

IMAGE DENOISING USING WAVELET EMBEDDED ANISOTROPIC DIFFUSION (WEAD)

Jeny Rajan*, M.R Kaimal †

*Network Systems & Technologies Ltd (NeST), Technopark Campus, Trivandrum INDIA, Email :jenyrajan@rediffmail.com

†Dept. Of Computer Science, University of Kerala, Trivandrum, INDIA , Email :mrkaimal@yahoo.com

Keywords: Anisotropic Diffusion, Bayesian Shrinkage, Denoising, Wavelets.

Abstract

In this paper a PDE based hybrid method for image denoising is introduced. The method is a bi-stage filter with anisotropic diffusion followed by wavelet based bayesian shrinkage. Here efficient denoising is achieved by reducing the convergence time of anisotropic diffusion. As the convergence time decreases, image blurring can be restricted and will produce a better denoised image than anisotropic or wavelet based methods. Experimental results based on PSNR, SSIM and edge analysis shows excellent performance of the proposed method.

1 Introduction

Image denoising has a significant role in image pre processing. As the application areas of image denoising are more, there is a big demand for efficient denoising algorithms. In this work, we developed a new method; Wavelet Embedded Anisotropic Diffusion (WEAD), and applied it to denoise images corrupted with additive Gaussian noise. The intention behind this method is to reduce the convergence time of anisotropic diffusion and thereby increase its performance. The proposed method produces excellent results when compared with various wavelet shrinkage [6], [7], [8] and non-linear diffusion methods [1], [2].

Second order partial differential equations have been used as efficient methods for removing noise from images. One of the most commonly used PDE based denoising technique since its introduction is the Perona-Malik method [1]. The Perona – Malik equation for an image u is given by

$$\frac{\partial u}{\partial t} = \text{div}[c(\nabla u)\nabla u] , u(x, y)|_{t=0} = u_0(x, y) \quad (1)$$

where ∇u is the gradient of the image u , div is the divergence operator and c is the diffusion coefficient. The diffusion coefficient c is a non-decreasing function and diffuses more on plateaus and less on edges and thus edges are preserved. Another objective for the selection of $c(\cdot)$ is to incur backward diffusion around intensity transitions so that edges are sharpened, and to assure forward diffusion in smooth

areas for noise removal [2]. Some of the commonly employed diffusivity functions are given in [12]. The equation (1) is studied as an efficient tool for noise removal and scale space analysis of images.

Wavelet based methods are always a good choice for image denoising and has been discussed widely in literatures for the past two decades [6]-[8],[11],[12]. Wavelet shrinkage permits a more efficient noise removal while preserving high frequencies based on the disbalancing of the energy of such representations [4]. The technique denoises image in the orthogonal wavelet domain, where each coefficient is thresholded by comparing against a threshold; if the coefficient is smaller than the threshold, it is set to zero, otherwise it is kept or modified.

Wavelet shrinkage depends heavily on the choice of a thresholding parameter and the choice of this threshold determines, to a great extent the efficacy of denoising. The denoising process is based on the fact that the wavelet transform compresses most of the L^2 energy of the signal in a restricted number of large coefficients. The procedure can be summarized in three steps

$$\begin{aligned} Y &= W(x) \\ Z &= T(Y, \lambda) \\ Y' &= W^{-1}(Z) \end{aligned} \quad (2)$$

where x is the affected signal, $W(\cdot)$ and W^{-1} is the forward and inverse wavelet transform operators. $T(Y, \lambda)$ denotes the denoising operator with soft or hard threshold λ . Of the various methods based on wavelet thresholding, VisuShrink[6], SureShrink[7], BayesShrink[8] and its variants are the most popular. VisuShrink uses one of the well known thresholding rules: the universal threshold. In addition, subband adaptive systems have superior performance, such as SureShrink, which is a data driven system. Recently, BayesShrink [8], which is also a data driven subband adaptive technique, is proposed and outperforms VisuShrink and SureShrink. In the proposed method BayesShrink is used along with anisotropic diffusion to get a better performance than stand alone Anisotropic diffusion or BayesShrink.

