TARGET TRACKING WITH AN ATTENTIVE FOVEAL SENSOR

Laura Li[†], *Douglas Cochran*[†], *and Ross Martin*[‡]

[†] Arizona State University; Tempe, AZ 85287-7206, USA
[‡] Lockheed-Martin Defense Systems; Goodyear, AZ 85338, USA

e-mail: laurali@asu.edu; cochran@asu.edu; ross@ross.interwrx.com

ABSTRACT

This paper addresses the problem of tracking a moving target using measurements taken by a foveal sensor; i.e., a sensor having a foveal region of high acuity surrounded by a periphery in which the acuity of measurements is much less. As in human vision, the foveal region is assumed to be movable and the algorithms developed attempt to exploit it by keeping the target within the foveal region as much as possible. When tracking a target whose motion is modeled by the state of a linear system driven by white gaussian noise, an approach for management of the foveal sensor motion is demonstrated that allows the foveal sensor to always outperform a sensor having no fovea (i.e., all periphery) and to approach the performance of a sensor having no periphery (i.e., all fovea) in some cases.

1. INTRODUCTION

Given a fixed direction of gaze, different areas of the visual field are analyzed at different resolutions by the human visual system. In particular, the resolution of the central or *foveal* region of the visual field is much greater than that of the surrounding *peripheral* region [1, 2]. Faced with the task of tracking a small moving object, human subjects adust their direction of *gaze* by movement of the eyes, the head, or the entire body in order to keep the object within the foveal region of the visual field. This process is known as *attentive vision* [3].

This paper proposes sensor management approaches for a one-dimensional foveal sensors and compares their performance in the context of tracking a moving target whose motion is modeled by a linear system driven by gaussian noise. The approaches proposed all estimate the target's position from the sensor data and attempt to adust the position of the foveal region to keep the target near its center. They are idealized models of the biological process of attentive vision in that no constraints are imposed on the motion of the foveal region. The approaches discussed are based in Kalman filtering and differ in the metods used to compensate for the sensor's nonlinearity in estimating the target's position from the sensor data.

2. SYSTEM MODELS

The target tracking system consists of a signal model, an attentive sensor, and an estimator with a feedback mechanism. The system model is shown in figure 1.

2.1. Target Motion Model

The target position is modeled by one state of a stable multidimensional discrete-time linear system driven by white noise. In particular, dynamical model addressed by the approach developed in this paper has the form

$$x_k = Ax_{k-1} + b\omega_k \tag{1}$$

where A is a stable $n \times n$ matrix and the (scalar) random variables ω_k are independent, zero-mean, and gaussian with variance σ_ω^2. The target's position at time k is the scalar $s_k = g^T x_k$ where g is a column n-vector having exactly one entry equal to one and all others equal to zero (i.e., s_k is a particular component of x_k). With A and g fixed, the agility of the target can be controlled by choice of σ_ω^2.

2.2. Attentive Sensor

To create the foveal effect, the sensor is modeled as a known memoryless nonlinear scalar function γ of the target position s having large slope near $s = 0$ and smaller slope away from $s = 0$. Zero mean gaussian noise of fixed variance σ_η^2 is added to the output of this function so that there is uncertainty in the target position after the measurement. If s falls where the slope of γ is large, the uncertainty in the target position due to the measurement noise will be smaller than if s falls where the slope of γ is small. Thus, it is this nonlinear output map with additive noise that models

