FILTERED VARIATION METHOD FOR DENOISING AND SPARSE SIGNAL PROCESSING

Kivanc Kose[⋆]   *Volkan Cevher*[†]   *A. Enis Cetin*[⋆]

[⋆]Department of Electrical and Electronics Engineering, Bilkent University
[†]Laboratory for Information and Inference Systems, Ecole Polytechnique Federale de Lausanne

ABSTRACT

We propose a new framework, called Filtered Variation (FV), for denoising and sparse signal processing applications. These problems are inherently ill-posed. Hence, we provide regularization to overcome this challenge by using discrete time filters that are widely used in signal processing. We mathematically define the FV problem, and solve it using alternating projections in space and transform domains. We provide a globally convergent algorithm based on the projections onto convex sets approach. We apply to our algorithm to real denoising problems and compare it with the total variation recovery.

Index Terms— Filtered variation, total variation, regularization, projection onto convex sets

1. INTRODUCTION

We introduce a new framework called Filtered Variation (FV) for denoising, compressive sensing (CS) and sparse signal processing applications. In this article, we use denoising as a running example to highlight the salient features of our approach. We first link the discrete time filtering with the optimization through the FV framework. We then solve the denoising problem using the proposed FV approach. In many denoising problems, we have access to a noisy version $\mathbf{y} = [y_1, \ldots, y_n]$ of the original signal $\mathbf{x} = [x_1 \ldots x_n]$ that can be represented as follows:

$$\mathbf{y} = \mathbf{x} + \mathbf{u} \qquad (1)$$

where $\mathbf{u}$ is the noise.

Total Variation (TV) based solutions are quite popular for denoising applications [1]-[8]. In discrete TV functional, the difference between neighboring samples are computed and the ℓ_1 or ℓ_2-norm of the difference vector is minimized. Hence, the TV method inherently assumes that the signal (or image) is low-pass signal and tries to minimize the high-pass energy. In signal processing, we have decades of experience on high-pass filter design. Therefore, instead of computing the difference between the samples we can filter the signal using an appropriate high-pass filter and minimize the ℓ_1 or ℓ_2 energy of the output. Furthermore, in image and video processing, it is possible to use diagonal and other directional high-pass filters in image denoising applications.

In FV based denoising, our goal is to find a solution to the following optimization problem:

$$\min \ \mathrm{FV}_p(\mathbf{x}) \qquad (2)$$
$$\text{s.t. } \|L\mathbf{x} - \mathbf{y}\| \leq \delta \qquad (3)$$

where FV stands for the filtered variation and it is defined as follows:

$$\mathrm{FV}_p(\mathbf{x}) = \|\mathbf{H}\mathbf{D}\mathbf{x}\|_p, \ p = 1, 2 \qquad (4)$$

where $\mathbf{X}, \mathbf{D}$ and $\mathbf{H}$ represent the signal, the signal transform (e.g., DCT, DHT, DFT) and the discrete-time filter in the transform domain, respectively and p denotes which ℓ_p-norm is used. In (3) and (4) the norm can be selected as ℓ_1 or ℓ_2 norms, which correspond to anisotropic and isotropic FV, respectively.

In our approach, denoising is achieved by minimizing the high-frequency energy of the observations, subject to the constraint given in (3). In (2)-(4) we posed the problem in frequency domain because for any given fixed transform, noise is typically in coherent with the transform, therefore it is spread out. By means of a proper filtering operation in the transform domain, one can exploit this fact to effectively denoise the signal. Besides, it is possible to solve the problem completely in time (or space) domain as well.

In article we solve this regularized signal denoising problem by applying several different time (space) and frequency domain constraints on filtered versions of the signal $\mathbf{x}$. This approach is similar to the methodology described in [9, 10, 11]. Since the FV cost function is convex it is also possible to solve FV based problems using convex programming. We provide a solution using the Projections onto Convex Sets (POCS) method. The following FV based constraints correspond to a class of convex sets:

$$C_i^p = FV_p(x) = \left\{ \|\mathbf{H}\mathbf{D}\mathbf{x}\|_p \leq \varepsilon \right\}, \ p = 1, 2 \text{ and } i = 1, \ldots, M. \qquad (5)$$

