Information Fusion 55 (2020) 220–244

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full Length Article

Image denoising review: From classical to state-of-the-art approaches

Bhawna Goyal[a,*], Ayush Dogra[b], Sunil Agrawal[a], B.S. Sohi[c], Apoorav Sharma[a]

[a] UIET, Panjab University, Chandigarh, India
[b] CBME, IIT Ropar, Punjab, India
[c] Chandigarh University, Gharuan, Punjab, India

ARTICLE INFO

Keywords:
Denoising
Spatial
Transform
Hybrid
Filters
PSNR

ABSTRACT

At the crossing of the statistical and functional analysis, there exists a relentless quest for an efficient image denoising algorithm. In terms of greyscale imaging, a plethora of denoising algorithms have been documented in the literature, in spite of which the level of functionality of these algorithms still holds margin to acquire desired level of applicability. Quite often noise affecting the pixels in image is Gaussian in nature and uniformly deters information pixels in image. Based on some specific set of assumptions all methods work optimally, however they tend to create artefacts and remove fine structural details under general conditions. This article focuses on classifying and comparing some of the significant works in the field of denoising.

1. Introduction

Due to inherent physical limitations of various recording devices, images become prone to manifestation of some random noise during image acquisition. Noise can be understood as a basic signal distortion which hinders the process of image observation and information extraction. Suppression of image noise forms elementary basis of image analysis and processing, hence any progress made in image denoising domain helps strengthening our understanding of basic image statistics and processing [1].

With massive increase in generation of digital image often captured in poor atmospheric/illumination conditions, image restoration methods have become indispensible tools in present era of computer aided analysis. From numerous types of noise prevailing in different kinds of images Additive White Gaussian Noise (AWGN), impulse noise (salt and pepper), quantisation noise, Poisson noise and speckle noise are most frequently discussed noises in literature [2]. AWGN primarily occurs in analog circuitry during image acquisition and transmission. The prevalence of other types of noises such as quantisation noise, impulse noise, speckle noise and Poisson noise occurs mainly due to faulty manufacturing, bit error and inadequate photon count during image acquisition.

There are a variety of digital images available which provide valuable information in diverse fields of application like medical imaging, remote sensing, military and surveillance, robotics and artificial intelligence. The contamination of these images irrevocably destroys image interpretability. Therefore image denoising methods find widespread use in the field of medical imaging, remote sensing, military and surveillance, biometrics and forensics, industrial and agricultural automation.

1.1. Noise suppression

An image in its digital form can be assumed to be an encoded form of a matrix with grey-level or colour pixel intensity values as its elements. Of course, this matrix has the third dimension of time in case of a video. In case of grey scale images it can be referred to as a signal in two dimensions i.e. *(x, u(x))*, where *u(x)* is the pixel intensity value at location *x*. In colour image, value $u(x)$ assumes a triplet of values for each of red, green and blue components. In order to maintain simplicity this review shall remain confine to grey-scale images except for brief mention of denoising in colour domain [1, 8]. The methodology opted for basic elementary grey-scale image denoising can be explicitly interpolated to colour scale images and 3-dimensional videos. In an orderly fashion an image noise model can roughly be approximated as:

$$f(x) = u(x) + n(x)\ x \subset X,\ \ X \subset Z^2$$

where *u(x)* denotes true signal value and *n(x)* denotes noise at location *x*. Noise can be additive or multiplicative in nature. Most often it is found to independent and identically distributed with zero mean and standard deviation σ. In CCD (charged coupled devices) cameras noise may be present in electron circuitry (thermal noise), due to inadequate photon count (photon noise) or it may be quantisation noise. However AWGN is mostly commonly encountered noise in real-time applications, hence the AWGN model is the primary focus of this article.

An image is an extremely interesting form of data which comprises of several entities, for instance smooth homogenous regions, edges, contours and spikes all resulting from gradually varying grey level values of pixel intensity. The low intensity pixel values are important to

* Corresponding author.
E-mail address: bhawnagoyal28@gmail.com (B. Goyal).

https://doi.org/10.1016/j.inffus.2019.09.003
Received 5 March 2019; Received in revised form 11 August 2019; Accepted 6 September 2019
Available online 6 September 2019

determine the overall shape and formation of objects in an image, whereas high intensity pixel values or high level variation in pixel intensity values of image forms edges and features. The fine feature details and edges are significant for the higher level information analysis, comprehension and extraction. The random manifestation of noise irrevocably corrupts this valuable pixel information [9,10].

In last two decades, the researchers continuously thrive to develop efficient denoising algorithms which aim at restoring a reasonable estimate from distorted image while preserving fine features and edges. Due to rising level of technical sophistication in image acquisition devices, digital sensors tend to increase pixel resolution; however, with similar aperture sizes these sensors become more susceptible to noise making the role image denoising algorithms more important. Software-oriented approaches are mostly device independent hence widely employed as compared to hardware and optical systems [3].

While exploring and exploiting the digital signal processing theory and tools, researchers have experimented with various diverse tools of mathematical sciences, in order to fabricate an algorithm to remove noise while maintaining as many feature details as possible. There exists a number of denoising methods originating from various disciplines like spectral and multi-resolution analysis, partial differential equations, probability theory and statistics etc. On the basis of the type of denoising algorithms, they can broadly be referred to as: spatial domain filtering, transform domain thresholding, random fields, statistical models, anisotropic diffusion methods, dictionary learning methods and hybrid methods [4]. Besides these other major denoising measures include spatial adaptive filters, statistical estimators of all sorts, stochastic analysis, morphological analysis and order statistics [5].

The primitive and well-established spatial domain methods include local and non-local filters which exploit the similarities between pixels or patches of an image. Both dictionary learning based and transform domain methods consider transforming images into frequency domain, in which similarities and dissimilarities of transformed coefficients are exploited. The difference between them is that transform domain approaches apply fixed basis functions for the representation of images, whereas learning-based methods use sparse representations on a redundant dictionary [6,7]. In this article, an attempt has been made to abridge the hierarchical layout of the development and establishment of various denoising schemes in case of removal of Additive White Gaussian Noise from basic averaging of the pixels in the neighbourhood to more sophisticated hybrid techniques. The broad categorisation of various image denoising methods is given in the Fig. 1.

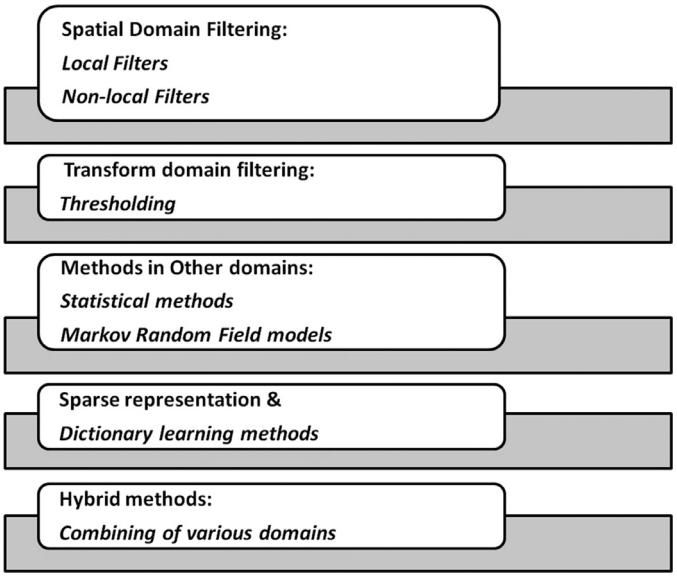

Fig. 1. Categorization of image denoising methods.

This article is structured as following:

- Section 2 covers the various spatial domain filtering methods.
- The transform domain methods and various thresholding schemes are discussed in Section 3.
- The methods in other domains including statistical and probability distribution based methods are discussed in Section 4.
- Section 5 discusses several hybrid domain methods.
- The sparse representation and dictionary learning methods are discussed in Section 6.
- Sections 7 and 8 constitutes the recent trends and applications for image denoising respectively.
- Section 9 entails the experimental setup for various state-of-the-art methods evaluated.
- The article is concluded in Section 10.

2. Spatial domain filtering

Ideally, one thinks as filtering to be the foremost solution to image noise removal i.e. suppression of the unwanted variation in pixel intensity values. Indeed, filtering of signals is fundamental to basic image processing and has long been in application for smoothing, sharpening, edge detection and contrast enhancement. In an algorithm in spatial domain, to denoise a single patch the correlation amongst the pixel candidates are exploited. In spatial domain, the set of operations is directly employed on image matrix, whereas in case of transform domain filtering, image matrix is first mapped into corresponding transformed coefficients and then further thresholding is carried out.

According to manner of selection of pixel candidates used in filtering process, a filter can be categorised as a local filter and a non-local filter. The underlying principle of an image denoising algorithm is that noise is uncorrelated amongst pixels and true signal pixel intensities are correlated to each other [1,8,11].

2.1. Local filters

A filter is considered to be local if the filter support for denoising a pixel is restricted by spatial distance. A non-local filter exploits the correlation amongst the entire range of pixels in an image. The popular local filters designed for noise reduction are Gaussian filter, Least Mean Square filters, Bilateral filter, Weiner filter, SUSAN filter, Anisotropic Diffusion filter, Rank filter, Steering Kernel Regression (SKR), Metric Steering Kernel Regression and Trained filters [6].

The simplest type of a Linear Translational Invariant filtering is averaging or mean or box filtering which generates output at each pixel as the average of neighbouring pixels in a given window. These linear filters tend to convolve image matrix with filter mask to generate a linear expansion of neighbourhood values. This type of spatial filtering is the most primitive and easiest form of noise removal, however they often generate undesirable amount of smoothing of edges, poor feature localisation and loss of details [1,6,8]. Fig. 2 depicts the layout of basic averaging filter.

Another types of linear filtering methods are Gaussian Filter (GF) [12] and Weiner filter [13,14] which do not employ mean of neighbouring pixels. Wiener filters are a class of optimum linear filters which involve linear estimation of a desired signal sequence from another related sequence. GF are typical linear filters which fall under the category of local filters and are isotropic in nature and have long being applied in the image denoising. Gaussian filtering is of particular significance in literature as the Fourier transform of Gaussian functions are real and their shapes are easily specified. GF invariably assumes that pixels in a neighbourhood are homogenous and have smooth spatial variations. However, in GF filter weights attenuates with increasing distance from centre and results in blurring of edges. Most of the local denoising filters have primarily been the improvisations of Gaussian filtering which have been proposed since then to provide better edge

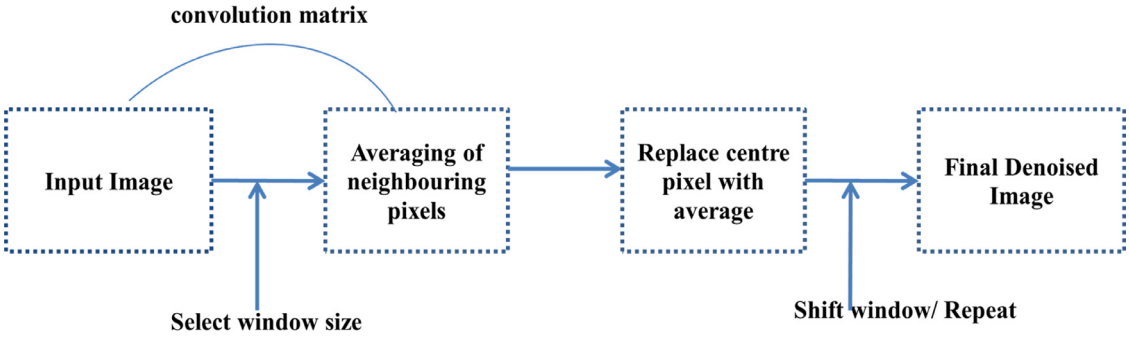

Fig. 2. Basic methodology of averaging filter.

preserving ability [6]. These LTI filters though form an easy and time efficient solutions, tend to remove important structural information and edge details which are vital to image analysis. Henceforth, various other non-linear filters have been proposed in order to address their limitations in a manner of their improvisations and innovative ideas. The age old median filtering [8,15], weighted median filter [16] and Rank filter [17] are basic and most trivial examples of non-linear filters which were formally developed and documented in literature. In median filtering, the neighbouring pixels are ranked according to brightness (intensity) and the median value becomes the new value for the central pixel.

With an aim to preserve edges, image details, image geometries and to overcome Gaussian blurring, the anisotropic diffusion filtering (ADF) was proposed, the idea of which goes back to Perona and Malik (PM) [18]. The ADF interpolated Gaussian filtering, by convolving the image u at x only in the direction which is orthogonal to its gradient $Du(x)$. In a simpler manner, ADF is a process of solving the anisotropic heat equation which is a second order partial differential equation. It is a non-linear diffusion process, which was proposed in order to overcome blurring and poor localisation issues posed by linear diffusion filtering [19,20].

ADF is an edge preserving image denoising algorithm which applies an inhomogeneous process to lessen the diffusivity at locations of edges and carries out a diffusion process in homogenous regions or regions with slight intensity variations. The modified heat equation (AD) given by Perona and Malik takes the following form [21]:

$$\frac{\partial u}{\partial s} = div(h|\nabla u(x,y,s)|)(\nabla u(x,y,s))$$

where $u_s = u(x, y, s)$, is the image obtained at an instance s, div is the divergence operator, ∇ is the gradient operator, and $h(.)$ is the diffusion coefficient or the edge-stopping function. The diffusion coefficient $h(\nabla u)$ is a local image gradient.

The magnitude of amount of diffusion is inversely proportional to gradient values or edges. Within homogenous or uniform regions, gradient value is small, hence $h(\nabla u)$ approaches 1, hence removes or smoothen away the uniform regions whereas, in case of near edges or boundaries, diffusion coefficient approaches zero and hence preserves the edges [21]. The diffusion coefficient can take either of the two following forms:

$$h(\nabla I) = e^{-\left(\frac{\nabla I}{r}\right)^2}$$

$$h(\nabla I) = \frac{1}{1+\left(\frac{\nabla I}{r}\right)^2}$$

$h(.)$ is positive and non-increasing function, r is the contrast parameter or gradient modulus threshold which controls the amount of denoising. The diffusion coefficient acts as a typical edge stopping function, which assumes zero value around edges or gradients.

It is a typical non-linear feature-preserving denoising algorithm, which is a standalone technique to address the issue of imaging artefacts. According to Dogra et al. [22], prevalence of artefacts which can disguise image interpretation especially in case of medical diagnosis can be effectively removed with anisotropic diffusion filtering.

Since its introduction by Perona and Malik, an extensive amount of literature has existed presenting a number of PDE-based anisotropic models which offer diverse modifications in order to attain a steady state solution and to address the issue of staircase effects [23–30]. The staircase effect presents a visually unpleasant artefactual display in images and often perplex as false edge. For instance, Qui et al. proposed a robust ridge detector, non-local derivative based feature preserving diffusion filtering process [31]. Besides this, in [32] a Gaussian filtering technique based on pre-denoising beforehand each iteration is proposed by Catte et al. as the modified PM model. The higher order partial differential equations have also been explored by scientists and have proved to be efficient in minimizing the staircase effects. Some of the work have also been carried out using fourth order partial differential equations (FPDE) and have produced satisfactory denoising results [33–35]. Lu and Tan in 2013 gave a fourth-order partial differential equation denoising model adopting a variety of conduction coefficients which results in better preservation of edges and features as compared to conventional second-order PDE based models [36]. Another extremely pioneering work based on FDPE was given by You–Kaveh employing a piecewise harmonic function to approximate noise-free image from noisy image [37]. Though a lot of research has been carried out based on FPDE, they often tend to generate speckle noise in filtered images and over-smooth the step edges as FPDE based models decay high frequency coefficients of the image much faster than PDE based models.

A PDE algorithm based on combination of local gradient, grey level variance and edge stopping function was given by Chao and Tsai in [38] to derive better edge preserving capability. Gilboa et al. [39] introduced an adaptive diffusion process which employed forward and backward diffusion process in order to introduce the aspect of better edge enhancement while suppressing noisy pixels in smoother sections. It gave an interesting review of the hierarchy of ADE based methodology along with its extended application in the field of image denoising and also demonstrated regularised ADE model which was well posed and has a steady-state non-trivial solution.

In [40] an adaptive thresholding based ADE model has been proposed. A modified version of gradient vector flow was given by Ghita et al. [41] to further improve denoising performance. Another domain of work, based on ADE filtering has gained a lot of importance which employs directional Laplacian and causes edge-directional diffusivity in original image [4,42].

An extensive amount of work has been documented in literature which derives its inspiration from anisotropic diffusion filtering-Perona Malik models in one way or other, few of which have been briefed out in this section. They are non-linear set of filtering techniques which design set levels of directions and gradients resulting in better edge preserving

ability as compared to other regularisation based filters. Despite the fact, that a lot of improvisations of ADE-PM model have established their applicability they still tend to sharpen boundaries with loss of textural details [4].

In 1992, Rudin and Osher proposed the total variation minimisation method which is another regularisation filtering algorithm to filter out noise from homogenous regions but not its edges. For a given noisy image $v(x)$, authors proposed to recover original image $u(x)$ as solution of a constrained minimisation idea such that, total variation of image is minimised, in turn generating a constant piecewise solution [43]. The restored image is simply obtained by minimizing the energy variation between the original and restored images. Some algorithms namely Chambolle's algorithm [44] and forward-backward algorithm [45] have been given as solutions to Total Variation (TV) based minimisation problems. However this technique was able to preserve information along straight edges but presented two major drawbacks: textural information was not preserved and it introduced staircase effect.

Most of the spatial domain filters are derived on basis of spatial proximity, whereas the neighbourhood filters considers grey level similarities to define a neighbourhood patch. The famous Yaroslavsky filter considers both spatial distance and grey-level similarities for averaging [46,47]. The Bilateral Filter (BF) was given by Tomasi and Munich as a modified form of neighbourhood filter, which weighs the distance to reference pixel instead of following a fixed neighbourhood [48].

Similarly, another spatial filtering approach based on neighbourhood filter, for providing increased edge retention is known as Smallest Univalue Segment Assimilating Nucleus (SUSAN) [49]. It constructs the average of all pixels which are equidistant from central pixel.

BF considers both grey level similarity and spatial closeness in the denoising process. It filters out noise in the regions belonging to reference pixels and preserves boundaries and sharp edges of the image. It has been established in [4,50], that BF is a form of discretisation of specific type of PDE-AD model. The pixels which are spatially close and have similar pixel intensities are averaged in a manner with weights decaying with the increasing distance from the reference pixel. Recently a lot of work is concentrated on providing improvisations of BF such as Weighted Bilateral filter, Robust Bilateral filter, Fast Bilateral filter, Multi-Resolutional bilateral filter [50].

It has been stated by Elad in [51] that bilateral filtering is similar to a Jacobi iteration of a weighted least squares minimisation. The process of bilateral filtering was simple and non-iterative, but its direct implementation was slow as BF has brute force implementation where computational complexity increases exponentially with increase in neighbourhood radius. In 2011, Chaudhary et al. gave a faster approach for BF by employing trigonometric range kernels instead of considering Euclidean spatial distance [52]. In 2013, another approach for an accelerated BF was given in [53]. Chaudhary et al. gave a sure and fast approach for image denoising using bilateral filtering in which Weighted Bilateral Filtering (WBF) is obtained by combining the original bilateral filter and its modification in a weighted fashion with an aim to minimise the MSE (Mean Square Error) [54]. Besides these several other accelerated forms of bilateral filtering have been given in [55-57].

According to Frabman and Durand, BF tends to create gradient reversal artefacts in high dynamic range detail decomposition and compression of images [55,58]. Also it can be clearly seen in [59] that BF does not work well at low noise levels. Though bilateral filtering is able to preserve edges, but they tend to retain some of the unwanted noisy coefficients as well. Also when standard deviation of noise exceeds edge contrast this method gives a poorly restored image [48].

In [60], Jin et al. introduced an algorithm which employs a weighting function to BF, where either centre pixel or vector median is chosen to be the part of filtering process. An empirical study based model for bilateral filter was given as multi-resolution bilateral filter in [61,62]. In a work given by Peng et al., an optimised approach for selection of vector parameters was given based on minimisation of Stein's unbiased risk estimate [63]. The popular joint bilateral filtering exploits global information from two correlated images [64]. Another filter, which was based on ideal interpretation of the bilateral filter, was guided filter [65]. This filter implementation was faster than conventional bilateral filter and derives its edge stopping function from a guidance image. This filter had an automatic time implementation, independent of window radius k and it freely chose arbitrary sizes of kernels in real time applications. It was also able to mitigate the issue of gradient reversal artefacts to some extent.

Besides these filters, various other filters namely steering kernel regression (SKR), Metric-SKR, KSPR and Trained filters will be briefed in the later section of this article.

This preceding section gives a brief insight into local spatial domain filters and their gradual progression in the field of denoising has been traced out. Since basic idea for implementation of local filters is exploitation of correlation amongst pixels, they tend to perform poorly at high noise levels, as the correlation amongst pixels is severely corrupted by elevated noise levels.

2.2. Non-local filters

In [2], Buades et al. besides classifying the then existing image denoising methods, based on experimental methodology introduced a filter, in order to exploit similarity amongst pixels in a non-local manner or in entire image, popularly known as Non-local mean (NLM) filter. This pioneering work clearly established that the self-similarity amongst characteristics of an image in a non-local manner is the biggest potential basis in the field of image denoising. NLM filter exploits presence of similar features or patterns in the image. In BF a weighted average of pixels in a given patch is generated and weights depend not only geometrical distances but also on contextual similarity. The method of non-local filtering can alternatively be understood as exploiting Euclidean distance between two patches around reference and the selected pixel. NLM algorithm is typical point-wise denoising approach, which generates a noise-free pixel for a single point only and similarity between the reference pixel neighbourhoods directly affects filter weights.

