

Symbolic dynamic filtering for image analysis: theory and experimental validation

Aparna Subbu · Abhishek Srivastav · Asok Ray · Eric Keller

Received: 24 January 2009 / Revised: 25 May 2009 / Accepted: 26 May 2009 / Published online: 23 June 2009
© Springer-Verlag London Limited 2009

Abstract Recent literature has reported the theory of symbolic dynamic filtering (SDF) of one-dimensional time-series data and its various applications for anomaly detection and pattern recognition. This paper extends the theory of SDF in the two-dimensional domain, where symbol sequences are generated from image data (i.e., pixels). Given the symbol sequence, a probabilistic finite state automaton (PFSA), called the *D*-Markov machine, is constructed on the principles of Markov random fields to incorporate the spatial information in the local neighborhoods of a pixel. The image analysis algorithm has been experimentally validated on a computer-controlled fatigue test apparatus that is equipped with a traveling optical microscope and ultrasonic flaw detectors. The surface images of test specimens, made of a polycrystalline alloy, are analyzed to detect and quantify the evolution of fatigue damage. The results of two-dimensional SDF analysis are in close agreement with those obtained from analysis of one-dimensional time-series data from the

ultrasonic sensor, which are simultaneously generated from the same test specimen.

Keywords Image analysis · Two-dimensional symbolic analysis · Statistical pattern recognition · Fatigue damage analysis

1 Introduction

Various methods of image analysis have been reported in both non-statistical and statistical signal processing literature on texture analysis, segmentation, edge detection and image understanding [1]. For example, Gabor functions and multiresolution representations via wavelets and other time-frequency analysis methods have been extensively used for image classification and image retrieval from large databases [2, 3]. Tools of hidden Markov models (HMM) also have been applied for image segmentation and compression [4].

This paper proposes and experimentally validates a new methodology of image analysis, where the underlying concept is built upon symbolic dynamic filtering (SDF) [5]. In the absence of a sufficiently accurate and reliable model of the dynamical system under consideration, SDF relies on the time-series data obtained from sensors that are used for condition monitoring and control. While large and abrupt deviations from the nominal condition are relatively easy to detect, the principal aim of SDF is to make inferences on gradually evolving anomalies (i.e., deviations from the nominal expected behavior of a dynamical system as a result of parametric or non-parametric changes). From these perspectives, SDF is viewed to be a data-driven analysis tool that constructs an abstract probabilistic model to represent the complex system being monitored. It recognizes changes and evolution in the system by adapting the model to observed

This work has been supported in part by the U.S. Army Research Office under Grant No. W911NF-07-1-0376, 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.

A. Subbu · A. Srivastav · A. Ray (✉) · E. Keller
Department of Mechanical Engineering,
The Pennsylvania State University,
University Park, PA 16802, USA
e-mail: axr2@psu.edu

A. Subbu
e-mail: avs127@psu.edu

A. Srivastav
e-mail: axs964@psu.edu

E. Keller
e-mail: eek105@psu.edu

data set. The statistical model is constructed on the principles of symbolic dynamics and fixed-structure, fixed-order Markov chains [6]. Recently, SDF has been applied for pattern recognition and anomaly detection in diverse applications (e.g., electronic circuits [7], induction motors [8,9], robot behavior recognition [10] and fatigue testing [11]). Comparisons with standard pattern recognition techniques (e.g., artificial neural networks, principal component analysis, kernel regression analysis, and particle filtering) have demonstrated superior performance of SDF for detection of anomalous behavior [12]. In all these applications, the derived data are in the form of an one-dimensional time-series.

It is the aim of this paper to extend the SDF methodology from one-dimensional to two-dimensional domains, an example of which is image data analysis. In this context, the theory of analytic signal space partitioning (ASSP) [13] is also extended to generate a two-dimensional array of symbols. Extension of SDF to image analysis allows not only detection of changes in the patterns of the observed image sequences, but also quantifies these changes in terms of an anomaly measure. This is potentially advantageous in applications where a quantification of the deviation from the nominal patterns facilitates decision making. In addition, SDF analysis of images can be used for traditional image classification, where the objective is to classify the observed image into one of the pre-defined categories. The concepts developed in this paper are validated for detection and estimation of surface damage in specimens of polycrystalline alloys. The results of image analysis for fatigue damage detection are found to be in close agreement with those derived from the (one-dimensional) time series of an ultrasonic sensor that also monitors fatigue damage in the same specimen.

The rest of the paper is organized into the following sections. Section 2 briefly reviews the underlying theories of SDF-based analysis as applied to one-dimensional time-series data. Section 3 develops the SDF methodology to address the two-dimensional nature of images. Section 4 provides a brief overview of the fatigue test apparatus, on which the image analysis algorithm is validated. It also formulates the image analysis problem used for validation of SDF analysis. Section 4.2 presents the results and compares them with time-series data obtained from ultrasonic sensors. The paper is summarized and concluded in Sect. 5 along with recommendations for future research.

2 Symbolic dynamic filtering (SDF) of one-dimensional data

While symbolic dynamic filtering (SDF) of time-series data has been reported in previous publications [5, 13, 14], the core concept is very briefly reviewed here to provide a succinct background for image analysis via two-dimensional SDF

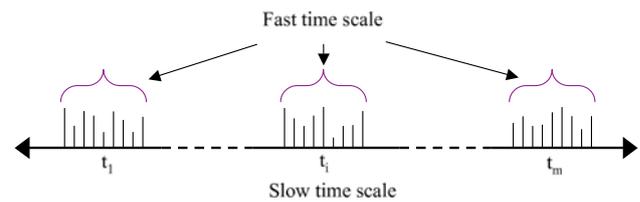


Fig. 1 Pictorial view of two-time-scales

that is addressed in Sect. 3. The reported SDF analysis of time-series data in a dynamical system has been formulated as a two-time-scale problem, wherein the system dynamics are assumed to have quasi-stationary behavior on the fast scale and any observable non-stationary behavior of the system occurs on the slow scale. In other words, the dynamical system is assumed to be statistically quasi-stationary in the fast scale, whereas nonstationarities (e.g., due to parametric or non-parametric changes in the system) occur on the slow scale. A pictorial view of the two-time-scale concept is presented in Fig. 1; however, the notion of these two scales is dependent on the dynamical system under observation.

Symbolic time-series analysis constructs probabilistic finite state automata (PFSA), known as the D -Markov machine [5], based on the statistical information in the symbol sequence generated from time-series data on the fast scale at individual slow-scale epochs. Thus, SDF tracks evolution of anomalies, on the slow scale by observing statistical changes in the PFSA.

