Appl. Math. and Comp. Sci., 1993, vol.3, No.3, 399–416

A SURVEY OF ROBUSTNESS PROBLEMS IN QUANTITATIVE MODEL–BASED FAULT DIAGNOSIS

RON J. PATTON*, J. CHEN*

Quantitative model–based fault diagnosis has become a popular issue in safety–critical systems, e.g., aircraft, spacecraft, chemical processes and nuclear plants. The use of dynamic system model information has been widely recognized as an important approach to fault detection and isolation for the case when there are no repeated hardware units. A prerequisite for reliability in model–based fault diagnosis is robust performance with respect to uncertainties. This paper gives a tutorial discussion of the different problems in robustness and surveys the state of the art in robust solutions for quantitative model–based fault diagnosis. The state observer with disturbance de–coupling design has been recommended as a good solution for robustness in fault diagnosis. Further research topics in robust fault diagnosis have also been outlined.

1. Introduction

Modern systems and equipment are often faced with unexpected changes, such as component faults and variations in operating conditions, that tend to degrade the overall system performance. In order to design a reliable, fault–tolerant control system, or to maintain a high level of performance for complex processes, e.g., spacecraft, aircraft, chemical processes and nuclear plants, etc., it is crucial that such changes are detected promptly and diagnosed so that corrective action can be taken to reconfigure the control system and accommodate the change (Frank, 1990; Patton, 1991; Patton *et al.*, 1989; Willsky, 1976).

A monitoring system which has the capability of detecting and locating a fault and diagnosing its characteristics, is called the *fault diagnosis system* (Patton, 1991). Such a system must consist of two main tasks — **fault detection and fault isolation** (FDI). The fault detection task simply consists of making a binary decision — either that something has gone wrong or that everything is fine. If necessary, this is followed by the next step, the fault isolation task — to determine the source of the fault, e.g., which sensor, actuator or system component has failed.

FDI can be achieved using a replication of hardware (e.g. computers, sensors, actuators and other components) in what is known as *hardware redundancy* in which outputs from identical components are compared for consistency (Patton,

* Department of Electronics, University of York, York, YO1 5DD, England

1991; Patton *et al.*, 1989). Alternatively, FDI can be achieved using analytical or functional information (in quantitative rather than qualitative form) about the process being monitored, i.e. based on a mathematical model of the system (Patton, 1991; Patton *et al.*, 1989). This latter approach is known as ***analytical redundancy*** which is sometimes also known as ***quantitative model-based*** FDI. FDI can also be achieved by ***knowledge-based*** approaches (Patton, 1991; Patton *et al.*, 1989) using qualitative (deep or shallow) information associated with heuristic reasoning. A general classification is shown in Figure 1.

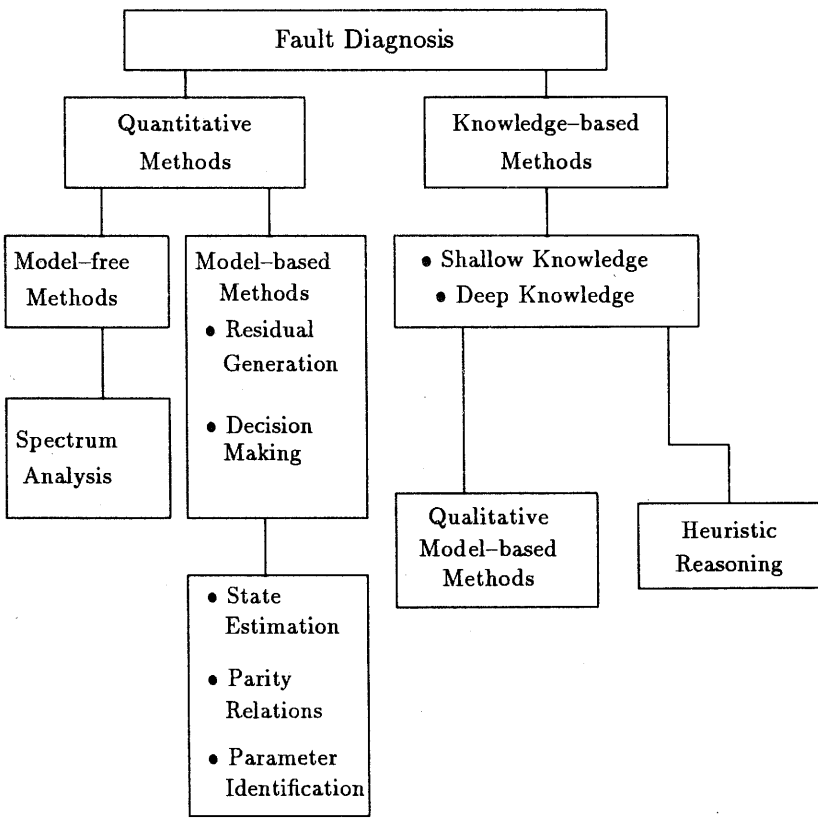

Fig. 1. Classification of FDI methods.

All model–based approaches to fault diagnosis employ mathematical models of the monitored system (Patton *et al.*, 1989). If the model is accurate and the characteristics of all the disturbances are known, FDI can be very straightforward. For a practical system, uncertainties are inevitable and may interfere seriously with diagnosis procedures. To improve the reliability and performance of quantitative model–based FDI, the discrimination of faults from uncertainties must be considered. This is the so–called ***robustness problem with respect to uncertainty*** in *quantitative model-based FDI.* As a general definition, the robustness means the degree to which the detection performance is unaffected by modelling errors and unknown (unmeasured) disturbances (Frank, 1990; Gertler, 1991; Patton, 1991; Patton *et al.*, 1989; Patton and Chen, 1991a).

This paper starts with a description and brief review of quantitative model-based FDI methods, followed by a tutorial discussion of ways in which the observer-based approach can be made robust against modelling errors and uncertainties.

