

DENOISING OF MEDICAL IMAGES USING TOTAL VARIATIONAL METHOD

V N Prudhvi Raj¹ and Dr T Venkateswarlu²

¹Associate Professor, VR Siddhartha Engineering College, Vijayawada, 520007, India
nagaprudhvi@yahoo.com

²Professor, University College of Engineering, SV University, Tirupati, India

ABSTRACT

Feature extraction and object recognition from images acquired by various imaging modalities are playing the key role in diagnosing the various diseases. These operations will become difficult if the images are corrupted with noise. So the need for developing the efficient algorithms for noise removal became an important research area today. Developing Image denoising algorithms is a difficult operation because fine details in a medical image embedding diagnostic information should not be destroyed during noise removal. In this paper the total variational method which had success in computational fluid dynamics is adopted to denoise the medical images. We are using split Bregman method from optimisation theory to find the solution to this non-linear convex optimisation problem. The present approach will outperform in denoising the medical images while compared with the traditional spatial domain filtering methods. The performance metrics we used to measure the quality of the denoised images is PSNR (Peak signal to noise ratio). The results showed that these methods are removing the noise effectively while preserving the edge information in the images.

KEYWORDS

Total Variation, Convex Optimisation, Split Bregman Iteration, PSNR.

1. INTRODUCTION

Medical information, composed of clinical data, images and other physiological signals, has become an essential part of a patient's care, during screening, in the diagnostic stage and in the treatment phase. Over the past three decades, rapid developments in information technology (IT) & Medical Instrumentation has facilitated the development of digital medical imaging. This development has mainly concerned Computed Tomography (CT), Magnetic Resonance Imaging (MRI), the different digital radiological processes for vascular, cardiovascular and contrast imaging, mammography, diagnostic ultrasound imaging, nuclear medical imaging with Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET). All these processes are producing ever-increasing quantities of images [17]. These images are different from typical photographic images primarily because they reveal internal anatomy as opposed to an image of surfaces.

In Natural monochromatic or colour images, the pixel intensity of the image corresponds to the reflection coefficient of natural light. Whereas images acquired for clinical procedures reflect very complex physical and physiological phenomena, of many different types, hence the wide variety of images. Each medical imaging modality (digital radiology, computerized tomography (CT), magnetic resonance imaging (MRI), ultrasound imaging (US)) has its own specific features corresponding to the physical and physiological phenomena studied, as shown in "Fig.1". These

medical images have their own unique set of challenges. Although our focus in this paper will be on two-dimensional images, three-dimensional (volume) images, time-varying two-dimensional images (movies), and time-varying three-dimensional images are commonly used clinically as imaging modalities are becoming more sophisticated [17].

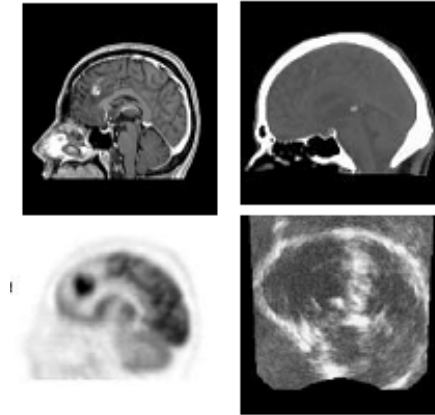


Figure1: Sagittal slices of the brain by different imaging modalities: a) magnetic resonance imaging (MRI), b) computed tomography (CT), c) positron emission tomography (PET), d) ultrasound (US)

1.1 Image Degradation and Restoration Model

Images are degraded in the process of acquisition and transmission through the communication media (Wired or Wireless). Acquisition modality may introduce some blur due to the poor dynamic range, poor calibration of instrument and if the object is in motion or the acquisition device is not handled by the operator while acquiring the images. During the transmission phase and acquisition phase some amount of noise is also added to these images. As a result of blur and noise the details of the image is going to be distorted which will create the problems in diagnostic stage [17]. So the process of removing blur and noise from the images (Image Restoration) is considered as a preprocessing step before going to the image analysis. The image degradation and restoration model is given in the “Fig.2”.

