Exploiting patch similarity for SAR image processing: the nonlocal paradigm

Charles-Alban Deledalle, Loïc Denis, Giovanni Poggi, Florence Tupin, Luisa Verdoliva

Exploiting patch similarity for SAR image processing: the nonlocal paradigm

C. Deledalle*, L. Denis†, G. Poggi‡, F. Tupin§, and L. Verdoliva‡

*IMB, CNRS – Université Bordeaux 1, France
†Laboratoire Hubert Curien, UMR 5516 CNRS / Université Jean Monnet, Université de Lyon, France
‡DIETI, University Federico II of Naples, Naples, Italy
§Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI, France

***Abstract*—Most current SAR systems offer high-resolution images featuring polarimetric, interferometric, multi-frequency, multi-angle, or multi-date information. SAR images however suffer from strong fluctuations due to the speckle phenomenon inherent to coherent imagery. Hence, all derived parameters display strong signal-dependent variance, preventing the full exploitation of such a wealth of information. Even with the abundance of despeckling techniques proposed these last three decades, there is still a pressing need for new methods that can handle this variety of SAR products and efficiently eliminate speckle without sacrificing the spatial resolution. Recently, patch-based filtering has emerged as a highly successful concept in image processing. By exploiting the redundancy between similar patches, it succeeds in suppressing most of the noise with good preservation of texture and thin structures. Extensions of patch-based methods to speckle reduction and joint exploitation of multi-channel SAR images (interferometric, polarimetric, or PolInSAR data) have led to the best denoising performance in radar imaging to date. We give a comprehensive survey of patch-based nonlocal filtering of SAR images, focusing on the two main ingredients of the methods: measuring patch similarity, and estimating the parameters of interest from a collection of similar patches.**

I. Introduction

Current SAR systems share two common characteristics: they provide a wealth of information thanks to polarimetric, wavelength or angle diversity, and they offer very high spatial resolutions that give access to the shape of man-made structures. In radar images, however, parameters of interest, like interferometric phase, coherence, polarimetric properties, or radiometry, are not directly accessible but must be estimated from unreliable data. It is essential that the estimation procedure be robust to the strong fluctuations in the measurements due to speckle without trading off the spatial resolution.

The simplest way to reduce speckle noise[1] is to average pixels in a rectangular window around the target pixel, so-called *spatial multilooking*. This process, already present in the first ERS satellite images and still widely used for interferometric or polarimetric data, leads to a uniform reduction of speckle in homogeneous areas, with a residual variance inversely proportional to the number of pixels averaged. However, by potentially mixing different signals, it also impairs such important signal features as region edges, man-made structures, and fine textures. The point is that multilooking is just a basic non-adaptive form of parameter estimation: to remove speckle without degrading fine features, local image content must be taken into account.

The design of efficient despeckling filters is a long-standing problem that has been the object of intense research since the advent of SAR technology [2], with first contributions dating back to the 80s [3]. However, research activity has accelerated significantly in recent years, reflecting both the success of SAR remote sensing in general and the lack of satisfactory methods for resolution-preserving speckle reduction. Some of the most successful methods proposed in the recent past for locally adaptive estimation are based on image models that enforce strong regularity constraints, either in the original domain (e.g., Markov random fields [4]), or in some transform domain (e.g., wavelet-based sparse representations [5]).

Very recently, *patches*, i.e., small rectangular image regions (typically squares of size between 3×3 and 11×11), have emerged as a powerful representation on which to build rich and flexible statistical models of natural images. Patches capture richer neighborhood configurations than first-order Markov random fields and are better localized than wavelets. Patch-based models do not enforce the solution to belong to a restricted class of signals, such as signals with bounded variations or with sparse transform coefficients. They exploit the self-similarity, typical of natural scenes, and look for similar patches not just in the immediate neighborhood of the target pixel but in an extended search area. Unlike a local method like [6] that considers only connected pixels, far apart pixels can be combined, thereby justifying the widespread *nonlocal* appellative. The evolution from explicit image models to the concept of patch redundancy corresponds to a true *methodological shift* in image processing. In particular, recent denoising methods most often rely on the notion of a patch.

Interestingly, the concept of nonlocal filtering originally emerged in image processing after the pioneering work of Lee [7] for SAR despeckling. However, it was only with the development of patch-based methods, following the seminal work [8] in the 2000s, and the definition of highly discriminative patch-based similarity measures, that nonlocal filtering could be successfully applied to the low-SNR speckle-corrupted SAR images. The paradigm of patch-based nonlocal estimation is particularly interesting for SAR image processing given the poor fit of classical models to SAR scenes, characterized by many very strong punctual targets and high-contrast structures that are poorly modeled with piecewise constants or

This work has been supported in part by the following grants: ANR DEFI 09-EMER-008-01, ANR-11-IDEX-0007, ANR-11-LABX-0063.

[1]Although speckle is itself a signal of possible interest, in the context of despeckling it is an undesired component, and hence customarily referred as "noise" with a slight abuse of terminology [1]

SAR IMAGING MODALITIES

Synthetic Aperture Radar systems are based on the emission of an electromagnetic wave which is then back-scattered by the ground level features, and finally recorded by the receiving antenna. Several modalities are used in SAR imaging (see [15] for a comprehensive survey):

amplitude: The simplest configuration provides, after SAR synthesis, an image of scattering coefficients $\boldsymbol{k}$ that are complex values with a magnitude (i.e., amplitude) that is representative of the radar cross-section. The square magnitude (i.e., intensity) is also often considered.

polarimetric: The use of different polarizations at emission and reception of the radar wave provides a deeper insight into the backscattering mechanisms inside the resolution cell. The scattering vector $\boldsymbol{k}$ is formed by the complex values corresponding to each combination of emission / reception polarizations. Usually, horizontal (H) and vertical (V) polarizations are used for emission and reception providing the scattering vector $\boldsymbol{k} = [z_{HH}, z_{HV}, z_{VH}, z_{VV}]^t$ (or, by reciprocity, $\boldsymbol{k} = [z_{HH}, z_{HV}, z_{VV}]^t$). This modality (PolSAR) is widely used to study vegetation growing or urban areas.

interferometric: The combination of two backscattered signals $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ measured by two close configurations of the acquisition systems can yield elevation information or ground displacement maps. In interferometric configurations, a composite signal $\boldsymbol{k}$ can be defined by the concatenation of the two received signals $\boldsymbol{k} = [\boldsymbol{k}_1^t, \boldsymbol{k}_2^t]^t$. These can be single (InSAR) or multi-polarization (PolInSAR) signals.

wavelets.

