

SENSITIVITY ANALYSIS OF A SPATIALLY-ADAPTIVE ESTIMATOR FOR DATA FUSION

K. Clint Slatton^{1,3}, Melba Crawford^{2,3}, and Brian L. Evans¹

1: Dept. of Electrical and Computer Engineering, The University of Texas at Austin

2: Dept. of Mechanical Engineering, The University of Texas at Austin

3: The Center for Space Research, The University of Texas at Austin

3925 W. Braker Ln., Suite 200, Austin, TX 78759 USA

Austin, TX 78712-1084 USA

{slatton, crawford}@csr.utexas.edu, bevans@ece.utexas.edu

Abstract - We analyze the parametric sensitivity of a spatially-adaptive multiscale data fusion method. The fusion problem is formulated as a recursive estimation problem in scale and space using a set of 1-D Kalman filters. The overall filter accommodates data acquired at different resolutions and missing data. The filter approaches optimal performance for data with spatially-varying statistics by adaptively updating the filter parameters using the innovation-correlation method. The contribution of this paper is the determination of the estimation error sensitivity to the process noise and measurement noise variances.

1. INTRODUCTION

Remote sensing applications often require analysis of multiple images of a single area that are acquired by different sensors. In general, the images will differ in spatial coverage, spatial resolution, and spectral coverage. It is possible to combine these data sets in a state-space framework so that indirect relationships between the observations and the estimated parameters, as well as the individual error characteristics of each sensor, may be taken into account.

One such application is the determination of topography using Interferometric Synthetic Aperture Radar (INSAR) and Laser Altimeter (LIDAR) sensors. Strip-map INSAR sensors can map topography over large areas with moderate accuracies, while LIDAR sensors can map topography over much smaller areas with very high accuracy. Combining data from the two sensors provides estimates of topography that are more accurate than those obtained from INSAR alone, yet have coverage that is both dense and extensive, which is difficult to obtain with LIDAR.

We apply the filter to high resolution LIDAR data and coarse resolution INSAR data acquired over a vegetated flood plain in Austin, Texas. We combine the two data types using a multiscale Kalman filter, which provides the estimates and estimate uncertainties at each pixel. We generalize the multiscale data fusion method in [1] by introducing a mechanism to detect and adapt to model errors and

non-stationarities in the observations. The 1-D Kalman filters that comprise the spatially adaptive component calculate the correlation of the innovations to update the process noise variance. We present the basic data fusion framework in Section 2 and our adaptive extension of it in Section 3. In Section 4 we describe the parametric dependence of the adaptive estimator. Finally, results and conclusions are presented in Sections 5 and 6.

2. MULTISCALE FILTER

We implemented a multiscale estimation of 2-D stochastic processes on a quadtree. The multiscale estimation is initiated with a fine-to-coarse sweep up the quadtree that is analogous to Kalman filtering with an added merge step. The fine-to-coarse sweep up the quadtree is followed by a coarse-to-fine sweep down the quadtree that corresponds to Kalman smoothing. We refer to this algorithm as the multiscale Kalman smoother (MKS). Using the scalar form for clarity, the linear coarse-to-fine model in [1] is given as

$$\begin{aligned}x(s) &= \Phi(s)x(Bs) + \Gamma(s)w(s) \quad \forall s \in \mathcal{S}, s \neq 1 \\y(s) &= H(s)x(s) + v(s) \quad \forall s \in \mathcal{T} \subseteq \mathcal{S}\end{aligned}\quad (1)$$

where s is the node index on the tree, and $s = 1$ denotes the root node. Here, x is the state variable, and y represents the observations. The stochastic forcing function w is assumed to be a Gaussian white noise process with unit variance, and the measurement error v is a Gaussian white noise process with sensor-dependent variance $R(s)$. \mathcal{S} represents the set of all nodes on the quadtree, and \mathcal{T} denotes those nodes at which an observation is available. B is a back-shift operator in scale, such that Bs is one scale coarser than s . Φ is the coarse-to-fine state transition operator, Γ is the coarse-to-fine process noise standard deviation, H is the measurement-state relation, and R represents the measurement variance of the observations. The scale is represented by the level in the quadtree, and is denoted by m . The support of the image at level m is $2^m \times 2^m$, where $m \in \{0, \dots, M\}$.

