1

Introduction

Associated with the increasing demands for higher system performance and product quality on the one side and more cost efficiency on the other side, the complexity and the automation degree of technical processes are continuously growing. This development calls for more system safety and reliability. Today, one of the most critical issues surrounding the design of automatic systems is the system reliability and dependability.

A traditional way to improve the system reliability and dependability is to enhance the quality, reliability and robustness of individual system components like sensors, actuators, controllers or computers. Even so, a fault-free system operation cannot be guaranteed. Process monitoring and fault diagnosis are hence becoming an ingredient of a modern automatic control system and often prescribed by authorities.

Originated in the early 70's, the model-based fault diagnosis technique has developed remarkably since then. Its efficiency in detecting faults in a system has been demonstrated by a great number of successful applications in industrial processes and automatic control systems. Today, model-based fault diagnosis systems are fully integrated into vehicle control systems, robots, transport systems, power systems, manufacturing processes, process control systems, just to mention some of the application sectors.

Although developed for different purposes by means of different techniques, all model-based fault diagnosis systems are common in the explicit use of a *process model*, based on which *algorithms* are implemented for *processing data* that are on-line collected and recorded during the system operation.

The major difference between the model-based fault diagnosis schemes lies in the form of the adopted process model and particular in the applied algorithms. There exists an intimate relationship between the model-based fault diagnosis technique and the modern control theory. Furthermore, due to the on-line requirements on the implementation of the diagnosis algorithms, powerful computer systems are usually needed for a successful fault diagnosis. Thus, besides the technological and economic demands, the rapid development of the computer technology and the control theory is another main reason why the model-based fault diagnosis technique is nowadays accepted as a powerful tool to solve fault diagnose problems in technical processes.

Among the existing model-based fault diagnosis schemes, the so-called observer-based technique has received much attention since 90's. This technique has been developed in the framework of the well-established advanced control theory, where powerful tools for designing observers, for efficient and reliable algorithms for data processing aiming at reconstructing process variables, are available. The focus of this book is on the observer-based fault diagnosis technique and the related topics.

1.1 Basic concepts of fault diagnosis technique

The overall concept of fault diagnosis consists in the following three essential tasks:

- *Fault detection*: detection of the occurrence of faults in the functional units of the process, which lead to undesired or intolerable behavior of the whole system
- *Fault isolation*: localization (classification) of different faults
- *Fault analysis or identification*: determination of the type, magnitude and cause of the fault.

A fault diagnosis system, depending on its performance, is called FD (for fault detection) or FDI (for fault detection and isolation) or FDIA (for fault detection, isolation and analysis) system, whose outputs are correspondingly alarm signals to indicate the occurrence of the faults or classified alarm signals to show which fault has occurred or data of defined types providing the information about the type or magnitude of the occurred fault.

The model-based fault diagnosis technique is a relatively young research field in the classical engineering domain *technical fault diagnosis*, its development is rapid and currently receiving considerable attention. In order to explain the essential ideas behind the model-based fault diagnosis technique, we first give a rough classification of the technical fault diagnosis technique, as sketched in Fig.1.1, and briefly review some traditional fault diagnosis schemes and their relationships to the model-based technique.

- *Hardware redundancy based fault diagnosis:* The core of this scheme, as shown in Fig.1.2, consists in the reconstruction of the process components using the identical (redundant) hardware components. A fault in the process component is then detected if the output of the process component is different from the one of its redundancy. The main advantage of this scheme is its high reliability and the direct fault isolation. The use of redundant hardware results in, on the other hand, high costs and thus the application of this scheme is only restricted to a number of key components.

