Multispectral Satellite Image Denoising via Adaptive Cuckoo Search-Based Wiener Filter

Shilpa Suresh, *Student Member, IEEE*, Shyam Lal, *Member, IEEE*, Chen Chen, *Member, IEEE*, and Turgay Celik, *Member, IEEE*

***Abstract*— Satellite image denoising is essential for enhancing the visual quality of images and for facilitating further image processing and analysis tasks. Designing of self-tunable 2-D finite-impulse response (FIR) filters attracted researchers to explore its usefulness in various domains. Furthermore, 2-D FIR Wiener filters which estimate the desired signal using its statistical parameters became a standard method employed for signal restoration applications. In this paper, we propose a 2-D FIR Wiener filter driven by the adaptive cuckoo search (ACS) algorithm for denoising multispectral satellite images contaminated with the Gaussian noise of different variance levels. The ACS algorithm is proposed to optimize the Wiener weights for obtaining the best possible estimate of the desired uncorrupted image. Quantitative and qualitative comparisons are conducted with 10 recent denoising algorithms prominently used in the remote-sensing domain to substantiate the performance and computational capability of the proposed ACSWF. The tested data set included satellite images procured from various sources, such as Satpalda Geospatial Services, Satellite Imaging Corporation, and National Aeronautics and Space Administration. The stability analysis and study of convergence characteristics are also performed, which revealed the possibility of extending the ACSWF for real-time applications as well.**

***Index Terms*— 2-D finite-impulse response (FIR) Wiener filter, adaptive cuckoo search (ACS) algorithm, metaheuristic optimization algorithms, satellite image denoising.**

I. Introduction

MULTISPECTRAL imagery (MSI) is steadily growing in its popularity as a digital means for remote sensing, terrain analysis, and detecting thermal signature. It is often used as a viable alternative for mapping applications when standard mapping and geodesy product become inadequate or outdated. The first and foremost attribute of the MSI is its capability to record spectral reflectance in different portions of the electromagnetic spectrum, which makes it useful in various applications [1], [2]. MS images embed the amount of reflectance and illumination of the scene procured, in a wide stretch of much narrower frequency bands than RGB images. The MSI system records such MS signal with the aid of multispectrum arrays of sensors. MS images thus convey an authentic representation of real-world scenes, and hence the performance measures of remote-sensing operations get enhanced. Compared with hyperspectral images (with 100–200 bands which record the acquired signal in a wide spectral range), MS images generally possess a less number of distinct spectral bands (i.e., 4–7) and hence is less bulky to process.

However, the MSI is often affected by noisy signals, thereby corrupting the original image. Such noise sources range from system calibration errors, recording equipment limitations, varying sensitivity of sensors, photon effects, and interfering natural phenomena [3]. Moreover, narrow-bandwidth and limited-radiance energy obtained via sensors increases the probability of thermal noises getting affected on the image pixels significantly. Such noises are inevitable in the satellite-based remote-sensing environment.

The characteristics of such interfering noises depend on the acquisition system as well as the type of the images to be processed. Mostly, this type of noise can be represented as a random process following the normal distribution (Gaussian) with zero mean. MS imaging systems consist of a large spectral redundancy, which implies that the obtained image with a range of frequency bands is correlated with each other. Hence, the noise removal task results in the elimination of minor spectral components. Thus, denoising of MS images remains as a challenging task because of lack of a robust approach.

The simplest way of denoising is to utilize the conventional 2-D denoising techniques to reduce noise in the MS image pixel-by-pixel or band-by-band. Such filtering approaches include the use of 2-D IIR and finite-impulse response (FIR) filters and adaptive filtering algorithms, such as 2-D least mean square (2-D-LMS) algorithm [4]–[6], 2-D normalized LMS (2-D-NLMS) algorithm [7], [8], and 2-D affine projection algorithm (2-D-APA) [6], [9]. Later, 2-D FIR adaptive Wiener filters were also developed which guaranteed an optimal trade-off between noise smoothing and inverse filtering [10]. The optimization of 2-D adaptive filter coefficients for the above-mentioned algorithms was performed using conventional optimization algorithms [7], [8].

Manuscript received September 25, 2017; revised December 15, 2017 and February 15, 2018; accepted March 7, 2018. Date of publication April 10, 2018; date of current version July 20, 2018. *(Corresponding author: Shyam Lal.)*

S. Suresh and S. Lal are with the Department of Electronics and Communication Engineering, National Institute of Technology Karnataka, Mangalore 575025, India (e-mail: shilparagesh89@gmail.com; shyam.mtec@gmail.com).

C. Chen is with the Center for Research in Computer Vision, University of Central Florida, Orlando, FL 32816-2365 USA (e-mail: chenchen870713@gmail.com).

T. Celik is with the School of Computer Science and Applied Mathematics, University of the Witwatersrand, Johannesburg 2000, South Africa, and also with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China (e-mail: celikturgay@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2018.2815281

Consequently, the advent of evolutionary and swarm intelligence-based algorithms paved the way to use its potential to solve similar iterative optimization problems [11]–[16]. For instance, Tzeng [14] proposed a 2-D adaptive FIR digital filter design using the genetic algorithm for filter coefficient optimization. Boudjelaba *et al.* conducted a comprehensive comparative study highlighting the merits and demerits of using evolutionary algorithms for designing 2-D adaptive FIR digital filters [15]. Such stochastic algorithms were then widely found use in the medical domain [17] as well as for denoising natural images [12], [13], [18], [19]. Recently, Bhandari *et al.* [20] extended the use of evolutionary algorithms for wavelet-domain-based MS satellite image denoising application. The authors used the adaptive differential evolution (JADE) algorithm to find the optimal subband thresholds to separate noise from image [20]. Although the authors claimed the algorithm to be superior compared with others, but it was less efficient in preserving the edges and other textural features.

The statistical estimation of the desired signal from the noisy signal using 2-D FIR adaptive Wiener filters is found to perform well, provided the filter coefficients are fine-tuned. The recent advancements in the field of artificial intelligence and the proven potential of metaheuristic algorithms motivated us to use the same for finding the optimal adaptive Wiener weights. A prior analytical study comparing the performance of different metaheuristic algorithms for solving nonlinear optimization problems was conducted before choosing cuckoo search (CS) algorithm in modeling the proposed denoising method [1], [21]. The CS algorithm was adopted accounting for its implementation simplicity and efficient solution exploitation and exploration strategies along its run. So as to further enhance the solution exploration strategy, with a commendable improvement in the algorithm's convergence capability, a computationally efficient adaptive CS (ACS) algorithm is proposed. Therefore, in this paper, an efficient 2-D FIR adaptive Wiener filtering method using the ACS algorithm (ACSWF) is proposed for denoising satellite images contaminated with additive white Gaussian noise (AWGN). The noise is assumed to be uniform in every band and the low-spectral correlations between bands are treated as self-evident. The major contributions of this paper are as follows.

1) A computationally efficient ACS algorithm is proposed by remodeling self-ACS (SACS) algorithm that was put forward by Li *et al.* for optimizing the weight vectors of an adaptive Wiener filter [22].
2) A robust ACS-based adaptive Wiener filtering (ACSWF) for denoising MS satellite images corrupted with the Gaussian noise is proposed.

Visual and numerical results highlight the superiority of the proposed method in restoring images corrupted with the Gaussian noise of different variance levels consistently. Performance stability and improved convergence capability of the proposed filter make it adaptable for other signal processing applications.

The rest of this paper is organized as follows. Section II presents a detailed study of various denoising techniques found in the literature for MS image denoising. Section III gives the details of the theory and implementation steps of the proposed ACSWF. Section IV presents the simulation results and discussions. Finally, Section V draws the conclusion.

II. Related Work

Unfortunately, the presence of noise in MS images not only affects the human interpretation but also limits the accuracy of the computational methods. The poor image quality also makes various quantitative measurements and computer-aided analysis challenging and unreliable. Hence, quality enhancement of such images by employing proper denoising methods becomes a prerequisite for all practical applications. In the last few decades, a large number of methods have been proposed for denoising MS and hyperspectral images. Focusing on the works done over the past decade for MS image denoising, we come across techniques using transform-domain-based schemes [23]–[26], nonlocal sparse models [27]–[30], anisotropic diffusion scheme [31], partial differential equations (PDEs) [28], nonlocal tensor based models [32], [33], and bilateral filtering [34], [35].

Scheunders and De Backer [23] proposed a Bayesian wavelet-based method for denoising MS satellite images using a prior noise-free image. The authors claimed that the proposed method is performing better compared with other MS image denoising algorithms. Chaux *et al.* put forward a nonlinear Stein-based estimator for wavelet denoising of multichannel data [24]. Experiments performed for denoising MS remote-sensing images significantly outperformed other wavelet-based methods. A wavelet-based MS image restoration technique was proposed by Duijster *et al.* based on an iterative expectation maximization algorithm, applying deconvolution and denoising steps alternately [36]. Experiments on Landsat and AVIRIS images highlighted the denoising efficiency of the proposed method over bandwise method [36]. Despite the aforementioned merits, wavelet-based methods are computationally complex because of domain transformation, and an inappropriate selection of basis functions or subband thresholds for denoising can cause blurring and ringing artifacts around edges.

