An Application of Minimum-Variance Smoothing
Longwall Mining Automation

Laine Mears (mears@clemson.edu) is an assistant professor at Clemson University, South Carolina, in the Department of Mechanical Engineering, and a member of the International Center for Automotive Research (CU-ICAR). He previously worked in the automotive manufacturing industry for over ten years, specializing in high-volume automated assembly. He received the B.S. in mechanical engineering from Virginia Tech and the M.S. and Ph.D. degrees in mechanical engineering from Georgia Institute of Technology. He can be contacted at Clemson University, International Center for Automotive Research, 343 Campbell Graduate Engineering Building, 4 Research Drive, Greenville, SC 29607 USA.

Jeannie Sullivan Falcon is a senior engineer in control and simulation at National Instruments, Austin, Texas. She has led research in control and mechatronics at the Air Force Research Laboratory, Midé Technology Corporation, and National Instruments. She received the B.S. in physics from Carnegie Mellon and the S.M. and Ph.D. in mechanical engineering from MIT.

Thomas Kurfess is the BMW professor of manufacturing, and director of the Campbell Graduate Engineering Center of CU-ICAR. He previously held academic positions at Carnegie Mellon and Georgia Institute of Technology. He received the S.B. and S.M. in mechanical engineering, the S.M. in electrical engineering, and the Ph.D. in mechanical engineering from MIT.

Longwall Mining Automation
An Application of Minimum-Variance Smoothing

GARRY A. EINICKE, JONATHON C. RALSTON, CHAD O. HARGRAVE, DAVID C. REID, and DAVID W. HAINSWORTH

Longwall mining is a method for extracting coal from underground mines. The mining technology involves a longwall shearer, which is a 15-m long, 100-t machine that has picks attached to two drums, which rotate at 30–40 rev/min. A longwall face is the mined area from which material is extracted. The shearer removes coal by traversing a face at approximately 25-min intervals. Traditionally, longwall mining equipment is controlled manually, where the face is aligned using a string line. Under manual control, the face meanders and wanders out of the coal seam, which causes rock to contaminate the coal and limits the rate of production. Coal production is maximized by maintaining a straight shearer trajectory and keeping the face within the seam. Therefore, precise estimates of the face locations are required so that the longwall equipment can be repositioned after each shear.

We are automating longwall equipment to improve production. A Northrop Grumman LN270 inertial navigation unit and an IEEE 802.11b wireless local area network client device are installed within a flame-proof enclosure, which, together with an odometer, are mounted on a shearer. The inertial navigation unit and odometer measure the shearer’s orientation and distance traveled across the face, respectively. The inertial navigation and odometer measurements are stored locally and subsequently forwarded when the shearer is near an access point of the wireless local area network. Upon completion of each shear, inertial navigation and odometer data are used to estimate the position of the face and control the roof support equipment for the next shear. In particular, minimum-variance, fixed-interval smoothing is applied to the inertial and odometer measurements to calculate face positions in three-dimensional space.

Filtering refers to the process of estimating the current value of a signal from noisy measurements up to the current time. In fixed-interval smoothing, measurements recorded over an interval are used to estimate past values of a signal. Compared to filtering, smoothing can provide improved estimation accuracy at the cost of twice the computational complexity. Smoothing is applicable wherever measurements are organized in blocks and retrospective data analysis is feasible. In the case of longwall mining, the...
position estimates and controls are calculated while the shearer is momentarily stationary at the ends of the face.

The fixed-interval smoothing techniques in use today are innovations of the 1950s and 1960s [1]–[3]. In [1], Wiener factorizes power spectral densities and constructs estimators that minimize the error variance. The Wiener solution is called “physically unrealizable by its very nature” [1] because it requires future data. In fact, the Wiener solution is a fixed-interval smoother, which can be implemented by both forward and backward processes. In practice, smoother realizations involve a forward pass over a measurement record, followed by a backward pass over quantities from the forward calculations.

A state-space approach and the minimum-variance filter of [4] are applied to the fixed-interval smoother problem in [2]. The smoother given in [2] maximizes a likelihood function under the assumption that the input noise processes are Gaussian. The structure of the smoother resembles the Kalman filter since the state-update calculation involves a smoother gain that multiplies the error between previous smoother and filter estimates. The fixed-interval smoother of [3] arises as a linear combination of states from a forward Kalman filter and a backward Kalman filter, weighted by the inverses of the error covariances.

The Wiener solution [1] applies to transfer function descriptions of time-invariant plants. A state-space generalization of this solution for output estimation problems is given in [5] and [6], which is applicable to time-varying, nonlinear, and uncertain plants. Unlike the techniques of [2] and [3], the smoother of [5] and [6] minimizes the variance of the output estimation error.

