Inpainting of Image Pixels Using $l1$ Norm Minimization

Oregon State University
Department of Electrical Engineering and Computer Science
ECE 569 - Convex Optimization Final Project
Opeyemi Ajibuwa

Abstract

This article outlines the previous works in which the inpainting problem was formulated as a total variation problem and solved by convex optimization. This work revisits the image restoration issue. The idea is that minimizing the $l1$ norm of the discrete gradient of a damaged image yields an output that is barely noticeable from the original non-degraded form. The inpainting problem was formulated to limit the high fidelity of known observations and minimize the overall norm of variability. By reviewing the proposed formulation, we identified and demonstrated how image pixel resolution is an important characteristic of inpainting.

1 Introduction

Image inpainting deals with restoring degraded images that may have been affected by scratches or text overlays, loss or impairment in transmission, object removal during editing, or disocclusion in rendering by cameras [3]. There is often a need to recover missing or damaged pixels of an image and restore them to their original form in many image processing applications. Specifically, inpainting is an interpolation problem that involves the recovery of missing pixels in a digital image based on pixels available in the observed image.

As image painting belongs to the class of ill-posed problems with no well-defined unique solution [3, 4], various methods have been proposed in the literature to solve the problem. Some of these techniques require implementing iterative numerical methods that are generally relatively slow and thus have limited use in practical applications. On the other hand, the total variation (TV) image model is widely used for inpainting problems due to its distinctive ability to preserve edges in recovered images [3, 4, 6, 7]. Under the total variation model, the image inpainting problem is formulated as an optimization problem with fidelity to image observations as a constraint.

The key component of the TV regularized inpainting methods is the famous TV norm, which has achieved great success in suppressing noise and preserving edges [7]. Based on this characteristic, we formulate the inpainting problem as a convex optimization problem described by [1] to estimate the pixel values in obscured/corrupted parts of an image by minimizing $l1$ total variation. The optimization problem was solved and evaluated on synthetic data under different image resolution sizes and noise conditions.

2 Problem Formulation

2.1 Background

In image inpainting, the task involves estimating an unknown region using information from the known regions. The corrupted image represented as u_0 in Figure 1 is made of the target region Ω, which represents the region to be inpainted, $\partial\Omega$ represents the boundary of Ω. The basic idea in most inpainting literature is to use information from the source region $(u_0 - \Omega)$ to estimate the unknown region with a high level of fidelity [5].

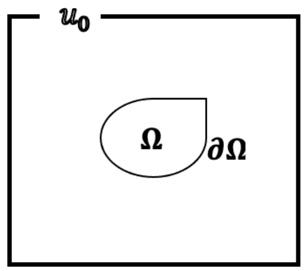

Figure 1: Image to be inpainted

Researchers proposed several variation-based image models to capture the inpainting problem. By formulating image inpainting as a total variation problem, the image is seen as a function of bounded variation (BV). The energy function is based on the total variational norm, hence the name of the TV image model [3]. The objective is to find a function Î of BV on the set Ω (BV(Ω)) on which the input image I is specified that minimizes a TV energy (defined as the integral of the gradient magnitude) of the picture within the hole while keeping the goal of smooth image intensity propagation in mind [3]. The optimization problem can thus be written as:

$$J_{TV}(\hat{I}) = \int_{\Omega} |\bigtriangledown \hat{I}(x)|dx + \lambda \int_{\Omega \setminus \cup} (I(x) - \hat{I}(x))^2 dx \tag{1}$$

Where the first integral term is the regularization term and the second term is a data term that measures the reconstructed image's fidelity to the input image for known samples.