Mairal *et al.* [27] introduced a new image restoration model combining the nonlocal means and sparse coding approaches. Quantitative and qualitative experiments on images corrupted with synthetic or real noise showed the effectiveness of the proposed model compared with other state-of-the-art denoising methods, in the expense of computational complexity. Prasath and Singh [31] proposed an MS image denoising scheme using coupled PDEs with anisotropic diffusion. The well posedness of the scheme guaranteed its stability with efficient prefiltering capacity, whereas the increased computational complexity made the scheme practically difficult to process images of bigger size and more than three channels [31]. A similar remote-sensing image denoising method using PDEs and auxiliary image priors (PDE-AIP) was proposed by Liu *et al.* [28] in 2012. Visual results and quantitative indicators proved the proposed method to be particularly suited for denoising images corrupted with high-variance noise. In 2014, Peng *et al.* [32] proposed a decomposable

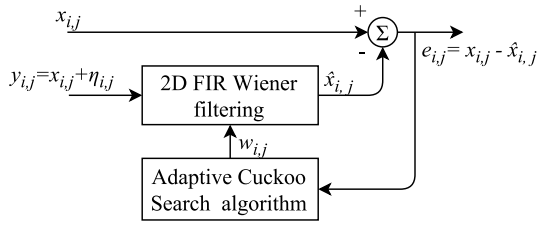

Fig. 1. Block schematic of the proposed ACSWF.

nonlocal tensor-based dictionary learning technique for MS image denoising. The proposed method ameliorated the image quality considerably but failed to preserve the structural details compared with others. Recently, Xie *et al.* [33] proposed a new MSI denoising model using a newly designed tensor-based sparsity measure. Experimental results substantiated the superiority of the proposed denoising scheme beyond the state of the arts in recovering the fine and course grained structures. At the same time, all tensor-based image processing techniques account for an increased computation time and memory. Peng *et al.* [34] used optimized vector bilateral filtering for MS image denoising which provided fair tradeoff between noise filtering and edge degradation. The authors selected the image-dependent filter parameters by optimization procedure based on Steins principle. In 2017, Papari *et al.* [35] introduced a bilateral filtering scheme for 3-D images. The filtering scheme was computationally efficient, but it resulted in reasonable image artifacts particularly near edges. Similarly, a plethora of literature is available on hyperspectral image denoising which can be roughly categorized as transform-based methods [37], filter-based methods [38], regularization-based methods [39]–[41], and kernel based methods [42].

III. Adaptive Cuckoo Search-Based Wiener Filter (ACSWF)

A. Theory Behind the Proposed ACSWF

This section presents the theory and fundamental concepts of the proposed ACS algorithm-based Wiener filtering for the denoising of satellite images corrupted with the AWGN. The key idea of the proposed method is to find the best possible estimate of the original image using 2-D FIR Wiener filtering. The 2-D FIR Wiener filter is modeled to adaptively modify its window weights, to minimize the mean square error (MSE) between the desired image and the filter output. The proposed denoising method uses the ACS algorithm for optimizing those filter weights ensuring the least possible mean squared error as compared with other similar metaheuristic algorithms. The block schematic of the proposed ACSWF for satellite image denoising is shown in Fig. 1.

The Wiener filtering theory assumes the signals to be stationary. The 2-D FIR Wiener filters are block-adaptive, wherein it calculates the filter coefficients periodically for a predefined block size of $n \times n$ samples [43].

Let $x_{i,j}$ be the original uncorrupted image pixel located at spatial location i, j and $y_{i,j}$ be its noisy counterpart contaminated with signal independent AWGN, $\eta_{i,j}$. Modeling a 2-D FIR Wiener filter for denoising this image, to obtain a linear estimate $\hat{x}_{i,j}$, requires minimizing the MSE value between $\hat{x}_{i,j}$ and $x_{i,j}$. It can be mathematically formulated as in

$$\text{MSE} = \sum_{i=1}^{M}\sum_{j=1}^{N}[\hat{x}_{i,j} - x_{i,j}]^2 \tag{1}$$

where $M$and$N$ denote the dimension of the input image to be processed. The linear estimate $\hat{x}_{i,j}$ of the desired signal obtained using Wiener filtering can be evaluated using [44], [45]

$$\hat{x}_{i,j} = \frac{\sigma^2_{x_{i,j}}}{\sigma^2_{x_{i,j}} + \sigma^2_{\eta_{i,j}}}[y_{i,j} - \mu_{x_{i,j}}] + \mu_{x_{i,j}} \tag{2}$$

where

$$y_{i,j} = x_{i,j} + \eta_{i,j}.$$

Parameters μ and σ^2 indicate the mean and variance of the signal, assuming $x_{i,j}$ to be a white Gaussian process and the noise mean to be zero. For estimating the linear estimate $\hat{x}_{i,j}$, we assume that the mean and variance values of the AWG noise ($\mu_{\eta_{i,j}}$, $\sigma^2_{\eta_{i,j}}$) are known [46], [47]. Hence, the prime focus is on estimating the mean and variance of the desired input signal, $\mu_{x_{i,j}}$ and $\sigma^2_{x_{i,j}}$, respectively. In our proposed method, the mean and variance measures of $x_{i,j}$ are usually estimated using the method devised by Kuan *et al.* [45]. The local signal statistics estimation using adaptive weight factors formulated by Kuan *et al.* [45] overcame the severe blurring around the edges of the images introduced using the local linear minimum MSE [44].

The authors proposed using monotonically decreasing functions such as the Gaussian function for calculating the filtering window weights, by asserting more confidence on variance estimate at the center of the window used. The use of adaptive window weights, rather than deterministic weights by Kuan *et al.* [45], seemed to be more appropriate and reliable for image filtering. Hence, we propound the use of adaptive Wiener weights in our proposed denoising method driven by the ACS algorithm (ACSWF). The optimal weights thus obtained are used for estimating the local statistics of the linear estimate of the desired signal. The estimated mean and variance measures using adaptive Wiener weights are formulated as given in

$$\begin{aligned}\hat{\mu}_{x_{i,j}} &= \sum_{a=i-n}^{i+n}\sum_{b=j-n}^{j+n} w_{i,j,a,b}y_{a,b} \\ \hat{\sigma}^2_{x_{i,j}} &= \sum_{a=i-n}^{i+n}\sum_{b=j-n}^{j+n} (w_{i,j,a,b}(y_{a,b} - \hat{\mu}_{x_{i,j}}))^2\end{aligned} \tag{3}$$

where $w_{i,j,a,b}$ is the adaptive Wiener weight vector.

B. Proposed Adaptive Cuckoo Search Algorithm

The CS is a stochastic metaheuristic algorithm evolved mimicking the obligate brood parasitic behavior shown by some cuckoo species [48]. The use of *Lévy* flight strategy rather than Brownian random walks for solution space exploration and its implementation simplicity because of the use of a

single control parameter p_a (switching parameter) are the key factors which make it superior compared with others. The flow of the CS algorithm is based on three idealized rules, wherein the single control parameter (p_a) denotes the probability of discovering alien eggs by the host species. Practical implementation of this scenario includes replacing p_a proportion of the current solution set with new ones [48]. Similarly, *Lévy* flight strategy for random walks follows variable step sizes ensuring least chances for oversampling, which eventually improved the convergence rate of the algorithm. Thus, the CS algorithm uses a balanced combination of local and global random walks controlled by a "switching parameter" p_a [48].

In order to further improve the solution searching capability of the CS algorithm, many CS variants were proposed for different optimization problems. A prior extensive study was undertaken comparing the potential of various ACS variants in the literature, for implementing the above-defined satellite image denoising problem [22], [49]–[52]. The SACS developed by Li *et al.* was found to perform better in the evaluated set, for our defined problem of satellite image denoising [22]. The SACS algorithm was developed by modifying and combining the mutation strategies used by two common variants of differential evolution (DE) algorithm [53]. The structure of any DE variants is generalized by denoting them as DE/x/y/z, where x is the mutation vector, y is the number of vectors used for mutation, and z indicates the crossover scheme [binomial (bin) or exponential (exp)] employed [2], [54]. The authors used a scale factor "ϕ" to control the pace of the algorithm to reach optimality. The scale factor "ϕ" was drawn in each iteration from a Gaussian distribution of mean 0.5 and standard deviation 0.1 [22].

Further investigations proved that the random initialization of "ϕ" values repeatedly in each iteration did not help in improving the convergence rate or enhancing the performance. At the same time, it resulted in an unwanted increase in the computational complexity of the algorithm. Hence, we remodeled the above-mentioned approach by replacing the scale factor "ϕ" with the switching parameter "p_a."

