# Texture Analysis of Mean Shift Segmented Low-Resolution Speckle-Corrupted Fractional SAR Imagery through Neural Network Classification

Gustavo D. Martín del Campo-Becerra,
Juan I. Yañez-Vargas, and Josué A. López-Ruíz

Centre for Advanced Research and Education of the National Polytechnic Institute
CINVESTAV-IPN, Unidad Guadalajara, México
{gmartin,jyanez,jalopez}@gdl.cinvestav.mx

**Abstract.** The novel proposal of this work is the application of the nonparametric mean shift technique, for image segmentation, to low-resolution (LR) speckle-corrupted imagery, acquired with conventional low-cost fractional synthetic aperture radar (Fr-SAR) systems; with aims of analyzing the resultant textures, related to the remotely sensed (RS) scenes, via neural network (NN) classification. The LR speckle-corrupted recovery of the spatial reflectivity maps, provided by Fr-SAR systems, is due to the fractional synthesis mode and the different model-level and system-level operational scenario uncertainties, peculiar to such systems operating in harsh remote sensing scenarios. The mean shift segmentation method delineates arbitrarily shaped regions in the treated LR image by locating the modes in the density distribution space, and by grouping all pixels associated with the same mode. Then, the textures extracted from the segmented image are classified through NN computing, to posteriorly be used for analysis and interpretation.

**Keywords:** Fractional synthetic aperture radar, remote sensing, mean shift segmentation, neural network classification, texture analysis.

## 1 Introduction

Conventional low-cost fractional synthetic aperture radar (Fr-SAR) systems inherently sacrifice spatial resolution due to their minor aperture synthesis mode; furthermore, since such Fr-SAR systems typically operate in harsh remote sensing (RS) environments, they suffer from different operational scenario uncertainties at model and system levels [1], [2]. These aspects cause the low-resolution (LR) and speckle corruption on the provided Fr-SAR imagery. Most Fr-SAR systems use the matched spatial filtering (MSF) methodology for image formation [1], [2]; the resultant LR speckle-corrupted MSF image represents an inaccurate estimate of the spatial spectrum pattern (SSP) of the backscattered field, which denotes a spatial map of the RS scene power reflectivity.

In this study, the LR speckle-corrupted Fr-SAR image, formed employing the conventional MSF processing method, is segmented through the proper mean shift

E. Bayro-Corrochano and E. Hancock (Eds.): CIARP 2014, LNCS 8827, pp. 998–1005, 2014.

procedure, for further texture analysis via neural network (NN) classification. Mean shift is a nonparametric tool for finding the modes (peaks) in a set of data samples, manifesting an underlying probability density function (pdf) [3]-[5]. The mean shift based image segmentation procedure consists on associating each pixel in the image with a mode located in its neighborhood, after nearby modes are pruned [3], [5].

The main perceptual characteristics in SAR imagery are tone, texture and edges [1]. Due to speckle and blurring, Fr-SAR image features with similar backscattering coefficients may express similar tone and texture, causing confusion. The mean shift segmentation process alleviates this confusion by delineating arbitrarily shaped regions associated to the same mode, regardless the additive and multiplicative speckle noise presence. The main advantage of such technique lies in achieving image segmentation without any assumptions concerning the underlying statistics involved in the images [3]. Once the Fr-SAR image is segmented, feature extraction from each texture in the segmented image is performed, in order to train a feed-forward back-propagation sigmoid NN, responsible of classifying the Fr-SAR image textures, associated to the remotely sensed scenes.

