

BAYESIAN RECONSTRUCTION OF HYPERSPECTRAL IMAGES BY USING COMPRESSED SENSING MEASUREMENTS AND A LOCAL STRUCTURED PRIOR

Yuri Mejía, Henry Arguello

Facundo Costa, Jean-Yves Tourneret, Hadj Batatia

Universidad Industrial de Santander

University of Toulouse, IRIT/INP-ENSEEIH

ABSTRACT

This paper introduces a hierarchical Bayesian model for the reconstruction of hyperspectral images using compressed sensing measurements. This model exploits known properties of natural images, promoting the recovered image to be sparse on a selected basis and smooth in the image domain. The posterior distribution of this model is too complex to derive closed form expressions for the estimators of its parameters. Therefore, an MCMC method is investigated to sample this posterior distribution. The resulting samples are used to estimate the unknown model parameters and hyperparameters in an unsupervised framework. The results obtained on real data illustrate the improvement in reconstruction quality when compared to some existing techniques.

Index Terms— Spectral imaging, compressive sampling, Bayesian reconstruction

1. INTRODUCTION

Compressive spectral imaging (CSI) has been recently receiving a lot of attention in the signal and image processing communities [1, 2, 3, 4, 5]. CSI consists of recovering the full spatial and spectral information of a scene from a significantly undersampled set of random projections acquired by a compressive spectral imager, such as the Coded Aperture Snapshot Spectral Imaging (CASSI) [6, 7, 8].

When using a significant undersampling rate, there is an infinite number of possible images that can be associated with a particular set of measurements. Thus, it is usual to promote realistic properties of the image to recover in order to regularize the problem and obtain a unique solution. One of the most common regularizations consists of enforcing the image of interest to have a sparse representation in a given basis [9, 10, 11, 12]. Although this technique has shown satisfactory results, the quality of the reconstruction can be further improved by exploiting additional properties of natural images such as their important degree of local structure similarity [13, 14].

In this work, we introduce a new hierarchical Bayesian model for the reconstruction of compressed hyperspectral images. This model promotes the solution to be sparse in a selected basis and smooth in the image domain. A Gibbs sampler is used to draw samples asymptotically distributed ac-

ording to its posterior distribution. The generated samples are then used to build estimators of the unknown model parameters.

The proposed model provides improved reconstruction quality due to the exploitation of the local structure similarity of natural images. It also allows the model hyperparameters to be estimated from the observed measurements in an unsupervised framework (instead of fixing them a-priori as required by most convex optimization methods [15, 16, 17]). Finally, the samples generated by the Gibbs sampler can be used to determine measures of uncertainty for the estimates such as the estimation variance of confidence intervals.

The paper is organized as follows: Section 2 introduces the proposed Bayesian model. Section 3 presents the Gibbs sampler used to generate samples asymptotically distributed according to the posterior of this model. Experiments conducted with real data are shown in Section 4. Conclusions and future work are reported in Section 5.

2. PROPOSED METHOD

The CSI inverse problem addressed in this paper can be formulated as follows

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{e} \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^{MNL}$ is the lexicographically ordered vectorization of an image of size $M \times N$ with L spectral bands, $\Phi \in \mathbb{R}^{P \times MNL}$ is the sparse measurement matrix that depends on the compressive sensing imager (see [1, 2, 3, 4] for more details), $\mathbf{e} \in \mathbb{R}^P$ is an additive white Gaussian noise and $\mathbf{y} \in \mathbb{R}^P$ is the observed measurement vector. In this paper, we concentrate on the CASSI sampler [6, 7, 8], whose sensing matrix Φ has the structure shown in Fig. 1. We propose to recover \mathbf{x} from \mathbf{y} by solving the following problem

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \left(\frac{1}{2\sigma_n^2} \|\mathbf{H}\boldsymbol{\theta} - \mathbf{y}\|^2 + \tau \|\boldsymbol{\theta}\|_1 + \lambda \|(\mathbf{B} - \mathbf{I})\boldsymbol{\Psi}\boldsymbol{\theta}\|^2 \right) \quad (2)$$

where $\mathbf{H} = \Phi\Psi$, $\boldsymbol{\theta} \in \mathbb{R}^{MNL}$ contains the coefficients of \mathbf{x} in a given basis $\Psi \in \mathbb{R}^{MNL \times MNL}$ (i.e., $\mathbf{x} = \Psi\boldsymbol{\theta}$) and $\mathbf{B} \in \mathbb{R}^{MNL \times MNL}$ is a weighting matrix (as the one considered in [14, 18]) associated with a low-pass filter. Note that the first term in (2) is the data-fidelity term, that the second term is a penalty enforcing sparsity of $\boldsymbol{\theta}$ and the third term promotes image smoothness by encouraging each pixel