The mutation strategies used by two DE variants given in (4) are modified and adopted in the proposed ACS algorithm [22]

$$\begin{aligned}&\text{DE/rand/1/bin}: w_{k,G+1} = w_{r_1,G} + F\cdot(w_{r_2,G} - w_{r_3,G})\\&\text{DE/best/1/bin}: w_{k,G+1} = w_{\text{best},G} + F\cdot(w_{r_1,G} - w_{r_2,G}\\&\qquad\qquad + w_{r_3,G} - w_{r_4,G}) \quad (4)\end{aligned}$$

where $k \in \{1,\ldots,N_p\}$, r_1, r_2, r_3, and r_4 are random integer indices selected from $k = \{1,\ldots,N_p\}$, x_{best} is the mutation vector, N_p is the population size, and F is the scaling factor used in the DE algorithm. $F \in [0, 1]$ helps in avoiding search stagnation of the algorithm by controlling the effect of the difference vector in mutation operation [54]. Equation (5) gives the modified form of (4) included in the proposed ACS algorithm

$$\begin{aligned}&\text{CS/rand/1/bin}: w_{k,G+1} = w_{r_1,G} + p_a\cdot(w_{r_2,G} - w_{r_3,G})\\&\text{CS/best/1/bin}: w_{k,G+1} = w_{\text{best},G} + p_a\cdot(w_{r_1,G} - w_{r_2,G}\\&\qquad\qquad + w_{r_3,G} - w_{r_4,G}) \quad (5)\end{aligned}$$

where p_a is the "switching parameter" in the CS algorithm. The pseudocode of this phase is as given in the following.

if $\text{rand}_k \geq 1-\frac{G}{G_{\max}}$ **then**
  CS/best/1/bin ;
else
  CS/rand/1/bin ;
end

In this pseudocode, rand_k, $k = \{1,\ldots N_p\}$ is a random number drawn from a uniform probability distribution, whose value lies within the range [0, 1]. The parameter G indicates the current iteration (generation) count and $G_{\max}$ denotes the total number of iterations included. Thus, in the ACS algorithm, the discovery and randomization stages of the CS algorithm are modified by adopting the two mutation strategies following the pseudocode given above. The randomization stage selects any of the two strategies by comparing the value of $(1 - (G/G_{\max}))$ with the $\text{rand}_k \in [0, 1]$ value, generated in each iteration for the entire population. It indicates that the probability of selecting any of the two search strategies is a function of the iteration count G. Hence, if rand_k is less than $(1 - (G/G_{\max}))$, CS/rand/1/bin is chosen, or else CS/best/1/bin is selected as the randomization strategy for the ACS algorithm. The selection between these two mutation strategies ensures an increased probability of exploration during the initial iterations and increased exploitation toward the final iterations. It can be easily analyzed, since it is evident that the value of $(1 - (G/G_{\max}))$ decreases from 1 to 0 as it proceeds through the iterations. Thus, the probability of selecting CS/rand/1/bin is more in the initial set of iterations, whereas the probability of choosing CS/best/1/bin increases toward the end. Hence, based on two new search strategies controlled by a linear decreasing probability rule, the ACS algorithm does a better balancing of its exploration and exploitation phases [22].

C. Implementation of Proposed ACSWF

The pseudocode and stepwise implementation details of the proposed ACSWF are given in the following.

Step 1: Initialize the solution space randomly using

$$w_{k,l} = w_{k,l}^{\min} + \text{rand}\left(w_{k,l}^{\max} - w_{k,l}^{\min}\right) \quad (6)$$

where $w_{(k,l)}$ denotes weights allotted for 2-D FIR Wiener filter. The subscripts $k = 1,\ldots,N_p$, where N_p is the population size, and $l = 1,\ldots,d$, where $d = n^2$ is the total number of coefficients required to form the filter weight matrix. The boundary constraints $[w^{\min}, w^{\max}]$ are set to $[-1, 1]$.

Step 2: The $n \times n$ weight matrix for 2-D FIR Wiener filtering is formed by the 2-D lexicographic conversion of each of the candidate solution vector from the entire population using

$$[w_{k,1}, w_{k,2},\ldots, w_{k,n}, w_{k,n+1},\ldots, w_{k,2n}, w_{k,n(n-1)+1}\ldots w_{k,d}] \Rightarrow \begin{bmatrix} w_{k,1} & \cdots & w_{k,n} \\ \cdots & \cdots & \cdots \\ w_{k,n(n-1)+1} & \cdots & w_{k,d} \end{bmatrix}. \quad (7)$$

Algorithm 1: Proposed ACSWF

// **Initialization**

Population initialization: $w_{k,l}$ where k = 1, 2,....N_p; l = 1, 2,.....d; N_p = population size; d = dimensionality;

Parameter initialization: Switching parameter (p_a), Maximum number of iterations (G_{max});

// **2D FIR Wiener filtering and MSE calculation**

2D lexicographic ordering of weight vectors w_k using equation (8);

Estimate $\hat{x}_{i,j}$ using equations (2) and (3) with $w_{i,j,a,b} = w_k$;

Calculate fitness (MSE) values (F_k) using equation (1) ;

Record the best fitness (MSE) value and the corresponding solution set (w_{best});

// **Adaptive Cuckoo Search algorithm**

for *(*$G \leq G_{max}$*)* **do**

  Lévy flight modeling of random walks to generate new solution sets following equation (9);

  Estimate $\hat{x}_{i,j}$ using updated solution set and compute their respective fitness values.;

  for *all k* **do**

    if $F_k^{new} > F_k$ **then**

      $F_k = F_{newk}$;

      $w_{\text{best}} = w_{new}$;

    end

  end

  for *all k* **do**

    if $\text{rand}_k \geq 1 - \frac{G}{G_{\max}}$ **then**

      CS/best/1/bin : $w_{k,G+1} = w_{r_1,G} + p_a \cdot \left(w_{r_2,G} - w_{r_3,G}\right)$;

    else

      CS/rand/1/bin : $w_{k,G+1} = w_{\text{best},G} + p_a \cdot \left(w_{r_1,G} - w_{r_2,G} + w_{r_3,G} - w_{r_4,G}\right)$;

    end

  end

  Output estimation and fitness value computation following equations (2), (3) and (1);

  Comparison of newly generated fitness values with the earlier set and recording the best (w_{best}) so far.;

end

Estimate $\hat{x}_{i,j}$ using the optimal adaptive Wiener weight vector as w_{best} following equations (2) and (3);

Step 3: Estimate the linear estimate of the desired signal $\hat{x}_{i,j}$ following (2) and (3) using the adaptive Wiener weight vector as w_k.

Step 4: Compute the fitness (MSE) values of the estimated filter outputs obtained using each possible weight matrix formed in Step 2 using (1). Repeat if number of iterations $G < G_{max}$.

Step 5: Retain the best possible weight matrix (solution) in the previous iteration and generate new random solutions by *Lévy* flights around the previous solution set [48]. The new population set thus formed follows:

$$w_{k,G+1} = w_{k,G} + \alpha L(s,\beta) \quad (0 < \beta \leq 2, \alpha = 0.01)$$
$$L(s,\beta) = \frac{\beta\Gamma(\beta)\sin(\frac{\pi\beta}{2})}{\pi} \cdot \frac{1}{s^{1+\beta}}; \quad (|\ s\ | \geq |\ s_0\ |) \tag{8}$$

where s represents the step size, s_0 is the smallest step (typically 0.1–1) and α represents the scaling factor for step size. G represents the iteration/generation count, i.e., $w_{k,G}$ and $w_{k,G+1}$ are the weight vectors formulated for the G^{th} and $G+1^{\text{th}}$ generation of the ACS algorithm, respectively.

Step 6: Estimate $\hat{x}_{i,j}$ using the new set of weight matrices formed, and compute the fitness value of the newly generated solutions using (1). Memorize the best solution.

Step 7: Apply mutative randomization to the existing solution set as explained in Section II-B and update the new solution set using (5).

Step 8: Compute the estimated output using the updated solution set and evaluate their respective fitness values.

Step 9: Increment the iteration count by 1, i.e., $G = G + 1$. Until $G = G_{\max}$ (maximum iterations).

Step 10: Obtain the optimal filter weights $w_{\text{best},G_{\max}}$ and denoise $y_{i,j}$ using them to get the best possible estimate of the desired signal.

TABLE I

QUANTITATIVE RESULT COMPARISON BETWEEN 2-D AWF (WITHOUT ACS), CSAWF, AND THE PROPOSED ACSWF FOR IMAGES 1 AND 2 WITH THREE DIFFERENT NOISE VARIANCE LEVELS

σ^2	Metric	2D AWF	CSAWF	SACSAWF	ACSWF	2D AWF	CSAWF	SACSAWF	ACSWF
		Image 1				Image 2			
10%	MSE	222.8147	128.4861	100.1683	**78.2289**	217.2714	209.313	205.4283	**191.3700**
	PSNR	24.6514	27.0422	28.1235	**29.1971**	24.7608	24.9228	25.0042	**25.3121**
	FSIM	0.9668	0.9756	0.9761	**0.9779**	0.9776	0.9776	0.9778	**0.9780**
	UQI	0.9903	0.9984	0.9985	**0.9987**	0.9021	0.9498	0.9502	**0.9516**
	NAE	0.1443	0.1014	0.0976	**0.0547**	0.1961	0.1924	0.1900	**0.1817**
	Time (s)	7.7730	7.5640	14.2640	10.5230	7.0300	7.9210	13.5440	9.8560
20%	MSE	255.5587	137.303	130.1459	**120.7792**	340.1693	232.4691	206.2720	**184.498**
	PSNR	24.0559	26.754	26.9865	**27.3109**	22.8139	24.4672	24.9864	**25.4709**
	FSIM	0.9440	0.9696	0.9702	**0.9842**	0.9599	0.9716	0.9802	**0.9817**
	UQI	0.9901	0.9984	0.9986	**0.9988**	0.8841	0.9452	0.9498	**0.9504**
	NAE	0.1542	0.1021	0.0985	**0.0701**	0.2455	0.1964	0.1846	**0.1155**
	Time (s)	8.9900	9.4360	14.2430	10.5240	8.9730	9.1020	16.5423	11.5236
30%	MSE	345.7862	137.6731	129.7479	**126.0928**	376.9581	232.0443	215.8591	**207.6128**
	PSNR	22.7427	26.7423	26.9998	**27.1239**	22.3679	24.4751	24.7891	**24.9583**
	FSIM	0.9324	0.9695	0.9701	**0.9760**	0.9478	0.9717	0.9751	**0.9776**
	UQI	0.9846	0.9983	0.9986	**0.9986**	0.875	0.9179	0.9199	**0.9221**
	NAE	0.2123	0.1022	0.0986	**0.0832**	0.2653	0.1963	0.1786	**0.1392**
	Time (s)	11.9860	12.5860	18.4533	14.0123	13.1010	13.4260	19.5640	15.1268

