Part 1 - Inverse Problems in Hyperspectral Imaging

José M. Bioucas Dias

Instituto de Telecomunicações
Instituto Superior Técnico
Universidade de Lisboa
Portugal

Société Française de Photogrammétrie et Télédétection
Grenoble, France, May, 2016
# Outline

## Part 1 - Brief overview of hyperspectral imaging in remote sensing
- The observation model (direct or forward problem)
- Degradation mechanisms (spatial blur and noise)
- Characterization of hyperspectral images (geometrical and statistical)
- Inverse problems in hyperspectral imaging (denoising, sharpening, unmixing)

## Part 2 - Inverse problems in a nutshell

## Part 3 - Denoising, sharpening, and unmixing
Measuring the radiation arriving the sensor with high spectral resolution over a sufficiently broad spectral band such that the acquired spectrum can be used to uniquely characterize and identify any given material.
Hyperspectral imaging: motivation

Landsat 7 TM bands

200 bands 10 mm
Remote sensing: basics

Radiance versus reflectance

\[ L(\lambda) = \frac{1}{\pi} E(\lambda) \rho(\lambda) \]

- \( E \) – Irradiance (W/m\(^2\))
- \( \rho \) – Reflectance
- \( L \) – Radiance (W/Sr/m\(^2\))
- \( \lambda \) – Wavelength (\(\mu\)m)
Remote sensing: the influence of atmosphere

Atmospheric molecules responsible for absorption [Lillesand & Keifer, 02]

(Energy blocked)

Photography

Hyperspectral scanners

Thermal scanners

Multispectral scanners

UV  VI  N-IR  MID-IR  TERMAL-IR  MIC

0.3μm  0.7μm  1μm  2.5μm  10μm  1mm
Spatial and spectral resolution trade-offs

The signal-to-ratio (SNR) associated with the Poissonian noise in a hyperspectral imaging system is given by ([Shaw & Burke 2003])

$$\text{SNR} \propto \frac{\Delta^2}{\text{ACR} \times R}$$

where $\Delta$ is the spatial resolution, $R$ is the number of bands, and ACR is the area coverage rate.

For the same SNR and ACR, we have

$$\frac{\Delta(R)}{\Delta(1)} = \sqrt{R}$$

In conclusion: Hyperspectral images tend to have low spatial resolution.
## Acquisition instruments

### Remote sensing

<table>
<thead>
<tr>
<th></th>
<th>HYDICE</th>
<th>AVIRIS</th>
<th>HYPERION</th>
<th>EnMAP</th>
<th>PRISMA</th>
<th>CHRIS</th>
<th>HyspIRI</th>
<th>IASI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude (Km)</td>
<td>1.6</td>
<td>20</td>
<td>705</td>
<td>653</td>
<td>614</td>
<td>556</td>
<td>626</td>
<td>817</td>
</tr>
</tbody>
</table>
| Spatial resolution (m) | 0.75           | 20              | 30              | 30              | 5-30            | 36              | 60              | V: 1-2 km  
H: 25 km        |
| Spectral resolution (nm) | 7-14           | 10              | 10              | 6.5-10          | 10              | 1.3-12          | 4-12            | 0.5 cm⁻¹        |
| Coverage (μm)    | 0.4-2.5         | 0.4-2.5         | 0.4-2.5         | 0.4-2.5         | 0.4-2.5         | 0.4-1.0         | 0.38-2.5        | 3.62-15.5 (645-2760 cm⁻¹) |
| Number of bands  | 210             | 224             | 220             | 228             | 238             | 63              | 217             | 8461             |
| Data cube size   | 200x320 x 210   | 512x614 X 224   | 660x256 x 220   | 1000x1000 x 228 | 400x880 X 238  | 748x748 X 63   | 620x512 X 210   | 765x120 X 8461  |
## Acquisition instruments

