

Bayesian Reconstruction of Hyperspectral Images by Using CS Measurements and a Local Structured Prior

ICASSP - New Orleans

Facundo Costa, Jean-Yves Tournet, Hadj Batatia ⁽¹⁾,
Yuri Mejía, Henry Arguello ⁽²⁾

(1) University of Toulouse, ENSEEIHT-IRIT-TéSA, France

(2) Universidad Industrial de Santander, Colombia

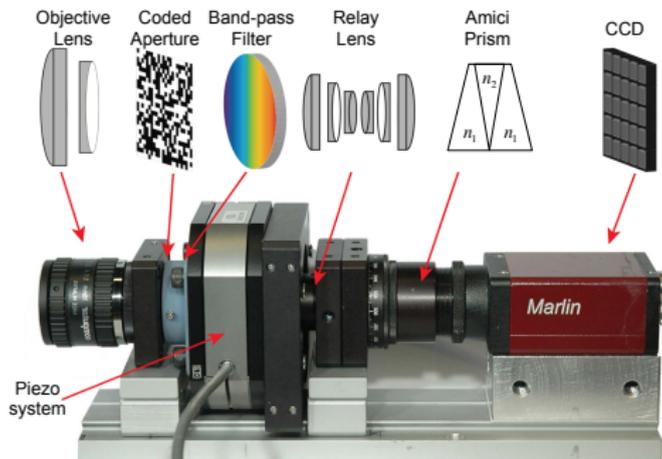
jyt@n7.fr

March 2017

Outline

- ▶ 1: Compressive Sensing Measurements
- ▶ 2: Bayesian Formulation
- ▶ 3: An MCMC Method
- ▶ 4: Spatial Regularization
- ▶ 5: Simulation Results

Compressive Spectral Imaging (CSI)



Coded Aperture Snapshot Spectral Imaging (CASSI)

Compressive Spectral Imaging (CSI): recovering the full spatial and spectral information of a scene from undersampled random projections acquired by a compressive spectral imager such as CASSI.

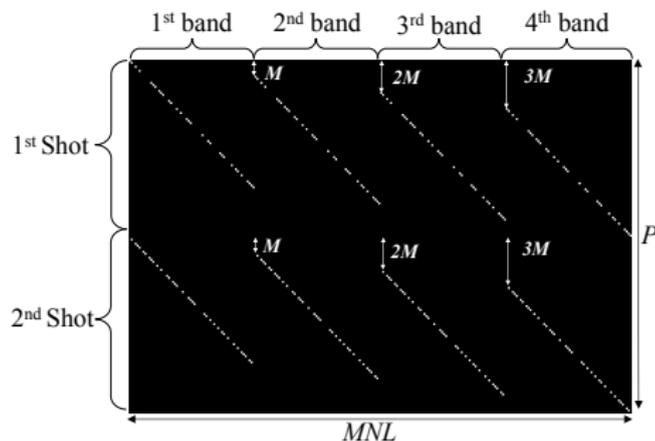
Compressive Sensing Measurements

Sensing matrix

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{n}$$

where Φ is fixed and \mathbf{n} is an additive Gaussian noise, i.e., $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_P)$

- ▶ Diagonal patterns related to the coded aperture
- ▶ Shifted patterns due to the prism effect
- ▶ Possible acquisition of multiple snapshots



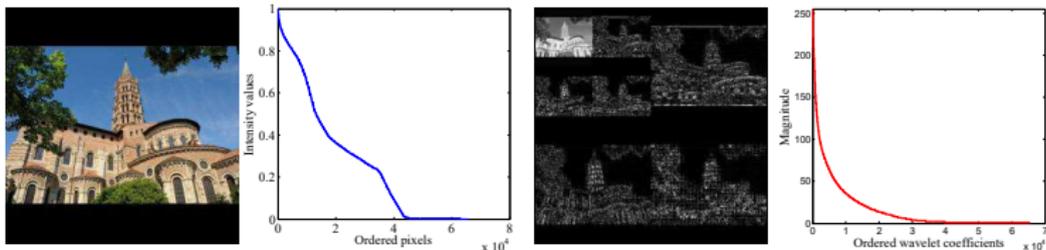
Compressive Sensing Measurements

Sparse representation of the image

$$x = \Psi\theta$$

where Ψ is constructed from predefined atoms

e.g., using the wavelet transform



Problem: how to estimate the unknown image x from compressed measurements $y = \Phi x + n$?

