Optimizing Automated Gas Turbine Fault Detection Using Statistical Pattern Recognition

E. Loukis

K. Mathioudakis

K. Papailiou

Laboratory of Thermal Turbomachines,
National Technical University of Athens,
Athens, Greece

A method enabling the automated diagnosis of gas turbine compressor blade faults, based on the principles of statistical pattern recognition, is initially presented. The decision making is based on the derivation of spectral patterns from dynamic measurement data and then the calculation of discriminants with respect to reference spectral patterns of the faults while it takes into account their statistical properties. A method of optimizing the selection of discriminants using dynamic measurement data is also presented. A few scalar discriminants are derived, in such a way that the maximum available discrimination potential is exploited. In this way the success rate of automated decision making is further improved, while the need for intuitive discriminant selection is eliminated. The effectiveness of the proposed methods is demonstrated by application to data coming from an industrial gas turbine while extension to other aspects of fault diagnosis is discussed.

1 Introduction

Gas turbine engine condition monitoring and fault diagnosis methods have found a wide use among gas turbine users recently. At the same time their philosophy and basic features have undergone a significant change. It is recognized today that such methods no longer rely on simple observations of directly measured quantities. They have become more sophisticated and incorporate processing of some level in order to provide information of direct significance about component conditions or faults. Also economical and practical reasons led to the requirement of as little as possible empirical background from the engine user. There is a tendency to minimize the required skills of the engine operator and to have computer procedures that perform most of the "skilled" work, in order to provide the user with information that needs minimal or no interpretation, and even with hints about required action (Doel, 1990). Expert system environments, available today even for personal computers, contribute to this direction, allowing the integration of data processing and decision making.

Another practical problem recognized by engine users, in the effort to perform efficient monitoring and fault diagnosis, is associated with the restrictions in instrumenting an engine. Installing and employing a large number of measuring instruments is not only difficult and costly but also requires special care in maintaining them and avoiding false alarms. From this point of view, it is desirable to use a minimal set of instruments, with maximum fault detection capabilities. The methodologies introduced in the present paper cover needs related to the above-mentioned desired features and requirements, at the same time being suitable for computerized fault detection.

Contributed by the International Gas Turbine Institute and presented at the 37th International Gas Turbine and Aeroengine Congress and Exposition, Cologne, Germany, June 1-4, 1992. Manuscript received by the International Gas Turbine Institute January 28, 1992. Paper No. 92-GT-29. Associate Technical Editor: L. S. Langston.

It has been established today that dynamic measurements allow the identification of minor gas turbine blade faults, which would remain hidden to performance monitoring techniques. Revealing blade faults, although they may not significantly disturb overall performance, becomes important since it can prevent the sudden occurrence of catastrophic failures (a typical example being the loss of a damaged rotating blade). Loukis et al. (1991) have examined a variety of possible dynamic measurements and discussed their suitability for the detection of blade faults. On the other hand, a set of data from implanted faults on an industrial gas turbine has served as a basis for a first development of automated fault detection methods by Loukis et al. (1992).

The work presented in this paper is based on the observations and the conclusions reached in the above-mentioned works of the authors. The purpose is to derive a methodology that gives the possibility to overcome practical problems encountered in the application of fault diagnosis methods and to provide tools that allow extension of the findings from the particular set of experimental information to other situations as well. In order to achieve these targets, principles that have already been used in other technological domains are for the first time applied to data from faults in an industrial gas turbine and are adapted to particular features of such data. In particular, principles of pattern recognition that have been applied in various other domains, as for example monitoring of machine tools (Emel and Kannatey-Asihm, 1988), are employed.

2 Prior Experience in Gas Turbine Fault Identification Using Measured Patterns

The possibility of establishing fault signatures from dynamic measurement data has been demonstrated by Mathioudakis et al. (1990), from tests performed on an industrial gas turbine and further discussed by Loukis et al. (1991, 1992). The test program included the study of various representative compressor blade faults by performing dynamic measurements with different measuring instruments. For reasons of completeness the description of the experiments is repeated here, in Appendix I. A first procedure for automated identification of faults, based on this data set, has already been described by Loukis et al. (1992). Before describing the present procedure it is useful to summarize the main points of that work. The procedure consists of two phases, the learning phase and the decision phase.

The *learning* phase includes all actions necessary to establish the background for fault identification. For each instrument the following steps are implemented:

(*a*) Measurements are performed on an engine in healthy condition as well as containing the faults of interest.

(*b*) The data sets for each fault are processed and fault indices are calculated from them by appropriate combination with data from healthy condition. Each fault index has the form of a vector $\mathbf{P}=(P_1, P_2, \ldots, P_N)$, whose components come from algebraic combinations of spectra from healthy and faulty engine at shaft rotational frequency harmonics. This vector is termed a reduced spectral difference pattern (or simply pattern).

(*c*) For each fault, a reference pattern is established using the patterns from all the data sets coming from experiments with this particular fault. The reference pattern constitutes a signature of the fault.

The *decision* phase includes all actions necessary for identifying an engine fault. The steps are:

(*a*) A data set is acquired from measurements on the engine, which has to be examined.

(*b*) A pattern is calculated, by appropriate processing of this data set and the corresponding data of healthy condition.

(*c*) The values of a discriminant are calculated for this pattern, with respect to available reference patterns of the faults, established during the learning phase.

