IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 15 January 2024, accepted 26 January 2024, date of publication 9 February 2024, date of current version 21 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3364397

RESEARCH ARTICLE

Multi-Scale Based Approach for Denoising Real-World Noisy Image Using Curvelet Thresholding: Scope and Beyond

SUSANT KUMAR PANIGRAHI[1], SANTOSH KUMAR TRIPATHY[2], ANIRBAN BHOWMICK[1], (Senior Member, IEEE), SANTOSH KUMAR SATAPATHY[3], (Member, IEEE), PAOLO BARSOCCHI[4], AND AKASH KUMAR BHOI[4,5], (Member, IEEE)

[1]School of Electrical and Electronics Engineering, VIT Bhopal University, Kothri Kalan, Sehore, Madhya Pradesh 466114, India
[2]School of Computing Science and Engineering, VIT Bhopal University, Kothri Kalan, Sehore, Madhya Pradesh 466114, India
[3]Information and Communication Technology, Pandit Deendayal Energy University, Gandhinagar, Gujarat 382426, India
[4]Institute of Information Science and Technologies, National Research Council, 56124 Pisa, Italy
[5]Directorate of Research, Sikkim Manipal University, Gangtok, Sikkim 737102, India

Corresponding authors: Santosh Kumar Tripathy (santoshtripathy1448@gmail.com) and Paolo Barsocchi (paolo.barsocchi@isti.cnr.it)

This work was supported by the CHIPS project ''Cyber and Human Intelligence for Physical Systems'', under Grant CUP: B53C24000490005.

ABSTRACT Naïve simulated additive white Gaussian noise (AWGN) may not fully characterize the complexity of real world noisy images. Owing to optimal sparsity in image representation, we propose a curvelet based model for denoising real-world RGB images. Initially, the image is decomposed in three curvelet scales, namely: the approximation scale (that retains low-frequency information), the coarser scale and the finest scale (that preserves high-frequency components). Coefficients in the approximation and finest scale are estimated using NLM filter, while a scale dependent threshold is adopted for signal estimation in the coarser scale. The reconstructed image in spatial domain is further processed using Guided Image Filter (GIF) to suppress the ringing artifacts due to curvelet thresholding. The proposed approach known as CTuNLM method is extended for color image denoising using uncorrelated YUV color space. Extensive experiments on multi-channel real noisy images are conducted in comparison with eight sate-of-the-art methods. With four encouraging qualitative and quantitative measures including PSNR and SSIM, we found that CTuNLM method achieves better denoising performance in terms of noise reduction and detail preservation. We further examined the potential of proposed approach by focusing only on the Finest scale curvelet Coefficients (FC). Features like small details, edges and textures always add up to improve the overall denoising performance, while minimizing spurious details. We studied ''The Curious Case of the Finest Scale'' and constructed ''Deep Curvelet-Net'': an encoder-decoder-based CNN architecture, as a pilot work. The encoder uses multiscale spatial characteristics from noisy FC, while the decoder processes de-noised FC under the supervision of encoder's multiscale spatial attention map. The ''Deep Curvelet-Net'' links encoder multiscale feature modeling with decoder spatial attention supervision to learn the most essential features for denoising. The CNN-based architecture only estimates FC, while all other CTuNLM stages are left unchanged to produce the denoised output. Results presented in this article validated the design of proposed CNN architecture in curvelet domain and motivated us to search beyond classical thresholding and/or filtering approaches.

INDEX TERMS Curvelet thresholding, deep Curvelet-Net, GIF, AWGN, skip-connection.

I. INTRODUCTION

The increasing demand of high (spatial) resolution images, with constant die size of CCD sensors, imaging systems

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan.

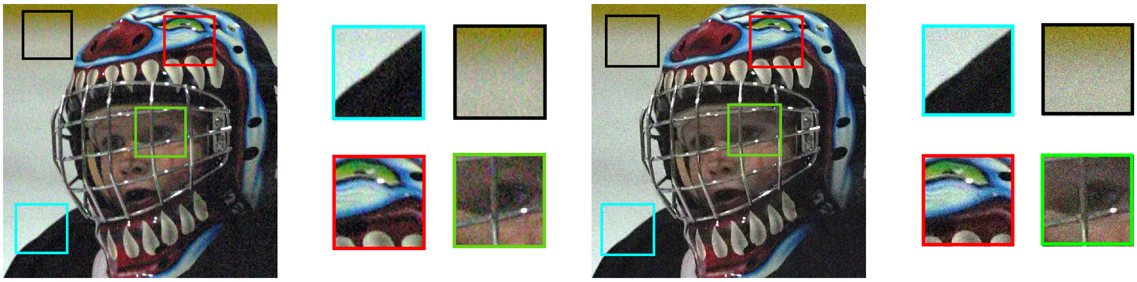


FIGURE 1. Comparison of (a) real noisy image with (c) simulated noisy image corrupted with Additive White Gaussian Noise (AWGN).

invariably add unwanted noise components while acquiring images. Higher pixel counts under limited sensor size damages the signal integrity at each pixel by receiving less photons (lights) and resulting less charges and lower signal to noise ratio [1], [2]. As modifying imaging systems is almost impractical, thus developing densoing algorithms is a key indispensable step in many image processing and computer vision tasks [3]. Furthermore, the general problem of image restoration can be solved through variable splitting by using sub-tasks of denoising algorithms [4], [5], [6], [7], [8], [9], [10].

Noise in the real-world images are generated from various sources of imaging pipeline including: short noise, amplifier noise and quantization noise. Moreover, real world images are scanned, quantized and also under gone through various lossy compression techniques. Therefore, the observed noise is signal dependent, correlated and can't be generalized as white and i.i.ds of Gaussian distributions [11]. As illustrated in Fig. 1, the noise statistics varies independently with different image patches for example the noise pattern in back-box is totally different from that of cyan box, although both patches are taken from same white object. In contrast, the noise pattern observed due to simulated additive noise, as shown in Fig. 1(d) is almost homogeneous for similar objects. Many literature attempts to recover a clean image from its noisy observation by using prior knowledge of both signal and degradation process. Neglecting the complexity of real-world noise, the degradation process can be formulated as the sum of a clean image, x and the additive white Gaussian noise (AWGN), n as: $y = x + n$, where $n \in \aleph(0, \sigma^2)$.

Image denoising may be formulated as an l_2-norm minimization problem assuming both the clean/ reference and noisy image laying in a higher dimensional space (see Fig.2). The task is to find an approximation of clean image such that the l_2-norm between the reference image and the noisy image is minimum. As there are infinite such points exist in the circumference of the circle, the most appropriate solution – without any new artifacts added due to the denoising algorithm assumptions – would be the point joining the straight line between the noisy and the clean image. Although, almost all the denoising algorithm invariably add some new artifacts in the recovered image, one must seek to develop an approach that produces denosing image with minimum visually annoying artifacts.

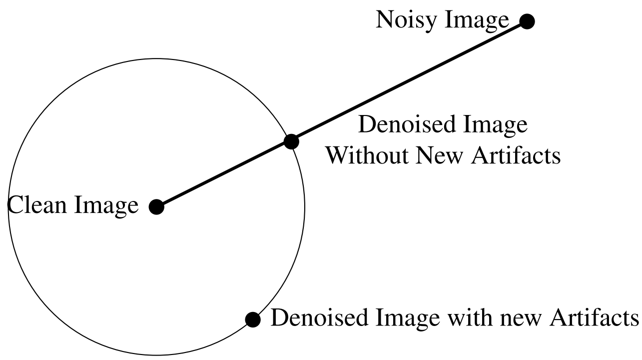


FIGURE 2. The two denoised images have the same ℓ_2-distance to the clean image, but only the denoised image lying on the path between the noisy image and the clean image contains no new artifacts.

With variety of techniques describing the essential qualitative and quantitative features of the image, the existing methods can be broadly categories into the model based [4], [5], [12], [13], [14], [15] and the discriminative learning based methods [8], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]. Model based methods can be categorized into two main sub-groups, spatial-domain-based dictionary learning methods [26], [27], [28], [29], [30] and transform-domain based multi-scale thresholding approaches [31], [32], [33], [34], [35], [36]. On the other hand, discriminative learning based approaches adopt the training set of degraded and ground truth image either in stage-wise (learning of image prior) [17] or convolutional neural network (CNN) [16], [18], [37] based methods to approximate the denoised image.

The model based methods rely on constructing an optimization scheme by adopting the distribution of noise or by exploiting the image priors as constraints or penalties. The non-local self-similarity (NSS) image priors search for similar patches in the whole image assuming natural images have structural similarity that may exist far from local patches [12]. Euclidean distance is commonly used as a measure of self-similarity. The NSS prior has been successfully utilized in many inverse problems, while obtaining state-of-the-art denoising performance for BM3D [31] and WNNM [4] methods. On the other hand, transform domain approaches construct a representative Euclidean space for sparse image representation (with a small number of significant coefficients) and therefore allow highly efficient image models. The transform-based methods can be grouped into data-adaptive and non-data-adaptive transform models. Data-adaptive transform models [27], [29] use image patches as basis functions. However, these types of modeling strategies might behave inconsistently with respect to outlier data or on the data which are not considered for training. Moreover, to create an abstract model with high quality performance, choosing the correct model parameters can be costly due to the amount of data they require. The fixed kernel based non-data adaptive, multiscale approaches are very popular and promising in exploiting line and Curve singularity for image denoising [38], [39]. The sparsity in image representation and the ability of fixed kernel basis functions to decorrelate the image signal and the noise subspace are key factors in the development of thresholding (or shrinkage) based image denoising approaches. Nevertheless, the sudden jump in coefficient magnitude can introduce ringing artifacts in the resulting denoised images. Notably, the recent adaptive soft-thresholding based method improves the image quality by exploiting the non-local correlation among the overlapping image patches; rather than considering a global image modeling [40]. Literature also supports combined approaches of feature selective thresholding and filtering in different frequency bands of transformed coefficients for image denoising [34], [41]. Recently these methods achieved competitive state-of-the-art results in many image restoration applications.

Recently, with availability of numerous image datasets, the discriminative learning based methods almost achieved the competitive performance of denoising under simulated additive Gaussian noise [3]. Without providing image priors manually, deep learning based denoising methods employ CNNs to develop models using a large set of clean and noisy image pairs. Notable CNN based denoising models like, DnCNN [18] and IrCNN [8] adopts deep residual network for image denoising, which effectively captured residual noise patterns. Flexibility to adopt unknown and varied noise level is still a challenging task for many neural network models. Zhang et al. [19] introduced FFDNet, which offered a fast and flexible solution using parallel feed-forward denoising blocks for adaptively control the trade-off between noise reduction and detail preservation. Data over-fitting for Gaussian noise and poor generalization to real-word noisy images (with more sophisticated noises) are two major road-blocks of CNN based approaches for image denoising [20].

