Modeling the distribution of patches with shift-invariance: application to SAR image restoration

Sonia Tabti, Charles-Alban Deledalle, Loïc Denis, Florence Tupin

MODELING THE DISTRIBUTION OF PATCHES WITH SHIFT-INVARIANCE: APPLICATION TO SAR IMAGE RESTORATION

Sonia Tabti[1] *, Charles-Alban Deledalle*[2]*, Loïc Denis*[3]*, Florence Tupin*[1]

[1]Institut Mines-Télécom, Télécom ParisTech CNRS-LTCI, Paris, France
[2]CNRS-Univ. Bordeaux, IMB, Talence, France
[3] CNRS-Univ. Saint-Etienne, Laboratoire Hubert Curien, Saint-Etienne, France

ABSTRACT

Patches have proven to be very effective features to model natural images and to design image restoration methods. Given the huge diversity of patches found in images, modeling the distribution of patches is a difficult task. Rather than attempting to accurately model all patches of the image, we advocate that it is sufficient that all pixels of the image belong to at least one well-explained patch. An image is thus described as a tiling of patches that have large prior probability. In contrast to most patch-based approaches, we do not process the image in patch space, and consider instead that patches should match well everywhere where they overlap. In-order to apply this modeling to the restoration of SAR images, we define a suitable data-fitting term to account for the statistical distribution of speckle. Restoration results are competitive with state-of-the art SAR despeckling methods.

1. INTRODUCTION

SAR images suffer from strong fluctuations due to the speckle phenomenon inherent to coherent imagery. The problem of speckle reduction has driven the development of numerous denoising methodologies these last 3 decades (see [1] for a recent review). Despite constant improvement of the methods, restored images still suffer from some defects like denoising artifacts (i.e., amplification of some spurious structures) and noise halos (i.e., regions where noise fluctuations remain).

Many recent and effective denoising approaches rely on the decomposition of the image into small rectangular areas called "*patches*", typically 8×8 pixels, that capture local information (geometry and texture). Two different strategies can be identified among the numerous patch-based denoising methods [2]: (i) methods that group similar patches and exploit the redundancy among selected patches to reject noise (by averaging [3], filtering in a transform space [4], or using principal component analysis [5, 6]); and (ii) methods that rely on a model of the prior distribution of patches.

The first family of methods relies on the assumption that several similar patches can be found within a reasonably small search area (typically 29×29 to 39×39 pixels). However, in SAR imagery many structures are rather rare, especially isolated bright targets and features corresponding to man-made structures. The second family of methods requires to model the distribution of patches and can be divided into the methods that use sparse coding with a redundant dictionary of patches [7, 8] and those that use Gaussian mixture models (GMM) [9, 10]. One of the main drawbacks of this family of methods is that they do not take into account some useful invariances. For instance in dictionary-based sparse decompositions, it is necessary to have atom combinations representing all shifted versions of a specific pattern. Similarly, GMM learning [10] does not exploit shift invariance. Epitomes [11, 12] offer an efficient way to encode dictionaries with many shifted versions of each atomes. In this paper, rather than explicitly modeling all patches in an image, which represents a huge variability, we consider describing only a fraction of them. To ensure a complete characterization of the image, we enforce that those patches described by our model form a tiling of the whole image.

This paper brings two contributions: the introduction of shift invariance in prior models of images (section 3) and the adaptation of the image restoration procedure to SAR imaging by accounting for speckle distribution (section 4).