(*d*) The diagnostic decision is made by algebraically comparing these discriminant values and classifying to the fault corresponding to the minimum distance.

An example of reference fault patterns derived and used as signatures in this procedure is shown in Fig. 1, where patterns for different instruments used and faults examined during the experimental investigation are shown.

This kind of procedure can be characterized as a "geometric" one, since the diagnostic decision is based on a minimum distance criterion from the reference patterns of the faults. The feasibility of applying this procedure has been demonstrated to be effective, with a certain degree of success for each individual instrument used as source of information.

There are, however, certain disadvantages of such a procedure for practical application in industrial environments:

(*a*) The statistical properties of the data used for establishing the signatures are not taken into account.

(*b*) The parameters used as discriminants are derived intuitively. This does not provide the possibility of improving the effectiveness of the method, unless the user has a better insight in choosing them.

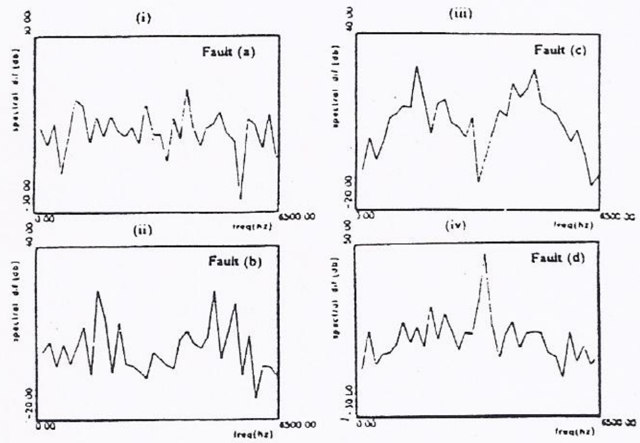

Fig. 1 Fault signatures derived from various dynamic measurements and faults: (*i*) accelerometer A5, (*ii*), (*iii*) pressure transducer PT2, (*iv*) microphone

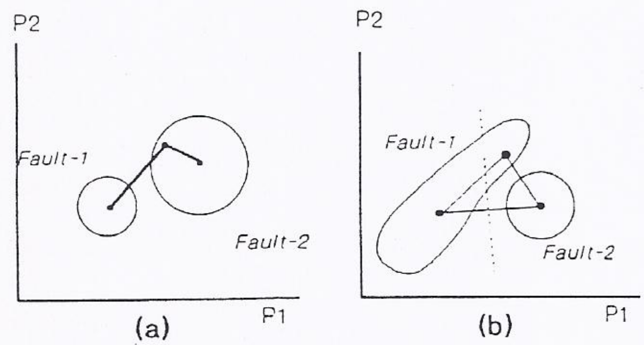

Fig. 2 Qualitative picture for effectiveness of geometric and statistical approaches: (*a*) both approaches succeed; (*b*) geometric procedure fails

(*c*) The conclusion is in a form such that it does not provide the possibility of assigning to it a level of confidence.

(*d*) Use of different discriminants does not lead always to the same conclusion for the kind of fault.

These drawbacks have the practical implication that in some cases false decisions can be made and there are no means of comparing the relative similarities of the examined case to the faults. This can be understood from the situations shown for the simplified case of two-dimensional patterns ($N=2$) in Fig. 2. In case (*a*) the regions occupied by points corresponding to patterns of each fault, in the two components P_1 and P_2 plane, are of such a shape and position, that using a geometric procedure leads for all points to a correct decision. In case (*b*), however, the regions are of such a shape and position, that some points of fault 2 can be assigned to fault 1 by a geometric procedure. Such points are for example the ones contained in the shaded region of the figure. In order to handle such situations successfully, a statistical approach to making the classification decision is required.

Nomenclature

C_i = covariance matrix for fault i, Eq. (5)

D_i = ith optimal linear discriminant

$D_{i,k}$ = discriminant i, calculated with respect to the reduced reference spectral difference pattern of fault k

N_i = the number of independent data sets available for fault i

N = number of components of reduced spectral difference patterns

M = number of faults

pr = conditional probability

$\mathbf{P}_i$ = reference spectral difference pattern for fault i

$\mathbf{X}$ = discriminant vector, Eq. (1)

$\mathbf{X}_{ij}$ = discriminant vector calculated from the jth data set of fault i

3 A Statistical Procedure for Fault Pattern Classification

The main feature of this procedure is that it is based not on a single discriminant calculated with respect to the available fault signatures, but on a vector **X**, which has as components a number of discriminants with respect to available signatures:

$$\mathbf{X} = [D_{1,1}, D_{2,1}, \ldots, D_{K,1}, \ldots, D_{1,M}, D_{2,M}, \ldots, D_{K,M}] \quad (1)$$

where D_{ij} is the ith discriminant calculated with respect to the signature of the jth fault, for a particular data set. This vector has $K \times M$ components when K discriminants are used for M faults.

