2010 IEEE International Geoscience and Remote Sensing Symposium
Honolulu, Hawaï, USA, July 2010

Polarimetric SAR estimation based on non-local means

Charles Deledalle[1], Florence Tupin[1], Loïc Denis[2]

[1] Institut Telecom, Telecom ParisTech, CNRS LTCI, Paris, France
[2] Observatoire de Lyon, CNRS CRAL, UCBL, ENS de Lyon, Université de Lyon, Lyon, France

July 28, 2010

Why non-local methods ?

(a) Noisy image

(b) Non-local means

Noise reduction + resolution preservation

Motivation

Why non-local methods ?

(a) Noisy image

(b) Non-local means

Noise reduction + resolution preservation

- Goal: to adapt non-local methods to PolSAR data,
- Method: take into account the statistics of PolSAR data,

Table of contents

Table of contents

Non-local means (NL means) (1/2)

Background

- Local filters (e.g. boxcar, Gaussian filter...) $\Rightarrow$ resolution loss,
- To combine similar pixels instead of neighboring pixels.

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{|s-t|^2}{\rho^2}} v_t$$

Gaussian filter

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{\mathbf{sim}(s,t)}{h^2}} v_t$$

[Yaroslavsky, 1985]

Non-local means (NL means) (1/2)

Background

- Local filters (e.g. boxcar, Gaussian filter...) $\Rightarrow$ resolution loss,
- To combine similar pixels instead of neighboring pixels.

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{|s-t|^2}{\rho^2}} v_t$$

Gaussian filter

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{\mathbf{sim}(s,t)}{h^2}} v_t$$

[Yaroslavsky, 1985]

Non-local means [Buades et al., 2005]

- Similarity evaluated using square patches Δ_s and Δ_t centered on s and t,
- Consider the redundant structure of images.

Non-local means (NL means) (1/2)

Background

- Local filters (e.g. boxcar, Gaussian filter...) $\Rightarrow$ resolution loss,
- To combine similar pixels instead of neighboring pixels.

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{|s-t|^2}{\rho^2}} v_t$$

Gaussian filter

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{\mathbf{sim}(s,t)}{h^2}} v_t$$

[Yaroslavsky, 1985]

Non-local means [Buades et al., 2005]

- Similarity evaluated using square patches Δ_s and Δ_t centered on s and t,
- Consider the redundant structure of images.

- Hyp. 1: similar neighborhood $\Rightarrow$ same central pixel,
- Hyp. 2: each patch is redundant (can be found many times).

Non-local means (NL means) (1/2)

Background

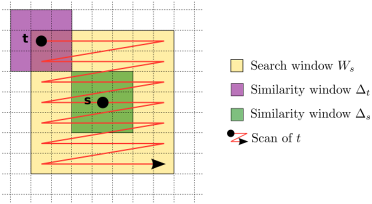
- Local filters (e.g. boxcar, Gaussian filter...) $\Rightarrow$ resolution loss,
- To combine similar pixels instead of neighboring pixels.

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{|s-t|^2}{\rho^2}} v_t$$

Gaussian filter

$$\hat{u}_s = \frac{1}{Z} \sum_t e^{-\frac{\mathbf{sim}(s,t)}{h^2}} v_t$$

[Yaroslavsky, 1985]



Algorithm of NL means

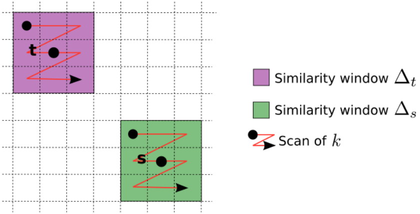
Non-local means (NL means) (2/2)

Similarity criterion [Buades et al., 2005]

- Euclidean distance:

$$\mathbf{sim}(s,t) = \sum_k |v_{s,k} - v_{t,k}|^2$$

with k the k-th respective pixel of Δ_s and Δ_t.



