IMAGE DE-NOISING TECHNIQUES: A REVIEW PAPER

[1] Pragati Sahu, [2] Shailesh Khaparkar, [3] Ritesh Beohar, [4] Pankaj Sahu
[1]Research Scholar, [2] H.O.D & Associate Professor, [3]Assistant Professor, [4]Assistant Professor
M.Tech. Communication Systems, Dept. of ECE, GGITS, Jabalpur

Abstract-* *Image de-noising is an important image processing task, both as a process itself, and as a component in other processes. In today's scenario transmission of information through images has become a major medium of communication. But during transmission of images they get affected by some external means called as noise. The search for effective and efficient image denoising methods is a great challenge for researchers. Different algorithms are available and each algorithm has its assumptions, advantages and limitations. This paper presents a review of some significant work in the field of Image De-noising. The brief introduction of some popular approaches is provided and discussed.

Keywords- Spatial filtering, linear filtering, non-linear filtering, DWT, Thresholding.

I. INTRODUCTION

Digital image processing is rapidly growing field of signal processing. It is concerned basically with extracting important and useful information from an image. Some important fields where image processing is being used are remote sensing, security monitoring, computer tomography, geographical survey etc. the data is collected from image sensors are affected from various types of noises. The main causes of generation of such noises are transmission errors or compression. Therefore there is a need of removal of these noisy errors which is called denoising process. Before processing of any image we must have to remove these noises from image. So for image restoration this would be very first step. Different noise affects the image in destructive manner in various levels. These noises can be categorized in various types like salt and pepper noise i.e. also known as impulse noise, Gaussian noise also called as uniform noise and random noise. Salt and pepper noise includes sparse light and dark disturbances. Pixels in the image are very different in intensity from the other ones. This type of noise will only affect a small number of image pixels. When viewed, the image contains dark and white dots, hence the term salt and pepper noise. This can have value either 0 or 255. Here 0 represents complete black and 255 represent complete white on gray scale image. The random valued impulse noise can have any value between 0 and 255; hence its removal is very important as well as difficult. In Gaussian noise each pixel in the image will be changed from its original value by a small amount. Random noise is a type of noise comprised of transient disturbances which occur at random times; its instantaneous magnitudes are specified only by probability distribution functions which give the fraction of the total time that the magnitude lies within a specified range. In image analysis image de-noising is a very important and essential pre-processing step. It basically recovers the true picture from the degraded one by different algorithms. This pre-processing technique does not affect the quality of image and do not alter any pictorial information. But just like any other process it also has some limitations which are making it a challenging task for researchers. Although it removes noise but also introduces some artifacts and blurring. In this paper different processes for various de-noising methods are being discussed.

This paper is organized as follows. Section II consists of different noise models. In section III a brief description of various techniques for evolution of image de-noising is given. Section IV gives classification and description of various de-noising methods. In section V conclusion for the work is given.

II. NOISE MODEL

Basically noise generated in any image is uncorrelated with image pixels. Impulse noise distribution is random over entire image. These noises can be categorized in Gaussian noise and impulse noise. Unlike Gaussian noise, impulse noise does not affect all pixels of images. Some of them will be noisy and some will be noiseless. In salt and pepper type of noise pixel will either take 255 or 0 values so it appears as white and black spots. So the probability of uncorrupted pixels will be P-1 and noisy pixel will be appeared with the probability P. In case of random valued impulse noise, noise is randomly distributed over the entire image and it can take up any gray level value from 255 to 0.

III. EVOLUTION OF IMAGE DE-NOISING TECHNIQUE

Image de-noising is a fundamental step of image acquisition and processing. Firstly spatial domain approach has been developed. Greatest advantage of such approach was its speed but along with this there was a major drawback i.e. discontinuities in image means it is unable to preserve edges. Then the focus was shifted to Wavelet domain from spatial and Fourier domain. Ever since the Donoho's wavelet based thresholding approach was published in 2003.

