國立臺灣大學電機資訊學院電信工程學研究所

碩士論文

Graduate Institute of Communication Engineering
College of Electrical Engineering and Computer Science
National Taiwan University

Master's Thesis

基於結構幾何先驗與貝葉斯非局部平均法之 SAR 影像散斑雜訊高效估計與抑制

Efficient Speckle Noise Estimation and Reduction for SAR Imagery Using Bayesian Nonlocal Means with Sketch-Based Geometric Priors

莫明勳

Ming-Hsun Mo

指導教授：丁建均 博士

Advisor: Jian-Jiun Ding, Ph.D.

中華民國 114 年 7 月

July, 2025

doi:10.6342/NTU202502541

誌謝

一篇論文的完成，不只是知識的累積，更是時光的堆疊與情感的沈澱。感謝每一位在我卡關時，像哆啦 A 夢一樣對我伸出圓手的人。

首先獻上最深的敬意與感謝，給我的指導教授丁建均老師。謝謝您在學術旅途上的耐心引領，如一盞始終明亮的燈火，靜靜照亮我在知識長河中時而迷失、時而躊躇的腳步。感謝老師願意支持我參與暑期實習，讓我得以在課堂之外獲得珍貴的歷練、體悟甚至是預聘。回想起與老師最初在大學推甄面試時討論 SVD 的畫面，時光彷彿轉瞬即逝。得遇嚴謹既溫厚的良師，是我學術路上最深的幸運。

特別感謝我摯愛的家人，在求學過程中對我的支持與鼓勵，使我能專心於課業而無後顧之憂，並順利完成碩士學位。你們是我心中的安慰與喜悅。爸媽的辛勞撐育與默默付出，不只是家庭穩定的重要支柱，對家庭與小孩的縝密投入，也一直守護著我走向正確的道路；哥哥的包容與關懷，讓身為弟弟的我，在你面前仍可保有天真及任性的一面，願往後的日子，你們繼續在各自的路上找到自己的平靜與幸福!

感謝實驗室的夥伴們:沅罡、雨婕、汶霑、登智，很幸運在這段旅程中，有你們這樣的夥伴。謝謝你們在研究上的建議與支持，也謝謝你們在生活裡的陪伴與笑聲。我們一起在小琉球的夜晚，聊些無厘頭的話題、說著很爛卻讓人開心的諧音梗，讓原本有些單調的研究生活，悄悄染上了色彩。

感謝衝浪社的好朋友們，與你們在台大校園奔波舉辦活動，在烏石港迎著日出日落追浪、一起嘻笑打鬧開玩笑的時光，就如夏天的海風，風起時恬淡舒適，如夢如醉；風靜時餘波盪漾，令人無限回味。只願那一陣風，也吹進你我的記憶，在往後的歲月中仍能被輕柔想起。

最後，謝謝虹筑，儘管我們曾經歷了風雨，陷入了低潮，仍然陪伴在彼此身旁。謝謝妳的陽光、開朗、陪伴與好笑的音效。在找自己的路途上，有愛的人相伴，是很幸福的事。很高興與妳共譜的這些青春回憶，將以這篇論文的方式，永存於臺灣碩博士論文網。

莫明勳 謹致

國立臺灣大學電信所

民國 114 年 7 月 1 日

doi:10.6342/NTU202502541

中文摘要

合成孔徑雷達（Synthetic Aperture Radar, SAR）影像由於雷達訊號的同調特性，固有地存在著乘性散斑雜訊，嚴重降低影像品質並影響後續影像分析與應用。傳統去噪技術雖具計算效率，但往往忽略結構與紋理資訊，導致無法有效地保留影像細節。此外，多數現有的雜訊估測方法依賴特定的參數分布假設或均質區域選取，在面對紋理複雜、結構多變的 SAR 場景時，這些假設難以成立，因而顯著降低了估測準確性。

本論文提出了一種基於貝葉斯非局部平均（Bayesian Nonlocal Means, BNLM）濾波器與草圖結構先驗（sketch-based geometric priors）的 SAR 影像散斑雜訊高效估測與抑制方法。在雜訊估測方面，研究結合了離散小波轉換（Discrete Wavelet Transform, DWT）、適應性局部變異數估測與多項式迴歸，成功地將雜訊估測的平均誤差大幅降低至 2.66%，顯著優於傳統的 Gaussian-Hermite 方法與 Generalized Gamma Distribution 方法。

在雜訊抑制方面，我們所提出的改良型 BNLM 框架包含兩項關鍵改進：(1)以 Wiener filter 為基礎之先驗估測方法，能有效降低估測偏差並更精確地反映局部變異數 (2) 透過草圖結構導引之異向性高斯核函數，能更精準地保留影像中的邊緣與紋理結構，本研究所提出之方法在量化指標 PSNR、SSIM、運算效率上均有明顯的提升。

總體來說，本研究所提出之方法，在公開的 SAR 機場資料集上進行的實驗驗證顯示，在 SAR 影像散斑雜訊估測與去噪中的準確性、視覺真實性及運算效率上皆有顯著提升，未來更可廣泛應用於衛星影像即時處理等多元化遙測領域。

關鍵字： 合成孔徑雷達、散斑雜訊、貝葉斯非局部平均法、雜訊估測、結構先驗

doi:10.6342/NTU202502541

ABSTRACT

Synthetic Aperture Radar (SAR) imagery inherently suffers from multiplicative speckle noise, which significantly deteriorates image quality and complicates subsequent analysis tasks. Conventional denoising approaches typically overlook critical structural and textural details, while prevalent noise estimation techniques are constrained by oversimplified parametric assumptions, leading to suboptimal accuracy in complex SAR scenes.

This study presents a robust framework for SAR speckle noise estimation and suppression, integrating Bayesian Nonlocal Means (BNLM) filtering with sketch-based geometric priors. In the noise estimation phase, a combination of Discrete Wavelet Transform (DWT), adaptive local variance estimation, and polynomial regression markedly reduces the estimation error to approximately 2.66%, surpassing conventional methodologies.

For noise suppression, the proposed enhanced BNLM method incorporates a Wiener filter-based prior to accurately model local variance and employs an anisotropic Gaussian kernel driven by geometric structural cues. Empirical evaluations utilizing public SAR datasets demonstrate substantial enhancements in Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and computational efficiency, underscoring the method's applicability for real-time satellite imagery processing.

Keywords: Synthetic Aperture Radar, speckle noise, Bayesian Nonlocal Means, noise estimation, structural priors.

doi:10.6342/NTU202502541

CONTENTS

doi:10.6342/NTU202502541

doi:10.6342/NTU202502541

LIST OF FIGURES

doi:10.6342/NTU202502541

LIST OF TABLES

doi:10.6342/NTU202502541

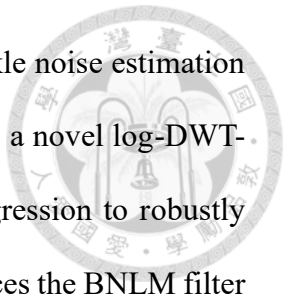
Chapter 1 Introduction

Synthetic Aperture Radar (SAR) imagery plays a vital role in remote sensing due to its ability to capture high-resolution data regardless of weather or lighting conditions. However, the inherent presence of speckle noise, originating from the coherent nature of radar signal reflection, severely degrades image quality, complicating tasks such as segmentation, classification, and object detection. Unlike additive Gaussian noise, speckle noise is multiplicative and signal-dependent, rendering conventional denoising techniques suboptimal for SAR applications.

To tackle this challenge, both speckle noise estimation and suppression must be addressed in a targeted, structure-aware manner. Accurate noise estimation serves as a cornerstone for adaptive filtering, yet existing approaches often rely on strong assumptions (e.g., local homogeneity or predefined distributions), which fail in heterogeneous or highly textured regions. Classical filters such as Lee and Frost provide efficient noise suppression but tend to oversmooth fine details and assume stationary statistics, limiting their performance in real-world SAR scenarios.

Recent advances in transform-domain processing and nonlocal methods have demonstrated improved performance by exploiting patch redundancy and multiscale decomposition. BM3D-based extensions and Bayesian nonlocal means (BNLM) filters exemplify this trend, but they either suffer from high computational complexity or lack structural adaptivity. Moreover, the reliance on noisy patches as priors introduces bias, particularly when structure and texture vary rapidly across the image.

doi:10.6342/NTU202502541



In this thesis, we propose an integrated framework for efficient speckle noise estimation and suppression, tailored for SAR imagery. The first stage introduces a novel log-DWT-based local variance estimation strategy, refined via polynomial regression to robustly capture noise statistics across diverse scenes. The second stage enhances the BNLM filter through two key contributions: a Wiener-based prior estimator and a sketch-based structural prior, both designed to improve denoising accuracy and edge preservation.

Extensive experiments on the publicly available RadarS SAR airport dataset confirm the superiority of our approach. The proposed method not only achieves significantly lower estimation error (average 2.66%) compared to state-of-the-art estimators (e.g., Gaussian-Hermite and GGD-based methods), but also demonstrates notable improvements in PSNR and SSIM for despeckling, while maintaining practical runtime performance. This work thus contributes a scalable and structure-aware solution for SAR image analysis, with promising applications in real-time remote sensing and downstream computer vision tasks.

doi:10.6342/NTU202502541

Chapter 2 Related Work

2.1 Noise Estimation

2.1.1 Estimation Using Bessel K-form PDF[15]

A significant contribution to statistical modeling of speckle noise is presented by Achim et al., who proposed using the Bessel K-form (BKF) probability density function to model log-transformed speckle noise coefficients in transform domains such as wavelet and Curvelet. Unlike traditional Gaussian or Laplacian assumptions, BKF effectively captures the heavy-tailed behavior of speckle noise, particularly in high-frequency subbands. The BKF PDF is defined as

$$f_X(x,p,c)=\frac{1}{\sqrt{\pi}\Gamma(p)c^{(2p+1)/4}}|x|^{p-1/2}K_{p-1/2}\left(\sqrt{\frac{2}{c}}|x|\right) \tag{2.1}$$

where

$\Gamma(p)$ is the Gamma function,

$K_v(\cdot)$ is the modified Bessel function of the second kind,

p is the shape parameter,

c is the scale parameter.

Special cases:

$p = 1$: reduces to Laplacian distribution,

$p \to \infty$: approximates Gaussian distribution,

$p < 1$: exhibits sharper peak and heavier tails.

To estimate these parameters, the authors initially proposed a moment-based estimator using the kurtosis β and variance σ^2 of log-domain coefficients:

doi:10.6342/NTU202502541

$$p = \frac{3}{\beta + 3}, \qquad \hat{c} = \frac{\sigma^2}{p} \tag{2.2}$$

However, to improve robustness, they developed a maximum likelihood estimation (MLE) framework based on the log-likelihood function of the BKF. Given samples $\{x_i\}_{i=1}^{n}$, the log-likelihood function is:

$$\begin{aligned} L(p,c) &= \sum_{i=1}^{n} \ln f_X(x_i, p, c) \\ &= n\left(\frac{1}{2}\ln 2 - \frac{p}{2}\ln c - \ln\Gamma(p)\right) + \left(p - \frac{1}{2}\right)\sum_{i=1}^{n}\ln|x_i| + \sum_{i=1}^{n}\ln K_{p-\frac{1}{2}}\left(\sqrt{\frac{2}{c}}|x_i|\right) \end{aligned} \tag{2.3}$$

By setting partial derivatives with respect to p and c to zero, two nonlinear equations are obtained. The paper solves them numerically using the secant method to obtain $\hat{p}$ and $\hat{c}$.

The BKF model was empirically validated using Kolmogorov–Smirnov tests on real ultrasound images. Results demonstrated that BKF significantly outperformed Gaussian and Normal Inverse Gaussian models in both wavelet and Curvelet domains, achieving the lowest KS distances and most consistent alignment in PP-plots.

This approach provides a more accurate statistical description of speckle noise in transform domains and is particularly relevant for our work, where precise modeling of the noise distribution is critical for effective noise variance estimation and subsequent denoising performance.