where $p = 1, 2$ corresponds to ℓ_1 and ℓ_2-norms respectively. Other closed and convex sets, described in Section 3 can be also imposed on the desired signal $\mathbf{x}$. The solution of the denoising problem is assumed to lie in the intersection of M different constraint sets as follows:

$$\mathbf{x} \in C = \bigcap_{i=1}^{M} C_i, \qquad (6)$$

where the constraint sets (C_i) are defined by the convex constraints as given at (5). Therefore, it is possible to reconstruct the original signal by performing successive orthogonal projections onto the closed and convex sets C_i [12, 13]. The POCS based iterative algorithm consists of making successive operations in time (or space) and transform domains, and it converges to a solution in the intersection of constraint sets C_i.

Extension to 2-D or higher dimensional signals is straightforward. Instead of a 1-D high-pass filter, 2-D or higher dimensional high-pass filters can be used in (6).

Organization of this paper is a follows. Section 2 explores the relationship between FV and TV models. Section 3 describes a class of useful filtering operations that can be used as constraints. Results are given in Section 4 followed by conclusions in Section 5.

2. FILTERED VARIATION BASED REGULARIZATION

In this section, we first review the TV approach and motivate FV for sparse signal recovery by using a high-pass filtering example.

TV functional was first introduced to signal and image processing problems by Rudin, Osher and Fatemi in 1990's [1]-[8]. For a 1-D signal $\mathbf{x}$ of length N, the discretized TV functional of $\mathbf{x}$ is defined as,

$$TV(\mathbf{x}) = \sum_{n=1}^{N} \sqrt{(x[n] - x[n+1])^2} \qquad (7)$$

where a discrete-gradient of the signal is the key component of the TV functional.

We note that the discrete gradient operation $v[n] = x[n] - x[n+1]$ in Equation (7) is a rough high-pass filtered version of $\mathbf{x}$. This filter is the high-pass filter used in Haar wavelet transform. Therefore, the relation between the signals $\mathbf{x}$ and $\mathbf{v}$ can be represented via convolution denoted by the operator $*$ as follows:

$$v[n] = h[n] * x[n] \qquad (8)$$

where $h[n] = \{-1, 1\}$ is the impulse response of Haar high-pass filter. In the DFT domain the same relationship can be represented by a multiplication operation as follows:

$$V[k] = H[k]X[k], \quad k = 1, 2, ..., N. \qquad (9)$$

In 9, $X[k], H[k], V[k]$ are the N-point DFT of the desired signal $x[n]$, high-pass filter $h[n]$ and the output $v[n]$, respectively. The TV cost function is equivalent to filtering the signal with a Haar high-pass filter and computing the ℓ_1 or ℓ_2 energy of the filtered output signal corresponding to anisotropic or isotropic cases, respectively.

The Haar filter has an ideal normalized angular cut-off frequency of $\frac{\pi}{2}$. It is possible to apply other high-pass filters and compute the output energy or it is possible to use the Parseval's relation and other Fourier domain relations to impose sparsity conditions on the desired signal. It is well-known [14] that:

$$\sqrt{\sum_{n} |v[n]|^2} = \sqrt{\sum_{k} \frac{1}{N} |V[k]|^2} \leq \max_{k} |V[k]| \leq \sum_{n} |v[n]| \ . \qquad (10)$$

In Section 3, based on the above relations, we define both time (space) and frequency domain FV constraints, which correspond to closed and convex sets for the CS problem.

It is also possible to define constraint sets on other transform domain representations, such as wavelets, but we focus on DFT and DCT domain in this article. The filtered output in transform domain $V[k] = H[k]X[k]$ is basically specified by the filter $\mathbf{H}$, which can be selected according to a given bandwidth specified by the user. In 2-D or higher dimensions, one is not restricted to horizontal or vertical high-pass filters. It is also possible to use directional high-pass filters.

3. FILTERED VARIATION ALGORITHM AND TRANSFORM DOMAIN CONSTRAINTS

In this section, we list six closed and convex constraints that can be used for image denoising. Each constraint qualifies different properties of the estimated signal such as; ℓ_1 or ℓ_2 energy of the high frequency band of the signal, local variations in the signal, the mean of the signal and the bit depth of the image. As in [3], all the constraints can be used at the same time, or any combination of these can be used together depending on the nature of the signal (or image) and the noise type.

