De-Blurring Wireless Capsule Endoscopy Images by Total Variation Minimization

Haiying Liu*†, W.-S. Lu†, and Max Q.-H. Meng*
* School of Control Science and Engineering
Shandong University, Jinan, China
† Dept. of Electrical and Computer Engineering
University of Victoria, Victoria, Canada
Email:{haiyingl, wslu}@ece.uvic.ca

Abstract

Wireless capsule endoscopy (WCE) finds wide applications in diagnosing diseases in human intestine on account of its convenience and non-invasiveness to the patients. Due to complicated environment of the intestine and intrinsic restrictions of the equipment in terms of image acquisition and transmission, however, raw WCE images are often corrupted by noise and degraded by within-channel or cross-channel blurs. In this paper, we present a total variation (TV) minimization framework for deblurring within-channel blurs for color WCE images. Specifically, the monotone fast iterative shrinkage/thresholding technique combined with the fast gradient projection algorithm, proposed recently by Beck and Teboulle, is extended to deal with multichannel images. In addition, the proposed algorithm is enhanced by incorporating a bisection technique to effectively identify a near optimal value for the regularization parameter of the TV-Frobenius objective function. Experimental results are presented to demonstrate the performance of the proposed algorithm.

1. Introduction

Wireless capsule endoscopy (WCE) is a state-of-the-art technology to diagnose gastrointestinal tract diseases with practically no invasiveness [1]. It acquires images during a slow squirm process and transmits them from inside of the body by a wireless transmitter. However, raw WCE images are often blurred due primarily to the complicated environment of the intestine and intrinsic restrictions of the equipment in terms of image acquisition and transmission. This in turn imposes difficulties for accurate and effective diagnosis. Image deblurring is an algorithmic procedure for restoring digital images degraded by one or more of various kinds of blurs. Of the rich variety of techniques for image restoration, it is most relevant to mention the celebrated Rudin, Osher and Fatemi (ROF) algorithm [2] that was the first to introduce the notion of total variation (TV) and treat image denoising problems in a TV-regularized minimization framework, and reference [3] was the first work on image deblurring based on TV minimization was reported. The last two decades have witnessed the growth of increasing research interest in this effective methodology and a large volume of literature covering a variety of image and video processing problems [4]. Dealing with blurs for multichannel images such as color images are more complicated than that for single-channel images because blurs may exist within and/or cross channels. In this paper, we report some new developments for the restoration of color WCE images with within-channel blurs in a TV minimization framework. Specifically, the monotone fast iterative shrinkage/thresholding algorithm (MFISTA) combined with the fast gradient projection algorithm, proposed recently in [5], is extended to deal with multichannel (e.g. color) images. In addition, the algorithm is enhanced by incorporating a bisection technique for tuning the regularization parameter of the TV-Frobenius (TV-F) objective function to its optimal value quickly and accurately. Experimental results using color WCE images are presented to demonstrate the performance of the proposed algorithm.

2. Previous work

2.1. Image model and TV-regularized minimization

Consider a blurring model for discrete images

$$\boldsymbol{u}_0 = \boldsymbol{\mathcal{A}}\boldsymbol{u} + \boldsymbol{w} \tag{1}$$

 978-1-4577-0253-2/11/$26.00

where $\mathcal{A}$ represents an affine map standing for a blurring operator, $\boldsymbol{u}_0$ denotes an observed noisy image, and $\boldsymbol{w}$ is normally distributed additive noise. Assuming noise $\boldsymbol{w}$ is Gaussian white with i.i.d. components of zero mean and variance σ^2, the restoration problem here is to estimate (recover) image $\boldsymbol{u}$ given the observation $\boldsymbol{u}_0$ and operator $\mathcal{A}$.