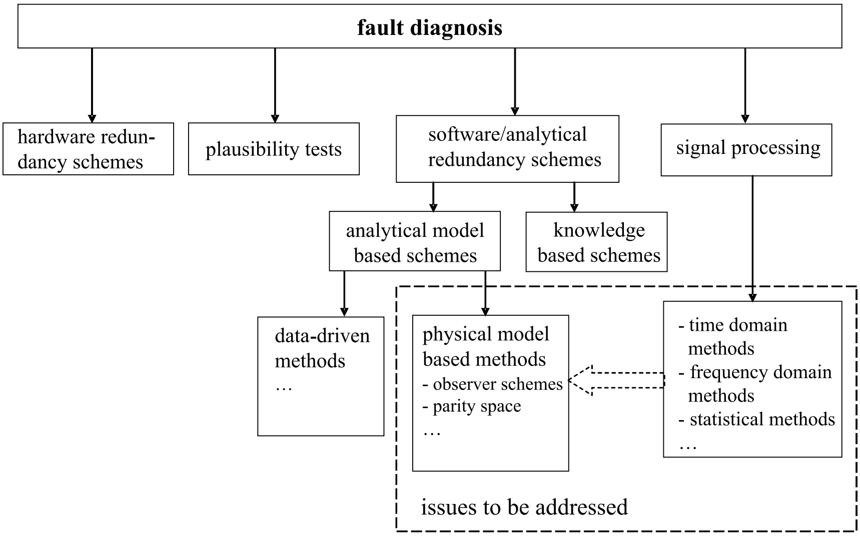

Fig. 1.1 Classification of fault diagnosis methods

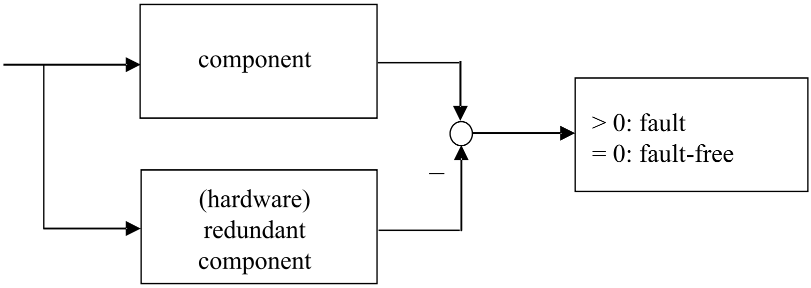

Fig. 1.2 Schematic description of the hardware redundancy scheme

- *Signal processing based fault diagnosis:* On the assumption that certain process signals carry information about the faults of interest and this information is presented in form of symptoms, a fault diagnosis can be achieved by a suitable signal processing. Typical symptoms are time domain functions like magnitudes, arithmetic or quadratic mean values, limit values, trends, statistical moments of the amplitude distribution or envelope, or frequency domain functions like spectral power densities, frequency spectral lines, ceptrum, etc. The signal processing based schemes are mainly used for those processes in the steady state, and their efficiency for the detection of faults in dynamic systems, which are of a wide operating range due to the possible variation of input signals, is considerably limited. Fig.1.3 illustrates the basic idea of the signal processing schemes.

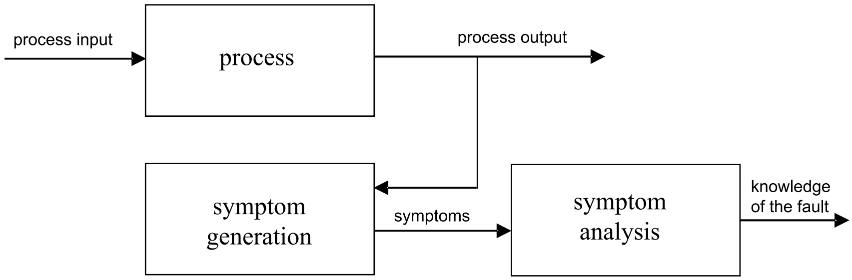

Fig. 1.3 Schematic description of the signal processing based scheme

- *Plausibility test:* As sketched in Fig.1.4, the plausibility test is based on the check of some simple physical laws under which a process component works. On the assumption that a fault will lead to the loss of the plausibility, checking the plausibility will then provide us with the information about the fault. The plausibility test is limited in its efficiency for detecting faults in a complex process or for isolating faults.

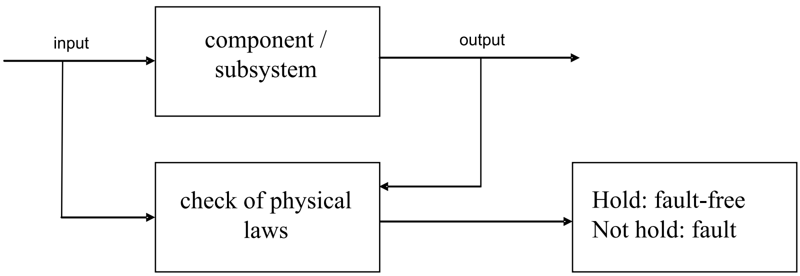

Fig. 1.4 Schematic description of the plausibility test scheme

The intuitive idea of the model-based fault diagnosis technique is to replace the hardware redundancy by a process model which is implemented in the software form on a computer. A process model is a quantitative or a qualitative description of the process dynamic and steady behavior, which can be obtained using the well-established process modelling technique. In this way, we are able to reconstruct the process behavior on-line, which, associated with the concept of hardware redundancy, is called *software redundancy concept.* Software redundancies are also called *analytical redundancies.*