2. The Basic Principles of Quantitative Model-Based FDI

The quantitative model-based FDI approach is based on the comparison of the actual system behaviour and the anticipated system behaviour which is generated by the mathematical model of the system being considered (Gertler, 1991; Willsky, 1976). The FDI process can be considered as two stages (see Fig. 2):

- **Residual generation:** In which, outputs and inputs of the system are processed by an appropriate algorithm (a processor) to generate residual signals.
- **Decision making:** The residuals are examined for the likelihood of faults, and a decision rule is then applied to determine if any fault has occurred. A decision process may consist of a simple threshold test on the instantaneous values or (alternatively) moving averages of the residuals, or methods of statistical decision theory, e.g., likelihood ratio testing and sequential probability ratio testing.

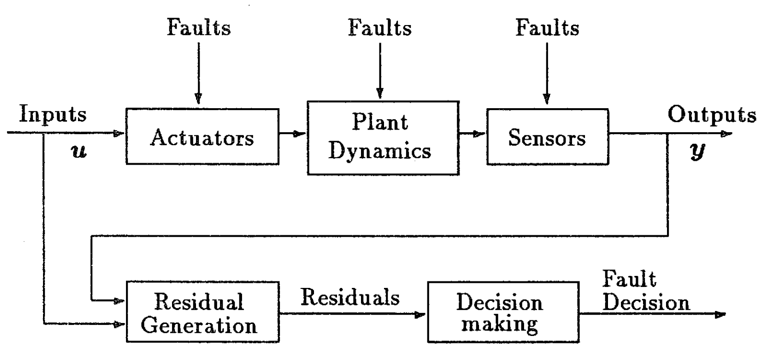

Fig. 2. Two stages structure of the fault diagnosis process.

The residuals are quantities that represent the inconsistency between the actual plant measurements and the mathematical model outputs. Multiple faults may occur and have to be isolated, a set of *structured* residuals is required, so that the different faults are reflected in the residuals in distinct ways and can be discriminated.

The state space (differential equation) model of the dynamic system to be diagnosed can be expressed as:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{R}_1\boldsymbol{f}(t) \tag{1}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t) + \boldsymbol{R}_2\boldsymbol{f}(t) \tag{2}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the state vector, $\boldsymbol{y}(t) \in \mathbb{R}^m$ is the output vector and $\boldsymbol{u}(t) \in \mathbb{R}^r$ is the input (or control) vector. $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ are known matrices with appropriate dimensions, whilst $\boldsymbol{f}(t) \in \mathbb{R}^q$ is a fault vector. Each element $f_i(t)(i = 1, 2, ..., q)$ corresponds to a special fault mode. From a practical point of view it is reasonable *not* to make further assumptions about the fault modes but consider these as *unknown time functions*. The matrices $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are known as *fault entry matrices* which represent the effect of faults on the system. Normally, these matrices are known.

The input–output description of the monitored system is:

$$\boldsymbol{y}(s) = \boldsymbol{G}_u(s)\boldsymbol{u}(s) + \boldsymbol{G}_f(s)\boldsymbol{f}(s) \tag{3}$$

where:

$$\boldsymbol{G}_u(s) = \boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D} \tag{4}$$

$$\boldsymbol{G}_f(s) = \boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{R}_1 + \boldsymbol{R}_2 \tag{5}$$

A traditional way of detecting faults is to use *limit checking* (Frank, 1987; Patton and Chen, 1991a), i.e. to compare plant variables with preset limits, for which the exceedance of a limit can indicate a fault situation. Whilst very simple, this method has a serious drawback, the plant variables vary significantly with different operating states of the plant. The limit thus has to be dependent on the operating state of the plant. Residual testing is a direct development of the limit checking method. The aim here is that residuals generated should be independent of the system operating state under nominal plant operating conditions. In the absence of faults, the residuals are only excited by unmodelled effects, such as parameter errors, noise and disturbances, and are nominally near zero. When a system fault occurs, the residuals deviate from zero in characteristic ways. Residuals are indicators of faults, and are ideally only affected by faults and not by other system changes. Hence, it is possible to use a fixed threshold so that when the residual signal exceeds this threshold a fault can be declared. A typical structure of a model–based residual generator is shown in Figure 3 which involves the processing of the input and output data of the system (Patton and Chen, 1991a).

A general mathematical description of the residual generator can now be expressed as:

$$\boldsymbol{r}(s) = [\boldsymbol{H}_u(s) \;\; \boldsymbol{H}_y(s)] \begin{bmatrix} \boldsymbol{u}(s) \\ \boldsymbol{y}(s) \end{bmatrix} = \boldsymbol{H}_u(s)\boldsymbol{u}(s) + \boldsymbol{H}_y(s)\boldsymbol{y}(s) \tag{6}$$

Here, $\boldsymbol{H}_u(s)$ and $\boldsymbol{H}_y(s)$ are transfer matrices which are realizable using stable linear systems.

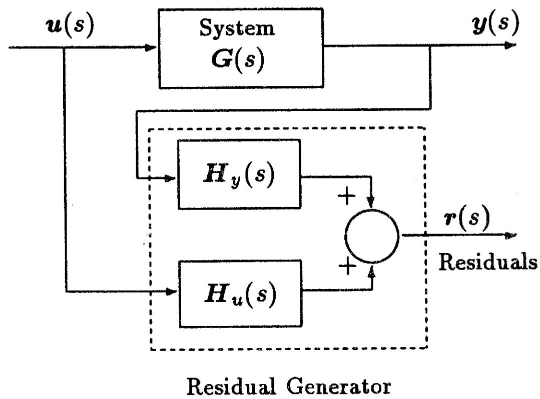

Fig. 3. The general structure of residual generation.

