# Nonparametric Learning of Dictionaries for Sparse Representation of Sensor Signals

Mingyuan Zhou, John Paisley and Lawrence Carin

Department of Electrical & Computer Engineering
Duke University, Durham, NC

# Outline

- Introduction
- The beta process for image inpainting
- Inpainting results for hyperspectral images

# Introduction

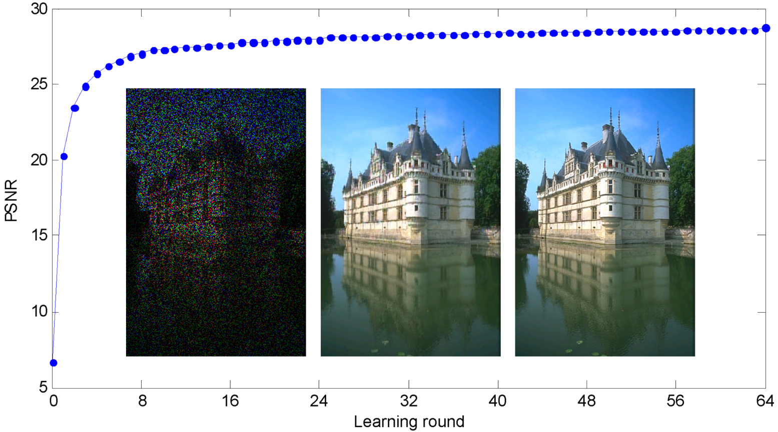
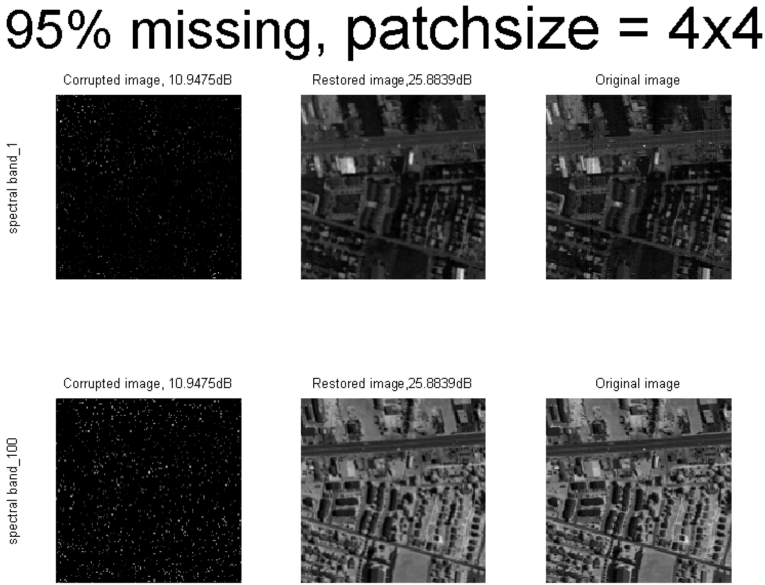


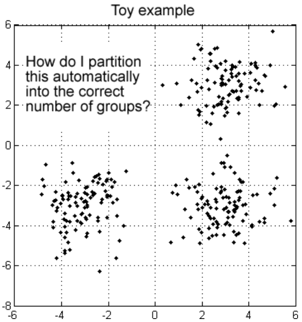

Figure: 480 $\times$ 320 image, 80% of data missing


95% missing, patchsize = 4x4
Corrupted image, 10.9475dB
Restored image,25.8839dB
Original image
spectral band_1
Corrupted image, 10.9475dB
Restored image,25.8839dB
Original image
spectral band_100

# Introduction

Nonparametric Bayesian Motivation: In mixture modeling, nonparametric Bayesian priors can be used to simplify a model's complexity. For example, in this Gaussian mixture model, only three Gaussians will be learned.

