Detection of Cracks in a Rotating Shaft Using Density Characterization of Orbit Plots

T. Haj Mohamad[1(✉)], A. A. Cavalini Jr.[2], V. Steffen Jr.[2], and C. Nataraj[1]

[1] Villanova Center for Analytics of Dynamical Systems, Villanova University, Villanova, PA 19085, USA
thajmoha@villanova.edu

[2] Structural Mechanics Laboratory (LMEst), Federal University of Uberlândia, Uberlândia, MG, Brazil
http://www.vcads.org, http://www.ufu.br

Abstract. The research focuses on fault detection and diagnostics of cracks in a rotating shaft by using the Extended Phase Space Topology approach (EPST). EPST is based on extracting features from the constructed density profile of the system vibration responses. The extracted features are ranked and the optimal set is selected by using mutual information. Finally, an artificial neural network is used as a classifier to distinguish between the different shaft conditions. The method was implemented on a laboratory scale rotor test rig that was seeded with two damage conditions produced by a crack propagator. As will be shown, the study demonstrates that the density distribution provides rich information about the shaft structural condition. Furthermore, results show that the innovative EPST procedure has outstanding performance in shaft crack detection with minimal knowledge about the dynamic responses of the system.

Keywords: Rotating machines · Crack detection · Machine learning

1 Introduction

Visual inspection, X-ray examinations, ultrasonic testing, and penetrating liquid are non-destructive evaluation techniques used in rotating shafts for crack detection purposes. These methods are traditionally applied during maintenance or intervention procedures. However, they have proved to be costly, since satisfactory results depend on detailed and, consequently, long inspections [1]. Thus, in recent years, experts have been developing modern structural health monitoring (SHM) techniques able to be applied in the machine during operation to reduce the costs involved and, at the same time, increase safety, performance, and working time.

SHM techniques based on vibration measurements are recognized as useful alternatives because they lead to satisfactory results even when the location of

K. L. Cavalca and H. I. Weber (Eds.): IFToMM 2018, MMS 61, pp. 90–104, 2019.
https://doi.org/10.1007/978-3-319-99268-6_7

the damage is not accessible or even unknown [2]. Even an incipient crack, for instance, is capable of changing the shaft stiffness and damping properties, which is reflected in the measured vibration responses. Thus, SHM techniques based on vibration signals have been proposed by several researchers [3–7].

Most SHM techniques devoted to crack detection monitor the synchronous vibration responses of the machine during normal or transient working conditions [8]. However, although widely used in industry, when applied under non-ideal conditions, such techniques eventually detect cracks that have already propagated significantly across the shaft cross-section [9–11]. It is worth mentioning that the influence of the crack on dynamic behavior of the shaft still depends on its location. It is maximum for cracks found in regions of maximum curvature and is nonexistent in regions of zero curvature. Therefore, the attention of researchers is being devoted to more sophisticated methodologies capable of identifying incipient cracks, which is usually difficult to observe in the directly measured vibration responses.

The Phase Space Topology (PST) and Extended Phase Space Topology (EPST) methods are examples of the mentioned sophisticated approaches. In earlier work, PST was introduced in our lab to characterize different phase space trajectories with quantitative measures [12,13]. The phase space represents all the states of the system evolving with time. Analyzing the topology of the phase space trajectory provides valuable information about the dynamics of a system in a qualitative fashion [14]. PST was applied to a nonlinear pendulum system and a nonlinear oscillator system in order to estimate the parameters of each system and to characterize different conditions of each system. As an extension of the previous work, EPST was introduced in [15–18] as a machinery diagnostics method and is based on characterizing the phase space topology of the system by determining some characteristic values. After extracting a feature set by using EPST, mutual information was used as a feature ranking and selecting technique. This technique was developed in our past work [19–22] to diagnose rolling bearing with various health conditions. Mutual information between the feature and the defect is used as a quantitative measure of quality.

In the present contribution, EPST was applied to the vibration responses of a horizontal rotating machine in order to detect and diagnose shaft faults. Three different shaft conditions were considered including non-defective rotor (HR), and a rotor with two damage conditions (DC1 and DC2) produced by a crack propagator over 24-h and 48-h time periods. The horizontal and vertical shaft displacements were measured by using proximity probes for each rotor structural condition. An optimal set of features were extracted from the vibrational signals to build an artificial neural network and to classify the rotor actual condition.

