Fast Algorithms for Restoration of Color Wireless Capsule Endoscopy Images

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) has been widely adopted to diagnose diseases in human intestine on account of its convenience and 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 blurred and noise contaminated. In this paper, we report several new developments for the restoration of color WCE images in a total variation (TV) minimization framework. Specifically, the monotone fast iterative shrinkage/thresholding algorithm combined with the fast gradient projection algorithm, proposed recently by Beck and Teboulle, is extended to deal with vector-valued (e.g. color) images. The new algorithm also incorporates a bisection technique to tune the weighting parameter of the TV-Frobenius objective function to its optimal value quickly and accurately. Experimental results are presented to demonstrate the performance of the proposed algorithm.**

I. Introduction

Wireless capsule endoscopy (WCE) is a state-of-the-art technology to diagnose gastrointestinal tract diseases with almost no invasiveness [1]. It acquires images during slow squirm process and transmits them from inside of the body by a wireless transmitter. However, raw WCE images are often blurred and noise contaminated 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 restoration is an algorithmic procedure for restoring digital images degraded by noise contamination or blurring, or both. 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 restoration problems in a TV-regularized minimization framework. 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 [3].

In this paper, we report some new developments for the restoration of color WCE images in a TV minimization framework. Specifically, the monotone fast iterative shrinkage/thresholding algorithm (MFISTA) combined with the fast gradient projection algorithm, proposed recently in [4], is extended to deal with vector-valued (e.g. color) images. The new algorithm also incorporate a bisection technique to tune the weighting parameter of the TV-Frobenius objective function to its optimal value quickly and accurately. Experimental results are presented to demonstrate the performance of the proposed algorithm.

II. Background and Related Work

A. Image Model and TV-Regularized Minimization

Consider a blurring model for discrete images

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

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. Note that the problem of image denoising also fits into model (1) by setting $\mathcal{A}$ to the identity operator. The restoration problem here is to estimate (recover) image $\boldsymbol{u}$ given the observation $\boldsymbol{u}_0$, operator $\mathcal{A}$ and some statistical properties of noise $\boldsymbol{w}$ (such as mean and variance of $\boldsymbol{w}$).

It is well known [2]-[4] 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, $\|g\|_F$ denotes the Frobenius norm of matrix, and $\|\boldsymbol{u}\|_{\text{TV}}$ denotes the two-dimensional (2-D) discrete total variation of $\boldsymbol{u}$, which for a matrix $\boldsymbol{u}$ of size n_1 by n_2 is defined by

$$\begin{aligned}\|\boldsymbol{u}\|_{\text{TV}} = &\sum_{i=1}^{n_1-1}\sum_{j=1}^{n_2-1}\sqrt{(u_{i,j}-u_{i+1,j})^2 + (u_{i,j}-u_{i,j+1})^2} \\ &+ \sum_{i=1}^{n_1-1}|u_{i,n_2}-u_{i+1,n_2}| + \sum_{j=1}^{n_2-1}|u_{n_1,j}-u_{n_1,j+1}|\end{aligned} \tag{3}$$

For the problem of removing noise from gray-scale images, the operator $\mathcal{A}$ in model (1) is set to the identity operator and, accordingly, the denoising problem is treated by solving the convex problem

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu\|\boldsymbol{u}\|_{\text{TV}} + \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_0\|_F^2 \tag{4}$$

where $\boldsymbol{u}_0$ denotes noisy observation.

B. 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

$$\|\boldsymbol{u}\|_{\text{CTV}} = \left(\sum_{i=1}^{3} \|\boldsymbol{u}^{(i)}\|_{\text{TV}}^2 \right)^{1/2} \tag{5}$$

where each $\|\boldsymbol{u}^{(i)}\|_{\text{TV}}$ is defined by (3). In the rest of the paper, we refer the TV-norm defined by (5) 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 (5) is a discrete counterpart of that proposed in [5] where images are treated as functions of continuous variables. The Euler-Lagrange (EL) equations associated with color-TV norm in (5) are found to be [5]

$$\frac{\|\boldsymbol{u}^{(i)}\|_{\text{TV}}}{\|\boldsymbol{u}\|_{\text{CTV}}} \nabla \circ \left(\frac{\nabla \boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|} \right) - \lambda \big(\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)} \big) = 0 \tag{6}$$

for $i = 1, 2, 3$, where $\boldsymbol{u}_0^{(i)}$ denotes the ith component of noisy observation $\boldsymbol{u}_0$, and λ is a Lagrange multiplier. Note that the three EL equations in (6) are not independent from each other but coupled via the ratio of the TV norms:

$$r_i(\boldsymbol{u}) = \frac{\|\boldsymbol{u}^{(i)}\|_{\text{TV}}}{\|\boldsymbol{u}\|_{\text{CTV}}} \qquad \text{for} \quad i = 1, 2, 3 \tag{7}$$

