Groupe hyperspectral

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

FCT, project UID/EEA/50008/2013

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)

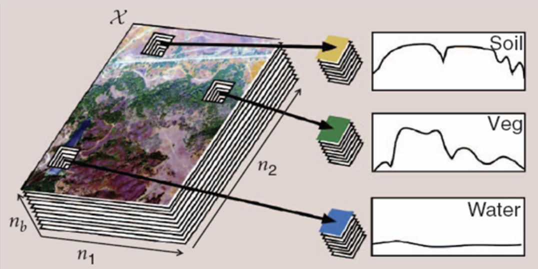
Part 2 - Inverse problems in a nutshell

Part 3 - Denoising, sharpening, and unmixing

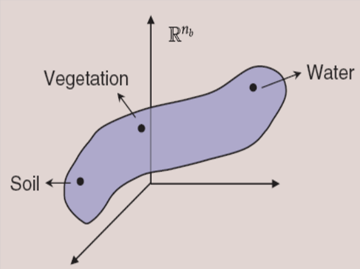
Hyperspectral imaging: motivation

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

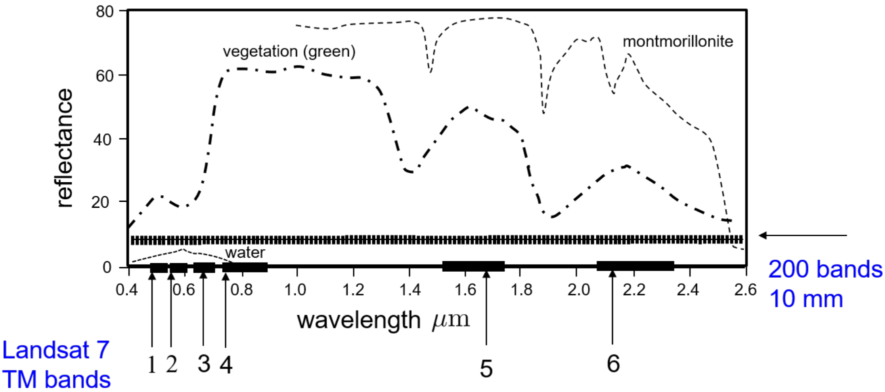
Hyperspectral imaging concept



Hyperspectral vectors in a low-dimensional manifold

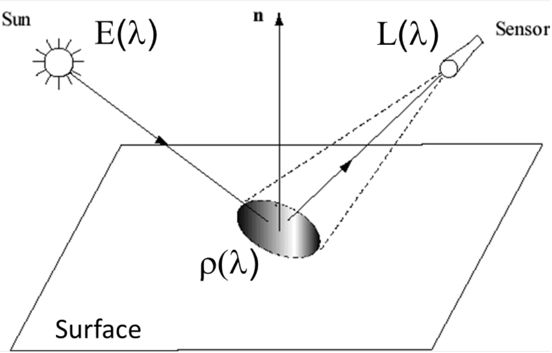


Hyperspectral imaging: motivation



Remote sensing: basics

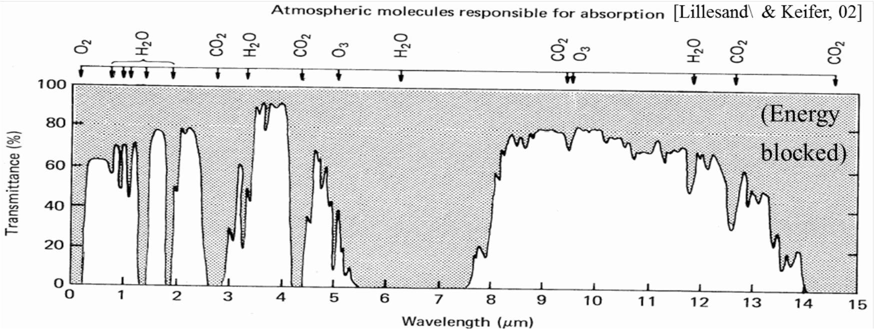
Radiance versus reflectance

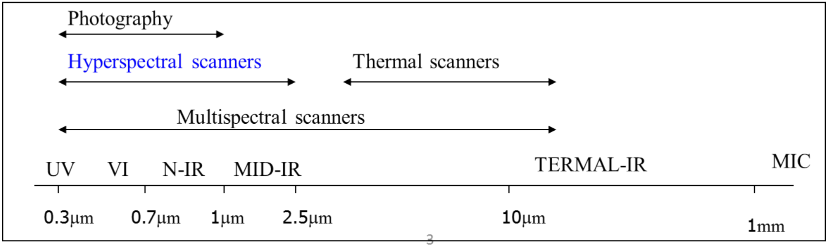


E – Irradiance (W/m^2)
ρ – Reflectance
L – Radiance (W/Sr/m^2)
λ – Wavelength (μm)

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

Remote sensing: the influence of atmosphere





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])

$$\mathsf{SNR} \propto \frac{\Delta^2}{\mathsf{ACR} * R}$$

where Δ 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

	HYDICE	AVIRIS	HYPERION	EnMAP	PRISMA	CHRIS	HyspIRI	IASI
Organization	NRL (USA , 1995)	NASA (USA , 1987)	USGS (USA, 2000)	GFC,OHB,DLR (GE, 2018)	ASI (I, 2017)	ESA (2001)	NASA (USA , 2022)	EUMETSAT 2006
Altitude (Km)	1.6	20	705	653	614	556	626	817
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 & 7.5-12	3.62-15.5 (645-2760 cm^{-1})
