Appl. Comput. Harmon. Anal. ••• (••••) •••–•••

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling

H. Omer, B. Torrésani *

Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form 17 June 2015
Accepted 6 October 2015
Available online xxxx
Communicated by Charles K. Chui

Keywords:
Stationary signals
Time-frequency
Time-scale
Warping
Modulation

ABSTRACT

A class of random non-stationary signals termed *timbre*×*dynamics* is introduced and studied. These signals are obtained by non-linear transformations of stationary random Gaussian signals, in such a way that the transformation can be approximated by translations in an appropriate representation domain. In such situations, approximate maximum likelihood estimation techniques can be derived, which yield simultaneous estimation of the transformation and the power spectrum of the underlying stationary signal.

This paper focuses on the case of modulation and time warping of stationary signals, and proposes and studies estimation algorithms (based on time-frequency and time-scale representations respectively) for these quantities of interest.

The proposed approach is validated on numerical simulations on synthetic signals, and examples on real life car engine sounds.

1. Introduction

Very often, statistical signal processing approaches and algorithms rest on stationarity assumptions placed on the signals of interest and/or on the noise, stationarity being understood as some statistical form of translation invariance. While in many signal processing problems of interest such stationarity assumptions make perfect sense (at least within sufficiently small segments of the signal of interest), this is not always the case. There are situations where departure from stationarity carries essential information on the studied phenomena, and in which one needs to measure non-stationarity accurately. This is an old problem, that has been considered in some specific situations. Most often, the reference stationary signals are modeled as (deterministic) sine waves or sums of sine waves (as is the case in modulation theory), and the object of interest takes the form of instantaneous frequency or group delay (see e.g. [1] and references therein). Recent developments in this domain involve approaches such as Empirical Mode Decompositions (see [2] for a review) and various instances of reassignment and synchrosqueezing methods [3], up to the shape function

---

* Corresponding author. Fax: +33 4 13 55 14 02.
E-mail address: bruno.torresani@univ-amu.fr (B. Torrésani).

http://dx.doi.org/10.1016/j.acha.2015.10.002
1063-5203/© 2015 Elsevier Inc. All rights reserved.

Please cite this article in press as: H. Omer, B. Torrésani, Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling, Appl. Comput. Harmon. Anal. (2015), http://dx.doi.org/10.1016/j.acha.2015.10.002

analysis which allows one to step away from the sine wave model [4]. All these approaches assume underlying stationarity in a deterministic setting (sine waves, or shape functions) and do not involve stochastic modeling of stationarity. However, there are several cases of interest where one cannot assume that the underlying signal is a sine wave or a sum of sine waves, because it originates from complex systems that involve fluctuations.

A good example of such situations is provided by the classical *shape from texture* problem of image processing (see for example [5,6]), where a regular (stationary) texture is projected on a curved surface, yielding images with non-stationary textures. When the underlying texture is periodic, deterministic approaches can be exploited, but this is no longer true when the underlying texture is random. Other examples, which we shall be interested in in this paper, come from audio signal processing, where non-stationary sounds can often be associated to non-stationary motions. Think for example of an accelerating car engine sound. Our ear is perfectly capable to detect regime changes from the sound, more precisely from the non-stationarity of the sound. That example is particularly interesting as in that case, the non-stationarity may be, in first approximation, modeled as a clock change (or time warping), i.e. a periodic system whose inner clock varies as a function of time. This example is the very motivation for the non-stationary sound models we are studying in this paper, which we term *timbre×dynamics* models: models in which a stationary random signal (whose power spectrum is interpreted as *timbre*) is modified by some nonlinear function, which encodes the dynamics of some underlying systems.

Such models have already been considered by M. Clerc and S. Mallat in the context of the shape from texture problem (see [6], and [7] for a more complete mathematical and statistical analysis of the approach). One main contribution was the observation that the non-linear transformation involved in the non-stationarity can in some situations approximately translate into a transport in some appropriate representation domain (time-scale, time-frequency). In addition, it was shown that characterizing the transport in question leads to estimates for the deformation.

The current paper builds on this approach, exploiting a slightly different point of view, namely explicit modeling of the underlying stationary signal as a (possibly complex) Gaussian random signal. Under such assumptions, the deformed signals are also complex Gaussian random signals, and the transport property in an appropriate representation domain alluded to above can be made quantitative using simple approximation techniques. This yields an approximate Gaussian (or complex Gaussian) model from which maximum likelihood estimation can be formulated.

More precisely, we consider two deformation models: modulations (which are conveniently studied in the short time Fourier transform domain), and time warping (studied in the wavelet domain). In both cases, we provide approximate expressions for the transform of the corresponding deformed stationary process, together with error estimates. We also provide sufficient conditions for the invertibility of the covariance matrix of the so-obtained complex Gaussian models, and propose estimation algorithms. The theoretical results are complemented by numerical results, both on synthesized signals and on real signal (car engine sounds), that show that the proposed approach is much more accurate and robust than simpler approaches based upon local frequency (or scale) averages of Gabor (or wavelet) transforms. In comparison, the approach is also more accurate and robust than the algorithm of Clerc and Mallat, that doesn't fully exploit the stationarity of the underlying signal.

This paper is organized as follows. After this introduction, we briefly account in Section 2 for the mathematical background this work rests upon, and introduce our notations. Section 3 is devoted to the study of stationary random signals deformed by modulations, presented in the finite-dimensional, discrete case and Section 4 develops the case of time warping. Numerical results are provided and discussed in Section 5, before the conclusion. More technical proofs are given in Section 6.

Part of this paper (namely the frequency modulation estimation) is an elaborated version, of a short paper published in a conference proceedings [8] (where results were announced without proofs). The original contributions of the present paper consist in the proofs of the results of [8] (in the finite-dimensional case),

Please cite this article in press as: H. Omer, B. Torrésani, Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling, Appl. Comput. Harmon. Anal. (2015), http://dx.doi.org/10.1016/j.acha.2015.10.002