Euclidean distance: comparison pixel by pixel

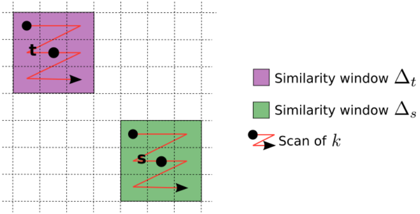
Non-local means (NL means) (2/2)

Similarity criterion [Buades et al., 2005]

- Euclidean distance:

$$\mathbf{sim}(s,t) = \sum_k |v_{s,k} - v_{t,k}|^2$$

with k the k-th respective pixel of Δ_s and Δ_t.



Euclidean distance: comparison pixel by pixel

- Hyp. 1: similar neighborhood ⇒ same central pixel,
- Hyp. 2: each patch is redundant (can be found many times),
- Hyp. 3: the noise is additive, white and Gaussian.

Table of contents

Weighted maximum likelihood for PolSAR data

Context and notations

- Searched data: $\mathbf{T}_s$ an $N \times N$ complex matrix,
- Noise model: $p(.|\mathbf{T}_s)$ a circular complex Gaussian,
- Noisy observations: $\mathbf{k}_s$ a N dimensional vector such that $\mathbf{k}_s \sim p(.|\mathbf{T}_s)$,
- To denoise: to search an estimate $\hat{\mathbf{T}}_s$ of $\mathbf{T}_s$.

Weighted maximum likelihood for PolSAR data

Context and notations

- Searched data: $\mathbf{T}_s$ an $N \times N$ complex matrix,
- Noise model: $p(.|\mathbf{T}_s)$ a circular complex Gaussian,
- Noisy observations: $\mathbf{k}_s$ a N dimensional vector such that $\mathbf{k}_s \sim p(.|\mathbf{T}_s)$,
- To denoise: to search an estimate $\hat{\mathbf{T}}_s$ of $\mathbf{T}_s$.

From weighted average to weighted maximum likelihood

- NL means perform a weighted average of noisy pixel values,
- For PolSAR data we suggest to use a weighted maximum likelihood:

$$\hat{\mathbf{T}}_s = \arg\max_{\mathbf{T}} \sum_t w(s,t) \log p(\mathbf{k}_t|\mathbf{T}) = \frac{\sum w(s,t)\mathbf{k}_t\mathbf{k}_t^\dagger}{\sum w(s,t)}$$

  where $w(s,t)$ is a weight approaching the indicator function of the set of redundant pixels (i.e with i.i.d values): $\{s,t|\mathbf{T}_s = \mathbf{T}_t\}$.
- Choice of $w(s,t)$: oriented, adaptive or non-local neighborhoods.

Context and notations

- Searched data: $\mathbf{T}_s$ an $N \times N$ complex matrix,
- Noise model: $p(.|\mathbf{T}_s)$ a circular complex Gaussian,
- Noisy observations: $\mathbf{k}_s$ a N dimensional vector such that $\mathbf{k}_s \sim p(.|\mathbf{T}_s)$,
- To denoise: to search an estimate $\hat{\mathbf{T}}_s$ of $\mathbf{T}_s$.

Example (Refined Lee - oriented neighborhood [Lee, 1981])

- Redundant pixels are located in one of these eight neighborhoods:

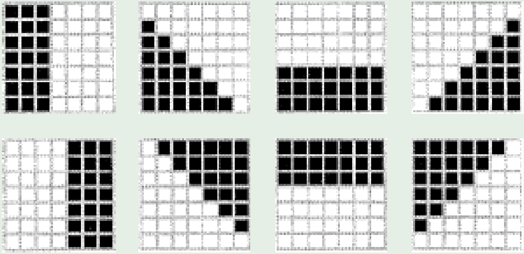


image extracted from [Lee, 1981]

Weighted maximum likelihood for PolSAR data

Context and notations

- Searched data: $\mathbf{T}_s$ an $N \times N$ complex matrix,
- Noise model: $p(.|\mathbf{T}_s)$ a circular complex Gaussian,
- Noisy observations: $\mathbf{k}_s$ a N dimensional vector such that $\mathbf{k}_s \sim p(.|\mathbf{T}_s)$,
- To denoise: to search an estimate $\hat{\mathbf{T}}_s$ of $\mathbf{T}_s$.