0-7803-6514-3/00/$10.00©2000 IEEE

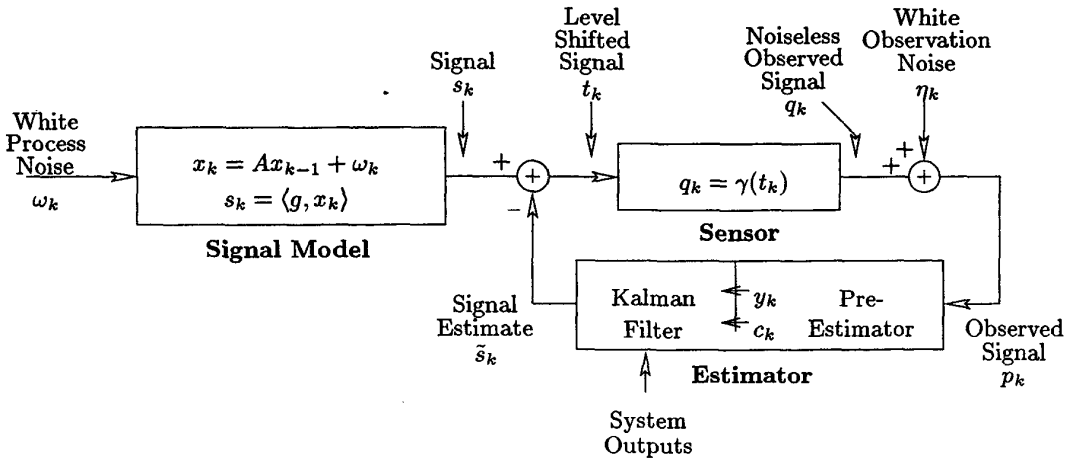

Figure 1: The overall model consists of a target motion model followed by an attentive sensing model. The former is a stable discrete-time linear system driven by white noise. The latter incorporates a noisy nonlinear observation function followed by an iterative estimator. The output of the estimator is used to shift the incoming signal values into the "foveal" region of the observation function.

the variable "resolution" property of the foveal sensor. Mathematically, the observation p_k is given by

$$p_k = \gamma(t_k) + \eta_k \tag{2}$$

where t_k is the input to the sensor and the random variables η_k are independent, zero-mean, and gaussian with variance σ_η^2.

In practice, γ might be chosen as a smooth function approximating, for example, the density of cones on the human retina. Since estimation problems for systems such as given in equation (1) are well understood in the presence of linear observation maps (i.e., via the Kalman filter), it is useful for experiments to consider the case of a piecewise-linear γ. Examples of both types of functions are depicted in figure 2.

To account for foveal motion, which is what makes a foveal sensor valuable, the input t_k is assumed to be formed as the difference of the target position s_k and a shift value chosen by the tracker. Since the fovea in this sensor model is centered at zero, the goal is to choose this offset value so that t_k is near zero. This is accomplished by setting $t_k = s_k - \tilde{s}_k$ where $\tilde{s}_k$ is a prediction of s_k prior to the measurement taken at time k.

2.3. Estimator

From the preceding discussion, it is evident that prediction the target position s_k at time k from past measurements $\{p_0, ..., p_{k-1}\}$, and hence the estimation s_{k-1} from these measurements is essential for effective exploitation of the fovea. With the motion model under

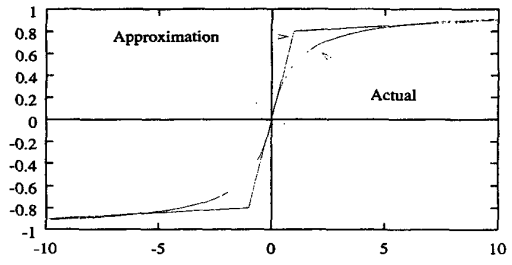

Figure 2: Nonlinear observation functions used to realize a sensor model having a high-acuity (foveal) region surrounded by a low-acuity (peripheral) region. A piecewise linear (PL) observation function as pictured function used for the simulations described in this paper because performance bounds can be readily given. A smooth output function, like the arctangent function pictured, provides a better approximation to a true biological system, however.

consideration, the Kalman filter can be adapted to this role by augmentation with a pre-estimator.