The residual must be independent of the normal operating state of the system. In the fault–free and case of no uncertainty, the residual is zero (or at least very small).

$$r(s) = 0 \text{ and } y(s) = G(s)u(s) \text{ for the fault–free case} \tag{7}$$

For satisfying this requirement, the transfer function matrices $H_u(s)$ and $H_y(s)$ must satisfy the equation:

$$H_u(s) + H_y(s)G_u(s) = 0 \tag{8}$$

Equation (6) is a *unified* and *generalized* representation of all residual generators. The design of the residual generator results simply in the choice of the transfer function matrices $H_u(s)$ and $H_y(s)$ which must satisfy equation (8). Different residual generation methods correspond to different parameterizations of $H_u(s)$ and $H_y(s)$. Observer and parity space residual structures are two examples of the parameterization of $H_u(s)$ and $H_y(s)$. One can obtain different residual generators using different forms for $H_u(s)$ and $H_y(s)$. The desired residual performance can be achieved by suitable designs of $H_u(s)$ and $H_y(s)$.

When faults occur in the monitored plant, the response of the residual vector is:

$$r(s) = H_y(s)G_f(s)f(s) \tag{9}$$

In order to detect the i–th fault in the residual $r(s)$, the i–th column $[H_y(s)\,G_f(s)]_i$ of the transfer function matrix $[H_y(s)\,G_f(s)]$ should be nonzero, and this is especially true for steady–state values, i.e.

$$[H_y(s)\,G_f(s)]_i \neq 0 \text{ and especially } [H_y(0)G_f(0)]_i \neq 0$$

The fault detection problem can then be stated in terms of some decision function, $J(\boldsymbol{r})$, and threshold J_{th} as follows:

$$J(\boldsymbol{r}) \le J_{\text{th}} \qquad \text{for} \qquad \boldsymbol{f}(t) = 0$$

$$J(\boldsymbol{r}) > J_{\text{th}} \qquad \text{for} \qquad \boldsymbol{f}(t) \ne 0$$

Evidently, the ideal case would be $J_{\text{th}} = 0$ which is however, normally impossible because of modelling errors.

The generation of residual quantities is a central issue in quantitative model based FDI schemes. A rich variety of methods proposed for residual generation can be classified as observer or filter based approaches (Frank, 1987), parity relations approaches (Chow and Willsky, 1984) and parameter estimation approaches (Isermann, 1984), as appropriate.

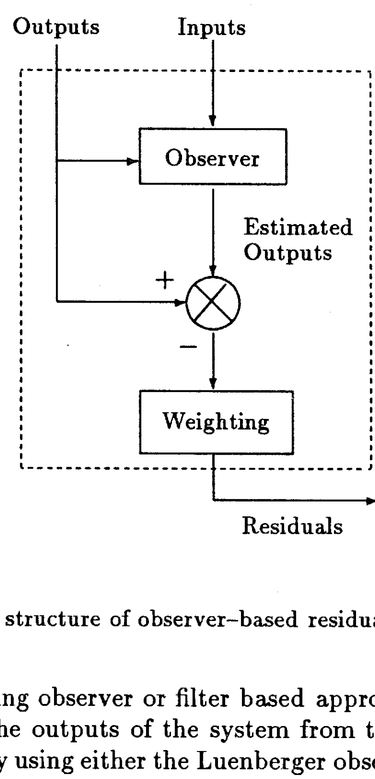

Fig. 4. General structure of observer–based residual generator.

The underlying idea in using observer or filter based approaches (as illustrated in Figure 4) is to estimate the outputs of the system from the measurements (or a subset of measurements) by using either the Luenberger observers in a deterministic setting (Frank, 1987) or the Kalman filter in a stochastic setting (Frank, 1987; Willsky, 1976). Then, the (weighted) output estimation error (or innovations) is used as a residual. The flexibility in selecting observer gains has been fully exploited in the literature yielding a rich variety of FDI schemes (Frank, 1987).

The residual generator based on the observer is described by the following:

$$\dot{\hat{\boldsymbol{x}}}(t) = (\boldsymbol{A} - \boldsymbol{KC})\hat{\boldsymbol{x}}(t) + (\boldsymbol{B} - \boldsymbol{KD})\boldsymbol{u}(t) + \boldsymbol{K}\boldsymbol{y}(t) \tag{10}$$

$$\hat{\boldsymbol{y}}(t) = \boldsymbol{C}\hat{\boldsymbol{x}}(t) + \boldsymbol{D}\boldsymbol{u}(t) \tag{11}$$

$$\boldsymbol{r}(t) = \boldsymbol{W}\boldsymbol{e}_y(t) = \boldsymbol{W}(\boldsymbol{y}(t) - \hat{\boldsymbol{y}}(t)) \tag{12}$$

where $\hat{\boldsymbol{x}}(t) \in \mathbb{R}^n$ is the state estimation vector, $\hat{\boldsymbol{y}}(t) \in \mathbb{R}^m$ the output estimation vector, $\boldsymbol{r}(t) \in \mathbb{R}^p$ is a residual vector and $\boldsymbol{W} \in \mathbb{R}^{p \times m}$ is a weighting matrix.

According to the residual generator structure of Figure 3, the transfer functions $\boldsymbol{H}_u(s)$ and $\boldsymbol{H}_y(s)$ for observer based approaches are:

$$\boldsymbol{H}_y(s) = \boldsymbol{WC}[s\boldsymbol{I} - (\boldsymbol{A} - \boldsymbol{KC})]^{-1}\boldsymbol{K} + \boldsymbol{W} \tag{13}$$

$$\boldsymbol{H}_u(s) = \boldsymbol{WC}[s\boldsymbol{I} - (\boldsymbol{A} - \boldsymbol{KC})]^{-1}(\boldsymbol{B} - \boldsymbol{KD}) + \boldsymbol{WD} \tag{14}$$

3. Robustness Problems in the Quantitative Model–Based Method

Clearly, the reliability of the FDI scheme must be higher than that of the monitored system. However, a model–based approach to FDI will require knowledge (i.e. a model) of the dynamics of the monitored system. The better the model used as a representation of the dynamic behaviour of the system, the better will be the FDI performance; this will be true up to a reasonable limit. The interesting question is just what is a reasonable model to enable good performance in FDI (Frank, 1990; Patton, 1991; Patton *et al.*, 1989; Patton and Chen, 1991a)? Furthermore, when this balance is reached, just what benefits can be gained by using *analytical* compared with *hardware* redundancy? These are amongst the most important challenges to be addressed.