### Remote sensing

#### Low spatial resolution

<table>
<thead>
<tr>
<th>Instrument</th>
<th>HYDICE</th>
<th>AVIRIS</th>
<th>HYPERION</th>
<th>EnMAP</th>
<th>PRISMA</th>
<th>CHRIS</th>
<th>HyspIRI</th>
<th>IASI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude (Km)</td>
<td>1.6</td>
<td>20</td>
<td>705</td>
<td>663</td>
<td>614</td>
<td>556</td>
<td>626</td>
<td>817</td>
</tr>
<tr>
<td>Spatial resolution (m)</td>
<td>0.75</td>
<td>20</td>
<td>30</td>
<td>30</td>
<td>5-30</td>
<td>36</td>
<td>60</td>
<td>V: 1-2 km H: 25 km</td>
</tr>
<tr>
<td>Spectral resolution (nm)</td>
<td>7-14</td>
<td>10</td>
<td>10</td>
<td>6.5-10</td>
<td>10</td>
<td>1.3-12</td>
<td>4-12</td>
<td>0.5 cm⁻¹</td>
</tr>
<tr>
<td>Coverage (µm)</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-1.0</td>
<td>0.38-2.5 &amp; 7.5-12</td>
<td>3.62-15.5 (64-2760 cm⁻¹)</td>
</tr>
<tr>
<td>Number of bands</td>
<td>210</td>
<td>224</td>
<td>220</td>
<td>228</td>
<td>238</td>
<td>63</td>
<td>217</td>
<td>8461</td>
</tr>
<tr>
<td>Data cube size (samples x lines x bands)</td>
<td>200x320 X210</td>
<td>512x614 X224</td>
<td>660x256 X220</td>
<td>1000x1000 X228</td>
<td>400x880 X238</td>
<td>748x748 X63</td>
<td>620x512 X210</td>
<td>765x120 X8461</td>
</tr>
</tbody>
</table>
## Remote sensing

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude (Km)</td>
<td>1.6</td>
<td>20</td>
<td>705</td>
<td>613</td>
<td>614</td>
<td>556</td>
<td>626</td>
<td>817</td>
</tr>
<tr>
<td>Spatial resolution (m)</td>
<td>0.75</td>
<td>20</td>
<td>30</td>
<td>30</td>
<td>5-30</td>
<td>36</td>
<td>60</td>
<td>V: 1-2 km H: 25 km</td>
</tr>
<tr>
<td>Spectral resolution (nm)</td>
<td>7-14</td>
<td>10</td>
<td>10</td>
<td>6.5-10</td>
<td>10</td>
<td>1.3-12</td>
<td>4-12</td>
<td>0.5 cm⁻¹</td>
</tr>
<tr>
<td>Coverage (μm)</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-2.5</td>
<td>0.4-1.0</td>
<td>0.38-2.5 &amp; 7.5-12</td>
<td>3.62-15.5 (645-2760 cm⁻¹)</td>
</tr>
<tr>
<td>Number of bands</td>
<td>210</td>
<td>224</td>
<td>220</td>
<td>228</td>
<td>238</td>
<td>63</td>
<td>217</td>
<td>8461</td>
</tr>
<tr>
<td>Data cube size (samples x lines x bands)</td>
<td>200x320 X210</td>
<td>512x614 X224</td>
<td>660x256 X220</td>
<td>1000x1000 X228</td>
<td>400x880 X238</td>
<td>748x748 X63</td>
<td>620x512 X210</td>
<td>765x120 X8461</td>
</tr>
</tbody>
</table>

**Low spatial resolution**

Large data volumes

Large data volumes
Contributions to the radiance measured by the sensor

\[ L(\lambda) = a(\lambda)\rho(\lambda) + b(\lambda) \]

\(a\) and \(b\) are complex functions of: viewing angles, sun irradiance, atmosphere transmittance and reflectance, and surface reflectance.
Processing flow of hyperspectral data cubes

- Radiance data cube
  - Atmospheric correction
- Reflectance data cube
  - Dimensionality reduction (optional)
- Reduced data cube

Inverse problems (this tutorial)

- Denoising
- Data fusion
- Unmixing
Observation model in RS hyperspectral imaging

\( X \in \mathbb{R}^{R \times N} \) denotes a hyperspectral reflectance image organized in a matrix with \( R \) spectral bands and \( N \) pixels per band.

\[
X = \begin{bmatrix} x^1 \\ \vdots \\ x^R \end{bmatrix} (R \text{ band images})
\]

\[
X = [x_1, \ldots, x_N] (N \text{ spectral vectors})
\]

\[
x = \text{vec}(X) := [x_1^T, \ldots, x_N^T]^T \in \mathbb{R}^n, n = RN
\]