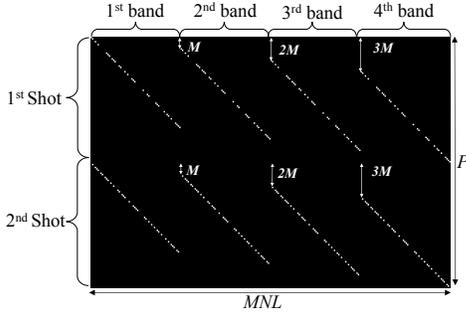


Fig. 1. Example of CASSI sensing matrix Φ for $M = N = 8$, $L = 4$ using two shots.

of \mathbf{x} to be close to a linear combination of its neighbors. The model hyperparameters τ and λ adjust the relative importance of the three terms. In the following section, we show that the inverse problem (2) is equivalent to the determination of the maximum a posteriori (MAP) estimator of a Bayesian model whose likelihood and priors are provided.

2.1. Likelihood

Assuming that the additive noise in (1) is white Gaussian with variance σ_n^2 , the distribution of the observation vector is a Gaussian distribution with mean vector $\mathbf{H}\boldsymbol{\theta}$ and covariance matrix $\sigma_n^2 \mathbf{I}_P$, i.e.,

$$f(\mathbf{y}|\boldsymbol{\theta}, \sigma_n^2) = \mathcal{N}(\mathbf{y}|\mathbf{H}\boldsymbol{\theta}, \sigma_n^2 \mathbf{I}_P) \quad (3)$$

where \mathbf{I}_P is the $P \times P$ identity matrix and \mathcal{N} denotes the Gaussian distribution. This likelihood is the Bayesian equivalent of the data-fidelity term previously shown in (2).

2.2. Prior distributions

2.2.1. Image coefficients $\boldsymbol{\theta}$

To promote sparsity in the wavelet domain and smoothness in the image domain, the following prior distribution is introduced for the image basis coefficients $\boldsymbol{\theta}$

$$f(\boldsymbol{\theta}|\lambda, \tau) \propto \mathcal{N}(\boldsymbol{\theta}|0, \mathbf{C}/\lambda) \prod_{i=1}^{NML} \exp(-\tau|\theta_i|) \quad (4)$$

with $\mathbf{C}^{-1} = \boldsymbol{\Psi}^T(\mathbf{B} - \mathbf{I})^T(\mathbf{B} - \mathbf{I})\boldsymbol{\Psi}$, and where λ and τ are two hyperparameters. This prior can be shown to be the Bayesian equivalent of the regularization terms in (2). To simplify the analysis, it is convenient to define the hyperparameter $a = \sigma_n^2 \tau^2$ which transforms the prior to

$$f(\boldsymbol{\theta}|\lambda, a, \sigma_n^2) \propto \mathcal{N}(\boldsymbol{\theta}|0, \mathbf{C}/\lambda) \prod_{i=1}^{NML} \exp\left(-\sqrt{\frac{a}{\sigma_n^2}}|\theta_i|\right). \quad (5)$$

It is then easy to show that (5) is the marginal distribution of the data-augmented density $f(\boldsymbol{\theta}, \boldsymbol{\delta}^2|\lambda, \tau)$, where

$\boldsymbol{\delta}^2 \in (\mathbb{R}^+)^{NML}$ is a vector of independent latent variables distributed according to gamma distributions such that (see [19] for details)

$$f(\boldsymbol{\delta}^2|a) = \prod_{i=1}^{NML} \mathcal{G}(\delta_i^2|1, \frac{a}{2}). \quad (6)$$

The advantage of using this data augmentation scheme is that the conditional distribution of $\boldsymbol{\theta}|\lambda, \boldsymbol{\delta}^2, \sigma_n^2$ is much easier to sample than the conditional distribution of $\boldsymbol{\theta}|\lambda, a, \sigma_n^2$ since

$$f(\boldsymbol{\theta}|\lambda, \boldsymbol{\delta}^2, \sigma_n^2) \propto \mathcal{N}(\boldsymbol{\theta}|0, \mathbf{C}/\lambda) \mathcal{N}(\boldsymbol{\theta}|0, \sigma_n^2 \boldsymbol{\Delta}) \quad (7)$$

where $\boldsymbol{\Delta} \in \mathbb{R}^{NML \times NML}$ is a diagonal matrix whose i th diagonal element is δ_i^2 .

