Published in
Proceedings of the Colloque interdisciplinaire en instrumentation, C2I, 443-450, 1998
which should be used for any reference to this work

Application of Kalman filtering to noise reduction on microsensor signals

Application de filtres de Kalman à la réduction de bruit de microcapteurs

C. Marselli, D. Daudet, H. P. Amann, F. Pellandini

Institute of Microtechnology, University of Neuchâtel
Rue A.-L. Breguet 2, CH-2000 Neuchâtel, Switzerland
Catherine.Marselli@imt.unine.ch

ABSTRACT. *Due to their low cost and small size, microsensors are of great interest for many applications. However, few are as accurate than their classical forerunners so that their output signals must be corrected. The present research work focuses on an error correction system designed to eliminate perturbations that are random in character. Provided they can be modelled, Kalman filters are well suited: they are used to estimate a desired signal from noisy measurements. In a second step, the basic method is modified to suit also applications where no model of the signal is available. The interest of this second method is outlined in the peculiar case of the offset correction for an angular rate microsensor. Both approaches are presented and illustrated through examples.*

RÉSUMÉ. *L'encombrement réduit et les coûts relativement bas des microcapteurs les rendent intéressants dans de nombreuses applications. Cependant, la plupart sont encore bien moins précis que des capteurs classiques de sorte que des méthodes de correction d'erreurs doivent être développées. Ce papier propose une méthode de réduction des bruits stochastiques basée sur un filtre de Kalman. Une fois qu'un modèle de l'erreur est défini, le filtre permet d'estimer le signal désiré à partir de mesures bruitées. Dans une version modifiée, l'algorithme peut être appliqué alors qu'aucun modèle du signal n'est disponible. L'intérêt de cette deuxième méthode est illustré dans le cas particulier de la correction d'offset pour un gyroscope micro-usiné. Les deux approches sont illustrées par plusieurs exemples.*

KEY WORDS: *stochastic error correction, stochastic error modelling, micromachined sensors, Kalman filtering, noise reduction.*

MOTS-CLÉS: *correction d'erreurs stochastiques, modélisation d'erreurs stochastiques, capteurs micro-usinés, filtres de Kalman, réduction du bruit.*

1. Introduction

Recent progress in semiconductor design and assembly techniques has led to an increasing symbiosis between mechanical and electronic devices. This evolution allowed the advent of microsensors. Due to their low cost and small size, they are of great interest in many applications. However, they are up to now far less accurate than classical devices so that correction methods must be investigated. The present research work focuses on an error correction system designed to eliminate perturbations that are random in character. Provided they can be modelled, a Kalman filter is well suited. It is yet rather used in sensor fusion applications than in correction problems for a single sensor. In the following, its use to estimate a desired signal from noisy measurements is presented and illustrated through simulations. The basic method is modified to suit applications where no model of the signal is available. The interest of this method is outlined in the peculiar case of the offset correction for an angular rate microsensor.

2. Reduction of sensor errors

2.1 Correction methods

Current microsensors suffer from errors due to internal causes (e.g., parasite vibration modes of the proof mass) as well as external causes (e.g., temperature). Since their accuracy is rather low, the problem of dealing with their errors becomes a crucial one whatever the level of accuracy required by the application. The correction methods depend on the characteristics of the perturbations. If the frequency range of the perturbations is different from the bandwidth of the signal, suitable filtering in the frequency domain eliminates undesired signals. This stage is usually necessary but does not suffice to suppress the sensor errors. In most cases, the key idea is to determine a model for the sensor behaviour and thereafter to estimate the error in the measurements with respect to the model.

For deterministic errors, corrections based on known mathematical relationships with respect to the varying parameters (e.g., quadratic dependence of the output of a pressure sensor in function of the temperature) are applied. Other undesired phenomena are random, but they can nevertheless be modelled. The present research work concerns those errors that can be corrected through the use of a Kalman filter. It focuses on the random description of the sensor *offset* (sometimes referred to as bias), which is the output of the sensor when no physical signal is measured. In the following, the output a_m of a sensor is simply described as:

$$a_m = k \cdot a + b$$

where a is the true signal to be measured, k is the sensitivity, and b is the offset.

2.2 Modelling procedure

Models of the sensor offset are determined using some identification procedure [JOH93]. The proposed models are *a posteriori* models since they are derived from empirical data, independently of underlying physical principles. They are also *black-box* models since only the output of the sensor - and not the internal structure - is considered. The first step of the identification procedure consists of designing an experiment to collect observations. Secondly, data are carefully examined and a model structure is selected accordingly. The parameters of the model are then identified using a linear regression approach. Finally, the model is validated through comparison with several batches of empirical data. The residual represents misfits between data and model. On one hand, the presence of any information remaining in the residuals is a clue that the model might be insufficiently complex or otherwise inappropriate. On the other hand, a major objective is to obtain a model of least possible complexity in order to reduce the computational load in subsequent processing. Therefore, despite a lack of fit with the data, the model may be accepted, provided the required accuracy is reached. In the current work, adaptive modelling has not been considered. The models are established during an off-line processing. Once they are validated, they are included in the system description and no more modified, whatever the results of the subsequent processing.

