowly varying target movements. For each class of motion, we identify a linear model of the fixating system from measurements on a real active vision system and analyze the range of target motions that can be handled with a given field of view. To illustrate the utility of such performance bounds, we sketch how the tracking performance can be maximized by dynamically adapting optical parameters of the system to the characteristics of the target motion.

The originality of our work arises from combining the theoretical analysis of a complete active vision system with rigorous performance measurements on the real system. We generate repeatable and controllable target motions with the help of two robot manipulators and measure the real-time performance of the system. The experimental results are used to verify or identify a linear model of the active vision system.

A major difference to related work lies in analyzing the *limits* of the linear models that we develop. Active vision systems have been modeled as linear systems many times before, but the performance limits at which the models break down and the system loses its target have not attracted much attention so far. With our work we hope to demonstrate how the knowledge of such limits can be used to actually extend the performance of an active vision system.

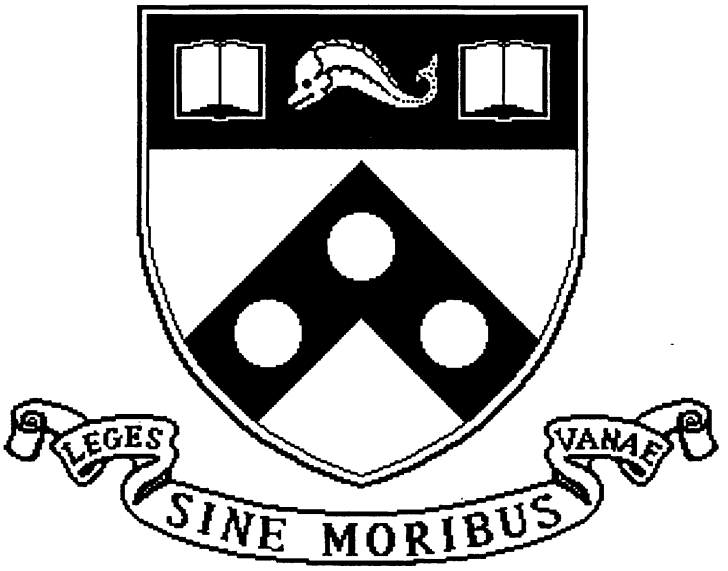
Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-97-22.

Extending Linear System Models to Characterize the Performance Bounds of a Fixating Active Vision System

MS-CIS-97-22 (GRASP LAB 419)

Ulf M. Cahn von Seelen, Ruzena Bajcsy



University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

1997

Extending Linear System Models to Characterize the Performance Bounds of a Fixating Active Vision System

Cahn von Seelen, Ulf M.
Sensar, Inc.
121 Whittendale Drive
Moorestown, NJ 08057 USA
Tel.: +1-609-222-9090
Fax: +1-609-722-1324
cahn@grip.cis.upenn.edu

Bajcsy, Ruzena
University of Pennsylvania
3401 Walnut Street, Suite 301C
Philadelphia, PA 19104-6228 USA
Tel.: +1-215-898-0370
Fax: +1-215-573-2048
bajcsy@grip.cis.upenn.edu

Abstract

If active vision systems are to be used reliably in practical applications, it is crucial to understand their limits and failure modes. In the work presented here, we derive, theoretically and experimentally, bounds on the performance of an active vision system in a fixation task. In particular, we characterize the tracking limits that are imposed by the finite field of view. Two classes of target motion are examined: sinusoidal motions, representative for targets moving with high turning rates, and constant-velocity motions, exemplary for slowly varying target movements. For each class of motion, we identify a linear model of the fixating system from measurements on a real active vision system and analyze the range of target motions that can be handled with a given field of view. To illustrate the utility of such performance bounds, we sketch how the tracking performance can be maximized by dynamically adapting optical parameters of the system to the characteristics of the target motion.

The originality of our work arises from combining the theoretical analysis of a complete active vision system with rigorous performance measurements on the real system. We generate repeatable and controllable target motions with the help of two robot manipulators and measure the real-time performance of the system. The experimental results are used to verify or identify a linear model of the active vision system.

A major difference to related work lies in analyzing the *limits* of the linear models that we develop. Active vision systems have been modeled as linear systems many times before, but the performance limits at which the models break down and the system loses its target have not attracted much attention so far. With our work we hope to demonstrate how the knowledge of such limits can be used to actually extend the performance of an active vision system.

Keywords: Active vision, performance evaluation, tracking, real-time image processing.

Extending Linear System Models to Characterize the Performance Bounds of a Fixating Active Vision System

Abstract

If active vision systems are to be used reliably in practical applications, it is crucial to understand their limits and failure modes. In the work presented here, we derive, theoretically and experimentally, bounds on the performance of an active vision system in a fixation task. In particular, we characterize the tracking limits that are imposed by the finite field of view. Two classes of target motion are examined: sinusoidal motions, representative for targets moving with high turning rates, and constant-velocity motions, exemplary for slowly varying target movements. For each class of motion, we identify a linear model of the fixating system from measurements on a real active vision system and analyze the range of target motions that can be handled with a given field of view. To illustrate the utility of such performance bounds, we sketch how the tracking performance can be extended by dynamically adapting some parameters of the system to the characteristics of the target motion.