This work does not attempt to investigate in deep the theoretical properties of the proposed model in general settings. Our primary goal is to demonstrate that how the performance of PDE based denoising methods can be improved by using the proposed hybrid method. The paper is

organized as follows. In section II Bayesian denoising technique is discussed. Section III and IV explains the proposed method, experimental results and comparison of the proposed method with other popular models. Finally conclusion and remarks are included in section V.

2 Denoising Using Bayesian Shrinkage

The Bayesian Shrinkage estimates a soft-threshold that minimizes the Bayesian risk. The Bayesian risk estimation is subband dependent. The threshold is mathematically derived in [8]. The generalized Gaussian distribution (GCD), following [9] is

$$GG_{\sigma X, \beta(x)} = C(\sigma X, \beta) \exp \{-[\alpha(\sigma X, \beta) |x|]^\beta\} \quad (5)$$

$-\infty < x < \infty, \sigma X > 0, \beta > 0$, where

$$\alpha(\sigma X, \beta) = \sigma_x^{-1} \left[\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)} \right]^{1/2} \quad (6)$$

and

$$C(\sigma X, \beta) = \frac{\beta \alpha(\sigma X, \beta)}{2\Gamma\left(\frac{1}{\beta}\right)} \quad (7)$$

and

$$\Gamma(t) = \int_0^\infty e^{-u} u^{t-1} du \quad (8)$$

is the gamma function. The parameter σX is the standard deviation and β is the shape parameter. Here the objective is to find a soft threshold T that minimizes the Bayes risk,

$$r(T) = E\left(\hat{X} - X\right)^2 = E_x E_{y|x} \left(\hat{X} - X\right)^2 \quad (9)$$

where $\hat{X} = \eta T(Y), Y | X \approx N(x, \sigma^2)$ and $X \approx GG_{\sigma X, \beta}$.

Denote the optimal threshold by T^*

$$T^*(\sigma X, \beta) = \arg \min_T r(T) \quad (10)$$

which is a function of the parameters σX and β . In [8] it is shown that for general σ , the threshold T_B can be written as

$$T_B(\sigma X) = \frac{\sigma^2}{\sigma X} \quad (11)$$

where σ^2 is the noise variance and σX^2 the signal variance. By using the threshold in (11), we can restore the image much better than by using VisuShrink or Sure Shrink, where VisuShrink uses a threshold choice, $\sigma\sqrt{2 \log M}$. This can be unwarrantedly large due to its dependence on the number of samples. SureShrink uses a hybrid of the universal threshold and the SURE threshold, derived from Stein's unbiased risk estimator.

3 Proposed Model

In the proposed model the Bayesian Shrinkage of the non-linearly diffused signal is taken. The equation can be written as

$$I_n = B_s(I_{n-1}) \quad (12)$$

where B_s is the Bayesian shrink and I_{n-1} is anisotropic diffusion as shown in (1) at $(n-1)^{th}$ time. Numerically (12) can be written as

$$I_n = B_s(I_{n-1} + \Delta t d_n) \quad (13)$$

where B_s can be calculated by finding T_b as mentioned in eqn (11) after taking wavelet transform of I_{n-1} .

The intention to develop this method is to decrease the convergence time of the anisotropic diffusion. It is understood that the convergence time for denoising is directionally proportional to the image noise level. In the case of anisotropic diffusion, as iteration continues, the noise level in image decreases (till it reaches the convergence point), but in a slow manner. But in the case of Bayesian shrinkage, it just cut the frequencies above the threshold and that in a single step. An iterative Bayesian Shrinkage will not incur any change in the detail coefficients from the first one. Now consider the proposed algorithm, here the threshold for Bayesian shrinkage is recalculated each time after anisotropic diffusion, and as a result of two successive noise reduction step, it approaches the convergence point much faster than anisotropic diffusion.