The standard Kalman formulation provides optimal estimates (in the mean squared sense) when there is perfect a

This work was supported by the National Aeronautics and Space Administration, under the Graduate Student Research Fellowship Program (Grant NGT-50239).

a priori knowledge of the state and measurement model parameters $\{\Phi(s), \Gamma(s), H(s), R(s)\}$. If errors exist in the assumed process or measurement noise variances, then the computed estimates and estimate uncertainty will be incorrect.

In previous work [2] we implemented a data fusion framework based on a multiscale Kalman filter presented in [1]. A quadtree data structure is created with the leaf nodes corresponding to LIDAR data at 1.25 m resolution, while INSAR data at 10 m resolution are used three levels up in the quadtree. The state process is assumed to follow a $1/f$ model known as fractional Brownian motion. Fractional Brownian motion models can be used to represent a wide range of natural process such as topography and atmospheric turbulence [3]. Using this model, the power spectrum of the state variable $x(s)$ can be represented by the multiscale model in (1) by using $\Phi(s)=1$ and $\Gamma(s) = \Gamma_0 2^{(1-\mu)m/2}$ [1]. The values of Γ_0 and μ are determined by matching the power spectra of the observations and $x(s)$.

The MKS approach provides reasonable estimates for the evolution of the state process in scale $\hat{x}(s)$. However, the resulting coarse-to-fine process noise variance $\Gamma^2(s)$ is constant at each scale; hence, it cannot accommodate non-stationarities in the imagery at a particular scale. The fine-to-coarse process noise variance $Q(s)$, used in the Kalman filter, is a function of $\Gamma^2(s)$ and is therefore also spatially uniform. An image that contains more than one landcover type, e.g. grassland and forest, is an example of a non-stationary 2-D process. Therefore, using a data model that is variable in scale but uniform in space will lead to suboptimal estimates in general.

3. ADAPTIVE MULTISCALE ESTIMATION

Several approaches to adaptive estimation have been reported in the literature. The *innovation-correlation* method is employed in this work [4] [5]. This method is theoretically based, and is computationally attractive because it is noniterative. It is also well suited to stochastic problems in which the precise dynamics of the process are not known. The prediction errors in the Kalman filter (innovations) comprise a zero-mean, white, Gaussian noise sequence when the model parameters are correct. Therefore, testing the innovation sequence for non-white behavior by computing the autocorrelation function (ACF) of the innovations indicates the presence of model errors.

For the 1-D estimation in [4], the innovation sequence of the Kalman filter is a function of the process noise variance Q . The ACF of the innovation sequence is computed, and if at least 5% of the ACF samples exceed a 95% confidence interval test threshold, the innovations are considered non-white. A new Q is then estimated using the ACF samples. The innovation-correlation algorithm was generalized in [5] to update Q locally. We apply the method of [5] in the spatial dimensions of the quadtree to produce asymptotically normal, unbiased, and consistent estimates of the fine-to-

coarse process noise variance $Q(m, i, j)$, where (i, j) are the image pixel coordinates. We refer to this algorithm as the adaptive MKS (AMKS) algorithm.

The multiscale filter initiates at the base of the quadtree according to the standard multiscale algorithm [1]. Upon reaching a level that contains dense observations m_c , two 1-D Kalman filters operate along the rows and columns of the image. The *a priori* estimates of the spatial filters include the multiscale estimates from the level below. The location-dependent estimates of $Q(m_c, i, j)$ from the row-wise filter and column-wise filter are averaged together. Q is updated by each filter in blocks of $1 \times N_b$ pixels. Letting $N_b = 2^{m_c}/4$ works well for images where $128 \geq 2^M \geq 512$. Here, M is the number of quadtree levels.