The originality of our work arises from combining the theoretical analysis of a complete active vision system with rigorous performance measurements on the real system. We generate repeatable and controllable target motions with the help of two robot manipulators and measure the real-time performance of the system. The experimental results are used to verify or identify a linear model of the active vision system.

A major difference to related work lies in analyzing the *limits* of the linear models that we develop. Active vision systems have been modeled as linear systems many times before, but the performance limits at which the models break down and the system loses its target have not attracted much attention so far. With our work we hope to demonstrate how the knowledge of such limits can be used to actually extend the performance of an active vision system.

1 Introduction

Active vision systems are often designed and analyzed as linear systems, whose behavior, by definition, scales arbitrarily. However, as physical systems they are subject to various nonlinearities that limit the kinds of input stimuli they can handle successfully. A tracking system, e.g., can follow targets only up to certain velocity and acceleration limits. If the target motion exceeds those limits, the tracking system will lose the target from its field of view and fail in its task.

If active vision systems are to serve reliably in practical applications, it is crucial to understand their limits and failure modes. In the work presented here, we derive bounds on the performance of an active vision system in a fixation task. In particular, we characterize the tracking limits that are imposed by the finite field of view. Two classes of target motion are examined: sinusoidal motions, representative for maneuvering targets, and constant-velocity motions, exemplary for more slowly varying target movements. For each class of motion, we verify or identify a linear model of the fixating system from measurements on a real active vision system and analyze the range of target motions that can be handled with a given field of view. To demonstrate the usefulness of such performance bounds, we sketch how the tracking performance can be extended by dynamically adapting an optical parameter of the system to the characteristics of the target motion.

After a brief review of related work in Section 2, Section 3 introduces our experimental setup. It consists of an active vision system and a robot manipulator that generates repeatable and controllable target motions. In Sections 4 and 5, we examine the behavior of our experimental system for sinusoidal and constant-velocity target motions. Finally, Section 6 sketches an application of the derived performance bounds.

2 Related work

Although active vision and visual servoing have a long history (see [11] for a recent review), experimental studies of the performance of the developed systems are rare. The tracking performance is often demonstrated with manually moved targets whose precise motion parameters are unknown.

The lack of systematic experiments partly lies in the need for a mechanism that generates controllable target motions. Probably the simplest form that such a mechanism can take are two light-emitting diodes that flash in counterphase and simulate a square wave trajectory [6]. For repeatable (but not precisely calibratable) motions in the plane, toy train sets are sometimes used [1, 9, 15]. Better calibrated planar trajectories are generated with rotating platforms or turntables [3, 5, 10, 7] and X-Y-tables [4].

Under some circumstances, a degree of freedom in the active vision system itself can be used. In [16], the tilt axis of the neck produces relative target motions that are tracked with the vergence axes. The authors in [3] mount the camera on a gantry robot whose six degrees of freedom are divided into degrees for camera tracking motions and degrees for simulating target motions. The advantage of using a shared motion platform is that target and head motions are executed under the same system clock and can be correlated with each other easily.

The largest design space for controlled experiments is achieved with a robot manipulator. The authors in [12] show the improvement in tracking performance of a series of controllers for a camera head by comparing the visual error in response to a light source moved by a robot arm. Likewise, the motion-based tracker of [8] pursues an object that is moved on a circular trajectory by a robot arm. The authors in [14] test the performance of an adaptive MIMO controller for a camera mounted on a robot arm by moving a target with a second robot arm and recording the relative transform between the end effectors.

In our work, we use two Puma 560 manipulators. The robots allow us to position the head and the target in repeatable positions and to generate precisely specified target trajectories. The next section describes our setup in more detail.

3 Experimental setup

The task in our experiments is fixation of a horizontally moving target by panning a camera. The input to the system is the target's visual angle (relative to the camera pan axis), and the controlled output is the camera's pan angle. The goal of fixation is to keep the target centered in the image by minimizing the difference between target angle and camera angle, the so-called visual error. The size of the visual error provides a measure of how well the system maintains fixation.

3.1 Active vision system

Our active vision configuration for the reported research consists of two principal components: a Helpmate Robotics BiSight binocular camera head with motorized 10:1 zoom lenses, and a network of four C40 digital signal processors for image processing and head/lens control in real-time. The head controller and the DSP network communicate with each other through their host workstations via TCP/IP connections.

The optics of the vision system, including lenses, cameras, and digitizer, has been completely calibrated so that it is known how the physical target motion translates into the observed target motion on the sensor plane. To locate the target in the image, a multi-scale normalized cross-correlation algorithm searches for a predefined target template within an image window. The search runs at video field rate (60Hz) on the DSP network. A simple proportional control law computes setpoints for the camera vergence axes from the recovered target location. A detailed description and analysis of the whole system can be found in [2].

3.2 Target motion generation

Besides a calibrated active vision system, performance measurements require a setup to generate controllable and repeatable target motions. We use two Puma 560 manipulators for this purpose. One robot arm statically holds the camera head in a repeatable position while the other arm moves the target (figure 1).