Number of bands	210	224	220	228	238	63	217	8461
Data cube size (samples x lines x bands)	200x320 X210	512x614 X224	660x256 X220	1000x1000 X228	400x880 X238	748x748 X63	620x512 X210	765x120 X8461

Acquisition instruments

Remote sensing

Low spatial resolution

	HYDICE	AVIRIS	HYPERION	EnMAP	PRISMA	CHRIS	HyspIRI	IASI
Organization	NRL (USA , 1995)	NASA (USA , 1987)	USGS (USA, 2000)	GFC,OHB,DLR (GE, 2018)	ASI (I, 2017)	ESA (2001)	NASA (USA , 2022)	EUMETSAT 2006
Altitude (Km)	1.6	20	705	653	614	556	626	817
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 & 7.5-12	3.62-15.5 (645-2760 cm^{-1})
Number of bands	210	224	220	228	238	63	217	8461
Data cube size (samples x lines x bands)	200x320 X210	512x614 X224	660x256 X220	1000x1000 X228	400x880 X238	748x748 X63	620x512 X210	765x120 X8461

Acquisition instruments

Remote sensing

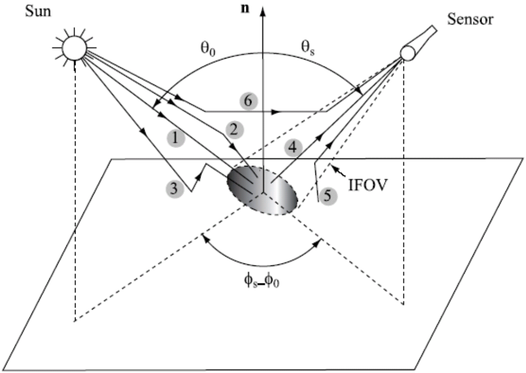
Low spatial resolution

	HYDICE	AVIRIS	HYPERION	EnMAP	PRISMA	CHRIS	HyspIRI	IASI
Organization	NRL (USA , 1995)	NASA (USA , 1987)	USGS (USA, 2000)	GFC,OHB,DLR (GE, 2018)	ASI (I, 2017)	ESA (2001)	NASA (USA , 2022)	EUMETSAT 2006
Altitude (Km)	1.6	20	705	653	614	556	626	817
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 & 7.5-12	3.62-15.5 (645-2760 cm^{-1})
Number of bands	210	224	220	228	238	63	217	8461
Data cube size (samples x lines x bands)	200x320 X210	512x614 X224	660x256 X220	1000x1000 X228	400x880 X238	748x748 X63	620x512 X210	765x120 X8461