Constraint-I - ℓ_1 FV Bound: The first constraint is based on the ℓ_1 energy of high frequency coefficients:

$$C_1 = \left\{ \mathbf{x} : \sum_{k=0}^{N-1} |H[k]X[k]| \leq \varepsilon_1 \right\}. \qquad (11)$$

It is possible to perform orthogonal projections onto this set in Discrete Time domain as described in [9]. Since, the DFT is a complex transform, it is easier to work with a real transform such as DCT or DHT. In this case the boundary hyperplanes of the region specified by the constraint set are real. The projection operation is essentially equivalent to making orthogonal projections onto hyperplanes forming the boundary, and it is similar to projection onto an ℓ_1 ball but it is on the transform domain and only high-frequency coefficients are updated. Since we perform projections onto an ℓ_1 ball type region, the solution that we obtain turns out to be sparse.

Constraint-II - Time and Space Domain Local Variational Bounds: The second constraint is based on the change in intensity between the consecutive samples of a signal (pixels of the image). In real-life, there is strong correlations between the samples of discrete-time signals (or images), and there is very little correlation between different parts of the signals (or images). Therefore, it is possible to remove the summation operator in TV or FV and consider regional TV or FV constraints on the signal. This leads to a high-pass constraint set for each sample of the signal (or pixel of the image):

$$C_{2,n} = \left\{ \left| \sum_{i=-l}^{l} h[i]x[n-i] \right| \leq P \right\}, \qquad (12)$$

where $h[i]$ is a high-pass filter with support length $2l + 1$ and P is a user defined bound. Projection onto hyperslabs $C_{2,n}$ do not correspond to low-pass filtering, because projections are essentially nonlinear operations. If the current iterate does not satisfy the bound, it is projected onto the hyperslab given in (12).

In practice, we cannot select a small bound value, unless we do not have a clear idea about the low-pass nature of the signal. The local FV constraint is especially useful to remove impulsive noise.

In image processing experiments we selected a very large bound ($P = 128$) for the high-pass filter $\mathbf{h} = \{\frac{-1}{4}, \frac{1}{2}, \frac{-1}{4}\}$ to avoid distorting the edges of the desired image. When there is an impulse within the analysis window of the filter, the filter output will be high and the samples within that window are modified by the projection. For example, the $C_{2,n}$ family of sets turn out to be useful for Laplacian noise. In image processing examples that we studied in the next section, it is also possible to apply filters in vertical and diagonal directions depending of the nature of the original image.

Constraint-III - Bound on High Frequency Energy: The following anisotropic constraint on high-frequency energy of the signal $\mathbf{x}$ is a closed and convex set:

$$C_{3a} = \left\{ \mathbf{x} : \sum_{k=k_0}^{N-k_0} |X[k]|^2 \leq \varepsilon_{3a} \right\} \qquad (13)$$

where ε is an upper bound. This corresponds to filtering the signal $\mathbf{x}$ with a high-pass filter whose cut-off frequency index is k_0 in the DFT domain:

$$H[k] = \begin{cases} 0, & \text{for } k < k_0 \text{ or } k > N - k_0 \\ 1, & \text{for } k_0 \leq k \leq N - k_0 \end{cases} \qquad (14)$$

where N is the size of the DFT. Although this filter suffers from the Gibbs phenomenon in time-domain, it is possible to use it in

denoising problems. The index k_0 is equal to $\frac{N}{4}$ for the normalized angular cut-off frequency of $\frac{\pi}{2}$, but any $0 < k_0 < \frac{N}{2}$ can be selected for a desired smoothness level. The set given in Eq. (13) is a convex set and it is easy to perform orthogonal projections onto this set. Let $s_o[n]$ be an arbitrary signal and $S_0[k]$ be its DFT. $S_p[k]$ of the projection $s_p[n]$ is given by

$$S_p[k]=\begin{cases}\sqrt{\frac{\varepsilon}{\varepsilon_o}}S_0[k]\text{ , if }\sum_{k=k_0}^{N-k_0}|S_0[k]|^2\geq\varepsilon,\ k_o\leq k\leq N-k_o\\[2mm] S_0[k],\text{ otherwise,}\end{cases} \tag{15}$$

where $\sum_{k=k_0}^{N-k_0}|S_o[k]|^2=\varepsilon_o$.