This paper is organized as follows. In Sect. 2, the mathematical details of EPST are introduced. Section 3 explains the steps of using mutual information to rank and select the optimal feature set. Section 4 describes the experimental test rig and measurement of data used in this study. In Sect. 5, the process of feature extraction and classifier training is described and the final prediction results are presented. Finally, Sect. 6 summarizes and concludes this paper.

2 Feature Extraction Method

The PST approach [12,13] is based on the transformation of phase space into the density space and characterizing the density with quantitative measures. It was shown that, depending on the geometry and shape of the phase space, the density diagram contains peaks of various heights and sharpness at multiple locations. This stems from the fact that the dynamical system occupies more time at specific regions of the space causing higher densities in those regions. The properties of the peaks in the density diagrams including the location, height, and sharpness of the peaks were used as features in the initial approach.

Despite the success of this approach, the need to search for the peaks in the density diagrams makes it difficult or sometimes even impractical to implement, especially for systems with noisy or more complex phase space patterns. The upgraded version of this approach, which will be described below, is based on approximating the density distribution with Legendre polynomials.

Regarding the kernel density estimation, let $X=(x_1, x_2, ..., x_n)$ be an independent and identically distributed sampled data drawn from a distribution with an unknown density function f. The shape of this function can be estimated by its kernel density estimator.

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{1}$$

where, $h > 0$ is a smoothing parameter called bandwidth and $K(.)$ is the kernel function which satisfies the following requirements. $\hat{}$ indicates that it is an estimate and h indicates that its value can depend on h.

$$\int_{-\infty}^{\infty} K(u)\mathrm{d}u = 1 \tag{2}$$

$$K(-u) = K(u)\ \forall u \tag{3}$$

There is a range of kernel functions that can be used, including uniform, triangular, biweight, triweight, Epanechnikov, normal, etc. Due to its conventional and convenient mathematical properties, the standard normal density function was used in the present contribution, defined as the following:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \tag{4}$$

Let x be a state of the system and $y_d = \hat{f}_h(x)$ its density computed using the kernel density estimator. y_d is then approximated with Legendre orthogonal polynomials. Legendre polynomials can directly be obtained from Rodrigues' formula which is given by:

$$P_m(x) = \frac{1}{2^m m!} \frac{d^m}{dx^m} \left[(x^2 - 1)^m\right] \tag{5}$$

where $m = 0, 1, 2, \ldots$, or can be obtained from the recursive definition by using Bonnet's recursion formula:

$$(m+1)P_{m+1}(x) = (2m+1)xP_m(x) - mP_{m-1}(x) \tag{6}$$

where the first two terms are given by:

$$P_0(x) = 1, \quad P_1(x) = x \tag{7}$$

The coefficients of the Legendre polynomials are obtained by using the linear least squares method assuming the following linear regression model:

$$f(x, \beta) = \sum_{j=0}^{m} \beta_j P_j(x) \tag{8}$$

Letting

$$X_{ij} = \frac{\partial f(x_i, \beta)}{\partial \beta_j} = P_j(x_i) \tag{9}$$

where, the estimated coefficients are given by:

$$\hat{\beta} = (X^T X)^{-1} X^T y_d \tag{10}$$

where the coefficients $\hat{\beta}$ constitute the features in our approach that can be used in classification or regression problems.

The approximated density based on Legendre Polynomials is calculated by using the following equation:

$$f = X\hat{\beta} \tag{11}$$

Root mean square error (RMSE) and Pearson's correlation coefficient (PCC) were calculated to compute the quality of the fit by using the following equations:

$$\text{RMSE} = \sqrt{\frac{1}{N} Z Z^T} \tag{12}$$

where, $Z = y_d - f$ is the residuals and N is the number of points in the density function.

$$\text{PCC} = \frac{a^T b}{\sqrt{(a^T a)(b^T b)}}; \tag{13}$$

where, $a = y_d - E\{y_d\}$ and $b = f - E\{f\}$, $E\{.\}$ is the expected value.