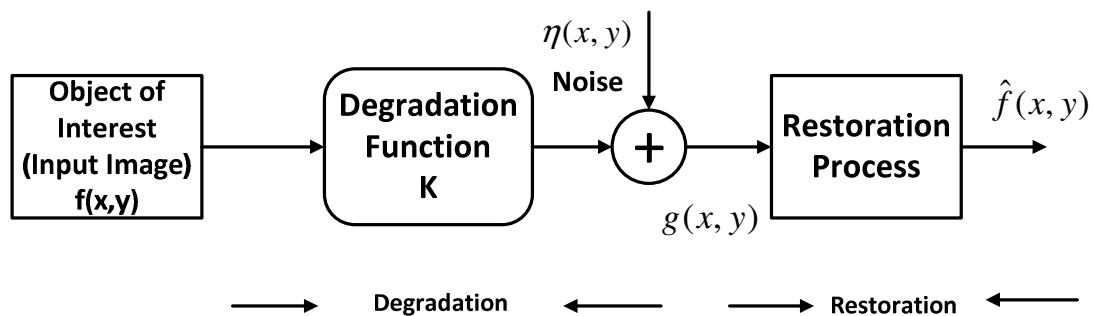


Figure 2: Image Degradation and Restoration model

The noise term may be additive, multiplicative or combination of both. In case of medical images we have both additive and multiplicative noise depending upon the modalities used for image acquisition. In general the noise generated due to the electronic components in an acquisition system is modeled with Gaussian noise which is an additive term. In the X-ray imaging the quantum noise or photon noise is present due to random generation of X-rays from the source at

any given time; this noise can be modeled using Poisson distribution [2]. In the case of ultrasound imaging images are degraded by the signal dependent noise known as speckle. The pattern of the speckle depends on the structure of the imaging tissue and various imaging parameters. The speckle noise is a multiplicative noise which can be modeled using Rayleigh distribution. The Magnetic resonance imaging (MRI) is suffering from Rician noise which can be modeled using Rician distribution [17]. In this paper we are trying to denoise the images corrupted with Gaussian noise, Laplacian noise and Poisson noise.

The mathematical modeling of degradation and restoration process is given as

$$\begin{aligned} g(x, y) &= f(x, y) * k(x, y) + \eta(x, y) \\ G(u, v) &= F(u, v) K(u, v) + N(u, v) \end{aligned} \quad (1)$$

Where K is the blurring kernel, $g(x, y)$ is the noisy and blurred observation and $f(x, y)$ is the signal we are recovering. In the case of denoising problem the blurring kernel will be dropped and the degradation model will be given as

$$\begin{aligned} g(x, y) &= f(x, y) + \eta(x, y) \\ G(u, v) &= F(u, v) + N(u, v) \end{aligned} \quad (2)$$

In the case of multiplicative noise the model is given as

$$g(x, y) = f(x, y) \cdot \eta(x, y) \quad (3)$$

Spatial filters are traditional means of removing noise from images and signals. Spatial filters usually smooth the data to reduce the noise, and also blur the data. Several new techniques have been developed in the last few years that improve on spatial filters by removing the noise more effectively while preserving the edges in the data [12] [13] [15]. Some of these techniques used the ideas from partial differential equations and computational fluid dynamics such as level set methods, non-linear isotropic and anisotropic diffusion. A Few techniques combined impulse removal filters with local adaptive filtering in the transform domain to remove not only white and mixed noise, but also their mixtures. In order to reduce the noise present in medical images many techniques are available like digital filters (FIR or IIR), adaptive filtering methods etc. However, digital filters and adaptive methods can be applied to signals whose statistical characteristics are stationary in many cases [12] [15]. Recently the wavelet transform has been proven to be useful tool for non-stationary signal analysis. Many denoising algorithms were developed on wavelet framework effectively but they suffer from four shortcomings such as oscillations, shift variance, aliasing, and lack of directionality [16]. In this paper we will present a different class of methods which are based on minimising the total variation of the image. These methods are performing denoising effectively by preserving the edge information in the images which will minimize the artifacts in the denoised data [6].