The spatial portion of the Kalman filter is applied only at levels in the quadtree where dense observations are present so that the innovation sequences have uniform support. The $Q(m_c, i, j)$ estimates from the spatial filters are then used to weight $Q(s)$ at each pixel upon returning to the multiscale filter at $m = m_c - 1$. The filter then proceeds up the remainder of the quadtree according to the standard multiscale algorithm.

4. PARAMETRIC DEPENDENCE OF THE ADAPTIVE FILTER

Because the spatial Kalman filters employed in the AMKS algorithm are scalar-valued and reach steady state quickly, we examine the solution to the scalar time-invariant discrete-time Riccati equation to understand the parametric dependence of the spatial filter performance.

The *a priori* error variance represents expected variance between the true and *a priori* estimated states

$$P_k^- = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T] \quad (2)$$

The scalar time-invariant Riccati difference equation describes the evolution of the discrete-time *a priori* error variance P_k^- as a function of $\{\Phi, Q, H, R\}$.

$$P_{k+1}^- = \Phi_k^2 \left\{ P_k^- - \frac{(P_k^- H_k)^2}{H_k^2 P_k^- + R_k} \right\} + Q_k \quad (3)$$

This difference equation can be transformed into a system of two simultaneous linear difference equations [6]:

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} = \begin{bmatrix} \Phi_k + \frac{Q_k H_k^2}{\Phi_k R_k} & \frac{Q_k}{\Phi_k} \\ \frac{H_k^2}{\Phi_k R_k} & \frac{1}{\Phi_k} \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix} \quad (4)$$

The general solution of (4) is

$$\begin{bmatrix} A_k \\ B_k \end{bmatrix} = \Psi^k \begin{bmatrix} P_0 \\ 1 \end{bmatrix} \quad (5)$$

where

$$\Psi = \begin{bmatrix} \Phi + \frac{H^2 Q}{\Phi R} & \frac{Q}{\Phi} \\ \frac{H^2}{\Phi R} & \frac{1}{\Phi} \end{bmatrix} \quad (6)$$

5. RESULTS

Table 1. Percent improvement in MSE obtained by the adaptive filter for different values of Q and $\Delta Q(Q/R)$.

Q	$\Delta Q(Q/R)$	Δ % MSE
10.0	0.10	52.4
1.0	0.01	38.5
10.0	1.00	53.7
1.0	0.10	44.7
10.0	10.00	16.6
1.0	1.00	30.5
10.0	100.00	2.2
1.0	10.00	18.8

The Ψ^k term in (5) dictates the rate of decay of the solution. The rate of convergence of the Kalman filter clearly depends on the magnitudes of Q and R and the ratio of noise variances Q/R . When Q/R is large, the solution decays rapidly. The filter becomes highly reactive to new input data because it weights the data more heavily than the signal model. This situation can arise when a process with large Q is observed using a sensor with a relatively small R . This is sometimes the case with near-range INSAR observations of terrain containing both forested and non-forested regions.

A series of 1-D simulations were performed to determine the impact of the adaptive estimation as a function of Q , Q/R , and ΔQ , where ΔQ is a multiplicative factor that represents an error in the *a priori* value of Q used in the spatial filter. One filter tracks the process with constant, but erroneous process noise variance $\Delta Q \cdot Q$. The other begins with the same noise variance, but employs the innovation-correlation method used in the AMKS algorithm to update its estimate of Q . The degree to which the adaptive filter improved upon the results of the non-adaptive filter was found to be a function of the magnitude of Q and the ratio $\Delta Q(Q/R)$. The resulting improvements in mean squared error (MSE) achieved by the adaptive filter are shown in Table 1. The smallest relative improvement in MSE of the adaptive estimator corresponds to the largest value of $\Delta Q(Q/R)$.

The relative reduction in MSE obtained from the adaptive filter was greater than 10% except in the case of $\Delta Q(Q/R) = 100$. In that case, the true $Q/R = 10$, and $\Delta Q = 10$. In other words, an already large ratio of process noise variance to measurement noise variance was exasperated by an overestimated *a priori* value for Q in the filter. The effective Q/R in the non-adaptive filter was 100, the largest value of the eight cases examined.