Example (IDAN - adaptive neighborhood [Vasile et al., 2006])

- Redundant pixels are located in an adaptive neighborhood (obtained by region growing algorithm):

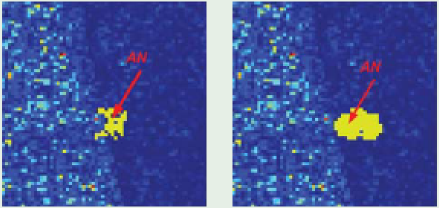


image extracted from [Vasile et al., 2006]

Weighted maximum likelihood for PolSAR data

Context and notations

- Searched data: $\mathbf{T}_s$ an $N \times N$ complex matrix,
- Noise model: $p(.|\mathbf{T}_s)$ a circular complex Gaussian,
- Noisy observations: $\mathbf{k}_s$ a N dimensional vector such that $\mathbf{k}_s \sim p(.|\mathbf{T}_s)$,
- To denoise: to search an estimate $\hat{\mathbf{T}}_s$ of $\mathbf{T}_s$.

Example (NL means - non-local neighborhood)

- Redundant pixels are located anywhere in the image:

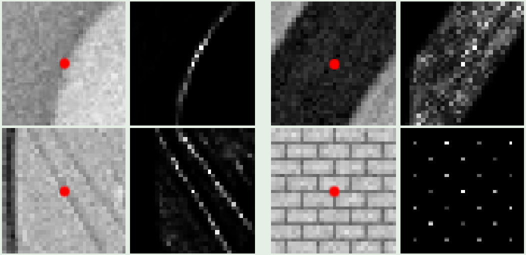


image extracted from [Buades et al., 2005]

Setting of the weights for PolSAR data (1/3)

- We search weights $w(s, t)$ such that:
  - $w(s, t)$ is high if $\mathbf{T}_s = \mathbf{T}_t$,
  - $w(s, t)$ is low if $\mathbf{T}_s \neq \mathbf{T}_t$,

Setting of the weights for PolSAR data (1/3)

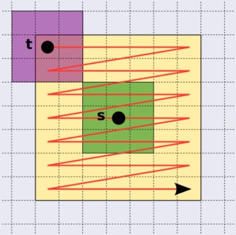
- We search weights $w(s,t)$ such that:
  - $w(s,t)$ is high if $\mathbf{T}_s = \mathbf{T}_t$,
  - $w(s,t)$ is low if $\mathbf{T}_s \neq \mathbf{T}_t$,

Statistical similarity between noisy patches [Deledalle et al., 2009]

- Using the hypothesis of NL means:

  Patches Δ_s and Δ_t similar $\Rightarrow$ central values s and t close.

- Weights are defined as follows:



$$w(s,t) = p(\mathbf{T}_{\Delta_{s,k}} = \mathbf{T}_{\Delta_{t,k}} | \mathbf{k}_{s,k}, \mathbf{k}_{t,k})^{1/h}$$
$$= \prod_k p(\mathbf{T}_{s,k} = \mathbf{T}_{t,k} | \mathbf{k}_{s,k}, \mathbf{k}_{t,k})^{1/h}$$

with $p(\mathbf{T}_{s,k} = \mathbf{T}_{t,k} | \mathbf{k}_{s,k}, \mathbf{k}_{t,k})$ statistical similarity

h regularization parameter.

