



Incorporate $TV - l^\infty$ model with Sparse Representations for Image Denoising, a post-processing approach

Tieyong Zeng

► To cite this version:

Tieyong Zeng. Incorporate $TV - l^\infty$ model with Sparse Representations for Image Denoising, a post-processing approach. 2006. <hal-00120876>

HAL Id: hal-00120876

<https://hal.science/hal-00120876v1>

Preprint submitted on 18 Dec 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

INCORPORATE $TV - L^\infty$ MODEL WITH SPARSE REPRESENTATIONS FOR IMAGE DENOISING, A POST-PROCESSING APPROACH

Tieyong Zeng

LAGA/L2TI, Université Paris 13
99, avenue Jean-Batiste Clément 93430 Villetaneuse, France
E-mail: zeng@math.univ-paris13.fr

ABSTRACT

Sparse representations of images have revoked remarkable interest recently. The assumption that natural images admit a sparse decomposition over a redundant dictionary leads to efficient algorithm for image processing. In particular, the K-SVD method has been recently proposed and shown to perform very well for gray-scale and color image denoising task ([1],[2]). Meanwhile, the $TV - l^\infty$ model with special choice of dictionary has been proved to be very effective for image restoration([3],[4]). In this paper, we propose a hybrid model which combines these two methods and may be regarded as a post-processing procedure for K-SVD. Due to the excellent work of K-SVD and the fact that $TV - l^\infty$ can reconstruct lost information quickly, this hybrid model lead to a new state-of-art denoising performance.

Index Terms— Sparse representation, total variation, image denoising, dictionary, K-SVD, post-processing

1. INTRODUCTION

In this paper, we consider the classic image denoising problem: an ideal image x is observed in the presence of an additive zero-mean Gaussian white noise of standard deviation σ . Thus the observed image y is:

$$y = x + b. \quad (1)$$

Our task is to design an algorithm to remove the noise from y , getting as close as possible to the original image x and in the mean time, getting a better visual quality if that is possible.

Recently, Michael Elad and Michal Aharon proposed an image denoising method via sparse and redundant representations over learned dictionaries([1]). This leads to state-of-the-art denoising performance. The importance of this method is that it can recover most of the information in the noisy image while there is very little wash-out effect(especially for the face region of image Barbara). That's to say, it is able to avoid the shortcoming of most total variation (TV) based denoising method.

The little drawback of this approach is that it has some so called "checkboard" effect and it sometimes still lost some texture information. It is well known that total variation model can avoid the checkboard effect and TV-Gabor([3]) has proved to be very effective for texture restoration. So in this paper, we try to incorporate total variation concept into K-SVD model. This will lead to a new state-of-art denoising performance.

1.1. K-SVD Denoising Algorithm

In this section, we will briefly review the main mathematical framework of the approach of Elad([1]), as this is one of starting point of our work. First let the clean image x be written as a column vector of length N . Given fixed-size patches $\sqrt{n} \times \sqrt{n}$, we assume that all such patches in the clean image x admit a sparse representation. Addressing the denoising problem as a sparse decomposition technique for each patch leads to the following energy minimization problem:

$$\{\alpha_{i,j}, \hat{D}, \hat{x}\} = \arg \min_{\mathcal{D}, \alpha_{i,j}, x} \lambda \|x - y\|_2^2 + \sum_{i,j} \mu_{i,j} \|\alpha_{i,j}\|_0 + \sum_{i,j} \|\mathcal{D}\alpha_{i,j} - R_{i,j}x\|_2^2. \quad (2)$$

In this equation, \hat{x} is the estimator of x , and the dictionary $\hat{D} \in \mathbb{R}^{n \times K}$ is an estimator of the best dictionary which gives the sparsest representation of the patches in the restoration image. The indices $[i, j]$ mark the position of the patch in the image(representing it's top-left corner). The binary matrix $R_{i,j}$ of $n \times N$ extracts the square $\sqrt{n} \times \sqrt{n}$ patch of coordinates $[i, j]$ from the image represented by a column vector of size N . $\mu_{i,j}$ is hidden parameter which is decided implicitly by the method of solving the above equation.