3 Feature Ranking and Selection by Using Mutual Information

Let v_i be the random variable with probability distribution function (PDF) $p(v_i)$ corresponding to the i^{th} feature. Let C be any classifier that maps the features into N_C classes and c_k the corresponding random variable with PDF $p(c_k)$, $k = 1, \ , 2, \ ..., \ N_C$. Note that c_k is a discrete random variable. The entropy and mutual information are defined as in Eqs. (14) and (15).

$$H(v_i) = -\int p(v_i) \log p(v_i) dv \tag{14}$$

$$I(v_i; c_k) = -\int p(v_i, c_k) \log \frac{p(v_i, c_k)}{p(v_i)p(c_k)} dv \tag{15}$$

Further, the entropy and mutual information are related by Eq. (16).

$$I(v_i; c_k) = H(c_k) - H(c_k|v_k) \tag{16}$$

In order to calculate the mutual information, it is necessary to find $p(c_k)$ and $p(c_k|x_i)$ from the data. However, it is easy to find $p(c_k)$ as it is a discrete random variable. By Bayesian rule we have

$$p(c_k|x_i) = \frac{p(x_i|c_k)p(c_k)}{p(x_i)} \tag{17}$$

The PDF of a continuous random variable v can be estimated using kernel density estimator in Eq. (1). From Eqs. (1) and (18), $p(v_i|c_k)$ can be calculated as follows:

$$p(v_i|c_k) = \frac{1}{N_k h} \sum_{i=1}^{N_k} K(\frac{v - v_{k_i}}{h}) \tag{18}$$

where N_k are the number of samples in the k^{th} class and v_{k_i} are the features belonging to k^{th} class.

Mutual information between a feature and a class can be calculated by using Eq. (19).

$$I(v_i, c) = \sum_{k=1}^{N_C} p(c_k) \log p(c_k) - \int \sum_{k=1}^{N_C} p(v_i|c_k)p(c_k) \log p(v_i|c_k) dv \tag{19}$$

However, in order to calculate the mutual information between a set of features, $v = [v_1 \ v_2 \ , ..., \ v_n]$ and a class, it is necessary to calculate the joint PDF

$p(v)$ of the feature set and the conditional joint PDF $p(v|c)$. Although it is possible to do this, it is cumbersome and often inaccurate. A simpler procedure is to use Eq. (20).

$$I(x;c) = \frac{1}{|S|}\sum_{v_i \epsilon S} I(v_i;c) - \frac{1}{|S-1|^2}\sum_{v_i, v_j \epsilon S} I(v_i;v_j)$$
$$v = \{v : v \epsilon S \subset V\} \tag{20}$$

The first part of the right hand side of Eq. (20) is the mean of the mutual information of each of the features and class; it is a measure of relevance of the set S. The second part consists of information between the features themselves; it is a measure of redundancy of the set S. Using this method it is necessary to only calculate the joint PDF of two features at a time. This method, when used in a sequential search, has similar performance to the actual value [23].

The feature selection process by using these measures is an optimization problem and can be formally defined as in Eq. (21).

$$\max I(v;c),\ v = \{v : v \epsilon S \subset V\} \tag{21}$$

where V is the set containing the features and S is some subset of it. If the size of S is equal to size of V then the solution of Eq. (21) will be an ordered set of features.

Equation (21) can be solved by using the greedy search technique. In the first step of this technique, set S is initialized to an empty set and a feature pool set defined as F is initialized to V. Next, S is populated iteratively with a feature from the feature pool such that it maximizes $I(v;c)$ at each stage. The selected feature is then removed from the feature pool. This process is continued till the feature pool is empty. The algorithm for ranking can be summarized as follows:

1. From the data find $p(c_k)$ and $H(c_k)$, $k = 1,\ 2,\ 3,\ ...,\ N_C$.
2. Set $S = \{\}$, $F = V$.
3. While F is not an empty set, DO
   (a) Set $i = 1$, Start Loop 1
   (b) Append the i^{th} element of F to S, i.e. $S_i = \{S,\ F_i\}$.
   (c) Set $j = 1$, Start Loop 2
   (d) Using Eq. 19 find $I(v_j, c)$.
   (e) Using Eq. 1 find $I(v_i, v_j)$.
   (f) If reached the end of S_i End Loop 2, else increment $j \rightarrow j + 1$ and go to Step d.
   (g) Estimate mutual information of set S_i, $I(S_i, c)$ using Eq. 20.
   (h) If reached the end of F End Loop 1, else increment $i \rightarrow i + 1$ and go to Step b.
   (i) Find the element v_i^* corresponding to Maximum $I(S_i, c)$.
   (j) Append v_i^* to S and remove it from F.
4. END WHILE
5. The final set S is the ordered feature set.