Setting of the weights for PolSAR data (2/3)

Bayesian decomposition

$$p(\mathbf{T}_1 = \mathbf{T}_2 | \mathbf{k}_1, \mathbf{k}_2) \propto \underbrace{p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2)}_{\text{similarity likelihood}} \times \underbrace{p(\mathbf{T}_1 = \mathbf{T}_2)}_{\textit{a priori}\ \text{similarity}}$$

Setting of the weights for PolSAR data (2/3)

Bayesian decomposition

$$p(\mathbf{T}_1 = \mathbf{T}_2 | \mathbf{k}_1, \mathbf{k}_2) \propto \underbrace{\boxed{p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2)}}_{\text{similarity likelihood}} \times \underbrace{p(\mathbf{T}_1 = \mathbf{T}_2)}_{\textit{a priori} \text{ similarity}}$$

Similarity likelihood

$$p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2) = \frac{\int p(\mathbf{k}_1 | \mathbf{T}_1 = t) p(\mathbf{k}_2 | \mathbf{T}_2 = t) p(\mathbf{T}_1 = t) p(\mathbf{T}_2 = t) \, dt}{\int p(\mathbf{T}_1 = t) p(\mathbf{T}_2 = t) \, dt}$$

- Since $p(\mathbf{T}_1 = t)$ is unknown, we define:

$$p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2) \triangleq p(\mathbf{k}_1 | \mathbf{T}_1 = \hat{\mathbf{T}}_{MV}) p(\mathbf{k}_2 | \mathbf{T}_2 = \hat{\mathbf{T}}_{MV})$$

Setting of the weights for PolSAR data (2/3)

Bayesian decomposition

$$p(\mathbf{T}_1 = \mathbf{T}_2 | \mathbf{k}_1, \mathbf{k}_2) \propto \underbrace{\boxed{p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2)}}_{\text{similarity likelihood}} \times \underbrace{p(\mathbf{T}_1 = \mathbf{T}_2)}_{\textit{a priori} \text{ similarity}}$$

Similarity likelihood

$$p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2) = \frac{\int p(\mathbf{k}_1 | \mathbf{T}_1 = t) p(\mathbf{k}_2 | \mathbf{T}_2 = t) p(\mathbf{T}_1 = t) p(\mathbf{T}_2 = t) \, dt}{\int p(\mathbf{T}_1 = t) p(\mathbf{T}_2 = t) \, dt}$$

- Since $p(\mathbf{T}_1 = t)$ is unknown, we define:

$$p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2) \triangleq p(\mathbf{k}_1 | \mathbf{T}_1 = \hat{\mathbf{T}}_{MV}) p(\mathbf{k}_2 | \mathbf{T}_2 = \hat{\mathbf{T}}_{MV})$$

- Since only two observations are available to estimate $\hat{\mathbf{T}}_{MV}$, the matrix is singular.
- By enforcing $\hat{\mathbf{T}}_{MV}$ to be diagonal, the following similarity criterion is obtained:

$$\log \left(\frac{|\mathbf{k}_{1,1}|}{|\mathbf{k}_{2,1}|} + \frac{|\mathbf{k}_{2,1}|}{|\mathbf{k}_{1,1}|} \right) + \log \left(\frac{|\mathbf{k}_{1,2}|}{|\mathbf{k}_{2,2}|} + \frac{|\mathbf{k}_{2,2}|}{|\mathbf{k}_{1,2}|} \right) + \log \left(\frac{|\mathbf{k}_{1,3}|}{|\mathbf{k}_{2,3}|} + \frac{|\mathbf{k}_{2,3}|}{|\mathbf{k}_{1,3}|} \right).$$

Setting of the weights for PolSAR data (2/3)

Bayesian decomposition

$$p(\mathbf{T}_1 = \mathbf{T}_2 | \mathbf{k}_1, \mathbf{k}_2) \propto \underbrace{\boxed{p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2)}}_{\text{similarity likelihood}} \times \underbrace{p(\mathbf{T}_1 = \mathbf{T}_2)}_{\textit{a priori} \text{ similarity}}$$

