Signal Processing 108 (2015) 259–271

Contents lists available at ScienceDirect

Signal Processing

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

Model based Bayesian compressive sensing via Local Beta Process

Lei Yu [a,*], Hong Sun [a], Gang Zheng [b], Jean Pierre Barbot [c]

[a] School of Electronic and Information, Wuhan University, 430072 Wuhan, China
[b] Non-A team, INRIA-Lille Nord Europe, 59650 Villeneuve d'Ascq Cedex, France
[c] ECS-Lab, ENSEA, 95014 Cergy-Pontoise Cedex, France

ARTICLE INFO

Article history:
Received 24 April 2014
Received in revised form
17 September 2014
Accepted 17 September 2014
Available online 30 September 2014

Keywords:
Compressive Sensing
Structured sparsity
Beta process
Variational Bayes

ABSTRACT

In the framework of Compressive Sensing (CS), the inherent structures underlying sparsity patterns can be exploited to promote the reconstruction accuracy and robustness. And this consideration results in a new extension for CS, called *model based CS*. In this paper, we propose a general statistical framework for model based CS, where both sparsity and structure priors are considered simultaneously. By exploiting the Latent Variable Analysis (LVA), a sparse signal is split into weight variables representing values of elements and latent variables indicating labels of elements. Then the Gamma-Gaussian model is exploited to describe weight variables to induce sparsity, while the beta process is assumed on each of the local clusters to describe inherent structures. Since the complete model is an extension of Bayesian CS and the process is for local properties, it is called Model based Bayesian CS via Local Beta Process (MBCS-LBP). Moreover, the beta process is a Bayesian conjugate prior to the Bernoulli Process, as well as the Gamma to Gaussian distribution, thus it allows for an analytical posterior inference through a variational Bayes inference algorithm and hence leads to a deterministic VB-EM iterative algorithm.

1. Introduction

Compressive Sensing (CS) provides a new sampling paradigm that allows signals to be sampled with sub-Nyquist rate (2 times of signals' Fourier bandwidth) [1]. In the framework of CS, samples are collected via a linear projection satisfying restricted isometry property, and provided that signals are sparse (itself or under transform), the exact recovery can be guaranteed theoretically [2]. The reconstruction of signals is often casted as the following regularized optimization problem:

$$\boldsymbol{\theta}^{\star} = \arg\min_{\theta} \tfrac{1}{2}\|\boldsymbol{y}-\Phi\boldsymbol{\theta}\|_2^2 + \lambda \cdot \psi(\boldsymbol{\theta}) \tag{1}$$

with $\boldsymbol{\theta} \in \mathbb{R}^n$ being the sparse signals, $\Phi \in \mathbb{R}^{m\times n}$ the sensing matrix and $\boldsymbol{y} \in \mathbb{R}^m$ the collected samples. $\psi: \mathbb{R}^n \to \mathbb{R}_+$ is a regularization term that induces the sparsity of solutions, e.g. the ℓ_p norm with $p \in [0,1]$. $\lambda > 0$ is a constant that balances distortion and sparsity.

Exploiting sparse regularization term, lots of algorithms have been proposed recently such as Linear Programming (Basis Pursuit) methods [3] and iterative soft thresholding algorithm [4] with $p=1$, greedy methods [5,6] and iterative hard thresholding algorithm [7] with $p=0$, iterative re-weighted least squares regression [8] and Bayesian methods [9,10] that can be considered to solve the regularized optimization problem with $p \in (0,1]$.

* Corresponding author.
E-mail addresses: ly.wd@whu.edu.cn (L. Yu), hongsun@whu.edu.cn (H. Sun), gang.zheng@inria.fr (G. Zheng), barbot@ensea.fr (J. Pierre Barbot).

http://dx.doi.org/10.1016/j.sigpro.2014.09.018
0165-1684/

On the other hand, besides the sparse prior, structural constraint on the support of the sparse signals has been found in many practical applications. The most common example is the wavelet coefficients that can be considered as a tree-structured sparse vector [11,12] due to the relationship between two conjuncted wavelet scales. Another application is ECG telemonitoring where Zhang et al. [13] exploited the block structure on the ECG signal to improve the compression performance. Similar block/cluster structure has also been exploited in ISAR imaging where the support constraints are due to the continuity of the target scene [14–16] and image denoising where the similar patches share the same sparsity patterns [17].

Theoretically, it can be proven that the introduction of structural priori largely improves the reconstruction performance [18,19]. Considering structured sparse signal model leads to an extension to CS, i.e. *Model-based Compressive Sensing* (MCS) [19]. Accordingly, many new algorithms are proposed by imposing structural sparse regularization on (1), for instance the mixed $\ell_{1,2}$ norm that has been widely employed to cope with problems with block structured sparsity [20] or multiple measurement vector model [17]. Other type of approaches has been exploited to cope with different structures. Group-structured extension to OMP [21], Block-OMP [22], Tree-based OMP [19], etc. are OMP-based approaches that proposed recently. However, the above methods are parametric approaches that require to set structure-related parameters at first. Another type of approaches is based on Bayesian CS framework [9], including Tree-based Bayesian CS [11], Cluss-MCMC/-VB [23,24], BSBL [25] and so on. The Bayesian approaches are non-parametric, but only specific to one type of structures. More recently, a general model for structured sparse signals is proposed via the Boltzmann machine [26,27], while the interaction matrix for the structural model should be set or pre-estimated.

