An iterative algorithm for L1-TV constrained regularization in image restoration.

K. Chen[1], **E. Loli Piccolomini**[2], **F. Zama**[2]

[1]Department of Mathematics, University of Liverpool
[2]Department of Mathematics, University of Bologna

E-mail: `K.Chen@liverpool.ac.uk, elena.loli@unibo.it, fabiana.zama@unibo.it`

Abstract. We consider the problem of restoring blurred images affected by impulsive noise. The adopted method restores the images by solving a sequence of constrained minimization problems where the data fidelity function is the ℓ_1 norm of the residual and the constraint, chosen as the image Total Variation, is automatically adapted to improve the quality of the restored images. Although this approach is general, we report here the case of vectorial images where the blurring model involves contributions from the different image channels (cross channel blur). A computationally convenient extension of the Total Variation function to vectorial images is used and the results reported show that this approach is efficient for recovering nearly optimal images.

1. Introduction

We consider here the problem of restoring a multichannel image Y^δ corrupted by blur and *salt and pepper* noise. The problem is solved by a variational approach, minimizing an objective functional where the data fidelity term is weighted by a suitable regularization term. This formulation requires the choice of the regularization parameter which is very critical. Therefore we consider the constrained formulation where the minimization of the data fidelity term is constrained by a suitably defined regularization term. *Salt and pepper* noise is considered an outlier which do not obey to the Gaussian noise model. Minimization of outliers effects can be obtained by replacing the ℓ_2 in the data fidelity term by the ℓ_1 norm [1, 2, 3, 4]. Moreover the combination of ℓ_1 data fidelity term with edge preserving regularizers has proven to be effective [2, 5, 6, 7, 8]. In this paper, we adopt the Total Variation regularizer extended to Multichannel images (MTV) [9]. We solve the deblurring-denoising problem by the following sequence of constrained minimization problems:

$$min_U \|\mathcal{H}U - Y^\delta\|_1 \quad s.t. \quad MTV(U) \leq \gamma_j, \quad j = 0, 1, \ldots \tag{1}$$

where $\mathcal{H}$ is the blurring kernel (here $\mathcal{H} \neq I$), U is the exact image (unknown) to be recovered. The initial smoothing term γ_0 is obtained by the input data Y^δ by a filtering procedure specific for multichannel images. The successive terms of the sequence $\{\gamma_j\}$, $j > 0$, are updated by an iterative method based on the residuals of (1). In this sense, the proposed method does not depend on any regularization or constraint parameter and only the recorded data Y^δ are necessary as input. The problems (1) are solved in Lagrangian dual form by the algorithm `CL1TV` presented in section 2. Some numerical experiments are reported in section 3, where the tests are performed on deblurring of color images, corrupted by cross-channel blur and salt and pepper noise. Finally the conclusions are given in section 4.

2. The Constrained L1-TV Algorithm (CL1TV)

We now discuss our iterative algorithm, named the Constrained L1-TV Algorithm (CL1TV), that computes a suitable set of values γ_j and solves the constrained problem (1) using the Lagrangian dual form. We shall present the algorithm for the case of multichannel images with cross channel blur.

2.1. Multichannel image restoration

Assume our observed p-channel image (with blur) is $Y = [Y_1, Y_2, \ldots, Y_p]$, which is the sum of convolutions of the blurring kernels $H_{n,k}$ with the image channels U_k of the original image $U = [U_1, \ldots, U_p]$:

$$Y_n = \sum_{k=1}^{p} H_{n,k} * U_k, \quad n = 1, \ldots, p \tag{2}$$

where each channel $Y_n \in \mathcal{R}^{M\times N}, \;\; n = 1, \ldots, p$. For example, the blurring matrix of truecolor images $\mathcal{H}$ has 3×3 blocks ($p = 3$) where the diagonal blocks $H_{k,k}$, $k = 1 \ldots, 3$ represent within-channel blur, while the blocks $H_{n,k}$ $n \neq k$ represent the cross-channel blur, i.e. how the blur on the n-th channel influences the k-channel. For a general p-channel image, we consider discrete Multichannel Total Variation, as extension of the Total Variation for a grayscale image, defined as follows:

$$MTV(U) = \sum_{j=1}^{M}\sum_{i=1}^{N} \left(|\nabla U|^2_{i,j,1} + |\nabla U|^2_{i,j,2} + \cdots + |\nabla U|^2_{i,j,p}\right)^{1/2}.$$

Forward difference formulas are used in $|\nabla U|^2 = U_x^2 + U_y^2$ to approximate the derivatives.