2. PATCH-BASED PRIORS

2.1. Approximating patch distribution with a GMM

Statistical modeling of images has a long history. While Markov random fields generally consider pairs of neighbor pixels, patch-based priors capture much richer information. Previous works [9] and [10] have shown that the distribution of patches in natural images can be well described by a GMM. The prior model for a patch $\boldsymbol{z}$ is then defined by:

$$\mathrm{p}(\boldsymbol{z}) = \sum_{k=1}^{K} \pi_k n_k \exp\{-\tfrac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu}_k)^t \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{z}-\boldsymbol{\mu}_k)\} \quad (1)$$

$$\approx \max_k \pi_k n_k \exp\{-\tfrac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu}_k)^t \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{z}-\boldsymbol{\mu}_k)\} \quad (2)$$

with π_k the k-th mixing weight, n_k a normalization given by $n_k = \det(2\pi\boldsymbol{\Sigma}_k)^{-1/2}$, $\boldsymbol{\mu}_k$ the k-th mode and $\boldsymbol{\Sigma}_k$ the corre-

This work is supported by DGA and CNRS funding through a PhD grant.

sponding k-th covariance matrix. Equation (2) locally approximates the GMM by the component with largest weight.

While many methods are defined in patch domain, we define the prior model in image domain as proposed in [10]:

$$\mathrm{p}(\boldsymbol{x}) = \prod_i \mathrm{p}(\boldsymbol{P}_i \boldsymbol{x}) \tag{3}$$

whith $\boldsymbol{P}_i$ the matrix extracting the i-th patch of image $\boldsymbol{x}$ and $\mathrm{p}(\boldsymbol{P}_i \boldsymbol{x})$ the distribution of patch $\boldsymbol{P}_i \boldsymbol{x}$ as defined in (1) or (2).

2.2. Expected Patch Log Likelihood (EPLL) method

We recall in this paragraph the principle of EPLL method [10] from which we derive our restoration method. The aim of the EPLL method is to reconstruct an image such that every patch in this reconstruction is likely under a specific prior while being close to the corrupted image. Under stationary white Gaussian noise, the MAP estimate of the image is given by:

$$\min_{\boldsymbol{x}} \ \frac{\lambda}{2}\|\boldsymbol{x}-\boldsymbol{y}\|^2 - \log \mathrm{p}(\boldsymbol{x}) \tag{4}$$

where $\boldsymbol{y}$ is a corrupted image, $\boldsymbol{x}$ the restored image, $\lambda > 0$ a parameter and the prior $\mathrm{p}(\boldsymbol{x})$ is defined by equation (3). The authors of [10] suggest solving the following optimization problem using the so-called half quadratic splitting method:

$$\min_{\boldsymbol{x},\{\boldsymbol{z}^i\}} \ \frac{\lambda}{2}\|\boldsymbol{x}-\boldsymbol{y}\|^2 + \sum_i \frac{\beta}{2}\|\boldsymbol{P}_i\boldsymbol{x}-\boldsymbol{z}^i\|^2 - \log \mathrm{p}(\boldsymbol{z}^i) \tag{5}$$

where β tunes the difference between $\boldsymbol{P}_i\boldsymbol{x}$ and the auxiliary variables $\{\boldsymbol{z}^i\}$. For a fixed β value, this problem can be solved alternatively for $\boldsymbol{x}$ and for the set of $\boldsymbol{z}^i$ (4 or 5 iterations are enough). Solving for $\boldsymbol{x}$ amounts to computing a linear combination of the noisy image $\boldsymbol{y}$ and patches $\boldsymbol{z}^i$ that project at a given pixel. Solving for $\boldsymbol{z}^i$ is also a quadratic problem whose solution is a Wiener filter. Their proposed prior is a GMM composed of 200 components learnt over a huge training basis of patches (10^6) extracted from natural images. The training took 30 hours of computation.

In the next section, we describe how the prior model can be made invariant to geometrical shifts, thus capturing only the remaining variability with each Gaussian component.