2.1.2 Estimation Using Generalized Gamma Distribution[16]

It is widely accepted that the observed intensity of an L-look SAR image conditioned on the underlying reflectivity is gamma distributed:

doi:10.6342/NTU202502541

$$P(F \mid I) = \left(\frac{L}{I}\right)^{L} \frac{F^{L-1}}{\Gamma(L)} \exp\left(-\frac{LF}{I}\right) \tag{2.4}$$

Under this model, the multiplicative noise is also gamma distributed.

However, such an assumption imposes restrictive constraints, particularly in heterogeneous or highly textured areas. To address this limitation, recent work has proposed adopting the Generalized Gamma Distribution (GGD) as a more flexible statistical model for speckle noise. The GGD is a three-parameter distribution defined as:

$$f(x;\alpha,p,\beta) = \frac{p}{\beta^{\alpha}\Gamma\left(\frac{\alpha}{p}\right)} x^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^{p}\right], \quad x > 0 \tag{2.5}$$

where $\alpha > 0$ and $p > 0$ are shape parameters, and $\beta > 0$ is a scale parameter. This formulation includes the exponential($\alpha = p = 1$), gamma($p = 1$) and Weibull($\alpha = 1$) distributions as special cases, making it a unifying framework for speckle modeling.

The first- and second-order moments of a GGD random variable $X \sim GG(\alpha, p, \beta)$ are derived using the generalized moment ratio:

$$G_k = \frac{\Gamma\left(\frac{\alpha+k}{p}\right)}{\Gamma\left(\frac{\alpha}{p}\right)} \tag{2.6}$$

which leads to the expressions $E[x] = \beta \cdot G_1$, $Var[x] = \beta^2(G_2 - {G_1}^2)$

Parameter estimation for the GGD is performed via Maximum Likelihood Estimation (MLE). Given samples $\{x_i\}_{i=1}^{n}$, the log-likelihood function is:

$$\ln L(\alpha,p,\beta) = n\ln p - n\alpha\ln\beta - n\ln\Gamma\left(\frac{\alpha}{p}\right) + (\alpha-1)\sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{p} \tag{2.7}$$

doi:10.6342/NTU202502541

This leads to a set of nonlinear equations in α, p, β, which are solved numerically.

Experimental validation in the referenced study was conducted using a high-resolution Ku-band SAR image divided into six homogeneous regions. The fit quality of GGD was compared against gamma, Weibull, and log-normal models using a chi-square goodness-of-fit test. The results consistently favored the GGD in terms of lower chi-square statistics, confirming its superior adaptability in modeling SAR speckle across diverse textures.

2.1.3 Estimation Based on Gaussian-Hermite Moments[17]

Ma et al. propose to exploit orthogonal Gaussian–Hermite moments to quantify speckle noise intensity in images. A 2D Gaussian smoothing kernel

$$G(x,y)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{2.8}$$

is paired with Hermite polynomials $H_p(t)$, defined by

$$H_n(t)=(-1)^n e^{t^2}\frac{d^n}{dt^n}e^{-t^2} \tag{2.9}$$

to form the 2D Gaussian–Hermite moment of order (p,q):

$$M_{pq}=\iint I(x,y)G(x,y)H_p\left(\frac{x}{\sigma}\right)H_q\left(\frac{y}{\sigma}\right)dxdy \tag{2.10}$$

Here $I(x,y)$ is the image intensity.

From these moments, the authors select four orders, $(1,0),(0,1),(3,0),(0,3)$, to form at each pixel (x,y) a feature vector

$$f(x,y)=\left[w_{10}M_{10}, w_{01}M_{01}, w_{30}M_{30}, w_{03}M_{03}\right]^T \tag{2.11}$$

where w_{pq} are weights to balance sensitivity to horizontal (M_{10}, M_{30}) and vertical

doi:10.6342/NTU202502541

(M_{01}, M_{03}) structures. Empirically, when speckle of variance σ^2 is added to a homogeneous image of gray level g_i, the distribution of these feature vectors forms a roughly circular cluster centered at the origin, whose radius grows with noise intensity and with g_i.

To capture dispersion of the cluster, a noise characteristic value M_{uv} is defined as the average radial distance:

$$M_{uv} = \frac{1}{N}\sum_{i=1}^{N}\left\| f\left(x_i, y_i\right)\right\|_1 \tag{2.12}$$

where N is the number of pixels in the selected region.

By synthetically varying σ^2 on a uniform image of $g_i = 50$, the authors record pairs (M_{uv}, σ^2) and fit a fourth-order polynomial $\sigma^2 = 0.0013x^4 - 0.1269x^3 + 0.8903x^2 - 10.896x + 321.468$ for $x = M_{uv}$.

This method is notable for requiring no a priori statistical model of the scene beyond a small homogeneous patch, and for its computational efficiency compared to exhaustive matching or sliding-window filters.

2.2 Noise Removal

2.2.1 Lee Filter[11]

Lee (1980) proposed a class of non-recursive, pixel-wise enhancement and denoising algorithms based on local statistics (local mean and variance). The methods are computationally simple, well-suited for real-time and parallel hardware implementation.

doi:10.6342/NTU202502541

Given an input image X, for a pixel $x_{i,j}$, the local statistics are computed over a neighborhood window $(2n+1)\times(2m+1)$ as

$$\mu_{i,j} = \frac{1}{(2n+1)(2m+1)} \sum_{k=i-n}^{i+n} \sum_{l=j-m}^{j+m} x_{k,l} \tag{2.13}$$

$$\sigma^2{}_{i,j} = \frac{1}{(2n+1)(2m+1)} \sum_{k=i-n}^{i+n} \sum_{l=j-m}^{j+m} \left(x_{k,l} - \mu_{i,j}\right)^2 \tag{2.14}$$

where $\mu_{i,j}$ and $\sigma_{i,j}^2$ denote the local mean and variance respectively.

Lee proposed a simplified gain-based method:

$$y_{i,j} = \mu_{i,j} + k\left(x_{i,j} - \mu_{i,j}\right), \quad k \geq 0 \tag{2.15}$$

when $k > 1$ means high pass sharpening, $0 < k < 1$ means smoothing and $k = 0$ means local averaging.

Multiplicative noise in this paper is modeled as:

$$z_{i,j} = x_{i,j}\mu_{i,j}, \quad E[u] = \mu_u, \quad Var(u) = \sigma^2{}_u \tag{2.16}$$

By Taylor expansion and local linearization, the model becomes:

$$\begin{aligned} \hat{x}_{i,j} &= \mu_{i,j} + K_{i,j}\left(z_{i,j} - \mu_u\mu_{i,j}\right) \\ K_{i,j} &= \frac{\sigma^2{}_{i,j}\mu^2{}_u}{\sigma^2{}_{i,j}\mu^2{}_u + \sigma^2{}_u\mu^2{}_{i,j}} \end{aligned} \tag{2.17}$$

This adaptive formulation enables per-pixel denoising while respecting spatial variations in both signal and noise.

Despite its efficiency and widespread use in speckle noise reduction, the Lee filter suffers from several critical limitations. Its reliance on fixed-size local windows often leads to over smoothing around edges and fine structures, especially in heterogeneous regions.

doi:10.6342/NTU202502541

Moreover, it assumes local stationarity within each window, which fails near sharp transitions, causing either noise residuals or structural blurring. Additionally, the filter ignores non-local redundancies and requires manual or global noise variance estimation, which limits adaptability to spatially varying noise. These shortcomings motivate the need for more adaptive and structure-preserving denoising frameworks, particularly those that incorporate non-local similarity, automatic noise estimation, and enhanced filtering control to better balance noise suppression and detail preservation.

2.2.2 Frost Filter[12][13]

Frost et al. (1982) derives a statistical model for coherent radar imagery, showing that the observed intensity is degraded by multiplicative speckle noise. Under the assumption of local stationarity, they derive a minimum-mean-square-error(MMSE) spatial filter $M(f)$ in the frequency domain:

$$\mathrm{M(f)}=\frac{S_r(f)H^*(f)}{S_r(f)S_n(f)}=\frac{1}{1+\dfrac{S_n(f)}{S_r(f)}}H^*(f) \tag{2.18}$$

where S_r and S_n are the power spectral densities of r and n, respectively, and H^* is the complex conjugate of the system transfer function. Assuming an autoregressive model for r with variance σ_r^2 and correlation length l, and a white-chi-square model for n with variance σ_n^2, they obtain in the spatial domain an exponential (Frost) kernel:

$$m(x,y)=K_0\exp\left(-\alpha\sqrt{x^2+y^2}\right),\quad \alpha=\sqrt{\frac{\sigma_r{}^2+l^2}{\sigma_r{}^2 l^2}+\frac{1}{\sigma_n{}^2}} \tag{2.19}$$

Here α is adapted at each pixel via local estimates of the sample mean $\hat{\mu}$ and variance $\hat{\sigma}^2$. This “Adaptive Frost” filter yields minimum-MSE smoothing in homogeneous regions while automatically narrowing its footprint near edges to preserve structure.

doi:10.6342/NTU202502541

Although the MMSE-derived Frost filter effectively suppresses speckle while preserving edges, several weaknesses remain. First, the filter relies on a fixed, isotropic window size and cannot simultaneously adapt to both large homogeneous regions and fine-scale details, often over-smoothing small features or under-filtering broad areas. Second, it assumes locally Gaussian backscatter and chi-square noise statistics, assumptions that break down in highly textured, multi-class, or non-Gaussian scenes, leading to biased estimates. Third, the exponential weighting kernel does not explicitly account for edge orientation or sub-pixel geometry, so it can blur along strong contours or introduce directional artifacts when edges change rapidly. Fourth, because it operates purely in the single-scale spatial domain, it lacks the ability to distinguish high-frequency speckle from genuine fine details, limiting its fidelity on multi-scale textures. Finally, the need to estimate local mean and variance at every pixel and compute an adaptive exponential weight, which imposes a significant computational burden, hindering real-time or large-scale applications. Addressing these issues may require adaptive, multi-scale frameworks and data-driven statistical models that better capture the nonstationary, anisotropic characteristics of real-world SAR imagery.

2.2.3 SAR-BM3D [9][10]

Despite the recent surge in deep learning methods, BM3D (Block Matching and 3D Filtering) remains a cornerstone in traditional image denoising due to its non-local, adaptive, and transform-based design. It is particularly effective for Gaussian noise removal and has even been extended to tasks such as non-blind deblurring.

First, we talk about the classical framework. BM3D operates in two sequential stages:

doi:10.6342/NTU202502541

Stage 1: Basic Estimate:

For each reference image patch R, the algorithm identifies a set of similar patches across the image using a block matching strategy. These patches are stacked to form a 3D group, which is then transformed using a 2D Discrete Cosine Transform (DCT) followed by a 1D Haar wavelet transform along the third dimension. The resulting coefficients undergo hard-thresholding, acting as a shrinkage operation to suppress noise. After applying the inverse 3D transform, each denoised patch is placed back into the image space, and overlapping regions are aggregated with adaptive weights.

Stage 2: Wiener Refinement

The second step reuses the first stage estimate as a guide. New patch groups are formed: one from the noisy image, the other from the denoised estimate. These are again processed via 3D transforms, but instead of hard thresholding, a Wiener filter is applied, using prior knowledge of noise statistics and the estimate to adaptively filter each coefficient. The process again concludes with inverse transforms and weighted aggregation.

Formally, BM3D filtering can also be interpreted through an optimization lens, solving coupled problems in the transform and spatial domains. The goal is to minimize the deviation between the observed signal and its reconstruction, both in image space and transform space:

$$y^* = \arg\min_y \left[\frac{1}{2\sigma^2} \|z - Ay\|_2^2 + \frac{1}{2\gamma} \|y - \psi\omega^*\|_2^2 \right] \tag{2.20}$$

$$\omega^* = \arg\min_\omega \left[\tau \|\omega\|_p + \frac{1}{2\xi} \|\omega - \phi y^*\|_2^2 \right] \tag{2.21}$$

where A is the degradation operator (e.g., blur kernel), ψ and ϕ are 3D

doi:10.6342/NTU202502541

transform/inverse transform matrices, and is the spectrum of the grouped patches. This optimization is solved iteratively, balancing spatial fidelity and transform-domain regularity.