Because the camera pan axes on the head are rotary degrees of freedom, suitable test stimuli are angular target motions, i.e., circular trajectories about the camera pan axis (for monocular tracking) or the head center (for binocular tracking). The circular trajectories

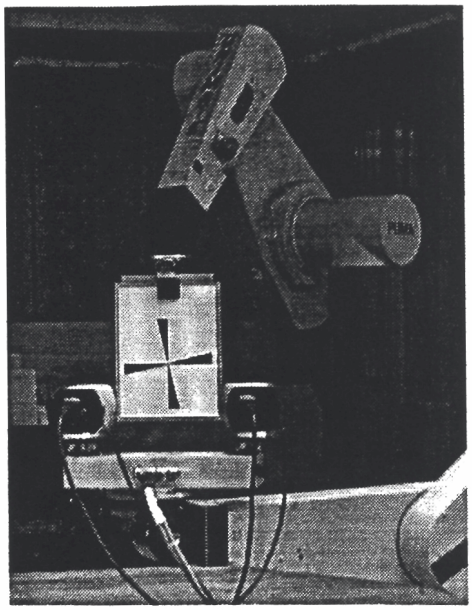


Figure 1: Robot setup for controlled target motion experiments. One robot holds the camera in a repeatable position while the other arm moves the target on an arc around the camera.

also minimize changes in the distance between target and cameras, preventing the target from going out of focus when higher focal lengths are used and the depth of field is small.

Two kinds of motion have been used in our performance measurements. The first type has a sinusoidal angular velocity profile whereas the second type of motion moves the target at a constant angular velocity, with short ramp-up and ramp-down segments at the beginning and end of the motion.

4 Sinusoidal motion

Angular target motion with a sinusoidal velocity profile provides a continuum of velocities and accelerations to the fixating active vision system and can serve as representative for targets maneuvering with high turning rates. Moreover, it is also a stimulus well-suited for identifying the transfer function of a linear system. We measured the response of the active vision system to target motions over a range of amplitudes and frequencies (see table 1). The amplitude of the motion was reduced with increasing frequency to keep the target motion within the velocity and acceleration bounds of the robot arm.

In the experiments, the target was tracked by a single camera at a distance of 134cm with a focal length of 24.5mm. The head controller recorded the camera axis position (as reported by the axis encoders) and the visual error (as reported by the image processing). In addition, the robot's host workstation logged the cartesian position of the target in the robot coordinate system at 50Hz so that the amplitude and frequency of the motion could be verified independently.

Figure 2 shows a typical experimental data set (1Hz target motion). It consists of camera position (top) and visual error (middle), both measured in degree and sampled

Amplitude (deg)	Period (sec)
5	10, 9, 8, 7, 6, 5
2.5	4, 3, 2
1	1, 0.9, 0.8, 0.7. 0.6
0.5	0.5, 0.4, 0.3

Table 1: Amplitudes and periods of sinusoidal target motions

at 565Hz (1.8ms). The visual error exhibits distinct steps because the image processing provides new values only at 60Hz. The camera, on the other hand, moves continously throughout the visual sampling interval. Note that although both camera position and visual error are recorded at the same time, they reflect system states at different times. The camera position measurement takes only a few microseconds whereas the visual error describes a system state 50ms ago.

The goal of the measurements is to estimate the gain and phase relationship between the input (the target motion) and the output (the camera motion) of the active vision system. However, the correspondence between the recorded camera motion and the target motion is prima facie unknown. The robot control, the head control, and the image processing all run

