Signal Processing 89 (2009) 981–988

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Statistical estimation of multiple parameters via symbolic dynamic filtering ☆

Chinmay Rao, Kushal Mukherjee, Soumik Sarkar, Asok Ray *

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

ARTICLE INFO

Article history:
Received 23 July 2008
Received in revised form
21 November 2008
Accepted 27 November 2008
Available online 16 December 2008

Keywords:
Parameter estimation
Symbolic dynamics
Time series analysis
Nonlinear dynamical systems

ABSTRACT

This paper addresses statistical estimation of multiple parameters that may vary simultaneously but slowly relative to the process response in nonlinear dynamical systems. The estimation algorithm is sensor-data-driven and is built upon the concept of symbolic dynamic filtering for real-time execution on limited-memory platforms, such as local nodes in a sensor network. In this approach, the behavior patterns of the dynamical system are compactly generated as quasi-stationary probability vectors associated with the finite-state automata for symbolic dynamic representation. The estimation algorithm is validated on nonlinear electronic circuits that represent externally excited Duffing and unforced van der Pol systems. Confidence intervals are obtained for statistical estimation of two parameters in each of the systems.

1. Introduction

Recent literature has reported various methods for estimation of multiple parameters, such as those based on joint state estimation [1], parity equations [2], generalized likelihood ratio [3], Karhunen–Loéve and Galerkin multiple shooting [4], similarity measures [5], and orthogonal Haar transform [6]. An application of parameter estimation is the detection and mitigation of evolving faults in interconnected dynamical systems [7]. Often evolution of gradual deviations from the nominal behavior in individual components of such systems may lead to cascaded faults because of strong input–output and feedback interconnections between the system components, and may eventually cause catastrophic failures and forced shutdown of the entire system. In such a scenario, the problem of degradation monitoring of the system reduces to simultaneous estimation of several slowly varying critical parameters.

Many conventional methods of parameter estimation are model-based and they are often inadequate for human-engineered complex systems due to unavailability of a reliable model of the process dynamics. To alleviate this problem, data-driven parameter estimation methods have been formulated in the setting of hidden Markov models (*HMM*) [8]. One such method is symbolic dynamic filtering (*SDF*) [9,10] that is based on the concept of symbolic time series analysis (*STSA*) [11]; *SDF* belongs to the class of data-driven statistical pattern recognition and enables compression of information into pattern vectors of low dimension for real-time execution on limited-memory platforms, such as small microprocessors in a sensor network. In a recent publication [12], performance of *SDF* has been shown to be superior to that of several pattern classification techniques such as principal component analysis (*PCA*), artificial neural networks (*ANN*), kernel regression analysis (*KRA*), particle filtering (*PF*) and unscented Kalman filtering (*UKF*), in terms of early

☆ This work has been supported in part by the US Army Research Office (ARO) 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.

* Corresponding author. Tel.: +1 814 865 6377.
E-mail addresses: crr164@psu.edu (C. Rao), kum162@psu.edu (K. Mukherjee), szs200@psu.edu (S. Sarkar), axr2@psu.edu (A. Ray).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2008.11.018

detection of changes, computation speed and memory requirements.

As an extension of the parameter estimation method of Tang et al. [13], which is also based on *STSA*, Piccardi [14] proposed multiple-parameter estimation in chaotic systems; although the symbolic analysis was performed on a probabilistic finite-state model, the parameter vector was estimated by (genetic algorithm) optimization in a deterministic setting. The present paper, which is built upon the concept of *SDF*, proposes an alternative approach to estimation of multiple parameters as described below.

The framework of *SDF* includes preprocessing of time series data by time–frequency analysis (e.g., wavelet transform [15] and Hilbert transform [16,17]). The transformed data set is partitioned using the maximum entropy principle [18] to generate the symbol sequences from the transformed data set without any significant loss of information. Subsequently, statistical patterns of the evolving system dynamics are identified from these symbol sequences through construction of probabilistic finite-state automata (*PFSA*). An additional advantage of transform space-based partitioning is reduction of spurious noise in the data set from which the *PFSA* is constructed; this feature provides additional robustness to *SDF* as discussed in [18]. The state probability vectors that are derived from the respective state transition probability matrices of *PFSA* serve as behavioral patterns of the evolving dynamical system under nominal and off-nominal conditions.

