

Review

BM3D Outperforms Major Benchmarks in Denoising: An Argument in Favor

¹Bhawna Goyal, ²Ayush Dogra and ²Apoorav Maulik Sharma

¹Department of ECE, Chandigarh University, Mohali, Punjab, India

²UIET, Panjab University, Chandigarh, India

Article history

Received: 14-04-2020

Revised: 24-05-2020

Accepted: 01-07-2020

Corresponding Author:

Bhawna Goyal

Department of ECE,

Chandigarh University, Mohali,

Punjab, India

Email: bhawnagoyal28@gmail.com

Abstract: The inherent physical limitations of imaging sensors lead to prevalence of additive white Gaussian noise in images which deters the feature extraction and analysis. There exists a number of denoising algorithms in literature, demonstrating their efficacy for removing noise while preserving feature details. At the crossing of functional and statistical analysis, one argues with new methods being devised quite frequently, whether the decade old BM3D is still efficient or not. While carrying out extended experimentation and evaluation for removal of Gaussian noise from natural images in terms peak signal to noise ratio, an argument in favor of BM3D has been presented in this manuscript.

Keywords: Denoising, PSNR, Shearlet, BM3D and Bitonic Filtering

Introduction

The digital data display and transmission has significantly propelled major fields of application such remote sensing, medical sciences, astronomy, surveillance and computer vision. Digital images are a two dimensional array of pixel intensity values (Buades *et al.*, 2005). The mechanism of image acquisition embodies the basic principle of illumination and projection of the object under investigation. Sensors or cameras undergo electronic and thermal fluctuations while image acquisition and transmission due to their inherent optical properties. Also, often objects are illuminated with inconsistent photon count. These factors results in the manifestation of additive white Gaussian noise in the image which irrevocably destroys the quality of image and hence hinders image interpretability (Zhang *et al.*, 2012). The other factors such as bit error rate and faulty manufacturing can lead to occurrence of other types of noises such as photon, quantum, impulse and speckle noise. It has been observed that the prevalence of these types of noise is seldom. However the presence of additive Gaussian noise is most often. It is spread across the image in a uniform manner and follows a normal distribution curve.

Gaussian function has a probability density function of the normal distribution. A random value calculated by the normal distribution is added to each of the clean pixel. The noise samples drawn are independent of each other and every pixel the variance and the mean are the same.

In statistical and functional analysis, the first approximation for any real valued random variable that

is uniformly distributed around a single mean are taken as normally distributed. The prime reasons for the normal distribution to be the most major PDF's are: According to central limit theorem, the mean for an expanded set of random variables non-dependently withdrawn from identical distribution is approximately distributed normally, irrespective of the form or the type of original distribution. Besides this, the results and analytical functions involving normal distribution can be derived in an explicit form

The noisy images are not only visually unpleasant, but also disrupt further image analysis and information extraction. Hence the removal of noise is necessary. The advancement in hardware based optical technology is able to mitigate such affects however software based algorithms are widely accepted due to their device independence (Zeng and Zhao, 2007).

The progression in the fields of digital signal processing, statistical methods and mathematical theories has resulted in the coining of technical algorithms for removal of noise. The efficacy of an image denoising algorithm is defined by noise removal amount while preservation of information pixel detail. The major concerns of researchers in the field of image denoising are (Tian *et al.*, 2020; Jifara *et al.*, 2019; Krull *et al.*, 2019):

- Adequate removal of noise
- Preservation of edges, features, details and textures
- Preservation of low contrast details
- Consistency in performance at varying noise levels
- Mitigation of occurrence of artifacts

The spectral and multi-resolution analysis, partial differential equations, probability theory and statistics etc., form the major disciplines for the origination of various denoising methods in literature. The era of image denoising algorithms starting from the basic averaging filtering has now marginalised its limits to advanced neural network and patch-based filtering. The popular methods include Bitonic filtering, Non-sub Sampled Shearlet Transform (NSST), Curvature filter, Support Vector Machines (SVM), anisotropic diffusion and Block Matching 3D collaborative filtering (BM3D). According to (Goyal *et al.*, 2020), the noise suppression can be categorised into spatial domain filtering, transform domain filtering, statistical methods, hybrid methods, machine learning based methods and patch based filtering. Broadly, spatial-domain methods directly work on image pixels and transform domain methods transforms the image into coefficients and then carries out thresholding (Portilla *et al.*, 2003; Zhang and Gunturk, 2008; Blu and Luisier, 2007). The discussion on these different domains remains beyond the purview of this manuscript; an effort has been made to highlight an important aspect of image denoising methods in the light of extended level experimentation and analysis. The convolution neural network based methods and machine learning has utterly revolutionised the field of image processing. There is a huge amount of literature demonstrating their outstanding performance (Shao *et al.*, 2013; 2008; Elad and Aharon, 2006). However it has been observed that, in case of medical imaging a typical ML algorithm recognises the significant features of region of interest which are “believed to be important” on the basis of the input data set (Erickson *et al.*, 2017).

ML techniques prerequisite a large amount of data for their training and validation which in turn raises the concerns regarding data sharing, computerised trust and privacy concerns. Besides this ML algorithm only works well in case of training but fails drastically when independent validation data processing is required. These algorithms rely pre-defined set of features which are of major concern in the field of medical imaging owing to the uniqueness of the relevant features. Therefore choosing the adequate features to correctly model the given research problem is challenging (Shao *et al.*, 2013). The ML algorithms involve huge memory requirement and high level of complexity due to online training iterations. Even if the data required is ready to choose, in case of medical imaging the extremely high effective computational cost of medical scans makes it an impractical tool for medical research.