This article reviews the development of the minimum-variance smoother and describes its use in longwall automation. We describe both continuous- and discrete-time smoother solutions. It is shown, under suitable assumptions, that the two-norm of the smoother estimation error is less than that for the Kalman filter. A simulation study is presented to compare the performance of the minimum-variance smoother with the methods of [2] and [3].

LONGWALL MINING

Longwall coal mining is an efficient mining method that can extract a high percentage of coal from a seam. In a longwall coal mine, two long, horizontal and permanent tunnels known as gate-roads are cut into a coal seam to form the boundaries for a large rectangular block of coal, known as a longwall panel [7]. Each gate-road is typically 5–6-m wide. A typical longwall panel is 200–400-m wide, and 2–5-km long, which is the length of the gate-roads. Figure 1 shows the ends of the gate-roads, where major pieces of mining equipment are installed across the longwall face.

Coal is extracted from the longwall panel at the face, which progressively moves back toward the start of the gate-roads [7]. This movement is known as retreat since the longwall panel is mined back toward the starting point of the gate-roads, where the gate-roads branch off at right angles from the main tunnel roadways. Increasingly, the
gate-roads are developed directly from portals in a final highwall of an open-cut operation.

The three main items of longwall mining equipment installed on the face are the shearer, which cuts the coal, the armored face conveyor, which guides and locates the shearer and transports the coal to a conveyor system in the gate-road, and the roof support system. The armored face conveyor has twin steel chains, which are sprocket driven by 800–1500-kW motors situated in the gate-roads. The chains are fitted with scraper bars, which run inside the conveyor and carry the coal along the face. Figure 2 shows the armored face conveyor and the roof support system. The shearer runs back and forth across the armored face conveyor, cutting coal from the face, while the mine roof is held up by the linked but individually hydraulically actuated roof supports, which shuffle forward as the coal is mined, causing the face to retreat. Behind the longwall roof support equipment, the roof collapses under its own weight in a region called the goaf. The goaf can contain unstable voids, which may collapse at any time and thus is off-limits to personnel [7].

Although longwall mining is traditionally a manual process, CSIRO Exploration and Mining is undertaking a Longwall Automation Project, sponsored by the industry-funded Australian Coal Association Research Program. This project is automating several aspects of the longwall mining operation, including a system that aligns the face of the retreating longwall panel perpendicularly to the gate-roads. The automation delivers significant productivity gains and also provides safety benefits since workers are removed from the hazardous face region.

Inertial navigation system position estimates exhibit drift rates of one nautical mile per hour at best and thus cannot be used to automate a longwall shearer. Instead, the milliradian-accuracy Euler angles reported by an inertial navigation system and independent odometer measurements are used in a dead-reckoning calculation to estimate the shearer positions. Optimal fixed-interval smoothing is applied to improve the dead-reckoning estimates, and then the longwall equipment is repositioned at the completion of each shear.

CONTINUOUS-TIME MINIMUM-VARIANCE SMOOTHING

Derivation of the Minimum-Variance Smoother

Smoothing problems can be solved in the continuous- and discrete-time domains (see “Smoothing Problem Definition”). In the continuous-time case, we consider a linear plant \( G \) that operates on the input \( w(t) \in \mathbb{R}^m \) to produce the output \( y(t) \in \mathbb{R}^p \) and has the state-space realization

\[
\begin{align*}
\dot{x}(t) &= A(t)x(t) + B(t)w(t), \quad x(0) = 0, \quad (1) \\
y(t) &= C(t)x(t), \quad (2)
\end{align*}
\]

where \( x(t) \in \mathbb{R}^n \) is the state vector. Suppose that observations

\[
z(t) = y(t) + v(t) \quad (3)
\]

are available, where \( v(t) \in \mathbb{R}^p \) is measurement noise. We desire a linear smoother \( \hat{H} \) that operates on \( z(t) \) and produces an estimate of \( y(t) \) that is optimal in a minimum-variance sense.

Assume that \( w(t) \) and \( v(t) \) are uncorrelated zero-mean white processes with positive-definite covariances \( Q(t) \) and \( R(t) \), respectively. Define an approximate Wiener-Hopf factor \( \Delta \) by

\[
\begin{bmatrix}
\dot{x}(t) \\
\delta(t)
\end{bmatrix} =
\begin{bmatrix}
A(t) & K(t)R^{1/2}(t) \\
C(t) & R^{1/2}(t)
\end{bmatrix}
\begin{bmatrix}
x(t) \\
z(t)
\end{bmatrix}, \quad x(0) = 0, \quad (4)
\]

where

\[
K(t) = P(t)C^T(t)R^{-1}(t) \quad (5)
\]

denotes the Kalman gain, in which \( P(t) \in \mathbb{R}^{n \times n} \) is the symmetric solution of the Riccati differential equation.
\[
\dot{P}(t) = A(t) P(t) + P(t) A^T(t) - K(t) R(t) K^T(t) + B(t) Q(t) B^T(t).
\] (6)