Generally, denoising real-noisy image is a two-step process: noise estimation (challenging for spatially varying uncorrelated noises) and feature attentive, non-blind denoising [3]. Noise Clinic [42] proved to be efficient in estimating the noise model depending on signal and frequency followed by non-local Bayes (NLB) model for image denoising. In contrast, the well-known software toolbox Neat Image [43] and few other methods [44], [45] are developed specifically for handling real-world noises. Interestingly, the benchmarking BM3D [31] (or CBM3D [32]) still demonstrates competitive performance compared to several denoising approaches [29]. We in this article try to amalgamate the concepts of signal sparsity and non-local self-similarity (NSS) to develop an algorithm for denoising real-world (RGB) noisy images. Inspired from [46], [47], [48], [49] and the seminal work presented in [50], we define the constrained minimization problem as:

$$\hat{x} = \underset{x}{\operatorname{argmin}} \frac{1}{2}\|y-x\|_2^2 + \lambda \mathbf{R}(x) \tag{1}$$

Note, the first term represents the data fidelity, whereas the second term depends on the image priors used. The regularization parameter λ effectively balances the trade-off between these two components. In this proposed work, we estimate $\hat{x}$ from the frequency domain coefficients, assuming the image is sparsely represented using curvelet basis functions. Thus a well-defined scale dependent threshold can be used to separate the signal components from its noisy observations, whereas an explicit prior is chosen using nonlocal constraint to estimate the denoising coefficients in the curvelet approximation and the finest scale (Section III). However, in the Deep-CNN based approach the adaptive moment estimation (ADAM) [51] is used to estimate the finest scale coefficients – instead of NLM filter – in the supervised deep-learning-based method (Section V-B).

A. SIGNIFICANT CONTRIBUTION

Owing to energy compact and linearity properties (due to tight frames), the multiresolution curvelet transform can represent any square integrable function and also obeys Parsevals' theorem. As a result the noise may remain additive in the transformed domain and NLM filter can be applied on the curvelet coefficients [2]. With a single parameter based image denoising framework, we highlight the main contribution of our work:

1) A fast and efficient model adopted to spatially varying noise; tunable at any known noise level is proposed for denoising real-world noisy images while demonstrating its potential for practical applications. Here, the hybrid approach is implemented using fast and improved NLM filter [12], [52] to speed-up the overall process.
2) We also examined the curious case of the Finest scale curvelet Coefficients (FC) and highlighted the scope of improvement for any method adopting curvelet based multiscale approach for image restoration.
3) Finally, an encoder-decoder based CNN architecture with spatial attention blocks (SAB) known as "Deep Curvelet-Net" is developed for denoising the curvelet finest scale coefficients.

II. PROPOSED DENOISING FRAMEWORK

In this section, we look closely at the sparse land model for image representation using curvelet transform and formulate the problem of image denoising with detailed block diagram in the subsequent sections.

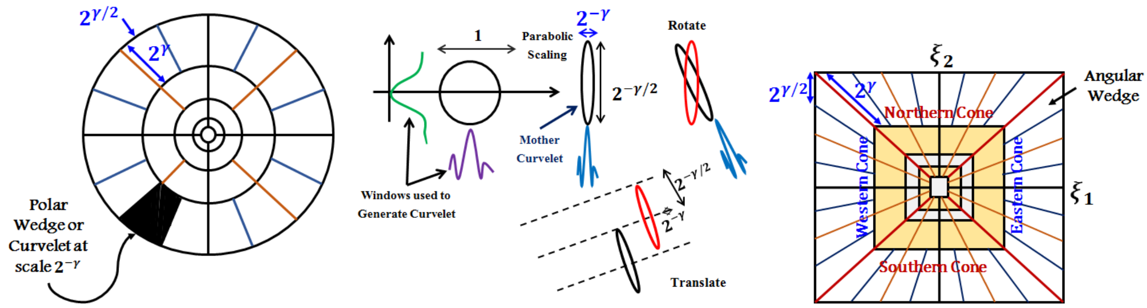


FIGURE 3. **(a) Construction of continuous curvelets. In Fourier space, curvelets are supported near a "parabolic" wedge. (b) A basic curvelet and the possible translations and orientations. (c) Discrete curvelet tilting with parabolic pseudopolar support in the frequency plane.**

A. SPARSE IMAGE REPRESENTATION IN THE CURVELET DOMAIN

The 2^{nd} generation, non-data-adaptive multiscale curvelet transform found its popularity in many application areas including image processing, seismic data exploration, fluid mechanics, and solving partial differential equations. By adopting an-isotropic scaling ($width = length^2$), curvelet efficiently represent curve singularity with minimum number of complex coefficients. Alike any other transformation, an image function $x \in \mathbf{L}^2(\mathbb{R}^2)$ can be represented as a linear combinations of curvelet basis or frame atoms $\phi_{\gamma,o,\tau} \in \mathbf{L}^2(\mathbb{R}^2)$, as:

$$x = \sum_{\gamma,\tau,o} C_{\gamma,\tau,o}(x)\phi_{\gamma,\tau,o} \tag{2}$$

where, $C_{\gamma,\tau,o} =< x, \phi_{\gamma,\tau,o} >$ are the curvelet coefficients and $< \cdot, \cdot >$ represents scalar product in $\mathbf{L}^2(\mathbb{R}^2)$. Here, ϕ, the basic curvelet is located at different scales, γ, translation, τ and rotations, o. In general formulation the curvelet is formed as a combination of two window functions $W(\cdot)$ (radial window) and $V(\cdot)$ (angular window); defined in frequency domain. Assuming $\xi = (\xi_1, \xi_2)^T$ representing frequency variable. Further, let $r = \sqrt{\xi_1^2, \xi_2^2}$, $\omega = arctan(\xi_1/\xi_2)$ be the coordinates in frequency domain. We define the "dilated basic curvelet" in polar coordinates as:

$$\tilde{\phi}_{\gamma,0,0}(r,\omega) = 2^{\frac{-3\gamma}{4}} W(2^{-\gamma} r)\tilde{V}_{N_\gamma}(\omega);\ r \geq 0, \omega \in [0, 2\pi) \tag{3}$$

As shown in Fig.3, the curvelet elements are locally supported near wedges; where the number of wedges defined at any scale $2^{-\gamma}$ as $N_\gamma = 4 \cdot 2^{\lceil \frac{\gamma}{2} \rceil}$ (i.e. wedges doubles at each ring). Now, we define the complete curvelet family in spatial domain with position index $\mathbf{p}$ as:

$$\phi_{\gamma,\tau,o}(\mathbf{p}) = \phi_{\gamma,0,0}\left(R_{\theta_{\gamma,o}}(\mathbf{p} - b_\tau^{\gamma,o})\right) \tag{4}$$

with parameters $\gamma \in \mathbb{N}_0$ and $\tau = (\tau_1, \tau_2)$. Note the rotation matrix at an angle θ, in Eq.4 is denoted as R_θ. Let us define the equidistance rotation angles $\theta_{\gamma,o}$ as:

$$\theta_{\gamma,o} = \frac{\pi\ o\ 2^{-\lceil \frac{\gamma}{2} \rceil}}{2}$$

and the positions as:

$$b_\tau^{\gamma,o} = b_{\tau_1,\tau_2}^{\gamma,o} = R_{\theta_{\gamma,o}}^{-1}\left((\tau_1/2^\gamma)(\tau_2/2^{\gamma/2})\right)^T$$

Finally, we redefine the curvelet transform of any $2D$ signal in spatial domain as [38]:

$$\begin{aligned} &C_{\gamma,\tau,o}(\mathbf{p}) \\ &= \int_{\mathbb{R}^2} x(\mathbf{p})\overline{\phi_{\gamma,\tau,o}(\mathbf{p})}d\mathbf{p} \\ &= \int_{\mathbb{R}^2} \hat{x}(\xi)\hat{\phi}_{\gamma,0,0}\left(R_{\theta_{\gamma,o}}\xi\right) e^{i<b_\tau^{\gamma,o},\xi>} d\xi \end{aligned} \tag{5}$$

where $\hat{x}(\xi)$ and $\hat{\phi}_{\gamma,0,0}(\xi)$ are the Fourier transform of $x(\mathbf{p})$ and $\phi(\mathbf{p})$.

For images being represented in Cartesian arrays; the curvelets are approximated (interpolated) to concentric squares instead of circular rings (Eq.3). Thus the rotation in replaced by shearing as shown in Fig.3(c).

$$\hat{\tilde{\phi}}_{\gamma,0,0}(\xi) = 2^{\frac{-3\gamma}{4}} W(2^{-\gamma}\xi_1) V\left(\frac{2^{\lfloor \frac{\gamma}{2} \rfloor}\xi_2}{\xi_1}\right) \tag{6}$$

Here, the basic curvelet $\hat{\tilde{\phi}}_{\gamma,0,0}$ determines the frequencies in trapezoid as:

$$(\xi_1, \xi_2) : 2^{\gamma-1} \leq \xi_1 \leq 2^{\gamma+1}, -2^{-\lfloor \frac{\gamma}{2} \rfloor}.\frac{2}{3} \leq \frac{\xi_1}{\xi_2} \leq 2^{-\lfloor \frac{\gamma}{2} \rfloor}.\frac{2}{3}$$

Now let us define the digital curvelet families in the Cartesian $2D$ grids at various scales, γ, translations, τ and orientations o as:

$$\tilde{\phi}_{\gamma,\tau,o}(\mathbf{p}) = \tilde{\phi}_{\gamma,0,0}\left(S_{\theta_{\gamma,o}}^T(\mathbf{p} - \tilde{b}_\tau^{\gamma,o})\right) \tag{7}$$

where, the shear matrix is:

$$S_\theta = \begin{pmatrix} 1 & 0 \\ -tan\theta & 1 \end{pmatrix}$$

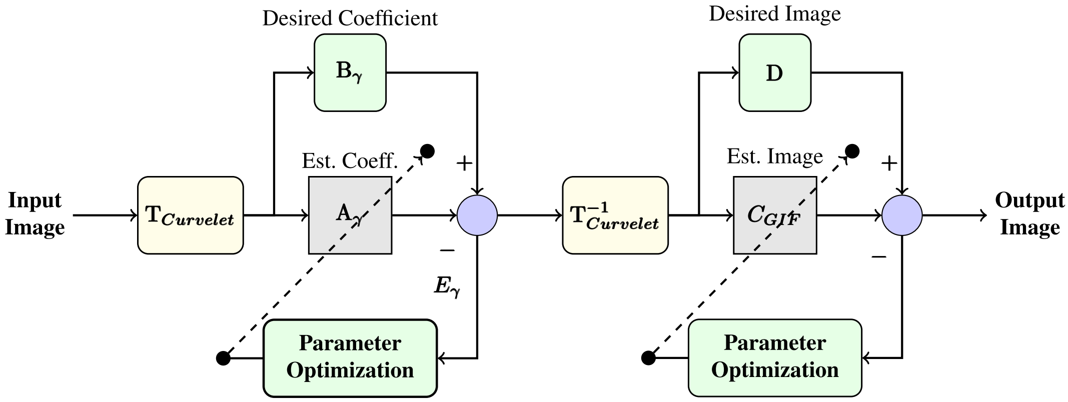


FIGURE 4. Proposed problem definition.(Est. = Estimated.)

and $\tilde{b}_\tau^{\gamma,o} = S_{\theta_{\gamma,o}}^{-T}\tau_\gamma$. Applying Fourier transform to Eq.7, we obtained the digital curvelet basis function as:

$$\begin{aligned}&\hat{\hat{\phi}}_{\gamma,\tau,o}\\&= e^{-i<\tilde{b}_\tau^{\gamma,o},\xi>}\hat{\hat{\phi}}_{\gamma,0,0}\left(S_{\theta_{\gamma,o}}^{-1}\xi\right)\\&= e^{-i<\tilde{b}_\tau^{\gamma,o},\xi>}2^{\frac{-3\gamma}{4}}W(2^{-\gamma}\xi_1)V\left(\frac{2^{\lfloor\frac{\gamma}{2}\rfloor}\xi_2}{\xi_1}+o\right)\end{aligned} \tag{8}$$

Thus, $\hat{\hat{\phi}}_{\gamma,\tau,o}$ is a compactly supported curvelet in the Fourier domain on sheared trapezoids.