It is well known [2–5] that the restoration at hand can be treated as an unconstrained convex optimization problem

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu\|\boldsymbol{u}\|_{\text{TV}} + \frac{1}{2}\|\mathcal{A}\boldsymbol{u} - \boldsymbol{u}_0\|_F^2 \tag{2}$$

where $\mu > 0$ is a regularization parameter, $\|\cdot\|_F$ denotes the Frobenius norm of matrix, and $\|\boldsymbol{u}\|_{\text{TV}}$ denotes the two-dimensional (2-D) discrete total variation of $\boldsymbol{u}$. For a matrix $\boldsymbol{u}$ of size n_1 by n_2, its TV norm is defined by

$$\|\boldsymbol{u}\|_{\text{TV}} = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\|\boldsymbol{D}_{i,j}\boldsymbol{u}\|_2, \quad \boldsymbol{D}_{i,j}\boldsymbol{u} = \begin{bmatrix}\boldsymbol{D}_{i,j}^{(h)}\boldsymbol{u}\\ \boldsymbol{D}_{i,j}^{(v)}\boldsymbol{u}\end{bmatrix} \tag{3a}$$

where

$$\boldsymbol{D}_{i,j}^{(h)}\boldsymbol{u} = \begin{cases} u_{i+1,j} - u_{i,j} & i < n_1,\\ 0 & i = n_1\end{cases}$$

$$\boldsymbol{D}_{i,j}^{(v)}\boldsymbol{u} = \begin{cases} u_{i,j+1} - u_{i,j} & j < n_2,\\ 0 & j = n_2\end{cases} \tag{3b}$$

2.2. A TV-norm for color images

Let $\boldsymbol{u} = \{\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)}\}$ be a color image of size n_1 by n_2, where $\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}$ and $\boldsymbol{u}^{(3)}$ are the components of $\boldsymbol{u}$ in red (R), green (G) and blue (B) channels, respectively. A natural extension of the TV norm for gray-scale images given in (3) to color (multi-channel) images is given by

$$\|\boldsymbol{u}\|_{\text{CTV}} = \left(\sum_{i=1}^{3}\|\boldsymbol{u}^{(i)}\|_{\text{TV}}^2\right)^{1/2} \tag{4}$$

where each $\|\boldsymbol{u}^{(i)}\|_{\text{TV}}$ is defined by (3). In the rest of the paper, we refer the TV-norm defined by (4) as color-TV norm, or briefly as CTV norm. Note that $\|\boldsymbol{u}\|_{\text{CTV}}$ is reduced to the conventional $\|\boldsymbol{u}\|_{\text{TV}}$ when $\boldsymbol{u}$ is a scalar-valued (e.g., gray-scale) image.

We remark that the CTV-norm defined in (4) is a discrete counterpart of a CTV norm defined in [6] for images as functions of *continuous variables*. In that case, the Euler-Lagrange (EL) equations are found to be

$$\frac{\|\boldsymbol{u}^{(i)}\|_{\text{TV}}}{\|\boldsymbol{u}\|_{\text{CTV}}}\nabla\circ\left(\frac{\nabla\boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|}\right) - \lambda\mathcal{A}^*\big(\mathcal{A}\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)}\big) = 0 \tag{5}$$

for $i = 1, 2, 3$, where $\mathcal{A}^*$ is the adjoint of $\mathcal{A}$, $\boldsymbol{u}_0^{(i)}$ denotes the ith component of noisy observation $\boldsymbol{u}_0$, and λ is a Lagrange multiplier. Note that the three EL equations in (5) are not independent from each other but coupled via the ratio of the TV norms:

$$r_i(\boldsymbol{u}) = \frac{\|\boldsymbol{u}^{(i)}\|_{\text{TV}}}{\|\boldsymbol{u}\|_{\text{CTV}}} \qquad \text{for} \quad i = 1, 2, 3 \tag{6}$$

Based on (5), the problem of deblurring color images can be carried out in the RGB space by numerically solving the partial differential equations (PDEs)

$$\frac{\partial\boldsymbol{u}^{(i)}}{\partial t} = r_i(\boldsymbol{u})\nabla\circ\left(\frac{\nabla\boldsymbol{u}^{(i)}}{\sqrt{\beta + \|\boldsymbol{u}^{(i)}\|^2}}\right) - \lambda\mathcal{A}^*\big(\mathcal{A}\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)}\big) \tag{7}$$

for $i = 1, 2, 3$, where β is a small positive scalar to prevent from dividing by zero. The reader is referred to [6] for details.

