Self-supervised learning of deep despeckling networks with MERLIN: ensuring the statistical independence of the real and imaginary parts

Emanuele Dalsasso, Frédéric Brigui, Loïc Denis, Rémy Abergel, Florence Tupin

Self-supervised learning of deep despeckling networks with MERLIN: ensuring the statistical independence of the real and imaginary parts

Emanuele Dalsasso[a], Frédéric Brigui[b], Loïc Denis[c,a], Rémy Abergel[d], and Florence Tupin[a]
[a]LTCI, Télécom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France
[b]DEMR, ONERA, Université Paris-Saclay, 91120 Palaiseau, France
[c]Université Jean Monnet Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Étienne, France
[d]Laboratoire MAP5, UMR CNRS 8145, Université Paris Cité, France

Abstract

Due to the wide variety of sensors, with different spatial resolutions, operating frequency bands, as well as acquisition modes (Stripmap, Spotlight, TOPS...), despeckling neural networks trained on a given type of SAR images do not perform well on other kinds of images. By considerably simplifying the building of training sets and directly including images from the sensor and acquisition mode of interest, self-supervised learning is a very appealing solution. This paper analyses the preprocessing requirements of the MERLIN strategy that assumes statistical independence of the real and imaginary parts of single look complex SAR images to perform the self-supervised training.

1 Introduction

Deep neural networks lead to unprecedented quality in several image restoration tasks (see for example the recent transformer-based methods [1, 2]). In the context of SAR despeckling, two difficulties impede the direct application of these techniques: the lack of speckle-free images (which can be only partially overcome by either averaging long time series or degrading the spatial resolution) and the poor generalization performance when moving from a training scenario in which spatially-uncorrelated synthetic speckle is considered to actual SAR images with correlations due to the system response [3]. Ideally, networks should be trained directly from SAR images corrupted by speckle, without requiring additional data to provide a ground truth. Several self-supervised learning strategies have been proposed these last years to overcome these limitations [4]. MERLIN [5] (coMplex sElf-supeRvised despeckLINg) is a self-supervised approach that requires only single look complex (SLC) images and that is robust to the spatial correlations of speckle. It therefore possesses attractive practical advantages to train networks tailored to specific sensors and/or acquisition modes. Yet, a key element for the application of MERLIN is the statistical independence of the real and imaginary parts of SLC images. While this independence condition is quite straightforward to obtain from satellite images captured in Stripmap mode, handling other SAR systems is more challenging.
This paper provides an in-depth analysis of the impact of a zero Doppler shift and describes pre-processing steps that make MERLIN applicable to Sentinel-1 images acquired with TOPS mode or airborne images with asymmetrical power spectral densities.

2 Learning to despeckle in a fully unspervised way: MERLIN

In this section, we first recall the self-supervised strategy MERLIN introduced in [5], then we analyze the impact of the SAR system response. In particular, we derive an analytical model that explains why a zero Doppler shift is detrimental to the despeckling process. In section 3, we then describe pre-processing strategies that are required for two specific SAR systems: Sentinel-1 with TOPS scanning mode and airborne SAR such as drones.

2.1 Recall of the self-supervised training strategy in MERLIN

Training a despeckling network in the absence of ground truth is possible using the self-supervised training strategy MERLIN [5, 6] that separates the real and imaginary parts of an SLC image and then trains a network fed with either the real part or the imaginary part to predict values as close as possible (in a statistical sense) to the imaginary or real parts, respectively. Under the assumption of fully developed speckle, the real part $\boldsymbol{a} \in \mathbb{R}^N$ and the imaginary part $\boldsymbol{b} \in \mathbb{R}^N$ of the complex amplitudes $\boldsymbol{z} \in \mathbb{C}^N$ are indeed statistically independent. Accessing only the real or imaginary component is not sufficient to be able to predict the fluctuations due to speckle in the other component and only the deterministic component of the signal is extracted: the scene reflectivity. It is proved in [6] that estimators of the reflectivity $\widehat{\boldsymbol{r}}(\cdot)$ that maximize, on average, the likelihoods $\mathrm{p}(\boldsymbol{b}|\widehat{\boldsymbol{r}}(\boldsymbol{a}))$ and $\mathrm{p}(\boldsymbol{a}|\widehat{\boldsymbol{r}}(\boldsymbol{b}))$ can recover the actual reflectivity, even though this reflectivity is not available for training, and that, in the presence of strong scatterers, the network recovers the sum of the intensity of the scatterer