4 Experimental Setup

This study was implemented on data collected from a rotor test rig shown in Fig. 1. This machine consists of a flexible steel shaft mounted on two roller bearings. Two orthogonal proximity probes oriented in the horizontal and vertical directions were used to measure the vibration response of the shaft. The tests were conducted with a balanced shaft operating at 1200 rpm. From a representative finite element model of the rotor, the first two critical speeds were determined, being approximately 1685 rev/min (28.1 Hz) and 5430 rev/min (90.5 Hz).

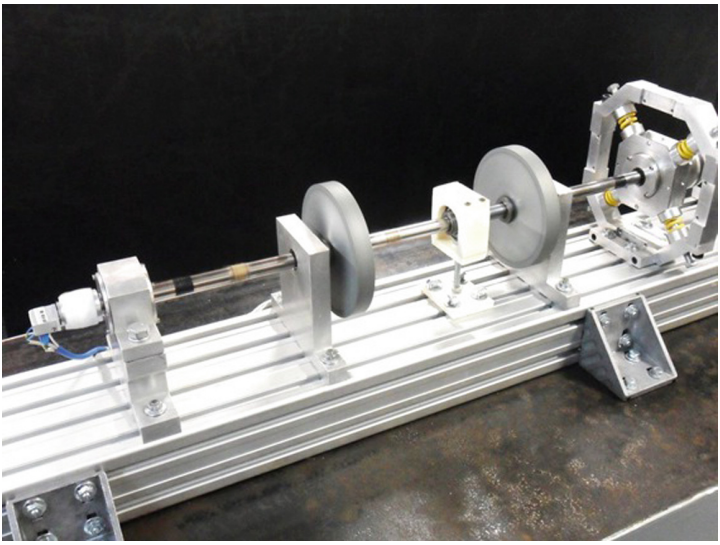

Fig. 1. The rotor test rig.

This experiment consists of three stages corresponding to three different structural conditions. The first stage was performed on the notched shaft, which is considered the healthy rotor condition (HR). Then, the crack propagator was used over a period of 24 h to produce a fatigue crack, which is the first damage condition (DC1). The third stage studied the rotor after using the crack propagator for another 24 h period to produce the second damage condition (DC2). All measurements were conducted on the rotor after removing the crack propagator device. In this case, a notch was first machined in the shaft. Then, a crack propagator was coupled to the shaft aiming to apply a constant downward force. Consequently, a fatigue crack could be nucleated, or propagated, at the notch position with the rotor in operation. Figure 2 shows the propagated crack that was produced by the crack propagator at an instant of time during the experiment.

Data was collected for the three rotor conditions including healthy rotor (HR), and rotor with the two damage conditions produced by a crack propagator

(DC1 and DC2) at a sampling rate of 4096 Hz and for a time length of 1 s. For each condition, 5 sets of data were collected for a total number of 15 sampled signals. Samples of the measured vibration for different rotor conditions in the vertical direction are shown in Fig. 3.

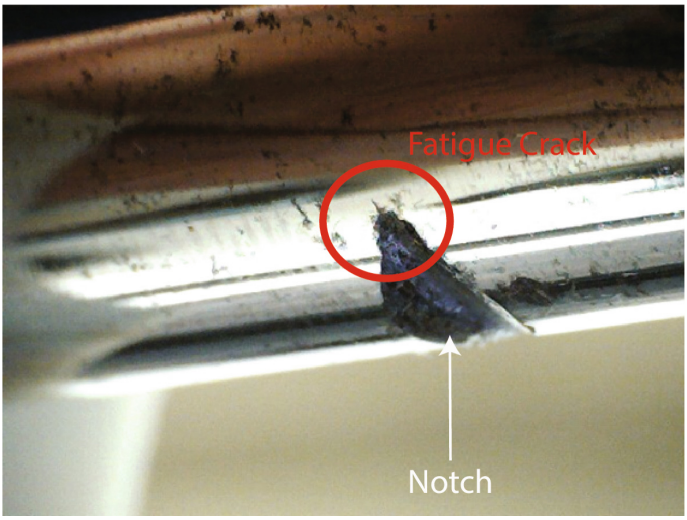

Fig. 2. A propagated crack during the experiment.

5 Analysis and Fault Classification

In order to analyze the vibration responses for various rotor conditions, each data set was divided into two segments. The total number of data segments that were obtained is 30. EPST is based on characterizing the phase space topology of the system by determining some characteristic values. A phase plane is the two-dimensional case of the phase space portrait and in the context of this problem, the orbit plot can be considered as a projection of the phase space into two dimensions. Therefore, the orbit plots of the shaft were investigated first.