However, the SAR-BM3D algorithm extends BM3D to address multiplicative speckle noise, particularly in SAR imagery, where traditional Gaussian noise assumptions fail. Instead of adapting BM3D through a log transform (as done in homomorphic filtering), SAR-BM3D directly modifies the block matching and collaborative filtering stages to handle the statistical properties of speckle.

Key improvements and adaptations in SAR-BM3D include:

1. Speckle-Adaptive Block Matching

Unlike the standard Euclidean distance used in BM3D, SAR-BM3D introduces speckle-specific similarity metrics:

First-pass distance metric (likelihood-based):

$$d^{(1)}\left(x,y\right)=\log\left(\frac{x}{y}+\frac{y}{x}\right)=\log\left(\frac{x^2+y^2}{2xy}\right) \tag{2.22}$$

Second-pass distance metric (refined with estimate):

$$d^{(2)}\left(x,y,\hat{u},\hat{v}\right)=\log\left(\frac{x^2+y^2}{2xy}\right)+\frac{\gamma L}{2L-1}\cdot\frac{\left(\hat{u}^2-\hat{v}^2\right)^2}{\hat{u}^2\hat{v}^2} \tag{2.23}$$

where L is the number of looks in SAR imaging, and γ is an empirical tuning parameter.

These metrics are grounded in the Nakagami or Gamma distributions modeling SAR speckle, replacing the naive squared difference used in classical BM3D.

doi:10.6342/NTU202502541

2. Transform Domain and Filtering

SAR-BM3D performs collaborative filtering on grouped patches using:

First pass: An Undecimated Wavelet Transform (UDWT) for robustness and translation-invariance, followed by hard-thresholding.

Second pass: A Wiener filter adapted to the speckle domain, using the basic estimate to guide filtering decisions.

Each group undergoes a 3D transform, thresholding or Wiener filtering, and is then inverse-transformed and aggregated, just like in BM3D.

2.2.4 Bayesian Nonlocal Means Filter[19]

Zhong et al. extend the Bayesian nonlocal means (BNLM) framework to synthetic aperture radar images by integrating sigma preselection to mitigate the bias inherent in using noisy patches as priors and to strengthen detail preservation. The proposed filter computes the estimate over a refined neighborhood $N(x)$, where weights derive from a Bayesian risk minimization under fully developed speckle assumptions.

$$\hat{u}(x) = \frac{\sum_{y \in N(x)} p\left(v(x) \mid u'(y)\right) p\left(u'(y)\right) u'(y)}{\sum_{y \in N(x)} p\left(v(x) \mid u'(y)\right) p\left(u'(y)\right)} \tag{2.24}$$

where the likelihood factorizes over the $M \times M$ patch pixels:

$$p\left(v(x) \mid u(y)\right) = \prod_{m=1}^{M \times M} p\left(v_m(x) \mid u_m(y)\right) \tag{2.25}$$

and each multiplicative-speckle term follows an L-look gamma distribution:

$$p\left(v_m(x) \mid u_m(y)\right) = \frac{v_m(x)^{L-1}}{\Gamma(L)} \left(\frac{L}{u_m(y)}\right)^L \exp\left(-\frac{L v_m(x)}{u_m(y)}\right) \tag{2.26}$$

Taking logarithms and grouping yields a weight proportional to

doi:10.6342/NTU202502541

$$\exp\left(-\frac{1}{h^2}\sum_{m=1}^{M\times M}\left(\frac{v_m(x)}{u_m(y)}\right)+\ln\left(u_m(y)\right)-\frac{L-1}{L}\ln\left(v_m(x)\right)\right) \tag{2.27}$$

with smoothing parameter $h = k\sigma$, $\sigma = 1/\sqrt{L}$, and $k \approx 2$. This formulation naturally reflects the multiplicative noise model.

To reduce bias from substituting $u(y)$ by the noisy patch value, the authors introduce an a priori mean $u'(y)$ computed via a simple classification: if the coefficient of variation (CV) in a 3×3 window is below the speckle standard deviation σ, use the local mean; otherwise, keep the raw noisy value. Pixel preselection then applies a sigma range (I_1, I_2) around $u'(x)$, obtained by solving

$$\xi=\int_{I_1}^{I_2} p_s(s)ds, \quad 0 < I_1 < 1 < I_2 < \infty \tag{2.28}$$

for a chosen inclusion probability ζ.

Although sigma preselection successfully reduces bias and enhances edge and point-scatter preservation, it depends on fixed hyperparameters (patch size, search window, thresholds ζ and Θ) that may not generalize across diverse scenes. The binary rule for a priori mean estimation can still blur strong reflectors if the 3×3 estimate is applied indiscriminately, and numerical determination of the sigma interval requires empirical calibration. Computational demands remain high for large-scale or real-time applications, and the single-scale formulation lacks the ability to distinguish fine speckle from genuine high-frequency textures. Future extensions might explore adaptive threshold learning, multiscale representations, or integration of learned texture priors to further improve both accuracy and efficiency.

doi:10.6342/NTU202502541

2.2.5 Diffusion Probabilistic Model[18]

Perera et al. propose SAR-DDPM, a novel application of denoising diffusion probabilistic models (DDPMs) to the task of speckle noise reduction in SAR images.

In DDPM, a forward noising process gradually corrupts a clean image x_0 into noise x_T through:

$$q(x_t \mid x_{t-1}) = N\left(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I\right) \tag{2.29}$$

Where β_t is a fixed variance schedule. The marginal at any timestep t is:

$$q(x_t \mid x_0) = N\left(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)I\right), \text{ with } \bar{\alpha}_t = \prod_{i=1}^{t}(1-\beta_i) \tag{2.30}$$

A noisy image x_t can be directly sampled as:

$$x_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\varepsilon, \ \varepsilon \sim \mathrm{N}(0, I) \tag{2.31}$$

The reverse process learns to recover x_0 by predicting the added noise:

$$p_\theta(x_{t-1} \mid x_t) = N\left(x_{t-1}; \mu_\theta(x_t, t), \sum_\theta(x_t, t)\right) \tag{2.32}$$

Here, denotes the probability density of a multivariate Gaussian with mean μ and covarianceΣ, evaluated at x.

Instead of training by minimizing the simplified mean squared error (MSE) loss between the true noise and the predicted noise, the network is trained conditioned on the speckled image x_S to make the model effective for SAR despeckling.

$$\mathrm{L}_{\text{SAR-DDPM}} = E_{t,x_0,x_S,\varepsilon}\left[\left\|\varepsilon - \varepsilon_\theta(t, x_t, x_S)\right\|^2\right] \tag{2.33}$$

This allows the network to infer the clean image by learning to denoise samples generated from x_S.

doi:10.6342/NTU202502541

Chapter 3 Proposed Noise Estimation

3.1 Speckle Noise Model

The multiplicative degradation model of a speckle-corrupted image:

$$\mathrm{I}(i,j) = x(i,j) * n(i,j) \tag{3.1}$$

where $x(i,j)$ and $n(i,j)$ denote the noise-free image and the speckle noise, respectively.

3.2 Methodology

The noise estimation framework relies on two core functions: speckle_var (Figure 3.1 The algorithm flowchart for speckle_var) and speckle_est (Figure 3.2 The algorithm flowchart for speckle_est), which work sequentially to accurately estimate speckle noise variance in SAR images.

For each sub-band, speckle_var calculates the local noise variance using sliding windows centered around each pixel. The local variance computation is further adjusted through an adaptive weighting mechanism that assigns higher weights to homogeneous regions and down-weights edges or high-contrast areas. This pixel-level noise map provides the initial spatial distribution of noise variance for each sub-band.

After the three passes of speckle_var on the LL, LH, and HL sub-bands, the results are combined using a weighted averaging scheme to form an overall noise variance estimate. The weights are determined based on the statistical characteristics of each sub-band,

doi:10.6342/NTU202502541

ensuring that sub-bands with more reliable noise information contribute more to the final estimate.

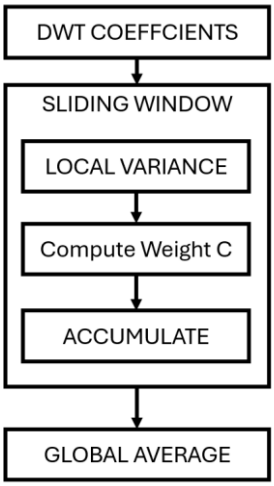


Figure 3.1 The algorithm flowchart for speckle_var

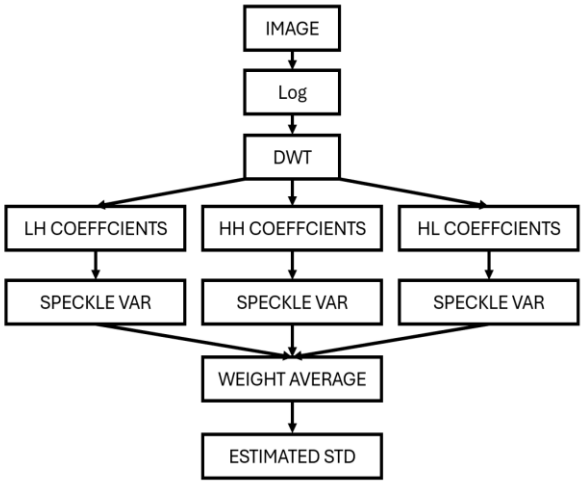


Figure 3.2 The algorithm flowchart for speckle_est

3.2.1 Log-Transform

Since speckle noise is modeled as (3.1). To convert this multiplicative relationship into an additive one, we apply the logarithmic operation:

$$I'(i,j) = Log\left(I(i,j)+1\right) \tag{3.2}$$

doi:10.6342/NTU202502541

where we add 1 inside the logarithm to avoid issues when $I(i,j)$ approaches zero. After the log-transform, the noise model approximately becomes:

$$I'(i,j) \approx Log\left(x(i,j)\right) + Log\left(n(i,j)\right) \tag{3.3}$$

Hence, the noise component transforms to an additive term $Log(n(i,j))$, making subsequent steps, like wavelet-based noise isolation, more straightforward.

3.2.2 Discrete Wavelet Transform

The two-dimensional Discrete Wavelet Transform (DWT) decomposes an image into four sub-bands by applying high-pass and low-pass filters in both horizontal and vertical directions. Assuming a signal x passes through a low-pass filter with impulse response g, the convolution operation is defined in (3.4).

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] \bullet g[n-k] \tag{3.4}$$

After down-sampling by a factor of two, the low-frequency (LP) and high-frequency (HP) components can be extracted using the low-pass filter g and high-pass filter h, as shown in (3.5) and (3.6).

$$y_{LP}[n] = \sum_{k=-\infty}^{\infty} x[k] \bullet g[2n-k] \tag{3.5}$$

$$y_{HP}[n] = \sum_{k=-\infty}^{\infty} x[k] \bullet h[2n-k] \tag{3.6}$$

This process results in four sub-bands: LL, LH, HL, and HH, each representing different spatial and frequency characteristics. The LL sub-band contains the low-frequency components, providing a coarse approximation of the image, while the LH, HL, and HH

doi:10.6342/NTU202502541

sub-bands capture high-frequency details in horizontal, vertical, and diagonal orientations, respectively. Since noise is often concentrated in the high-frequency components, DWT enables the extraction of noise elements for further processing.

