Signal Processing 125 (2016) 290–303

Contents lists available at ScienceDirect

Signal Processing

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

Extraction of instantaneous frequencies from ridges in time–frequency representations of signals

D. Iatsenko, P.V.E. McClintock, A. Stefanovska *

Department of Physics, Lancaster University, Lancaster LA1 4YB, UK

ARTICLE INFO

Article history:
Received 24 November 2015
Received in revised form
22 January 2016
Accepted 31 January 2016
Available online 10 February 2016

Keywords:
Ridge analysis
Wavelet ridges
Time–frequency representations
Wavelet transform
Windowed Fourier transform
Instantaneous frequency
Synchrosqueezing

ABSTRACT

In signal processing applications, it is often necessary to extract oscillatory components and their properties from time–frequency representations, e.g. the windowed Fourier transform or wavelet transform. The first step in this procedure is to find an appropriate ridge curve: a sequence of amplitude peak positions (ridge points), corresponding to the component of interest and providing a measure of its instantaneous frequency. This is not a trivial issue, and the optimal method for extraction is still not settled or agreed. We discuss and develop procedures that can be used for this task and compare their performance on both simulated and real data. In particular, we propose a method which, in contrast to many other approaches, is highly adaptive so that it does not need any parameter adjustment for the signal to be analyzed. Being based on dynamic path optimization and fixed point iteration, the method is very fast, and its superior accuracy is also demonstrated. In addition, we investigate the advantages and drawbacks that synchrosqueezing offers in relation to curve extraction. The codes used in this work are freely available for download.

1. Introduction

A recurring problem in many areas of science is that of identifying curvilinear structures in noisy data and, in many cases, following them as the system evolves in time. The object of study may be spatial, as in automated screening for diabetic retinopathy [1] and in astronomy [2], or it may be a wavelet transform as in the identification of substances through terahertz tomography measurements [3], or it can be a prehistory probability density whose ridge represents the most probable fluctuational path for a nonlinear system undergoing a large fluctuation [4]. It has recently been shown that time-dependent dynamics can be of particular importance, and effective methods have been devised for characterizing the time-dependent amplitudes and phases [5,6]. In these and enumerable other cases the basic problem is that of putting a best-fit curve through a set of points, typically tracing a sequence of extrema, in a digital object. Difficulties to be overcome, in addition to the noise, may include possible crossings, self-crossings or closure of the extracted curves. The approaches that have been proposed include, for example, a variety of ridge-based methods based on locally defined principal curves [7–11] and a method based on an adaptive short-time Fourier transform [12]. In what follows we will focus on the problem as it arises in the analysis of recorded signals.

Separation of the amplitude and frequency-modulated components (AM/FM components) in a given signal, and estimation of their instantaneous characteristics, is a classical problem of signal analysis. It can be approached

* Corresponding author.
E-mail addresses: dmytro.iatsenko@gmail.com (D. Iatsenko), p.v.e.mcclintock@lancaster.ac.uk (P.V.E. McClintock), aneta@lancaster.ac.uk (A. Stefanovska).

http://dx.doi.org/10.1016/j.sigpro.2016.01.024

by projecting the signal onto the time–frequency plane, on which the changes of its spectral content can be followed in time. Such projections are called time–frequency representations (TFRs), typical examples being the windowed Fourier transform (WFT) and the wavelet transform (WT). If the construction of the TFR is well-matched to the signal's structure, then each AM/FM component will appear as a "curve" in the time–frequency plane, formed by a unique sequence of TFR amplitude peaks – ridge points. Based on the properties of these curves, one can estimate the time-varying characteristics of the corresponding components (such as amplitude, phase and instantaneous frequency), an idea that was first expressed in [13]. In other words, to separate a signal into its AM/FM components one can: (a) trace the ridge curves corresponding to the individual components in the signal's TFR; (b) feed these curves into a chosen reconstruction method in order to recover the components and their characteristics (see [14] for a detailed study of the different reconstruction methods together with evaluations of their performance).

However, the first step of such an approach, namely finding the TFR peak sequences associated with the individual signal components, is not a trivial issue. In real cases there are often many peaks in the TFR amplitude at each time, and their number often varies. In such circumstances it can be unclear which peak corresponds to which component, and which are just noise-induced artifacts.

