RAPID COMMUNICATION

Symbolic dynamic analysis of surface deformation during fatigue crack initiation

Dheeraj Sharan Singh, Shalabh Gupta and Asok Ray

Department of Mechanical Engineering, The Pennsylvania State University, PA, USA
E-mail: dss240@psu.edu, szg107@psu.edu and axr2@psu.edu

Received 22 December 2009
Published 15 March 2010
Online at stacks.iop.org/MST/21/043003

Abstract
This rapid communication addresses early detection of fatigue damage evolution in polycrystalline alloys, based on the observation of surface deformation (e.g. roughness, linings and incisions). This method is well suited for calibration of other model-based and experimental tools for damage analysis and prediction in the fatigue crack initiation phase. To this end, the existing theory of symbolic dynamics-based feature extraction from time-series data is extended to the analysis of two-dimensional surface images. The resulting algorithms are experimentally validated on a fatigue-testing machine and a surface interferometer in the laboratory environment. The experiments have been conducted for analysis of statistical changes in the surface profiles due to gradual evolution of deformation in specimens, made of the 2024-T6 aluminum alloy.

Keywords: symbolic dynamics, surface deformation, fatigue crack initiation

1. Introduction

Fatigue life of structural materials is broadly classified into two phases [1]: (i) crack initiation and (ii) crack propagation. This classification assumes that 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 [2]. However, the estimation of remaining life in the crack initiation period is still an open issue because of the difficulties in measuring microstructural changes during early stages of fatigue damage. Since crack initiation predominantly forms a significant portion of the total life [1, 3], especially in the high cycle fatigue, the estimation of fatigue damage during crack initiation is of paramount importance for safety, reliability and maintenance of mechanical and aerospace structures.

Several methods have been proposed in recent literature for detection of fatigue crack initiation. Makabe et al [4] used a strain waveform for detection of short cracks in the order of 0.5–1.0 mm on test samples with a small hemispherical pit. Katayama et al [5] detected small surface cracks using an ac potential method on notched test samples. Tohmyoh et al [6] detected small cracks using surface shear-horizontal waves in test samples with a surface slit subjected to rotating and bending loads. Zilberstein et al [7] performed studies on the application of meandering winding magnetometer eddy current sensors for detection of crack initiation. Recent studies by Gupta et al [8, 9] on polycrystalline alloys have shown that ultrasonic signals can be used to capture small defects inside the material surface that occur during the early stages of fatigue damage. Since material characteristics (e.g. voids, dislocations and short cracks) influence ultrasonic impedance,
accumulation of small faults in test specimens changes the signal characteristics at the receiver end.

This paper presents an alternative method for detection of fatigue damage evolution by the analysis of surface deformation in polycrystalline alloys. The surface conditions, such as roughness, linings and incisions, have strong effects on fatigue strength and life of structural components [1]. Since surface deformation plays a significant role in the fatigue crack initiation phase, detection of changes in the surface profiles provides an early warning of forthcoming widespread fatigue damage. In this paper, a surface interferometer has been used to measure the surface profiles of specimens that are subjected to oscillating load. The surface profiles are generated as two-dimensional images where each pixel on an image represents the height of the local region. Although the current state of the art of instrumentation for this surface analysis method is not suitable for in situ applications, it supplements the information generated from alternative sensing methods (e.g. ultrasونics) by providing a means for calibration and validation of model-based and experimental tools of fatigue damage analysis and prediction in the crack initiation phase.

Symbolic dynamics-based methods have been shown to be useful for the analysis of sensor time series data [10]. The recent literature has introduced an information-theoretic method of feature extraction from time series, called symbolic dynamic filtering [8, 11] that is built upon the concepts of Symbolic Dynamics [12] and Statistical Mechanics [13]. Key advantages of this symbolic dynamics-based method of feature extraction are robustness to measurement noise, capability for early detection of damage, information compression as low-dimensional statistical patterns and real-time execution on commercially available computational platforms. These feature extraction algorithms, based on (one-dimensional) time series data, have been experimentally validated for early detection of fatigue damage in different applications such as mechanical and aerospace systems [8, 9] and electronic systems [11]. Along this line, major contributions of the paper are listed below:

- extension of the theory of symbolic dynamics-based feature extraction in time-series data to the analysis of two-dimensional images;
- experimental validation of the surface-deformation-analysis method for detection of evolving damage in the 2024-T6 aluminum alloy in the crack initiation phase.

