A Note on Boosting Algorithms for Image Denoising

Cory Falconer, C. Sean Bohun, and Mehran Ebrahimi(✉)

Faculty of Science, University of Ontario Institute of Technology,
2000 Simcoe Street North, Oshawa, ON L1H 7K4, Canada
{cory.falconer,sean.bohun,mehran.ebrahimi}@uoit.ca

Abstract. In recent years, non-local methods have been among most efficient tools to address the classical problem of image denoising. Recently, Romano et al. have proposed a novel algorithm aimed at "boosting" of a number of non-local denoising algorithms as a "black-box." In this manuscript, we consider this algorithm and derive an analytical expression corresponding to successive applications of their proposed "boosting scheme." Mathematically, we prove that such successive application does not always enhance the input image and is equivalent to a re-parameterization of the original "boosting" algorithm. We perform a set of computational experiments on test images to support this claim. Finally, we conclude that considering the blind application of such boosting methods as a general remedy for all denoising schemes is questionable.

Keywords: Image denoising · Nonlocal methods · Boosting

1 Introduction

Regardless of the acquisition process, digital images always contain undesired variation of pixel intensity, causing an image degradation in the form of noise, unavoidably reducing the quality of the image. Given a corrupted image $\mathbf{y}$, the goal of image denoising algorithms is to recover the original signal $\mathbf{x}$. Treating the noise to be additive, the image degradation takes the following form

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \tag{1}$$

in which we assume $\mathbf{n}$ is a zero-mean additive white noise that shares no dependence on $\mathbf{x}$. A wide variety of powerful algorithms have been proposed to address the classical problem of image denoising. This includes TV denoising [1,2], bilateral filtering [3], non-Local means (NLM) [4], and block matching 3D (BM3D) [5] to list a few. It is well known that patch based denoising algorithms are quite successful removing noise, though these methods also may remove relevant image content. Naturally, multiple methods have been proposed with intent to enhance the ability of these algorithms to remove noise while retaining image content.

F. Karray et al. (Eds.): ICIAR 2017, LNCS 10317, pp. 134–142, 2017.
DOI: 10.1007/978-3-319-59876-5_16

More specifically, we focus on algorithm enhancement schemes that utilize a sequential approach such as [6–8], classified by their usage of "post-processing."

The first class of methods entails recycling of noisy residuals with intent to import any "stolen" image content back into the estimation. The latter strengthens the approximation using a "cleaned" signal $\hat{\mathbf{x}}^k$, bypassing any unnecessary reintroduction of noise back into our image [7].

In [7], the authors demonstrate the effectiveness of their "boosting" method with a clear improvement in both SNR and visual clarity regardless of image structure or noise. In this paper, we raise the question if it is possible to apply this boosting algorithm sequentially, with a goal to iteratively improve the overall image quality at each iteration. We focus specifically on the Non-Local means algorithm for two reasons; (i) it is a fundamental algorithm utilizing a patch based approach and (ii) the NLM was validated as a method in which the "Boosting" scheme was applied successfully [7].

2 Non-local Means Denoising and Its Associated Boosting Algorithm

This section focusses on two separate aspects. In the first part, a review of the NLM denoising algorithm [4] is presented and in the second part, the concept of "boosting" in image denoising as introduced in [7] is reviewed.

2.1 Non-local Means Denoising

The NLM algorithm [4] exploits the innate redundancy of natural images, that replaces every pixel in the image with a weighted average of all pixels in the image. The weights are determined using neighborhood similarity of the pixels. In the process "similar" neighborhoods are assigned a high weight and dissimilar neighborhoods take low weights. Following the formation of the NLM algorithm, the neighborhood of size d is defined for $\forall\ x \in \Omega$, i.e., every point in the discrete image domain Ω, as

$$\mathcal{N}^d\{x\} = \{x + r|\ \|r\|_\infty \leq d\}\,.$$

The NLM algorithm denoted by $\mathbf{f}(\cdot)$ denoises the intensity of every pixel $x \in \Omega$ of an image $\mathbf{y}$ via

$$\mathbf{f}\left(\mathbf{y}(x)\right) = \frac{1}{C(x)} \sum_{y \in \Omega} w(x, y)\mathbf{y}(y), \tag{2}$$