Large data volumes

Large data volumes

Contributions to the radiance measured by the sensor



Rays

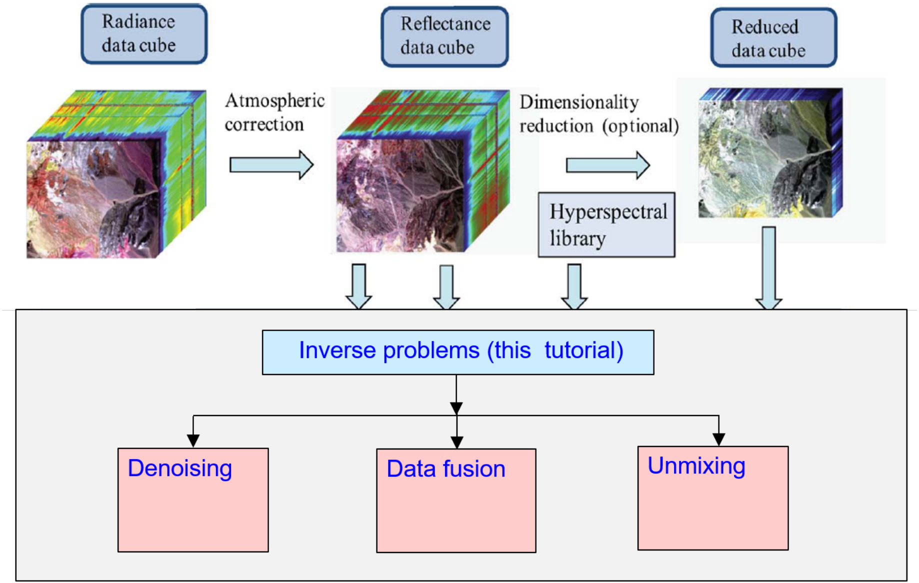
1. Sunlight
2. Skylight
3. Adjacency effect
4. Surface scattering
5. Atmosphere scattering
6. Path radiance

$$L(\lambda) = a(\lambda)\rho(\lambda) + b(\lambda)$$

---

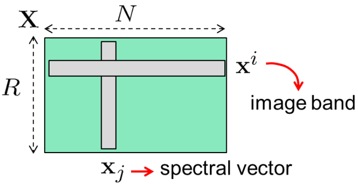
a and b are complex functions of: viewing angles, sun irradiance, atmosphere transmitance and reflectance, and surface refectance

Processing flow of hyperspectral data cubes



Observation model in RS hyperspectral imaging

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



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

$$\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N] \; (N \text{ spectral vectors})$$

$$\mathbf{x} = \mathrm{vec}(\mathbf{X}) := [\mathbf{x}_1^T, \ldots, \mathbf{x}_N^T]^T \in \mathbb{R}^n, \; n = RN$$

Linear observation model with additive noise

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

where $\mathbf{y}, \mathbf{n} \in \mathbb{R}^m$, $\mathbf{n}$ is an additive perturbation, and the matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ accounts for the spectral and spatial sensor blurring and downsampling mechanisms

Linear observation model

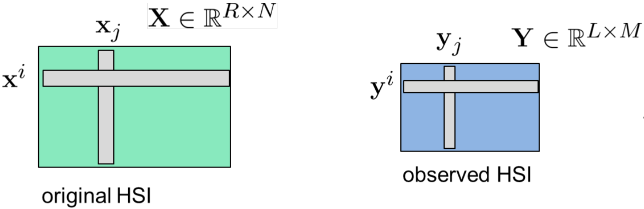
Often the action of $\mathbf{A}$ is separable with respect to the columns and rows of $\mathbf{X}$:

$$\boxed{\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \quad \Leftrightarrow \quad \mathbf{Y} = \mathbf{A}_\lambda \mathbf{X} \mathbf{A}_x + \mathbf{N}}$$

where

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

Unknowns:
$n = RN$



Observations:
$m = LM$

Degradation mechanisms: noise

The noise is dominated by two components: $\mathbf{y} = \mathcal{P}(\mathbf{A}\mathbf{x}) + \mathbf{n}$