The first term of the above equation demands a proximity between \hat{x} and y . The second and the third term both pose the image prior. This regularization term assumes that good-behaved natural images are to have a sparse representation for every patch, and from every position in the image, over the learned dictionary \hat{D} . The second term ensures the sparsest representation, and the third term forces the consistency of the decomposition.

1.2. $TV - l^\infty$ Algorithm for Denoising

Our second starting point is the works of ([3]). They investigate the solution provided by the following model:

$$\begin{cases} \text{minimize } TV(x) \\ \text{under the constraint } \|x - y\|_{\mathcal{D}_0, \infty} \leq \tau \end{cases} \quad (3)$$

where $\|\cdot\|_{\mathcal{D}_0, \infty}$ is defined by

$$\|x\|_{\mathcal{D}_0, \infty} = \sup_{\psi \in \mathcal{D}_0} |\langle x, \psi \rangle|,$$

for a finite dictionary $\mathcal{D}_0 \subset \mathbb{R}^N$ and the choice of τ depends on the noise level σ . For an explicitly definition of discrete definition of TV of image x , please see [3] and references cited therein. Typical choice of \mathcal{D}_0 is but not limited to wavelet basis, wavelet packets, translate-invariant Gabor dictionary or translate of small special patches. Using the idea of penalization, (3) is solved by minimizing the unconstrained energy

$$\hat{x} = \arg \min_x TV(x) + \gamma \sum_{\psi \in \mathcal{D}_0} \varphi_\tau(\langle x - y, \psi \rangle), \quad (4)$$

with

$$\varphi_\tau(t) = (\sup(|t| - \tau, 0))^2,$$

and for a large number γ . This model assumes the original image x has small total variation and forces the residual $x - y$ to be smaller than τ in every special direction $\psi \in \mathcal{D}_0$. This leads to an efficient restoration method which preserves edges and texture very well and has little checkboard effect.

Now before presentation of our new hybrid model, let's do some comparisons for these two models. Comparing to $TV - l^\infty$ model, K-SVD model gives higher $PSNR$ and has very few washout effect, especially for the face region of Barbara; comparing to K-SVD model, $TV - l^\infty$ model has more chance of avoiding the checkboard effect. Both methods can recover most part of the texture and sometime they can be regarded as complementary. And Do NOT forget that both methods use energy-minimization idea. So why not combine these two?!

This will lead to our new approach.

Meanwhile, we should point out that incorporate TV into sparse representation is not a totally new idea. In fact, the authors of [5] have considered similar model together with special choosing dictionaries which leads to an efficient image decomposition method.

2. TV-SPARSEST HYBRID MODEL

From the above discussion, we now propose our new approach. In order to solve the denoising task, we try to use the following new energy minimization problem:

$$\{\hat{\alpha}_{i,j}, \hat{D}, \hat{x}\} = \arg \min_{\mathcal{D}, \alpha_{i,j}, x} \lambda \|x - y\|_2^2 + \sum_{i,j} \mu_{i,j} \|\alpha_{i,j}\|_0 +$$

$$\sum_{i,j} \|\mathcal{D}\alpha_{i,j} - R_{i,j}x\|_2^2 + \beta TV(x) + \gamma \sum_{\psi \in \mathcal{D}_0} \varphi_\tau(\langle x - y, \psi \rangle). \quad (5)$$

Now in the new Eq.5, the first three terms are just the same as the Eq.2 and the last two are from Eq.3. The term $TV(x)$ forces the estimator image \hat{x} to be total variation limited and $\sum_{\psi \in \mathcal{D}_0} \varphi_\tau(\langle x - y, \psi \rangle)$ makes the energy of residual $\hat{x} - y$ on the special interesting direction no higher than τ . With all these constraints, better result is expected.

3. NUMERICAL ASPECTS

To approximate a solution for Eq.5, we first use the K-SVD procedure of ([1]) to minimize (2) and later we use another TV iteration to minimize the last two terms. Basically, this algorithm can also be regarded as initialization with K-SVD result and then TV -iterations as post-processing procedure. We present the details in Table 1.