The mechanism by which the large $\Delta Q(Q/R)$ value impacts the performance of the adaptive estimator involves the ACF of the innovations. The large $\Delta Q(Q/R)$ reduces the energy of the non-zero lag terms in the ACF of the innovations relative to the zero-lag term, which reduces the apparent correlation of the innovations. As a result, the innovation-correlation algorithm is less likely to detect non-white statistics in the innovations of the filter due to the erroneous *a priori* value of Q .

Figure 1 shows a dense set of LIDAR data that were acquired over the test site. The imaged area covers 300×300 m² of a river flood plain in Austin, Texas. Coarse INSAR data were acquired over the same area by the NASA/JPL TOPSAR sensor. Vegetation heights were calculated from the observations [2]. A sparse subset of the LIDAR data is used to represent the incomplete coverage of terrain that is typical for LIDAR sensors. A uniform sparse grid of LIDAR data representing 15% of the total data set was used in the data fusion.

The $Q(i, j)$ arrays are uniform at every scale in the MKS algorithm, but are updated in the AMKS algorithm at the coarse observation scale m_c (see Fig. 2). Figure 3 shows the INSAR and LIDAR observations of vegetation height used in the data fusion, the resulting estimates, and the corresponding estimate uncertainty.

The process noise associated with the actual vegetation height varies because the image contains both grass fields and forest. The *a priori* estimates for the fine-to-coarse process noise variance Q in the MKS algorithm are spatially uniform and therefore necessarily incorrect in some places. The adaptive algorithm is able to detect this error and correct for it. The resulting fused estimates cover the entire imaged area, including areas where LIDAR data were withheld. Yet, the high-resolution LIDAR information is preserved where it is available because of its small measurement uncertainty. The INSAR data is effectively sharpened through the fusion with the LIDAR data. The estimate uncertainty in Fig. 3 is the squareroot of the Kalman smoother error variance P^s [2]. The estimate uncertainty is lowest where the LIDAR data are available.

With error defined as the difference between the estimates and the full set of LIDAR data, the AMKS algorithm achieves smaller MSE than the MKS algorithm. The percent reduction in global MSE from using just the coarse data to using the multiscale fusion is 70%. The additional percent reduction in MSE from the standard multiscale fusion to the spatially-adaptive multiscale fusion is 0.03%. This small reduction in MSE of the AMKS algorithm, relative to the standard MKS algorithm, is consistent with the 1-D simulation results because $(Q/R) > 1$ and the *a priori* Q was larger than the true Q , as indicated by a downward revision of the Q estimates.

The adaptive estimator impacts the multiscale data fusion results less than the simple 1-D estimation results. This is because it is implemented on the dense data only, which correspond to coarse scales. In this example, the INSAR data are eight times coarser than the LIDAR data, $m_c = M - 3$. The impact of the updated Q is greatest when $m_c = M - 1$.

The Q estimated by the adaptive filter is closer to the true Q than the standard multiscale result, which indicates that the adaptive estimation performance is closer to optimal. The adaptive estimation also provides additional insight into the spatial variability of the observed process. The AMKS algorithm requires approximately 15% more computation

time than the MKS algorithm.

6. CONCLUSIONS

The multiscale estimation represents an improvement over simply splicing in high-resolution data where ever it is available because it accounts for the fact that observations acquired at different resolutions actually observe different processes. The improvement over the standard multiscale estimation provided by the spatially-adaptive estimation is small in magnitude in this particular case, but represents in important extension nonetheless. It addresses the error that results in the standard multiscale estimation when the process noise is not constant or the *a priori* estimate of the process noise variance is incorrect and therefore delivers performance that is closer to optimal.