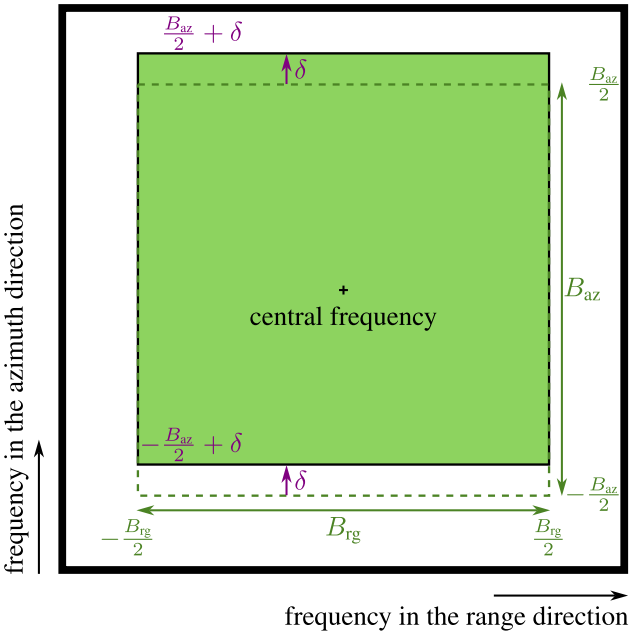

Figure 1 A simple model of the transfer function of a SAR system with a zero Doppler shift δ.

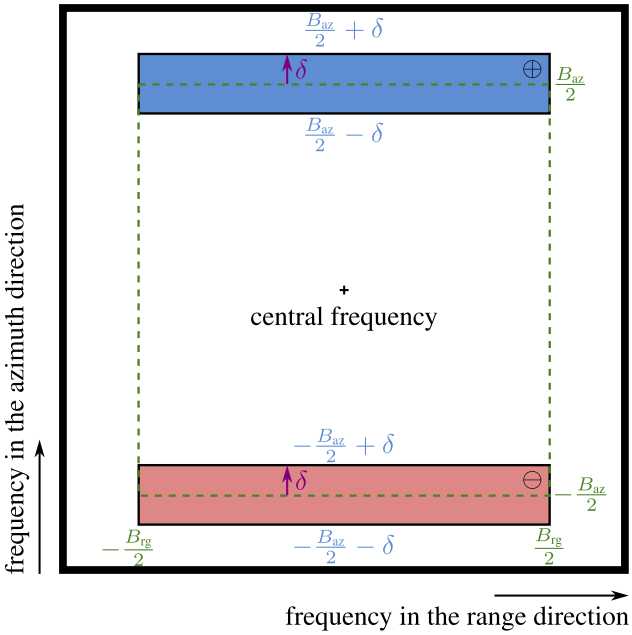

Figure 2 The cross spectral density between the real and imaginary parts of speckle-corrupted complex amplitudes induced by the transfer function shown in Figure 1. The blue rectangle is positive while the red rectangle is negative.

and of the reflectivity of the background.

2.2 Statistical independence of real and imaginary parts and the impact of the SAR system response

The statistical independence of the real and imaginary parts of the SLC images used to train the despeckling network is essential. The likelihood $\mathrm{p}(\boldsymbol{b}|\boldsymbol{r})$ is, indeed, maximal for $\boldsymbol{r}=\boldsymbol{b}$. If $\boldsymbol{a}$ and $\boldsymbol{b}$ were statistically dependent, the network would learn a mapping $\widehat{\boldsymbol{r}}(\boldsymbol{a})$ that recovers at best $\boldsymbol{b}$, thereby keeping most of the speckle fluctuations.
Let $\boldsymbol{r}\in\mathbb{R}_+^N$ be the vector containing the reflectivities of the N pixels of the scene. Under Goodman's fully developed speckle model, the real part $\boldsymbol{a}\in\mathbb{R}^N$ and the imaginary part $\boldsymbol{b}\in\mathbb{R}^N$ of the vector of complex amplitudes $\boldsymbol{z}\in\mathbb{C}^N$ are jointly Gaussian, centered, and have the following covariance matrix:

$$\mathrm{Cov}\left[\begin{pmatrix}\boldsymbol{a}\\ \boldsymbol{b}\end{pmatrix}\right]=\begin{pmatrix}\frac{1}{2}\mathrm{diag}(\boldsymbol{r}) & \mathbf{0}\\ \mathbf{0} & \frac{1}{2}\mathrm{diag}(\boldsymbol{r})\end{pmatrix}. \quad (1)$$