Fusion as an Inverse Problem

Data fidelity term

$$\frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 = \frac{1}{2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|_2^2$$

Sparse regularization

$$\varphi_1(\boldsymbol{\theta}) = \|\boldsymbol{\theta}\|_1$$

Spatial regularization

$$\varphi_2(\boldsymbol{\theta}) = \|(\mathbf{B} - \mathbf{I})\Psi\boldsymbol{\theta}\|_2^2$$

where \mathbf{B} is an appropriate weighting matrix (low-pass filter)

Conclusion

$$\arg \min_{\boldsymbol{\theta}} \left[\frac{1}{2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|_2^2 + \tau\varphi_1(\boldsymbol{\theta}) + \lambda\varphi_2(\boldsymbol{\theta}) \right]$$

Outline

- ▶ 1: Compressive Sensing Measurements
- ▶ 2: Bayesian Formulation
- ▶ 3: An MCMC Method
- ▶ 4: Spatial Regularization
- ▶ 5: Simulation Results

Bayesian LASSO¹

Observation model

$$\mathbf{y} = \mathbf{H}\boldsymbol{\theta} + \mathbf{n}$$

where $\boldsymbol{\theta}$ is sparse and $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_P)$

Optimization problem

$$\arg \min_{\boldsymbol{\theta}} \left[\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2 + \tau \|\boldsymbol{\theta}\|_1 \right]$$

Problem: how to adjust the regularization parameter τ ?

Equivalent problem

$$\arg \max_{\boldsymbol{\theta}} \left[\exp \left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2 \right) \exp(-\tau \|\boldsymbol{\theta}\|_1) \right]$$

¹ T. Park and G. Casella, "The Bayesian Lasso," Journal of the American Statistical Association, vol. 103, no. 482, pp. 681-686, 2008.

Bayesian LASSO

Observation model

$$\mathbf{y} = \mathbf{H}\boldsymbol{\theta} + \mathbf{n}, \quad \text{sparse } \boldsymbol{\theta}, \mathbf{n} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_P)$$

Bayesian formulation

▶ Gaussian likelihood

$$f(\mathbf{y}|\boldsymbol{\theta}) = \mathcal{N}(\mathbf{H}\boldsymbol{\theta}, \sigma_n^2 \mathbf{I}_N) \propto \exp\left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2\right)$$

▶ Independent Laplacian priors

$$f(\boldsymbol{\theta}|\tau) = \prod_{i=1}^P \exp(-\tau|\theta_i|) = \exp(-\tau\|\boldsymbol{\theta}\|_1)$$

▶ Posterior

$$f(\boldsymbol{\theta}|\mathbf{y}) \propto \exp\left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2\right) \exp(-\tau\|\boldsymbol{\theta}\|_1)$$

Hierarchical Bayesian Model

- ▶ Gaussian likelihood

$$f(\mathbf{y}|\boldsymbol{\theta}, \sigma_n^2) = \mathcal{N}(\mathbf{H}\boldsymbol{\theta}, \sigma_n^2 \mathbf{I}_P) \propto \exp\left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2\right)$$

- ▶ Independent Laplacian priors

$$f(\boldsymbol{\theta}|\tau) = \prod_{i=1}^p \exp(-\tau|\theta_i|) = \exp(-\tau\|\boldsymbol{\theta}\|_1)$$

- ▶ Joint noise variance and hyperparameter prior

$$\pi(\tau, \sigma_n^2)$$

- ▶ Posterior

$$f(\boldsymbol{\theta}, \sigma_n^2, \tau|\mathbf{y}) \propto \exp\left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2\right) \exp(-\tau\|\boldsymbol{\theta}\|_1) \pi(\tau, \sigma_n^2)$$

How can we estimate $\boldsymbol{\theta}, \sigma_n^2, \tau$?

Outline

- ▶ 1: Compressive Sensing Measurements
- ▶ 2: Bayesian Formulation
- ▶ 3: An MCMC Method
- ▶ 4: Spatial Regularization
- ▶ 5: Simulation Results

The Bayesian LASSO

Posterior

$$f(\boldsymbol{\theta}, \sigma_n^2, \tau | \mathbf{y}) \propto \exp\left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2\right) \exp(-\tau \|\boldsymbol{\theta}\|_1) \pi(\tau, \sigma_n^2)$$

Completion

- Scale mixture of a Gaussians distributions

$$\tau e^{-\tau|\theta_i|} = \int_0^\infty \frac{1}{\sqrt{2\pi s}} e^{-\frac{\theta_i^2}{2s}} \frac{\tau^2}{2} e^{-\frac{\tau^2 s}{2}} ds$$

