DYNAMIC ESTIMATION WITH SELECTABLE LINEAR MEASUREMENTS

Dana Sinno and Douglas Cochran

Arizona State University, Tempe, AZ 85287-7206, USA

ABSTRACT

This paper is concerned with a class of dynamic estimation problems in which the estimator has the ability to dynamically select, from among a temporally evolving set of possibilities, the source of the data on which the estimate will be based. After motivating and formulating this class of "attentive estimation" problems in some generality, the paper focuses on the special case in which the state of a linear discrete-time dynamical system driven by gaussian noise is to be estimated using linear measurements corrupted by additive gaussian noise. This differs from the standard Kalman filtering problem in that the measurement map at each time step is selectable from a pre-determined set of such maps. When the system dynamics and noise statistics are known, the problem admits a "sensor scheduling" solution, i.e., a criterion for measurement selection that can be used to determine an optimal sequence of output functions in an open-loop fashion prior to the onset of estimation. When the noise statistics or other parameters are unknown, however, closed-loop adaptive strategies for measurement selection can improve estimator performance.

1. INTRODUCTION

The problem of deducing information about a physical environment from measurements obtained by one or more sensors arises in a myriad of engineering applications, including numerous defense and surveillance applications. Consequently, techniques of estimation in both static and dynamic situations and incorporating a wide variety of mathematical models of the physical processes involved have formed a central theme in engineering literature for several decades.

A "rule of thumb" regarding this kind of estimation problem is that more data never hurts; data that is known to contain no information about the parameters of interest can always be ignored. Faced with deciding how to select and configure sensors to measure a physical environment, this more-is-better perspective suggests using as many sensors with as much bandwidth, et cetera, as possible. In practical situations, however, sensor configurations compatible with constraints on availability, cost, communication bandwidth, processing power, and other limiting factors must be selected.

It is noteworthy that biological sensory systems, many of which exhibit performance far surpassing their engineering counterparts, have evolved toward the use of a small number of configurable sensors. Rather than having several eyes, cats have only two eyes – each encompassing a small high-acuity region (retina) surrounded by a field of much lower acuity (periphery). Nevertheless, these animals are exceptionally able in tracking and pursuit of maneuvering prey through the use of eye and head motions that capitalize on the capabilities of their sparse optical sensor suites. In loose analogy with the phenomenon of "attention" in biological systems, the term "Attentive Sensing" is proposed to capture the broad class of dynamic problems in which a subset of (in principle) available sensor data must be selected from which information about an environment or system is to be deduced.

This paper focuses on a particular attentive sensing situation: estimation of the state of a linear discrete-time dynamical system driven by gaussian noise using linear measurements corrupted by additive gaussian noise. This is similar to the standard Kalman filtering problem except that the measurement map at each time step is selectable from a pre-determined set of such maps. When the system dynamics and noise statistics are known, the problem is shown to admit a "sensor scheduling" solution, i.e., a criterion for measurement selection that can be used to determine an optimal sequence of output functions in an open-loop fashion prior to the onset of estimation. When the noise statistics or other parameters are unknown, however, closed-loop adaptive strategies for measurement selection can improve estimator performance.

2. ATTENTIVE ESTIMATION

2.1. Classical iterative estimation

A classical discrete-time iterative estimation problem involves estimating the state x_k of a linear stochastic system

$$x_{k+1} = A_k x_k + \omega_k \tag{1}$$

in which A_k and B_k are matrices and the ω_k are independent vectors of gaussian noise with known covariance matrix Q_k. The estimate is to be based on noisy linear measurements of the state

$$y_k = C_k x_k + \nu_k \tag{2}$$

where C_k is a matrix and the ν_k are independent vectors of gaussian noise having known covariance matrix R_k which are independent of the ω_k.

The optimal estimate of x_n given $y_0,\ldots,y_n$ (in most commonly accepted senses [7]) is $\hat{x}_n = \mathsf{E}[x_n|y_0,\ldots,y_n]$. This estimate is provided iteratively by the Kalman filter.

This work was supported in part by the Tactical Surveillance Systems Division, Defence Science and Technology Organisation, Salisbury 5108, South Australia.