For the classification, known principles of statistical pattern recognition (e.g., Sing-Tze Bow, 1984) are employed. The procedure consists in evaluating the M conditional probabilities of having the particular vector **X**, which is calculated using the data set acquired from the examined engine, if each of the faults exists. The decision is made classifying the vector to the fault that corresponds to the maximum probability, while the rest of the faults are ranked according to their relative probabilities:

$\mathbf{x} \in$ fault-i IF:

$$pr(\mathbf{X}/\text{fault } i) = \sum_{j=1}^{M} [pr(\mathbf{X}/\text{fault } j)] \quad (2)$$

We can see that for the implementation of a statistical procedure the possibility of calculating the conditional probability functions of having a given vector **X** $pr(\mathbf{X}/\text{fault } i)$ $i = 1, 2, \ldots, M$, for each of the faults is required. In order to obtain this possibility, a learning phase has to be carried out. During this phase, using a number of data sets for each of the faults, the corresponding vectors **X** are derived. From these, it is selected which kind of statistical distribution function is the most appropriate for statistically modeling them and then its parameters are calculated. In our case a multidimensional normal distribution has been selected. Therefore the probability of having a given vector **X** when fault i is present is expressed by an equation of the form:

$$pr(\mathbf{X}/\text{fault } i) = \frac{1}{(2\pi)^{n/2} |\mathbf{C}_i|^{1/2}} e^{-(\mathbf{X}-\underline{\mathbf{X}}_i)\cdot C_i^{-1}\cdot(\mathbf{X}-\underline{\mathbf{X}}_i)^T} \quad (3)$$

where $\underline{\mathbf{X}}_i$ and $\mathbf{C}_i$, $i = 1, 2, \ldots M$ are the average vector and the covariance matrix for each of the faults.

During the learning phase the average vectors $\underline{\mathbf{X}}_i$, $i = 1, 2, \ldots M$ are established for each fault by arithmetic averaging:

$$\underline{\mathbf{X}}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{X}_{ij} \quad (4)$$

where $\mathbf{X}_{ij}$ is the vector calculated from the jth data set of the ith fault and N_i are the numbers of the independent data sets, which are available for the ith fault. Also the covariance matrices $\mathbf{C}_i$, for each fault, are evaluated through the equation:

$$\mathbf{C}_i = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (\mathbf{X}_{ij} - X_i)^T (\mathbf{X}_{ij} - X_i) \quad (5)$$

From the above we can see that for the application of the statistical procedure, a data base is required, containing for each fault not only the healthy spectra and the reference patterns $\underline{\mathbf{P}}_i$, but also the corresponding discriminant vectors $\underline{\mathbf{X}}_i$ and covariance matrices $\mathbf{C}_i$. This data base is created during the learning phase. The flow chart of both the learning and decision phases of the statistical procedure is shown in Fig. 3.

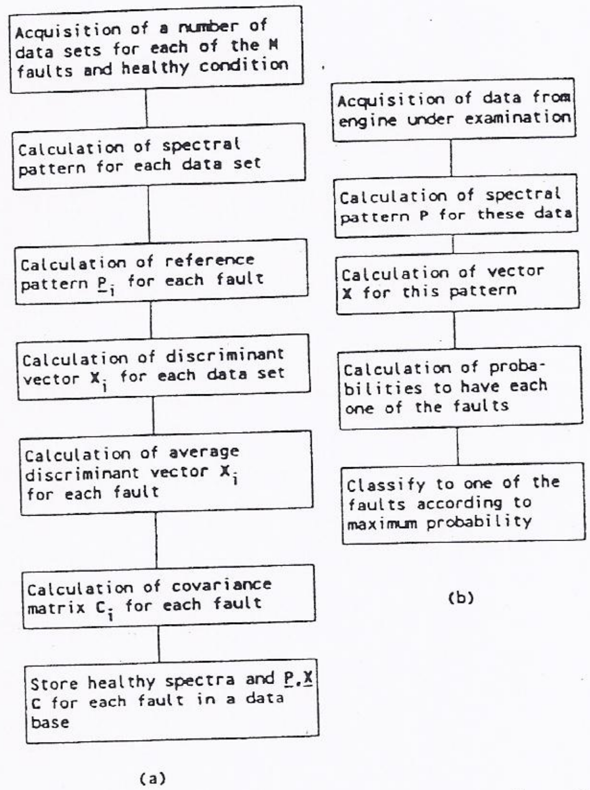

Fig. 3 Flow chart of statistical procedure: (*a*) learning phase; (*b*) decision phase

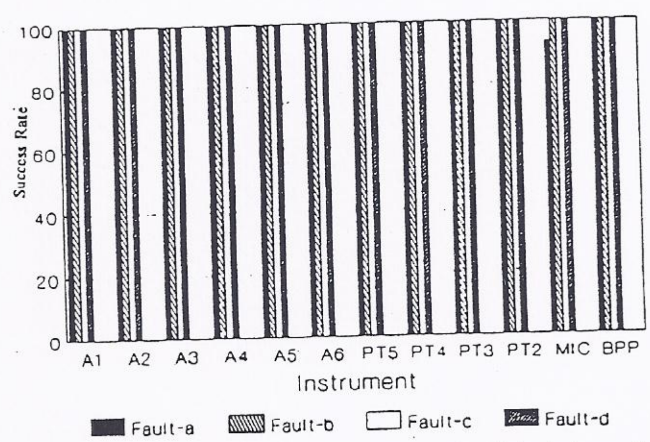

Fig. 4 Rate of success for statistical procedure for all the instruments

The statistical procedure has been applied to data that were acquired from a number of different instruments, during the experiments with four compressor blade faults on an industrial gas turbine described in Appendix I. Two discriminants were used, the ones described by Loukis et al. (1992), whose definitions are repeated in Appendix II. Therefore the dimension of vector **X** was $M \cdot N = 2 \times 4 = 8$. For the learning phase 16 data sets (4 data sets for each of the operating points A, B, C, D) were used. The testing was performed during the decision phase using the same data sets employed during the learning phase. The rate of success for each fault was calculated for each of the instruments. These rates of success are shown in Fig. 4. It can be seen that the rates of success are very high and reach 100 percent for almost all the cases, which is much higher compared with the corresponding results of the geometric procedure (Loukis et al., 1992).