- Hierarchical representation

$$\begin{aligned} \mathbf{y} &\sim \mathcal{N}(\mathbf{y}; \mathbf{H}\boldsymbol{\theta}, \sigma_n^2 \mathbf{I}_N) \\ \boldsymbol{\theta} | \sigma_n^2, s_1^2, \dots, s_p^2 &\sim \mathcal{N}(\boldsymbol{\theta}; \mathbf{0}_p, \sigma_n^2 \mathbf{D}_p), \quad \mathbf{D}_p = \text{diag}(s_1^2, \dots, s_p^2) \\ s_1^2, \dots, s_p^2 | \tau &\sim \prod_{j=1}^p \left(\frac{\tau^2}{2} e^{-\frac{\tau^2 s_j^2}{2}} \right), \quad \pi(\tau) \sim 1/\tau \\ \pi(\sigma_n^2) &\sim 1/\sigma_n^2 \quad (\text{Jeffreys prior}) \end{aligned}$$

Generalized Inverse Gaussian Distribution

$$\pi(x) = \left(\frac{a}{b}\right)^{1/4} K_{1/2}^{-1}(\sqrt{ab}) \frac{1}{\sqrt{x}} \exp\left[-\frac{1}{2}\left(\frac{b}{x} + ax\right)\right] I_{\mathbb{R}^+}(x)$$

where $K_{1/2}$ is a modified Bessel function, hence

$$\begin{aligned} \int_0^\infty \frac{1}{\sqrt{x}} \exp\left[-\frac{1}{2}\left(\frac{b}{x} + ax\right)\right] dx &= \left(\frac{b}{a}\right)^{1/4} K_{1/2}(\sqrt{ab}) \\ &= \sqrt{\frac{\pi}{2}} \frac{1}{\sqrt{a}} \exp(-\sqrt{ab}). \end{aligned}$$

can be used to demonstrate that the Laplace distribution is a scale mixture of Gaussian distributions.

Gibbs Sampler

Algorithm 1 Gibbs sampler

- 1: Initialize τ and σ_n^2
 - 2: Sample θ from its prior distribution
 - 3: **repeat**
 - 4: **for** $i = 1$ to p **do**
 - 5: Sample s_i^2 from $f(s_i^2|\theta_i, \sigma_n^2, \tau)$
 - 6: **end for**
 - 7: Sample θ from $f(\theta|\mathbf{y}, \sigma_n^2, \mathbf{s}^2)$
 - 8: Sample τ from $f(\tau|\theta)$
 - 9: Sample a from $f(a|\delta^2)$
 - 10: Sample σ_n^2 from $f(\sigma_n^2|\mathbf{y}, \theta, \delta^2)$
 - 11: **until** convergence
-

Outline

- ▶ 1: Compressive Sensing Measurements
- ▶ 2: Bayesian Formulation
- ▶ 3: An MCMC Method
- ▶ 4: Spatial Regularization
- ▶ 5: Simulation Results

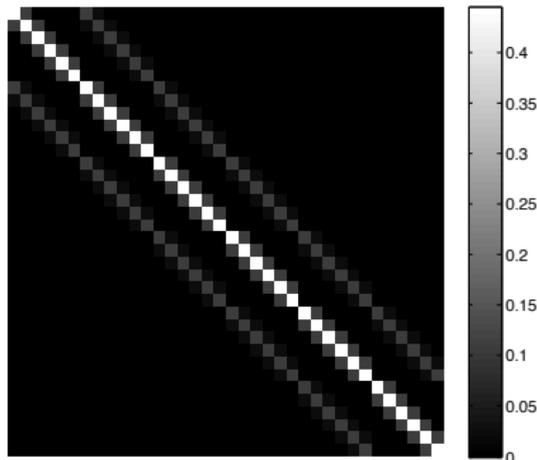
Include Spatial Regularization into Bayesian LASSO

Optimization problem

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



(a)



(b)

(a) Zero mean Gaussian filter of size 3×3 with $\sigma = 0.6$, (b) matrix B created by using the Gaussian filter of (a).

Include Spatial Regularization into the Bayesian LASSO

Bayesian formulation

- ▶ Equivalent problem

$$\arg \max_{\boldsymbol{\theta}} \left[\exp \left(-\frac{1}{2\sigma_n^2} \|\mathbf{y} - \mathbf{H}\boldsymbol{\theta}\|^2 - \tau \|\boldsymbol{\theta}\|_1 \right) \exp(-\lambda \|(\mathbf{B} - \mathbf{I})\boldsymbol{\Psi}\boldsymbol{\theta}\|^2) \right]$$

- ▶ Our proposal

Gibbs Sampler

Algorithm 2 Gibbs sampler

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

Conditional Distributions of $f(\sigma_n^2, \boldsymbol{\theta}, a, \lambda, \delta_i^2 | \mathbf{y})$

Full conditionals $f(\delta_i^2 | \theta_i, \sigma_n^2, a)$, $f(\boldsymbol{\theta} | \mathbf{y}, \sigma_n^2, \boldsymbol{\delta}^2, \lambda)$, $f(\lambda | \boldsymbol{\theta})$, $f(a | \boldsymbol{\delta}^2)$ and $f(\sigma_n^2 | \mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\delta}^2)$ associated with the posterior distribution of interest.