2. DENOISING USING TOTAL VARIATIONAL APPROACH

2.1 Introduction

It is observed that the noise will be at high frequencies and the signals and images with excessive & spurious detail will have the high total variation i.e. the integral of the absolute gradient of

those signals and images is high. Based on these observations it is proposed to reduce the total variation of the signal or image subject to it for getting a close match to the original signal. This is the key idea behind the denoising using total variational method [6].

The total variational technique has advantages over the traditional denoising methods such as linear smoothing, median filtering, Transform domain methods using Fast Fourier transform and Discrete Cosine Transform which will reduce the noise in medical images but also introduce certain amount of blur in the process of denoising which will damage the texture in the images in lesser or greater extent. The Total Variational approach will remove the noise present in flat regions by simultaneously preserving the edges in the medical images which are very important in diagnostic stage [1] [2] [3].

The total variation (TV) of a signal measures how much the signal changes between signal values. Specifically, the total variation of an N-point signal $x(n), 1 \leq n \leq N$ is defined as

$$TV(x) = \sum_{n=2}^N |x(n) - x(n-1)| \quad (4)$$

Given an input signal x_n , the aim of total variation method is to find an approximation signal call it, y_n , which is having smaller total variation than x_n but is "close" to x_n . One of the measures of closeness is the sum of square errors:

$$E(x, y) = \frac{1}{2} \sum_n (x_n - y_n)^2 \quad (5)$$

So the total variation approach achieves the denoising by minimizing the following discrete functional over the signal y_n :

$$E(x, y) + \lambda V(y) \quad (6)$$

By differentiating the above functional with respect to y_n , in the original approach we will derive a corresponding Euler-Lagrange equation which is numerically integrated with x_n (the original signal) as initial condition. Since this problem is a convex functional, we can use the convex optimization techniques to minimize it to find the solution y_n .

2.2 Denoising Algorithm

The problem of image denoising or noise removal is, given a noisy image $g : \Omega \rightarrow \mathbf{R}$, to estimate the clean underlying image f . For Gaussian noise (additive white), the degradation model describing the relationship between $g(x, y)$ and $f(x, y)$ is

$$g(x, y) = f(x, y) + \eta(x, y) \quad (7)$$

Where $\eta(x, y)$ is i.i.d zero mean Gaussian distributed. In this paper we are testing the proposed denoising method for three types of noises which are very common in medical images. They are Gaussian noise, Laplacian Noise and Poisson noise which are normally encountered in the process of image acquisition. Getting the good denoising results depend on using a good noise

model which will accurately describe the noise in the given image. The following table gives the three different noise models which are used in our paper:

Table 1: Noise models for Gaussian, Laplacian and Poisson Noise

Noise Model	PDF of Noise Model
Gaussian	$P(g(x, y) z(x, y)) = \frac{1}{z} \exp\left(-\frac{(z(x, y) - g(x, y))^2}{2\sigma^2}\right)$
Laplacian	$P(g(x, y) z(x, y)) = \frac{1}{z} \exp\left(-\frac{ z(x, y) - g(x, y) }{2\sigma^2}\right)$
Poisson	$P(g(x, y) z(x, y)) = \frac{1}{z} \exp(-z(x, y)) z(x, y)^{g(x, y)}$

Where $z(x, y) = f(x, y)$ for denoising and $\frac{1}{z}$ is the normalisation such that densities sum to one.

- The Gaussian model is a reasonable approximation for true noise distribution; the original Gaussian model was introduced by Rudin, Oscher and Fatemi [3].
- The Laplace model is better for salt-and-pepper and dark shot noise which will have fat tail noise distributions was developed by chan and Esedoglu [1].
- The Poisson model will describe the noise introduced due to low-light image acquisition and also this model is a rough approximation for multiplicative noise and the Poisson model was developed by Le et al. [2].

