Groupe hyperspectral

Part 3.1 - Hyperspectral denoising

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 3 - Inverse problems in hyperspectral imaging

- Denoising
- Hyperspectral sharpening
- Hyperspectral unmixing

Denoising

Observation model: $\mathbf{y} = \mathbf{x} + \mathbf{n}$ (or $\mathbf{Y} = \mathbf{X} + \mathbf{N}$)

- $\mathbf{x}, \mathbf{n} \in \mathbb{R}^n$, $n = RN$
- $\mathbf{N} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{C}_x \otimes \mathbf{C}_\lambda)$

Denoising is arguably the simplest inverse problem. But is us a fundamental one.

Any model for $\mathbf{X}$ (prior, regularizer, constraints) that works well in a denoising problem is very likely to work well in other applications

Relevant approaches to hyperspectral denoising

- 3D wavelet-based [Rasti et al., 12, 13]
- non-local patch-based methods [Maggioni et al., 12]
- ST-TV and ST-NLTV regularization [Bresson & Chan, 08], [Yuan et al., 12], [Cheng et al., 14], [Chiercia et al., 15]
- tensor decompositon [Karami et al. 11]
- low rank and self-similarity [Zhuang & B-D, 16]

3D Wavelet-based denoising

Let $\mathbf{W} \in \mathbb{R}^{d \times n}$ be a wavelet transform and $\mathbf{w} := \mathbf{W}\mathbf{x}$ (the wavelet coefficients)

Fundamental property: 3D wavelet coefficients of HSIs are sparse or compressible

- $\mathbf{w}$ is sparse means that $\|\mathbf{w}\|_0 \ll d \qquad (\|\mathbf{w}\|_0 = \{|w_i : w_i \neq 0|\})$
- $\mathbf{w}$ is compressible means that its elements have a fast decaying tails

Convex variational formulation to denoising

$$\min_{\mathbf{x}} (1/2)\|\mathbf{y} - \mathbf{x}\|_Q^2 + \tau\|\mathbf{W}\mathbf{x}\|_1 \tag{1}$$

where $\|\mathbf{x}\|_Q^2 := \mathbf{x}^T\mathbf{Q}\mathbf{x}$ is a weighted ℓ_2-norm (usually, $\mathbf{Q} = \mathbf{C}_x^{1/2} \otimes \mathbf{C}_\lambda^{1/2}$)

The ℓ_1 norm, jointly with the quadratic data fidelity term, promotes sparsity on the vector $\mathbf{w} = \mathbf{W}\mathbf{x}$

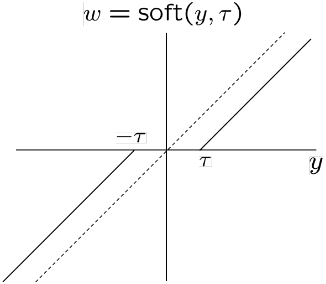
3D Wavelet-based denoising

If $\mathbf{Q} = \mathbf{I}$ and $\mathbf{W}$ is orthogonal ($\mathbf{W}\mathbf{W}^T = \mathbf{W}^T\mathbf{W} = \mathbf{I}$), then

$$\min_{\mathbf{x}} (1/2)\|\mathbf{y} - \mathbf{x}\|_Q^2 + \tau\|\mathbf{W}\mathbf{x}\|_1 \quad \Leftrightarrow \quad \min_{\mathbf{w}} \|\widetilde{\mathbf{y}} - \mathbf{w}\|_Q^2 + \tau\|\mathbf{w}\|_1$$

where $\widetilde{\mathbf{y}} := \mathbf{W}\mathbf{y}$ and

$$\widehat{\mathbf{x}} = \mathbf{W}^T\widehat{\mathbf{w}} \qquad \widehat{\mathbf{w}} = \text{soft}(\widetilde{\mathbf{y}}, \tau)$$

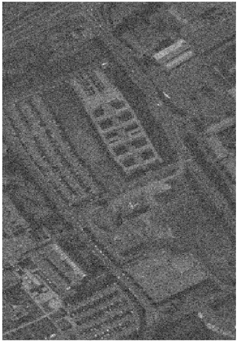


Insight: $(\widetilde{y}_i = \widetilde{x}_i + \widetilde{n}_i)$

- $\widetilde{x}_i$ is heavy tailed
- $\widetilde{n}_i \sim \mathcal{N}(0, \sigma_n^2)$
- $\widetilde{w}_i = 0$ if $\widetilde{y}_i$ is dominated by the noise
- $\widetilde{w}_i = \widetilde{y}_i \pm \tau$ if $\widetilde{y}_i$ is dominated by the signal

