

Optimal Partitioning of Ultrasonic Data for Fatigue Damage Detection ★

Dheeraj Sharan Singh Soumik Sarkar Shalabh Gupta Asok Ray
dss240@psu.edu szs200@psu.edu szg107@psu.edu axr2@psu.edu

*Department of Mechanical Engineering
The Pennsylvania State University
University Park, PA 16802, USA*

Keywords: *Symbolic Dynamics, Optimal feature extraction, Fatigue Crack Initiation, Pattern Identification, Damage Classification*

Abstract—This paper presents an analytical tool for online fatigue damage detection in polycrystalline alloys that are commonly used in mechanical structures. The underlying theory is built upon symbolic dynamic filtering (SDF) that optimally partitions time series data for feature extraction and pattern classification. The proposed method has been experimentally validated on a fatigue test apparatus that is equipped with ultrasonics sensors and a traveling optical microscope for fatigue damage detection.

I. INTRODUCTION

Fatigue damage is one of the most commonly encountered source of structural degradation of mechanical structures, made of polycrystalline alloys. The process of fatigue damage accumulation in such materials normally takes place in two stages:

- Stage I (Grain Level): Development of microscopic cracks along grain boundaries or along crystal shear planes within a critical grain.
- Stage II (Macro Level): Transformation or coalescence of short cracks into a microcrack potentially leading to a normal breakage crack.

The life of a machine component in the above two stages depends on many factors such as loading conditions, material geometry, material strength, initial distribution of flaws within material. Given the random nature of loading and initial distribution of flaws, model-based analysis is infeasible for such materials under fatigue loads. However, many model-based techniques have been reported in recent literatures related to structural health monitoring [1]. These models are dependent on initial (randomly distributed) defects tool that would be reliable for real-time decision making. This can be accomplished by appropriate sensors combined with time series data analysis along with knowledge of current status of fatigue damage.

★ This work has been supported in part by NASA under Cooperative Agreement No. NNX07AK49A, and by the U.S. Army Research Laboratory and the U.S. Army Research Office under Grant No. W911NF-07-1-0376. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

A variety of damage detection techniques, based on different sensing devices (e.g., ultrasonics, acoustic emission, and eddy currents), have been proposed in recent literature for fatigue damage monitoring [2][3][4]. These sensing methods are not very suitable for real-time monitoring of damage. The other technique to detect microstructural changes within the materials used in complex electromechanical structures is ultrasonic sensing. Ultrasonic impedance is very sensitive to small microstructural changes occurring during early stages of fatigue damage evolution. Therefore, it is logical to detect the incipient damage from changes in statistical patterns of ultrasonic data due to gradual evolution of anomalies (i.e., deviations from the nominal behavior) in material micro structures [5].

Several techniques of pattern identification are available [6], but only very few of these tools (e.g., artificial neural networks, and principal component analysis) have been applied for online damage detection. Also, Such applications are largely restricted to the crack propagation regime after a substantial part of the useful service life has already been expended. Recently symbolic time series analysis(STSA) [7] [8] has been proposed and demonstrated successfully on ultrasonic time series data for early anomaly detection of fatigue damage. In this method ultrasonic signals are converted from the time domain to quasi-stationary symbolic sequences by symbolic dynamic encoding [9]. This procedure enables noise suppression due to symbolization by coarse graining [10], extraction of relevant information by *partitioning* and information compression into low-dimensional probability vectors. A major step in SDF-based feature extraction is partitioning of time-series data to generate symbols that are subsequently converted to feature vectors by use of the probabilistic finite state automata (PFSA) [7]. In this context, the major contributions of this article are listed below.

- Formulation of sensor data based fatigue damage detection in crack initiation and propagation regime.
- Optimization of partitioning in the SDF-based feature extraction setting.

The paper is organized into seven sections including the present one. Section II formulates the damage classification and detection problem in two different stages of fatigue evolution. Section III introduces the experimental setup and describe the test and data acquisition details. Section IV presents a brief background and formulation of SDF-based feature extraction for the current problem and Section V describes the partitioning optimization methodology. Section VI presents the results of validating the proposed classification scheme on the experimental data. Section VII summarizes the paper and makes major conclusions along with recommendations for future research.