2.2. Computation of the sequence γ_j

The starting value γ_0 of the sequence of regularization constraints $\gamma_j, j = 1, \ldots$, is defined as the Multichannel Total Variation of a low pass filtered version $U^{(F)}$ of the blurred noisy data. Such a filtered image is computed by solving the following unconstrained problem:

$$U^{(F)} = argmin_U \|\mathcal{H}U - Y^\delta\|_1 + \lambda_F MTV(U), \tag{3}$$

where λ_F is obtained by undersampling the sum of the spectra of the blurring matrices $H_{n,k}$ acting on each channel and taking the minimum of the Fourier coefficients. In the case of cross channel blur, we assume that each submatrix $H_{i,j}$ can be diagonalized by a unitary Fourier matrix F: $H_{i,j} = F^*\Psi^{(i,j)}F, \;\; i,j = 1, \ldots p$ and we approximate $\mathcal{H}$ by a block diagonal matrix $\bar{\mathcal{H}}$ where each diagonal block is

$$\bar{H}_{i,i} = F^*\bar{\Psi}^{(i)}F, \quad \bar{\Psi}^{(i)} = \frac{1}{p}\sum_{j=1}^{p} \Psi^{(i,j)}, \quad i = 1, \ldots, p.$$

By undersampling the power spectrum we compute the filter parameters β_i as:

$$\beta_i = \min_{1\le k\le N} \left| \frac{(FY_i^\delta)_{k\cdot M}}{\bar{\Psi}^{(i)}_{k\cdot M}} \right|, \quad i = 1, \ldots, p \tag{4}$$

and define $\lambda_F = \min_{1\le i\le p} \beta_i$. For each channel i, we have that $|(FY_i^\delta)_\ell|$ decay faster than $|(\bar{\Psi}^{(i)})_\ell|$ ($\ell = 1, \ldots, k\cdot M$) until they level off when the noise starts to dominate [10]. Hence we take the minimum value in (4) to limit as much as possible the influence of noise. Then $\gamma_0 = MTV(U^{(F)})$.

We experimentally observe that the residual function $r(\gamma) = \|\mathcal{H}U(\gamma) - Y^\delta\|_1$ decreases for increasing values of γ (for γ less than an optimal value $\tilde{\gamma}$). In particular $r(\gamma)$ steeply decreases when $\gamma \ll \tilde{\gamma}$ and becomes flat around $\tilde{\gamma}$. We can define an increasing sequence $\gamma_j, j = 1, \dots r$, and compute γ_j by applying a kind of cross validation method, minimizing the residual $r(\gamma)$. In practice, checking the slope of the residual curve, we compute γ_j with the following relation:

$$\gamma_j = \gamma_{j-1}(1+P), \quad \texttt{if } |r(\gamma_j) - r(\gamma_{j-1})| > P_{tol}|(\gamma_j - \gamma_{j-1})| \tag{5}$$

where P_{tol} is an assigned tolerance. Let γ_s be the exit value of (5) then, if the slope is still negative, the adaptation procedure is stopped ($\gamma_r = \gamma_s$). Otherwise, if the slope in γ_s is greater than P_{tol}, then a backtracking procedure (`BCKTRK`) is performed in the interval $[\gamma_{s-1}, \gamma_s)$, by using (5) with $P = P/2$ and $j = s$.

Algorithm 2.1 (Output: γ_r).

Compute λ_F *and compute* $U^{(F)}$ *by solving* (3). $\gamma_0 = MTV(U^{(F)})$. *j=0;*
`repeat`
    $j = j + 1; \quad \gamma_j = \gamma_{j-1}(1 + P)$
    $D_j = (r(\gamma_j) - r(\gamma_{j-1}))/(\gamma_j - \gamma_{j-1})$
`until` $D_j < -P_{tol}$
`if` $D_j < P_{tol}$
    $\gamma_r = \gamma_j$
`else`
    $a = \gamma_{j-1}$, $b = \gamma_j$
    compute $\gamma_r =$ `BCKTRK`$(a,\ b)$.
`end`

In our experiments, $P_{tol} = 10^{-4}$, $P = 0.1$.