## 2 Imaging Radar Problem Phenomenology

The imaging radar problem model treated in this section is structurally similar to the previous studies [6]. The model of the observation field $u$ is defined through the vector-form equation of observation given by [6]

$$\mathbf{u} = \tilde{\mathbf{S}}\mathbf{v} + \mathbf{n} = \left(\mathbf{S} + \mathbf{S}_{\Delta}\right)\mathbf{v} + \mathbf{n}, \tag{1}$$

where the vectors $\mathbf{u}$, $\mathbf{n}$ and $\mathbf{v}$ are treated as Gaussian zero-mean vectors composed of the coefficients $\{u_m\}_{m=1}^{M}$, $\{n_m\}_{m=1}^{M}$ and $\{v_l\}_{l=1}^{L}$ of the finite-dimensional approximations of the observation $u$, noise $n$ and complex random reflectivity $v$ fields, respectively; and $\tilde{\mathbf{S}}$ is the $M \times L$ matrix-form approximation of the integral perturbed signal formation operator (SFO). Vectors $\mathbf{u}$, $\mathbf{n}$ and $\mathbf{v}$ are characterized by the correlation matrices $\mathbf{R}_{\mathbf{u}} = \left\langle \tilde{\mathbf{S}}\mathbf{R}_{\mathbf{v}}\tilde{\mathbf{S}}^{+} \right\rangle + N_0\mathbf{I}$, $\mathbf{R}_{\mathbf{n}} = N_0\mathbf{I}$ and $\mathbf{R}_{\mathbf{v}} = \mathbf{D}(\mathbf{b}) = \mathrm{diag}(\mathbf{b})$, correspondingly, where superscript $^{+}$ stands for Hermitian conjugate and $N_0$ is the white observation noise power. Vector $\hat{\mathbf{b}} = \mathcal{L}\{\hat{\mathbf{B}}\}$ is a lexicographically ordered vector-form approximation of the SSP map $\mathbf{B} = \{b(l_x, l_y)\}$ over the $L_x \times L_y$ pixel-framed 2-D scene $\{l_x = 1,\ldots, L_x;\ l_y = 1,\ldots, L_y;\ l = 1,\ldots, L = L_xL_y\}$ [6].

The radar imaging problem discrete representation is expressed as [6]

$$\hat{\mathbf{b}} = \mathrm{est}\{\mathbf{b} \mid \mathbf{u} = \tilde{\mathbf{S}}\mathbf{v} + \mathbf{n}\}. \tag{2}$$

This is interpreted as the SSP estimation from the available data recordings $\mathbf{u}$ degraded by composite additive multiplicative noise with perturbation statistics $\left\langle \tilde{\mathbf{S}}\mathbf{R}_{\mathbf{v}}\tilde{\mathbf{S}}^{+} \right\rangle$ unknown to the observer. The nonlinear inverse problem at hand is ill posed, due to the violation of the existence condition. Besides, the unavoidable presence of noise in the data and problem model imprecisions, add statistical uncertainty to the SSP estimation problem.

## 3 Matched Spatial Filtering SSP Estimate

The matched spatial filtering (MSF) technique is the most commonly used approach to estimate the SSP from the available recordings of data $\mathbf{u}$. The MSF technique is roughly described as the application of the solution operator, the scaled conjugate transpose (adjoint) SFO matrix $\mathbf{S}^+$, to the measured data $\mathbf{u}_{(j)}$, recorded from several independent observations $\{j = 1,...,J\}$; the squared detection to the MSF filter outputs; and the averaging of all approximations acquired from each independent data observation [6]:

$$\hat{\mathbf{b}}_{\text{MSF}} = \{\mathbf{S}^+\mathbf{Y}\mathbf{S}\}_{\text{diag}}, \tag{3}$$

where the data matrix $\mathbf{Y} = \frac{1}{J}\sum_{j=1}^{J}\mathbf{u}_{(j)}\mathbf{u}^+_{(j)}$ is formed by the averaging of the multiple independent observed data realizations $\{\mathbf{u}_{(j)}; j = 1,...,J\}$.

Synthetic aperture radar (SAR) systems hardware quality has a direct impact on the resolution of the imagery acquired employing the MSF processing method [1], [2]. An example of an image acquired via a high-cost high-resolution wide focus SAR system, using the MSF image formation technique, is depicted by **Fig. 1(a)**. In the other hand, **Fig. 1(b)** corresponds to the image obtained via a simulated low-cost LR Fr-SAR system, employing also the MSF image formation methodology.