Figure 4 shows the obtained orbit plots for all three rotor conditions. As can be seen, the shapes of the orbit plots vary with the rotor condition and this variation becomes manifest for the HR and DC1 conditions. Note that the rotor response of the second damage condition (DC2) has a smaller orbit plot compared to the first damage condition (DC1). This counter/intuition response is often reported [24], and can be associated with the shape of the crack surface changing from DC1 to DC2 conditions. Additionally, a small bow could be generated in the shaft by the crack propagator during the tests explaining the variation of the 1X vibration component. These observation indicate that

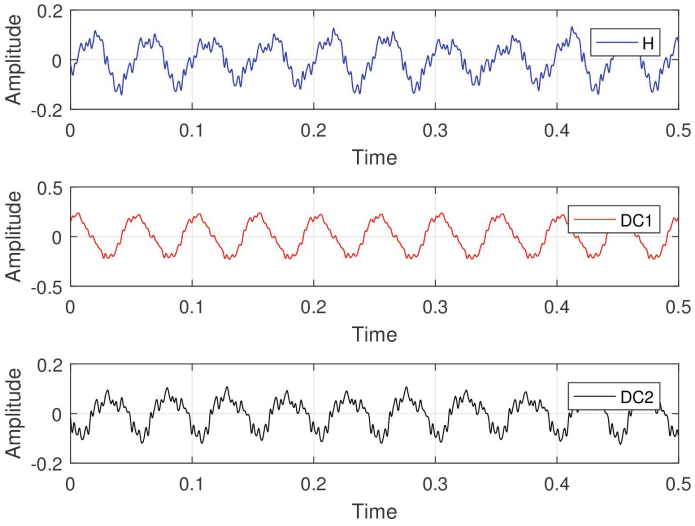

Fig. 3. Samples of the time responses of the vertical proximity probe for the three rotor conditions.

condition monitoring of rotors can be performed by characterizing the rotation profile of the shaft.

The density distribution of the vibration data was estimated by using the kernel density estimator. Samples of the estimated density distributions for different rotor conditions in the horizontal and vertical directions (direction of the orbit plots axis) are shown in Figs. 5 and 6, respectively.

Inspecting the orbit plots and their corresponding density distributions, several conclusion can be drawn. First, every loop in the orbit plot has two peaks in the density distribution profile due to the higher concentration points at the returning points. Additional loops in the orbit plots produces additional peaks in the density distribution profile. Second, the estimated density distributions of three rotor conditions are distinguishable in the horizontal and the vertical directions. The density distributions are bimodal which indicates having two frequently occurring states. Finally, The density distribution of the first damage condition (DC1) has a wider range and lower frequency compared with the other rotor conditions. Furthermore, the density distribution of the second crack damage has the highest frequency. These observations indicate that the density distribution of the vibration data provides valuable information for characterizing the dynamic response of different rotor conditions.

After estimating the density distribution for each sample set including both directions, the density distribution profiles were approximated by using Legendre polynomials of order 30 (see Eq. (8)). This order was chosen based on the fit quality between the actual and the approximated densities, which was measured by calculating the RMSE and PCC values. The coefficients β of the Legendre polynomials were computed for each of the 30 sampled signals using the linear

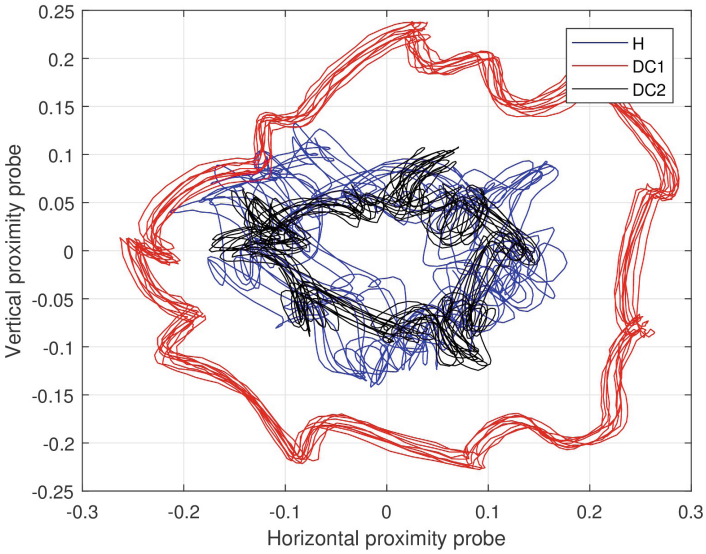

Fig. 4. Orbit plots for the three rotor conditions.

least squares method as shown in Eq. (10). The extracted features for horizontal and vertical data were 62 features in total for each sample set. By using the mutual information as explained in Sect. 3, the feature set was ranked and shown in Table 1. In β_i^s, i represents the coefficient number and s is the proximity probe direction (horizontal or vertical).