We can also use a DCT domain high-pass energy constraint on the desired signal using the following set

$$C_{3b}=\left\{\mathbf{x}:\sum_{k=k_0}^{N-1}(X_{DCT}[k])^2\leq\varepsilon_{3b}\right\}, \tag{16}$$

which is also a convex set. In (16), $\mathbf{X}_{DCT}$ represents the DCT of the signal $\mathbf{x}$. It is straightforward to make orthogonal projections onto the DCT domain set C_{3b} as in Equation (15).

Constraint-IV - User Designed High-pass Filter: In this case, we do not assume a specific cut-off frequency but use the frequency response of a given high-pass filter:

$$C_4=\left\{\mathbf{x}:\sum_{k=0}^{N-1}|H[k]X[k]|^2\leq\varepsilon_4\right\}. \tag{17}$$

The set C_4 is also a closed and convex set. Orthogonal projection onto this set is not as easy as Condition-I, because the set is a closed ellipsoid. It can be implemented using numerical methods, [15, 16].

Constraint-V - The Mean Constraint: The fifth constraint is actually proposed in [3]. It is based on the mean of the original signal. Typically this information can be estimated from a pool of similar types of images (e.g. satellite images, images of hand-writing, faces etc.) A constraint based on the mean information can be defined as follows:

$$C_5=\left\{\mathbf{x}:\sum_{n=1}^{N}\frac{x[n]}{N}=\mu_x\right\} \tag{18}$$

where N is the number of the pixels in the image and μ_x is the mean of the original image.

Constraint-VI: Image bit-depth constraint In general, the users know the color (bit) depth of the original image. Due to this fact, it is possible to define a constraint on the bit depth of the reconstructed image as follows:

$$0\leq x[i,j]\leq(2^M-1) \tag{19}$$

where M is the number of the bit planes used in the original representation. This constraint is also proposed in [3].

4. EXPERIMENTAL RESULTS

We first present a denoising example from [3]. Combettes and Pesquet used the image shown in Fig.1-(a) in [3], to test their TV based denoising algorithm. They added i.i.d. Laplacian noise to the original 128x128 grayscale image. The signal-to-noise ratio is 1dB. To compare the FV algorithm to the TV denoising, we cropped the original image (Fig. 1-(a)) from their paper and added Laplacian noise to the image. In [3] the pixel range was [-261,460]. In our case the pixel range turns out to be [-391,511].

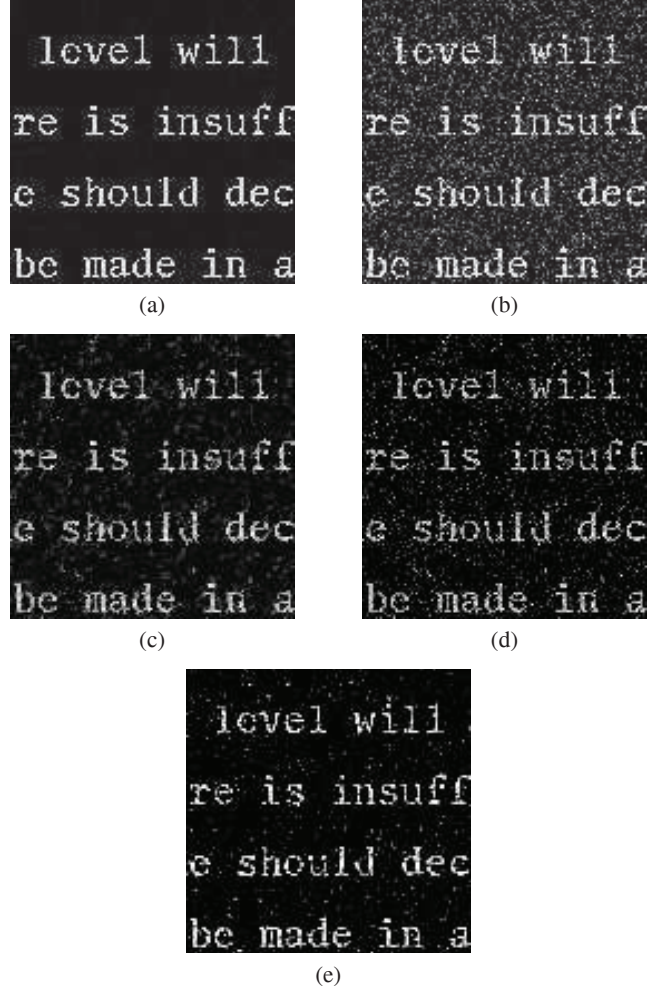As shown in Fig. 1 the characters in the image that are recovered