Under this simple modeling, the real and imaginary parts of the complex amplitude are not only statistically independent but also spatially uncorrelated. In practice, the response of the SAR system is not perfect but involves some amount of smoothing, i.e., the complex amplitudes after the SAR synthesis $\tilde{\boldsymbol{z}}$ are related to the ideal complex amplitudes $\boldsymbol{z}$ through the spatial-domain SAR system response $\mathbf{H}\in\mathbb{C}^{N\times N}$: $\tilde{\boldsymbol{z}}=\mathbf{H}\boldsymbol{z}$. The decomposition $\mathbf{H}=\mathbf{M}+j\mathbf{N}$ of the SAR system response into real and imaginary components leads to the expression of the covariance matrix of the real and imaginary parts $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$ of the SLC image $\tilde{\boldsymbol{z}}$:

$$\mathrm{Cov}\left[\begin{pmatrix}\tilde{\boldsymbol{a}}\\ \tilde{\boldsymbol{b}}\end{pmatrix}\right]= \begin{pmatrix}\mathbf{M} & -\mathbf{N}\\ \mathbf{N} & \mathbf{M}\end{pmatrix}\begin{pmatrix}\frac{1}{2}\mathrm{diag}(\boldsymbol{r}) & \mathbf{0}\\ \mathbf{0} & \frac{1}{2}\mathrm{diag}(\boldsymbol{r})\end{pmatrix}\begin{pmatrix}\mathbf{M}^{\mathrm{t}} & \mathbf{N}^{\mathrm{t}}\\ -\mathbf{N}^{\mathrm{t}} & \mathbf{M}^{\mathrm{t}}\end{pmatrix}, \quad (2)$$

which gives:

$$\mathrm{Cov}\left[\tilde{\boldsymbol{a}}\right]=\mathrm{Cov}\left[\tilde{\boldsymbol{b}}\right]=\tfrac{1}{2}\mathbf{M}\mathrm{diag}(\boldsymbol{r})\mathbf{M}^{\mathrm{t}}+\tfrac{1}{2}\mathbf{N}\mathrm{diag}(\boldsymbol{r})\mathbf{N}^{\mathrm{t}} \quad (3)$$

and

$$\mathrm{Cov}\left[\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}}\right]=\tfrac{1}{2}\mathbf{M}\mathrm{diag}(\boldsymbol{r})\mathbf{N}^{\mathrm{t}}-\tfrac{1}{2}\mathbf{N}\mathrm{diag}(\boldsymbol{r})\mathbf{M}^{\mathrm{t}}. \quad (4)$$

This shows that the SAR system response not only introduces spatial correlations in $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$ (equation (3)) but can also break the initial statistical independence between real and imaginary parts (equation (4) when $\mathbf{M}\mathrm{diag}(\boldsymbol{r})\mathbf{N}^{\mathrm{t}}\neq\frac{1}{2}\mathbf{N}\mathrm{diag}(\boldsymbol{r})\mathbf{M}^{\mathrm{t}}$). In the next paragraph we show that this is in particular the case of SAR systems with zero Doppler shifts.

2.3 A model of the impact of zero Doppler shifts

When considering sufficiently small areas, the SAR system can be considered shift-invariant (similarly to the isoplanatism domain of optical systems) and characterized by a transfer function in Fourier domain. In order to derive closed-form expressions of the correlations between the real and imaginary parts of a SAR image, we consider the following simplifying assumptions: (i) the area is homogeneous (and $\boldsymbol{r}=\mathbf{1}$ to shorten the notations), (ii) the SAR