Linear observation model with additive noise

\[
y = Ax + n
\]

where \( y, n \in \mathbb{R}^m \), \( n \) is an additive perturbation, and the matrix \( A \in \mathbb{R}^{m \times n} \) accounts for the spectral and spatial sensor blurring and downsampling mechanisms.
Linear observation model

Often the action of $A$ is separable with respect to the columns and rows of $X$:

$$y = Ax + n \iff Y = A_\lambda XA_x + N$$

where

- $Y, N \in \mathbb{R}^{L \times M}$ and $y = \text{vec}(Y)$
- $A = A_T \otimes A_\lambda$ ($\otimes$ denotes kronecker product)
- $A_\lambda \in \mathbb{R}^{L \times R}$ acts on the rows (spectral domain) of $X$
- $A_x \in \mathbb{R}^{N \times M}$ acts on the columns (spatial domain) of $X$

Unknowns: $n = RN$

Observations: $m = LM$
Degradation mechanisms: noise

The noise is dominated by two components: \( y = \mathcal{P}(Ax) + n \)

1. **Non-additive Poissonian** noise due to the photon counting process (\( \mathcal{P}(Ax) \))
   - Recall: \( y \sim \mathcal{P}(x) \)
   - \( P(y = k) = \frac{e^{-x}x^k}{k!} \), \( \mathbb{E}[y] = x \), \( \mathbb{V}[y] = x \), \( \text{SNR} = \frac{\mathbb{E}^2[y]}{\mathbb{V}[y]} = x \)

2. **Additive Gaussian** noise due to electronic circuits (\( n \))
   - Accurate statistical modeling of the noise having into account the Gaussian and the Poissonian components is a challenging task ([B-D & Nascimento, 08], [Acito et al., 11], [Jezierska, 14], [Chouzenoux et al., 15])
   - Atmospheric correction process introduces further complications

In this tutorial, we often assume that the noise is Gaussian additive pixelwise independent with band-dependent variance
Example of Gaussian and Poissonian noise (ROSIS, band 60)

Gaussian Noise:
\[ y = x + n, \sigma = 0.03 \]

Poissonian Noise:
\[ y = \mathcal{P}(\gamma x), \gamma = 100 \]

Anscombe transform:
\[ \sqrt{y + a} \approx \sqrt{\gamma x + a + w} \]
Example: noise estimation

HySime: dim = 20
([B-D, & Nascimento, 08])
Characterization of the hyperspectral images

Hyperspectral data cubes are highly correlated in the spectral-spatial domain

⇒ Live in low (or in the union of low) dimensional manifolds or subspaces ([B-D & Nascimento, 08], [B-D et al. 12], [Ma et al., 14], [Heylen et al. 14])

\[ X = EZ, \quad E \in \mathbb{R}^{R \times p}, \quad p \ll R \]

⇒ Sparsely represented by 3D wavelets (multiresolution representations) ([Rasti et al., 12], [Fowler & Rucker, 07])

\[ w = Wx \in \mathbb{R}^d \quad \text{(wavelet coefficients)} \quad \|w\|_0 \ll d \quad \text{(}\|w\|_0 = \{ |w_i : w_i \neq 0| \}) \]

⇒ Exhibit self-similarity, thus suited to non-local dictionary based techniques ([Castrodad et al., 11], [Elad et al., 06])

\[ \text{Patch}(X) = D\alpha, \quad \alpha \text{ is parse} \]
Example of subspace identification [B-D & Nascimento, 08]

Pavia University (ROSIS, \(R = 103, N = 610 \times 340\))

\[
\text{SNR}_{i} = \frac{e_i^T R_n e_i}{e_i^T R_e e_i}
\]

SNR = 3 dB

dimension of the signal subspace = 10
Example of 3D wavelet decomposition

Pavia University
(ROSIS, $R = 103$, $N = 610 \times 340$)

Reconstruction from 3% of the 3D wavelet coefficients
PSNR = 35 dB
PSNR = 32 dB (1%)

Coefficients of the dual-tree 3D complex wavelets [Kingsbury, 02]
Inverse problems in hyperspectral imaging