## 4 Mean Shift Segmentation

This section firstly reviews the results described in [3], which are posteriorly applied to Fr-SAR imagery. Kernel density estimation is one of the most popular density estimation methods. Let $\{\mathbf{a}_i\}_{i=1}^{I}$ be an arbitrary set of $I$ data points in the $d$ -dimensional space $\mathbb{R}^d$, the kernel density estimator for the sample point $\mathbf{a}$, with kernel $K(\mathbf{a})$ and window radius (bandwidth value) $h$, is defined by [3]

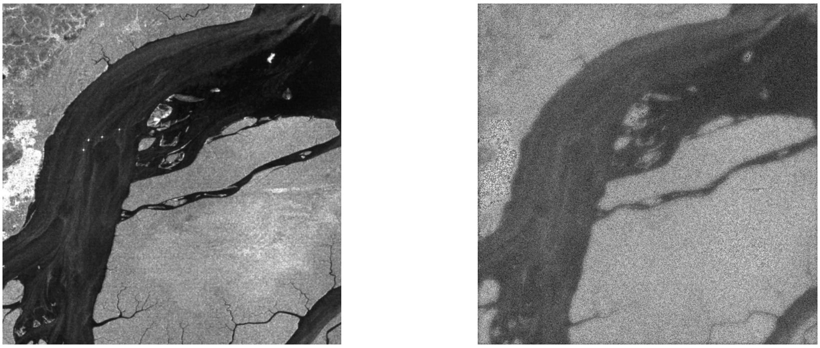
$$\hat{f}_K(\mathbf{a}) = \frac{1}{Ih^d}\sum_{i=1}^{I} K\left(\frac{\mathbf{a}-\mathbf{a}_i}{h}\right). \tag{4}$$

For the case of radially symmetric kernels satisfying [3]

$$K(\mathbf{a}) = c_K k\left(\|\mathbf{a}\|^2\right), \tag{5}$$

it suffices to define the profile $k(a)$ of the kernel, only for $a \geq 0$. The normalization constant $c_K$ assures $K(\mathbf{a})$ integrates to one. The next step is to find the modes of the underlying density $f(\mathbf{a})$, located among the zeros of the gradient $\nabla f(\mathbf{a}) = \mathbf{0}$.

Function $g(a) = -k'(a)$ is defined when the derivative of the kernel profile $k(a)$ exists. Taking the latter into consideration, the gradient of the kernel density estimator (4) is expressed by [3]



(a) (b)

**Fig. 1.** (a) HR 512×512 pixel-framed test scene borrowed from the real-world high resolution SAR imagery [7]; **(b)** LR speckle-corrupted MSF image of the same scene formed with a simulated Fr-SAR system; model system parameters: triangular range point spread function (PSF), the width (at ½ of the peak value) $l_y = 10$ pixels; Gaussian bell azimuth PSF, the width (at ½ of the peak value) $l_x = 15$ pixels; the worst case single-look scenario with fully developed speckle, (SNR = 0 dB).

$$\nabla \hat{f}_K(\mathbf{a}) = \frac{2c_K}{Ih^{d+2}} \left[ \sum_{i=1}^{I} g\left( \left\| \frac{\mathbf{a}-\mathbf{a}_i}{h} \right\|^2 \right) \right] \left[ \frac{\sum_{i=1}^{I} \mathbf{a}_i g\left( \left\| \frac{\mathbf{a}-\mathbf{a}_i}{h} \right\|^2 \right)}{\sum_{i=1}^{I} g\left( \left\| \frac{\mathbf{a}-\mathbf{a}_i}{h} \right\|^2 \right)} - \mathbf{a} \right] \tag{6}$$