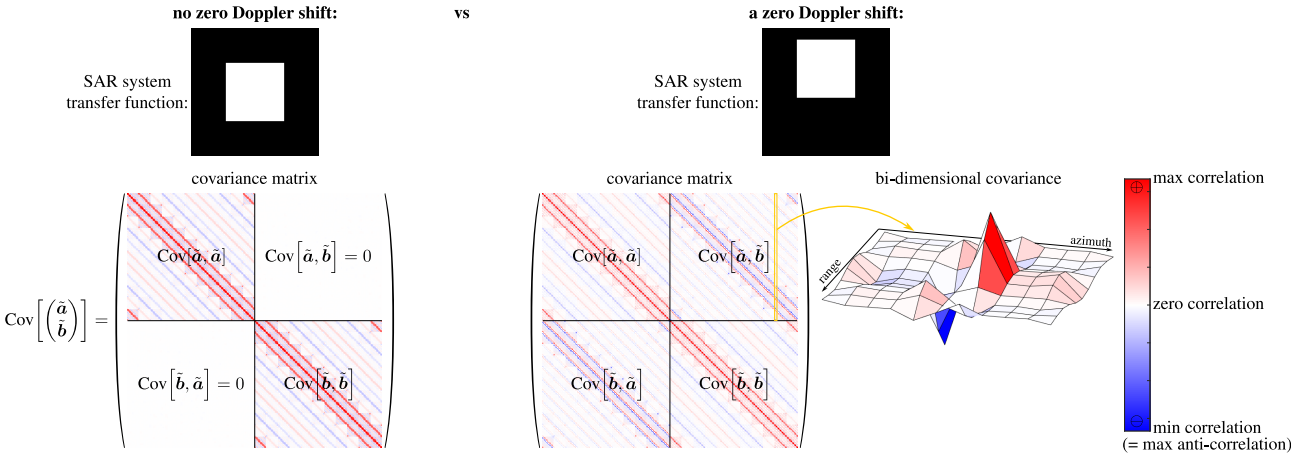

Figure 3 Numerical experiment illustrating the theoretical derivation of the covariance of real and imaginary components of an SLC image in case of a zero-Doppler shift: in the absence of zero-Doppler shift (left-hand side), there are spatial correlations among the real and imaginary parts but the imaginary part is independent from the real part. This is no longer the case when a zero-Doppler shift is present (right-hand side of the figure).

system is shift-invariant and its transfer function is a rectangular function in the range and azimuth frequency directions (see Figure 1). The continuous-domain expression of the transfer function of such system is:

$$t(\nu_{\text{rg}}, \nu_{\text{az}}) = \Pi_{[-\frac{B_{\text{rg}}}{2}, \frac{B_{\text{rg}}}{2}]}(\nu_{\text{rg}}) \cdot \Pi_{[-\frac{B_{\text{az}}}{2}+\delta, \frac{B_{\text{az}}}{2}+\delta]}(\nu_{\text{az}}), \quad (5)$$

with $\Pi_{[a,b]}(\nu)$ the rectangular function such that $\Pi_{[a,b]}(\nu) = 1$ for all ν such that $\nu \in [a, b]$ and $\Pi_{[a,b]}(\nu) = 0$ otherwise. In space-domain, the impulse response of the SAR system is:

$$h(x, y) = B_{\text{rg}}\text{sinc}(\pi x B_{\text{rg}}) B_{\text{az}}\text{sinc}(\pi y B_{\text{az}}) \exp(-2j\pi y\delta), \quad (6)$$

with $\text{sinc}(u) = (\sin u)/u$ for $u \neq 0$ and $\text{sinc}(0) = 1$ the cardinal sine function, and x and y the range and azimuth coordinates, respectively.

Owing to the shift-invariance property of the SAR system, the spatial-domain response is convolutive and writes: $\mathbf{H} = \text{conv}(\boldsymbol{h})$, where conv() builds a convolution matrix and $\boldsymbol{h}$ is the sampled impulse response. Note that the real part of h is even in both variables x and y while the imaginary part of h is even with respect to x and odd with respect to y. The operator $\mathbf{M}$ thus corresponds to the convolution with an even kernel and $\mathbf{M}^{\text{t}} = \mathbf{M}$, while $\mathbf{N}$ is a convolution with an odd kernel and $\mathbf{N}^{\text{t}} = -\mathbf{N}$. Since we assumed that $\text{diag}(\boldsymbol{r})$ corresponds to the identity matrix, we get that $\text{Cov}\left[\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}\right] = -\mathbf{MN}$ (equation (4) gives $-\frac{1}{2}(\mathbf{MN} + \mathbf{NM})$ and operators $\mathbf{M}$ and $\mathbf{N}$ being discrete convolutions, they commute). The composition of convolution operators $\mathbf{M}$ and $\mathbf{N}$ corresponds to a single convolution that we characterize in Fourier domain in the following. Since $\mathbf{M}$ and $\mathbf{N}$ are the real and imaginary part of $\mathbf{H}$, their transfer function are given by the decomposition of $\mathbf{H}$ into a symmetrical and an anti-symmetrical component. Under our continuous-domain modeling, we can decompose:

$$t(\nu_{\text{rg}}, \nu_{\text{az}}) = \underbrace{\frac{t(\nu_{\text{rg}}, \nu_{\text{az}}) + t(\nu_{\text{rg}}, -\nu_{\text{az}})}{2}}_{\text{even}} + \underbrace{\frac{t(\nu_{\text{rg}}, \nu_{\text{az}}) - t(\nu_{\text{rg}}, -\nu_{\text{az}})}{2}}_{\text{odd}}, \quad (7)$$

where the even component gives the real part of h in the spatial domain and the odd component gives j times its imaginary part. We represent in Figure 2 the product of these transfer functions, i.e., the cross spectral density (CSD) of the real and imaginary parts of $\tilde{z}$:

$$S_{\tilde{a}\tilde{b}} = \tfrac{-1}{4j}\Pi_{[-\frac{B_{\text{rg}}}{2}, \frac{B_{\text{rg}}}{2}]}(\nu_{\text{rg}}) \cdot \Big(\Pi_{[\frac{B_{\text{az}}}{2}-\delta, \frac{B_{\text{az}}}{2}+\delta]}(\nu_{\text{az}}) - \Pi_{[-\frac{B_{\text{az}}}{2}-\delta, -\frac{B_{\text{az}}}{2}+\delta]}(\nu_{\text{az}})\Big). \quad (8)$$

The inverse Fourier transform of this CSD finally gives the cross-correlation function:

$$\begin{aligned}[\tilde{a} \star \tilde{b}](x, y) &= B_{\text{rg}}\text{sinc}(\pi x B_{\text{rg}})\delta\text{sinc}(2\pi y\delta)\sin(\pi y B_{\text{az}}) \\ &= \frac{B_{\text{rg}} B_{\text{az}}}{2}\text{sinc}(\pi x B_{\text{rg}})\text{sinc}(\pi y B_{\text{az}})\sin(2\pi y\delta)\end{aligned} \quad (9)$$

which is the continuous-domain version of the covariance $\text{Cov}\left[\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}\right]$. The expression (9) shows that the cross-correlation (and the covariance) between the real and imaginary parts is zero when $\delta = 0$, which was expected since h is then real-valued (i.e., $\mathbf{N} = \mathbf{0}$). When $\delta \neq 0$, the cross-correlation is still zero for $y = 0$. This indicates that, under our model of a SAR system with a rectangular transfer function and a zero Doppler shift, the real and imaginary parts of the complex amplitudes at any given pixel are statistically independent. Yet, there are statistical dependencies between the real part at a pixel and the imaginary parts of the pixels with neighboring azimuth coordinates.

This precludes the use of the real and imaginary parts of an SLC image with a zero Doppler shift to train a neural network with MERLIN: the network would then exploit the statistical dependency of neighboring azimuth locations to recover the speckle in the other component, leaving a significant amount of speckle in its output.

Figure 3 illustrates on a small numerical experiment these results. Images with constant reflectivities and fully-developed speckle were simulated. A different SAR transfer function has been applied to the images: on the left-hand side of the figure the rectangular transfer function has no zero-Doppler shift, while on the right-hand side there is a zero-Doppler shift. The bottom part of the figure displays the covariance matrix of the vector formed by concatenating all values of the real part and of the imaginary part of an image, as defined in equation (2). This 2×2 block matrix is block-diagonal in the absence of zero-Doppler shift, which indicates the independence[1] of the real and imaginary parts. There are some correlations between the real part values of neighboring pixels due to the low-pass filtering effect of the SAR system (and, similarly, correlations between imaginary part values): the blocks on the diagonal are full[2]. When a zero-Doppler shift is introduced, correlations between the real and imaginary parts appear (see bottom-right of the figure). A bi-dimensional representation of the spatial structure of this covariance shows the expected shape derived in equation (9) with a positive correlation and a negative correlation due to the sine function. This theoretical analysis and numerical experiments indicate that special care must be taken to ensure that the real and imaginary parts of the SLC image are actually independent by applying an adequate pre-processing step.