where the weight $w(x, y)$ and normalization $C(x)$ are defined as

$$w(x, y) = \exp\left(-\frac{1}{h^2}\left\|\mathbf{y}(\mathcal{N}^d\{x\}) - \mathbf{y}(\mathcal{N}^d\{y\})\right\|_{2,a}^2\right), \quad C(x) = \sum_{y \in \Omega} w(x, y)$$

respectively. Note that $\|.\|_{2,a}$ denotes the Gaussian-weighted-semi norm, for any patch $\mathbf{P}$

$$\|\mathbf{P}\|_{2,a} = \|G \star \mathbf{P}\|_{2,a}$$

where G is a Gaussian kernel with a variance of a^2, and $\star$ is the convolution operator.

2.2 Boosting of Image Denoising Algorithms

As proposed in [7], the so called "**SOS** boosting" of an image denoising algorithm $\mathbf{f}$, to recover a superior approximation of the image $\mathbf{x}$ given a noisy realization $\mathbf{y}$, consists of iterations of the following three steps: **S**trengthen, **O**perate, and **S**ubtract. In detail for a single iteration:

1. **S**trengthen the signal by adding the previously denoised image and original noisy input;
2. **O**perate an image denoising algorithm on the new enhanced signal;
3. **S**ubtract the previously denoised image from the result of 2.

Treating the denoising operator $\mathbf{f}(\cdot)$ as a "black-box," the generalized boosting method proposed in [7] is formulated as

$$\hat{\mathbf{x}}^{k+1} = \tau \mathbf{f}(\mathbf{y} + \rho\hat{\mathbf{x}}^k) - (\tau\rho + \tau - 1)\hat{\mathbf{x}}^k \tag{3}$$

where $\hat{\mathbf{x}}^k$ denotes the approximate solution at iteration k, and $\hat{\mathbf{x}}^0$ is an initial approximation of $\mathbf{x}$. It is necessary to mention that parameter ρ determines the steady state solution and τ controls the rate of convergence. Since the boosting operation is sequential and fixed for all iterations, we pose the assumption $\hat{\mathbf{x}}^{k+1} = \hat{\mathbf{x}}^k = \hat{\mathbf{x}}^*$ to find a steady-state solution satisfying

$$\hat{\mathbf{x}}^* = \left((\rho + 1)\mathbf{I} - \rho\mathbf{f})\right)^{-1} \mathbf{f} \cdot \mathbf{y}. \tag{4}$$

Note that here we considered a linearization of the discretized denoising operator $\mathbf{f}$ as $\mathbf{f}(\mathbf{y}) = \mathbf{f} \cdot \mathbf{y}$, as in [7], and $\mathbf{I}$ denotes the identity matrix. As proposed by [7] the solution given by (4) is the "boosted" image that yields a superior approximation compared to the $\mathbf{f}(\mathbf{y})$, i.e., denoising of $\mathbf{y}$ using the denoising operator $\mathbf{f}$.

From the structure of (4) we define $\mathcal{B}(\mathbf{f})$ as

$$\mathcal{B}(\mathbf{f}) = \left((\rho + 1)\mathbf{I} - \rho\mathbf{f})\right)^{-1} \mathbf{f} \tag{5}$$

which is referred to as the "boosted" form of $\mathbf{f}$ and is envisioned as a "new" denoising algorithm compared to $\mathbf{f}$. The question is raised if it would be possible to further enhance the "boosting" by considering it as a new operator and successively applying the "boosting" process.

3 Iterative "Boosting" of "Boosting" Operators

Since we are trying to apply a boosting operation onto itself, we begin the derivation by substitution $\mathcal{B}(\mathbf{f})$ in for the denoising filter matrix $\mathbf{f}$ and repeat the process for a fixed parameter ρ. The following theorem summarizes the result of n consecutive boosting operations.