The role of the pre-estimator is to linearize the output map as necessary to apply the Kalman filter; i.e., at each time k, the true observed value p_k and the output map γ must be replaced by a linear gain value c_k and a "measurement" y_k for use by the Kalman filter. Several ways of accomplishing this were investigated, including:

- *High-gain linear approximation:* With a PL output map as described above, c_k is given as the slope of γ near zero and y_k is taken to be p_k. This amounts to assuming the measurement is always taken from within the foveal region.
- *Low-gain linear approximation:* With a PL output map c_k is given as the slope of γ far from zero and y_k is taken to be p_k. This is a suitable approximation if the measurement is rarely taken from within the foveal region.
- *High-gain exact inverse:* With a PL output map, c_k is given as the slope of γ near zero but now y_k is taken to be $\gamma^{-1}(p_k)$.
- *Low-gain exact inverse:* With a PL output map, c_k is given as the slope of γ far from zero but now y_k is taken to be $\gamma^{-1}(p_k)$.
- *Statistical inverse:* With any invertible γ, y_k is taken to be $\gamma^{-1}(p_k)$ and c_k is the value that would yield the same variance for $c_k y_k$ as for the true measurement p_k.
- *Improved statistical inverse:* In this scheme, which achieved the best performance over a wide range of values for the driving signal and measurement noise variances, histograms are used to adaptively estimate the conditional mean and variance of s_k given p_k at each step. Then y_k and c_k are defined to match these statistics.

Regardless of which of these methods is used in the pre-estimator, the Kalman filter is applied to the linearized system

$$\begin{aligned} x_k &= Ax_{k-1} + b\omega_k \\ y_k &= g^T x_k + c_k^{-1}\eta_k. \end{aligned} \tag{3}$$

to provide the pre-measurement target position estimate used in foveal control.

3. SIMULATIONS

To illustrate the utility of the foveal sensor in tracking within the mathematical setting described above,

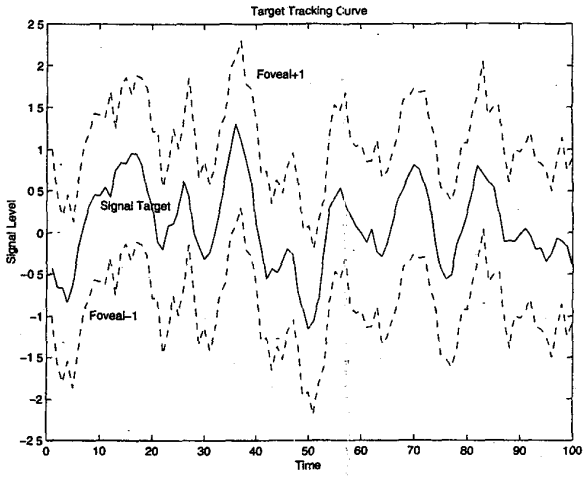

Figure 3: Target position (solid) and foveal edge position (dashed) as functions of time for a system with small measurement noise.

several simulations were performed. All of these simulations used the following system parameters:

$$A = \begin{bmatrix} 0.75 & 0.20 \\ -0.20 & 0.75 \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$g = \begin{bmatrix} 1 & 0 \end{bmatrix}^T \tag{4}$$

The output map was the PL function

$$\gamma(s) = \begin{cases} .1s - 1.9 & s < -1 \\ 2s & s \in [-1, 1] \\ .1s + 1.9 & s > 1 \end{cases} \tag{5}$$

3.1. Target Tracking Curves

Figures 3 and 4 show target position s_k and foveal region bounds $(\tilde{s}_k - 1, \tilde{s}_k + 1)$ as functions of time for scenarios with small measurement noise and larger measurement noise, respectively. The proposed tracking scheme is able to exploit the foveal region for all measurements in the first case and for most of the measurements in the second case.