Although this approach did not requires tracking and correlation of wavelet maximx and minima across different scales as proposed by Mallat.[3] There was renewed interest in wavelet approach since Donoho's. [4]. Data adaptive threshold were introduced to achieve optimum threshold[6]. Translation invariant method can improve the quality of perception. More researches were Gaussian scale mixtures, hidden markovo models also Bayesian de-noising. Different statistical models are focused to model the statistical properties of wavelet coefficients and its neighbors. Future trend will be to find more probabilistic model for non-orthogonal wavelet coefficients distribution.

IV. NOISE CLASSIFICATION

For image de-noising two basic methods are popular termed as special filtering method and transform domain filtering method.

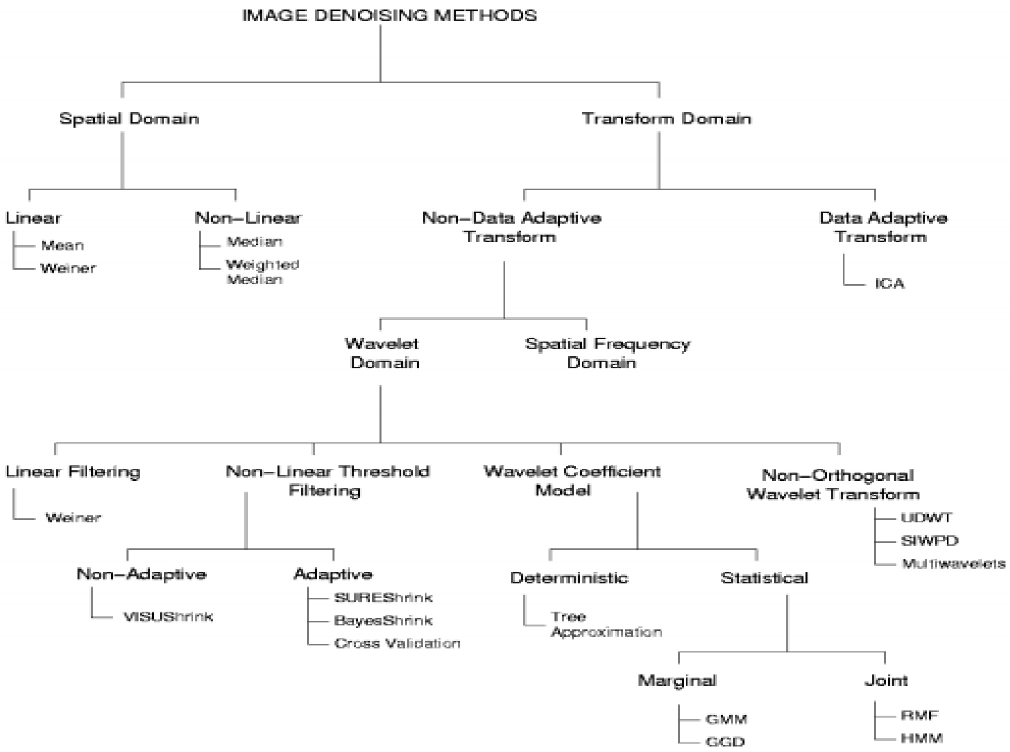

Figure 1: Image denoising techniques classification

4.1. Spacial Filtering-

It is specially used for image enhancement. It performs many tasks like image sharpening etc. it works on neighboring pixels and filtered image pixels are assigned to a corresponding location in a new image. Spatial filters are further classified as non-linear and linear filters.

4.1.1. Non-Linear Filters-

Special filters employ a low pass filtering on groups of pixels with an assumption that noise occupies the higher frequency region of the spectrum. Special filters remove noise from an image in considerable level but it generates various unwanted effect like blurring in image. Various types of special filters are available here like median filter, adaptive median filter etc.

(A). Median Filter-

It considers each pixel in image in turn and looks at its nearby neighbor to decide whether or not it is representative of its surrounding. It replaces the pixel's present value with median of neighbor pixel values.

(B). Adaptive Median Filter-

It performs special processing to determine which pixels in an image have been affected by noise. It classifies pixels as noise by comparing each pixel in image to its surrounding neighbor pixels. The size of neighborhood is adjustable. A pixel that is different from majority of its neighborhood as well as not structurally aligned with those pixels are then replaced by median pixel values of pixels.

(C). Weighted Median Filter-

Centre median filter is easy to implement, it gives more weight to some values within the window. One most important type of weighted filter is centre weighted median filter which gives more weight to the central value of the window.