As the convergence time decreases, image blurring can be restricted, and as a result image quality increases. The whole process is illustrated in Fig.2. Fig.2(a) shows the convergence of the image processed by Perona-Malik anisotropic diffusion. The convergence point is at P . ie. at P we will get the better image, with the assumption that the input image is a noisy one. If this convergence point P can be shifted towards y-axis, its movement will be as shown in Fig 2 (b).ie. if we pull the point P towards y-axis, it will move in a left-top fashion. Here the Bayesian shrinkage is the catalyst, which pulls the convergence point P of the anisotropic diffusion towards a better place. The method can be extended to other PDE based methods like fourth order PDEs, Total-variation minimization, Complex diffusion etc.

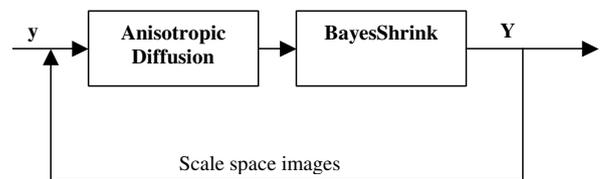


Fig 1: Block diagram of the proposed denoising algorithm. The iteration process will continue till the input signal y is converged to Y .

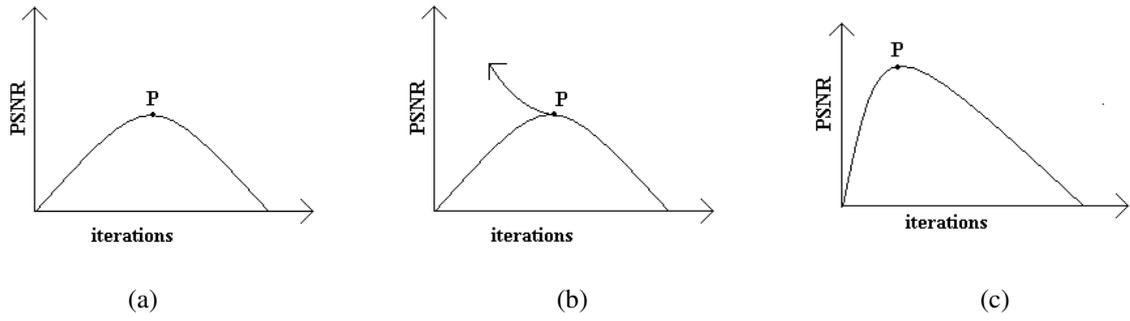
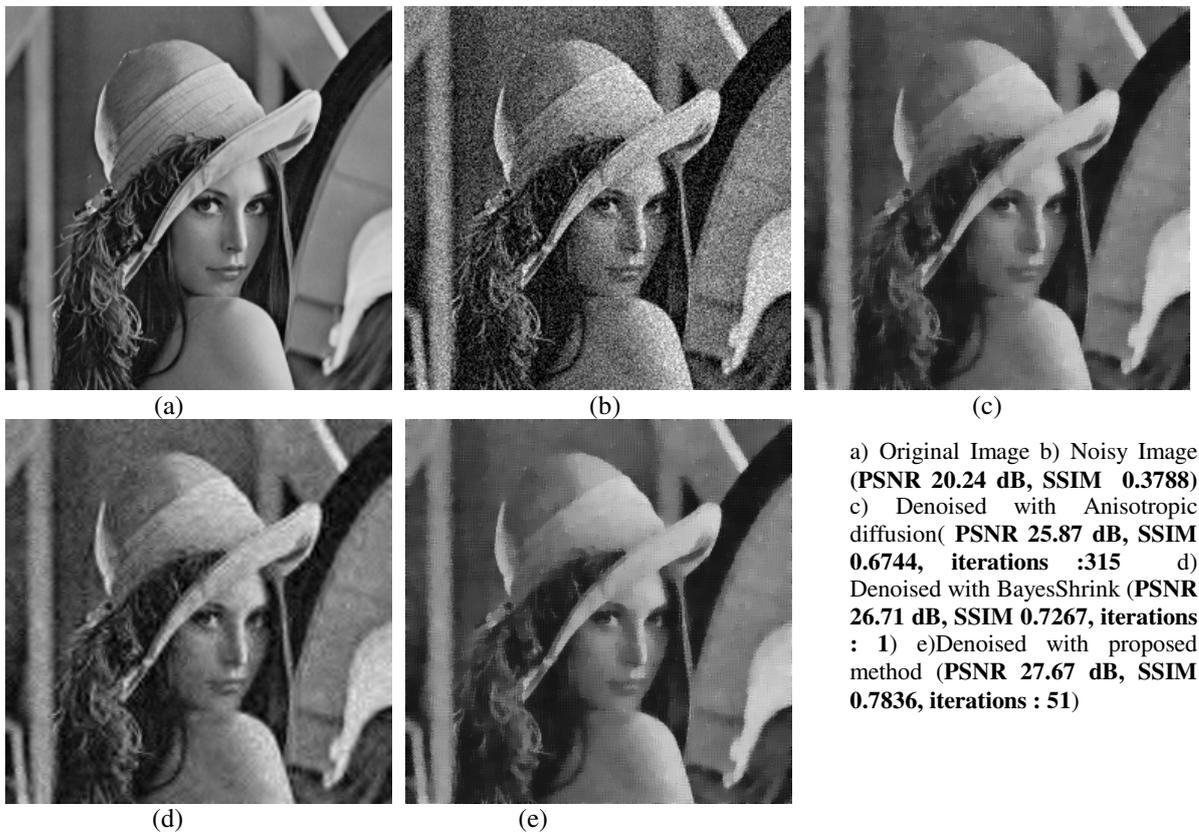