D. Illustration of Proposed ACSWF for Image Denoising

Experiments were conducted for denoising test images corrupted with the Gaussian noise of three different noise variance levels using 2-D adaptive Wiener filter (2-D AWF), CS-based AWF (CSAWF), SACS-based AWF (SACSAWF), and the proposed ACSWF. The results obtained were compared to substantiate the effect of ACS algorithm in optimizing Wiener weights compared with the other. The qualitative and quantitative results obtained for denoising images contaminated with the Gaussian noise is presented in Fig. 2 and Table I, respectively. The quantitative metrics compared include MSE [55], peak signal-to-noise ratio (PSNR) [56], feature similarity index (FSIM) [57], universal quality index (UQI) [58], normalized absolute error (NAE) [59], and CPU running time. Lower MSE and NAE values along with higher PSNR, FSIM,

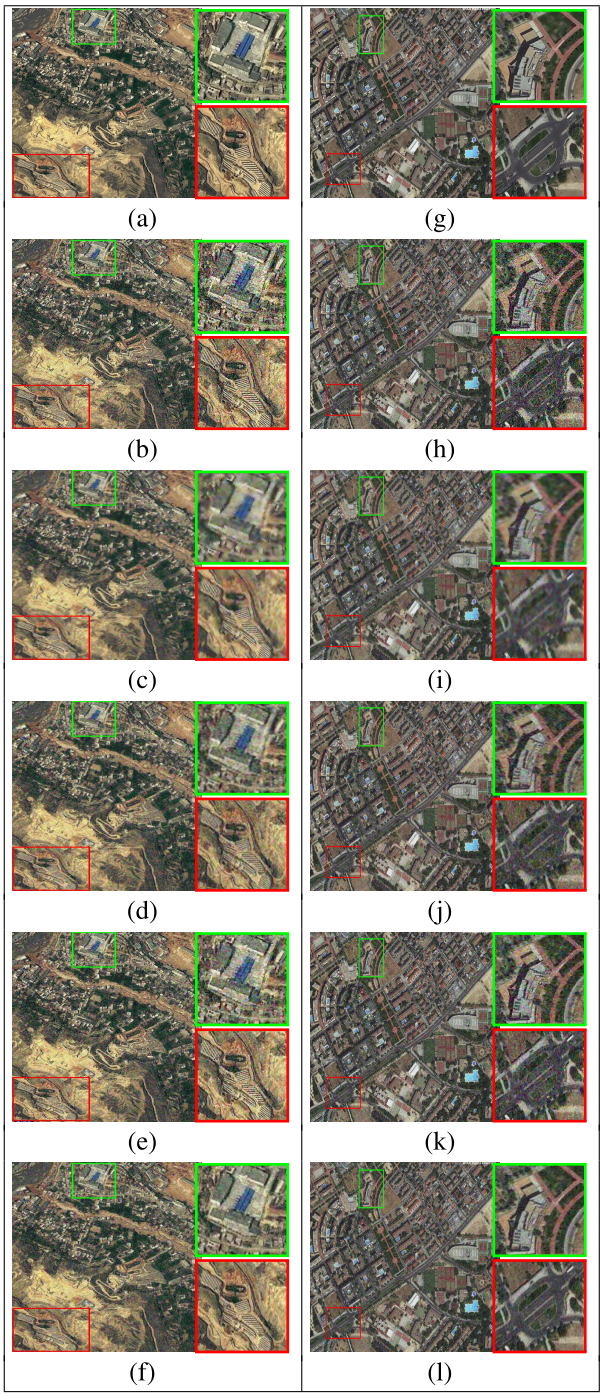

Fig. 2. Simulation results of denoising test images corrupted with Gaussian noise of 30% variance level using 2-D AWF, CSAWF, SACSAWF, and ACSWF. (a)–(f) Image 1. (g)–(l) Image 2. Best zoomed on screen.

and UQI values are the preferable conditions for evaluating the performance of the denoising algorithms. The quantitative analysis given in Table I provides two major inferences: 1) with weight search algorithms, such CS and SACS [22], the image denoising performance can be improved and 2) as compared with the other search algorithms, such as CS and SACS [22], our proposed ACS can achieve superior performance, demonstrating its effectiveness. The visual results also indicate that the use of ACS algorithm in optimizing Wiener weights has a great impact in estimating the desired image. The ACSWF proved to be very efficient in removing the AWGN from the corrupted image, preserving the relevant edges and other features of the image.

The 2-D AWF selects the weight vectors for a pixel (i, j) adaptively by biasing the estimated statistical measures in favor of pixels with values similar to $y_{i,j}$

$$w(i, j, a, b) = \frac{A(i, j)}{1 + \xi(\max\left[\epsilon^2, (y_{i,j} - y_{a,b})^2\right]}$$

$$A(i, j) = \left\{ \sum_{a,b} \frac{1}{1 + \xi(\max\left[\epsilon^2, (y_{i,j} - y_{a,b})^2\right]} \right\}^{-1} \quad (9)$$

where $\xi > 0$, $\epsilon = 2.5\sigma_\eta$, $w(i, j, a, b) = 0$, and $A(i, j)$ is the normalization constant. Parameter ξ was chosen such that $\xi\epsilon^2 \gg 1$ so as to limit outliers [60]. Since our test images are complex scenes with profound edge and textual information, this adaptive weight selection constraint will tend to fuse the closely laid edges. It reflects as blurring along the edges and minor textural features for such images, whereas this approach works effectively for preserving edges of less complex scenes.

The use of the metaheuristic algorithm effectively enhances this adaptive weight vector selection problem. The use of the CS algorithm for optimizing the weights (CSAWF) works well in preserving the edges, since the selection of weight factors are solely based on the fitness function, which is to reduce the MSE value between the estimated pixels and the original image. The use of uncorrupted image prior helps in optimizing the wiener weights to estimate a close enough approximation of the original image. The proposed ACSWF approach includes a better balance between solution exploration and exploitation strategies, which efficiently explores every single possibility for the optimal weight vectors which reduces the calculated MSE value. It eventually reduces the chances of getting clogged in a near optimal solution set as seen using the CS algorithm [2]. Hence, the selection of the best possible Wiener weight vector using the proposed ACSWF approach results in a better denoising result with less image artifacts.

IV. Results and Discussion

A. Data Set Description and System Specifications

The tested data set includes satellite images procured from various sources, such as Satpalda Geospatial Services, Satellite Imaging Corporation, and NASA. All the test images included in the study are MS images with four bands (blue: 430–550 nm; green: 500–620 nm; red: 590–710 nm; near IR: 740–940 nm). The three images included in this paper are referred as Image 1: landslide in Zhouqu, China, 1263 × 1261, WorldView-2, MS-4 50 m Res. (http://earthobservatory.nasa.gov), Image 2: Madrid, Spain, 1000 × 1000, WorldView-3, MS-4 40 cm Res. (http://www.satimagingcorp.com), Image 3: real noisy image, Sentinel-2, 1034 × 1030, (computed noise variance 19.88%), (https://eros.usgs.gov/sentinel-2). Two more images with their simulation results are included in the Supplementary Material. Simulations were carried out using MATLAB R2015a software running on an Intel Core i7-3770