The statistical methods are also quite limited in terms of their complexity and large dependency on unpredictable model behaviour. This narrows down our debate to the performance comparison of spatial domain and transform-domain methods. While presenting our argument we have chosen three benchmark methods representative of their respective domains namely BM3D (Hybrid method), NSST (transform domain method) and Bitonic filtering (spatial domain filtering). The various works presented in (Goyal *et al.*, 2020) backs the selection of these methods. While discussing these methods we will initially lay a basic understanding of these algorithms.

BM3D

It has already been extensively argued that exploitation of non-local image similarity forms the major grounds for large improvement in denoising algorithms i.e., to consider or analysis similar pixels which are not confined to a local neighbourhood. Deriving its basics from non-local grouping the block matching 3-D collaborative filtering was proposed in (Dabov *et al.*, 2006). The technique was divided into three steps: Grouping; Group matching and collaborative filtering. BM3D employs an enhanced level of sparsity by group similar 2D arrays in 3d data arrays and are labelled as groups. These 3d data arrays so obtained are subjected to transformation. The transformation of 2-D signals leads to generation of a transform spectrum where high valued coefficients define the signal and low valued coefficients contains noise and peaks. Further these spectrum obtained are threshold in order to obtain the noise shrinkage. This results in formation of a 3d estimate that contains of filtered grouped image blocks. Then this 3d transformation is inverse transformed. This collaborative filtering where preserves the essential features of each unique block, also removes noise. The filtered blocks so obtained placed back to their original positions. Then this process of group and filtering is repeated with carrying out the shrinkage of noise using Weiner filtering. In between these overlapping blocks, number of estimates is generated for each pixel which is combined together. The significant improvement in the performance of this method is attributed to collaborative filtering and block matching. This method is able to preserve even the finest details in the image while adequately removing the noise (Burger *et al.*, 2012; Dabov *et al.*, 2006). BM3D is the hybridisation of spatial and transform domain. Presently it is the era of hybridization. It is observed that where spatial domain filters are able to preserve edges they lack on preserving low contrast details, however in case of transform domain filters, low contrast details are preserved but there is emergence of ringing artifacts

around edges. Hence hybridization is indeed necessary to harness the attributes of both these domains while overcoming their limitations

NSST

In the year 2008, shearlet transform was introduced in the field of image processing. Since the shearlets were shift variant and resulted in Gibbs phenomena (prevalence of artefacts), non sub sampled shearlet transform was proposed which was shift invariant and provided a representational format of the multidimensional data. NSST is multi-scale and multidimensional tool that is able to represent the image features in all directions significantly. The entire implementation of NSST consisted of employment of non-sub sampled shearing filters and non-sub sampled pyramid filter banks. The key feature is that NSST is able induce shift invariance which is accomplished by omitting up and down sampling. These results in non-decimation of coefficients amongst levels and the size of the sub-bands remain same as the original input image. The NSST can be understood as combination of scaling and sheering filter in all directions. The most striking advantage of this method is preservation of fine feature details and mitigation of ringing artefacts which is attributed to its highly directional representation format. This method is able to exhibit exceptional performance in terms of image denoising (Easley *et al.*, 2008; Lim, 2013).

Bitonic Filtering

A signal in its definition can be modelled as function of second order. Its imperative a signal is made of smooth curves, singularities, maximas and minimas. A simple concept can be stated that a continuous periodic signal contains only one maxima or minima within a given range. However, a signal with numerous peaks and dips can be modelled as noise. Conceptualising the above stated fact, (Treece, 2016), designed the Bitonic filtering which filters signal with only one maxima or minima in a given range. The filter so designed is independent of noise estimation. It is basically morphological based filtering which employs rank filtering in order to carry out the filtering of the true signal. This simple yet effect 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 (Treece, 2016; Goyal *et al.*, 2018).

Experiments, Results and Discussion

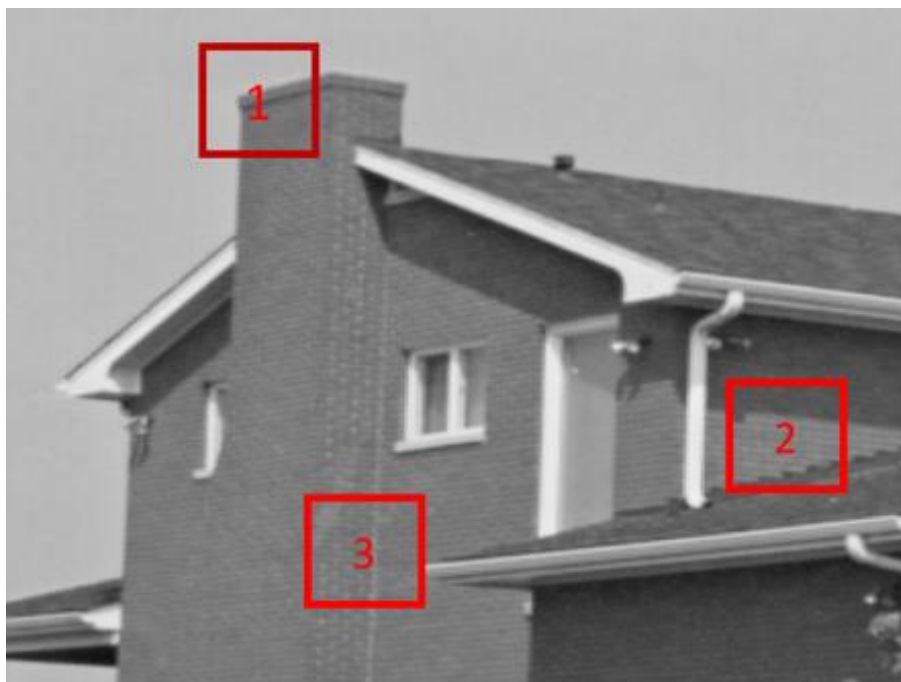
The experiments have been performed on the “house” image infected with Gaussian noise with different standard deviations viz. 10, 20, 30, 40 and 50, using MATLAB version R2013b. The house images are of size 256×256 and have file extension “.png”.