After preposition of original NLM filter in [2], several improvisations have been developed afterwards, in order to accelerate filter implementation and to improvise the qualitative and quantitative results. Zhang et al. introduced a two-directional non-local variation model for image denoising which exploits an idea that if similar patches are arranged as matrices then rows and columns exhibit similarities [66]. In order to accelerate NLM, refined set of candidate patches was generated by determining mean and gradient for pre-selection of the contributing neighbourhoods [67]. In another work by Wang et al. in 2012, Gabor feature non-local algorithm was proposed in order to denoise textural images [68]. An improved NLM (INLM) has been given in [69], which considerably improves efficacy of original NLM by assuming symmetry in the weighting function and by computing the Euclidean distance via symmetrical movement of recursive average filter (bi-square robust function). The execution of the method in an iterative manner results in better grouping of patches, based on pre-stage processing. This modified bi-square based robust weight calculation results in reduced loss of details. Though this filter removes extra noise, it tends to blur details due to post-processing.

Inspired by some modified versions of NLM, Xu et al. proposed a patch-grouping based NLM (NLMPG) algorithm for denoising of remote sensing images [70]. This algorithm mainly exploits Block-wise NLM algorithm given by Coupe et al. [71] by exhibiting improvisation in two aspects: it cuts out the redundancy by selecting most similar patches and customises filtering constant for centre patches in accordance with the variance of image patches. This algorithm showed elevated performance than some existing NLM improvisations and K-SVD (k means singular value decomposition).

Various other improvisations of NLM have been given in [72,73]. Though NLM is the first filter which exploits self-similarity in the entire range of pixels, it introduces artefacts and performance degradation

when the image lacks in similar patches. The conceptual exploitation of non-local image features to preserve edges and details while removing noise have remained the potential idea for most of the state-of-the-art image denoising algorithms which will be discussed in the later section of this article.

Besides this, there are many other mathematical concepts which have been widely employed in literature in context of spatial domain image denoising. Morphological filtering is one such topic. Morphological operations deal with tools to extract image features that are useful in description and representation of a shape. The morphological filtering comprises of opening and closing operations for smoothing of images and removal of noise. The opening operation smoothes contours of an object, breaks narrow dips and suppresses bright peaks, whereas closing generally fuses the narrow dips, eliminates small holes and diminishes darker details [74].

However, a very recent filter based on morphological filtering has been developed by Graham Treece in order to adaptively remove the amount of noise present in images. The bitonic filter follows the principle of bitonicity i.e. it preserves the signal content which is locally bitonic (contains only one maxima or minima in a given range) and removes continuously altering noise pixels. The filter has shown promising results in low to high noise values. Bionic filter is a typical non-local filter which constructs the weighted Gaussian output of opening and closing filter weights. It is an adaptive image denoising which preserves edges while removing noise, without need of prior knowledge of amount of noise. This filter is shown to have better denoising performance than Gaussian, Median and Opening–Closing and Closing–Opening filters in case of AWGN and Impulse noise [75].

The primary objective of any denoising algorithm is retention of fine feature details as much as possible, while denoising the image to a significant level of clarity, representation and interpretation. Most of the filters work on the principle of simply disabling the mechanism of noise reduction, at the point of detection of edges. However, these filters tend to follow similar criterion at fairly higher noise values, where intensity variation in noise exceeds edge discontinuity. Therefore, though spatial domain filters have achieved high levels of performance, there is always preservation of features at the cost of prevalence of noise. The noise is random and has been presumed to have higher frequency, and true signal has lower frequency and is repetitive in nature, therefore the higher frequency can easily be removed in order to achieve denoising. However the edges and features are often not repetitive and have higher frequency and their removal causes edge corruption and signal blurring. Therefore, these spatial domain filters continuously strive to remove high frequency noise and not high frequency signal, which remains the single most challenging problem in spatial domain filtering.

Besides this, spatial domain filters are prone to gradient reversal and halo effects near the edges. Also due to large window size of filters and convolution process, spatial domain filters are slow and computational ineffective. In some cases they cause staircase and rare-patch effect and it some cases they tend to cause over blurring of edges. Therefore there is a need to design a generalised framework which combines more than one technique which not only solves the problem of image denoising but also is able to retain contours, edges, features and details.

Due to high computational complexity and inability in differentiating between noise and sharp features in spatial domain, researchers have moved to transform domain techniques for image denoising.

The transform domain filters depend on various types of thresholds in order to differentiate between edge and noise and have sustained paramount importance in the field of digital signal processing and multi-resolution analysis. The next section is dedicated to various transform based image denoising methods and their merits and demerits.

3. Transform domain filtering

In contrast to approaches targeted in the spatial domain techniques, transform domain methods exploit the property of sparsity i.e. the signal could be represented by fewer number of non-zero coefficients or an image could be represented as a linear expansion of few high valued coefficient. This property has made them an extremely applicable digital signal processing tool both in 1-D and 2-D domain. The representational attributes like localisation, isotropy, multi-resolution and orientation of basis functions at a variety of directions are major properties of image transform methods. There are a large number of variations available in this category such as Fourier transform (FT), Fast Fourier Transform (FFT), discrete cosine transform (DCT), wavelets, curvelets, steerable wavelets, ridgelets, ripplet, curvelet, contourlets, wedgelets, bandlets, directionlets, shearlets and directional filter banks [1,6,8].

FT and DCT have been widely employed in image denoising and loss less compression. Though image restoration process was simpler with the help of DCT, the reconstructed images showed blocking artefacts and failed drastically at preserving edge features and details of images. The major disadvantage associated with FT was its long basis functions, whereas the multi-resolution tool- Wavelet transform exhibits sparsity (few number of non-zero coefficients). The major reason which made wavelet more popular than FT is that wavelets provide localisation both in space and time, whereas FT is localised in frequency only. Also in comparison to DCT, Discrete wavelet transform (DWT) employs higher number of optimal set of functions to represent sharp edges than cosines. Besides this in case of high pass sub-bands, DCT is able to provide higher frequency resolution and lower spatial resolution. A considerable amount of literature shows the superior performance of wavelet transform as compared FT and DCT, in terms of efficient image restoration and lossless compression [76].

The advent of theory of multi-resolution signal decomposition with wavelet transform by Mallat in 1989 brought a revolution in the field of image processing and computer vision [77].

A hierarchical framework for interpretation of image information is efficiently provided by multiresolution representation tools (MRT). At coarser resolutions, image can be viewed through larger structures which mainly identify the context of the image. With increasing resolutions, fine detailed structures are analysed. This particular property of MRT is useful in pattern recognition and computer vision [6].

DWT (Discrete Wavelet Transform) is a mathematical tool which is obtained by discretisation of continuous wavelet transform and is constitutes a translation and dilation factor thereby encompassing entire image at different scales in a sliding window manner. The energy compaction property of wavelet transform i.e. most of the information in image is encoded in few high valued coefficients is one of the major reason for its widespread application [78,79].

DWT constructs an approximation level of the images at coarser levels using low-pass filter banks and detailed level by employing high pass filter banks. Each level of approximation is further decomposed into image representation at next scale

The DWT uses a number of mother wavelets which provide basis function for signal decomposition. The well-known wavelet families are Biorthogonal, Coifflet, Daubechies, Discrete Meyer, Haar wavelet, reverse biorthogonal wavelet and Symlet [80–84]. The efficiency and suitability of mother wavelets are governed by factors like length of filters, orthogonality, compact support and smoothness. Daubechies or DbN, is most commonly employed mother wavelet function as they prove to be most suited for textural and feature analysis.

The denoising methods based on the wavelet usually transform the image content into multiple sub-bands at different resolution and scales. Larger frequency coefficients contain the low frequency image information (approximation level) and noise and details exist in high frequency sub-bands. By thresholding smaller coefficients noise can be removed and finally the image is restored by inverse transforming the coefficients into spatial domain. This can further be understood by Fig. 3.

In the multi-resolution subspace, dissimilarity between noise and true signal is accentuated on the observation that at low resolutions, signal components i.e. the edges or significant geometrical structures over-exceeds noisy pixels [85]. The basic idea behind removal of noise

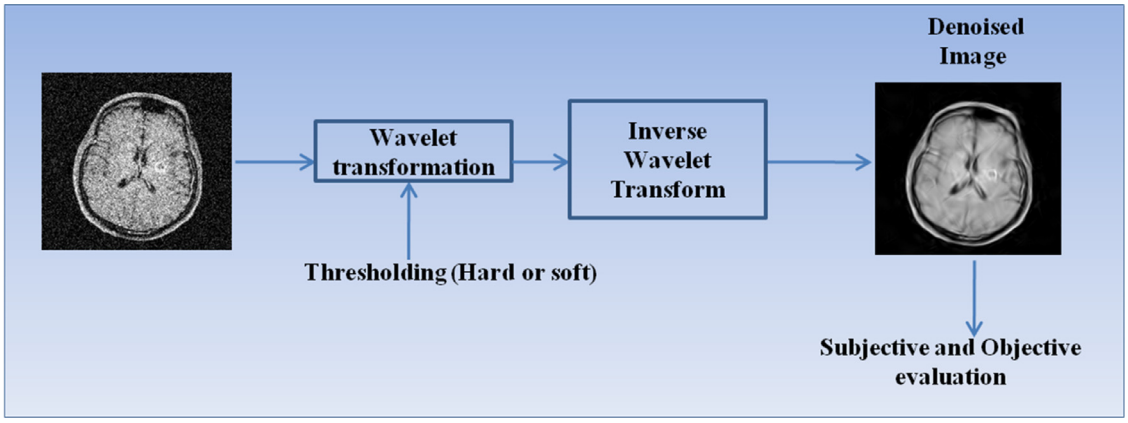

Fig. 3. Image denoising with wavelet transform.

is that smaller wavelet coefficients in the high frequency band can be thresholded and high magnitude coefficients in the lower frequency band are preserved which are lesser likely to be noise. The performance of wavelet thresholding depends on two major criteria: one is the selection of the wavelet basis or mother wavelet i.e. Haar, Daubechies etc. and other is the choice of threshold at all sub-bands for all levels.

3.1. Thresholding

Rather bigger challenge in field of image denoising resides with the choice of threshold. For instance, a larger threshold value will shrink the signal features, resulting in over-smoothing or blurring of signal, whereas a smaller threshold value might leave higher number of coefficients associated with noise information. There are mainly two types of thresholding functions given by Donoho and Johnstone in 1995: hard thresholding (HT) where each value is compared against threshold value and lower value is replaced by zero and other one is soft thresholding (ST); where coefficients larger than threshold value are modified by subtraction with the threshold value [86,87]. The HT leads to abrupt changes which leads to the generation of artefacts whereas ST sometimes tends to over smoothen the restored image. As far as visual appearance of the images is concerned, ST is better than HT. However it has been reported in literature than in some cases hard thresholding gives better results by preserving edge information [88].

ST often induces error and bias in the restored images. To address this problem the literature has contained several improvisations for the design of optimal thresholding function. For instance, hyperbola function [89], optimal linear interpolation shrink algorithm [90], were proposed in order to overcome limitations posed by ST and HT. Besides this, a thresholding similar to ST, semi-soft thresholding and Garrotte thresholding was proposed by Fodor and Kamath in 2001 [91]. These thresholds have a higher degree of sensitivity towards small variations in signal and generated results with lesser mean squared error. With the selection of suitable threshold value, wavelet coefficients are modified according to the shrinkage function. The restored image is obtained by inverse transforming the thresholded coefficients.

To analyse the progressive advancement in the field of design of threshold values, the existing methods are being categorised in three types [92,93]:

- Non-Adaptive Universal threshold: A unique threshold is selected for all the sub-bands
- Sub-band Adaptive thresholding: A threshold is selected for all the coefficients in the sub-band
- Spatially Adaptive thresholding: Each wavelet coefficient has its own threshold.

Under these categories the various types of thresholding methods existing in the literature are listed in Table 1.

Table 1
List of thresholding schemes.

S. no.	Thresholding methods
1	Universal Threshold
2	VISUShrink
3	SURE-Shrink
4	BayesShrink
5	Prob-Shrink
6	SURELET
7	Neigh Shrink Sure
9	Block Shrink

3.1.1. Universal threshold

To begin with, a universal threshold (UT) can be given as:

$$\lambda = \sigma\sqrt{2\log(L)}$$

where L is length of the signal, σ is the noise variance. In case of large number of samples UT gives a fair estimate of the restored image [86].

3.1.2. VISUShrink

VISUShrink is a non-adaptive universal threshold given by Donoho and Johnstone in 1995. The threshold value can be given as:

$$T_{visu} = \sigma_n\sqrt{2\log(k)}$$

where ${\sigma_n}^2$ is the noise variance of the AWGN, and K corresponds to range of the image. For larger values of K, T_{visu} is also high, therefore it generates a fairly smooth estimate. This type of threshold does not adapt well to the discontinuities in images however, it throughputs better performance than universal threshold. In case of high probability where the number of pixels reaches infinity it gives the best asymptotic MSE (Mean Squared Error) performance. However it often leads to over-smoothing of the images resulting in higher MSE [87].

3.1.3. SURE-Shrink

SURE-Shrink was the very first attempt in terms of design of an adaptive threshold where a different threshold is calculated for each of the detail sub-band. It is a hybridisation of Universal and SURE threshold and depends on energy of the particular sub-band. It employs SURE threshold in case of high energy coefficients and UT in case of sparse sub-bands and gives better performance than VISUShrink. The threshold value at each resolution level is given by:

$$\lambda = \min(t,\ \sigma\sqrt{2\log(k)}$$

where K refers to the number of wavelet coefficients in the specific detail sub-band, t is employed to minimise SURE and σ is the noise variance. SURE-Shrink minimises MSE, delivers a better performance than

VISUShrink and is smoothness adaptive: i.e. it preserves the boundaries and discontinuities in the images.

According to the principle of wavelet energy compaction, most of the signal's energy is concentrated in few wavelet coefficients and noise is spread over all the coefficients and λ depends on the number of wavelet coefficients in the specific band. In case of extreme sparsity, the information contributed by the coefficients where signal is zero, divulge the information extended by the non-zero signal wavelet coefficients [94,95].

3.1.4. BayesShrink

BayesShrink was given by Chang et al. and is a mathematical framework which proposes a data driven adaptive threshold for generalised Gaussian distribution in each of the detail sub-band. The threshold so devised aims at minimizing the Bayesian risk.

According to its mathematical framework [88,92,93].The variance of a noisy image can be given as:

$$\sigma_y^2 = \sigma_x^2 + \sigma_n^2$$

where σ_y^2 is the variance of the noisy image, σ_x^2 is the variance of the original signal and σ_n^2 is the noise variance.

Then Bayesian Threshold is defined as [88]:

$$T_{Bay} = \frac{\sigma_n^2}{\sigma_x^2}$$

$$\sigma_x = \sqrt{\min(\sigma_y^2 - \sigma_n^2,\ 0)}$$

$$\sigma_y^2 = \frac{1}{K}\sum_{k=1}^{K} B_k^2$$

B_k denotes the wavelet coefficients at each scale and K is the total number of sub-band coefficients and σ_n is computed with the help of Robust Median Estimator.This type of threshold generates better visual results than SURE-Shrink.

3.1.5. Prob-Shrink

Given by Pižurica and Philips, Prob-Shrink is a probalistic shrinkage function which calculates the probability of the event: the given coefficients contain signal features of interest. This method assumes a generalised Laplacian prior for the noise free data and computes the product of the probability and wavelet coefficient. In Prob-shrink framework the signal of interest or signal wavelet coefficients exceeds the threshold which leads to the shrinking of the smaller coefficients and preservation of larger coefficient values [94,96].

3.1.6. SURELET

In a work in [97] an orthonormal wavelet approach was given towards denoising. In this method, the denoising criterion was directly parameterised as non-linear expansion of the elementary processes with weights unknown.

This technique aimed at minimizing MSE between the clean and noisy signal. SURE (Stein's Unbiased Risk Estimator) is a statistically unbiased MSE estimator which only depends on the noisy image and minimisation of which leads to solutions of linear set of equations. Luiser et al. further improvised the work given in [97] by appending the principle of SURE. It was suggested that a typical denoising process can be expressed as linear combination of elementary denoising processes- Linear Expansion of Thresholds (LET). This approach is popularly called as SURELET. This method has been applied for grey scale and colour images and for mixed Poisson–Gaussian denoising [98].

3.1.7. Neigh Shrink Sure (NSS)

Neigh Shrink (NS) is a data driven sub-band adaptive thresholding in which a squaring window is centred on the noisy coefficient to be shrunk. This method was further improvised to introduce Neigh-Shrink SURE. NS employed a non-optimal universal threshold and for all the wavelet sub-bands same neighbouring window size was used. However, NSS derived an optimal threshold value and SURE based neighbouring window size for each of the sub-band. The mathematical formula for NSS can be given as:

$$\lambda^s K^s = \arg\min SURE\left(W_s, \lambda,\ K\right)$$

λ^s is the optimal threshold for sub-band S, K^s is the optimal window size for sub-band S and W_s governs the minimisation of SURE [99].

3.1.8. Block Shrink

In this approach an optimal size of block and threshold value is calculated for each wavelet sub-band in order to minimise SURE. This type of thresholding preserves or eliminates the coefficients at the same time. Hence it is easier to implement and gives better performance than SURE Shrink and Neigh Shrink [100].

The thresholding schemes play a vital role in image denoising process and each one has its own pros and cons. This is a brief account of various thresholding methods existing in literature. By exploiting the interscale and intrascale correlations, these methods have been further improvised. BLS-GSM is also one such example which is discussed in an elaborative way in the later section owing to its widespread acclamation in image denoising. The transform domain methods have been widely employed in image denoising, as they provide favourable assumptions for filtering out noisy information from image features and details.

The following properties of wavelets have made them an extremely popular method for image transformation [6]:

- Sparsity: most of the information is contained in few high magnitude coefficients and rest of coefficients are small.
- Multiresolution: Different sub-bands at different levels, contains varied type of information. This facilitates different operations on sub-bands for increased precision.

In spite of the fact that wavelets have achieved remarkable success in providing optimal approximation for one dimensional piecewise continuous functions they are not very effective in optimally representing multivariate functions in an image that are typically governed by anisotropic features such as edges [101]. WT (Wavelet transform) is better to represent 1-D signals but when it comes to 2-D signals where edges and contours are significant image features vital to comprehension and readability of images, these transforms did not work well.

The limited performance of the DWT is attributed to its lack of shift invariance i.e. a small variation in input signal can develop significant variations in distribution of energy amongst wavelet coefficients at different scales; poor directional sensitivity towards edges as the DWT coefficients are separable and real. Furthermore, in case of DWT, if number of pixels in an image patch exceeds the number of coefficients in the patch, ringing phenomena would occur.

To overcome these issues associated with DWT, popular undecimated discrete wavelet transform was proposed in [102] for denoising. This method was able to alleviate issues posed by DWT to some extent, but it resulted in increased redundancy in throughputs and was computationally inefficient.

Further, keeping in mind the shift variance and reduced directional sensitivity of DWT, Kingsbury in 1998, proposed the Dual-Tree Complex Wavelet transform (DT-CWT) which provided shift-invariance and directional sensitivity with the help of Gabor filters. Besides this it gave better reconstruction, limited redundancy and was computationally efficient [103].

On similar lines an adaptive thresholding of wavelet packet was performed in order to suppress the noise in medical images by Fathi et al. in [90]. The technique showed an increased amount of performance at varied noise levels for medical images as compared to basic wavelet thresholding methods. Besides this, in adaptive directional lifting schemes, the ADL transform is applied in predicted directions besides horizontal and vertical directions [104].

In a rather interesting denoising approach an idea of selective wavelet shrinkage was employed. It determined that selection or rejection of wavelet coefficients is better than probabilistic methods as former can statistically facilitate identification of a narrow interval for estimated parameter which eventually adjusts wavelet coefficient with a high level of certainty. This method chooses wavelet coefficients based on image features like their absolute value, regulatory across multi-resolution scales and spatial regularity. The few large magnitude coefficients are grouped together. The first threshold chooses the coefficients of large magnitude and second chooses a subset which exhibits spatially regular behaviour. This two-threshold method being non-iterative in nature provided a computationally simple solution facilitating real time processing [248].

The traditional orthogonal, maximally decimated wavelet based denoising methods exhibited visual artefacts like Gibbs phenomena due to lack of shift invariance. Coifman gave the concept of “cycle spinning” in [105], to average out translational dependence thereby, addressing visual artefacts. For a given set of shifts, once the data is shifted, it is denoised and it is then unshifted. By carrying this for entire range of shifts and averaging the shifts, results so obtained exhibited much lower number of visual artefacts. Another interesting improvisation in the wavelet domain was introduction of steerable pyramids [106].

The agenda of improvisation of wavelet performance was three-fold: First to design adaptive thresholding methods in order to preserve more number of high magnitude signal coefficients, second was to design directionally sensitive image transforms which could boost denoising performance and third was to hybrid wavelet transform with various methods in other domains in order to make optimal use of non-local image features.

The various threshold based methods have been discussed above. Besides this there are various other approaches in order to enhance the performance of wavelet transform by employing data-adaptive and edge preserving thresholding methods.