The potential of this new paradigm for SAR imaging has quickly been recognized, with more than 30 papers published since 2009 that describe patch-based methods applied to SAR, including most state-of-the-art despeckling techniques [9], [10], [11]. In addition, nonlocal patch-based methods are very flexible and can readily be extended to different SAR modalities [12], [13], [14].

This paper reviews the underlying ideas and principles of nonlocal estimation methods proposed in SAR imaging. We consider despeckling as a familiar and important case study, but we also address the more general point of view of parameter estimation, looking ahead at extensions to more challenging SAR modalities and estimation problems. Before diving into the core of patch-based methods, we begin by describing the classical speckle model in SAR imaging and the major families of estimation methods that have emerged these last 3 decades.

II. A SHORT OVERVIEW OF SAR DESPECKLING

Depending on the modality, SAR systems can record up to 6 channels of complex valued signals (see box "SAR imaging modalities"). All these signals present highly varying fluctuations because SAR is a *coherent* imaging system (see box "Speckle fluctuations in radar images"). The simplest way to reduce these fluctuations and estimate the values of the physical parameters is to average several independent samples from the data. This operation, called multi-looking, was applied in various forms from the very beginning of the SAR era. However, such simple averaging that applies equally to every region of the image, regardless of the local heterogeneity, strongly degrades the spatial resolution.

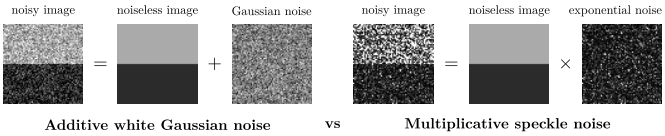
Improved approaches have long been proposed to enhance this basic estimation method by better taking into account the image information. Many efforts have been devoted to the case of amplitude images, corresponding to the modulus of single polarization data. Most of these "filtering" methods are described in the review paper [2], and the very recent tutorial [1]. The first attempts were derived according to estimation theory: Lee [3] proposed a minimum mean square error estimator in the class of linear filters, while Lopes et al. [16] considered the maximum *a posteriori* estimator. Both works used a statistical model limited to local distributions, and it is worth noting that these two filters were popular because of the clever analysis of the local context, using window splitting, or edge and target detection [16]. The idea of selecting the most relevant samples in the window has been further developed in [6] and is the main motivation of patch-based approaches. The following generation of filtering approaches introduced stronger priors to guide the solution.

A first family includes Markovian and variational approaches which impose smoothness or regularity constraints on the solution through a suitable prior model. These approaches usually lead to minimizing an energy function composed of two terms. The first term reflects the data distribution and is related to statistical models of speckle (see the box "Speckle fluctuations in radar images"). Due to the heavy tail of the distributions of speckle-corrupted images, classical least-squares data fitting must be replaced by a more relevant criterion derived from speckle distributions. The second term relies on some prior on the solution. Although regularization models such as gradient sparsity have been investigated [4], they do not fit well SAR signal properties. The Markovian formalism can be easily extended to deal with different SAR modalities like interferometric data. However, the specific nature of SAR signals is poorly captured by simple models and more complex ones lead to very hard optimization problems.

A second large family of approaches is based on wavelet transforms. Thanks to their spatially localized and multi-resolution basis functions, wavelets yield sparse yet accurate representations of natural images in the transform domain. Sharp discontinuities and point-like features, so common in SAR images, are well described by a small number of basis functions, just like the large homogeneous regions between them. This compact representation was quickly recognized as a powerful tool for denoising. In fact, while the signal is projected on a relatively small number of large wavelet coefficients, the white noise remains white after the transform, and hence evenly distributed on all coefficients. Signal and noise can be, therefore, efficiently separated by means of an appropriate non-linear processing, such as hard/soft thresholding or more sophisticated shrinkages. The encouraging results provided by early SAR despeckling techniques spawned an intense research to overcome the shortfalls of soft/hard thresholding. A popular approach considered wavelet shrinkage as a Bayesian estimation problem, possibly expressed after application of a homomorphic transform in order to reduce speckle to an independent additive perturbation. The major problem in this context becomes the modeling of signal and noise by suitable distributions, and a number of parametric models have indeed been proposed [5], [17]. Further improvements come from *joint* modeling of wavelet coefficients in and across subbands, like in [18]. Despite its potential, the wavelet transform cannot deal by itself with the high heterogeneity of SAR scenes. A number of spatially adaptive techniques were therefore proposed, based on some

SPECKLE FLUCTUATIONS IN RADAR IMAGES

Coherent signals like SAR data present strong fluctuations. The waves backscattered by elementary scatterers inside each resolution cell are not "in phase" but arbitrarily "out-of-phase", which results in constructive and destructive interferences. In a SAR intensity image, fluctuations due to speckle follow a heavy tail distribution (large deviations occur often) and are signal dependent (standard deviation is proportional to the radiometry), which departs from the usual additive Gaussian noise model:

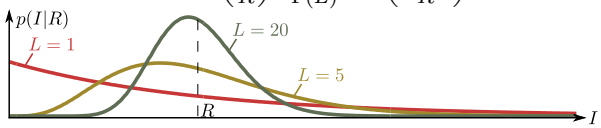


Measured SAR intensity $I > 0$ in untextured areas deviates from the radiometry $R > 0$ according to an exponential distribution:

$$p(I|R) = \frac{1}{R}\exp(-\frac{I}{R})\,. \tag{1}$$

Multi-look SAR images, obtained by averaging L intensity images, follow a gamma distribution:

$$p(I|R) = \left(\frac{L}{R}\right)^L \frac{I^{L-1}}{\Gamma(L)}\exp\left(\frac{-L\,I}{R}\right). \tag{2}$$



In the more general case of a K-dimensional scattering vector as encountered in polarimetric and interferometric modalities $\boldsymbol{k} = [z_1, \ldots, z_K]^t$, the observed complex vector $\boldsymbol{k}$ follows a K-dimensional circular complex Gaussian distribution under Goodman's fully developped speckle model (valid for untextured areas with physically homogeneous and rough surfaces):

$$p(\boldsymbol{k}|\boldsymbol{\Sigma}) = \frac{1}{\pi^K|\boldsymbol{\Sigma}|}\exp\left(-\boldsymbol{k}^\dagger\boldsymbol{\Sigma}^{-1}\boldsymbol{k}\right) \tag{3}$$

where $\boldsymbol{\Sigma} = \mathbb{E}\{\boldsymbol{k}\boldsymbol{k}^\dagger\}$ is a $K \times K$ complex covariance matrix characteristic of the imaged surface, $\mathbb{E}$ is the expectation, and † the hermitian transpose. Some radar images are not available in the form of scattering vectors, but rather as empirical covariance matrices:

$$\boldsymbol{C} = \frac{1}{L}\sum_{i=1}^{L}\boldsymbol{k}_i\boldsymbol{k}_i^\dagger\,, \tag{4}$$

where the sum is carried over several scattering vectors for each pixel. In the case of amplitude images, $\boldsymbol{k}$ is a scalar and $\boldsymbol{C}$ corresponds to an L-looks intensity image.

