Mechanical Systems and Signal Processing 40 (2013) 208–221

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

# Detection, identification, and quantification of sensor fault in a sensor network

Jyrki Kullaa *

*Department of Applied Mechanics, Aalto University School of Engineering, P.O. Box 14300, FI-00076 Aalto, Finland*

ARTICLE INFO

*Article history:*
Received 24 August 2012
Received in revised form
8 March 2013
Accepted 5 May 2013
Available online 2 June 2013

*Keywords:*
Sensor fault
Structural health monitoring
Minimum mean square error estimation
Likelihood ratio test
Hardware redundancy

ABSTRACT

In structural health monitoring (SHM) and control, the structure can be instrumented with an array of sensors forming a redundant sensor network, which can be utilized in sensor fault diagnosis. In this study, the objective is to detect, identify, and quantify a sensor fault using the structural response data measured with the sensor network. Seven different sensor fault types are investigated and modelled: bias, gain, drifting, precision degradation, complete failure, noise, and constant with noise. The sensor network is modelled as a Gaussian process and each sensor in the network is estimated in turn using the minimum mean square error (MMSE) estimation The sensor fault is identified and quantified using the multiple hypothesis test utilizing the generalized likelihood ratio (GLR). The proposed approach is experimentally verified with an array of accelerometers assembled on a wooden bridge. Different sensor faults are simulated by modifying a single sensor. The method is able to detect a sensor fault, identify and correct the faulty sensor, as well as identify and quantify the fault type.



## 1. Introduction

With emerging sensor technology, an increasing number of sensors can be installed to structures for monitoring and control. Due to the high number of low-cost sensors, sensor faults become more frequent compared to the structure's lifetime. Because the monitoring or control applications utilize the sensor data for decision, it is important that the data acquired are accurate and reliable. A faulty sensor cannot perform its function properly but instead may provide false information for decision, thus making the system unreliable. Therefore, it is necessary to detect such failures and adapt to the new situation with correcting actions. With a high number of sensors, the measurement system is redundant, and removing a sensor will result in no loss of information. This fact can be utilized in detecting, isolating and correcting a faulty sensor. Sometimes it is also important to identify the type and magnitude of sensor fault, for example to investigate the long-term behaviour of the sensor network in monitoring systems, to make a decision upon sensor replacement, to find the origin of the fault, or to design more robust sensors or data acquisition systems. In this paper, an automatic method is proposed to detect and isolate the faulty sensor in a network, and then to identify the type and magnitude of the fault.

A brief review of sensor validation research is given in the following. There are two main approaches: (1) hardware redundancy and (2) analytical redundancy [1]. The first approach uses the fact that several sensors measure the same quantity. The second approach utilizes a mathematical model of the system, for example a finite element model, and the

* Permanent address: Metropolia University of Applied Sciences, P.O. Box 4021, FI-00079 Metropolia, Finland. Tel.: +358 207836176; fax: +358 207836102.
*E-mail address:* jyrki.kullaa@metropolia.fi

0888-3270/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ymssp.2013.05.007

redundancy is provided by the model. Both approaches have their advantages and disadvantages. For hardware redundancy, extra sensors are required, and for analytical redundancy, an accurate mathematical model must be created. In many cases, it is impractical or too expensive to build a precise model using the finite element method or experimental modal analysis. In this paper, we restrict to hardware redundancy.

The sensor network is assumed to monitor a linear time-invariant system, e.g. a structure or a process, with simultaneous sampling. There are basically two types of redundancy: direct or static redundancy and temporal or dynamic redundancy [2]. Static redundancy corresponds to spatial correlation between different sensors at the same time instant, whereas dynamic redundancy takes into account temporal correlation between adjacent time increments. Using hardware redundancy, static redundancy is possible if the number of states of the system is less than the number of sensors. In this paper, the application is typically a structural system, e.g. a bridge. The number of states is roughly equal to the number of active natural modes. If dynamic redundancy is utilized, the number of sensors may be lower, but stationarity of the process is often assumed.