4 Optimal Selection of Fault Discriminants

The question posed next is whether it is possible to improve the effectiveness of a classification procedure by choosing different discriminants. In order to try to give an answer to this problem, we rely on previous experience about fault patterns.

In the work of Loukis et al. (1991) it was established that

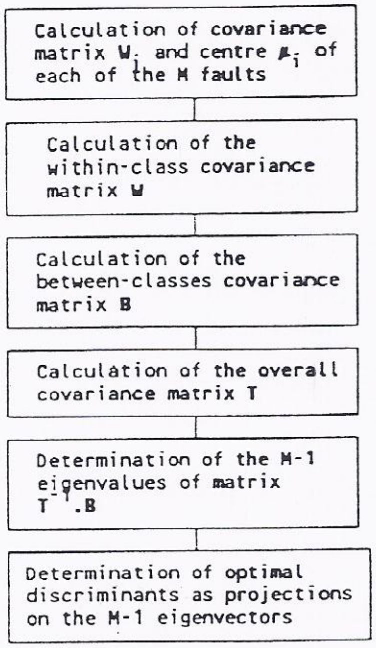

Fig. 5 Flow chart of optimal selection procedure

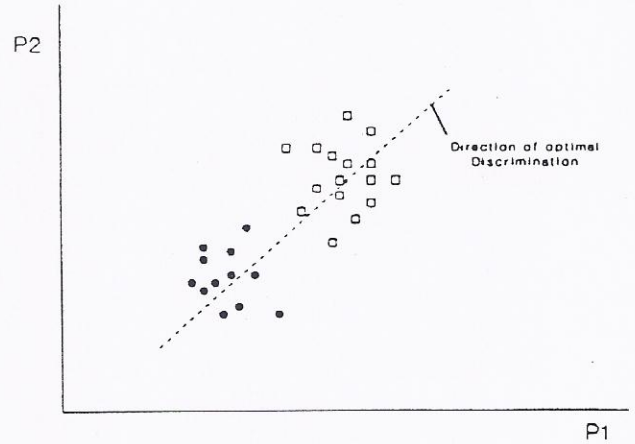

Fig. 6 Qualitative picture of the physical meaning of direction with maximum discriminability

the diagnostic information for blade faults lies mainly in a number of discrete frequencies, which are harmonics of the shaft rotational frequency. This information is incorporated in the reduced spectral difference pattern **P**, consisting of values of a fault index calculated at these shaft harmonics. However, the information at the different shaft harmonics may not be of the same value with respect to the different faults of interest.

In order to have an efficient and reliable diagnosis, the whole of this information has to be exploited. One approach to achieving this target was the one previously mentioned, based on discriminants that are derived intuitively. This approach, however, is empirical and it is not certain that it will be successful in the general case. For this reason a systematic approach for selecting discriminants is needed, according to predefined criteria.

Establishing the methods of statistical classification of fault signatures from measurement data, enables the introduction of such criteria. From the nature of the problem two criteria are employed for the selection. If the pattern **P** coming from each data set is represented as a point in the discriminant space, the two criteria have the following form:

(*i*) Points coming from experiments with the same fault should be very close together.

(ii) The clusters of points corresponding to different faults should be far apart.

The problem of selecting the discriminants according to these criteria can be formulated mathematically using the principles of optimal linear feature selection, from pattern recognition theory (see, for example, Karagiannis and Steinhauer, 1988). According to these principles, parameters that are linear combinations of the pattern **P** components are considered as possible discriminants:

$$u = \sum_{k=1}^{N} \lambda_k P_k \qquad (6)$$

The values of the linear multipliers λ_k, $k = 1, 2, \ldots, N$ are then calculated during a learning phase, by using the available data, in such a way that a quantity expressing the discrimination between the faults becomes maximum. We term such a parameter a Discrimination Potential Indicator (DPI). From these values of the linear multipliers, the optimal linear discriminant is determined. Similarly the second, the third best, etc., discriminants can be determined.

In our case we found that a suitable DPI according to the above two criteria is the Fisher Ratio (Appendix II). It is known that when M reference patterns are available, then the maximization of the Fisher Ratio leads to $M-1$ independent linear combinations $D_1, D_2, \ldots D_{M-1}$ (their DPI values are in descending order). These linear combinations are then chosen as discriminants. The flow chart of such a procedure is shown in Fig. 5.

A qualitative picture of the underlying idea for this procedure can be given through Fig. 6. In this figure a simplified case of two-dimensional patterns ($N=2$) is presented, each pattern being represented by a single point in a two-dimensional space, with coordinates the values of the pattern components. In this space we can see two groups of points, each group corresponding to data from one particular fault. Each linear combination of pattern components corresponds to a direction in this space. From this figure we can see that different directions are characterized by different discrimination potential. For example, the directions of P_1 and P_2 are characterized by low discrimination potential and the clusters of the two faults overlap, when projected on these directions. On the contrary, we can see that there are directions, along which there is no overlapping of the clusters, one of them giving the maximum discrimination potential, as indicated on the figure.

