Journal of Statistical Planning and Inference 166 (2015) 126–137

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Sparse Bayesian representation in time–frequency domain

Gwangsu Kim [a], Jeongran Lee [b,*], Yongdai Kim [c], Hee-Seok Oh [c]

[a] Department of Statistics, Korea University, Republic of Korea
[b] Bell Labs, Alcatel-Lucent, Republic of Korea
[c] Department of Statistics, Seoul National University, Republic of Korea

ARTICLE INFO

Article history:
Received 25 November 2013
Received in revised form 25 January 2015
Accepted 13 February 2015
Available online 5 March 2015

MSC:
62G20
94A12

Keywords:
Bayesian inference
Beta-Bernoulli prior
Gabor frames
Overcomplete dictionaries
Regularization
Sparsity
Time–frequency analysis

ABSTRACT

We consider a Bayesian time–frequency surfaces modeling of sound signals. The model is based on decomposing a signal into time–frequency domain using Gabor frames, which requires a careful regularization through appropriate variable selection to cope with the overcompleteness. We propose to impose a time-line beta-Bernoulli prior on the time–frequency coefficients of Gabor frames to create dependency structures coupled with the stochastic search variable selection to achieve sparsity. Theoretical aspects of the prior specification are investigated and an efficient MCMC algorithm is developed. Performance of the proposed model with other popularly used models is compared through analyzing simulated and real signals.

1. Introduction

This paper is concerned with a Bayesian variable selection (BVS) for time–frequency surface estimation. A Markov chain Monte Carlo (MCMC) method for BVS is proposed by George and McCulloch (1993), Ghosal (1999), and Jiang (2007) studied asymptotic properties of the posterior distribution of when the number of covariates diverges, and BVS approaches have been applied to various problems (Li and Zhang, 2010; Richardson et al., 2010).

A BVS method for time–frequency surface estimation of nonstationary signals, in which the inverse modeling of chosen energy coefficients is important, has been proposed by Wolfe et al. (2004). Though the hierarchical mixture prior used in Wolfe et al. (2004) is flexible enough to yield an efficient sampling algorithm, it would be worthy of investigation to adopt a more structured prior if prior information about parameter structures is available. Moreover, in case of high-dimensional data, automatic adaptation of the sparsity is also an important consideration. In this paper, we propose a new BVS method to explore these aspects in a time–frequency surface modeling of sound signals. In particular, the proposed prior incorporates well the time-line dependency of signals.

Time–frequency decompositions of signals are closely related to our perception of sounds because when listening to music we perceive sounds that have frequencies varying in time. Gabor (1946) proposed a method of decomposing signals over dictionaries of elementary waveforms termed 'time–frequency atoms' in the time–frequency surface. Since

* Corresponding author.
E-mail addresses: s88012@gmail.com (G. Kim), goodran@gmail.com (J. Lee), ydkim@stats.snu.ac.kr (Y. Kim), heeseok@stats.snu.ac.kr (H.-S. Oh).

http://dx.doi.org/10.1016/j.jspi.2015.02.008
0378-3758/

then numerous modelings and analyses of sound signals have been developed, which results in a production of various dictionaries. The key issue in constructing dictionaries with time–frequency atoms is to make them adapted to structural properties of signals. For example, orthonormal bases of time–frequency atoms do not have redundancy and so they are not appropriate for time–frequency analysis since Balian–Low theorem implies that redundancy is necessary for good time–frequency localization (Grochenig, 2001). With overcomplete dictionaries, the decomposition is not unique because some elements in the dictionary can be represented by other elements. Thus, we need a way to choose a proper one among many representations. In this paper, we use Gabor frames as an overcomplete dictionary and sparse representation of a signal in this dictionary—the one with fewer significant coefficients, as a criterion to identify the model.