3 Experimental results

3.1 SAR-Light drone data

3.1.1 Data pre-processing

Drone data has been acquired by the ONERA SAR drone SAR-Light [7] in X-band at high resolution (over a frequency bandwidth of 720MHz) and in stripmap configuration. Azimuth resolution is achieved with a constant integration angle for the central frequency of the emitted signal. The spectral support of the SLC image is then larger for high frequencies and smaller for small frequencies. This yields to a transfer function with a trapezoidal support which cannot be made symmetric without modifying its shape. To do so, the support of the transfer function t is first defined by the binary mask $w(\nu_{\text{rg}}, \nu_{\text{az}})$:

$$w(\nu_{\text{rg}}, \nu_{\text{az}}) = \begin{cases} 1 & \text{if } |t(\nu_{\text{rg}}, \nu_{\text{az}})| > \lambda \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where the threshold λ has been chosen empirically. When estimating the mask w from the power spectral density of an SLC image, mathematical morphology operations are generally necessary to fill small holes and obtain a regular support. This support can be made fuzzy by additionally performing a convolution with a Gaussian kernel. Assuming that the azimuth spectrum has been centered on 0Hz, a symmetrical support w_{sym} is obtained by performing horizontal and vertical symmetries:

$$w_{\text{sym}}(\nu_{\text{rg}}, \nu_{\text{az}}) = w(\nu_{\text{rg}}, \nu_{\text{az}}) \cdot w(-\nu_{\text{rg}}, \nu_{\text{az}}) \cdot w(\nu_{\text{rg}}, -\nu_{\text{az}}) \cdot w(-\nu_{\text{rg}}, -\nu_{\text{az}}). \tag{11}$$

The pre-processed image is obtained by applying a filter with a transfer function corresponding to the product of the symmetric mask w_{sym} and a spectral apodization (introduced to reduce the sidelobes associated to the response of bright targets). The apodization function that is applied is a 2D separable Hamming function, see Figure 4. The transfer function of the pre-processed data has the Hermitian symmetry, implying the independence between real and imaginary parts of the pre-processed SLC image. Moreover, the apodization function reduces the smearing of strong point-like scatterers, making the image easier to interpret.

Table 1 Description of the hyperparameters used to train a residual U-Net with MERLIN.

	SAR-light	Sentinel-1 TOPS
# images	1	10
patch size	256×256	256×256
batch size	12	12
stride size	32	128
# batches	101	1365
# epochs	30	30
gradient norm	2.0	1.0
learning rate	10^{-2}	10^{-2}
	10^{-3} after 4 epochs	10^{-3} after 4 epochs
	10^{-4} after 20 epochs	10^{-4} after 20 epochs

3.1.2 Results

MERLIN has been trained on a single image acquired by the drone system SAR-light operated by the ONERA. Image size is 931×2000 pixels and it has been cropped into overlapping patches as described in Table 1. Figure 5 shows the result on a crop of size 500×500 pixels. It can be observed that the reflectivity of the strong scatterer in the middle of the scene is well restored by MERLIN. Compared to MuLoG+BM3D, MERLIN can handle spatial correlations in the speckle, thus preserving the original image resolution. As a consequence, MuLoG+BM3D has failed to reconstruct the dark and low-contrasted linear structure in the top left of the image, while MERLIN has effectively preserved it. Moreover, fine textures characterizing the agricultural field are well restored by MERLIN, while MuLoG+BM3D has oversmoothed such areas.

3.2 Sentinel-1 TOPS data

3.2.1 Data pre-processing

Bursts composing TOPS SLC products present a linearly varying Doppler frequency in the azimuth direction, which is due to the steering of the antenna [8]. Inverting such lin-

[1] note that since the real and imaginary parts of SLC images are jointly Gaussian, decorrelation and statistical independence are equivalent

[2] since the SAR system is shift-invariant, these covariance blocks have a special structure: Block Toeplitz Matrix with Toeplitz Blocks (BTTB), i.e., it corresponds to a 2D convolution matrix