Based on (6), the problem of denoising of color images can be carried out in the RGB space by numerically solving the partial differential equations (PDEs) [5]

$$\frac{\partial \boldsymbol{u}^{(i)}}{\partial t} = \frac{\|\boldsymbol{u}^{(i)}\|_{\text{TV}}}{\|\boldsymbol{u}\|_{\text{CTV}}} \nabla \circ \left(\frac{\nabla \boldsymbol{u}^{(i)}}{\sqrt{\beta + \|\boldsymbol{u}^{(i)}\|^2}} \right) - \lambda \big(\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)} \big) \tag{8}$$

for $i = 1, 2, 3$, where β is a small positive scalar to prevent from dividing by zero.

C. Gradient-Based Algorithms for Problems (4) and (2)

Several authors have proposed methods for addressing image denoising and deblurring problems by solving optimization problems (4) and (2), respectively [3]. Of particular interest and relevance to the work reported here is a dual-based approach developed in [4] which yields fast gradient projection algorithms. Due to the limitation in space, the reader is referred to Sections 4 and 5 of [4] for the algorithmic details of the monotone fast iterative shrinkage-thresholding algorithm (MFISTA) and fast gradient projection (FGP) algorithm.

III. A FAST ALGORITHM FOR DENOISING COLOR IMAGES

A. Analysis

Within the TV-minimization framework, there are two denoising approaches for color images, namely the channel-by-channel (CBC) TV minimization and Color TV (CTV) minimization. If one treats color images as functions of continuous variables, then the CBC-TV approach leads to the EL equations ([3])

$$\nabla \circ \left(\frac{\nabla \boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|} \right) - \lambda \big(\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)} \big) = 0 \quad \text{for} \quad i = 1, 2, 3 \tag{9}$$

which are evidently independent from each other, while the CTV approach leads to the EL equations in (6) which are coupled via the TV norm ratios $r_i(\boldsymbol{u})$ in (7). Since high correlation between channel components $\boldsymbol{u}^{(i)}$ often exists, hence the EL equations they are associated with are expected to be related to each other. For this reason a denoising technique based on CTV minimization is expected to outperform its counterpart based on CBC-TV minimization. This was confirmed to be the case in [5] for the continuous image model.

For discrete images, the formulation corresponding to the decoupled EL equations (9) is given by

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu \|\boldsymbol{u}^{(i)}\|_{\text{TV}} + \frac{1}{2} \|\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)}\|_F^2 \tag{10}$$

for $i = 1, 2, 3$, where parameter μ is inversely related to Lagrange multiplier λ , i.e., $\mu \propto 1/\lambda$. Fast algorithms exist ([3], [4]) for solving (10).

For the reason stated above, however, we are more interested in a discrete formulation that is related to the CTV-based EL equations (6). To this end, we use (6) and (7) to write the EL equations as

$$\nabla \circ \left(\frac{\nabla \boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|} \right) - \frac{\lambda}{r_i(\boldsymbol{u})} \big(\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)} \big) = 0 \tag{11}$$

for $i = 1, 2, 3$. On comparing (11) with (9), 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 $\lambda_i(\boldsymbol{u}) = \lambda / r_i(\boldsymbol{u})$, (11) becomes

$$\nabla \circ \left(\frac{\nabla \boldsymbol{u}^{(i)}}{\|\boldsymbol{u}^{(i)}\|} \right) - \lambda_i(\boldsymbol{u}) \big(\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)} \big) = 0 \tag{12}$$

where each "generalized" Lagrange multipliers $\lambda_i(\boldsymbol{u})$ is a function of the (entire) color image $\boldsymbol{u}$, through which the three EL equations are coupled. From (12), a counterpart of formulation (10) follows as

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu_i(\boldsymbol{u}) \|\boldsymbol{u}^{(i)}\|_{\text{TV}} + \frac{1}{2} \|\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)}\|_F^2 \tag{13}$$

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

$$\mu_i(\boldsymbol{u}) = \mu \cdot r_i(\boldsymbol{u}) = \mu \cdot \frac{\|\boldsymbol{u}^{(i)}\|_{\text{TV}}}{\|\boldsymbol{u}\|_{\text{CTV}}} \quad \text{for} \quad i = 1, 2, 3 \tag{14}$$