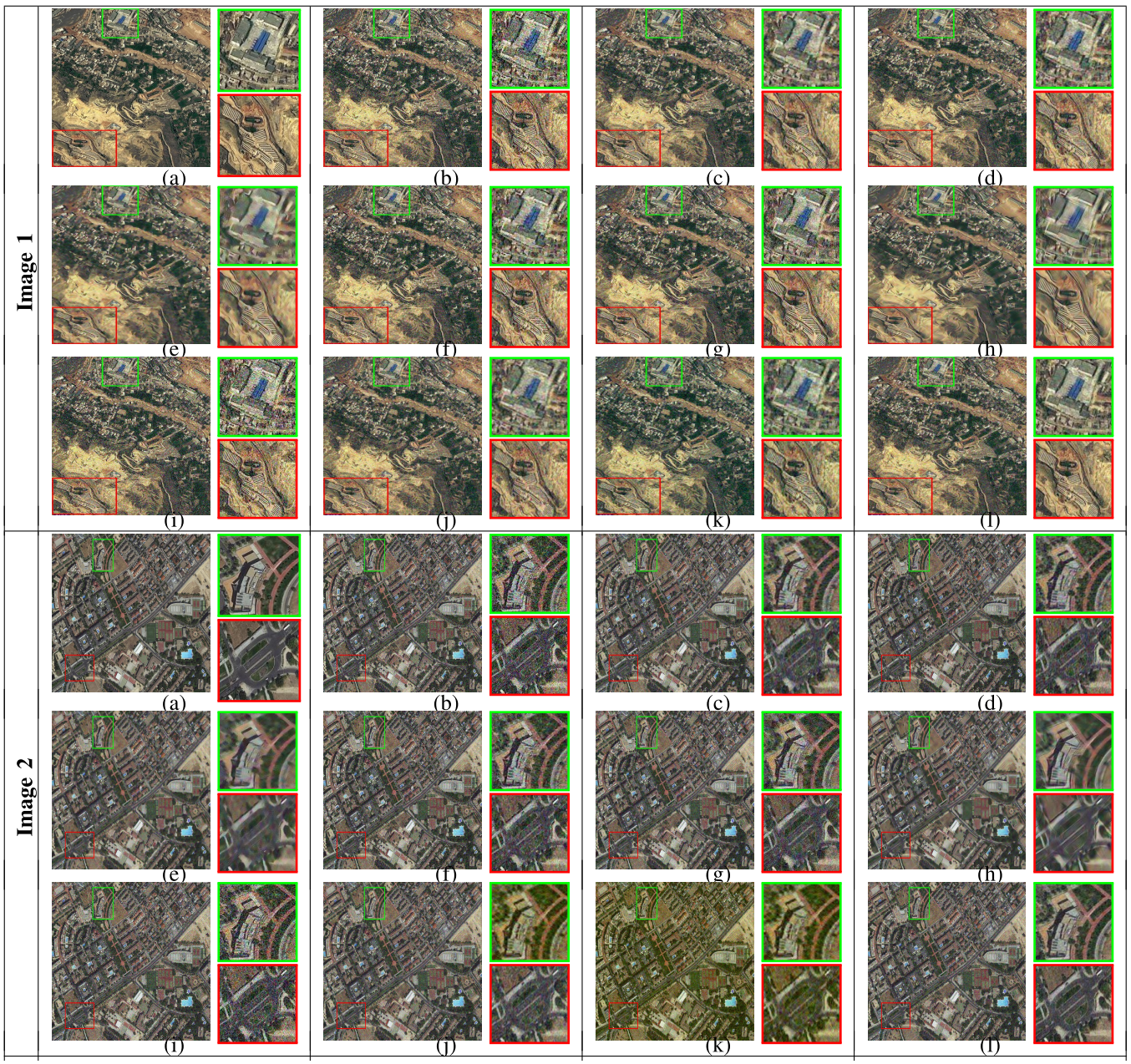

Fig. 3. Simulation results of denoising Images 1 and 2 corrupted with Gaussian noise of variance 30% using state-of-the-art denoising algorithms. (a) Original. (b) Noisy. (c) 2-D-NLMS. (d) 2-D-APA. (e) BM3-D-PCA. (f) PDE-AIP. (g) CII-NLM. (h) DST. (i) SSTV. (j) 2-D ABC adaptive filtering. (k) JADE. (l) Proposed ACSWF.

system with 3.40-GHz CPU, 8-GB RAM, and 64-bit operating system. The optimal value for the "switching parameter" (p_a) used in the proposed algorithm is fixed as 0.5, by conducting a prior empirical study [2]. This value ensures a proper balancing between the exploitation and exploration stages of the optimization algorithm.

B. Simulation Results

This section includes experimental results obtained using three high-spatial resolution MS satellite images of which one is a real noisy image obtained. Two more such MS images with their qualitative and quantitative results are provided in Fig. 3 and Tables II–IV of the Supplementary Material. The performance analysis of the proposed ACSWF is carried out by comparing it with basic 2-D adaptive filtering methods, such as 2-D-NLMS algorithm [19], [61] and 2-D-APA [19], [62]. We have also included most recent state-of-the-art denoising methods used with MS satellite images which employs the principal component analysis combined with block-matching 3-D (BM3-D-PCA) [63], [64], PDE-AIP [28], nonlocal cosine integral images (CII-NLM) [25], discrete shearlet transform (DST) [26], and spatiospectral total variation (SSTV) [30]. For ensuring a fair comparison, we have included comparison with recent denoising algorithms using metaheuristics, such as the JADE algorithm [20] and the 2-D artificial bee colony (ABC)

TABLE II

QUANTITATIVE COMPARISON BETWEEN DIFFERENT ALGORITHMS FOR DENOISING IMAGES 1 AND 2 CORRUPTED WITH GAUSSIAN NOISE OF VARIANCE 10%

Image	Met. \ Algo.	2D-NLMS	2D-APA	BM3D-PCA	PDE-AIP	CII-NLM	DST	SSTV	2D ABC	JADE	ACSWF
Image 1	MSE	138.1648	176.9961	280.6502	261.0002	271.2083	288.3612	168.913	176.0388	196.0059	**78.2289**
	PSNR	26.7268	25.6512	23.6492	23.9644	23.7978	23.5314	25.8542	25.6747	25.2081	**29.1971**
	FSIM	0.9691	0.9711	0.9889	0.9725	0.9719	0.9727	0.9668	0.9718	0.9693	**0.9779**
	UQI	0.9983	0.9980	0.9621	0.9977	0.9966	0.9974	0.9983	0.9979	0.9967	**0.9987**
	NAE	0.1024	0.1229	0.1458	0.1253	0.1381	0.1368	0.1143	0.1176	0.1250	**0.0547**
	Time (s)	9.4690	18.4520	10.1230	22.3920	9.6130	10.1130	14.5560	16.2520	12.5430	10.5230
Image 2	MSE	231.7467	230.6078	279.6011	239.5899	252.4129	280.3517	274.178	217.3529	237.0807	**191.3700**
	PSNR	24.4807	24.5021	23.6654	24.3361	24.1097	23.6538	23.7505	24.7591	24.3818	**25.3121**
	FSIM	0.9714	0.9765	0.9866	0.9741	0.9773	0.9763	0.9716	0.9770	0.9764	**0.9780**
	UQI	0.9392	0.9084	0.1999	0.9081	0.9391	0.9383	0.9301	0.9034	0.9188	**0.9516**
	NAE	0.1962	0.2016	0.1999	0.1927	0.1965	0.1968	0.2168	0.1959	0.1955	**0.1817**
	Time (s)	9.5540	20.5540	12.5520	23.3350	10.3310	11.4320	14.9960	17.8010	14.8930	9.8560

TABLE III

QUANTITATIVE COMPARISON BETWEEN DIFFERENT ALGORITHMS FOR DENOISING IMAGES 1 AND 2 CORRUPTED WITH GAUSSIAN NOISE OF VARIANCE 20%

Image	Met. \ Algo.	2D-NLMS	2D-APA	BM3D-PCA	PDE-AIP	CII-NLM	DST	SSTV	2D ABC	JADE	ACSWF
Image 1	MSE	251.6535	195.4102	394.3092	362.1446	480.2542	326.9010	217.7484	230.8015	244.4655	**120.7792**
	PSNR	24.1228	25.2213	22.1724	22.5420	21.3161	22.9866	24.7513	24.4984	24.2486	**27.3109**
	FSIM	0.9559	0.9598	0.9817	0.9671	0.9569	0.9594	0.9540	0.9577	0.9530	**0.9842**
	UQI	0.9966	0.9 969	0.9875	0.9954	0.9976	0.9954	0.9969	0.9967	0.9965	**0.9988**
	NAE	0.1422	0.1303	0.1785	0.1506	0.1056	0.1418	0.1306	0.1374	0.1391	**0.0701**
	Time (s)	10.9910	22.4340	12.4660	24.6340	10.7150	12.3390	16.7730	20.3330	16.5430	10.5240
Image 2	MSE	340.9504	287.5738	385.6581	341.2975	460.8813	302.2861	336.2252	308.0926	384.2303	**184.4980**
	PSNR	22.8039	23.5433	22.2688	22.7995	21.4949	23.3266	22.8645	23.2440	22.2849	**25.4709**
	FSIM	0.9547	0.9623	0.9790	0.9705	0.9652	0.9661	0.9584	0.9610	0.9574	**0.9817**
	UQI	0.8787	0.8864	0.9342	0.9167	0.9412	0.9268	0.8823	0.8947	0.8600	**0.9504**
	NAE	0.2479	0.2313	0.1986	0.1978	0.1995	0.2004	0.2464	0.2356	0.2569	**0.1155**
	Time (s)	11.4520	22.4860	14.5620	25.6890	11.3320	12.3160	15.3350	18.4230	16.5840	11.5236

TABLE IV

QUANTITATIVE COMPARISON BETWEEN DIFFERENT ALGORITHMS FOR DENOISING IMAGES 1 AND 2 CORRUPTED WITH GAUSSIAN NOISE OF VARIANCE 30%