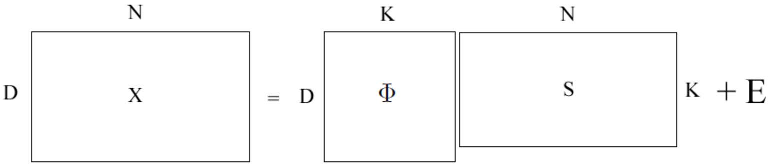
$$
\begin{aligned}
x_i &\sim \mathcal{N}(\mu_{c_i}, \Sigma_{c_i}) \\
c_i &\sim \sum_{k=1}^{K} \pi_k \delta_k \\
\pi &\sim \mathrm{Dirichlet}(\alpha/K, \ldots, \alpha/K) \\
\{\mu_k, \Sigma_k\} &\sim \mathrm{Normal-Wishart}
\end{aligned}
$$




The nonparametric prior used here is the Dirichlet Process.

# Introduction

The same issue arises with latent factor models:

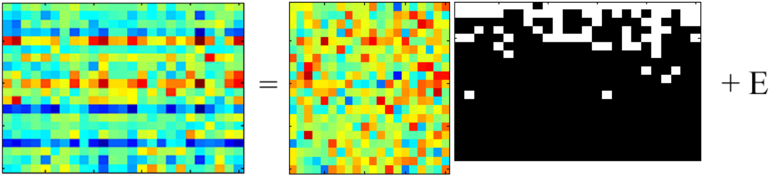



We want a nonparametric prior that will find a subset of the $K$ columns of $\Phi$ that are necessary to represent the data. The beta process prior achieves this end.

# Nonparametric latent factor models

An example of a draw from the prior to be discussed.

$$X = \Phi Z + E$$



# Nonparametric latent factor models

The generative process for the beta process factor analysis (BPFA) model is,

$$
\begin{aligned}
x_i &\sim \mathcal{N}\left(\Phi(w_i \circ z_i), \sigma_n^2 I\right) \\
w_{ik} &\sim z_{ik}\mathcal{N}\left(0, \sigma_{w_k}^2\right) + (1 - z_{ik})\delta_0 \\
z_{ik} &\sim \mathrm{Bernoulli}(\pi_k) \\
\pi_k &\sim \mathrm{Beta}\left(\frac{\alpha\gamma}{K}, \alpha(1 - \frac{\gamma}{K})\right) \\
\phi_k &\overset{iid}{\sim} H_0 \\
\sigma_n^{-2} &\sim \mathrm{Gamma}(a, b) \\
\sigma_{w_k}^{-2} &\sim \mathrm{Gamma}(c, d)
\end{aligned}
$$

where $H_0$ is a prior on a $D$-dimensional vector, e.g., $\mathcal{N}(0, \Sigma)$. For this prior, $S = W \circ Z$.

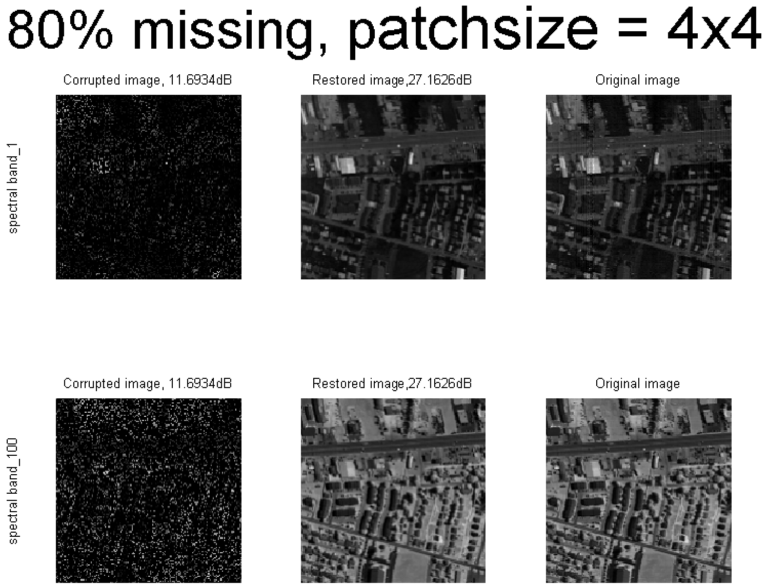
# The Indian Buffet Story

A particular representation of the beta process gives rise to a sequential process, which has the following story.