For instance, in a work given by Silva et al., image is initially divided into a set of blocks and then coefficients were obtained employing wavelet transform. An edge strength based adaptive thresholding was used to suppress noisy coefficients [107]. In 2013, Qiu et al. proposed explicit filter known as LLSURE. In this method a simple affine transform for input signal was considered for the filtered output in a local window. Further, optimal transform coefficients are constructed by minimizing SURE. This method was motivated from SURELET and Guided Filter [108]. Further there were methods developed on statistical noise models, MAP and Markov random fields which will be discussed in another section of this article.

In order to address the limitations posed by wavelet at higher dimensions; a new representational system called Ridgelet was pioneered by Candes and Donoho in 1999, which would effectively represent the singularities associated with lines in 2 dimensions [109,110]. The idea behind ridgelets was to enforce the wavelet with effective handling of point singularities in Radon domain. The implementation in radon transform domain facilitates mapping of the line singularity into point singularity [111]. Initially, ridgelets were proposed for continuous space, however the discrete implementations were a challenging problem, as straightforward implementation by enabling the discretisation of the continuous framework, would result in interpolation in spatial coordinates thereby leading to redundancy and imperfect reconstruction.

Therefore, in 2003 [112], a finite discrete ridgelet transform (DRT) was proposed that was both invertible and redundant. The proposed transform was found to have numerical stability and included adaptive schemes along with directional and orthonormal bases for digital signal in two dimensions. The properties of this transform were validated in several applications including image denoising. The DRT had built-in linear geometrical structures, which facilitated direct thresholding of the ridgelet coefficients in contrast to wavelet domain where adjacent coefficients are initially chained together before thresholding them over the contours. In this algorithm as well, the basic building black was radon transform [113] which resulted in wrap around effect or aliasing and geometrical infidelity. The issue of wrap-around effect was effectively removed in a work given in [114]. The author simply employed the finite radon transform with size equal to three. A multi-scale image denoising algorithm was developed in this work which effectively preserved edges while removing noise. This method employed a moving window pyramid which yielded better approximation of lines and curves.

Further, an adaptive digital ridgelet transform (ADR) was proposed in 2016, as a multi-scale image decomposition algorithm for image denoising. The improvised methodology resulted in better representation of line and curve information by taking into consideration the underlying structures of image. The experimental results for this algorithm were able to depict PSNR improvement as compared to ridgelet and curvelet transform [115]. An orthonormal finite ridgelet transform was also proposed in literature for image denoising [116]. Besides this, there were other image denoising algorithms documented in literature which employed ridgelet transform in conjunction with other methods in order to increase its efficacy. These methods fall into typical hybrid domain methods, which will be discussed in the later section of this article.

Contemporary to ridgelet, curvelet transform has also been pervasively studied in literature for image denoising. The curvelet and ridgelet transforms were infact the major breakthroughs in the field of contour based image representation.

The prime motivation behind development of curvelet and ridgelet was to expand inherent limit plaguing the denoising of images with wavelet transform. This limit arises from the fact that along the significant edges of image, wavelet exhibits large magnitude coefficients even at finer scales, such that within the map of large magnitude coefficients, one may find edges repeated at each and every scale. This intuitively means that a large number of wavelet coefficients are required in order to represent image edges in an efficient manner. This leads to higher computational complexity as there are so many wavelet coefficients to process. Besides this the MSE achieved by wavelet based denoising methods was of the order ϵ, noise parametric measure. However, the theoretical limits suggested it could be achieved uptil $\epsilon^{4/3}$.Therefore it was required to achieve this ideal MSE by developing new expansions which could efficiently represent smooth functions as well as edges using few non-zero coefficient values.

It is a general understanding that edges are more of ‘*curved*’ parameters rather than being straight! At very fine scales, curved edges appear to be straight, so one should deploy ridgelet in a more localised manner in order to capture edges. Therefore, Curvelet transform, a multi-scale variant of ridgelet transform was developed by Starck et al. based on isolation of varied scale using sub-band filtering [117]. The authors reported that likewise ridgelets, curvelets occurred at all scales and locations as well as orientations. Besides the properties of variable length and width, curvelets possessed patchy anisotropy. The length and width were related by parabolic scaling and anisotropy increases at higher scales. It could be witnessed from experimental analysis given in [117], that thresholding of curvelet coefficients results in near-optimal solutions of smooth regions as well as discontinuities along curves.

In work given in [118], curvelet coefficients were thresholded using soft-thresholding based on Bayesian theory in order to reconstruct ionograms from the noisy ones. In another piece of work, a new transform based on combination of wavelet, curvelet and ridgelet transform known VST (Variance stabilizing transform) was given for removal of Poisson noise. This transform is also popularly called extension of Anscombe transform [119].

The primary limitation of Curvelet transform is that it is not built directly in the discrete domain and also is not able to provide a multiresolution geometry. Also mathematical analysis and implementation in discrete domain are more involved and less efficient as compared to other directional transforms. Besides this, curvelets were lacking in orthogonality and weren’t critically sampled. Also thresholding in curvelet

domain results in visual artefacts in the manner of ghost effects alongside strong linear features.

While invigorating battle to design multi-scale transforms prevailed in beginning of 20th century, some imaging transform like Brushlets [120], Wedgelets [121], Beamlets [122] benefited the field of image processing in an immense manner by enabling the efficient singularity analysis. Further ahead, by considering the noteworthy chronological order, after curvelet transform, arose the era of contourlet transform. While witnessing, the failed performance of wavelet transform for representing 2-D discontinuities, it was progressively thought that besides precise localisation and multi-resolution analysis, new 2-dimensional should be designed with elongated shapes oriented at different directions and with varied aspect ratio.

The curvelet transform was able to achieve all this, but the discretisation of curvelet transform for critical sampling was rather challenging. Therefore, the improvement of the curvelet transform over wavelet to represent 2-D piecewise smooth signals with smooth and discontinuous curves inspired Donoho and Vetterli, to design Contourlet transform (CT). The contourlet transform is a double iterated filter bank structure which solves the anisotropy scaling equations for curves resulting in a fast curvelet like implementation. The main idea behind construction of contourlet was to design a sparse representation, thereby providing smooth curves across discontinuities. In the basic understanding, edges are referred to as gathering of smooth boundaries along shapes of physical objects and discontinuities are sharp transitions in intensities. The contourlet transform gave improved performance over wavelet and curvelet transform in terms of denoising. CT can be understood as deployment of wavelet-like transform for edge detection, and localisation of contour segments using directional transform. In double filter bank structure, Laplacian pyramid is used to acquire discontinuities and further directional filter banks are employed to expand basis images as contours or in other words to frame point discontinuities into linear structures [123,124].

According to Do and Vetterli [123], Contourlets are constructed with elongated basis function have lesser directional features and denoising with contourlet transform tends to introduce artefacts and Gibbs phenomena as they are translation invariant. Henceforth in a work given in [125], Eslami and Radha, employed cycle spinning based contourlet transform to induce directional sensitivity in terms of image denoising. This technique gave promising results for removal of AWGN as compared to contourlet transform over extended range of standard deviation of noise.

Due to sub-sampling at varied resolution levels, there is a shear lack of translational invariance, which is vital to image denoising. Therefore in another attempt by Eslami and Radha, translational-invariant contourlet transform was designed which employed modified versions of directional filter banks and algorithm *a trous* for 1-D wavelet transform. Along with bivariate shrinkage thresholding their scheme presented an incredible denoising approach [126].

In a work given by Cunha et al., the concept of non-sub-sampled contourlet transform (NSCT) was floated which was fully multi-scale, shift invariant and had a fast directional expansion. It employed non-sub-sampled pyramid banks and filter banks which dispensed the need of 2-D factorisation thereby eliminating Gibbs phenomena. The NSCT was coupled with hard thresholding and local adaptive threshold to generate considerably higher PSNR than contourlet transform [127]. Some other significant works based on contourlet transform for image denoising could be found in [128–130].

The study of generalisation of parabolic scaling law result in designing of another interesting MRA (multiresolution analysis) tool i.e. Ripplet Type-1 transform. This transform is a higher order generalisation of curvelet transform with the addition of two more parameter; support c and degree d. RT converges to curvelet transform when $c = 1$ and $d = 2$. Ripplet transform was shown to have exhibited promising results in image denoising at par with curvelet transform [131,132]. Ridgelet resolves the discontinuities along lines where as Counterlet and Ripplet are confined to resolve discontinuities along smooth curves. But these transforms are incapable to resolve discontinuities along edges and contours. While devising a transform one has to use a basis element with much higher directional sensitivity, having more shapes and directions than the classical wavelets. In this relentless progressive evolution in the designing of new transform some of the major transform are bandlets [133], haarlet [134], ranklet [135], morphlet [136], tetrolet [137] and many more which could be found in [138]. Also to remove Poisson noise, Luiser et al. proposed PURELET (Poisson unbiased risk estimate linear expansion of thresholds). It was devised to do two functions, one is signal preservation and other is noise suppression in presence of shot noise. PURE denoising function uses unnormalised Haar wavelet transform [139]. Many methods have been reported in literature which target both Gaussian and Poisson noise simultaneously.

In the process of continuous progression of LET's (Linear expansion of thresholds) several image transforms were devised. The significant transforms in the field of denoising have been discussed at some stretch in this article, while some are just mentioned.

Having transit through a gulf of transforms, there were a whole other era of highly directionally sensitive transforms in literature, few of them are Shearlet transform, Directionlet transform and Framelet transform. Since these transforms have exhibited exceptional performances in image denoising in terms of PSNR, SSIM and visual quality, therefore they are being discussed in brief in the following section.

In the linear expansion of thresholds; multi-directionality (M-DIR) and Directional vanishing moments (DVM) (zero at $w = o$; sparse representation) are major properties which make them suitable for any image processing task. The efficiency of 2-D wavelet transform is limited by its spatial isotropy and process of convolution and non-separable filter in both horizontal and vertical directions. Curvelets and contourlet provided anisotropy to some extent but they require oversampling during implementation and are highly complex. Velisavljevic et al. [140] proposed a multidirectional anisotropic transform based on integer lattices named it Directionlet. DT is an anisotropic, multi-scale geometric transform with directional vanishing moments along any two directions. Besides being computationally efficient, it has perfect reconstruction and critically sampled basis functions. In a work given in [141] several thresholding schemes have been exploited in Directionlet domain to exploit intrascale and interscale dependencies.

In order to substantiate the need of efficient edge representation while denoising of images, an increased sparsity based algorithm was proposed in [142] using directionlet transform. For each pixel value multiple plausible estimates based on constructed DT representation of directions of edges were obtained. By computing weighted average of each of these estimates final denoising image was constructed and has shown to depict better performance that other multi-directional wavelet based transforms [142]. Besides this a novel denoising algorithm based on Cauchy's density function and direction-let transform was designed to remove speckle noise from SAR (synthetic aperture radar) images [143]. Another exemplary and rather recent work on removal of speckle noise from ultrasound images is given by Shahdoosti can be found in [250] and is proposed in 2018. In this method general speckle method is employed and log likehood function is maximised in order to obtain noise free pixels. Further, distribution of each noisy pixel is weighted in accordance with statistical similarity between the patch of noisy pixel and that of the pixel being processed. The subjective as well numerical observations depicted the superiority of algorithm over state-of-the-art using spleen, pancreas and liver ultrasound images.

In the era of multi-scale directional representation of images, a fairly new representation scheme has been presented by Easley et al. and is called Shearlet. It is found to be very effective in providing optimally sparse representations for a variety of multidimensional data and has showed improvements in variety of image processing applications such as image enhancement, image fusion and image denoising. Shearlets have been extensively used in context of image denoising [144]. It

Table 2
Classification of image denoising methods.

Domain	Spatial domain filtering	Transform domain filtering	Other domains	Hybrid methods	Sparse representation
Image denoising methods	Local Filters & Non-Local Filters: Average filter, LTI Filters [1, 6, 8], Gaussian Filter [12], Weiner Filter [13,14], Rank Filter [17], Anisotropic Diffusion Filtering- PM-Model [18], Yaroslavasky Filter [47], Bilateral Filter [48], SUSAN filter [49], WBF, RBF [54], Guided Filter [65], Joint bilateral Filter [64], fast bilateral filter [55–57], Bitonic filter [75], Co-occurrence [217], NLM [6], INLM [69], Morphological filters [74].	Fourier Transform [76], Wavelet transform [77], DWT [6], Dual Tree complex wavelet transform [103], Ridgelet [115], Curvelet [117], Ripplet [131], Contourlet [123], NSCT [128–130], Shearlet, NSST [144], Directionlet [140], Framelet [150], BLS-GSM (Bayesian Least Square Gaussian Scale Mixtures [155], BM3D (Block Matching 3-D collaborative filtering) [151], LPG-PCA (Local Pixel Grouping-Principal Component Analysis) [153], Thresholding methods	Bivariate Laplacian Distribution [159], Maximum *a* prosteriori [160], Markov Random Field [163], MSB-CRF [166], Sylvester Lyapunov Equation [168], RADL [169], Hidden Markov Models [170]	Bilateral Filtering+ Short Term Fourier Transform [172], Dual Domain Image Denoising [173], Progressive Image Denoising [3], Perona-Malik-Isotropic Diffusion [176], Trivariate wavelet+ Spatial Domain filtering [177], Non-local mean +Partial Differential equation [179], BM3D+Brushlet [180], NLTV + BM3D [181], TV+NLM [182], DTCWT+ Ridgelet [192]	K-Singular Value Decomposition [197], K-LLD [201], LSSC (learned simultaneous sparse coding) [202], CSR (convolution sparse representation) [4], M-SKR (metric steering kernel regression) [203], PLOW (patch Based locally optimal wiener filter
Applications	Natural Image denoising, medical Imaging, Remote sensing, Infrared Image denoising, Under water image denoising, SAR de-speckling				

provides a more flexible tool and is more natural in implementation. The major underlying property of Shearlet is its shift invariance. It is band limited and compactly supported and is basically a type of composite wavelets having a ‘;directional parameter as well. It helps to provide the multi-resolutional and multi-scale representations due to the introduction of shear parameter and anisotropic dilation. However Shearlet based denoising algorithms should be optimised using adaptive thresholding schemes which tend reinforce the image features and details initially preserved by Shearlet transform [145]. Non-sub sampled Shearlet transform is a special type of discrete Shearlet transform. The major underlying property of NSST which has made it very popular for image denoising is its invariance to shift of the input signal.

In context of denoising it is well known that shift variance introduces Gibbs phenomena around singularities i.e. artefacts which are considerably reduced by NSST. The NSST is implemented using Non-sub sampled Pyramid filter banks (NSP) and Non sub sampled shearing filters (NSS) [146]. With an aim to obtain near optimal estimates and to remove Gibbs type artefacts, shearlet transform was coupled with total variation method for denoising of images. The shearlet coefficients were subjected to adaptive total variation minimisation which resulted in outperforming related methods in curvelet domain form highly complex images [147]. While witnessing the edge-preserving multi-scale and multi-directional representation of the NSST, a technique based on NSST and twin support vector machines (T-SVM) along with sub-band adaptive thresholding was proposed in [4]. This work besides outperforming the techniques like probshrink and BLS-GSM (Bayesian Least Square Gaussian Scale Mixtures), presents a very brief, concise and informative literature survey on image denoising. The technique was typically able to preserve edges while suppressing the noise. Some of other effective image denoising algorithms exploiting the edge preserving properties of NSST are documented in [148,149]. After extensive literature survey, it is found that NSST is a prospective image denoising algorithm when coupled with effective thresholding schemes and various spatial domain filters.

A major pioneering work in the field of image denoising includes adaptive directional lifting schemes for WT. This method is proposed as translational invariant directional framelet transform which provides substantially superior redundant representation along with novel orientation estimation based on Gabor filters [150].

Besides these above mentioned transform domain techniques, the interscale and intrascale correlations of the wavelet coefficients have been exploited in BM3D [151], BLS-GSM [152] and LPG-PCA [153]. These three techniques infact beg the medal for the state-of-the-art denoising techniques. The hierarachy of various methods under specific domains is summarised in Table 2.

3.2. BLS-GSM (Bayesian Least Square Gaussian Scale Mixtures). It is a well-established fact that there exist interscale and intrascale dependencies amongst wavelet coefficients. It can be defined that if a (parent) wavelet coefficient is small or large, its adjacent (child) coefficient is also small or large and these coefficients tend to spread across scales and hence can be modified. Since the wavelet coefficients were critically sampled and can lead to visually unpleasant artefacts there was need to develop overcomplete wavelets. This idea was exploited in BLS-GSM, where wavelet coefficient’s neighbourhoods at different positions and scales are modelled as Gaussian Scale mixture (GSM). The noise free coefficients were computed using Bayesian least square estimator. In this algorithm, first the AWGN contaminated image is transformed into wavelet coefficients. The each local neighbourhood is modelled as GSM and centre of each of the neighbourhood is estimated by employing local Weiner estimator. In order to obtain the full optimal estimate of coefficients BLS solutions are computed. The denoised image can be obtained by performing the inverse wavelet transform. This denoising algorithm is by far the most proficient image denoising algorithm whose efficiency is matchless. However, since this method requires strict estimate of power spectrum density, this makes it non-adaptive [152].

Some of other methods based on exploitation of interscale and intrascale dependencies are as follows: Bi-shrink is a wavelet shrinkage technique designed by Sendur et al. using MAP (Maximum *a Posteriori)* and is based on a non-Gaussian bivariate distribution for modelling interscale dependencies [154]. In another work intrascale variability of the wavelet coefficients is taken into account [155]. A method based on Gaussian Mixture Model (GMM) and Generalised Gaussian Distribution (GGD) for wavelet coefficients was given in [156].

3.3. BM3D (Block Matching 3-D collaborative filtering). Motivated by the idea of non-local grouping in NLM filter and redundant or overcomplete representation in BLS-GSM, Dabov et al. designed an effective filtering in 3D transform domain in 2007. It was designed by combining sliding window process with block matching. BM3D can be denoted as multipoint approach which comprises of three steps: The single noisy image is processed in a sliding manner, where blocks that exhibit similarity (minimum Euclidean distance) with currently processed block are searched and grouped together. The resulting matched blocks are piled together to form a 3D array. The 3D groups so computed are redun-

dant and hence facilitates more effective representation than WT. The correlation amongst the blocks is exploited by applying a 3D decorrelating unitary transform which results in a sparse representation of the original signal in 3D transform domain. After the efficient and sparse representation of noisy image, noise attenuation is done by applying hard thresholding or Weiner filtering on transform coefficients. BM3D comprises of these two steps in a sequential manner except for the fact that hard thresholding is carried out in the first step and Weiner filtering is carried out in the second step. The local estimation of matched blocks constructed by inverse 3D transform, results in improved denoising performance and effective detail preservation. The grouping provides extremely reliable statistical data and Weiner filtering of 3D arrays is quite effective.

The similarity amongst overlapping patches and correlation between wavelet coefficients tends to deliver high performance with BM3D. However, in case of presence of unique patches, BM3D ceases to provide optimal solution. At low noise levels, BM3D has been established to provide optimal denoising solutions; however, at standard deviation above 40, performance of this algorithm drops as correlation amongst blocks is highly disturbed by noisy pixels, even though pre-filtering is carried out. Also sometimes instead of wavelet, in order to enhance performance at standard deviation 40, DCT transform is employed which have fixed basis functions but results in periodic artefacts. However in an article [158], it has been commented that this replacement is unnecessary and performance improvement can be established by adjusting some numerical parameters and same has been shown in that work. The BM3D approach was primarily designed taking into consideration the contamination of the images by Additive White Gaussian Noise [151]. To further increase sparsity of true signal in 3D domain Dabov et al. proposed a generalisation of BM3D filter by making use of mutually similar shape adaptive neighbourhoods. These adaptive shape neighbourhoods display local adaptivity to image details and features so that the true signal becomes mostly homogenous. Further ahead in 2009 to enhance adaptive capacity of pre-existing shape adaptive BM3D Dabov et al. introduced PCA (principal component analysis) as a part of the 3D domain. The overall 3D transform is a separable composition of PCA and a fixed orthogonal 1D transform in third dimension. The technique can be further improvised by addressing novel shrinkage criteria which are highly adaptive to the utilised transform [157,158].

3.4. LPG-PCA (Local Pixel Grouping-Principal Component Analysis). Using classic idea of exploitation of non-local features and sparsity, Zhang et al. proposed LPG-PCA for image denoising in order to mitigate the issue with fixed basis function in DCT transform. The PCA introduced the concept of adaptive basis functions.

In this method, a two-step approach is followed where each of noisy pixel and its typical neighbourhood is grouped together into a vector variable. These vectors so obtained are transformed into well-known PCA domain or KL transform using eigen values and vectors. The shrinkage is carried out in PCA domain. The same process is followed in the second step. The local fine grain-like edges are prone to have incorrect representation in BM3D due to fixed basis function. LPG-PCA is able to overcome this issue with help of local adaptive basis functions. However, input to second stage is filtered patches from the first stage. The second stage is exact replica of first stage except for different noise levels. This factor due to persistence of errors at first stage degrades the denoising performance at the second stage as well [153].

In the field of transform domain image denoising the primitive efforts were centred on thresholding of orthogonal wavelet coefficients of noisy pixels. Then further the idea of translational invariant schemes came into picture with the use of undecimated wavelet transform. More recently several experiments have been carried out with tree-based methods and modifications of various thresholding schemes like Bayesian thresholding, level-dependent thresholding and so on. Then focus substantially shifted towards highly directional transforms and their combination with adaptive thresholding methods, the non-local pixel similarity clustering like BM3D, Local pixel grouping, BLS-GSM and NSST and so on. Various conventional properties like multiresolution, sparsity, edge detection and anisotropy made transform domain methods an extremely popular choice for image denoising.