2.3. Gradient-based algorithms for problem in (2)

Several authors [3–5] have proposed methods to address image deblurring problems by solving optimization problems (2). Of particular interest and relevance to the work reported here is a dual-based approach developed in [5] which yields a monotone fast iterative shrinkage-thresholding algorithm (MFISTA) where each iteration requires to solve a denoising subproblem which is carried out using a fast gradient projection (FGP) algorithm. Due to the limitation in space, the reader is referred to Sections 4 and 5 of [5] for the algorithmic details of the MFISTA/FGP technique.

3. An algorithm for deblurring color images

3.1. Analysis

Multichannel images may suffer from either within-channel or cross-channel blur. Within a TV-minimization framework, there are two deblurring approaches for multichannel images, namely the channel-by-channel (CBC) TV minimization and Color-TV (CTV) minimization. If one treats a color image as a vector-valued function of two continuous variables with three differentiable scalar function components, then the CBC-TV approach leads to the EL equations [3]

$$\nabla\circ\left(\frac{\nabla\boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|}\right) - \lambda\mathcal{A}^*\big(\mathcal{A}\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)}\big) = 0 \quad \text{for} \quad i = 1, 2, 3 \tag{8}$$

which are evidently independent from each other in the sense that these equation can be dealt with individually and the solution of one equation does not effect the solutions of the other two equations. On the other hand, the CTV approach leads to the EL equations in (5) which are coupled via the TV norm ratios $r_i(\boldsymbol{u})$ in (6). Since high correlation between channel components $\boldsymbol{u}^{(i)}$ often exists, the associated EL equations are expected to be coupled with each other. For this reason a deblurring technique based on CTV minimization is expected to outperform its counterpart based on CBC-TV minimization. This expectation was positively confirmed in [6] for image denoising problems.

For discrete images, the formulation corresponding to the CBC-TV based EL equations (8) is given by

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu\|\boldsymbol{u}^{(i)}\|_{\text{TV}} + \frac{1}{2}\|\boldsymbol{\mathcal{A}u}^{(i)} - \boldsymbol{u}_0^{(i)}\|_F^2 \qquad (9)$$

for $i = 1, 2, 3$, where parameter μ is inversely related to Lagrange multiplier λ, i.e., $\mu \propto 1/\lambda$. Fast algorithms exist ([4], [5]) for solving (9).

For the reason stated earlier, however, we are more interested in a discrete formulation that is related to the CTV-based EL equations (5). To this end, we use (5) and (6) to write the steady-state EL equations (7) (with time *t* approaching infinity) as

$$\nabla \circ \left(\frac{\nabla \boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|} \right) - \frac{\lambda}{r_i(\boldsymbol{u})} \boldsymbol{\mathcal{A}}^* \left(\boldsymbol{\mathcal{A}u}^{(i)} - \boldsymbol{u}_0^{(i)} \right) = 0 \qquad (10)$$

for $i = 1, 2, 3$. On comparing (10) with (8), we see that the effect of employing CTV rather than CBC-TV is essentially to modify the *constant* Lagrange multiplier λ to an *image-dependent* Lagrange multiplier $\lambda/r_i(\boldsymbol{u})$. By letting $\lambda_i(\boldsymbol{u}) = \lambda/r_i(\boldsymbol{u})$, (10) becomes

$$\nabla \circ \left(\frac{\nabla \boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|} \right) - \lambda_i(\boldsymbol{u}) \boldsymbol{\mathcal{A}}^* \left(\boldsymbol{\mathcal{A}u}^{(i)} - \boldsymbol{u}_0^{(i)} \right) = 0 \qquad (11)$$

where each "generalized" Lagrange multipliers $\lambda_i(\boldsymbol{u})$ is a function of the entire image $\boldsymbol{u}$, through which the three EL equations are coupled. From (11), a counterpart of formulation (9) follows as

