Abstract

In this paper, we propose a method for blind image inpainting, in which the locations of the pixels which are missing or corrupted with impulse noise are not known. A logarithmic transformation is applied to convert the multiplication between the image and binary mask into an additive problem. The image and mask terms are then estimated iteratively with total variation regularization applied on the image, and $\ell_0$ regularization on the mask term. The resulting alternating minimization scheme simultaneously estimates the image and mask, in the same iterative process. Experimental results show that the proposed method can deal with a larger fraction of missing pixels than two phase methods which first estimate the mask and then reconstruct the image.

1 Introduction

It is known that faulty sensors or bit errors during transmission can cause some pixels in an image to be lost or corrupted by impulse noise [5, 10]. In the case of missing pixel values, the corrupted pixels are assumed to have a value equal to zero, and the problem of estimating the complete image is called the inpainting problem.

We represent the image to be estimated with $n$ pixels as a vector, say in lexicographic ordering, $x \in \mathbb{R}^n$. Let $m < n$ be the number of observed pixels or pixels free from impulse noise. The mapping from $x$ to the partially observed image $y$ is given by,

$$y = A(x + n_\text{s}),$$

where the mask operator $A$ is a $n \times n$ identity matrix with the diagonal elements equal to zero, corresponding to missing pixels, and $n_\text{s}$ is an additive Gaussian noise term corrupting the observed pixel values.

When the index set of the observed pixels or the observation mask $A$ is known, the inpainting problem can be solved by one of several existing methods for image reconstruction from a sparse set of observations [1, 2].

Adaptive filtering [10], [9] can be used to detect outliers and remove impulse noise for low levels of missing/damaged pixels. Some two phase methods such as [6] have a mask detection step in which the mask is detected through the use of one of these filters and the image is estimated in a separate step.

1.1 Contributions

In this paper, we propose a method to estimate the image $x$ without knowing the observation mask $A$ apriori, i.e., we simultaneously estimate the image and the mask. We formulate the masking operation as a summation after logarithmic compression, and apply a total variation (TV) [8, 12] regularizer on the image term, and an $\ell_0$-norm regularizer on the term corresponding to the mask. The TV regularizer encourages the estimate of $x$ to be piece-wise smooth, while the $\ell_0$-norm regularizer encourages the mask term to be sparse. Experimental results show that our proposed method can deal with as many as 95% of the pixels missing, which is higher than reported in literature.

2 Proposed Method

Given the observation $y$, our goal is to estimate the image $x$ and the mask $A$. Since $A$ is a diagonal matrix $A = \text{diag}(a)$ and the masking operation is element-wise multiplication, we will optimize over the vector of diagonal elements $a \in \{0, 1\}^n$. When a pixel with index $i$ is observed, the corresponding mask element $a_i = 1$, and when pixel $i$ is lost, $a_i = 0$. Thus, a pixel $k$ in vector $y$ is defined as, $y_k = x_k \times a_k$.

We do not know apriori if a given pixel $y_k$ corresponds to an observed one ($a_k = 1$) or not ($a_k = 0$). Rather than have $a_k = 0$ when the pixel is not observed, we can define $a_k$ to be a small value in the order of $10^{-K}$ or smaller, $K$ being a positive integer greater than or equal to 3. Defining $v_i = \log(a_i)$, we have,

$$v_i = \begin{cases} 0, & \text{if } i \text{ is observed} \\ -K, & \text{otherwise} \end{cases}$$

(2)

If the maximum possible pixel value is 255, the value of $K$ must satisfy $K > \log 255$. Assuming that $y$ and $x$ are always positive, applying a logarithmic transformation on (1) converts it into an additive model, $g_i = u_i + v_i$, where $u_i = \log(x_i)$, and $g_i = \log(y_i + \delta)$. A small positive bias term $\delta > 0$ is added to $y$ to guarantee positivity.

Our problem is now estimating the vectors $u$ and $v$, given the log transformed observation $g$. We assume that our image $x$, and therefore its logarithmic transformation $u$ are piece-wise smooth. We therefore apply a TV regularizer on the log transformed image $u$, and the $\ell_0$-norm regularizer on the log-transformed mask $v$. Our estimation problem is therefore,

$$(\hat{u}, \hat{v}) = \arg \min_{u,v} \frac{1}{2} \|g - u - v\|_2^2 + \frac{\lambda_1}{2} TV(u) + \frac{\lambda_2}{2} \|v\|_0.$$ (3)

where $\lambda_1, \lambda_2 > 0$ are the respective regularization parameters.