3. Application of Kalman filtering to noise cancellation problems

3.1 Kalman filter equations

The Kalman filter estimates the *state* of a dynamic system having certain types of random behaviour. The system must be described in a *state space* form:

$$x_{k+1} = \Phi_k \cdot x_k + w_k$$
$$z_k = H_k \cdot x_k + v_k$$

x (n x 1) is called the *state vector*. It is composed of any set of variables sufficient to completely describe the unforced motion of a dynamic system. z (l x 1) is called the *observation vector*. It concerns data that can be known through measurements. w_k and v_k are the state and measurement white noise with known covariance matrices Q_k and R_k, respectively. They are mutually not correlated. Φ_k is the *state transition matrix*, H_k the *observation matrix*.

The Kalman filter is based on a recursive algorithm. At time t_k, the optimum combination of measured and estimated results is given by:

$$\hat{x}_k = \hat{x}_k^- + K_k \cdot \left(z_k - H_k \cdot \hat{x}_k^-\right)$$

where $\hat{x}_k^-$ denotes the *a priori* estimate. The 'hat' denotes an estimate and the superscript minus indicates that this is the best estimate prior to assimilating the measurement at t_k. The Kalman filter gain K_k can be written as:

$$K_k = P_k \cdot H_k^T \cdot \left(H_k \cdot P_k^- \cdot H_k^T + R_k\right)^{-1}$$

where the error covariance matrix P_k associated with the optimal estimate is obtained from:

$$P_k = \left(I - K_k \cdot H_k\right) \cdot P_k^-$$

To recursively compute the Kalman filter gain for the next step, the predictions for the state estimate and covariance at the next step are given by:

$$\hat{x}^{-}_{k+1} = \Phi_k \cdot \hat{x}_k$$

$$P^{-}_{k+1} = \Phi_k \cdot P_k \cdot \Phi^T_k + Q_k$$

The derivation of the *extended* Kalman filter allows the estimation of a non-linear system state. Details on these algorithms as well as on state space description of random processes can be found in many textbooks (e. g., [GRE93]).

3.2 Design of the filter

Let us now consider how a Kalman filter is designed in order to separate two random processes in the case where an input combination of both signal and noise is available. To this purpose, the state vector is composed of all the variables describing the models of the random processes. Additional state variables are appended to account for either non-white state or measurement noise. Φ_k and Q_k are straightforwardly deduced from the models. The observation vector is simply the measured signal that is a mix of signal and noise. Therefore, the observation matrix is usually a combination of unit matrices and zeros matrices according to the state vector. Once the system is properly designed, the algorithm of the discrete Kalman filter is used to estimate the signal.

3.2.1 Example: Signal and noise purely random

Here, the signal is assumed an exponentially correlated process (correlation time $\alpha = 0.01$); a single state variable is used to model it:

$$x_{k+1} = e^{-\alpha} x_k + \alpha\sqrt{1 - e^{-2\alpha}}\, w_k$$

The measurement noise is assumed white of strength 0.01. Hence, no additional state variable is needed. The state matrix is reduced to the scalar $e^{-\alpha}$, and the observation matrix equals simply 1. Figure 1 next page shows that the filter is able to separate the signal from the noise.

3.2.2 Example: Signals with deterministic structure

Let us now consider that the signal has a deterministic structure but that some parameters (e.g., slope of a ramp, amplitude of a sine) are unknown. A Kalman filter can be used to estimate them. In the following test, the signal is assumed a ramp. Two state variables are needed: x_1 is the signal and x_2 is the slope. Since the signal is assumed purely deterministic, no white noise is added in the state equations:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_k$$

The corrupting noise is assumed an exponentially correlated process, though any random model (e.g., ARMA) would have been convenient, leading to the addition of a state variable to the state vector. The state matrix is completed accordingly. White noise of strength 0.1 is then superposed and included in the measurement equations. An arbitrary initial value x_2 = 10 for the unknown slope is defined.