According to above Fig. 1, the areas in the original image can be identified as “small scale details”, “large scale details” and “mixed details”. In small scale details the transition between the pixel intensities is very small and these areas are represented by number 3 used in the image. Secondly, in large scale details the transition between the pixel intensities is very large hence they represent strong edges present in the image and these areas are represented by number 1 in the image. Finally, the mixed transitions are those in which the small scale and large scale pixel transitions are in the proximity of each other and these areas are represented by number 2 in the image. This analogy is also supported by the pixel intensity map in which the small scale details have very low intensity transitions. The number 1 represents the strong edges where the transitions are very high and there are areas, represented jointly by number 2 and 3 where small scale and significantly larger scale transitions are present. In a noisy image, the small scale details get deteriorated very badly so, it becomes very difficult for a de-noising algorithm to differentiate between the noise and these details. In other words, it is very difficult to preserve these details hence, any de-noising algorithm which could be able to preserve these small scale details would be the most efficient algorithm to de-noise the images. We would take the help of the pixel intensity maps to demonstrate the efficiency of three de-noising algorithms used to de-noise the image and to choose the most efficient one.

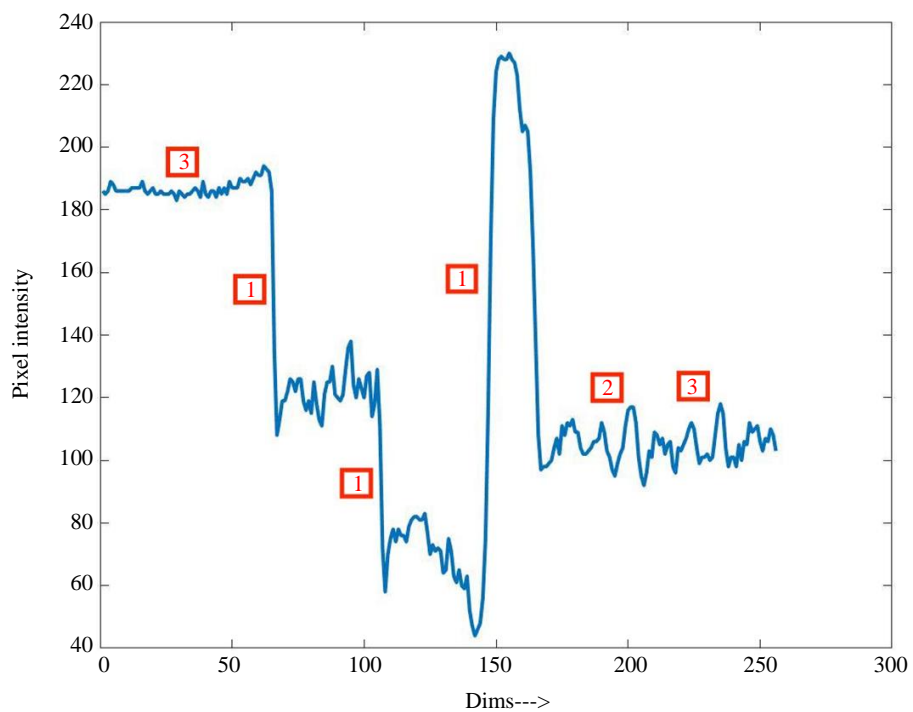
First of all, let us observe the house images corrupted by noise of different sigma (Fig. 2a to 2e) and their corresponding pixel intensity maps (Fig. 2f to 2j). When the noise standard deviation is low (i.e., 10 and 20) the small scale details are affected while the large scale details or the strong edges are still intact. When the noise standard deviation is increased beyond 20, it also starts affecting the strong edges. It is further justified by the corresponding pixel intensity maps which are placed right below the concerned images. The main requirement from a de-noising algorithm is to remove noise by minimally affecting the details. In this manuscript, we have denoised the images using three algorithms viz. Bitonic filter, Shearlet transform and Block matching and 3D filtering (BM3D). We will analyze the performance of these three algorithms also on the basis of pixel intensity maps so as to provide better explanation and judgment about their performances. For the sake of minimalism, we will include the pixel intensity maps of the de-noised images from two standard deviations viz. 10 and 50. Also, for a comprehensive representation of the results images are named according to the standard deviation of the noise added for instance, image corrupted by noise of standard deviation 20 is named house 20 and so on, up to house 50.

