Abstract. Removal of noise is an important step in the image restoration process, and remains a challenging problem in spite of the sophistication of recent research. Popular approaches can be categorized in various ways, a classical one of which involves minimization of variation in an image. This report presents a comparison of three such classical image denoising methods, all minimizing the variation of an image - $H^1$ Regularization, Bounded Variation (BV) Regularization and a Primal Dual Hybrid Gradient (PDHG) scheme. These algorithms are compared based on their assumptions and shortcomings using a methodology that examines the rate of convergence, signal-to-noise ratio (SNR), and overall quality of the denoised image.

1 Introduction

This section provides the mathematical premise of image denoising algorithms that attempt to minimize the total variation in an image. The goal of these methods is to preserve the underlying structure in the image, while removing noise. Mathematically, the general problem can be stated as the following minimization:

$$\min_u \int_\Omega |\nabla u| \text{ such that } \|u - f\|^2_2 \leq |\Omega| \sigma^2$$  \hspace{1cm} (1)

where $u$ is the resultant denoised image, $f$ is the original image (ground truth, without any noise), $\Omega$ is the domain and $\sigma^2$ is the estimate of the variance of the noise in the image. Eq. 1 can be transformed into an unconstrained minimization problem using the Lagrange multiplier $\lambda$, as follows:

$$\min_u \int_\Omega |\nabla u| + \|u - f\|^2_2$$  \hspace{1cm} (2)

In this paper, we explore three algorithms that approach the above problem in different ways, and provide a three-way comparison between them. Experiments with a synthetic phantom are presented, and the algorithms are evaluated based on the quality of the result, SNR, and convergence criteria.

2 Methods and Implementation

This section provides an overview of the theory and implementation for each of the three methods being compared. For further details, the reader is referred to the explanation given in the class notes [1] and the paper by Zhu [2].
2.1 $H^1$ Regularization

This algorithm aims to minimize the $L_2$ norm of the image gradient. The final minimization function can be given as,

$$E = \int_{\Omega} |\nabla u|^2 + \|u - f\|^2_2$$  \hspace{1cm} (3)

Discretizing Eq. 3 leads to a simple gradient descent (GD) scheme approximating the solution. With an assumption of zero-gradient at the image boundaries, and given a step size $\epsilon$, the $k^{th}$ iteration of this scheme at each pixel location is given as:

$$u^k \leftarrow u^{k-1} + \epsilon(\Delta u^{k-1} - \lambda(u^{k-1} - f))$$

Improvements such as the Conjugate Gradient (CG) method and Fast Fourier Transform (FFT) have been proposed to improve the speed of the iterative minimization scheme described above. Even though FFT is the fastest of the three, we will not include it in the evaluation since it is difficult to compare this one-step solution with the other iterative schemes.

2.2 Bounded Variation (BV) Regularization

$H^1$ regularization is an $L_2$ minimization algorithm and results in significant smoothing of edges in the resulting image. Such smoothing results from the fact that the $L_2$ norm is infinite at discontinuities in the image. BV regularization avoids this problem by aiming to minimize the total variation in the image, making it an $L_1$ minimization scheme. The minimization function for BV regularization can be written as follows:

$$E = \int_{\Omega} |\nabla u| + \|u - f\|^2_2$$  \hspace{1cm} (4)

A gradient descent solution to Eq. 4 results in the following update at the $k^{th}$ iteration:

$$u^k \leftarrow u^{k-1} + \delta t |\nabla u^{k-1}| \left(div \left( \frac{\nabla u^{k-1}}{|\nabla u^{k-1}|} \right) - 2\lambda(u^{k-1} - f) \right)$$

The above update is more complex since we have an $L_1$ norm, instead of the $L_2$ norm used in the $H^1$ regularization.

2.3 Primal Dual Hybrid Gradient (PDHG) Scheme

In [2] the authors note the deficiencies in the primal gradient descent and dual gradient descent algorithms, and propose a hybrid solution that uses both the primal and dual spaces to overcome shortcomings of the earlier algorithms. A discrete version of the PDHG algorithm entails the following two steps:

$$x^{k+1} \leftarrow P_X(x^k + \tau_k \lambda A^T y^k)$$
\[ y^{k+1} \leftarrow (1 - \theta_k) y^k + \theta_k \left( \frac{1}{\lambda} A x^{k+1} \right) \]

where \( x \) is the vectorized form of the solution to the dual problem \( \omega \), \( y \) is the vectorized form of \( u \) and \( z \) is the vectorized form of \( f \). Also, \( A^T y \equiv \nabla u \) and \( A x \equiv -\nabla \cdot \omega \); while \( P_X \) is a normalization term and the time steps \( \tau_k, \theta_k \) are defined in [2].