The empirical covariance $\boldsymbol{C}$ follows a Wishart distribution given by:

$$p(\boldsymbol{C}|\boldsymbol{\Sigma}, L) = \frac{L^{LK}|\boldsymbol{C}|^{L-K}}{\Gamma_K(L)|\boldsymbol{\Sigma}|^L}\exp\left(-L\ \mathrm{Tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{C})\right)\,, \tag{5}$$

where $\mathrm{Tr}(\cdot)$ is the matrix trace. The equivalent number of looks L acts as the shape parameter of the Wishart distribution.

prior classification of the image, typically in homogeneous, heterogeneous (e.g., textures) and highly heterogeneous (e.g., point targets) regions, in order to tune filtering parameters or strategies to the different regions [19].

The approaches reviewed above generally try to estimate the signal at a certain pixel from the noisy observations at pixels close to it. However, with plenty of data to estimate the signal, that is, the observations over the whole image, why restricting attention only to a small neighborhood of the target? The obvious answer is that not all image pixels carry valuable information on the target, and only close pixels are used because they are expected to be more similar to it, and hence, better estimators. The above consideration makes clear that the image denoising problem may be performed in two separate steps: *i)* selecting good predictors, which carry useful information on the target, and then *ii)* using them to perform the actual estimate. Lacking any other hints, spatial closeness is taken as a surrogate of signal similarity relying heavily on the fact that natural images are predominantly low-pass. Needless to say, sophisticated filters go much beyond a simple distance-based weighting of contributions, but this basic criterion remains a founding paradigm of local filters.

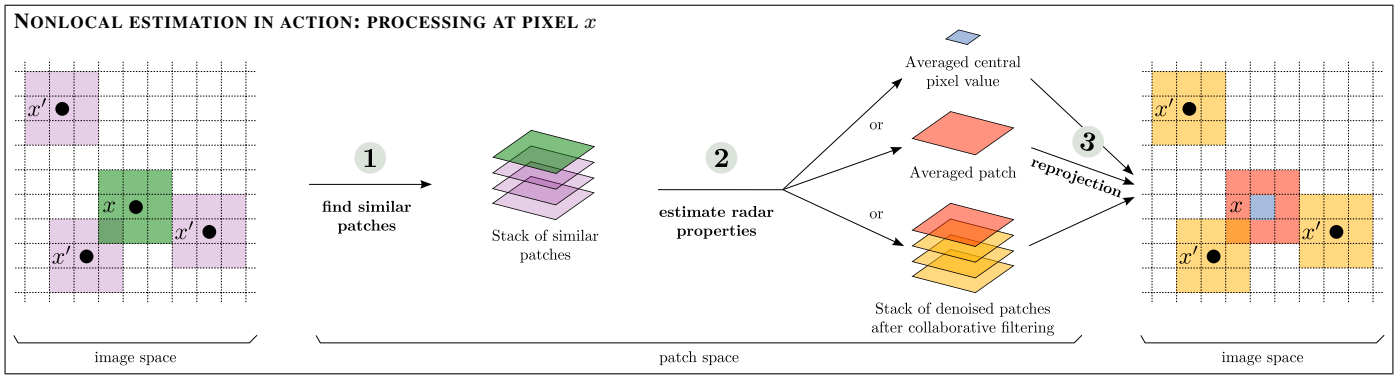
The patch-based nonlocal approach avoids the potentially dangerous identification between closeness and similarity and goes back to the original problem, trying to identify the pixels more similar to the target, irrespective (to a certain extent) of their spatial distance from it. In next Section we explore in more depth the fundamental steps involved in nonlocal SAR despeckling and review the current state of the art.

III. NONLOCAL APPROACHES: EXPLOITING PATCH REDUNDANCY

At the core of the nonlocal approach stands the selection of suitable predictors based on their similarity with the target.

This idea began to gain some popularity with the bilateral filter, proposed for additive white Gaussian noise (AWGN) denoising, with predictor weights depending not only on their spatial distance from the target, but also on their similarity with it, measured by the difference between observed values. Despite its simplicity, this filter gave a surprisingly good denoising performance, reducing the annoying edge smearing phenomenon. However, the pixel-wise estimation of similarity was very rough, and happened to reinforce observed values affected by strong noise, justifying the need for the spatial-distance term. The fundamental step towards nonlocal filtering was then the introduction, in NL-means [8], of a *reliable* patch-wise measure of similarity. When a relatively large patch is taken into account, it is very unlikely that pixels characterized by a signal much different from the target be accepted as good predictors. Thanks to the improved reliability, predictors can be weighted based *only* on similarity, with no reference to spatial information. Unlike local filters, non-local ones do not impose any specific structure (connectivity, shape) or smoothness, but only exploit that patches recur more or less frequently, a *self-similarity* property common to most images. Recurring patches are easily found in smooth regions, but just as well around region boundaries, textures, artificial structures, etc., as shown in Fig.1. Hence, for most patches, several other patches can be found with similar content. This form of stationarity in the space of patches is central to nonlocal approaches.

Nonlocal estimation methods generally follow a three-step scheme, summarized in the box "Nonlocal estimation in action", with many possible variations at each step and, possibly, pre-processing steps and/or iterative refinement of results by repeated nonlocal estimations. The first step identifies similar patches (patch size is generally set from 3×3 to 11×11 pixels). It must reliably find, within an extended search window (typically 21×21 to 39×39), patches that are close to the reference central patch. Once several patches have been selected (from a few tens to all the hundreds of patches within



the search window), they are assigned relative weights. The second step combines patches, according to their weights, to form an estimate of either the central pixel, the central patch, or all selected patches. The estimates computed from all possible reference patches are then merged in a last step to produce the final image.