1. Non-additive Poissonian noise due to the photon counting process ($\mathcal{P}(\mathbf{A}\mathbf{x})$)
   Recall: $y \sim \mathcal{P}(x)$

$$P(y=k) = \frac{e^{-x}x^k}{k!}, \quad \mathbb{E}[y] = x, \quad \mathbb{V}\mathrm{ar}[y] = x, \quad \mathrm{SNR} = \frac{\mathbb{E}^2[y]}{\mathbb{V}\mathrm{ar}[y]} = x$$

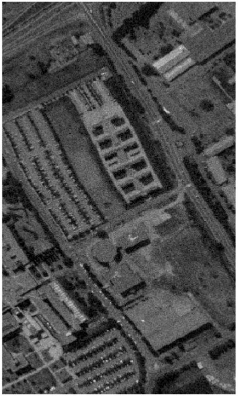
2. Additive Gaussian noise due to electronic circuits ($\mathbf{n}$)

- 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)

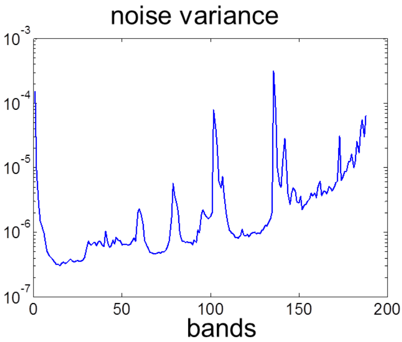


Gaussian Noise:
$\mathbf{y} = \mathbf{x} + \mathbf{n}$, $\sigma = 0.03$

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

Anscombe transform:
$\sqrt{\mathbf{y} + a} \simeq \sqrt{\gamma \mathbf{x} + a} + \mathbf{w}$

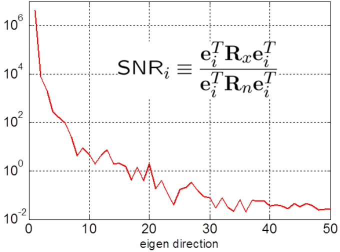
Example: noise estimation



HySime: dim = 20

([B-D, & Nascimento, 08])

SNR-SD



Characterization of the hyperspectral images

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

$\Rightarrow$ 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])

$$\mathbf{X} = \mathbf{E}\mathbf{Z} \quad \mathbf{E} \in \mathbb{R}^{R \times p} \quad p \ll R$$

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

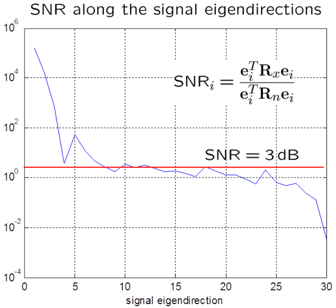
$$\mathbf{w} = \mathbf{W}\mathbf{x} \in \mathbb{R}^d \quad \text{(wavelet coefficients)} \quad \|\mathbf{w}\|_0 \ll d \qquad (\|\mathbf{w}\|_0 = \{|w_i : w_i \neq 0|\})$$

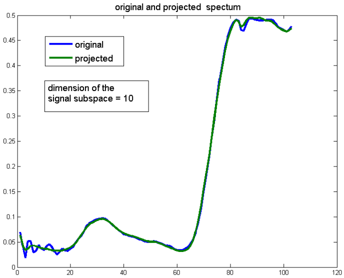
$\Rightarrow$ Exhibit self-similarity, thus suited to non-local dictionary based techniques
([Castrodad et al., 11], [Elad et al., 06])

$$\text{Patch}(\mathbf{X}) = \mathbf{D}\boldsymbol{\alpha} \quad \boldsymbol{\alpha} \quad \text{is parse}$$

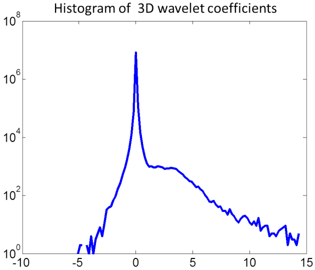
Example of subspace identification [B-D & Nascimento, 08]