The amount of improvement of the adaptive estimation over the non-adaptive estimation depends on the measurement error variance R of the dense observations (INSAR in this case) and the process noise variance Q . R is determined by sensor characteristics and the data. The *a priori* value of Q is also determined from that data by matching power spectra. Therefore, it is not generally possible to control the values of Q and R . But in cases where the resulting $Q/R < 1$, greater relative reduction in MSE from the AMKS algorithm can be expected.

7. REFERENCES

- [1] P. W. Fieguth, W. C. Karl, A. S. Willsky, and C. Wunsch, "Multiresolution optimal interpolation and statistical analysis of TOPEX/POSEIDON satellite altimetry," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 280–292, Mar. 1995.
- [2] K. C. Slatton, M. M. Crawford, and B. L. Evans, "Fusing interferometric radar and laser altimeter data to estimate surface topography and vegetation heights," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 2470–2482, Nov. 2001.
- [3] G. W. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proc. IEEE*, vol. 81, pp. 1428–1450, Oct. 1993.
- [4] R. K. Mehra, "Approaches to adaptive filtering," *IEEE Trans. Automatic Control*, pp. 693–698, Oct. 1972.
- [5] G. Noriega and S. Pasupathy, "Adaptive estimation of noise covariance matrices in real-time preprocessing of geophysical data," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 1146–1159, Sep. 1997.
- [6] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

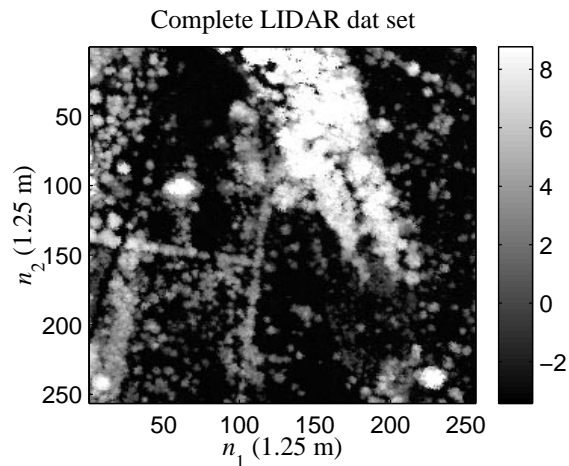


Fig. 1. Calculated vegetation heights from a dense set of LIDAR data.

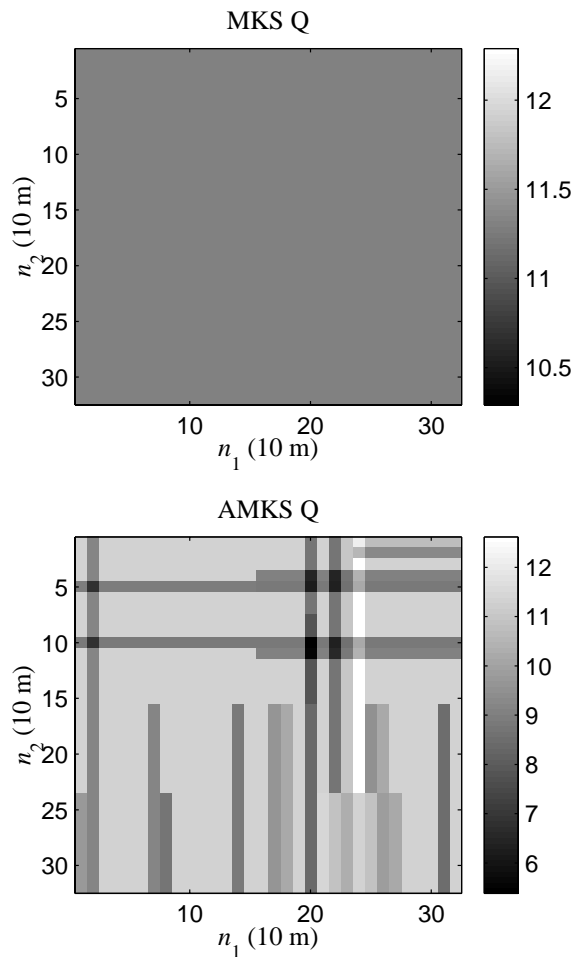


Fig. 2. $Q(m, i, j)$ at $m = m_c$ in the MKS algorithm and AMKS algorithm.