We now illustrate these steps for the special case of NL-means [8], indicating by $v(x)$ and $\underline{\boldsymbol{v}}(x)$, respectively, the value at pixel x, and the image patch centered on x in the observed image. In NL-means, each pixel x' in a large search area around the target pixel x is considered and a similarity measure $\Delta(x, x')$ is computed by comparing the two patches centered on x and x'. Under AWGN, the sum of squared differences is a natural criterion to evaluate similarity:

$$\Delta(x, x') = \|\underline{\boldsymbol{v}}(x') - \underline{\boldsymbol{v}}(x)\|^2 \,. \tag{6}$$

This measure is used to compute the weight for each predictor pixel x', with large weights associated with similar patches and negligible weights with dissimilar ones. An exponential kernel is used to this end:

$$w(x, x') = \exp\left(-\Delta(x, x')/h\right) \,, \tag{7}$$

with the bandwidth parameter h governing the weight distribution. Finally, the target pixel x is estimated through the weighted average of all pixels within the search area:

$$\widehat{u}^{\mathrm{NL}}(x) = \frac{\sum_{x'} w(x, x') v(x')}{\sum_{x'} w(x, x')} \,, \tag{8}$$

where $\widehat{u}$ denotes an estimate of the unknown signal value u.

In the NL-means, therefore, only one estimate is obtained for each pixel x, which corresponds to the top branch depicted in the box "Nonlocal estimation in action". Generalizations provide several estimates for each pixel that must be properly combined in the last step.

The NL-means filter and its numerous variants are known to perform well under AWGN. However, the extension to SAR imaging, and in particular to speckle, is by no means trivial,

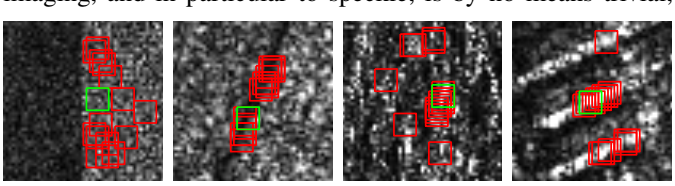


Fig. 1. Typical fragments of SAR images (51×51 pixels), from the left: a homogeneous region, a line, a texture and a structure. For each target patch (green) several similar 8×8 patches (red) are easily found in the same fragment.

and has been the object of several recent papers. One has to understand the foundation of such filters, and provide a flexible formulation suitable for different modalities and models of noise. In the following, we focus on the main concepts that have been proposed in SAR imaging, and try to provide some insight into the problems that emerge when noise departs from the Gaussian distribution, and possible solutions to them. The key point of all these extensions is the consideration of the specific distributions of radar data.

Step 1: Defining patch similarities – The first step of nonlocal estimation methods is the identification of similar patches through a (dis)similarity criterion $\Delta(x, x')$. This criterion quantifies, in a principled manner, by how much the unknown patches $\underline{\boldsymbol{u}}(x)$ and $\underline{\boldsymbol{u}}(x')$ differ. Based on the similarity to the reference patch at x, predictor patches at x' can be either *soft-assigned* or *hard-assigned* to the set of similar patches. In the case of soft-assignment, a weight reflecting the level of similarity is associated to each patch within the search area, otherwise the most similar patches are included and used to perform the prediction.

Under AWG noise, (6) is a natural criterion to evaluate similarity between two patches. Fluctuations created by speckle are multiplicative and non-Gaussian, see the box "Speckle fluctuations in radar images". Specific criteria must be derived for the comparison of patches in SAR imaging.

The similarity between two patches is generally defined as the sum of the similarity δ of each pair of corresponding pixels in the two noisy patches: $\Delta(x, x') = \sum_\tau \delta(v(x+\tau), v(x'+\tau))$. In order to improve the discrimination power of the similarity criterion, several authors suggest using a pre-estimate $\widetilde{u}$ computed either over the whole patch [20], obtained after a first iteration [11] or at the previous iteration [9] of the nonlocal method. The similarity then takes the form: $\Delta(x, x') = \sum_\tau \delta(\widetilde{u}(x+\tau), \widetilde{u}(x'+\tau))$. The following paragraphs describe possible approaches to derive the pixel-wise similarity δ.

Detection approach – Dissimilarity can be defined based on the detection of the difference between the underlying values u_1 and u_2 [21]. This detection problem can be formulated as an hypothesis test where the null hypothesis corresponds to no difference (H_0: $u_1 = u_2 = u_{12}$) and the alternative one to a difference (H_1: $u_1 \neq u_2$). Among several criteria considered in [21], the generalized likelihood ratio (GLR) is shown to

perform best and fullfil several invariance properties:

$$\delta^{\text{GLR}}(v_1, v_2) = -\log \frac{\sup_{u_{12}}[\,p(v_1|u_{12})\,p(v_2|u_{12})\,]}{\sup_{u_1}[\,p(v_1|u_1)\,]\,\sup_{u_2}[\,p(v_2|u_2)\,]}\,. \quad (9)$$

Criteria specific to SAR imaging can be derived from (9) by using the statistical speckle models recalled in the box "Speckle fluctuations in radar images". Both single channel (intensity) and multi-channel (polarimetric and/or interferometric) configurations lead to a criterion defined as the log of the ratio between arithmetic and geometric means [22]:

$$\delta^{\text{GLR}}(\boldsymbol{C}_1, \boldsymbol{C}_2) = 2L \log \frac{\frac{1}{2}|\boldsymbol{C}_1 + \boldsymbol{C}_2|}{\sqrt{|\boldsymbol{C}_1| \cdot |\boldsymbol{C}_2|}}\,, \quad (10)$$

where $\boldsymbol{C}$ may represent either the L-looks intensity (single-channel images) or the L-looks empirical covariance matrices (multi-channel images). This is a natural criterion to compare patches corrupted by a multiplicative noise such as speckle since the criterion is invariant to a multiplicative change of contrast. It has been successfully used in nonlocal filtering of intensity, interferometric and polarimetric SAR images [12], [14]. Detection-based criteria using similar expressions are also at the heart of [9], [13], [23], [24].