However transform methods also had some drawbacks. The wavelet transform was ineffective in representing the smooth transitions and DCT could not represent sharp transitions or singularities. The characterisation of natural images with different patterns became difficult as most of the transforms had fixed basis functions. The core idea of the wavelet denoising stemed from the fact, that most of the signal energy is compacted in few wavelet coefficients while the noisy pixels are spread throughout the wavelet coefficients. However the efficiency of these methods is strongly limited by the estimation of appropriate threshold. Besides this, some methods resulted in ringing artefacts. Methods like BLS-GSM overcome these issues by exploiting redundancy in local neighbourhoods. BM3D accomplishes better denoising results by exploring redundancy in the non-local manner. However when the assumption of self-similarity amongst the image pixels is not guaranteed well, the performance of these methods is affected by occurrence of image artefacts and residual noise.

4. Methods in other domains

Other than spatial and transform domain methods, researchers have designed image denoising methods based on statistical models of multiscale coefficients and random fields. In statistical model schemes, the noiseless coefficients are estimated from prior information based on Bayesian estimator. These methods aim at constructing spatially homogenous models adaptive to varying levels of signals with help of parameterisation of the variables which are self-random in nature. For instance, Yin et al. proposed an algorithm for removal of noise with the help anisotropic bivariate Laplacian distribution function based modelling of wavelet coefficients [159].

In another interesting set of works, Rabbani proposed a noise suppression method which constituted of thresholding of coefficients using steerable pyramids and Laplacian probability distribution function. This method was further improvised in [160]. In [161], an approach towards denoising of images is given by maximizing the MAP (maximum a prosteriori) measure with help of Gaussian Scale mixtures. Besides this, Fathi and Nilchi developed an adaptive wavelet coefficient thresholding algorithm which was statistically optimum. Despite the widespread acclaim of the statistical methods, they are observed to introduce ringing artefacts in the images [162].

The methods based on random fields is another very popular domain which is readily exploited in the field of image denoising as well as other low level processing such as image classification and segmentation. In these methods, the value of intensity of a pixel depends on neighbouring pixels. These methods are based on an observation that a global representation of an image can be obtained from its local physical structures which is accomplished with the help of a conditional probability distribution function popularly called as Markov random Field (MRF). Further under a Bayesian framework, a maximum a priori (MAP) probability is defined which states that a particular clean and sharp image could have been made noisy in order to obtain a specific degraded image. These MRF models establish relationship between the neighbouring pixels and between observed and estimated values at each pixel [163].

In [164], a three layered hierarchical MRF model has been proposed for image denoising where each of three layers depict the basic textures and smooth regions, the noisy image and observed noisy data respectively. In a work given by Chen et al., another MAP estimation based denoising solution has been employed using MRF's for SAR despeckling [165]. Zhong and Wang exploited both spectral and spatial dependencies and proposed denoising algorithms in the multiple-spectral-band-

conditional random fields (MSB-CRF) for hyperspectral images [166]. In an interesting work given by Ho and Hwang, separate Bayesian networks were obtained for different input images with the help of a single Bayesian network estimation of wavelet coefficients [167].

In another efficient denoising algorithm, Bayesian and MAP criterion were put in use in order to induce preservation of edges while removing AWGN and multiplicative noise having Rayleigh and Poisson distribution respectively from CT images [168]. The significant contribution of this work was development of a denoising algorithm using Sylvester Lyapunov equation in Bayesian Framework. In a work given by Wang et al. in [169], a robust adaptive directional lifting (RADL) algorithm was performed at each pixel level by estimating the noise idea of pixel pattern classifier.

While understanding the concept of inter-scale and intra-scale dependencies of the wavelet coefficients, Romberg et al. [170], established that where MRF models are more efficient at capturing intra-scale correlations, Hidden Markov models (HMM) work well for inter-scale correlations amongst edges and features. Their method employed the attributes of HMM to capture the higher order statistics, coefficient correlation at same scale and local structures. In order to acquire the interaction amongst coefficients across scales the model was constructed using Hidden Markov Trees. Further the MAP estimation maximisation technique is employed in order to restore the denoised image. The method has been shown to work except for higher computational and training burden. In order to address this issue, Malfait proposed a GSM model for neighbourhood estimation of wavelet coefficients with the help of a random scalar multiplier [171].

Exploiting a model based on intrascale and interscale dependencies of coefficients a method has been proposed in [252], for denoising for grey scale images. It has already been explained in concept of Gaussian scale mixtures that image can be estimated with the help of parent coefficients. The method comprises of two stages in which first stage comprises of pre-filtering with Kalman filter [KL] and second stage constitutes denoising of noisy wavelet coefficients with KF only. In fact employment of KF facilitates to estimate coefficients of high-frequency subbands from coefficients of coarser scales and noisy observations of neighbouring coefficients.

For the first time, Hidden Markov tree with mixture of one-sided exponential density has been exploited to remove noise in images in a work given in [253]. Further this step was combined with Dual contourlet transform which induced shift invariance and mitigated Gibbs phenomena.

The methods based on MRF field have achieved considerable success in image denoising however it has been found that, in real time applications, the energy of the random fields should be computationally feasible in lesser time and it is always extremely exigent to derive near optimal energy functions which could provide desired solutions [4].

5. Hybrid denoising techniques

The field of denoising has witnessed the design of large number of efficient and widely applicable denoising algorithms. Each of them has its own share contributions, applications and limitations. This abundant amount of denoising algorithms has exploited various domains of digital signal processing and relevant theory. Initially various types of spatial filtering techniques with exploitation of local and non-local similarities were designed. Besides this, various wavelet based adaptive shrinkage methods were employed. Other methods included directionally sensitive transforms and statistical methods exploiting the inter-scale and intra-scale dependencies. Several algorithms have been deployed based on sparse representation, random fields and dictionary learning. These methods have shown to be progressively improved in one manner or the other in terms of quantitative measurements and visual quality.

However the current state-of-the-art techniques combine more than one domain in order to exploit advantages of multi-domain while overcoming limitations of each of them. For instance, various spatial domain filtering methods are able to preserve the edges but they tend to overblur low contrast details in the homogenous regions. On the other hand, transform domain methods are able to represent the textures and low contrast information but they tend to introduce ringing artefacts around the edges.

Based on this approach Knaus and Zwicker combined two popular filtering techniques namely bilateral filtering from spatial domain and Short Term Fourier transform (STFT) with wavelet shrinkage from transform domain. With the help of this hybrid approach authors were able to preserve low contrast details along with preservation of edges. This method is popularly called Dual-Domain Image denoising (DDID) and was given in 2013 [172]. This work was further extended as Dual Domain Filtering (DDF) by employing wavelet shrinkage and bilateral filtering as robust noise estimators in two different domains. The extension of DDID was employed by introducing guided filtering using a second guiding image. This method outperformed the various others competitive methods such BM3D, NLM and PLOW with much lower computational complexity [173].

PID (Progressive Image Denoising) is rather a quite recent and widely acclaimed work of Knaus and Zwicker, in which the connection between wavelet shrinkage and robust estimation of differential noise was exploited. However this method required higher number of iterations via guided filtering as compared to its prior version named DDF. Contrary to DDID it does not require to differentiate between noisy image and guided image. This method extends the vision of rejecting bias by introducing signals as outliers in the frequency domain employing, wavelet shrinkage as robust noise estimation. And finally these methods employ DA (Deterministic annealing) and SA (Stimulated Annealing). This method unifies the spatial domain and frequency domain resulting in a much simpler but effective denoising algorithm competitive with most state-of-the-art techniques in case of synthetic images [3]. Further a much faster version of DDID (approximately three times) was designed by Pierazzo et al. [174]. This method was shown to mitigate the artefacts created by DDID in the frequency domain.

A progressive dual-domain filter was proposed in order to enhance and denoise low quality optical remote sensing images. Initially high contrast images, obtained with the help of bilateral filter were enhanced using histogram modification method. Further the low contrast image approximation structures were captured using STFT and enhanced by employing a corrective parameter which was adaptive [175]. These continuous variations in the dual domain image denoising strongly witness the influence of hybrid domain in image denoising.

The isotropic diffusion (ID) and PM (Perona–Malik Diffusion) has been widely employed independent methods in denoising and have shown to give good performance. However ID causes edge blurring, whereas PM introduces staircase effect leading to loss of the local details in image. Therefore in [176], the authors proposed a hybrid image denoising model by improving the ID and PM models by employing second order directional derivative (directional Laplacian). The integrated approach of modified PM and ID model resulted in suppression of noise in homogenous regions along with preservation of fine details and textures besides mitigating the stair case effects.

In an extremely interesting work, a recent method is given by Shahdoosti and Hazavei in which compressed sensing, block-matching and sparse representation based learned dictionaries are coupled together in bandlet domain. In this method first noisy image is transformed in bandlet domain and noise coefficients are threshold using hard thresholding. This method assumes that image is sparsely represented over a dictionary rather than basis elements. Also block matching technique is employed to represent dictionary elements such that identical sparse vectors are used for all patches in a group. The given method was tested on eight well known grey scale images and it outperformed several state-of-the-art denoising techniques [251].

For removal of Gaussian Noise from images, Yu et al. proposed an algorithm by employing a trivariate wavelet shrinkage in the spatial do-

main bilateral filtering. This method was motivated from the fact that wavelet based methods are prone to low frequency noise and ringing artefacts and spatial domain methods are computationally expensive. The wavelet coefficients were modelled as trivariate Gaussian shrinkage based on maximum a priori estimator. The overall method resulted in cost cutting and joint bilateral filtering by using wavelet transform [177].

In another work given in [178], learning based algorithm is embedded with curvelet transform tiling. The optimal tiling for best representation of a given class of images was decided on the basis of maximizing PSNR and minimizing MSE by variation of parameters like angular and scale decompositions. The method was shown to have robust performance as compared to curvelet transform and NLM filter. Another method under hybrid domain methods is a combination of PDE and generalised cross validation (GCV) theory in tetrolet transform. The image was decomposed using tetrolet transform and an adaptive threshold was constructed using the GCV theory. The method has shown to give outstanding performance in terms of PSNR values [179].

The well-known BM3D algorithm for denoising is no less example of hybridisation of spatial and transform domain filtering. Inspired by this idea of clustering 3D groups of similar patches, brushlet thresholding was employed to remove the noise in the frequency domain. This method depicted improved visual quality of the ultrasound images as compared to BM3D alone [180].

An intriguing approach of integrating non-local total variation (NLTV) and BM3D is given in [181] for image denoising. The method obtained a pre-denoised image with the help of BM3D. Further while constructing the fidelity term of NLTV model, the pre-processed image was put to use for formulating energy functions and weighing functions in the regularisation term. The energy compaction equations were solved using split Bregman algorithm. The proposed method demonstrated exceptionally high visual quality with removal of artefacts and outperformed NLTV and BM3D.

It has been discussed earlier that TV models are known to have given good image smoothing properties; however they tend to propel staircase effect, by overly smoothing the flat areas. Also NLM filter, shows good performance in reconstructing low frequency smooth regions but fail at edges as there are not enough similar patches which leads prevalence of noise around edges (rare patch effect). This method employs a non-local data fidelity term instead of non-local regularisation term. The proposed method was able to perform efficient denoising free form staircase and rare patch effect [182].

Further ahead, in literature, combining wavelet or curvelet domain with TV and diffusion methods was an extremely common approach to address the issue of Gibbs phenomena. In a study given in [183], authors established that replacing these representations with Shear let transform is much better approach when combined with total variation minimisation for removal of artefacts and noise with preservation of edges. This method is one other exemplary instance of hybrid domain methods. In one other method, local Weiner filtering with the help of elliptic directional windows is employed in differently oriented sub-bands of wavelet representation [184]. A hybrid denoising method based on replacing of 2-D separable wavelet transform with Bandlet transform in NSCT domain is given in [185]. The proposed method was employed to remove Gaussian and Poisson noise [185]. Another space-transform domain image denoising includes combining of non-local mean algorithm with hard thresholding in the wavelet domain. The resultant images are processed with guided filter in order to get rid of the ringing artefacts introduced at the first stage [186]. Besides this a band-specific shearlet based image denoising technique is given in [187] which employ different thresholding techniques in different sub-bands containing varied amount of noise. In a significant piece of work, Candes et al. proposed to combine the total variation minimisation model with curvelet and ridgelet expansions in which sub-bands coefficients are approximately threshold [188]. A lot of authors have rather employed the total variation minimisation schemes in combination with other transform domain representations. In this category of image denoising methods, [189] is another work given by Durand and Forment, in which after the decomposition of the noisy image into the sub-bands, the total variation minimisation is employed in order to threshold the noisy coefficients.

In a rather recent piece of work, in order to address the issue of Gibbs-like artefacts while improving PSNR, Shahdoosti et al. has proposed an image denoising algorithm based on ripplet formulation of the total variation minimisation [190]. In their work smooth and textured regions were obtained using twin support vector machines. Instead of minimisation of variation of noisy images, this method tried to minimise variation in the image obtained from non-textured regions of the ripplet sub-bands. Besides this another work proposes to combine the non-linear anisotropic diffusion with curvelet shrinkage [191]. In this method of post shrinkage via ripplet transform, the reconstructed image was processed with total variation diffusion by which only high frequency insignificant coefficients were suppressed with the help of a constrained projection. By combining shift invariance of DTCWT and directional sensitivity of ridgelet transform a new hybrid approach was presented in order to suppress the AWGN [192].

In 2013 Shreyamsha proposed a denoising scheme based on Gaussian/Bilateral filtering and noise thresholding. In this scheme the idea of method noise has been presented. The method noise here refers to the difference between original image and image denoised by certain algorithm. The difference in itself is noise and is called method noise. In GBFMT (Gaussian/Bilateral Filter Method Noise Thresholding), a noisy image is assumed to be corrupted with Gaussian white additive noise with zero mean and a known variance which can be varied for testing at high and low noise values. The corrupted image is filtered using bilateral filtering and the residual image so obtained is thresholded in the wavelet domain using hard thresholding [193]. In a similar work bilateral filter was replaced by non-local filter to exploit the concept of method noise thresholding (NLMNT) [194]. Since then many different combinations have been tried in context of method noise thresholding [195]. Similarly in [196] the concept of denoising has been extended to hybridisation of Weiner filter method and Shear let transform. The multi-scale and multi-resolution representations have led to the improvement in the PSNR and MSE values. Besides this a lot many hybridisation image denoising techniques have been employed using sparse representation and dictionary learning methods. In order to abridge higher level comprehension of these methods, some of the hybrid methods are briefed under sparse representation based methods and some are listed under recent trends in image denoising.

Through extensive literature survey it has been observed that properties like multi-resolution, sparsity and edge detection have made transform methods very popular methods for denoising. However, these methods cannot represent smooth transitions very well and tend to lose on intricate image details and features. On the other hand, edge preserving spatial domain filters tends to retain noise features and artefacts. To overcome such issues various improvisations are being made by exploiting the properties of both spatial and transform domain filters. These combinational methods with intuitive adaptive thresholding hold a lot of scope in context of efficient image denoising in terms of both objective and subjective evaluation.

6. Sparse representation and dictionary learning methods

The redundant and sparse representations have long been the basic dynamic foundation for researchers in image denoising. One major reason for requirement of redundancy arises from its ability to induce shift invariance. The concepts of sparsity and redundancy have been widely made applicable by use of various multi-scale and directional transforms like curvelets, ridgelets, brushlets, and directionlet. On the very same approach of introducing sparsity and redundancy in the image representation methods, the sparse representation based dictionary learning methods were designed. This approach varied in the manner that it possesses an example based restoration. In this category, the

common successful approach is BLS-GSM (Bayes Least Square Gaussian Scale Mixture) which formulates as image denoising problems as inverse problems using Bayesian minimisation and hence need of an image-prior based on entropy, smoothness or sparsity in some transform domain becomes indispensible. While these methods tend to derive their image estimate from a mathematical expression representative of the image prior, these example based methods tend to derive their image prior from pre-existing images. While incorporating the idea of prior learning in sparse and redundant representations, most important thing is dictionary which provides a learned set of parameters. These methods eliminate need for fixed basis functions as is the case in curvelet or contourlets transforms [197].

The introduction of orthogonal matching pursuit (OMP) laid the foundation of commencement of image denoising problem as a direct sparse decomposition technique with redundant dictionaries [198–200]. In this category of methods quality of denoised image largely depends on type of dictionary which can either be learned from patches of noise-free images or adaptive learning of noisy image itself.

In other words it can be stated that these methods adapt a dictionary. These methods are usually employed as patch based denoising methods and averaging is carried out in case of overlapping patches. The dictionary constructed from learned examples contains atoms which can be understood as approximating the patches using a sparse linear combination of atoms.

In this section we review only some of the major dictionary learning methods which are K-Singular Value decomposition (K-SVD), Expected Patch Log Likelihood (EPLL), K-Locally Learned Dictionaries (K-LLD), Learned Simultaneous Sparse Coding (LSSC), Patch based Locally Optimal Wiener (PLOW), Trained Filters, Metric Steering Kernel Regression (SKR) filters, Convolution Sparse Representation (CSR), Locally adaptive Regression Kernels (LARKs) and various others [4].

Elad et al. with an aim to embed local over-complete dictionaries in Bayesian global estimator proposed K-SVD as an iterative denoising method which adaptively learns from the entire range of pixels of noisy image. Each of the iteration comprises of employing OMP to estimate coefficients for each patch (an initial dictionary is used for computing sparse approximations of all patches) and updating the dictionary using singular value decomposition for one column at a time. Each patch can be estimated from a series of patches from the dictionary. K-SVD simulated the era of denoising with learned dictionaries which restores image information using a more adaptive model. However this method still needs improvement in case of large patches due to its computational burden and limited size of the dictionary [197]. In another method known as K-LLD, in order to avoid selection of patches with different structures to be mistaken as similar matches, SKR classification is exploited [201].

The LSSC method is quite similar to K-SVD where similar patches are denoised using same sparse decomposition. This method exploits self-similarity criteria of images in a non-local manner [202]. The motivation behind this method was that, in case of K-SVD slight change in input can lead to significant change in dictionary atoms which is undesirable. Therefore LSSC was proposed as an improvisation to K-SVD by exploiting similar patches in a non-local manner. This method certainly sped up the process of search of atoms in an unstructured dictionary. This method has found to be giving absolute state-of-the-art denoising performance but it has a high computational complexity. Although LSSC has achieved satisfactory denoising results by employing clustering in sparse decompositions however its performance largely depends on initial dictionary which has to be trained offline with high quality images. The CSR denoising method was able to achieve much higher performance than LSSC along with lower complexity [4].

Another class of methods which exploit dictionaries and are similar to spatial domain filtering are trained filters. These filters are either nonparametric or semi-parametric and have given excellent denoising performance. These filters are similar to spatial domain filters except for the fact that they carry out a non-parametric process in order to calculate weights of filter by employing offline training using abundance of images. These methods ensure an adaptive image restoration process as fixed kernel coefficients are replaced by trained coefficients. For each noisy patch, same classification is employed and filtered coefficients are obtained from the look up table (LUT) constructed in previous training process. These methods are computationally fast as LUT has to be trained only once offline. Some of the major examples of these methods are M-SKR (metric- steering kernel regression) [203] and K-SPR (semi-parametric regularisation) [204].

In image denoising approaches the major issues lies in extraction of appropriate features capable of determining the relation between adjacent coefficients. Most of denoising algorithms are not able to effectively suffice this purpose. DNN is class of machine learning algorithms which proposes to learn flexible models devoid of any pre-requisite assumptions and data distribution. In this class of image denoising algorithms Shahdoosti and Remza proposed a novel image denoising algorithm based on CNN model and shearlet transform which can perform much better in terms feature classification then various existing state-of-the-art algorithms for denoising [244]. Support vector machines (SVMs) is an another important dictionary learning algorithm is introduced by Vapnik [245] in 1998 and have been extensively applied in function estimation and classification problems. They are developed in context of statistical learning theory and aimed at structural risk minimisation. Least square support vector machine is a computationally effective and faster version of SVM owing to obtaining of solutions to linear equations instead of quadratic programming.

In a work proposed by Wang et al., in 2010 LS-SVM were combined with undecimated wavelet transform in order to denoise images while preserving the edges [246]. Then again in 2013, he employed LS-SVM to classify NSCT coefficients into noisy coefficients and edge coefficients. With the help of LS-SVM, NSCT coefficients at different scales are classified into smooth regions and areas containing texture regions. For the maximisation of texture information, edge coefficients are left unmodified, whereas noisy coefficients are threshold using Bayesian threshold. The combination of NSCT (multi-scale, multi-direction, shift-invariant) and LS-SVM, this method was able to outperform methods like Probshrink and BLS-GSM objectively [247].

The major problem with the dictionary learning methods is their computational complexity as they require several iterations and have unstructured dictionaries which makes them unsuitable for real time applications [4].

7. Recent trends in image denoising

From local filtering to non-local self-similarity exploration and from adaptive thresholding to learned dictionaries, the field of denoising has evolved in an inexplicable manner. The study of these large numbers of methods of reported in literature is overwhelming, it is nearly impossible to discuss each of these pioneering methods. In this arriving section of literature we try to embark upon recent trends and main areas of plan of action witnessed in the field of denoising.