Parameter estimation algorithms, based on *SDF*, have been experimentally validated for real-time execution in different applications, such as degradation monitoring in electronic circuits [12] and fatigue damage monitoring in polycrystalline alloys [19]. While these applications of *SDF* have focused on estimation of only a single parameter, the work reported here addresses statistical estimation of multiple parameters. Specifically, this paper is an extension of the earlier work [20] on single-parameter estimation to estimation of multiple parameters that may vary simultaneously. The resulting algorithms are validated on the same test apparatus as [20] for the following electronic systems:

(1) Externally excited Duffing system [21]:

$$\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2}+\beta\frac{\mathrm{d}x}{\mathrm{d}t}+\alpha_1x(t)+x^3(t)=A\cos(\omega_e t) \tag{1}$$

where the amplitude $A=22.0$, excitation frequency $\omega_e=5.0$, and nominal values of the parameters, to be estimated, are $\alpha_1=1.0$ and $\beta=0.1$.

(2) Unforced van der Pol system [22]:

$$\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2}-\mu(1-x^2(t))\frac{\mathrm{d}x(t)}{\mathrm{d}t}+\omega^2x(t)=0 \tag{2}$$

where nominal values of the parameters, to be estimated, are $\mu=1.0$ and $\omega=1.0$.

While the parameter estimation algorithm is tested on an experimental apparatus, a system model is generally used for the purpose of training. Therefore, model reliability or statistic of the modeling error is crucial for robustness of the algorithm and should be known *a priori*. However, this issue is not within the scope of this paper as both training and testing are carried out on similar experimental devices.

2. Review of SDF and single-parameter estimation

This section succinctly reviews the theory of *SDF* [9] and explains the underlying concept of single-parameter estimation [20] in the *SDF* framework.

Extraction of statistical behavior patterns from time series data is posed as a two-scale problem. The *fast scale* is related to response time of the process dynamics. Over the span of data acquisition, dynamic behavior of the system is assumed to remain invariant, i.e., the process is quasi-stationary at the fast scale. In other words, variations in the statistical behavior of the dynamical system are assumed to be negligible on the fast scale. The *slow scale* is related to the time span over which deviations (e.g., parametric changes) may occur and exhibit non-stationary dynamics. The parameters are estimated based on the information generated by *SDF* of the data collected over the fast scale at a slow scale epoch. This method is also applicable to estimation of slowly varying parameters. The rationale is that, since the parameters vary slowly, they are treated as invariants at a given slow scale epoch; accordingly, the fast-scale statistical behavior of the dynamical system may change at different slow scale epochs (that are simply referred to as epochs in the sequel).

2.1. Forward problem in the symbolic dynamic setting

This subsection summarizes the *forward problem* for detection of deviation patterns in the *SDF* setting:

(1) *Time series data acquisition on the fast scale from sensors and/or analytical measurements* (i.e., *outputs of a physics-based or an empirical model*). Data sets are collected at the parameter values as a set $\{s^0, s^1, \ldots, s^k, \ldots\}$, where s^k denotes the value of the parameter at the epoch k.

(2) *Generation of wavelet transform coefficients with an appropriate choice of the wavelet basis and scales*. The wavelet transform largely alleviates the difficulties of phase-space partitioning and is particularly effective with noisy data from high-dimensional dynamical systems.

(3) *Maximum entropy partitioning of the wavelet space at a reference condition*. Each segment of the partitioning is assigned a particular symbol from the symbol alphabet Σ. This step enables transformation of time series data from the continuous domain to the symbolic domain [23].

(4) *Construction of a probabilistic finite-state automaton* (*PFSA*) *at the reference condition*. The structure of the finite-state machine is fixed for subsequent parameter values until a new reference is selected.

(5) *Computation of the reference pattern vector* $\mathbf{p}(s^0)$ *whose elements represent state occupation probabilities of the PFSA at the reference condition*. Such a pattern vector is recursively computed as an approximation of the