Table 1: PSNR values for denoising of house images corrupted with noise of $\sigma = 10, 20, 30, 40$ and 50

Technique	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$	$\sigma = 40$	$\sigma = 50$
Bitonic filter	31.89	29.59	27.29	25.51	24
Shearlet transform	34.86	31.98	30.30	28.97	27.89
BM3D	36.71	33.77	32.08	30.64	29.69



(a)



(b)

Fig. 1: (a) Original house image (b) Pixel intensities of house image

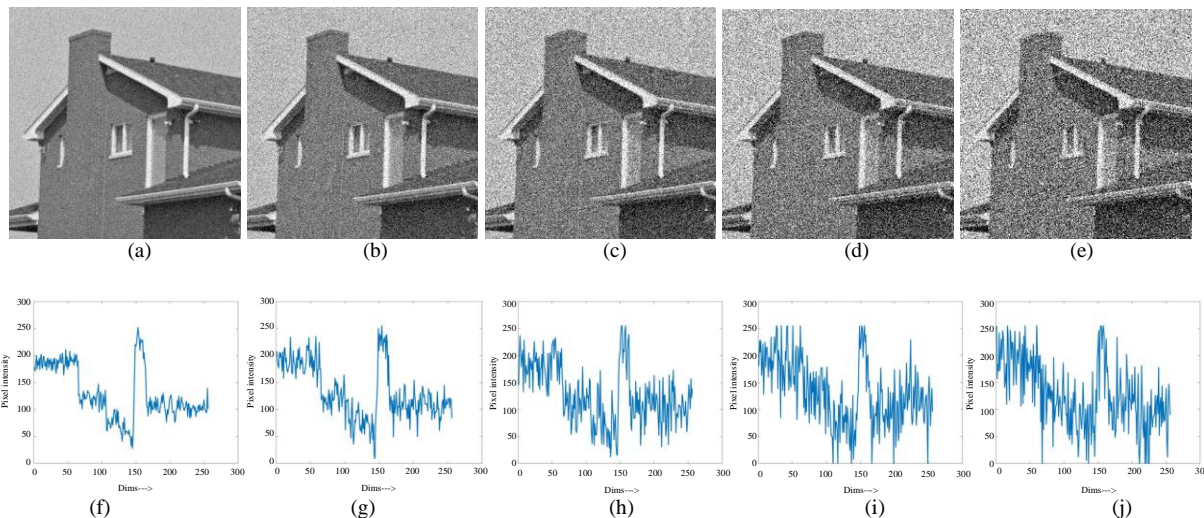


Fig. 2: Images corrupted by noise (a) to (e), (a) house10 (b) house20 (c) house30 (d) house40 (e) house50 and corresponding pixel intensity maps (f) to (j), for (f) house10 (g) house20 (h) house30 (i) house40 (j) house50

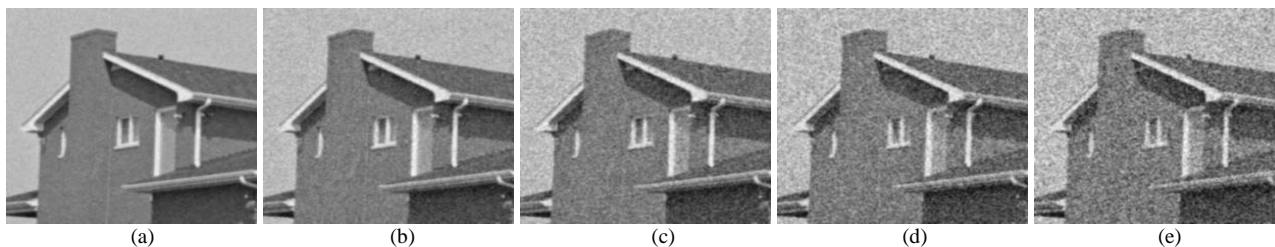
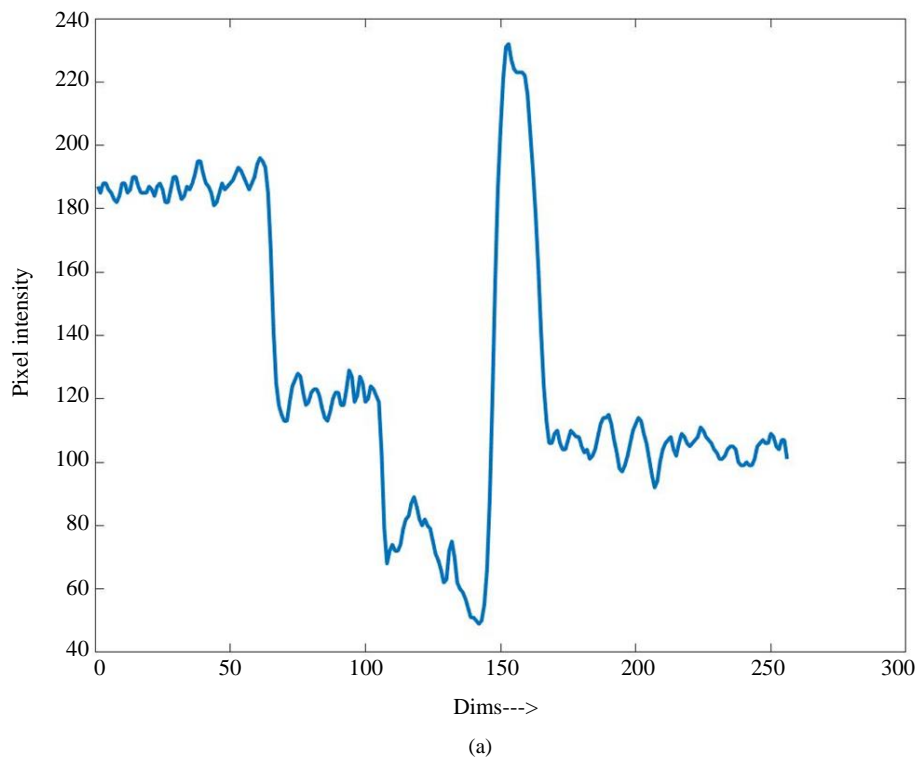