Sensor validation has been studied e.g. in [3,4] using the parity space approach. Other methods include principal component analysis (PCA) [2,5–10], independent component analysis (ICA) [10], factor analysis [11], and minimum mean square estimation (MMSE) [12,13]. Comparisons of different methods have been performed in [10,11]. MMSE has also been applied to damage detection and localization in structures [14] and in distinguishing between different sources of changes in vibration data: environmental influences, sensor faults, and structural damage [15].

The analysis can be based on the parity space approach which uses a measurement model $\mathbf{y}=\mathbf{H}\mathbf{x}$ to generate a set of residuals that are sensitive to faults [3]. More specifically, the anomalies in the data are found from the null-space of $\mathbf{H}^T$. Here $\mathbf{y}$ is a vector of the measurements (sensors), $\mathbf{H}$ is the measurement matrix and $\mathbf{x}$ is a vector of the unknown state variables. For a redundant system, the number of columns in $\mathbf{H}$ must be higher than the number of rows, i.e. the dimension of $\mathbf{y}$ is larger than the dimension of $\mathbf{x}$. The disadvantage of this approach is that the measurement matrix $\mathbf{H}$ is often unknown. One could try to identify it or alternatively use a direct statistical method in which the aforementioned measurement model is not explicitly used. In [2], this type of model was applied using principal component analysis (PCA).

There are few studies identifying the types or magnitudes of sensor fault. The objective of this paper is to identify the most common sensor faults and assess the severity of the fault, or more specifically, estimate the fault parameters of the sensor. In a large sensor network, it would be most useful and informative to automatically infer the type and severity of the fault. For example, some fault types may occur recurrently. The severity of the fault gives information for possible replacement of the sensor.

The primary assumption is that initially all sensors in the network are functioning. Training data are available from this network to build the model. The measurement matrix $\mathbf{H}$ is not explicitly formed.

This paper is organized as follows. The sensor network model is presented in Section 2. This model is then used in Section 3 to detect a sensor fault with a composite hypothesis test using the generalized likelihood ratio test (GLRT). Different sensor fault models are derived in Section 4. An algorithm is proposed in Section 5 to identify and quantify the sensor fault using the multiple hypothesis test. Experimental results are presented in Section 6 to validate the proposed method. Finally, concluding remarks are given in Section 7.

## 2. Sensor network model

The dynamic response $\mathbf{x}$ of a linear system comprises of the modal contribution of the $d$ lowest modes and the static correction term [16]:

$$\mathbf{x}(t)=\sum_{i=1}^{d}\phi_i q_i(t)+\left[\mathbf{K}^{-1}-\sum_{i=1}^{d}\mathbf{F}_i\right]\mathbf{B}f(t) \tag{1}$$

where $\phi_i$ is the mode shape vector of mode $i$ and $q_i(t)$ is the response of mode $i$, The term in the brackets is a constant matrix, where $\mathbf{K}$ is the stiffness matrix of the system and $\mathbf{F}_i$ is

$$\mathbf{F}_i=\frac{\phi_i\phi_i^T}{\phi_i\mathbf{K}\phi_i^T} \tag{2}$$

$f(t)$ is the vector of load amplitude functions and $\mathbf{B}$ is the load distribution, or input, matrix with a number of columns equal to the number of load amplitude functions.

If the load distribution is constant with $\mathbf{B}$ having $r$ columns, the number of states is $d+r$. For example, in case of a single concentrated force, a high number of modes are excited, but if $f(t)$ consists of lower frequencies only, the higher modes respond statically, which can be taken into account with just a single term as $\mathbf{B}$ consists of one column only. The load distribution can also be such that it only excites the lowest modes and the static correction term is negligible. For example, the distribution of wind excitation is often smooth exciting only the lowest modes and consequently the static correction term is small. The wind distribution is also often relatively constant and $\mathbf{B}$ can be approximated with a single column. Traffic excitation may need more sensors as the load distribution is not constant and consequently $\mathbf{B}$ can have several columns. In such a case the sensor network model may not be able to model all the response, but some additional noise may be