In practice, common wavelet bases include Haar and Daubechies wavelets. Haar wavelet is widely used due to its simplicity and low computational cost. However, for this study, which focuses on SAR images, the sym4 wavelet from the Symlets family was selected as the decomposition basis. The choice of sym4 is motivated by its symmetry, which provides better boundary handling and feature preservation compared to other wavelet bases. Moreover, sym4 retains the high-order vanishing moments characteristic of Daubechies wavelets, ensuring effective noise suppression while maintaining image details. This makes it particularly well-suited for SAR image processing, where preserving high-frequency details and achieving smooth, stable decompositions are crucial.

3.2.3 Local Variance Estimation and Accumulation

For each pixel (i, j), a $(2n+1) \times (2n+1)$ neighborhood (or "local window") around that pixel is considered. The variance of the intensities in this local window is computed as:

$$\mathrm{var}_{local} = \frac{1}{|\Omega|} \sum_{(u,v)\in\Omega} \left(I(u,v) - \overline{I_{\Omega}} \right)^2 \tag{3.7}$$

where Ω denotes the set of coordinates in the $(2n+1) \times (2n+1)$ neighborhood around (i, j).

Next, a parameter σ_{noise} serves as an initial guess or prior estimate of the noise variance.

doi:10.6342/NTU202502541

Using var_{local} and σ_{noise}, we define:

$$C = \frac{\text{var}_{local}}{\text{var}_{local} + \sigma_{noise}} \tag{3.8}$$

In this study, the initial noise variance σ_{noise} is estimated based on the Equivalent Number of Looks (ENL), which characterizes the level of speckle noise in SAR images. Given that our target images are urban airports, we rely on prior knowledge from literature, where typical ENL values for urban or complex terrain are generally low due to high structural variability and limited spatial averaging. Specifically, we set the ENL to 0.5 for initial estimation. This choice reflects the high noise level typically observed in such environments, and the corresponding noise variance is calculated using:

$$\sigma_{noise} = \frac{1}{\sqrt{ENL}} \tag{3.9}$$

The intuition is as follows:

If $var_{local} \gg \sigma_{noise}$, then C will be close to 1. This suggests that the local variance is quite large, likely due to edges or texture rather than just noise.

If $var_{local} \approx \sigma_{noise}$, then C will be moderate, indicating the local variance is in the same range as the noise variance.

If $var_{local} \ll \sigma_{noise}$, then C will be close to 0, indicating that noise dominates, and the algorithm will be more inclined to calculate that region.

To refine the overall noise estimate, we accumulate a weighted version of each local variance into a running total, noise_var_sum. The weight chosen here is $(1 - C)^2$, giving

$$\text{noise\_var\_sum} = (1 - C)^2 \times \text{var}_{local} \tag{3.10}$$

doi:10.6342/NTU202502541

When local variance is much larger than σ_{noise}, $C \approx 1$, so $(1-C)^2 \approx 0$. This means such pixels contribute very little to noise_var_sum, since they likely represent structure or edges rather than just noise.

When local variance is on the order of σ_{noise}, C is smaller, so $(1-C)^2$ is larger. These areas are more indicative of genuine noise, and thus they add more weight to the accumulated noise estimate.

After summing over all pixels in the image (or over all valid pixels where the local window can be computed), the global noise variance estimate is obtained by dividing the accumulated sum by the number of pixels, mathematically:

$$\hat{\sigma}_{init}^{\ 2} = \frac{\sum_{i,j}\left[\left(1-C(i,j)\right)^2 \times \mathrm{var}_{local}(i,j)\right]}{\text{total\_pixel\_count}} \tag{3.11}$$

This final step averages the local noise-indicating contributions across the whole image. The result is an overall estimate of the noise variance $\hat{\sigma}_{init}^2$.

3.2.4 Final Noise Estimation Using Polynomial Regression

Polynomial regression is a technique employed to model a set of data points by constructing a mathematical function or graph that best fits the given data.

This modeling process may include specific constraints or parameters, which can potentially influence the overall pattern of the resulting function or graph. Widely recognized within the research community, polynomial regression serves as a prominent

doi:10.6342/NTU202502541

statistical approach for addressing non-linear system outputs.

Therefore, we derive the polynomial regression coefficients from a set of experimental speckle images specifically chosen for their diverse real-world texture variations. We introduce known noise levels σ_{added} into these images, then calculate initial noise estimates σ_{init} using (3.11). Repeating this process over 100 trials and a broad range of noise levels produces a robust σ_{init} dataset for each image. By averaging these results elementwise, we obtain a single reference vector. This reference vector, in conjunction with the corresponding σ_{added} values, is used to build a polynomial mapping function.

3.3 Experiment and Results

The experiments were conducted using SAR airport images from the publicly available RadarS SAR Dataset[6].To maintain consistency and fair comparison with the Gaussian-Hermite approach[7] and the Generalized Gamma method[8], we introduced controlled levels of speckle noise and measured the estimation errors using the same metric.

3.3.1 Evaluation Metric

To assess the effectiveness of the proposed algorithm, Speckle noise with Matlab "imnoise" function was applied to all tested images at varying levels. The noise standard deviation ranged from 10 to 100 in increments of 10. The performance evaluation was conducted using an error metric derived from:

$$Error(\%) = \frac{\left|\sigma_{estimation} - \sigma_{added\_noise}\right|}{\sigma_{added\_noise}} \times 100\% \tag{3.12}$$

doi:10.6342/NTU202502541

3.3.2 Experimental Results of Speckle Noise Estimation

We first aggregated the errors across all noise levels (std = 10 to 100) and all images in our test suite, arriving at an average estimation error. As seen in Table 3.1, our proposed method achieves an average noise estimation error of 2.6644%, significantly outperforming both the GH-based approach (63.5502%) and the GGD-based approach (69.9423%).

Table 3.1 Comparison of average noise estimation error rates (%)
for all levels of added speckle noise on all images

	Proposed	GH-based	GGD-based
Average Error	2.6644	63.5502	69.9423

This pronounced difference highlights the effectiveness of our local variance-driven weighting strategy combined with DWT- based analysis, which enables better capture of the multiplicative nature of speckle noise. The GH-based and GGD-based methods, on the other hand, struggle with highly textured and heterogeneous regions, leading to substantially higher estimation errors.

A more granular view of the estimation errors for each image is provided in Table 3.2. Our approach consistently delivers low estimation errors, even in images with complex textures, whereas both the GH-based and GGD-based techniques exhibit highly variable errors. Notably, in datasets like BEIJING_shahejichang_2, our proposed method attains a near-ideal error of 1.3142%, whereas the GH-based method produces a much larger 85.636% error, and the GGD-based method is also highly inaccurate at 78.745%. Similar trends emerge across all images, reinforcing the robustness and adaptability of our local

doi:10.6342/NTU202502541

variance estimation and polynomial regression framework.

Table 3.2 Comparison of average noise estimation error rates (%) for all levels of added speckle noise for individual datasets

	Proposed	GH-based	GGD-based
BEIJING_shahejichang_1	3.1332	36.4373	45.109
BEIJING_shahejichang_2	1.3142	85.636	78.745
BEIJING_shahejichang_3	2.5696	46.7041	76.493
BEIJING_shahejichang_4	4.1386	38.6677	72.049
BEIJING_shahejichang_5	3.1850	35.466	45.254
BEIJING_shahejichang_6	1.5565	66.3589	44.299
BEIJING_xijiaojichang_1	1.4174	22.1305	41.655
BEIJING_xijiaojichang_2	4.6972	24.8504	77.572
BEIJING_xijiaojichang_3	1.5758	22.8935	36.875
Guanghan_jichang_1	6.4368	50.182	42.762
Guanghan_jichang_2	1.7572	226.2633	70.485
Guanghan_jichang_3	1.0650	104.5253	71.67
Guanghan_jichang_4	1.7643	66.038	69.13

Figure 3.3 and Figure 3.4 illustrate the noise estimation trends for selected datasets, where our proposed method closely tracks the ground truth across all noise levels. In contrast, the GH-based and GGD-based approaches consistently exhibit substantial deviations. In Figure 3.5, we provide an aggregated performance comparison across all tested datasets, further confirming the superior tracking accuracy and stability of our proposed solution.

doi:10.6342/NTU202502541

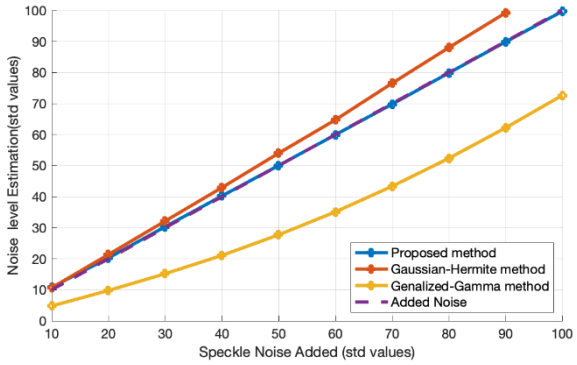


Figure 3.3 Benchmarking of three methods on BEIJING_xijiaojichang_1 dataset

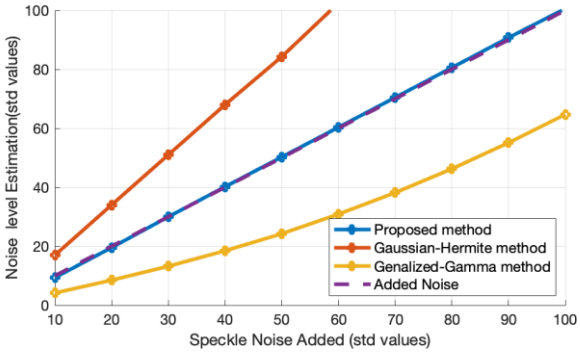


Figure 3.4 Benchmarking of three methods on BEIJING_shahejichang_2 dataset

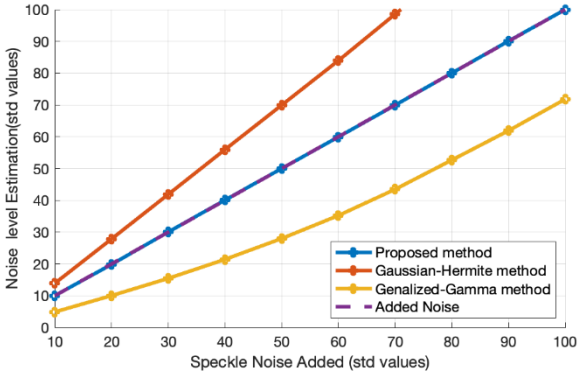


Figure 3.5 Aggregated comparison of all three methods across datasets

doi:10.6342/NTU202502541

3.3.3 Summary of Experimental Findings

Our proposed DWT-based speckle noise estimation method achieves an average estimation error of only 2.6644%, significantly outperforming both the Gaussian-Hermite method (63.5502%) and the Generalized Gamma method (69.9423%). The results demonstrate that our local variance-driven weighting strategy and polynomial regression refinement provide a substantial advantage in accurately capturing the multiplicative nature of speckle noise. These findings suggest that our approach is a more reliable solution for real-world SAR image analysis, improving noise estimation accuracy and enhancing subsequent image processing tasks.

This significant reduction in error is evident both in aggregate evaluations and on a per-image basis, especially in challenging scenarios with complex textures. From a technical standpoint, the method offers a robust framework for accurately capturing the multiplicative nature of speckle noise, ensuring that subsequent image processing tasks such as feature extraction and classification can be performed with higher fidelity. On a commercial level, the high precision and reliability of our approach translate into substantial cost efficiencies by reducing computational overhead and improving processing times, and offer a competitive edge for deployment in advanced imaging systems. Moreover, the versatility of the method suggests promising applications beyond SAR imaging, including ultrasound and other high-demand fields where noise reduction is critical.

In summary, this work not only sets a new benchmark in speckle noise estimation but also paves the way for future innovations in image processing technologies, ensuring both technical excellence and significant market impact.

doi:10.6342/NTU202502541

Chapter 4 Proposed Speckle Noise Removal

4.1 A New Form BNL Filter

Let v(x) represent the noisy observation at pixel, and let u(x) denote the noise-free reflectance (intensity or amplitude) at the corresponding pixel in a SAR image. Additionally, the notations v(x) and u(x) are also used to represent the vectorized patches centered at pixel x with dimensions. The multiplicative speckle noise observed in SAR images can thus be mathematically formulated as: $v(x) = u(x) \cdot s(x)$, where s(x) represents the multiplicative speckle component, typically modeled by a Gamma distribution governed by the number of looks L.

