Wavelet based Multilevel Sub-Band Adaptive Thresholding for Image Denoising By PSO

BABITA SAXENA, PHD[1], DR. R. K. ARYA SR. SCIENTIST[2], DR. ASMITA M. MOGHE[3]
[1]*Scholar UIT-RGPV*
[2]*MPCST Bhopal*
[3]*Head (IT), UIT-RGPV*

Abstract— The image de-noising is one of the most widely studied area in the field of image processing. There are many ways (like communication channel, imperfect sensors, interference etc.) by which the noise may affect the image. Depending upon the nature of noise and the image many techniques has been already proposed. However, it is difficult for any technique to operate on wide of noises over different kind of images (like SAR images, X-ray images, Ultrasound images etc.). The best possible solution for such cases is to use adaptive techniques. In this paper we are presenting a multilevel wavelet decomposition based adaptive thresholding technique which utilizes the modified Particle Swarm Optimization (PSO) algorithm to find out the optimal values for thresholds and level of decompositions for given objective function. The modification of PSO is done through random perturbation in particle velocities which induces small randomness in new particle position estimation. This randomness can effectively increase the particle search space, which ultimately provide a much better solution than the conventional PSO. Finally, the proposed algorithm is validated by testing it over different kind of images corrupted by different values of noise.

Indexed Terms- Image De-noising, Wavelet Decomposition, Adaptive Thresholding, Particle Swarm Optimization (PSO).

I. INTRODUCTION

The recent development in digital imaging, multimedia technology and social networking has promoted the research in the field of image processing. Including the many image processing applications, such as compression, enhancement, and recognition, the noise removal preprocessing functions are essential in each application. It is the most common and essential processing steps in imaging systems. However, the image de-noising, is a tradeoff to be found between noise reduction and preserving important image details. To achieve a good tradeoff, a de-noising algorithm has to be capable to handle image discontinuities. The wavelet decomposition have the inherent characteristics which makes the construction of such spatially adaptive algorithms feasible. The wavelet transform arranges the essential information contained by image pixels into a number of bands called LL, LH, HL and HH. These coefficients can be further refined into different resolution scales. Because to such properties, the additive Gaussian noise can be effectively removed even by simple thresholding and shrinkage of the wavelet coefficients. In this paper, we further extend the adaptive thresholding approach. The main improvements are (1) sub-division of the image into smaller blocks which does not have too much pixel variations. (2) A joint level of decomposition and threshold search for each block and (3) a modified PSO algorithm for searching the optimal values for decomposition levels and thresholds. The paper is organized as follows. In Section II, similar literatures and concepts are briefly reviewed. A brief description of wavelet decomposition and thresholding is presented in Section III, Section IV discusses the PSO and the modified PSO algorithm. The image quality evaluation measures are presented in Section IV. The proposed technique and its practical implementation is described in Section V. The simulation results are presented and discussed in Section VI. Finally, concluding comments are given in Section VIII.

