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

ICASSP - New Orleans

Facundo Costa, Jean-Yves Tourneret, 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)
Compressive Sensing Measurements

Sensing matrix

\[ y = \Phi x + n \]

where \( \Phi \) is fixed and \( n \) is an additive Gaussian noise, i.e., \( n \sim \mathcal{N}(0, \sigma_n^2 I_P) \)

- 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 \( \Psi \) 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

\[ \frac{1}{2} \| y - \Phi x \|_2^2 = \frac{1}{2} \| y - H\theta \|_2^2 \]

Sparse regularization

\[ \varphi_1(\theta) = \| \theta \|_1 \]

Spatial regularization

\[ \varphi_2(\theta) = \| (B - I) \Psi \theta \|_2^2 \]

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

Conclusion

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

- 1: Compressive Sensing Measurements
- 2: Bayesian Formulation
- 3: An MCMC Method
- 4: Spatial Regularization
- 5: Simulation Results
Bayesian LASSO\textsuperscript{1}

Observation model

\[ y = H\theta + n \]

where \( \theta \) is sparse and \( n \sim \mathcal{N}(0, \sigma_n^2 I_P) \)

Optimization problem

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

Problem: how to adjust the regularization parameter \( \tau \)?

Equivalent problem

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

Bayesian LASSO

Observation model

\[ y = H\theta + n, \quad \text{sparse } \theta, \ n \sim N(0, \sigma_n^2 I_P) \]

Bayesian formulation

- **Gaussian likelihood**
  
  \[ f(y|\theta) = \mathcal{N}(H\theta, \sigma_n^2 I_P) \propto \exp\left(-\frac{1}{2\sigma_n^2}||y - H\theta||^2\right) \]

- **Independent Laplacian priors**
  
  \[ f(\theta|\tau) = \prod_{i=1}^{p} \exp(-\tau |\theta_i|) = \exp(-\tau ||\theta||_1) \]

- **Posterior**
  
  \[ f(\theta|y) \propto \exp\left(-\frac{1}{2\sigma_n^2}||y - H\theta||^2\right) \exp(-\tau ||\theta||_1) \]
Hierarchical Bayesian Model

- **Gaussian likelihood**
  \[
  f(y|\theta, \sigma_n^2) = \mathcal{N}(H\theta, \sigma_n^2 I_P) \propto \exp\left(-\frac{1}{2\sigma_n^2}||y - H\theta||^2\right)
  \]

- **Independent Laplacian priors**
  \[
  f(\theta|\tau) = \prod_{i=1}^{p} \exp(-\tau|\theta_i|) = \exp(-\tau||\theta||_1)
  \]

- **Joint noise variance and hyperparameter prior**
  \[
  \pi(\tau, \sigma_n^2)
  \]

- **Posterior**
  \[
  f(\theta, \sigma_n^2, \tau|y) \propto \exp\left(-\frac{1}{2\sigma_n^2}||y - H\theta||^2\right) \exp(-\tau||\theta||_1)\pi(\tau, \sigma_n^2)
  \]

How can we estimate $\theta, \sigma_n^2, \tau$?
Outline

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

Posterior

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

Completion

- Scale mixture of a Gaussians distributions

\[ \frac{\tau}{2} e^{-\tau|\theta_i|} = \int_0^{\infty} \frac{1}{\sqrt{2\pi s_i^2}} e^{-\frac{\theta_i^2}{2s_i^2}} \frac{\tau^2}{2} e^{-\frac{\tau^2 s_i^2}{2}} ds_i^2 \]

- Hierarchical representation

\[ y \sim \mathcal{N}(y; H\theta, \sigma_n^2 I_P) \]
\[ \theta|\sigma_n^2, s_1^2, ..., s_p^2 \sim \mathcal{N}(\theta; 0_p, \sigma_n^2 D_p), \quad D_p = \text{diag}(s_1^2, ..., s_p^2) \]
\[ s_1^2, ..., s_p^2|\tau \sim \prod_{i=1}^{p} \left( \frac{\tau^2}{2} e^{-\frac{\tau^2 s_i^2}{2}} \right), \quad \pi(\tau) \sim 1/\tau \]
\[ \pi(\sigma_n^2) \sim 1/\sigma_n^2 \quad (\text{Jeffreys prior}) \]
Generalized Inverse Gaussian Distribution

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

where \( K_{1/2} \) is a modified Bessel function, hence

\[ \int_{0}^{\infty} \frac{1}{\sqrt{x}} \exp \left[ -\frac{1}{2} \left( \frac{b}{x} + ax \right) \right] dx = \left( \frac{b}{a} \right)^{1/4} K_{1/2} \left( \sqrt{ab} \right) = \sqrt{\frac{\pi}{2}} \frac{1}{\sqrt{a}} \exp \left( -\sqrt{ab} \right). \]