It is clear that the main problem which can obstruct the progress and improvement in reliability of FDI schemes is that of robustness. All model–based methods for FDI involve a residual generation procedure (Frank, 1990; Patton, 1991; Patton *et al.*, 1989; Patton and Chen, 1991a), the residual being a *fault indicator* signal which is zero–valued for the fault–free situation when no uncertainty acts upon the detection system. Both faults and uncertainties affect the residual, and discrimination between these two effects is difficult. The main task in the design of a robust FDI system is thus to generate residuals which are insensitive to uncertainties, whilst at the same time sensitive to faults, and therefore robust.

3.1. The Robustness Problem with Respect to Parameter Uncertainties

A major problem in the field of FDI schemes is caused by uncertainties in the values of physical parameters of the monitored system. In general, model–based methods of FDI are essentially based on the goodness of fit of mathematical models for the monitored system. They therefore depend critically upon the values of the many physical characteristics of the system, such as properties of mass, moments of inertia, electric circuit parameters, aerodynamic or hydrodynamic forces and moments, heat transfer properties, etc. If these are all known with precision the residuals will be accurate and the FDI scheme may display remarkable sensitivity to soft or incipient faults and immunity to false alarm. However, in most systems, even those that are modelled accurately as linear and time invariant (the simplest case of a dynamic system), some physical parameter values are only known approximately. Consequently, the state or parameter estimators must be designed using only nominal values for the uncertain parameters or using some accommodation mechanism to compensate for uncertainty. The resulting residuals are always in error, the severity of the errors depending upon the manoeuvres of the monitored system in ways which are not easily determined.

In the presence of process parameter variations, the state equations of the process can be written as:

$$\dot{\boldsymbol{x}}(t) = (\boldsymbol{A} + \Delta\boldsymbol{A})\boldsymbol{x}(t) + (\boldsymbol{B} + \Delta\boldsymbol{B})\boldsymbol{u}(t) + \boldsymbol{R}_1\boldsymbol{f}(t) \tag{15}$$

$$\boldsymbol{y}(t) = (\boldsymbol{C} + \Delta\boldsymbol{C})\boldsymbol{x}(t) + (\boldsymbol{D} + \Delta\boldsymbol{D})\boldsymbol{u}(t) + \boldsymbol{R}_2\boldsymbol{f}(t) \tag{16}$$

where $\Delta\boldsymbol{A}$, $\Delta\boldsymbol{B}$, $\Delta\boldsymbol{C}$ and $\Delta\boldsymbol{D}$ represent modelling errors in the form of matrix parameter perturbations.

When the residual generator (10)–(12) is applied to the system of equations (15) and (16), the state estimation error and residual equations are:

$$\dot{e}(t) = (\boldsymbol{A} - \boldsymbol{K}\boldsymbol{C})e(t) + \boldsymbol{R}_1\boldsymbol{f}(t) - \boldsymbol{K}\boldsymbol{R}_2\boldsymbol{f}(t) + \Delta\boldsymbol{A}\boldsymbol{x}(t)$$
$$+\Delta\boldsymbol{B}\boldsymbol{u}(t) - \boldsymbol{K}\Delta\boldsymbol{C}\boldsymbol{x}(t) - \boldsymbol{K}\Delta\boldsymbol{D}\boldsymbol{u}(t) \tag{17}$$

$$\boldsymbol{r}(t) = \boldsymbol{W}\boldsymbol{C}e(t) + \boldsymbol{W}\boldsymbol{R}_2\boldsymbol{f}(t) + \boldsymbol{W}\Delta\boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{W}\Delta\boldsymbol{D}\boldsymbol{u}(t) \tag{18}$$

Now, the terms $\Delta\boldsymbol{A}\boldsymbol{x}(t)$, $\Delta\boldsymbol{B}\boldsymbol{u}(t)$, $\Delta\boldsymbol{C}\boldsymbol{x}(t)$ and $\Delta\boldsymbol{D}\boldsymbol{u}(t)$ will force the residual vector $\boldsymbol{r}(t)$ to wander away from, zero even if no faults occur in the system. Indeed, the effects of parameter uncertainties and faults are mixed up together and it is difficult to distinguish their separate effects on the residual. Clearly, as the perturbation terms are usually not known for a real application, it may be difficult to solve this robustness problem completely. However, it is possible to minimize the effect of the unwanted terms in residuals to yield predominantly the fault information.

3.2. The Robustness Problem with Respect to Estimator Initial Conditions

From equations (10)–(12), we can get the residual

$$\boldsymbol{r}(s) = \boldsymbol{W}\boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A}_c)^{-1}[\boldsymbol{R}_1\boldsymbol{f}(s) - \boldsymbol{K}\boldsymbol{R}_2\boldsymbol{f}(s)]$$
$$+\boldsymbol{W}\boldsymbol{R}_2\boldsymbol{f}(s) + \boldsymbol{W}\boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A}_c)^{-1}\boldsymbol{e}(0) \tag{19}$$

where $A_c = A - KC$. If a fault occurs during the transient phase, nonzero initial conditions $e(0)$ will cause the residual to be non–zero. If $e(0)$ is known, the residual generation problem becomes trivial. However, in general $e(0)$ is unknown so that this is also a difficult robustness problem to solve.

3.3. The Robustness Problem with Respect to Nonlinearities

Most dynamic plants are nonlinear, whilst many behave almost linearly, provided that they are not required to deviate from a narrow region around a nominal operating condition. An FDI scheme based on linear models might be quite satisfactory for these conditions. However, outside of this region of operation the nonlinearities of the plant produce signals which are not modelled accurately by the FDI scheme, and these may then be mis–interpreted as faults. Hence, a fault monitoring system must be tested over the entire operating range of the plant being monitored.