Information approach – In line with the detection approach, the authors of [9], [20] consider a similar hypothesis test involving pre-estimates of the parameters. Good criteria to perform such a hypothesis test are provided by the h-ϕ divergences where h and ϕ refer to pre-defined functions. The h-ϕ divergences measure the quantity of information shared by the distributions parametrized by $\widetilde{u}_1$ and $\widetilde{u}_2$: they evaluate the proportion of samples from one distribution that can be explained by the other. Specific choices of h and ϕ lead for instance to the Hellinger divergence or the Kullback-Leibler divergence. Again, taking into account specific SAR distributions provides well-founded criteria. The symmetrical version of Kullback-Leibler divergence (sKL) gives, in the case of Wishart-distributed empirical covariance matrices:

$$\delta^{\text{sKL}}(\widetilde{\boldsymbol{\Sigma}}_1, \widetilde{\boldsymbol{\Sigma}}_2) = L\,\text{Tr}(\widetilde{\boldsymbol{\Sigma}}_1^{-1}\widetilde{\boldsymbol{\Sigma}}_2 + \widetilde{\boldsymbol{\Sigma}}_2^{-1}\widetilde{\boldsymbol{\Sigma}}_1) + \text{const.} \quad (11)$$

again involving ratios, as customary in SAR imagery.

Geometric approach – The similarity can be defined by deriving a metric suitable to the specificities of SAR data. D'Hondt *et al* [25] suggest using Hermitian semi-definite positive matrices and propose a metric connected to geodesic and Riemannian distances:

$$\delta^{\text{Geo}}(\widetilde{\boldsymbol{\Sigma}}_1, \widetilde{\boldsymbol{\Sigma}}_2) = \|\log(\widetilde{\boldsymbol{\Sigma}}_2^{-1/2}\widetilde{\boldsymbol{\Sigma}}_1\widetilde{\boldsymbol{\Sigma}}_2^{-1/2})\|_F^2 \quad (12)$$

where $\log$ is the matrix logarithm and $\|\cdot\|_F$ the Froebenius norm. This similarity offers some interesting invariance properties. In the case of intensity images with pre-estimated radiometry $\widetilde{R}$, this criterion boils down to $(\log(\widetilde{R}_2) - \log(\widetilde{R}_1))^2$ which is the square distance between observations after applying a homomorphic transform on the pre-estimates. It has been shown in the framework of detection theory that criteria based on variance-stabilization (such as the log transform in SAR) enjoy good properties [21]. It is interesting to see that such a criterion can be derived using different interpretations (a geometric founding or a statistical reasoning).

Estimation approach – Lee *et al* [26] have shown that the estimation of radar properties may suffer from a systematic bias arising from the procedure that detects similar noisy values (hard-assignment). For SAR intensity images, they suggest using a pre-selection rule of the form $I_2 \in [R_1\zeta, R_1\zeta']$ – called *sigma range* – for which the pre-selected samples I_2 do not introduce bias in the subsequent estimation when the radiometry is identical ($R_1 = R_2$). Values of ζ and ζ' are computed by an iterative method, while the unknown R_1 is replaced by a pre-estimate $\widetilde{R}_1$ called *a priori* mean [26], [10]. By rather considering a pre-selection rule of the form $\delta(I_1, I_2) \leq \gamma$, numerical integrations over I_1 and I_2 show that the subsequent estimation is unbiased as soon as the rule can be rewritten as $I_2 \in [I_1/\zeta, I_1\zeta]$. Detection rules such as GLR (10) verify this property.

Some of the estimators used in the second step of nonlocal estimation methods use the similarity $\Delta(x, x')$ in order to weight the importance of the patch $\underline{\boldsymbol{v}}(x')$ (i.e., soft-assigment). The similarity $\Delta(x, x')$ is mapped into a weight $w(x, x')$ using a function termed *kernel*. Many different kernels have been proposed in the literature, from simple thresholding $w(x, x') = \mathbb{1}[\Delta(x, x') < h]$, exponential kernels as in (7), to more sophisticated ones [20], [24], [14]. The shape of the kernel (*e.g.*, smooth, discontinuous or trapezoidal) changes the contributions of patches that may correspond to false detections and thus controls the bias / variance trade-off. In [20], [14], suitable kernels have been defined to guarantee the same bias / variance trade-off, irrespective of the modality, the noise statistic or the patch size.

Step 2: Estimation of radar properties – After selection of a stack of patches and/or computation of relative weights during the first step of the nonlocal estimation method, these patches are combined in a second step to form an estimate of the radar properties. This combination can be a simple (weighted) averaging as in the NL-means, or a more evolved estimator.

In SAR imaging, observations $\boldsymbol{v}$ may denote a collection of intensities I, amplitudes $\sqrt{I}$ or log-transformed values $\log(I)$, a vector of noisy coefficients in a transformed domain (DCT, Fourier, wavelets…), a collection of scattering vectors $\boldsymbol{k}$ or of empirical covariance matrices $\boldsymbol{C}$. The associated collection of parameters of interest $\boldsymbol{u}$, are generally the radiometry R or covariance matrice $\boldsymbol{\Sigma}$. The estimation step computes radar parameters from the collection of observations $\boldsymbol{v}$ and/or pre-estimations $\widetilde{\boldsymbol{u}}$ gathered during the first step. The most common estimators for patch-based denoising of SAR images are:

- SME/WSME – NL-means and many of its successors combine similar patches into a weighted average (weighted sample mean estimator, WSME), where the weights are derived from the similarities: see equation (8). By using weights, the estimation relies more heavily on samples that are more similar (thus, more reliable), which reduces bias. Compared to an estimation based solely on a collection of patches detected as highly similar (i.e., by hard-assignment), the variance can also be reduced by considering a larger number of samples.

The sample mean estimator corresponds to the conditional expectation $\mathbb{E}[\boldsymbol{v}|\boldsymbol{u}]$. This expectation may not directly be equal to the parameter of interest $\boldsymbol{u}$, but require a debiasing step. A

notable example is the case of (weighted) averaging of log-transformed SAR images.

The estimation of covariance matrices from empirical ones can be performed by WSME [12], [25], [20], [24], [14]:

$$\widehat{\boldsymbol{\Sigma}}^{\text{WSME}}(x) = \frac{\sum_{x'} w(x,x')\boldsymbol{C}(x')}{\sum_{x'} w(x,x')}. \tag{13}$$

with $w(x,x')$ the weight depending on the similarity $\Delta(x,x')$ between patches extracted at pixel locations x and x'.