II. LITERATURE REVIEW

A wide range of image de-noising algorithms are based on the Discrete Wavelet Transform (DWT). In [7] shows that additive white Gaussian noise can be effectively removed even by simple thresholding of the wavelet coefficients. Other relatively simple shrinkage in presented in [2] which uses Bayes estimation assuming independent wavelet coefficients. Wavelet based de-noising techniques using image characteristics, inter-scale dependencies or intra-scale (spatial) correlations between image wavelet coefficients are presented respectively in [1][5][11]. Combined inter and intra-scale dependencies in a decimated, orthonormal wavelet basis is presented in [15] however it only sows minor improvement. Although, in a non-decimated wavelet basis [17], this technique gives a clear advantage in terms of quantitative image quality measures as well as in visual quality of the results. The technique presented in [17] combines the inter-scale and intra-scale dependencies. It uses the bi-level Markov Random Field (MRF) model to encode prior information about spatial clustering of wavelet coefficients. The inter-scale dependencies are estimated through measure of inter-scale ratios. The statistical properties of these measures are conveyed through a conditional probability density model, and combined with the prior model to form a Bayesian network. Since the conditional model from [17] is heuristic and parameterized, which makes it very complex for its practical implementation. Another Bayesian network-based approach in presented in [23] which investigated the statistical characterization of interscale ratios of wavelet coefficients, and different local criteria for distinguishing useful coefficients from noise are evaluated; after that a joint conditional model is introduced, and finally anisotropic Markov Random Field prior model used to de-noise the image. In [11] [12] a modified approach to [17] is presented. The authors in [11] and [12] did not used inter-scale statistics, but instead used the magnitude of wavelet coefficient as its significance estimate. The low complexity wavelet transform based image de-noising is proposed in [21] their model is inspired by simple wavelet image compression algorithm uses the Estimation Quantization coder. The proposed model assumes the wavelet image coefficients distribution as zero mean Gaussian with high local correlation. Assuming a marginal prior distribution on wavelet coefficients variances they try to estimate them using an approximate Maximum A Posteriori Probability rule. Finally an approximate Minimum Mean Squared Error estimation procedure is applied to restore the noisy wavelet image coefficients. Despite the simplicity of the presented method, it achieves the significantly good results. An expectation–maximization (EM) algorithm in wavelet domain for image restoration is presented in [22] the algorithm achieves the regularization by promoting a reconstruction with low complexity, wavelet coefficients, taking advantage of the sparsity of wavelet decomposition. The EM algorithm herein proposed to solve the resulting criteria more accurately then general approximation.

.

III. WAVELET THRESHOLDING AND THRESHOLD SELECTION

Let the signal be $\{f_{ij}, i, j = 1,2,3 \ldots\ldots, N\}$, where N can be integer power of 2, has been corrupted by additive white Gaussian noise (AWGN) and one observes

$$g_{ij} = f_{ij} + \epsilon_{ij}, \qquad i, j = 1,2,3 \ldots., N. \qquad (1)$$

Where $\{\epsilon_{ij}\}$ are independent and identically distributed (iid) as normal $N(0, \sigma^2)$ and independent of $\{f_{ij}\}$. The goal is to remove the noise, or "de-noise" $\{g_{ij}\}$, and to obtain an estimate $\{\hat{f}_{ij}\}$ of $\{f_{ij}\}$ which minimizes the mean squared error (MSE),

$$MSE(\hat{f}) = \frac{1}{N^2} \sum_{i,j=1}^{N} \left(\hat{f}_{ij} - f_{ij}\right)^2. \qquad (2)$$

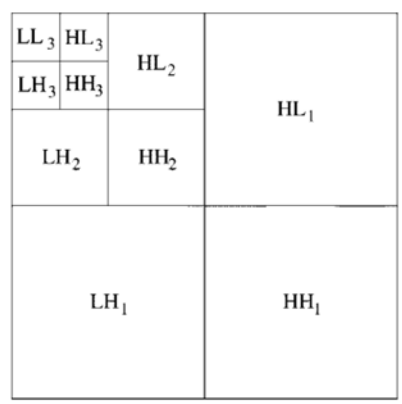

Figure 1. Sub-bands of the 2-D orthogonal wavelet transform.

Let $g = \{g_{ij}\}_{i,j}$, $f = \{f_{ij}\}_{i,j}$, and $\epsilon = \{\epsilon_{ij}\}_{i,j}$; that is, the boldfaced letters will denote the matrix representation of the signals under consideration. Let $Y = \mathbb{w}g$ denote the matrix of wavelet coefficients of g, where $\mathbb{w}$ is the two-dimensional dyadic orthogonal wavelet transform operator, and similarly