Example: 3D Wavelet-based denoising (3D-DWT)

Example with 3D-DWT $N = 640 \times 340, R = 103,\ \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$



Original HSI

Noisy observation
$\text{PSNR}_y = 20\,\text{dB}$

Estimated band 60
$\text{PSNR}_{\widehat{x}} = 31\,\text{dB}$

Wavelet-based denoising with orvercomplete representations

If $d > n$ ($\mathbf{W}^T$ is overcomplete) or $\mathbf{Q} \neq \alpha \mathbf{I}$, then

$$\min_{\mathbf{x}} (1/2)\|\mathbf{y} - \mathbf{x}\|_Q^2 + \tau \|\mathbf{W}\mathbf{x}\|_1 \quad \not\Leftrightarrow \quad \min_{\mathbf{w}} \|\widetilde{\mathbf{y}} - \mathbf{w}\|_Q^2 + \tau \|\mathbf{w}\|_1$$

Analysis formulation

$$\min_{\mathbf{x}} (1/2)\|\mathbf{y} - \mathbf{x}\|_Q^2 + \tau \|\mathbf{W}\mathbf{x}\|_1$$

Synthesis formulation ($\mathbf{W}$ is a Parseval frame: $\mathbf{W}^T\mathbf{W} = \mathbf{I}$)

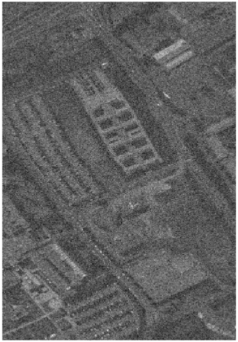
$$\min_{\mathbf{w}} (1/2)\|\mathbf{y} - \mathbf{W}^T\mathbf{w}\|_Q^2 + \tau \|\mathbf{w}\|_1, \quad \text{s.t.:} \quad \mathbf{W}\mathbf{W}^T\mathbf{w} = \mathbf{w}$$

Both, the analysis and the synthesis optimizations are easily solved with SALSA and FBPD algorithms.

Example: 3D Wavelet-based denoising (3D-DT-COMP)

Example with 3D-DT-COMP $N = 640 \times 340, R = 103,\ \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$

Original HSI



Noisy observation
$\mathrm{PSNR}_y = 20\,\mathrm{dB}$

Estimated band 60
$\mathrm{PSNR}_{\widehat{x}} = 33\,\mathrm{dB}$

Non-local patch(cube)-based methods

Real world images are self-similar: given an image patch (cubes in volume), there are similar patches at different locations and scales.

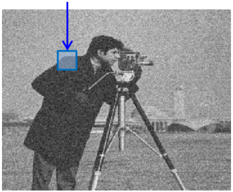
Self-similarity has been mainly exploited in two ways:

- Non-local (generalized) mean: for each patch find similar ones in the image and produce a patch estimate based on the found patches ([Buades et al., 2005], [Dabov et al., 07], [Maggioni et al., 12])
- Dictionary learning: express each patch as sparse representation in a given dictionary, which may be learned from the data ([Elad & Aharon, 05] [Mairal et. al., 08,10])

Patch-based image holds the state-of-the-art in image denoising ([Dabov et al., 07], [Maggioni et al., 12], [Chatterjee, P. Milanfar, 12])

Non-local patch(cube)-based methods

$\mathbf{y}_i$(noisy patch)

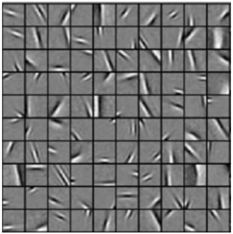


Denoising algorithm

1. Dictionary learning: estimate $\mathbf{D} := [\mathbf{d}_1, \ldots, \mathbf{d}_K]$ from the (overlapped) patches $\mathbf{x}_i$, $i = 1, \ldots$ by solving the matrix factorization problem

$$\min_{\mathbf{D}, \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{N_p}} \sum_{i=1}^{N_p} \|\mathbf{y}_i - \mathbf{D}\boldsymbol{\alpha}_i\|_2^2 + \lambda \|\boldsymbol{\alpha}_i\|_1$$