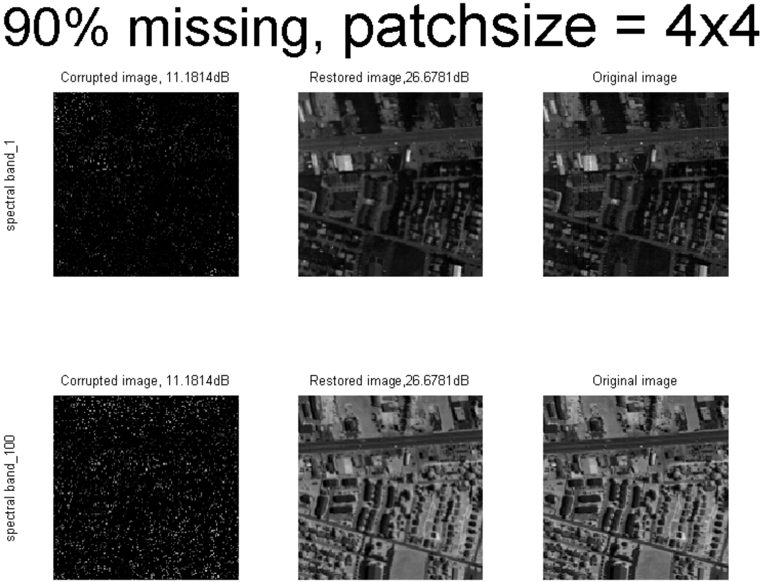
The Indian buffet process story:

1. The first customer walks into an Indian buffet and samples $c_1 \sim \mathrm{Poisson}(\gamma)$ dishes.
2. The $N^{th}$ customer samples a previously tasted dish with probability $\frac{m_k}{\alpha+N-1}$, and then samples $c_N \sim \mathrm{Poisson}\left(\frac{\alpha\gamma}{\alpha+N-1}\right)$ new dishes.

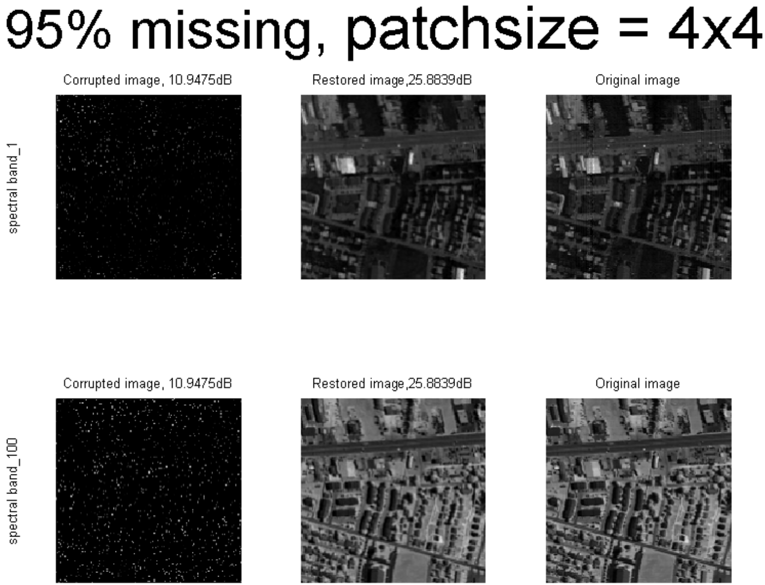
# Results


80% missing, patchsize = 4x4
spectral band_1
Corrupted image, 11.6934dB
Restored image,27.1626dB
Original image
spectral band_100
Corrupted image, 11.6934dB
Restored image,27.1626dB
Original image