$= \mathbb{W}f$ and $V = \mathbb{W}\epsilon$. The readers are referred to references such as [23], [31] for details of the two-dimensional orthogonal wavelet transform. It is convenient to label the sub-bands of the transform as in Fig. 1. The sub-bands $HH_k, HL_k, LH_k, k = 1,2,3,\dots,J$ are called the details, where k is the scale, with J being the largest (or coarsest) scale in the decomposition, and a sub-band at scale k has size $N/2^k \times N/2^k$. The sub-band LL_j is the low-resolution residual, and J is typically chosen large enough such that $\frac{N}{2^J} \ll N$ and $\frac{N}{2^J} > 1$. Note that since the transform is orthogonal, $\{V_{ij}\}$ are also $iid\ N(0,\sigma^2)$.

The wavelet thresholding de-noising method filters each coefficient Y_{ij} from the detail subbands with a threshold function (to be explained shortly) to obtain $\hat{X}_{ij}$ (estimated de-noised coefficients). The de-noised estimate is then $\hat{f} = \mathbb{W}^{-1}\hat{X}$, where $\mathbb{W}^{-1}$ is the inverse wavelet transform operator. There are two thresholding methods frequently used.
The soft-threshold function (also called the shrinkage function)

$$\eta_T(x) = sgn(x) \cdot Max(|x| - T, 0) \qquad (3)$$

Takes the argument and shrinks it toward zero by the threshold T. The other popular alternative is the hard-threshold function

$$\psi_T(x) = x \cdot 1\{|x| > T\} \qquad (4)$$

Which keeps the input if it is larger than the threshold T ; otherwise, it is set to zero.

The wavelet thresholding procedure removes noise by thresholding only the wavelet coefficients of the detail sub-bands, while keeping the low-resolution coefficients unaltered. The soft-thresholding rule is chosen over hard-thresholding for several reasons. First, soft-thresholding has been shown to achieve near-optimal mini-max rate over a large range of Besovspaces [12], [14]. Second, for the generalized Gaussian prior assumed in this work, the optimal soft-thresholding estimator yields a smaller risk than the optimal hard-thresholding estimator. Lastly, in practice, the soft-thresholding method yields more visually pleasant images over hard-thresholding because the latter is discontinuous and yields abrupt artifacts in the recovered images, especially when the noise energy is significant. In what follows, soft-thresholding will be the primary focus.

While the idea of thresholding is simple and effective, finding a good threshold is not an easy task. For one-dimensional (1-D) deterministic signal of length M, Donoho and Johnstone [14] proposed for VisuShrink the universal threshold, $T_U = \sigma\sqrt{2logM}$, which results in an estimate asymptotically optimal in the minimax sense (minimizing the maximum error over all possible -sample signals). One other notable threshold is the SURE threshold [15], derived from minimizing Stein's unbiased risk estimate [30] when soft-thresholding is used. The SureShrink method is a hybrid of the universal and the SURE threshold, with the choice being dependent on the energy of the particular subband [15]. The SURE threshold is data-driven, does not depend on explicitly, and SureShrink estimates it in a subband-adaptive manner. Moreover, SureShrink has yielded good image de-noising performance and comes close to the true minimum MSE of the optimal soft-threshold estimator (cf. [4], [12]), and thus will be the main comparison to our proposed method.

IV. PSO AND THE MODIFIED PSO ALGORITHM

4.1 Standard PSO Algorithm

The PSO algorithm is inspired by the natural swarm behavior of birds and fish. It was introduced by Eberhart and Kennedy in 1995. In PSO each particle in the population represents a possible solution of the optimization problem, which is defined by its objective/cost/fitness function. In each iteration, a new position of particles are calculated based on its last location and velocity.

Initially, the PSO algorithm deploy the particles randomly within the search space, then it simply uses the objective function to estimate the fitness of each particle (solution). Each particle maintains its position, fitness, velocity, and the best fitness value it has achieved also known as particle best or individual best solution. Finally, the PSO algorithm estimates the global best solution (particle position which gives minimum/maximum fitness value among all particles in the swarm).