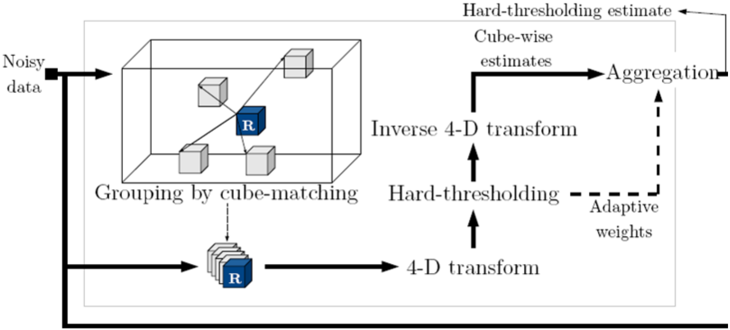
$\mathbf{D}$ dictionary



2. Patch composition: compute an estimate of the image by averaging the estimated patches $\widehat{\mathbf{x}}_i = \mathbf{D}\boldsymbol{\alpha}_i$

Block matching 4D (BM4D) ([Maggioni et al., 12])

BM4D is an extension to 3D images of BM3D ([Dabov et al., 07])

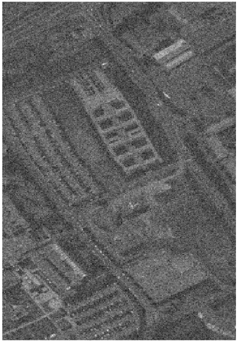


From ([Maggioni et al., 12])

Example: Cube-based denoising with BM4D

Example with BM4D $N = 640 \times 340, R = 103,\ \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$

Original HSI



Noisy observation
PSNR$_y = 20\,\mathrm{dB}$

Estimated band 60
PSNR$_{\widehat{x}} = 36\,\mathrm{dB}$ (600 sec)

Low rank + self-similarity ([Zhuang & B-D, 16])

Low rank: $\mathbf{X} = \mathbf{E}\mathbf{Z}$ $\quad$ $\mathbf{E} \in \mathbb{R}^{R\times p}$ holds an orthonormal basis for range$(\mathbf{X})$

$$\begin{aligned}\widehat{\mathbf{Z}} &= \arg\min_{\mathbf{Z}} \frac{1}{2}\|\mathbf{E}\mathbf{Z} - \mathbf{Y}\|_F^2 + \lambda\phi(\mathbf{Z}) \\ &= \arg\min_{\mathbf{Z}} \frac{1}{2}\|\mathbf{Z} - \mathbf{E}^T\mathbf{Y}\|_F^2 + \lambda\phi(\mathbf{Z}),\end{aligned}$$

Regularizer ϕ is decoupled

$$\phi(\mathbf{Z}) = \sum_{i=1}^{k} \phi_i(\mathbf{Z}^i)$$

Solution:

$$\widehat{\mathbf{Z}} = \psi_{\lambda\phi}(\mathbf{E}^T\mathbf{Y}) = \begin{bmatrix} \psi_{\lambda\phi_1}(\mathbf{e}_1^T\mathbf{Y}) \\ \vdots \\ \psi_{\lambda\phi_k}(\mathbf{e}_k^T\mathbf{Y}) \end{bmatrix}$$

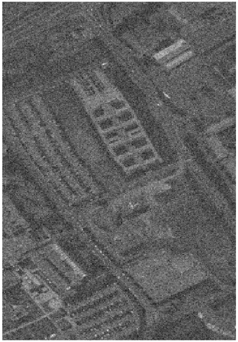
where

$$\psi_{\lambda\phi_i} = \arg\min_{\mathbf{w}} \frac{1}{2}\|\mathbf{y} - \mathbf{w}\|_F^2 + \lambda\phi_i(\mathbf{w})$$

is the so-called denoising operator, or Moreau proximity operator (MPO) of ϕ

Example: Low rank + self-similarity ([Zhuang & B-D, 16])

MPO for ϕ: BM3D $N = 640 \times 340, R = 103,\ \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$



Original HSI

Noisy observation
$\text{PSNR}_y = 20\,\text{dB}$

Estimated band 60
$\text{PSNR}_{\widehat{x}} = 39\,\text{dB}$ (8 sec)

Example: Low rank + self-similarity ([Zhuang & B-D, 16])

Denising + inpainting

MPO for ϕ: BM3D $N = 640 \times 340, R = 103,\ \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$

Original HSI

Noisy observation
PSNR$_y = 20\,$dB

Estimated band 60
PSNR$_{\widehat{x}} = 39\,$dB (8 sec)