0-7803-4428-6/98 $10.00

2.2. Formulation of the attentive estimation problem

A related attentive estimation problem arises when the state of the system (1) is to be estimated using noisy measurements that are selectable from among a set of linear observation maps; i.e., C_k in (2) is selectable from a set $\mathcal{C}_k$. The goal in this situation is to choose a sensing strategy $\{C_0, ..., C_n\}$ that provides an optimal estimate of $\hat{x}_k$ of x_k from y_0, ..., y_k at each stage k. Throughout the remainder of this paper, the optimization criterion will be taken to be mean-squared error $\mathsf{E}[(x_k - \hat{x}_k)^T(x_k - \hat{x}_k)] = \mathrm{tr}\ \mathsf{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]$. The approach described below is equally applicable to optimizing any function of the estimation error covariance $P_k \triangleq \mathsf{E}[(x_n - \hat{x}_n)(x_n - \hat{x}_n)^T]$, however.

The Kalman filter propagates pre-measurement error covariance S_k and post-measurement error covariance P_k according to [7]

$$\begin{aligned} S_{k+1} &= AP_kA^T + Q \\ P_k &= (S_k^{-1} + C_k^T R^{-1} C_k)^{-1} \end{aligned} \quad (3)$$

Examination of these equations makes the solution to this problem evident: at each time step k, C_k should be selected to minimize the trace of P_k.

Previously published work on this type of problem (e.g., [8, 10]) has emphasized the problem of calculating an observation map C_k that minimizes some given cost function of P_k. The form of the second equation in (3) suggests the formidability of such a calculation. Here, the collection $\mathcal{C}_k$ is assumed to be finite or parameterized in such a way to allow an optimal or nearly optimal C_k to be found by exhaustive search or perhaps some efficient search strategy on the parameterized search space (e.g., a gradient or genetic algorithm).

It is important to note that the solution described here is an open-loop strategy. "Sensor scheduling" can be undertaken based on knowledge of the system parameters before any data are actually collected.

2.3. Examples of attentive estimation

This section presents two examples of attentive estimation. In both cases, the state vectors are three-dimensional and the observation maps are selectable to allow measurement of exactly one of the states at each stage k (i.e., $\mathcal{C}_k = \{(1\,0\,0), (0\,1\,0), (0\,0\,1)\}$).

Example 1:

The system is of the form (1) with $A_k = 0.1I$ (I is the 3×3 identity matrix) and is driven by noise ω_k with covariance matrix $Q = \mathrm{diag}\,(0.1\ 12.0\ 12.0)$ for all k. The variance of the measurement noise is $R = 1$.

The behavior of the attentive estimator is depicted in figure 1. The attentive strategy collects data from sensors 2 and 3 at equal rates, but ignores sensor 1. This result is intuitively appealing in that one would expect to devote more attention to the states about whose behavior one has less certainty (i.e., those with the largest variances should be observed most often). The estimation error variances on states 2 and 3 are bounded below by a bound depending on the stability of A. In this case, the uncertainty about these states is always greater than the uncertainty about state 1, so sensor 1 is never sampled.

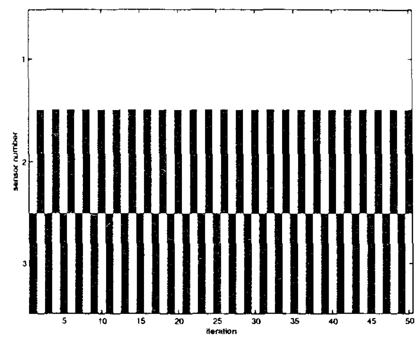

Figure 1: Attentive sensing with a decoupled and stable system. The plot shows sensor number (1, 2, and 3) versus iteration. Black bars indicate when each sensor is "on." Sensor 1 is never measured because its associated state's uncertainty is small.

Figure 2 shows smoothed estimation error as a function of time k for five estimation strategies: (i) measure sensor 1 only, (ii) measure sensor 2 only, (iii) measure sensor 3 only, (iv) measure sensors 1, 2, and 3 in a round-robin fashion, and (v) attentive sensing. In this case, attentive sensing is the best strategy with a mean-square error of 12.96 versus 24.67 for (i), 13.42 for (ii), 13.27 for (iii), and 17.08 for (iv). For obvious reasons, however, strategies (ii) and (iii) are nearly as good as the attentive sensing strategy.

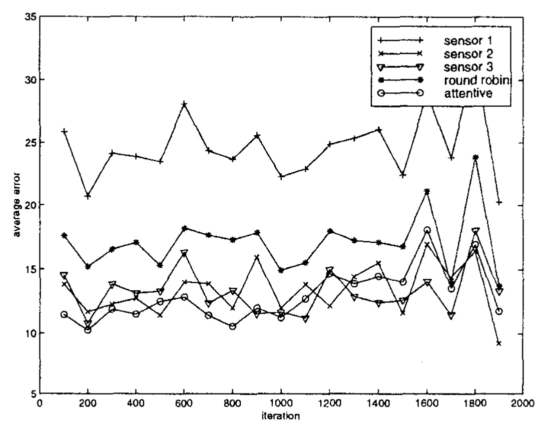

Figure 2: Smoothed estimation error as a function of time for five sensing strategies using the system of Example 1.