2.2.2. Noise variance σ_n^2

The noise variance is assigned a Jeffrey's prior

$$f(\sigma_n^2) \propto \frac{1}{\sigma_n^2} \mathbf{1}_{\mathbb{R}^+}(\sigma_n^2) \quad (8)$$

where $\mathbf{1}_{\mathbb{R}^+}(\xi) = 1$ if $\xi \in \mathbb{R}^+$ and 0 otherwise (indicator function on \mathbb{R}^+). This choice is very classical when no information about a scale parameter is available (see [20]).

2.3. Hyperparameter priors

In order to be able to estimate the hyperparameters a and λ of a Bayesian model, one can define a hierarchical Bayesian model defined by hyperparameter priors. The hyperpriors considered in this work are summarized in this section.

2.3.1. Hyperprior of a

A Jeffrey's prior is assigned to a to keep the system as non-informative as possible

$$f(a) \propto \frac{1}{a} \mathbf{1}_{\mathbb{R}^+}(a). \quad (9)$$

2.3.2. Hyperprior of λ

A conjugate gamma hyperprior is assigned to λ

$$f(\lambda) = \mathcal{G}(\lambda|\alpha_\lambda, \beta_\lambda). \quad (10)$$

The values of α_λ and β_λ are chosen to make the hyperprior non-informative. The values we used for our experiments are specified in Section 4.

2.4. Posterior distribution

Using the likelihood as well as the parameter and hyperparameter priors defined in the previous sections, the posterior distribution associated with the proposed hierarchical Bayesian model is

$$f(\sigma_n^2, \boldsymbol{\theta}, a, \lambda, \boldsymbol{\delta}^2|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\theta}, \sigma_n^2) f(\boldsymbol{\theta}|\lambda, \boldsymbol{\delta}^2, \sigma_n^2) f(\boldsymbol{\delta}^2|a) f(\sigma_n^2) f(a) f(\lambda). \quad (11)$$

3. GIBBS SAMPLER

The posterior distribution (11) is intractable, in the sense that it does not allow closed-form expressions for the Bayesian estimators of the its parameters and hyperparameters to be derived. Thus we propose to draw samples distributed according to (11) using a Markov chain Monte Carlo (MCMC) method. The generated samples are then used to estimate the unknown image jointly with the other model parameters and hyperparameters. More precisely, after an appropriate burn-in period, the samples associated with a given parameter generated by the MCMC method are averaged in order to approximate the minimum mean square error estimator of this parameter. The MCMC method considered in this paper is a Gibbs sampler, which generates the unknown variables sequentially according to their conditional distributions, as shown in Algorithm 1 (see [20]). Note that the full vector θ can be sampled in a single step since it is more efficient than sampling each element of θ separately (it allows convergence to be reached with a smaller number of iterations). The corresponding conditional distributions are shown in Table 1 where \mathcal{GIG} , \mathcal{N} , \mathcal{G} , and \mathcal{IG} are the generalized inverse Gaussian, normal, gamma, and inverse gamma distributions (see [19] for details about the \mathcal{GIG} distribution).

Algorithm 1 Gibbs sampler.

```

Initialize  $a, \sigma_n^2$  and  $\lambda$ 
Sample  $\theta$  from its prior distribution
repeat
  for  $i = 1$  to  $N$  do
    Sample  $\delta_i^2$  from  $f(\delta_i^2|\theta_i, \sigma_n^2, a)$ 
  end for
  Sample  $\theta$  from  $f(\theta|\mathbf{y}, \sigma_n^2, \delta^2, \lambda)$ 
  Sample  $\lambda$  from  $f(\lambda|\theta)$ 
  Sample  $a$  from  $f(a|\delta^2)$ 
  Sample  $\sigma_n^2$  from  $f(\sigma_n^2|\mathbf{y}, \theta, \delta^2)$ 
until convergence