II. FATIGUE DAMAGE CLASSIFICATION

Fatigue is broadly classified into two phases; namely crack initiation and crack propagation. As damage mechanism of these two phases are significantly different, similar detection methods may not work effectively to classify different damage levels in these two phases. Damage evolution in crack initiation phase is much slower, resulting in smaller change in ultrasonic signals as compared to that in the crack propagation phase. So to get maximum information from the ultrasonic data sequences, different partitioning is required in these two phases. Damage classification scheme for these phases are defined in the sequel.

A. Crack initiation

Since crack initiation predominantly forms a significant portion of the total life [11], especially in the high cycle fatigue, the quantification of fatigue damage during crack initiation is of paramount importance for safety, reliability, and maintenance of mechanical and aerospace structures. As this phase is caused by multiple small cracks, dislocations and other defects, any direct physical parameter to define damage in this regime is difficult. In the present classification method different classes are defined based on the fraction of crack initiation life. This is defined as the number of load cycles that a specimen sustains before appearance of the crack at surface. In this paper crack initiation is divided into three classes :

Class	Used crack initiation life fraction
1	0 to 0.5
2	0.5 to 0.8
3	0.8 to 1.0

Identifying the class of damage is very critical for the maintenance as well as the control operation. If the component is in first class of damage it means more than 50% life is still remaining, so it is safe to continue the operation. While the third class suggests that more than 80% is consumed so crack may appear soon, sending alarm to controller to take the appropriate action for safe operation. This can also be a basis for maintenance schedule.

B. Crack Propagation

The phase transition from crack initiation to crack propagation occurs when several small microcracks coalesce together to develop a single large crack that propagates under oscillating load. Several crack propagation models have been developed based on the inherent stochastic nature of fatigue damage evolution for prediction of the remaining useful life. Due to stochastic nature of material properties and

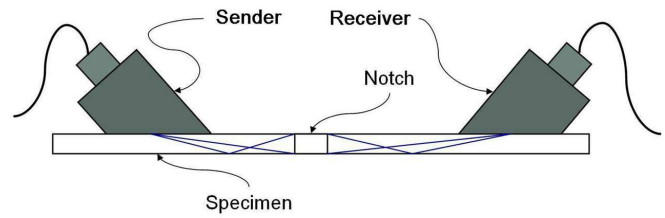


Fig. 1. Ultrasonic flaw detection scheme

operating conditions, any physical model requires knowledge of some parameter associated with the particular specimen or component. These parameters are random and can not be known beforehand. Crack propagation rate is a function of crack length; after a certain length, crack propagates unstably leading to catastrophic failure. So appropriate action must be taken before crack attains a critical size. In this classification problem damage is classified on basis of crack length. The crack propagation stage is divided into four classes:

Class	Crack Length
1	0.5 to 1.75 mm
2	1.75 to 3.5 mm
3	3.5 to 5.5 mm
4	more than 5.5 mm

III. EXPERIMENTAL PROCEDURE

The fatigue tests were conducted at a constant amplitude sinusoidal load for low-cycle fatigue, where the maximum and minimum loads were kept constant at 87MPa and 4.85MPa, respectively. For low cycle fatigue studied in this paper, the stress amplitude at the crack tip is sufficiently high to observe the elasto-plastic behavior in the specimens under cyclic loading. A significant amount of internal damage caused by multiple small cracks, dislocations and microstructural defects alters the ultrasonic impedance, which results in signal distortion and attenuation at the receiver end.

The optical images were collected automatically at every 200 cycles by the optical microscope which is always focussed in the crack tip. As soon as crack is visible by the microscope, crack length is noted down after every 200 cycles. Ultrasonic waves with a frequency of 5 MHz were triggered at each peak of the sinusoidal load to generate data points in each cycle. Ultrasonic data were collected at the peak of each sinusoidal load cycle, where the stress is maximum and the crack is open causing maximum attenuation of the ultrasonic waves. The slow time epochs for data analysis were chosen to be 1000 load cycles (i.e., ~80 sec) apart. To generate training and test data sample multiple experiments are conducted on different specimen. For each specimen all ultrasonic signals are labeled with either life fraction (Crack initiation) or crack length (crack propagation).

IV. SYMBOLIC DYNAMIC FILTERING (SDF)-BASED FEATURE EXTRACTION

This section reviews symbol dynamic filtering approach for anomaly detection [7]. The methodology of this symbolic feature extraction tool is reported in recent literature [12]; a brief outline of the procedure is succinctly presented here for completeness of the paper.