2.1 Analytic signal space partitioning (ASSP)

The first step in SDF is symbol sequence generation from a symbol alphabet Σ via partitioning of the observed sequence of time-series data, which is encoded as a sequence of coarse-grained symbols. The encoding process involves construction of a partition $\mathcal{P} \triangleq (P_0, P_1, \dots, P_{|\Sigma|})$ on a compact region Ω in the state space of the dynamical system, which consists of $|\Sigma|$ mutually exclusive and exhaustive subsets of Ω , where $|\Sigma|$ is the cardinality of the alphabet Σ ; each partition segment P_i is denoted by a unique symbol $\sigma \in \Sigma$. Thus, the observed time-series data is encoded into a sequence of symbols, each of which belongs to the alphabet Σ .

Several partitioning techniques have been suggested in the literature for symbol generation, prominent among which is symbolic false nearest neighbors (SFNN) [15]. In the context of SDF, a major shortcoming of SFNN is that it becomes computation-intensive if the dimension of the underlying dynamical system is large; furthermore, if the time series becomes noise-corrupted, then the symbolic false neighbors rapidly grow in number and may erroneously require a large symbol alphabet to capture pertinent information. To circumvent these difficulties, analytic signal space partitioning (ASSP) has been proposed as an alternative to SFNN-based partitioning [13]. The objective of ASSP is to capture

the relevant statistical patterns in real time and its underlying theory is built upon Hilbert transform of the observed real-valued data sequence to generate the complex-valued analytic signal [16] as explained below.

Let $x(t)$ be a real-valued function whose domain is the real field $\mathbb{R} = (-\infty, +\infty)$. Then, the Hilbert transform [17] of $x(t)$ is defined as

$$\tilde{x}(t) = \mathcal{H}[x](t) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{x(\tau)}{t - \tau} d\tau \tag{1}$$

That is, $\tilde{x}(t)$ is the convolution of $x(t)$ with $\frac{1}{\pi t}$ over \mathbb{R} , which is represented in the Fourier domain as

$$\mathcal{F}[\tilde{x}](\xi) = -i \operatorname{sgn}(\xi) \mathcal{F}[x](\xi) \tag{2}$$

where $\operatorname{sgn}(\xi) = \begin{cases} +1 & \text{if } \xi > 0 \\ -1 & \text{if } \xi < 0 \end{cases}$ and $i \triangleq \sqrt{-1}$.

Given the Hilbert transform of a real-valued signal $x(t)$, the corresponding complex-valued analytic signal is defined as

$$\mathcal{A}[x](t) = x(t) + i\tilde{x}(t) \tag{3}$$

$$\mathcal{A}[x](t) = a(t) \exp(i\varphi(t)) \tag{4}$$

where $a(t)$ and $\varphi(t)$ are called the instantaneous amplitude and instantaneous phase of $\mathcal{A}[x](t)$, respectively. The construction of Eq. (3) is based on the fact that the values of Fourier transform of a real-valued function at negative frequencies are redundant due to their Hermitian symmetry imposed by the transform. A mapping of the analytic signal in the complex domain \mathbb{C} onto the real space \mathbb{R}^2 results in a two-dimensional pseudo-phase plot. The pseudo-phase plot generates a two-dimensional representation of the n -dimensional manifold, where n could be an arbitrarily large positive integer. The analytic signal of the time-series data can be viewed as a trajectory in the two-dimensional pseudo-phase space. Figure 2 shows an example of ASSP for a typical one-dimensional time-series data set.

The trajectory of an analytic signal in the two-dimensional domain \mathbb{R}^2 is based on its magnitude and phase at each data point. The partition segments are determined from the magnitude and phase of the analytic signal by either maximum entropy partitioning or uniform partitioning or their combination [14]. The choice between maximum entropy partitioning and uniform partitioning is based on the distribution of the analytic signal in the two-dimensional pseudo-phase space and the patterns that need to be detected. This flexibility facilitates adaptation of ASSP to data sets from various distributions. Once the partition segments are determined, a data point in the analytic signal sequence is assigned the symbol, corresponding to the segment. Thus, the symbol sequence is naturally derived from a (complex-valued) sequence of the analytic signal.

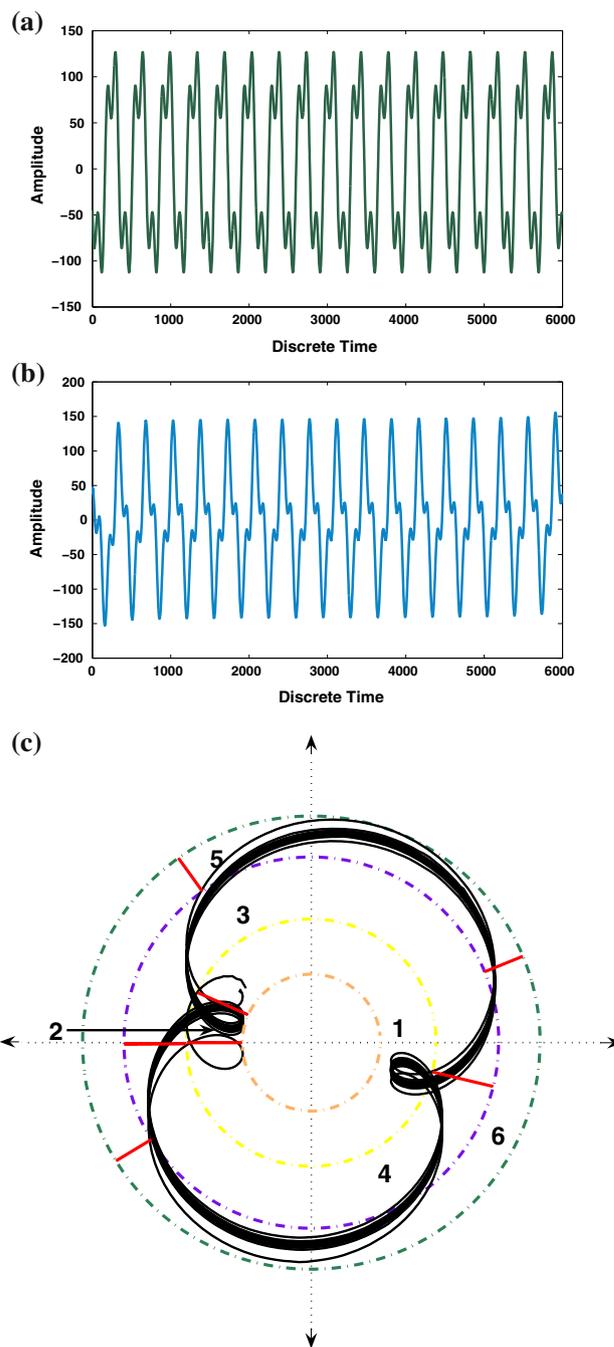


Fig. 2 **a** The time-series data from sensor, **b** the Hilbert transform of the time-series data, **c** the mapping of the complex analytical signal into the \mathbb{R}^2 domain and the derived partition. The number indicates the symbol assigned to each partition

2.2 Construction of probabilistic finite state automata (PFSA)

Construction of a probabilistic finite state automata (PFSA), called D -Markov machines, is originally motivated from an anomaly detection perspective [5]. In the sequel, the words, PFSA and D -Markov machines are used synonymously.