Similarity likelihood

$$p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2) = \frac{\int p(\mathbf{k}_1 | \mathbf{T}_1 = t) p(\mathbf{k}_2 | \mathbf{T}_2 = t) p(\mathbf{T}_1 = t) p(\mathbf{T}_2 = t) \, dt}{\int p(\mathbf{T}_1 = t) p(\mathbf{T}_2 = t) \, dt}$$

- Since $p(\mathbf{T}_1 = t)$ is unknown, we define:

$$p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2) \triangleq p(\mathbf{k}_1 | \mathbf{T}_1 = \hat{\mathbf{T}}_{MV}) p(\mathbf{k}_2 | \mathbf{T}_2 = \hat{\mathbf{T}}_{MV})$$

- Since only two observations are available to estimate $\hat{\mathbf{T}}_{MV}$, the matrix is singular.
- By enforcing $\hat{\mathbf{T}}_{MV}$ to be diagonal, the following similarity criterion is obtained:

$$\log\left(\frac{|\mathbf{k}_{1,1}|}{|\mathbf{k}_{2,1}|} + \frac{|\mathbf{k}_{2,1}|}{|\mathbf{k}_{1,1}|}\right) + \log\left(\frac{|\mathbf{k}_{1,2}|}{|\mathbf{k}_{2,2}|} + \frac{|\mathbf{k}_{2,2}|}{|\mathbf{k}_{1,2}|}\right) + \log\left(\frac{|\mathbf{k}_{1,3}|}{|\mathbf{k}_{2,3}|} + \frac{|\mathbf{k}_{2,3}|}{|\mathbf{k}_{1,3}|}\right).$$

NB: the diagonal assumption is only used to derive a similarity criterion between two noisy pixels. Full covariance matrices are then estimated via WMLE.

Setting of the weights for PolSAR data (2/3)

Bayesian decomposition

$$p(\mathbf{T}_1 = \mathbf{T}_2 | \mathbf{k}_1, \mathbf{k}_2) \propto \underbrace{p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2)}_{\text{similarity likelihood}} \times \underbrace{\boxed{p(\mathbf{T}_1 = \mathbf{T}_2)}}_{\textit{a priori} \text{ similarity}}$$

Refined similarity

- Refining the weights is necessary when the signal to noise ratio is low.
- Using a pre-estimate $\hat{\mathbf{T}}$ at pixel 1 and 2 provides two estimations of the noise models:

$$p(.|\hat{\mathbf{T}}_1) \qquad \text{and} \qquad p(.|\hat{\mathbf{T}}_2).$$

- Assuming $p(\mathbf{T}_1 = \mathbf{T}_2)$ depends on the proximity of $p(.|\hat{\mathbf{T}}_1)$ to $p(.|\hat{\mathbf{T}}_2)$, then we define

$$p(\mathbf{T}_1 = \mathbf{T}_2) \triangleq \exp(-\frac{SD_{KL}(\hat{\mathbf{T}}_1 || \hat{\mathbf{T}}_2)}{T})$$

$$\text{with } SD_{KL}(\hat{\mathbf{T}}_1, \hat{\mathbf{T}}_2) \propto tr\left(\hat{\mathbf{T}}_1^{-1}\hat{\mathbf{T}}_2\right) + tr\left(\hat{\mathbf{T}}_2^{-1}\hat{\mathbf{T}}_1\right) - 6.$$

- The Kullback-Leibler divergence provides a statistical test of the hypothesis $\mathbf{T}_1 = \mathbf{T}_2$ [Polzehl and Spokoiny, 2006]

Setting of the weights for PolSAR data (3/3)