In 2010, in [205] Chatterjee and Milanfar eluded the inherent question of final bounds of denoising. With so many methods pouring in, one ponders is there a final limit to denoising. In this work the authors estimated lower bound on image denoising in terms of MSE. It was depicted that despite the phenomenal growth in recent techniques there is still a room for improvement and the field of denoising is not dead yet.

Further in 2012, these authors proposed the PLOW filter, motivated by the statistical observations listed in the above mentioned work. They employed a patch based redundant Weiner filter where parametric model was designed considering both geometric and photometric similarity. The method has shown to have given an extremely high visual quality of images, outperforming the famous BM3D [206].

Following the lead of intense focus on hybrid domain image denoising Zhang et al. in 2013, proposed an image restoration method

based on joint statistical modelling (JSM) in the space-transform domain. This method derives a novel method for exploiting both local smoothness and nonlocal self-similarity to generate high fidelity image reconstruction. JSM is hybrid platform which generates a powerful amalgam of regularisation based (total variation models) and non-local methods. The effectiveness of proposed algorithm is demonstrated with the help of rigorous experiments for Gaussian as well as impulse noise [207].

The performance of patch based methods for instance BM3D is inherently limited due to insufficient number of similar patches and due to increasing complexity in case of larger images. Therefore, Talabi and Milanfar in 2014, contributed significantly to this line of research by designing global image denoising filter (GLIDE) which takes into account all similar or informative parts of an image. In this method, after the pre-filtering of the input image, a small number of pixels are sampled and are fed as inputs to Nystrom equation method. The final image is estimated by thresholding the eigen values of filter matrix [208].

Another work based on hybridisation of interscale and intrascale dependencies in Direction let domain was proposed in [209] in 2015. The image is first transformed into the Direction let coefficients and then the coefficients are modelled using bivariate PDF to exploit the in-scale and cross-scale dependence. This method gives far better performance than bivariate shrinkage in the wavelet domain.

Another motivating approach towards image denoising is found in [210]. This work on addressed a natural question to adaptively replace images using prior of an image scene. It exploits the external and internal correlations thereby clustering the similar patches of noisy and true images available on web. This method suggested a two-stage image denoising using graph based optimisation and transform domain filtering. The method was able to outperform BM3D over wide range of noise values. Recently in 2018, some authors studied the image denoising problem as deduction of a non-linear diffusion model into a cross-diffusion problem [211].

Patch based processing has emerged as the core idea in image restoration problems. This process consists of decomposing the image into similar overlapping patches and restoring them by plain averaging. This method, besides image denoising, has been widely employed in areas of image inpainting and deblurring. However, major flaw in these patch based methods is imposition of prior on intermediate results, rather than processing the final outcome. The EPLL was designed in order to address this issue. In a recent work in 2016, a method to extend and improve EPLL has been proposed [212]. This method not only imposes prior on final image but constructs a multi-scale prior for different scale patches extracted from the input image. This method achieved considerable improvement over the major state-of-the-art image restoration algorithms including BM3D.

Another optimised pixel based weighted averaging approach has been designed for patch based image denoising in [213]. The various image denoising algorithms restores clean image while leaving behind the residual. The residual image left behind should contain uncorrelated contaminating noise, but it contains some remnants from the clean image as well. In a patch based image denoising algorithm, a regularisation approach was proposed to render the residual patches as uncorrelated as possible. With this the authors generated an analytical solution for sparse coding and a new online dictionary [214].

It has been discussed that NSST has depicted exceptional performance in terms of multi-resolution analysis and shift invariance. This approach was further exemplified in a work, which constructs a feature vector based spatial regularity in NSST domain using SUnSAL. The SUnSAL classifier distinguishes between edge based and noise based coefficients which are threshold using Bayes Shrink method [215].

In 2016, pioneers, Shahdoosti and Khayat presented an innovative and effective method for denoising of images by the hybridisation of NSST, SUnSAL and ansitropic diffusion [249]. To obtain several scale and directional components, NSST is applied to noisy image. Then SUnSAL classifier is employed in order to distinguish noisy and edge related coefficients. The prime idea is to minimise energy of noise coefficients using diffusion equations. The noise free image obtained from anisotropic diffusion filtered coefficients and textured coefficients (which remained unchanged) is evidently artefact free, constitutes edge information outperforming methods like BLS-GSM, Probshrink and work given in [4].

Over wide range of seamless efforts for image denoising the major underlying approach is edge recovery. This approach has been worked upon in a recent work in [216]. This method consists of an edge detection and fusion model based on anisotropic diffusion model. A denoised image is obtained and edges of denoised images are replaced by edges extracted by various edge detection and regularisation based methods.

Redefining the boundaries of well-known bilateral filter, a new boundary preserving filter named co-occurrence filter has been recently designed by weighing pixels which occur frequently rather than considering the Gaussian range. The low-occurring pixels on other hand would have lower weights in the co-occurrence matrix. This approach resulted in boundary preservation besides recovery of edges. This work introduced the definition of co-occurrence matrix and its application to denoising and pioneers a direction for designing various new methods which learns co-occurrence matrix [217].

In case of stochastic approach towards image denoising, an alternative approach based on Markov-Chain Monte Carlo sampling has been given in literature. This method dynamically adapts itself to image characteristics and noise statistics providing high performance in terms of image denoising while exhibiting lower complexity. This method formulates the problem on the basis of AWGN, computing prior from the noisy image [218]. This additive image denoising problem is formulated as a Bayesian least squares problem, where the goal is to estimate the denoised image given the noisy image as the measurement and an estimated posterior. The posterior is estimated using a nonparametric importance-weighted Markov-Chain Monte Carlo sampling approach based on an adaptive Geman–McClure objective function [219]. Some approaches improvisations of this method have been given in [220–223].

Out of many approaches towards denoising the primary and current focus of researchers is exploiting non-local self-similarity, combining them with other structure preserving techniques or improvising the already existing self-similarity based approaches like BM3D and NLM.

A new patch matching method employing nearest neighbour collaborative filtering (NN-CF) is demonstrated in [254]. This work presents two variations of CF based patch making criterion. By employing new patch making approach to various patch matching based denoising methods like BM3D and NLM, authors depicted robustness to patch making and new formulation to integrate external and internal denoising. An image denoising approach using generalised Cauchy filter is illustrated in [255] by using the particle swarm optimisation.

A detail preserving cluster wise denoising method providing a stable noise level estimation is given in [256]. This method finely clusters non local patches with diverse features using adaptive clustering method. The noise in this technique is removed by employing three steps namely progressive PCA thresholding; by SVD thresholding based on Marchenko–Pastur law and coefficient wise PCA domain LMMSE filtering.

Out of numerous improvisations of NLM denoising a work that effectively exploits Affine Invariant Self Similarity (AV-SS) prevalent in natural images is given in [257]. This method enables selection of more similar patches by adapting to different size and shapes of patches by performing appropriate patch comparison using AV-SS. It has been shown that this method achieves top-tier performance in terms of quantitative analysis.

In contrast to assuming fixed noise problem, targeting real time denoising problems authors in [258], proposed a blind denoising method which permits estimation of noise from a single image which is signal and frequency dependent. The efficient performance of the method

is demonstrated on real JPEG images where formation model was unknown and it depicted state-of the-art performance.

Besides BM3D, PLOW, LPG-PCA and Shape adaptive SVD, a method employing non-local self-similarity and low rank approximation is given in [259] and is proposed as a modification of patch based methods. Initially similar patches are classified using block matching technique. Then SVD is employed to factorise each group of similar patches and only singular vector and largest singular values are selected. The final image is generated by aggregating processed patches. This method besides effectively removing noise provided optimal energy compaction and is shown to be computationally efficient.

Having attained a tremendous progress image representation, limiting the computational burden of algorithm became the principal niche in denoising. Non-local approach has been evidently depicted as most efficient criterion for denoising. However it was found that some of the nearest neighbour approaches introduced bias in results. Therefore Frosio and Kautz proposed a statistical neighbour approach in 2019 which depicted that only fewer statistical nearest neighbours are required to generate images of better quality both in case of grey scale and coloured images [260].

In another work outlying limiting performance of BM3D on seismic data, authors designed an improvisation to BM3D by including a factor, local similarity indicator, between noisy signal and true signal. This method prevented drainage of signal energy by recovering it and adding back to the denoised result. This intriguing approach has been proven to ensure fidelity of structure representation in final results [261].

Palmprint recognition is a reliable identification method in biometrics and is an important image processing application. However presence of noise in low resolution palmprints significantly degrades biometrics methods based on them. In [262] a generative adversial network method has been proposed to mitigate presence of noise which reduced error rate from 10.841% to 1.532% after denoising [262]. In order to preserve complementary contextual information and to reconstruct fine patterns and textures, Li et al. proposed multi-scale gated fusion network (MGFN) as an improvisation of deep convolutional network for image denoising. This method was able to put forth a simple but effective loss function to train the network [263].

With a means to remove noise from different spectral bands of hyperspectral images a denoising problem is presented as band fusion problem. In [264], first a target noisy band is selected and in order to denoise it, superior quality bands related to it are fused together. The method has been depicted to shown as superior to various existing HSI denoising methods in terms of visual and objective analysis.

With an objective to harness the attributes of various techniques a hybridised denoised method is proposed in [265], which integrates TV regularisation and group sparsity related to self-similarity of image blocks. This concatenated technique consists of four steps i.e. matching of blocks, updating of basis vectors, sparsity regularisation and TV smoothing. In contrast to various other which employ fixed basis functions this method used updated local basis vectors derived from SVD to induce sparsity. This method has been shown to be outperforming methods such BM3D itself [265]. Another recently published denoising method combines the nuclear norm, TV norm, and L1 norm, which facilitates to exploit the low- rank property of natural images, enhancing smoothness and removing sparse noise [266].

While exploiting the concept of method noise in deep learning, a deep feed forwarding denoising convolutional neural network based method has been proposed in [267] to denoise medical images. Unlike other methods, this method carries outs learning of residual noise from noisy image where denoised images are obtained by subtracting learned residual from noisy image. Experimental results reveal that this approach has better performance than some other methods. Another residual correlation based regularisation for image denoising method is given [268]. In this method analytical and a novel online dictionary update is performed generating solution for sparse coding i.e. atom selection and coefficient calculation. This method exploits the approach of residual image containing the remnants of clean image.

8. Applications of image denoising

Image denoising is fundamental image processing problem which finds applications in various fields of technology. With voluptuous growth of science and technology, scope of human action is expanding, creating higher demand of information which results in high resolution images. The noise disturbs this information and image denoising becomes increasingly important for subsequent image analysis. Image denoising not only generates visually pleasant images, but also “uncovers” the corrupted information pixel for further image analysis and information extraction. The process of image denoising has left no area of application unaffected which includes medical imaging, biometrics, remote sensing, HVS (Human Visual System), military surveillance and infrared image denoising.

8.1. Medical image denoising

Magnetic resonance imaging is used for the visualisation of soft tissues of various organs of the body. MRI’s are invariably corrupted with various noise factors like variable field strength, radio frequency pulses or due to receiver bandwidth. In [223], motivated by radon transform, a wavelet based multi-scale MRI noise suppression method has been given. The noise present was modelled as the Gaussian Noise and translational invariant wavelet transform was used to decompose noisy MR image into multiple scales. The noise was estimated using Rician noise distribution and final denoised image was obtained by using inverse Radon transform. The method was evaluated using Phantom, simulation and human brain MR images. The results for simulation brain images are depicted in [223] have been given Fig. 4.

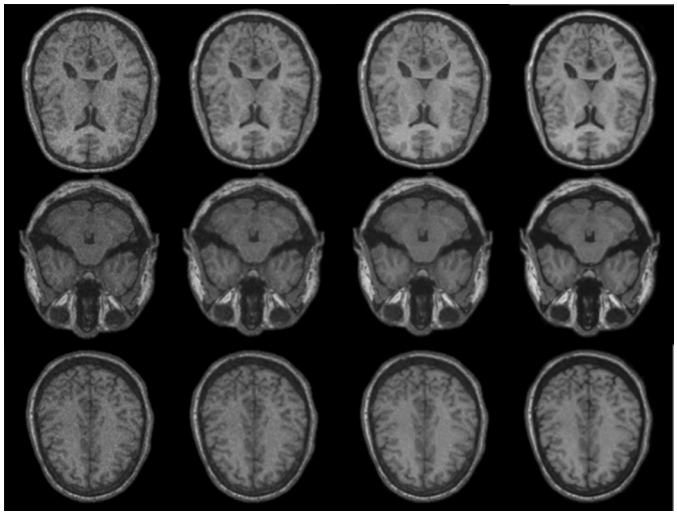

Fig. 4. Noisy data set for simulation brain images (first column); denoised images for Weiner filter (second column); wavelet (third column); (fourth column) [223].

The Computed Tomography is yet another form of imaging technologies which suffers from image degradation due to noise owing to the inherent problem of hardware restriction. Therefore a lot of methods have been designed in order denoise the CT images [269]. In [224], authors propose to preserve edges, corners, textures and sharp edges with the help of Tetrolet transform based locally shrinkage rule to denoise high frequency coefficients. The proposed method in this work revealed good performance in terms of noise removal and structural preservation. The method has been evaluated in terms of MSE and PSNR and the visual

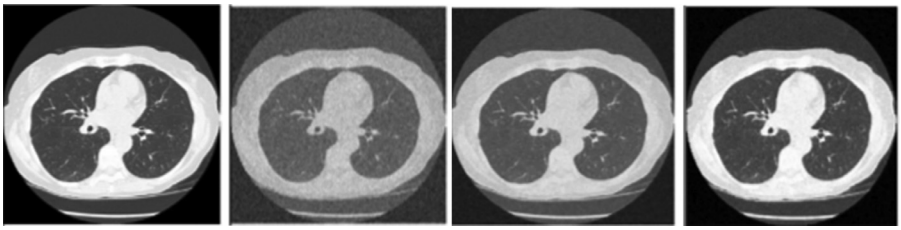

Fig. 5. Original CT image; noisy image at $\sigma = 20$; denoised image with biorthogonal wavelet thresholding; tetrolet shrinkage rule (Left to Right) [224].

results for one of the data sets are reproduced from 224 and are given in Fig. 5.

8.2. Underwater image denoising

Underwater laser imaging has been rapidly employed in order to acquire information regarding underwater regions. However the image resolution degrades due to the corruption of acquisitions with noise due to watery medium. The major reasons of occurrence of noise in underwater imaging is attenuation and back scattering of laser waves by water resulting in poor image quality. A work in [225], analyses the application of MATLAB encoded image denoising algorithms for underwater images. The underwater source images have been restored using median filter, average filter and a hybrid method based on wavelet transform. The results for grey scale images have been given in Fig. 6.

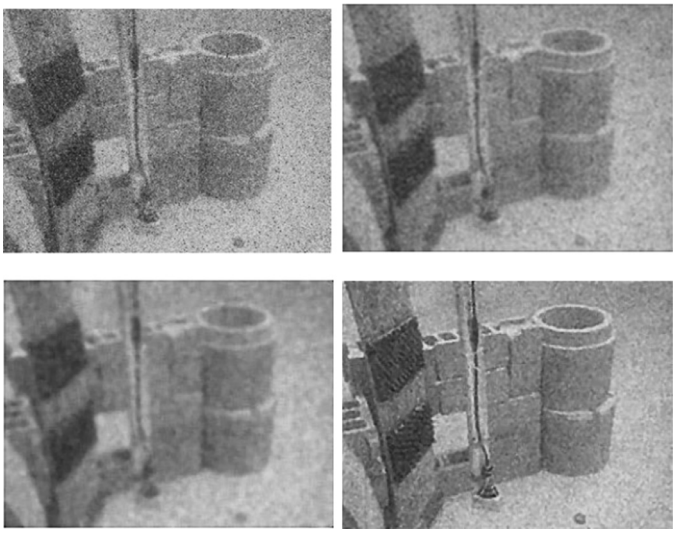

Fig. 6. Underwater noisy image; denoised images using averaging filter (first row L-R); median filter; wavelet method (second row L-R) [225].

8.3. Remote sensing-image denoising

Remotely sensed satellite images are invariably corrupted with noise either during signal transmission or acquisition. The removal of high frequency noisy components becomes inevitable in order to improve the visual appearance, extract and analyse the satellite imaging information. The exploitation of directional correlation in spatial and transform domain has been a common idea for image denoising. However the success of these methods largely depends on the amount of orientation correlation measurement. Therefore in [226], author's proposed a hybrid directional lifting scheme for denoising of satellite images. This method employed pixel classification and orientation estimation both. To deal with denoising of the land remote-sensing satellite images (LANDSAT). The results for Kochi dataset have been depicted in Fig. 7.

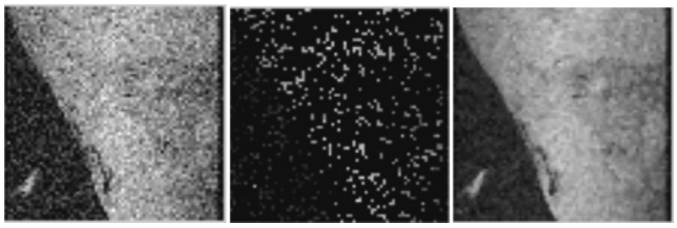

Fig. 7. Kochi LANDSAT image; direction estimation; denoised image with HDL (left-right) [226].

8.4. SAR image denoising

To despeckle the SAR (synthetic aperture radar) images, is one of the important research problems in the field of remote sensing. This imaging technology finds wide range of applications in military surveillance. To overcome the limitations of spare-representation based image denoising methods, a novel image denoising algorithm based on sparse representation in shearlet domain and continuous cycle spinning has been given in [227]. After the decomposition of the noisy image using shearlet transform, an optimal denoising model was constructed using cycle spinning theory in the sparse representational domain. The method not only improved PSNR but also considerably improved the visual results which have been shown in Fig. 8 for reader comprehension.

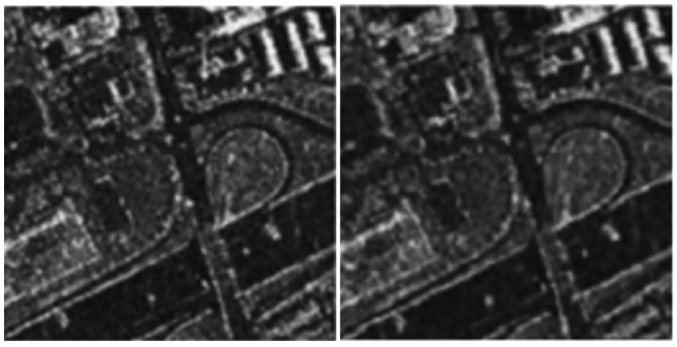

Fig. 8. Noisy SAR image; SAR denoised image [227].

8.5. Infrared image denoising

The infrared imaging technology is employed in order to capture the thermal radiation emitted by the metal objects and it finds extensive application in the night mode imaging and military surveillance. The infrared images are corrupted with Poisson noise due to low-light conditions and atmospheric perturbations. In [228] Anscombe transform has been used in order generalise the Poisson distribution into Gaussian distribution and to denoise the infrared images. The method was comprised of total variation regularisation based improvisation in the wavelet domain in order to remove heavy Poisson noise from low-light infrared images and same is being depicted in Fig. 9.

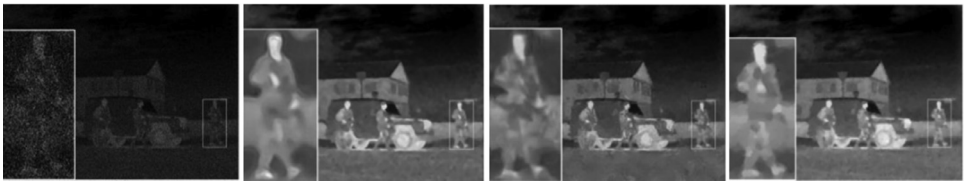

Fig. 9. Noisy INFRARED image; denoised results by NLM; BM3D; method given in [228].

9. Experimental setup

In order to present a performance comparison various existing popular denoising techniques have been analysed on the basis of (i) Visual Analysis; (ii) Objective Analysis. The source image dataset is chosen as House Image; MR (Magnetic Resonance Image) Image; PAN (Panchromatic Image) [229]; so as to reflect adequate amount of diversity in the complexity of image content. These images include the natural image, medical image and remote sensing image. The experiments are carried out on Intel® Core(TM) i5-72000U- CPU @2.50 GHz processor and 8 GB memory using MATLAB^{TM}. The resolution of these images is 256×256.

9.1. Image contamination model

The AWGN noise of standard deviation $\sigma = 10, 20, 30, 40$ and 50 has been added to source images, so as to check consistency in performance of denoising methods.

9.2. Objective evaluation metrics

In order to quantify numerically the performance of a denoising algorithm in terms of its extent of noise removal and ability of edge preservation various objective metrics have been devised and are being readily employed in literature. The major performance metrics employed in field of denoising are (in the order of their popularity): PSNR (Peak Signal-to-Noise Ratio), MSE (Mean Square Error) SSIM (structural similarity), VIF (Visual Image Fidelity), Entropy, AD (Average difference), MD (Maximum Difference), Normalised Absolute Error (NAE) and Normalised Correlation [4,230–242] The mathematical formulas of these metrics along with their significance is summarised in Table 3.