The objective is to identify small changes in the critical parameters of a dynamical system as early as possible before it manifests into a catastrophic disruption (e.g., onset of a chaos in the dynamical system sense, or a phase transition in the thermodynamic sense) in the behavior of the dynamical system. The core assumption in the D -Markov machine construction is that the process is a D th order Markov chain, where $D \in \mathbb{N}$.

Definition 2.1 A stochastic stationary symbol sequence $\dots \sigma_{-2}\sigma_{-1}\sigma_0\sigma_1\sigma_2\dots$ is said to be a D th order Markov chain if the following holds true.

$$\begin{aligned} P(\sigma_i/\sigma_{i-1}\sigma_{i-2}\dots\sigma_{i-D}\sigma_{i-D-1}\dots) \\ = P(\sigma_i/\sigma_{i-1}\sigma_{i-2}\dots\sigma_{i-D}) \end{aligned} \quad (5)$$

The D -Markov machine starts with a state space structure, where the states of the machine represent all sequences of symbols $\sigma_i\sigma_{i+1}\sigma_{i+2}\dots\sigma_{i+D-1}$ that may occur in the symbol sequence. Thus, with cardinality $|\Sigma|$ of the symbol alphabet and depth D , the total maximum number of states in the machine is given by $|\Sigma|^D$. The partitioning methodology also affects the choice of depth D in the finite state machine. A higher D is required if there are signal patterns such that the current symbol is dependent on $D - 1$ preceding symbols in the sequence. While a relatively large D ensures that small features are captured through long-term relational inter-pixel information, the required computational load may increase significantly, thus hampering real-time applications. Analytic signal space partitioning (ASSP) reduces the need for a large depth, because Hilbert transform is a convolution operation that helps ASSP to capture the memory in a time-series data. Each symbol in the derived symbol sequence has some memory of the time-series points around it, thus alleviating the requirement of a large depth. It may suffice to have $D = 1$ at the expense of a relatively large alphabet size. This property of ASSP is very attractive because the number of PFSA states could be exponential in D .

Having determined the structure of the PFSA, the transitional probabilities π_{ij} between the states (i.e., upon occurrence of the symbol $\sigma \in \Sigma$) are determined. As the system trajectory evolves, different states are visited with different frequencies. The number of times a state is visited as well as the number of times a particular symbol is received, is counted. The state probabilities as well as the state–state transition probabilities are calculated for each state in this way. The probabilities are obtained by frequency counts. Thus $\pi_{ij} \approx \frac{n_{ij}}{N_i}$, where n_{ij} is the number of times a transition takes place from state i to state j and N_i is the total number of visits of state i , i.e., $N_i = \sum_j n_{ij}$, where N_i is determined by a stopping rule [12]. The state transition matrix $\Pi \triangleq [\pi_{ij}]$, which is constructed as an irreducible stochastic matrix, represents the pattern observed in the time-series data. Alternatively,

the state probability vector \mathbf{p} that is the left eigenvector corresponding to the unique unity eigenvalue of the irreducible matrix Π could be considered as a pattern instead of Π itself [5, 12].

2.3 Anomaly measure

To detect variations in patterns of the time-series data evolving on the slow scale due to an anomaly, a PFSA is constructed based upon the (reference) nominal condition. The partitioning and PFSA structure derived for the nominal condition serves as a frame of reference for subsequent epochs. Under anomalous conditions, the patterns in the time-series data may change; in order to quantify these changes in the patterns, an anomaly measure ψ is derived from the pattern vector \mathbf{p} . There are a variety of options for the selection of anomaly measure [5]; for example, the anomaly measure could be derived in terms of the Euclidean distance function between two pattern vectors, i.e.,

$$\psi = d(\mathbf{p}, \tilde{\mathbf{p}}) = \sqrt{(\mathbf{p} - \tilde{\mathbf{p}})^T (\mathbf{p} - \tilde{\mathbf{p}})} \quad (6)$$

Since the analysis is performed on a finite-dimensional vector space resulting from a PFSA, norm equivalence is naturally enforced and selection of a specific norm as a candidate for anomaly measure is not critical.

The anomaly measure derived by using the SDF methodology, described above, has the following advantages compared to other pattern identification methods as discussed in [12].

- Sensitivity to small deviations in the signal while remaining robust to measurement noise and spurious disturbances.
- Real-time implementation due to short execution time and small memory requirements on a commercially available microcomputer system.

3 Two-dimensional symbolic dynamic filtering

This section extends the concept of SDF from one-dimensional time-series data to two-dimensional image data and focuses on pattern change detection in images. This requires interpretation and quantification of two-dimensional patterns in the image domain into a real number to develop a measure of the observed change. This section also elaborates partitioning and D -Markov machine generation for the two-dimensional signal domain, which are the two major procedures in two-dimensional SDF.

3.1 Two-dimensional analytic signal space partitioning

The input to the PFSA for image analysis is symbols that encapsulate relevant patterns of images instead of the raw pixel values. Hence, the partitioning methodology becomes much more crucial than that for (one-dimensional) time-series data as the structure and size of the PFSA are described in terms of the alphabet size $|\Sigma|$ and depth D . In this context, Hilbert-transform-based ASSP is an appropriate choice for SDF as it helps reduction of the depth D in the construction of the PFSA.

