CCIT Report # 426 June 2003

Robust System Identification Using Speech Signals

Israel Cohen

Abstract

An important component of a multichannel hands-free communication system is the identification of the coupling between sensors in response to a desired source signal. In this paper, a robust system identification approach adapted to speech signals is proposed. A weighted least-squares optimization criterion is introduced, which includes the probability that the desired signal is present in the observed signals. An asymptotically unbiased estimate for the system's transfer function is derived, and a corresponding recursive on-line implementation is presented. We show that compared to a competing nonstationarity-based method, a significantly smaller error variance is achieved and generally shorter observation intervals are required. Furthermore, in case of a time-varying system, faster convergence and higher reliability of the system identification are obtained. Evaluation of the proposed system identification approach is performed under various noise conditions, including simulated stationary and nonstationary white Gaussian noise, and car interior noise in real pseudo-stationary and nonstationary environments. The experimental results confirm the advantages of proposed approach.

Index Terms

Array signal processing, system identification, signal detection, acoustic noise measurement, speech enhancement, spectral analysis, adaptive signal processing.

I. Introduction

An important component of a multichannel hands-free communication system is the identification of the coupling between sensors in response to a desired source signal [1], [2], [3]. This coupling, often referred to as the acoustical transfer function (ATF) ratio, represents the relation between the impulse responses of the sensors to the desired source. In reverberant and noisy environments, the coupling identification enables to construct an adaptive blocking channel, for an accurate derivation of a reference noise signal, and an adaptive noise canceller, for eliminating directional or coherent noise sources [4]. Furthermore, it also facilitates multichannel signal detection and postfiltering techniques, which employ the transient power ratio between the beamformer output and the reference signals [5], [6].

The author is with the Department of Electrical Engineering, Technion - Israel Institute of Technology, Technion City, Haifa 32000, Israel (email: icohen@ee.technion.ac.il; tel.: +972-4-8294731; fax: +972-4-8323041).