A General model for TV-regularized denoising, Deblurring, and Inpainting is to find an image $f(x, y)$ that minimizes

$$\min_{f \in BV(\Omega)} \int_{\Omega} |\nabla f(x, y)| dx dy + \int_{\Omega} \lambda(x, y) F(Kf(x, y), g(x, y)) dx dy \quad (8)$$

Where ∇ denotes the gradient, ∇^2 denotes Laplacian, and $\|\cdot\|_p$ denotes the L_p norm on Ω . Variable (x, y) will be used to denote a point in two-dimensional space. $f(x, y)$ is an original image, $g(x, y)$ is an observed noisy image. The integrals are over a two-dimensional bounded set $\Omega \subset \mathbf{R}^2$ and $|\nabla f(x, y)|$ denotes the gradient magnitude of $f(x, y)$ at $(x, y) \in \mathbf{R}^2$. Function $g(x, y)$ is the given noise and blur corrupted image, K is the blur operator, $\lambda(x, y)$ is a nonnegative function specifying the regularization strength, BV stands for Bounded variation and F determines the type of data fidelity:

$$F(Kf(x, y), g(x, y)) = \begin{cases} \frac{1}{2}(Kf(x, y) - g(x, y))^2 & \text{Gaussian noise} \\ |Kf(x, y) - g(x, y)| & \text{Laplace noise} \\ Kf(x, y) - g(x, y) \log Kf(x, y) & \text{Poisson} \end{cases} \quad (9)$$

For simplicity, $\lambda(x, y)$ is usually specified as a positive constant, $\lambda(x) = \lambda$

Spatially-varying fidelity weight

The parameter lambda (λ) is a positive value specifying the fidelity weight which controls the amount of denoising. This parameter must be tuned for good results, since choosing a large λ removes a limited amount of noise while a small λ removes more noise but smooths out the signal [6]. A spatially-varying fidelity weight can be specified by setting lambda (λ) as a $M \times N$ matrix which is denoted by $\lambda(x, y)$.

Role of Regularization Parameter λ

In the denoising process the regularization parameter λ is having a critical role. The denoising is zero when $\lambda = 0$, therefore the result is same as the input signal. As the regularization parameter λ increases the amount of denoising is also increases, when $\lambda \rightarrow \infty$ the total variation term plays strong role, which will produce the smaller total variation, at the expense of being less like the input (noisy) signal [6]. So the regularization parameter choice is critical for achieving the right amount of noise removal.

In total variational approach, we propose to solve the general problem

$$\hat{f}(x, y) = \arg \min_{f(x, y)} J(f(x, y)) \quad (10)$$

$$\text{With } J\{f(x, y)\} = \frac{1}{2} \|g(x, y) - K\{f(x, y)\}\|_2^2 + \lambda J_{reg}\{f(x, y)\} \quad (11)$$

Where K is the blurring kernel, $g(x, y)$ is the noisy and blurred observation and $f(x, y)$ is the signal we are recovering. The generic regularisation term J_{reg} is weighted by λ . In this paper we are considering J_{reg} from the class of convex regularisers.

The Total variation approach is to search over all possible functions to find a function $f : \Omega \rightarrow \mathbf{R}$ that minimizes (8) [4] [5] [6]. In this paper we are using the split Bregman method to solve the minimization problem by operator splitting and then solving split problem by applying Bregman iteration [10]. For (8), the split problem is

$$\min_{\vec{d}, z, f} \int_{\Omega} |\vec{d}(x, y)| dx dy + \int_{\Omega} \lambda(x, y) F(z(x, y), g(x, y)) dx dy \quad (12)$$

subject to $\vec{d} = \nabla f$, $z = Kf$

The split problem is not different from the original (8). The point is that the two terms of the objective have been split: The first term $\int_{\Omega} |\vec{d}|$ only depends on \vec{d} and the second term $\int \lambda F(z, g)$ only on z . Still \vec{d} and z are indirectly related through the constraints $\vec{d} = \nabla f$, $z = Kf$.