B. Minimization Problem (13)

The algorithm we propose for solving (13) is iterative. In its kth iteration, the $\mu_i(\boldsymbol{u})$ in (13) is set to $\mu_i(\boldsymbol{u}_{k-1}) = \mu \cdot r_i(\boldsymbol{u}_{k-1})$ where $\boldsymbol{u}_{k-1}$ is an iteration 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 (13) becomes

$$\underset{\boldsymbol{u}}{\text{minimize}} \quad \mu_i(\boldsymbol{u}_{k-1})\|\boldsymbol{u}^{(i)}\|_{\text{TV}} + \frac{1}{2}\|\boldsymbol{u}^{(i)} - \boldsymbol{u}_0^{(i)}\|_F^2 \qquad (15)$$

for $i = 1, 2, 3$, which is a standard TV-F minimization problem that can be solved using the conventional MFISTA/FGP [4]. An important difference of the proposed algorithm from that of [4] is that unlike the MFISTA in [4] 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*.

C. 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 denoising 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. Below, we use model (1) (with $\boldsymbol{\mathcal{A}}$ set to the identity operator) and formulation (4) to illustrate the technique. It follows from (1) that if a solution $\boldsymbol{u}$ from a denoising algorithm is in perfect agreement with the original noise-free and non-blurred image, it should satisfy

$$\|\boldsymbol{u} - \boldsymbol{u}_0\|_F^2 = \|\boldsymbol{w}\|_F^2 \approx n_1 n_2 \sigma^2 \qquad (16)$$

From (4), we see that parameter μ controls the trade-off between the TV-norm of the image and the closeness of $\boldsymbol{u}$ to $\boldsymbol{u}_0$ in Frobenius norm: if μ is set to be too large, then (4) would put a heavier weight on the TV-norm term and, as a result, the second term $\frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_0\|_F^2$(known as fidelity term) gets too large, exceeding $\frac{1}{2}n_1 n_2 \sigma^2$ and violating (16); if μ is too small, then (4) would weigh the second term too heavy, leading to a $\|\boldsymbol{u} - \boldsymbol{u}_0\|_F^2$ considerably smaller than $n_1 n_2 \sigma^2$ that violates (16) again. Consequently, $\|\boldsymbol{u} - \boldsymbol{u}_0\|_F^2$ 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 (16). Based on the above analysis, a bisection technique is incorporated into the algorithm for efficiently identifying a near optimal value of μ and the enhanced denosing algorithm 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 (15) for $i = 1, 2, 3$. Form iterate $\boldsymbol{u}_k = \{\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \boldsymbol{u}^{(3)}\}$.

Step 3: If $\|\boldsymbol{u}_k - \boldsymbol{u}_0\|_F^2 > n_1 n_2 \sigma^2$, 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 (15) for $i = 1, 2, 3$ to get a solution $\boldsymbol{u}_W$. 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 (15), 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 $\boldsymbol{u}_W$ grows exponentially, hence an upper bound μ_U can be secured rather quickly.

IV. PERFORMANCE EVALUATION

A color WCE image of size 140×122 was used for performance evaluation of the proposed algorithm, see Fig.1(a). The image corrupted by additive Gaussian white noise with $\sigma = 0.05$ is shown in Fig.1(b). The interval $[\mu_L, \mu_U]$ required by the bisection technique (see III. C) was identified as $\mu_L = 0$ and $\mu_U = 0.2$. The modified MFISTA/FGP algorithm based on the CTV formulation (15) for color images, equipped with the bisection technique, was applied to the above 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. It took the algorithm 10 iterations to converge with 1.2419 seconds of elapsed time. The restored image is shown in Fig.1(c). The signal-to-noise (SNR) of the noise-corrupted image was 18.0937 dB, while the SNR of the denoised image was found to be 25.1278 dB.

Figure 2 shows three profiles of $\{\mu_i(\boldsymbol{u}_{k-1}) = \mu_k \cdot r_i(\boldsymbol{u}_{k-1}),$ for $k = 1, 2, ..., 10\}$ where μ_k is determined in the kth iteration of the algorithm, see Step 2 in Sec.3.C. It is observed that each sequence reaches its steady-state level after ten iterations, but the individual steady-state levels are different from each other across the three channels.

To justify the bisection technique described in Sec.3.C, Fig.3 depicts the profile of the values of μ_k generated in the iterations. We see μ_k converges to a value of 0.0667 after 10 iterations. The SNR of solution $\boldsymbol{u}$ with respect to parameter μ over interval $[\mu_L, \mu_U]$ is shown in Fig.4. Notice that the SNR of the solution produced by the proposed algorithm with $\mu = 0.0667$ is very close to the maximum SNR achievable by a solution of (15), which is 25.1352 dB when μ is taken to be 0.0640.

