An Algorithmic Approach to the Optimal Extraction of Signals from Intelligent Sensors

PJ Boltryk*, CJ Harris* and NM White*

* School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK. pb@ecs.soton.ac.uk

ABSTRACT

This paper describes the development of an intelligent sensor architecture, where signal conditioning is performed onboard the sensor itself, in software. Our proposed architecture uses data-based models of the sensor for signal conditioning and fault detection, so that the sensor is robust to degradation and its processed output includes an estimate of uncertainty with each measurement value for higher level sensor management processes such as data fusion. We use a data-based kernel representation for the signal conditioning system, which avoids deriving physical models of the sensor from first principles. A sparse realisation of the kernel model provides fast predictions and opportunities for efficient updating of the sensor model to enable reconfiguration of the sensor model based on incoming data. We show that these techniques have the ability to detect degradation in a MEMS sensor, using elevated temperatures in laboratory conditions.

Keywords: Intelligent sensor, condition monitoring, novelty detection, kernel density estimation

1 INTRODUCTION

An intelligent sensor can be defined as a sensor that incorporates the facility for autonomous on-board signal processing, to enable real-time processes such as fault detection, fault isolation and signal conditioning to occur within the sensor itself. Whereas fault detection and fault identification tasks are traditionally implemented at the higher-level sensor management level, we propose that an intelligent sensor should implement this signal conditioning in software modules at the sensor level as shown in figure 1, using similar approaches as the research on self-validating sensors at the University of Oxford [1], [2].

The motivation behind this approach is that the sensor management can assume higher confidence in output signals from the sensor, thus reducing the requirements for redundancy in sensor networks. Such a sensor should also be able to detect and tolerate drift and ageing effects of the sensing element, which is of considerable importance for devices such as MEMS chemical sensors which are prone to long-term drift due to poisoning.

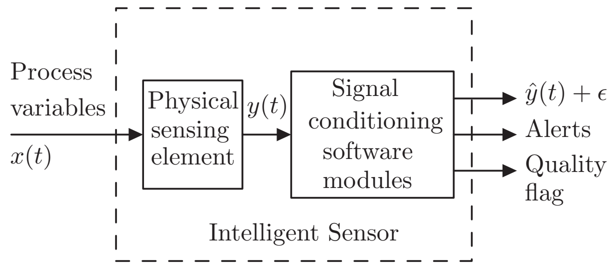

Figure 1: Software modules required for an intelligent sensor

The conditioned output from this sensor should provide a measurement value $\hat{y}(t)$ in all situations, together with an estimate of the uncertainty ϵ attached to the prediction for use in data fusion processes at the sensor management level. Further sensor outputs include a quality flag describing whether the sensory data is valid, and an estimate of the likely reason for poor data.

A generic intelligent sensor architecture can be applied to all types of sensor, with device-specific data-based models performing real-time signal conditioning tasks. Pre-processing converts the raw signals $x(t)$ in the sensor modality (e.g. ^{o}C or acoustic intensity) into an engineering signal, such as voltage, and calibrates the signal (which might be updated periodically to account for drift of the sensor), and provides basic filtering to eliminate noise and thus improve the SNR of the sensor system.

Condition monitoring algorithms compare sensor signals with those predicted by a sensor model. A significant residual error between these two values may be indicative of sensor drift or a fault. Nonlinear data-based models such as support vector machines (SVM) are used to model the sensor for calculating the residual using time series prediction. Unlike analytical models, these techniques use example data to train a model of the system and therefore avoid having to derive a mathematical model of the system from first principles. Due to their sparsity, SVM models can produce parsimonious sensor models, which is advantageous for fast real-time signal processing. The data-based model can also be updated to reconfigure the sensor for drift or ageing mechanisms.

Kernel-based models such as SVM provide direct estimates of prediction uncertainty ϵ, which is a prerequisite for data fusion processes. The same SVM architecture can be used for novelty detection using a density estimation approach, whereby abnormal sensory data can be identified using the distribution of healthy sensor data. In the current study, the kernel-based density estimation algorithms are tested on a MEMS-based accelerometer system that is subjected to a high temperature environment.