The paper is organized into five sections including the present one. Section 2 presents the symbolic dynamics-based method for feature extraction from two-dimensional sensor data. Section 3 describes the test apparatus and experimental procedure for validation of the underlying algorithms. Section 4 presents the results obtained from the analysis of surface profiles of an aluminum specimen. The paper is concluded in section 5 along with recommendations for future research.

2. Symbolic dynamics and encoding

This section presents the underlying concepts and salient properties of the symbolic dynamic filtering (SDF) method. While the details have been reported in previous publications for (one-dimensional) time series analysis [8, 11, 13], this section extends the underlying concepts of SDF for analysis of (two-dimensional) surface images.

In the symbolic dynamics literature [12], the primary objective is to represent time series data from a dynamical system as a symbol sequence. Let \( \Omega \in \mathbb{R}^n \) be a compact (i.e. closed and bounded) region, within which the trajectory of the dynamical system is circumscribed. The region \( \Omega \) is partitioned into a finite number of (mutually exclusive and exhaustive) cells. Let the cell, visited by the trajectory, be denoted as a random variable taking a symbol value from the alphabet \( \Sigma \). The trajectory of the dynamical system described as \( \{x_0, x_1, \ldots, x_k, \ldots\} \), where \( x_k \in \Omega \), passes through or touches one of the cells of the partition and the corresponding symbol is assigned to it. The resulting symbol sequence is denoted as \( \{\sigma_0, \sigma_1, \ldots, \sigma_k, \ldots\} \), where \( \sigma_k \in \Sigma \). In this way, symbolic dynamics is viewed as coarse graining of the data space, which is subjected to (possible) loss of information due to granular imprecision of the partitioning cells; however, the essential robust features need to be preserved in the symbol sequences by an appropriate partitioning of the data space.

This paper aims to extend the concepts of symbolic pattern analysis for the purpose of feature extraction from two-dimensional data (i.e. surface images). In this regard, the essential concepts of data partitioning for symbolization and construction of probabilistic finite-state automata (PFSA) for feature extraction are described below.

2.1. Symbolization of two-dimensional surface profiles

The surface profiles generated from an interferometer, described in section 3, are two-dimensional images where each pixel represents the height of the local region. The \( x-y \) coordinates of the surface image denote the size of the image and the \( z \)-coordinate denotes the pixel values (i.e. the surface height). Symbolization of the surface profile is performed by partitioning the \( z \)-axis such that each pixel is transformed into a symbol that corresponds to the relevant cell of the partition. For example, if the alphabet is chosen as \( \Sigma = \{a, b, c, d\} \), i.e. \( |\Sigma| = 4 \), then there are three partitioning planes and these planes partition the ordinate along the \( z \)-axis of the image into three mutually exclusive and disjoint regions, and a symbol from the alphabet \( \Sigma \) is assigned to each pixel corresponding to the region where its intensity falls (see figure 1). In general, the choice of alphabet size depends on specific data and experiments. The partitioning can be performed by different methods [8] such as the maximum entropy partitioning (MEP) and the uniform partitioning (UP) to generate a symbol image.

The two-dimensional data of surface image are partitioned such that the ordinates between the maximum and minimum of the data along the \( z \)-axis are divided into regions by several parallel planes. These mutually disjoint regions form a partition, where each region is labeled with one symbol from the alphabet \( \Sigma \). If the data point lies in a particular region, it is coded with the symbol associated with that region. Thus, a two-dimensional picture of symbols is created from a
given surface profile. If the partitioning planes are separated by equal-sized intervals, then the partition is called uniform partitioning (UP). Intuitively, it is more reasonable if the information-rich regions of a data set are partitioned finer and those with sparse information are partitioned coarser. To achieve this objective, the maximum entropy partitioning (MEP) method has been adopted in this paper such that the entropy of the generated symbol sequence is maximized. The procedure for selection of the alphabet size $|\Sigma|$, followed by generation of a MEP, has been reported in [8].

For the purpose of feature extraction and damage detection, the partitioning is performed at the nominal condition (i.e. time epoch $t_0$), and subsequently it is kept constant for all time epochs $\{t_1, t_2, \ldots, t_k, \ldots\}$, i.e. the structure of the partition is fixed at the nominal condition. In other words, the partitioning structure generated at the nominal condition serves as the reference frame for data analysis at subsequent time epochs [11]. The method of symbolization of two-dimensional surface data is described below.