In this paper, we propose a general statistical framework for model based CS, where both sparsity and structure priors are considered simultaneously. By exploiting the Latent Variable Analysis (LVA), a sparse signal is split into weight variables representing values of elements and latent variables indicating labels of elements. Then we assume that weight variables obey a Gamma-Gaussian model to induce the sparsity. On the other hand, according to the inherent structures, the latent variables can be described by a graph with local clusters, thus a beta-Bernoulli process is assumed on each of the local clusters to describe the properties of structures. Since the complete model is an extension of Bayesian CS and the process is for local properties, it is called Model based Bayesian CS via Local Beta Process (MBCS-LBP). Moreover, the beta process is a Bayesian conjugate prior to the Bernoulli Process, as well as the Gamma to Gaussian distribution, thus it allows for an analytical posterior inference through a variational Bayes inference algorithm and hence leads to a deterministic VB-EM iterative algorithm.

The rest of paper is organized as follows. In Section 2, we briefly review the framework of Bayesian CS. Then the proposed structured sparsity model is presented in Section 3. After that, the posterior inference through the variational Bayes approach is given in Section 4. Experiments are carried out in Section 5 to verify superior performance of the proposed algorithm. The paper ends up with a conclusion.

2. Bayesian compressive sensing

The canonical form of CS could be written as follows:

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{2}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ is the sensing matrix satisfying the so-called RIP [2], $\boldsymbol{\theta} \in \mathbb{R}^n$ is the original sparse signal, $\boldsymbol{y} \in \mathbb{R}^m$ is the compressed measurement and ϵ is the possible noise or perturbations. Note that $m \ll n$ to ensure the sufficient compression, thus the reconstruction for $\boldsymbol{\theta}$ is degenerated into an underdetermined linear inverse problem. Assuming a white noise for $\boldsymbol{\epsilon}$ with variance $\sigma_0 = \alpha_0^{-1}$, one can easily obtain a Gaussian likelihood model on the measurements $\boldsymbol{y}$, written as

$$p(\boldsymbol{y}|\boldsymbol{\theta}) \propto \exp\left(-\frac{\alpha_0}{2}\|\boldsymbol{y}-\boldsymbol{\Phi}\boldsymbol{\theta}\|^2\right) \tag{3}$$

On the other hand, a sparsity-inducing prior is imposed via a generalized Gaussian distribution on original signals [3,28,8], such as

$$p(\boldsymbol{\theta}) \propto \exp\left(-\sum_{i=1}^{n}|\theta_i|^p\right) \tag{4}$$

where $p \in [0,1]$. Thus the Maximum a Posteriori (MAP) solutions to (2) estimator for $\hat{\boldsymbol{\theta}}$ could be formulated as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \lambda\|\boldsymbol{y}-\boldsymbol{\Phi}\boldsymbol{\theta}\|^2 + \sum_{i=1}^{n}|\theta_i|^p \tag{5}$$

where $\lambda = \alpha_0$ is a trade-off parameter balancing sparsity with quality of fit.

Solving the MAP solution is possible to find the true sparse reconstruction for CS. Nevertheless, the trade-off parameter λ in (5) reminds to be estimated via other algorithms or fixed in advance according to prior knowledge. A nonparametric method to solve the sparse linear inverse problem will be applausive.

Considering the process of CS measurement as a hierarchical Bayesian model as shown in Fig. 1, it provides a new extension to CS [29,9,10], called *Bayesian CS* and leads to a nonparametric sparse linear inverse solver. Instead of imposing a generalized Gaussian prior for sparse signals, Bayesian CS injects the sparse constraint through a conditional Gaussian prior (6) with its inverse variance (precision) guided by a hyperprior of Gamma distribution, or called a *Gamma-Gaussian* model.

$$p(\boldsymbol{\theta}|\boldsymbol{\alpha}) \propto \prod_{i=1}^{n} \exp\left(-\frac{\alpha_i}{2}\theta_i^2\right) \tag{6}$$

where α_i obeys a Gamma distribution $p(\alpha_i|a,b) \propto \alpha_i^{a-1}e^{-b\alpha_i}$ with $a,b>0$.

Integrating out the hyperparameters $\boldsymbol{\alpha} \triangleq \{\alpha_i\}_{1:n}$, the implicit prior for sparse signals $\boldsymbol{\theta}$ can be formulated via

$$\begin{aligned} p\left(\boldsymbol{\theta}\middle|a,b\right) &= \int p(\boldsymbol{\theta}|\boldsymbol{\alpha})p(\boldsymbol{\alpha}|a,b)\, d\boldsymbol{\alpha} \\ &\propto \prod_{i=1}^{n}\left(b+\frac{\theta_i^2}{2}\right)^{-(a+1/2)} \end{aligned} \tag{7}$$