- The *fast scale* is related to the response time of process dynamics. Over the span of a given time series data sequence, the behavioral statistics of the system are assumed to remain invariant.
- The *slow scale* is related to the time span over which the process may exhibit non-stationary dynamics due to (possible) evolution of anomalies.
- Sensor time series data, generated from a physical system or its dynamical model, are collected at a slow-scale epoch and let it be denoted as \mathbf{q} . A compact (i.e., closed and bounded) region $\Omega \in \mathbb{R}^n$, where $n \in \mathbb{N}$, within which the stationary time series is circumscribed, is identified. Let the space of time series data sets be represented as $\mathcal{Q} \subseteq \mathbb{R}^{n \times N}$, where $N \in \mathbb{N}$ is sufficiently large for convergence of statistical properties within a specified threshold.
- Encoding of Ω is accomplished by introducing a partition $\mathbb{B} \triangleq \{B_0, \dots, B_{(|\Sigma|-1)}\}$ consisting of $|\Sigma|$ mutually exclusive (i.e., $B_j \cap B_k = \emptyset \forall j \neq k$), and exhaustive (i.e., $\bigcup_{j=0}^{|\Sigma|-1} B_j = \Omega$) cells, where each cell is labeled by symbols $\sigma_j \in \Sigma$ and $\Sigma = \{\sigma_0, \dots, \sigma_{|\Sigma|-1}\}$ is called the alphabet. This step enables transformation of the time series data $\{\mathbf{q}\}$ to a symbol sequence $\{\mathbf{s}\}$, consisting of the symbols σ_j in the alphabet Σ .
- A probabilistic finite state machine (*PFSA*), consisting of r states, is then constructed and the symbol sequence $\{\mathbf{s}\}$ is run through the *PFSA*. Thus, an $(r \times r)$ state transition matrix $\Pi \equiv [\pi_{jk}]$ is obtained at the slow-scale epoch, where $\pi_{jk} \geq 0$ is the transition probability from state j to state k of the *PFSA*. State probability vector $\mathbf{p} = [p_1 \dots p_r]$ is the left eigenvector corresponding to the unique unity eigenvalue of Π . The $(1 \times r)$ vector \mathbf{p} could be treated the extracted feature vector.

For anomaly detection using SDF, the nominal time-series is partitioned by one of the classical schemes (e.g., Uniform Partitioning (UP) or Maximum Entropy partitioning (MEP)) [7], [13], [14]. Then using the steps described before, a low dimensional feature vector \mathbf{p}^{nom} is constructed for the nominal slow-scale epoch. Similarly, from a time-series at a possibly anomalous epoch, feature vector $\mathbf{p}^{\text{off-nom}}$ is constructed using the same partitioning.

V. OPTIMIZATION OF PARTITIONING

Properties and variations of transformation from the symbol space to the feature space have been extensively studied in the Mathematics, Computer Science and especially Data Mining literature. Apparently, similar efforts have not been expended to investigate partitioning of time series data to optimally generate symbol blocks for pattern classification and anomaly detection. Wavelet space partitioning (WSP) is particularly effective for noisy data for large-dimensional dynamical systems [7]; maximum entropy partitioning (MEP) was used to generate symbol blocks from time series data by WSP. Subbu and Ray [14] introduced Hilbert-transform-based analytic signal space partitioning (ASSP) as an alternative to WSP, and Sarkar et. al [15] generalized ASSP

for symbolic analysis of noisy signals. Nevertheless, these partitioning techniques primarily provide a symbolic representation of the underlying dynamical system under a given quasi-stationary condition, rather than capturing the data-evolution characteristics due to a fault in the system. The partitioning optimization methodology elaborated in this section endeavors to overcome this shortcoming to make SDF, a robust data-driven feature extraction tool for pattern classification and fault detection.

In this paper, the wrapper method is adopted (i.e., the classification error is minimized on the training set for optimization) primarily because of the non-binary nature of the problem at hand and the possible non-Gaussian distribution of training samples in the feature space.