From (6), the first term, on the right side of the equation, is proportional to the density estimate at $\mathbf{a}$ computed with kernel $G(\mathbf{a}) = c_G g\left(\|\mathbf{a}\|^2\right)$; and the second term, the difference between the weighted mean, is the mean shift, which uses kernel $G$ for weights and $\mathbf{a}$ as the kernel (window) center. Expression (6) becomes [3]

$$\nabla \hat{f}_K(\mathbf{a}) = \hat{f}_G(\mathbf{a}) \frac{2c_K}{h^2 c_G} \mathbf{m}_G(\mathbf{a}), \tag{7}$$

yielding

$$\mathbf{m}_G(\mathbf{a}) = \frac{1}{2} h^2 c \frac{\nabla \hat{f}_K(\mathbf{a})}{\hat{f}_G(\mathbf{a})}, \tag{8}$$

which reveals that the mean shift vector (8) computed with kernel $G$ at location $\mathbf{a}$ is proportional to the normalized density gradient estimate acquired with kernel $K$. The mean shift vector points towards the maximum increase of density direction.

An image is represented as a two-dimensional lattice of $p$-dimensional vectors, $p=1$ in the gray-level case of the considered LR speckle-corrupted Fr-SAR imagery. The lattice is represented in the spatial domain, whereas the gray-level information is represented in the range domain, assuming Euclidean metric for both domains. The kernel (5) is defined as the product of two radially symmetric kernels [3]

$$K(\mathbf{a}) = \frac{C}{h_s^2 h_r^p} k\left(\left\|\frac{\mathbf{a}^s}{h_s}\right\|^2\right) k\left(\left\|\frac{\mathbf{a}^r}{h_r}\right\|^2\right), \tag{9}$$

where $\mathbf{a}^s$ and $\mathbf{a}^r$ are the spatial part and range part of a feature vector $\mathbf{a}$ correspondingly, $k(a)$ is the profile, common to both domains, $h_s$ and $h_r$ are the employed kernel bandwidths, which control the kernel size and determines the mode detection resolution [3], and $C$ is the corresponding normalization constant.

In order to describe the mean shift segmentation process applied to Fr-SAR imagery, consider $\{\mathbf{x}_i\}_{i=1}^I$ to be the pixels of the SSP map matrix-form approximation $\hat{\mathbf{B}}$ in the joint domain, $\{\mathbf{z}_i\}_{i=1}^I$ are the mean shift technique points of convergence and $\{H_i\}_{i=1}^I$ are a set of labels (scalars). Represent by $\{\mathbf{y}_j\}_{j=1}^J$ the successive positions of kernel $G$, where, from (6) [3],

$$\mathbf{y}_{j+1} = \sum_{i=1}^I \mathbf{x}_i g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right) \Big/ \sum_{i=1}^I g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right); j = 1, ..., J; \tag{10}$$

is the weighted mean at $\mathbf{y}_j$ with $\mathbf{y}_1$ as the kernel center initial location. For each $\{i = 1, ..., I\}$ [3]-[5]:

i. Initialize $j = 1$ and $\mathbf{y}_j = \mathbf{x}_i$.
ii. Calculate $\mathbf{y}_{j+1}$ (10) until convergence, $\mathbf{m}_G(\mathbf{y}_J) = \mathbf{y}_J - \mathbf{y}_{J-1} \approx \mathbf{0}$. Convergence occurs simultaneously in both domains.
iii. Assign $\mathbf{z}_i = (\mathbf{x}_i^s, \mathbf{y}_J^r)$; superscripts $s$ and $r$ stand for spatial and range domain respectively.
iv. Once all $\{\mathbf{z}_i\}_{i=1}^I$ are reached, group together all of them closer than $h_s$ and $h_r$ in the range and spatial domain correspondingly, in order to delineate the clusters $\{\mathbf{C}_p\}_{p=1}^P$ in the joint domain.
v. For each $\{i = 1, ..., I\}$ assign $H_i = \{p \mid \mathbf{z}_i \in \mathbf{C}_p\}$.