One way of initializing (6) is to use a Hamiltonian matrix solver that employs the results of [8] to find

\[
P(0) = \max_{\hat{P} \in [0, T]} \{ P : 0 = A(t) P + P A^T(t) - P C^T(t) R^{-1}(t) C(t) P + B(t) Q(t) B^T(t) \}.
\] (7)

A smoother \( \mathcal{H} \) can be realized by substituting \( \hat{\Delta} \) into the optimal solution (see “Smoothing Problem Definition”), which yields the system

\[
\mathcal{H} = I - R(t)(\hat{\Delta}^H)^{-1}\hat{\Delta}^{-1},
\] (8)

where \( \hat{\Delta}^H \) denotes the adjoint of \( \hat{\Delta} \) [5], [6]. Suppose that (6) is suitable initialized, and a positive-definite solution \( P(t) \) exists for all \( t \in (0, T) \). The smoother (8) requires the inverse of the approximate Wiener-Hopf factor \( \hat{\Delta}^{-1} \), which is found by taking the inverse of the system (4), resulting in

\[
\begin{bmatrix}
\dot{x}(t) \\
\sigma(t)
\end{bmatrix} =
\begin{bmatrix}
A(t) - K(t) C(t) & K(t) & R^{-1/2}(t) \\
-R^{-1/2}(t) C(t) & R^{-1/2}(t) & 0
\end{bmatrix}
\begin{bmatrix}
x(t) \\
\sigma(t) \\
z(t)
\end{bmatrix},
\] \( x(0) = 0 \). (9)

Similarly, \( (\hat{\Delta}^H)^{-1} \) is realized by

\[
\mathcal{H} = I - R(\Delta^H)^{-1} = I - R(R + \hat{G} Q \hat{G}^H)^{-1}.
\]

Hence knowledge of the plant \( \hat{G} \) together with the second-order noise statistics \( Q \) and \( R \) is sufficient to construct an optimal smoother.

---

**Smoothing Problem Definition**

Consider Figure S1 in which the inputs \( w \) and \( v \) are uncorrelated zero-mean white noises with covariances \( Q \) and \( R \), respectively. The time dependence is omitted so that the notation is applicable to both continuous-time and discrete-time domains. Assume that the linear plant \( g \) operates on \( w \in \mathbb{R}^m \) to produce outputs \( y \in \mathbb{R}^p \). Assuming that measurements \( z = y + v \) are available, we seek a linear system \( \mathcal{H} \) that operates on \( z \) and generates an estimate \( \hat{y} \) of \( y \). Consider the fictitious reference signal indicated by the dotted lines in Figure S2. It follows that the output estimation error \( e = \hat{y} - y \) is generated by \( e = R_i, \) where \( R = [\mathcal{H} \mathcal{H}^2 - G] \) is the linear map from the input noises \( i = \begin{bmatrix} v \\ w \end{bmatrix} \) to the error \( e \).

Let \( \mathcal{G}^H \) denote the adjoint of \( \mathcal{G} \). The problem is to design a system \( \mathcal{H} \) that minimizes \( \| R \mathcal{H} \|_2^2 \), that is, a solution is sought that minimizes the output estimation error covariance. Define

\( \Delta \Delta^H = \mathcal{G} Q \mathcal{G}^H + R \), where the linear operator \( \Delta : \mathbb{R}^p \to \mathbb{R}^p \) is a Wiener-Hopf factor. From the definition of \( R \), it follows that

\[
\mathcal{R} \mathcal{H}^H = \mathcal{G} Q \mathcal{G}^H - \mathcal{G} Q \mathcal{G}^H \mathcal{H} \mathcal{G} Q \mathcal{G}^H + \mathcal{H} \Delta \Delta^H \mathcal{H}^H.
\]

Completing the square leads to

\[
\mathcal{R} \mathcal{H}^H = \mathcal{R}_1 \mathcal{H}^2 + \mathcal{R}_2 \mathcal{H}^2,
\]

where \( \mathcal{R}_1 \mathcal{H}^2 = \mathcal{G} Q \mathcal{G}^H - \mathcal{G} Q \mathcal{G}^H (\Delta \Delta^H)^{-1} \mathcal{G} Q \mathcal{G}^H \) and \( \mathcal{R}_2 = \mathcal{H} \Delta - \mathcal{G} Q \mathcal{G}^H (\Delta \Delta^H)^{-1} \mathcal{G} Q \mathcal{G}^H \). Since \( \| \mathcal{R}_1 \|_2^2 \) excludes \( \mathcal{H} \), this quantity defines the lower bound for the output estimation error covariance.