Fig. 1. (a) Original image. (b) noisy image. (c) ℓ^p denoising with bounded total variation and additional constraints [3] (Fig. 15 from [3]) (p=1.1). (d) ℓ^p denoising without the total variation constraint [3] (Fig. 16 from [3]). (e) Denoised image using the FV method using C_2, C_4 and C_5.

by our algorithm (Fig.1-(e)) are visually sharper compared to Fig.1-(c) and the impulsive noise is significantly reduced compared to ℓ_1 denoising.

The progress of the decrease in Normalized Root Mean Square Error (NRMSE), which is defined as $||x-x_o||/||x_o||$ in [3], is shown in Fig. 2. Our algorithm converges to an NRMSE level of -9 dB in 10-to-12 iterations. On the other hand the time-domain TV algorithm takes around 100 iterations to converge as shown in Fig. 18 in [3]

We also performed an experiment about estimating ℓ_1 and ℓ_2 high-frequency energy bounds ε_1 and ε_3, respectively. In this experiment, we try to estimate these bounds from the noisy image. Bounds are selected as 80% of ℓ_1 (ε_{1a}), 60% of ℓ_1 (ε_{1b}) and 80% of the ℓ_2 (ε_{3a}) energies of the noisy image, respectively. ε_{1o} corresponds to the ℓ_1 energy of the original image. Experimental results indicate that estimating ε_1 and ε_3 are possible from flat portions of the image and the FV algorithm is not sensitive to the ε_1 and ε_3 values. As
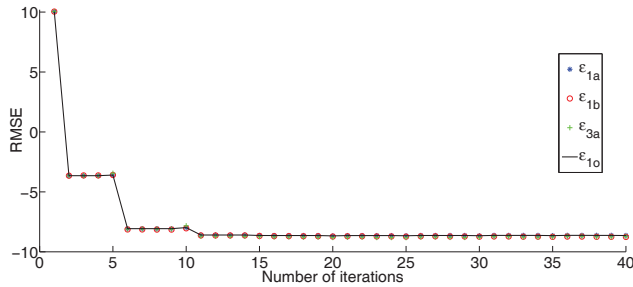
Fig. 2. NRMSE vs. iteration curves for FV denoising the image shown in Fig. 1. ε_{1o} and ε_{3o} correspond to the ℓ_1 and ℓ_2 energy of the original image. Bounds are selected $\varepsilon_{1a} = 0.8\varepsilon_{1o}$, $\varepsilon_{1b} = 0.6\varepsilon_{1o}$, and $\varepsilon_{3a} = 0.8\varepsilon_{3o}$

shown in Fig. 2, in all cases restored images are very close to each other. Convergence graphs closely overlap with each other as shown in Fig.1
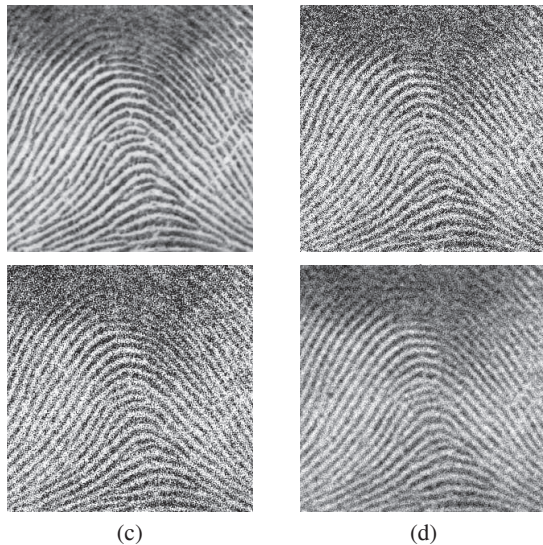