In a multi-class problem, ideally one should jointly minimize all the off-diagonal elements of the confusion matrix, while maximizing the diagonal elements. However, in that case, the dimension of the objective space blows up with increase in the number of classes which is obviously impractical. Therefore, two costs may be defined on the confusion matrix by using another penalty weighting matrix, elements of which denote the relative penalty values for different confusions in the classification process. Formally, let there be Cl_1, \dots, Cl_n classes of labeled time-series data given as the training set. A partitioning \mathbb{B} is employed to extract features from each sample and a k-NN classifier \mathbb{K} is used to classify them. After the classification process, the confusion matrix \mathbf{C} is obtained, where the value of its element c_{ij} denotes the frequency of data from class Cl_i being classified as data from Cl_j . Let \mathbf{W} be the weighting matrix, where the value of its element w_{ij} denotes the penalty incurred by the classification process for classifying a data set from Cl_i as a data set from class Cl_j . With these definitions, the cost due to expected classification error, $Cost_E$ is defined as:

$$Cost_E = \frac{1}{N_s} \left(\sum_i \sum_j w_{ij} c_{ij} \right) \quad (1)$$

where N_s is the total number training samples including all classes. The outer sum in the above equation sums the total penalty values for misclassifying each class Cl_i . Thus $Cost_E$ is related to the expected classification error.

It is implicitly assumed in many supervised learning algorithms that the training data set is a statistically similar representation of the whole data set. However, this assumption may not be very accurate in practice. A natural solution to this problem is to choose a feature extractor that minimizes the worst-case classification error [16] as well. In the present setting, the cost due to worst-case classification error, $Cost_W$ can be defined as:

$$Cost_W = \max_i \left(\frac{1}{N_i} \sum_j w_{ij} c_{ij} \right) \quad (2)$$

where N_i is the number of training samples in class Cl_i . Finally, in accordance with multi-objective optimization literature, an overall cost $Cost_O$ is defined as a linear

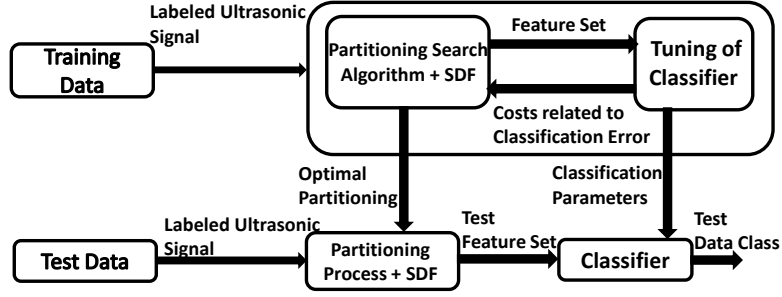


Fig. 2. General Framework for optimization of feature extraction

combination of $Cost_E$ and $Cost_W$ with parameter $\alpha \in [0, 1]$,

$$Cost_O = \alpha Cost_E + (1 - \alpha) Cost_W \quad (3)$$

Figure 2 depicts the general outline of the classification process. Labeled time series data from the training set are partitioned and the generated low-dimensional feature vectors (via symbolization and PFSA construction) are fed to the classifier. After classification, the training error cost, defined in Eq. (3), is computed and fed back to the feature extraction block. During classification, the classifier may be tuned to obtain better classification rates. For example, for k-NN classifiers [17], choice of neighborhood size or the distance metric can be tuned. The partitioning is updated to reduce the cost based on the feedback. The iteration is continued until the set of optimal partitioning in a multi-objective scenario and the correspondingly tuned classifier are obtained.

Ideally, the optimization procedure involves construction of the Pareto front by minimizing $Cost_O$ for different values of α and the user should choose a particular value of α as the operating point. Thus, the optimal partitioning \mathbb{B}^* is the solution to the following optimization problem:

$$\mathbb{B}^* = \arg \min_{\mathbb{B}} Cost_O(\mathbb{B}) \quad (4)$$

Optimization Procedure

This paper adopts a sequential search based optimization technique of partitioning, previously proposed in [18]. As the continuity of the partitioning function with respect to the range space of classification error-related costs may not exist or at least are not adequately analyzed, gradient-based optimization methods are not explored in this paper. Suppose, the number of cells of the partitioning \mathbb{B} is $|\Sigma|$ and the region $\Omega \in \mathbb{R}^1$ (see Section IV) that circumscribes the one dimensional times series data space, is identified. To construct the search space, a suitably fine grid size depending on the data characteristics needs to be assumed. Each of the grid boundaries denotes a possible position of a partitioning cell boundary. Let the data space region Ω be divided into G grid cells, i.e., there are $G - 1$ grid boundaries excluding the boundaries of Ω . Thus, there are $|\Sigma| - 1$ partitioning boundaries to choose among $G - 1$ possibilities, i.e., the number of elements (i.e., $|\Sigma|$ -dimensional partitioning vectors) in the space \mathcal{P} of all possible partitioning is: $G - 1 C_{|\Sigma| - 1}$. It is clear from this analysis that the partitioning space \mathcal{P} may get significantly large with increase in values of G and $|\Sigma|$ (e.g., for $G \gg |\Sigma|$, computational complexity increases