A linear representation is often used to model a non–linear system for small perturbations around the operating point. Consider a *linearization error* $\Delta f(x, u)$ as follows:

$$\dot{x}(t) = Ax(t) + Bu(t) + R_1 f(t) + \Delta f(x, u) \tag{20}$$

$$y(t) = Cx(t) + Du(t) + R_2 f(t) \tag{21}$$

Thus

$$\dot{e}(t) = (A - KC)e(t) + R_1 f(t) - KR_2 f(t) + \Delta f(x, u) \tag{22}$$

$$r(t) = WCe(t) + WR_2 f(t) \tag{23}$$

The linearization error will affect the residuals which may cause false alarm.

3.4. The Robustness Problem with Respect to Disturbances and Noise

Dynamic systems are normally subjected to inputs other than those intended by the system designer. These inputs, called disturbances, are usually random functions originating in the environment, such as fluctuations in the wind. Furthermore, the sensors usually have electronic noise superimposed on their signals. This noise is also random, but originates from a different source and is usually uncorrelated with the disturbances. Most signal processing techniques used by designers to account for random fluctuations assume that noise and (sometimes) disturbances are stationary Gaussian processes having known parameters. If the actual disturbances and noise are non–stationary, non–Gaussian or correlated, then this stochastic approach to FDI will not perform well.

Consider the system:

$$\dot{x}(t) = Ax(t) + Bu(t) + G\xi(t) + R_1 f(t) \tag{24}$$

$$y(t) = Cx(t) + Du(t) + \epsilon(t) + R_2 f(t) \tag{25}$$

where $\xi(t)$ is a disturbance vector and $\epsilon(t)$ represents additive noise on the sensors. Thus, the estimation error and residual equation are:

$$\dot{e}(t) = (A - KC)e(t) + R_1 f(t) - KR_2 f(t) + G\xi(t) - K\epsilon(t) \quad (26)$$

$$r(t) = WCe(t) + WR_2 f(t) + W\epsilon(t) \quad (27)$$

The Kalman filter can be used to estimate the state in an optimal way when the noise statistics are known (in the case of stationary, Gaussian process and measurement noise), this will then minimize the effect of the noise on the residuals. Also, by applying signal–averaging to the residuals, the effect of the noise can be decreased.

3.5. The Robustness Problem with Respect to Unmodelled Dynamics

Most systems can have significantly higher order dynamics than their models. In the design of an FDI scheme, we only use a low order model to approximate the high order system. Hence, the dynamic errors appear in the model.

$$\dot{x}(t) = Ax(t) + Bu(t) + A_1 x_h(t) + R_1 f(t) \quad (28)$$

$$y(t) = Cx(t) + Du(t) + R_2 f(t) \quad (29)$$

where $A_1 x_h(t)$ is a *model reduction error* term.

$$\dot{e}(t) = (A - KC)e(t) + R_1 f(t) - KR_2 f(t) + A_1 x_h(t) \quad (30)$$

$$r(t) = WCe(t) + WR_2 f(t) \quad (31)$$

This robustness problem is valid in model–based approaches in which observers are used as a part of the residual–generation mechanism, particularly when reduced–order models are used (In practice, a reduced–order model of some forms is always in use for a real application).

3.6. The Robustness Problem with Respect to Fault Types

A component in a process can malfunction in many ways according to the following Classification of faults:

(i) According to the *location* at which each fault affects the system:

- sensor fault
- computer fault
- actuator fault
- communication fault
- a fault in a component other than a sensor or actuator (e.g. a fault in a feed–pump or valve)

(ii) According to the type of fault signals:

- bias
- abrupt changes
- drift
- stochastic faults
- slow varying fault

For example, a sensor can suffer a change of scale factor, a bias which may not be constant, a nonlinearity due to wear or friction, a measured value stuck at a particular level within its dynamic range, excess noise, or hysteresis. Some FDI schemes are designed to detect only specific types of faults, and these become more cumbersome as the number of faults in the *fault repertoire* is increased. Clearly, if a malfunction should occur which is not in this repertoire or *fault diagnosis knowledge base*, then the FDI scheme will not recognise it. It is better to have a scheme which detects any fault and identifies the faulty component, even though the specific type of fault is not identified. It is, however, still important to isolate a fault event from events caused by internal process uncertainties or external disturbances.

Fault types mainly affect the decision making process, and this is especially true for stochastic systems. A fault monitor which is robust to fault types can include *hypothesis-generation* and *hypothesis-testing* (Basseville and Benveniste, 1986; Patton *et al.*, 1989). The hypothesis generation procedure is to build up a repertoire of known or hypothesised possible malfunctions or faults in system components or instruments (sensors). The most powerful approaches to hypothesis generation should be based on the combination of qualitative and quantitative reasoning to enhance the advantages of each approach, whilst minimizing their disadvantages.

4. Robustness Solutions for Quantitative Model–Based FDI

4.1. Structured Uncertainty

Up to now, the most powerful and successful ways to achieve robustness in FDI make use of *disturbance de-coupling* ideas (Frank, 1991; Frank and Wünnenberg, 1989; Gertler, 1991; Patton *et al.*, 1987; Patton *et al.*, 1989; Patton and Chen, 1991a; Patton and Chen, 1991b; Patton and Chen, 1992; Watanabe and Himmelblau, 1982). In these approaches, all uncertainties are summarized as disturbance terms acting on the dynamic model. The effect of this uncertainty, on the system can be illustrated as follows.