The digital curvelet in Cartesian grid (Eq.8) is defined in the Fourier domain and simply for an image in $2D$ the curvelet transform can be calculated as:

$$C(Image) = IFFT\,[FFT(Curvelet) \times FFT(Image)] \tag{9}$$

With two digital implementations proposed in [38] Fast Discrete Curvelet Transform (FDCT) via. Unequispaced FFT (USFFT) corresponds to the most faithful and exact implementation strategy with computational complexity approximately close to FFT implementation: $\mathcal{O}(M^2log(M))$. In this article we have considered FDCT algorithm for sparse image representation and noise reduction via. multiscale filtering and hard thresholidng.

B. PROBLEM FORMULATION

Fig. 4 depicts the problem formulation of the proposed approach in the curvelet domain. We define a operator A_γ and reformulate the problem with a fast limited (three) scale curvelet decomposition method to analyze both the coarser scale (low-frequency) and the finest scale (high-frequency) noise, distinctively. Assuming a scale dependent threshold λ_γ, we defined the operator, A_γ in the curvelet domain $\mathbf{T}_{Curvelet}$ as:

$$\text{A}_\gamma = \left[\text{A}_{NLM,\gamma_l} \mid \bar{\mathbf{I}}_{C_\gamma \geq \lambda_\gamma} \mid \text{A}_{NLM,\gamma_h}\right] \tag{10}$$

The image is initially transformed to curvelet domain but only decomposed in three curvelet scales. Here, A_{NLM,γ_l} and A_{NLM,γ_h} are two operators depicting the non-local means filter applied independently in the approximation (low-frequency) scale and the Finest (high frequency) scale, respectively. On the other hand in the coarser scale, we applied adaptive threshold, λ_γ and the identity matrix operator $\bar{\mathbf{I}}$ retained the coefficients, $\mathbf{C}_\gamma \geq \lambda_\gamma$. As shown in Fig.4, the parameters of operator A_γ are tuned by minimizing the MSE, $\|E_\gamma\|^2$ between estimated and desired coefficients denoted as B_γ.

In the last and final step, we processed the reconstructed image (known as initial denoised image obtained by inverse curvelet transform $T_{Curvelet}^{-1}$ using Guided image Filter (GIF). C_{GIF} in the spatial domain is adopted to mitigate the distortions introduced due to thresholding. Moreover, the edge-aware filter also aides in preserving local structures like: edges, textures and small details. Literature suggests that with proper tuning parameters of C_{GIF}, improves the denoising performance up to $1.0dB$ [41]. Moreover, compared to its earlier versions proposed in [35], for computational tractability and ease of usage a single input parameter, σ_{est} (estimated noise standard deviation) is used to tune all other parameters in Eq. 10.

III. METHODOLOGY

The proposed image denoising framework, as shown in Fig. 5 is designed specifically to handle the real-world noisy image with a single tuning parameter. By adopting several changes compared to the initial algorithm as in [35], we developed a fast, efficient and flexible, denoising method using a standalone model that can handle both spatially variant and invariant noise when the noise standard-deviation is known or unknown for real-time applications. Moreover, we purely focused on multi-channel (RGB) images contaminated with natural sensor noise especially due to low-light conditions.

The noise variance estimator in [53] is used to calculate noise standard deviation, σ_{est} as the median absolute deviation (MAD) of wavelet coefficients at the HH-scale. The wavelet based noise estimator, $\sigma_{est} = \frac{MAD}{0.6745}$ proved to very efficient and robust [54]. In the proposed multiscale based NLM filtering approach the single unknown parameter,

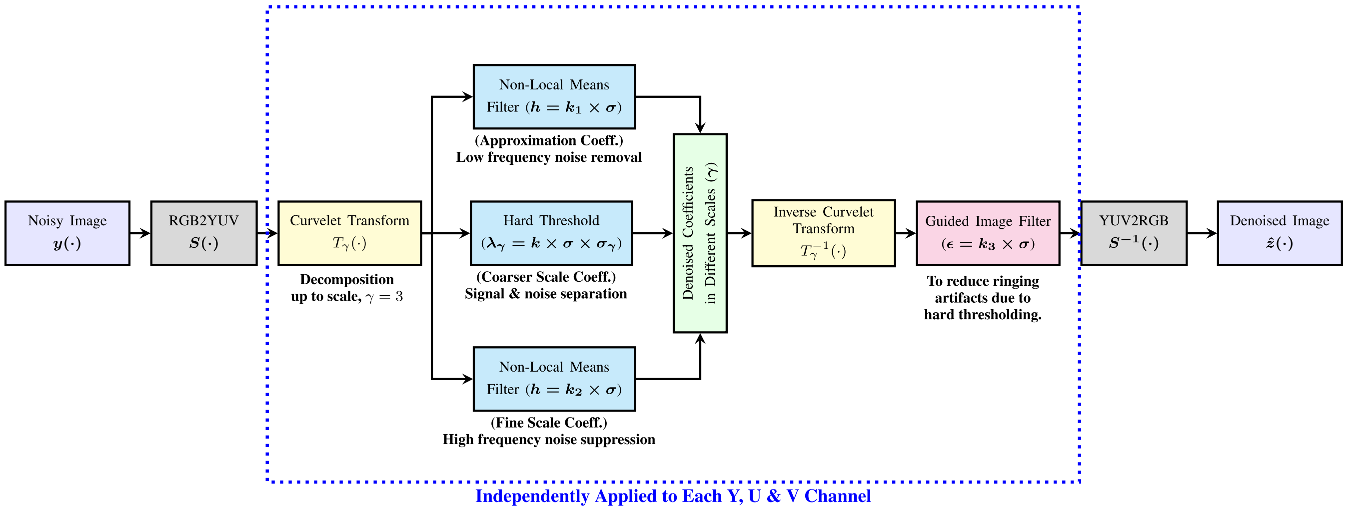


FIGURE 5. **Block diagram of the proposed CTuNLM Framework. Here $S(\cdot)$ and $S^{-1}(\cdot)$ are forward and inverse color transform matrix.**

noise standard deviation is estimated as the first indispensable step. Out of two solutions listed in [35], and inspired from the seminal work of Dabov et al. [32], we considered luminance /color-difference based de-correlated space for image denoising, as shown in Fig. 5. Literature also suggest that a better subjective quality – with an improvement up to 1.2dB – can be obtained for algorithms implemented in YUV color space compared to correlated RGB-space [55].

We particularly exploit the knowledge that multiscale transforms like curvelet, constructed for detecting correlations in images lead to sparse representations in the transform domain with a small number of significant coefficients at various scales/frequency-bands having high intra-scale correlation. Therefore, it allows highly efficient image models [39]. The FDCT via. USFFT [38] is initially applied in the YU & V-channels to perform a limited scale decomposition and to fully exploit the non-local self-similarity (NSS) in the approximation and finest scale. Unlike [12], the weights of NLM filter as defined in Eq. 11 is optimized by taking the advantages of symmetry property (roughly halves the computation time) and adopting a look-up table for speeding-up the computation process. Moreover, the NLM filter (known as FNLM [52]) is partly *c*-coded to improve the overall run-time complexity. For any two given noisy coefficients, $\mathcal{C}_Y$ at positions ξ_i and ξ_j, the similarity between two (non-local) patches (in either approximation/the finest scale) $\mathbf{N}(\xi_i)$ and $\mathbf{N}(\xi_j)$ is reformulated in the curvelet domain as:[1]

$$W(\xi_i, \xi_j) = exp\left(-\frac{\| \mathcal{C}_Y\left(\mathbf{N}(\xi_i)\right) - \mathcal{C}_y\left(\mathbf{N}(\xi_j)\right) \|_{2,a}^2}{h^2}\right) \quad (11)$$

The Gaussian modulated similarity measure using Eucledian distance is denoted as $\| \cdot \|_{2,a}^2$ with a represented as the standard deviation of Gaussian function. The parameters, $h_1 = k_1 \cdot \sigma_{est}$ and $h_2 = k_2 \cdot \sigma_{est}$ control the smoothness of the NLM filter in the approximation and finest scales, respectively. As the smoothing parameter is mathematically related to estimated noise σ_{est}, it is capable of adapting varying noise power.

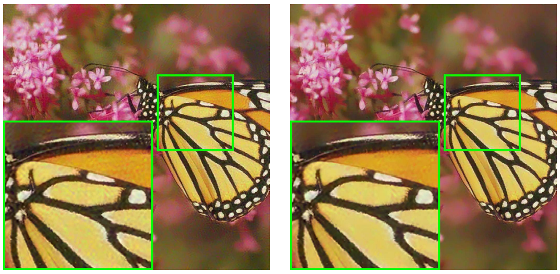
Unlike, the approximation and the finest scale: the courser scale coefficients are decomposed in various angles (or orientations), o. As defined in Eq. 10, we defined the scale dependent threshold as:

$$\lambda_\gamma = k\sigma_{est}\sigma_\gamma \quad (12)$$

The scale dependent variance σ_γ^2 is estimated using Monte-Carlo simulation and the parameter, $k = 1.5$ is obtained using empirical method. The hard-thresholding in the coarser scale although separates noise from the signal components, introduces ringing artifacts. Cycle spinning of curvelet coefficients [56] and post-processing filtering techniques [41] are only few solutions suggested in the literature for suppressing ringing artifacts around the edges. The reconstructed image obtained from inverse curvelet transform (in YUV Space) is further processed using Fast Guided Image Filtering (GIF) [57] to retain small image details like textures and edges. Moreover, the fast GIF is computationally efficient and improves the speed up to $O(M)$ for M pixels. Fig. 6 illustrates the visual and quantitative improvement of denoised image before and after the applications of GIF. The results validate the application of post-processing Guided Image filtering for suppression of ringing artifacts and preservation of image small details like textures and edges.[2]

[1]The NLM filter is applied in the curvelet domain by decomposing each channel of color image in YUV-space, separately. Therefore, we formulate NLM filter for single channel $2D$ coefficient sub-band.

[2]The code implemented and tested in MATLAB@2022 will be available in our official gitHub page: https://github.com/susant146/CTuNLM_Image-Denoising/



(a) Before GIF (b) After GIF

FIGURE 6. Effect of post-processing GIF filter. (a) PSNR = 29.735, SSIM = 0.8927. (b) PSNR = 30.585, SSIM = 0.9306.