The refined BNLM estimate $\hat{u}(x)$ [1] is computed pixel-wise as the weighted average of all values u(y) in the neighborhood $\Delta(x)$ around pixel x. This is formally given by

$$\hat{u}(x) = \frac{\sum_{y \in \Delta(x)} p\big(v(x) \mid u(y)\big)\, p\big(u(y)\big) u(y)}{\sum_{y \in \Delta(x)} p\big(v(x) \mid u(y)\big)\, p\big(u(y)\big)} \tag{4.1}$$

where $\hat{u}(x)$ is obtain pixelwise as the weighted average of all gray values $u(y)$ in the neighborhood $\Delta(x)$ of x. The term $p(v(x)|u(y))p(u(y))$ acts as the similarity measure between $v(x)$ and $u(y)$

We illustrate the formulation of the modified BNL filter using an intensity image as our example, noting that the amplitude scenario yields a comparable formulation. Under the assumption of fully developed, statistically independent speckle, the conditional distribution $p(v(x)|u(y))$ introduced in equation (4.1) can be rewritten as

doi:10.6342/NTU202502541

$$p\left(v(x)\,|\,u(y)\right)=\prod_{n=1}^{N\times N}p\left(v_n(x)\,|\,u_n(y)\right) \tag{4.2}$$

where $u_n(y)$ and $v_n(x)$ noting the n th pixel in the corresponding patches, respectively. Assuming that $u_n(y)$ belongs to the set of potential reflectance values corresponding to $v_n(x)$, the conditional probability density function $p(v_n(x)|u_n(y))$ for an intensity SAR image with L-looks can be formulated as shown in [2][3]

$$p\left(v_n(x)\,|\,u_n(y)\right)=\frac{v_n(x)^{L-1}}{\Gamma(L)}\left(\frac{L}{u_n(y)}\right)^{L}\exp\left(-\frac{Lv_n(x)}{u_n(y)}\right) \tag{4.3}$$

where $\Gamma(\cdot)$ is the gamma function. Based on (4.3), (4.2) can be rewritten as [4]

$$p\left(v(x)\,|\,u(y)\right)\propto\exp\left(-\frac{1}{h^2}\sum_{n=1}^{N\times N}\left(\frac{v_n(x)}{u_n(y)}\right)+\ln\left(u_n(y)\right)-\frac{L-1}{L}\ln\left(v_n(x)\right)\right) \tag{4.4}$$

where h is the smoothing parameter.

While the original BNLM formulation employs a similarity metric defined in the spatial domain, this formulation incurs substantial computational overhead due to the use of explicit division and logarithmic operations on raw pixel values. In contrast, by transitioning to the logarithmic domain via variable substitution and carefully deriving the transformed likelihood, we obtain a new distance metric that retains the statistical properties of the original model but enables much more efficient computation.

Taking the logarithm and summing across all patch elements yields the linear-domain distance measure based on (4.4):

$$d_{lin}(x,y)=\sum_{n\in patch}\left[\frac{v_n(x)}{u_n(y)}+\ln u_n(y)-\frac{L-1}{L}\ln v_n(x)\right] \tag{4.5}$$

Let $s_n=\log v_n$, $t_n=\log u_n$. Then $v_n=e^{s_n}$, $u_n=e^{t_n}$. The transformation of

doi:10.6342/NTU202502541

variables yields:

$$p\left(s_n \mid t_n\right) = p\left(v_n = e^{s_n} \mid u_n = e^{t_n}\right)\left|\frac{dv_n}{ds_n}\right| = p\left(v_n \mid u_n\right)e^{s_n} \tag{4.6}$$

Based on (4.6), (4.4) can be written as:

$$\begin{aligned} p\left(s_n \mid t_n\right) &= \frac{L^L}{\Gamma(L)}\frac{e^{s_n(L-1)}}{e^{t_n L}}\exp\left(-L\frac{e^{s_n}}{e^{t_n}}\right)e^{s_n} \\ &= \frac{L^L}{\Gamma(L)}\exp\left[s_n\left(L-1\right)+s_n - Lt_n - Le^{s_n-t_n}\right] \\ &= \frac{L^L}{\Gamma(L)}\exp\left[L\left(s_n - t_n\right) - Le^{s_n-t_n}\right] \\ &\propto \exp\left[-L\left(e^{s_n-t_n} - \left(s_n - t_n\right)\right)\right] \end{aligned} \tag{4.7}$$

Assuming independence across patch pixels, the joint conditional density leads to a log-domain distance based on (4.7):

$$\begin{aligned} d_{\log}\left(x,y\right) &= -\ln\prod_{n\in patch} p\left(s_n\left(x\right)\mid t_n\left(y\right)\right) \\ &= L\sum_{n\in patch}\left[e^{s_n(x)-t_n(y)} - \left(s_n\left(x\right)-t_n\left(y\right)\right)\right] \end{aligned} \tag{4.8}$$

This becomes the dedicated distance metric for SAR despeckling under the log-domain formulation.

The result is constructed identically to the BNLM framework:

$$p_{\log}\left(v\left(x\right)\mid u\left(y\right)\right) \propto \exp\left[-\frac{1}{h^2}L\sum_{n\in patch}\left[e^{s_n(x)-t_n(y)} - \left(s_n\left(x\right)-t_n\left(y\right)\right)\right]\right] \tag{4.9}$$

Let σ^2 denote the variance of speckle noise. As commonly adopted in related work such as[5], the smoothing parameter h is modeled to scale linearly with the noise level, i.e., $h = k\sigma$, where $k \approx 2$ is found to be effective for intensity SAR imagery. At the same time, the speckle standard deviation is estimated by our previous noise estimation algorithm.

doi:10.6342/NTU202502541

Previous works such as **錯誤! 找不到參照來源。** commonly assume a uniform prior distribution over $u(y)$ to simplify the Bayesian estimator, i.e., $p\big(u(y)\big) = 1/|\Delta(x)|$, where |Δ(x)| denotes the number of candidates in the search region. Additionally, earlier BNLM implementations tend to approximate both the unknown reflectance $u(y)$ and the target $u(x)$ by their noisy counterparts $v(y)$ and $v(x)$, respectively. While this substitution simplifies computation, it inevitably introduces estimation bias, especially in high-variance or heterogeneous regions.

In contrast, our method proposes a refined estimation strategy where the prior estimate $u'(y)$ is not derived from local mean reflectance, but rather from the output of a preliminary Wiener filter applied to the speckled image. This choice provides a more stable and informative priori estimate that leverages local variance characteristics to enhance noise suppression while retaining signal structures.

Further, we integrate a sketch-based structure-aware selection mechanism to guide the computation of similarity weights. By analyzing gradient orientation and edge strength, this mechanism designates whether a pixel belongs to a structural region. If so, anisotropic weighting based on geometric alignment is applied, enabling better preservation of directional features such as edges and textures. Critically, rather than assuming a uniform prior $p(u'(y))$, our approach treats $p(u'(y)) \propto G(x, y)$, where $G(x, y)$ is the sketch-based anisotropic Gaussian kernel, thus embedding structural information directly into the prior. The aggregation subset $N(x) \subset \Delta(x)$ is thereby refined not only by intensity similarity but also by structural conformity.

doi:10.6342/NTU202502541

Together, these two contributions (1) replacing the prior mean with Wiener-based prior estimation, and (2) applying sketch-based structural referencing, yield the improved estimator:

$$\hat{u}(x)=\frac{\sum_{y\in N(x)} p\big(v(x)\,|\,u'(y)\big)G(x,y)\,u'(y)}{\sum_{y\in N(x)} p\big(v(x)\,|\,u'(y)\big)G(x,y)} \tag{4.10}$$

where

$p(v(x)|u'(y))$ is the usual log-domain likelihood term

$G(x,y)$ is the sketch-based anisotropic Gaussian kernel (so that structural patches are weighted more heavily), and

$N(x)$ is the preselected subset of $\Delta(x)$.

4.2 A Wiener-based Prior Estimation

In the original BNLM framework, the a priori mean $u'(y)$ is computed via a local σ-filter that aggregates neighboring pixel intensities to approximate the true reflectance before speckle corruption. This local mean plays a critical role in reducing estimation bias by supplying a more accurate representation of $u(y)$ when evaluating the likelihood $p(v(x)|u(y))$. However, σ-filter–based means are themselves susceptible to residual speckle and may not optimally balance noise suppression with detail preservation in heterogeneous SAR regions. In contrast, our proposed method replaces this σ-filter mean with the output of a Wiener filter applied to the raw speckled image. Specifically, by first executing a spatially adaptive Wiener filter, parameterized by local variance estimates, we obtain a smoothed image $u_w(y)$ that better approximates the underlying noise-free reflectance. This Wiener-filtered image serves as the new prior estimate:

doi:10.6342/NTU202502541

$$u'(y) = u_w(y) = Wiener(v(y)) \tag{4.11}$$

Because the Wiener filter incorporates both local mean and variance information, its output suppresses speckle more effectively than a simple mean filter while retaining fine structural details. Consequently, when computing nonlocal weights, we substitute $u'(y) = u_w(y)$ into the likelihood term $p(v(x)|u'(y))$, thereby reducing bias and improving denoising accuracy. This Wiener-based prior estimation not only aligns with the Bayesian rationale of using the best available local estimate but also demonstrates significantly enhanced performance in preserving texture and edge information compared to the original σ-filter approach.

4.3 Improved Preselection Based on the Sigma Range

In the original BNLM formulation, pixel-level preselection is used to restrict the candidate set $N(x) \subset \Delta(x)$ by exploiting the statistical properties of multiplicative speckle. Specifically, after computing the local a priori mean $u'(y)$ (obtained via a σ-filter), one defines a threshold

$$T = \frac{v_{\max}}{2} \tag{4.12}$$

where v_{max} is the maximum possible SAR intensity. If $u'(y) > T$ the algorithm determines a multiplicative interval $[I_1,\ I_2]$ such that

$$\int_{I_1}^{I_2} p_s(s)ds = \xi, \quad 0 < I_1 < 1 < I_2 < \infty \tag{4.13}$$

and $p_s(s)$ the Gamma-distributed speckle model with L looks. In practice, ξ is chosen close to 0.90~0.95 to cover the bulk of the speckle distribution without including outliers **錯誤！找不到參照來源。**. Each candidate pixel y is included in $N(x)$ only if its

doi:10.6342/NTU202502541

observed intensity $v(y)$ satisfies

$$u'(x)I_1 < v(y) < u'(x)I_2 \tag{4.14}$$

i.e., if $\frac{v(y)}{u'(x)} \in [I_1, I_2]$. Conversely, when $u'(x) \leq T$, the prior mean is deemed too small to reliably estimate a σ-interval, so one simply retains all $y \in \Delta(x)$, which means no preselection, to avoid "dark-area" overpruning.

We refine this preselection step by substituting the σ-filter–based mean with our Wiener-filtered prior $u_w(x)$. Since the Wiener output generally exhibits reduced speckle and more accurate local variance estimates, the resulting interval $[I_1, I_2]$ computed from $u_w(x)$ is tighter around the true reflectance. In other words, for $u_w(x) > T$, we solve equation (4.13) in exactly the same way, but now $u'(x) = u_w(x)$. Consequently, candidate pixels y are required to satisfy equation (4.14).

4.4 A Sketch-Based Structural Prior

In the original BNLM framework, the prior distribution over candidate patches $u(y)$ is assumed uniform

$$p(u(y)) = \frac{1}{|\Delta(x)|} \tag{4.15}$$

where $|\Delta(x)|$ denotes the total number of pixels in the search window. Under this assumption, all pixels in $\Delta(x)$ are treated equally, and the nonlocal weight reduces to the likelihood term alone:

$$w_{ori}(x, y) = p(v(x) \mid u'(y)) \tag{4.16}$$

While this simplification reduces the computational burden, it completely ignores any

doi:10.6342/NTU202502541

geometric or structural information, leading to over-smoothing along edges or thin linear features.