The state space form:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{E}\boldsymbol{d}(t) + \boldsymbol{R}_1\boldsymbol{f}(t) \tag{32}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t) + \boldsymbol{R}_2\boldsymbol{f}(t) \tag{33}$$

The input–output form:

$$\boldsymbol{y}(s) = \boldsymbol{G}_u(s)\boldsymbol{u}(s) + \boldsymbol{G}_d(s)\boldsymbol{d}(s) + \boldsymbol{G}_f(s)\boldsymbol{f}(s) \tag{34}$$

and

$$\boldsymbol{G}_d(s) = \boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{E} \tag{35}$$

where the term $\boldsymbol{E}\boldsymbol{d}(k)$ is used to represent uncertainties acting upon the system. The disturbance $\boldsymbol{d}(t)$ is unknown, but its distribution matrix $\boldsymbol{E}$ is assumed to be known (i.e. the directions represented by the columns of these matrices are known); this is an example of ***structured uncertainty.*** On substituting equation (34) into equation (6), the s–domain residual vector is:

$$r(s) = H_y(s)G_f(s)f(s) + H_y(s)G_d(s)d(s) \tag{36}$$

If the residual generator has been designed to satisfy:

$$H_y(s)G_d(s) = 0 \tag{37}$$

the disturbance is totally de–coupled from the residual. That is to say, the residual can be made independent of all disturbances to enable robust FDI. This is the principle of ***disturbance de–coupling.*** Disturbance de–coupling designs can be achieved by using ***the unknown input observer*** (Frank and Wünnenberg, 1982, 1989; Watanabe and Himmelblau, 1982) or ***alternatively, eigenstructure assignment*** (Gertler, 1991; Patton and Chen, 1991a, 1991b, 1992; Patton *et al.*, 1987) approaches. As far as the design of robust residuals is concerned, these methods are formally equivalent and use different mathematical tools to achieve the same goal in robustness (Gertler, 1991).

For observer–based residual generation approach, it can be shown that:

$$H_y(s)G_d(s) = WC[sI - (A - KC)]^{-1}E \tag{38}$$

In order to make the residual $r(t)$ to be independent of the uncertainty, it is necessary that:

$$WC[sI - (A - KC)]^{-1}|E = 0 \tag{39}$$

This is a special case of the ***ouput–zeroing problem*** (Patton *et al.*, 1989). Once E is known, the remaining problem is to choose the matrices K and W to satisfy (39), in addition to choosing the suitable eigenvalues to optimize the FDI performance. The eigenstructure assignment is a direct way to design disturbance de–coupled residual generators. By suitable assignment of the eigenstructure of the observer, the residual can be designed to provide disturbance de–coupling (Gertler, 1991; Patton and Chen, 1991a, 1991b, 1992; Patton *et al.*, 1987).

Theorem 1. *If* $WCE = 0$*, and all rows of the matrix* $H = WCE$ *are left eigenvectors of* $(A - KC)$ *corresponding to any eigenvalues, equation (39) holds true.*

Theorem 2. *If* $WCE = 0$*, and all columns of the matrix* E *are right eigenvectors of* $(A - KC)$ *corresponding to any eigenvalues, equation (39) holds true.*

If condition (37) does not hold, perfect (accurate) de–coupling cannot be obtained. One can consider approximate de–coupling. The best that can be done is to find an ***optimal*** solution by minimizing a performance index containing a measure of the effects of both disturbances and faults. An appropriate choice of performance index can thus be defined as:

$$J = \frac{||\boldsymbol{H}_y(j\omega)\boldsymbol{G}_d(j\omega)||}{||\boldsymbol{H}_y(j\omega)\boldsymbol{G}_f(j\omega)||} \tag{40}$$

By minimizing the performance index J for a specified frequency range, an approximate de–coupling design can be achieved (Ding and Frank, 1991; Frank, 1991).

4.2. Unstructured Uncertainty

Clearly, a necessary assumption for disturbance de–coupling is that the disturbance distribution matrix must be known *a priori*. In more general problems, the uncertainty structure is unknown (i.e. the *unstructured* uncertainty) and for such cases, the robustness problems are more difficult to solve. In this case, the system can be described as:

$$\boldsymbol{y}(s) = (\boldsymbol{G}_u(s) + \Delta\boldsymbol{G}_u(s))\boldsymbol{u}(s) + (\boldsymbol{G}_f(s) + \Delta\boldsymbol{G}_f(s))\boldsymbol{f}(s) \tag{41}$$

And, the residual is:

$$\boldsymbol{r}(s) = \boldsymbol{H}_y(s)(\boldsymbol{G}_f(s) + \Delta\boldsymbol{G}_f(s))\boldsymbol{f}(s) + \boldsymbol{H}_y(s)\Delta\boldsymbol{G}_u(s)\boldsymbol{u}(s) \tag{42}$$

For *unstructured* uncertainties as defined by $\Delta\boldsymbol{G}_u(s)$ and $\Delta\boldsymbol{G}_f(s)$, it is very difficult to achieve robust residual generation. One way of achieving this is to obtain an approximate structure for the uncertainties. As this approximate structure is used to design disturbance de–coupling residual generators, the a suitably robust FDI is achievable. Within the framework of International Research on robust fault detection, a generalised approach to obtain the structure of uncertainties has hitherto been lacking. Recent work by Patton and Chen (1991a, 1991c, 1992) shows that, even for a system as complex as a jet engine, the uncertainty distribution can be estimated either for one operating point or over a wide dynamic range of operation of the engine. The results show that, for a complex and very non–linear thermodynamic engine system, a simple estimation procedure can be used to estimate the uncertainty distribution at some operating points in demanded high compressor shaft speed (and hence thrust). Once the matrix $\boldsymbol{E}$ corresponding to individual operating points is known, an optimization procedure can then yield the optimum uncertainty distribution to enable a robust FDI design. This approximate structured description is considered an optimal description of the uncertainty in the jet engine system and can even be updated on–line. Patton and Chen (1991a, 1991c, 1992) have demonstrated that this powerful approach works well over a wide range of operation of a simulated non–linear engine system.

Robust fault diagnosis by robust design of residuals is defined as the class of *active methods* (Gertler, 1991; Patton and Chen, 1991a). By this, we mean that the effect of uncertainties on residuals has been minimized, or on the other hand the effect of faults on residuals has been maximized.