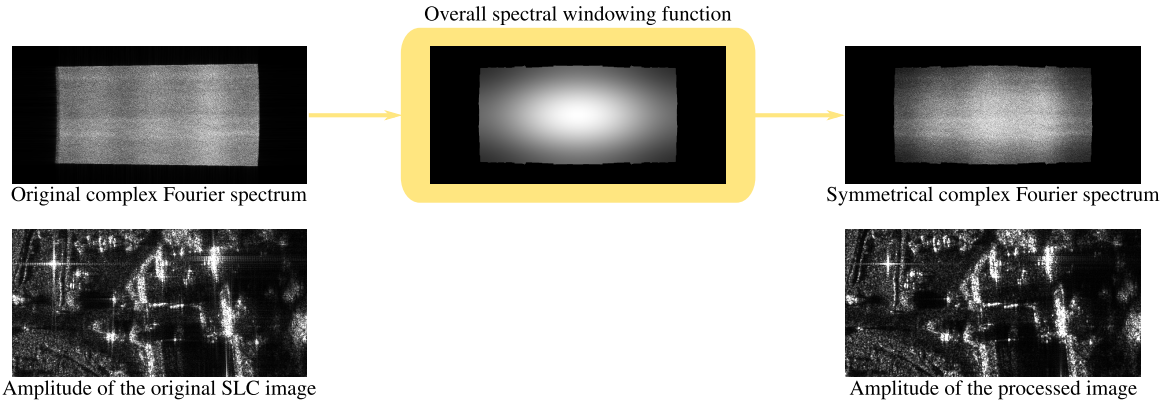

Figure 4 Pre-processing of drone SLC data to make its spectrum symmetrical.

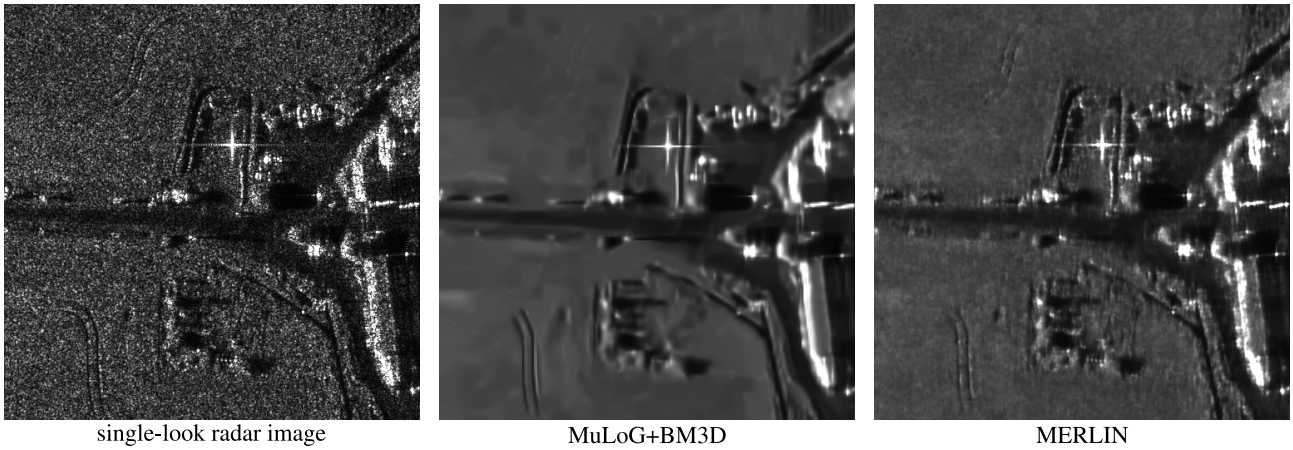

Figure 5 Results on SAR-light data. Left: the speckle corrupted amplitude image. Center: the image filtered with MuLoG+BM3D. Right: estimation produced with MERLIN.

ear frequency modulation is necessary to ensure symmetry of the spectrum. This operation is referred to as *deramping*. Depending on the acquisition process, even after deramping the azimuth spectrum might not be centered on 0Hz. The operation consisting in estimating the Doppler centroid frequency and applying a global shift to the complex spectrum is called *demodulation*. Retrieving the deramped and demodulated SLC image $\bar{z}$ from the original TOPS SLC image $\tilde{z}$ can be done in a single step as follows [9]:

$$\bar{z} = \mathrm{diag}\,(\mathbf{\Phi}) \cdot \mathrm{diag}\,(\mathbf{\Psi}) \cdot \tilde{z}\,, \tag{12}$$

where $\mathbf{\Phi}$ and $\mathbf{\Psi}$ correspond respectively to the vectors of deramping and demodulation values, as provided by the SLC product metadata, see [8].

3.2.2 Results

To train a residual U-Net with MERLIN we applied the pre-processing described above to a set of 10 TOPS Sentinel-1 bursts and used them as training data, with the hyperparameters set as indicated in Table 1. Results obtained with MERLIN are compared to those produced by SAR2SAR [10], which is a semi-supervised despeckling algorithm that exploits temporal acquisitions to fine-tune a pre-trained model on real SAR intensities. Thus, both MERLIN and SAR2SAR handle spatially correlated speckle. MERLIN is very effective at filtering out the speckle. The complex scene of Figure 6 demonstrates that field frontiers are well respected, linear structures and bright targets are preserved and no residual speckle fluctuations seem to affect the filtered image. In terms of quality of the results, SAR2SAR and MERLIN are comparable. What makes MERLIN advantageous is the simplicity to set up a training.