approximately by a factor of $G/|\Sigma|$ with increase in value of $|\Sigma|$ by one). In that case, usage of a direct search approach becomes infeasible for evaluation of all possible partitioning. Therefore, this paper develops a sub-optimal solution for the multi-objective optimization problem described above.

The objective space consists of the scalar valued cost $Cost_O$, while decisions are made in the space \mathcal{P} of all possible partitioning. In the case of one-dimensional time series data, a partitioning consisting of $|\Sigma|$ cells may be succinctly represented by the $|\Sigma| - 1$ points that separate the cells. In the sequel, a Σ -cell partitioning \mathbb{B} is expressed as $\Lambda_{|\Sigma|} \triangleq \{\lambda_1, \lambda_2, \dots, \lambda_{|\Sigma| - 1}\}$, where, $\lambda_i, \forall i \in \{1, 2, \dots, |\Sigma| - 1\}$ denotes a partitioning boundary. The overall cost is dependent on a specific partitioning Λ and is denoted by $Cost_O(\Lambda)$. This sub-optimal partitioning scheme involves sequential estimation of the elements of the partitioning Λ .

The partitioning process is initiated by searching the optimal cell boundary to divide the data set into two cells, i.e., $\Lambda_2 = \{\lambda_1\}$, where λ_1 is evaluated as

$$\lambda_1^* = \arg \min_{\lambda_1} Cost_O(\Lambda_2) \quad (5)$$

Now, the two-cell optimal partitioning is given by $\Lambda_2^* = \{\lambda_1^*\}$. The next step is to partition the data into three cells as Λ_3 by dividing either of the two existing cells of Λ_2^* with the placement of a new partition boundary at λ_2 , where λ_2 is evaluated as

$$\lambda_2^* = \arg \min_{\lambda_2} Cost_O(\Lambda_3) \quad (6)$$

where $\Lambda_3 = \{\lambda_1^*, \lambda_2\}$. The optimal 3-cell partitioning is obtained as $\Lambda_3^* = \{\lambda_1^*, \lambda_2^*\}$. In this (local) optimization procedure, the cell that provides the largest decrement in $Cost_O$ upon further segmentation ends up being partitioned. Iteratively, this procedure can be extended to obtain the m cell partitioning as follows.

$$\lambda_{|\Sigma| - 1}^* = \arg \min_{\lambda_{|\Sigma| - 1}} Cost_O(\Lambda_{|\Sigma|}) \quad (7)$$

where $\Lambda_{|\Sigma|} = \Lambda_{|\Sigma| - 1}^* \cup \{\lambda_{|\Sigma| - 1}\}$ and the optimal $|\Sigma|$ cell partitioning is given by $\Lambda_{|\Sigma|}^* = \Lambda_{|\Sigma| - 1}^* \cup \{\lambda_{|\Sigma| - 1}^*\}$

This optimization procedure is monotonically decreasing in the cost function with every additional sequential operation, i.e., $Cost_O(\Lambda_{|\Sigma| - 1}^*) \geq Cost_O(\Lambda_{|\Sigma|}^*)$. This is evident from the following argument.

Let $\Lambda_{|\Sigma| - 1}^*$ be the $(|\Sigma| - 1)$ -cell partitioning that minimizes $Cost_O$. Based on the algorithm, $\Lambda_{|\Sigma|} = \Lambda_{|\Sigma| - 1}^* \cup \{\lambda_{|\Sigma| - 1}\}$.