To address these shortcomings, we introduce a sketch-based structural prior that replaces the uniform assumption. Let $G(x)$ denote an anisotropic Gaussian kernel constructed from local gradient orientation $\theta(x)$ (extracted via Sobel) and edge strength (thresholded by Otsu). We now set $p(u'(y)) \propto G(x,y)$, which biases the prior in favor of pixels y aligned with the same local structure as x. Hence, the combined nonlocal weight becomes

$$w(x,y) = p\big(v(x) \mid u'(y)\big) G(x,y) \tag{4.17}$$

where $u'(y)$ is the Wiener-filtered prior estimate of the unknown reflectance at y. Equivalently, one may normalize $G(x,y)$ over the entire search region:

$$p\big(u(y)\big) = \frac{G(x,y)}{\sum_{y' \in \Delta(x)} G(x,y')} \tag{4.18}$$

So that

$$\begin{aligned} w(x,y) &= p\big(v(x) \mid u'(x)\big)\, p\big(u(y)\big) \\ &\propto p\big(v(x) \mid u'(x)\big) G(x,y) \end{aligned} \tag{4.19}$$

4.4.1 Sketch Map Extraction

First, we compute Sobel gradient magnitude $G_{mag}(x)$ and orientation $\theta(x)$ at each pixel. Apply Otsu's method to threshold $G_{mag}(x)$, yielding a binary edge mask $E(x)$. Then we define the sketch map $S(x)$ as

$$S(x) = \big[G_{mag}(x) > \tau \big] \vee E(x) \tag{4.20}$$

where τ is a chosen gradient threshold, ensuring strong structural cues are captured.

doi:10.6342/NTU202502541

4.4.2 Anisotropic Gaussian Kernel

Inspired by [14], we want one axis of the Gaussian to lie along the structure (the "major" axis, with larger variance), and the other axis across the structure (the "minor" axis, with smaller variance). To do this: Let $\Delta x = x_x - y_x$ and $\Delta y = x_y - y_y$ be the relative offsets from pixel x to candidate y. We can rotate these offsets by the local orientation $\theta = \theta(x)$:

$$f_1 = -\Delta y \sin\theta + \Delta x \cos\theta,\ f_2 = \Delta y \cos\theta + \Delta x \sin\theta \tag{4.21}$$

Choose a minor-axis standard deviation σ_s (e.g., 1–2 pixels) and an elongation factor $\lambda > 1$.

Then we define

$$G(x,y) = \exp\left(-\left(\frac{f_1^2}{\sigma_s^2} + \frac{f_2^2}{(\lambda\sigma_s)^2}\right)\right) \tag{4.22}$$

Here, the major axis (lengthened by λ) aligns with the local structure, allowing smoothing along edges, while the minor axis remains narrow to avoid averaging across the edge.

In other words, across-structure weight (term with f_1) penalizes large offsets perpendicular to the edge and along-structure weight (term with f_2) decays more slowly along the edge direction encouraging stretching of the kernel along the line or edge.

If candidate y lies along the same structure direction as x, then $f_1 \approx 0$ and $|f_2|$ may be moderate, yielding a relatively large $G(x,y)$. Such pixels receive increased weight in the BNLM average.

doi:10.6342/NTU202502541

Conversely, if y rosses a structural boundary, f_1 becomes large, causing $G(x,y)$ to become exponentially small. These cross-boundary pixels thus contribute negligibly, preserving edge sharpness.

By embedding this sketch-based prior into the BNLM algorithm, we ensure that $w(x,y)$ strongly favors pixels that both statistically align in intensity (via the likelihood) and geometrically align in structure (via $G(x,y)$). The combined estimator $u'(x)$ in (4.10) thus achieves a superior balance between despeckling and edge/line retention, outperforming the original uniform-prior BNLM in preserving fine geometric details.

4.5 Experimental Results

In order to quantitatively assess despeckling performance, we use two widely adopted image-quality metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). PSNR evaluates the pixel-wise fidelity between the denoised output and the original, pre-noise image (i.e., the dataset image before synthetic speckle was added). Mathematically,

$$\text{PSNR}=10\log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{4.23}$$

where MAX is the maximum possible pixel value (e.g., 1.0 in normalized images) and MSE is the mean squared error between the filtered result and the original image. A higher PSNR indicates closer approximation to the original appearance. SSIM, on the other hand, measures perceived structural similarity by combining local luminance, contrast, and structural comparisons; it ranges from 0 to 1, with larger values indicating better preservation of edges and textures. Together, PSNR and SSIM provide complementary

doi:10.6342/NTU202502541

perspectives: PSNR quantifies overall error reduction relative to the original image, while SSIM gauges how well geometric and textural features are maintained.

To rigorously validate the proposed sketch-based BNLM framework, we evaluated performance on a set of 208 synthetic SAR airport images from the publicly available RadarS SAR Dataset. All images were processed under identical environment and parameters. We compared our method against four well-accepted despeckling techniques:

1. Lee Filter
2. Frost Filter
3. SAR-BM3D
4. Original Bayesian Nonlocal Means Filter
5. Diffusion Model

The following subsections present quantitative metrics (PSNR, SSIM), visual comparisons, and runtime analysis.

4.5.1 Quantitative Metrics

We first report the average PSNR and SSIM values obtained by each method over all 208 test images.

Table 4.1 Average PSNR and SSIM across all test images

Method	PSNR (dB)	SSIM
Noisy Input	15.053	0.3437
Lee Filter	20.833	0.5366
Frost Filter	21.049	0.5593
SAR-BM3D	21.017	0.5297
Original BNLM	22.2286	0.6987

doi:10.6342/NTU202502541

Diffusion Model	15.0546	0.3511
Proposed	23.1713	0.7797

Next, Table 2 isolates the PSNR improvement (Δ_{PSNR}) of each method relative to the noisy baseline:

Table 4.2 Average PSNR Gain over Noisy Input

Method	Δ_{PSNR} (dB)
Lee Filter	+5.7800
Frost Filter	+5.9960
SAR-BM3D	+5.9640
Original BNLM	+7.1756
Diffusion Model	+0.0016
Proposed	+8.1183

4.5.2 Visual Results

Quantitative metrics provide objective comparisons, but visual inspection remains essential to gauge perceptual quality. Six example images from the dataset are shown to as below.

In the figures, Lee filters oversmooth the edges, causing noticeable blurring. Frost better preserves edges a little bit, and SAR-BM3D introduces block-like texture artifacts along straight lines. Original BNLM maintains reasonable edge sharpness yet slightly blurs thin lines due to its isotropic prior.

Our method retains the boundaries and subtle textural variations, owing to the anisotropic

doi:10.6342/NTU202502541

kernel alignment guided by the sketch map. The annotated SSIM of 0.7770 confirms superior structure preservation.

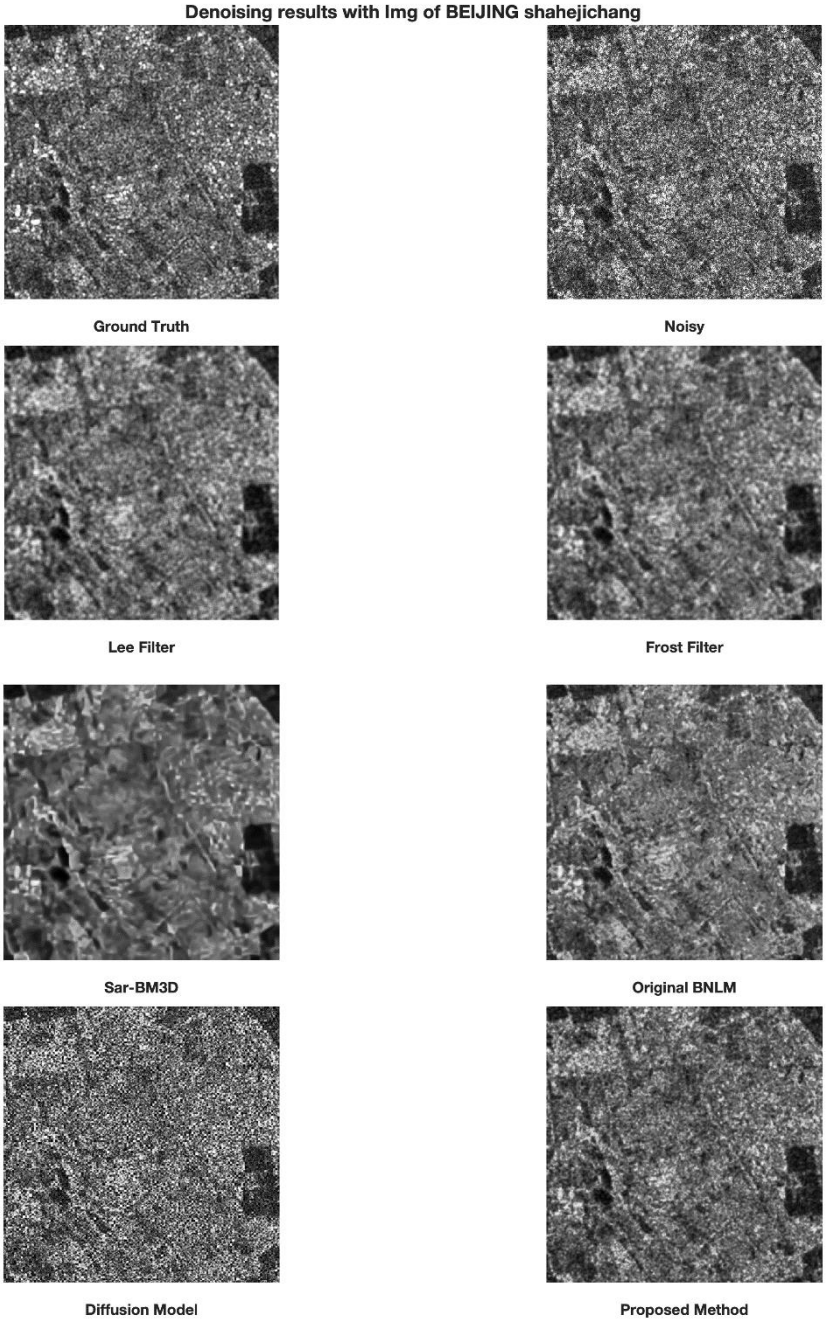


Figure 4.1 Denoising results of Image “BEIJING_shahejichang_2_11”