2.3. Solution of the L1-TV subproblem

For each $\gamma = \gamma_j$ the constrained problem (1) is solved using its Lagrangian dual form:

$$\max_\lambda \min_U \mathcal{L}(U,\lambda), \quad \mathcal{L}(U,\lambda) \equiv \|\mathcal{H}U - Y^\delta\|_1 + \lambda\left(MTV(U) - \gamma\right). \tag{6}$$

Imposing the first order conditions $\nabla_\lambda \mathcal{L}(U,\lambda) = 0$, we compute the solution $(\hat{\lambda}, \hat{U})$ of (6) by solving:

$$\begin{gathered} \text{find } \hat{\lambda} \ \text{ s.t. } \ MTV(\hat{U}(\lambda)) - \gamma = 0 \quad \text{where } \hat{U} \equiv U(\hat{\lambda}) \text{ is the solution of} \\ \min_U \|\mathcal{H}U - Y^\delta\|_1 + \lambda\left(MTV(U) - \gamma\right) \end{gathered} \tag{7}$$

The monotonicity property of $MTV(U(\lambda)) - \gamma$ allows us to solve the nonlinear equation:

$$MTV(U(\lambda)) - \gamma = 0 \tag{8}$$

by a hybrid bisection+secant method [11] yielding a sequence $\{\lambda_k\}$ converging to the root $\hat{\lambda}$. Usually few bisection iterations are necessary ($k_s \simeq 3$) to guarantee the convergence of the secant iterations and globally less than 10 iterations are performed by the hybrid method when stopped by the following criterion:

$$|MTV(U^{(k)}) - \gamma| < \tau_r|MTV(U^{(0)}) - \gamma| + \tau_a \quad \text{or} \quad |\lambda_k - \lambda_{k-1}| < \tau_a \quad \text{or} \quad k > \texttt{maxit} \tag{9}$$

with $\tau_r = \tau_a = 10^{-3}$ and `maxit` $= 15$ in our experiments. By solving (7) with $\lambda = \lambda_k$, we obtain a sequence $U^{(k)} = U(\lambda_k)$ converging the solution $\hat{U}$. Using MTV an efficient and fast method for solving (7) is the fast alternating minimization algorithm proposed in [9] (FTVD4). However, different solvers could be applied to (7), related to different extensions of the TV to multichannel images. The procedure for solving a subproblem (1) is outlined in Algorithm 2.2.

Algorithm 2.2 (CL1TV – Input: γ_j, $\mathcal{H}, Y^\delta$, λ_0,k_s; output: U, λ).

compute $U^{(0)}$ *solving* (7) *with* $\lambda = \lambda_0$ *using FTVD4 function*
$k = 0$ *% Solution Computation*
`repeat`
    $k = k + 1$
    Apply a hybrid bisection-secant method for solving $MTV(U(\lambda_k)) - \gamma = 0$
    compute $U^{(k)}$ *solving* (7) *with* $\lambda = \lambda_k$ *using FTVD4 function*
`until` *exit condition* (9)
$U = U^{(k)}$.