Theorem 1. *Given some denoising operator* $\mathbf{f}$, *and some fixed parameter* $\rho, \forall n \in \mathbb{N}$

$$\mathcal{B}^n(\mathbf{f}) = \left((\rho+1)^n(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)^{-1}\mathbf{f}. \tag{6}$$

Proof. By induction, taking $n = 1$,

$$\begin{aligned}\mathcal{B}^1(\mathbf{f}) &= \left((\rho+1)^1(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)^{-1}\mathbf{f} = \left((\rho+1)\mathbf{I}-\rho\mathbf{f}-\mathbf{f}+\mathbf{f}\right)^{-1}\mathbf{f}\\ &= \left((\rho+1)\mathbf{I}-\rho\mathbf{f}\right)^{-1}\mathbf{f}.\end{aligned}$$

Hence, $\mathcal{B}^1(\mathbf{f}) = \mathcal{B}(\mathbf{f}) = \left((\rho+1)\mathbf{I}-\rho\mathbf{f}\right)^{-1}\mathbf{f}$. Now, assuming Eq. (6) holds for some positive integer $n = k$, that is

$$\mathcal{B}^k(\mathbf{f}) = \left((\rho+1)^k(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)^{-1}\mathbf{f},$$

we find that

$$\begin{aligned}\mathcal{B}^{k+1}(\mathbf{f}) &= \mathcal{B}(\mathcal{B}^k(\mathbf{f}))\\ &= \left((1+\rho)\mathbf{I}-\rho\mathcal{B}^k(\mathbf{f})\right)^{-1}\mathcal{B}^k(\mathbf{f})\\ &= \left((1+\rho)\mathbf{I}-\rho\left((1+\rho)^k(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)^{-1}\mathbf{f}\right)^{-1}\left((1+\rho)^k(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)^{-1}\mathbf{f}\\ &= \left(\left((1+\rho)^k(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)\left((1+\rho)\mathbf{I}-\rho\left((1+\rho)^k(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)^{-1}\mathbf{f}\right)\right)^{-1}\mathbf{f}\\ &= \left((1+\rho)^{k+1}(\mathbf{I}-\mathbf{f})+(1+\rho)\mathbf{f}-\rho\mathbf{f}\right)^{-1}\mathbf{f}.\end{aligned}$$

This means that the theorem holds for $n = k + 1$, and this completes the induction. ■

As illustrated by Theorem 1, a sequential application of $\mathcal{B}$ onto denoising filter matrix $\mathbf{f}$, could be performed in one step. Such sequential application is nothing but a re-parametrization of the expression. That is $(\rho+1) \mapsto (\rho+1)^n$, or equivalently $\rho \mapsto (\rho+1)^n - 1$. We can also observe that if $\mathbf{I}-\mathbf{f}$ is invertible and $\rho > 0$, then

$$\lim_{n\to\infty}\mathcal{B}^n(\mathbf{f}) = \lim_{n\to\infty}\left((\rho+1)^n(\mathbf{I}-\mathbf{f})+\mathbf{f}\right)^{-1}\mathbf{f} = \lim_{n\to\infty}((\rho+1)^n(\mathbf{I}-\mathbf{f}))^{-1}\mathbf{f} = 0.$$

4 Experiments

In this section we investigate the effect of boosting of the NLM operator as well as its consecutive boosting on test images.