Iterative scheme

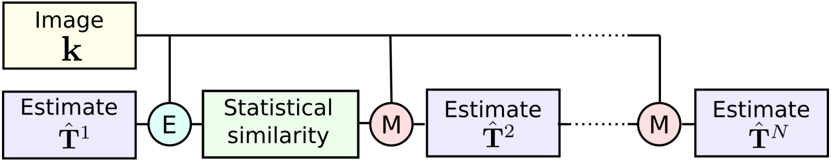
- The refined similarity involves an iterative scheme in two steps:
  1. Estimate the weights from $\mathbf{k}$ and $\hat{\mathbf{T}}^{i-1}$:

$$w(s,t) \leftarrow \prod p(\mathbf{k}_1, \mathbf{k}_2 | \mathbf{T}_1 = \mathbf{T}_2)^{1/h} p(\mathbf{T}_1 = \mathbf{T}_2 | \hat{\mathbf{T}}^{i-1})^{1/T},$$

  2. Maximize the weighted likelihood:

$$\hat{\mathbf{T}}_s^i \leftarrow \frac{\sum w(s,t)\mathbf{k}_t\mathbf{k}_t^\dagger}{\sum w(s,t)}.$$

- The procedure converges in about ten iterations.



Scheme of the iterative filtering process

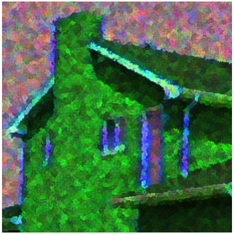
Table of contents

Results of NL-PolSAR (Simulations)

(a) Noise-free

(b) Noisy

(c) Boxcar 7×7



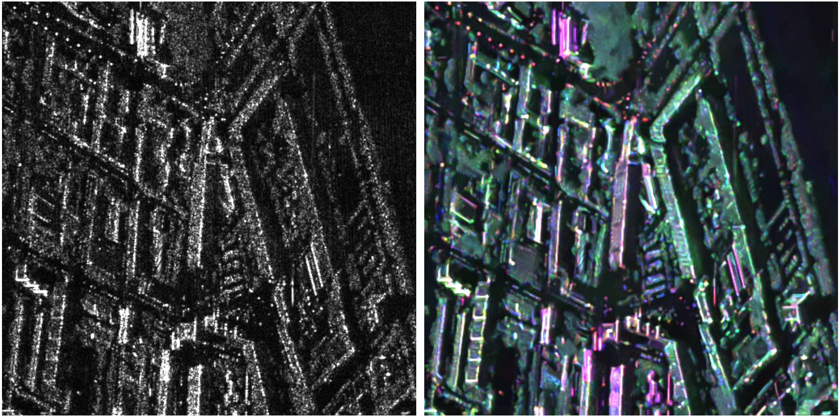
(d) Lee

(e) IDAN

(f) **NL-PolSAR**

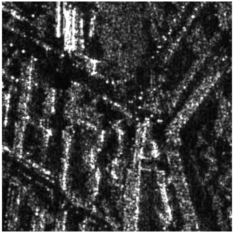
$$\underbrace{|HH + VV|, |HH - VV|, |HV|}_{RVB}$$

Results of NL-PolSAR (Dresden, -Band E-SAR data ©DLR)

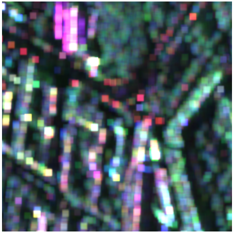


(a) Noisy HH (b) **NL-PolSAR**

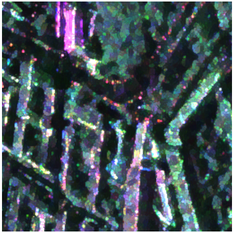
$$\underbrace{|HH + VV|, |HH - VV|, |HV|}_{RVB}$$