4.3. Adaptive Threshold

Efforts to enhance the robustness of FDI can also be made at the decision–making stage and we call this the *passive approach* (Gertler, 1991; Patton and Chen, 1991a).

Due to the great number of parameter uncertainties, disturbances and noise encountered in practical application, one will rarely find a situation where the conditions for perfect robust residual generation are met. This is especially true for unstructured uncertainties. It is therefore necessary to provide sufficient robustness not only in the residual generation stage but also in the stage of *residual evaluation* (a step in decision–making). The goal of robust residual evaluation is to enable reliable decision–making in the sense that the false alarm and missing alarm rate due to uncertainties of residuals become satisfactorily small. This can be achieved in several ways, e.g. by statistical data processing, averaging, correction, or by finding and using the most effective threshold.

The common approach to fault decision–making is to define the non–zero threshold, at which the decision functions are compared. Normally, fixed thresholds are used. If a decision signal exceeds the threshold, the occurrence of a fault is assumed. If, on the other hand, the decision function remains below the threshold, the monitored process is considered free of faults. The problem with the fixed threshold approach is that the sensitivity to faults will be intolerably reduced if the threshold is chosen too high, whereas the rate of false alarms will be too large when the threshold is chosen too low. The proper choice of the threshold is a delicate problem. The idea is based on the perception that in the case of system uncertainties, the residual fluctuates with the changing system inputs even if no fault occurs. Walker (1989) has proposed the determination of the optimal threshold via Markov theory. In the case of large manoeuvres these changes might be large enough so that there is *no fixed threshold* giving satisfactorily fault detection at a tolerable false alarm rate. In order to increase the robustness in such a situation it is possible to use *adaptive* thresholds (Clark, 1989; Ding and Frank, 1991; Emami–Naeini *et al.*, 1988; Frank, 1991), where thresholds are varied according to the control activity of the plant.

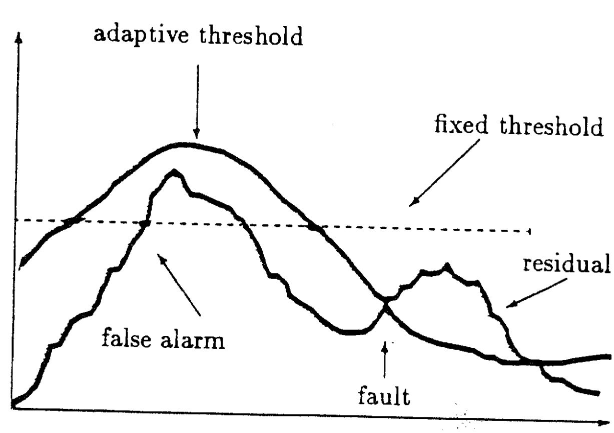

Fig. 5. The adaptive threshold and the residual.

Figure 5 shows the typical shape of an adaptive threshold for direct residual evaluation. The question is how to determine the threshold adaptive law? Clark (1989) used an empirical adaptive law. Emami-Naeini *et al.*, (1988) proposed the ***threshold selector*** method, by which the adaptive threshold can be obtained in a systematic way. The threshold selector presents a new and innovative tool for analysis and synthesis of FDI algorithms. The optimal threshold is shown to be a function of the bound on modelling errors, the noise properties, the speed of the FDI observers, and the types (classes) of reference and fault signals. This approach was generalized by Ding and Frank (1991) and this is shown in Figure 6.

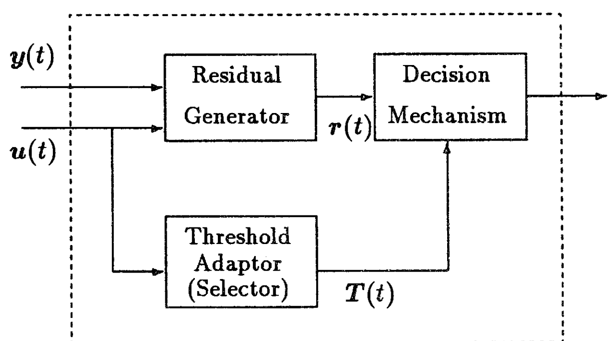

Fig. 6. FDI scheme with threshold adaptor (or selector).

The basis of determining the adaptive threshold is to assume that:

$$\|G_u(s)\| \leq \delta \tag{43}$$

From this assumption and equation (42), we can see that the bound of the fault-free residual is:

$$\|r(s)\| = \|H_y(s)\Delta G_u(s)u(s)\| \leq \|\delta H_y(s)u(s)\| \tag{44}$$

Hence, the adaptive threshold is defined as:

$$T(s) = \delta H_y(s)u(s) \tag{45}$$

We call the adaptive threshold approach a *passive* method for robust FDI (Patton and Chen, 1991a). By this we mean that the reliable decision-making under uncertain residuals situation. The passive approaches to robust FDI design, by their very nature deal with a problem of bounds on uncertainty; bounds are expressed in terms of the thresholds for; these detection required to maintain robustness. Hence, a combination of active and passive approaches can offer potential for real robustness, especially when considering practical applications.

It is believed that the success of an FDI scheme depends on the accurate and appropriate modelling of the monitored process. Hence, some attention to the field and issues of robustness must be paid to ensure that sufficient modelling of the monitored process is achieved.

5. Conclusion

In this paper the robustness problem in quantitative model–based FDI has been discussed. It has been pointed out that the critical issue of FDI is the robustness with respect to uncertainties.

So far the majority of studies in robustness for model–based FDI have focused on Robust Residual–Generation problems. Observer–based approaches to residual generation have stimulated a significant interest in the development of methods related to the robust control problem; this is possible as the observer is the dual of the controller design problem. These active approaches tackle the robustness problem in a more direct way, by considering the real mechanisms of uncertainty (structured or unstructured). De–coupling ideas are used whenever the uncertainty can be considered as structured and this can be achieved for example, by using either eigenstructure design or the so–called unknown input observer.