Several approaches [18–20] have been reported for construction of two-dimensional Hilbert transform. The Hilbert transform product theorem that was first proposed by Stark [21] has been extended for the construction of two-dimensional ASSP in the present context. The two-dimensional Hilbert transform of a real-valued function $x(t_1, t_2)$ with a two-dimensional domain \mathbb{R}^2 , i.e., $t_1, t_2 \in \mathbb{R}$, is defined as

$$\begin{aligned} \tilde{x}(t_1, t_2) &= \mathcal{H}[x](t_1, t_2) \\ &= \frac{1}{\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{x(\tau_1, \tau_2)}{(t_1 - \tau_1)(t_2 - \tau_2)} d\tau_1 d\tau_2 \end{aligned} \tag{7}$$

similar to the two-dimensional Fourier transform of $x(t_1, t_2)$ that is defined as

$$\begin{aligned} \hat{x}(\xi_1, \xi_2) &\triangleq \mathcal{F}[x](\xi_1, \xi_2) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} x(t_1, t_2) e^{-j\xi_1 t_1} e^{-j\xi_2 t_2} dt_1 dt_2 \end{aligned} \tag{8}$$

Two-dimensional Hilbert transform [21] in Eq. (2) yields

$$\begin{aligned} \hat{\tilde{x}}(\xi_1, \xi_2) &\triangleq \mathcal{F}[\tilde{x}](\xi_1, \xi_2) \\ &= (-i \operatorname{sgn}(\xi_1))(-i \operatorname{sgn}(\xi_2))\mathcal{F}[x](\xi_1, \xi_2) \end{aligned} \tag{9}$$

as the frequency response of the two-dimensional Hilbert transform in Eq. (7) implies the following relationship between a signal and its Hilbert transform in the Fourier domain:

$$\hat{\tilde{x}}(\xi_1, \xi_2) = -\operatorname{sgn}(\xi_1) \operatorname{sgn}(\xi_2) \hat{x}(\xi_1, \xi_2) \tag{10}$$

As pixels belong to a class of functions with the discrete two-dimensional domain, the Hilbert transform operator needs to be extended for discrete signals. The discrete Hilbert transform (DHT) of a one-dimensional sequence $\{x(k), k = 0, 1, \dots, N - 1\}$ is defined as [22]

$$\tilde{x}(k) = \text{IDFT}[\mathcal{H}(\ell)\hat{x}(\ell)] \tag{11}$$

where IDFT denotes inverse discrete Fourier transform and $\{\hat{x}(\ell)\}$ is the discrete Fourier transform (DFT) of $\{x(k)\}$ and $\{\mathcal{H}(\ell)\}$ is the periodic discrete representation of $-i \operatorname{sgn}(\xi)$ in Eq. (2).

$$\mathcal{H}(\ell) = \begin{cases} -i, & \ell = 1, 2, \dots, \frac{N}{2} - 1, \\ 0, & \ell = 0, \frac{N}{2}, \\ i, & \ell = \frac{N}{2} + 1, \dots, N - 1. \end{cases} \tag{12}$$

Following the preceding definitions of DHT and two-dimensional continuous Hilbert transform (see Eq. 10), DHT of a two-dimensional discrete sequence $\{x(k_1, k_2)\}, k_1 = 0, 1, \dots, N_1 - 1$ and $k_2 = 0, 1, \dots, N_2 - 1$ is defined as follows:

$$\tilde{x}(k_1, k_2) = \text{IDFT}[\mathcal{H}(\ell_1, \ell_2)\hat{x}(\ell_1, \ell_2)] \tag{13}$$

where

$$\mathcal{H}(\ell_1, \ell_2) = \begin{cases} -1 & \text{for } \ell_1 = 1, \dots, \frac{N_1}{2} - 1, \\ & \ell_2 = 1, \dots, \frac{N_2}{2} - 1, \\ & \text{for } \ell_1 = \frac{N_1}{2} + 1, \dots, N_1 - 1 \text{ and} \\ & \ell_2 = \frac{N_2}{2} + 1, \dots, N_2 - 1. \\ 0 & \text{for } \ell_1 = 0, \frac{N_1}{2}, \\ & \text{for } \ell_2 = 0, \frac{N_2}{2}. \\ 1 & \text{for } \ell_1 = 1, \dots, \frac{N_1}{2} - 1 \text{ and} \\ & \ell_2 = \frac{N_2}{2} + 1, \dots, N_2 - 1, \\ & \text{for } \ell_1 = \frac{N_1}{2} + 1, \dots, N_1 - 1 \text{ and} \\ & \ell_2 = 1, \dots, \frac{N_2}{2} - 1. \end{cases} \tag{14}$$

The Hilbert transform, as defined above, is constructed as a separable transform, i.e., the two-dimensional DHT is a consequence of applying the one-dimensional DHT along one direction, followed by one-dimensional DHT in the other direction. In that case, the two-dimensional complex analytic signal can no longer be defined by simply extending Eq. (3) because the unique properties of the analytic signal with respect to its phase and magnitude, as defined by Gabor [16], no longer bear a physical significance. Nevertheless, such a construction of an analytic image is reported in literature [18] and the resulting two-dimensional DHT is derived as a directional transform [23] (i.e., with order-dependent transformation of rows and columns) to maintain the unique properties of analytic signal.

The above construction of an analytic image has not been adopted for two-dimensional ASSP due to its order dependence. Instead, the signal, as defined by the following relation, is used and is now termed as the two-dimensional pseudo-analytic signal.

$$\mathcal{A}(k_1, k_2) \triangleq x(k_1, k_2) + i \tilde{x}(k_1, k_2) \tag{15}$$

The pseudo-analytic signal is used for ASSP since it is constructed by combining the image data and its Hilbert transform. This construction makes the patterns generated

from the PFSA sensitive to small pattern variations in an image. It is observed by numerical simulation that this construction is reasonably robust to additive noise and spurious disturbances; rigorous mathematical analysis in this direction is a topic of future research. Although the domain of pseudo-analytic signal is the two-dimensional space, its range is the one-dimensional complex space. By taking advantage of this fact, the analytic signal is raster-scanned as a one-dimensional complex signal and the partitioning is made on the values of the pseudo-analytic signal. The symbol matrix is then derived from the one-dimensional complex signal via the partitioning method described in the previous section, where each symbol corresponds to a pixel location; thereby the PFSA structure follows from the standard irreducible stochastic matrix and the resulting state probability vector is treated as the pattern vector as it was done in the analysis of one-dimensional time series.

3.2 Markov representation and anomaly measure

The Markov property in one-dimensional time-series sequences is now extended to the two-dimensional image space [1] as follows. Let a random field \mathbb{F} be constructed as a collection of random variables on a lattice \mathbf{L} , as defined below; and an image is considered as a random field on a lattice set of dimension equal to the size of the image and whose variables are the pixels.