Recasting this TV inpainting optimization problem as functions of bounded variation (BV) and TV norm. The TV model for image inpainting can be re-expressed as [7]:

$$\min_{u} || \bigtriangledown u||_1 \quad s.t \; ||Hu - g||_2 \leq \varepsilon, \tag{2}$$

$u \in \mathrm{R}^N$ denotes the vector form of an unknown sharp image, $g \in \mathrm{R}^M$ denotes the vector form of the damaged image, $H \in \mathrm{R}^{M \times N}$ denotes a linear operator that causes the pixels in u to be randomly missing or damaged, $|| \cdot ||_1$ and $|| \cdot ||_2$ denote the l_1 and l_2 norms, respectively, and $\bigtriangledown \in \mathrm{R}^{L \times N}$ is the gradient operator, and $|| \bigtriangledown u||_1$ is the TV norm of u. Most picture inpainting approaches equate equation(2) to the following optimization problem to make it easier to deal with:

$$\min_{u} \frac{1}{2} ||Hu - g||_2^2 \; + \; \mu || \bigtriangledown u||, \tag{3}$$

where $\mu > 0$, and $0.5 \times ||Hu - g||_2^2$ +$|| \bigtriangledown u||_1$ is referred to as cost functional. Although unconstrained problem (3) is simpler to solve than problem (2), it ignores parameter ε, which represents the noise level. Furthermore, searching for the correct value for parameter u results in solving equation (3) repeatedly, which reduces the efficiency of the method. As a result, problem (3) is most likely not the best equivalent transformation of problem (1).

To avoid the high-frequency artifacts in the estimated solution of equation (3) and reduce the computational demands of finding the exact value of μ, The author in [2] proposed a TV discretization strategy based on finding the discrete derivative of the image and solving the resulting optimization problem using the split Bregman algorithm. Their approach to discrete derivatives use one-sided differences and half-sample symmetric extensions at points near the inpainted object boundaries.

The discrete derivative ∂ of a uniformly-sampled signal $f_0, f_1, \cdots, f_{N-1}$ in one dimension is defined as a forward difference,

$$\begin{pmatrix} \partial f_0 \\ \partial f_1 \\ \vdots \\ \partial f_{N-2} \\ \partial f_{N-1} \end{pmatrix} = \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ & & & & 0 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-2} \\ f_{N-1} \end{pmatrix} \tag{4}$$

Because the half-sample symmetric extension is $2N$-periodic, the discrete gradient can be thought of as a cyclic convolution of the reflected signal ($f_0, \cdots, f_{n-1}, f_{n-1} \cdots, f_0$) with the filter h_{-1} = 1 and h_0 = -1 and h is zero otherwise. The discrete gradient of an $N \times N$ image u in two dimensions is defined by the discrete derivative as $\bigtriangledown u$ = (∂_xu, ∂_yu)^T, where the subscript on ∂ denotes the dimension along which the difference is applied. Thus, the authors proposed that summing the vector magnitude $| \bigtriangledown u_{i,j}|$ over all pixels of the image to be inpainted approximates the TV as expressed in equation to obtain the discretization of the TV norm.

$$||u||_{TV(\Omega)} \approx \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} | \bigtriangledown u_{i,j}|, \tag{5}$$

where $\bigtriangledown u$ is the above-mentioned discrete gradient The split Bregman algorithm was employed to solve the final constrained optimization problem expressed as

$$\arg\min_{d,u} \sum_{i,j} |d_{i,j)}| + \frac{1}{2} \sum_{i,j} \lambda_{i,j}(f_{i,j} - u_{i,j})^2 \quad s.t \;\; d = \bigtriangledown u \tag{6}$$

2.2 Proposed Method

Assume we have an image $X \in R^{m \times n}$ with known pixel values $X_{i,j}$ for indices $i, j \in \mathcal{K}$. We use the assumption that natural images have small total variation, which quantifies local deviations between neighboring pixels, to recover the values of missing pixels whose indices are not in $\mathcal{K}$. We want a reconstruction Y that minimizes the l_1 total variation while preserving the known pixel values. This results in an optimization problem of the form:

$$minimize \sum_{i=1}^{m-1} \sum_{i=1}^{n-1} \left\| \begin{bmatrix} [Y_{i+1,j} - Y_{i,j}] \\ [Y_{i,j+1} - Y_{i,j}] \end{bmatrix} \right\|_1 \tag{7}$$

subject to $Y_{i,j}$ = $X_{i,j}$ $\forall$ $(i, j) \in \mathcal{K}$

Where $\left\| \begin{bmatrix} [Y_{i+1,j} - Y_{i,j}] \\ [Y_{i,j+1} - Y_{i,j}] \end{bmatrix} \right\|_1$ represents l_1 norm of the discrete gradient at all $(i, j)th$ pixel location of the image.