If $\lambda_{|\Sigma|-1}$ is chosen such that it already belongs to $\Lambda_{|\Sigma|-1}^*$, then there would be no change in the partitioning structure and $Cost_O(\Lambda_{|\Sigma|}) = Cost_O(\Lambda_{|\Sigma|-1}^*)$. Since $Cost_O(\Lambda_{|\Sigma|}^*) \leq Cost_O(\Lambda_{|\Sigma|}) \forall \Lambda_{|\Sigma|}$, it follows that $\min(Cost_O(\Lambda_{|\Sigma|-1})) \geq \min(Cost_O(\Lambda_{|\Sigma|}))$. The monotonicity in the cost function allows formulation of a rule for termination of the sequential optimization algorithm. The process of creating additional partitioning cells is stopped if the cost decrease falls below a specified positive scalar threshold η_{stop} and the stopping rule is: $\Lambda_{|\Sigma|-1}^*$ is the optimal partitioning if

$$Cost_O(\Lambda_{|\Sigma|-1}^*) - Cost_O(\Lambda_{|\Sigma|}^*) \leq \eta_{stop}. \quad (8)$$

VI. RESULTS AND DISCUSSION

This section presents the classification results for the two stages of fatigue damage evolution : (i) crack initiation and (ii) crack propagation. As damage mechanism is entirely different in these two regimes, which is evident from big change in the ultrasonic signal in these two phases. To start with the classification process first time series data is converted to wavelet transform at a particular scale, time shift and basis function selected appropriately. Each transformed signal is normalized with the maximum amplitude of transformed signal obtained at the beginning of experiment, when there is no damage in the specimen. The normalization is done to remove the effect of variability in placement of ultrasonic sensors during different experiments.

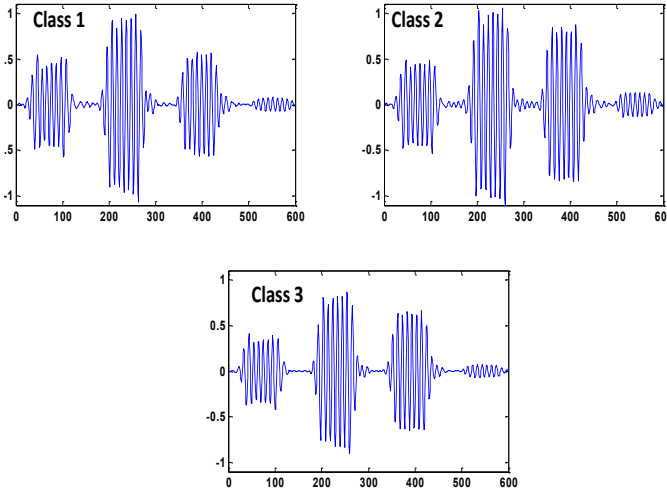


Fig. 3. Representative signal from different classes in crack initiation phase

A. Crack Initiation

The fatigue crack initiation phase is divided into three classes based on the fraction of crack initiation life. The three classes represent different life fraction as defined in section II-A. Fig. 3 shows the representative test data sample for each class. For partitioning the data, a weighing matrix \mathbf{W} needs to be defined to calculate the costs $Cost_E$ and $Cost_W$ from the confusion matrix for the training data set. The weights w_{ij} are defined as penalty for classifying a data sample originally from Cl_i is classified as a member of Cl_j . The weighing matrix has to follow two rules in the case of fatigue damage classification: (i) $w_{ii} = 0 \forall i \in \{1, 2, 3\}$, i.e.

there is no penalty for correct classification and (ii) $w_{ij} > w_{ji}$; if $i > j$, i.e. penalty for classifying a higher level damage as a lower level damage is much more than vice-versa. \mathbf{W} is given as:

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 0 & 1 \\ 6 & 3 & 0 \end{pmatrix}$$

The data space region Ω is divided into 100 grid cells, i.e., 99 grid boundaries excluding the boundaries of Ω . The sequential partitioning optimization procedure described in Section V is then employed to identify the optimal partitioning. The threshold value η_{stop} for stopping the algorithm is chosen to be 0.001 and the optimal alphabet size is found to be, $|\Sigma| = 7$. For SDF analysis, the depth for constructing *PFSA* sates is taken to be, $D = 1$ and features are classified by a k-NN classifier (with $k = 5$) using the Euclidean distance metric.