3. INTRODUCING THE SHIFT-INVARIANCE

To prevent from encoding all shifted versions of the same structure into the GMM, we require that only *some* patches from $\boldsymbol{x}$ be well described by our dictionary. Well-explained patches must cover all the image $\boldsymbol{x}$, i.e., $\boldsymbol{x}$ should be close to a tiling of the image domain with patches drawn from the dictionary. Our prior model for image $\boldsymbol{x}$ is thus defined by:

$$\mathrm{p}(\boldsymbol{x}) = \prod_i \max_{j \in \mathcal{N}(i)} \mathrm{p}(\boldsymbol{P}_j \boldsymbol{x}) \tag{6}$$

where $\mathcal{N}(i)$ is the set of patch indexes that are in the neighborhood of patch i, i.e., all patches that cover pixel i. Note

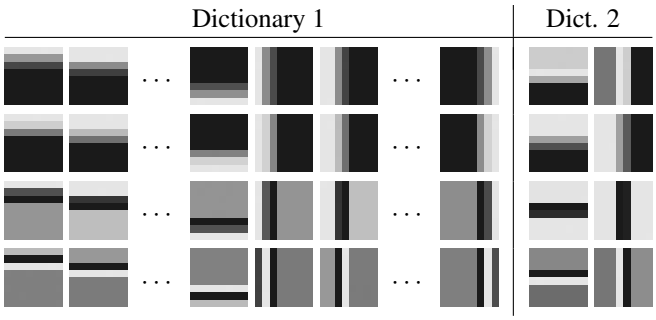


Fig. 1. Each column represents the first 4th eigenvectors of the covariance matrices of two dictionaries. The first dictionary models horizontal and vertical edges (with a precision of ± 1 pixel) for all possible shifts in the patch with a mixture of 10 zero-mean Gaussians. The second one models only one shift of these edges with a mixture of 2 zero-mean Gaussians.

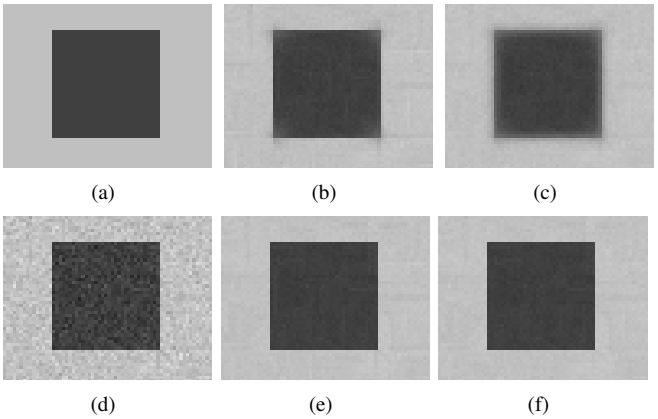


Fig. 2. (a) Noise free image. (d) Noisy image. (b-c) Non-shift invariant reconstructions and (e-f) shift invariant reconstructions. (b,e) are obtained with the 1st dictionary of 10 Gaussians and (c,f) with the 2nd one of 2 Gaussians (see Fig.1).

that if the probabilities of all patches are stored as an image[1], then the maximum operation in eq.(6) corresponds to the dilation operator from mathematical morphology.

The adaptation of EPLL method to our shift-invariant prior leads to the following minimization problem:

$$\min_{\boldsymbol{x}} \ \frac{\lambda}{2}\|\boldsymbol{x}-\boldsymbol{y}\|^2 + \sum_i \min_{j \in \mathcal{N}(i), \boldsymbol{z}^i} \left\{ \frac{\beta}{2}\|\boldsymbol{P}_j\boldsymbol{x}-\boldsymbol{z}^i\|^2 - \log \mathrm{p}(\boldsymbol{z}^i) \right\} \tag{7}$$

For a fixed $\boldsymbol{x}$, solving for j and $\boldsymbol{z}^i$ amounts to comparing the probability of each patch $\boldsymbol{P}_j\boldsymbol{x}$ that cover pixel i according to each of the K Gaussian components. Once the component with largest probability is identified, $\boldsymbol{z}^i$ is computed by Wiener filtering as with EPLL method.