Denoising

Observation model: \( Y = X + N \)
- \( X, N \in \mathbb{R}^{R \times N} \)
- \( N \) is Gaussian with matrix normal distribution: \( N \sim \mathcal{MN}(0_{R \times N}, C_\lambda, C_x) \) (this is equivalent to say that \( \text{vec}(N) \sim \mathcal{N}(0_{RN}, C_x \otimes C_\lambda) \))

Objective: estimate \( X \)

Unmixing (linear mixing model - LMM)

Observation model: \( Y = ES + N \)
- \( E \in \mathbb{R}^{p \times N} \) (endmember matrix)
- \( S \in \mathbb{R}^{p \times N} \) (abundance matrix)
- \( N \sim \mathcal{MN}(0_{R \times N}, C_\lambda, C_x) \)

Objective: estimate \( E, S \)
Matrix normal distribution

Let \( X \in \mathbb{R}^{R \times N} \). A matrix normal distribution

\[
    X \sim \mathcal{M}\mathcal{N}(M, C_\lambda, C_x)
\]

is a generalization of the multivariate normal distribution if and only if

\[
    \text{vec}(X) \sim \mathcal{N}(\text{vec}(M), C_x \otimes C_\lambda)
\]

This implies that

\[
    p(X|M, C_\lambda, C_x) = \frac{\exp \left( \frac{1}{2} \text{tr} \left[ C_x^{-1} (X - M)^T C_\lambda^{-1} (X - M) \right] \right)}{(2\pi)^{RN}|C_\lambda|^{R/2}|C_x|^{N/2}}
\]

- \( M := E[X] \)
- \( C_\lambda \) - among-row covariance
- \( C_x \) - among-column covariance
Inverse problems in hyperspectral imaging

Denoising

Observation model: \( Y = X + N \)

- \( X, N \in \mathbb{R}^{R \times N} \)
- \( N \) is Gaussian with matrix normal distribution: \( N \sim \mathcal{MN}(0_{R \times N}, C_\lambda, C_x) \) (this is equivalent to say that \( \text{vec}(N) \sim \mathcal{N}(0_{RN}, C_x \otimes C_\lambda) \))

Objective: estimate \( X \)

Unmixing (linear mixing model - LMM)

Observation model: \( Y = ES + N \)

- \( E \in \mathbb{R}^{p \times N} \) (endmember matrix)
- \( S \in \mathbb{R}^{p \times N} \) (abundance matrix)
- \( N \sim \mathcal{MN}(0_{R \times N}, C_\lambda, C_x) \)

Objective: estimate \( E, S \)
Inverse problems in hyperspectral imaging

Hyperspectral sharpening (deblurring, superresolution, fusion)

Observation model: \( Y_h = X A_x M + N_h \quad Y_m = A\lambda X + N_m \)

- \( X \in \mathbb{R}^{R \times N} \)
- \( Y_h \in \mathbb{R}^{R \times M} \) - observed hyperspectral image 
  \((M = N/d^2 \text{ - } d \text{ is the downsampling factor})\)
- \( Y_m \in \mathbb{R}^{L \times N} \) - observed multispectral image
- \( A_x \in \mathbb{R}^{N \times N} \) - (usually a convolution)
- \( A\lambda \in \mathbb{R}^{L \times R} \) - (spectral responses of the MS sensor)
- \( M \in \mathbb{R}^{N \times M} \) - (downsampling matrix)
- \( N_h \sim \mathcal{MN}(0_{R \times M}, C_{h\lambda}, C_{hx}) \)
- \( N_m \sim \mathcal{MN}(0_{L \times N}, C_{m\lambda}, C_{mx}) \)

Objective: estimate \( X \)
## Hyperspectral image compressive sensing

**Observation model:** \( y = Ax + n \)

- \( x \in \mathbb{R}^n, \, n = RN \)
- \( y \in \mathbb{R}^m, \, m \ll n \)
- \( A \in \mathbb{R}^{m \times n} - \text{measurement matrix (often } A = A_x^T \otimes A_\lambda) \)
- \( n \sim \mathcal{N}(0_m, C_n) \)

**Objective:** estimate \( x \) (equivalently, \( X = \text{vec}^{-1}(x) \))