2 CONDITION MONITORING OF SENSORY DATA

The use of sensors to perform condition monitoring is commonplace, for example, in mechanical systems to detect the onset of wear in bearings and mechanical components. Such systems may rely on a novelty detection approach, where a feature vector of the current behaviour is compared with the mode of operation during periods of operation when the machinery was in a healthy state. Ypma [3] provides a comprehensive background to condition monitoring using novelty detection of machinery. In a rotational system where accelerometers are used to detect wear in roller bearings, the feature vector may be derived from the frequency spectra of the measured vibration. Degradation in the performance of the bearing can be detected by observing frequency spectra components that are not apparent during trouble-free operation. *A priori* knowledge about specific failure mechanisms of the system and their usual influence on the feature vector signature may enable the condition monitoring system to isolate the device-specific faults once novelty detection has detected abnormal spectra.

The optimal approach to performing condition monitoring of the sensor itself is less obvious. Whereas a rotating machine might have a well-defined frequency spectrum corresponding to normal operation, an ideal accelerometer is expected to sense across a wide range of excitation amplitudes and frequency spectra that are dependent on the vibration source rather than the health of the sensor. Moving the accelerometer to a new environment should therefore not cause the intelligent sensor to predict that it is suffering a fault. Deriving the feature vector for the novelty detection from the frequency spectra alone is likely to lead to high levels of false alarms if the sensor is attached to an unusual vibration source. For a given sensor type, a feature vector is required to be chosen that is representative of a healthy mode of operation, and which highlights the onset of changes in sensor behaviour due to a failure or defect.

2.1 Density estimation

A density estimation scheme is used for the novelty detection process. Density estimation may be defined as the process of estimating the underlying probability density function (pdf) for a set of data observations [4]. In the condition monitoring scenario, this technique provides a method for detecting whether a given set of new data belongs to the same underlying data distribution as the pdf deduced from example data observed during healthy operation. A dataset that fails this hypothesis may be classed as abnormal or novel, which may be representative of a failure mode of the system under consideration.

Given a set of example data during which the sensor is operating in a healthy state, the non-parametric density estimator estimates the data's pdf directly from the data without any necessity for *a priori* information about the distribution. Whilst density estimation using non-parametric kernel techniques such as Parzen windows is well advanced (see for example, [5]), making predictions using this approach is computationally demanding because all training data is used with the model. Therefore, for this application where minimising processing requirements is crucial, density estimation is achieved using SVM which allows for a sparse solution. A further difference is that the density estimation process chosen for this study is to convert the problem to a regression problem closely following the procedure in Chen *et al* [6], briefly summarised below.

If a set of data $\mathbf{x}$ belongs to an underlying data distribution with a true density $p(\mathbf{x})$ then the objective is to estimate this density $\hat{p}(\mathbf{x})$ using (1), subject to the constraints that the weights β should sum to unity and all of the weights should be non-negative. For this work, K is a Gaussian kernel function.

$$\hat{p}(\mathbf{x}) = \sum_{k=1}^{N} \beta_k K(\mathbf{x}, \mathbf{x}_k), \tag{1}$$

From statistics, the distribution function $F(x)$ of a continuous random variable is related to the underlying density function by

$$F(x) = \int_{-\infty}^{x} p(v)\mathrm{d}v. \tag{2}$$

$F(x)$ can be approximated from a sample of data drawn from the distribution by the empirical distribution function (EDF)

$$F(x) \approx f(\mathbf{x}; N) = \frac{1}{N} \sum_{k=1}^{N} \prod_{j=1}^{m} \theta(x_j - x_{j,k}), \tag{3}$$

which is known to be a good approximation of the true cumulative distribution function. $\theta(z)$ is known as the indicator function, and takes a value of 1 if $z > 0$ and 0 otherwise. The weights β are solved using regression of the EDF in the image space, using a mean field method [7]. A particular advantage of this algorithm is that training reduces to a linear programming optimisation, rather than the more usual quadratic programming burden. Fast model updates make it feasible to autonomously reconfigure the sensor's models.

3 DETECTING FAULTS IN MEMS SENSOR

The novelty detection algorithms were tested by subjecting a dual-axis ADXL203 accelerometer to high temperature conditions in an environmental chamber. A sheathed cable was used to transmit vibrations from an external mechanical shaker to the accelerometer, allowing data to be captured at the high temperature conditions without risking damage to the shaker, and limiting the effect of the high temperature on the characteristics of the vibration.

Data corresponding to healthy sensor operation was recorded, for a range of excitation frequencies and across the specified temperature range of the sensor (-40^{o}C to 125^{o}C). Each channel was sampled at 20kHz, and included an anti-alias filter and a high pass filter to remove unwanted mains pickup.