Fig. 3: Bitonic filter de-noising results for (a) house10 (b) house20 (c) house30 (d) house40 (e) house50



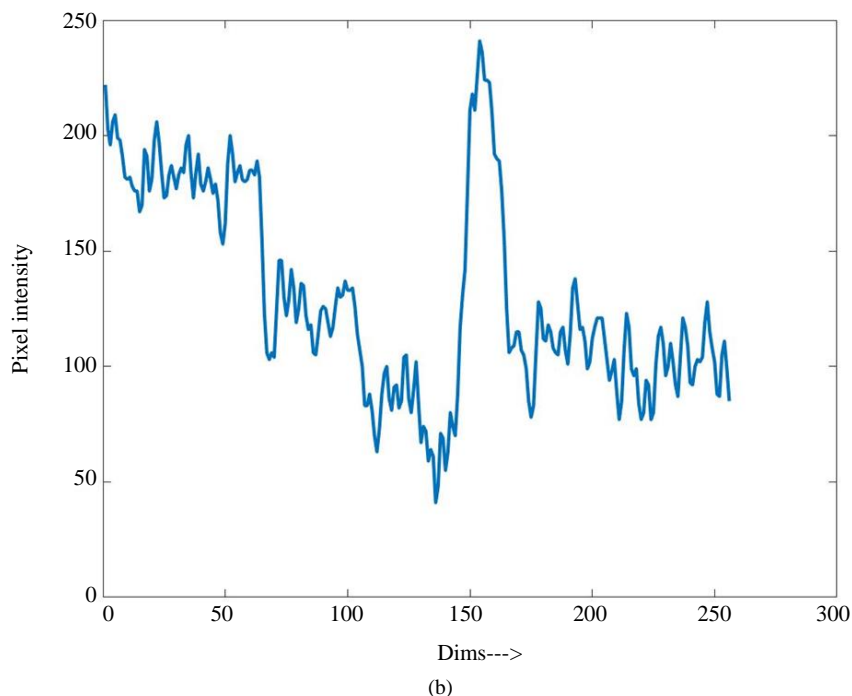


Fig. 4: Pixel intensity maps for (a) house10 and (b) house50, de-noised images using Bitonic filter

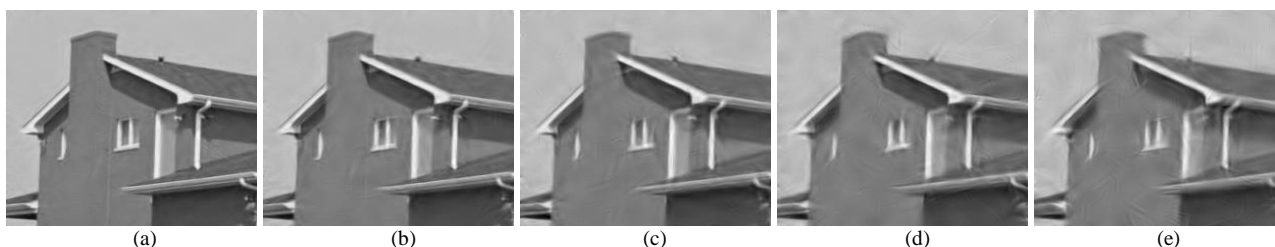


Fig. 5: Shearlet transform de-noising results for (a) house10 (b) house20 (c) house30 (d) house40 (e) house50

A de-noising algorithm must recover enough details from a noise corrupted image to make sense of a scene for human visual system. Keeping this in mind, we will start our discussion with bitonic filter. The bitonic filter is a linear filter which performs better only when noise sigma is low, the noise is not uniformly distributed and in the presence of impulse noise. When noise sigma is increased to 20 and beyond, the bitonic filter starts performing very poorly which is evident from the visual results shown in Fig. 3. Also, as the noise distribution is not uniform in this case, bitonic filter fails to perform. According to pixel intensity maps shown in Fig. 4, when noise sigma is low, small details as well as large details are well preserved but as soon as noise sigma is increased to 50 the strong edges are partially recovered while the small scale details are completely ignored.

In case of Shearlet transform, the denoising results are quite satisfactory. Even when noise sigma is increased to 50, strong details are well preserved after de-noising as shown in Fig. 5. Although, the small scale

details are completely vanished which is also quite evident from pixel intensity maps shown in Fig. 6. Strong edges are well recovered but mixed scales are not recovered properly. The highlighted areas show the actual situation as the curves are flattened. The intensity of noise in image at standard deviation 50 is large amount of noise which destroys the intricate image details to a large extent. Bitonic filter is though able to preserve more fine details in the image; however, it stops the denoising at and around edges. In case of Shearlet filter while removing noise, fine textures details are also compromised because at increasing noise levels the intensity variations between true signal and noise diminishes, hence it becomes difficult to distinguish between noise and signal.