5 Application to Experimental Data

The optimal selection procedure has been applied to the data set acquired during the experiments with the four faults on an industrial gas turbine, using a number of different instruments as described in Appendix I. The number of optimal discriminants derived was $M-1=3$. The same 16 different measurements sets were used here as well.

The points from the different experiments represented on the plane D_1-D_2 for the data from pressure transducer PT2 are shown in Fig. 7. It is observed from this figure that points from experiments with the same fault are grouped in clusters. These clusters are well separated from each other and therefore provide the possibility of making diagnostic decisions with a high confidence level.

The points from the different experiments represented on the plane D_1-D_2, for the data from accelerometer A1 are shown in Fig. 8. If we compare these results to the results of Fig. 7, it can be noticed that the clusters of points are still discrete, but (*a*) a larger scattering of the points within each cluster is observed, and (*b*) the clusters are generally closer to each other. This indicates that the discrimination on the basis of acceler-

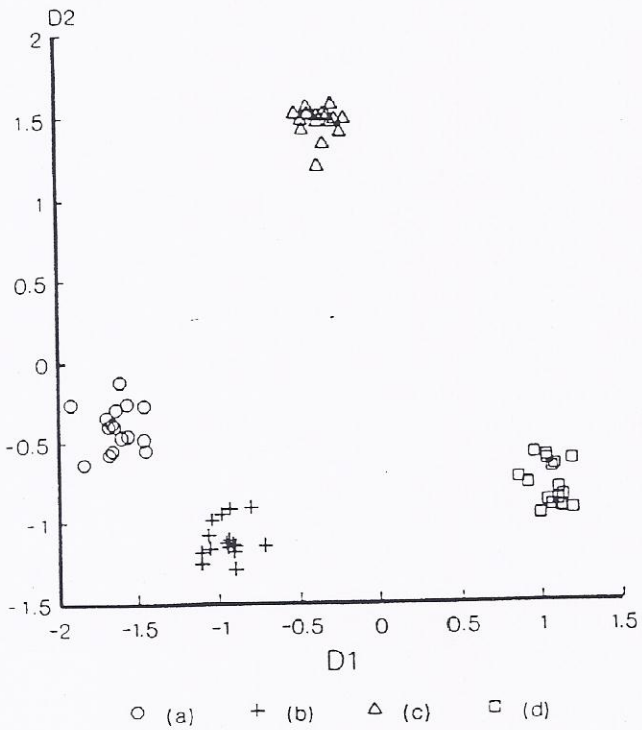

Fig. 7 Points corresponding to data from faults (*a*)–(*d*) on the plane of the discriminants D_1, D_2; measurement data from pressure transducer PT2

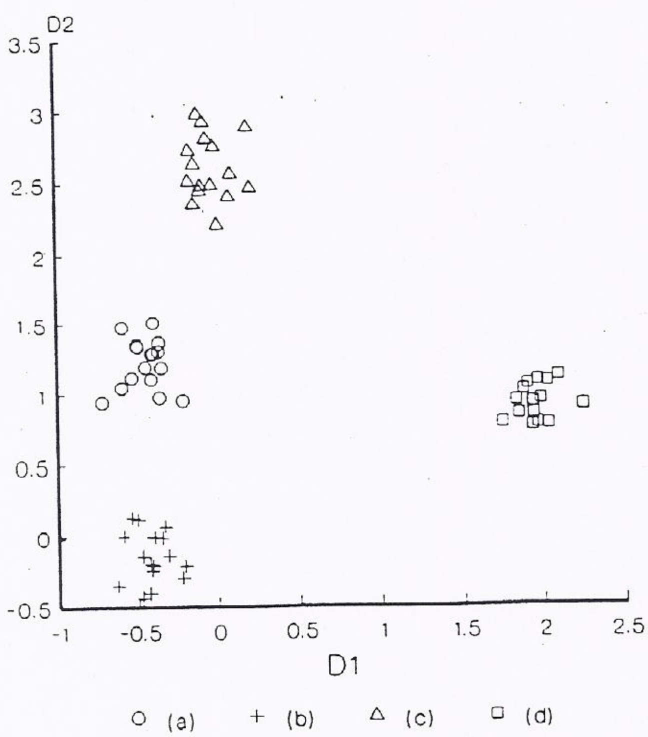

Fig. 8 Points corresponding to data from faults (*a*)–(*d*) on the plane of the discriminants D_1, D_2; measurement data from accelerometer A1