**Definition 2.1 (Surface profile).** Let $\mathcal{H} \triangleq \{(i, j) : i, j \in \mathbb{N}, 1 \leq i \leq m, 1 \leq j \leq n\}$ be the set of coordinates of a surface profile consisting of $(m \times n)$ pixels. Let $\mathcal{R}$ denote the interval that spans the range of surface heights. Then, a surface profile is defined by a map $S : \mathcal{H} \rightarrow \mathcal{R}$.

**Definition 2.2 (Symbolization).** Given $\Sigma$ to be the alphabet, let the partitioning of the interval $\mathcal{R}$ be defined by a map $P : \mathcal{R} \rightarrow \Sigma$. Then, the symbolization of an image is defined by a map $S_\Sigma \equiv P \circ S$ such that $S_\Sigma : \mathcal{H} \rightarrow \Sigma$ that maps each pixel of the image to a symbol in $\Sigma$.

The left-hand plate in figure 1 shows the intensity plot of a surface profile measured by the interferometer and the right-hand plate in figure 1 shows the corresponding symbol image. In this way, a symbol is assigned to each pixel on the surface profile according to the partitioning and the surface profile is converted to the symbol image. The partitioning of sensor data to generate symbolic representations enables robust feature extraction in a dynamical system [8]. Furthermore, symbolization significantly reduces the memory requirements.

2.2. Construction of PFSA and feature extraction

This section presents the construction of a probabilistic finite-state automaton (PFSA) for feature extraction based on the symbol image generated from the surface profile. States of the PFSA represent different combinations of a group of symbols on the symbol image and the edges represent the transition probabilities between these states. Note that, for one-dimensional (i.e. time series) data analysis, the states denote symbol blocks (i.e. words) within a window of certain length.

Let us now extend the notion of states for the analysis of two-dimensional data sets via construction of a ‘state image’ from a ‘symbol image’.

**Definition 2.3 (State).** Let $\mathcal{W} \subset \mathcal{H}$ be a two-dimensional window of size $(\ell \times \ell)$ and its size is denoted as $|\mathcal{W}| = \ell^2$. Then, the state of a symbol block located in the window $\mathcal{W}$ is defined as the configuration $q = S_\Sigma(\mathcal{W})$.

Let the set of all possible states in a window $\mathcal{W} \subset \mathcal{H}$ be denoted as $\mathcal{Q} \triangleq \{q_1, q_2, \ldots, q_{|\mathcal{Q}|}\}$, where $|\mathcal{Q}|$ is the number of (finitely many) states. Let us denote $\mathcal{W}_{i,j} \subset \mathcal{H}$ to be the window at a pixel location $(i, j) \in \mathcal{H}$ that represents, for example, the coordinates of the northwest corner pixel of the window $\mathcal{W}_{i,j}$. In this notation, $q_{i,j} = S_\Sigma(\mathcal{W}_{i,j})$ denotes the state at pixel $(i, j) \in \mathcal{H}$. Thus, every pixel $(i, j) \in \mathcal{H}$ corresponds to a particular state $q_{i,j} \in \mathcal{Q}$ on the image. Then, $|\mathcal{Q}|$ is bounded as $|\mathcal{Q}| \leq |\Sigma|^{|\mathcal{W}|}$; the inequality is due to the fact that some of the states might have zero probability of occurrence.

Every pixel in the image $\mathcal{H}$ is mapped to a state (i.e. a two-dimensional word or block of symbols), excluding the pixels that lie at the periphery depending on the window size. This concept of state formation facilitates capturing of long range dynamics (i.e. word to word interactions) on a symbol image. Figure 2 shows an illustrative example of the transformation of a ‘symbol image’ to the ‘state image’ based on a sliding window $\mathcal{W}$ of size $(2 \times 2)$.

In general, a large number of states would require a high computational capability and hence might not be feasible for real-time applications. The number of states, $|\mathcal{Q}|$, increases with the window size $|\mathcal{W}|$ and the alphabet size $|\Sigma|$. For example, if $\ell = 2$ and $|\Sigma| = 4$, then the total number of states is $|\mathcal{Q}| \leq |\Sigma|^{|\mathcal{W}|} = 256$. Therefore, for computational efficiency, it is necessary to compress the state set $\mathcal{Q}$ to an effective smaller set $\mathcal{O} \triangleq \{o_1, o_2, \ldots, o_{|\mathcal{O}|}\}$ [8] that enables mapping of two or more different configurations in a window $\mathcal{W}$ to a single state. State compression must preserve sufficient