### 3 Results and Discussion

This section details experiments designed to compare the image denoising methods described above. The discussion is based on rate of convergence, SNR and image quality. We present experiments with the synthetic Shepp-Logan phantom to examine the effect of the parameter \( \lambda \) and also compare results of all the denoising algorithms under consideration.

The in-built function within Matlab was used to generate a 256 × 256 image of the Shepp-Logan phantom. Gaussian noise with \( \mu = 0 \) and \( \sigma = 0.05 \) was added to this image to create the test image for comparing all algorithms. Fig. 1 shows the original and noisy images used in the experiments.

![Fig. 1. Synthetic phantom image: original image (l) and noisy image (r)](image)

The GD and CG schemes were tested for the \( H^1 \) regularization algorithm. CG was found to be order of magnitude faster than the gradient descent, but at the same time the image quality degraded faster with changes in the Lagrangian weight \( \lambda \). Fig. 2 shows the changes in the resulting image for various values of \( \lambda \). It is clear that a smaller value of \( \lambda \) results in an image with blurred edges in both implementations. However, it must be noted that for a given \( \lambda \), the result of the CG scheme appears more blurred than that for the GD scheme. Thus, there is an inherent trade-off between image quality and speed when choosing one of the \( H^1 \) regularization schemes.

In order to compare the other algorithms on an even footing, we choose \( \lambda = 0.1 \) and a tolerance (or stopping criterion for the iterative scheme) of \( 10^{-6} \).
Fig. 2. $H^1$ Regularization: results after 500 iterations for various values of $\lambda$ for the gradient descent (top) and CG (bottom) schemes.

Fig. 3 shows the results of the various approaches for the Shepp-Logan phantom. Qualitatively, both the $H^1$ regularization schemes are out-performed by the BV regularization and the PDHG algorithms. However, it is difficult to say which of the latter performs better. The result of the BV algorithm displays a local smoothing which gives the image an artificial look, whereas the PDHG result appears to preserve the inherent structure much better.

Fig. 3. Denoised Image: (a) $H^1$-GD, (b) $H^1$-CG, (c)BV, (d)PDHG schemes respectively
Quantitatively, the CPU times (seconds) and resultant signal-to-noise ratio (SNR) for the various algorithms are tabulated in Tab. 3. Here again, we notice a trade-off between speed and quality. BV regularization is the slowest, but provides the best SNR. However, the SNR does not reflect the shortcomings of the algorithm, and may not be a practical metric for comparison (e.g. $H^1$ regularization methods have an SNR comparable to BV and PDHG, but they result in blurred edges).

Table 1. Execution times and resultant SNR for the various algorithms acting on the Shepp-Logan phantom

<table>
<thead>
<tr>
<th></th>
<th>$H^1$-GD</th>
<th>$H^1$-CG</th>
<th>BV</th>
<th>PDHG</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU-time</td>
<td>3.7378</td>
<td>0.6756</td>
<td>15.4298</td>
<td>7.4070</td>
</tr>
<tr>
<td>SNR</td>
<td>7.3105</td>
<td>7.2273</td>
<td>10.0228</td>
<td>8.7167</td>
</tr>
</tbody>
</table>

4 Conclusion

This paper compares and contrasts three algorithms for image denoising. These algorithms are tested on a synthetic phantom, and a comparison is made based on image quality, SNR and convergence. $H^1$ regularization methods suffer due to the fact that they minimize the $L_2$ norm, and result in images with blurred edges. Of the two implementations tested, CG was found to be an order of magnitude faster than GD, but the resulting image quality did not differ very much. Both these schemes are also sensitive to the Lagrangian weight $\lambda$, and may benefit from the use of a spatially varying weight rather than a fixed value, as used in the experiments here. Even though both GD and CG result in SNRs comparable to BV regularization and PDHG method, the resulting images are of poor quality. BV regularization preserves edges better than $H^1$ regularization, but results in local smoothing which makes the result look artificial. This method is also the slowest of those tested. However, it uses a gradient descent scheme, and will benefit from an adaptive step size. PDHG emerges as the best algorithm in terms of the trade-off between rate of convergence and image quality.

5 Acknowledgements

Thanks to Srikant for information on the function to generate the Shepp-Logan phantom within Matlab.

References