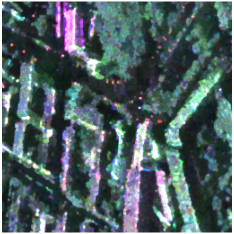
(a) Noisy HH



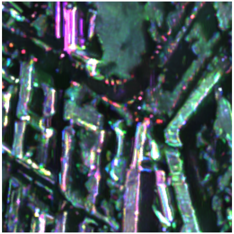
(b) Boxcar 7×7



(c) Lee



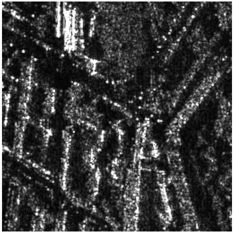
(e) IDAN



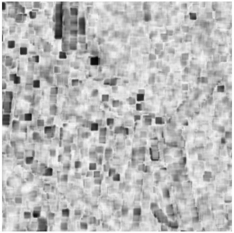
(f) **NL-PolSAR**

$$\underbrace{|HH+VV|, |HH-VV|, |HV|}_{RVB}$$

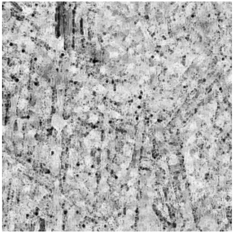
Results of NL-PolSAR (Dresden, -Band E-SAR data ©DLR)



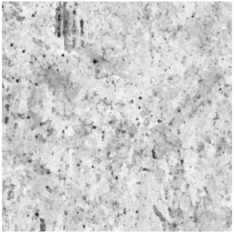
(a) Noisy HH



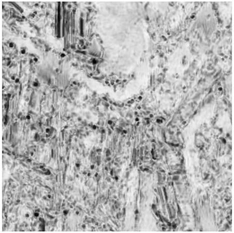
(b) Boxcar 7×7



(c) Lee



(e) IDAN



(f) **NL-PolSAR**

Entropy H

Conclusion

- We proposed an efficient estimator of local covariance matrices for PolSAR data based on non-local approaches,
- The idea is to search iteratively the most suitable pixels to combine,
- Similarity criteria:
  - joint similarity between the noisy observations (HH, HV, VV) of surrounding patches, and
  - joint similarity between the pre-estimated covariance matrices of surrounding patches.

Conclusion

- We proposed an efficient estimator of local covariance matrices for PolSAR data based on non-local approaches,
- The idea is to search iteratively the most suitable pixels to combine,
- Similarity criteria:
  - joint similarity between the noisy observations (HH, HV, VV) of surrounding patches, and
  - joint similarity between the pre-estimated covariance matrices of surrounding patches.
- Results on simulated and L-Band E-SAR data:
  - Good noise reduction without significant loss of resolution,
  - Preservation of the inter-channel information,
  - Outperforms the state-of-the-art estimators,
  - Appealing for classification purposes.

Conclusion

- We proposed an efficient estimator of local covariance matrices for PolSAR data based on non-local approaches,
- The idea is to search iteratively the most suitable pixels to combine,
- Similarity criteria:
  - joint similarity between the noisy observations (HH, HV, VV) of surrounding patches, and
  - joint similarity between the pre-estimated covariance matrices of surrounding patches.
- Results on simulated and L-Band E-SAR data:
  - Good noise reduction without significant loss of resolution,
  - Preservation of the inter-channel information,
  - Outperforms the state-of-the-art estimators,
  - Appealing for classification purposes.
- More details in:

[Deledalle et al., 2009] Deledalle, C., Denis, L., and Tupin, F. (2009).
Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights.
IEEE Transactions on Image Processing, 18(12):2661–2672.

website: `http://perso.telecom-paristech.fr/~deledall`

[Buades et al., 2005] Buades, A., Coll, B., and Morel, J. (2005).
A Non-Local Algorithm for Image Denoising.
Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2.

[Deledalle et al., 2009] Deledalle, C., Denis, L., and Tupin, F. (2009).
Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights.
IEEE Transactions on Image Processing, 18(12):2661–2672.