Figure 2 illustrates the gain brought by the proposed shift-invariant modeling for a denoising task, compared to EPLL method. A synthetic image of a square, corrupted by Gaus-

[1] patches of an image are in bijection with the pixels of the image, if we omit boundary issues that require proper handling

sian noise is restored using two different dictionaries built specifically for this task (see Fig.1). The first dictionary encodes all possible shifts of vertical and horizontal edges while the second one encodes a single location of the vertical and horizontal edges. The EPLL method fails to restore the square with both dictionaries since neither corners (1st dictionary) nor shifted versions of the edges (2nd dictionary) are encoded. Our proposed shift-invariant procedure succeeds in restoring the square with both dictionaries and in particular with the 2-atoms dictionary. Note that corners do not need to be explicitly encoded in the dictionary since, around each pixels, at least one surrounding patch is well explained. Much more compact dictionaries can thus be used.

4. ADAPTATION TO SAR IMAGERY

We discussed so-far about prior models $\mathrm{p}(\boldsymbol{x})$ based on patches. To apply such models to a speckle reduction task, EPLL method defined in eq. (5) cannot be directly applied. Indeed, the data fidelity term $\|\boldsymbol{x}-\boldsymbol{y}\|^2$ does not consider the specificity of speckle fluctuations. In a Bayesian interpretation one should replace this term by the log-likelihood given by statistical speckle models.

SAR images have two main distinctive features compared to "natural" images. First, they have a very high dynamic range due to strong specular reflexions of the incident electromagnetic wave on man-made structures. Second, due to the use of coherent illumination, interference phenomenon occur which create random fluctuations proportional to the back-scattered signal. For both reasons we propose to process log-transformed data. GMM modeling is made easier with data of reduced dynamic range. Besides, the log transform stabilizes the variance, i.e., log-transformed speckle is independent from the radiometry of the radar scene. Since log-transformed speckle does not follow a Gaussian distribution, using a quadratic penalty is not adapted. In particular, the quadratic penalty does not account for the asymmetry of the distribution of log-transform data, resulting in images with several isolated darker pixels.

Under Goodman's speckle model [13], log-transformed SAR images[2] follow a Fisher-Tippet distribution [14], which leads to the following log-likelihood:

$$-\log \mathrm{p}(\boldsymbol{y}|\boldsymbol{x}) = \lambda \sum_{i=1}^{N} \left(e^{y_i - x_i} + x_i - y_i \right) \tag{8}$$

where x_i is the i-th element of $\boldsymbol{x}$. When replacing the quadratic data fidelity with this likelihood in eq. (5), the image restoration problem can be solved by alternating the minimization for $\boldsymbol{x}$ with fixed $\boldsymbol{z}^i$, and with respect to $\boldsymbol{z}^i$ and j with fixed $\boldsymbol{x}$.

Considering $\boldsymbol{z}^i$, solving for $\boldsymbol{x}$ can be done efficiently with an iterative numerical scheme. Remark that due to the specific structure of matrices $\boldsymbol{P}_j$, the sum can be rewritten:

$$\sum_i \|\boldsymbol{P}_{j_i^\star}\boldsymbol{x} - \boldsymbol{z}^i\|^2 = \sum_i c_i (x_i - \bar{z}_i)^2$$

where $j_i^\star$ is the index of the patch covering pixel i that best fits the prior, $\bar{\boldsymbol{z}} = \mathrm{diag}(\boldsymbol{c}) \sum_i \boldsymbol{P}_{j_i^\star}^t \boldsymbol{z}^i$ is the uniform reprojection of patches $\boldsymbol{z}^i$ in the image domain, i.e. the image obtained by averaging patch values that project onto each pixel, and c_i is the number of patches $\boldsymbol{P}_{j_i^\star}\boldsymbol{x}$ that project onto pixel i. Hence, solving the minimization problem involving the data term of (8) with respect to $\boldsymbol{x}$ boils down to minimizing the following separable cost function:

$$\sum_{i=1}^{N} \left[\lambda (e^{y_i - x_i} + x_i - y_i) + \frac{\beta}{2} c_i (x_i - \bar{z}_i)^2 \right] \tag{9}$$