IV. RESULTS AND DISCUSSION

A detail analysis and comparison of image denoising performance with various state-of-the-art techniques are presented in this section. As illustrated earlier the proposed method is investigated for denoising real (RGB) noisy image corrupted with natural noises. For uniformity in comparison, the image dimension was kept either fixed to default size as provided in the database or cropped to $M = 512 \times 512$ rows and columns. Three bench-marking real-noisy image datasets: PolyU Real-noisy Image Database [58], Cross Channel (CC) image database [59] and Renoir Image Dataset [60] were considered in this article for testing and validation of proposed denoising algorithm. Being the simplest inverse problem, a plethora of research has been carried out in this area, still it remains an open problem in image processing [33], [61]. With various denoising methods available, authors of this article only focused on seven recent and/or benchmarking techniques including: CBM3D [32], MC-WNNM [13], DDID2 [34], FFDNet [19], GSRC-NLP [30], TWSC [28] and MSI Color-tSVD [29] for comparison. The selected methods are specifically designed for multi-channel real noisy image and according to literature these methods achieved best performance for both spatially invariant and variant noises. Moreover, the authors of this article either used MATLAB/Python codes available publicly under default parameter settings for un-biased and faithful comparison.[3]

A. IMPLEMENTATION DETAILS

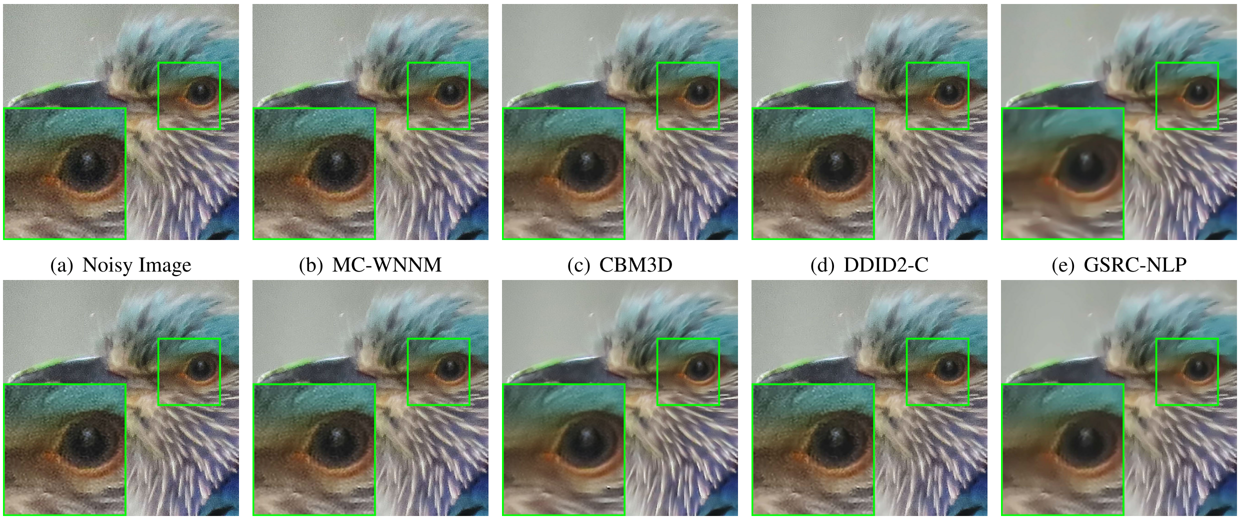
The proposed CTuNLM based denoising algorithm has four tunable parameters including, k_1 and k_2, the weight kernel parameters of NLM filter (see Eq. 11) applied on the curvelet coefficients in the approximation and the finest scale, respectively. Similarly, k, the scale dependent constant in Eq. 12 and k_3 ($\epsilon = k_3 \times \sigma$) the smoothing parameter of GIF [57] are the other two tunable quantities of the proposed algorithm. We used a similar approach as mentioned in [41] to obtain the optimal value for these parameters. Note, in this well-engineered approach, the parameters are tuned once, using TID2013 image database [63]. $k_1 = 0.4$, $k_2 = 0.6$, $k = 1.5$ and $k_3 = 2.1$. As mentioned earlier, our algorithm only takes two inputs, noisy image (in RGB scale) and the estimated noise standard deviation σ_{est} with all other parameters being co-dependent on σ_{est}.

B. QUANTITATIVE AND VISUAL ASSESSMENT

Visual quality provides subjective assessment, while the quantitative measures provide a numerical interpretation, that is more objective than subjective. In absence of reference images, we adopted subjective assessment to study the perceptual quality of the denoised image in terms of the correct preservation of edges and textures and non-presence of artifacts. Fig. 7 and 8 illustrate the denoised output obtained from various methods including the commercially available Neat Image (NI) denoising software [43]. Assuming BM3D and FFDNet are widely considered as bench-marking image denoising algorithms, we choose to compare our resultant images with these methods for visual assessment. CBM3D, MC-WNNM and DDID2 exhibit few visible artifacts manifesting as low frequency and structural noise. As one can see that Neat Image [43], reduces much noises, while preserving most of the image details. GSRC-NLP [30], TWSC [28] and MSI Color-tSVD [29] induced many algorithm-based artifacts. On the other hand, FFDNet [19], a well-engineered approach for handling spatially varying non-Gaussian noise selects from many outputs from set of noise levels excels in optimizing noise reduction and detail preservation. The proposed method without needing any training and entirely based on image/signal representation technique, proved to be very efficient in retaining image features both around the edges and in the flat regions.

Table 1 and 2 presents the quantitative results in terms of PSNR and SSIM measure. We provide three best and three competitive results with average value of the PSNR and SSIM measure in the last row of each table. The results were obtained from two different datasets containing real world noisy image and the corresponding ground truth or clean image. Noise estimation and adopting variant, non-Gaussian noise is an important task – while it is very implausible – but unless otherwise specified, we assumed invariant noise standard deviation for real-noisy image. The noise variance is calculated using well known wavelet based MAD estimator [53] for our proposed method. For all the other methods, images were denoised at default settings without changing the authors suggested algorithms/parameters. One can see that the proposed algorithm provides competitive results compared to CBM3D, MC-WNNM and DDID2 methods for all the images. An increase in PSNR and SSIM measure indicates both noise reduction and structural preservation of denoised image with respect to the available ground truth image. The hybrid approaches of CBM3D and DDID2 combines spatial filtering and wavelet thersholding for image denoising and widely considered as the bench-marking denoising approach. However, these methods are not yet perfect, while introducing visible artifacts in the homogenous region, manifesting as low-frequency noise. The weight based and the group level correlation employed for enhancing sparsity based methods in TWSC [28] and MSI

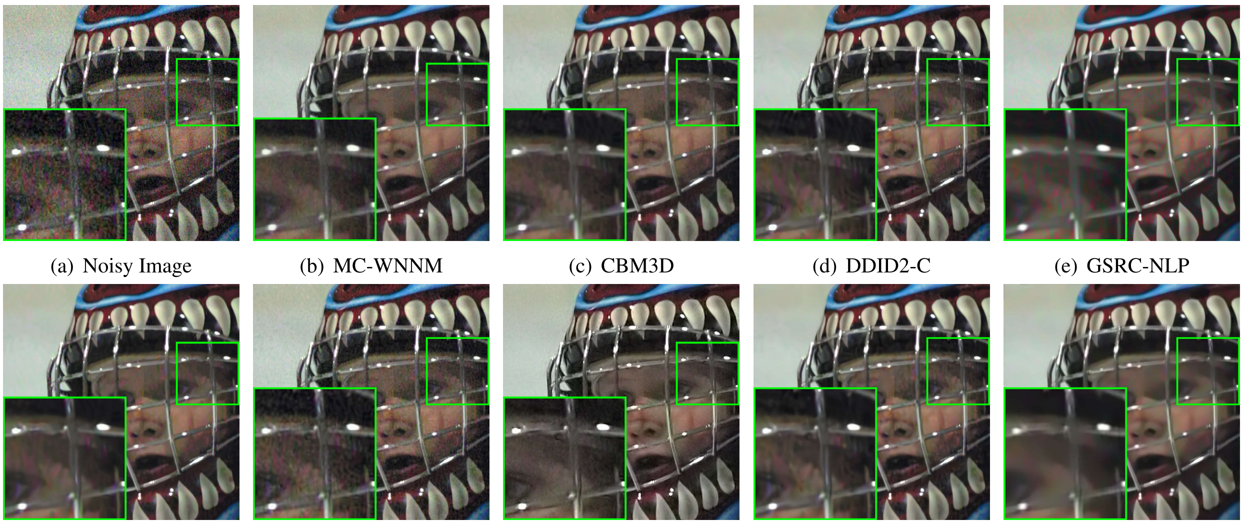
[3] We would like to thank all authors for sharing their codes.



(a) Noisy Image (b) MC-WNNM (c) CBM3D (d) DDID2-C (e) GSRC-NLP

(f) TWSC (g) MSI Color-tSVD (h) Neat Image (i) FFDNet (j) CTuNLM

FIGURE 7. Visual comparison 01.



(a) Noisy Image (b) MC-WNNM (c) CBM3D (d) DDID2-C (e) GSRC-NLP

(f) TWSC (g) MSI Color-tSVD (h) Neat Image (i) FFDNet (j) CTuNLM

FIGURE 8. Visual comparison 02.

Color-tSVD [29] proved efficient in denoising real-noisy image as these models are specifically developed for handling non-Gaussian color noises. The CNN based FFDNet [19] picks the best resultant image from multiple outputs with varying noise map seems to take the betterment at denoising for few real-noisy images, while proposed method excels in removing noise with preserving significant details for other images.

To complement the quantitative results as shown in Table 1 and 2 best denoising output from each database are shown in Fig. 9 and 10 with corresponding PSNR and SSIM values. The results demonstrate CBM3D and DDID2 generate some noise-caused color artifacts across the whole image, while MC-WNNM, TWSC and MSI Color-tSVD tend to over-smooth the resultant image. However, our method provides promising results in comparison with the sate-of-the-art FFDNet algorithm.

C. ARTIFACT STUDY

The problem of image denoising formulated as l_2 norm minimization (Section I), not only requires the point (corresponds to noisy image represented in higher dimension) to move as close as to the reference point (corresponds to ground-truth clean image) but also it is indispensable to lie on the line joining the two points, as shown in Fig. 2. Finding such a method from infinitely possible solutions is practically implausible. Every denoising algorithm, while dealing with real-noisy image (including proposed CTuNLM method) makes some initial assumption about the image model – assuming image is sparse in the curvelet domain – and the noise model. Moreover, almost-all denoising algorithm estimate noise, while assuming the spatially varying noise as invariant quantity, for real noisy scenario [19]. We also highlight the fact that, the denoising method $\mathcal{D}_h$ always estimates the parameter h that are based on some assumptions

TABLE 1. PSNR and SSIM measures on test-set images of CC-Image dataset [59]. First three-rows indicate the best results obtained from the selected images, whereas the next three rows illustrate comparable results and the last row presents the average measure on the images of entire dataset.

MC-WNNM [13]	CBM3D [32]	DDID2_C [34]	GSRC-NLP [30]	TWSC [28]	MSI Color-tSVD [29]	FFD-Net [19]	Proposed CTuNLM
Peak Signal to Noise Ratio (PSNR) in dB							
40.355	43.485	42.904	41.954	41.664	42.939	43.814	44.171
41.242	43.711	43.141	41.867	41.860	43.223	44.070	44.611
38.511	42.464	41.953	41.806	42.789	41.334	42.551	44.556
38.312	40.856	40.169	40.409	40.575	40.134	41.215	40.798
36.384	37.088	37.220	34.262	37.241	37.496	37.034	36.518
37.948	41.640	41.234	39.620	40.987	41.164	41.728	40.279
38.621	41.272	40.891	39.732	41.098	40.709	41.626	41.536
Structural Similarity Index Measure (SSIM) [62]							
0.9587	0.9848	0.9846	0.9838	0.9738	0.9775	0.9850	0.9894
0.9340	0.9754	0.9731	0.9710	0.9636	0.9666	0.9764	0.9901
0.9426	0.9718	0.9656	0.9672	0.9727	0.9647	0.9698	0.9858
0.9593	0.9795	0.9806	0.9799	0.9696	0.9521	0.9783	0.9651
0.9227	0.9435	0.9432	0.9457	0.9153	0.8945	0.9472	0.9120
0.9684	0.9863	0.9871	0.9902	0.9849	0.9596	0.9841	0.9805
0.9501	0.9784	0.9785	0.9607	0.9793	0.9742	0.9824	0.9856

TABLE 2. PSNR and SSIM measures on test-set images of PolyU-image dataset [58]. First three-rows indicate the best results obtained from the selected images, whereas the next three rows illustrate comparable results and the last row presents the average measure on the images of entire dataset.