Image denoising is an extremely challenging field. With the availability of large number of sophisticated denoising algorithms none of them is able to mark the marginal limits of denoising. Some preserve fine details at cost of residual noise and some smooth the

image for a pleasant view however at comprised image information.

The third case of BM3D de-noising is the strongest one. This method is most efficient in comparison to the other two methods as it is successful in recovering useful details from noisy images not only when noise sigma is low but also when noise sigma value is high (Fig. 7 and 8). The BM3D algorithm also provides highest PSNR among all as shown in Table 1. In another observation, it is quite evident from below Fig. 9 that where, bitonic filter is unable to

remove the noise at all, the shearlet transform introduces ringing artifacts into the denoised image. These artifacts become more profound with higher standard deviation of noise. These artifacts are introduced when the inter-scale correlation between the decomposed levels is high. The down-sampling operation involved in transforms also facilitates these artifacts to exist. However, these artifacts can be removed by following certain steps before the reconstruction of the concerned signal. The result obtained by BM3D is clean in this aspect.

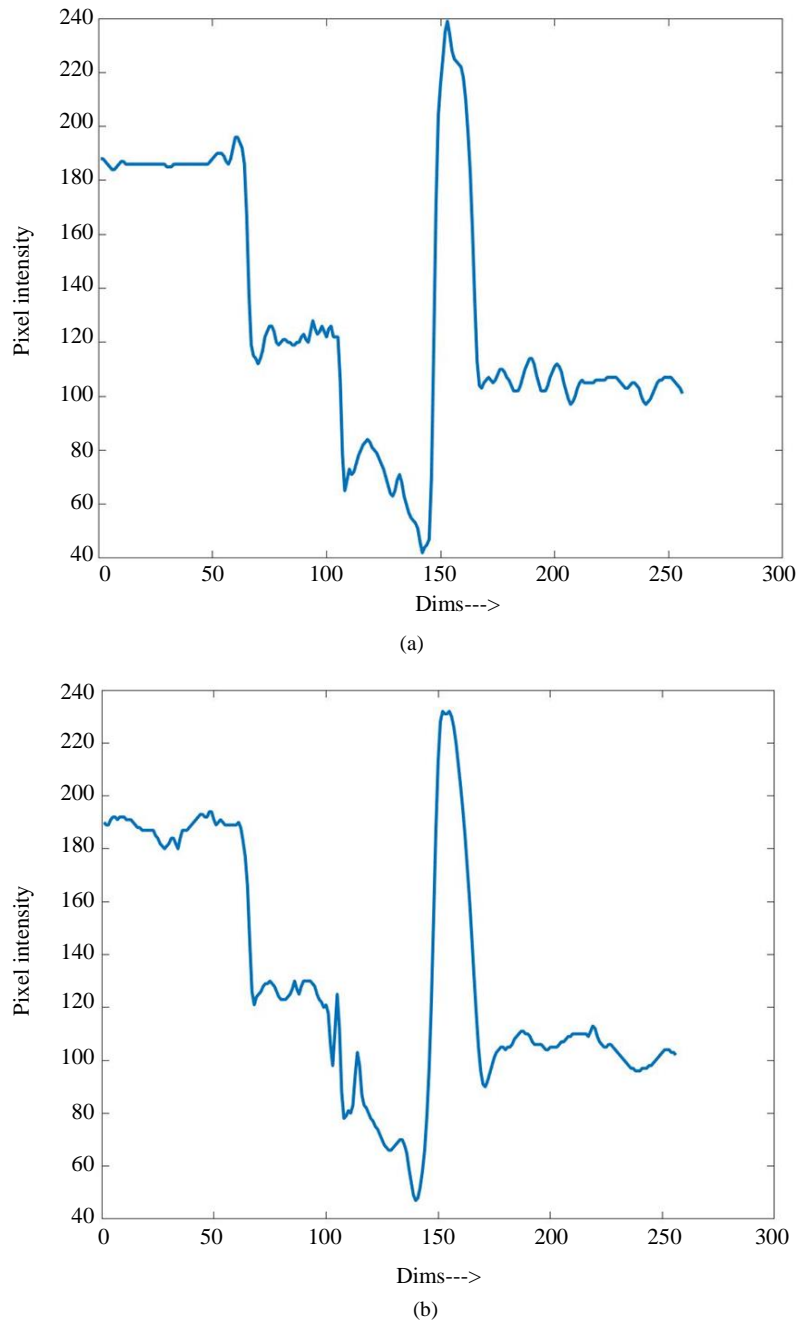


Fig. 6: Pixel intensity maps for (a) house10 and (b) house50, de-noised images using Shearlet transform



Fig. 7: BM3D de-noising results for (a) house10 (b) house20 (c) house30 (d) house40 (e) house50