In the above denoising procedure, follow with ([1]), we assume that σ is known and then we set $J = 10$, $\lambda = 30/\sigma$, $C = 1.15$ (another choice is 0.93 which is based on Rayleigh law, see ([2])). As our algorithm only depends on γ/β , we can fix β as 1.0. Being a penalization parameter, the typical choice of γ is between 9999 and 999999. $\tau = 3.5\sigma$ is empirically from ([3]). We discuss a little about the choice of k .

After K-SVD procedure, we can already get a fairly good estimator \hat{x} . Now if we aim to obtain an estimator \hat{x} which is as near as possible to the original image, i.e. we need a higher $PSNR$, we can stop the TV iteration when:

- $TV(\hat{x})$ reaches its top

This is not strange, when the estimator is near the original image, the higher the TV , the higher possibility of presence of recovered structures, then the higher the $PSNR$. Typical choice is $k = 2$ or 3.

On the other side, if we want better visual quality (less artifact, better texture, better structure), we can iterate more (but not too much else this will cause wash out effect) to get a smaller TV . But in the same time, we should note that this will surely make $PSNR$ a slightly lower. Our tests show that this time $k = 15$ is a good choice.

4. EXPERIMENTS RESULTS

We report our experiments for $\sigma = 20, 30$ on Barbara image. As the same in [1], we assume that σ is already known or could be estimated from elsewhere. We use Gabor dictionary (Gabor II of ([3]), large size), as we think that this one can recover better high-frequency information which has higher possibility than low-frequency to have been lost during the K-SVD procedure. The sum of FFT of the 145 filters in this dictionary is shown as Fig.1.

Task: Denoise a given image y from Gaussian noise of standard deviation σ via solve Equation (5).

Parameters: n -block size, K -size of first dictionary, J -number of K-SVD iterations, λ -Lagrange multiplier, C -noise gain, β -penalization parameter for TV , γ -penalization parameter for special direction of residual $x - y$, k -number of TV iterations.

1. Set $\hat{x} = y$, \mathcal{D} =overcomplete DCT dictionary.

2. Repeat J times:

- Sparse Coding Stage:

Use OMP pursuit algorithm to compute the representation vectors $\alpha_{i,j}$ for every fixed patch $R_{i,j}\hat{x}$, by approximating the solution of

$$\min_{\alpha_{i,j}} \|\alpha_{i,j}\|_0 \text{ subject to } \|R_{i,j}\hat{x} - \mathcal{D}\alpha_{i,j}\|_2^2 \leq (C\sigma)^2.$$

- Dictionary Update Stage:

For each column $l = 1, 2, \dots, k$ in \mathcal{D} , update it by

- Find the set of patches that use this atom, $\omega_l = \{(i, j) | \alpha_{i,j}(l) \neq 0\}$
- Find every index $(i, j) \in \omega_l$, compute its representation error

$$e_{i,j}^l = R_{i,j}\hat{x}_{i,j} - \sum_{m \neq l} d_m \alpha_{i,j}(m)$$

- Set E_l as the matrix whose columns are $\{e_{i,j}^l\}_{(i,j) \in \omega_l}$
- Apply SVD decomposition $E_l = U\Delta V^T$. Choose the updated dictionary column \tilde{d}_l to be the first column of U . Update the coefficient values $\{\alpha_{i,j}(l)\}_{i,j \in \omega_l}$ to be entries of V multiplied by $\Delta(1, 1)$.