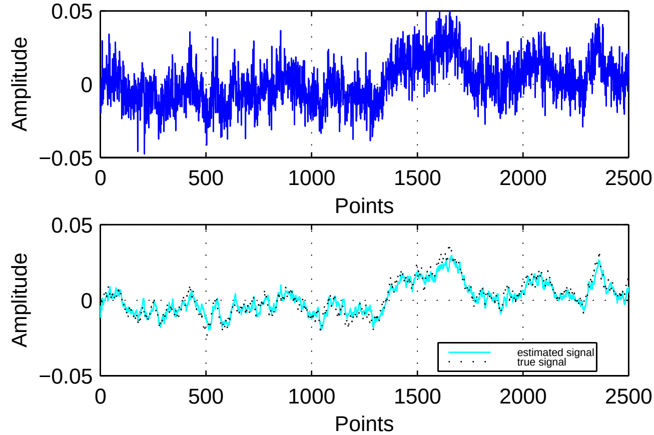

Figure 1. *Estimation of an exponentially correlated process: Measurement (top) - Estimated signal compared with true signal (bottom).*

Figure 2 shows the ability of the filter to separate the ramp from the noise and to recover the true value of the slope (0.001) despite the wrong initial value.

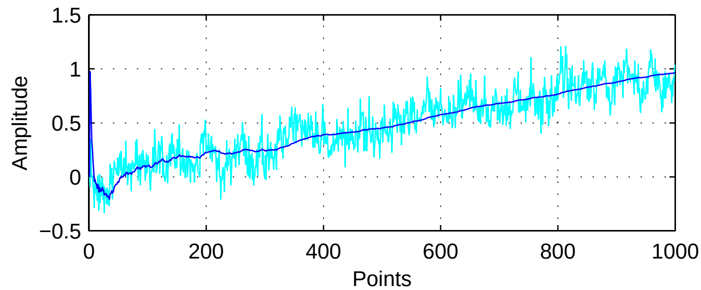

Figure 2. *Estimation of the slope of a deterministic ramp signal: Sum of a ramp, an exponentially correlated process and white noise (grey) - Estimated ramp (black).*

In the second example, the measurement is a combination of two sine waves whose amplitudes are corrupted by white noise. The amplitude of each sine is described as a random constant:

$$x_{k+1} = x_k + w_k$$

Figure 3 shows the ability of the filter to track the value of the sine amplitudes.

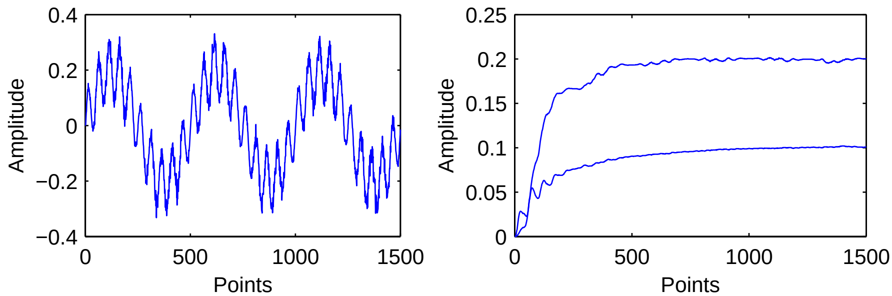

Figure 3. *Sum of two sine waves with noisy amplitudes and their estimations.*

4. Pseudo-model

4.1 Principle

From the previous discussion, it results that the Kalman filter allows recovering a signal from a noisy measurement, provided a model for both signal and noise is defined. The processes can be purely or partially random, stationary or non-stationary. The method has therefore an interest in a wide range of applications, since many types of signals are concerned. However, in some applications involving sensors, no assumptions on the physical signal to be measured may be available, e.g., in navigation problems, where accelerometers and angular rate sensors (gyros) are used to compute the trajectory of a vehicle, though no suitable model of the motion may be definable [BRO72]. Therefore, an alternative design, called *pseudo-model*, has been studied. The principle is briefly explained below, for more detailed equations refer to [MAR97]. The state is chosen as containing sensitivity and offset variations as well as the physical signal:

$$X = \begin{pmatrix} k & b & a \end{pmatrix}^T$$

The number of variables to describe k and b depends on the models found during the identification stage. The main feature of the design is that only one state variable is attributed to the signal, whatever its real shape. The state equation corresponding to the signal is simply obtained from:

$$a = \frac{a_m - b}{k} = f(k, b)$$

Since the function f is non-linear, the equations of the extended filter are used. The observation of the system is still the measurement of the sensor and the non-linear measurement equation is:

$$a_m = k \cdot a + b = h(k, b, a)$$

4.2 Simulation results

The behaviour of the filter using the pseudo-model has been studied through simulations. Sensor measurements are generated as the sum of a true signal and an offset. In the first case shown, the true signal is a step, in the second example, it is the sum of 2 sine waves. The offset is assumed an ARMA noise defined by 2 poles (p_1 = -1, p_2 = 0.5) and 1 zero (z_1 = 0.1). It is described with two state variables:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & 1 \\ -0.5 & 0 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}_k + \begin{pmatrix} 1.1 \\ -0.5 \end{pmatrix} \cdot w_k$$

The sensor offset o is given by the first state variable o_1, the white noise w_k has a variance of $\sigma^2 = 1e^{-2}$, the sensitivity is assumed constant and modelled accordingly with one state variable. Obviously, different models for the offset or for the scale factor would lead to different set of variables. The physical signal, in turn, is always described with one state variable, whatever its nature. Figure 4 shows the ability of the filter to recover the real signal, although no signal model has been provided.