Definition 3.1 Let M_x and M_y be positive integers. Then, a lattice is a collection of discrete points; in the context of images it can be visualized as a rectangular-ordered array of sites as follows.

$$\mathbf{L} = \{(i, j) | 1 \leq i \leq M_x, 1 \leq j \leq M_y\} \tag{16}$$

The one-dimensional Markov property is defined by the conditional dependence of the “future” only on the knowledge of the immediate “past”. This notion of future and past instants may not directly translate into the two-dimensional space. Instead, a random field is said to be Markovian if a point is conditionally dependent on the knowledge of its local neighborhood $N_{i,j}$ and not on the entire field \mathbb{F} . In other words,

$$P(\mathbf{x}_{i,j} / \mathbf{x}_{m,n}; \{(m, n) \in \mathbf{L} \setminus (i, j)\}) \tag{17}$$

$$= P(\mathbf{x}_{i,j} / \mathbf{x}_{m,n}; \{(m, n) \in N_{i,j}\}) \tag{18}$$

The shape of the neighborhood characterizes the random field; various definitions of this neighborhood exist in the literature [24]. In this paper, a neighborhood is defined as a square window of size $(2D + 1) \times (2D + 1)$ at a distance D encapsulating the current pixel, where $D \in \mathbb{N}$. This concept of a neighborhood is illustrated in Fig. 3.

Based on the above definition of the Markov property in random fields, the structure and construction of a D -Markov

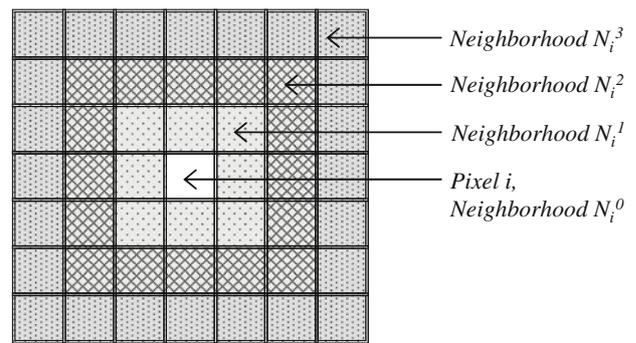


Fig. 3 The neighborhoods at different depths for a variable in the random field are depicted by the different textured regions. The neighborhood N_i^0 signifies just the current pixel, i . N_i^1, N_i^2, N_i^3 correspond to square regions at a distance 1, 2, 3, respectively, from the center pixel

machine are formulated in the two-dimensional domain of images as follows. The Hilbert-transform-based partitioning scheme, when applied to an image, generates a symbol corresponding to each random variable on the lattice. Given the cardinality $|\Sigma|$ of the symbol alphabet and the depth D , the machine starts with a fixed state space structure exactly like the one described in Sect. 2. Each state in the machine represents one of the possible symbol combinations of depth D , $\sigma_i \sigma_{i+1} \sigma_{i+2} \dots \sigma_{i+D-1}$, occurring in the image, where

$$\begin{aligned} \sigma_i &\in N_i^0 \\ \sigma_{i+1} &\in N_i^1 \\ &\vdots \\ \sigma_{i+D-1} &\in N_i^{D-1} \end{aligned} \tag{19}$$

where N_i^k denotes neighborhood k of the lattice point $\sigma_i \in \Sigma$.

The state probabilities and state–state transition probabilities are computed by centering a window of radius D around a symbol on a lattice point. The current state is determined by Eq. (19) and its count is incremented as a transition is encountered. The possible states that a state can make a transition to is determined by the symbols that occur in the N^D -neighborhood of the current symbol and the state–state transition counts are incremented accordingly. This procedure is followed for the states that are visited in the N^{D-1} -neighborhood of the current symbol. The boundary of the image where a symbol does not have a uniform neighborhood, is dealt with by counting the transition probabilities of states that occur in the truncated neighborhoods. Figure 4 shows an example of construction of a machine, based on the D -Markov random field with depth $D = 1$ and symbol size $|\Sigma| = 2$.

Once the PFSA construction is defined, the anomaly measure is obtained in the same way as it was done for one-dimensional signals in Sect. 2.3. The partitioning and PFSA structure are derived for the nominal image to serve as a reference frame for comparison with all subsequent image

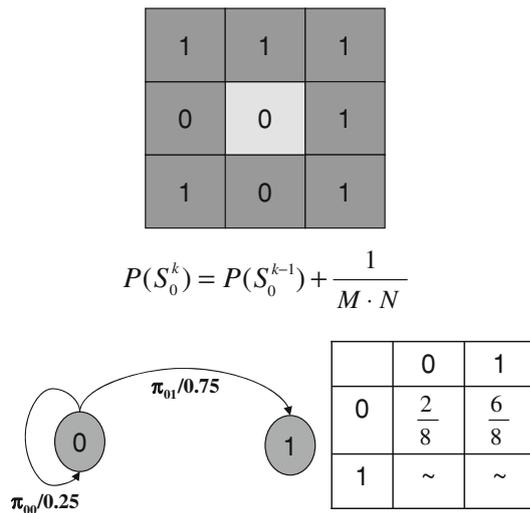


Fig. 4 Example of construction of the Π -matrix and the state probability vectors for a sample Markov random field. Here the symbol in the *center light gray* lattice point is the starting state at N^0 . M and N indicate the size of the whole random field. $M \times N$ is thus the cardinality of the random field

frames. The patterns in the subsequent images may evolve with time, which is reflected in the changed parameters of their corresponding state transition matrices. The distance between the patterns of these representative images serves as a measure of the change observed in the images evolving with time.

4 Experimental validation of the algorithm

The SDF-based image analysis algorithm is validated by analyzing the experimental data from an integrated fatigue test apparatus that is computer-controlled and is equipped for in situ monitoring of fatigue damage in specimens of polycrystalline alloys.

4.1 Description of the test apparatus

As shown in Fig. 5, the test apparatus consists of three main subsystems for fatigue damage propagation, damage surface monitoring and internal flaw detection. The three subsystems for fatigue testing, ultrasonic flaw detection and damage surface monitoring communicate using the TCP/IP protocol for data collection and synchronization.

- *Fatigue testing machine (FTM):* The fatigue testing machine is a MTS 831.10 Elastomer Test System suitable for dynamic testing of specimens both in load and displacement control. It can apply a force of up to ± 25 kN and has a displacement range of ± 50 mm. Dynamic testing can be done at rate ranging from 0.01 to 200 Hz.