Formulations of the form (3) with different regularizers have appeared in the context of image decomposition [13] and deblurring with a sum of regularizers [3]. Alternating minimization schemes involving a sum of the $\ell_0$-norm and a convex term were also used in [4, 11] for sparse image recovery. Since (3) is a separable problem, we can apply a simple iterative alternating method as in [3].

Isolating the terms in each variable keeping the other fixed, leads to a Gauss-Seidel scheme. Solving for $u$ at iteration $t$,

$$u^{(t)} = \arg \min_{u} \frac{1}{2} \|g - u - v^{(t)}\|_2^2 + \frac{\lambda_1}{2} TV(u).$$

This is a TV regularized denoising problem, the solution of which can be computed efficiently using Chambolle’s algorithm [7].

Similarly, for $v$ at iteration $t$ we have,

$$v^{(t)} = \arg \min_{v} \frac{1}{2} \|g - u^{(t)} - v\|_2^2 + \frac{\lambda_2}{2} \|v\|_0.$$ (5)

This problem although non-convex, has a solution given by the hard threshold,

$$\hat{v}^{(t)} = H_{\lambda_2}(g - u^{(t)}),$$ (6)

where $H_{\lambda_k}(\cdot)$ is the hard threshold operator and is defined element-wise as,

$$\hat{v}_i^{(t)} = \begin{cases} 0, & \text{if } (g_i - u_i^{(t)}) \leq \lambda_2 \\ (g_i - u_i^{(t)}), & \text{otherwise} \end{cases}$$ (7)

This iterative process is run until the relative difference between successive estimates falls below a given threshold. The estimates of the image and mask are computed by inverting the logarithmic transformation, $\hat{x} = 10^\hat{u}$ and $\hat{a} = 10^{\hat{v}}$.

3 Experimental Results

All experiments were performed on MATLAB on an Ubuntu Linux based server with 64 GB of RAM. To test our proposed method, we generate a random binary mask with a fraction of its elements equal to zero and multiply it element-wise to our image corrupted with additive Gaussian noise. The criteria used to evaluate the accuracy of estimation are the Improvement in Signal to Noise Ratio (ISNR), which is defined as $10 \log_{10} \left( \frac{\|x - \hat{x}\|_2^2}{\|x - \hat{a}\|_2^2} \right)$.
In the results presented in this paper, the values of the regularization parameters used were \( \lambda_1 = 0.008 \) and \( \lambda_2 = 0.01 \), which were found to work well. Figure 1 shows the worst case results obtained with the proposed method for the 512 × 512 Lena image. The binary mask has only 5% of its pixels equal to 1, which means that 95% of the pixels were randomly discarded. Figures 1(b) and 1(c) show the observed images obtained with this mask, with additive Gaussian noise with Signal to Noise Ratios (SNR) of 5 dB and 20 dB, respectively.

Table 1 compares the proposed method with other methods for blind inpainting, for Gaussian noise with \( \sigma = 15 \). The fraction of missing pixels is varied from 0.1 to 0.9. The methods, fast two phase deblurring with TV [6], Adaptive Outlier Pursuit (AOP) [16], and K-ALS [14] were run for the ranges of fractions of missing pixels, as reported by their authors. We can see from these tables that [16] and [14] produce better ISNRs with fewer missing pixels, but our proposed method produces a better ISNR for higher percentages of missing pixels. While [6] is more accurate in estimating the observation mask, it produces a lower ISNR.

### 4 Conclusions

We have presented an iterative method for inpainting an image without knowing the locations of the missing pixels, based on alternating minimization to simultaneously estimate the image and observation mask. The method can deal better with a larger fraction of missing pixels than existing methods. Preliminary results show that the method is promising for the removal of impulse noise. Current and future work include optimizing the regularization parameters, a detailed convergence analysis, and extension of the framework to non-Gaussian observation models.

### References


### Acknowledgements

This work was supported by Fundação para a Ciência e Tecnologia (FCT), Portuguese Ministry of Science and Higher Education, through a Post-doctoral fellowship (contract no. SFRH/BPD/79011/2011) and FCT project (PEst-OE/EEI/LA0009/2013).