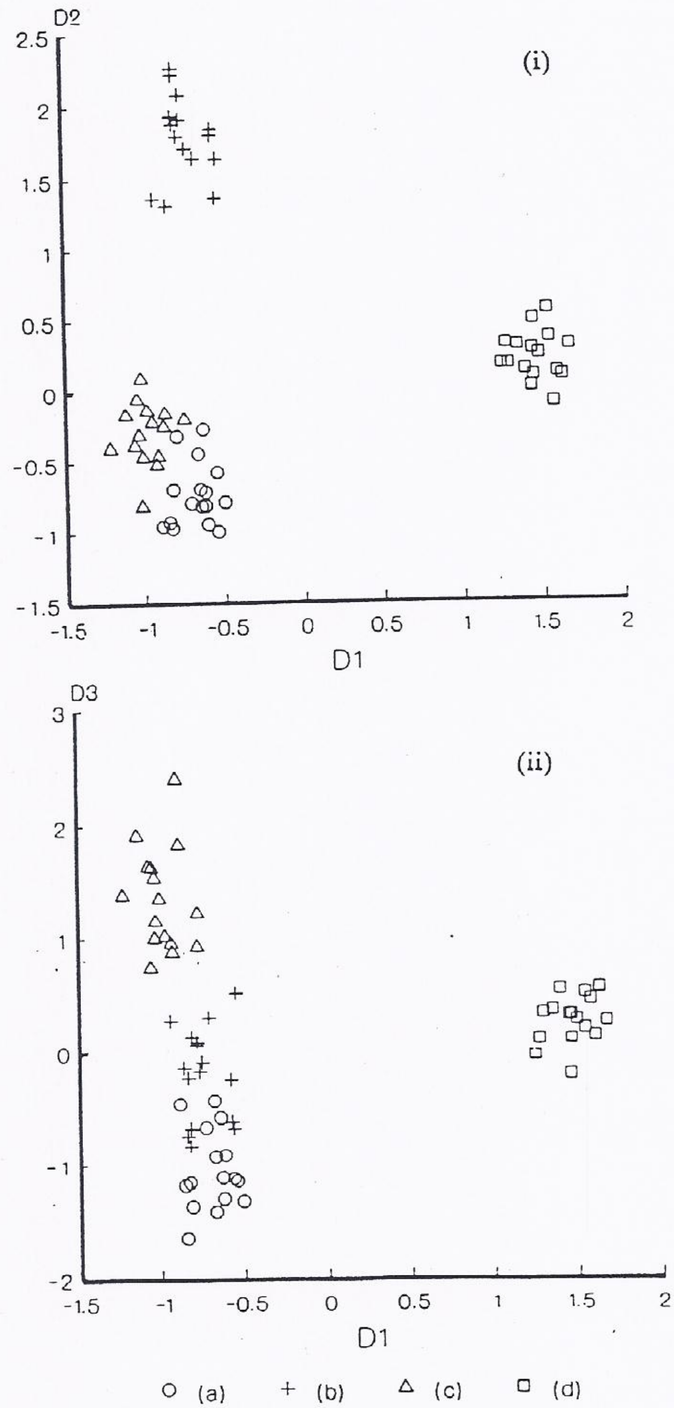

Fig. 9 Points corresponding to data from faults (*a*)–(*d*), (*i*) on the plane of the discriminants D_1, D_2, (*ii*) plane D_1, D_3; measurement data from microphone

ometer data should be less effective than on the basis of pressure transducer data.

The same picture for microphone data is shown in Fig. 9 on the D_1–D_2 and D_1–D_3. It is interesting to observe here that the scattering of points in each cluster is larger than it is for the accelerometer and the pressure transducer. It can also be seen that faults *a* and *c* clusters overlap on the D_1–D_2 plane while they are separated on the D_1–D_3 plane. A similar observation is valid for faults *a* and *b*. These observations show that if the decision is based on only two discriminants it can lead to a low success rate, while the success rate will be much higher when all three discriminants are used. On the other hand the clusters of faults *a*, *b*, *c* are close to each other but all of them well separated from the one of fault *d*. This constitutes a quantitative verification of observations based on visual inspect of fault signatures, reported by Loukis et al. (1991). (The visual inspection had shown that stator faults are easily discriminable from microphone measurements, while no visual discrimination was possible for rotor blade faults.)

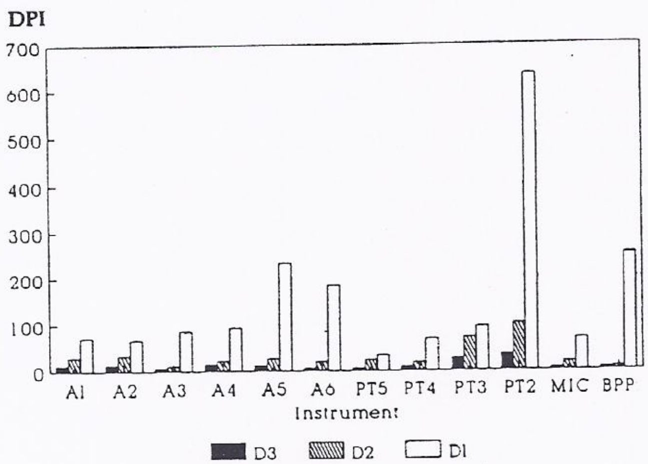

Fig. 10 DPI values for optimal discriminants D_1, D_2, D_3, for all measuring instruments

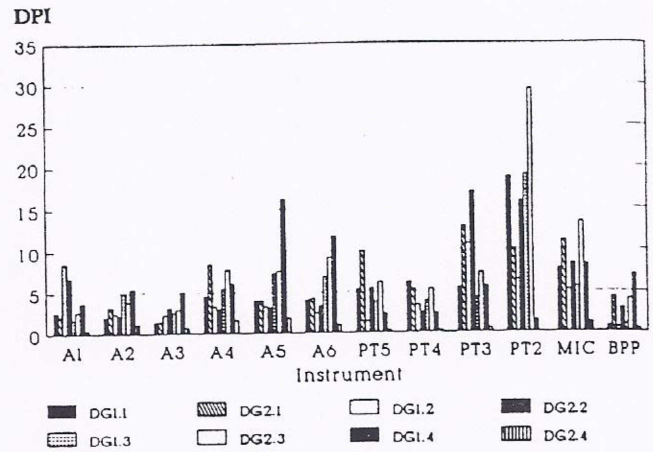

Fig. 11 DPI values for the discriminants DG_{1k}, DG_{2k}, calculated with respect to the four fault signatures for all instruments