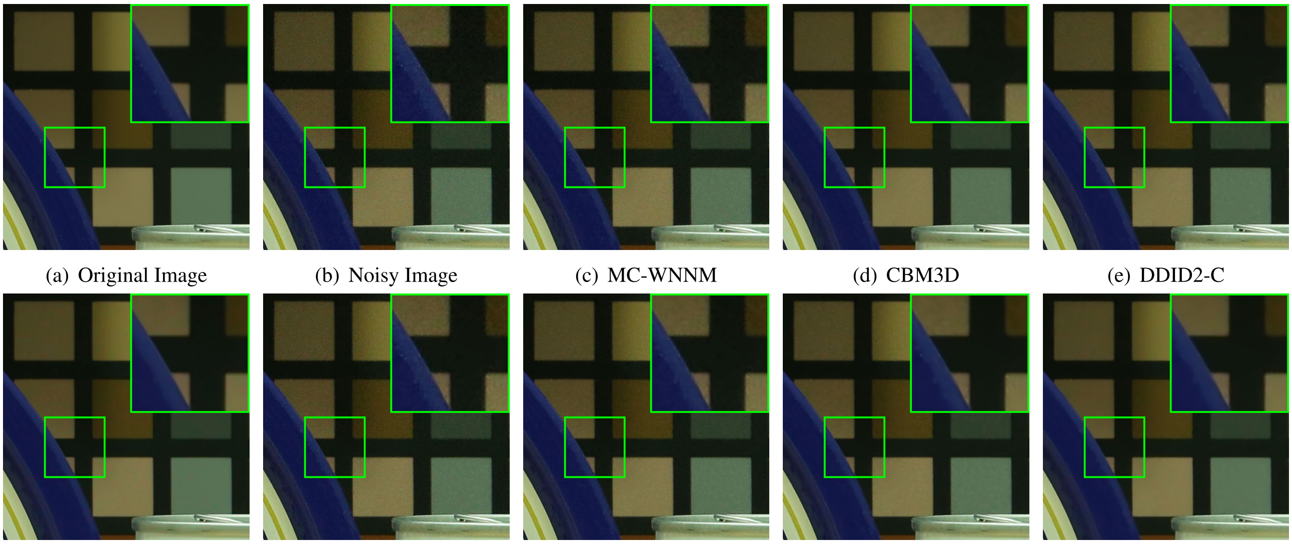
MC-WNNM [13]	CBM3D [32]	DDID2_C [34]	GSRC-NLP [30]	TWSC [28]	MSI Color-tSVD [29]	FFD-Net [19]	Proposed CTuNLM
Peak Signal to Noise Ratio (PSNR) in dB							
40.288	40.799	40.649	40.743	40.281	40.403	40.962	41.496
37.803	40.739	40.764	40.969	40.739	39.808	41.139	41.311
33.369	34.398	34.386	34.179	33.969	34.187	34.498	34.635
38.594	39.759	39.645	39.023	38.929	39.502	39.853	39.763
35.292	35.956	35.675	35.601	35.996	35.887	36.051	35.876
37.544	39.896	39.581	39.888	39.956	39.290	40.197	39.855
36.458	37.711	37.535	37.494	37.691	37.389	37.832	37.865
Structural Similarity Index Measure (SSIM) [62]							
0.9723	0.9802	0.9804	0.9801	0.9765	0.9768	0.9820	0.9848
0.9632	0.9828	0.9851	0.9822	0.9724	0.9788	0.9856	0.9867
0.9733	0.9785	0.9783	0.9766	0.9720	0.9767	0.9791	0.9804
0.9494	0.9595	0.9561	0.9432	0.9603	0.9626	0.9512	0.9377
0.9647	0.9776	0.9791	0.9722	0.9765	0.9757	0.9795	0.9759
0.9346	0.9629	0.9639	0.9523	0.9592	0.9584	0.9642	0.9586
0.9547	0.9741	0.9750	0.9705	0.9725	0.9701	0.9755	0.9765

on either signal or noise. According to [12], the "method noise" – as the difference between the clean ground truth image and the denoised image – should be as similar as to the noise, without any visible structural distortions, that are not part of the latent image. For any approximation prospective, a denoising approach with a smaller value of method noise adds less visual artifacts.

More often or not it is very hard to locate and identify any artifacts in the denoised image. This is because that (sometimes) artifacts increase the preserved realism of the denoised image, as structures in the image are recognized as details. We have conducted a few experiments using Renoir image dataset [60] and selected flat/homogeneous regions with almost constant image pixel intensity, as shown in Fig. 11. The selected image patches would highlight the denoising performance of each approach while yielding the notable artifacts manifesting as method noises. Fig. 12 illustrates the characteristic of artifacts for each denoising method with respective quantitative values. Most of the denoised outputs in Fig. 12 exhibit low-frequency noises and additionally suffer from loss of contrast. The use of GIF almost eliminates structural artifacts and flattens the low-frequency noise, while preserving the essential details. To complement the visual assessment, we provide the average PSNR and SSIM measures (in Table 3) to indicate the competitiveness of proposed approach in suppressing visual artifacts while favoring the image and noise modeling.

D. RUN-TIME COMPLEXITY

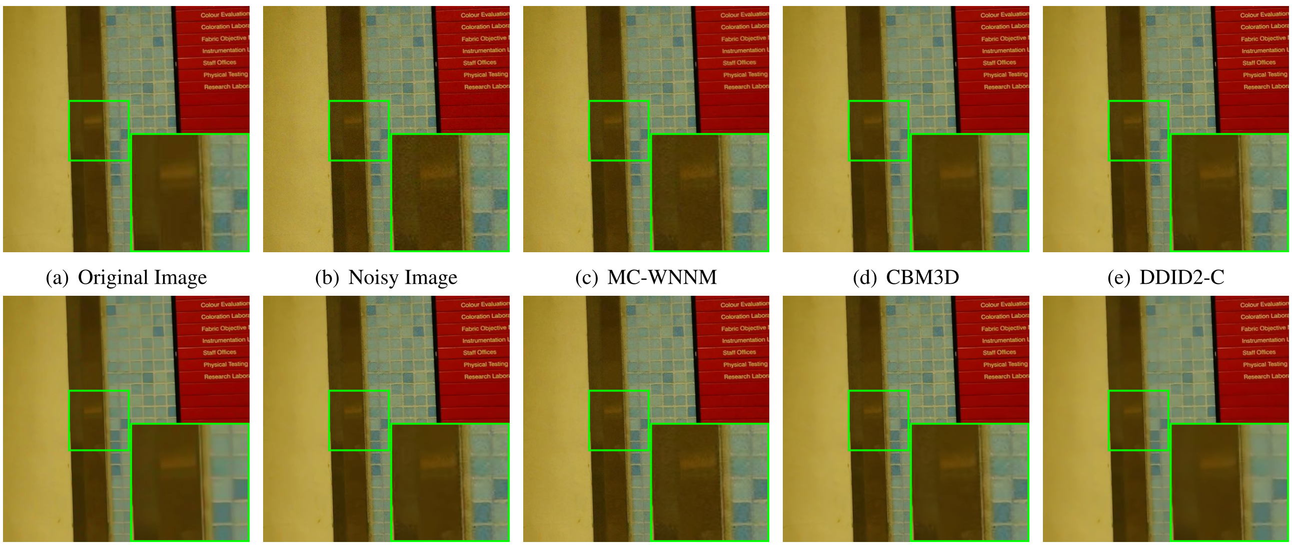
Table 4 compares the run-time complexity (in seconds) of all competing methods. Here, experiments were conducted using MATLAB 2022b environment on a machine with Intel(R) Core(TM) $i5-3210MCPU$@2.50GHz and 4GB-RAM. The average CPU run-time (in seconds) of different



(a) Original Image (b) Noisy Image (c) MC-WNNM (d) CBM3D (e) DDID2-C

(f) GSRC-NLP (g) TWSC (h) MSI Color-tSVD (i) FFDNet (j) CTuNLM

FIGURE 9. Visual and quantitative comparison 01. (b) PSNR = 37.9748, SSIM = 0.9458; (c) PSNR = 41.2228, SSIM = 0.9798; (d) PSNR = 42.8671, SSIM = 0.9847; (e) PSNR = 42.2852, SSIM = 0.9846; (f) PSNR = 42.4273, SSIM = 0.9837; (g) PSNR = 41.7750, SSIM = 0.9738; (h) PSNR = 41.7191, SSIM = 0.9775; (i) PSNR = 42.7998, SSIM = 0.9850; (j) PSNR = 42.9445, SSIM = 0.9894.



(a) Original Image (b) Noisy Image (c) MC-WNNM (d) CBM3D (e) DDID2-C

(f) GSRC-NLP (g) TWSC (h) MSI Color-tSVD (i) FFDNet (j) CTuNLM

FIGURE 10. Visual and quantitative comparison 02. (b) PSNR = 34.9890, SSIM = 0.9253; (c) PSNR = 36.7342, SSIM = 0.9459; (d) PSNR = 37.1164, SSIM = 0.9605; (e) PSNR = 36.7269, SSIM = 0.9596; (f) PSNR = 37.0585, SSIM = 0.9548; (g) PSNR = 36.9569, SSIM = 0.9523; (h) PSNR = 36.7353, SSIM = 0.9547; (i) PSNR = 37.1293, SSIM = 0.9613; (j) PSNR = 37.1813, SSIM = 0.9653.

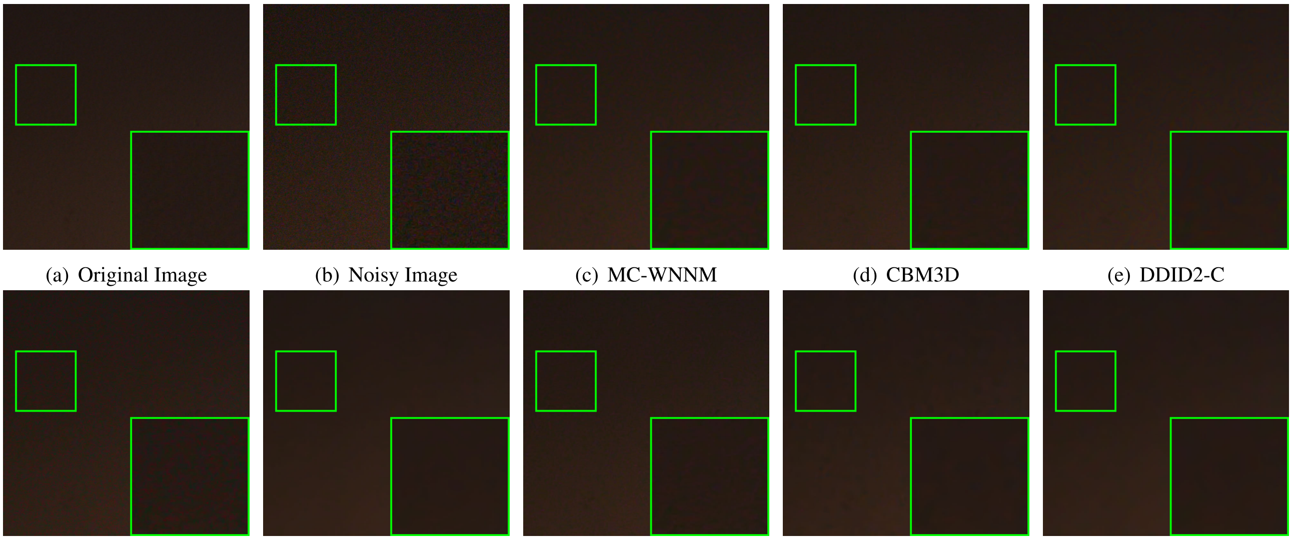
algorithms implemented on 500 images are shown in Table 4. We highlight the computational complexity of each background methods used for proposed algorithm: FDCT (USFFT) = $\mathcal{O}(M^2 log(M))$ (with three scale decomposition). Fast NLM filter is used in the approximation scale with $\frac{M}{2}$ coefficients and the finest scale with M coefficients. The use of look up-table for weight calculation and the partially C-codded algorithm accelerate the overall speed of the algorithm. Similarly, the GIF is computationally efficient with complexity equals to $\mathcal{O}(M)$ for M-pixels. However, as the proposed CTuNLM needs extra time to denoise the each YUV components for multi-channel implementation compared to the following methods CBM3D, MSI Color tSVD and FFDNet. Note although CBM3D is implemented using C++, mex-function and parallelization, the proposed approach is competitive with purely MATLAB implementation.