Now the Bregman iteration is used to solve the split problem. In every iteration, it calls for the solution of the following problem:

$$\min_{\vec{d}, z, f} \int_{\Omega} |\vec{d}(x, y)| dx dy + \int_{\Omega} \lambda(x, y) F(z(x, y), g(x, y)) dx dy + \frac{\gamma_1}{2} \|\vec{d} - \nabla f - \vec{b}_1\|_2^2 + \frac{\gamma_2}{2} \|z - Kf - b_2\|_2^2 \dots\dots\dots (13)$$

Additional terms in the above expression are quadratic penalties enforcing the constraints and b_1, b_2 are the variables connected to the Bregman iteration algorithm [7] [10][11].

The solution of (13), which minimizes jointly over \vec{d}, z, f is approximated by alternately minimizing one variable at a time, that is, fixing z and f minimising over \vec{d} then fixing \vec{d} and f minimising over z and so on. This method leads to three variable subproblems.

The \vec{d} subproblem

Variables z and f are fixed and the sub problem is

$$\min_{\vec{d}} \int_{\Omega} |\vec{d}(x, y)| dx dy + \frac{\gamma_1}{2} \|\vec{d} - \nabla f - \vec{b}_1\|_2^2 \tag{14}$$

Its solution decouples over x and is known in closed form:

$$\vec{d}(x, y) = \frac{\nabla f(x, y) + \vec{b}_1(x, y)}{|\nabla f(x, y) + \vec{b}_1(x, y)|} \max\{|\nabla f(x, y) + \vec{b}_1(x, y) - 1| - 1/\gamma_1, 0\} \tag{15}$$

This is the important subproblem which drives the Total Variation minimization.

The z Sub problem

Variables \vec{d} and f are fixed and the sub problem is

$$\min_z \int_{\Omega} \lambda F(z, g) dx dy + \frac{\gamma_2}{2} \|z - Kf - b_2\|_2^2 \tag{16}$$

The solution decouples over x . The optimal z satisfies

$$\lambda \partial_z F(z, g) + \gamma_2 (z - Kf - b_2) = 0 \quad (17)$$

The f sub problem

Variables \vec{d} and z are fixed and the sub problem is

$$\min_f \frac{\gamma_1}{2} \|\vec{d} - \nabla f - \vec{b}_1\|_2^2 + \frac{\gamma_2}{2} \|z - Kf - b_2\|_2^2 \quad (18)$$

For denoising K is identity and the optimal f satisfies

$$\frac{\gamma_2}{\gamma_1} f - \Delta f = \frac{\gamma_2}{\gamma_1} (z - b_2) - \text{div}(\vec{d} - \vec{b}_1) \quad (19)$$

This is a sparse, symmetric positive definite linear system. The solution f can be efficiently approximated by Gauss-Seidel iteration [5] [6].

The full algorithm

The minimization (8) is solved with the following iteration:

Initialise $f = z = b_2 = 0, \vec{d} = \vec{b}_1 = 0$

While "not converged"

Solve the \vec{d} subproblem

Solve the z Sub problem

Solve the f sub problem

$$\vec{b}_1 := \vec{b} + \nabla f - \vec{d}$$

$$b_2 := b_2 + Kf - z$$

While solving these subproblems, the x^{th} sub problem solution is computed from the current values of all other variables and overwrites the previous value of variable x . Convergence will be checked by testing the maximum difference from the previous iterate: $\|f^{\text{cur}} - f^{\text{prev}}\|_2 < \text{ToI}$.

3. RESULTS & DISCUSSIONS

The performance of this method is evaluated by applying the method on monochrome colon image with various λ values for the images corrupted with Gaussian, Laplacian, and Poisson noise of various variance levels. From the results we can say that the total variation method will preserve the edges effectively compared to the spatial filters. The denoising strength will decrease with the increase of regularization parameter.