An additional feature of the method is that the DPI can be used as a quantitative measure for the discrimination capability of discriminants. As an application first a quantitative assessment of the discrimination capability of the three optimal discriminants D_1, D_2, D_3 was done for all measuring instruments. The DPI values for D_1, D_2, D_3 and each measuring instrument are presented in Fig. 10. It can be observed that D_1 performs better than D_2 and D_2 better than D_3. This was expected since D_1, D_2, D_3 are by definition in descending order of DPI. On the other hand, DPI values for the discriminants of Appendix II with respect to the four fault signatures, are shown in Fig. 11. These results provide a comparative picture of the potential of the intuitively chosen discriminants. Comparison to Fig. 10 shows that the intuitive discriminants are characterized by much poorer performance than the systematically selected ones, since they exhibit much lower DPI values. This fact verifies the usefulness of the above systematic selection procedure.

The discriminants D_1, D_2, and D_3 have been used for the implementation of a statistical procedure of fault classification. This procedure has been based on a vector **X** of the form $\mathbf{X} = [D_1, D_2, D_3]$. The learning phase and the decision phase were according to Section 3. The rate of success for each fault, calculated for each of the instruments, was 100 percent for all cases. These results are a further confirmation of the improved performance achieved when systematic selection of discriminants is used.

6 Discussion

The procedures presented in the above sections possess features that make them very useful for implementation in practical industrial situations of fault diagnosis. Their most useful feature is that they incorporate data processing and decision making, while they can be implemented by a computer system.

The statistical procedure of classifying fault signatures contributes to the modern requirements mentioned in the introduction by increasing the efficiency and the reliability of diagnosis and giving the advantage of associating probabilities to the fault diagnosis. This feature makes the procedure more realistic for industrial application, in comparison with the geometric method.

The procedure of systematic discriminant selection offers the flexibility to define parameters that fulfill predefined discrimination criteria, increasing in this way the efficiency and reliability of the diagnosis as well. This systematic selection ensures that the classification procedure will be organized in accordance with the peculiarities of the specific problem, instead of a priori choosing discriminants intuitively and finding whether they are effective or not. On the other hand this procedure offers the possibility to redefine the discriminants when new data are available or new faults have to be introduced. This means that when extension of the data base to include new faults is effected, the procedure can be adapted to the new data, ensuring that discrimination capabilities will still be optimal.

Another useful product of the optimal selection procedure is the possibility to rank not only the discriminants, but also the various measuring instruments as well, according to the discrimination possibility they offer. Figure 10 can be interpreted in such a way if the values of DPI for the same discriminant are compared for the various instruments. We see that some instruments have higher DPI values than others, indicating better discrimination capabilities. Pressure transducer PT2 is characterized by the maximum DPI from all the instruments. This agrees with the conclusion reported by Loukis et al. (1991), from observations of various instrumental data, that pressure transducers are best suited for blade fault identification. The procedure above is therefore very useful in selecting the minimal and most appropriate set of instruments.

Finally, it should be mentioned that the procedure described above has been derived from one set of experiments on one engine. In order to examine generalization aspects, more data from other faults and other engines should be used.

7 Conclusions

In order to develop a sound procedure for automated gas turbine blade fault diagnosis the principles of pattern recognition theory have been employed. It was shown that when the statistical properties of fault signatures are taken into account advanced possibilities of decision making are provided.

A procedure for the systematic discriminant selection was developed. This procedure gives the possibility to select discriminants according to a priori set criteria and eliminates the need of selecting them intuitively. It was shown that this procedure further improves the decision making.

The procedure above also provides a possibility of assessing quantitatively the diagnostic effectiveness of (*a*) parameters used as discriminants, (*b*) instruments used as a source of information. Such a possibility is very useful in diagnostic systems design.

References

Carchedi, F., and Wood, G. R., 1982, "Design and Development of a 12:1 Pressure Ratio Compressor for the Ruston 6-MW Gas Turbine," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 104, pp. 823–830.

Doel, D., 1990, "The Role of Expert Systems in Commercial Gas Turbine Engine Monitoring," ASME Paper No. 90-GT-374,.

Emel, E., and Kannatey-Asibu, E., Jr., 1988, "Tool Failure Monitoring in Turning by Pattern Recognition Analysis of AE Signals," ASME *Journal of Engineering for Industry*, Vol. 110, pp. 137–145.

Karagiannis, G., and Steinhauer, C., 1988, "Pattern Recognition," National Technical University of Athens, Department of Mechanical Engineering, Athens.

Loukis, E., Wetta, P., Mathioudakis, K., Papathanasiou, A., and Papailiou, K., 1991, "Combination of Different Unsteady Quantity Measurements for Gas Turbine Blade Fault Diagnosis," ASME Paper No. 91-GT-201.

Loukis, E., Mathioudakis, K., and Papailiou, K. D., 1992, "A Procedure for Automated Gas Turbine Blade Fault Identification Based on Spectral Pattern Analysis," ASME JOURNAL OF ENGINEERING FOR GAS TURBINE AND POWER, Vol. 114, pp. 201–208.