• MLE/WMLE – The sample mean is not a correct estimator of some parameters, e.g., the amplitude $\sqrt{I}$. The sample mean should then be replaced by an estimator adjusted to the specific noise distribution of the observations, such as the maximum likelihood estimator. Similarly to the WSME, sample weights can be introduced in the estimation method. Maximum likelihood can be generalized into the weighted maximum likelihood estimator (WMLE) [9]:

$$\widehat{u}^{\text{WMLE}}(x) = \arg\max_{u} \sum_{x'} w(x,x') \log p[v(x')|u]\,. \tag{14}$$

WMLE leads to (13) for Wishart-distributed covariance matrices, but improvements are obtained for matrices with particular structures as in interferometric SAR imaging [13].

• MMSE – Kervrann *et al.* proposed a method called Bayesian NL-means [27] that estimates the parameters $\boldsymbol{u}$ as a linear combination of pre-estimated patches $\widetilde{\boldsymbol{u}}$, with weights defined by the likelihood of each pre-estimated patch with respect to the observation $\boldsymbol{v}$:

$$\widehat{\boldsymbol{u}}^{\text{MMSE}}(x) \approx \frac{\sum_{x'} p[\boldsymbol{v}(x)|\widetilde{\boldsymbol{u}}(x')]\,\widetilde{\boldsymbol{u}}(x')}{\sum_{x'} p[\boldsymbol{v}(x)|\widetilde{\boldsymbol{u}}(x')]} \tag{15}$$

This estimator can be interpreted as an approximation of the minimum mean square error estimator, i.e., the posterior mean. Pre-estimates are usually obtained thanks to a two-step or iterative pre-filtering [10]. Since the pre-estimation step provides only a coarse estimation $\widetilde{\boldsymbol{u}}$ of the patches $\boldsymbol{u}$, a smoothing parameter is introduced to reduce the selectivity of the likelihood function and thus avoid weighting too much patches that are very close to the observation [27].

• LMMSE – The linear minimum mean square error estimator has long been used in SAR imaging [26]. This estimator restricts the form of the solution to linear transforms of $\boldsymbol{v}$, which is most efficient when noise and signal are well separated. In SAR-BM3D [11], the LMMSE is computed after an undecimated discrete wavelet transform (UDWT) is applied to the stack of similar patches. Expectation and variance of observations and parameters can be obtained from band-wise statistics and later refined using pre-filtered patches after a first restoration has been performed.

The family of homomorphic approaches transform the multiplicative noise into an additive one by taking the logarithm of the observed intensity / amplitude. Several papers have derived nonlocal estimators in this sense, *e.g.*, [28] where an "adjust mean" step is used to deal correctly with the bias arising from the Gaussian assumption (i.e., debiasing step described for SME). Note that estimators that process directly SAR data (i.e., without log-transforms) are often preferable since the debiasing step is then unnecessary.

Step 3: Reprojection to image space – The second step of nonlocal methods provides estimates either for a single pixel (pixel-wise estimation), for a single patch (patch-wise estimation), or for the whole stack of patches (stack-wise estimation). The first option corresponds to basic NL means, already described above, so let us focus on the other two strategies.

The difference in patch-wise filtering is that all pixels in the patch, not just the central one, are estimated at once. Since each pixel is estimated several times, a suitable aggregation phase is necessary to combine all such estimates. In particular, we need to define a reprojection function, $g(\cdot)$, to get an estimate in the pixel domain: $\widehat{u}(x) = g(\underline{\widehat{\boldsymbol{u}}}_1(x), \ldots, \underline{\widehat{\boldsymbol{u}}}_K(x))\,.$ What is more important, these estimates refer to different patches, with different reliability levels, an information that can be exploited to improve results. The reprojection can be performed through a weighted average of the K estimators:

$$\widehat{u}(x) = \sum_{k=1}^{K} \alpha_k\, \underline{\widehat{\boldsymbol{u}}}_k(x)\,. \tag{16}$$

The simplest form of aggregation is to consider uniform weights α_k, as done in the blockwise NL-means [8]. Another strategy is to set the weight associated with each estimate as inversely proportional to its variance [29].

To illustrate why patch-wise estimation improves performance, let us consider the special case of a pixel near the boundary between two homogeneous regions. Since the patch centered on it is strongly heterogeneous, most other patches of the search area, coming from homogeneous regions on either side of the boundary, are markedly dissimilar from it, and contribute very little to the average. The estimate, thus, involves only a small effective number of predictors, those along the edge, which results in a high variance. As a result, a visible "halo" of residual noise is observed near edges, a phenomenon well-known in NL-means, also referred to as the rare patch effect. The target pixel, however, belongs to a large number of patches, not just the patch *centered* on it, many of them drawn from the homogeneous region the pixel belongs to. In patch-wise re-projection *all* these patches are included in the average reducing the estimate variance, especially if suitable weights are used to take into account the reliability of each contribution.

Let us now consider the third strategy, with stack-wise filtering. A first difference w.r.t. patch-wise filtering is that now all patches collected in the stack are *collaboratively* filtered before reprojecting them to their original position. The major improvement is that the stack is filtered in three dimensions, that is, not only along the stack but also in the spatial domain. In BM3D [29] and, with necessary adjustments to the SAR domain, in SAR-BM3D [11], the whole stack, formed by just a limited number of similar patches, is wavelet transformed, Wiener filtered, and back transformed. By so doing, strong spatial structures are emphasized through filtering while random noise is efficiently suppressed. As a matter of fact, these techniques exhibit significant improvements especially in highly structured areas (edges, point reflectors, textures). The efficiency of collaborative filtering comes from the full

Fig. 2. Single-look TerraSAR-X image of Barcelona (image courtesy ©Infoterra GmbH) and the radiometry estimated by the nonlocal method [11]. Images are 511×1043 pixels with a spatial resolution of 2.5 m.