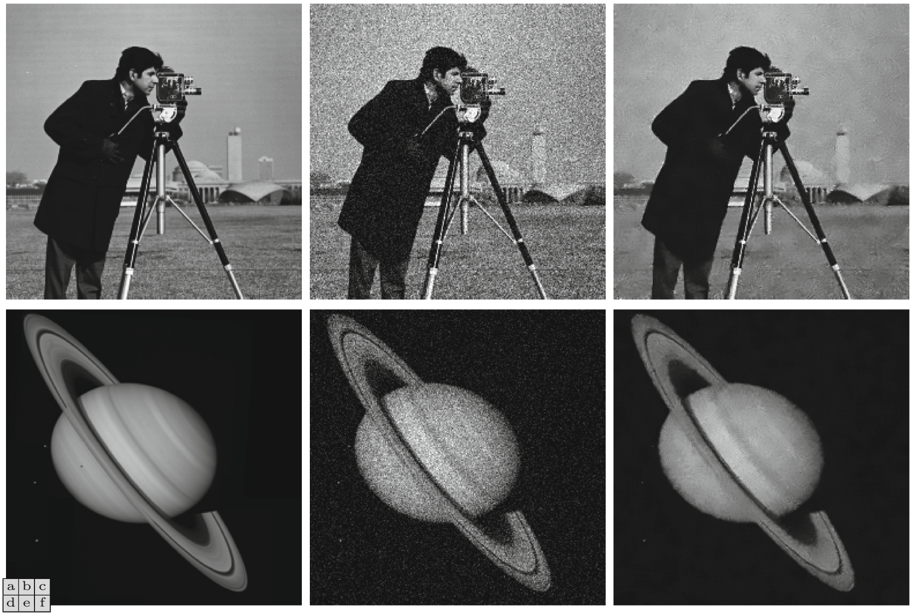

Fig. 1. (a) Ground truth $\mathbf{x}$, (b) Corrupted with AWGN ($\sigma = 0.01$), $\mathbf{y}$, (c) NLM restored $\mathbf{f} \cdot \mathbf{y}$, (d) Ground truth $\mathbf{x}$, (e) Corrupted with AWGN ($\sigma = 0.01$), $\mathbf{y}$, (f) NLM restored $\mathbf{f} \cdot \mathbf{y}$

We implemented a sparse representation of the NLM algorithm as the denoising operator $\mathbf{f}$, with a search window of size 5×5 and smoothing parameter $h = 10\sigma$. Note that σ is the standard deviation of the additive Gaussian white noise corrupting the ideal image $\mathbf{x}$.

As it can be seen in Fig. 1 after an application of the NLM on noisy Cameraman and Saturn images with Additive White Gaussian Noise (AWGN) of $\sigma = 0.01$, an increase in overall quality is achieved. To quantitatively measure the performance at each application of the boosting $\mathcal{B}^n(\mathbf{f})$, we measure the SNR (signal-to-noise ratio) as a function of boosting iteration count n. We also define $\mathcal{B}^0(\mathbf{f}) = \mathbf{f}$ i.e., the NLM restoration operator.

Curves relating to the computed SNR of $\mathcal{B}^n(\mathbf{f}) \cdot \mathbf{y}$ are given in Figs. 2 and 3 respectively for Cameraman and Saturn. In each figure two different values of noise standard deviation, namely $\sigma = 0.01$ and $\sigma = 0.05$, are considered. The parameter ρ is varied in the range of 0.1 to 1.5 for these curves.

It is interesting to note that for both images, when the noise standard deviation is $\sigma = 0.05$, no boosting can outperform the original NLM algorithm. This fact can be verified by considering the family of curves in Figs. 2(a) and 3(a) that are all decreasing.

For the smaller noise standard deviation $\sigma = 0.01$, the peak of the curves is observed for Figs. 2(b) and 3(b). A zoomed-in version of both images for $\sigma = 0.01$ can be seen on Figs. 2(c) and 3(c). It can be observed that $\rho = 1.3$ and