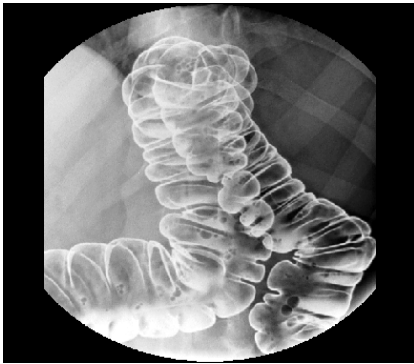
GAUSSIAN:

Table 2: Noise level vs PSNR for Images corrupted with Gaussian Noise

$\lambda = 1$				$\lambda = 5$		
Noise level (σ)	0.03	0.06	0.09	0.03	0.06	0.09
PSNR	21.5712	21.5832	21.5575	27.5057	27.4987	27.4811
$\lambda = 10$				$\lambda = 15$		
Noise level (σ)	0.03	0.06	0.09	0.03	0.06	0.09
PSNR	30.8364	30.6805	30.9470	32.7724	32.4796	32.1117

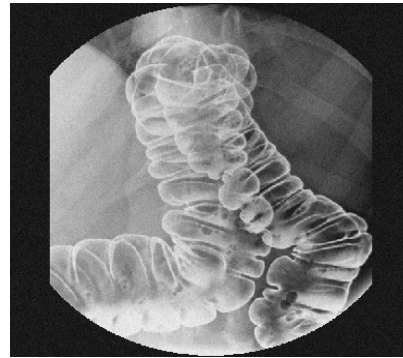
For sigma=0.03

Original image

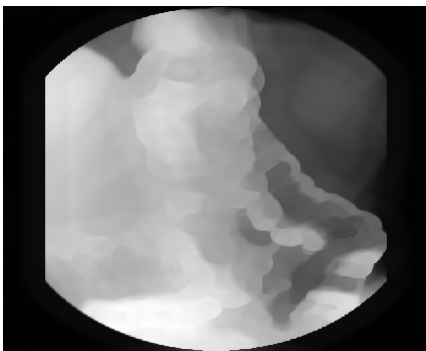


$\lambda = 1$

noisy image



$\lambda = 5$



$\lambda = 10$



$\lambda = 15$

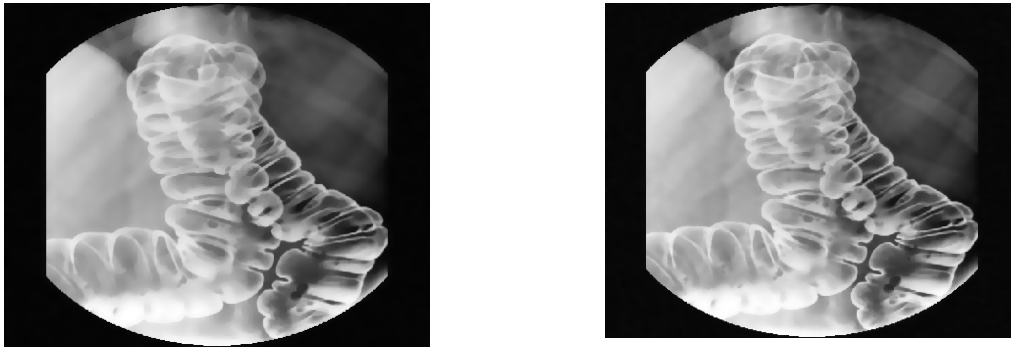


Figure 3: Denoising of Image with noise variance 0.03 and for different values of λ . As the λ value increases the denoising will be effective and it will closely approximate the original signal.