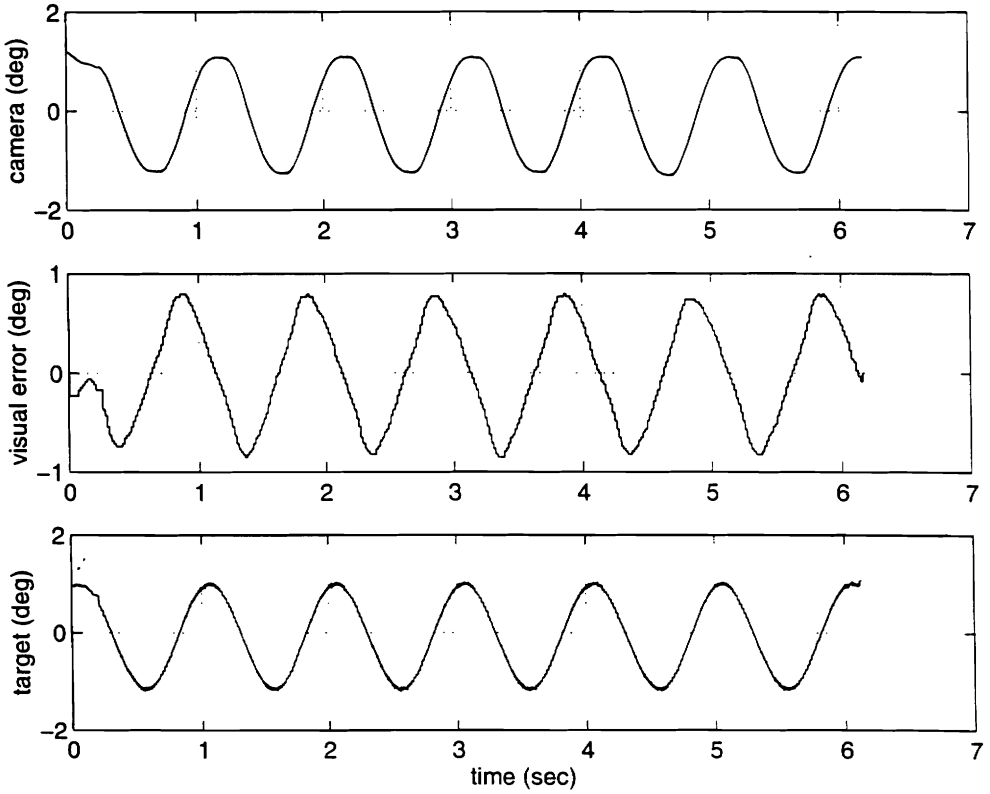


Figure 2: Camera position. visual error, and reconstructed target position for a 1Hz target motion. The camera lags the target by about 37° and is overshooting by 15% (1.2dB).

on different workstations, under independent system clocks, so that the temporal relation of the target motion and camera motion records is uncertain. Therefore it is impractical to work with the target motion as recorded by the robot's host workstation. Instead, the target motion can be reconstructed from the data captured by the camera head controller. The target angle at a given point in time equals the sum of the camera angle reported at that time and the visual error reported 50ms later. Figure 2 (bottom) shows an example of reconstructed target motion. Fitting sinusoids to the recorded camera position data and the reconstructed target position data finally yields the amplitude and phase values necessary to compute the gain and the phase delay of the system response (shown in the next section in figure 4).

4.1 System identification

The basic structure of our active vision system applied to a fixation task is shown in figure 3. Because fixation strives to minimize positional deviation, input and output of the system are position measurements, namely the angular target position θ_T (input) and the camera pan angle θ_C (output).

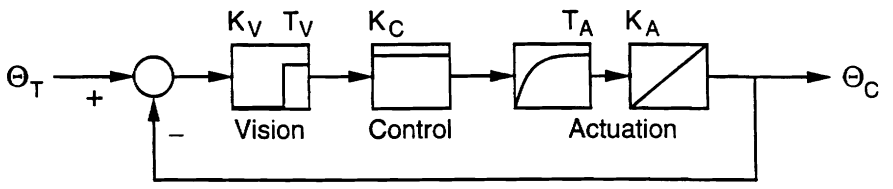


Figure 3: Model of the active vision system for a fixation task

The difference of target position and camera position determines the angular position of the target on the image plane. This position is measured by the vision component of the active vision system; the angular distance from the image center is the visual error. The computation causes a delay T_V of about 50ms (light integration in the sensor (16.7ms), digitization (16.7ms), image processing (14.9ms) and transport of the visual error (1.8ms)). In addition, the offset of the lens nodal point from the camera pan axis introduces a scaling K_V of the visual error. In the reported experiments, the approximate scaling gain is $K_V = 0.9$.

The visual error is converted by the control component into a camera position increment. The current implementation uses a simple proportional control scheme with a gain $K_C = 0.2$.

The actuation component mainly acts as an integrator for the position increments. Since the mechanical motor time constant for the pan axes is fairly small (tens of milliseconds), the motor dynamics is basically invisible at low target frequencies. At higher target frequencies, the motor dynamics needs to be taken into account. A first-order model with time constant T_A and gain K_A proves to be adequate for the system identification, as we will see.

Putting together the models for the individual components, we arrive at the following

closed-loop transfer function $G(s)$ for the active vision system:

$$G(s) = \frac{K_V K_C K_A \, \mathrm{e}^{-T_V s}}{s(T_A s + 1) + K_V K_C K_A \, \mathrm{e}^{-T_V s}}$$

The only unknown parameters in the transfer function are the actuation constants T_A and K_A. A nonlinear least-squares fit of the transfer function to the measured system response yields $T_A = 0.025$ and $K_A = 55$. With these values, the transfer function describes the experimental data well except for the highest examined frequencies (figure 4). There are several sources of corruption of the experimental response. One basic problem is variable friction of the pan axes which changes the mechanical parameters of the system. Second, the time for the transmission of the visual error from the image processing platform to the head controller varies to some extent, and with it the loop delay T_V. Finally, the experimental data becomes less reliable with increasing frequency as the mechanical imperfections start to deform the system response.

The system response indicates two different behaviors of the system. For small target frequencies, the system basically has unit gain and linear phase. It will track the target with a constant delay of approximately 87ms. At high target frequencies, the first-order motor dynamics in the loop give the system a resonant peak around 2Hz. The system will first overshoot and, as the target frequency grows, will increasingly lag behind the target, eventually moving in counterphase to it.