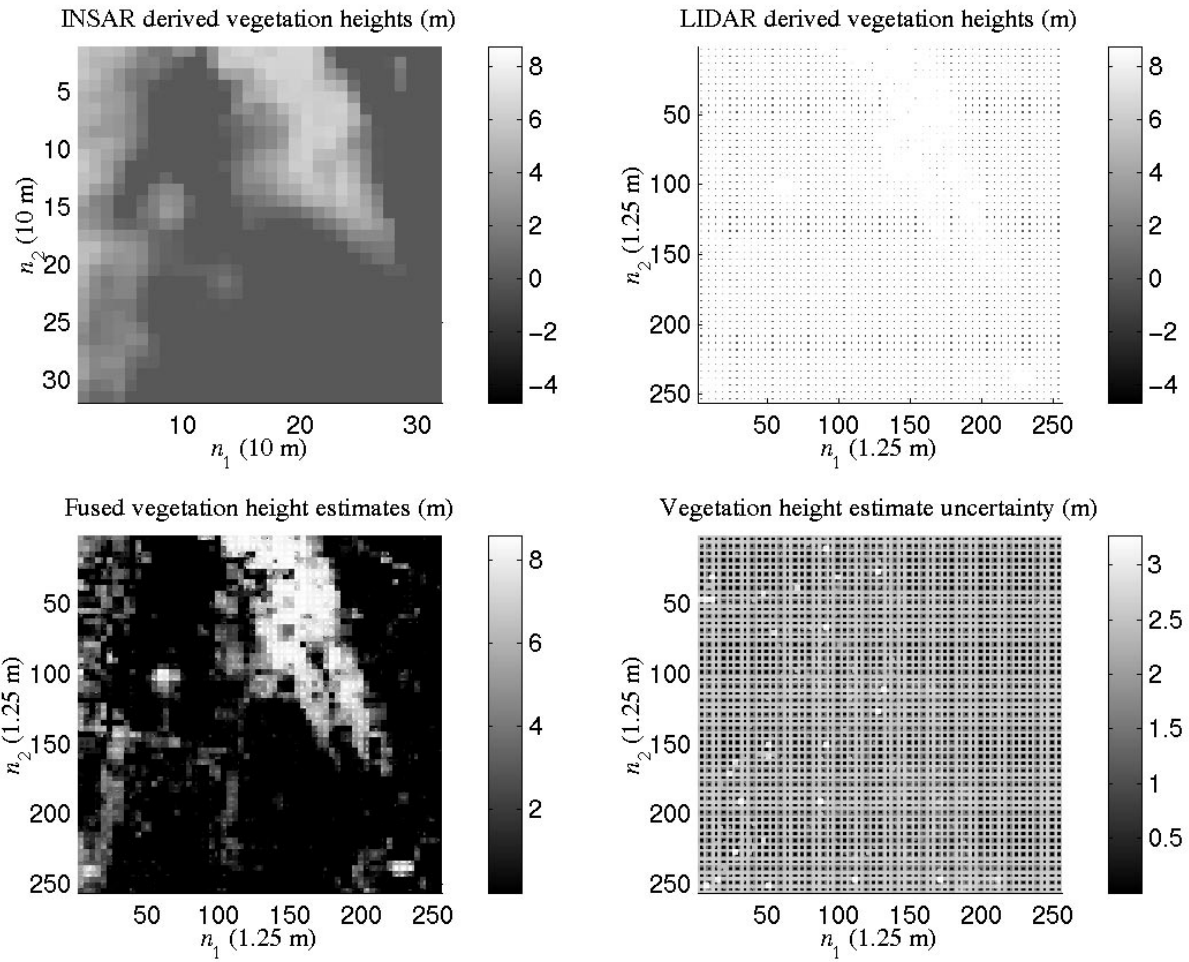


Fig. 3. AMKS output for fused estimates of vegetation heights.