4 Conclusion

The application of the self-supervised training strategy MERLIN requires the statistical independence of the real and imaginary parts of the SLC images. A theoretical analysis of the effect of an asymmetry of the power spectral density of the SLC image due to the presence of a zero Doppler shift shows that correlations appear between the real and imaginary parts. Interestingly, these correlations do not show up when computing the correlation between the real and imaginary values at a given pixel but rather between real and imaginary values of pixels with nearby azimuth coordinates. Compensating this zero Doppler shift requires deramping Sentinel-1 images acquired in TOPS mode. Beyond shifts of the spectrum, other asymmetries must also be suppressed. This is especially the case of airborne images where variations of the size of the synthetic antenna between near range and far range lead to

Figure 6 Results on Sentinel-1 TOPS data. Left: the speckle corrupted amplitude image. Center: the image filtered with SAR2SAR. Right: estimation produced with MERLIN.

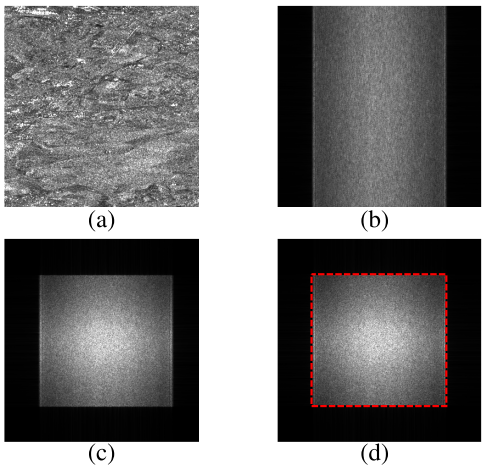

Figure 7 Effect of deramping and demodulation on a TOPS Sentinel-1 SLC product. (a) the amplitude of the SLC image; (b) the complex Fourier spectrum; (c) the deramped spectrum; (d) the deramped and demodulated spectrum

a trapezoidal transfer function. We proposed a dedicated preprocessing to enforce the symmetry of spectra and have shown that the self-supervised MERLIN training strategy then performs satisfyingly on these images. The wide applicability of self-supervised learning strategies is promising to learn rich models directly from SAR data.

5 Acknowledgements

This project has been funded by ANR (the French National Research Agency) and DGA (Direction Générale de l'Armement) under ASTRAL project ANR-21-ASTR-0011.

6 Literature

[1] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *CVPR*, 2022, pp. 5728–5739.

[2] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general U-shaped transformer for image restoration," in *CVPR*, 2022, pp. 17 683–17 693.

[3] S. Vitale, G. Ferraioli, and V. Pascazio, "Analysis on the building of training dataset for deep learning SAR despeckling," *IEEE GRSL*, vol. 19, pp. 1–5, 2021.

[4] E. Dalsasso, L. Denis, M. Muzeau, and F. Tupin, "Self-supervised training strategies for sar image despeckling with deep neural networks," in *EUSAR 2022*. VDE, 2022, pp. 1–6.

[5] E. Dalsasso, L. Denis, and F. Tupin, "As if by magic: self-supervised training of deep despeckling networks with MERLIN," *IEEE TGRS*, vol. 60, pp. 1–13, 2021.

[6] I. Meraoumia, E. Dalsasso, L. Denis, R. Abergel, and F. Tupin, "Multitemporal Speckle Reduction With Self-Supervised Deep Neural Networks," *IEEE TGRS*, vol. 61, pp. 1–14, 2023.

[7] F. Brigui, S. Angelliaume, N. Castet, X. Dupuis, and P. Martineau, "Sar-light - first sar images from the new onera sar sensor on uav platform," in *IGARSS 2022*, 2022, pp. 7721–7724.

[8] N. Miranda, "Definition of the TOPS SLC deramping function for products generated by the S-1 IPF," *ESA Technical Note*, 2014.

[9] R. Abergel, L. Denis, S. Ladjal, and F. Tupin, "Subpixellic methods for sidelobes suppression and strong targets extraction in single look complex SAR images," *IEEE JSTARS*, vol. 11, no. 3, pp. 759–776, 2018.

[10] E. Dalsasso, L. Denis, and F. Tupin, "SAR2SAR: A semi-supervised despeckling algorithm for SAR images," *IEEE JSTARS*, vol. 14, pp. 4321–4329, 2021.