Fig 2: Working of WEAD (a) Shows the convergence of a noisy image (convergence at P). If this P can be shifted towards left, image quality can be increased and time complexity can be reduced. Illustrated in (b). (c) shows the signal processed by WEAD. It can be seen that the convergence point is shifted to left and moved upwards.



a) Original Image b) Noisy Image (PSNR 20.24 dB, SSIM 0.3788)
 c) Denoised with Anisotropic diffusion(PSNR 25.87 dB, SSIM 0.6744, iterations :315
 d) Denoised with BayesShrink (PSNR 26.71 dB, SSIM 0.7267, iterations : 1)
 e)Denoised with proposed method (PSNR 27.67 dB, SSIM 0.7836, iterations : 51)

Fig 3: Noise removal with Anisotropic Diffusion, BayesShrink and Propose method

4 Experimental Results & Comparative Analysis

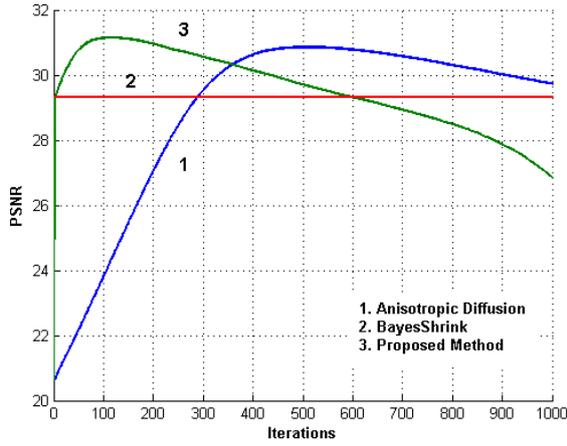
Experiments were carried out on various types of images. Comparisons and analysis were done on the basis of MSSIM (Mean Structural Similarity Index Matrix) and PSNR (Peak Signal to Noise Ratio)