doi:10.6342/NTU202502541

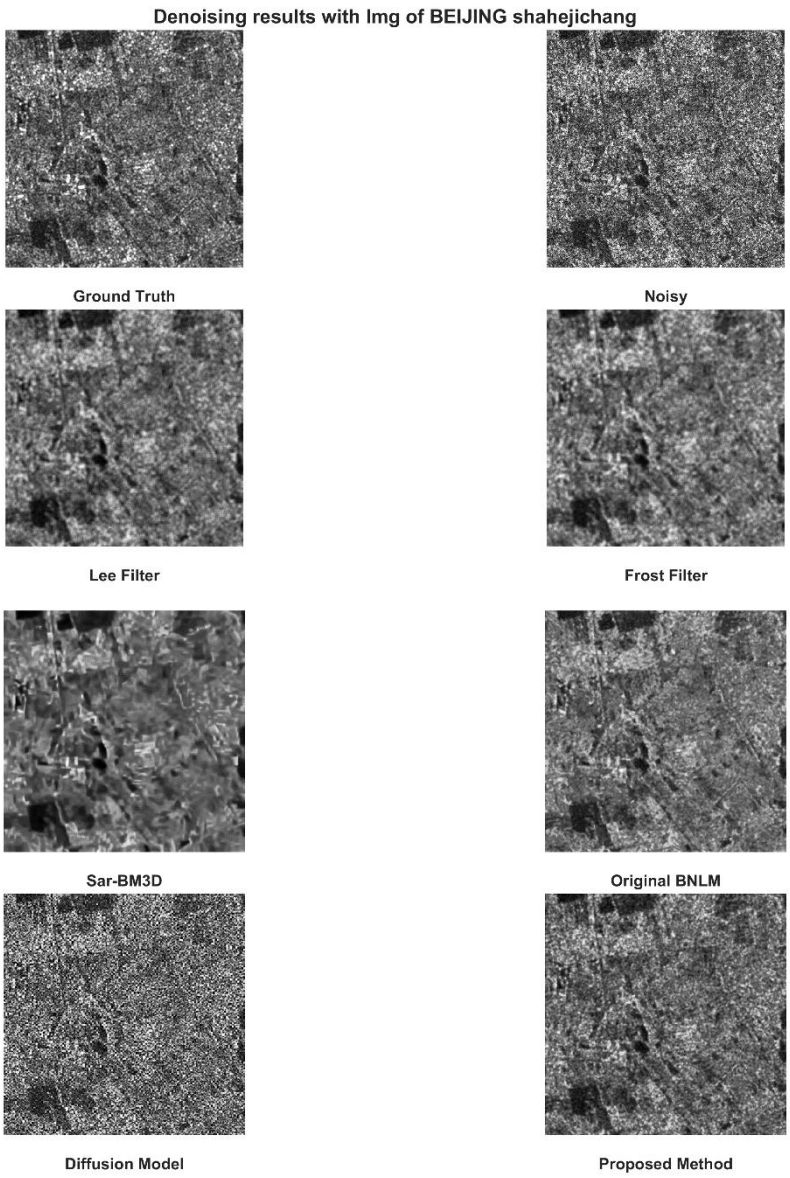


Figure 4.2 Denoising results of Image “BEIJING_shahejichang_1_16”

doi:10.6342/NTU202502541

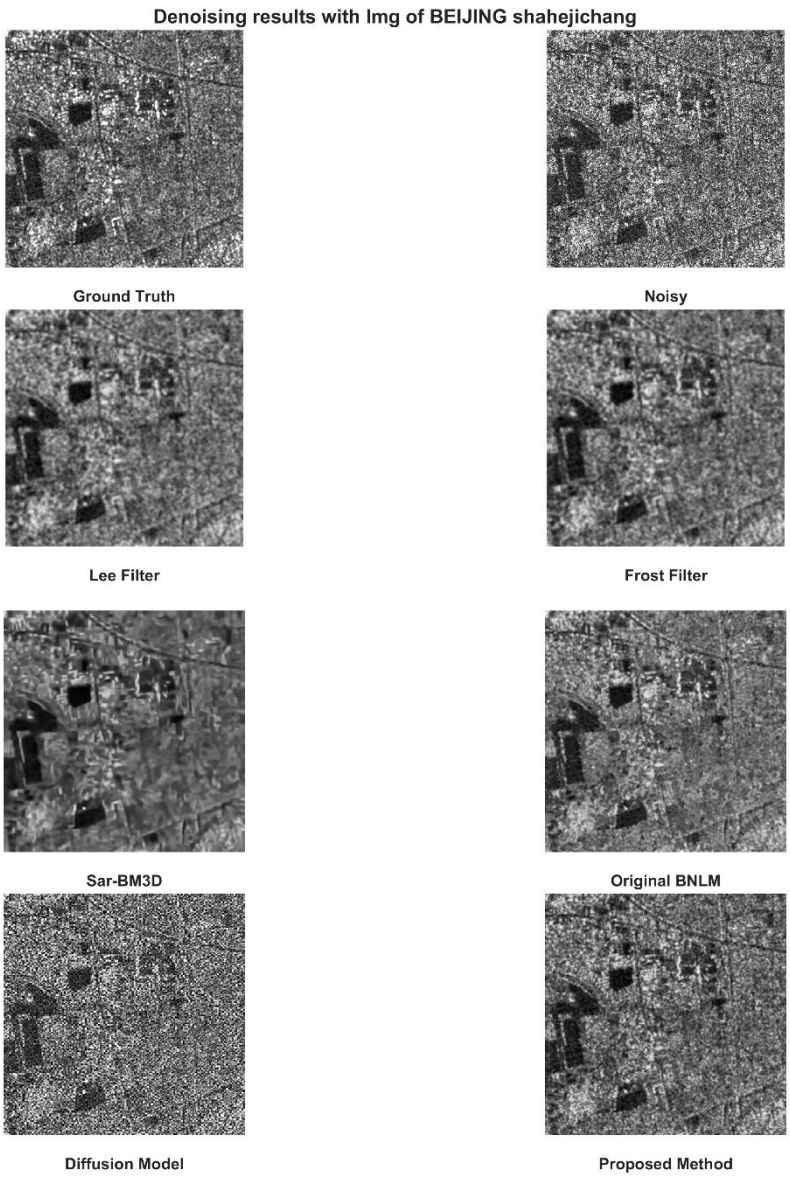


Figure 4.3 Denoising results of Image “BEIJING_shahejichang_1_4”

doi:10.6342/NTU202502541

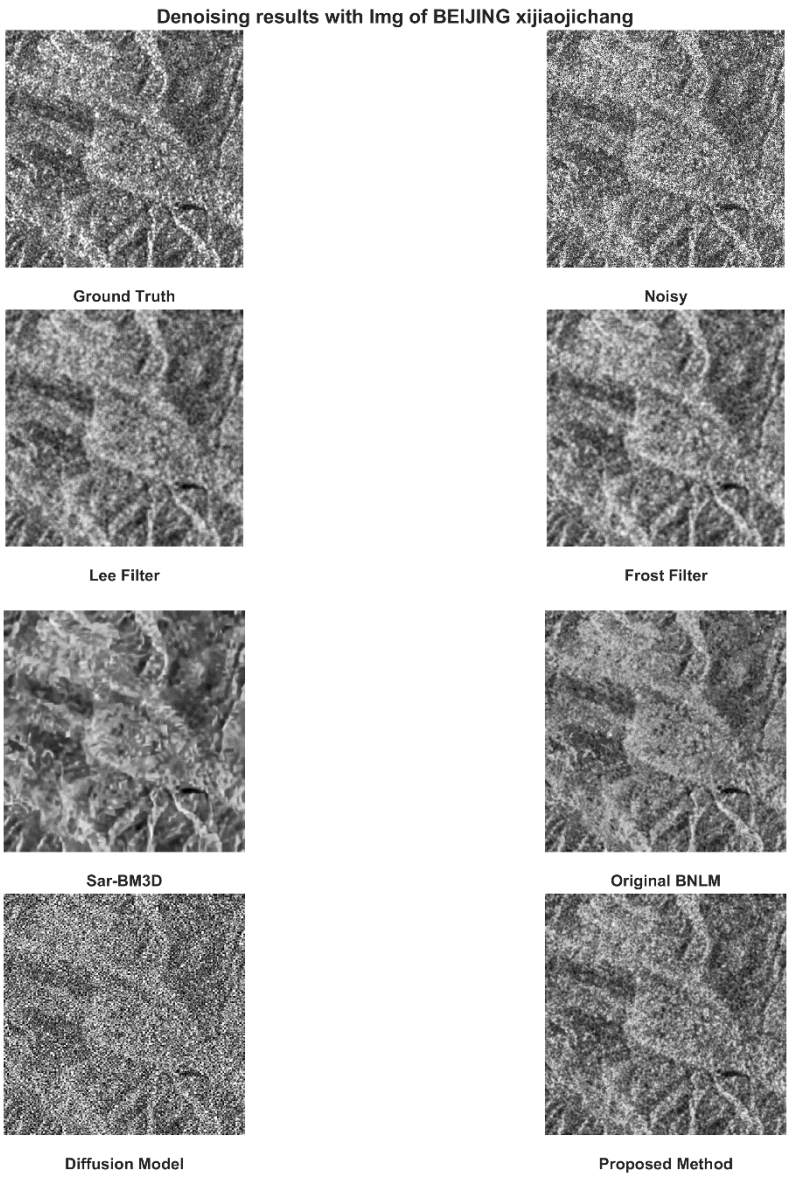


Figure 4.4 Denoising results of Image “BEIJING_shahejichang_3_5”

doi:10.6342/NTU202502541

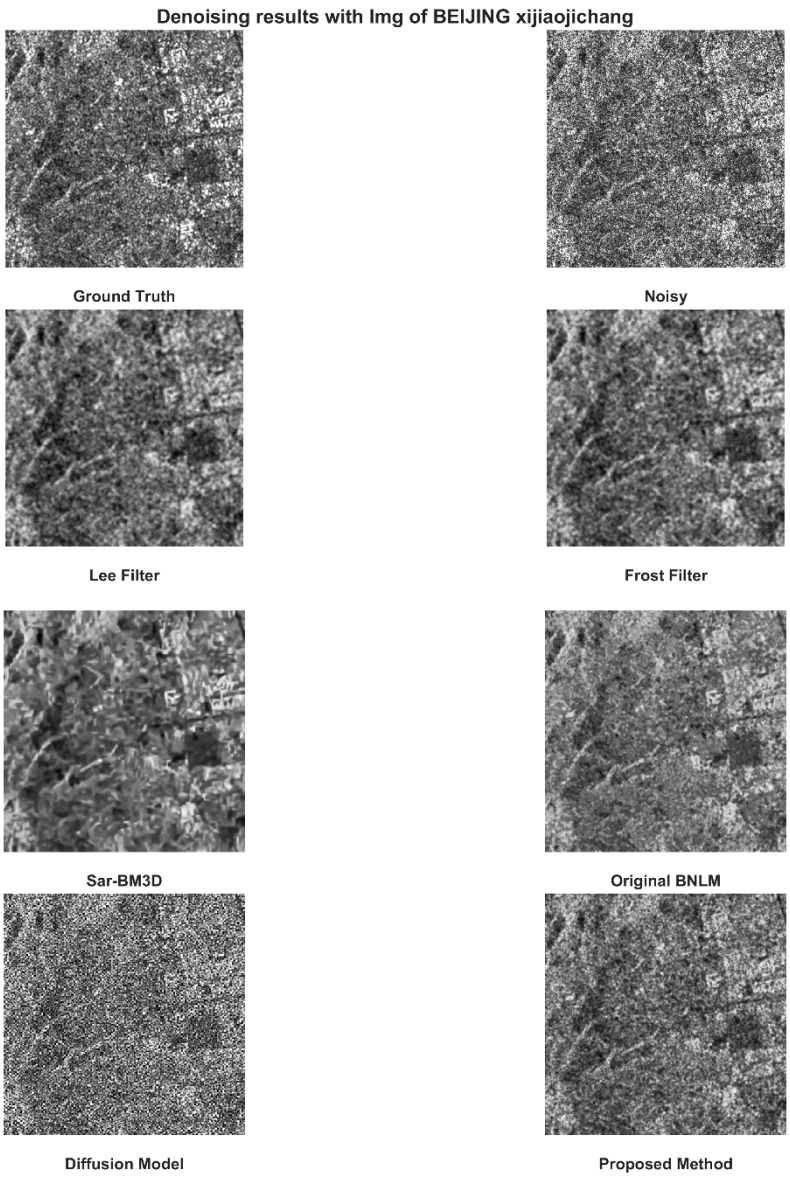


Figure 4.5Denoising results of Image “BEIJING_ xijiaojichang_3_6”