Similar to the hardware redundancy schemes, in the framework of the software redundancy concept the process model will run in parallel to the process and be driven by the same process inputs. It is reasonable to expect that the re-constructed process variables delivered by the process model will well follow the corresponding real process variables in the fault-free operating states and show an evident derivation by a fault in the process. In order to receive this information, a comparison of the measured process variables (output signals) with their estimates delivered by the process model will then be made. The difference between the measured process variables and their estimates is called *residual.* Roughly speaking, a residual signal carries the most important message for a successful fault diagnosis:

$$\text{if residual } \neq 0 \text{ then fault, otherwise fault-free.}$$

The procedure of creating the estimates of the process outputs and building the difference between the process outputs and their estimates is called *residual generation.* Correspondingly, the process model and the comparison unit build the so-called *residual generator,* as shown in Fig.1.5.

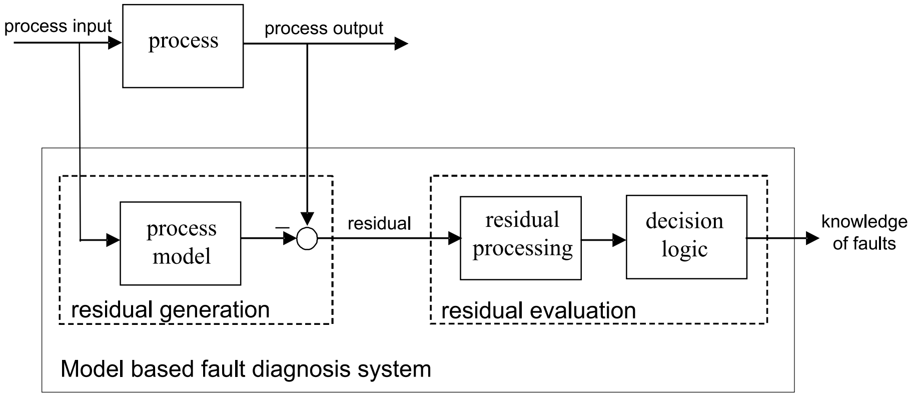

Fig. 1.5 Schematic description of the model-based fault diagnosis scheme

Residual generation can also be considered as an extended plausibility test, where the plausibility is understood as the process input-output behavior and modelled by an input-output process description. As a result, the plausibility check can be replaced by a comparison of the real process outputs with their estimates.

Since no technical process can be modelled exactly and there often exist unknown disturbances, in the residual signal the fault message is corrupted with model uncertainties and unknown disturbances. Moreover, fault isolation and identification require an additional analysis of the generated residual to distinguish the effects of different faults. A central problem with the application of model-based fault diagnosis technique can be expressed as filtering/extracting the needed information about the faults of interests from the residual signals. To this end, two different strategies have been developed:

- designing the residual generator to achieve a decoupling of the fault of interests from the other faults, unknown disturbances and model uncertainties

- extracting the information about the fault of interests from the residual signals by means of post-processing of the residuals. This procedure is called *residual evaluation.*

The first strategy has been intensively followed by many of the research groups working on model-based fault diagnosis technique. One of the central schemes in this area is the so-called *observer-based fault diagnosis* technique, which is also the focus of this book. The basic idea behind the development of the observer-based fault diagnosis technique is to replace the process model by an observer which will deliver reliable estimates of the process outputs as well as to provide the designer with the needed design freedom to achieve the desired decoupling using the well-established observer theory.

In the framework of residual evaluation, the application of the signal processing schemes is the state of the art. Among a number of evaluation schemes, the *statistical methods* and the so-called *norm based evaluation* are the most popular ones which are often applied to achieve optimal post-processing of the residual generated by an observer. These two evaluation schemes are common in that both of them create a bound, the so-called *threshold,* regarding to all possible model uncertainties, unknown inputs and the faults of no interests. Exceeding the threshold indicates a fault in the process and will release an alarm signal.