Table 1. The ranked features using mutual information

Class	Ranked orthogonal function coefficients								
	1	2	3	4	5	6	7	8	9
β_i^s	β_{13}^v	β_0^h	β_{11}^v	β_2^v	β_{15}^v	β_2^h	β_9^v	β_1^h	β_0^v

A neural network of 25 hidden neurons was trained by using 50% of the data samples and backpropogation algorithm. Additionally, 50% of the data samples were used for testing the trained classifier. Confusion matrices that document the performance of the training and testing classification model are shown in Fig. 7. The training and the testing confusion matrix show a virtually perfect result. The classifier was able to correctly predict all the cases with 100% accuracy. Note that in this classification model all the extracted features were used and the ranked features were not used.

In order to select the best feature subset that represents the dynamic system, the mutual information techniques was used to rank the training data set. Then, an iterative process creates a neural network of 25 hidden neurons and k inputs

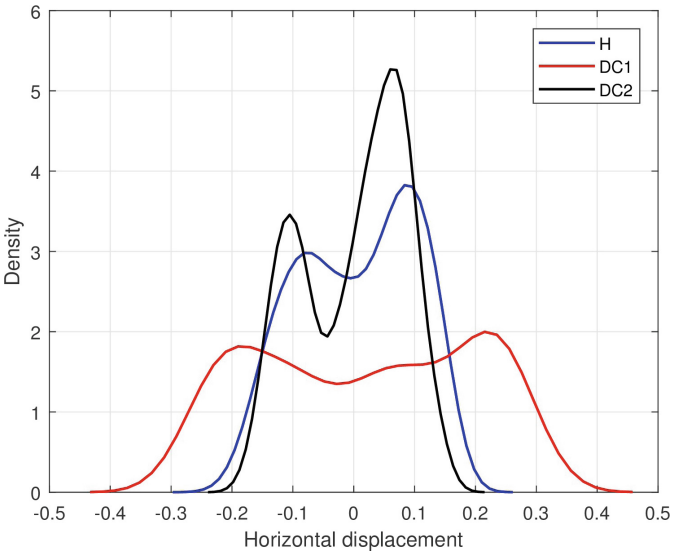

Fig. 5. Density distribution of horizontal displacement for the three rotor conditions.

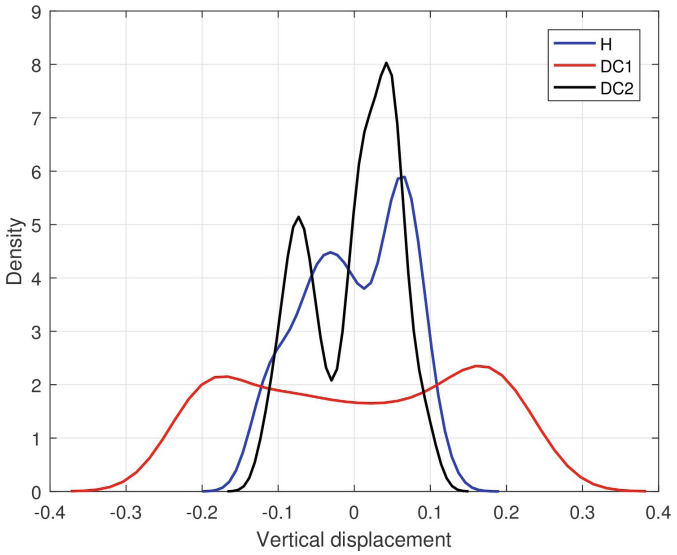

Fig. 6. Density distribution of vertical displacement for the three rotor conditions.

(features). The k features are first k subset in the ranked feature set. Finally, the testing set is used to calculate the accuracy of prediction and the optimal subset is selected based on the highest classification accuracy. The feature selection

process is summarized in Fig. 8. Prediction results of the first three iterations are presented in Table 2.