3 Algorithm

By the formulation of the inpainting problem in equation (7), the total variation objective function (TV_{l1}) only depends on the discretization of the l_1 norm. The problem is convex and can be inpainted. The objective function of total variation in equation (7) can find the discrete gradient and find the missing pixels in the corrupted image by minimizing its l1 norm. However, the restored image must match the image whose pixels are damaged with a valid index.

Note that for any true digital image, its pixel can attain only a fine number of values. Hence, it is natural to require all pixel values of the restored image to lie in a certain interval $[a_l, a_u]$. Such a constraint is called the box constraint. For instance, the images considered in this review are all 8-bit images, and as we would like to restore them in a dynamic range, we constrained our optimization problem to the range of [0, 255].

Table 1: **Algorithm**

Step 1: Select the area on the image to be inpainted, the mask
Step 2: Compute the total variation of the corrupted image
a. Find the discrete gradient $\bigtriangledown(\mathrm{Y})_{i,j}$
b. Compute the l_1 norm of the discrte gradient
Step 3: Use CVX to solve Equation (3) until an optimal solution is found
Step 4: Display the inpainted image

4 Experiment Details

Because a large dataset of damaged painting images is unavailable, existing methods in the literature generate their own image inpainting datasets by adding some artificial distortion such as noise [4], text [2], scratch [3], objects (shapes) [2], and masks. In this review, we synthesize images that have been distorted by text and noise in order to evaluate the performance of the proposed algorithm. We put it through its paces on three RGB images with varying image resolution sizes, as well as under conditions of text obfuscation and varying degrees of extreme sparsity caused by noise.

The resolution of the image is divided into three categories: full-size 512×512 pixels, half-size 256×256 pixels, and third resolution 1024×1024 pixels. Finally, the algorithm was implemented in python and the tests were carried out on an Intel Core i5-7267U processor running at 3.10GHz with 8GB of RAM.

4.1 Inpainting in the presence of noise

This subsection considers three interesting cases of image inpainting, namely the recovery of an image in the presence of additive noise. The *cat* image in three different image resolutions is used to test the proposed method's performance under noisy conditions. We first examined the inpainting characteristics of a 256 $\times$ 256 pixels image with 70% and 90% of the pixel values corrupted.

As shown in Figures 2 and 3, the visual quality of the recovered image degrades as the image becomes more sparse. This blurring effect is caused in part by the inpainting algorithm's tendency to over-smoothen the edges of the degraded regions in order to recover the image when there is extreme sparsity. Furthermore, image resolution impacts the overall quality of the recovered image in terms of the blur effects generated.

In the case of a 256 $\times$ 256 resolution with 90% of the pixels noisy, our method fails to properly recover the original image's texture, resulting in smoothing and blurring artifacts in the synthesized region. As expected, the 512 $\times$ 512 image in Figures 4 and 5 has a noticeable improvement in visual quality when compared to the previous case. The recovered image has a better visual quality with an extreme sparsity of 90% because the higher pixels resolution gradually improves the recovered output, gaining on the smoothening effects experienced with the lower 256 $\times$ 256 pixels resolution image. Finally, the 1024 $\times$ 1024 image resolution in Figure 6 and 7 produces cleaner and sharper edges and textures for our inpainting algorithm, as the visual quality improves significantly over lower resolution images under both sparsity conditions considered.