We seek a solution \( \mathcal{H} \) that minimizes \( \| \mathcal{R} \mathcal{H} \|_2^2 \). Minimizing \( \| \mathcal{R}_2 \mathcal{H} \|_2^2 \) yields the optimum solution

\[
\mathcal{H} = \mathcal{G} \mathcal{G}^H (R + \mathcal{G} Q \mathcal{G}^H)^{-1} = (\Delta \Delta^H - R(\Delta \Delta^H)^{-1}.
\] (S1)

The above smoother is of order \( n^4 \) complexity. An equivalent \( n^2 \) order smoother can be found by simplifying (S1), which yields

---

**FIGURE S1** The output estimation problem. The output of the linear system \( y = Gw \) is contaminated with additive measurement noise \( v \). An estimate \( \hat{y} = H \hat{G} \) is desired, where \( H \) is a linear operator to be designed.

**FIGURE S2** Construction of the output estimation error. The error \( e \) is constructed by taking the difference between the estimate \( \hat{y} \) and the fictitious reference signal \( y \). It can be seen that \( e = H v + (\hat{G} - G) w \).
Longwall coal mining is an efficient mining method that can extract a high percentage of coal from a seam.

\[
\begin{bmatrix}
-\dot{\lambda}(t) \\
\beta(t)
\end{bmatrix} = \begin{bmatrix}
A^T(t) - C^T(t)K^T(t) & -C^T(t)R^{-1/2}(t) \\
K^T(t) & R^{-1/2}(t)
\end{bmatrix}
\begin{bmatrix}
\dot{\lambda}(t) \\
\sigma(t)
\end{bmatrix}, \quad \lambda(0) = 0.
\]

(10)

The estimate \(\hat{\gamma}(t|T)\) of \(y(t)\), given the measurements \(z(t)\) on \((0, T]\), is calculated as

\[
\hat{\gamma}(t|T) = z(t) - R(t)\beta(t).
\]

(11)

Note that \(\hat{\gamma}(t|T)\) estimates only the output of the plant, whereas the smoothers in [2] and [3] estimate the entire state.

The smoother (9)–(11) can be implemented by the following two-step procedure. First, integrate (9) from \(t = 0\) to \(t = T\), to obtain the output \(\sigma(t)\). Second, integrate (10) from \(t = T\) to \(t = 0\), to obtain \(\beta(t)\) for use within (11) to obtain the smoothed estimate \(\hat{\gamma}(t|T)\).

The output \(\beta(t)\) results from the adjoint system \((\hat{A}^H)\), which operates backward on \(\sigma(t)\). In practice, \((\hat{A}^H)\) can be realized by operating the forward recursion \((\hat{A}^T)\) on the time-reversed \(\sigma(t)\) and taking the time reverse of the result. That is, \(\beta(t)\) can be calculated from

\[
\begin{bmatrix}
\dot{\lambda}(\tau) \\
\beta(\tau)
\end{bmatrix} = \begin{bmatrix}
A^T(\tau) - C^T(\tau)K^T(\tau) & -C^T(\tau)R^{-1/2}(\tau) \\
K^T(\tau) & R^{-1/2}(\tau)
\end{bmatrix}
\begin{bmatrix}
\dot{\lambda}(\tau) \\
\sigma(\tau)
\end{bmatrix}, \quad \lambda(0) = 0.
\]

(12)

which evolves forward in the time-to-go variable \(\tau = T - t\). Thus, the smoother can be implemented by integrating (12) from \(\tau = 0\) to \(\tau = T\), computing \(\beta(t)\), and then using it within (11) to obtain the smoothed estimate \(\hat{\gamma}(t|T)\).

**Comparison of Smoother and Filter Performance**

Let \(\mathcal{R} = [\mathcal{H} \ | \ \mathcal{G} - \mathcal{G}]\) denote the linear map from the input noises

\[
i(t) = \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}
\]

to the output estimation error \(e(t) = \hat{\gamma}(t|T) - y(t)\). The error covariance can factored into two components, namely, \(\mathcal{R}^H = \mathcal{R}_1 \mathcal{R}_1^H + \mathcal{R}_2 \mathcal{R}_2^H\), in which \(\mathcal{R}_1 \mathcal{R}_1^H\) excludes \(\mathcal{H}\) and defines a lower bound for \(\mathcal{R}^H\) (see “Smoothing Problem Definition”). It can be shown [6] that the smoother (9)–(11) satisfies

\[
\mathcal{R}_2 = R(t)[(\Delta \hat{A}^H) - (\hat{\Delta} \hat{H} - e(t))^{-1}],
\]