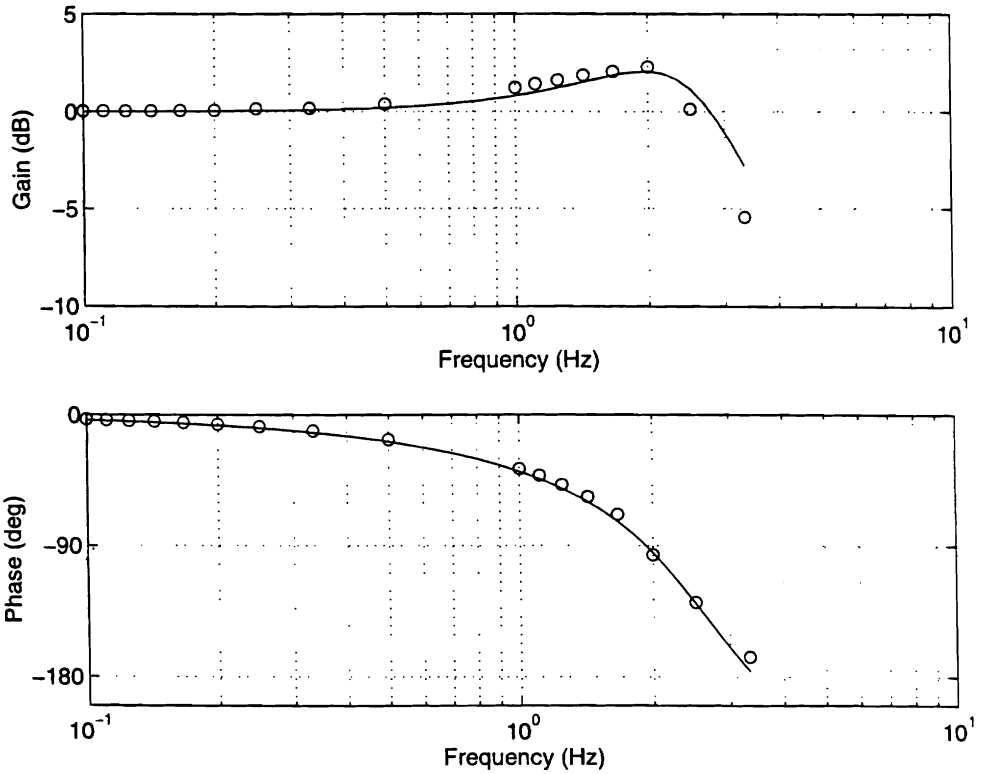


Figure 4: Frequency response of the active vision system. The circles mark the measured data points. Overlayed is the fitted transfer function $G(s)$.

4.2 Limits of the fixation performance

The transfer function $G(s)$ describes the response of our fixating system to a moving target in terms of gain and phase. In particular, the response to a target moving sinusoidally with frequency ω is a sinusoidal camera motion, amplified by $|G(i\omega)|$ and delayed by $\angle G(i\omega)$. The visual error $E(t)$ of the active vision system, i.e., the angular deviation of the target from the image center, is the difference of the target and camera motion, which is again a sinusoid of the same frequency and with amplitude A' and phase Φ:

$$E(t) = A\sin\omega t - |G(i\omega)| \cdot A\sin(\omega t - \angle G(i\omega)) = A'(\omega)\sin(\omega t + \Phi(\omega))$$

The maximum of the visual error, $E(\omega, A)$, is its amplitude A':

$$E(\omega, A) = A' = A \cdot \sqrt{1 + |G(i\omega)|^2 - 2|G(i\omega)|cos\angle G(i\omega)}$$

If the active vision system were truly linear, the target motion could be scaled indefinitely, and the system would still respond to it, namely with an equally scaled visual error. In practice, of course, the system will lose the target as soon as the visual error exceeds one half the field of view, and the linear behavior will break down.

We can now characterize the system limits by plotting the visual error over the amplitude and frequency of the target motion (figure 5). The system will track all motions whose parameters (ω, A) lie below the level contour for the visual error corresponding to one half the field of view. It is apparent that amplitude and frequency of the target motion can be traded off against each other in order to keep the visual error within the bounds of the system capabilities.

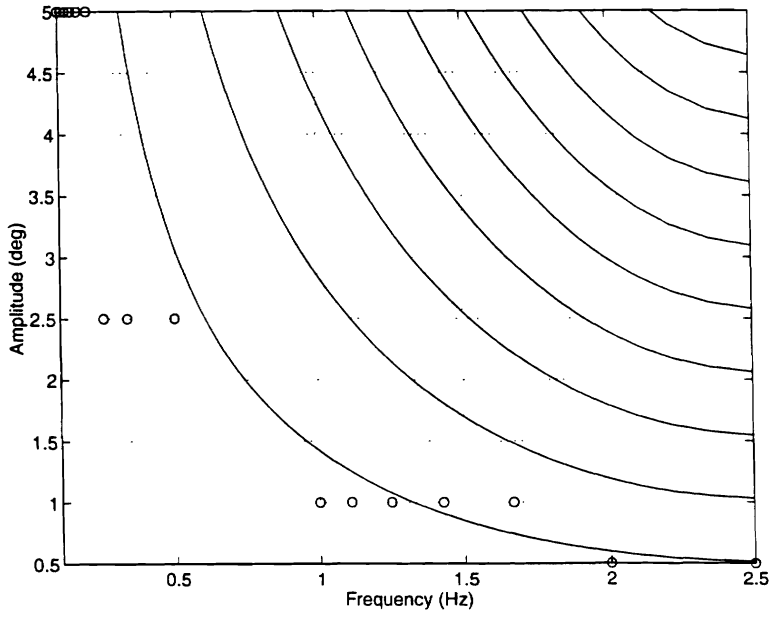


Figure 5: Lines of constant visual error for a range of target frequencies and amplitudes. The left-most line represents 1° visual error. The error increases by 1° with each line. Overlayed as circles are the target motion parameters used during system identification.