One needs to be aware of the fact that PSNR improvement or any other objective metric's comparative increment is not the ultimate objective of image denoising. The primary goal of image denoising is to generate denoised images with high visual quality, better edge preservation and pattern identification which form an integral part of the Human Visual System (HVS). Studies related to HVS reveals that human eye can hardly comprehend slight differences in orientations, scales and texture surface. However it is a lot more sensitive to sharp contours, deformed faces, words and symbols, artefacts in smooth regions. Also visual analysis is very important for comparison of different image denoising methods, because human is usually the end user Therefore the visual analysis is weighed more for discussion of any type of image denoising technique.

In this section, the objective evaluation metric Peak Signal-Noise Ratio has been used for performance comparison amongst various methods. The experimental analysis section is vital to comprehensive strength of the review. We have selected methods from various classes of denoising. All the methods chosen for experimentation are tested on our own versatile dataset (medical image, natural image and remote sensing image) for additive white Gaussian Noise. In this section we analyse the actual performance of various methods on varying noise levels and different types of images. A lot of methods have been experimentally demonstrated in literature on variety of dataset. In this manuscript we have tried to choose some representative methods from each class of denoising. Some of these methods have been reviewed in the text above and some are cited for reader reference as discussing all the methods is nearly impossible. Moreover these methods will be covered at a stretch in terms of objective and subjective evaluation. Under the category of spatial domain filters methods employed are SBF [48], RBF [54], WBF [54], NLM [2], bitonic filter [75], Diffusion filter assortments (DF(A-B)) [242], Gaussian filter [8], Median Filter

Table 3
Quantitative performance metrics.

Metric	Mathematical formula	Significance
MSE	$\frac{\sum M,N[I_1(m,n)-I_2(m,n)]^2}{M\times N}$	Lower MSE depicts better image restoration
PSNR	$10log_{10}(\frac{R^2}{MSE})$, $R = 255$	Higher PSNR means better quality of the denoised image; frequently applied image quality metric in literature
SSIM (x, y)	$\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2}{(\mu_x^2+\mu_y^2+C_1)\ (\sigma_x^2+\sigma_y^2+C_2)}$ C_1 and C_2 are constants, μ_x *and* μ_y are the mean intensities, σ is the standard deviation	Metric correlated to human perception; high SSIM indicates more restoration of original information
VIF	$\frac{\sum_{k\in subbands} I(D^{m,k};T^{m,\ k}\|s^{m,k}}{\sum_{k\in subbands} I(D^{m,k};R^{m,\ k}\|s^{m,k}}$ where $D^{m,\ k}$ represents m elements of the random field that describes the coefficients from subband k, $T^{m,k}$ and $R^{m,k}$ depicts the visual signal at the output of the HVS model for reference and test images	Information fidelity criterion; quantifies the Shannon information which is shared between the distorted and the reference image, fidelity aspect correlating with visual quality
Entropy (EN)	SE $=\sum_i A_i^2\log(A_i^2)$; EN= SE(A)-SE(B), where SE stands for Shannon entropy	EN is a statistical computation of randomness which conveys the amount of texture information;
AD	$\frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(A(x,y)-B(x,y)}{M\times N}$	Gives the average difference between the clean and denoised image
MD	$MD = MAX\|A(x,y) - B(x,y)\|$	Lower value of MD means better image quality
NAE	$NAE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(A(x,y)-B(x,y))}{\sum_{x=1}^{M}\sum_{y=1}^{N}(A(x,y)}$	Gives the normalized rate of error; lower NAE gives better denoised images
NCC	$\frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(A(x,y)\times(B(x,y)}{\sum_{x=1}^{M}\sum_{y=1}^{N}(A(x,y)^2}$	Depicts the amount of correlation between the clean and denoised image

Assumption: A and B represent the clean and denoised images respectively and $M\times N$, is the size of the image.

Table 4
Performance comparison of various state-of-the-art denoising techniques.

Denoising methods	PSNR														
	MRI					HOUSE					PAN				
	$\sigma = 10$	20	30	40	50	10	20	30	40	50	10	20	30	40	50
SBF	32.18	25.78	20.79	17.44	15.13	33.23	26.11	20.95	17.60	15.14	28.62	23.44	19.60	16.91	14.74
RBF	30.20	29.59	28.54	27.13	25.63	31.75	30.98	29.69	28.19	26.37	19.54	19.37	19.14	18.84	18.40
WBF	32.55	30.18	28.64	27.15	25.64	33.64	31.30	29.74	28.19	26.19	28.66	24.03	21.56	20.02	19.22
INLM	35.01	31.18	29.37	27.79	26.49	35.63	32.88	30.93	29.44	28.04	28.32	24.10	21.61	19.87	18.73
Bitonic	29.39	27.27	25.25	23.25	21.70	31.89	29.59	27.29	25.51	24.00	20.16	19.88	19.48	16.91	14.74
DF(A)	25.33	24.57	23.65	22.51	21.46	24.66	24.47	24.13	23.73	23.28	17.12	17.02	16.91	16.76	16.60
DF(B)	23.08	22.64	22.07	21.33	20.56	22.17	22.12	22.00	21.82	21.63	15.67	15.62	15.55	15.46	15.36
DF(C)	28.40	26.98	25.48	23.93	22.62	22.21	21.99	21.64	21.12	20.63	17.84	17.69	17.45	17.22	16.90
DF(D)	28.45	26.84	25.16	23.44	22.01	21.75	21.59	21.32	20.98	20.61	18.38	18.18	17.92	17.59	17.26
vi-hard	29.11	25.26	22.86	21.20	19.92	30.69	27.58	25.96	24.53	23.45	23.14	19.12	17.15	16.02	15.32
vi-soft	26.14	22.76	20.82	19.48	21.53	28.05	25.23	23.67	22.45	21.64	19.79	16.85	15.45	14.60	14.03
GF	26.38	25.38	24.21	22.84	21.64	29.67	28.88	28.75	26.72	25.65	18.13	18.01	17.84	17.63	17.39
MF	24.35	23.74	22.96	21.98	21.04	27.80	27.48	26.93	26.39	25.70	16.92	16.84	16.74	16.61	16.46
SH3D	29.51	23.53	20.05	17.63	15.81	28.10	22.10	18.65	16.29	14.55	28.31	22.40	19.01	16.67	14.93
BM3D	35.84	32.25	30.28	28.63	27.77	36.71	33.77	32.08	30.64	29.69	29.32	24.62	22.25	20.70	19.51
SAPCA	35.82	32.30	30.29	28.89	27.78	37.01	33.89	32.12	30.74	29.52	29.62	24.95	22.25	20.70	19.51
SURELET	34.55	31.14	29.05	27.35	26.25	35.29	32.06	30.03	28.47	27.19	28.91	24.14	21.85	20.37	19.30
GBFMT	29.56	23.58	20.10	17.69	15.87	28.14	22.14	18.68	16.33	14.60	28.37	22.45	19.07	16.72	14.97
NLFMT	29.52	23.55	20.08	17.66	15.84	28.13	22.12	18.67	16.31	14.58	28.33	22.41	19.02	16.68	14.94
NSST	34.37	30.83	29.05	27.55	26.74	34.86	31.98	30.30	28.97	27.89	26.89	22.40	20.52	19.34	18.47
MRF	30.64	26.61	20.80	17.99	16.32	32.80	26.39	19.15	16.51	14.93	26.88	23.27	19.24	16.77	15.18
NLGRTV	24.12	23.71	20.14	22.54	22.27	26.87	26.62	25.48	25.24	25.67	17.43	17.26	19.00	16.65	16.73
LPGPCA	35.56	31.77	29.63	28.10	26.94	36.12	33.07	31.20	29.72	28.50	28.88	24.08	21.68	20.16	19.08
NCSR	31.48	30.43	27.73	25.00	18.26	36.78	33.86	32.11	30.52	29.05	24.85	24.49	21.99	201.7	18.83
Ker. Reg.	28.45	27.63	26.01	22.86	18.91	29.94	30.89	30.57	24.44	17.52	23.90	23.16	21.50	18.61	15.76
GFOE	33.77	28.91	25.90	23.29	21.46	34.01	30.10	27.19	24.83	22.57	23.09	22.09	20.92	20.05	19.27
PGPCA	34.43	30.14	26.43	23.97	22.30	35.18	32.00	30.13	28.88	27.56	28.95	23.88	21.27	19.64	18.50
PLPCA	34.54	30.37	26.65	24.27	22.60	35.81	32.39	30.46	28.88	27.76	29.20	24.05	21.44	19.82	18.65
PHPCA	34.47	30.33	26.65	24.27	22.54	35.67	32.38	30.39	29.08	27.84	29.11	23.96	21.36	19.67	18.55
TV1	28.71	27.04	25.96	24.99	24.21	31.78	29.93	28.49	27.39	26.47	19.45	18.94	18.39	17.85	17.41
Deep NN	34.19	30.63	29.43	27.68	26.45	36.02	32.96	31.78	30.11	29.61	29.30	23.67	21.56	19.63	19.37

[8,15], Gaussian Field of Experts (GFOE) [270], Total variation minimisation [276], non-local version of general relative total variation (NLGRTV) [271]; the different transform domain methods include Vi-hard [94], Vi-soft [95], SURELET [97], NSST [59, 144]. Under the category of statistical methods we have included Markov Random Fields (MRF) [163]. To illustrate an idea of different sparse representation and dictionary learning methods we have included Nonlocally Centralised Sparse Representation (NCSR) [272], Locally Adaptive Kernel Regression (LARK) [273], deep neural networks (NN) [274] in experimental analysis. Besides this some of the other representative methods include LPGPCA [153], BM3D-SH3D (joint image sharpening and denoising) [243], BM3D [151], BM3D-SAPCA [157], PGPCA (Patch based Global PCA), PLPCA (Patch based local PCA), PHPCA (Patch Based Hierarchical PCA) [275]. The methods employed under hybrid domain are: GBFMT [193], NLFMT [194]. The objective results for these methods in terms of Peak Signal to Noise ratio at low to high noise levels have been given in the Table 4. In Fig. 10, the visual results at standard deviation of 30 are being shown for House image only due to space constraint. The experimental results given in Table 4 and Fig. 10 above summarise the subjective and objective evaluation of various image denoising techniques.

9.3. Objective analysis

In order to quantify and analyse the performance of different domains of image denoising, we have simulated MR image, House image and PAN image with AWGN at standard deviation 10, 20, 30, 40 and 50. The noise from these images is removed using 31 image denoising methods representative of their domains and PSNR is calculated. The parameter settings for these methods have been selected as given by authors in reference papers. It is clearly evident in Table 4 that in case of MRI, BM3D and its Shape adaptive principal component analysis version gives highest value of PSNR. Same is the case for House image as PSNR for BM3D, SAPCA and NCSR is the highest. However in case of PAN image only SAPCA is able to give highest amount of PSNR while NCSR and BM3D give a comparatively lower value. It can be seen that the type of data set plays a significant role in determining the performance of the specific algorithm; as SAPCA gives PSNR 32 for house image and 22 for PAN at standard deviation 30.

Further the second level for PSNR is begged by methods like DNN, LPGPCA, NSST, SURELET, and INLM for MRI; NCSR LPGPCA NSST, SURELET, and INLM for House image; and WBF, DNN, PLPCA, PGPCA, NCSR and LPGPCA for PAN image at standard deviation 30.

Other methods like Bitonic filter, SBF and DF (A-D) gives lowest values of PSNR for all three types of data sets.

9.4. Visual analysis

For the sake of minimalism, visual results for denoised images at standard deviation 30 have been shown in the manuscript. Referring to Fig. 10, it is quite apparent that non-local similarity and patch based methods such as BM3D, SAPCA, INLM, LPGPCA and NCSR demonstrates top tier performance as in the resultant image: edges are well preserved; visual quality is high, homogenous regions are free from noise, low-contrast objects are visible, textures are also well preserved and there is no emergence of artefacts. This conforms the fact, that the idea of non-local grouping forms the major reform in image denoising. These visual results resonate with the objective evaluation for these best performing techniques.

Further methods such as NSST, DNN, PLPCA, PHPCA and TV1 gives considerable amount of denoising and corner and edges are also preserved. However these methods induce some ringing artefacts around

Fig. 10. House image denoised using various methods at standard deviation 30.

edges. Some methods like PGPCA, Ker. Reg., NLFMT and SURELET causes over-smoothing and generates artificial textures. Rest of the methods are rather performing poorly both in case of denoising ability and edge preservation.

The SURELET transform gives higher PSNR value than shearlet transform, however, it can be seen that in case of visual results the amount of denoising performed by SURELET is lower which may be due to poor preservation of edge feature and the noise still persists. The WBF and RBF show improved results than SBF in terms of their denoising performances at higher noise level. This can be attributed to the fact SBF simply disables denoising at higher noise levels, whereas WBF adequately removes the high levels of noise along with the preservation of edges. Some of the hybrid methods like GBFMT and NLFMT are also able to achieve fair amount of denoising comparable to vi-hard and vi-soft. Besides, these hybrid methods give much better visual results than the conventional wavelet thresholding methods. The details and edges are drastically lost in the visual results given by vi-hard and vi-soft.

Besides this there is a striking observation that although NCSR gives high quality visual results, its quantitative analysis shows a lower level of PSNR as compared to top tier assessments. On the other hand SURELET method is depicting considerable level of PSNR, it gives low visual quality. Methods like MRF model are performing moderately both objectively and subjectively.

Therefore it can be concluded that methods like BM3D and SAPCA preserves both edges and gives smooth homogeneous regions. Methods like NSST and DNN are able to preserve low contrast details but they tend to induce artefact around edges and methods like WBF, RBF and Bitonic filter are able to preserve edges but they tend to give poor denoising performance in homogenous regions.

Besides the varied level of performance of different methods it can be seen that almost all methods depicts degradation in performance with increasing levels of noise.

10. Conclusion

In this article, an earnest effort has been made in order to compare, classify and evaluate various image denoising methods. Extensive efforts by a huge number of researchers have generated a structural literature which exhibits substantial progressive growth attained by a chain of sequential incremental improvements. While it is nearly impossible to cover all of them, we have covered each domain of image denoising with several methods representative of each category. These methods have been divided into five categories namely spatial domain methods, transform domain methods, methods in statistical domains, hybrid methods and sparse representation and dictionary learning methods. The basic idea of image denoising is edge preservation while getting rid of noisy pixels. The era of image denoising started from basic averaging filter and later on transformed into edge preserving filters. In spatial domain, improved non-local means filter perform well. Through various analytical procedures we observed that exploitation of non-local image similarity or non-local grouping stems the major reform in improvement of results for image denoising for instance NLM filters and BM3D. In transform domain methods, transforms with overcomplete basis functions,

increased number of scales, directions and orientations generated better results. The properties like multiresolution, sparsity, shift invariance and anisotropy are highly needed for an image transform to work well for image denoising. The exploration of overcomplete basis functions in BLS-GSM resulted in improved performance than wavelets with fixed basis functions. Multi-resolution based algorithms perform far better than single resolution ones.

The spatial domain filters is capable of eliminating high frequency noise but at the cost of blurring of low contrast information, whereas transform domain methods were able to preserve details in homogeneous regions but tend to introduce ringing artefacts around the edges. The dictionary learning and sparse representation methods possesses redundant and overcomplete dictionaries, but they often lead to computational burden.

It is observed that an ideal denoising algorithm should possess overcomplete representation, non-local grouping, adaptive basis functions and be computationally efficient. A single denoising algorithm is not able to exhibit desired level of applicability which is true even for BM3D (non-local grouping leads to search complexity). Therefore it was suggested to combine various methods in different domains to harness the attributes of various domains while overcoming limitations of each other. The hybrid domain methods have attained brilliant levels of visual quality of denoised images besides substantial increment in PSNR. Besides this, it can be stated that HVS is primary evaluation factor for performance evaluation of these methods. It is clear from detailed survey that each category of methods is important and has its own specific set of applications. For instance, time sensitive applications require local filtering or transform domain methods as non-local grouping leads to computational burden. Though dictionary learning methods significantly boost the performance but they are not fit where memory and complexity are major concerns and also online training is not viable in practical applications.

Moreover, it has been observed that most of the methods tend to leave behind residual noise and are not able to sustain their performance at higher noise levels. The era of denoising is seemingly approaching its bounds. However it still holds margin of improvement in terms of performance consistency, high visual quality, and preservation of fine feature details, artefact mitigation, more structured dictionaries and adaptive basis functions.

References

[1] M. Lebrun, M. Colom, A. Buades, J.-M. Morel, Secrets of image denoising cuisine, Acta Numer. 21 (2012) 475–576.
[2] A. Buades, B. Coll, J.-M. Morel, A review of image denoising algorithms, with a new one, Multiscale Model. Simul. 4 (2) (2005) 490–530.
[3] C. Knaus, M. Zwicker, Progressive image denoising, IEEE Trans. Image Process. 23 (7) (2014) 3114–3125.
[4] H.-Y. Yang, X.-Y. Wang, P.-P. Niu, Y.-C. Liu, Image denoising using nonsubsampled shearlet transform and twin support vector machines, Neural Netw. 57 (2014) 152–165.
[5] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Trans. Image Process. 15 (Dec (12)) (2006) 3736–3745.
[6] L. Shao, R. Yan, X. Li, Y. Liu, From heuristic optimization to dictionary learning: a review and comprehensive comparison of image denoising algorithms, IEEE Trans. Cybern. 44 (7) (2014) 1001–1013.
[7] T. Buades, Y. Lou, J.-M. Morel, Z. Tang, A note on multi-image denoising, in: Local and Non-Local Approximation in Image Processing, 2009. LNLA 2009. International Workshop on, 2009, pp. 1–15.
[8] R.C. Gonzalez, R.E. Woods, Image processing, Digit. Image Process. 2 (2007).
[9] K. Irie, A.E. McKinnon, K. Unsworth, I.M. Woodhead, A model for measurement of noise in CCD digital-video cameras, Meas. Sci. Technol. 19 (4) (2008) 045207.
[10] E. Jakeman, On the statistics of K-distributed noise, J. Phys. A 13 (1) (1980) 31.
[11] R. Yan, L. Shao, L. Liu, Y. Liu, Natural image denoising using evolved local adaptive filters, Signal Process. 103 (2014) 36–44.
[12] L. Shapiro, G. Linda, Stockman, Computer Vision, Prentice-Hall, 2001.
[13] J. Benesty, J. Chen, Y. Huang, Study of the widely linear Wiener filter for noise reduction, in: Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on, 2010, pp. 205–208.
[14] A.K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, Upper Saddle River, 1989.
[15] I. Pitas, A. Venetsanooulos, Nonlinear Digital Filters: Principles and Applications, Kluwer, Boston, 1990.
[16] R. Yang, L. Yin, M. Gabbouj, J. Astola, Y. Neuvo, Optimal weighted median filters under structural constraints, in: ISCAS, 1993, pp. 942–945.
[17] R.C. Hardie, K.E. Barner, Rank conditioned rank selection filters for signal restoration, IEEE Trans. Image Process. 3 (2) (1994).
[18] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Mach. Intell. 12 (7) (1990) 629–639.
[19] P. Perona, J. Malik, Scale space and edge detection using anisotropic diffusion, in: Proc. IEEE Computer Society Workshop on Computer Vision, Dec., IEEE Computer Society Press, Washington, 1987, pp. 16–22.
[20] P. Perona, T. Shiota, J. Malik, Anisotropic diffusion, in: Geometry-Driven Diffusion in Computer Vision, Springer, 1994, pp. 73–92.
[21] C. Tsiotsios, M. Petrou, On the choice of the parameters for anisotropic diffusion in image processing, Pattern Recognit. 46 (5) (2013) 1369–1381.
[22] A. Dogra, B. Goyal, S. Agrawal, Osseous and digital subtraction angiography image fusion via various enhancement schemes and Laplacian pyramid transformations, Future Gener. Comput. Syst. 82 (2018) 149–157.
[23] Y.-L. You, W. Xu, A. Tannenbaum, M. Kaveh, Behavioral analysis of anisotropic diffusion in image processing, IEEE Trans. Image Process. 5 (11) (1996) 1539–1553.
[24] O. Scherzer, J. Weickert, Relations between regularization and diffusion filtering, J. Math. Imaging Vis. 12 (1) (2000) 43–63.
[25] L. Alvarez, P.-L. Lions, J.-M. Morel, Image selective smoothing and edge detection by nonlinear diffusion. II, SIAM J. Numer. Anal. 29 (3) (1992) 845–866.
[26] S. Osher, L.I. Rudin, Feature-oriented image enhancement using shock filters, SIAM J. Numer. Anal. 27 (4) (1990) 919–940.
[27] C.R. Vogel, M.E. Oman, Iterative methods for total variation denoising, SIAM J. Sci. Comput. 17 (1) (1996) 227–238.
[28] L. Alvarez, L. Mazorra, Signal and image restoration using shock filters and anisotropic diffusion, SIAM J. Numer. Anal. 31 (2) (1994) 590–605.
[29] G. Gilboa, N. Sochen, Y.Y. Zeevi, Forward-and-backward diffusion processes for adaptive image enhancement and denoising, IEEE Trans. Image Process. 11 (7) (2002) 689–703.
[30] G. Gilboa, N. Sochen, Y.Y. Zeevi, Image enhancement and denoising by complex diffusion processes, IEEE Trans. Pattern Anal. Mach. Intell. 26 (8) (2004) 1020–1036.
[31] Z. Qiu, L. Yang, W. Lu, A new feature-preserving nonlinear anisotropic diffusion for denoising images containing blobs and ridges, Pattern Recognit. Lett. 33 (3) (2012) 319–330.
[32] F. Catté, P.-L. Lions, J.-M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, SIAM J. Numer. Anal. 29 (1) (1992) 182–193.
[33] M.R. Hajiaboli, An anisotropic fourth-order diffusion filter for image noise removal, Int. J. Comput. Vis. 92 (2) (2011) 177–191.
[34] M. Lysaker, A. Lundervold, X.-C. Tai, Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time, IEEE Trans. Image Process. 12 (12) (2003) 1579–1590.
[35] M.R. Hajiaboli, A self-governing hybrid model for noise removal, in: Pacific-Rim Symposium on Image and Video Technology, 2009, pp. 295–305.
[36] W. Zeng, X. Lu, X. Tan, Non-linear fourth-order telegraph-diffusion equation for noise removal, IET Image Process. 7 (4) (2013) 335–342.
[37] Y.-L. You, M. Kaveh, Fourth-order partial differential equations for noise removal, IEEE Trans. Image Process. 9 (10) (2000) 1723–1730.
[38] S.-M. Chao, D.-M. Tsai, An improved anisotropic diffusion model for detail-and edge-preserving smoothing, Pattern Recognit. Lett. 31 (13) (2010) 2012–2023.
[39] D. Chen, S. MacLachlan, M. Kilmer, Iterative parameter-choice and multigrid methods for anisotropic diffusion denoising, SIAM J. Sci. Comput. 33 (5) (2011) 2972–2994.
[40] J. Xu, Y. Jia, Z. Shi, K. Pang, An improved anisotropic diffusion filter with semi-adaptive threshold for edge preservation, Signal Process. 119 (2016) 80–91.
[41] O. Ghita, P.F. Whelan, A new GVF-based image enhancement formulation for use in the presence of mixed noise, Pattern Recognit. 43 (8) (2010) 2646–2658.
[42] Y.Q. Wang, J. Guo, W. Chen, W. Zhang, Image denoising using modified Perona–Malik model based on directional Laplacian, Signal Process. 93 (9) (2013) 2548–2558.
[43] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.
[44] A. Chambolle, An algorithm for total variation minimization and applications, J. Math. Imaging Vis. 20 (1–2) (2004) 89–97.
[45] P.L. Combettes, J.-C. Pesquet, Proximal splitting methods in signal processing, in: Fixed-Point Algorithms for Inverse Problems in Science and Engineering, 49, Springer, 2011, pp. 185–212.
[46] L. Yaroslavsky, M. Eden, Fundamentals of Digital Optics: Digital Signal Processing in Optics and Holography, Birkhäuser, Boston, 1996.
[47] L.P. Yaroslavsky, Digital Picture Processing: An Introduction, 9, Springer Science & Business Media, 2012.
[48] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: Computer Vision, 1998. Sixth International Conference on, 1998, pp. 839–846.
[49] S.M. Smith, J.M. Brady, SUSAN—a new approach to low level image processing, Int. J. Comput. Vis. 23 (1) (1997) 45–78.
[50] M. Zhang, B.K. Gunturk, Multiresolution bilateral filtering for image denoising, IEEE Trans. Image Process. 17 (12) (2008) 2324–2333.
[51] M. Elad, On the origin of the bilateral filter and ways to improve it, IEEE Trans. Image Process. 11 (10) (2002) 1141–1151.
[52] K.N. Chaudhury, D. Sage, M. Unser, Fast O(1) bilateral filtering using trigonometric range kernels, IEEE Trans. Image Process. 20 (12) (2011) 3376–3382.
[53] K.N. Chaudhury, Acceleration of the shiftable O(1) algorithm for bilateral filtering and nonlocal means, IEEE Trans. Image Process. 22 (4) (2013) 1291–1300.