(13)

where \(e(t) = C(t)g(t)\), and \(\hat{\Delta} \hat{H}\) is the operator whose realization is \(\hat{\chi}(t) = A(t)x(t) + B(t)\) and \(\hat{y}(t) = x(t)\). It follows that the two-norm of the output error covariance \(\|\mathcal{R}^H\|_2\) is minimized when the solution of (6) is nonincreasing. Suppose that there exists \(\delta > 0\) such that \(P(t) \geq P(t + \delta)\) for \(0 \leq t \leq \delta\), and

\[
\begin{bmatrix}
B(t)Q(t)B(t) \\
A(t)
\end{bmatrix} \geq \begin{bmatrix}
A^T(t + \delta) & -C^T(t + \delta)R^{-1/2}(t + \delta)C(t + \delta)
\end{bmatrix}
\]

for all \(t \in (0, T]\). Under these conditions, it follows that \(P(t) \geq P(t + \delta)\) for all \(t \in (0, T]\),

\[
\lim_{t \to \infty} \|\mathcal{R}_2 \mathcal{R}_2^H\|_2 = 0,
\]

(14)

and \(e(t) \in L_2\), and thus the smoother error system is asymptotically stable [6]. That is, the smoother achieves the best possible performance whenever the Riccati equation solution reaches steady-state conditions.

The causal part \(\mathcal{H}_+\) of the minimum-variance smoother \(\mathcal{H}\) is given by

\[
\mathcal{H}_+ = \{I - R(\hat{A}^H) \hat{A}^{-1}\}^+
\]

\[
= \{I - R(\hat{A}^H) \hat{A}^{-1}\}^+, \quad \mathcal{A}^H
\]

(15)

which has the realization

\[
\hat{x}(t) = A(t)x(t) + K(t)[z(t) - C(t)x(t)]
\]

(16)

and

\[
\hat{y}(t) = C(t)x(t).
\]

(17)

It can be seen that \(\mathcal{H}_+\) is the minimum-variance filter. Suppose that either the plant \(G\) is time invariant or the solution of (6) is nonincreasing, so that the smoother achieves (14). Let \(\mathcal{R}\) denote the map from the input noises to the filter error \(e(t) = \hat{y}(t) - y(t)\). The filter error covariance can be written as \(\mathcal{R}^H = \mathcal{R}_1 \mathcal{R}_1^H + \mathcal{R}_2 \mathcal{R}_2^H\), in which \(\mathcal{R}_1 \mathcal{R}_1^H\) excludes \(\mathcal{H}_+\) and
\[ R_2 = H_2 \Delta - GQ(t)g^H(\Delta^H_2)^{-1} \]
\[ = R(t)[(\Delta^H_2)^{-1} - (\Delta R^T(2(t))^{-1})\Delta], \quad (18) \]

which implies \( \|R_2 R^T_2\|_2 > 0 \). Thus, from (14),
\[ \|R R^T\|_2 < \|R_2 R^T_2\|_2; \quad (19) \]

that is, the smoother (9)–(11) outperforms the Kalman filter (16), (17).

**Comparison with Alternative Fixed-Interval Smoothers**

We now show that the minimum-variance smoother differs from the two-filter smoother [3] and the Hamiltonian smoother [9], [10], which is equivalent to the maximum-likelihood smoother [2]. It is then demonstrated that the minimum-variance solution outperforms both the maximum-likelihood smoother [9], [10], which is equivalent to the maximum-likelihood smoother [2] and the two-filter smoother [3].

The Hamiltonian smoother [9], [10] is given by
\[
\begin{bmatrix}
\dot{x}(t|T) \\
\dot{\lambda}(t)
\end{bmatrix} =
\begin{bmatrix}
A(t) & B(t)Q(t)B^T(t) \\
-C^T(t)R^{-1}(t)C(t) & A^T(t)
\end{bmatrix}
\begin{bmatrix}
x(t|T) \\
\lambda(t)
\end{bmatrix}
\]
\[ + \begin{bmatrix} 0 \\ C^T(t)R^{-1}(t) \end{bmatrix} z(t), \quad \lambda(T) = 0, \quad (20) \]

and
\[ x(t|T) = x(t) + P(t)\lambda(t). \quad (21) \]

Substituting (21) and its derivative into the first row of (20) yields the Kalman filter (16). Substituting \( \lambda(t) = P^{-1}(t)(x(t|T) - x(t)) \) into the second row of (20) yields the maximum-likelihood smoother [2]
\[ \dot{x}(t|T) = A(t)x(t) + G(t)[x(t|T) - x(t)], \quad (22) \]

where \( G(t) = B(t)Q(t)B^T(t)P^{-1}(t) \).