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

Algorithm 1 Gibbs sampler

1: Initialize $\tau$ and $\sigma_n^2$
2: Sample $\theta$ 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 $\theta$ from $f(\theta | y, \sigma_n^2, s^2)$
8: Sample $\tau$ from $f(\lambda | \theta)$
9: Sample $a$ from $f(a | \delta^2)$
10: Sample $\sigma_n^2$ from $f(\sigma_n^2 | y, \theta, \delta^2)$
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

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

(a) Zero mean Gaussian filter of size $3 \times 3$ 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

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

- Our proposal
Gibbs Sampler

**Algorithm 2** Gibbs sampler

Initialize $a$, $\sigma_n^2$ and $\lambda$
Sample $\theta$ from its prior distribution

repeat
  for $i = 1$ to $p$ do
    Sample $\delta_i^2$ from $f(\delta_i^2 | \theta_i, \sigma_n^2, a)$
  end for
  Sample $\theta$ from $f(\theta | y, \sigma_n^2, \delta^2, \lambda)$
  Sample $\lambda$ from $f(\lambda | \theta)$
  Sample $a$ from $f(a | \delta^2)$
  Sample $\sigma_n^2$ from $f(\sigma_n^2 | y, \theta, \delta^2)$
until convergence
Conditional Distributions of $f(\sigma^2_n, \theta, a, \lambda, \delta^2_i | y)$

Full conditionals $f(\delta^2_i | \theta_i, \sigma^2_n, a)$, $f(\theta | y, \sigma^2_n, \delta^2, \lambda)$, $f(\lambda | \theta)$, $f(a | \delta^2)$ and $f(\sigma^2_n | y, \theta, \delta^2)$ associated with the posterior distribution of interest.

<table>
<thead>
<tr>
<th>$\delta^2_i$</th>
<th>$\mathcal{GIG} \left( \frac{1}{2}, a, \frac{\theta^2_i}{\sigma^2_n} \right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>$\mathcal{N} \left( \frac{\Sigma H^T y}{\sigma^2_n}, \Sigma \right)$, $\Sigma^{-1} = \frac{1}{\sigma^2_n} (H^T H + \Delta^{-1}) + \lambda C^{-1}$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>$\mathcal{G} \left( \frac{NML}{2} + \alpha \lambda, \frac{</td>
</tr>
<tr>
<td>$a$</td>
<td>$\mathcal{G} \left( NML, \frac{</td>
</tr>
<tr>
<td>$\sigma^2_n$</td>
<td>$\mathcal{IG} \left( \frac{NML+P}{2}, \frac{1}{2} \left(</td>
</tr>
</tbody>
</table>

Sampling $\theta$ using a perturbation-optimization algorithm\(^2\)

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) SpaRSA$^3$ Smooth, (top right) Bayesian LASSO and (bottom right) SpaRSA LASSO.

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
- \( \Psi_3 \): cosine basis.
PSNRs for Different Reconstruction Algorithms

<table>
<thead>
<tr>
<th>Compression ratio</th>
<th>13%</th>
<th>26%</th>
<th>40%</th>
<th>53%</th>
<th>66%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td>24.4</td>
<td>27.1</td>
<td>28.6</td>
<td>29.6</td>
<td>30.4</td>
</tr>
<tr>
<td>Bayesian LASSO</td>
<td>22.9</td>
<td>26.0</td>
<td>27.5</td>
<td>28.4</td>
<td>28.4</td>
</tr>
<tr>
<td>SpaRSA smooth</td>
<td>25.2</td>
<td>27.1</td>
<td>28.8</td>
<td>29.7</td>
<td>30.6</td>
</tr>
<tr>
<td>SpaRSA LASSO</td>
<td>23.5</td>
<td>26.8</td>
<td>28.5</td>
<td>29.4</td>
<td>30.4</td>
</tr>
</tbody>
</table>

Table: PSNRs [dB] obtained by the different algorithms.
## Computational Cost

<table>
<thead>
<tr>
<th>Computational cost</th>
<th>Iterations</th>
<th>Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td>500</td>
<td>$20 \times 10^3$</td>
</tr>
<tr>
<td>Bayesian LASSO</td>
<td>750</td>
<td>$18 \times 10^3$</td>
</tr>
<tr>
<td>SpaRSA smooth</td>
<td>300</td>
<td>316</td>
</tr>
<tr>
<td>SpaRSA LASSO</td>
<td>300</td>
<td>42</td>
</tr>
</tbody>
</table>

**Table:** Computational costs for a 53% compression ratio.