ABSTRACT

We propose a new approach to SAR despeckling, based on the combination of multiple alternative estimates of the same data. The many despeckling methods proposed in the literature possess different and often complementary strengths and weaknesses. Given a reliable pixel-wise classification of the image, one can take advantage of this diversity by selecting the more appropriate combination of estimators for each image region. We implement a simplified version of this approach, using soft classification and two state-of-the-art despeckling tools, with opposite properties, as basic estimators. Experiments on real-world high-resolution SAR images prove the effectiveness of the proposed technique and confirm the potential of the whole approach.

Index Terms— SAR despeckling, soft classification, nonlocal techniques.

1. INTRODUCTION

A huge number of remote-sensing images of the Earth are available nowadays, but extracting useful information from such data is not easy. SAR (Synthetic Aperture Radar) images, in particular, are corrupted by intense speckle noise which makes it difficult to recover the data of interest. Therefore, despeckling is very often a necessary step before undertaking higher-level data processing tasks. Unfortunately, it is also a very challenging problem especially for single-look products, which guarantee high spatial resolution but exhibit highest speckle intensity. A SAR despeckling technique should pursue different goals with reference to different image features, in particular it should (i) suppress speckle as much as possible in homogeneous regions; (ii) preserve textures, region boundaries, and other linear structures (roads, waterways, etc.); (iii) avoid filtering natural or man-made strong permanent scatterers; (iv) avoid introducing artifacts. It is utterly implausible that a single despeckling filter can satisfy all these requirements, since all filters are based on some necessarily simplified model of the SAR image.

Indeed, the trade-off between speckle suppression and detail preservation is well-known in the SAR despeckling literature. Several efforts have been undertaken to adapt filtering to local image features or, more ambitiously, to classify pixels in advance so as to use different models and methods for the different areas to be processed. In [1], for example, the area to be filtered is subject to an explicit heterogeneity test based on the coefficient of variation: only pixels belonging to homogeneous areas are then filtered through a MAP approach. This procedure can also be adopted in the wavelet domain by a proper measure of texture energy in order to perform an ad hoc MAP estimate [2].

Recently nonlocal approaches have gained much popularity in SAR despeckling [3]. Despite their inherent adaptivity, they present the same dichotomy of their predecessors [4]. Some filters have strong despeckling capacity [5] and other ensure good detail preservation [6]. To address this problem some forms of adaptivity have been proposed in the literature. In [7] for example a local selection method of the best parameters is inglobed in the denoising algorithm. Also in [8], based on the estimate of the local image activity, the main algorithm parameters were tuned to better fit the characteristics of the area under analysis.

In this paper we follow a different approach to deal with such issues: rather than adapting a single technique to the varying local image statistics, we combine several of them, with suitable adaptive weights. We follow the idea proposed in [9], where two different filters are used to handle homogeneous and heterogeneous areas, and where a stack-wise processing is proposed to improve the classification map that guides the whole process. However, a major difference is in the combination phase, where we depart from the sharp two-class model, and combine linearly the output of both engines with weights varying in the whole [0,1] interval. Given the very low signal-to-noise ratio characterizing our source images, the use of multiple estimators is very likely to reduce the variance of the overall estimate. A solid argument in support of this thesis is the success and widespread use of the collaborative filtering, which is a fundamental ingredient of many state-of-the-art tools.

2. PROPOSED DESPECKLING STRATEGY

We consider the general framework described in Fig.1(a). Several despeckling tools are supposed to be available, and we filter the input image by means of all of them. Therefore, we obtain a certain number of alternative estimates for each pixel of the image. In parallel, the input image is also analyzed in order to extract descriptive point-wise features,
which are then fed to a classifier. This tool instructs the combiner on the best way to take advantage of the various estimates in order to provide the final despeckled image. Assuming that state-of-the-art despeckling tools have been chosen, with different strengths and weaknesses, the core of the method is the combination strategy. This is not a trivial task, and work is currently under way to adapt some machine learning tools.

Here, we focus on a simplified implementation of this approach where a single feature is used to measure the pixel homogeneity with the neighborhood, and hence to modulate the combination of the two state-of-the-art filters. We consider the SAR image as composed by just two types of regions, homogeneous: characterized by uniform electrical and geometrical properties; and heterogeneous: including all remaining areas, such as boundaries, textures, man-made objects, and so on. By so doing, we cluster more specific classes in larger groups characterized by common needs.

2.1. Soft Classification

For single-look SAR images, even the simple homogeneous/heterogeneous classification turns out to be quite challenging, and it has received considerable attention in the SAR literature, with the coefficient of variation [10] and the ratio, $A/G$, of the arithmetic to geometric means [11] as the most popular statistics used to this end. Both statistics, although associated with a single target pixel, are computed on a suitably large window centered on the target and sliding over the image, so as to reduce the disruptive effects of noise.

In particular, the $A/G$ ratio arises readily, under some simplifying hypotheses, as the solution of a Generalized Likelihood Ratio (GLR) test. We consider the amplitude model $x = u\phi$, where the observed pixel value, $x$, is the product of an underlying “clean” signal, $u$, and a Nakagami distributed speckle sample $\phi$

$$p_\Phi(\phi) = \frac{2L^L}{\Gamma(L)} \phi^{2L-1} e^{-L \phi^2} u(\phi)$$  \hspace{1cm} (1)

where $\Gamma(\cdot)$ is the Eulerian Gamma function and $u(\cdot)$ the step function.