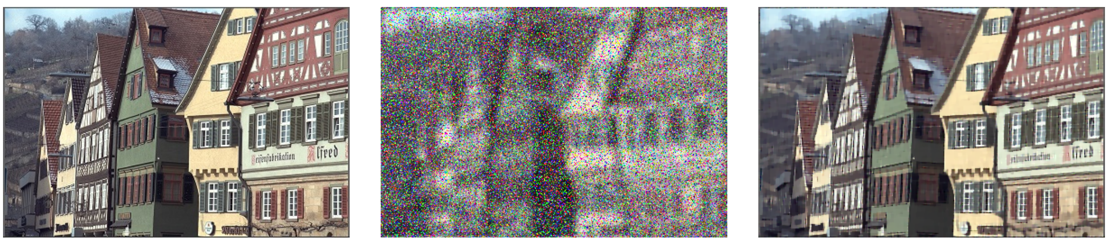
3. Numerical Results

In this section we present some numerical results obtained with the proposed `CL1TV` method in the deblurring of color images. The tests have been performed in Matlab 2012a. In all the experiments, the maximum number of allowed `CL1TV` iterations is 5 and the maximum number of bisection-secant iterations is 10. The blurred noisy images Y^δ are obtained by applying the blurring kernel $\mathcal{H}$ to the test image U^* and by adding salt and pepper noise. In our experiments, we consider two levels of noise: medium noise and high noise, where the number of corrupted pixels in Y^δ is 40% and 80%, respectively. The following cross-channel blurring matrix has been used:

$$\mathcal{H} = \begin{pmatrix} .7G(21,11) & .15G(21,11) & .15G(21,11) \\ .1G(21,11) & .8G(21,11) & .1G(21,11) \\ .2G(21,11) & .2G(21,11) & .6G(21,11) \end{pmatrix}. \tag{10}$$

where $G(h, \sigma)$ is a Gaussian function of size $h = 21$ pixels and variance $\sigma = 11$ pixels. The quality of the reconstructed images is evaluated through the relative error: $E = \|U^* - \tilde{U}\|_2 / \|U^*\|_2$ where U^* and $\tilde{U}$ are the exact and reconstructed image, respectively. We show the results on two color images:

- the I1 photographic image in figure 1;
- the I2 synthetic image of figure 2.



(a) Original (b) Noisy blurred (high noise) (c) Restored

Figure 1: Original, blurred (high noise) and restored images I1.

In figure 3 we report the relative errors (figures (a) and (c)) and the L1 residuals (figures (b) and (d)) as functions of the parameter γ for I1. We have obtained the plots by considering a

(a) Original (b) Noisy blurred (medium noise) (c) Restored

Figure 2: Original, blurred (medium noise) and restored images I2.

set of equispaced values γ_i in the interval $[8.10^3, 2.10^4]$ (sufficiently larger than $[\gamma_0, \gamma_s]$) and by solving problems of the form (1) with the proposed `CL1TV` method. From the plots (a) and (c) (medium and high noise, respectively) we see that the computed solution is very close to the best possible with the considered model. This means that the estimate of γ made by our algorithm is good. Figures (b) and (d) show that the residual really decreases and then becomes flat near the best value of γ; confirming the validity of method (5) based on the computed residual $r(\gamma)$. We remind that the relative errors reported in figures (a) and (c) are not used in the algorithm for choosing γ, but they are shown only to confirm the validity of the proposed method. In table 1 we report some numerical results obtained with the `CL1TV` method. It is evident that the relative error is strongly reduced from the recorded image ($E0$) to the computed image (Er). The computational time is proportional to the number of FTVD iterations (implementing the Alternating Directions algorithms).

Test (nl)	$E0$	Er	it_γ	it_λ	it_{FTVD}
I1 40%	0.477	0.139	5	33	15768
I1 80%	0.600	0.145	5	24	10941
I2 40%	0.306	0.050	5	37	5670
I2 80%	0.4166	0.051	5	27	4021

Table 1: `CL1TV` algorithm results. $E0$ is the relative error of the recorded image, Er the relative error of the restored image, it_γ the number of elements of the sequence γ_i, it_λ the total number of iterations of the bisection and secant methods, it_{FTVD} the total number of iterations of the FTVD method.

4. Conclusions

In this paper we proposed the `CL1TV` iterative method for automatic restoration of blurred noisy images corrupted by impulsive noise. The method only requires the recorded image and the blurring kernel as input. The algorithm solves a sequence of constrained minimization problems where the objective function is the ℓ_1-norm of the residual and the constraint function is the Total Variation (extended for vectorial images). The experiments carried on some images proved that the method efficiently recovers nearly optimal results; hence they are encouraging for further tests in deblurring and other linear inverse problems in imaging, such as tomographic