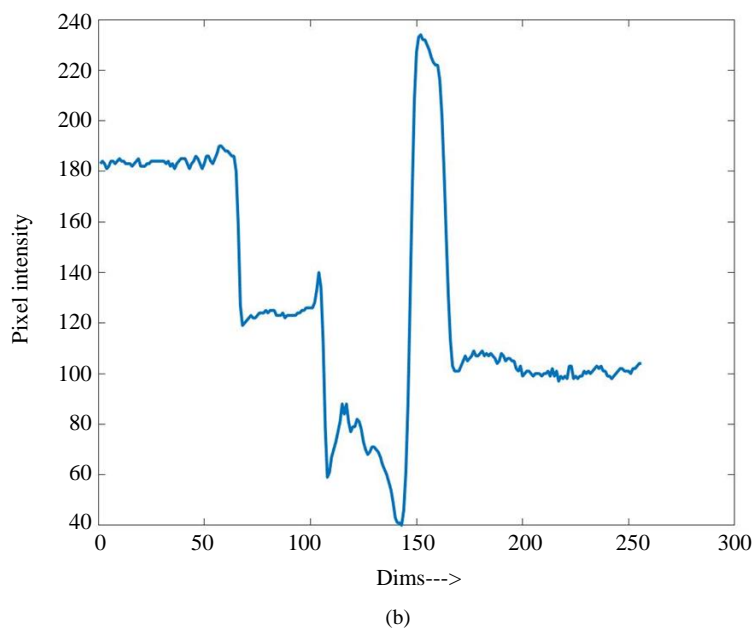
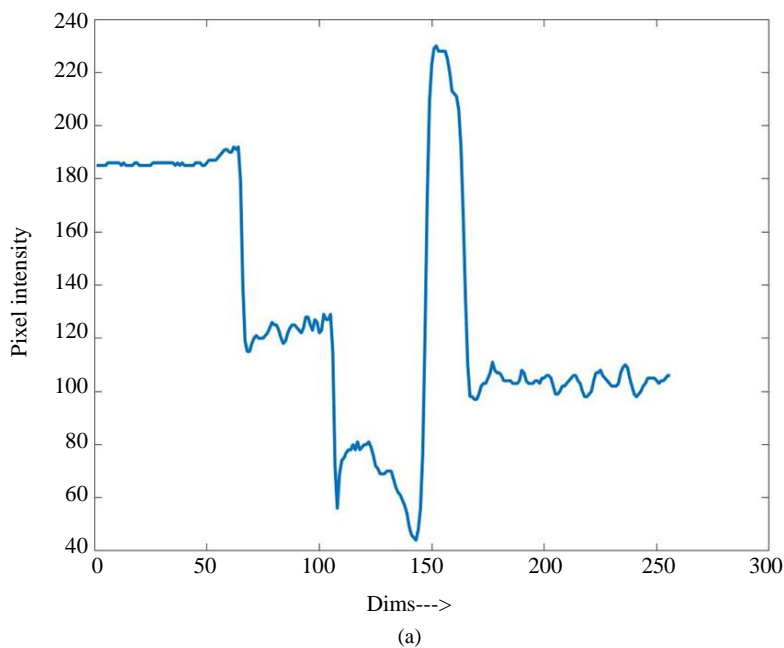


Fig. 8: Pixel intensity maps for (a) house10 and (b) house50, de-noised images using BM3D algorithm

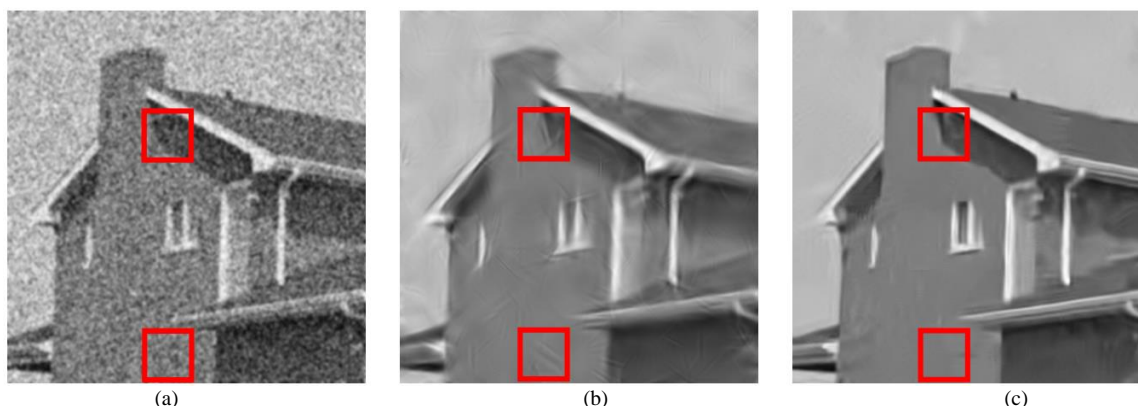


Fig. 9: Denoising results for house50 image using (a) Bitonic filter (b) Shearlet transform (c) BM3D. The highlighted areas of image (b) shows ringing artifacts

Conclusion

In this article three pioneer techniques in the field of denoising are reviewed and compared. An effort has been made to present these techniques in a comprehensive way such that it establishes their achievements and highlights their drawbacks as well. It is well known that, an efficient de-noising algorithm must have adaptive basis function, over-completeness and grouping based on non-local features. After proper analysis of the experiment results it can be successfully concluded that, BM3D is the most efficient, state-of-the-art technique providing excellent results not only at lower levels of noise but at higher levels of noise also. Also, it is quite evident from the discussions that, even the most efficient de-noising algorithms in the current scenario, are not able to recover small-scale details. A small margin of improvement remains there.