[54] K.N. Chaudhury, K. Rithwik, Image denoising using optimally weighted bilateral filters: a sure and fast approach, in: Image Processing (ICIP), 2015 IEEE International Conference on, 2015, pp. 108–112.
[55] F. Durand, J. Dorsey, Fast bilateral filtering for the display of high-dynamic-range images, in: ACM Transactions on Graphics (TOG), 21, 2002, pp. 257–266.
[56] F. Porikli, Constant Time O(1) Bilateral Filtering, 2008.
[57] Q. Yang, K.-H. Tan, N. Ahuja, Real-time O(1) bilateral filtering, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 557–564.
[58] Z. Farbman, R. Fattal, D. Lischinski, R. Szeliski, Edge-preserving decompositions for multi-scale tone and detail manipulation, in: ACM Transactions on Graphics (TOG), 27, 2008, p. 67.
[59] B. Goyal, A. Dogra, S. Agrawal, B.S. Sohi, A three stage integrated denoising approach for grey scale images, J. Ambient Intell. Humaniz. Comput. (2018) 1–16.
[60] L. Jin, C. Xiong, H. Liu, Improved bilateral filter for suppressing mixed noise in color images, Digit. Signal Process. 22 (6) (2012) 903–912.
[61] M. Zhang, B.K. Gunturk, Multiresolution bilateral filtering for image denoising, IEEE Trans. Image Process. 17 (12) (2008) 2324–2333.
[62] W. Shi, J. Li, M. Wu, An image denoising method based on multiscale wavelet thresholding and bilateral filtering, Wuhan Univ. J. Nat. Sci. 15 (2) (2010) 148–152.
[63] H. Peng, R. Rao, S.A. Dianat, Multispectral image denoising with optimized vector bilateral filter, IEEE Trans. Image Process. 23 (1) (2014) 264–273.
[64] H. Yu, L. Zhao, H. Wang, Image denoising using trivariate shrinkage filter in the wavelet domain and joint bilateral filter in the spatial domain, IEEE Trans. Image Process. 18 (10) (2009) 2364–2369.
[65] K. He, J. Sun, X. Tang, Guided image filtering, in: European Conference on Computer Vision, 2010, pp. 1–14.
[66] X. Zhang, X. Feng, W. Wang, Two-direction nonlocal model for image denoising, IEEE Trans. Image Process. 22 (1) (2013) 408–412.
[67] M. Mahmoudi, G. Sapiro, Fast image and video denoising via nonlocal means of similar neighborhoods, IEEE Signal Process. Lett. 12 (12) (2005) 839–842.
[68] S. Wang, Y. Xia, Q. Liu, J. Luo, Y. Zhu, D.D. Feng, Gabor feature based nonlocal means filter for textured image denoising, J. Vis. Commun. Image Represent. 23 (7) (2012) 1008–1018.
[69] B. Goossens, H. Luong, A. Pizurica, W. Philips, An improved non-local denoising algorithm, in: 2008 International Workshop on Local and Non-Local Approximation in Image Processing (LNLA 2008), 2008, pp. 143–156.
[70] S. Xu, Y. Zhou, H. Xiang, S. Li, Remote sensing image denoising using patch grouping-based nonlocal means algorithm, IEEE Geosci. Remote Sens. Lett. 14 (12) (2017) 2275–2279.
[71] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, C. Barillot, An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images, IEEE Trans. Med. Imaging 27 (4) (2008) 425–441.
[72] D. Tschumperlé, L. Brun, Non-local image smoothing by applying anisotropic diffusion PDE's in the space of patches, in: Image Processing (ICIP), 2009 16th IEEE International Conference on, 2009, pp. 2957–2960.
[73] S. Grewenig, S. Zimmer, J. Weickert, Rotationally invariant similarity measures for nonlocal image denoising, J. Vis. Commun. Image Represent. 22 (2) (2011) 117–130.
[74] Y. Teng, Y. Zhang, Y. Chen, C. Ti, Adaptive morphological filtering method for structural fusion restoration of hyperspectral images, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (2) (2016) 655–667.
[75] G. Treece, The bitonic filter: linear filtering in an edge-preserving morphological framework, IEEE Trans. Image Process. 25 (11) (2016) 5199–5211.
[76] Z. Xiong, K. Ramchandran, M.T. Orchard, Y.-Q. Zhang, A comparative study of DCT-and wavelet-based image coding, IEEE Trans. Circuits Syst. Video Technol. 9 (5) (1999) 692–695.
[77] S. Mallat, A Wavelet Tour of Signal Processing, Elsevier, 1999.
[78] M. Misiti, Y. Misiti, G. Oppenheim, J.P. Michel, Wavelet Toolbox: For Use With Matlab, 1996.
[79] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1989) 674–693.
[80] C.K. Chui, Wavelets: a tutorial in theory and applications, First and Second Volume of Wavelet Analysis and Its Applications, 1992.
[81] J. Kovacevic, W. Sweldens, Wavelet families of increasing order in arbitrary dimensions, IEEE Trans. Image Process. 9 (3) (2000) 480–496.
[82] G. Pajares, J.M. De La Cruz, A wavelet-based image fusion tutorial, Pattern Recognit. 37 (9) (2004) 1855–1872.
[83] A. Dogra, Performance comparison of different wavelet families based on bone vessel fusion, Asian J. Pharm. 10 (04) (2017).
[84] A. Dogra, B. Goyal, S. Agrawal, C.K. Ahuja, Efficient fusion of osseous and vascular details in wavelet domain, Pattern Recognit. Lett. 94 (2017) 189–193.
[85] V. Bruni, B. Piccoli, D. Vitulano, A fast computation method for time scale signal denoising, Signal Image Video Process. 3 (1) (2009) 63–83.
[86] D.L. Donoho, J.M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81 (3) (1994) 425–455.
[87] D.L. Donoho, De-noising by soft-thresholding, IEEE Trans. Inf. Theory 41 (3) (1995) 613–627.
[88] S.G. Chang, B. Yu, M. Vetterli, Adaptive wavelet thresholding for image denoising and compression, IEEE Trans. Image Process. 9 (9) (2000) 1532–1546.
[89] B. Vidakovic, Nonlinear wavelet shrinkage with Bayes rules and Bayes factors, J. Am. Stat. Assoc. 93 (441) (1998) 173–179.
[90] A. Fathi, A.R. Naghsh-Nilchi, Efficient image denoising method based on a new adaptive wavelet packet thresholding function, IEEE Trans. Image Process. 21 (9) (2012) 3981–3990.
[91] I.K. Fodor, C. Kamath, Denoising through wavelet shrinkage: an empirical study, J. Electron. Imaging 12 (1) (2003) 151–161.
[92] S. Sulochana, R. Vidhya, Image denoising using adaptive thresholding in framelet transform domain, Int. J. Adv. Comput. Sci. Appl. 3 (9) (2012).
[93] A. Dixit, P. Sharma, A comparative study of wavelet thresholding for image denoising, Int. J. Image Graph. Signal Process. 12 (2014) 39–46.
[94] D.L. Donoho, I.M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, J. Am. Stat. Assoc. 90 (432) (1995) 1200–1224.
[95] D.L. Donoho, I.M. Johnstone, Threshold selection for wavelet shrinkage of noisy data, in: Proceedings of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1, IEEE, 1994, pp. A24–A25.
[96] A. Pi˘zurica, W. Philips, Estimating the probability of the presence of a signal of interest in multiresolution single and multiband image denoising, IEEE Trans. Image Process. 15 (3) (2006) 654–665.
[97] F. Luisier, T. Blu, M. Unser, A new sure approach to image denoising: interscale orthonormal wavelet thresholding, IEEE Trans. Image Process. 16 (3) (2007) 593–606.
[98] F. Luisier, T. Blu, SURE-LET multichannel image denoising: interscale orthonormal wavelet thresholding, IEEE Trans. Image Process. 17 (4) (2008) 482–492.
[99] G.Y. Chen, T.D. Bui, A. Krzyzak, Image denoising using neighbouring wavelet coefficients, Integr. Comput. Aided Eng. 12 (1) (2005) 99–107.
[100] D. Zhou, X. Shen, Image denoising using block thresholding, in: Image and Signal Processing, 2008. CISP'08. Congress on, 3, 2008, pp. 335–338.
[101] J.-L. Starck, F. Murtagh, E.J. Candès, D.L. Donoho, Gray and color image contrast enhancement by the curvelet transform, IEEE Trans. Image Process. 12 (6) (2003) 706–717.
[102] M. Lang, H. Guo, J.E. Odegard, C.S. Burrus, R.O. Wells, Noise reduction using an undecimated discrete wavelet transform, IEEE Signal Process. Lett. 3 (1) (1996) 10–12.
[103] N. Kingsbury, The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement, in: Signal Processing Conference (EUSIPCO 1998), 9th European, 1998, pp. 1–4.
[104] W. Ding, F. Wu, X. Wu, S. Li, H. Li, Adaptive directional lifting-based wavelet transform for image coding, IEEE Trans. Image Process. 16 (2) (2007) 416–427.
[105] R.R. Coifman, D.L. Donoho, Translation-invariant de-noising, in: Wavelets and Statistics, 103, Springer, 1995, pp. 125–150.
[106] E.P. Simoncelli, W.T. Freeman, The steerable pyramid: a flexible architecture for multi-scale derivative computation, in: Image Processing Proceedings., International Conference, 3, 1995, pp. 444–447.
[107] R.D. Da Silva, R. Minetto, W.R. Schwartz, H. Pedrini, Adaptive edge-preserving image denoising using wavelet transforms, Pattern Anal. Appl. 16 (4) (2013) 567–580.
[108] T. Qiu, A. Wang, N. Yu, A. Song, LLSURE: local linear SURE-based edge-preserving image filtering, IEEE Trans. Image Process. 22 (1) (2013) 80–90.
[109] E.J. Candes, Ridgelets: Theory and Applications, Stanford University, Stanford, 1998.
[110] E.J. Candès, D.L. Donoho, Ridgelets: a key to higher-dimensional intermittency? Philos. Trans. R. Soc. Lond. Ser. A 357 (1999) 2495–2509.
[111] S.R. Deans, The Radon Transform and Some of Its Applications, Wiley, New York, 1983.
[112] M.N. Do, M. Vetterli, The finite ridgelet transform for image representation, IEEE Trans. Image Process. 12 (1) (2003) 16–28.
[113] E.D. Bolker, The finite Radon transform, Contemp. Math. 63 (1987) 27–50.
[114] X. Wang, Wrap-around effect removal finite ridgelet transform for multiscale image denoising, Pattern Recognit. 43 (11) (2010) 3693–3698.
[115] Q. Huang, B. Hao, S. Chang, Adaptive digital ridgelet transform and its application in image denoising, Digit. Signal Process. 52 (2016) 45–54.
[116] M.N. Do, M. Vetterli, Image denoising using orthonormal finite ridgelet transform, in: Wavelet Applications in Signal and Image Processing VIII, 4119, 2000, pp. 831–843.
[117] J.-L. Starck, E.J. Candès, D.L. Donoho, The curvelet transform for image denoising, IEEE Trans. Image Process. 11 (6) (2002) 670–684.
[118] Z. Chen, S. Wang, G. Fang, J. Wang, Ionograms denoising via curvelet transform, Adv. Space Res. 52 (7) (2013) 1289–1296.
[119] B. Zhang, J.M. Fadili, J.-L. Starck, Wavelets, ridgelets, and curvelets for Poisson noise removal, IEEE Trans. Image Process. 17 (7) (2008) 1093–1108.
[120] F.G. Meyer, R.R. Coifman, Brushlets: a tool for directional image analysis and image compression, Appl. Comput. Harmon. Anal. 4 (2) (1997) 147–187.
[121] D.L. Donoho, Wedgelets: nearly minimax estimation of edges, Ann. Stat. 27 (3) (1999) 859–897.
[122] D.L. Donoho, X. Huo, Beamlets and multiscale image analysis, in: Multiscale and Multiresolution Methods, Springer, 2002, pp. 149–196.
[123] M.N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106.
[124] M.N. Do, M. Vetterli, Contourlets: a directional multiresolution image representation, in: International Conference on Image Processing Proceedings 2002, 1, 2002 I–I.
[125] R. Eslami, H. Radha, The contourlet transform for image denoising using cycle spinning, in: Signals, Systems and Computers, 2004. Conference Record of the Thirty–Seventh Asilomar Conference on, 2, 2003, pp. 1982–1986.
[126] R. Eslami, H. Radha, Translation-invariant contourlet transform and its application to image denoising, IEEE Trans. Image Process. 15 (11) (2006) 3362–3374.
[127] A.L. Da Cunha, J. Zhou, M.N. Do, The nonsubsampled contourlet transform: theory, design, and applications, IEEE Trans. Image Process. 15 (10) (2006) 3089–3101.
[128] L. Huang, H. Wang, B. Zhu, Adaptive thresholds algorithm of image denoising based on nonsubsampled contourlet transform, in: Computer Science and Software Engineering, 2008 International Conference on, 6, 2008, pp. 209–212.