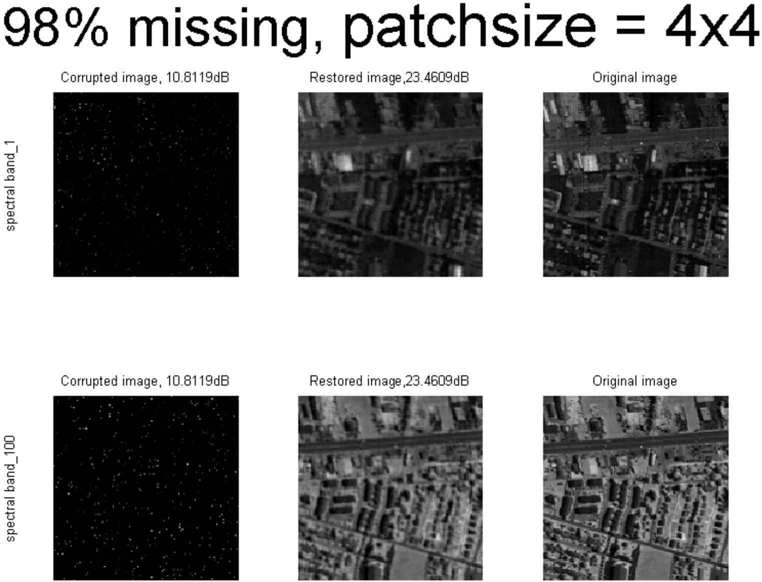
# Results


90% missing, patchsize = 4x4
Corrupted image, 11.1814dB
Restored image,26.6781dB
Original image
spectral band_1
Corrupted image, 11.1814dB
Restored image,26.6781dB
Original image
spectral band_100

# Results


95% missing, patchsize = 4x4
Corrupted image, 10.9475dB
Restored image,25.8839dB
Original image
spectral band_1
Corrupted image, 10.9475dB
Restored image,25.8839dB
Original image
spectral band_100

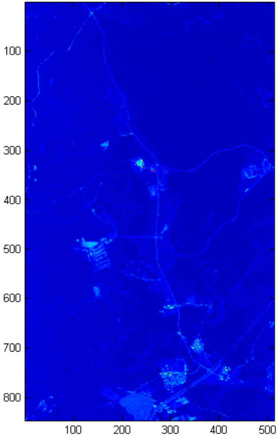
# Results



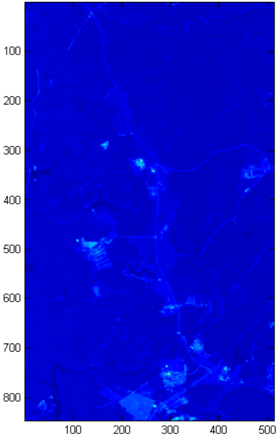
98% missing, patchsize = 4x4
Corrupted image, 10.8119dB
Restored image,23.4609dB
Original image
spectral band_1
Corrupted image, 10.8119dB
Restored image,23.4609dB
Original image
spectral band_100

# Results

98% missing, patchsize = 4x4, channel=1



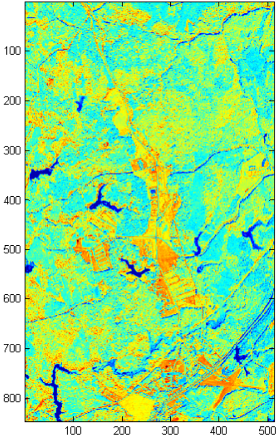

Original image



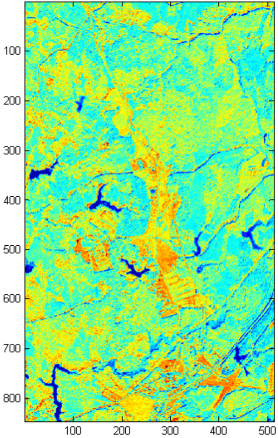

Restored image

# Results

98% missing, patchsize = 4x4, channel=50

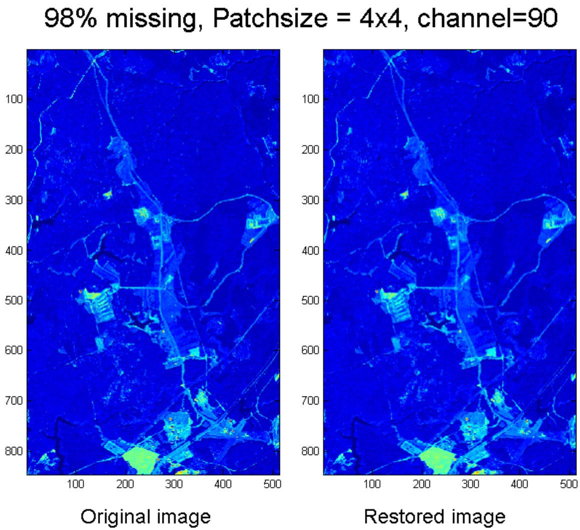



Original image




Restored image

# Results



98% missing, Patchsize = 4x4, channel=90
100
200
300
400
500
600
700
800
100 200 300 400 500
Original image
Restored image