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





Example of 3D wavelet decomposition



Coefficients of the dual-tree 3D complex wavelets [Kingsbury, 02])

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

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

Inverse problems in hyperspectral imaging

Denoising

Observation model: $\mathbf{Y} = \mathbf{X} + \mathbf{N}$

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

Objective: estimate $\mathbf{X}$

Unmixing (linear mixing model - LMM)

Observation model: $\mathbf{Y} = \mathbf{ES} + \mathbf{N}$

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

Objective: estimate $\mathbf{E}, \mathbf{S}$

Matrix normal distribution

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

$$\mathbf{X} \sim \mathcal{MN}(\mathbf{M}, \mathbf{C}_\lambda, \mathbf{C}_x)$$

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

$$\mathsf{vec}(\mathbf{X}) \sim \mathcal{N}(\mathsf{vec}(\mathbf{M}), \mathbf{C}_x \otimes \mathbf{C}_\lambda)$$

This implies that

$$p(\mathbf{X}|\mathbf{M}, \mathbf{C}_\lambda, \mathbf{C}_x) = \frac{\exp\left(\frac{1}{2}\mathsf{tr}\left[\mathbf{C}_x^{-1}(\mathbf{X}-\mathbf{M})^T\mathbf{C}_\lambda^{-1}(\mathbf{X}-\mathbf{M})\right]\right)}{(2\pi)^{RN}|\mathbf{C}_\lambda|^{R/2}|\mathbf{C}_x|^{N/2}}$$

- $\mathbf{M} := \mathbb{E}[\mathbf{X}]$
- $\mathbf{C}_\lambda$ - among-row covariance
- $\mathbf{C}_x$ - among-column covariance

Inverse problems in hyperspectral imaging

Denoising

Observation model: $\mathbf{Y} = \mathbf{X} + \mathbf{N}$

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

Objective: estimate $\mathbf{X}$

Unmixing (linear mixing model - LMM)

Observation model: $\mathbf{Y} = \mathbf{ES} + \mathbf{N}$

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

Objective: estimate $\mathbf{E}, \mathbf{S}$

Inverse problems in hyperspectral imaging

Hyperspectral sharpening (deblurring, superresolution, fusion)

Observation model: $\mathbf{Y}_h = \mathbf{X}\mathbf{A}_x\mathbf{M} + \mathbf{N}_h \qquad \mathbf{Y}_m = \mathbf{A}_\lambda\mathbf{X} + \mathbf{N}_m$

- $\mathbf{X} \in \mathbb{R}^{R\times N}$
- $\mathbf{Y}_h \in \mathbb{R}^{R\times M}$ - observed hyperspectral image
  ($M = N/d^2$ - d is the downsampling factor)
- $\mathbf{Y}_m \in \mathbb{R}^{L\times N}$ - observed multispectral image
- $\mathbf{A}_x \in \mathbb{R}^{N\times N}$ - (usually a convolution)
- $\mathbf{A}_\lambda \in \mathbb{R}^{L\times R}$ - (spectral responses of the MS sensor)
- $\mathbf{M} \in \mathbb{R}^{N\times M}$ - (downsamplimg matrix)
- $\mathbf{N}_h \sim \mathcal{MN}(\mathbf{0}_{R\times M}, \mathbf{C}_{h\lambda}, \mathbf{C}_{hx})$
- $\mathbf{N}_m \sim \mathcal{MN}(\mathbf{0}_{L\times N}, \mathbf{C}_{m\lambda}, \mathbf{C}_{mx})$

Objective: estimate $\mathbf{X}$

Inverse problems in hyperspectral imaging

Hyperspectral image compressive sensing

Observation model: $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$

- $\mathbf{x} \in \mathbb{R}^n$, $n = RN$
- $\mathbf{y} \in \mathbb{R}^m$, $m \ll n$
- $\mathbf{A} \in \mathbb{R}^{m \times n}$ - measurement matrix (often $\mathbf{A} = \mathbf{A}_x^T \otimes \mathbf{A}_\lambda$)
- $\mathbf{n} \sim \mathcal{N}(\mathbf{0}_m, \mathbf{C}_n)$

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