The test is carried out on a block of $N$ pixels centered on the target. Given the two hypotheses

$$\begin{align*}
H_0: & \quad u_1 = u_2 = \cdots = u_N = u \\
H_1: & \quad 1 - H_0
\end{align*}$$  \hspace{1cm} (2)

that all signal amplitudes are equal ($H_0$), or else that they are not ($H_1$), the generalized likelihood ratio reads as

$$\Lambda(x) = \sup_{u_1,\ldots,u_N} \left\{ \frac{f_{X_1|u_1}(x_1) \cdots f_{X_N|u_N}(x_N)}{\sup_u \{f_{X_1|u}(x_1) \cdots f_{X_N|u}(x_N)\}} \right\}$$  \hspace{1cm} (3)

With the given speckle model, the GLR statistic in logarithmic form reads eventually as

$$\lambda(x) = NL \ln \left( \frac{A}{G} \right)$$  \hspace{1cm} (4)

where $G = \sqrt[N]{\prod_{i=1}^{N} x_i^2}$ and $A = \sum_{i=1}^{N} x_i^2/N$ are the geometric and arithmetic means of the observed intensities.

Despite this elegant formulation, classification results based on the $A/G$ ratio are often disappointing and depend strongly on the window size $N$. Using a small window, the decision statistic is very noisy, even in perfectly homogeneous regions. On the other hand, with a larger window there is a marked loss of resolution, and isolated outliers may cause blockiness. To tackle these problems we adopt stack-wise processing, that we found to be a good solution in our previous work [9]. The idea is inspired by nonlocal filtering. In fact, when estimating pixel statistics on the surrounding
window we implicitly assume that nearby pixels are homogeneous with the target. For a small window, this is very likely to happen, while it becomes increasingly unlikely as we add more pixels. Therefore, rather than enlarging the window, we select new patches that are more likely to be homogeneous with the target, based on the same patch similarity measure used in the nonlocal filtering. In more detail, we build a stack by collecting the $K$ patches most similar to the target patch in a suitable search area surrounding the target, and compute the $A/G$ statistics on the averaged patch. By so doing, we carry out a sort of synthetic temporal multilooking, reducing noise with little resolution loss.

Note that, since classification is not our ultimate goal, but only an intermediate step towards better despeckling, we do not need to separate sharply the two classes but can resort to a soft classification, allowing any combination of the the two despeckled images. Given the limited classification reliability, this is a valuable guarantee against high errors.

2.2. Despeckling Tools

We use two despeckling tools, SAR-BM3D [6] and the homomorphic version of LSSC [12]. Therefore, we focus only on their strengths and weaknesses, referring the reader to the original papers for more detail. SAR-BM3D combines nonlocal filtering with other effective denoising tools, like wavelet shrinkage and Wiener filtering. Its major asset, typical of nonlocal filters, is the ability to faithfully preserve relevant details, in particular region boundaries, thin structures, man-made objects. Instead, it provides only a limited speckle rejection in homogeneous areas.

The Learned Simultaneous Sparse Coding, LSSC [12] belongs to the class of sparse coding techniques, where each patch of the noisy image is represented as a combination of reference patches drawn from a suitable dictionary. The dictionary is itself built from the observed image patches by solving a minimization problem. Since the number of dictionary patches is much smaller than the number of image patches from which they are synthesized, they represent a smoothed version of them, granting a strong noise suppression. On the other hand, LSSC tends to oversmooth high-frequency details, and therefore has qualities complementary w.r.t. SAR-BM3D. We apply the original AWGN version of LSSC in the tails, and therefore has qualities complementary w.r.t. SAR-BM3D. We apply the original AWGN version of LSSC in the homomorphic domain having care not to introduce biases in the process, as done also in some papers using sparse representations for SAR despeckling [13, 14].

3. EXPERIMENTAL RESULTS

To test the proposed method we carried out experiments on several real-world single-look SAR images. Given space constraints, here we present and discuss results only for the TerraSAR-X image taken over Rosenheim, in Germany, shown in Fig.3(a). In Fig.2 we show various classification maps. The soft classification maps of parts (a)-(c) are obtained using the A/G statistic as input of a suitable logistic nonlinearity, with output in [0,1]. Values close to 1(0) correspond to large[small] A/G ratios Computing A/G by the stack-wise processing, with stacks of 64 patches (c), increased accuracy significantly w.r.t. to the single-patch case, with no increase in blockiness. Finally, in part (d), we show a hard classification map obtained by thresholding the last A/G ratio field. The sharp transitions are bound to produce visible artifacts in a hard combination.

In Fig.3 we show the despeckled versions of the original image (a) obtained using SAR-BM3D (b), homomorphic LSSC (c), their soft combination (d) through the map of Fig.2(c) (all images should be inspected on a computer screen with suitable zooming). As expected, the SAR-BM3D image provides a limited speckle rejection, while the H-LSSC image exhibit some undesired smoothing on man-made areas. The soft combination image, instead, represents a better result, especially for subsequent automatic analysis tools, which will not be hampered by speckle. We also compare results with some state-of-the-art nonlocal filters: FANS (e) [8], PPB (f) [5], NL-SAR (g) [7], and a recent technique based on a variational model proposed by Chen et al. (h) [15]. NL-SAR appears to provide a despeckled image with little or no artifacts, like the proposed method, but the latter has a better edge and structure preservation. Given these promising results, we are confident to obtain further improvements with a full-fledged implementation of the approach.

4. REFERENCES