Image	Met. \ Algo.	2D-NLMS	2D-APA	BM3D-PCA	PDE-AIP	CII-NLM	DST	SSTV	2D ABC	JADE	ACSWF
Image 1	MSE	368.1637	261.7553	360.9084	477.7469	690.3150	425.4046	255.4553	277.3758	450.2817	**126.0928**
	PSNR	22.4704	23.9518	22.5568	21.3388	19.7403	21.8428	24.0577	23.7001	21.5960	**27.1239**
	FSIM	0.9378	0.9454	0.9769	0.9692	0.9642	0.9606	0.9446	0.9445	0.9315	**0.9760**
	UQI	0.9857	0.9951	0.9867	0.9938	0.9967	0.9932	0.9953	0.9953	0.9553	**0.9986**
	NAE	0.1772	0.1521	0.1897	0.1748	0.1076	0.1587	0.1434	0.1523	0.1996	**0.0832**
	Time (s)	11.4560	26.1140	13.8950	28.3150	11.6630	13.5500	18.9910	22.2330	29.7720	14.0123
Image 2	MSE	376.0204	366.7047	343.8434	456.3029	670.3104	387.6814	385.2724	374.6892	647.4976	**207.6128**
	PSNR	22.3787	22.4876	22.7672	21.5383	19.868	22.2461	22.2731	22.3941	20.0184	**24.9583**
	FSIM	0.9461	0.9492	0.9739	0.9641	0.9689	0.9661	0.9478	0.9465	0.9380	**0.9776**
	UQI	0.8807	0.8693	0.9021	0.8955	0.9114	0.8952	0.8705	0.8792	0.7248	**0.9221**
	NAE	0.2650	0.2648	0.1988	0.1997	0.1993	0.2021	0.2692	0.2628	0.3403	**0.1392**
	Time (s)	13.4520	25.8420	17.4650	26.9910	12.3350	14.6030	16.7810	22.5480	28.5630	15.1268

adaptive filtering algorithm [19]. The aforementioned metaheuristic-based denoising algorithms and the proposed ACSWF were executed for 31 independent trials, and the best results among them are furnished.

Experiments were conducted for denoising images corrupted with three different Gaussian noise variance levels, i.e., 10%, 20%, and 30%, to substantiate the efficiency of the proposed algorithm. Since satellite images are most prone to the Gaussian noise of small variance levels, we have fixed the window size to be 3×3 for the entire set of experiments. Parameter initialization phase of the proposed algorithm includes manual assignment of parameter values for N_p, $w^{\min}$, $w^{\max}$, and $G_{\max}$. We have fixed their values to be 50, -1, 1, and 100, respectively. The dimensionality d of the optimization problem denotes the total number of elements included as filter coefficients for the defined filter. Since we have chosen the filter window size $n \times n$ to be 3×3, the value of $d = n^2$ is set as 9. All other parameters used in the algorithms compared are chosen from their respective references.

TABLE V

QUANTITATIVE COMPARISON BETWEEN DIFFERENT ALGORITHMS FOR DENOISING SENTINEL 2 (REAL NOISY) IMAGE CORRUPTED WITH COMPUTED NOISE VARIANCE 19.88%

Image	Algo. / Met.	2D-NLMS	2D-APA	BM3D-PCA	PDE-AIP	CII-NLM	DST	SSTV	2D ABC	JADE	ACSWF
Image 3	MSE	229.1087	242.7652	294.5684	239.5899	363.8362	238.3903	280.3517	297.6898	368.2286	**149.3287**
	PSNR	24.5304	24.2789	23.4389	24.3361	22.5217	24.3579	23.6538	23.3932	22.4696	**26.3894**
	FSIM	0.8872	0.8899	0.9107	0.9741	0.8995	0.9207	0.9763	0.8885	0.9008	**0.9809**
	UQI	0.9949	0.9951	0.9236	0.9081	0.9077	0.9972	0.9383	0.9949	0.9437	**0.9979**
	NAE	0.0998	0.1004	0.2294	0.1927	0.2003	0.0992	0.1968	0.1138	0.1392	**0.0687**
	Time (s)	11.4440	20.1244	18.4421	24.2454	11.5634	12.5453	17.5564	19.5635	16.8565	**12.9856**

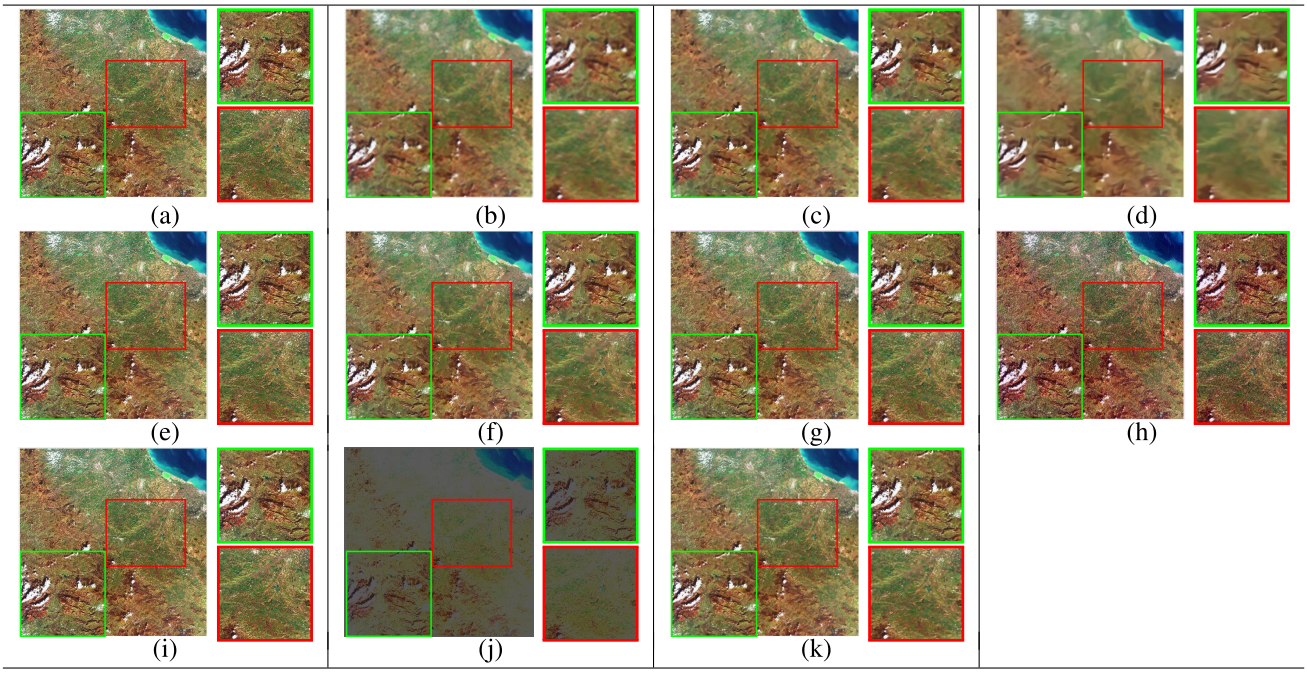

Fig. 4. Simulation results of denoising Image 3. Sentinel 2 (real noisy image, computed noise variance 19.88%) using state-of-the-art denoising algorithms. (a) Noisy. (b) 2-D-NLMS. (c) 2-D-APA. (d) BM3-D-PCA. (e) PDE-AIP. (f) CII-NLM. (g) DST. (h) SSTV. (i) 2-D ABC adaptive filtering. (j) JADE. (k) Proposed ACSWF.

Tables II–VI present the quantitative results obtained by comparing the 10 algorithms for denoising test images corrupted with the Gaussian noise of three different variance levels. The test images are processed band-by-band, assuming the noise to be uniform in every band and finally combined to form the denoised image. The quantitative metrics are computed by averaging the bandwise metric values and are reported in Tables II–VI. The proposed ACSWF yielded the least possible MSE value with a minimum NAE factor. These measures indicate the closeness of denoised images to the original uncorrupted image. A high value of PSNR, FSIM, and UQI for the proposed algorithm among the compared set highlights the visual quality of the estimated output image.

Subjective evaluation of denoised images obtained for 30% Gaussian noise variance level case using different algorithms compared is presented in Figs. 3 and 4. The proposed ACSWF proved to be very effective in preserving the information bearing structures like terrain edges and other significant textural features, especially for high noise variance levels. It also resulted in very less blurring effect to the estimated image compared with others for all the three noise variance levels investigated.

TABLE VI

STATISTICAL ANALYSIS COMPARING MSE VALUES OBTAINED USING JADE, 2-D ABC, AND ACSWF ALGORITHMS FOR DENOISING IMAGES 1 AND 2 CORRUPTED WITH GAUSSIAN NOISE VARIANCE 30% AND IMAGE 3 WITH COMPUTED NOISE VARIANCE 19.88%

Image	Algorithm	STD	Mean	Best MSE	Worst MSE
1	JADE	17.5790	473.5087	450.2817	502.2817
	2D ABC	8.7461	289.5990	277.3758	308.3758
	ACSWF	**4.5563**	**132.7582**	**126.0928**	**139.4540**
2	JADE	29.0564	627.3001	582.4976	682.4976
	2D ABC	12.7946	376.1842	354.6892	394.6892
	ACSWF	**3.7335**	**212.9852**	**207.6128**	**218.8975**
3	JADE	10.0663	383.8186	368.2286	398.8541
	2D ABC	5.0695	306.8256	297.6898	314.9296
	ACSWF	**3.3941**	**154.2131**	**149.3287**	**159.6937**

The stability and converging capability of the three metaheuristic-based denoising algorithms were analyzed by

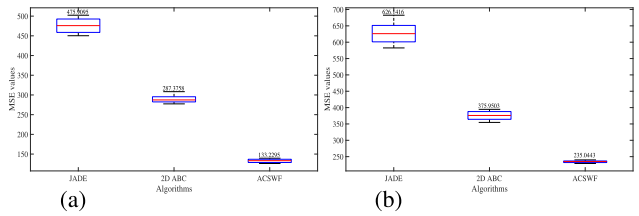

Fig. 5. Stability comparison between JADE, 2-D ABC, and the proposed ACSWF using box-and-whisker plots. (a) Image 1 and (b) Image 2 for 30% noise variance level.