Example: Low rank + self-similarity ([Zhuang & B-D, 16])

MPO for ϕ: BM3D $N = 640 \times 340, R = 103$, Poissonian noise $\mathbf{y} = \mathcal{P}(\gamma \mathbf{x})$

Original HSI

Noisy observation

Estimated band 60

$\text{PSNR}_{\widehat{x}} = 47\,\text{dB}$ (11 sec)

Structured tensor TV-based regularization

ST-TV (Vector TV, Hyperspectral TV)
([Bressom & Chan, 2008], [Yuan et al., 2012])

$$\min_{\mathbf{X}} (1/2)\|\mathbf{Y} - \mathbf{X}\|_F^2 + \lambda\phi_{\textsf{ST-TV}}(\mathbf{X})$$

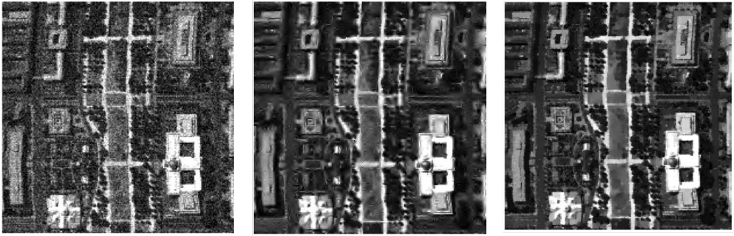
where

$$\phi_{\textsf{ST-TV}}(\mathbf{X}) = \sum_{i=1}^{N} \tau_i \|\mathbf{x}_i - \mathbf{x}_{h_i}, \mathbf{x}_i - \mathbf{x}_{v_i}\|_F$$

- ST-TV promotes localized step gradients within the image bands and align the "discontinuities" across the bands
- by controlling the amount of spatial regularization, parameters τ_i mitigates the well known undesirable staircasing effects associated to TV regularization
- ST-TV optimization problem is easily solved by SALSA and FBPD algorithms

Example: ST-TV denoising ([Yuan et al., TGRS 12])

HYDICE band 108 of Washington DC Mall



HYDICE band of Urban data sets

Left: Noisy band; Middle Wavelet; Right: ST-TV (SSAHTV))

Structured tensor NLTV-based regularization

ST-NLTV denoising and inpainting

$$\min_{\mathbf{X}} (1/2)\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_Q^2 + \lambda\phi_{\mathsf{ST\text{-}NLTV}_p}(\mathbf{X})$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a diaginal operator (mask), $\mathbf{x} = \mathrm{vec}(\mathbf{X})$, and

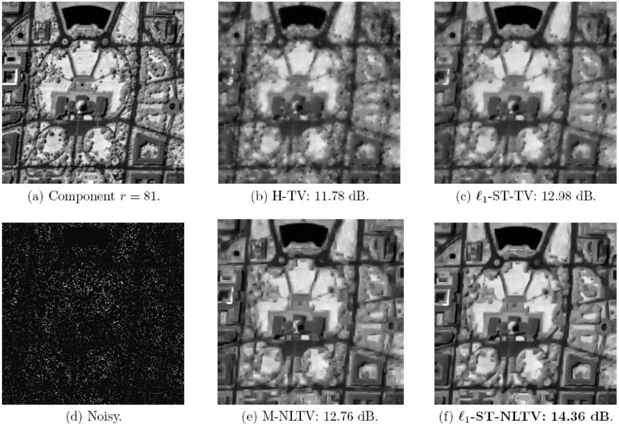
$$\phi_{\mathsf{ST\text{-}NLTV}_p}(\mathbf{X}) = \sum_{i=1}^{N} \tau_i \big\|[(\omega_{i,j}(\mathbf{x}_i - \mathbf{x}_j), j \in \mathcal{N}_i]\big\|_{\mathcal{S}_p}$$

is the ST-NLTV regularizer ([Chiercia et. al., 15], multichannel-NLTV for $p = 2$) [Cheng, et al. 14]

The weights $\omega_{i,j}$, learned from the observed data, measure the similarity between pixels i and j

- ST-NLTV improves over ST-TV regarding the preservation of textures, details, and fine structures
- ST-NLTV optimization problem is solved by SALSA and FBPD algorithms

HYDICE band 81 (Urban area)



Left: Inpainting example from [Chiercia et al., 12]. Degradation: additive zero-mean white Gaussian noise with $\sigma = 5$ and 90% of decimation ($N = 65536$, $R = 191$, $M = 6553$ and $L = 191$)