4.1.2 Linear Filter

Linear filters are generally of two types: mean filter and wiener filter. These filters execute poorly in presence of noise, which results in form of loss of image information.

(A). Mean Filter

Mean filter is a simple sliding window special filter which replaces the value of central window with the average value of all nearby pixel itself. It is implemented with the convolution mask, generally 3×3 mask is used.

(B) Weiner Filter

Weiner filtering requires the information on the spectra of noise and original signal it works better when the signal is smooth. To overcome such problems wavelet based denoising techniques are being used.

4.2. Transform Domain

Transform domain can be classified depending on the function. It can be further subdivided into non-adaptive data transform and adaptive data transform.

4.2.1. Non Adaptive Data Transform

(A). Spatial Frequency Filter

It uses a low pass filter with fast fourier transform. Here we have to assign a cut-off frequency to the filter when the noise is decorrelated with useful signal. Drawback of such transform method is that they are time consuming and dependent on cut-off frequency. Also this may cause artificial frequency in new processed images.

(B). Wavelet Domain-

Wavelet Domain process is again subdivided into following techniques:

(a). Linear Filter

If the signal corruption can be modelled as gaussian process, Linear filters such as Weiner filter can give the optimal result and mean square error (MSE) is the accuracy criterion. Wiener filtering is used where data corruption can be modeled as a Gaussian process and accuracy criterion is mean square error. However, if we design a filter on this assumption, this results in a filtered image which is very displeasant than the original noisy signal even though it considerably reduces the MSE. In a wavelet domain spatially adaptive Weiner Filtering is proposed in which intrascale filtering is not allowed in any case.

(b). Non-Linear Threshold Filtering

Non-Linear threshold filtering is the most investigated domain in denoising using wavelet transform. It basically uses the property of wavelet transform and the fact that wavelet transform maps noise in signal domain to that of noise in transform domain. Thus while signal energy becomes more concentrated into fewer coefficients in transform domain noise energy does not. The method where small coefficients are removed leaving other coefficients untouched is known as Hard Thresholding. However this method produces spurious blips known as artifacts. To overcome these demerits soft thresholding was introduced where coefficients above the threshold are shrunk by the absolute value of threshold itself.

The procedure in which small coefficients are removed while others are left untouched is called Hard Thresholding [5]. But the method generates spurious blips, better known as artifacts, in the images as a result of unsuccessful attempts of removing moderately large noise coefficients. To overcome the demerits of hard thresholding, wavelet transform using soft thresholding was also introduced in [5]. In this scheme, coefficients above the threshold are shrunk by the absolute value of the threshold itself. Similar to soft thresholding, other techniques of applying thresholds are semi-soft thresholding and Garrote thresholding [6]. Most of the wavelet shrinkage literature is based on methods for choosing the optimal threshold which can be adaptive or non-adaptive to the image.

Non-Adaptive thresholds

i. VISUShrink [12] is non-adaptive universal threshold, which depends only on number of data points. It has asymptotic equivalence suggesting best performance in terms of MSE when the number of pixels reaches infinity. VISUShrink is known to yield overly smoothed images because its threshold choice can be unwarrantedly large due to its dependence on the number of pixels in the image.

Adaptive Thresholds

ii. SUREShrink [12] uses a hybrid of the universal threshold and the SURE [Stein's Unbiased Risk Estimator] threshold and performs better than VISUShrink. BayesShrink [17, 18] minimizes the Bayes' Risk Estimator function assuming Generalized Gaussian prior and thus yielding data adaptive threshold. BayesShrink outperforms SUREShrink most of the times. Cross Validation [19] replaces wavelet coefficient with the weighted average of neighborhood coefficients to minimize generalized cross validation (GCV) function providing optimum threshold for every coefficient.

The assumption that one can distinguish noise from the signal solely based on coefficient magnitudes is violated when noise levels are higher than signal magnitudes. Under this high noise circumstance, the spatial configuration of neighboring wavelet coefficients can play an important role in noise-signal classifications. Signals tend to form meaningful features (e.g. straight lines, curves), while noisy coefficients often scatter randomly.