[Lee, 1981] Lee, J. (1981).
Refined filtering of image noise using local statistics.
Computer graphics and image processing, 15(4):380–389.

[Polzehl and Spokoiny, 2006] Polzehl, J. and Spokoiny, V. (2006).
Propagation-separation approach for local likelihood estimation.
Probability Theory and Related Fields, 135(3):335–362.

[Vasile et al., 2006] Vasile, G., Trouvé, E., Lee, J., and Buzuloiu, V. (2006).
Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation.
IEEE Transactions on Geoscience and Remote Sensing, 44(6):1609–1621.

[Yaroslavsky, 1985] Yaroslavsky, L. (1985).
Digital Picture Processing.
Springer-Verlag New York, Inc. Secaucus, NJ, USA.

(a) Noise-free

(b) Noisy

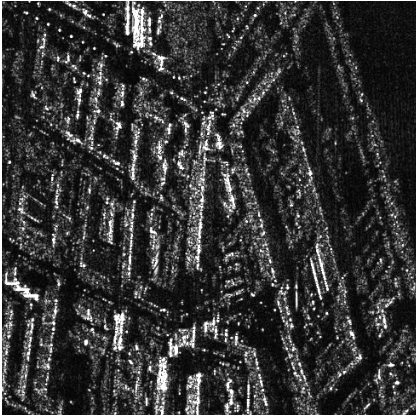
(c) Boxcar 7×7

(d) Lee

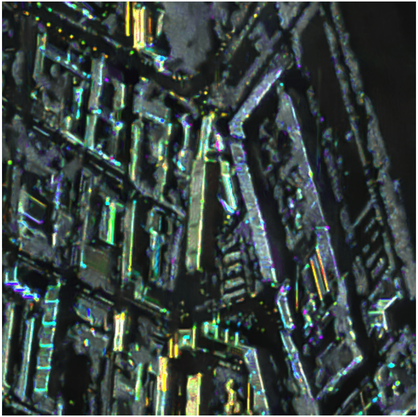
(e) IDAN

(f) **NL-PolSAR**

$$\underbrace{|1-\alpha/\pi|}_{H}, \underbrace{|1-H|}_{S}, \underbrace{|HH|+|VV|+|HV|}_{V}$$



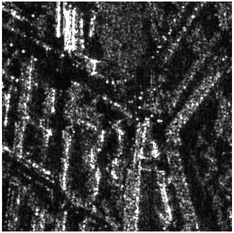
(a) Noisy



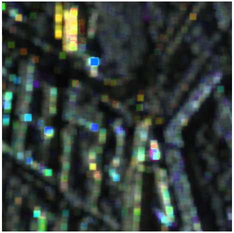
(b) NL-PolSAR

$$\underbrace{|1-\alpha/\pi|}_{H}, \underbrace{|1-H|}_{S}, \underbrace{|HH|+|VV|+|HV|}_{V}$$

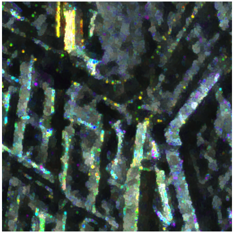
Results of NL-PolSAR (Dresden, L-Band E-SAR data ©DLR)



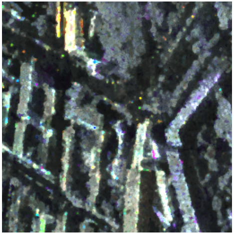
(a) Noisy HH



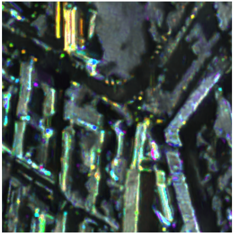
(b) Boxcar 7 × 7



(c) Lee



(e) IDAN



(f) **NL-PolSAR**

$$\underbrace{|1-\alpha/\pi|}_{H}, \underbrace{|1-H|}_{S}, \underbrace{|HH|+|VV|+|HV|}_{V}$$