Integrated application of the both strategies, as shown in Fig.1.3 as well as in Fig.1.5, marks the state of the art of the model and observer-based fault diagnosis technique.

1.2 Historical development and some relevant issues

The study on model-based fault diagnosis began in the early 1970s. Strongly stimulated by the newly established observer theory at that time, the first model-based fault detection method, the so-called failure detection filter, was proposed by Beard and Jones. Since then, the model-based FDI theory and technique went through a dynamic and rapid development and is currently becoming an important field of automatic control theory and engineering. As shown in Fig.1.6, in the first twenty years, it was the control community that made the decisive contribution to the model-based FDI theory, while in the last decade, the trends in the FDI theory are marked by enhanced contributions from

- the computer science community with knowledge and qualitative based methods as well as the computational intelligent techniques
- the applications, mainly driven by the urgent demands for highly reliable and safe control systems in the automotive industry, in the aerospace area, in robotics as well as in large scale, networked and distributed plants and processes.

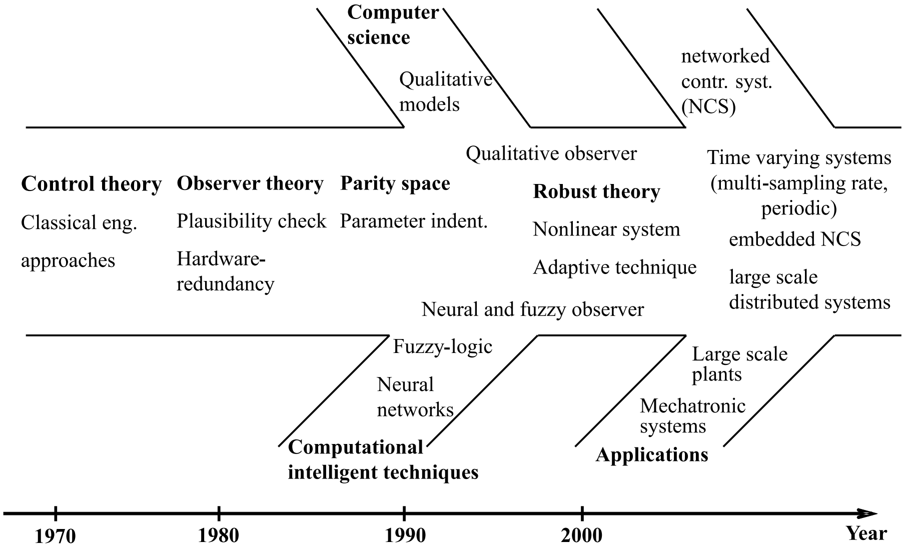

Fig. 1.6 Sketch of the historic development of model-based FDI theory

In the first decade of the short history of the model-based FDI technique, various methods were developed. During that time the framework of the model-based FDI technique had been established step by step. In his celebrated survey paper in *Automatica* 1990, Frank summarized the major results achieved in the first fifteen years of the model-based FDI technique, clearly sketched its framework and classified the studies on model-based fault diagnosis into

- observer-based methods
- parity space methods and
- parameter identification based methods.

In the early 90's, great efforts have been made to establish relationships between the observer and parity relation based methods. Several authors from different research groups, in parallel and from different aspects, proven that the parity space methods lead to certain types of observer structures and are therefore structurally equivalent to the observer-based ones, even though the design procedures differ. From this viewpoint, it is reasonable to include the parity space methodology in the framework of the observer-based FDI technique. The interconnections between the observer and parity space based FDI residual generators and their useful application to the FDI system design and implementation build one of the central topics of this book. It is worth to point out that both observer-based and parity space methods only deal with residual generation problems.

In the framework of the parameter identification based methods, fault decision is performed by an on-line parameter estimation, as sketched in Fig.1.7.

In the 90's, there was an intensive discussion on the relationships between the observer and parameter estimation FDI schemes. Comparisons between these two schemes have been made on different benchmarks. These efforts lead to a now widely accepted point of view that both schemes have advantages and disadvantages in different respects, and there are arguments for and against each scheme.