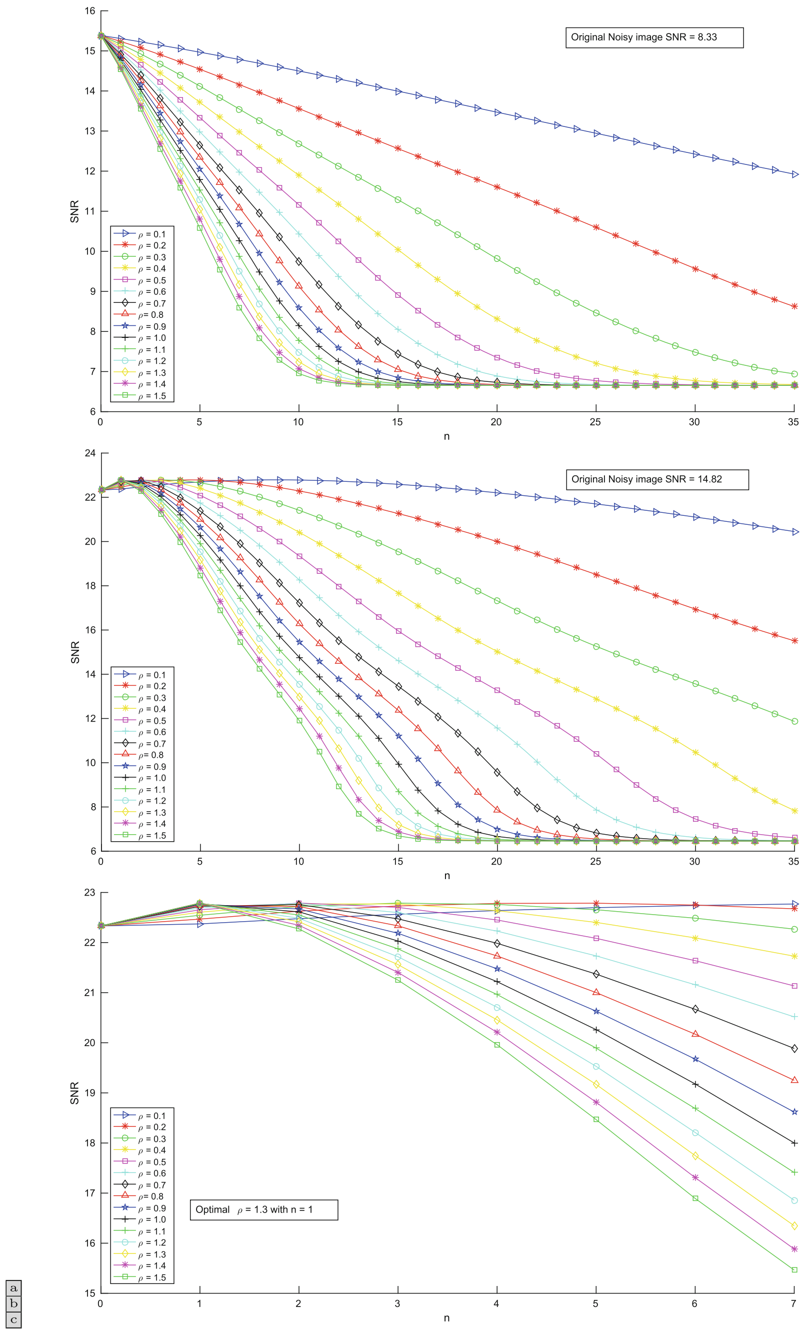

Fig. 2. Sequential boosting of NLM algorithm under operator $\mathcal{B}^n(\mathbf{f})$ applied on noisy Cameraman (a) AWGN ($\sigma = 0.05$), (b) AWGN ($\sigma = 0.01$), (c) Zoomed AWGN ($\sigma = 0.01$) and corresponding SNR