(c) Non-orthogonal Wavelet Transforms

Undecimated Wavelet Transform (UDWT) has also been used for decomposing the signal to provide visually better solution. Since UDWT is shift invariant it avoids visual artifacts such as pseudo-Gibbs phenomenon. Though the improvement in results is much higher, use of UDWT adds a large overhead of computations thus making it less feasible. In [20] normal hard/soft thresholding was extended to Shift Invariant Discrete Wavelet Transform. In [21] Shift Invariant Wavelet Packet Decomposition (SIWPD) is exploited to obtain number of basis functions. Then using Minimum Description Length principle the Best Basis Function was found out which yielded smallest code length required for description of the given data.

Then, thresholding was applied to denoise the data. In addition to UDWT, use of Multiwavelets is explored which further enhances the performance but further increases the computation complexity. The Multiwavelets are obtained by applying more than one mother function (scaling function) to given dataset. Multiwavelets possess properties such as short support, symmetry, and the most importantly higher order of vanishing moments. This combination of shift invariance & Multiwavelets is implemented in [22] which give superior results for the Lena image in context of MSE.

(d) Wavelet Coefficient Model

This approach focuses on exploiting the multiresolution properties of Wavelet Transform. This technique identifies close correlation of signal at different resolutions by observing the signal across multiple resolutions. This method produces excellent output but is computationally much more complex and expensive. The modeling of the wavelet coefficients can either be deterministic or statistical.

Deterministic

The Deterministic method of modeling involves creating tree structure of wavelet coefficients with every level in the tree representing each scale of transformation and nodes representing the wavelet coefficients. This approach is adopted in [23]. The optimal tree approximation displays a hierarchical interpretation of wavelet decomposition. Wavelet coefficients of singularities have large wavelet coefficients that persist along the branches of tree. Thus if a wavelet coefficient has strong presence at particular node then in case of it being signal, its presence should be more pronounced at its parent nodes. If it is noisy coefficient, for instance spurious blip, then such consistent presence will be missing. Lu et al. [24], tracked wavelet local maxima in scalespace, by using a tree structure. Other denoising method based on wavelet coefficient trees is proposed by Donoho [25].

Statistical Modeling of Wavelet Coefficients

This approach focuses on some more interesting and appealing properties of the Wavelet Transform such as multiscale correlation between the wavelet coefficients, local correlation between neighborhood coefficients etc. This approach has an inherent goal of perfecting the exact modeling of image data with use of Wavelet Transform. A good review of statistical properties of wavelet coefficients can be found in [26] and [27]. The following two techniques exploit the statistical properties of the wavelet coefficients based on a probabilistic model.

Marginal Probabilistic Model

A number of researchers have developed homogeneous local probability models for images in the wavelet domain. Specifically, the marginal distributions of wavelet coefficients are highly kurtotic, and usually have a marked peak at zero and heavy tails. The Gaussian mixture model (GMM) [28] and the generalized Gaussian distribution (GGD) [29] are commonly used to model the wavelet coefficients distribution. Although GGD is more accurate, GMM is simpler to use. In [30], authors proposed a methodology in which the wavelet coefficients are assumed to be conditionally independent zero-mean Gaussian random variables, with variances modeled as identically distributed, highly correlated random variables.

An approximate Maximum A Posteriori (MAP) Probability rule is used to estimate marginal prior distribution of wavelet coefficient variances. All these methods mentioned above require a noise estimate, which may be difficult to obtain in practical applications. Simoncelli and Adelson [33] used a twoparameter generalized Laplacian distribution for the wavelet coefficients of the image, which is estimated from the noisy observations. Chang et al. [34] proposed the use of adaptive wavelet thresholding for image denoising, by modeling the wavelet coefficients as a generalized Gaussian random variable, whose parameters are estimated locally (i.e., within a given neighborhood).