The MSSIM[10] is used to evaluate the overall image quality and is defined as

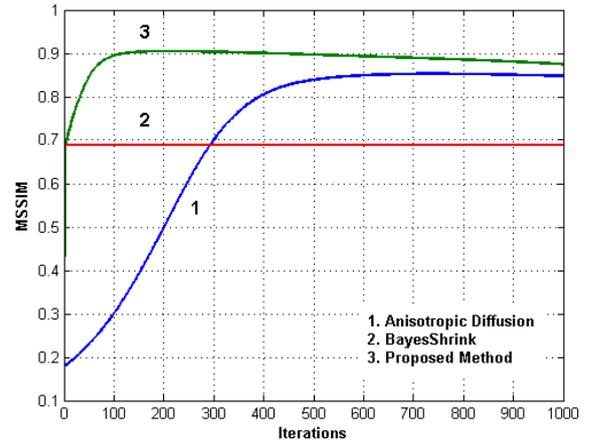
$$MSSIM(X, Y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j) \quad (14)$$

where X and Y are the reference and the distorted images respectively, M is the number of local windows in the image, SSIM is the Structural Similarity Index Matrix, x_j and y_j are the image contents at the j^{th} local window. The Structural Similarity Index Matrix (SSIM) is defined as

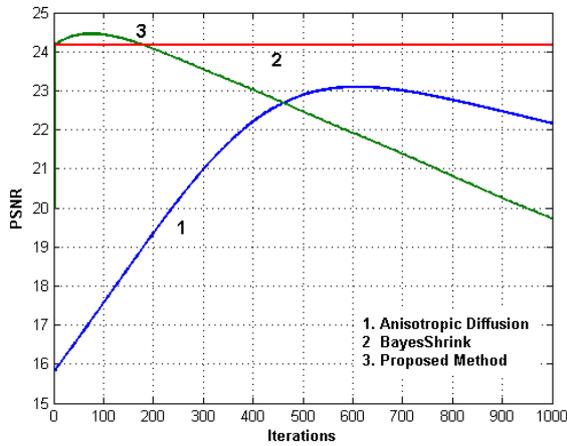
$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (15)$$



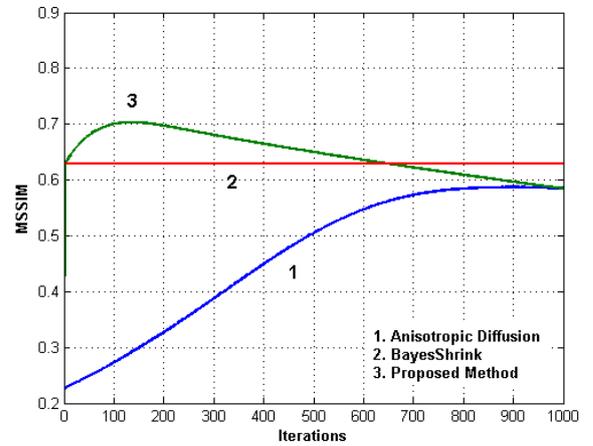
(a)



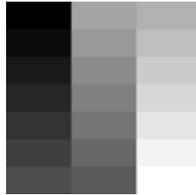
(b)



(c)



(d)



(e)



(f)

Fig 4. Comparative Analysis of Anisotropic Diffusion, BayesShrink and Proposed Method. (a) and (c) based on PSNR and (b) and (d) based on Mean SSIM. For (a) and (b) the image used is gray (shown in (e)) and for (c) and (d) the image used is Lena (shown in (f)). It can be seen that in both cases the proposed method performs better than the other two.

where μ_x and μ_y are the estimated mean intensity along x and y directions and σ_x and σ_y are the standard deviation respectively. σ_{xy} can be estimated as

$$\sigma_{xy} = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \right) \quad (16)$$

C_1 and C_2 in (15) are constants and the values are given as

$$C_1 = (K_1 L)^2 \quad (17)$$

and

$$C_2 = (K_2 L)^2 \quad (18)$$

where $K_1, K_2 \ll 1$ is a small constant and L is the dynamic range of the pixel values (255 for 8 bit grayscale images).