Table 2. The accuracy of the neural network using k ranked features

k	Training accuracy	Testing accuracy
1	66.6%	66.6%
2	66.6%	66.6%
3	100%	100%

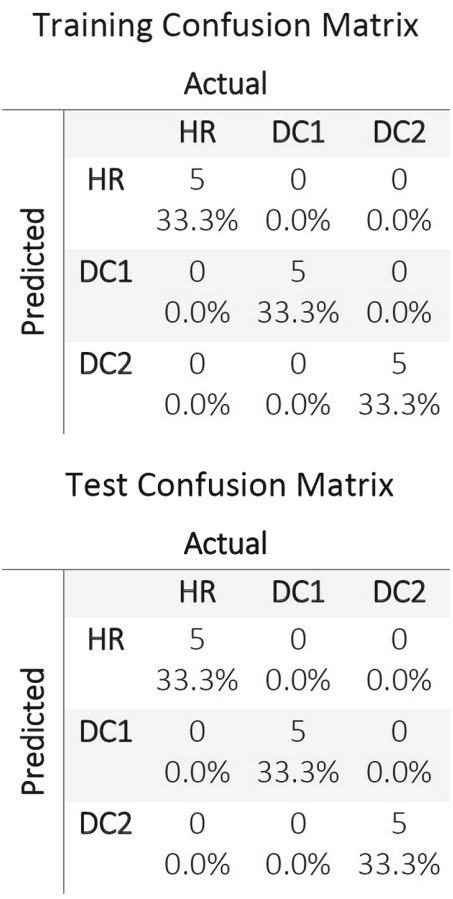

Fig. 7. Confusion matrices for training and testing data using all extracted features.

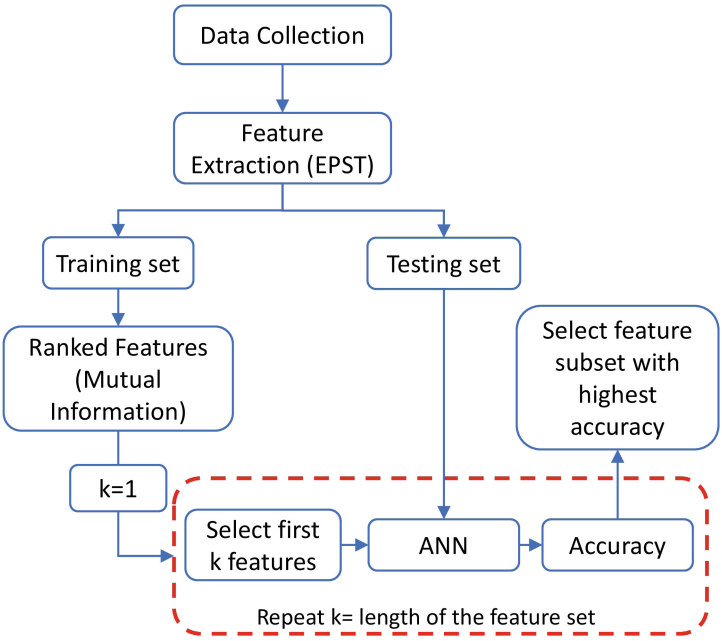

Fig. 8. Process diagram of Feature selection algorithm.

6 Conclusion

The present contribution showed that phase space and the orbit plots have valuable information regarding the condition of the rotor system. Additionally, it was demonstrated that characterizing the density distribution of the orbit plots by using orthogonal functions basis can be useful in describing the behavior of the dynamic system. A horizontal rotating machine operating at a constant speed was studied in order to detect fatigue cracks with different levels. Outstanding performance was achieved by using the EPST features with 100% classification accuracy. EPST is based on the density distribution of the projection of the phase space topology. Furthermore, feature selection was performed in order to select an optimal subset. Remarkable classification results were accomplished by only using three features. We feel that the EPST method is overwhelmingly successful in its capability to be applied to incipient crack detection with minimal knowledge about the rotating machine. It is worth mentioning that in real applications, without any knowledge about the crack presence, any malfunctions can be detected by using the conveyed methodology.

Acknowledgments. THM and CN are supported by the US Office of Naval Research under the grant ONR N0014-15-1-2311 with Capt. Lynn Petersen as the Program Manager. We deeply appreciate this support and are humbled by ONR's enthusiastic recognition of the importance of this research. AAC and VS are thankful for the

financial support provided to the present research effort by CNPq (574001/2008-5) and FAPEMIG (TEC-APQ-3076-09/TEC-APQ-02284-15) through INCT-EIE.