The PSO algorithm can be explained in following steps:

Step 1: Let the position and velocity of particles at k^{th} iteration is given by

$$P^k = \{p_1^k, p_2^k, p_3^k, \dots .. p_N^k\}$$
$$V^k = \{v_1^k, v_2^k, v_3^k, \dots .. v_N^k\}$$

Where N shows the number of particles.

Step 2: Then fitness of each particle at k^{th} iteration

$$F^k = \{f_1^k, f_2^k, f_3^k, \dots \dots \dots f_N^k\}$$
$$f_i^k = fun_{objective}(x_i^k)$$

Step 3: calculate the particle best (P_{pest}^k) and global best (G_{best}^k) as follows:

$$P_{best}^k = \{p_{best,1}^k, p_{best,2}^k, \dots \dots \dots \dots, p_{best,N}^k\}$$
$$p_{best,i}^k = \min\{p_{best,i}^0, p_{best,i}^1, \dots \dots, p_{best,i}^k\}$$
$$G_{best}^k = \min\left\{\begin{matrix} f_1^0, f_2^0, \dots, f_N^0, f_1^1, f_2^1, \dots, f_N^1, \dots \dots, \\ f_1^k, f_2^k, \dots ., f_N^k \end{matrix}\right\}, \quad i = 1,2,3 \dots, N.$$

Step 4: Update velocity and position of each particle

$$v_i^k = w^{k-1} * v_i^{k-1} + c_1 \cdot r_1\left(p_{best,i}^{k-1} - x_i^{k-1}\right) + c_2 \cdot r_2\left(G_{best}^{k-1} - x_i^{k-1}\right)$$

Where w represents the inertia weight. While r_1 and r_2 are the two uniformly distributed random numbers having mean of 0 and variance of 1.

Step 5: Update inertia and positions of each particle

$$w^k = w_{start} - \frac{\left((w_{start} - w_{end}) * k\right)}{iter_{max}}$$
$$x_i^k = x_i^{k-1} + v_i^k$$

The value of the inertial coefficient w_{start} and w_{end} are typically chosen between 0.8 and 1.2.

Step 6: Repeat the whole process till the stopping criteria meets.

4.1 Modified PSO Algorithm

In the modified algorithm the calculation of inertia component w^k is modified by the following equation:

$$w^k = R^k * \left(w_{start} - \frac{\left((w_{start} - w_{end}) * k\right)}{iter_{max}}\right)$$

Where R^k is calculated as

$$R^k = \mu R^{k-1}(1 - R^{k-1}), R^0 = 0.63, \mu = 4$$

Which represents the logistic map of degree 2, and known for its complex, chaotic behavior from very simple non-linear dynamical equations.

The modification increases the particle search space, which ultimately provide a much better solution than the conventional PSO.

V. IMAGE QUALITY MEASURES

The image quality measures are very important factor in this particular problem because this term is used as fitness value in proposed algorithm. Hence if improper measure is set then the possibility of finding good results can greatly fade. Hence in this paper we prefer the use of most common measure known as Mean Squared Error (MSE), which is defined as:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X(i,j) - Y(i,j)\right)^2$$

Where M and N are rows and columns in image respectively. X and Y represents the original and De-noised images.

Another measure known as Peak Signal to Noise Ratio ($PSNR$) could also be taken although it is derived from MSE

$$PSNR(dB) = 10\log_{10}\left(\frac{255^2}{MSE}\right)$$

However we prefer the MSE because it minimizes with better quality and give a much larger scale resolution.

VI. PROPOSED TECHNIQUE

Since the variables in sub-band adaptive thresholding image de-noising problem are level of decompositions and threshold for each block $L = \{l_1, l_2, l_3, \dots \dots l_M\}, T = \{t_1, t_2, t_3, \dots \dots t_M\}$, where M is total number of blocks. Hence the vector of a particle is group of elements corresponding to $[L, T]$. Therefore, the position of any particle at iteration k can be represented as the vector $p_i^k = \{l_1^k, l_2^k, \dots .., l_M^k, t_1^k, t_2^k, \dots \dots . t_M^k\}$ which sets the dimention of the vector equal to $2M$.