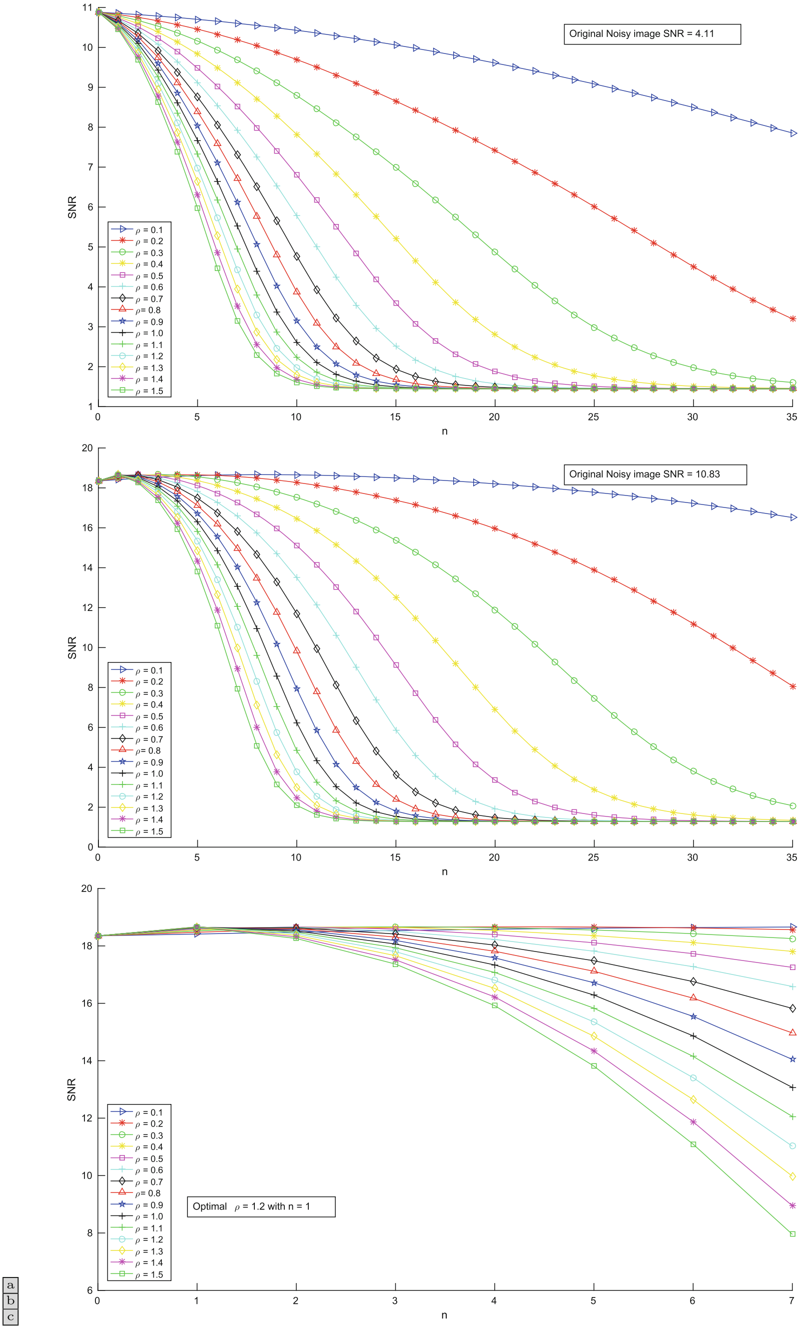

Fig. 3. Sequential boosting of NLM algorithm under operator $\mathcal{B}^n(\mathbf{f})$ applied to noisy Saturn (a) AWGN ($\sigma = 0.05$), (b) AWGN ($\sigma = 0.01$), (c) Zoomed AWGN ($\sigma = 0.01$) and corresponding SNR