Joint Probabilistic Model

Hidden Markov Models (HMM) [35] models are efficient in capturing inter-scale dependencies, whereas Random Markov Field [36] models are more efficient to capture intrascale correlations. The complexity of local structures is not well described by Random Markov Gaussian densities whereas Hidden Markov Models can be used to capture higher order statistics. The correlation between coefficients at same scale but residing in a close neighborhood are modeled by Hidden Markov Chain Model where as the correlation between coefficients across the chain is modeled by Hidden Markov Trees. Once the correlation is captured by HMM, Expectation Maximization is used to estimate the required parameters and from those, denoised signal is estimated from noisy observation using well-known MAP estimator. In [31], a model is described in which each neighborhood of wavelet coefficients is described as a Gaussian scale mixture (GSM) which is a product of a Gaussian random vector, and an independent hidden random scalar multiplier. Strela et al. [32] described the joint densities of clusters of wavelet coefficients as a Gaussian scale mixture, and developed a maximum likelihood solution for estimating relevant wavelet coefficients from the noisy observations. Another approach that uses a Markov random field model for wavelet coefficients was proposed by Jansen and Bulthel [37]. A disadvantage of HMT is the computational burden of the training stage. In order to overcome this computational problem, a simplified HMT, named as uHMT [27], was proposed.

4.2.2. Independent Component Analysis (ICA)

Under the category of data adaptive transformation independent component analysis (ICA) is most widely used technique for finding or extracting individual signal from mixtures. Main application of ICA is in blind source separation. It is also helpful for denoising of gaussian and non-gaussian distribution. Because it uses sliding window method, its cost of computation is very high. Also it requires samples which are free of noise but it is difficult to find in some applications.

V. CONCLUSION

This paper reviews the existing denoising algorithms, such as filtering approach; wavelet based approach. Different noise models including additive and multiplicative types are used. They include Gaussian noise, salt and pepper noise, speckle noise and Brownian noise. The filtering approach seems to be a better choice when the image is corrupted with salt and pepper noise. The wavelet based approach finds applications in denoising images corrupted with Gaussian noise. Selection of the denoising algorithm is application dependent.

Performance of denoising algorithms is measured using uantitative performance measures such as peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR) as well as in terms of visual quality of the images. Many of the current techniques assume the noise model to be Gaussian. In reality, this assumption may not always hold true due to the varied nature and sources of noise. An ideal denoising procedure requires a priori knowledge of the noise, whereas a practical procedure may not have the required information about the variance of the noise or the noise model. Thus, most of the algorithms assume known variance of the noise and the noise model to compare the performance with different algorithms. Gaussian Noise with different variance values is added in the natural images to test the performance of the algorithm. Not all researchers use high value of variance to test the performance of the algorithm when the noise is comparable to the signal strength. Use of FFT in filtering has been restricted due to its limitations in providing sparse representation of data. Wavelet Transform is the best suited for performance because of its properties like sparsity, multiresolution and multiscale nature. In addition to performance, issues of computational complexity must also be considered. Thresholding techniques used with the Discrete Wavelet Transform are the simplest to implement. Non-orthogonal wavelets such as UDWT and Multiwavelets improve the performance at the expense of a large overhead in their computation. HMM based methods seem to be promising but are complex.

When using Wavelet Transform, Nason [40], emphasized that issue such as choice of primary resolution (the scale level at which to begin thresholding) and choice of analyzing wavelet also have a large influence on the success of the shrinkage procedure. When comparing algorithms, it is very important that researchers do not omit these comparison details. Several papers did not specify the wavelet used neither the level of decomposition of the wavelet transform was mentioned. It is expected that the future research will focus on building robust statistical models of non-orthogonal wavelet coefficients based on their intra scale and inter scale correlations. Such models can be effectively used for image denoising and compression.

VI. REFERENCES

[1] Anestis Antoniadis, Jeremie Bigot, "Wavelet Estimators in Nonparametric Regression: A Comparative Simulation Study," Journal of Statistical Software, Vol 6, I 06, 2001.

[2] Castleman Kenneth R, Digital Image Processing, Prentice Hall, New Jersey, 1979.

[3] S. Grace Chang, Bin Yu and Martin Vetterli, "Adaptive Wavelet Thresholding for Image Denoising and Compression," IEEE Trans. Image Processing, Vol 9, No. 9, Sept 2000, pg 1532-1546.

[4] David L. Donoho, "De-noising bysoft-thresholding,"http://citeseer.nj.nec.com/cache/papers/cs/2831/http:zSzzSzwwwstat.stanford.eduzSzreportszSzdonohozSzdenoiserelease3.pdf/donoho94 de noising.pdf, Dept of Statistics, Stanford University, 1992.