3. Set:

$$\hat{x} = (\lambda I + \sum_{i,j} R_{i,j}^T R_{i,j})^{-1} (\lambda y + \sum_{i,j} R_{i,j}^T \mathcal{D} \alpha_{i,j})$$

4. Repeat k times:

- Calculate direction of gradient:

$$w = \text{div}\left(\frac{\nabla \hat{x}}{|\nabla \hat{x}|}\right) - \frac{\gamma}{\beta} \sum_{\psi \in \mathcal{D}_0} \varphi'_\tau(\langle \hat{x} - y, \psi \rangle) \psi$$

- Find the optimal step :

$$t = \arg \min_{s \in \mathbb{R}} \beta TV(\hat{x}) + \gamma \sum_{\psi \in \mathcal{D}_0} \varphi_\tau(\langle \hat{x} - y, \psi \rangle)$$

- Update \hat{x} :

$$\hat{x} = \hat{x} + tw.$$

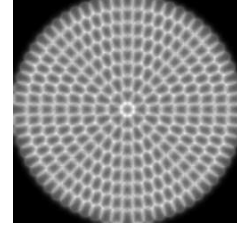


Fig. 1. Sum of Fourier transforms of the 145 filters in the Gabor II dictionary (big size). This image is larger size version of Bottom-Right of Fig.1 in [3].

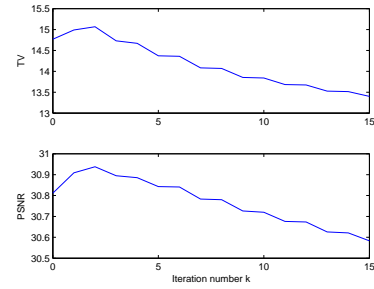


Fig. 2. The relationship $k-TV(\hat{x})$ and $k-PSNR(\hat{x})$ where k is the TV iteration number of Table 1. Note that $k = 0$ is result of K-SVD of Elad.

4.1. Noise level of $\sigma = 20$.

Our first experiment is to denoise the Barbara image with noise level $\sigma = 20$, the $PSNR$ of noisy image is 22.0977.

Fig.2 shows the relationship between $TV(\hat{x})$ and $PSNR(\hat{x})$ with the iteration number k in the main algorithm (see step 4 of Table 1). This Fig. tells us that after about 2 or 3 TV iteration, $TV(\hat{x})$ and $PSNR(\hat{x})$ reach their tops at the same time. So if we aim to get a higher $PSNR$ denoising result, we should stop at the iteration where $TV(\hat{x})$ reaches its top.

For our experiment, the highest $PSNR$ is 30.9376 of our new approach when $k = 2$, this is a slightly higher than K-SVD of Elad (30.8113) which claimed state-of-art denoising performance and much higher than the classical Rudin-Osher-Fatemi method (24.6759).

Meanwhile, if we want to a better visual quality, we can continue the TV iteration and $k = 10$ to 15 is a typical choice. Fig.3 displays a slice of left-bottom part of the Barbara with $k = 15$ from which we clearly know our approach is more effective.

4.2. Noise level of $\sigma = 30$.

In our experiment, when $\sigma = 30$ the $PSNR$ of the slice is 18.5448. Fig.4 shows the result for the same slice as Fig.3.

Table 1. General form of the algorithms. σ is known. Step 1 to 4 is typical K-SVD procedure([1]) to minimize (2). This algorithms can be regarded as initialization with K-SVD result and then a TV -iteration as post-processing.

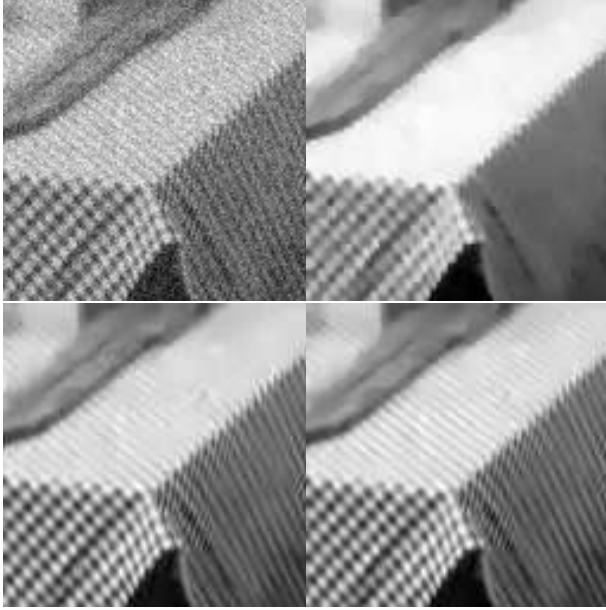


Fig. 3. Denoising a 128×128 slice of Barbara. From Left to right and from top to bottom: Noisy image ($\sigma = 20$), PSNR 22.0896; Rudin-Osher-Fatemi, PSNR 24.2663; K-SVD of Elad, PSNR 28.9013; Our new approach, PSNR 29.1148.