The application of the mean shift segmentation procedure to Fr-SAR imagery is exemplified by Fig. 2. The mean shift segmentation technique is applied to the image depicted by Fig. 1(b) using a uniform kernel, achieving convergence in a finite number of steps [3] with $(h_s, h_r) = (20, 32)$.

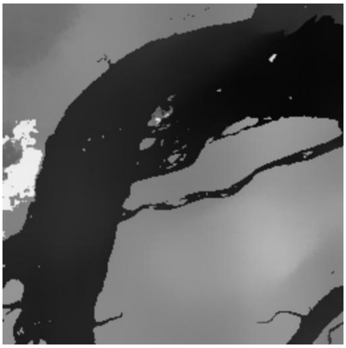


**Fig. 2.** Mean shift segmented Fr-SAR image with a 512×512 pixel frame

## 5 Neural Network Classification

The supervised NN model consists of a multi-layer perceptron (MLP) which uses a back-propagation learning methodology and a sigmoid activation function; it is summarized as follows [8], [9]:

i. Select an input vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_J)^{\mathrm{T}}$ and a desired output vector $\mathbf{d} = (\rho_1, \ldots, \rho_I)^{\mathrm{T}}$ ; superscript $^{\mathrm{T}}$ stands for transpose.

ii. Initialize the synaptic weights between layers $\{l = 1, \ldots, L\}$ with small random values.

iii. Compute the activation of each neuron. The activation $\rho_i^l$ of neuron $i$ in layer $l$ is expressed by

$$\rho_i^l = \varphi\left( \sum_{j=1}^{J} \rho_j^{l-1} w_{ij}^l + \lambda_i^l \right), \tag{11}$$

where summation is done over all $J$ neurons in the $(l-1)$ layer, and $\varphi(\cdot)$ is the sigmoid activation function. The weight for the connection from the $j$ neuron in the $(l-1)$ layer to the $i$ neuron in the $l$ layer is denoted by $w_{ij}^l$; the bias of the $i$ neuron in layer $l$ is represented by $\lambda_i^l$.

iv. Find the error between the layer $L$ output vector $\boldsymbol{\rho}$ and the target vector $\mathbf{d}$ via

$$E = \frac{1}{2} \sum_{i=1}^{I} (\boldsymbol{\rho}_i - \mathbf{d}_i)^2 . \tag{12}$$

The calculated error is then propagated backward in order to obtain the change in the synaptic weights between layers $\{l = 1, \ldots, L\}$, through $\Delta w_{ij}^l = -\alpha \frac{\partial E}{\partial w_{ij}^l}$, where $\alpha$ is a training rate coefficient [8].

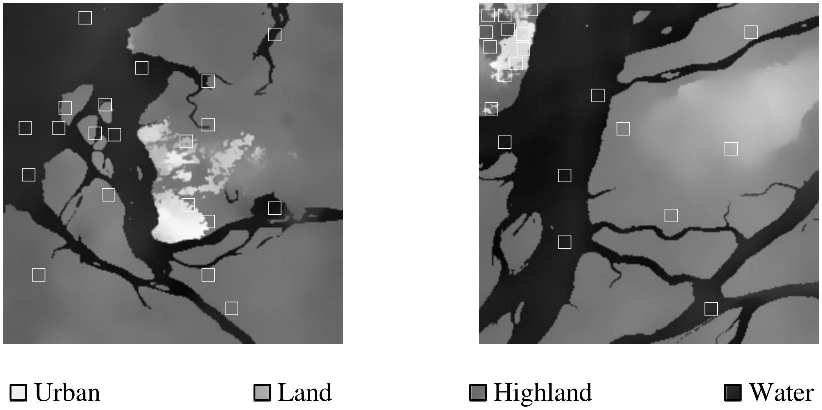



**Fig. 3.** Feature vectors extraction from 512×512 pixel-framed mean shift segmented Fr-SAR imagery, used to train the neural network (NN) classifier.