The two-filter smoother [3] uses the Kalman filter states from (16) within the backward recursion
\[ -\dot{\lambda}(t) = A(t)\lambda(t) + \tilde{K}(t)[x(t) - C(t)\lambda(t)], \quad (23) \]
\( \lambda(T) = 0 \), and the linear combination
\[ x(t|T) = [P^{-1}(t) + \tilde{P}^{-1}(t)]^{-1}[P^{-1}(t)x(t) + \tilde{P}^{-1}(t)\lambda(t)], \quad (24) \]

where \( \tilde{K}(t) = \tilde{P}(t)C^T(t)R^{-1}(t) \), in which
\[ -\tilde{P}(t) = A(t)\tilde{P}(t) + \tilde{P}(t)A^T(t) - \tilde{K}(t)R(t)\tilde{K}^T(t) + B(t)Q(t)B^T(t). \quad (25) \]

The Riccati differential equation (25) can be initialized in many ways, such as (7). It can be seen that the optimum minimum-variance solution (9)–(11) and the smoothers (20)–(24) have similar implementation costs since they all rely on forward and backward processes of order \( n^2 \) complexity. A state estimate can be obtained from (10) using
\[ x(t|T) = C^T(t)[z(t) - R(t)\beta(t)]. \quad (26) \]

where \( C^T(t) = [C^T(t)C^{-1}(t)C^T(t) \) is the Moore-Penrose pseudoinverse. Note that the optimal smoother (9)–(11) is applicable to output estimation and not applicable to state estimation when \( C(t) \) has insufficient rank.

**Scalar Example**

Let \( A(t) = a \in \mathbb{R}, B(t) = 1, \) and \( C(t) = c \in \mathbb{R} \) denote time-invariant parameters of the plant \( G \). Also, let \( p \) and \( k \in \mathbb{R} \) denote the solution of (6) and the Kalman gain (5), respectively.

The transfer function of the Kalman filter (15) is
\[ H(s) = k(s - a + kc)^{-1}, \]

where \( s \) denotes the Laplace transform variable. The transfer function of the maximum-likelihood smoother for output estimation from (22) is
\[ \tilde{H}(s) = cg(-s - a + g)^{-1}H(s) = cgk(-s - a + g)^{-1}(s - a + kc)^{-1}, \]

where \( g = \sigma_a^2p^{-1} \). Applying the two-filter formula (24) results in

![FIGURE 3 Mean-square error versus input signal-to-noise ratio (SNR) for the case of Gaussian measurement and process noises. This plot compares the performance of (i) the optimum minimum-variance smoother, (ii) the two-filter smoother [3], (iii) the maximum-likelihood smoother [2], and (iv) the Kalman filter. Note that the fixed-interval smoothers outperform the Kalman filter. This example demonstrates that the optimum minimum-variance smoother exhibits the lowest mean-square error. At high SNR, however, the difference in smoother performance becomes inconsequential.](image-url)
Gaussian, outperforms the maximum-likelihood smoother.

The optimal minimum-variance smoother differs from the fixed-interval smoother given in [2], [3], [9], and [10].

Consider an ideal scenario in which the system parameters are known precisely and the noise processes are Gaussian. Simulations are conducted for the case $a = -1$, $c = 1$, $\sigma_{w}^{2} = 1$, $T = 100$ s, and $\delta t = 1$ ms, using 1000 realizations of zero-mean noise processes. Figure 3 shows the resulting mean-square error versus input signal-to-noise ratio (SNR). The figure illustrates (19), namely, that the minimum-variance smoother outperforms the causal component. Although the optimal smoother exhibits improved mean-square error, at high SNR the difference in smoother performance becomes inconsequential. For this scalar example, the two-filter formula (24) results in poles common to the optimal solution and outperforms the maximum-likelihood smoother.

The minimum-variance smoother (9)–(11) does not assume that the underlying noise processes are Gaussian. In contrast, the development of the maximum-likelihood smoother in [3] assumes that the noises are Gaussian. Suppose now that the process noise is the unity-variance deterministic signal $w(t) = (\sin(t)/\sigma_{\sin(t)})$, where $\sigma_{\sin(t)}^{2}$ denotes the sample variance of $\sin(t)$. In this case, Figure 4 shows that the optimal smoother can improve on the performance of maximum-likelihood smoother by several decibels.

THE DISCRETE-TIME MINIMUM-VARIANCE SMOOTHER

In the discrete-time case, assume that $G$ has the state-space realization

$$x_{k+1} = A_{k}x_{k} + B_{k}w_{k},$$

$$y_{k} = C_{k}x_{k},$$

where $x_{k} \in \mathbb{R}^{n}$, $A_{k} \in \mathbb{R}^{n \times n}$, $B_{k} \in \mathbb{R}^{n \times m}$ and $C_{k} \in \mathbb{R}^{p \times n}$ with $E[w_{k}w_{k}^{T}] = Q_{k} \in \mathbb{R}^{m \times m}$. Suppose that the observations

$$z_{k} = y_{k} + v_{k}$$

are available, where $v_{k}$ is measurement noise and $E[v_{k}v_{k}^{T}] = R_{k} \in \mathbb{R}^{p \times p}$. The development of the discrete-time, minimum-variance smoother mirrors the continuous-time