Figure 2: 256 $\times$ 256 image with 70% missing data

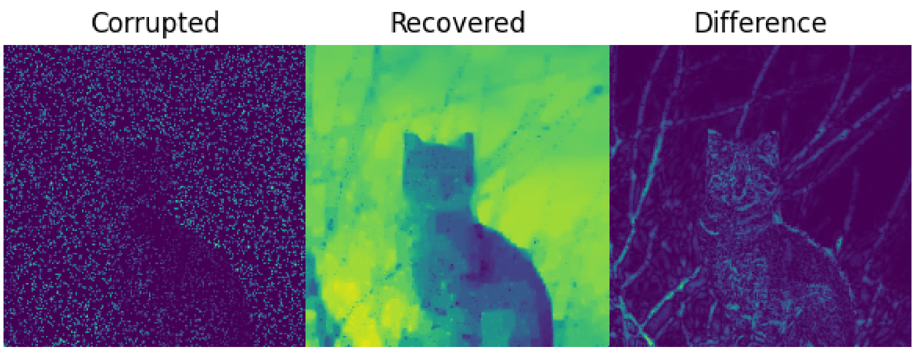

Figure 3: 256 $\times$ 256 image with 90% missing data

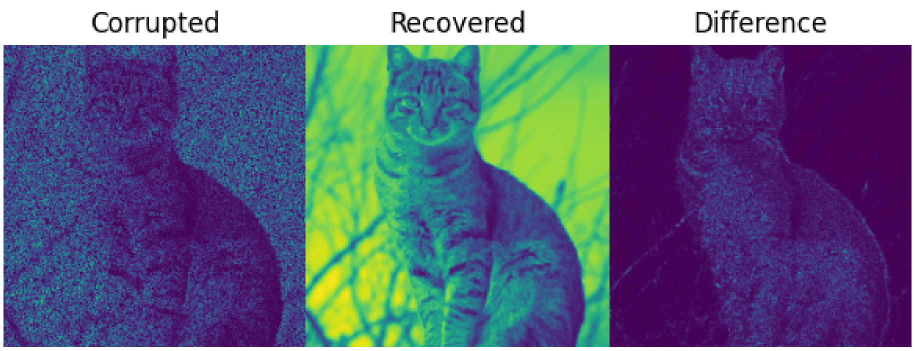

Figure 4: 512 $\times$ 512 image with 70% missing data

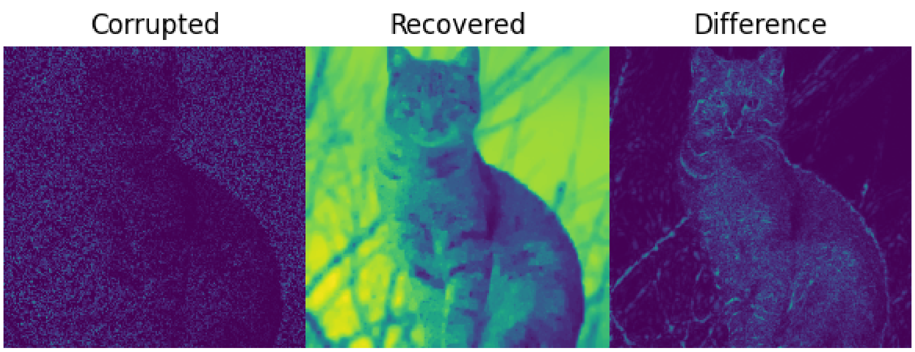

Figure 5: 512 $\times$ 512 image with 90% missing data

Figure 6: 1024 $\times$ 1024 image with 70% missing data

Figure 7: 1024 $\times$ 1024 image with 90% missing data

4.2 Inpainting for text removal

To deal with the text overlay problem, we examine the performance of our algorithm on two images with different pixel resolutions. The Monalisa image and its text overlay in Figure 8 have 512 $\times$ 512 dimensions, whereas the cat image and text overlay have 1024 $\times$ 1024 pixels resolution. The visual output clearly shows that our TV inpainting algorithm is successful in removing the overprinted text on the two test cases while leaving no residual blur effects. However, closer inspection of the outputs (on a zoom-in) reveals that the recovered output of the cat (with a higher resolution) has a higher visual quality, owing to the higher resolution of the pixel values.