Mathioudakis, K., Papathanasiou, A., Loukis, E., and Papailiou, K., 1991, "Fast Response Wall Pressure Measurement as a Means of Gas Turbine Blade Fault Identification," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 113, pp. 269–275.

Sing-Tze Bow, 1984, *Pattern Recognition*, Marcel Dekker Inc., New York-Basel.

APPENDIX I

Experimental Faults and Instrumentation

The present investigation was based upon dynamic measurements on an industrial gas turbine into which a number of different representative blade faults were introduced. The test engine was the Ruston Tornado, described by Charchedi and Wood (1982). The instrumentation layout and measurement conditions described by Loukis et al. (1991).

During the experiments four categories of measurements were performed simultaneously:

(*i*) Unsteady internal wall pressure, using two unsteady pressure transducers facing the first four rotors of the compressor, named PT2–PT5 respectively.

(*ii*) Casing vibration, with the accelerometers A1 to A6 mounted at the outside compressor casing.

(*iii*) Shaft displacement at compressor inlet bearings, with a Bently Nevada system (named BPP).

(*iv*) Sound pressure levels, with a double-layer microphone (named MIC).

Five experiments were performed testing the datum healthy engine and engines with the following four faults:

Fault a. Rotor fouling; all stage 2 rotor blades were coated with textured paint in order to roughen blade surface and alter their contour, in order to simulate a fault of all the blades of a rotor.

Fault b. Individual rotor blade fouling; two blades of stage 1 rotor separated by five intact blades, were coated by textured paint in order to simulate a slight individual rotor blade fault.

Fault c. Individual rotor blade twisted; a single blade of stage 1 rotor was twisted in order to simulate a severe individual rotor blade fault.

Fault d. Stator blade restaggering; three stage 1 stator blades were mistuned in order to simulate a stator fault.

All numbered references to faults and instruments in the text will correspond to the nomenclature above.

Tests were performed at four different engine loads; full load, half load, quarter load, and no load (termed operating points A, B, C, D, respectively, hereafter).

APPENDIX II

Formulas for Evaluation of Pattern Properties

Definition of Discriminants. The discriminants used for pattern classification by Loukis et al. (1992) are defined through the following relations:

$$DG_{1,k}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(P(i)-P_K(i))^2}\ \text{Euclidean distance}$$

$$DG_{2,k}=\frac{\sum_{i=1}^{N}(P(i)-Av(\mathbf{P}))(P_k(i)-Av(\mathbf{P_K}))}{\sqrt{\sum_{i=1}^{N}(P(i)-Av(\mathbf{P})^2)}\ \sqrt{\sum_{i=1}^{N}(\mathbf{P_K}(i)-Av(P_K)^2)}}$$

(correlation coefficient)

$$Av(\mathbf{P})=\frac{1}{N}\sum_{i=1}^{N}P(i)$$

Formulas for Optimal Linear Feature Selection. The principles we employ for optimal linear feature selection are briefly described below. Extensive descriptions can be found in textbooks (e.g., Karagiannis and Steinhauer, 1988).

The Fisher ratio expresses the classification capabilities for data if they are projected on one direction in the feature space. For a given direction, defined by its unit vector **u**, it is defined by a quotient that has as numerator the variance of the centers of the classes, if data points are projected on this direction. Its denominator is the average internal variance of the classes, if data points are projected on this direction. It is calculated, through the relation:

$$F(\mathbf{u})=\frac{\mathbf{u}^T\cdot\mathbf{B}\cdot\mathbf{u}}{\mathbf{u}_T\cdot\mathbf{W}\cdot\mathbf{u}}$$

where:

- $\mathbf{u}$ = the unit vector along the directional examined
- $\mathbf{B}$ = between-classes covariance matrix
- $\mathbf{B}=\sum_{k=1}^{M}\frac{N_k}{N}(\underline{\mu}_k-\mathbf{m})(\underline{\mu}_k-\mathbf{m})^T$
- M = number of classes
- N_k = number of available patterns of the kth class
- μ_k = mean pattern of the kth class
- N = total number of available patterns (for all the classes)
- $\underline{\mathbf{m}}$ = overall mean pattern (of all the classes)
- $\mathbf{W}$ = within-class covariance matrix, equal to the arithmetic average of the covariance matrices $\mathbf{W}_k$, $k=1, \ldots, M$, of all the classes:

$$\mathbf{W}=\sum_{k=1}^{M}\mathbf{W}_k$$

$$\mathbf{W}_k=\frac{1}{N_k}\sum_{i=1}^{N_k}(\mathbf{X}_i-\mu_k)(\mathbf{X}_i-\mu_k)^T$$

It can be shown that the sum of the within-classes covariance matrix **W** and the between-classes covariance matrix **B** gives the overall covariance matrix **T**, which takes into account the patterns of all classes, defined by the following equation:

$$\mathbf{T}=\frac{1}{N}\sum_{i=1}^{N}(\mathbf{X}_i-\mathbf{m})(\mathbf{X}_i-\mathbf{m})^T$$

It can be shown that the direction characterized by the maximum value of the above Fisher Ratio is the eigenvector of the $\mathbf{T}^{-1}\cdot\mathbf{B}$ matrix corresponding to its highest eigenvalue. Generally there are $M-1$ linearly independent directions maximizing the Fisher ratio, which are the eigenvectors of the $\mathbf{T}^{-1}\cdot\mathbf{B}$ matrix. Their Fisher ratios decrease with the corresponding eigenvalues.