5 Constant-velocity motion

So far, we have analyzed the fixation performance of the active vision system for input stimuli that represent targets maneuvering with high turning rates. Another common class of motions are constant-velocity profiles, encountered, e.g., in tracking targets that pass the camera at a distance. In this section, we develop an alternative linear model of our fixating active vision system and measure experimentally the system response to constant-velocity target motions.

5.1 A linear discrete-time model

In the previous section, we identified a continous transfer function that describes the system behavior over a large range of frequencies. Tracking constant-velocity motions, however, involves mostly lower frequencies[1], and we can work with a simpler model of the system that neglects the first-order dynamics of the actuation component. It models the pan axes as pure integrators that achieve the desired position increment within a single sampling period (one video field, 16.7ms). Since under this assumption all system components complete their work in multiples of the sampling period, a discrete-time model can be used.

The input to the system is the angular target position θ_T, the output is the pan angle θ_C of the camera. The image processing is represented by a delay element with a delay of 3 sampling periods (light integration in the sensor (16.7ms), digitization (16.7ms), image processing (14.9ms), and transport of the visual error (1.8ms)) and a scaling gain K_V. The P-control corresponds to a proportional gain K_C, and the actuation is represented by an integrator. Figure 6 shows the complete model, combining K_V and K_C into a single gain K. The closed-loop transfer function of the model looks as follows:

$$G(z) = \frac{\Theta_C}{\Theta_T} = \frac{K}{z^3 - z^2 + K}$$

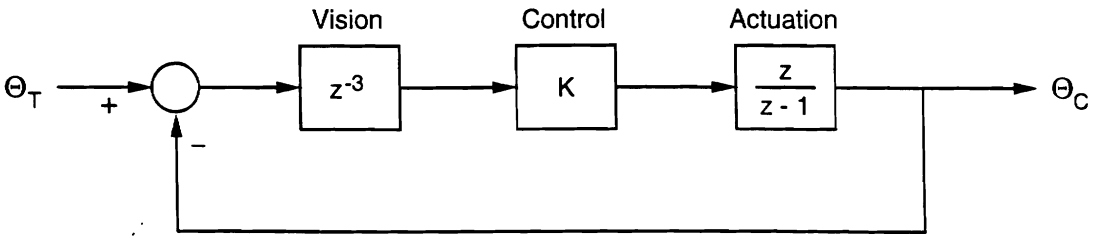


Figure 6: Linear discrete-time model of the active vision system

To analyze the stability of the system as a function of K, we have to determine the poles of $G(z)$, which are the roots of the third-order polynomial in the denominator, $z^3 - z^2 + K$. The polynomial has three real roots for $K \leq 4/27$ and one real and two complex conjugate roots otherwise. Therefore the system is critically damped for a gain $K = 4/27 \approx 0.148$. Simulations in Matlab show that the system remains stable until $K \approx 0.618$.

[1]The Laplace transform of a ramp falls off with $1/s^2$.

5.2 Visual error

A PI-position control system such as our implementation is a first-order system because the pan motor acts as an integrator of the position error. If the input to a first-order system is a position ramp, the steady-state error is nonzero and constant [13]. For the active vision system this means that a target moving with constant velocity will be tracked with a constant visual error.

The previous subsection introduced a linear discrete-time model of the active vision system and analyzed its stability as a function of the control gain K. The steady-state error is defined as the value to which the output of the stable system converges. Let $G(z) = K/(z^3 - z^2 + K)$ be the closed-loop transfer function of the tracker and $R(z) = Tz/(z-1)^2$ be the z-transform of the unit-ramp input sequence $r[n] = Tn$. T is the sampling period of the system (1/59.94s $\approx$ 16.7ms). Then the system response in the z-domain, $Y(z)$, is the product of $G(z)$ and $R(z)$:

$$Y(z) = G(z) \cdot R(z) = \frac{K}{z^3 - z^2 + K} \cdot \frac{Tz}{(z-1)^2}$$

The steady-state ramp response error is defined in the time domain as $\lim_{n\to\infty}(r[n] - y[n])$. To transform the system response $Y(z)$ back into the time domain, we perform a partial-fraction expansion and, neglecting any decaying step and/or sinusoid terms, arrive at the steady-state response $y_{ss}[n]$:

$$y_{ss}[n] = Tn - \frac{T}{K}$$

The steady-state error for the unit ramp response is found as the difference of the input $r[n]$ and the steady-state output $y_{ss}[n]$:

$$\lim_{n\to\infty}(r[n] - y[n]) = r[n] - y_{ss}[n] = Tn - (Tn - \frac{T}{K}) = \frac{T}{K}$$

The steady-state error for the system response to a target moving with velocity v is simply the scaled error $v \cdot T/K$.

5.3 Experimental system response

We have performed extensive experiments with constant-velocity target motions to verify the predicted tracking behavior. In these experiments, the target moved with a constant velocity through an arc while the visual error of the actively tracking camera was recorded. Five sets of experiments were run, using different focal lengths. A range of target velocities was examined for each focal length. Since different focal lengths translate the same visual error given in degrees into different errors when measured in pixels, the trackable target velocities are smaller for high focal lengths. Table 2 lists the examined combinations of focal length and target velocity.