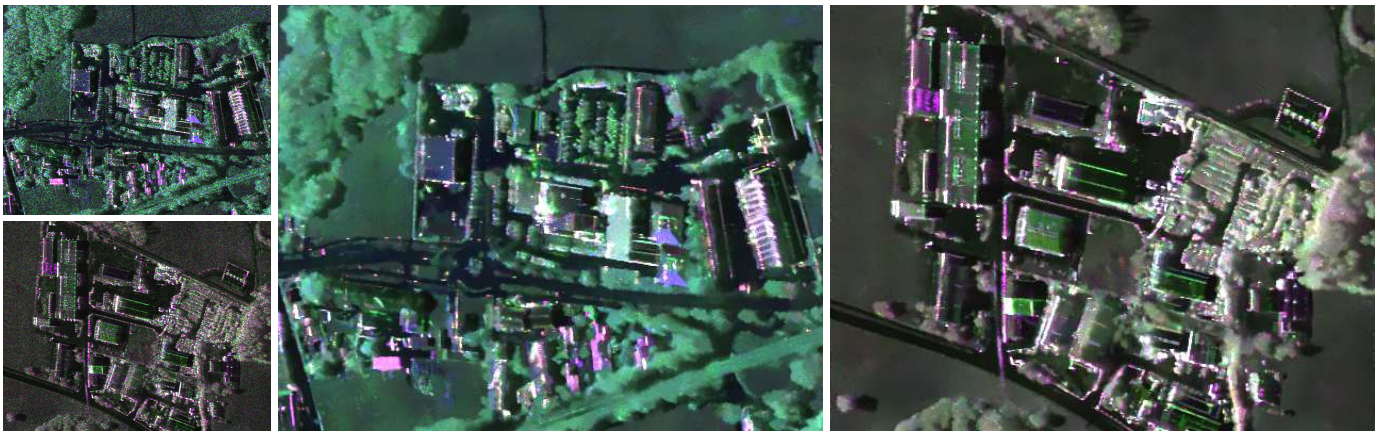


Fig. 3. On the left, two PolSAR images of the area of Kaufbeuren (Germany) sensed by F-SAR, single-look, X-band (top), S-band (bottom) (image courtesy ©DLR). On the center and right images, estimated polarimetric signatures using the nonlocal patch-based method [14]. Images are 400×512 pixels with a spatial resolution of 0.5 m.

exploitation of the redundancy of information in a stack of similar patches.

The box titled "Overview of nonlocal estimation methods in SAR imaging" gives a synthetic view of some of the main methods developed for nonlocal estimation in SAR imaging. Needless to say, performance depends on the setting of several parameters, like patch size and search area size, which should be related to image resolution, smoothing strength, and balance between original and pre-estimated data. In most of the nonlocal approaches these parameters are set by hand. Few works have considered semi-supervised setting or automatic setting with spatial adaptation [14].

Some sample experimental results that confirm the potential of nonlocal methods are presented in figure 2 for an amplitude image with SAR-BM3D [11] and in figure 3 for a polarimetric data with NL-SAR [14]. Visual inspection shows the performance of the approach in reducing strong fluctuations while preserving important features like targets and lines. Note that no systematic artefacts should be observed with these approaches unless the parameters are not set properly (*e*.g., no consideration of over-sampled data). A rigorous performance evaluation of nonlocal despeckling techniques is beyond the scope of this paper. Some frameworks for SAR despeckling evaluation and comparison are proposed in [30] and [1].

IV. Open issues and future trends

Patch-based approaches provide the best performance to-date for speckle reduction in SAR intensity images [30] and polarimetric or interferometric SAR images [14]. Beyond speckle suppression, they improve parameter estimation, drastically enhancing radar measures. Therefore, they can be expected to have a strong impact, in the near future, on major applications of radar imaging, improving the biomass estimation with polarimetric - interferometric data; increasing the spatial resolution in urban monitoring with radar tomography; enabling more reliable phase unwrapping methods for interferometric SAR.

As a by-product of patch-based speckle reduction methods, several similarity criteria especially suited for SAR imaging have been established. These criteria are central to many applications: for classification and indexing using patch clustering; for change detection; for movement monitoring by patch tracking.

Patch-based methods are at their beginning and many open issues have yet to be solved. The speckle model consider so far (simple complex Gaussian) is known to inaccurately describe very high resolution images or textured areas. Introduction of more accurate models (e.g., Weibull, Fisher or generalized gamma distributions for amplitude, and corresponding matrix-variate polarimetric distributions) raises questions about the increased complexity of estimators and the possible loss of robustness with the increase of degrees of freedom. Another limit is the geometric deformations appearing on SAR images with elevated objects. Such deformations should be considered to perform joint restoration of images taken from multiple incidence angles.

OVERVIEW OF NONLOCAL ESTIMATION METHODS IN SAR IMAGING

Nonlocal estimation methods perform the three steps depicted in the box "Nonlocal estimation in action". A great variety of methods dedicated to SAR images have been proposed over the last few years. These methods follow different paths to implement each of the three steps:

1) similar patches identification (choice of the similarity criterion);
2) estimation of radar properties (choice of the estimator);
3) reprojection of estimates onto the image space (choice of the domain).

Below an overview of some of the main methods devoted to SAR imaging.

Method (oldest-first)	**Domain**	**Estimator**	**SAR modality**	**Self-similarity domain**	**Similarity criterion**	**Scheme**
PPB / NL-InSAR	Pixel-wise	WMLE	SAR (any L) / InSAR ($L=1$)	Patch-wise	Detection + Information	Iterative
[9], [13]	*Originality:* iteratively refines the weights by comparing patches of previous estimates and patches of the noisy image.					
Pretest NLM	Pixel-wise	WSME	PolSAR ($L \geq 3$)	Patch-wise	Detection	One step
[12]	*Originality:* direct extension of [8] with a selection based on a GLR (referred to as pretest step).					
Bayesian NLM	Pixel-wise	MMSE	SAR (any L)	Patch-wise	Estimation	One step
[10]	*Originality: prior* patches are extracted in a multi-looked image. A sigma-range pre-selection is used and darker pixels are discarded.					
SAR-BM3D	Stack-wise	LMMSE	SAR (any L)	Patch-wise	Detection + Information	Two steps
[11]	*Originality:* works on the UDWT of stacks: 1st step uses statistics of each subband; 2nd one uses statistics provided by the 1st iteration.					
Bilateral NLM	Pixel-wise	WSME	PolSAR (any L)	Pixel-wise	Geometric	Iterative
[25]	*Originality:* images are multi-looked, then iteratively updated by nonlocal averaging based on geometrical comparison of pixel values (no patch).					
Stochastic NLM	Pixel-wise	WSME	PolSAR (non-stationnary L)	Stats-wise	Information	One step
[20]	*Originality:* lets L vary in the image. Patch similarity is based on the divergence between MLE estimates of $\mathbf{\Sigma}$ and L within the patch.					
Discriminative NLM	Pixel-wise	WSME	PolSAR	Patch-wise	Detection	Iterative
[24]	*Originality:* iteratively refines weights based on the ratio of diagonal elements of empirical covariance matrices and the span of the previous iterate.					
NL-SAR	Pixel-wise	WMLE	(Pol)(In)SAR (any L)	Patch-wise	Detection	Adaptive
[14]	*Originality:* fully automatic: patch sizes, search windows and pre-filtering strenghts are spatially tuned to provide improved results.					