![Figure 1. Symbolization of the surface profile.](image-url)
local information on lattice spin systems while still capturing signal dynamics [8]. Sensitive while a small alphabet could miss the information of tradeoff between robustness to noise and capability to detect available computational power, and is made by an appropriate depends on specific applications and noise level as well as the principle of sliding block code [12]. As shown in figure 4, the region enclosed within a box in the upper-left corner contains three adjacent pairs, namely, states \{o_1, o_2\}, \{o_1, o_3\} and \{o_1, o_4\}. Therefore, in this neighborhood, the corresponding counters of occurrences, i.e. \(N(o_1, o_2)\), \(N(o_1, o_3)\) and \(N(o_1, o_4)\) are increased by 1. The window slides to the right and to the bottom to cover the entire state image, and the transition probabilities \(\varphi(o_1|o_k)\) \(\forall o_1, o_k \in \mathcal{O}\) are computed using equation (1). Thus, PFSA are constructed by sliding a window on the symbol image as depicted in figure 4. This procedure generates a stochastic matrix \(\Pi\) as

\[
\Pi = \begin{bmatrix}
\varphi(o_1|o_1) & \cdots & \varphi(o_1|o_4) \\
\vdots & \ddots & \vdots \\
\varphi(o_4|o_1) & \cdots & \varphi(o_4|o_4)
\end{bmatrix},
\]

(2)

where \(\Pi \equiv [\pi_{jk}]\) with \(\pi_{jk} = \varphi(o_j|o_k)\). Note \(\pi_{jk} \geq 0 \forall j, k\) and \(\sum_k \pi_{jk} = 1 \forall j\).

In order to construct a low-dimensional feature vector, the stationary state probability vector \(\mathbf{u}\) is obtained as the left eigenvector corresponding to the (unique) unity eigenvalue of the (irreducible) stochastic transition matrix \(\Pi\). The state probability vector \(\mathbf{u}\) is called the ‘pattern vector’ and is generated at different time epochs from the corresponding state transition matrices. The scalar measures generated by calculating distances \(d(\mathbf{u}, \mathbf{u}')\) between the current pattern vector and pattern vector at the nominal condition are used to denote the damage measures. Examples of \(d(\mathbf{u}, \mathbf{u}')\) are Euclidean norm and Kullback–Leibler divergence [14].

3. Description of the experimental setup

This section briefly describes the experimental apparatuses used for validation of the underlying concepts presented in
The vertical scan speeds up to 10 μm per pixel, takes images with a resolution of 2 μm per pixel at a distance of ~20 mm. The primary objective of this optical microscope is to monitor the specimen for detection of a surface crack that grows and leads to crack propagation. However, small surface deformations during crack initiation cannot be detected by this microscope.

A three-dimensional profile of the specimen surface is generated by a NewView 5000 surface interferometer apparatus that measures the surface heights of the scanned area ranging from 1 nm to 5000 μm with a resolution of 0.1 nm at the vertical scan speeds up to 10 μm s⁻¹. It can scan areas up to 50 mm × 50 mm using its unique image stitching capabilities. This interferometer uses a non-contact scanning method using light interferometry to acquire ultrahigh resolution images. It uses a closed-loop piezoelectric scanner employing low-noise capacitive sensors to ensure accurate and repeatable linear motion over the full range of a scanned area. The surface interferometer is equipped with the software MetroPro that is used to operate the interferometer and to store the surface profile data in a convenient format for further data processing. Figure 5 shows a plot of the surface profile of a scanned area on the specimen surface.

The test specimen, made of the 2024-T6 aluminum alloy, is 3 mm thick and 50 mm wide and has a slot of 1.58 mm × 4.57 mm at the center of one edge. The side notch is constructed in the specimen to act as a stress riser that helps the crack to be initiated at its end. The specimen is designed to fail within a relatively short period so that the failure characteristics from crack initiation to crack propagation can be conveniently analyzed in the laboratory environment.

Surface deformations are most dominant in the notch region due to higher stress concentration. In the current experiments, the interferometer is configured to measure surface profile heights from 1 nm to 100 μm. The surface profile data are acquired and stored for analysis by symbolic dynamic filtering (see section 2). This surface profile corresponds to the nominal condition when no damage has accumulated in the specimen. After acquiring the surface profile, the specimen is mounted on the fatigue damage testing apparatus, where it is subjected to a constant amplitude sinusoidal load at a frequency of 50 Hz with the maximum load of 14 000 N and the minimum load of 500 N. Upon completion of 10 000 cycles, the specimen is dismounted from the fatigue damage testing apparatus and remounted on the surface interferometer to monitor the changes in its surface profile due to fatigue loading. The surface profile data are acquired and this procedure is repeated after every 10 000 cycles until a crack becomes clearly visible on the specimen surface by the optical microscope. These data sets are processed by the SDF method to detect and quantify the evolution of fatigue damage during crack initiation.