For each combination of focal length and target velocity, the target moved forward and backward through a 20°-arc with a radius of approximately 137cm. The active vision system tracked the target at 60Hz and recorded the visual error.

f (mm)	target velocities (deg/s)
19.9	2, 4, 5, 6, 8, 10, 14, 16, 18, 22, 26, 30, 32
39.8	2, 4, 5, 6, 8, 10, 12, 14, 16
60.6	2, 4, 5, 6, 8, 10
80.5	2, 3, 4, 5, 6, 7, 8
100.8	1, 2, 3, 4, 5

Table 2: Combinations of focal length and target velocity examined in the constant-velocity experiments

Figure 7 shows the steady-state visual error for the various velocities and focal lengths. The superimposed lines are a linear least-squares fit to the data. The fitted slopes agree to within 5% with the predicted steady-state ramp response factor $T/K = (1/59.94)/0.2 = 0.083$.

5.4 Limits of the fixation performance

Unlike the visual error for sinusoidal motions, the visual error for constant-velocity motions is related to the motion parameters in a very simple way: it grows linearly with the target velocity v. The maximum trackable target velocity is therefore proportional to the field of view. The basic requirement for tracking a target moving at a constant velocity v is that the visual error $v \cdot T/K$ be kept smaller than one half the field of view.

For the values $K = 0.2$ and $T = 1/59.94$s of our active vision system, v must be smaller than six times the field of view per second, restricting the target motion to one tenth of the field of view during a single video field or to half the field of view over a period of five video fields. This velocity limit makes sense in the light of an earlier result. The continuous transfer function developed in Section 4 characterizes the system behavior at low frequencies as a pure delay of 87ms, i.e., the system lags the target by approximately five sampling periods. If the target moves from the image center more than half the field of view during this time, the system will lose it from view.

6 Conclusions

In this paper we have shown how to predict the performance bounds of a fixating active vision system for two classes of motion, sinusoidal (representing maneuvering targets) and constant-velocity (representing steady targets). For each class of motion, we have related the visual error to the field of view and derived bounds on the maximum target motion that can be tracked.

The knowledge of the tracking limits can be used to maximize the performance of the active vision system while guaranteeing that the system does not lose the target. An example for this kind of edge-of-the-envelope operation is the task of maximizing the spatial resolution of the target image with the help of a zoom lens. Ideally, we would like to increase the focal length as much as possible to maintain a maximally magnified target