REFERENCES

[1] F. Argenti, A. Lapini, L. Alparone, and T. Bianchi. A tutorial on Speckle Reduction in Synthetic Aperture Radar Images. *IEEE Geosci. Remote Sens. Mag.*, pages 6–35, 2013.

[2] R. Touzi. A review of speckle filtering in the context of estimation theory. *IEEE Trans. Geosci. Remote Sens.*, 40(11):2392–2404, 2002.

[3] J.S. Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2(2), 1980.

[4] L. Denis, F. Tupin, J. Darbon, and M. Sigelle. Joint Regularization of Phase and Amplitude of InSAR Data: Application to 3D reconstruction. *IEEE Trans. Geosci. Remote Sens.*, 2009.

[5] A. Achim, P. Tsakalides, and A. Bezerianos. SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling. *IEEE Trans. Geosci. Remote Sens.*, 41(8), 2003.

[6] G. Vasile, E. Trouvé, and J. Lee. Intensity-Driven Adaptive-Neighborhood Technique for Polarimetric and Interferometric SAR Parameters Estimation. *IEEE Trans. Geosci. Remote Sens.*, 44(6), 2006.

[7] J.S. Lee. Digital image smoothing and the sigma filter. *Computer Vision, Graphics, and Image Processing*, 24(2):255–269, 1983.

[8] A. Buades, B. Coll, and J. M. Morel. A Review of Image Denoising Algorithms, with a New One. *Multiscale Mod. & Simu.*, 4(2), 2005.

[9] C.-A. Deledalle, L. Denis, and F. Tupin. Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. *IEEE Trans. Image Process.*, 18(12), 2009.

[10] H. Zhong, Y. Li, and L. Jiao. SAR Image Despeckling Using Bayesian Nonlocal Means Filter With Sigma Preselection. *IEEE Geosci. Remote Sens. Lett.*, 8(4), 2011.

[11] S. Parrilli, M. Poderico, C.V. Angelino, and L. Verdoliva. A Nonlocal SAR Image Denoising Algorithm Based on LLMMSE Wavelet Shrinkage. *IEEE Trans. Geosci. Remote Sens.*, 50(2), 2012.

[12] J. Chen, Y. Chen, W. An, Y. Cui, and J. Yang. Nonlocal filtering for polarimetric SAR data: A pretest approach. *IEEE Trans. Geosci. Remote Sens.*, 49(5), 2011.

[13] C.-A. Deledalle, L. Denis, and F. Tupin. NL-InSAR: Nonlocal interferogram estimation. *IEEE Trans. Geosci. Remote Sens.*, 49(4), 2011.

[14] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigbers, and M. Jäger. NL-SAR: a unified Non-Local framework for resolution-preserving (Pol)(In)SAR denoising. Preprint HAL – hal-00844118, 2013.

[15] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou. A tutorial on Synthetic Aperture Radar. *IEEE Geosci. Remote Sens. Mag.*, pages 6–43, 2013.

[16] A. Lopes, R. Touzi, and E. Nezry. Adaptative speckle filters and scene heterogeneity. *IEEE Trans. Geosci. Remote Sens.*, 28(6), 1990.

[17] F. Argenti, T. Bianchi, and L. Alparone. Multiresolution MAP Despeckling of SAR Images Based on Locally Adaptive Generalized Gaussian pdf Modeling. *IEEE Trans. Image Process.*, 15(11):3385–3399, 2006.

[18] H. Xie, L.E. Pierce, and F.T. Ulaby. SAR speckle reduction using wavelet denoising and Markov random field modeling. *IEEE Trans. Geosci. Remote Sens.*, 40(10):2196–2212, January 2002.

[19] S Foucher, G B Bénié, and J M Boucher. Multiscale MAP filtering of SAR images. *IEEE Trans. Image Process.*, 10(1):49–60, January 2001.

[20] L. Torres, S.J.S. Sant'Anna, C. da Costa Freitas, and A.C. Frery. Speckle reduction in polarimetric SAR imagery with stochastic distances and nonlocal means. *Pattern Recognition*, in press, 2013.

[21] C. Deledalle, L. Denis, and F. Tupin. How to compare noisy patches? Patch similarity beyond Gaussian noise. *International Journal of Computer Vision*, 99(1), 2012.

[22] K. Conradsen, A. Aasbjerg Nielsen, J. Schou, and H. Skriver. A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.*, 41(1):4–19, 2003.

[23] G. Poggi, F. Sica, L. Verdoliva, G. Fornaro, D. Reale, and S. Verde. Nonlocal methods for filtering interferometric SAR datasets. In *Proceedings of TyWRRS: From Earth Observation to Homeland Security*, 2012.

[24] G. Liu and H. Zhong. Nonlocal Means Filter for Polarimetric SAR Data Despeckling Based on Discriminative Similarity Measure. *IEEE Geosci. Remote Sens. Lett.*, PP(99), 2013.

[25] O. D'Hondt, S. Guillaso, and O. Hellwich. Iterative Bilateral Filtering of Polarimetric SAR Data. *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, 6(3):1628–1639, 2013.

[26] J.-S. Lee, J.-H. Wen, T. L Ainsworth, K.-S. Chen, and A. J Chen. Improved sigma filter for speckle filtering of SAR imagery. *IEEE Trans. Geosci. Remote Sens.*, 47(1):202–213, 2009.

[27] C. Kervrann, J. Boulanger, and P. Coupé. Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal. In *Scale Space and Variational Methods in Computer Vision*, pages 520–532. Springer, 2007.

[28] M. Makitalo, A. Foi, D. Fevralev, and V. Lukin. Denoising of single-look SAR images based on variance stabilization and nonlocal filters. In *International Conference on Mathematical Methods in Electromagnetic Theory*. IEEE, 2010.

[29] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.*, 16(8), 2007.

[30] G. Di Martino, M. Poderico, G. Poggi, D. Riccio, and L. Verdoliva. Benchmarking framework for SAR despeckling. *IEEE Trans. Geosci. Remote Sens.*, 52(3), 2014.