Now if the initial population of particle be N then the complete set can be presented as:

$$P^k = \{p_1^k, p_2^k, p_3^k, \dots .. p_N^k\}$$
$$p_i^k = \{l_1^k, l_2^k, \dots .., l_M^k, t_1^k, t_2^k, \dots \dots . t_M^k\}$$
$$l_{min} \le l_i^k \le l_{max}, \ for\ \forall i, k$$
$$t_{min} \le t_i^k \le t_{max}, \ for\ \forall i, k$$

Where l_{min}, t_{min} and l_{max}, t_{max} are the user defined lower and upper bounds for level of decompositions and threshold value.

Using the above mapping we can call the PSO algorithm to find the best values for L and T as follows:

VII. SIMULATION RESULTS

In this section the following algorithms are tested over different types of images as shown in figure 3,

- The Conventional Genetic Algorithm
- The Modified PSO Algorithm

The simulation of all algorithms is performed using MATLAB. The population size N_P and maximum iteration number $iter_{max}$ are set as 100 and 100, respectively. w_{max} and w_{min} are set to 0.9 and 0.1 respectively because these values are widely accepted and verified in solving various optimization problems. The list of all values used for the system are shown in the table below

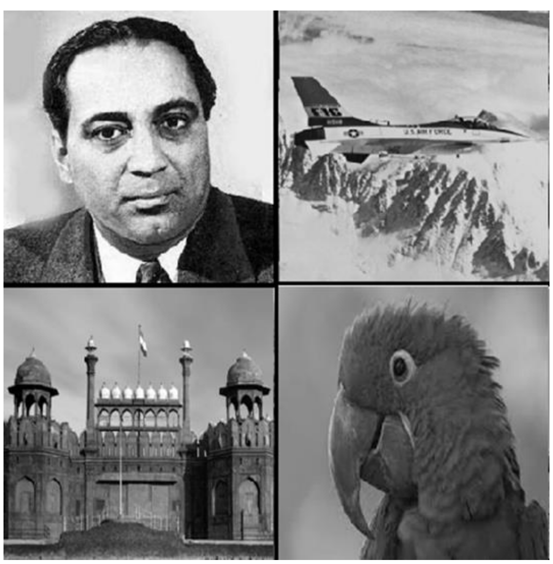

Figure 4: images used as Watermark Homi Bhabha, Airplane, Redfort and Bird.

CONCLUSION

This paper presents a Modified PSO based sub-band multilevel adaptive thresholding technique for image denoising. The proposed method exploits strength of Wavelet domains transform and capability of MPSO to obtain the good filtration results. The idea of searching of threshold value in the different decomposition component is based on the fact that property of wavelet coefficients depends upon the nature of image. Implementation results show that the impact of is greatly reduced as compare to previous method. Presented method is also tested for most of the common images, and the simulation results show that proposed method is more efficient and robust compare to previous method.

REFERENCES

[1] Tomáš Filler, Jan Judas, and Jessica Fridrich "Minimizing Additive Distortion in Watermarking using Syndrome-Trellis Codes", Information Forensics and Security, IEEE Transactions on (Volume:6 , Issue: 3),Sept. 2011

[2] H.A. Chipman, E.D. Kolaczyk, and R.E. McCulloch, "Adaptive Bayesian Wavelet Shrinkage," J. of the Amer. Statist. Assoc., vol. 92, pp. 1413–1421, 1997.

[3] A. Cohen and J. Kovaˇcevi´c, "Wavelets: The Mathematical Back-ground," Proc. of the IEEE, vol. 84, pp. 514–522, 1996.