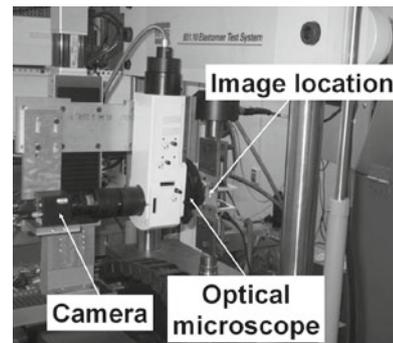


Fig. 5 Fatigue test apparatus

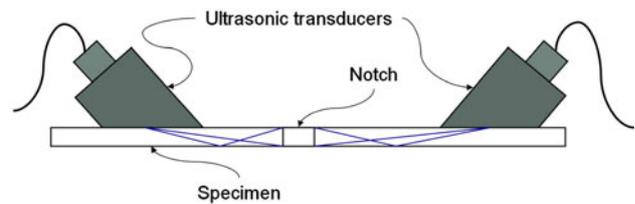


Fig. 6 Ultrasonic transducers sensing in through mode

- *Ultrasonic flaw detection:* The ultrasonic flaw detection system consists of an ultrasonic pulser/receiver board that is used to send ultrasonic waves through the specimen using a piezoelectric transducer as shown in Fig. 6. The test is performed in pulse-echo or in pitch-catch (through) mode. In the pulse-echo mode, the same piezoelectric transducer acts as a sensor collecting reflected ultrasonic from the internal flaws or damage. While in the pulse-echo mode ultrasonic waves send in by one piezoelectric transducer are received by another piezoelectric sensor.
- *Specimen surface monitoring:* The test-bed is equipped with a traveling microscope for in situ monitoring of damage surface of the specimen. The Olympus BX series microscope is mounted on a tri-axial motion stage set-up that gives it the capability to locate and track damage propagation. The microscope is fitted with a camera and a frame-grabber to capture specimen images in real time. Together with the microscope and the camera, the maximum resolution of the damage surface monitoring system is ≈ 2 micron per pixel.

4.2 Experimental results and discussion

To validate the image analysis algorithm, experiments have been conducted on the fatigue testing machine with specimens, made of Aluminum alloy 6061-T6511. Each specimen is 3-mm thick and 50-mm wide and has a slot of 1.58 mm \times 4.5 mm at the center. The central notch is made to increase the stress concentration factor that ensures crack initiation and propagation at the notch ends. The test specimens

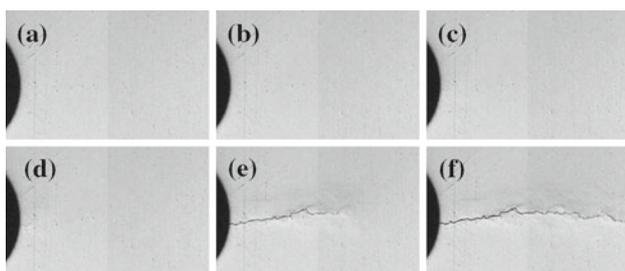


Fig. 7 A subset of the images used in the SDF analysis for detection of crack propagation on the surface of a test specimen under observation; **a** at 1,000 cycles, **b** at 40,000 cycles, **c** at 93,000 cycles when crack appearance on the specimen surface is first detected, **d** at 97,800 cycles when the crack is still very small, **e** at 139,000 cycles when the crack is fully developed, and **f** at 165,000 cycles when the specimen breaks

have been subjected to sinusoidal loading under the tension–tension mode. The stress amplitude was varied from 93.3 to 6.67 MPa at a frequency of 20 Hz. Ultrasonic sensors of 3.5 MHz peak frequency are used in the through mode to monitor internal material damage around the notch. The ultrasonic data, sampled at 50 MHz Images of the intended damage site, have been taken at regular intervals using the surface damage monitoring system. Figure 7 shows some of the images obtained from the optical microscope from a specimen under fatigue load. These images show the specimen at different cycles during the execution of the experiment. To be consistent, the images are taken under similar load on the specimen. However, it is noted that, even though all the images are all taken at a fixed load on the specimen, there is relative motion between the microscope and the specimen during the experiment. Hence, successive images have a certain amount of pixel displacements.

The focus of the experiments is to extract and represent the intuitive spatial relations between the pixels of an image through SDF analysis. In the context of image analysis via SDF, the aim is to accurately detect the transition from initiation to propagation of the crack on the surface due to fatigue damage in the material. It is noted that the inter-pixel relations do not change within successive images under normal conditions, irrespective of the relative motion between the images. However, as a flaw occurs, the inter-pixel relations are altered leading to a change in the derived statistical information. For an image where a surface crack in the specimen is not visible by the optical microscope, the detected pattern may not appreciably vary. As the flaw deteriorates further, the crack propagation is correlated with a sharp change in the representation via SDF analysis, while remaining immune to the relative motion between successive samples.

4.3 Image data analysis

A statistics-based image analysis technique is explored in this paper. To validate the image analysis algorithm, the derived

results of image analysis are compared with experimental data obtained via ultrasonic sensors. The fatigue tests have been conducted on the same Al6061-T6511 aluminum specimens for image analysis and ultrasonic analysis. The specimen has a notch in the center to ensure that a region of high stress exists. Due to this stress concentration, the crack occurs at the notch ends. The experiment was performed at 20 Hz frequency under sinusoidal loading. The images were taken every 2,000 cycles at the mean load.

A sample of images from a typical experiment is shown in Fig. 7. Images are taken after 3,000 cycles so that all transient harmonics have died out and the experimental apparatus is at a steady-state operating condition. Thereon, images are taken every 2,000 cycles. The first image after 3,000 cycles serves as the nominal condition. As is observed in Fig. 7c, the surface crack appears around the central notch end at the region of maximal stress. This a priori knowledge of the physical process enhances computational efficiency of the SDF algorithm by focusing the analysis on the region of interest around the notch edge, which is marked as a rectangular window of height 200 pixels around the center of the notch and width equal to the image. In addition to the computational efficiency, the region of interest is selected to mitigate the effects of inter-frame motion by focusing on the region where the patterns are most important.

An image is captured at 3,000 cycles when the transients die down; this is representative of the nominal condition of the experiment. This image’s pseudo-analytic signal is computed and the distribution of the complex-valued variables is mapped onto the \mathbb{R}^2 domain. This distribution is used for partitioning which results in a rectangular array whose elements are symbols corresponding to the pixels at the equivalent locations of the original image. For the images in the current experiment, the amplitude of the signal is partitioned with four segments using uniform partitioning, and the phase is partitioned using uniform partitioning into two segments per amplitude partition. Uniform partitioning is chosen to mitigate the effects of the inter-frame motion in the experiment. Thus, the partitioning process converts an image whose pixel values range from 0–255 (0 depicting black, 255 representing white) into a symbol matrix whose cardinality is eight. Since the two-dimensional discrete Hilbert transform (DHT) is used for partitioning, each symbol not only contains the pixel information, but also encapsulates long-term pixel dependencies of that pixel with its neighbors in the horizontal and vertical directions due to the convolution operation associated with the Hilbert transform. The partition derived at the nominal condition is preserved and used as a template for all subsequent images. The PFSA are constructed for each image as described in Sect. 3.2. The Euclidean-norm distance between the pattern vectors of the nominal image and the current image (see Eq. 6) is used to quantify the change in the image of the specimen surface.