Acknowledgement

All the authors contributed to conception, drafting and critical proof reading of this manuscript. This work has been done in collaboration with UIET, Panjab University Chandigarh during the year 2018-19.

Author's Contributions

Bhawna Goyal: Conceived of the presented idea. She drafted the introductory theory along with supervision of the findings.

Ayush Dogra and Apoorav Maulik Sharma: conducted the experiments and jointly drafted the parts of the "Experiments, results and discussions" section.

All authors discussed the results and contributed to the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

References

- Blu, T. and F. Luisier, 2007. The SURE-LET approach to image denoising. *Trans. Image Proc.*, 16: 2778-2786. DOI: 10.1109/TIP.2007.906002
- Buades, A., B. Coll and J.M. Morel, 2005. A non-local algorithm for image denoising. *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition, (VPR' 05)*, pp: 60-65.
- Burger, H.C., C.J. Schuler and S. Harmeling, 2012. Image denoising: Can plain neural networks compete with BM3D? *Proceedings of the Conference on Computer Vision and Pattern Recognition, Jun. 16-21, IEEE Xplore Press, Providence, RI, USA*, pp: 2392-2399. DOI: 10.1109/CVPR.2012.6247952
- Dabov, K., A. Foi, V. Katkovnik and K. Egiazarian, 2006. Image denoising with block-matching and 3D filtering. *Proc. Int. Society Opt. Eng.*, 6064: 354-365. DOI: 10.1117/12.643267
- Easley, G., D. Labate and W.Q. Lim, 2008. Sparse directional image representations using the discrete shearlet transform. *Applied Comput. Harmonic Anal.*, 25: 25-46. DOI: 10.1016/j.acha.2007.09.003
- Elad, M. and M. Aharon, 2006. Image denoising via sparse and redundant representations over learned dictionaries. *Trans. Image Proc.*, 15: 3736-3745. DOI: 10.1109/TIP.2006.881969
- Erickson, B.J., P. Korfiatis, Z. Akkus and T.L. Kline, 2017. Machine learning for medical imaging. *Radiographics*, 37: 505-515. DOI: 10.1148/rg.2017160130
- Goyal, B., A. Dogra, S. Agrawal and B.S. Sohi, 2018. Two-dimensional gray scale image denoising via morphological operations in NSST domain and bitonic filtering. *Future Generat. Comput. Syst.*, 82: 158-175. DOI: 10.1016/j.future.2017.12.034
- Goyal, B., A. Dogra, S. Agrawal, B.S. Sohi and A. Sharma, 2020. Image denoising review: From classical to state-of-the-art approaches. *Inform. Fus.*, 55: 220-244. DOI: 10.1016/j.inffus.2019.09.003

- Jifara, W., F. Jiang, S. Rho, M. Cheng and S. Liu, 2019. Medical image denoising using convolutional neural network: A residual learning approach. *J. Supercomput.*, 75: 704-718.
DOI: 10.1007/s11227-017-2080-0
- Krull, A., T.O. Buchholz and F. Jug, 2019. Noise2void-learning denoising from single noisy images. *Proceedings of the Conference on Computer Vision and Pattern Recognition*, Jun. 15-20, IEEE Xplore Press, Long Beach, CA, USA, pp: 2129-2137.
DOI: 10.1109/CVPR.2019.00223
- Lim, W.Q., 2013. Nonseparable shearlet transform. *Trans. Image Proc.*, 22: 2056-2065.
DOI: 10.1109/TIP.2013.2244223
- Portilla, J., V. Strela, M.J. Wainwright and E.P. Simoncelli, 2003. Image denoising using scale mixtures of Gaussians in the wavelet domain. *Trans. Image Proc.*, 12: 1338-1351. DOI: 10.1109/TIP.2003.818640
- Shao, L., H. Zhang and G. De Haan, 2008. An overview and performance evaluation of classification-based least squares trained filters. *Trans. Image Proc.*, 17: 1772-1782. DOI: 10.1109/TIP.2008.2002162
- Shao, L., R. Yan, X. Li and Y. Liu, 2013. From heuristic optimization to dictionary learning: A review and comprehensive comparison of image denoising algorithms. *Trans. Cybernet.*, 44: 1001-1013.
DOI: 10.1109/TCYB.2013.2278548
- Tian, C., Y. Xu and W. Zuo, 2020. Image denoising using deep CNN with batch renormalization. *Neural Networks*, 121: 461-473.
DOI: 10.1016/j.neunet.2019.08.022
- Treece, G., 2016. The bitonic filter: Linear filtering in an edge-preserving morphological framework. *Trans. Image Proc.*, 25: 5199-5211.
DOI: 10.1109/TIP.2016.2605302
- Zeng, G. and R. Zhao, 2007. Image denoising using least squares wavelet support vector machines. *Chinese Opt. Lett.*, 5: 632-635.
- Zhang, G.D., X.H. Yang, H. Xu, D.Q. Lu and Y.X. Liu, 2012. Image denoising based on support vector machine. *Proceedings of the Spring Congress on Engineering and Technology*, May 27-30, IEEE Xplore Press, Xian, China, pp: 1-4.
DOI: 10.1109/SCET.2012.6341928
- Zhang, M. and B.K. Gunturk, 2008. Multiresolution bilateral filtering for image denoising. *Trans. Image Proc.*, 17: 2324-2333.
DOI: 10.1109/TIP.2008.2006658