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu_i(\boldsymbol{u})\|\boldsymbol{u}^{(i)}\|_{\text{TV}} + \frac{1}{2}\|\boldsymbol{\mathcal{A}u}^{(i)} - \boldsymbol{u}_0^{(i)}\|_F^2 \qquad (12)$$

where $\mu_i(\boldsymbol{u}) \propto 1/\lambda_i(\boldsymbol{u})$ and assumes the form

$$\mu_i(\boldsymbol{u}) = \mu \cdot r_i(\boldsymbol{u}) = \mu \cdot \frac{\|\boldsymbol{u}^{(i)}\|_{\text{TV}}}{\|\boldsymbol{u}\|_{\text{CTV}}} \quad \text{for} \quad i = 1, 2, 3 \quad (13)$$

3.2. Solving minimization problem in (12)

The algorithm we propose for solving (12) is iterative. In its *k*th iteration, the $\mu_i(\boldsymbol{u})$ in (12) is set to $\mu_i(\boldsymbol{u}_{k-1}) = \mu \cdot r_i(\boldsymbol{u}_{k-1})$ where $\boldsymbol{u}_{k-1}$ is an iterate obtained from the preceding iteration and constant μ is determined by a bisection technique to be described in detail in Sec.3.3. In this way, the problem in (12) becomes

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu_i(\boldsymbol{u}_{k-1})\|\boldsymbol{u}^{(i)}\|_{\text{TV}} + \frac{1}{2}\|\boldsymbol{\mathcal{A}u}^{(i)} - \boldsymbol{u}_0^{(i)}\|_F^2 \quad (14)$$

for $i = 1, 2, 3$, which is a standard TV-F minimization problem that can be solved using the conventional MFISTA/FGP [5]. An important difference of the proposed algorithm from that of [5] is that unlike the MFISTA in [5] where the regularization parameter is kept invariant in the entire iteration process, the proposed algorithm updates $\mu_i(\boldsymbol{u})$ from $\mu_i(\boldsymbol{u}_{k-1})$ to $\mu_i(\boldsymbol{u}_k)$ once iterate $\boldsymbol{u}_k$ is obtained. The iteration continuous until the difference of two consecutive iterates in norm is less than a prescribed tolerance, or the number of iterations reaches a prescribed integer *K*.

3.3. A bisection technique for determining an optimal regularization parameter

Needless to say, using a good or, whenever possible, optimal value of regularization parameter μ is crucial as it affects the deblurring performance directly and significantly. The bisection technique described below is designed to determine a near optimal value of μ quickly. The technique is based on the fact that μ is related to the variance of noise $\boldsymbol{w}$ (see (1)) in a simple manner. For the sake of simplicity, we use model (1) and formulation (2) to illustrate the technique. It follows from (1) that if a solution $\boldsymbol{u}$ from a deblurring algorithm is in perfect agreement with the original noise-free and non-blurred image, it should satisfy

$$\|\boldsymbol{\mathcal{A}u} - \boldsymbol{u}_0\|_F^2 = \|\boldsymbol{w}\|_F^2 \approx n_1 n_2 \sigma^2 \qquad (15)$$

From (2), we see that parameter μ controls the trade-off between the TV-norm of the image and the closeness of $\boldsymbol{\mathcal{A}u}$ to $\boldsymbol{u}_0$ in Frobenius norm: if μ is set to be too large, then (2) would put a heavier weight on the TV-norm term and, as a result, the second term $\frac{1}{2}\|\boldsymbol{\mathcal{A}u} - \boldsymbol{u}_0\|_F^2$ (known as fidelity term) gets too large, exceeding $\frac{1}{2}n_1 n_2 \sigma^2$ and violating (15); if μ is too small, then (2) would weigh the second term too heavy, leading to a $\|\boldsymbol{\mathcal{A}u} - \boldsymbol{u}_0\|_F^2$ considerably smaller than $n_1 n_2 \sigma^2$ that violates (15) again. Consequently, $\|\boldsymbol{\mathcal{A}u} - \boldsymbol{u}_0\|_F^2$ as a function of μ is a monotonic function that increases with μ, and a near optimal value of μ can be identified as one that approximately satisfies (15). Based on the above analysis, a bisection technique for efficiently identify a near optimal value of μ is developed for the problem in (14) and is outlined as follows.

Step 1: Set an initial iterate, say, to the noisy observation $\boldsymbol{u}_0$ and identify an interval $[\mu_L, \mu_U]$ which contains the optimal value of μ. Set tolerance ε, and $k = 1$.