[129] M.F. Hossain, M.R. Alsharif, K. Yamashita, LMMSE-based image denoising in non-subsampled contourlet transform domain, in: International Conference on Image and Signal Processing, 2010, pp. 36–43.
[130] S. Satheesh and K. Prasad, "Medical image denoising using adaptive threshold based on contourlet transform," arXiv:1103.4907, 2011.
[131] J. Xu, L. Yang, D. Wu, Ripplet: a new transform for image processing, J. Vis. Commun. Image Represent. 21 (7) (2010) 627–639.
[132] D. Gupta, R.S. Anand, B. Tyagi, Ripplet domain non-linear filtering for speckle reduction in ultrasound medical images, Biomed. Signal Process. Control 10 (2014) 79–91.
[133] E. Le Pennec, S. Mallat, Non-Linear Image Approximation with Bandelets, CMAP/École Polytechnique, 2003 Tech. Rep.
[134] H.B. Kekre, S.D. Thepade, J. Jain, N. Agrawal, Iris recognition using texture features extracted from haarlet pyramid, Int. J. Comput. Appl. 11 (12) (2010) 1–5.
[135] F. Smeraldi, Ranklets: orientation selective non-parametric features applied to face detection, in: Object recognition supported by user interaction for service robots, 3, IEEE, 2002, pp. 379–382.
[136] J.R. Kaplan, D.L. Donoho, The morphlet transform: a multiscale representation for diffeomorphisms, in: DEFORM'06-Proceedings of the Workshop on Image Registration in Deformable Environments, 2006, p. 21.
[137] J. Krommweh, Tetrolet transform: a new adaptive haar wavelet algorithm for sparse image representation, J. Vis. Commun. Image Represent. 21 (4) (2010) 364–374.
[138] http://www.laurent-duval.eu/siva-wits-where-is-the-starlet.html (Accessed on 23. 11November .2016).
[139] F. Luisier, T. Blu, M. Unser, Undecimated Haar thresholding for Poisson intensity estimation, in: Image Processing (ICIP), 2010 17th IEEE International Conference on, 2010, pp. 1697–1700.
[140] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, P.L. Dragotti, Directionlets: anisotropic multidirectional representation with separable filtering, IEEE Trans. Image Process. 15 (7) (2006) 1916–1933.
[141] R. Sethunadh, T. Thomas, Spatially adaptive image denoising using inter-scale dependence in directionlet domain, IET Image Process. 9 (2) (2014) 142–152.
[142] J. Liu, Y. Wang, K. Su, W. He, Image denoising with multidirectional shrinkage in directionlet domain, Signal Process. 125 (2016) 64–78.
[143] Q. Gao, Y. Lu, D. Sun, Z.-L. Sun, D. Zhang, Directionlet-based denoising of SAR images using a Cauchy model, Signal Process. 93 (5) (2013) 1056–1063.
[144] G. Easley, D. Labate, W.-Q. Lim, Sparse directional image representations using the discrete shearlet transform, Appl. Comput. Harmon. Anal. 25 (1) (2008) 25–46.
[145] W.-Q. Lim, The discrete shearlet transform: a new directional transform and compactly supported shearlet frames, IEEE Trans. Image Process. 19 (5) (2010) 1166–1180.
[146] K. Guo, D. Labate, W.-Q. Lim, Edge analysis and identification using the continuous shearlet transform, Appl. Comput. Harmon. Anal. 27 (1) (2009) 24–46.
[147] P. Jain, V. Tyagi, A survey of edge-preserving image denoising methods, Inf. Syst. Front. 18 (1) (2016) 159–170.
[148] B. Goyal, A. Dogra, S. Agrawal, B.S. Sohi, Two-dimensional gray scale image denoising via morphological operations in NSST domain & bitonic filtering, Future Gener. Comput. Syst. 82 (2018) 158–175.
[149] B. Hou, X. Zhang, X. Bu, H. Feng, SAR image despeckling based on nonsubsampled shearlet transform, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (3) (2012) 809–823.
[150] Y. Shi, X. Yang, Y. Guo, Translation invariant directional framelet transform combined with Gabor filters for image denoising, IEEE Trans. Image Process. 23 (1) (2014) 44–55.
[151] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Trans. Image Process. 16 (8) (2007) 2080–2095.
[152] J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain, IEEE Trans. Image Process. 12 (11) (2003) 1338–1351.
[153] L. Zhang, W. Dong, D. Zhang, G. Shi, Two-stage image denoising by principal component analysis with local pixel grouping, Pattern Recognit. 43 (4) (2010) 1531–1549.
[154] L. Sendur, I.W. Selesnick, Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency, IEEE Trans. Signal Process. 50 (11) (2002) 2744–2756.
[155] L. Sendur, I.W. Selesnick, Bivariate shrinkage with local variance estimation, IEEE Signal Process. Lett. 9 (12) (2002) 438–441.
[156] P. Moulin, J. Liu, Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors, IEEE Trans. Inf. Theory 45 (3) (1999) 909–919.
[157] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, BM3D image denoising with shape-adaptive principal component analysis, SPARS'09-Signal Processing with Adaptive Sparse Structured Representations, 2009.
[158] Y. Hou, C. Zhao, D. Yang, Y. Cheng, Comments on image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Trans. Image Process. 20 (1) (2011) 268–270.
[159] S. Yin, L. Cao, Y. Ling, G. Jin, Image denoising with anisotropic bivariate shrinkage, Signal Process. 91 (8) (2011) 2078–2090.
[160] H. Rabbani, R. Nezafat, S. Gazor, Wavelet-domain medical image denoising using bivariate laplacian mixture model, IEEE Trans. Biomed. Eng. 56 (12) (2009) 2826–2837.
[161] G. Yu, G. Sapiro, S. Mallat, Solving inverse problems with piecewise linear estimators: from Gaussian mixture models to structured sparsity, IEEE Trans. Image Process. 21 (5) (2012) 2481–2499.
[162] A. Fathi, A.R. Naghsh-Nilchi, Efficient image denoising method based on a new adaptive wavelet packet thresholding function, IEEE Trans. Image Process. 21 (9) (2012) 3981–3990.
[163] A. Rangarajan, R. Chellappa, Markov random field models in image processing, in: M.A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, 1995, pp. 564–567.
[164] Y. Cao, Y. Luo, S. Yang, Image denoising based on hierarchical Markov random field, Pattern Recognit. Lett. 32 (2) (2011) 368–374.
[165] S. Chen, M. Liu, W. Zhang, J. Liu, Edge preserving image denoising with a closed form solution, Pattern Recognit. 46 (3) (2013) 976–988.
[166] P. Zhong, R. Wang, Multiple-spectral-band CRFs for denoising junk bands of hyperspectral imagery, IEEE Trans. Geosci. Remote Sens. 51 (4) (2013) 2260–2275.
[167] J. Ho, W.-L. Hwang, Image denoising using wavelet Bayesian network models, in: Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, 2012, pp. 1105–1108.
[168] J.M. Sanches, J.C. Nascimento, J.S. Marques, Medical image noise reduction using the Sylvester–Lyapunov equation, IEEE Trans. Image Process. 17 (9) (2008) 1522–1539.
[169] X.T. Wang, G.M. Shi, Y. Niu, L. Zhang, Robust adaptive directional lifting wavelet transform for image denoising, IET Image Process. 5 (3) (2011) 249–260.
[170] J.K. Romberg, H. Choi, R.G. Baraniuk, Bayesian wavelet-domain image modeling using hidden Markov trees, in: Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on, 1, 1999, pp. 158–162.
[171] M. Malfait, D. Roose, Wavelet-based image denoising using a Markov random field a priori model, IEEE Trans. Image Process. 6 (4) (1997) 549–565.
[172] C. Knaus, M. Zwicker, Dual-domain image denoising, in: 20th IEEE International Conference on Image Processing (ICIP), 2013, 2013, pp. 440–444.
[173] C. Knaus, M. Zwicker, Dual-domain filtering, SIAM J. Imaging Sci. 8 (3) (2015) 1396–1420.
[174] M. Rais, Fast and Accurate Image Registration. Applications to On-Board Satellite Imaging, Université Paris-Saclay, 2016.
[175] Z. Huang, Y. Zhang, Q. Li, T. Zhang, N. Sang, H. Hong, Progressive dual-domain filter for enhancing and denoising optical remote-sensing images, IEEE Geosci. Remote Sens. Lett. 15 (5) (2018) 759–763.
[176] N. Wang, et al., A hybrid model for image denoising combining modified isotropic diffusion model and modified Perona-Malik model, IEEE Access (2018).
[177] H. Yu, L. Zhao, H. Wang, Image denoising using trivariate shrinkage filter in the wavelet domain and joint bilateral filter in the spatial domain, IEEE Trans. Image Process. 18 (10) (2009) 2364–2369.
[178] H. Al-Marzouqi, G. AlRegib, Curvelet transform with learning-based tiling, Signal Process. 53 (2017) 24–39.
[179] C. Zhang, Y. Chen, C. Duanmu, Y. Yang, Image denoising by using PDE and GCV in tetrolet transform domain, Eng. Appl. Artif. Intell. 48 (2016) 204–229.
[180] Y. Gan, E. Angelini, A. Laine, C. Hendon, BM3D-based ultrasound image denoising via brushlet thresholding, in: Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on, 2015, pp. 667–670.
[181] Y. Bai, Y. Liu, Q. Zhang, L. Jia, Z. Gui, Image denoising via an improved non-local total variation model, J. Eng. 2018 (8) (2018) 745–752.
[182] C. Sutour, C.-A. Deledalle, J.-F. Aujol, Adaptive regularization of the NL-means: application to image and video denoising, IEEE Trans. Image Process. 23 (8) (2014) 3506–3521.
[183] G.R. Easley, D. Labate, F. Colonna, Shearlet-based total variation diffusion for denoising, IEEE Trans. Image Process. 18 (2) (2009) 260–268.
[184] P.-L. Shui, Image denoising algorithm via doubly local wiener filtering with directional windows in wavelet domain, IEEE Signal Process. Lett. 12 (10) (2005) 681–684.
[185] X. Wang, J. Gao, Image denoising method based on nonsubsampled contourlet transform and bandelet transform, in: Information Science and Engineering (ICISE), 2009 1st International Conference on, 2009, pp. 1278–1281.
[186] S.K. Panigrahi, S. Gupta, P.K. Sahu, Curvelet-based multiscale denoising using non-local means & guided image filter, IET Image Process. 12 (6) (2018) 909–918.
[187] A. Karami, R. Heylen, P. Scheunders, Band-specific shearlet-based hyperspectral image noise reduction, IEEE Trans. Geosci. Remote Sens. 53 (9) (2015) 5054–5066.
[188] E.J. Candès, F. Guo, New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction, Signal Process. 82 (11) (2002) 1519–1543.
[189] S. Durand, J. Froment, Reconstruction of wavelet coefficients using total variation minimization, SIAM J. Sci. Comput. 24 (5) (2003) 1754–1767.
[190] H.R. Shahdoosti, S.M. Hazavei, Combined ripplet and total variation image denoising methods using twin support vector machines, Multimedia Tools Appl. 77 (6) (2018) 7013–7031.
[191] J. Ma, G. Plonka, Combined curvelet shrinkage and nonlinear anisotropic diffusion, IEEE Trans. Image Process. 16 (9) (2007) 2198–2206.
[192] G.Y. Chen, B. Kégl, Image denoising with complex ridgelets, Pattern Recognit. 40 (2) (2007) 578–585.
[193] B.S. Kumar, Image denoising based on gaussian/bilateral filter and its method noise thresholding, Signal Image Video Process. 6 (7) (2013) 1159–1172.
[194] B.S. Kumar, Image denoising based on non-local means filter and its method noise thresholding, Signal Image Video Process. 7 (6) (2013) 1211–1227.
[195] G. Gao, Image denoising by non-subsampled shearlet domain multivariate model and its method noise thresholding, Optik 124 (22) (2013) 5756–5760.
[196] P. Xu, Q. Miao, X. Tang, J. Zhang, A denoising algorithm via wiener filtering in the shearlet domain, Multimedia Tools Appl. 71 (3) (2014) 1529–1558.
[197] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Trans. Image Process. 15 (12) (2006) 3736–3745.

[198] S. Mallat, Z. Zhang, Matching Pursuit With Time-Frequency Dictionaries, Courant Institute of Mathematical Sciences, New York, United States, 1993.
[199] Y.C. Pati, R. Rezaiifar, P.S. Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, in: Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, 1993, 1993, pp. 40–44.
[200] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (1) (2001) 129–159.
[201] P. Chatterjee, P. Milanfar, Clustering-based denoising with locally learned dictionaries, IEEE Trans. Image Process. 18 (7) (2009) 1438–1451.
[202] J. Mairal, F. Bach, J. Ponce, G. Sapiro, A. Zisserman, Non-local sparse models for image restoration, in: 12th International Conference on Computer Vision, 2009 IEEE, 2009, pp. 2272–2279.
[203] X. Zhu, P. Milanfar, Automatic parameter selection for denoising algorithms using a no-reference measure of image content, IEEE Trans. Image Process. 19 (12) (2010) 3116–3132.
[204] P. Bouboulis, K. Slavakis, S. Theodoridis, Adaptive kernel-based image denoising employing semi-parametric regularization, IEEE Trans. Image Process. 19 (6) (2010) 1465–1479.
[205] P. Chatterjee, P. Milanfar, Is denoising dead? IEEE Trans. Image Process. 19 (4) (2010) 895–911.
[206] P. Chatterjee, P. Milanfar, Patch-based near-optimal image denoising, IEEE Trans. Image Process. 21 (4) (2012) 1635.
[207] J. Zhang, D. Zhao, R. Xiong, S. Ma, W. Gao, Image restoration using joint statistical modeling in a space-transform domain, IEEE Trans. Circuits Syst. Video Technol. 24 (6) (2014) 915–928.
[208] H. Talebi, P. Milanfar, Global image denoising, IEEE Trans. Image Process. 23 (2) (2014) 755–768.
[209] R. Sethunadh, T. Thomas, Spatially adaptive image denoising using inter-scale dependence in directionlet domain, IET Image Process. 9 (2) (2014) 142–152.
[210] H. Yue, X. Sun, J. Yang, F. Wu, Image denoising by exploring external and internal correlations, IEEE Trans. Image Process. 24 (6) (2015) 1967–1982.
[211] G. Galiano, J. Velasco, On a cross-diffusion system arising in image denoising, Comput. Math. Appl. 26 (5) (2018) 984–996.
[212] V. Papyan, M. Elad, Multi-scale patch-based image restoration, IEEE Trans. Image Process. 25 (1) (2016) 249–261.
[213] J. Feng, L. Song, X. Huo, X. Yang, W. Zhang, An optimized pixel-wise weighting approach for patch-based image denoising, IEEE Signal Process. Lett. 22 (1) (2015) 115–119.
[214] G. Baloch, H. Ozkaramanli, R. Yu, Residual correlation regularization based image denoising, IEEE Signal Process. Lett. 25 (2) (2018) 298–302.
[215] H.R. Shahdoosti, O. Khayat, Image denoising using sparse representation classification and non-subsampled shearlet transform, Signal Image Video Process. 10 (6) (2016) 1081–1087.
[216] H. Ma, Y. Nie, An edge fusion scheme for image denoising based on anisotropic diffusion models, J. Vis. Commun. Image Represent. 40 (2016) 406–417.
[217] R.J. Jevnisek, S. Avidan, Co-occurrence filter, in: CVPR, 2017, pp. 3816–3824.
[218] A. Wong, A. Mishra, W. Zhang, P. Fieguth, D.A. Clausi, Stochastic image denoising based on Markov-chain Monte Carlo sampling, Signal Process. 91 (8) (2011) 2112–2120.
[219] L. Xu, M. Shaifee, A. Wong, D. Clausi, Fully-connected continuous conditional random field with stochastic cliques for dark-spot detection in SAR imagery, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (7) (2016) 2882–2890.
[220] A. Wong, J. Scharcanski, Monte Carlo despeckling of transrectal ultrasound images of the prostate, Digit. Signal Process. 22 (5) (2012) 768–775.
[221] F. Li, L. Xu, A. Wong, D.A. Clausi, QMCTLS: quasi Monte Carlo texture likelihood sampling for despeckling of complex polarimetric SAR images, IEEE Geosci. Remote Sens. Lett. 12 (7) (2015) 1566–1570.
[222] L. Xu, F. Li, A. Wong, D.A. Clausi, Hyperspectral image denoising using a spatial–spectral Monte Carlo sampling approach, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (6) (2015) 3025–3038.
[223] X. Yang, B. Fei, A wavelet multiscale denoising algorithm for magnetic resonance (MR) images, Meas. Sci. Technol. 22 (2) (2011) 025803.
[224] M. Kumar, M. Diwakar, CT image denoising using locally adaptive shrinkage rule in tetrolet domain, J. King Saud Univ.-Comput. Inf. Sci. 30 (1) (2018) 41–50.
[225] S. Jian, W. Wen, in: Study on underwater image denoising algorithm based on wavelet transform, 806, 2017, p. 01. 2006.
[226] T.S. Sharmila, K. Ramar, Efficient analysis of hybrid directional lifting technique for satellite image denoising, Signal Image Video Process. 8 (7) (2014) 1399–1404.
[227] S. Liu, M. Liu, P. Li, J. Zhao, Z. Zhu, X. Wang, SAR image denoising via sparse representation in shearlet domain based on continuous cycle spinning, IEEE Trans. Geosci. Remote Sens. 55 (5) (2017) 2985–2992.
[228] Y. Shen, et al., Improved anscombe transformation and total variation for denoising of lowlight infrared images, Infrared Phys Technol 93 (2018) 192–198.
[229] https://www.mathworks.com/matlabcentral/fileexchange/67703-image-processing-dataset-for-color-grey-image-fusion–image-blending–image-denoising–enhancement (accessed Accessed on 9. 10October. 2017)
[230] A. Dogra, S. Agrawal, B. Goyal, Efficient representation of texture details in medical images by fusion of Ripplet and DDCT transformed images, Trop. J. Pharm. Res. 15 (9) (2016) 1983–1993.
[231] A. Dogra, B. Goyal, S. Agrawal, Bone vessel image fusion via generalized Reisz wavelet transform using averaging fusion rule, J. Comput. Sci. 21 (2017) 371–378.
[232] A. Dogra, S. Agrawal, B. Goyal, N. Khandelwal, C.K. Ahuja, Color and grey scale fusion of osseous and vascular information, J. Comput. Sci. 17 (2016) 103–114.
[233] A. Dogra, S. Agrawal, N. Khandelwal, C. Ahuja, Osseous and vascular information fusion using various spatial domain filters, Asian J. Res. Chem. 9 (7) (2016) 937.
[234] B. Goyal, S. Agrawal, B.S. Sohi, A. Dogra, Noise reduction in MR brain image via various transform domain schemes, Asian J. Res. Chem. 9 (7) (2016) 919.
[235] A. Dogra, B. Goyal, S. Agrawal, From multi-scale decomposition to non-multi-scale decomposition methods: a comprehensive survey of image fusion techniques and its applications, IEEE Access 5 (2017) 16040–16067.
[236] J. Yadav, A. Dogra, B. Goyal, S. Agrawal, A review on image fusion methodologies and applications, Res. J. Pharm. Technol. 10 (4) (2017) 1239.
[237] B. Goyal, A. Dogra, S. Agrawal, B.S. Sohi, Dual way residue noise thresholding along with feature preservation, Pattern Recognit. Lett. 94 (2017) 194–201.
[238] A. Dogra, S. Kadry, B. Goyal, S. Agrawal, An efficient image integration algorithm for night mode vision applications, MultimedIS Tools Appl. (2018) 1–18.
[239] A. Dogra, B. Goyal, S. Agrawal, Current and future orientation of anatomical and functional imaging modality fusion, Biomed. Pharmacol. J. 10 (4) (2017) 1661–1663.
[240] A. Dogra, M.S. Patterh, CT and MRI brain images registration for clinical applications, J. Cancer Sci. Ther. 6 (2014) 018–026.
[241] A. Dogra, P. Bhalla, CT and mri brain images matching using ridgeness correlation, Biomed. Pharmacol. J. 7 (2) (2014) 20.
[242] R. Kumar, 2010 (https://in.mathworks.com/matlabcentral/fileexchange/28112-diffusion-filtering-for-image-denoising).
[243] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Joint image sharpening and denoising by 3D transform-domain collaborative filtering, in: Proc. 2007 Int. TICSP Workshop Spectral Meth. Multirate Signal Process., SMMSP, 2007, Citeseer, 2007.
[244] H.R. Shahdoosti, Z. Rahemi, Edge-preserving image denoising using a deep convolutional neural network, Signal Process. 159 (2019) 20–32.
[245] V. Vapnik, The support vector method of function estimation, in: Nonlinear Modeling, Springer, Boston, MA, 1998, pp. 55–85.
[246] X.Y. Wang, H.Y. Yang, Z.K. Fu, A new wavelet-based image denoising using undecimated discrete wavelet transform and least squares support vector machine, Expert Syst. Appl. 37 (10) (2010) 7040–7049.
[247] X.Y. Wang, H.Y. Yang, Y. Zhang, Z.K. Fu, Image denoising using SVM classification in nonsubsampled contourlet transform domain, Inf. Sci. 246 (2013) 155–176.
[248] E.J. Balster, Y.F. Zheng, R.L. Ewing, Feature-based wavelet shrinkage algorithm for image denoising, IEEE Trans. Image Process. 14 (12) (2005) 2024–2039.
[249] H.R. Shahdoosti, O. Khayat, Combination of anisotropic diffusion and non-subsampled shearlet transform for image denoising, J. Intell. Fuzzy Syst. 30 (6) (2016) 3087–3098.
[250] H.R. Shahdoosti, Z. Rahemi, A maximum likelihood filter using non-local information for despeckling of ultrasound images, Mach. Vis. Appl. 29 (4) (2018) 689–702.
[251] H.R. Shahdoosti, S.M. Hazavei, A new compressive sensing based image denoising method using block-matching and sparse representations over learned dictionaries, Multimedia Tools Appl. 78 (9) (2019) 12561–12582.
[252] H.R. Shahdoosti, Two-stage image denoising considering interscale and intrascale dependencies, J. Electron. Imaging 26 (6) (2017) 063029.
[253] H.R. Shahdoosti, S.M. Hazavei, Image denoising in dual contourlet domain using hidden Markov tree models, Digit. Signal Process. 67 (2017) 17–29.
[254] S. Parameswaran, E. Luo, T.Q. Nguyen, Patch matching for image denoising using neighborhood-based collaborative filtering, IEEE Trans. Circuits Syst. Video Technol. 28 (2) (2016) 392–401.
[255] A. Karami, L. Tafakori, Image denoising using generalised Cauchy filter, IET Image Process. 11 (9) (2017) 767–776.
[256] W. Zhao, Y. Lv, Q. Liu, B. Qin, Detail-preserving image denoising via adaptive clustering and progressive PCA thresholding, IEEE Access 6 (2017) 6303–6315.
[257] V. Fedorov, C. Ballester, Affine non-local means image denoising, IEEE Trans. Image Process. 26 (5) (2017) 2137–2148.
[258] M. Lebrun, M. Colom, J.-M. Morel, Multiscale image blind denoising, IEEE Trans. Image Process. 24 (10) (2015) 3149–3161.
[259] Q. Guo, C. Zhang, Y. Zhang, H. Liu, An efficient SVD-based method for image denoising, IEEE Trans. Circuits Syst. Video Technol. 26 (5) (2015) 868–880.
[260] I. Frosio, J. Kautz, Statistical nearest neighbors for image denoising, IEEE Trans. Image Process. 28 (2) (2018) 723–738.
[261] H. Wang, S. Cao, K. Jiang, H. Wang, Q. Zhang, Seismic data denoising for complex structure using BM3D and local similarity, J. Appl. Geophys. (2019).
[262] S. Chen, S. Chen, Z. Guo, Y. Zuo, Low-resolution palmprint image denoising by generative adversarial networks, Neurocomputing 358 (2019) 275–284.
[263] S. Li, Y. Chen, R. Jiang, X. Tian, Image denoising via multi-scale gated fusion network, IEEE Access 7 (2019) 49392–49402.
[264] X. Zheng, Y. Yuan, X. Lu, Hyperspectral image denoising by fusing the selected related bands, IEEE Trans. Geosci. Remote Sens. 57 (5) (2018) 2596–2609.
[265] J. Liu, S. Osher, Block matching local SVD operator based sparsity and TV regularization for image denoising, J. Sci. Comput. 78 (1) (2019) 607–624.
[266] H. Wang, Y. Li, Y. Cen, Z. He, L. Zhang, Multi-matrices low-rank decomposition with structural smoothness for image denoising, IEEE Trans. Circuits Syst. Video Technol. (2019).
[267] W. Jifara, F. Jiang, S. Rho, M. Cheng, S. Liu, Medical image denoising using convolutional neural network: a residual learning approach, J. Supercomput. 75 (2) (2019) 704–718.
[268] G. Baloch, H. Ozkaramanli, R. Yu, Residual correlation regularization based image denoising, IEEE Signal Process. Lett. 25 (2) (2018) 298–302.
[269] M. Diwakar, M. Kumar, A review on CT image noise and its denoising, Biomed. Signal Process. Control 42 (2018) 73–88.
[270] Y. Weiss, W.T. Freeman, What makes a good model of natural images? in: 2007 IEEE Conference on Computer Vision and Pattern Recognition, June, IEEE, 2007, pp. 1–8.
[271] Q. Liu, B. Xiong, M. Zhang, Adaptive sparse norm and nonlocal total variation methods for image smoothing, Math. Probl. Eng. (2014) (2014).

[272] W. Dong, L. Zhang, G. Shi, X. Li, Nonlocally centralized sparse representation for image restoration, IEEE Trans. Image Process. 22 (4) (2012) 1620–1630.
[273] H. Takeda, S. Farsiu, P. Milanfar, Deblurring using regularized locally adaptive kernel regression, IEEE Trans. Image Process. 17 (4) (2008) 550–563.
[274] H.C. Burger, C.J. Schuler, S. Harmeling, Image denoising: can plain neural networks compete with BM3D? in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2012, pp. 2392–2399.
[275] C.-A. Deledalle, J. Salmon, A.S. Dalalyan, Image denoising with patch based PCA: local versus global, in: BMVC, 81, 2011, pp. 425–455.
[276] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, T. Pock, An introduction to total variation for image analysis, Theor. Found. Numer. Methods Sparse Recov. 9 (263–340) (2010) 227.