V. BEYOND CURVELET THRESHOLDING: DEEP CURVELET-NET

Curvelet transform is a multiscale and multidirectional image representation technique that has proven to be valuable in many image restoration tasks. However, we need to look beyond multiscale thresholding, while preserving fine details in the restored image. In this section, a pilot work demonstrating both the power of Convolutional Neural

(a) Noisy Image1 (b) Noisy Image2 (c) Noisy Image3 (d) Noisy Image4 (e) Noisy Image5

FIGURE 11. A sample of noisy flat/homogeneous regions with almost constant image pixel intensity. Images are obtained from Renoir image dataset [60].



(a) Original Image (b) Noisy Image (c) MC-WNNM (d) CBM3D (e) DDID2-C

(f) GSRC-NLP (g) TWSC (h) MSI Color-tSVD (i) FFDNet (j) CTuNLM

FIGURE 12. Artifact Study. The highlighted patches illustrate the method noise present in each denoised image. (b) PSNR = 30.584, SSIM = 0.6775; (c) PSNR = 37.016, SSIM = 0.8868; (d) PSNR = 38.501, SSIM = 0.9223; (e) PSNR = 38.409, SSIM = 0.9225; (f) PSNR = 39.0802, SSIM = 0.9277; (g) PSNR = 40.386, SSIM = 0.9533; (h) PSNR = 39.490, SSIM = 0.9308; (i) PSNR = 39.961, SSIM = 0.9495; (j) PSNR = 40.326, SSIM = 0.9542.

TABLE 3. Artifacts study.

MC-WNNM [13]	CBM3D [32]	DDID2_C [34]	GSRC-NLP [30]	TWSC [28]	MSI Color-tSVD [29]	FFD-Net [19]	Proposed CTuNLM
Peak Signal to Noise Ratio (PSNR) in dB							
29.796	31.510	30.512	29.777	31.561	31.153	32.151	32.443
Structural Similarity Index Measure (SSIM) [62]							
0.6772	0.7473	0.7482	0.7398	0.7618	0.7413	0.7716	0.7846

TABLE 4. Average and standard deviation of run-time complexity of various methods, implemented on 500 different color images.

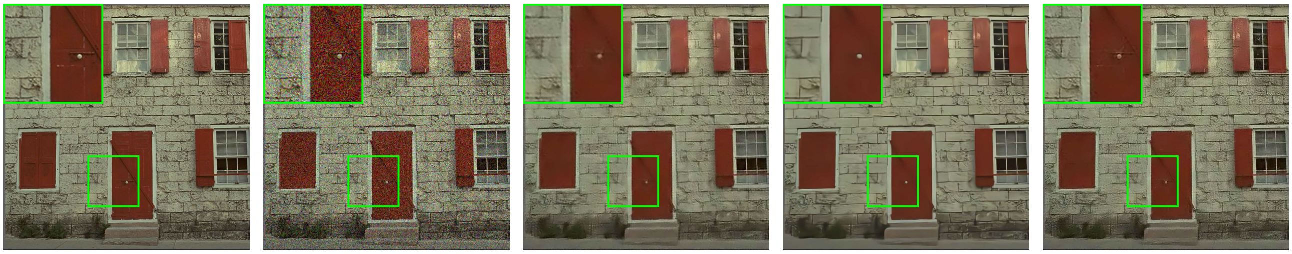
CPU Run-Time in Seconds							
MC-WNNM [13]	CBM3D [32]	DDID2_C [34]	GSRC-NLP [30]	TWSC [28]	MSI Color-tSVD [29]	FFD-Net [19]	Proposed CTuNLM
516.072 ± 18.656	**4.379 ± 0.181**	694.823 ± 44.884	625.904 ± 57.32	548.398 ± 27.694	9.372 ± 0.707	37.449 ± 1.362	26.642 ± 3.673

Networks (CNNs) and the discriminative ability of curvelet feature in the finest scale is investigated for the general problem of image denoising.[4] Prior to the application of deep-CNN model, we discus the curious-case of the finest scale coefficients and its importance in image denoising problems.

[4]While authors are still working on fully exploring the concept for other curvelet scales.

A. THE CURIOUS CASE OF THE FINEST SCALE

By focusing on the finest scale, curvelet transform enables the isolation and targeted denoising or artifact removal, leading to improved image restoration results. At the finest scale,



(a) Original Image (b) Noisy Image (σ = 30) (c) CTuNLM (d) FFDNet (e) CTuNLM with Original Finest Scale Coef.

FIGURE 13. The curious case of the Finest Scale. Illustration of denoising improvement, while preserving the original Finest scale coefficients. (b) PSNR = 18.617, SSIM = 0.6839; (c) PSNR = 27.493, SSIM = 0.8891; (d) PSNR = 27.811, SSIM = 0.8945; (e) PSNR = 30.347, SSIM = 0.9182.

$Y_{\gamma,h}$, coefficients are precisely aide in representing textures, and other significant image structures, making it easier to enhance or restore these specific features without affecting the rest of the image. To study the significance of finest scale coefficients in image restoration, we conducted several experiments on the TID2013 image dataset [63]. While considering the proposed CTuNLM method applied to all other curvelet scales but retaining the original image finest scale coefficients, we obtained the denoised image as shown in Fig. 13. From simulated AWGN with standard deviation σ ranging between $[5, 50]$, we found that on an average of 0.075 to 0.3 improvement in SSIM index and 1.5 to $3dB$ improvement in PSNR measure of the denoised images. Results demonstrated the importance of curvelet finest scale coefficients in preserving most indispensable latent image feature.

B. DEEP CURVELET-NET

Multiscale or scale-invariant feature modelling and spatial attention mechanisms using deep learning played a crucial role in shaping several research objectives by providing a way to exploit enhanced features from the input data. One way to utilize multiscale features is to design a multilayer CNN with varying receptive fields [64]. The conceptual denoising encoder-decoder model proposed in this study is depicted in Fig. 14. The proposed model takes advantage of both multiscale feature modelling and spatial attention mechanisms in a unique manner. Specifically, the encoder is meticulously designed using multilayers of depth-separable convolution layers (DSCL). Each layer incorporates three columns of DSCL. Thus, the proposed encoder consists of six layers with three columns of DSCL. Each layer's features are fused and surpassed by the next layer in order to transfer the multiscale features to the next layer.

On the other hand, the decoder, which is designed using single-column multilayers of depth-separable convolution layers, has the advantage of surpassing the multiscale spatial attention maps from each encoder layer to its corresponding decoder layer of a particular scale. Such decoder module we call as multiscale spatial attentive decoder. The preference for depth-separable convolutions over conventional convolutions is driven by the objective of reducing computational overhead during convolution processes [65]. During decoding the denoised FC, the fusion of multiscale spatial attention maps from the encoder at each scale of the decoding layer acts as additional supervision to improve the feature modelling capability. Such an attention map is obtained by using the Spatial Attention Block (SAB) [66] and applying it to the features at different scales of the encoder. The architecture details of SAB are presented in Fig. 14(b). Note that such spatially attentive features are used in the respective scale of the decoder through a skip connection, as shown in Fig. 14. We highlight the layer details of the proposed CNN architecture in Table 5.

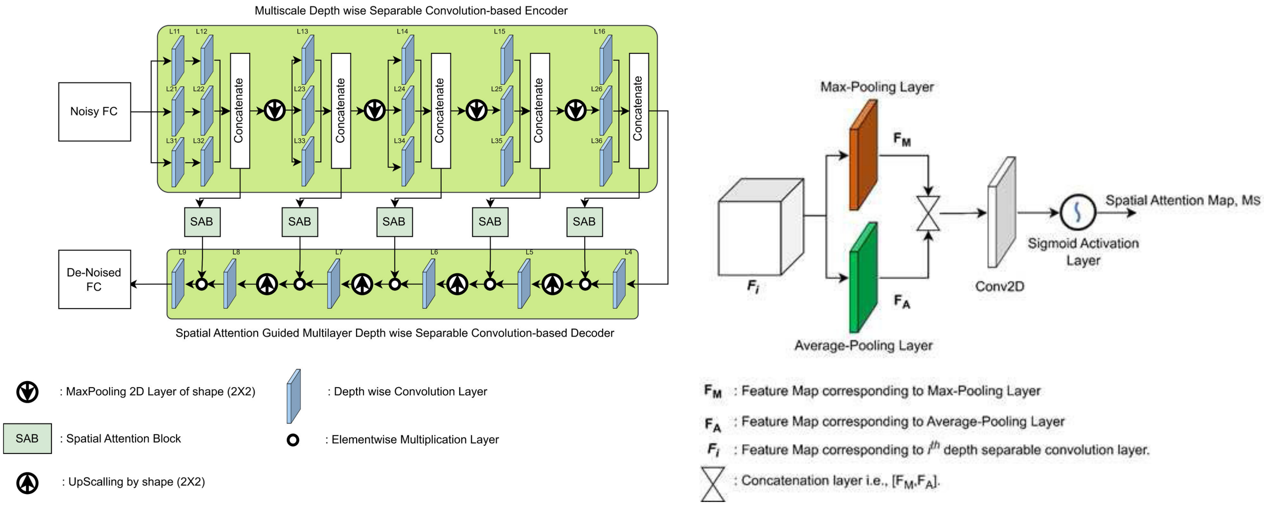
1) OPTIMIZATION

As illustrated in Eq. 10, the finest curvelet scale is denoted as $\lambda_{\gamma h}$ and from Eq. 1, we define set of latent coefficients as $X_{\lambda_{\gamma h}} = \{x1_{\lambda_{\gamma h}}, x2_{\lambda_{\gamma h}}, \cdots xn_{\lambda_{\gamma h}}\} \in \mathbb{R}^{H\times W\times N}$. Similarly, the noisy finest curvelet coefficients are represented by a set $Y_{\lambda_{\gamma h}} = \{y1_{\lambda_{\gamma h}}, y2_{\lambda_{\gamma h}}, \cdots yn_{\lambda_{\gamma h}}\} \in \mathbb{R}^{H\times W\times N}$. We formulate the minimization function as the squared error between the input and output denoised coefficients as stochastic objective function with parameters θ [51].

$$\underset{\theta}{\operatorname{argmin}} Loss = \underset{\theta}{\operatorname{argmin}}(\frac{1}{N} \times \sum_{i=1}^{N}(X_{\lambda_{\gamma h}} - Y_{\lambda_{\gamma h}})^2) \quad (13)$$

2) NETWORK TRAINING

The training process is carried out using the TID2013 image dataset [63], which encompasses 24 images representing a diverse range of natural scenes. To ensure uniformity, all reference images in the dataset are resized to 512 × 512 pixels. The noisy images are generated using simulated additive Gaussian noise (AWGN) of zero mean and standard deviations of $\sigma = [1, 70]$. The training leverages the curvelet finest scales from both reference and noisy images, employing the ADAM optimizer and Mean Squared Error (MSE) loss function. Specifically, the ADAM algorithm utilizes hyperparameters $\alpha = 0.01$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$, alongside a mini-batch size of 24. The learning rate is exponentially decayed from 0.001 to 0.0001 over



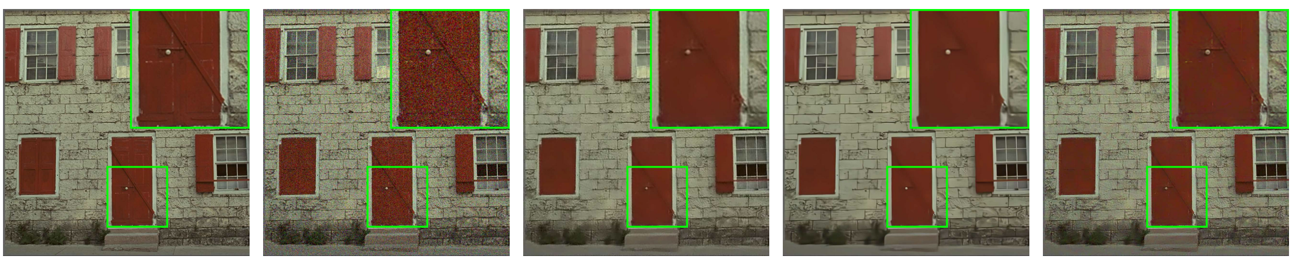
(a) Proposed AutoEncoder Model for Denoising. Note: The Curvelet Finest Scale Coefficients are denoted as FC.