doi:10.6342/NTU202502541

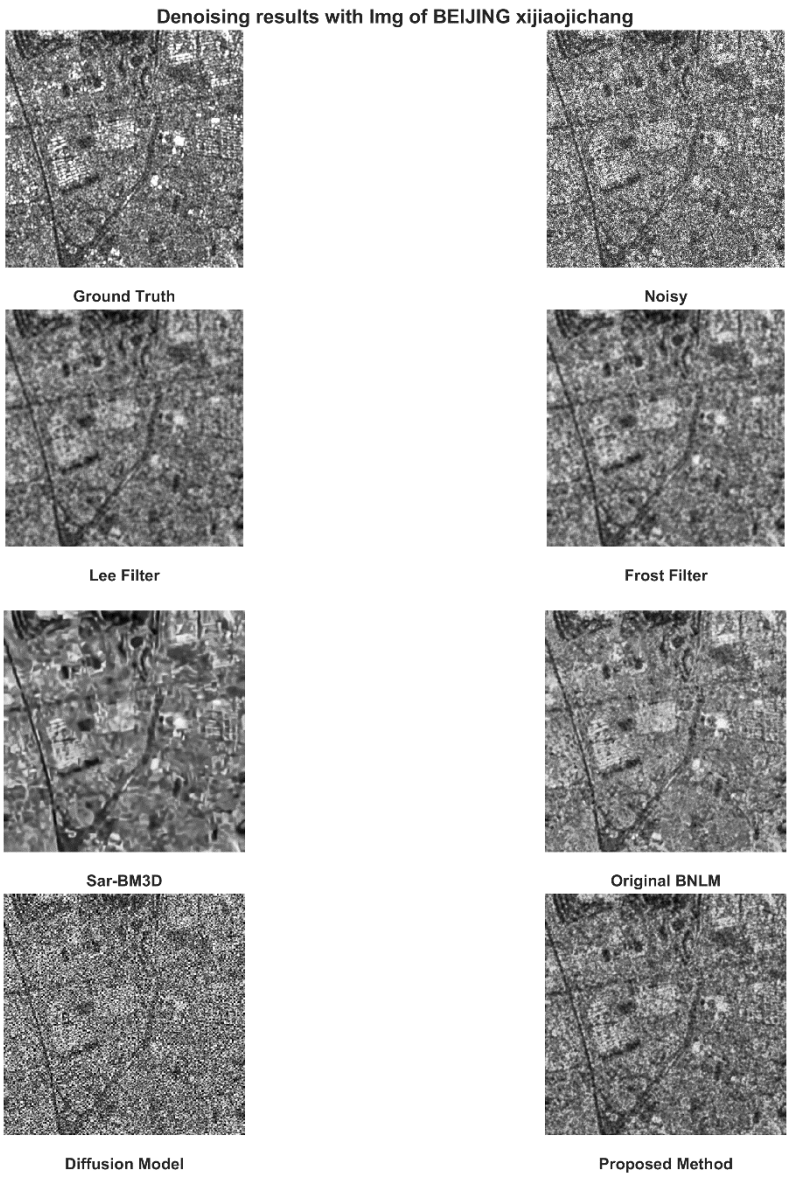


Figure 4.6Denoising results of Image “BEIJING_ xijiaojichang_3_7”

doi:10.6342/NTU202502541

4.5.3 Processing Time and Efficiency

Beyond denoising quality, computational efficiency is crucial for large-scale or real-time SAR image analysis. Table 3 compares average processing times per 512×512 image. While classical local filters (Lee: 0.066s, Frost: 0.7248s) remain fastest, they compromise visual quality. SAR-BM3D (1.633s) improves structural fidelity but introduces considerable computational overhead. Original BNLM, despite its strong despeckling performance, exhibits impractically high runtime (123.8218s).

Table 4.3 Average Processing Time

Method	Runtime (sec)
Lee Filter	0.066
Frost Filter	0.7248
SAR-BM3D	1.633
Original BNLM	123.8218
Diffusion Model	59.03
Proposed	3.2434

Lee Filter: 0.066 s, fastest local method but limited quality.

Frost Filter: 0.7248 s, moderate speed with improved smoothing.

SAR-BM3D: 1.633 s, advanced patch grouping at moderate cost.

Original BNLM: 130.04 s, exhaustive spatial-domain matching.

Diffusion Model: 59.03s, second longest time with lowest performance.

Proposed Method: 3.2434 s, $\sim 40\times$ speedup over original BNLM.

doi:10.6342/NTU202502541

This acceleration arises from:

Log-domain distance computation, which replaces repeated divisions and logarithms on raw intensities with more efficient exponentials on precomputed log-patches.

Sketch-based preselection, which prunes candidates outside the structural σ-range before exhaustive patch matching.

Anisotropic kernel weighting, which focuses computations on fewer, more relevant patches aligned with local geometry. The proposed method thus achieves near-real-time performance without compromising denoising accuracy, making it well-suited for onboard satellite processing and large-batch SAR enhancement.

4.5.4 Summary of Experimental Findings

PSNR & SSIM: Proposed method achieves 23.17 dB PSNR and 0.7797 SSIM in average, surpassing all baselines in both mean and consistency.

Visual Quality: Demonstrates superior edge preservation and speckle suppression across varied SAR scenes.

Efficiency: Balances high-quality despeckling with practical runtimes (3.24 s/image) suitable for large-scale applications.

These results validate the sketch-based BNLM framework as an effective and efficient SAR despeckling solution.

doi:10.6342/NTU202502541

Chapter 5 Conclusion

Synthetic Aperture Radar imagery plays a vital role in remote sensing due to its ability to capture high-resolution data regardless of weather or lighting conditions. However, the inherent presence of speckle noise, originating from the coherent nature of radar signal reflection, severely degrades image quality, complicating tasks such as segmentation, classification, and object detection. Unlike additive Gaussian noise, speckle noise is multiplicative and signal-dependent, rendering conventional denoising techniques suboptimal for SAR applications.

To tackle this challenge, both speckle noise estimation and suppression must be addressed in a targeted, structure-aware manner. Accurate noise estimation serves as a cornerstone for adaptive filtering, yet existing approaches often rely on strong assumptions, which fail in heterogeneous or highly textured regions. Classical filters such as Lee and Frost provide efficient noise suppression but tend to oversmooth fine details and assume stationary statistics, limiting their performance in real-world SAR scenarios.

Recent advances in transform-domain processing and nonlocal methods have demonstrated improved performance by exploiting patch redundancy and multiscale decomposition. BM3D-based extensions and Bayesian nonlocal means (BNLM) filters exemplify this trend, but they either suffer from high computational complexity or lack structural adaptivity. Moreover, the reliance on noisy patches as priors introduces bias, particularly when structure and texture vary rapidly across the image.

In this thesis, we propose an integrated framework for efficient speckle noise estimation

doi:10.6342/NTU202502541

and suppression, tailored for SAR imagery. The first stage introduces a novel log-DWT-based local variance estimation strategy, refined via polynomial regression to robustly capture noise statistics across diverse scenes. The second stage enhances the BNLM filter through two key contributions: a Wiener-based prior estimator and a sketch-based structural prior, both designed to improve denoising accuracy and edge preservation.

Extensive experiments on the publicly available RadarS SAR airport dataset confirm the superiority of our approach. The proposed method not only achieves significantly lower estimation error compared to state-of-the-art estimators, but also demonstrates notable improvements in PSNR and SSIM for despeckling, while maintaining practical runtime performance. This work thus contributes a scalable and structure-aware solution for SAR image analysis, with promising applications in real-time remote sensing and downstream computer vision tasks.

doi:10.6342/NTU202502541

REFERENCE

[1] C. Kervrann, J. Boulanger, and P. Coupe, “Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal,” in Proc. Int. Conf. Scale Space Methods Variational Methods Comput. Vis., 2007, pp. 520–532.

[2] M. Dai, C. Peng, A. K. Chan, and D. Loguinov, “Bayesian wavelet shrinkage with edge detection for SAR image despeckling,” *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1642–1648, Aug. 2004.

[3] 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., vol. 47, no. 1, pp. 202–213, Jan. 2009.

[4] H. Zhong, Y. W. Li, and L. C. Jiao, “Bayesian nonlocal means filter for SAR image despeckling,” in *Proc. Asia-Pacific Conf. Synthetic Aperture Radar*, Xian, China, Oct. 2009, pp. 1096–1099

[5] J. A. Buades, B. Coll, and J. M. Morel, “A review of image denoising algorithms, with a new one,” SIAM Interdisc. J.: Multiscale Model. Simul., vol. 4, no. 2, pp. 490–530, 2005.

[6] Fan Zhang, Fei Ma, Yongsheng Zhou. A Benchmark Sentinel-1 SAR Dataset for Airport Detection (SAR-Airport-1.0)[OL]. Journal of Radars, 2024. https://radars.ac.cn/web/data/getData?dataType=SAR-Airport.

[7] M. Miao, Z. Xue and P. Zhao, "A Blind Estimation for Speckle Noise Based on Gaussian-Hermite Moments," 2016 International Symposium on Computer, Consumer and Control (IS3C), Xi'an, China, 2016, pp. 829-832, doi: 10.1109/IS3C.2016.211.

[8] S. Intajag and S. Chitwong, "Speckle Noise Estimation with Generalized Gamma

doi:10.6342/NTU202502541

Distribution," 2006 SICE-ICASE International Joint Conference, Busan, Korea (South), 2006, pp. 1164-1167, doi: 10.1109/SICE.2006.315296.

[9] D. Cozzolino, S. Parrilli, G. Scarpa, G. Poggi and L. Verdoliva, "Fast Adaptive Nonlocal SAR Despeckling," in IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 2, pp. 524-528, Feb. 2014, doi: 10.1109/LGRS.2013.2271650.

[10] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," in IEEE Transactions on Image Processing, vol. 16, no. 8, pp. 2080-2095, Aug. 2007, doi: 10.1109/TIP.2007.901238.

[11] J. -S. Lee, "Digital Image Enhancement and Noise Filtering by Use of Local Statistics," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-2, no. 2, pp. 165-168, March 1980, doi: 10.1109/TPAMI.1980.4766994.

[12] V. S. Frost, J. A. Stiles, K. S. Shanmugan and J. C. Holtzman, "A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-4, no. 2, pp. 157-166, March 1982

[13] Guozhong Chen, Xingzhao Liu and Zhixin Zhou, "Modified frost speckle filter based on anisotropic diffusion," *2007 IET International Conference on Radar Systems*, Edinburgh, UK, 2007, pp. 1-4, doi: 10.1049/cp:20070566.

[14] J. Wu, F. Liu, L. Jiao, X. Zhang, H. Hao and S. Wang, "Local Maximal Homogeneous Region Search for SAR Speckle Reduction With Sketch-Based Geometrical Kernel Function," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5751-5764, Sept. 2014, doi: 10.1109/TGRS.2013.2292081.

[15] S. M. Kabir and M. I. H. Bhuiyan, "Speckle noise modeling using the Bessel K-

doi:10.6342/NTU202502541

Form PDF," 2012 7th International Conference on Electrical and Computer Engineering, Dhaka, Bangladesh, 2012, pp. 268-271, doi: 10.1109/ICECE.2012.6471537.

[16] S. Intajag and S. Chitwong, "Speckle Noise Estimation with Generalized Gamma Distribution," *2006 SICE-ICASE International Joint Conference*, Busan, Korea (South), 2006, pp. 1164-1167, doi: 10.1109/SICE.2006.315296.

[17] M. Miao, Z. Xue and P. Zhao, "A Blind Estimation for Speckle Noise Based on Gaussian-Hermite Moments," 2016 International Symposium on Computer, Consumer and Control (IS3C), Xi'an, China, 2016, pp. 829-832, doi: 10.1109/IS3C.2016.211.

[18] M. V. Perera, N. G. Nair, W. G. C. Bandara and V. M. Patel, "SAR Despeckling Using a Denoising Diffusion Probabilistic Model," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 4005305, doi: 10.1109/LGRS.2023.3270799.

[19] H. Zhong, Y. Li and L. Jiao, "SAR Image Despeckling Using Bayesian Nonlocal Means Filter With Sigma Preselection," in IEEE Geoscience and Remote Sensing Letters, vol. 8, no. 4, pp. 809-813, July 2011, doi: 10.1109/LGRS.2011.2112331.

doi:10.6342/NTU202502541