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

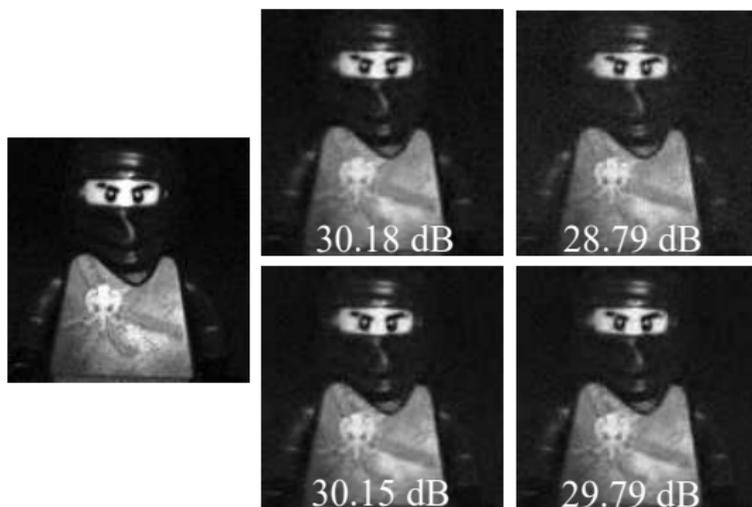
Sampling $\boldsymbol{\theta}$ using a **perturbation-optimization algorithm**²

²F. Orieux, O. Feron and J. F. Giovannelli, "Sampling High-Dimensional Gaussian Distributions for General Linear Inverse Problems," IEEE Signal Processing Letters, vol. 19, no. 5, pp. 251-254, May 2012.

Outline

- ▶ 1: Compressive Sensing Measurements
- ▶ 2: Bayesian Formulation
- ▶ 3: An MCMC Method
- ▶ 4: Spatial Regularization
- ▶ 5: Simulation Results

Qualitative Results



Seventh spectral band of the image: (Left) Ground truth. **Reconstruction results** for: (top center) the proposed method, (bottom center) SpaRSA³ Smooth, (top right) Bayesian LASSO and (bottom right) SpaRSA LASSO.

³ 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.

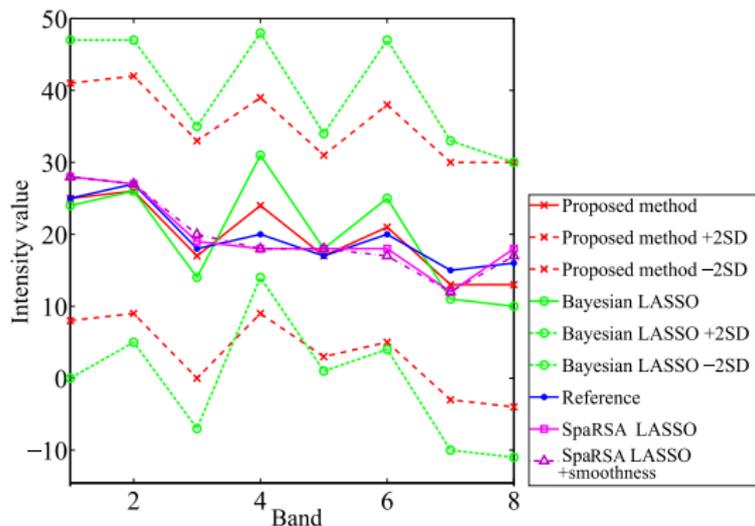


Figure: Spectral signature for pixel #(20, 33).

- ▶ The estimates obtained using the smoothing term are closer to the ground truth.
- ▶ Bayesian methods provide confidence measures for the estimates

Conclusions and Future Work

Conclusions

- ▶ **Hierarchical Bayesian model** solving 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** sampling the full image in a single step using a **perturbation optimization algorithm**
- ▶ Including **a spatial smoothing term** can improve the PSNR of the recovered image **up to 2dB**.

Prospects

- ▶ Other regularization terms: **Total Variation?** **l_p regularization?**
- ▶ Analyze the effects of the **sensing matrix** on the reconstruction performance and design an **optimal sensing matrix**

Thanks

Basis Representation

Basis representation

$$\Psi = \Psi_1 \otimes \Psi_2 \otimes \Psi_3$$

- ▶ $\Psi_1 \otimes \Psi_2$: 2D-Wavelet Symmlet 8 basis
- ▶ Ψ_3 : cosine basis.

PSNRs for Different Reconstruction Algorithms

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: PSNRs [dB] obtained by the different algorithms.

Computational Cost

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: Computational costs for a 53% compression ratio.