(b) Spatial Attention Block (SAB)

FIGURE 14. **Proposed deep curvelet-net model.**

TABLE 5. **Layer details of proposed CNN architecture shown in Fig. 14.**

Layer Names	Kernel Shape	Depth Multiplier	Layer Names	Kernel Shape	Depth Multiplier
L11	11×11	2	L4	3×3	1
L12	9×9	2	L5	3×3	1
L13	7×7	2	L6	3×3	1
L14	5×5	2	L7	3×3	1
L15	4×4	2	L8	3×3	1
L16	3×3	2	L9	3×3	1
L21	9×9	2	L31	3×3	2
L22	7×7	2	L32	3×3	2
L23	4×4	2	L33	3×3	2
L24	3×3	2	L34	3×3	2
L25	3×3	2	L35	2×2	2
L26	3×3	2	L36	2×2	2

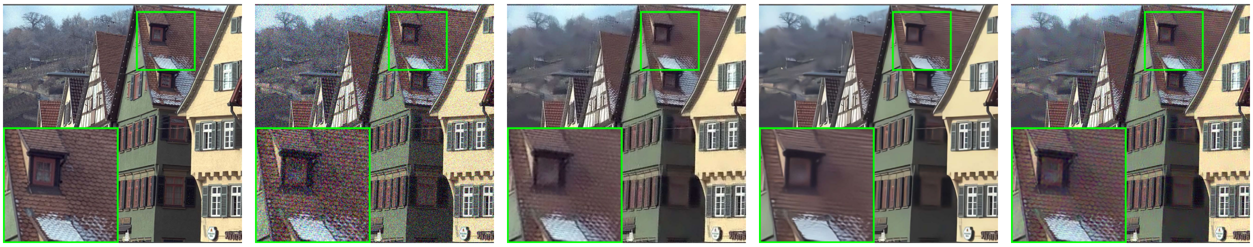


(a) Original Image (b) Noisy Image (σ = 20) (c) CTuNLM (d) FFDNet (e) CurveletNet

FIGURE 15. **Illustration of proposed deep CurveletNet Results applied on the Finest curvelet Coefficients. (b) PSNR = 22.150, SSIM = 0.8077; (c) PSNR = 28.733, SSIM = 0.9187; (d) PSNR = 28.836, SSIM = 0.9158; (e) PSNR = 29.711, SSIM = 0.9246.**

30 epochs. For the denoising process, training and testing involve 70% and 30% of coefficients, respectively, offering a comprehensive evaluation of the model's effectiveness.

We conducted a few initial experiments to demonstrate the effectiveness of proposed CNN architecture in image denoising. Fig. 15 & 16 illustrate the results obtained using deep CurveletNet architecture applied on the noisy curvelet finest scale, while keeping the estimation method same for other scales as shown in Fig. 5. Note the improvement in denoising quality, while maintaining both structural and



(a) Original Image (b) Noisy Image (σ = 30) (c) CTuNLM (d) FFDNet (e) CurveletNet

FIGURE 16. Illustration of proposed deep CurveletNet Results applied on the Finest curvelet Coefficients. (b) PSNR = 18.622, SSIM = 0.7115; (c) PSNR = 27.367, SSIM = 0.9046; (d) PSNR = 27.635, SSIM = 0.9077; (e) PSNR = 28.059, SSIM = 0.9103.

textural details. The quantitative measures PSNR and SSIM also validates the pilot attempt of curevlet coefficient estimation from its noisy observation using CNN architecture, as shown in Fig. 14. However, the performance of the proposed approach may be limited by the limitations of signal dependent noise representation/filtering in Curvelet domain for uncertain real-world photographs other than contaminated under low-light conditions. Therefore, authors are still working on developing a standalone CNN architecture using Curvelet coefficient for general inverse problems in image processing.

VI. CONCLUSION

Designing a standalone model for denoising real-world images contaminated from multiple sources of noise is always challenging; yet an indispensable problem in image processing. By adopting optimal sparsity in image representation, we propose a curvelet based denoising model that offers an efficient way to analyze both low-frequency and high-frequency noises separately. Multiscale filtering (fast NLM filter) in the approximation and the finest scale, while threshoding the coarser scales, the proposed CTuNLM algorithm is extended for denoising multi-channel real-world noisy images (with single input parameter). The use of Guided Image filter (GIF as post-processing operation in spatial domain) further enhances the quality of denoised images by suppressing ringing artifacts due to curvelet thresholding. The performance of proposed CTuNLM algorithm is compared with several recent methods including state-of-the-art BM3D and FFDNet algorithm. Experimental results validate the rationale of multiscale combined approach and exhibit superior denoising performance in terms of objective and subjective evaluation metrics on multi-channel real-world noisy images. In the second approach, by looking beyond multiscale curvelet filtering, we studied ''The Curious Case of the Finest Scale'', in search for further improvement in restoration quality. Thus, in a pilot work, an encoder-decoder based deep leaning CNN architecture with spatial attention block (SAB) known as ''Deep CurveletNet'' was developed for denoising the finest curvelet coefficients. Plugging these denoised coefficients in the CTuNLM logarithm, we observed a significant improvement in performance for images corrupted with simulated Gaussian noises. This innovative approach highlights the significance of supervised learning based CNN methods in estimating curvelet coefficients, while opening up new possibilities for further improvement in the broad domain of inverse problems in image processing.

REFERENCES

[1] C. Chen, Q. Chen, J. Xu, and V. Koltun, ''Learning to see in the dark,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3291–3300.

[2] S. K. Panigrahi, ''Image denoising by edge preserved curvelet thresholding,'' Ph.D. dissertation, Dept. Elect. Eng., Nat. Inst. Technol., Rourkela, Odisha, 2019.

[3] S. Anwar and N. Barnes, ''Real image denoising with feature attention,'' in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 3155–3164.

[4] S. Gu, L. Zhang, W. Zuo, and X. Feng, ''Weighted nuclear norm minimization with application to image denoising,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2862–2869.

[5] D. Zoran and Y. Weiss, ''From learning models of natural image patches to whole image restoration,'' in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 479–486.

[6] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pająk, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian, J. Kautz, and K. Pulli, ''FlexISP: A flexible camera image processing framework,'' *ACM Trans. Graph.*, vol. 33, no. 6, pp. 1–13, 2014.

[7] Y. Romano, M. Elad, and P. Milanfar, ''The little engine that could: Regularization by denoising (RED),'' *SIAM J. Imag. Sci.*, vol. 10, no. 4, pp. 1804–1844, Jan. 2017.

[8] K. Zhang, W. Zuo, S. Gu, and L. Zhang, ''Learning deep CNN denoiser prior for image restoration,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3929–3938.

[9] J. Zhang, Y. Cao, S. Fang, Y. Kang, and C. Wen Chen, ''Fast haze removal for nighttime image using maximum reflectance prior,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 7418–7426.

[10] Y. Liu, Z. Yan, J. Tan, and Y. Li, ''Multi-purpose oriented single nighttime image haze removal based on unified variational retinex model,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 4, pp. 1643–1657, Apr. 2023.

[11] N. Iqbal, M. Deriche, G. AlRegib, and S. Khan, ''Blind curvelet-based denoising of seismic surveys in coherent and incoherent noise environments,'' *Arabian J. Sci. Eng.*, vol. 48, pp. 10925–10935, Apr. 2023.

[12] A. Buades, B. Coll, and J.-M. Morel, ''A non-local algorithm for image denoising,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun. 2005, pp. 60–65.

[13] J. Xu, L. Zhang, D. Zhang, and X. Feng, ''Multi-channel weighted nuclear norm minimization for real color image denoising,'' in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 1096–1104.

[14] Y. Liu, Z. Yan, A. Wu, T. Ye, and Y. Li, ''Nighttime image dehazing based on variational decomposition model,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2022, pp. 640–649.

[15] Y. Liu, Z. Yan, T. Ye, A. Wu, and Y. Li, ''Single nighttime image dehazing based on unified variational decomposition model and multi-scale contrast enhancement,'' *Eng. Appl. Artif. Intell.*, vol. 116, Nov. 2022, Art. no. 105373.

[16] H. C. Burger, C. J. Schuler, and S. Harmeling, ''Image denoising: Can plain neural networks compete with BM3D?'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2392–2399.

[17] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1256–1272, Jun. 2016.

[18] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[19] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.

[20] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, "Toward convolutional blind denoising of real photographs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 1712–1722.

[21] R. S. Thakur, R. N. Yadav, and L. Gupta, "State-of-art analysis of image denoising methods using convolutional neural networks," *IET Image Process.*, vol. 13, no. 13, pp. 2367–2380, 2019.

[22] V. Murali and P. Sudeep, "Image denoising using DnCNN: An exploration study," in *Advances in Communication Systems and Networks: Select Proceedings of ComNet*. Singapore: Springer, 2020, pp. 847–859.

[23] W. Li, H. Liu, and J. Wang, "A deep learning method for denoising based on a fast and flexible convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5902813.

[24] Y. Mansour and R. Heckel, "Zero-shot noise2noise: Efficient image denoising without any data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2023, pp. 14018–14027.

[25] Y. Liu, Z. Yan, S. Chen, T. Ye, W. Ren, and E. Chen, "NightHazeFormer: Single nighttime haze removal using prior query transformer," in *Proc. 31st ACM Int. Conf. Multimedia*, Oct. 2023, pp. 4119–4128.

[26] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

[27] Z. Zha, X. Liu, Z. Zhou, X. Huang, J. Shi, Z. Shang, L. Tang, Y. Bai, Q. Wang, and X. Zhang, "Image denoising via group sparsity residual constraint," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 1787–1791.

[28] J. Xu, L. Zhang, and D. Zhang, "A trilateral weighted sparse coding scheme for real-world image denoising," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 20–36.

[29] Z. Kong and X. Yang, "Color image and multispectral image denoising using block diagonal representation," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4247–4259, Sep. 2019.

[30] Z. Zha, X. Yuan, B. Wen, J. Zhou, and C. Zhu, "Group sparsity residual constraint with non-local priors for image restoration," *IEEE Trans. Image Process.*, vol. 29, pp. 8960–8975, 2020.

[31] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[32] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, Sep./Oct. 2007, pp. I–313.