3.1 Feature vector

The feature vector is assigned derived features from a window of the time history recorded for each channel of the dual-axis accelerometer. Three independent features were chosen to represent the operation of the sensor. The first two dimensions include the mean value of the windowed time history for each of the x and y channels. Changes in these values over time indicate a drift in the sensor's behaviour, perhaps causing bias effects. The third dimension is the correlation coefficient between the measured frequency spectra for the x and y channels, over the same time window. The reasoning for this measure is that for a mechanical system, it is highly likely that there is some degree of cross-coupling between accelerations in the x and y directions, which would result in common frequency components in the spectra for the two channels. It is therefore argued that if there is negligible correlation between the vibration spectra for the two channels then at least one of the channels is malfunctioning. For FFT calculation efficiency, all three features were calculated for time windows of 1024 samples.

3.2 Sensor performance beyond 125^{o}C

The accelerometer package is rated to a maximum operating temperature of 125^{o}C, and a maximum storage temperature of 150^{o}C. Referring to figure 2, a strong reduction in sensitivity is apparent once the accelerometer is subjected to temperatures in excess of the 125^{o}C limit. This is most obvious when comparing the peak at 1kHz (corresponding to the fundamental drive frequency

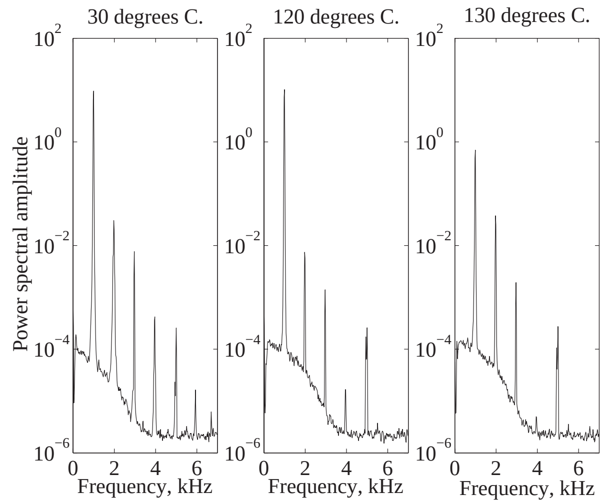

Figure 2: Comparison between the power spectral amplitudes for the accelerometer when subjected to temperatures of 30, 120 and 130^{o}C. Note the strong reduction in sensitivity at the fundamental drive frequency for operation at 130^{o}C.

of the shaker) across the three spectra plots. The spectral amplitude at 1kHz is comparable for both 30^{o}C and 120^{o}C time histories; however, there is a tenfold reduction in the sensitivity when the accelerometer's ambient temperature increases to 130^{o}C, despite there being no change in the acceleration amplitude provided by the shaker system. This reduction in sensor sensitivity is also apparent at higher temperatures, and throughout the operating frequency range. It is not obvious how to use this feature in novelty detection because it is difficult to differentiate between a genuine reduction in source vibration amplitude and a reduction in the sensor's sensitivity. Future work will address this issue because incorporation of adaptive sensor gain to compensate for this loss of sensitivity may provide a useful method for reconfiguring the sensor according to environmental conditions. Further tests are also required to determine whether there is a significant change in the device's linearity at this temperature, which would also be compensated in sensor software.

The environmental chamber's temperature was increased beyond the 150^{o}C temperature storage limit as specified by the manufacturer, in an attempt to accelerate the effect of ageing of the sensor element. Accelerometer data was recorded in 5^o steps, up to the environmental chamber's maximum temperature of 180^{o}C. After maintaining 180^{o}C for 30 minutes, one of the two channels of the accelerometer was observed to change its characteristics dramatically, as shown in figure 3. A malfunctioning electronic component is suspected rather than mechanical damage to the sensing element, because

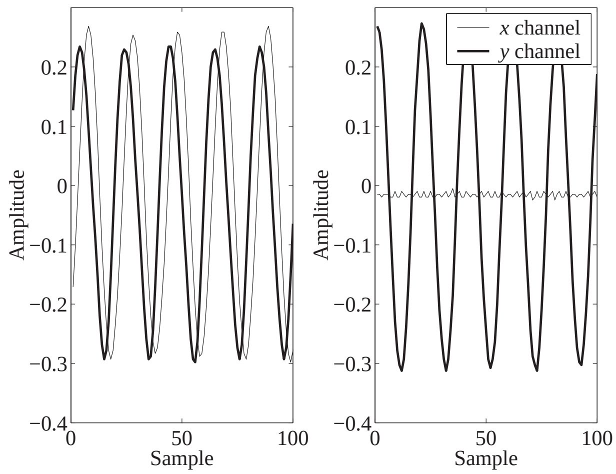

Figure 3: Left-hand plot shows time histories of x and y channels before damage to sensor. Right-hand plot shows same environmental conditions after high temperature damage.