Example 2:

The system is identical to that in Example 1 except for *the driving* noise covariance matrix, which is

$$Q = \begin{bmatrix} 1.0 & 0.8 & 0 \\ 0.8 & 1.0 & 0 \\ 0 & 0 & 1.4 \end{bmatrix}$$

As shown in figure 3, the attentive strategy now measure all three states, but emphasizes the more uncertain state 3 over the others.

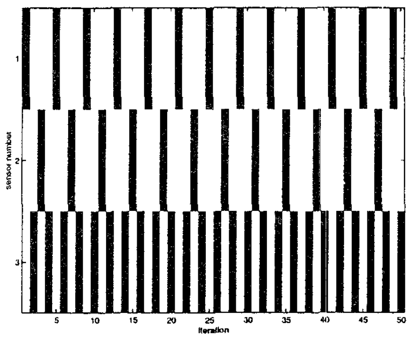

Figure 3: Attentive sensing with a stable system driven by colored noise. Sensor 3 is measured more frequently than sensors 1 and 2 because its associated state uncertainty is greater.

The attentive strategy also provided the lowest mean-square estimation error among strategies (i)–(v) in this example, as expected.

3. ADAPTIVE ATTENTIVE ESTIMATION

The attentive estimation problem changes considerably when the noise covariances are unknown and must be estimated adaptively along with the state. The attentive estimation algorithm's choice of measurement at each stage depends explicitly on the covariance matrices Q_k and R_k, as is evident from (3). This section develops an approach for attentive estimation in the case of system and measurement noise whose covariance matrices are unknown, but assumed to be constant and diagonal. The approach is based on innovation monitoring in the Kalman filter and is similar to published adaptive Kalman filtering ideas (e.g., [6, 9]). The prototype version of this algorithm presented here also assumes that the collection of observation maps available at each time step consists of measurements of each individual component of the system state. The algorithm's performance is demonstrated in a few simple examples.

3.1. Problem formulation

Consider the linear system and measurement models of equations (1) and (2) in which the system noise has constant diagonal covariance matrix (i.e., $Q_k = Q = \mathrm{diag}\,(q_1, ..., q_n)$), the measurement noise is zero, and the observation maps C_k are selectable from among the set of n-vectors with all zero entries except for a single entry of one (i.e., exactly one state can be measured at step k).

Consider the innovation sequence η_k given by

$$\eta_k = y_k - C_k\hat{x}_{k|k-1} = y_k - C_k A\hat{x}_{k-1}$$

where $\hat{x}_{k|k-1}$ denotes the (pre-measurement) estimate of x_k based on $y_0, ..., y_{k-1}$ and $\hat{x}_k$ is the (post-measurement) estimate of x_k based on $y_0, ..., y_k$. The covariance matrix D_k of η_k belongs to the set $\{D^1, ..., D^n\}$ where D^r denotes the r^{th} diagonal element of $AP_{k-1}A^T$ plus $q_r + R$.

With this notation, an algorithm for adaptive attentive estimation is outlined as follows: (1) Start with a diagonal estimate of Q_i of Q with the entries of Q_i chosen sufficiently large to ensure that they overestimate the entries of Q. (2) Run the Kalman filter with an attentive strategy, as described above. (3) Compute estimates of $D^1, ..., D^n$ from the data. (4) Compare the sample statistics of the innovations to those predicted theoretically assuming Q is correct. (5) Update Q by forming a convex sum of the current estimate $\hat{q}_r$ with $\hat{D}^r$ minus the r^{th} diagonal element of $AP_{k-1}A^T - R$ (the weight of the update value in this sum is controlled by an adaptation constant $0 < \alpha < 1$).

3.2. Examples of adaptive attentive estimation

An example of adaptive attentive estimation using the algorithm outlined above is as follows. The system has three states, as in the earlier (non-adaptive) examples, and the observation maps are again selectable to allow measurement of exactly one of the states at each stage k (i.e., $C_k = \{(1\,0\,0), (0\,1\,0), (0\,0\,1)\}$). The system matrix is $A = 0.25\,I$, the actual value of Q is $Q_a = \mathrm{diag}\,(1\ 1\ 2)$, and the initial estimate of Q is $Q_i = \mathrm{diag}\,(6\ 3\ 4)$.