Unlike the case of a quadratic fidelity term, the solution of (9) does not have a simple closed-from solution. Indeed, the authors of [15] show that this solution is expressed with the Lambert W function [16], which can be time consuming to evaluate. Instead, since eq. (9) is strictly convex and differentiable in $\boldsymbol{x}$, it can be solved efficiently with an iterative numerical scheme based on Newton's method that performs at iteration $t+1$ the following update:

$$x_i^{(t+1)} = x_i^{(t)} - \frac{\lambda(1 - e^{y_i - x_i^{(t)}}) + \beta c_i (x_i^{(t)} - \bar{z}_i))}{\lambda e^{y_i - x_i^{(t)}} + \beta c_i} \tag{10}$$

for all pixels i. Since the problem is separable, several iterations can be performed with a cost negligeable compared to the optimization of the MAP problem for all $\boldsymbol{z}^i$. We notice that 50 iterations are enough to reach an accurate solution.

5. RESULTS ON SAR IMAGES

We now illustrate the interest of shift-invariance and an adapted data fidelity term for the denoising of high resolution SAR images. The results presented in Figure 3 and 4 are obtained on a 2-looks SAR image using the GMM prior learnt in [10]. In Figure 3, we compare the EPLL applied on the logarithmically transformed image, with our data term and with an additional shift-invariance. The resulting images preserve most of the back-scattering targets which is appreciable in SAR imagery applications. We observe that our method does not present the common artefact encontered when using a logarithmically transformed method (pixels darker than their neighbors circled in red) and preserves edges which are blurry without the shift invariance property (areas circled in orange). Figure 4 compares our method with state of the art methods such as [18], [19] and confirms the gain of a shift invariant model.

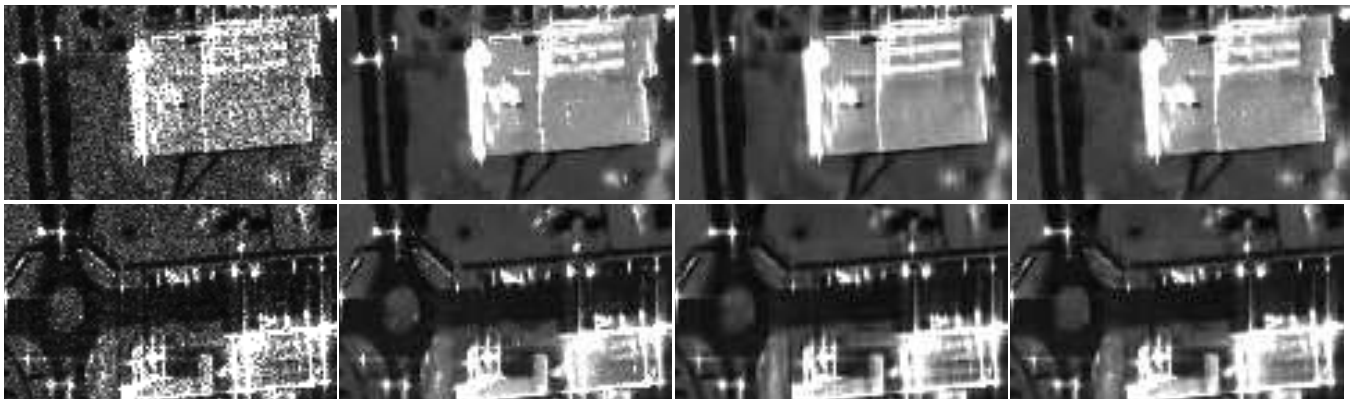
6. CONCLUSION

Gaussian Mixture Models are very effective to model the distribution of patches in natural images. Although covariance

[2] both intensity and amplitude data

(a) (b) (c) (d)

Fig. 3. (a) 2-look SAR image of Toulouse (France) ©ONERA ©CNES (the two images at bottom are zooms of the top one). (b) Result of EPLL applied after log-transform and with a post debiaising step following [17]. (c-d) Result of our approach with a Fisher-Tippet based data fidelity term on the log-transformed data respectively without and with invariance.