$$H(s) = 1 - \frac{\sigma_{w}^{2}}{\sigma_{w}^{2}}(1 - k(s - a + k))^{-1}.$$
case. The same smoother structure arises, but the Kalman gain is calculated differently. The optimum output estimation solution (see “Smoothing Problem Definition”) suggests a smoother of the form

$$\hat{\lambda} = I - K_k(\hat{\lambda}H)^{-1} \hat{\lambda}^{-1},$$  \hspace{1cm} (30)$$

in which $\hat{\lambda}$ is an approximate Wiener-Hopf factor. A state-space realization of (30) is given by

$$\begin{bmatrix} x_{k+1} \\ \alpha_k \\ \beta_k \end{bmatrix} = \begin{bmatrix} A_k - K_kC_k \\ -\Omega_k^{-1/2}C_k \\ -K_k \end{bmatrix} \begin{bmatrix} x_k \\ \alpha_k \\ \beta_k \end{bmatrix} + \begin{bmatrix} 0 \\ \Omega_k^{-1/2} \end{bmatrix}, \hspace{1cm} x_0 = 0, \hspace{1cm} (31)$$

$$\begin{bmatrix} \lambda_{k-1} \\ \beta_{k-1} \\ \lambda_k \\ \beta_k \end{bmatrix} = \begin{bmatrix} A_k^T - C_k^T K_k^T \\ -K_k \end{bmatrix} \begin{bmatrix} \lambda_k \\ \beta_k \end{bmatrix} + \beta_{k-1}, \hspace{1cm} \lambda_N = 0, \hspace{1cm} (32)$$

$$\tilde{y}_{k/N} = z_k - R_k\beta_k, \hspace{1cm} (33)$$

where $K_k = A_kP_kC_k^T \Omega_k^{-1}, \hspace{1cm} \Omega_k = C_kP_{k-1}C_k^T + R_k, \hspace{1cm}$ and $P_k = P_k^T > 0$ is the solution of the Riccati difference equation

$$P_{k+1} = A_kP_kA_k^T - K_k\Omega_kK_k^T + B_kQ_kB_k^T. \hspace{1cm} (34)$$

The Riccati equation (34) can be initialized analogously to (7). A block diagram of the smoother is shown in Figure 5.

LONGWALL MINE POSITION ESTIMATION

The nominal drift rate of a high-quality military inertial navigation system is typically one nautical mile per hour. This high drift error precludes the use of inertial position estimates alone to control an underground longwall shearer. Instead, the Euler angles reported by an inertial navigation system mounted on a longwall shearer. Yaw refers to the horizontal heading of the shearer across the face. Pitch and roll are rotations around the shearer’s transverse and longitudinal axes, respectively.

In the absence of underground truth position references, offline simulations demonstrate the error performance of smoother-based position estimation. The Euler angles $\theta_k, \varphi_k,$ and $\phi_k$ are generated by

$$\begin{bmatrix} \theta_{k+1} \\ \varphi_{k+1} \\ \phi_{k+1} \end{bmatrix} = A \begin{bmatrix} \theta_k \\ \varphi_k \\ \phi_k \end{bmatrix} + w_k,$$

where

$$A = \begin{bmatrix} 0.95 & 0 & 0 \\ 0 & 0.95 & 0 \\ 0 & 0 & 0.95 \end{bmatrix},$$

and $w_k \in \mathbb{R}^3$ with

$$Q = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}.$$

It is assumed that measurements of the Euler angles are available from gyro sensors. The rectangular position coordinates $x_k, y_k,$ and $z_k$ are propagated by

FIGURE 6 The Euler angles measured by the inertial navigation system mounted on a longwall shearer. Yaw refers to the horizontal heading of the shearer across the face. Pitch and roll are rotations around the shearer’s transverse and longitudinal axes, respectively.

FIGURE 7 Mean-square error of the position estimate versus input signal-to-noise ratio. The plot shows the performance of (i) dead reckoning, (ii) the Kalman filter, and (iii) the minimum-variance smoother. Dead reckoning provides the worst performance because the Euler angle measurements are not filtered. The optimal fixed-interval smoother exhibits the best mean-square error since it exploits all of the data available during the measurement interval.
\[
\begin{bmatrix}
    x_{k+1} \\
    y_{k+1} \\
    z_{k+1}
\end{bmatrix}
= \begin{bmatrix}
    x_k \\
    y_k \\
    z_k
\end{bmatrix}
+ (d_k - d_{k-1}) \begin{bmatrix}
    \sin(\theta_k) \\
    \sin(\phi_k) \\
    \sin(\phi_k)
\end{bmatrix},
\]

where it is assumed that distance measurements \(d_k\) are available from an odometer sensor. Figure 7 shows the mean-square error versus input SNR resulting from 1000 realizations of zero-mean Gaussian noise processes, under the simplifying assumption that \(d_k - d_{k-1} = 1\). In Figure 7 the dead-reckoning estimates are calculated from (35). Figure 7 demonstrates that Kalman filtering \([11]\) of the dead-reckoning estimates can yield several decibels improvement in mean-square error. The mean-square error exhibited by the minimum-variance smoother (31)–(33) shows that the optimal smoother outperforms the optimal filter.