The adaptive attentive estimation strategy is depicted in figure 4. Sensor 1 receives attention initially due to the large initial estimate of its associated driving noise variance. As the estimate of Q improves, the algorithm recognizes that sensor 3 has the highest uncertainty and shifts attention to that sensor. After 300 steps, the estimate of Q in the adaptive attentive algorithm is $Q_{\text{attentive}} = \mathrm{diag}\,(1.46\ 1.41\ 2.09)$. Note that sensors 1 and 2 still receive some attention after a large number of steps.

Figure 5 compares the smoothed estimation error as a function of step k for the adaptive attentive algorithm and a round-robin sensing strategy. Though its state estimation performance is notably better, the attentive algorithm only estimates Q well enough to determine the optimal sensing strategy. The round-robin approach provides a better estimate of Q after 300 steps: $Q_{rr} = \mathrm{diag}\,(1.09\ 1.06\ 2.14)$.

4. CONCLUDING REMARKS

This paper has described and outlined the importance of a broad class of "attentive sensing" problems about which relatively little research has been published. Approaches to a particular type of attentive estimation problem have been introduced and demonstrated in both non-adaptive and adaptive cases.

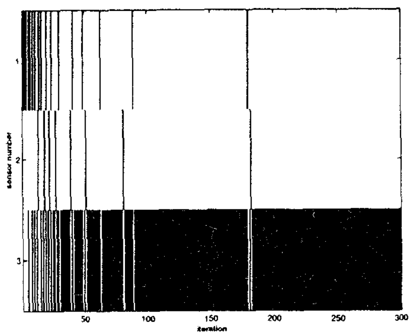

Figure 4: Sensing strategy produced by the adaptive attentive algorithm for the system of Example 1. Sensor 1 receives initial attention due to the large initial estimate of its associated driving noise variance. As the estimate of Q improves, attention shifts to sensor 3.

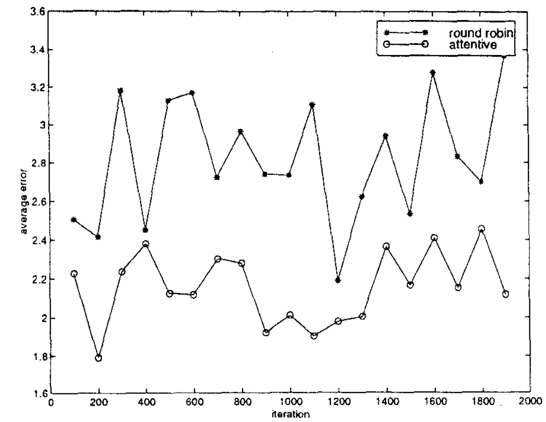

Figure 5: Smoothed estimation error as a function of step k for the adaptive attentive algorithm and a round-robin sensing strategy.

Research currently underway is examining numerous aspects of attentive sensing including, in particular, development of a more general adaptive attentive estimation algorithm that eliminates simplifying assumptions made in this paper.

5. REFERENCES

[1] M. Athans, "On the determination of optimal costly measurement strategies for linear stochastic systems," Automatica, vol. 8, no. 4, pp. 397-412, July 1972.

[2] J.S. Baras and A. Bensoussan, "Optimal sensor scheduling in nonlinear filtering of diffusion processes," SIAM Journal of Control and Optimization, vol. 27, no. 4, pp. 786-813, July 1989.

[3] W.D. Blair and Y. Bar-Shalom, "Tracking maneuvering targets with multiple sensors: Does more data always mean better estimates?" IEEE Transactions on Aerospace and Electronic Systems, vol. 32, no. 1, pp. 450-456, January 1996.

[4] A. Blake and A. Yuille, Active Vision. MIT Press, 1992.

[5] D. Cochran and R. Martin, "Nonlinear filtering models of attentive vision," Proceedings of the IEEE International Conference on Circuits and Systems, Supplementary volume, pp. 27-30, May 1996.

[6] A.H. Jazwinski, "Adaptive filtering," Automatica, vol. 5, no. 4, pp. 475-485, July 1969.

[7] F.L. Lewis, Optimal Estimation. Wiley Interscience, 1986.

[8] D. Marr, Vision. W. H. Freeman and Company, 1982.

[9] R.M. Mehra, "On the identification of variances and adaptive Kalman filtering," IEEE Transactions on Automatic Control, vol. AC-15, no. 2, pp. 175-184, April 1970.

[10] Y. Oshman, "Optimal sensor selection strategy for discrete-time state estimators," IEEE Transactions on Aerospace and Electronic Systems, vol. 30, no. 2, pp. 307-314, April 1994.