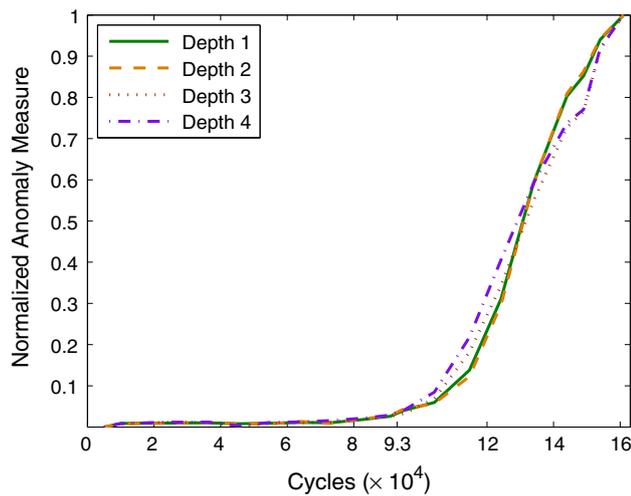


Fig. 8 The anomaly measure derived using SDF on the images from Fig. 7 for different depths of the D -Markov machine. Deformation was first observed at 93,000 cycles

The test specimen is persistently excited under sinusoidal stress except when the experiment is paused periodically to observe the specimen surface. As the test is paused at the mean load, it is not possible to ensure that the specimen settles at the same physical co-ordinates as the loading is resumed. Due to this changing frame of reference between the camera and the specimen, there is a relative motion between the images in the experiment. This is the reason for a trivial but, non-zero anomaly measure value before crack appearance on the surface of the specimen, even though there is no visible surface deformation. The SDF analysis is performed on the images to observe the effects of increasing the depth of the D -Markov machine for anomaly detection, as seen in Fig. 8. Increasing the depth D allows the inclusion of pixel interactions with its neighbors into the pattern recognition problem. These interactions are incorporated to a certain degree into the symbols by the Hilbert transform.

In the image sequence, the effects of relative motion between the camera and specimen are mitigated to a large extent in the partitioning phase by choosing uniform partitioning with a relatively small alphabet size. An increase in the depth D does not affect the pre-crack part of the curve. But, statistical information about the neighboring symbols improves the slope of the anomaly profile once the crack has appeared due to increased information about the length of the crack. This effect is observed in the anomaly measure curve of Fig. 8, where the part of the curve before the appearance of a crack does not change much but the detection profile improves marginally with an increase in the depth D . Since the improvements in the results are insignificant with an increased depth, a depth of $D = 1$ is deemed to be sufficient in the current application of SDF-based image analysis. The Hilbert transform enhances sensitivity to a change in the patterns due to its sensitivity to high frequencies (edges) in

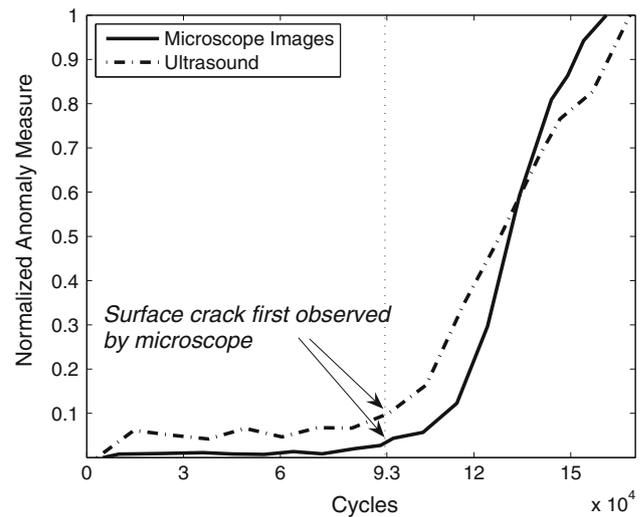


Fig. 9 The normalized anomaly measure curve for Ultrasound sensor data and Microscope Images. Though the crack appears at 93,000 cycles, the ultrasonic sensor shows a rise in the anomaly measure significantly earlier due to its ability to detect internal deformities on the specimen. Once the crack appears, the image analysis is significantly more sensitive to the evolution of the pattern as seen by the steeper rise in its anomaly measure

the image. This is observed by the relatively low value of the anomaly measure (~ 0) before the appearance of a crack, especially when compared to the value of the anomaly once the crack is propagating across the surface of the specimen. Though, increasing the depth D increases the rate of detection of crack growth, it is not significantly high enough to use a depth greater than 1.

Figure 9 exhibits a pair of anomaly measure profiles that are generated with the image and ultrasonic data for performance comparison and validation while the respective sets of image and analysis were collected simultaneously from the same test specimen during a test. The (one-dimensional) ultrasonic data, obtained from the sensors mounted on the surface of the specimen, was analyzed using the SDF algorithm. A total of seven symbols were used for SDF analysis with the depth parameter $D = 1$; higher values of D did not improve the results. Unlike images, the ultrasonic sensors are able to monitor structural changes within the specimen and not merely on the surface of the specimen. This feature of the ultrasonic detection method can be observed in the gradual accumulation of the anomaly measure before the crack appears on the surface of the specimen. In this part of the experiment, the stress causes material deformities within the specimen; these deformities on the specimen surface are not immediately visible by the optical microscope, but the ultrasonic sensor can detect them. Hence, an accumulation in the anomaly measure can be seen in the crack initiation phase of the ultrasonic anomaly measure, whereas the image-based analysis remains relatively flat for the anomaly

measure before the appearance of the cracks on the surface of the specimen.

Once the crack appears on the surface of the specimen, the SDF algorithm is able to effectively track the growth of the crack with time using images. In fact, it is very sensitive to the growth in the crack as seen by the significantly higher value for the slope of the anomaly measure curve compared with the ultrasonic data based anomaly measure. The optical microscope is unable to observe the internal changes within the specimen and hence may not be effective for damage prediction in the crack initiation phase.

5 Summary and conclusions

This paper presents a novel method of image analysis based on the principle of symbolic dynamic filtering (SDF). In this approach, representative symbols are generated based on the Hilbert-transform-based signal representation of the two-dimensional pixel arrays. The pseudo-analytic signal extracts relevant information from the higher frequencies of an image while remaining sensitive to the smoother regions (low frequencies). The concept of Markov random fields is used to define the D -Markov machine construction [5] for two-dimensional signals. The resulting algorithm incorporates the spatial interpixel relationships intrinsic to images. The Hilbert-transform-based partitioning allows D -Markov machines of depth $D = 1$ to be able to detect the pattern changes by incorporating neighborhood information in the symbols; however, in pattern identification problems where more accuracy might be required, higher values of depth (i.e., $D > 1$) could be used.