[5] David L. Donoho and Iain M. Johnstone, "Adapting to Unknown Smoothness via Wavelet Shrinkage," Journal of American Statistical Association, 90(432):1200-1224, December 1995.

[6] 1/f noise, "Brownian Noise," http://classes.yale.edu/9900/math190a/OneOverF.html, 1999.

[7] B.M.Gammel, "Multifractals," http://www1.physik.tumuenchen. de/~gammel/matpack/html/Mathematics/Multifractals.html, September 1996.

[8] Langis Gagnon, "Wavelet Filtering of Speckle Noise-Some Numerical Results," Proceedings of the Conference Vision Interface 1999, Trois- Riveres.

[9] Amara Graps, "An Introduction to Wavelets," IEEE Computational Science and Engineering, summer 1995, Vol 2, No. 2.

[10] David Harte, Multifractals Theory and applications, Chapman and Hall/CRC, New York, 2001.
[11] Matlab 6.1, “Image Processing Toolbox,” http://www.mathworks.com/access/helpdesk/help/toolbox /images/im ages. shtml
[12] Reginald L. Lagendijk, Jan Biemond, Iterative Identification and Restoration of Images, Kulwer Academic, Boston, 1991.
[13] J.N. Lin, X. Nie, and R. Unbehauen, “Two-Dimensional LMS Adaptive Filter Incorporating a Local-Mean Estimator for Image Processing,” IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing, Vol 40, No.7 July 1993, pg. 417-428.
[14] Jacques Lèvy Vèhel and Evelyne Lutton, “Evolutionary signal enhancement based on Hölder regularity analysis,” Project Fractales- INRIA, 2001.
[15] Mandelbrot, B., and Wallis, J., "Noah, Joseph and operational hydrology," Water Resources Research 4, 909-918, 1968.
[16] Mallat S.G, “A theory for multiresolution signal decomposition: The wavelet representation,” IEEE Trans. Pattn Anal. Mach. Intell., 11, 674- 693, 1989.
[17] Peter R. Massopust, Fractal Functions, Fractal Surfaces, and Wavelets, Academic press, San Diego, 1994.
[18] Matlab6.1, “Matlab,” http://www.mathworks.com/, May 2001.
[19] Wayne Niblack, An Introduction to Digital Image Processing, Prentice Hall, New Jersey, 1986.
[20] Rudolf H. Riedi, “Multifractals and Wavelets: A potential tool in Geophysics,” SEG Expanded Abstracts, Rice University, Houston, Texas, 1998.
[21] Image Processing Fundamentals-Statistics, “Signal to Noise ratio,” http://www.ph.tn.tudelft.nl/courses/FIP/noframes/fip-Statisti.html, 2001.
[22] Carl Taswell, “The What, How, and Why of Wavelet Shrinkage Denoising,”, http://www.toolsmiths.com/docs/CT199809.pdf, Technical Report, Stanford, CA, 1999.
[23] Tim Edwards, “Discrete Wavelet Transforms: Theory and Implementation,” Discrete Wavelet Transforms, Stanford University, Draft #2, June 4, 1992
[24] S. Theodoridis and K. Koutroumbas, Pattern Recognition, Academic Press, San Diego, CA, 1999.
[25] Scott E Umbaugh, Computer Vision and Image Processing, Prentice Hall PTR, New Jersey, 1998.
[26] Jacques Lèvy Vèhel, Bertrand Guiheneuf, “Multifractal image denoising,” Project Fractales-INRIA, April, 1997.
[27] Jacques Lèvy Vèhel, “Signal Enhancement Based on Hölder Regularity Analysis,” Project Fractales-INRIA, 2001.
[28] Jacques Lévy Véhel, “Fraclab,” www-rocq.inria.fr/fractales/, May 2000
[29] B. Vidakovic, Statistical modeling by wavelets, John Wiley and Sons, Inc. New York, 1999.
[30] Matlab 6.1, “Wavelet tool http://www.mathworks.com/access/helpdesk/help/toolbox /wavelet/w avelet .shtml