the ADXL203 was subjected to a maximum temperature of only 180^{o}C. Figure 4 plots the density predicted by the density estimation procedure, for the data recorded as the temperature is ramped from 150^{o}C up to 180^{o}C, held at this temperature for 30 minutes, and then cooled to 90^{o}C. A strong change in the density beyond the 604$^{\text{th}}$ dataset corresponds to the beginning of data after the 30 minute hold at 180^{o}C and therefore coincides with the visually observed damage. A software threshold in the sensor can successfully classify this data as abnormal. Since abnormal data is detected for all subsequent datasets, the sensor should recognise that the data is not rogue, but that the sensor is partially damaged.

4 CONCLUSIONS

A generic architecture for intelligent sensors has been proposed that is applicable for all types of sensors. Software modules in the sensor perform tasks such as condition monitoring, providing the sensor management with more reliable measurements. Provision of estimates of measurement uncertainty make sensory data useful for data fusion processes. Data-based models are selected for the sensor models within the software modules, because these avoid physical models of the sensor which are generally difficult to derive and limited to linear behaviour. An SVM model is chosen in the current work, which offers parsimonious sensor models, and low processing requirements for predictions and periodic updates of the model to allow for changes in sensor behaviour over time. Density estimation for novelty detection uses an SVM approach formulated as a regression problem. The algorithms have been successfully tested on a MEMS accelerometer that suffered an electronic fault whilst being subjected to high temperatures.

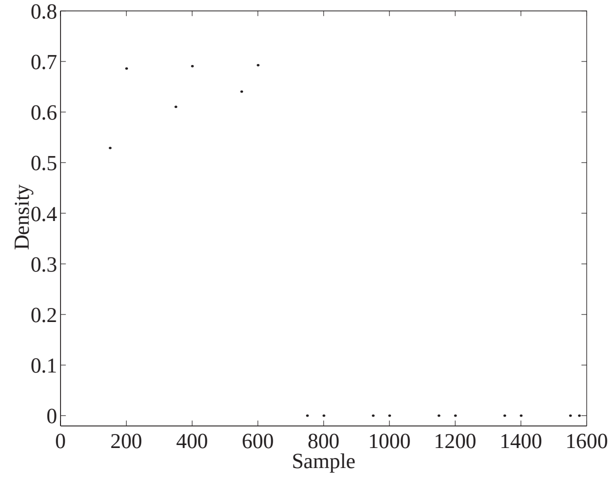

Figure 4: Plot of the density predicted by the SVM-based density estimation for the incoming accelerometer data.

Future work aims to investigate methods for making the novelty detection sensitive to the observed loss in sensitivity of the accelerometer at temperatures beyond the specified operating regime of the sensor. Automatic software reconfiguration using adaptive gain may increase the useful temperature range of this device.

REFERENCES

[1] J Yang and D Clarke, "A Self-Validating Thermocouple", IEEE Trans. on Control Systems Technology, vol. 5, no. 2, March 1997.

[2] M Henry, "Recent developments in self-validating (SEVA) sensors", Sensor Review, vol. 21, no. 1, pp16-22, 2001.

[3] A Ypma, "Learning methods for machine vibration analysis and health monitoring", PhD thesis, Delft University of Technology, 2001

[4] M Desforges, P Jacob and A Ball, "Fault detection in rotating machinery using kernel-based probability density estimation", International Journal of Systems Science, vol.31, no.11, pp1411-1426, 2000,

[5] M Wand and M Jones, "Kernel Smoothing", Chapman and Hall, pp10-110, 1995.

[6] S Chen, X Hong and C Harris. "Sparse kernel density construction using orthogonal forward regression with leave-one-out test score and local regularization", IEEE Trans. Systems, Man, and Cybernetics Part B, vol 34, issue. 4, pp1708-1717, 2004.

[7] J Gao, S Gunn and C Harris "Mean field method for the support vector machine regression", Neurocomputing, vol 50, pp 391-405, 2003.