[33] C. Knaus and M. Zwicker, "Dual-domain image denoising," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 440–444.

[34] C. Knaus and M. Zwicker, "Dual-domain filtering," *SIAM J. Imag. Sci.*, vol. 8, no. 3, pp. 1396–1420, 2015.

[35] S. K. Panigrahi, S. Gupta, and P. K. Sahu, "Curvelet thresholding with multiscale NLM filtering for color image denoising," in *Proc. IEEE Region 10 Conf. (TENCON)*, Nov. 2017, pp. 2220–2225.

[36] V. Kiani, A. Harati, and A. Vahedian Mazloum, "Iterative wedgelet transform: An efficient algorithm for computing wedgelet representation and approximation of images," *J. Vis. Commun. Image Represent.*, vol. 34, pp. 65–77, Jan. 2016.

[37] V. Jain and S. Seung, "Natural image denoising with convolutional networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 21, 2008, pp. 1–8.

[38] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, Jan. 2006.

[39] J. Ma and G. Plonka, "The curvelet transform," *IEEE Signal Process. Mag.*, vol. 27, no. 2, pp. 118–133, Mar. 2010.

[40] H. Liu, R. Xiong, J. Zhang, and W. Gao, "Image denoising via adaptive soft-thresholding based on non-local samples," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 484–492.

[41] S. K. Panigrahi, S. Gupta, and P. K. Sahu, "Curvelet-based multiscale denoising using non-local means & guided image filter," *IET Image Process.*, vol. 12, no. 6, pp. 909–918, 2018.

[42] M. Lebrun, M. Colom, and J.-M. Morel, "The noise clinic: A blind image denoising algorithm," *Image Process. On Line*, vol. 5, pp. 1–54, Jan. 2015.

[43] ABSoft. *Neat Image*. Accessed: Aug. 20, 2023. [Online]. Available: https://ni.neatvideo.com/home

[44] Y. Hou, J. Xu, M. Liu, G. Liu, L. Liu, F. Zhu, and L. Shao, "NLH: A blind pixel-level non-local method for real-world image denoising," *IEEE Trans. Image Process.*, vol. 29, pp. 5121–5135, 2020.

[45] Y. Liu, S. Anwar, Z. Qin, P. Ji, S. Caldwell, and T. Gedeon, "Disentangling noise from images: A flow-based image denoising neural network," *Sensors*, vol. 22, no. 24, p. 9844, Dec. 2022.

[46] S. Wali, Z. Liu, C. Wu, and H. Chang, "A boosting procedure for variational-based image restoration," *Numer. Math., Theory, Methods Appl.*, vol. 11, no. 1, pp. 49–73, 2018.

[47] S. Wali, C. Li, A. Basit, A. Shakoor, R. Ahmed Memon, S. Rahim, and S. Samina, "Fast and adaptive boosting techniques for variational based image restoration," *IEEE Access*, vol. 7, pp. 181491–181504, 2019.

[48] S. Wali, A. Shakoor, A. Basit, L. Xie, C. Huang, and C. Li, "An efficient method for Euler's elastica based image deconvolution," *IEEE Access*, vol. 7, pp. 61226–61239, 2019.

[49] S. Wali, C. Li, M. Imran, A. Shakoor, and A. Basit, "Level-set evolution for medical image segmentation with alternating direction method of multipliers," *Signal Process.*, vol. 211, Oct. 2023, Art. no. 109105.

[50] Z. Xie, L. Liu, Z. Luo, and J. Huang, "Image denoising using nonlocal regularized deep image prior," *Symmetry*, vol. 13, no. 11, p. 2114, Nov. 2021.

[51] D. Kingma and J. B. Adam, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, vol. 5, 2015, p. 6.

[52] D.-J. Kroon. *Fast Non-Local Means 1D, 2D Color and 3D*. Accessed: Aug. 20, 2023. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/27395-fast-non-local-means-1d-2d-color-and-3d

[53] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sep. 1994.

[54] S. Sari, H. Roslan, and T. Shimamura, "Noise estimation by utilizing mean deviation of smooth region in noisy image," in *Proc. 4th Int. Conf. Comput. Intell., Model. Simul.*, Sep. 2012, pp. 232–236.

[55] M. Podpora, G. P. Korbas, and A. Kawala-Janik, "YUV vs RGB-choosing a color space for human-machine interaction," in *Proc. FedCSIS (Position Papers)*, 2014, pp. 29–34.

[56] Y. Cheng, Y. Li, and D. Xue, "Image denoising method based on curvelet cycle spinning," in *Proc. 4th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Oct. 2008, pp. 1–3.

[57] K. He and J. Sun, "Fast guided filter," 2015, *arXiv:1505.00996*.

[58] J. Xu, H. Li, Z. Liang, D. Zhang, and L. Zhang, "Real-world noisy image denoising: A new benchmark," 2018, *arXiv:1804.02603*.

[59] S. Nam, Y. Hwang, Y. Matsushita, and S. J. Kim, "A holistic approach to cross-channel image noise modeling and its application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1683–1691.

[60] J. Anaya and A. Barbu, "Renoir—A dataset for real low-light image noise reduction," *J. Vis. Commun. Image Represent.*, vol. 51, pp. 144–154, Feb. 2018.

[61] P. Chatterjee and P. Milanfar, "Is denoising dead?" *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 895–911, Apr. 2010.

[62] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[63] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. J. Kuo, "Color image database TID2013: Peculiarities and preliminary results," in *Proc. Eur. Workshop Vis. Inf. Process. (EUVIP)*, Jun. 2013, pp. 106–111.

[64] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 589–597.

[65] L. Kaiser, A. N. Gomez, and F. Chollet, "Depthwise separable convolutions for neural machine translation," 2017, *arXiv:1706.03059*.

[66] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 3–19.

SUSANT KUMAR PANIGRAHI received the M.Tech. and Ph.D. degrees from the Department of Electrical Engineering, National Institute of Technology (NIT) Rourkela, in 2012 and 2019, respectively. During his tenure as a Ph.D. Student, he was a Teaching Assistant and was involved in the development of embedded systems and real-time laboratory and curriculum design of a few B.Tech. and M.Tech. courses. He is currently an Assistant Professor with the School of Electrical and Electronics Engineering, VIT Bhopal University, Madhya Pradesh, India. He has an academic experience of more than five years. He has more than ten SCI and Scopus publications in various peer-reviewed conferences and journals. His research interests include inverse problems in image denoising, sparse signal representation and analysis, computer vision, and embedded system design and programming. He received the Best Paper Award in the IEEE Students' Conference at MNNIT Allahabad. He is a Reviewer with notable journals, such as *Optical Engineering*, *Journal of Electronics Imaging* (SPIE), and *IETE Journal of Research* (Taylor).

SANTOSH KUMAR TRIPATHY received the M.Tech. degree in CSE from S.I.E.T Dhenkanal under B.P.U.T Odisha and the Ph.D. degree in CSE from the Indian Institute of Technology (BHU) Varanasi, Uttar Pradesh. He has an academic experience of more than ten years and has taught various undergraduate courses and taken various academic responsibilities. He has five publications in SCI/SCIE-indexed and two publications in ESCI-indexed peer-reviewed reputed international journals. He also has six publications in various international conferences indexed in SCOPUS held in India and abroad. He is also active in reviewing articles from various reputed international journals, such as Springer, Elsevier, World-Scientific Publisher, and IOS Press. His research interests include computer vision, image processing, machine learning, deep learning, visual surveillance, medical image analysis, and image security.

ANIRBAN BHOWMICK (Senior Member, IEEE) received the Ph.D. degree from the Birla Institute of Technology, Mesra, with a focus on speech processing, in 2018. He was a Postdoctoral Fellow with the Sound and Audio Group, IISc Bengaluru, for one and a half years, where he worked on a project of audio analytics funded by DRDO. He is currently an Assistant Professor Senior with the School of EEE, VIT Bhopal University, Bhopal, Madhya Pradesh, India. He has published more than 25 papers in international journals and conferences. He is the Co-PI of SERB funded project and has almost nine years of teaching experience. His research interests include signal and image processing and communication and related fields. He is a Reviewer of many reputed journals namely, *Multidimensional System and Signal Processing*, *Applied Acoustics*, *IET Signal Processing*, *Computers and Electrical Engineering*, and IEEE Access.

SANTOSH KUMAR SATAPATHY (Member, IEEE) received the bachelor's degree in IT and the M.Tech. degree in CSE from BPUT, Rourkela, Odisha, India, in 2009 and 2012, respectively, and the Ph.D. degree in computer science and engineering from Pondicherry University (Central University), Puducherry, India, in 2022. He was an Assistant Professor with different colleges in Odisha for eight years. He is a regular reviewer of reputed journals, namely IEEE, Springer, Elsevier, and Taylor and Francis. He has published several papers in national and international journals and conferences and more than 50 publications in the Scopus database. He has also served on numerous organizing panels for international conferences and workshops. He is currently editing several books with Springer Nature, Elsevier, and CRC Press. His research interests include biomedical technologies, the Internet of Things, computational intelligence, and EEG analysis. He is a Reviewer of many reputed journals, such as *Scientific Reports* (Springer Nature), *Biocybernetics and Biomedical Engineering* (Elsevier), *Multimedia Tools and Applications* (Springer), *Journal of Healthcare Engineering* (Hindawi), *Sleep Oxford University* (Press), and *Neuro Computing* (Elsevier).

PAOLO BARSOCCHI received the M.Sc. and Ph.D. degrees in information engineering from the University of Pisa, in 2003 and 2007, respectively. Since 2017, he has been the Head of the Wireless Networks Research Laboratory. He is currently a Researcher with the Information Science and Technologies Institute (ISTI), National Research Council (CNR), Pisa, Italy. He has coauthored more than 150 articles and has published in international journals and conference proceedings. His research interests include the IoT, cyber-physical systems, indoor localization, and radio channel signal processing. He is a member of numerous program committees and editorial board of international journals and program chair of several conferences.

AKASH KUMAR BHOI (Member, IEEE) received the B.Tech., M.Tech., Ph.D. degrees. He was appointed as the honorary title of "Adjunct Fellow" with the Institute for Sustainable Industries and Liveable Cities (ISILC), Victoria University, Melbourne, Australia, from August 2021 to July 2022. He was the University Ph.D. Course Coordinator for research and publication ethics (RPE) with Sikkim Manipal University (SMU). He is currently an Adjunct Research Faculty with the Directorate of Research, SMU, and a Research Associate with the Wireless Networks (WN) Research Laboratory, Institute of Information Science and Technologies, National Research Council (ISTI-CRN), Pisa, Italy. He is also the former Assistant Professor (SG) with the Sikkim Manipal Institute of Technology and served for about ten years. He has published several papers in national and international journals and conferences and more than 150 publications registered in the Scopus database. He has also served on numerous organizing panels for international conferences and workshops. He is currently editing several books with Springer Nature, Elsevier, Routledge, and CRC Press. His research interests include biomedical technologies, the Internet of Things, computational intelligence, antenna, and renewable energy. He is a member of ISEIS and IAENG, an Associate Member of IEI and UACEE, and an editorial board member reviewer of Indian and international journals. He is a regular reviewer of reputed journals, namely IEEE, Springer, Elsevier, Taylor and Francis, and Inderscience. He is also serving as a guest editor for special issues of the journals, such as Springer Nature, Wiley, Hindawi, and Inderscience. He is listed in the World's Top 2% Scientists for single-year impact for the year 2022 (compiled by John P. A. Ioannidis, Stanford University, and published by Elsevier BV).

Open Access funding provided by 'Consiglio Nazionale delle Ricerche-CARI-CARE-ITALY' within the CRUI CARE Agreement