Through usage of surface-mounted ultrasonic sensors, the SDF analysis has been used for anomaly detection in polycrystalline alloys under sinusoidal stress cycles. The SDF analysis algorithm, proposed in this paper, has been applied to the above experiment by analyzing the images obtained from a microscopic camera that observes the surface of the test specimen. Although detection of crack initiation is not possible due to the inability of the optical microscope to observe internal deformities of a specimen, the SDF algorithm is able to reliably detect and quantify the propagation of the crack on the specimen surface. The experiment introduces interpixel motion between successive images, but due to the statistical nature of the analysis, the pre-processing step of registration is not required. The anomaly measure remains at a trivially small (but non-zero) value before the appearance of a surface flaw. Once such a flaw appears on the specimen surface, the crack is reliably detected.

The results derived from (two-dimensional) image data were found to be close to those obtained from (one-dimensional) time-series data obtained from an ultrasonic sensor on the same test specimen. It is concluded that the SDF

algorithm is able to detect and track the appearance and growth of a surface crack in a specimen of polycrystalline alloy. The extension of the SDF algorithm to images helps in relating the crack length in the two-dimensional image domain to a quantifiable metric, anomaly measure, which increases in proportion to the length of the crack.

The algorithm, to detect deviations from normally observed images, is potentially useful for automated visual inspection in industrial applications [25, 26]. Specifically, the SDF algorithm has several customizable features (e.g., symbol alphabet size and depth of the D -Markov machine) that allow extraction and representation of inter-pixel dependencies in an image as a finite-state machine, which can be applied to several pattern classification problems in image processing. The image analysis algorithm could also be useful for sensor fusion, where decision making may require inputs from heterogeneous sensors whose outputs are not necessarily of the same dimension and scale.

Further theoretical, computational, experimental research is necessary before the proposed image analysis algorithm is put for industrial use. From this perspective, avenues of future research include:

- Theoretical research in two-dimensional Hilbert transform to establish robustness of symbolic dynamic filtering to noise and spurious disturbances while retaining sensitivity to changes in the signal.
- Comparison of the results obtained through infrared imaging with the ultrasonic sensors data for early detection of any anomalous behavior.
- Usage of the directional analytic image transform where the relevant information is directionally oriented.

References

1. Bovik, A. (ed.): Handbook of Image and Video Processing. Academic Press, New York (2000)
2. Lee, T.: Image representation using 2D Gabor wavelets. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(10), 959–971 (1996)
3. Manjunath, B., Ma, W.: Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 837–842 (1996)
4. Li, J., Gray, R., Olshen, R.: Multiresolution image classification by hierarchical modeling with two dimensional hidden Markov models. *IEEE Trans. Inform. Theory* **46**(5), 1826–1841 (2000)
5. Ray, A.: Symbolic dynamic analysis of complex systems for anomaly detection. *Signal Process.* **84**(7), 1115–1130 (2004)
6. Lind, D., Marcus, M.: An Introduction to Symbolic Dynamics and Coding. Cambridge University Press, Cambridge (1995)
7. Rajagopalan, V., Chakraborty, S., Ray, A.: Estimation of slowly-varying parameters in nonlinear systems via symbolic dynamic filtering. *Signal Process.* **89**(2), 339–348 (2008)
8. Samsi, R., Ray, A., Mayer, J.: Early detection of stator voltage imbalance in three-phase induction motors. *Electr. Power Syst. Res.* **79**(1), 239–245 (2009)

9. Chakraborty, S., Ray, A., Subbu, A., Keller, E.: Analytic signal space partitioning and symbolic dynamic filtering for degradation monitoring of electric motors. *Signal Image Video Process.* (2008, to appear)
10. Mallapragada, G., Chattopadhyay, I., Ray, A.: Automated behavior recognition in mobile robots using symbolic dynamic filtering. *Proceedings of the I Mech E Part I: Journal of Systems & Control Engineering*, vol. 222, no. 6, pp. 409–424 (2008)
11. Gupta, S., Ray, A., Keller, E.: Symbolic time series analysis of ultrasonic signals for fatigue damage monitoring in polycrystalline alloys. *Meas. Sci. Technol.* **17**(7), 1963–1973 (2006)
12. Rao, C., Ray, A., Sarkar, S., Yasar M.: Review and comparative evaluation of symbolic dynamic filtering for detection of anomaly patterns. *Signal Image Video Process.* (2008). doi:[10.1007/s11760-008-0061-8](https://doi.org/10.1007/s11760-008-0061-8)
13. Subbu, A., Ray, A., Space partitioning via hilbert transform for symbolic time series analysis. *Appl. Phys. Lett.* **92**(8), 084107-1–084107-3 (2008)
14. Rajagopalan, V., Ray, A.: Symbolic time series analysis via wavelet-based partitioning. *Signal Process.* **86**(11), 3309–3320 (2006)
15. Buhl, M., Kennel, M.: Statistically relaxing to generating partitions for observed time-series data. *Phys. Rev. E* **71**(4), 046213 (2005)
16. Gabor, D.: Theory of communication. *IEE J. Commun. Eng.* **93**, 429–457 (1946)
17. Cohen, L.: *Time-Frequency Analysis*. Prentice Hall PTR, Englewood Cliffs (1995)
18. Havlicek, J., Havlicek, J., Bovik, A.: The analytic image. *Image Processing, 1997. Proceedings., International Conference on*, October 1997, vol. 2, pp. 446–449
19. Hahn, S.: Multidimensional complex signals with single-orthant spectra. *Proc. IEEE* **80**(8), 1287–1300 (1992)
20. Zhu, Y.M., Peyrin, F., Goutte, R.: The use of a two-dimensional hilbert transform for wigner analysis of 2-dimensional real signals. *Signal Process.* **19**(3), 205–220 (1990)
21. Stark, H.: An extension of the Hilbert transform product theorem. *Proc. IEEE* **59**(9), 1359–1360 (1971)
22. Bose, N.: *Digital Filters: Theory and Applications*. Krieger Publishing Company, New York (1993)
23. Havlicek, J., Harding, D., Bovik, A.: The multicomponent AM-FM image representation. *IEEE Trans. Image Process.* **5**(6), 1094–1100 (1996)
24. Geman, S., Geman, D.: Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(6), 721–741 (1984)
25. Vitria, J., Bressan, M., Radeva, P.: Bayesian classification of cork stoppers using class-conditional independent component analysis. *Syst. Man Cybern. Part C IEEE Trans. Appl. Rev.* **37**(1), 32–38 (2007)
26. Kumar, A.: Computer-vision-based fabric defect detection: a survey. *IEEE Trans. Ind. Electron.* **55**(1), 348–363 (2008)