3.2. Performance Curves

Results obtained using the various pre-estimation strate gies described are summarized in figure 5. The figure shows curves of constant mean-square estimation error (=1.0) for the different strategies as a function of driving signal variance σ_ω^2 and observation noise variance σ_η^2. The upper and lower bound curves come from linear systems with output map slopes corresponding

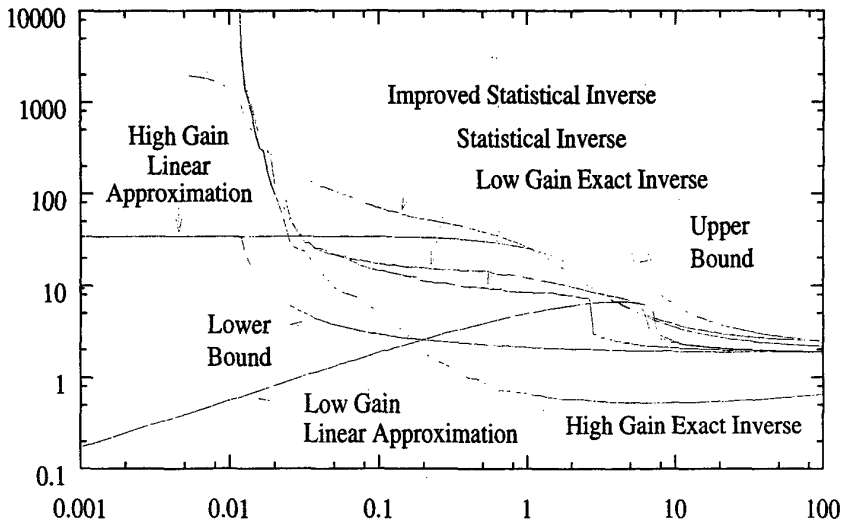

Figure 5: Curves of constant mean-square error obtained using a foveal sensor in target tracking. Various curves correspond to the different pre-estimators described. The upper and lower bound curves come from linear systems with output map slopes corresponding to an all-foveal sensor and to a sensor with no fovea, respectively.

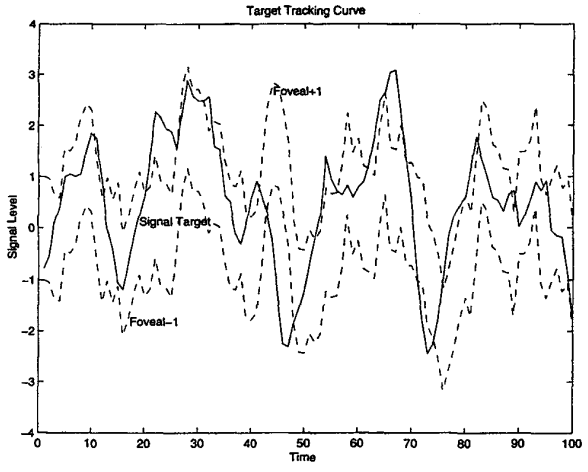

Figure 4: Target position (solid) and foveal edge position (dashed) as functions of time for a system with larger measurement noise that the one used to obtain the result in figure 3.

to an all-foveal sensor and to a sensor with no fovea, respectively.

4. CONCLUSIONS

This paper only demonstrates the viability of exploiting a foveal sensor in target tracking in one fairly specific scenario. Ongoing and future work will consider other target motion models, smooth foveal models, and will seek to develop analytical performance results.

5. REFERENCES

[1] D. Marr, *Vision.* W.H. Freeman and Co., 1982.

[2] J.E. Dowling, *The Retina: An Approachable Part of the Brain.* Harvard University Press, 1992.

[3] A. Black and A. Yuille, *Active Vision.* MIT Press, 1992.

[4] G. Sandini, P. Questa, D. Scheffer, and A. Mannucci, "A Retina-Like CMOS Sensor and Its Applications," *Proceedings of the First IEEE Sensor Array and Multichannel Signal Processing Workshop*, Cambridge MA, March 2000.

[5] A.E. Bryson and Y.C. Ho, *Applied Optimal Control.* Hemisphere Publishing Corp., 1975.