From this Fig. we obviously see that our new approach performs best. Both Rudin-Osher-Fatemi method and K-SVD of Elad fail to recover the texture of tablecloths on the desk, while our approach still can recover most of the information. And for the left part this slice, Rudin-Osher-Fatemi lost the texture information and K-SVD can recover some of this information but the texture is a little disordered. But for our new approach, this is a different story, we have recovered more information and the texture is still arranged.

Globally, for this level of noise, the *PSNR* of noisy image and the result of Rudin-Osher-Fatemi, K-SVD, and our approach are respectively 18.5867, 24.0429, 28.5947, and 28.8376.

5. CONCLUSION

In this paper, we have proposed a hybrid model which combines the sparse representation and total variation concept. In our current methodology of solving the model, this can also be regarded as initialization $TV - l^\infty$ model with K-SVD or as a post-processing for K-SVD. Solving the hybrid model directly or choosing \mathcal{D}_0 adaptively to noisy image may lead to better result and thus could be future work. Due to the leading performance of K-SVD and the fact that $TV - l^\infty$ with Gabor dictionary procedure can reconstruct the lost information quickly, our hybrid model attains a new state-of-art denoising performance, equivalent or surpassing recently published leading alternative methods including K-SVD itself.

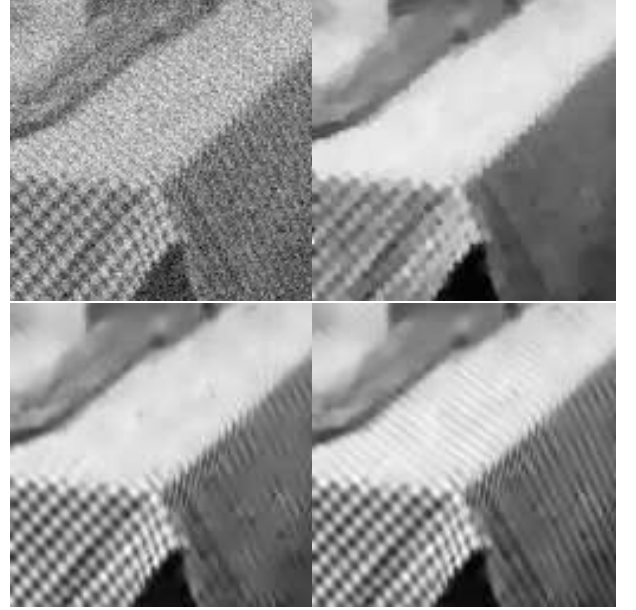


Fig. 4. Denoising the same slice of Barbara as Fig.3. From Left to right and from top to bottom: Noisy ($\sigma = 30$), PSNR 18.5448; Rudin-Osher-Fatemi, PSNR 23.4331; K-SVD of Elad, PSNR 26.4467; Our new approach, PSNR 27.032.

Acknowledgement

The author would like to thank for Professeur M. Elad for his great help.

6. REFERENCES

- [1] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. on Image Processing*, to appear.
- [2] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. on Image Processing*, to appear.
- [3] T. Zeng and F. Malgouyres, "Using gabor dictionaries in a $tv - l^\infty$ model, for denoising," in *Proceedings of ICASSP, Toulouse, France, 2006*, vol. II, pp. 865–868.
- [4] S. Lintner and F. Malgouyres, "Solving a variational image restoration model which involves l^∞ constraints," *Inverse Problem*, vol. 20(3), pp. 815–831, June 2004.
- [5] J.-L. Starck, M. Elad, and D.L. Donoho, "Image decomposition via the combination of sparse representation and a variational approach," *IEEE Trans. on Image Processing*, vol. 14(10), pp. 1570–1582, 2005.