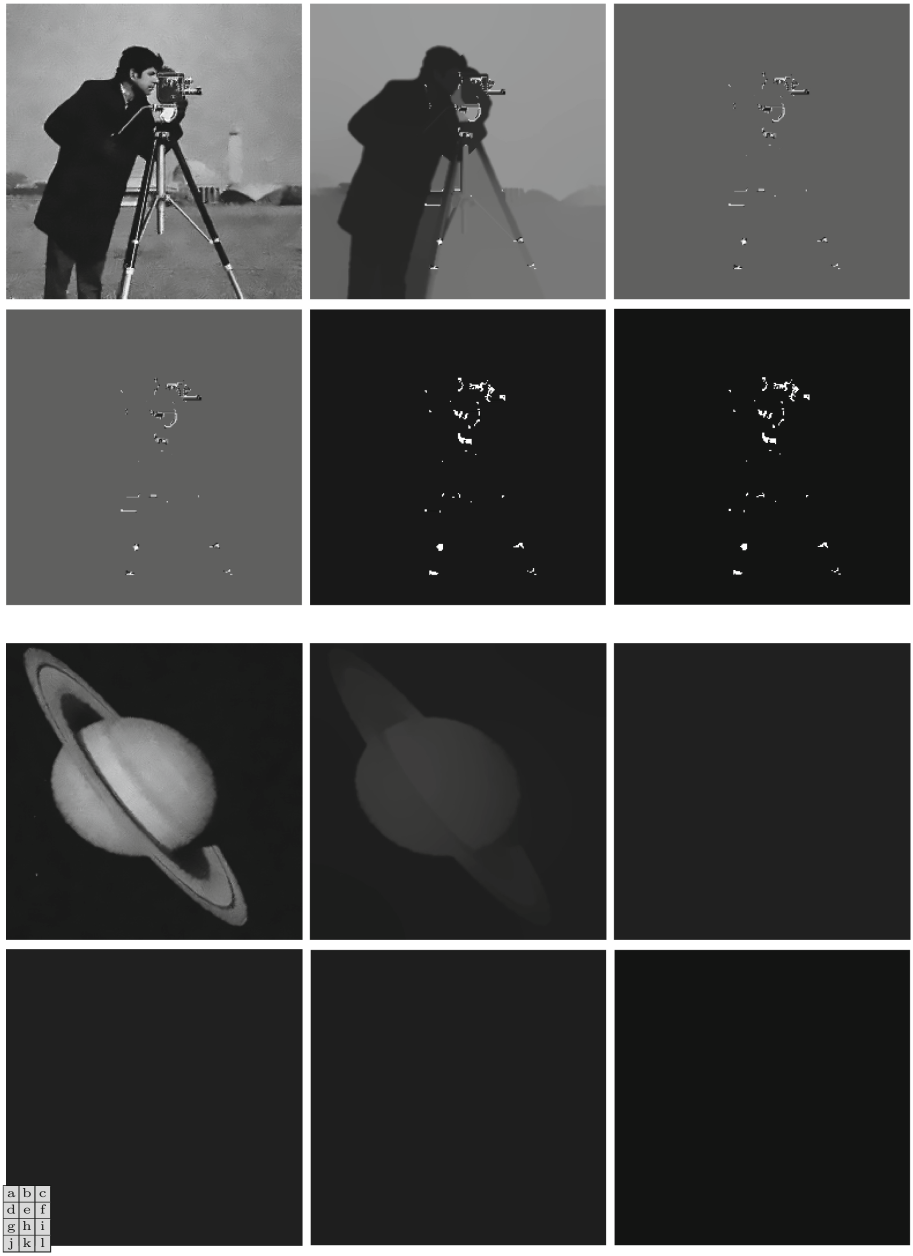

Fig. 4. Resulting $\mathcal{B}^n(\mathbf{f}) \cdot \mathbf{y}$ on noisy image $\mathbf{y}$ with AWGN ($\sigma = 0.01$) for (a) $n = 1$, (b) $n = 10$, (c) $n = 20$, (d) $n = 40$, (e) $n = 50$, (f) $n = 60$, corresponding to the optimal $\rho = 1.3$, (g) $n = 1$, (h) $n = 10$, (i) $n = 20$, (j) $n = 40$, (k) $n = 50$, (l) $n = 60$, corresponding to the optimal $\rho = 1.2$

$\rho = 1.2$ for $n = 1$ correspond to the peak of SNR corresponding the two images of Cameraman and Saturn for the noise standard deviation of $\sigma = 0.01$. Finally, Fig. 4 shows the resulting images $\mathcal{B}^n(\mathbf{f}) \cdot \mathbf{y}$ for a range of values of n fixing the corresponding peak ρ values of $\rho = 1.3$ and $\rho = 1.2$.

5 Discussion and Conclusions

Based on our experiments, choosing an optimal value of ρ can be a challenging task for the SOS boosting algorithms. We conclude that for some higher noise levels, SOS boosting does not improve the NLM algorithm, regardless of the choice of ρ. Assuming SOS boosting as a new denoising operator, we asked the question whether we can improve its performance by its iterative applications. We proved that iterative applications of the SOS boosting is equivalent to a re-parametrization of ρ. In addition, we showed that iterative applications of the SOS boosting for a fixed positive ρ converges to a zero image. This proves that considering the blind application of the SOS boosting as a general remedy for all denoising schemes is incorrect. We believe more research is required to further study the domain of applicability of the proposed boosting operator.

Acknowledgments. This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

1. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D: Nonlinear Phenom. **60**(1–4), 259–268 (1992)
2. Selesnick, I.: Total variation denoising (an MM algorithm). NYU Polytechnic School of Engineering Lecture Notes (2012)
3. Zhang, M., Gunturk, B.K.: Multiresolution bilateral filtering for image denoising. IEEE Trans. Image Process. **17**(12), 2324–2333 (2008)
4. Buades, A., Coll, B., Morel, J.-M.: A non-local algorithm for image denoising. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, vol. 2, pp. 60–65. IEEE (2005)
5. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007)
6. Charest, M.R., Elad, M., Milanfar, P.: A general iterative regularization framework for image denoising. In: 2006 40th Annual Conference on Information Sciences and Systems, pp. 452–457. IEEE (2006)
7. Romano, Y., Elad, M.: Boosting of image denoising algorithms. SIAM J. Imaging Sci. **8**(2), 1187–1219 (2015)
8. Talebi, H., Zhu, X., Milanfar, P.: How to SAIF-ly boost denoising performance. IEEE Trans. Image Process. **22**(4), 1470–1485 (2013)