References

1. Saavedra, P.N., Cuitiño, L.A.: Crack detection and vibration behavior of cracked beams. Comput. Struct. **79**, 1451–1459 (2001)
2. Carden, E.P., Fanning, P.: Vibration based condition monitoring: a review. Struct. Health Monit. **3**, 355–377 (2004)
3. Doebling, S.W., Farrar, C.R., Prime, M.B.: A summary review of vibration-based damage identification methods. Shock Vibr. Dig. **30**, 91–105 (1998)
4. Darpe, A.K., Gupta, K., Chawla, A.: Coupled bending, longitudinal and torsional vibrations of a cracked rotor. J. Sound Vibr. **269**, 33–60 (2004)
5. Sawicki, J.T., Storozhev, D.L., Lekki, J.D.: Exploration of NDE properties of AMB supported rotors for structural damage detection. J. Eng. Gas Turbines Power **133**, 1–9 (2011)
6. Bachschmid, N., Pennacchi, P., Tanzi, E.: Cracked rotors: a survey on static and dynamic behaviour including modelling and diagnosis. Springer-Verlag, Germany (2010)
7. Cavalini Jr., A.A., Sanches, L., Baschmid, N., Steffen Jr., V.: Crack identification for rotating machines based on a nonlinear approach. Mech. Syst. Signal Process. **79**, 72–85 (2016)
8. Sawicki, J.T., Friswell, M.I., Pech, A.H., Wroblewski, A.: Condition monitoring of rotor using active magnetic bearing. In: ASME Turbo Expo 2008 (2008)
9. Eisenmann, R.C., Eisenmann Jr., R.C.: Machinery Malfunction: Diagnosis and Correction. Prentice Hall, NJ (1998)
10. Bently, D.E., Hatch, C.T.: Fundamentals of Rotating Machinery Diagnostics. Bently Pressurized Bearing Company, NV (2002)
11. Muszynska, A.: Rotordynamics. CRC Press, FL (2005)
12. Samadani, M., Kwuimy, C.A.K., Nataraj, C.: Model-based fault diagnostics of nonlinear systems using the features of the phase space response. Commun. Nonlinear Sci. Numer. Simul. **20**, 583–593 (2015)
13. Samadani, M., Kwuimy, C.A.K., Nataraj, C.: Diagnostics of a nonlinear pendulum using computational intelligence. In: ASME 2013 Dynamic Systems and Control Conference (2013)
14. Eckmann, J.P., Ruelle, D.: Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. **57**, 617–656 (1985)
15. Samadani, M., Mohamad, T.H., Nataraj, C.: Feature extraction for bearing diagnostics based on the characterization of orbit plots with orthogonal functions. In: 28th Conference on Mechanical Vibration and Noise (2016)
16. Mohamad, T.H., Kwuimy, C.A.K., Nataraj, C.: Discrimination of multiple faults in bearings using density-based orthogonal functions of the time response. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (2017)
17. Mohamad, T.H., Nataraj, C.: Gear fault diagnostics using extended phase space topology. In: Annual Conference of the Prognostics and Health Management Society (2017)
18. Mohamad, T.H., Samadani, M., Nataraj, C.: Rolling element bearing diagnostics using extended phase space topology. J. Vibr. Acoust. **140**(6) (2018)

19. Kappaganthu, K., Nataraj, C., Samanta, B.: Feature selection for bearing fault detection based on mutual information. In: Proceedings of the IUTAM Symposium on Emerging Trends in Rotor Dynamics, New Delhi, India, 23–26 March (2009)
20. Kappaganthu, K., Nataraj, C.: Feature selection for outer race bearing fault detection based on mutual information. Emerg. Trends Rotor Dyn. (2010)
21. Kappaganthu, K., Nataraj, C.: Feature selection for fault detection in rolling element bearings using mutual information. J. Vibr. Acoust. **133**, 1–11 (2011)
22. Kappaganthu, K., Nataraj, C.: Mutual information based feature selection from data driven and model based techniques for fault detection in rolling element bearings. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (2011)
23. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information: criteria of max-dependency, max-relevance and min-redundancy. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1226–1238 (2005)
24. Nelson, H.D., Nataraj, C.: Dynamics of a rotor system with a cracked shaft. J. Vibr. Acoust. Stress Reliab. Des. **108**, 189–196 (1986)