```

δ_i^2	$\mathcal{GIG}\left(\frac{1}{2}, a, \frac{\theta_i^2}{\sigma_n^2}\right)$
θ	$\mathcal{N}\left(\frac{\Sigma \mathbf{H}^T \mathbf{y}}{\sigma_n^2}, \Sigma\right), \Sigma^{-1} = \frac{1}{\sigma_n^2}(\mathbf{H}^T \mathbf{H} + \mathbf{\Delta}^{-1}) + \lambda \mathbf{C}^{-1}$
λ	$\mathcal{G}\left(\frac{NML}{2} + \alpha_\lambda, \frac{\ (\mathbf{B}-\mathbf{I})\Psi\theta\ ^2}{2} + \beta_\lambda\right)$
a	$\mathcal{G}\left(NML, \frac{\ \delta\ ^2}{2}\right)$
σ_n^2	$\mathcal{IG}\left(\frac{NML+P}{2}, \frac{1}{2}\left(\ \mathbf{y} - \mathbf{H}\theta\ ^2 + \sum \frac{\theta_i^2}{\delta_i^2}\right)\right)$

Table 1. Full conditionals $f(\delta_i^2|\theta_i, \sigma_n^2, a)$, $f(\theta|\mathbf{y}, \sigma_n^2, \delta^2, \lambda)$, $f(\lambda|\theta)$, $f(a|\delta^2)$ and $f(\sigma_n^2|\mathbf{y}, \theta, \delta^2)$ associated with (11).

3.1. Sampling considerations

The variables δ_i^2 , λ , a and σ_n^2 can be easily sampled using standard generators of random variables. However, sampling

all the elements of θ jointly in a direct manner is not straightforward, since the inversion of the precision matrix $\Sigma^{-1} \in \mathbb{R}^{NML \times NML}$ becomes intractable even for small hyperspectral images. In order to solve this problem, we propose to use the sampling technique introduced by Orioux *et al* [21] that was proposed to sample from high dimensional multivariate Gaussian distributions. Note that the method bypasses the problem of inverting the precision matrix Σ^{-1} by using a perturbation-optimization algorithm. A conjugate gradient method has been used in this paper to solve the perturbation-optimization problem.

4. SIMULATION RESULTS

In order to evaluate the performance of the proposed algorithm, experiments were performed on a real hyperspectral image. The image was acquired using a monochromator with wavelengths separated by 1nm, more precisely with a CCD camera AVT Marlin F033B, with 656×492 pixels, 24 spectral bands and a pixel pitch of $9.9\mu\text{m}$. The experiments were conducted on a section of 128×128 pixels and 8 spectral bands (from 461nm to 596nm) that was extracted from the acquired image. Five matrices Φ were calculated, each corresponding to a random realization of a CASSI sensing matrix with a different compression ratio $\frac{P}{MNL}$, more precisely 13%, 26%, 40%, 53%, and 66%. Compressed images were then generated by multiplying the hypercube by each of the matrices Φ and adding Gaussian noise to obtain SNR = 25dB.

Four different algorithms were used to recover the hypercube from each set of measurements: *i*) The SpaRSA optimization algorithm for the LASSO problem (SpaRSA LASSO) [10], *ii*) the SpaRSA algorithm for solving the proposed problem (2) (SpaRSA smooth), *iii*) the Bayesian LASSO [22] and *iv*) the proposed method. Note that algorithms *ii* and *iv* estimate the image by solving (2) whereas algorithms *i* and *iii* do not use the smoothing term. Note also that algorithms *i* and *ii* require to set a priori values of σ_n^2 , τ and λ , e.g., by using cross-validation, whereas algorithms *iii* and *iv* are implemented in a Bayesian framework, that estimates its own hyperparameters from the observed data.

The basis representation Ψ used in all experiments is defined as the Kronecker product of three bases $\Psi = \Psi_1 \otimes \Psi_2 \otimes \Psi_3$, where the combination $\Psi_1 \otimes \Psi_2$ is the 2D-Wavelet Symmlet 8 basis and Ψ_3 is the cosine basis. For algorithms *ii* and *iv*, \mathbf{B} was chosen as a low-pass Gaussian filter with radius 3 and standard deviation 0.6. Finally, the hyperprior parameters were fixed by cross validation to $\alpha_\lambda = 1 \times 10^{-3}$ and $\beta_\lambda = 10 \times 10^{-6}$.