Step 2: Set $\mu_k = (\mu_L + \mu_U)/2$, solve (14) for $i = 1, 2, 3$ and form $\boldsymbol{u}_k = \{\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)}\}$.

Step 3: If $\|\boldsymbol{\mathcal{A}}\boldsymbol{u}_k - \boldsymbol{u}_0\|_F^2 > n_1 n_2 \sigma^2$, set $\mu_U = \mu$; otherwise set $\mu_L = \mu$.

Step 4: If $\mu_U - \mu_L < \varepsilon$, output solution $\boldsymbol{u}_k$ and stop; otherwise set $k = k + 1$ and repeat from Step 2.

Remarks:

(1) The technique possesses an exponential convergence rate of $1/2^k$.

(2) It requires an interval $[\mu_L, \mu_U]$ containing the solution μ to start.

Since μ is always positive, the initial lower bound can be set to $\mu_L = 0$. To get an initial upper bound, we use an initial $\mu_W > 0$ and solve (14) for $i = 1, 2, 3$ to get a solution $\boldsymbol{u}$. If $\|\boldsymbol{u}_W - \boldsymbol{u}_0\|_F^2 > n_1 n_2 \sigma^2$, then set $\mu_U = \mu_W$. Otherwise, double the value of μ_W, solve (14) for an updated $\boldsymbol{u}_W$, and check to see if $\|\boldsymbol{u}_W - \boldsymbol{u}_0\|_F^2 > n_1 n_2 \sigma^2$ is satisfied. The process continues until an appropriate upper bound $\mu_U > 0$ is identified. We remark that this is a procedure where a candidate upper bound μ_W grows exponentially, hence an upper bound μ_U can be secured rather quickly.

4. Performance evaluation

A color WCE image of size 140×122 was used for performance evaluation of the proposed algorithm, see Fig.1(a). The same image but suffering a 7×7 averaging blur plus a small amount of additive Gaussian white noise with $\sigma = 10^{-4}$ is shown in Fig.1(b). The modified MFISTA/FGP algorithm based on the CTV formulation (14) for color images, equipped with the bisection technique, was applied to the above blurred and noise-corrupted WCE image. The algorithm was implemented in MATLAB. A Windows XP laptop PC with an Intel Core Duo CPU P8700@2.53 GHz with 2.0 GB of RAM, equipped with MALAB 7.8.0, was used to run the code. The restored image is shown in Fig.1(c). The peak-signal-to-noise (PSNR) of the blurred and noise-corrupted image was 18.5717 dB. A profile of the PSNR of the deblurred image in the final round of iterations using the optimized value of μ (see below) is depicted in Fig.2. We see that the PSNR of the deblurred image after 10^4 iterations was 31.5638 dB that offers 12.9921 dB improvement, and the PSNR was found to be 32.8658 dB after 3×10^4 iterations.

As described in Sec.3.3, within a single iteration (with fixed preceding iterate $\boldsymbol{u}_{k-1}$), a bisection procedure was performed to identify a near optimal value of μ. In the first iteration ($k = 1$) of the algorithm, for example, the interval $[\mu_L, \mu_U]$ required by the bisection algorithm was identified

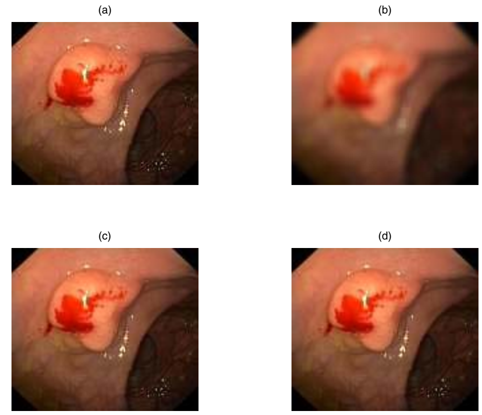

Figure 1: (a) Original WCE image, (b) Blurred and noise-corrupted image, (c) Restored image by CTV-TV minimization, (d) Restored image by CBC-TV minimization.