In the present paper, we concentrate solely on the problem of ridge curve identification, which is of great importance in time–frequency signal processing. Ridge analysis is widely used, e.g. machine fault diagnosis [15], fringe pattern analysis [16], studies of cardiovascular dynamics [17] and system classification [18,19]. Although curve extraction has been addressed explicitly in the past [17,20–23], there seems to be no agreement as to the optimal procedure to be used for this task. Here we discuss and generalize some existing algorithms, present new ones, and compare their performance. We end up with a method that is accurate and of almost universal applicability, so that it works well for a large class of signals and, in most cases, does not require adjustment by the user; this is the main contribution of the work. The effects of synchrosqueezing [23–26] on curve extraction are also studied.

The plan of the work is as follows. After reviewing the background and notation in Section 2, we discuss different schemes for curve extraction in Section 3. In Section 4 we compare the performance of these schemes, while the advantages and drawbacks of synchrosqueezing in relation to curve extraction are studied in Section 5, and the limitations of the proposed methods are discussed in Section 6. We draw conclusions and summarize the work in Section 7. A dynamic programming algorithm for fast optimization of a path functional of particular form over all possible peak sequences is discussed in the Appendix.

2. Background and notation

In what follows, we denote by $\hat{f}(\xi)$ and $f^+(t)$, respectively, the Fourier transform of the function $f(t)$ and its positive frequency part:

$$\hat{f}(\xi)=\int_{-\infty}^{\infty} f(t)e^{-i\xi t}\,dt \quad\Leftrightarrow\quad f(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(\xi)e^{i\xi t}\,d\xi,$$
$$f^+(t)\equiv\frac{1}{2\pi}\int_{0}^{\infty}\hat{f}(\xi)e^{i\xi t}\,d\xi. \tag{2.1}$$

Next, by an AM/FM component (or simply component) we will mean a signal of the form:

$$x(t)=A(t)\cos\phi(t)\ \ (\forall t: A(t)>0,\quad \phi'(t)>0), \tag{2.2}$$

which is additionally required to satisfy $A(t)e^{i\phi(t)}\approx 2[A(t)e^{i\phi(t)}]^+$, so that $A(t)$ and $\phi(t)$ are determined uniquely and, in the case of a single component, can be found using the analytic signal approach; more detailed discussions of issues related to the definition and estimation of the amplitude $A(t)$, phase $\phi(t)$ and instantaneous frequency $\nu(t)\equiv\phi'(t)$ of the component can be found in [14,27–30].

In real cases, a signal usually contains many components $x_i(t)$ of the form (2.2), as well as some noise $\zeta(t)$ (that can be of any form, and is not necessarily white and Gaussian [30]):

$$s(t)=\sum_i x_i(t)+\zeta(t). \tag{2.3}$$

The goal of ridge analysis is to extract these components, either all or only those of interest, from the signal's TFR.

The two main linear TFRs suitable for components extraction and reconstruction are the windowed Fourier transform (WFT) $G_s(\omega,t)$ and the wavelet transform (WT) $W_s(\omega,t)$. Given a signal $s(t)$, they can be constructed as

$$\begin{aligned}
G_s(\omega,t)&\equiv\int_{-\infty}^{\infty}s^+(u)g(u-t)e^{-i\omega(u-t)}\,du\\
&=\frac{1}{2\pi}\int_0^{\infty}e^{i\xi t}\hat{s}(\xi)\hat{g}(\omega-\xi)\,d\xi,\\
W_s(\omega,t)&\equiv\int_{-\infty}^{\infty}s^+(u)\psi^*\left(\frac{\omega(u-t)}{\omega_\psi}\right)\frac{\omega\,du}{\omega_\psi}\\
&=\frac{1}{2\pi}\int_0^{\infty}e^{i\xi t}\hat{s}(\xi)\hat{\psi}^*(\omega_\psi\xi/\omega)\,d\xi,
\end{aligned} \tag{2.4}$$

where $s^+(t)$ is the positive frequency part of the signal (as defined in (2.1)), $g(t)$ and $\psi(t)$ are respectively the window and wavelet functions chosen, and $\omega_\psi\equiv\operatorname{argmax}|\hat{\psi}(\xi)|$ denotes the wavelet peak frequency (for the WFT we assume $\operatorname{argmax}|\hat{g}(\xi)|=0$). Note that the WT is commonly defined through the scales $a=\omega_\psi/\omega$, but that in (2.4) we have already transformed to frequencies.

The main difference between the two TFRs mentioned is that the WFT distinguishes the components on the basis of their frequency differences (linear frequency resolution), while the WT does so on the basis of ratios between their frequencies (logarithmic frequency resolution). In effect, while the time-resolution of the WFT is fixed, for the WT it is linearly proportional to frequency, so that the time-modulation of the higher frequency components is represented better than that for the components at lower frequencies.