For comparison purpose, classical partitioning schemes, such as, Uniform Partitioning (UP) and Maximum Entropy Partitioning (MEP) are also used with the same alphabet size, $\Sigma = 7$. Finally, the confusion matrices for the Optimal, Uniform and Maximum Entropy Partitioning on the test data set are given by \mathbf{C}_{test}^{OptP} , \mathbf{C}_{test}^{UP} and \mathbf{C}_{test}^{MEP} respectively.

$$\mathbf{C}_{test}^{UP} = \begin{pmatrix} 92 & 8 & 0 \\ 3 & 87 & 10 \\ 0 & 10 & 90 \end{pmatrix}, \mathbf{C}_{test}^{MEP} = \begin{pmatrix} 95 & 5 & 0 \\ 3 & 90 & 7 \\ 0 & 11 & 89 \end{pmatrix}$$

$$\mathbf{C}_{test}^{OptP} = \begin{pmatrix} 95 & 5 & 0 \\ 4 & 91 & 5 \\ 0 & 9 & 91 \end{pmatrix}$$

Table I compares the values of $Cost_E$ and $Cost_W$ for OptP, UP and MEP on the test set. It is interesting to notice that both $Cost_E$ and $Cost_W$ have reduced compared to the classical partitioning schemes. Also observing the confusion matrices, it is evident that confusion between different classes are least for optimal partitioning as compared to other two partitioning methods.

TABLE I
COMPARISON OF CLASSIFICATION PERFORMANCES OF DIFFERENT PARTITIONING SCHEMES ON TEST DATA SET

Partitioning	$Cost_E$	$Cost_W$
OptP	0.1222	0.1866
UP	0.1489	0.2000
MEP	0.1357	0.2267

B. Crack Propagation

During crack propagation damage is defined by crack length. Training and test data set is classified based on crack length as all the ultrasonic signal is labeled with the corresponding crack length. This regime has been partitioned into four classes as defined in section II-B. The weighing matrix \mathbf{W} is defined as per rules described in section VI-A. \mathbf{W} matrix is defined as

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 3 & 0 & 1 & 2 \\ 6 & 3 & 0 & 1 \\ 9 & 6 & 3 & 0 \end{pmatrix}$$

Sequential optimization of partitioning has been carried on the training data set to find optimal partitioning. The optimal alphabet size is 6 with stopping threshold value $\eta_{stop} = 0.005$. The confusion matrices for Optimal, Uniform and Maximum Entropy Partitioning on the test data set are given by C_{test}^{OptP} , C_{test}^{UP} and C_{test}^{MEP} respectively. Table II shows the comparison of classification performances using different partitioning processes.

TABLE II
COMPARISON OF CLASSIFICATION PERFORMANCES OF DIFFERENT
PARTITIONING SCHEMES ON TEST DATA SET

Partitioning	$Cost_E$	$Cost_W$
OptP	0.255	0.5
UP	0.40333	0.68
MEP	0.31333	0.52

The confusion matrices obtained from the three different partitioning schemes are:

$$C_{test}^{UP} = \begin{pmatrix} 94 & 5 & 1 & 0 \\ 15 & 74 & 11 & 0 \\ 0 & 9 & 85 & 6 \\ 1 & 5 & 11 & 83 \end{pmatrix} \quad C_{test}^{MEP} = \begin{pmatrix} 93 & 7 & 0 & 0 \\ 12 & 82 & 6 & 0 \\ 0 & 7 & 89 & 4 \\ 1 & 3 & 7 & 89 \end{pmatrix}$$

$$C_{test}^{OptP} = \begin{pmatrix} 93 & 7 & 0 & 0 \\ 5 & 89 & 6 & 0 \\ 0 & 4 & 92 & 4 \\ 0 & 3 & 7 & 90 \end{pmatrix}$$

It is observed that both costs $Cost_E$ and $Cost_W$ have reduced for optimal partitioning as compared to maximum entropy and uniform partitioning. Also, from confusion matrix it is evident that optimal partitioning has improved the classification results. By close observation of confusion matrix we find that chances of higher damage level data samples being classified as a lower level damage is reduced by the current choice of weighing matrix.

VII. CONCLUSIONS AND FUTURE WORK

This paper presents symbolic dynamic filtering (SDF)-based feature extraction and pattern classification for fatigue damage detection in polycrystalline alloys that are common used in mechanical structures. Sequential optimal partitioning has been used for feature extraction to improve pattern classification rate and reducing classification errors. Different partitioning methods have been tested for detection and classification of the two phases of fatigue damage, namely, initiation and propagation. The classification results using optimal partitioning is compared with the other classical partitioning scheme (Uniform partitioning and Maximum entropy partitioning). Optimal partitioning results in better classification than the other classical partitioning schemes by reducing costs related to classification. The proposed method of fatigue damage classification has been validated to yield low classification error rates in presence of sensor noise and variability in the material properties that have a significant bearing on the ultrasonic impedance.