LAPLACIAN:

Table 3: Noise level vs PSNR for Images corrupted with Laplacian Noise

$\lambda = 1$				$\lambda = 5$		
Noise level (σ)	0.03	0.06	0.09	0.03	0.06	0.09
PSNR	33.1720	30.9347	29.4693	31.3277	26.1588	23.0863
$\lambda = 10$				$\lambda = 15$		
Noise level (σ)	0.03	0.06	0.09	0.03	0.06	0.09
PSNR	31.3548	26.3370	23.442	31.2177	26.2236	23.6758

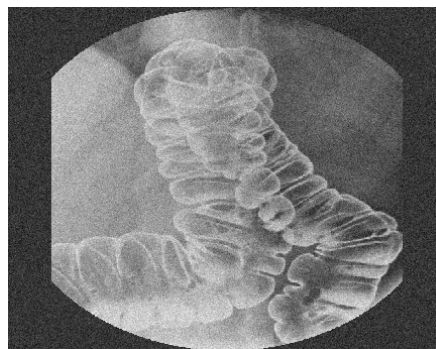
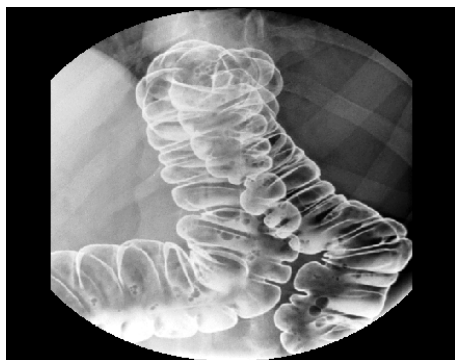
POISSON:

Table 2: Noise level vs PSNR for Images corrupted with Poisson Noise

$\lambda = 1$				$\lambda = 5$		
Noise level (σ)	0.03	0.06	0.09	0.03	0.06	0.09
PSNR	23.7205	23.6232	23.4706	23.4706	28.9153	27.3584
$\lambda = 10$				$\lambda = 15$		
Noise level (σ)	0.03	0.06	0.09	0.03	0.06	0.09
PSNR	32.4672	29.8744	27.7309	33.5382	29.7888	26.4509

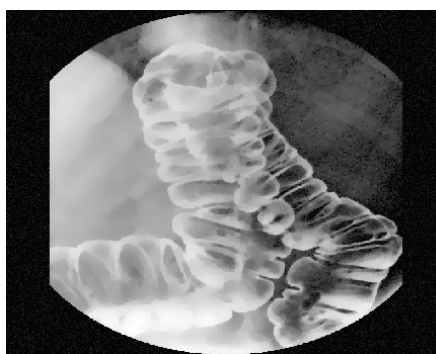
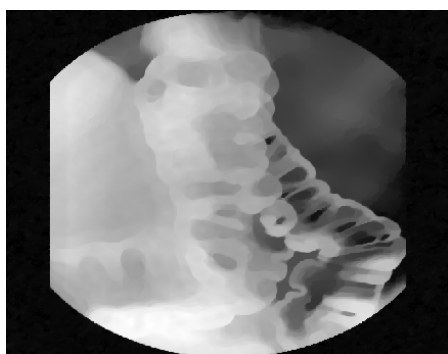
Original image

Noisy image (Poisson Noise)



$\lambda = 1$

$\lambda = 5$



$\lambda = 10$

$\lambda = 15$

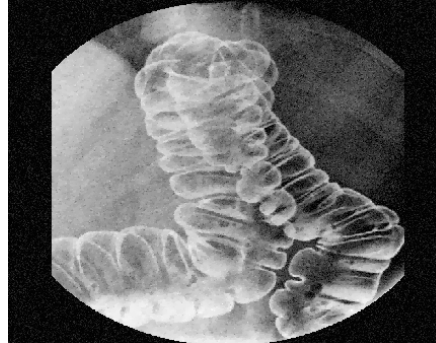
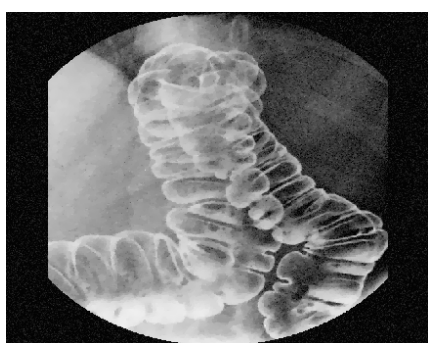


Figure 4: Denoising of Image with noise variance 0.09 and for different values of λ .