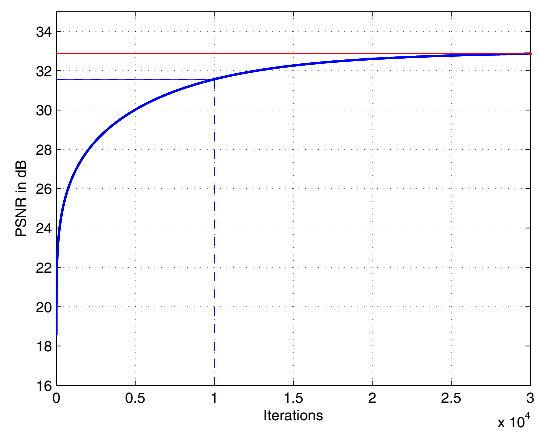

Figure 2: Profile of the PSNR of the deblurred image.

as $\mu_L = 0$ and $\mu_U = 10^{-4}$. With tolerance ϵ set to 5×10^{-6} (see step 4 of the bisection algorithm), five values of μ were produced and optimal μ was found to be $\mu = 3.125 \times 10^{-6}$. Fig.2 depicts a profile of these five μ's. As a result, the optimal values of $\mu_i(\boldsymbol{u}_0))$ for each channel were found to be $2.1534 \times 10^{-6}, 1.7334 \times 10^{-6}$ and 1.4575×10^{-6}. As expected, these μ's remain small primarily because of the low level of the additive noise in the observed image data $\boldsymbol{u}_0$.

For comparison purposes, a modified MFISTA/FGP algorithm based on CBC-TV formulation (9), equipped also with the bisection technique, was also applied to the above blurred and noise-corrupted image, the PSNR achieved by the deblurring algorithm was found to be 31.8996 dB, a gain that was 0.9692 dB less than that obtained by the CTV-TV

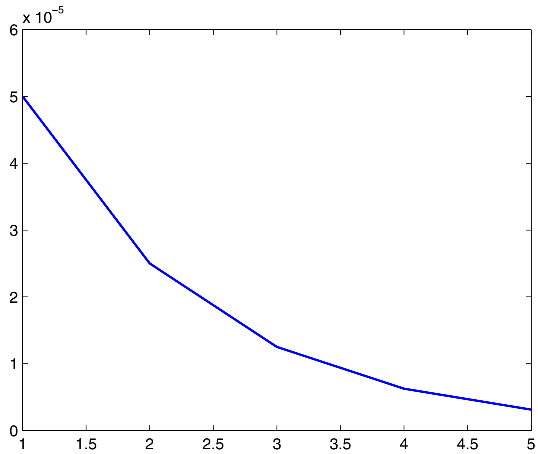

Figure 3: Profile of parameter μ produced by the bisection step.

based algorithm. The deblurred image obtained is shown in Fig.1(d). Visual inspection of Fig.1 clearly indicates the success of both CTV-TV and CBC-TV based algorithms in removing the blur from the image, with the CTV-TV based algorithm having the edge on its CBC-TV counterpart.

5. Conclusions

We have proposed an algorithm for removing within-channel blurs as well as random noise from multichannel images. The algorithm is built on a concept of color total variation in a MFISTA/FGP algorithmic framework. In addition, the algorithm is enhanced by incorporating a bisection technique into the algorithm that helps identify a near optimal value for a key regularization parameter. Simulation results have demonstrated the effectiveness of the proposed algorithm for restoring color WCE images.

6. Acknowledgment

The work of Haiying Liu has been supported by the Chinese Scholarship Council during her visit to University of Victoria.

References

[1] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, p. 417, 2000.

[2] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[3] L. Rudin and S. Osher, "Total variation based image restoration with free local constraints," in *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, vol. 1. IEEE, 1994, pp. 31–35.

[4] UCLA Computational and Applied Mathematics Reports, a web site maintained by the Math Dept., http://www.math.ucla.edu/applied/cam/.

[5] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *Image Processing, IEEE Transactions on*, vol. 18, no. 11, pp. 2419–2434, 2009.

[6] P. Blomgren and T. Chan, "Color TV: total variation methods for restoration of vector-valued images," *Image Processing, IEEE Transactions on*, vol. 7, no. 3, pp. 304–309, Mar. 1998.