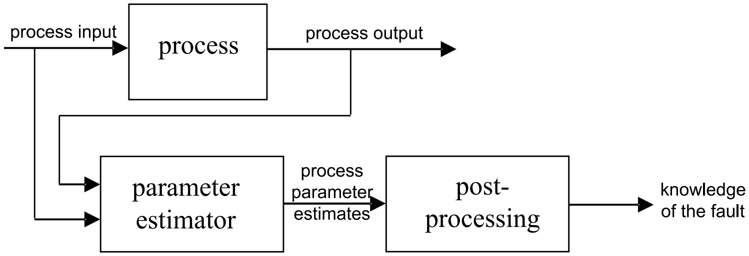

Fig. 1.7 Schematic description of the parameter identification scheme

It is interesting to notice that the discussion at that time was based on the comparison between an observer as residual generator and an parameter estimator. In fact, from the viewpoint of the FDI system structure, observer and parameter estimation FDI schemes are more or less common in residual generation but significantly different in residual evaluation. The residual evaluation integrated into the observer-based FDI system is performed by a feedforward computation of the residual signals, as shown in Fig.1.5, while a recursive algorithm is used in the parameter estimation methods to process the residual signals aiming at a parameter identification and the resulted parameter estimates are further fed back to the residual generator, as illustrated in Fig.1.8. Viewing from this aspect, the parameter identification based fault diagnosis system is structured in a feedback closed-loop, and in against the observer-based FD system is open-loop structured.

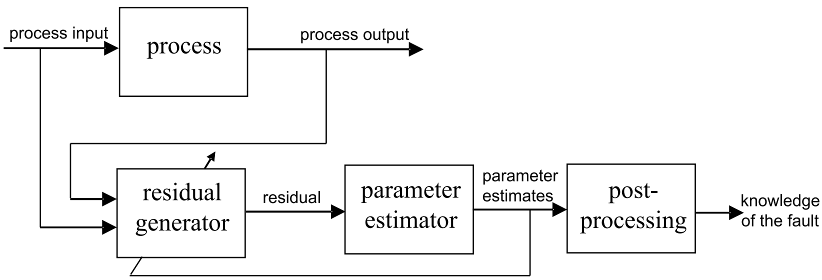

Fig. 1.8 An alternative view of the parameter identification scheme

The application of the well-developed adaptive observer theory to the fault detection and identification in the recent decade is the result of a reasonable combination of the observer-based and parameter identification FDI schemes.

The major difference between the adaptive observer-based and parameter identification FDI schemes lies in the residual generation. In other words, the adaptive observer-based FDI schemes differ from the regular observer-based ones in the way of residual evaluation.

In this book, our focus in on the residual generation and evaluation issues in the framework of the observer and parity space based strategies. Besides of the introduction of basic ideas, special attention will be paid to those schemes and algorithms which are devoted to the analysis, design and synthesis of FDI systems.

1.3 Notes and references

To author's knowledge, the first book on the model-based fault diagnosis technique with a strong focus on the observer and parity space based FDI schemes was published 1989 by Patton et al. [116]. For a long time, it was the only reference book in this area and has made a decisive contribution to the early development of the model-based FDI technique.

The next two monographs, published by Gertler in 1998 [64] and by Chen and Patton in 1999 [21], address different issues of the model-based FDI technique. While [64] covers a wide spectrum of the model-based FDI technique, [21] is dedicated to the robustness issues in dealing with the observer-based FDI schemes. There are numerous books that deal with model-based FDI methods in part, for instance [10, 13, 69] or address a special topic in the framework of the model-based fault diagnosis technique like [100, 133]. In two recent books by Patton et al. [117] and Isermann [81], the latest results on model-based FDI technique achieved in the last decade are well presented.

In the last three decades, numerous survey papers have been published. We divide them into three groups, corresponding to the different development phases of the model-based FDI technique, and give some representative ones from each group:

- introduction and establishment of the observer, parity space and parameter identification based FDI schemes [50, 67, 79, 146]
- robustness issues [51, 52, 55, 114]
- nonlinear, adaptive FDI schemes, application of computational intelligence [53, 90, 140].

Representative study on the relationships between the observer and parity relation based methods can be found, for instance, in [28, 62, 74]. For the comparison study on parameter identification and observer-based FDI schemes the reader is referred to [1, 26, 63].