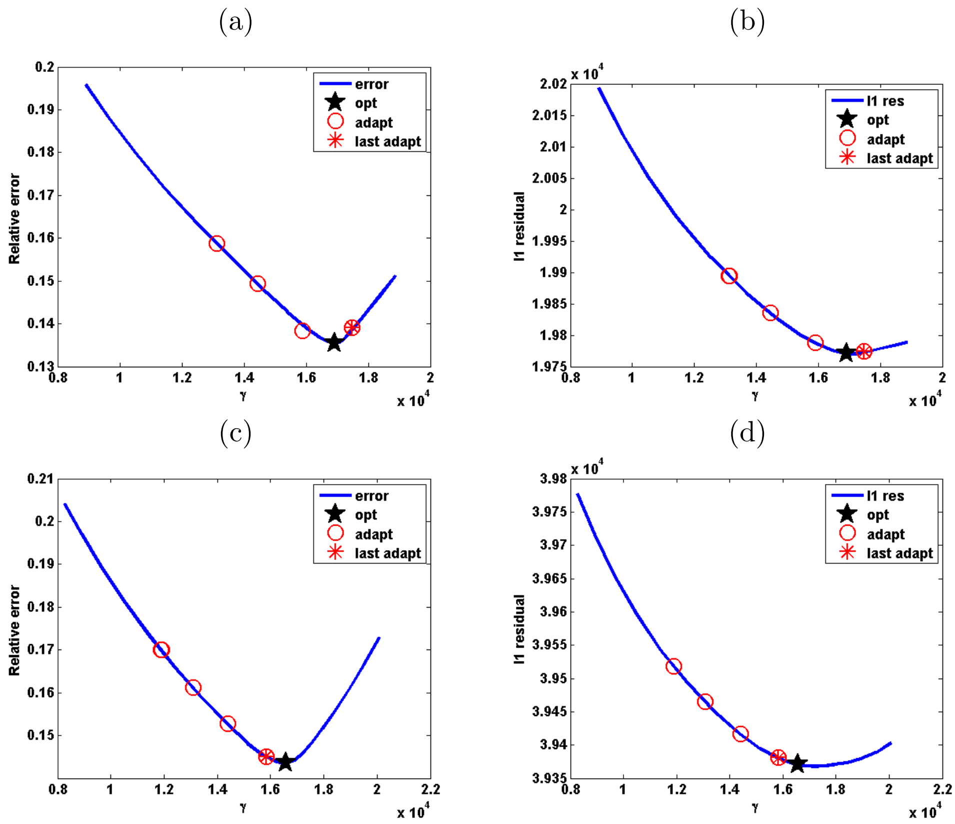


Figure 3: Relative errors (a)(c) and L1 residuals (b)(d) for the tests with medium 40% noise (a)(b) and high 80% noise (c)(d).

reconstruction. In future work we will compare our approach to other methods proposed in the literature and test different algorithms for the solution of the unconstrained problem (7).

Acknowledgement
The authors would like to thank the reviewers for their comments and suggestions.

References
[1] Nikolova M 2004 *Journal of Mathematical Imaging and Vision* **20** 99–120
[2] Chan R, Ho C W and Nikolova M 2005 *Image Processing, IEEE Transactions on* **14** 1479–1485
[3] Bar L, Brook A, Sochen N and Kiryati N 2007 *Image Processing, IEEE Transactions on* **16** 1101–1111
[4] Chan R H and Chen K 2008 *SIAM Journal on Scientific Computing* **30** 1474–1489
[5] Chan T and Esedoglu S 2006 *Siam J. Appl. Mathem.* **66** 1632–1648
[6] Karkkainen T, Kunish K and Majava K 2005 *Computing* **74** 353–376
[7] Clason C, Jin B and Kunish K 2010 *Siam J. Sci. Comput.* **32** 1484–1505
[8] Condat L and Mosaddegh S 2012 Joint demosaicking and denoising by total variation minimization *Image Processing (ICIP), 2012 19th IEEE International Conference on* pp 2781–2784
[9] Yang J, Zhang Y and Yin W 2009 *Siam J. Sci. Comput.* **31** 2482–2865
[10] Hansen P C, Nagy J and O'Leary D 2006 *Deblurring Images: Matrices, Spectra, And Filtering* (SIAM)
[11] Chen K, Piccolomini E L and Zama F 2014 *Numerical Algorithms* **67** 73–92