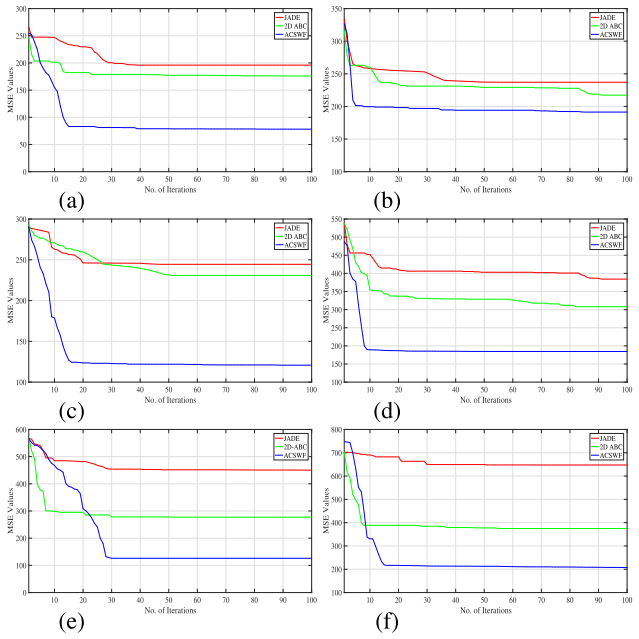

Fig. 6. Convergence rate analysis between JADE, 2-D ABC, and ACSWF algorithms. (a), (c), and (e) Convergence plots for Image 1 for noise variance level of 10%, 20%, and 30%, respectively. (b), (d), and (f) Convergence plots for Image 2 for noise variance level of 10%, 20%, and 30%, respectively.

investigating their box-and-whisker plots, statistical parameters, and convergence characteristics plots. Stability analysis was carried out by comparing the box-and-whisker plots of the three metaheuristic-based denoising algorithms. The random initialization phase followed by metaheuristic algorithms, such as JADE, 2-D ABC, and the proposed ACSWF used for image denoising, lead to slight discrepancies in the final optimal solutions.

Box-and-whisker plots give a graphical representation of the fitness value obtained in each repetition of the same experiment. Hence, it helps in assessing the repeatability of an algorithm over time, subject to the same experimental conditions. A highly stable algorithm gives more or less the same experimental results on each repetition and hence can be used for real-time applications as well. Figs. 5 and 7(a) present the box-and-whisker plots comparing the three metaheuristic algorithms included in our study, by repeating each experiment for 31 independent trials. The red line across each box indicates their respective median value and is showed on the top of each box. Comparison of the statistical parameters, such as standard deviation, mean, best value, and worst value of

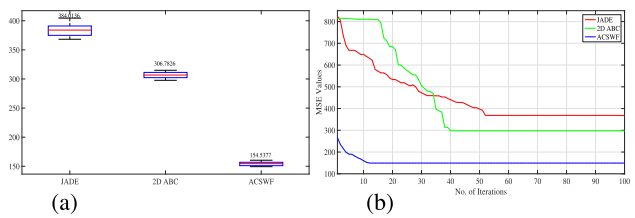

Fig. 7. (a) Stability and (b) convergence plot comparison between JADE, 2-D ABC, and the proposed ACSWF using box-and-whisker plots for Image 3.

the fitness function (MSE) obtained, after 31 trials of each experiment is presented in Table VI. The comparison also proves the stability and efficiency of the proposed ACSWF for denoising satellite images.

The test images were denoised using all the three algorithms by fixing the maximum number of iterations to be 100. The fitness values obtained in each run were plotted against the iteration count to analyze the convergence characteristics of each. The random population initialization phase and the solution exploration and exploitation strategies followed by different metaheuristic algorithms have a high impact on governing its convergence characteristics. Figs. 6 and 7(b) show the convergence plots, wherein the proposed ACSWF emerged to be converging fast to the least possible fitness (MSE) value. The reasonably fast convergence rate makes ACSWF adaptable for real-time applications too.

V. Conclusion

In this paper, a 2-D FIR Wiener filter based on the ACS algorithm (ACSWF) was proposed for denoising satellite images corrupted with AWGN. The ACS algorithm was proposed to optimize the adaptive Wiener filter weights to obtain the best possible estimate of the desired input image. Performance assessment included quantitative and qualitative comparisons with most studied and state-of-the-art denoising algorithms, such as 2-D NLMS, 2-D-APA, BM3-D-PCA, PDE-AIP, CII-NLM, DST, SSTV, JADE, and 2-D-ABC. Simulation experiments were conducted for denoising satellite images corrupted with three different Gaussian noise variance levels to substantiate the performance of the proposed filter. Robustness of the proposed filter was evaluated by testing it across a wide range of MS satellite image data set.

Comparisons between the evaluated performance metrics quantify the efficiency of the proposed filter in preserving significant image features with the LMS error. Subjective comparisons between the resultant images also highlight the filtering capability of the proposed ACSWF in denoising satellite images with reasonably fewer image artifacts. Stability and convergence capability analysis performed between JADE, 2-D-ABC, and ACSWF using box-and-whisker plots, statistical parameters, and convergence characteristics plots also highlights the performance of the proposed denoising filter. The ACSWF emerged with the least MSE value in a less number of iterations among the three. Substantial improvement in the quantitative and qualitative results along its stable nature makes it highly adaptable for other real-time image processing applications.

As a part of the future study, the performance of the proposed filter can be assessed for other 2-D signal processing applications, such as channel estimation and system identification.

ACKNOWLEDGMENT

The authors would like to thank the editors and anonymous reviewers for their valuable suggestions and comments which helped to improve the quality of this paper.

REFERENCES

[1] S. Suresh and S. Lal, "An efficient cuckoo search algorithm based multilevel thresholding for segmentation of satellite images using different objective functions," *Expert Syst. Appl.*, vol. 58, pp. 184–209, Oct. 2016.

[2] S. Suresh, S. Lal, C. S. Reddy, and M. S. Kiran, "A novel adaptive cuckoo search algorithm for contrast enhancement of satellite images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3665–3676, Aug. 2017.

[3] X.-Y. Wang, Y.-C. Liu, and H.-Y. Yang, "An efficient remote sensing image denoising method in extended discrete shearlet domain," *J. Math. Imag. Vis.*, vol. 49, no. 2, pp. 434–453, 2014.

[4] M. M. Hadhoud and D. W. Thomas, "The two-dimensional adaptive LMS (TDLMS) algorithm," *IEEE Trans. Circuits Syst.*, vol. CS-35, no. 5, pp. 485–494, May 1988.

[5] Z.-W. Li, X.-L. Ding, D.-W. Zheng, and C. Huang, "Least squares-based filter for remote sensing image noise reduction," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2044–2049, Jul. 2008.

[6] S. Haykin and B. Widrow, *Least-Mean-Square Adaptive Filters*. Hoboken, NJ, USA: Wiley, 2003, vol. 31.

[7] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated Gaussian data," *IEEE Trans. Inf. Theory*, vol. IT-34, no. 4, pp. 680–691, Jul. 1988.

[8] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 132–135, Feb. 2004.

[9] G.-O. Glentis, "An efficient affine projection algorithm for 2-D FIR adaptive filtering and linear prediction," *Signal Process.*, vol. 86, no. 1, pp. 98–116, 2006.

[10] H. Zhang, A. Nosratinia, and R. O. Wells, "Image denoising via wavelet-domain spatially adaptive FIR Wiener filtering," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 4. Jun. 2000, pp. 2179–2182.

[11] V. Soni, A. K. Bhandari, A. Kumar, and G. K. Singh, "Improved sub-band adaptive thresholding function for denoising of satellite image based on evolutionary algorithms," *IET Signal Process.*, vol. 7, no. 8, pp. 720–730, 2013.

[12] J. L. de Paiva, C. F. M. Toledo, and H. Pedrini, "An approach based on hybrid genetic algorithm applied to image denoising problem," *Appl. Soft Comput.*, vol. 46, pp. 778–791, Sep. 2016.

[13] M. Malik, F. Ahsan, and S. Mohsin, "Adaptive image denoising using cuckoo algorithm," *Soft Comput.*, vol. 20, no. 3, pp. 925–938, 2016.

[14] S.-T. Tzeng, "Design of 2-D FIR digital filters with specified magnitude and group delay responses by GA approach," *Signal Process.*, vol. 87, no. 9, pp. 2036–2044, 2007.

[15] K. Boudjelaba, D. Chikouche, and F. Ros, "Evolutionary techniques for the synthesis of 2-D FIR filters," in *Proc. Statist. Signal Process. Workshop (SSP)*, Jun. 2011, pp. 601–604.