4. Results and discussion

Following the SDF procedure (see section 2), the symbol image is obtained by maximum-entropy partitioning (MEP) of the surface profile data. The data range of two-dimensional surface profiles is partitioned into |Σ| = 6 segments. For ℓ = 2 and m = 1, i.e., the total number of states equal to |Σ|m = 6, the SDF method is able to capture the gradual evolution of surface profiles. The six subplots in figure 6 show the evolution of fatigue damage at different time epochs that are separated by 10 000 cycles. The top plate in each subplot of figure 6 shows the surface profile obtained from the interferometer from a rectangular region (1 mm × 1.14 mm) near the notch on the specimen. The corresponding bottom plate shows the histogram of probability distribution generated using symbolic dynamic filtering of the surface data. The top plate as seen in figure 6(a) shows the surface profile image at the nominal condition before the start of the fatigue damage experiment. At this condition, the damage measure is taken to be zero, which is considered as the reference point with the available information on potential damage being negligible. This condition is represented as the uniform distribution (i.e., maximum entropy) in the histogram at the bottom plate of figure 6(a). The top plate in figure 6(b) shows small changes in the surface profile of the specimen at ~10 000 cycles, and the corresponding bottom plate exhibits deviations from the nominal probability distribution. This proves that the surface profile, measured by the interferometer, produces credible damage information during crack initiation. Subsequently, the top plates in subplots (c), (d) and (e) of figure 6 at ~20, ~30 and ~40 kilocycles, respectively, exhibit noticeable evolution of surface deformation; this is also in agreement with evolving

---

Figure 5. A three-dimensional surface profile generated from the interferometer.
deviations in the corresponding histograms of probability distributions. At ~50 kilocycles, the optical microscope in the fatigue test apparatus detected the appearance of a single large crack on the specimen surface, which indicates the onset of crack propagation phase. It is hypothesized that multiple small cracks would have developed underneath before their coalescence and appearance as a single large crack on the surface. Although the small cracks underneath the surface were not visible from the optical microscope, the interferometer detected small deformations that are precursors of a surface crack. The top plate in figure 6(f) shows the surface profile at ~50 kilocycles and the bottom plate shows further deviation of the probability distribution from the nominal condition. The crack propagation stage starts at ~50 kilocycles and thereon the damage information is easily available from the optical microscope and other sensors.
The gradual evolution of surface deformation at different time epochs is represented by the corresponding changes in the histograms of probability distributions that are generated using symbolic dynamic filtering of the surface profiles. It is evident that the surface profile images provide significant information during early stages of fatigue crack initiation. Figure 7 shows the damage measure curve that indicates gradual evolution of fatigue damage based on the surface profile data generated from the interferometer. The damage measure provides relative information that quantifies damage during crack initiation with respect to the nominal condition.

5. Conclusions and future work

This paper presents the concept of detection and quantification of damage during fatigue crack initiation in mechanical and aerospace structures, made of polycrystalline alloys. The surface profile data, obtained from the interferometer, are combined with the two-dimensional SDF to detect and quantify early damage in structural materials. This concept is validated on a fatigue test apparatus and a surface interferometer in the laboratory environment. The results demonstrate that the proposed method is suitable for detection of fatigue damage at the crack initiation stage. It is also capable of quantifying the damage rate and of revealing the characteristics of fatigue damage evolution in polycrystalline alloys.

The proposed method is a step toward building a reliable instrumentation system for early detection and estimation of remaining useful life of stressed structures. While there are many other research issues, the following tasks are recommended for future work.

- Calibration of ultrasonic sensors in the crack initiation phase by comparison with the proposed surface-deformation method.
- Damage estimation based on the ultrasonic sensing and surface profile measurements.
- Quantification of structural durability of critical components for the assessment of plant operational reliability.
- Development of real-time instrumentation for in situ measurements of surface profile of machinery components.

Acknowledgments

This work has been supported in part by the US Office of Naval Research (ONR) under grant no N00014-09-1-0688 and by NASA under Cooperative Agreement no NNX07AK49A. 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.

References