The second parameter used for evaluation is the PSNR which is defined as

$$PSNR = 20 \log_{10} \left(\frac{b}{rms} \right) \quad (19)$$

where b is the largest possible value of the signal (typically 255 or 1), and rms is the root mean square difference between two images.

Fig 3 shows the image denoised with anisotropic diffusion, Bayesshrink and proposed method. It is observed that the proposed method reduces the number of iterations of anisotropic diffusion from 315 to 51 and improves the image quality. It can be seen that there is 10% improvement over anisotropic diffusion and around 5% over Bayesshrink in preserving image structure. Based on PSNR also it can be seen that the proposed method performs better than the other two. The graphs in Fig.4 shows comparative analysis of anisotropic diffusion, bayesshrink and proposed method. It is clear that the performance of the methods depends on image type and noise levels. But in both cases, whether anisotropic diffusion or bayesshrink gives better results, the performance of the proposed method seems to be much better than the other two. It can be seen that the proposed method preserves image structures much better than anisotropic diffusion and bayesshrink. Also the number of iterations required for the proposed method to produce the better image is much less than that of anisotropic diffusion. The experiment is repeated for various types of images with varying noise levels and seems that the method proposed is giving better results than anisotropic diffusion and bayesshrink.

Conclusion

A method to improve the performance of nonlinear anisotropic diffusion is proposed. The method produces a converged image with less number of iterations preserving image edges better than anisotropic diffusion.

Acknowledgements

The authors would like to thank the Scientists of ADRIN (Hyderabad, India) for helpful discussions. They would also like to thank the authorities of NeST (Trivandrum, India) for providing the necessary facilities for doing the experiments.

References

- [1] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion", *IEEE Trans. Pattern Analysis and Machine Intelligence.*, vol 12. No. 12, pp 629-639, July 1990.
- [2] Yu-Li You, Wenguan Xu, Allen Tannenbaum and Mostafa Kaveh, "Behavioral Analysis of Anisotropic Diffusion in Image Processing", *IEEE Trans. Image Processing*, vol. 5, no. 11, pp 1539-1553, November 1996.
- [3] W. Rudin, "Real and Complex Analysis", New York : McGraw-Hill, 1996.
- [4] Gabriel Cristobal, Monica Chagoyen, Boris Escalante-Ramirez, Juan R Lopez, "Wavelet-based denoising methods, A comparative study with applications in microscopy (Unpublished work style)," Proc. SPIE's International Symposium on Optical Science, Engineering and Instrumentation, Wavelet Applications in Signal and Image Processing IV, Vol. 2825, Denver, CO. 1996.

- [5] Raghuram Rangarajan, Ramji Venkataramanan, Siddharth Shah, "Image Denoising Using Wavelets", ---, 2002.
- [6] David L Donoho, "Ideal spatial adaptation by wavelet shrinkage", *Biometrika*, 81(3) : 425-455, August 1994.
- [7] 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
- [8] 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.
- [9] R.L. Joshi, V.J. Crump and T.R. Fisher, "Image subband coding using arithmetic and trellis coded quantization", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp 515-523, Dec. 1995.
- [10] Zhou Wang, Alan Conard Bovik, Hamid Rahim Sheik and Erno P Simoncelli, "Image Quality Assessment : From error visibility to structural similarity", *IEEE Trans. Image Processing*, Vol. 13, No. 4, 2004.
- [11] R.R Cofiman and David L. Donoho, "Translation invariant denoising", *Wavelets and Statistics*, Lecture Notes in Statistics, Springer Verlag, 1995.
- [12] Pavel Mrazek, Joachim Weickert and Gabriele Steidl, "Correspondence between Wavelet Shrinkage and Nonlinear Diffusion", *Scale-Space 2003*, LNCS 2695, pp. 101-116, 2003