[16] S. K. Sarangi, R. Panda, and M. Dash, "Design of 1-D and 2-D recursive filters using crossover bacterial foraging and cuckoo search techniques," *Eng. Appl. Artif. Intell.*, vol. 34, pp. 109–121, Sep. 2014.

[17] F. LatifoğLu, "A novel approach to speckle noise filtering based on Artificial Bee Colony algorithm: An ultrasound image application," *Comput. Methods Programs Biomed.*, vol. 111, no. 3, pp. 561–569, 2013.

[18] S. Kockanat, N. Karaboga, and T. Koza, "Image denoising with 2-D FIR filter by using artificial bee colony algorithm," in *Proc. Int. Symp. Innov. Intell. Syst. Appl. (INISTA)*, Jul. 2012, pp. 1–4.

[19] S. Kockanat and N. Karaboga, "A novel 2D-ABC adaptive filter algorithm: A comparative study," *Dig. Signal Process.*, vol. 40, pp. 140–153, May 2015.

[20] A. K. Bhandari, D. Kumar, A. Kumar, and G. K. Singh, "Optimal sub-band adaptive thresholding based edge preserved satellite image denoising using adaptive differential evolution algorithm," *Neurocomputing*, vol. 174, pp. 698–721, Jan. 2016.

[21] H. Rakhshani and A. Rahati, "Snap-drift cuckoo search: A novel cuckoo search optimization algorithm," *Appl. Soft Comput.*, vol. 52, pp. 771–794, Apr. 2017.

[22] X. Li and M. Yin, "Modified cuckoo search algorithm with self adaptive parameter method," *Inf. Sci.*, vol. 298, pp. 80–97, Mar. 2015.

[23] P. Scheunders and S. D. Backer, "Wavelet denoising of multicomponent images using Gaussian scale mixture models and a noise-free image as priors," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1865–1872, Jul. 2007.

[24] C. Chaux, L. Duval, A. Benazza-Benyahia, and J. C. Pesquet, "A nonlinear stein-based estimator for multichannel image denoising," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3855–3870, Aug. 2008.

[25] B. Xue, Y. Huang, J. Yang, L. Shi, Y. Zhan, and X. Cao, "Fast nonlocal remote sensing image denoising using cosine integral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1309–1313, Nov. 2013.

[26] T. S. Anju and N. R. N. Raj, "Satellite image denoising using shearlet transform," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2016, pp. 0571–0575.

[27] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2272–2279.

[28] P. Liu, F. Huang, G. Li, and Z. Liu, "Remote-sensing image denoising using partial differential equations and auxiliary images as priors," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 358–362, May 2012.

[29] H. Yue, X. Sun, J. Yang, and F. Wu, "Image denoising by exploring external and internal correlations," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1967–1982, Jun. 2015.

[30] H. K. Aggarwal and A. Majumdar, "Hyperspectral image denoising using spatio-spectral total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 442–446, Mar. 2016.

[31] V. B. S. Prasath and A. Singh, "Multispectral image denoising by well-posed anisotropic diffusion scheme with channel coupling," *Int. J. Remote Sens.*, vol. 31, no. 8, pp. 2091–2099, 2010.

[32] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2949–2956.

[33] Q. Xie *et al.*, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1692–1700.

[34] H. Peng, R. Rao, and S. A. Dianat, "Multispectral image denoising with optimized vector bilateral filter," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 264–273, Jan. 2014.

[35] G. Papari, N. Idowu, and T. Varslot, "Fast bilateral filtering for denoising large 3D images," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 251–261, Jan. 2017.

[36] A. Duijster, P. Scheunders, and S. De Backer, "Wavelet-based EM algorithm for multispectral-image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3892–3898, Nov. 2009.

[37] G. Chen and S.-E. Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 973–980, Mar. 2011.

[38] X. Liu, H. Shen, Q. Yuan, X. Lu, and C. Zhou, "A universal destriping framework combining 1-D and 2-D variational optimization methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 808–822, Feb. 2018.

[39] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.

[40] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4009–4018, Jul. 2013.

[41] X. Liu, X. Lu, H. Shen, Q. Yuan, Y. Jiao, and L. Zhang, "Noise removal from hyperspectral image with joint spectral–spatial distributed sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5425–5439, Sep. 2016.

[42] Y. Yuan, X. Zheng, and X. Lu, "Spectral–spatial kernel regularized for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3815–3832, Jul. 2015.

[43] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Hoboken, NJ, USA: Pearson, Oct. 2008.

[44] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 165–168, Feb. 1980.

[45] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, no. 2, pp. 165–177, Feb. 1985.
[46] S. G. Chang, B. Yu, and M. Vetterli, "Image denoising via lossy compression and wavelet thresholding," in *Proc. Int. Conf. Image Process.*, Oct. 1997, pp. 604–607.
[47] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1522–1531, Sep. 2000.
[48] X.-S. Yang and S. Deb, "Cuckoo search via Lévy flights," in *Proc. World Congr. Nature Biol. Inspired Comput. (NaBIC)*, Dec. 2009, pp. 210–214.
[49] P. Ong, "Adaptive cuckoo search algorithm for unconstrained optimization," *Sci. World J.*, vol. 2014, Sep. 2014, Art. no. 943403, doi: http://dx.doi.org/10.1155/2014/943403
[50] Z. Zhang and Y. Chen, "An improved cuckoo search algorithm with adaptive method," in *Proc. 7th Int. Joint Conf. Comput. Sci. Optim. (CSO)*, Jul. 2014, pp. 204–207.
[51] M. Naik, M. R. Nath, A. Wunnava, S. Sahany, and R. Panda, "A new adaptive cuckoo search algorithm," in *Proc. 2nd Int. Conf. Recent Trends Inf. Syst. (ReTIS)*, Jul. 2015, pp. 1–5.
[52] M. K. Naik and R. Panda, "A novel adaptive cuckoo search algorithm for intrinsic discriminant analysis based face recognition," *Appl. Soft Comput.*, vol. 38, pp. 661–675, Jan. 2016.
[53] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, 1997.
[54] X.-S. Yang, *Nature-Inspired Optimization Algorithms*. Amsterdam, The Netherlands: Elsevier, Feb. 2014.
[55] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. 20th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2010, pp. 2366–2369.
[56] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electron. Lett.*, vol. 44, no. 13, pp. 800–801, Jun. 2008.
[57] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
[58] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
[59] C. P. Loizou, C. S. Pattichis, M. Pantziaris, T. Tyllis, and A. Nicolaides, "Quality evaluation of ultrasound imaging in the carotid artery based on normalization and speckle reduction filtering," *Med. Biol. Eng. Comput.*, vol. 44, p. 414, May 2006.
[60] M. K. Ozkan, M. I. Sezan, and A. M. Tekalp, "Adaptive motion-compensated filtering of noisy image sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 4, pp. 277–290, Aug. 1993.
[61] M. S. E. Abadi and S. N. Aali, "The novel two-dimensional adaptive filter algorithms with the performance analysis," *Signal Process.*, vol. 103, pp. 348–366, Oct. 2014.
[62] M. Muneyasu, T. Hinamoto, and H. Yagi, "A realization of 2-D adaptive filters using affine projection algorithm," *J. Franklin Inst.*, vol. 335, no. 7, pp. 1185–1193, 1998.
[63] 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.
[64] G. Chen, S.-E. Qian, and S. Gleason, "Denoising of hyperspectral imagery by combining PCA with block-matching 3-D filtering," *Can. J. Remote Sens.*, vol. 37, no. 6, pp. 590–595, 2012.

Shilpa Suresh (S'17) received the M.E. degree in communication systems from Hindustan University, Chennai, India, in 2012. She is currently pursuing the Ph.D. degree with the National Institute of Technology Karnataka, Mangalore, India.

Her research interests include satellite image processing using stochastic metaheuristic algorithms.

Shyam Lal (M'11) received the M.Tech. degree from the National Institute of Technology, Kurukshetra, India, in 2007, and the Ph.D. degree from the Birla Institute of Technology, Mesra, Ranchi, India, in 2013.

He is currently an Assistant Professor with the Department of Electronics and Communication Engineering, National Institute of Technology Karnataka, Mangalore, India. He has authored or co-authored over 50 papers in refereed journals and conferences in his research areas. His current research interests include digital image processing, remote sensing, and medical image processing.

Chen Chen (M'17) received the B.E. degree in automation from Beijing Forestry University, Beijing, China, in 2009, the M.S. degree in electrical engineering from Mississippi State University, Starkville, MS, USA, in 2012, and the Ph.D. degree from the Department of Electrical Engineering, The University of Texas at Dallas, Richardson, TX, USA, in 2016.

He currently holds a post-doctoral position with the Center for Research in Computer Vision, University of Central Florida, Orlando, FL, USA. He has authored or co-authored over 50 papers in refereed journals and conferences in his research areas. His research interests include signal and image processing, computer vision, and deep learning.

Turgay Celik (M'15) received the second Ph.D. degree from the University of Warwick, Coventry, U.K., in 2011.

He is currently a Professor of computer science and the Director of the National e-Science Postgraduate Teaching and Learning Platform at the University of Witwatersrand, Johannesburg, South Africa. His research interests include visual and cognitive computing, machine learning, data science, artificial intelligence, cloud computing, cybersecurity, and remote sensing.

Dr. Celik is an Associate Editor of *IET Electronics Letters*, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and *Signal, Image and Video Processing* (Springer).