The advantages of the proposed fatigue detection scheme include real-time implementation where the classification information can be readily used to make maintenance &

control decisions, which is critical for the safe operation of machine or structural components. The following topics are suggested for future research.

- Use change in surface profile as a basis for classification in crack initiation regime.
- Usage of other types of classifiers (e.g., Support Vector Machines) and feature extraction methods, and performance comparison among different classifiers and feature extraction methods.
- Study the effect of phase transition on complexity of the fatigue process to get new insight about fatigue damage evolution phenomenon.
- Development of a control strategy based on the classifier information to increase the fatigue life;

REFERENCES

- [1] M. A. Meggiolaro and J. T. P. Castro, "Statistical evaluation of strain-life fatigue crack initiation predictions," *International Journal of Fatigue*, vol. 26, pp. 463–476, 2004.
- [2] S. Grondel, C. Delebarre, J. Assaad, J. P. Dupuis, and L. Reithler, "Fatigue crack monitoring of riveted aluminium strap joints by lamb wave analysis and acoustic emission measurement techniques," *NDT&E international*, vol. 35, pp. 339–351, 2006.
- [3] V. Zilberstein, K. Walrath, D. Grundy, D. Schlicker, N. Goldfine, E. Abramovici, and T. Yentzer, "Mwm eddy-current arrays for crack initiation and growth monitoring," *International Journal of Fatigue*, vol. 25, pp. 1147–1155, 2003.
- [4] J. Baram, "Fatigue-life prediction by an order statistics treatment of acoustic-emission signals," *Exper. Mech.*, vol. 33, pp. 189–194, 1993.
- [5] S. I. Rokhlin and J. Y. Kim, "In situ ultrasonic monitoring of surface fatigue crack initiation and growth from surface cavity," *International journal of fatigue*, vol. 25, pp. 41–49, 2003.
- [6] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. John Wiley & Sons Inc., 2001.
- [7] A. Ray, "Symbolic dynamic analysis of complex systems for anomaly detection," *Signal Processing*, vol. 84, no. 7, pp. 1115–1130, 2004.
- [8] S. Gupta, A. Ray, and E. Keller, "Symbolic time series analysis of ultrasonic data for early detection of fatigue damage," *Mechanical Systems and Signal Processing*, vol. 21, pp. 866–884, 2007.
- [9] D. Lind and M. Marcus, *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, United Kingdom, 1995.
- [10] R. Badii and A. Politi, *Complexity hierarchical structures and scaling in physics*. Cambridge University Press, United Kingdom, 1997.
- [11] J. C. Newman, "Fastran-ii - a fatigue crack growth structural analysis program," *NASA Technical Memorandum 104159*, Langley Research Center, 1992.
- [12] S. Gupta, A. Ray, S. Sarkar, and M. Yasar, "Fault detection and isolation in aircraft gas turbine engines: Part i - underlying concept," *Proceedings of the I Mech E Part G: Journal of Aerospace Engineering*, vol. 222, no. 3, pp. 307–318, May 2008.
- [13] V. Rajagopalan and A. Ray, "Symbolic time series analysis via wavelet-based partitioning," *Signal Processing*, vol. 86, no. 11, pp. 3309–3320, 2006.
- [14] A. Subbu and A. Ray, "Space partitioning via hilbert transform for symbolic time series analysis," *Applied Physics Letters*, vol. 92, p. 084107, February 2008.
- [15] S. Sarkar, K. Mukherjee, and A. Ray, "Generalization of Hilbert transform for symbolic analysis of noisy signals," *Signal Processing*, vol. 89, no. 6, pp. 1245–1251, 2009.
- [16] R. Alaiz-Rodriguez, A. Guerrero-Curieses, and J. Cid-Sueiro, "Minimax classifiers based on neural networks," *Pattern Recognition*, vol. 38, no. 1, pp. 29 – 39, 2005.
- [17] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [18] X. Jin, S. Sarkar, K. Mukherjee, and A. Ray, "Suboptimal partitioning of time-series data for anomaly detection," in *Conference on Decision and Control*, (Shanghai, China), December 2009.