(a) (b) (c) (d)

Fig. 4. (a) Two zooms of the same SAR image as in Figure 3. Comparisons of denoising results between (b) the NL-SAR filter [18] (c) the BM3D filter on log-tranformed data [19], and (d) the proposed approach.

matrices used by EPLL algorithm *implicitly* encode the diversity due to shifts of a given pattern, our approach *explicitly* accounts for shifts by requiring that only one of the representative patches containing a pixel be well modeled by the mixture of Gaussians. This model seems better suited to preserve some sharp structures such as edges. We applied this method on high resolution SAR images taking into account the noise distribution of this coherent imagery and obtained improved denoising results.

As future work, we plan to introduce in our framework invariance to radiometric changes. In fact many patches can be very similar up to a contrast change and taking advantage of this property could make the model more accurate since it will capture more intricate variability than simple geometrical and radiometric changes. These invariances pave the way to the learning of compact models of the distribution of patches, requiring fewer Gaussian components to reach the same accuracy. Applications of this work concerns the learning of shift-invariant dictionaries dedicated to SAR images in order to preserve better specific structures such as bright targets.

7. 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.*, pp. 6–35, 2013.

[2] M. Lebrun, A. Buades, and J.-M. Morel, "A non-local Bayesian Image Denoising Algorithm," *SIAM J. Imaging Sciences*, vol. 6, no. 3, 2013.

[3] A. Buades, B. Coll, and J. M. Morel, "Image Denoising Methods . A New Nonlocal Principle," *SIAM Review*, vol. 52, no. 1, 2010.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering.," *IEEE Transactions on Image Processing*, vol. 16, no. 8, 2007.

[5] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognition*, vol. 43, no. 4, pp. 1531–1549, 2010.

[6] C-A. Deledalle, J. Salmon, and A. S. Dalalyan, "Image denoising with patch based PCA: local versus global," in *BMVC*, 2011.

[7] M. Aharon, M. Elad, and A.M. Bruckstein, "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[8] B.A. Olshausen et al., "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, 1996.

[9] G. Yu, G. Sapiro, and S. Mallat, "Solving Inverse Problems With Piecewise Linear Estimators: From Gaussian Mixture Models to Structured Sparsity," *IEEE Transactions on Image Processing*, vol. 21, pp. 2481–2499, 2012.

[10] D. Zoran and Y. Weiss, "From Learning Models of Natural Image Patches to Whole Image Restoration," *ICCV*, 2011.

[11] N. Jojic, B.J. Frey, and A. Kannan, "Epitomic analysis of appearance and shape," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 34–41.

[12] L. Benoît, J. Mairal, F. Bach, and J. Ponce, "Sparse image representation with epitomes," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 2913–2920.

[13] J. W. Goodman, "Some fundamental properties of speckle," *Journal of optical society of America*, 1976.

[14] H. Xie, L.E Pierce, and F.T. Ulaby, "Statistical properties of logarithmically transformed speckle," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 3, pp. 721–727, 2002.

[15] J.M. Bioucas-Dias and M. A. Figueiredo, "Multiplicative noise removal using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, 2010.

[16] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the lambert W function," *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.

[17] H. Xie, L.E. Pierce, and F.T. Ulaby, "SAR speckle reduction using wavelet denoising and Markov random field modeling," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2196–2212, Jan. 2002.

[18] 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.

[19] M. Makitalo, A. Foi, D. Fevralev, and V. Lukin, "Denoising of single-look SAR images based on variance stabilization and nonlocal filters," in *Mathematical Methods in Electromagnetic Theory (MMET), 2010 International Conference on*. IEEE, 2010, pp. 1–4.