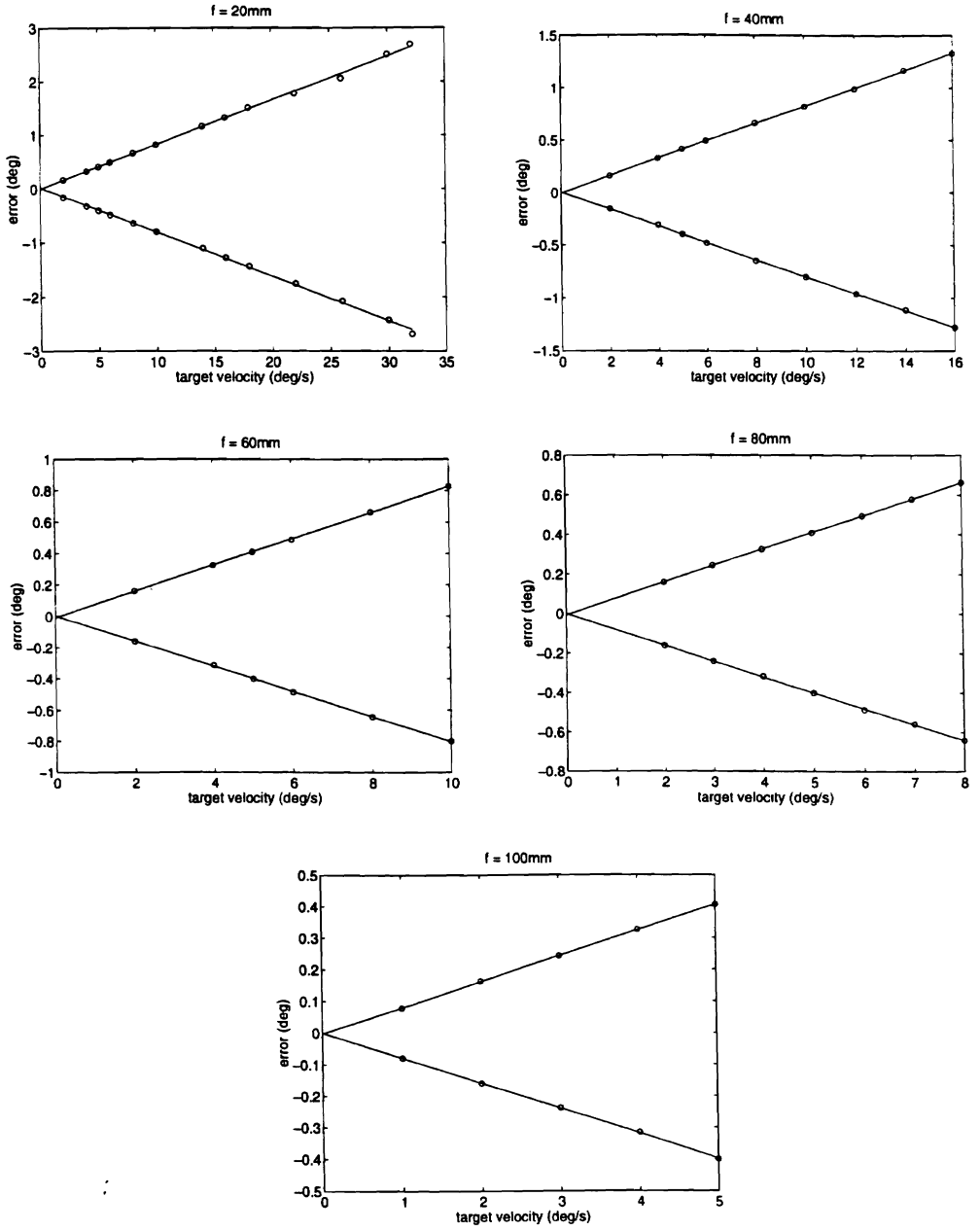


Figure 7: Steady-state ramp-response error as a function of target velocity, observed with five different focal lengths. The overlayed lines show linear least-squares fits. The bottom part of each graph corresponds to the forward motion, the top part to the backward motion.

image. However, as the image magnification grows, the velocity of the target in the image increases, and the visual error with it.

Knowing the performance limits, we can now adapt the focal length to the characteristics of the target motion. The zoom control can be done dynamically, responding to changes in the target motion and assuring that the motion is always safely contained in the resulting field of view.

Future work will address the analysis of active vision systems with more complex control laws. Predictive tracking, e.g., can reduce the visual error and extend the performance limits of the system. Other interesting questions arise from the motors in an active vision system. Besides the finite field of view, they constitute another nonlinear, saturating component whose influence on the system performance remains to be examined.

The performance of an active vision system is a complex notion that goes beyond traditional system measures and depends very much on the task. Performance is also laborious to measure rigorously. It requires a well-calibrated apparatus to precisely establish the value of all parameters that affect the performance. With our work we hope to identify these parameters and illustrate the challenges and benefits of a systematic performance evaluation.

References

[1] Peter K. Allen, Aleksandar Timcenko, Billibon Yoshimi, and Paul Michelman. Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Transactions on Robotics and Automation*, 9(2):152–165, April 1993.

[2] *(Author's PhD thesis.)*

[3] François Chaumette and A. Santos. Tracking a moving object by visual servoing. In *12th World Congress, International Federation of Automatic Control*, pages 409–414, Sydney, Australia, 18–23 July 1993.

[4] Henrik I. Christensen, Jakob Horstmann, and Torben Rasmussen. A control theoretical approach to active vision. In *Asian Conference on Computer Vision*, pages 201–210, Singapore, 5–8 December 1995.

[5] David Coombs and Christopher Brown. Real-time binocular smooth pursuit. *International Journal of Computer Vision*, 11(2):147–164, October 1993.

[6] Peter I. Corke and Malcolm C. Good. Dynamic effects in high-performance visual servoing. In *International Conference on Robotics and Automation*, pages 1838–1843, Nice, France, 12–14 May 1992.

[7] Peter I. Corke and Malcolm C. Good. Dynamic effects in visual closed-loop systems. *IEEE Transactions on Robotics and Automation*, 12(5):671–683, October 1996.

[8] Kostas Daniilidis, Christian Krauss, Michael Hansen, and Gerald Sommer. Real time tracking of moving objects with an active camera. Technical Report 9509, Computer Science Institute, Christian-Albrechts University Kiel, Kiel, Germany, October 1995.

[9] Joachim Denzler and Dietrich W. R. Paulus. Active motion detection and object tracking. In *International Conference on Image Processing*, volume 3, Austin, TX, 13–16 November 1994.

[10] John C. Fiala, Ronald Lumia, Karen J. Roberts, and Albert J. Wavering. TRICLOPS: A tool for studying active vision. *International Journal of Computer Vision*, 12(2/3):231–250, April 1994.

[11] Seth Hutchinson, Gregory D. Hager, and Peter I. Corke. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5):651–670, October 1996.

[12] J. E. W. Mayhew, Y. Zheng, and S. A. Billings. Layered architecture for the control of micro saccadic tracking of a stereo camera head. In *British Machine Vision Conference*, pages 377–386, Leeds, UK, 22–24 September 1992.

[13] Katsuhiko Ogata. *Modern Control Engineering.* Prentice Hall, Englewood Cliffs, NJ, 2 edition, 1990.

[14] Nikolaos P. Papanikolopoulos, Bradley J. Nelson, and Pradeep K. Khosla. Six degree-of-freedom hand/eye visual tracking with uncertain parameters. *IEEE Transactions on Robotics and Automation*, 11(5):725–732, October 1995.

[15] I. D. Reid and D. W. Murray. Active tracking of foveated feature clusters using affine structure. *International Journal of Computer Vision*, 18(1):41–60, April 1996.

[16] Albert J. Wavering and Ronald Lumia. Predictive visual tracking. In *Intelligent Robots and Computer Vision XII, SPIE Proceedings vol. 2056*, pages 86–97, Boston, MA, 8–9 September 1993.