However, based on the almost imperceptible difference in the outputs recovered in the two results, we can draw the same conclusions about the two cases. This demonstrates that the proposed algorithm (and TV inpainting in general) is efficient when the inpainting region is thin, as in our test with the superimposed texts.

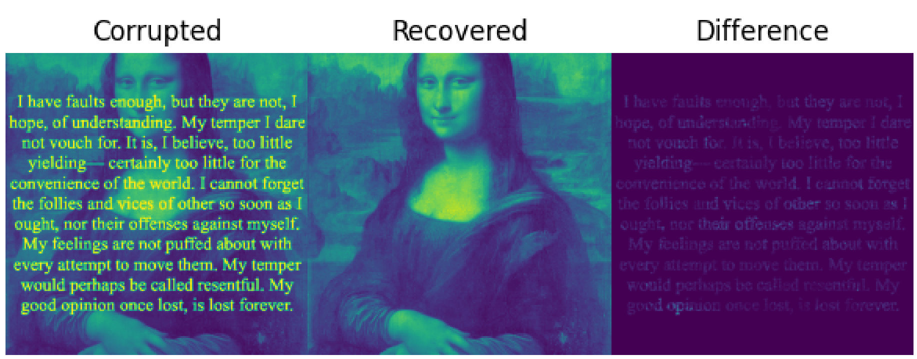

Figure 8: 512 $\times$ 512 image text removal

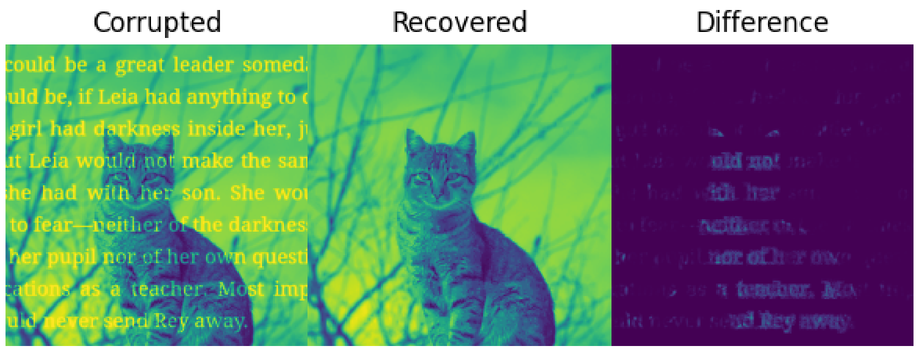

Figure 9: 1024 $\times$ 1024 image text removal

5 Conclusions

This project presented a reformulation of the image inpainting problem. This formulation allows us to investigate and understand inpainting issues and identify appropriate properties based on total minimization of variability. The effectiveness of the proposed method is demonstrated by testing it under different image resolution sizes and different noise conditions. Further work can be considered to explore an effective approach to validate the proposed method against the latest image inpainting techniques.

6 Image credits

Cat by 24 Warriors (Pinterest) and Monalisa by Ali Al Araimi (Twitter)

References

[1] Stephen Boyd and Steven Diamond. Consensus optimization and machine learning. *Available at: https://www.slideshare.net/0xdata/h2o-world-consensus-optimization-and-machine-learning-stephen-boyd*, 2015.

[2] Pascal Getreuer. Total variation inpainting using split bregman. *Image Processing On Line*, 2:147–157, 2012.

[3] Christine Guillemot and Olivier Le Meur. Image inpainting: Overview and recent advances. *IEEE signal processing magazine*, 31(1):127–144, 2013.

[4] Jun Liu, Ting-Zhu Huang, Ivan W Selesnick, Xiao-Guang Lv, and Po-Yu Chen. Image restoration using total variation with overlapping group sparsity. *Information Sciences*, 295:232–246, 2015.

[5] S Padmavathi, N Archana, and KP Soman. Hierarchical approach for total variation digital image inpainting. *arXiv preprint arXiv:1207.3576*, 2012.

[6] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.

[7] Su Xiao. Image inpainting method based on total variation regularization. *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, 10(3):242–247, 2017.