(c)   (d)

Fig. 3. (a) Original fingerprint image, (b) fingerprint image with 15 db AWGN. (c) Image Restored by TV constraint (SNR=7.45dB). (d) Image restored by the proposed algorithm using C_2, C_4 and C_5 (SNR=12.75 dB)

We also conduct an experiment with an image of a fingerprint as shown in Fig.3-(a). We added white Gaussian Noise to the original signal and obtained the noisy image shown in Fig. 3-(b) with SNR=4.9dB. Using FV constraints we obtained a reconstructed version of the signal whose SNR=12.75 dB (Fig. 3-(d)). On the other hand, TV constraint leads to an image with SNR=7.45dB (Fig. 3-(c)).

5. CONCLUSION

Filtered variations framework is developed for denoising and sparse signal processing applications. In this method, regularization is achieved by using discrete-time high-pass filters instead of taking the difference of neighboring signal samples as in the TV method. The FV based denoising problem is solved by making alternating projections in space and transform domains. It is experimentally observed FV approach provides better denoising results compared to the TV approach. If some prior knowledge about the original

signal exists, it is possible to design high-pass filters according to the signal and incorporate it to the FV framework.

6. REFERENCES

[1] Leonid I. Rudin, Stanley Osher, and Emad Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, November 1992.

[2] R. Davidi, G.T. Herman, and Y. Censor, "Perturbation-resilient block-iterative projection methods with application to image reconstruction from projections," *International Transactions in Operational Research*, vol. 16, no. 4, pp. 505–524, 2009.

[3] Patrick L. Combettes and Jean-Christophe Pesquet, "Image restoration subject to a total variation constraint," *IEEE Trans. Image Processing*, vol. 13, pp. 1213–1222, 2004.

[4] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin, "An iterative regularization method for total variation-based image restoration," *Simul*, vol. 4, pp. 460–489, 2005.

[5] T. F. Chan, S. Esedoglu, F. Park, and A. Yip, "Mathematical models in computer vision: The handbook, ch. recent developments in total variation image restoration," *Springer*, pp. 17–32, 2005.

[6] Dan Butnariu, Ran Davidi, Gabor T. Herman, and Ivan G. Kazantsev, "Stable convergence behavior under summable perturbations of a class of projection methods for convex feasibility and optimization problems," 2007.

[7] F. Malgouyres, "Minimizing the total variation under a general convex constraint for image restoration," *Image Processing, IEEE Transactions on*, vol. 11, no. 12, pp. 1450 – 1456, dec 2002.

[8] M Persson, D Bone, and H Elmqvist, "Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography," *Physics in Medicine and Biology*, vol. 46, no. 3, pp. 853, 2001.

[9] A. E. Çetin, "An iterative algorithm for signal reconstruction from bispectrum," in *IEEE Transactions on Signal Processing*, 1991, vol. 39, pp. 2621–2628.

[10] H.J. Trussell and M. R. Civanlar, "The landweber iteration and projection onto convex set," in *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1985, vol. 33, pp. 1632–1634.

[11] K. Kose and A.E. Cetin, "Low-pass filtering of irregularly sampled signals using a set theoretic framework," *Signal Processing Magazine, IEEE*, vol. 28, no. 4, pp. 117 –121, july 2011.

[12] L. M. Bregman, "The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 200–217, 1967.

[13] D. C. Youla and H. Webb, "Image restoration by the method of convex projections, part i-theory," *IEEE Transactions on Medical Imaging*, vol. MI-I-2, pp. 81–94, 1982.

[14] A. E. Çetin, "Reconstruction of signals from fourier transform samples," *Signal Processing*, vol. 16, pp. 129–148, 1989.

[15] Michael Elad and Arie Feuer, "Restoration of a single super-resolution image from several blurred, noisy, and undersampled measured images," 1997.

[16] Yu-Hong Dai, "Fast algorithms for projection on an ellipsoid," vol. 16, no. 4, pp. 986–1006, 2006.