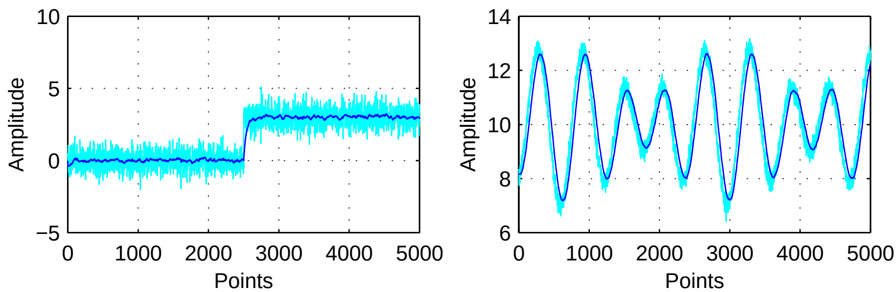

Figure 4. *Estimation of a step shaped (left) and sine shaped signal (right) signal using the pseudo-model approach. The offset is assumed an ARMA signal.*

5. Correction of an angular rate microsensor

The first generation of an in-house designed silicon gyro [PAO96] shows a large offset drift leading quickly to unacceptable result. The error can be described as a low-frequency drift with a superposed noise of higher frequency. Since attempts to find a single model failed, it has been decided to model both effects separately. The unbounded trend is modelled with an integrated random walk, a process obtained by integrating white noise twice:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_k + \begin{pmatrix} 0 \\ \sigma \end{pmatrix} \cdot w_k$$

Once the DC trend of the offset is removed, the model for the residual noise is estimated with an ARMA structure. The Kalman filter corresponding to the pseudo-model design is now used to correct the output of this gyro during a test. The sensor output is recorded during 35 minutes and, in the middle of the experiment, the angular velocity is changed from 0°/s to 100°/s. After correction, the high-frequency noise is attenuated whereas the DC drift is removed; the accuracy is drastically enhanced.

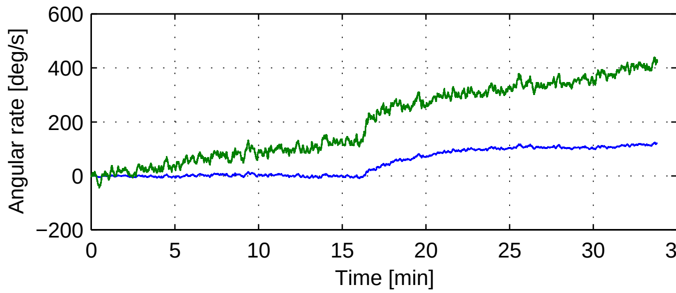

Figure 5. *Correction of the gyro signal for a step change of angular velocity.*

However, a low-pass effect appears when the signal shows a sudden change, due to an adaptation effect of the filter. Tests have shown that the choice of the noise measurement matrix has a straightforward influence on this low-pass effect and that a trade-off has to be found between the attenuation obtained and the adaptation of the filter.

6. Conclusion

A Kalman filter is well suited to eliminate time-dependent sensor errors, which are either partially or totally random, provided adequate state space models could be defined to describe the undesired phenomena. Thus, it has a straightforward interest in microsensor correction. In our design, the state vector contains the models of both the signal and the errors. However, the need for a model for the signal may be too restrictive for some applications. Therefore, an alternative approach, referred to as the pseudo-model, is proposed where the model of the signal to be estimated is derived from the measurement equations. The algorithm was successfully applied for the offset correction of a gyro. However, good working of the filter is tied to the adequacy of the model. Hence, possible extensions of the work should also consider investigations on adaptive models.

7. References

[BRO72] R. G. Brown, "Integrated Navigation Systems and Kalman Filtering: A Perspective", Navigation: Journal of the Institute of Navigation, Vol. 19:4, pp. 355-362, Winter 1972-73.

[GRE93] M. Grewal, "Kalman Filtering: Theory and Practice", Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[MAR97] C. Marselli, "Error Modelling of a Silicon Angular Rate Sensor", Symposium Gyro Technology, Stuttgart (D), pp. 4.0-4.9, September 16-17 1997.

[PAO96] F. Paoletti, "A Silicon Vibrating Gyroscope with Piezoresistive Detection and Electromagnetic Excitation", IEEE MEMS Wokshop 96, San Diego, pp.162-167, 1996.