It is important to stress that most effective robustness methods work well even when the uncertainty structure (e.g. directions of disturbances or unknown inputs) is not known. However, even for these methods some knowledge of uncertainty e.g. in terms of frequency distribution or likely magnitudes and point of excitation in the system must be known *a priori*. For the case of unstructured uncertainty, some authors have used knowledge of the bounds of uncertainty to solve the robustness design problem via frequency domain methods, which are well known in robust control (e.g. H^∞ and related methods). In some cases robustness design methods for structured uncertainty have been applied to systems with unstructured uncertainty using the time–domain technique, for example by estimating the disturbance model through state estimation.

Acknowledgements

The authors acknowledge the funding support for this research from the UK Science & Engineering Research Council and the UK Ministry of Defence (grant ref. No. GR/G2586.3).

References

Basseville M. and Benveniste A. (1986): *Detection of abrupt changes in signals and dynamics systems.* — . LNCIS No.77, Berlin: Springer–Verlag.

Chow E.Y. and Willsky A.S. (1984): *Analytical redundancy and the design of robust detection systems.* — IEEE Trans. Aut. Control, v.AC–29, No.7, pp.603–614.

Clark R.N. (1978): *Instrument fault detection.* — IEEE Trans. Aero. and Electron. Syst., v.AES–14, No.3, pp.456–465.

Clark R.N. (1989): *State estimation schemes for instrument fault detection.* — Chapter 2, In: Patton *et al.*, (Eds.) (1989): *Fault Diagnosis in Dynamic Systems, Theory and Application.* — London: Prentice Hall (Control Engineering Series), pp.22–45.

Ding X. and Frank P.M. (1991): *Frequency domain approach and threshold selector for robust model-based fault detection and isolation.* — Proc. IFAC/IMACS Symp. *SAFEPROCESS'91*, Baden-Baden, Germany, Sept 10–13, v.1, pp.307–312.

Emami–Naeini A. E., Akhter M.M. and Rock S.M. (1988): *Effect of model uncertainty on failure detection: the threshold selector.* — IEEE Trans. Aut. Control, v.AC–33, No.2, pp.1106–1115.

Frank P.M. (1987): *Fault diagnosis in dynamic system via state estimation – A survey.* — In: System Fault Diagnostics, Reliability and Related Knowledge–Based Approaches, Tzafestas *et al.*, (Eds.), v.1, pp.35–98, Dordrecht, D. Reidel Press.

Frank P.M. (1990): *Fault Diagnosis in dynamic system using analytical and knowledge based redundancy – A survey and some new results.* — Automatica, v.26, No.3, pp.459–474.

Frank P.M. (1991): *Enhancement of robustness in observer–based fault detection.* — Proc. IFAC/IMACS Symp. *SAFEPROCESS'91*, Baden–Baden, Germany, Sept 10–13, v.1, pp.275–287.

Frank P.M. and Wünnenberg J. (1989): *Robust fault diagnosis using unknown input schemes.* — In: Patton *et al.*, (Eds.), (1989), *Fault Diagnosis in Dynamic Systems, Theory and Application.* — London: Prentice Hall (Control Engineering Series), pp.47–98.

Gertler J. (1991): *Analytical redundancy methods in failure detection and isolation.* — Proc. IFAC/IMACS Symp. *SAFEPROCESS'91*, Baden–Baden, Germany, Sept.10–13, v.1, pp.9–21.

Isermann R. (1984): *Process fault detection based on modelling and estimation methods: A survey.* — Automatica, v.20, No.4, pp.387–404.

Patton R.J. (1991): *Fault detection and diagnosis in aerospace systems using analytical redundancy.* — IEE Computing and Control Engineering J., v.2, No.3, pp.127–136.

Patton R.J. and Chen J. (1991a): *A review of parity space approaches to fault diagnosis.* — Proc. IFAC/IMACS Symp. *SAFEPROCESS'91*, Baden–Baden, Germany, Sept.10–13, v.1, pp.239–255.

Patton R.J. and Chen J. (1991b): *Optimal selection of unknown input distribution matrix in the design of robust observers for fault diagnosis.* — Proc. IFAC/IMACS Symp. *SAFEPROCESS'91*, Baden–Baden, Germany, Sept.10–13, v.1, pp.221–226.

Patton R.J. and Chen J. (1991c): *Detection of faulty sensors in aero jet engine systems using robust model-based methods.* — IEE Colloquium on "Condition Monitoring for Fault Diagnosis", 18 October.

Patton R.J. and Chen J. (1992): *Robust fault detection of jet engine sensor systems by using eigenstructure assignment.* — Proc. *AIAA Guidance, Navigation and Control Conf.*, New Orleans, August, AIAA–91–2797, 1991, Also in: J. of Guidance, Control and Dynamics, v.15, No.6, pp.1491–1497.

Patton R.J., Frank P.M. and Clark R.N. (Eds.) (1989): *Fault Diagnosis in Dynamic Systems, Theory and Application.* — London: Prentice Hall (Control Engineering Series).

Patton R.J., Willcox S.W. and Winter S.J. (1987): *A parameter insensitive technique for aircraft sensor fault analysis.* — AIAA J. of Guidance, Control and

Dynamics, v.10, No.3, pp.359–367, (Translated into Russian and re–published, In: J. of Aeronautics and Space Technology, 82–91, August 1988).

Walker B. (1989): *Fault detection threshold determination using markov theory.* — Chapter 14, In: Patton *et al.*, (Eds.) (1989): *Fault Diagnosis in Dynamic Systems, Theory and Application.* — London: Prentice Hall (Control Engineering Series), pp.477–508.

Watanabe K. and Himmelblau D.M. (1982): *Instrument fault detection in systems with uncertainties.* — Int. J. Systems Science, v.13, No.2, pp.137–158.

Willsky A.S. (1976): *A survey of design methods for failure detection in dynamic systems.* — Automatica, v.12, No.6, pp.601–611.

Received March 10, 1993
Revised June 24, 1993