For comparison purposes, a modified MFISTA/FGP algorithm based on CBC-TV formulation (9), equipped also with

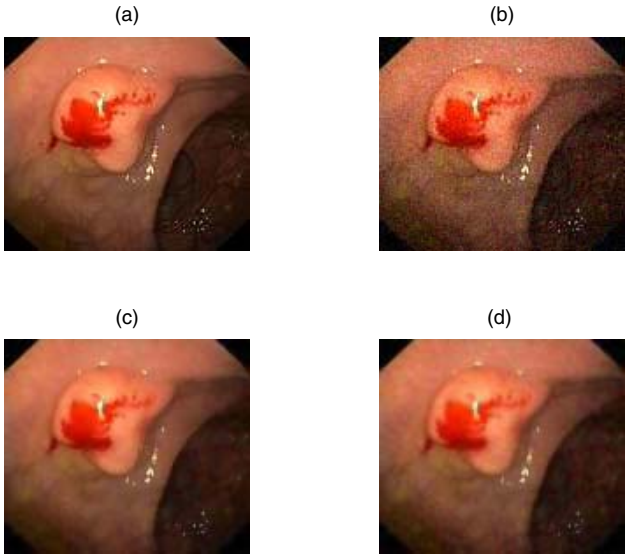

Fig. 1: (a) Original WCE image, (b) Noise-corrupted image, (c) Image restored by CTV-TV minimization, (d) Image restored by CBC-TV minimization.

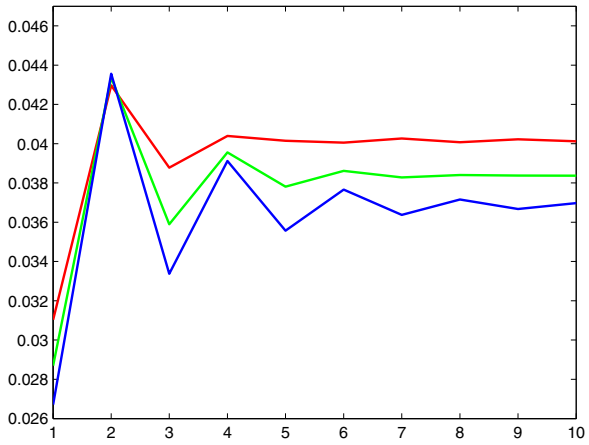

Fig. 2: Profiles of three parameter sequences for $i = 1$ (red channel), 2 (green channel) and 3 (blue channel).

the bisection technique, was also applied to the above noise-corrupted image. It took the algorithm 10 iterations to converge to a solution. The SNR achieved by the denoising algorithm was found to be 24.3148 dB, a gain that was 0.8130 dB less than that obtained by the CTV-based algorithm. The restored image is shown in Fig.1(d).

V. CONCLUSIONS

We have proposed an algorithm for removing within-channel 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

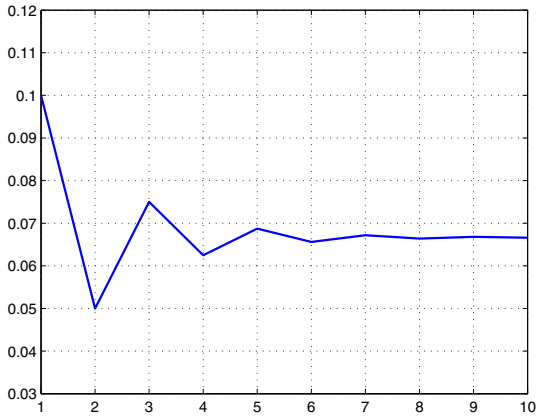

Fig. 3: $\{\mu_k, k=1,2,...,10\}$ generated by the proposed algorithm.

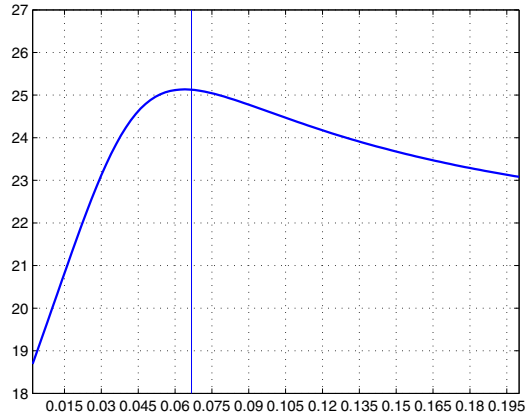

Fig. 4: SNR at the solution of (10) with respect to parameter μ with $r_i(\boldsymbol{u}_{k-1})$ fixed to $r_i(\boldsymbol{u}_0)$.

for a key regularization parameter. Simulation results have demonstrated the effectiveness of the proposed algorithm for restoring color WCE images.

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

[4] 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.

[5] 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.