[4] R.R. Coifman and D.L. Donoho, "Translation-invariant denoising," in Wavelets and Statistics, A. Antoniadis and G. Oppenheim, editors, pp. 125–150, New York: Springer Verlag, 1995.

[5] M. S. Crouse, R. D. Nowak, and R. G. Baranuik, "Wavelet-Based Statistical Signal Processing Using Hidden Markov Models," IEEE Trans. on Signal Proc., vol. 46, pp. 886–902, Apr 1998.

[6] I. Daubechies, Ten Lectures on Wavelets, Philadelphia: SIAM,1992.

[7] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," J. Amer. Stat. Assoc., vol. 90, pp. 1200–1224, 1995.

[8] R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons, 1973.

[9] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," IEEE Trans.on Pattern Anal. and Machine Intel., vol. 6 pp. 721–741, Nov 1984.

[10] T.-C. Hsung, D. P.-K Lun, and W.-C. Siu. “Denoising by singularity detection,” IEEE Trans. on Signal Proc., vol. 47, pp. 3139–3144, Nov 1999.

[11] M. Jansen and A. Bultheel, “Geometrical Priors for Noise-free Wavelet Coefficient Configurations in Image De-noising,” in Bayesian inference in wavelet based models, P. M¨uller and B. Vidakovi´c, editors, Springer Verlag 1999, pp. 223–242.

[12] M. Jansen and A. Bultheel, “Empirical Bayes approach to improve wavelet thresholding for image noise reduction,” J. of the Amer. Statist. Assoc., vol. 96, no. 454, pp. 629-639, 2001.

[13] M. Kivan¸c Mih¸cak, I. Kozintsev, K. Ramchandran, and P. Moulin, “Low-Complexity Image Denoising Based on Statistical Modeling of Wavelet Coefficients,” IEEE Signal Proc. Letters, vol. 6, pp. 300–303, Dec 1999.

[14] S. Z. Li, Markov Random Field Modeling in Computer Vision, Springer-Verlag, 1995.

[15] J. Liu and P. Moulin, “Analysis of interscale and intrascale dependencies between image wavelet coefficients”, in Proc. Int. Conf. on Image Proc., (ICIP), Vancouver, Canada, Sep. 2000.

[16] D.J.C. MacKay, “Introduction to Monte Carlo Methods,” in Learning in Graphical Models, M. I. Jordan, editor, MIT Press, 1999.

[17] M. Malfait and D. Roose, “Wavelet-Based Image Denoising Us-ing a Markov Random Field a Priori Model,” IEEE, Trans. On Image Proc., vol. 6, pp. 549–565, April 1997.

[18] S. Mallat, “A theory for Multiresolution Signal Decomposition: The Wavelet Representation,” IEEE Trans. Pattern Anal. And Machine Intel., vol. 11, no. 7, pp. 674–693, July 1989.

[19] S. Mallat and S. Zhong, “Characterization of Signals from Multi-scale Edges,” IEEE Trans. on Pattern Anal. and Machine Intel., vol. 14, pp. 710–732, July 1992.

[20] S. Mallat and W. L. Hwang, “Singularity Detection and Pro-cessing with Wavelets,” IEEE Trans on Inform. Theory, vol. 38,pp. 617–643, Mar 1992.

[21] M. Kıvanc Mihcak, Igor Kozintsev, Kannan Ramchandran and Pierre Moulin “Low-complexity image denoising based on statistical modeling of wavelet coefficients”, Signal Processing Letters, IEEE (Volume:6 , Issue: 12) DEC. 1999.

[22] Mário A. T. Figueiredo, and Robert D. Nowak “An EM Algorithm for Wavelet-Based Image Restoration”, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 12, NO. 8, AUGUST 2003.

[23] Aleksandra Pizurica, Wilfried Philips, Ignace Lemahieu, and Marc Acheroy “A Joint Inter- and Intrascale Statistical Model for Bayesian Wavelet Based Image Denoising”, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 11, NO. 5, PAGES 545–557, MAY 2002.