Mechanical Systems and Signal Processing 115 (2019) 720–735

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Two robust approaches to multicomponent signal reconstruction from STFT ridges

Xiangxiang Zhu [a], Zhuosheng Zhang [a,*], Jinghuai Gao [b], Wenting Li [a]

[a] *School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China*
[b] *National Engineering Laboratory for Offshore Oil Exploration, Xi'an 710049, China*

ARTICLE INFO

Article history:
Received 30 January 2018
Received in revised form 20 June 2018
Accepted 23 June 2018

Keywords:
Multicomponent signal decomposition
Ridges
Half-quadratic optimization
Self-paced learning
Noise robustness

ABSTRACT

The problem of how to accurately reconstruct the multicomponent signal from the ridges of the short-time Fourier transform (STFT) is considered. Especially when time-frequency representations contain strong noise corruption and crossed components, exact reconstruction becomes more difficult. In this paper, we propose two robust ridge reconstruction approaches, i.e., weighted ridge reconstruction (WRR) and self-paced ridge reconstruction (SPRR). The former is aim to select a more robust loss function to eliminate the influence of the strong noise and outliers. A half-quadratic optimization algorithm is developed to solve the proposed problem efficiently. The latter incorporates self-paced learning (SPL) method into ridge reconstruction model to sequentially include ridge points into signal reconstruction from easy to complex, which not only can suppress noise and outliers, but also can avoid a bad result in the presence of missing observations. Simulation and real-life signals are employed to show the effectiveness and practicability of the proposed approaches.

1. Introduction

Many real-life signals, such as radar signals, seismic signals and mechanical fault signals, can be modeled as superpositions of amplitude- and frequency-modulated (AM-FM) waves [1,2]. Containing several modes they are often called multicomponent signals. How to accurately and efficiently extract the constituent components from the multicomponent signal has been a hot point in the signal processing community. Especially when the components of such signals are corrupted by strong noise and overlap in time-frequency (TF) plane, exact reconstruction becomes more challenging [3,4].

In the past decades, many multicomponent signal decomposition/reconstruction methods have been proposed, and generally can be classified into three classes. The first one directly extracts each signal mode in time domain, among which the empirical mode decomposition (EMD) [5] and its extensions [6–8] are the most well-known. Although these methods have been widely applied in various fields, some challenging issues, like lack of mathematical theory, poor resolution to separate close modes and sensitive to strong noise, till remain to be addressed [2,12]. Another popular method is the sparsification approach [9,10], which tries to find sparse representations of signal modes over a parameterized dictionary such as the Fourier or wavelet-based dictionary. However, if no priori information is available, it is rather difficult to determine the size of the dictionary which balances the tradeoff between the approximation accuracy and the computation load. The second class proposes to extract components in frequency domain, which involves empirical wavelet transform [11], variational

* Corresponding author.
E-mail address: zszhang@mail.xjtu.edu.cn (Z. Zhang).

https://doi.org/10.1016/j.ymssp.2018.06.047
0888-3270/

mode decomposition [12] and so on. This kind of methods assumes that the signal modes have narrow-band properties and have distinct spectrum support, thus they are powerless for the wide-band modes, particularly when overlapping spectrum happens.

The third class is focused on TF analysis methods, several of which include short-time Fourier transform (STFT), wavelet transform (WT), synchrosqueezing transform [13] and synchroextracting transform [14], etc. To recover each component in TF domain, the most obvious way is by inverse STFT or inverse WT [3,15]. However, such methods are easily disturbed by the noise, especially the interference between crossed components, because it is difficult to define a rational area of each mode to achieve a tradeoff between energy lost and noise introducing. Another way is the parametric methods, which use a pre-defined model, such as polynomial [4,16], piecewise polynomial [17], and sinusoidal/Fourier models [18,19], to characterize the instantaneous frequencies (IFs) of the AM-FM modes. After estimating the model parameters, various post-processing methods, such as singular value decomposition method [4], joint-optimization technique [20] and non-parametric Gaussian latent feature model [21] are employed for signal recovery. These methods usually involve multidimensional searches in the parameter space, which will be relatively time consuming. And the predefined signal model may only adapt to specific situations. Rather than using the parametric methods, a popular way which has attracted considerable interest is to use the ridges, being the curve at the TF plane along which the signal energy is locally maximum, for recovery, which involves penalization approach [22,23], skeleton method [24,25] and wavelet ridge signal decomposition (WRSD) method [26]. For the penalization approach, it is by minimizing a suitably chosen quadratic function and using the values of the skeleton (i.e., the TF representations evaluated on the ridge) as linear constraints to reconstruct. This method requires a lot of time to build the penalty functions and to solve the optimization problem. For the skeleton and WRSD, they are based on the fact that the ridges give a reasonably good estimate for the IFs of the components. However, this estimation is biased and inaccurate when estimating the signals which contain stronger modulations [27], let alone the overlapped non-stationary signals. Besides, S. Mallat [28] considered the TF atoms on the ridge as a frame and used the available frame algorithm to reconstruct. Despite this method achieves good performance in term of accuracy and efficiency, it is sensitive to strong interference, especially in the presence of outliers and missing observations (see Section 4 for some numerical experiments).

To better eliminate the influence of strong interference in TF plane and obtain a more robust reconstruction, in this paper, we give two robust ridge reconstruction approaches, i.e., weighted ridge reconstruction (WRR), self-paced ridge reconstruction (SPRR). For WRR, more robust loss functions are employed and adaptive weights are assigned to different ridge points via half-quadratic minimization [29], which is much more insensitive to heavy noise and outliers. For SPRR, self-paced learning (SPL) method [30] is incorporated into the ridge reconstruction model to sequentially include ridge points into signal reconstruction from easy to complex, which not only can effectively eliminate the influence of the strong noise, but also helpful to avoid a bad result in the presence of missing observations.

The remainder of the paper is organized as follows. In Section 2, we give a brief description of the STFT ridges of the multicomponent signal, half-quadratic optimization and self-paced learning. In Section 3, the ridge reconstruction problem and two robust reconstruction approaches are introduced. The simulation test is carried out in Section 4. In Section 5, a sound signal and a vibration signal are utilized to show the effectiveness of the proposed methods. Finally, the conclusions are given in Section 6.

2. Notations and background

2.1. STFT ridges

Consider a multicomponent AM-FM signal as

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} a_k(t) e^{j\phi_k(t)}, \tag{1}$$

where K is a positive integer representing the number of AM-FM components, $\sqrt{j} = -1, a_k(t) > 0$ is the instantaneous amplitude of the k-th component (or mode), and $\phi_k(t)$ is the instantaneous phase of the k-th component satisfying $\phi'_k(t) > 0$. Generally, $a_k(t)$ is considered as a continuously differentiable function and $\phi_k(t)$ a two times continuously differentiable function.

Let $g(t)$ be a real even window function and it satisfies $\|g\| = 1$. Define $g_{\mu,\xi}(t)$ as

$$g_{\mu,\xi}(t) = g(t-\mu)\exp(j\xi t). \tag{2}$$

The STFT of signal $f(t)$ can be represented as

$$S_f(\mu,\xi) = \langle f(t), g_{\mu,\xi}(t)\rangle = \int_{-\infty}^{+\infty} f(t) g(t-\mu) e^{-j\xi t} dt. \tag{3}$$

The general form for the STFT of a multicomponent signal can be written in the following [23,28]:

$$S_f(\mu,\xi) = \sum_{k=1}^{K} a_k(\mu) e^{j(\phi_k(\mu)-\mu\xi)} \hat{g}(\xi - \phi'_k(\mu)) + r(\mu,\xi), \tag{4}$$