Figure 8 shows some shearer trajectories calculated from underground-mine inertial and odometer measurements. The blue and black curves indicate the estimated roof and floor positions, respectively, for 45 shears of a 250-m coalface. The vertical changes in the trajectories, which can be seen in the figure, reflect the horizon adjustments made by the operators to steer the shearer within the naturally undulating coal seam.

Inertial navigation-system position estimates exhibit drift rates of one nautical mile per hour at best. Consequently, if used without additional sensors, the error range of an inertial navigation system increases by one nautical mile per each hour of operation. Therefore, these estimates cannot be used to automate a longwall shearer without additional sensor inputs. We instead combine the yaw, pitch, and roll from the inertial system, which are accurate to better than 1 mrad, together with odometer measurements. The odometer has a resolution of 65,536 counts per revolution that has a gear and ratchet coupling to the shearer rail, resulting in
approximately ±50-cm error over 250 m. From anecdotal observations, obtained when the shearer reaches surveyed points at the gate-roads, the combination of dead reckoning and smoothing yields an average three-dimensional positioning error of about ±20 cm. The estimated shearer positions are used to keep the face straight and in the seam.

Compared to manual control of the mine equipment, the automated system yields improved production rates of 140 tons per hour. In addition to productivity gains, automating longwall equipment leads to safety benefits. The coalface is a hazardous area because methane and carbon monoxide are present, while the area is hot and humid since water is sprayed over the face to minimize the likelihood of sparks occurring when the shearer picks strike rock. By automating manual processes, face workers can be removed from these hazardous areas.

CONCLUSIONS

Longwall mine equipment is traditionally operated manually, which results in the face wandering out of the seam and coal product being contaminated by rock. Since drift error rates from an inertial navigation unit are greater than one nautical mile per hour, they cannot be used alone to control underground mine equipment. We thus use Euler angles from an inertial system along with odometer measurements within fixed-interval smoothing to estimate the face positions and control longwall equipment at the completion of each shear. The automated system keeps the face straight and thus within the seam, while minimizing creep into the gate-roads.

Smoothers can provide performance improvement compared to filtering at the cost of twice the implementation complexity. The minimum-variance smoother involves a cascade of Kalman and adjoint Kalman predictors within the Wiener solution. For time-varying plants, this smoother minimizes the error variance. In the time-invariant case, the smoother is equivalent to the unrealizable Wiener solution, and minimizes the mean-square error. It is demonstrated by an example that minimum-variance solutions can outperform other fixed-interval smoothers. Under ideal conditions, where the system parameters are known precisely and the noises are white Gaussian processes, the performance benefit provided by the optimal smoother can be inconsequential. It is demonstrated by example that, when the noise processes are not Gaussian, the reduction in mean-square error can be significant. The minimum-variance smoother described herein is an output estimator rather than a state estimator, and may not be applicable to state estimation problems in which the output mappings are of insufficient rank.

REFERENCES


AUTHOR INFORMATION

Garry A. Einicke (garry.einicke@csiro.au) received the bachelor’s, master’s, and doctoral degrees, all in electrical and electronic engineering, from the University of Adelaide, in 1979, 1991, and 1996, respectively. He is an adjunct associate professor at the University of Queensland and a principal research scientist in CSIRO Exploration and Mining, where he is the leader of the signal processing team. He chairs the signal processing and communications chapter in the Queensland section of the IEEE. He can be contacted at CSIRO, Technology Court, Pullenvale 4069, Australia.

Jonathon C. Ralston received the B.Eng. degree in electronics and computing and the Ph.D. degree in signal processing from the Queensland University of Technology, Brisbane, in 1991 and 1996, respectively. He is a principal research engineer in CSIRO Exploration and Mining, where he is the leader of the mine automation group.

Chad O. Hargrave received the B.Eng. degree in communications and electronics and the B.A. degree from the University of Queensland, Brisbane, in 1994 and 1999, respectively. He is a Ph.D. candidate at the University of Queensland and a senior research engineer in CSIRO Exploration and Mining, where he leads mine automation projects.

David C. Reid received the B.Eng. degree in electronics and computing and the Ph.D. in signal processing from the Queensland University of Technology, Brisbane, in 1991 and 1996, respectively. He is a principal research engineer in CSIRO Exploration and Mining, where he is the leader of the longwall automation team.

David W. Hainsworth received the B.Eng. degree in communications and electronics and the Ph.D. in signal processing from the University of Queensland, Brisbane, in 1971 and 1976, respectively. He is a senior principal research engineer and leader of the science and engineering program within CSIRO Exploration and Mining.