4. CONCLUSIONS & FUTURE WORK

TV-based regularization is particularly well-suited to image restoration in certain cases. The underlying assumption of TV is that images under consideration belong to the class of piecewise-constant signals. Hence, the recovered images resulting from the application of this model in the presence of noise are subject to the so-called staircase effect. Total variation regularization is well capable of preserving edges of uniform, small curvature, it does strongly smooth, and may even

destroy, small scale structures with high curvature edges. Therefore, the use of TV may generate undesirable artifacts and compromise the quality of the recovered image. This problem can be minimised using second (or higher)-order regularization methods.

REFERENCES

- [1] T.F. Chan and S. Esedoglu. Aspects of total variation regularized L1 function approximation." UCLA,CAM Report 04-07, 2004.
- [2] T. Le, R. Chartrand, and T. Asaki. A Variational Approach to Constructing Images Corrupted by Poisson Noise," JMIIV, vol. 27(3), pp. 257-263, 2007.
- [3] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms." Physica D, vol. 60, p. 259-268, 1992.
- [4] J.-B. Hiriart-Urruty and C. Lemar_echal. Convex Analysis and Minimization Algorithms I. Springer-Verlag, 1993.
- [5] A. Chambolle. An algorithm for total variation minimization and applications. J. of Math. Imaging and Vision, 20:89-97, 2004.
- [6] Nonlinear Total Variation Based Noise Removal Algorithms L. Rudin, S. Osher, E. Fatemi, Physica D: Nonlinear Phenomena, Vol. 60, Issues 1-4, 1992.
- [7] Iterative methods for total variation denoising, C.R. Vogel, M.E. Oman, SIAM J. on Scientific Computing 17 (1): 227-238, 1996.
- [8] Aspect of Total Variation Regularized L1 Function Approximation, Chan T., Esedoglu S., SIAM J. Appl. Math., 65(5):1817-1837, 2005.
- [9] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, And Wotao Yin "An Iterative Regularization Method For Total Variation-Based Image Restoration", International Journal of Multiscale Modelling and Simulation, Vol 4, No.2, PP 480-489.
- [10] Wotao Yin , Stanley Osher, Donald Goldfarb And Jerome Durbon, " Bregman Iterative Algorithms for Minimization with Applications to Compressed Sensing", vol. 1, No.1, pp143-168.
- [11] Tobias Lindstrøm Jensen, "First-order Convex Optimization Methods for Signal and Image processing", Ph.D Thesis, Department of Electronic Systems, Aalborg University, Denmark.
- [12] R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y.Neuvo, "Optimal weighted median filters under structural constraints," IEEE Trans. Signal Processing, vol. 43, pp. 591–604, Mar. 1995.
- [13] A. Ben Hamza, P. Luque, J. Martinez, and R. Roman, "Removing noise and preserving details with relaxed median filters," J. Math. Imag. Vision, vol. 11, no. 2, pp. 161–177, Oct. 1999.
- [14] A.K.Jain,Fundamentals of digital image processing. Prentice-Hall, 1989.
- [15] A.Buades, B.Coll,J. M.Morel"A Review of Image Denoising Algorithms, with a new one", Journal of Multiscale Modeling and Simulation, Vol.4, No.2, pp. 490-530, 2005.
- [16] D L Donoho and M. Jhonstone, "Wavelet shrinkage: Asymptopia? ", J.Roy.Stat.Soc., SerB, Vol.57, pp. 301-369, 1995.
- [17] Isaac Bankman, "Handbook of Medical Imaging", Academic Press, 2000.