**Table 1.** Confusion matrix

| | Urban | Land | Highland | Water | Total |
|---|---|---|---|---|---|
| Urban | 49 | 0 | 0 | 0 | 49 |
| Land | 1 | 48 | 0 | 0 | 49 |
| Highland | 0 | 2 | 50 | 1 | 53 |
| Water | 0 | 0 | 0 | 49 | 49 |
| Total | 50 | 50 | 50 | 50 | 200 |

v. Update the synaptic weights using

$$w^{l}_{ij_{[m+1]}} = w^{l}_{ij_{[m]}} + \Delta w^{l}_{ij}; \tag{13}$$

where $\{m = 1,...,M\}$ are the iterations required for achieving the desired NN output.

With aims of training the texture classifier, a set of 200 feature vectors are used. Each feature vector is formed from a lexicographically ordered $20\times 20$ pixels squared window, taken from a set of $10$ mean shift segmented Fr-SAR images. Each feature vector is associated with one of four textures (classes): urban, land, highland and water. The feature vectors extraction process is exemplified by Fig. 3.

Confusion matrix is presented in Table 1. The diagonal elements refer to the correctly identified textures; out of 200 feature vectors considered, correctly identified textures sum to an overall of 196, meaning a 98% of classification accuracy.

## 6 Concluding Remarks

The reported results show that notwithstanding the low resolution, blurring, speckle and additive noise presence in the considered Fr-SAR imagery, thanks to the applied mean shift segmentation process and the appropriate feature vectors extraction, the classification accuracy obtained is satisfactory for distinguishing the provided Fr-SAR imagery texture samples between the four contemplated classes.

The mean shift segmentation procedure technique is convenient because of its simplicity and applicability. Nevertheless, due to the pixel by pixel processing, the computational cost of this methodology is high. On the other hand, the feature vectors extraction step might be suppressed if using unsupervised classification; however, unclassified areas may arise.

## References


1. Cutrona, L.G.: Synthetic Aperture Radar. In: Skolnik, M.I. (ed.) Radar Handbook., 2nd edn. McGraw-Hill, MA (1990)
2. Cumming, I.G., Wong, F.H.: Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation, 1st edn. Artech House, MA (2005)
3. Comaniciu, D., Meer, P.: Mean Shift: A Robust Approach toward Feature Analysis. IEEE Trans. on Pattern Analysis and Machine Intelligence 24(5), 603–619 (2002)
4. Lang, F., Yang, J., Li, D., Wei, J.: Mean-Shift-Based Speckle Filtering of Polarimetric SAR Data. IEEE Trans. on Geoscience and Remote Sensing 52(7), 4440–4454 (2014)
5. Jarabo, P., Rosa, M., de la Mata, D., Vicen, R., Maldonado, S.: Spatial-Range Mean-Shift Filtering and Segmentation Applied to SAR images. IEEE Trans. on Instrumentation and Measurement 60(2), 584–597 (2011)
6. Shkvarko, Y.V.: Unifying experiment design and convex regularization techniques for enhanced imaging with uncertain remote sensing data. –Part I: Theory; –Part II: Adaptive implementation and performance issues. IEEE Trans. Geoscience and Remote Sensing 48(1), 82–111 (2010)
7. COSMO-SkyMed Website for Institutional and Scientific Users, `http://www.cosmo-skymed.it`
8. Kulkarni, A.: Computer Vision and Fuzzy-Neural Systems, 1st edn. Prentice Hall PTR, NJ (2001)
9. Jie, Y., Yan, L., Zhong, Z., Jing, J.: Research on supervised classification of fully polarimetric SAR image using BP neural network trained by PSO. In: 8th World Congress on Intelligent Control and Automation (WCICA), pp. 6152–6157 (2010)