Table 2 shows the PSNRs of the reconstructed images obtained for different compression ratios of Φ and for the four algorithms. We can observe that the algorithms using the smoothness term have a performance that is up to 2dB higher than the ones that do not. Note that the optimization algorithms yield slightly better results than their Bayesian

counterparts. We believe that this improved performance is due to the fact that the hyperparameters in the optimization algorithms have been manually set to optimize the PSNR whereas these hyperparameters are estimated from the data by the Bayesian algorithms. Fig. 2 illustrates the seventh spectral band recovered by the algorithms when applied to the measurements corresponding to a 53% compression ratio. The algorithms that include the smoothness term clearly provide results that are visually less noisy.

Compression ratio	13%	26%	40%	53%	66%
Proposed method	24.4	27.1	28.6	29.6	30.4
Bayesian LASSO	22.9	26.0	27.5	28.4	28.4
SpaRSA smooth	25.2	27.1	28.8	29.7	30.6
SpaRSA LASSO	23.5	26.8	28.5	29.4	30.4

Table 2. PSNRs [dB] obtained by the different algorithms.

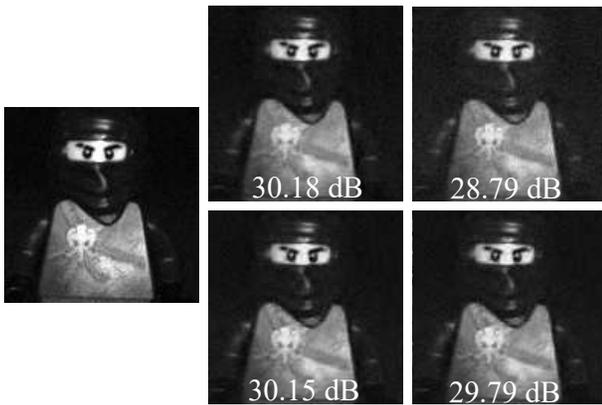


Fig. 2. Seventh spectral band of the image: (Left) Ground truth. Reconstruction results for: (top center) the proposed method, (bottom center) SpaRSA Smooth, (top right) Bayesian LASSO and (bottom right) SpaRSA LASSO.

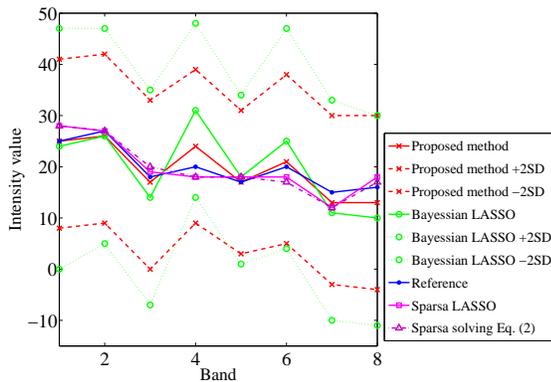


Fig. 3. Spectral signature for pixel #(20, 33).

Examples of reconstructions along the spectral axis for spatial pixel #(20, 33) (selected randomly) are compared in Fig. 3. The methods that use the smoothing term are closer

Compression ratio	13%	26%	40%	53%	66%
Proposed method	10.85	10.51	9.79	9.21	9.05
Bayesian LASSO	25.90	18.21	15.06	13.10	13.10

Table 3. Mean standard deviations of the estimations of x .

Computational cost	Iterations	Seconds
Proposed method	500	20×10^3
Bayesian LASSO	750	18×10^3
SpaRSA smooth	300	316
SpaRSA LASSO	300	42

Table 4. Computational costs for a 53% compression ratio.

to the ground truth. In addition, the Bayesian methods are able to calculate the standard deviation (SD) of their estimations, which is included in the graph showing that the reference is within their 2 SD error margin. Table 3 shows the values of SD for both Bayesian approaches. The proposed method presents significantly lower values of SD, providing more confident estimates than the Bayesian LASSO.

Finally, Table 4 shows the computational costs obtained with the algorithms for measurements associated with a 53% compression ratio. All algorithms were ran on a personal computer with an Intel Core i7-4790 CPU 3.60GHz processor and 32GB of memory. Algorithms were implemented using MATLAB with MEX files written in C. The algorithms based on MCMC methods are significantly slower than the ones based on optimization, as usual. However, it is important to note that optimization algorithms require to have their hyperparameters manually adjusted, and thus have to be run several times to find the optimal solution. Conversely, Bayesian methods can estimate their hyperparameters jointly with the recovered image coefficients using a single run.