To find sparse representations under overcomplete systems, pursuit algorithms (Chen et al., 1998; Mallat and Zhang, 1993) search for sparse dictionary approximations. Especially, the basis pursuit (BP) by Chen et al. (1998), which is related to the LASSO (Tibshirani, 1996) in regression problem, focuses on the sparsity in redundant packet dictionaries by a principle of decomposing a signal into an optimal superposition of dictionary elements having the smallest ℓ_1-norm of coefficients among all such decompositions. However, the coefficients in the time–frequency surface obtained by the ℓ_1-penalty with an overcomplete dictionary can lose local constancy of the coefficients profile in features. In the meanwhile, matching pursuit (MP) algorithm of Mallat and Zhang (1993) uses greedy approaches to the function approximation that are guaranteed to provide a maximally sparse representation. The maximally sparse representation, however, loses inherited dependency structures in time-varying frequencies. In Bayesian framework, Wolfe et al. (2004) employed the idea of the spike-and-slab prior (Mitchell and Beauchamp, 1988) to control the sparsity of the model. Three variations of the spike-and-slab prior are proposed to achieve the persistence of energies in the time–frequency surface, which are related to conditioning the local neighborhood of a time–frequency surface point corresponding to the atom. These priors, however, do not control the sparsity and redundancy properly resulting in many spurious isolated components.

To overcome shortcomings in the aforementioned methods and to improve time–frequency representation and its inverse reconstruction, we propose a novel prior to model the dependency and sparsity of atoms across the time–frequency surface. To favor the persistence of meaningful signal traits and trends in the time–frequency surface, it is necessary to model the correlated structure of a signal in that domain since the frequency at one time point is affected by the other frequencies near by. For this purpose, we use a prior which consists of hierarchical parameter structures coupled with a beta-Bernoulli prior, which is a modification of the Indian buffet process (IBP; Griffiths and Ghahramani, 2005). The main contribution of this paper is to find a desirable specification of the prior parameters by investigating theoretical properties of the posterior which detects well structured sparsities existing in signals.

The remaining part of this article is organized as follows. In Section 2, we specify the model and prior. In Section 3, we take into account theoretical properties of the posterior to choose a proper specification of the prior parameters and develop an efficient MCMC algorithm. Results of simulation studies as well as real data analysis are presented in Section 4, and concluding remarks follow in Section 5.

2. Material and methods

2.1. An overcomplete dictionary and the frame theory

A time–frequency dictionary $\mathcal{D} = \{\psi_\gamma\}_{\gamma\in\Gamma}$ with a collection of parameters Γ is composed of atoms of unit norm $\|\psi_\gamma\| = 1$, which are localized in time and frequency. Constructing a dictionary of localized time–frequency atoms can be thought of as covering the time–frequency surface by resolution cells having various time- and frequency-widths under Heisenberg's uncertainty principle (Heisenberg, 1927).

The frame theory is used to construct redundant dictionaries that define complete, stable, and localized signal representations. Suppose that a dictionary $\{\psi_\gamma\}_{\gamma\in\Gamma}$ has been selected to approximate a signal f. The dictionary $\{\psi_\gamma\}_{\gamma\in\Gamma}$ which consists of elements in a Hilbert space $\boldsymbol{V}$ is said to be a *frame* if there exists $B \geq A > 0$ such that, $\forall h \in \boldsymbol{V}$,

$$A\|h\|^2 \leq \sum_{\gamma\in\Gamma} |\langle h, \psi_\gamma\rangle|^2 \leq B\|h\|^2.$$

The frame theory gives energy equivalence conditions that results in stable operators since any perturbation of frame coefficients affects h with a similar magnitude. Frames are more general than orthonormal bases and yield more flexibility.

2.2. Gabor frames

Gabor frames formalize the notion of the windowed Fourier transform and the concept of valid tilings of the time–frequency surface (see, *e.g.*, Wolfe et al. (2004) and Grochenig (2001)). For a L-periodic sequence $\{f(t)\}_{t=0}^{L-1} \in \ell_2(\mathbb{Z})$, Gabor expansion is defined as

$$f(t) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} c_{m,n} g_{m,n}(t), \quad (t = 0, 1, \ldots, L-1),$$