5. CONCLUSION

This paper introduced a hierarchical Bayesian model to solve the compressive spectral imaging problem by promoting the image to be sparse in a given basis and smooth in the spatial domain. A Gibbs sampler was developed to draw samples asymptotically distributed according to the corresponding posterior, sampling the full image in a single step to accelerate the convergence speed. The generated samples were used to calculate the Bayesian estimators of the unknown image. The resulting algorithm was compared to other reconstruction methods for a hyperspectral image compressed with different compression ratios. Our experiments showed that including a spatial smoothing term can improve the PSNR of the recovered image up to 2dB. Future work includes the introduction of different regularization terms to promote smoothness of the recovered image, such as total variation. It would also be interesting to analyze the effect of small uncertainties affecting the sensing matrix on the performance of the image reconstruction algorithm.

6. REFERENCES

- [1] C. V. Correa, H. Arguello, and G. R. Arce, "Snapshot colored compressive spectral imager," *Journal of the Optical Society of America A*, vol. 32, no. 10, pp. 1754–1763, 2015.
- [2] X. Lin, Y. Liu, J. Wu, and Q. Dai, "Spatial-spectral encoded compressive hyperspectral imaging," *ACM Trans. Graph.*, vol. 33, no. 6, pp. 233:1–233:11, 2014.
- [3] L. Wang, Z. Xiong, D. Gao, G. Shi, and F. Wu, "Dual-camera design for coded aperture snapshot spectral imaging," *Applied optics*, vol. 54, no. 4, pp. 848–58, 2015.
- [4] H. Arguello and G. R. Arce, "Colored Coded Aperture Design by Concentration of Measure in Compressive Spectral Imaging," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1896–1908, 2014.
- [5] O. Eches, N. Dobigeon, C. Mailhes, and J. Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model. application to hyperspectral imagery," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1403–1413, June 2010.
- [6] H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, "Higher-order computational model for coded aperture spectral imaging," *Appl. Opt.*, vol. 52, no. 10, pp. D12–D21, Apr 2013.
- [7] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–B51, Apr 2008.
- [8] R. M. Willett, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, "Sparsity and structure in hyperspectral imaging : Sensing, reconstruction, and target detection," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 116–126, Jan 2014.
- [9] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [10] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, July 2009.
- [11] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: From theory to applications, a survey," *Journal of Communications and Networks*, vol. 15, no. 5, pp. 443–456, Oct 2013.
- [12] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Applied and Numerical Harmonic Analysis. Springer New York, 2013.
- [13] J. Fessler, "Image reconstruction," in *Image Reconstruction*, chapter 1, pp. 1–18. 2016.
- [14] U. Feige, "Why are images smooth?," in *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, New York, NY, USA, 2015, ITCS '15, pp. 229–236, ACM.
- [15] Y. Qiu, J. Yan, and F. Xu, "Nonmonotone adaptive Barzilai-Borwein gradient algorithm for compressed sensing," *Abstract and Applied Analysis*, vol. 2014, 2014.
- [16] G. Chen, D. Li, and J. Zhang, "Iterative gradient projection algorithm for two-dimensional compressive sensing sparse image reconstruction," *Signal Processing*, vol. 104, pp. 15–26, 2014.
- [17] B. Stephen, J. Bobin, and E. J. Candes, "Nesta: A fast and accurate first-order method for sparse recovery," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.
- [18] K. Zhang, X. Gao, D. Tao, and X. Li, "Single Image Super-Resolution With Non-Local Means and Steering Kernel Regression," *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4544–4556, nov 2012.
- [19] S. Raman, T. J Fuchs, P. J Wild, E. Dahl, and V. Roth, "The Bayesian group-lasso for analyzing contingency tables," in *Proc. 26th ACM Annu. Int. Conf. Mach. Learn. (ICML)*, Montreal, Quebec, Jun. 2009.
- [20] G. Casella and C. P Robert, *Monte Carlo Statistical Methods*, New York: Springer-Verlag, 1999.
- [21] F. Orieux, O. Féron, and J-F Giovannelli, "Sampling high-dimensional Gaussian distributions for general linear inverse problems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 251–254, 2012.
- [22] T. Park and G. Casella, "The Bayesian Lasso," *J. Amer. Stat. Soc.*, vol. 103, no. 482, pp. 681–686, 2008.