PATCH-DISAGREEMENT AS A WAY TO IMPROVE K-SVD DENOISING

Yaniv Romano

Department of Electrical Engineering
Technion, Haifa 32000, Israel
yromano@tx.technion.ac.il

Michael Elad

Department of Computer Science
Technion, Haifa 32000, Israel
elad@cs.technion.ac.il

ABSTRACT

In this paper we propose a way to improve the K-SVD image denoising algorithm. The suggested method aims to reduce the gap that exists between the local processing (sparse-coding of overlapping patches) and the global image recovery (obtained by averaging the overlapping patches). Inspired by game-theory ideas, we define a disagreement-patch as the difference between the intermediate locally denoised patch and its corresponding part in the final outcome. Our algorithm iterates the denoising process several times, applied on modified patches. Those are obtained by subtracting the disagreement-patches from their corresponding input noisy ones, thus pushing the overlapping patches towards an agreement. Experimental results demonstrate the improvement this algorithm leads to.

Index Terms— Image restoration, denoising, disagreement, sparse representation, K-SVD.

1. INTRODUCTION

Denoising is a central and long studied problem in image processing. Consider a degradation model of the form

$$\mathbf{y} = \mathbf{x} + \mathbf{v}, \tag{1}$$

where $\mathbf{y}$ is a noisy image (measurement), $\mathbf{x}$ is a clean image, and $\mathbf{v}$ is a zero-mean additive Gaussian noise (uncorrelated to $\mathbf{x}$), all of size $r \times c$. The denoising process seeks for an approximation $\hat{\mathbf{x}}$ of the unknown signal $\mathbf{x}$. Note that $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{v}$ are held as column vectors after lexicographic ordering.

Patch processing has become popular in recent years. Many state-of-the-art denoising algorithms, which are build upon different image models, are essentially patch-based, e.g., the NLM [1], K-SVD [2], BM3D [3], LSSC [4], and others [5–7]. In this paper we give a special attention to the K-SVD image denoising [2], which restores the image using an adaptive sparsity model imposed on its patches [8, 9]. This model has been shown to be very effective in predicting the underlying signal. It assumes that each image patch can be represented as a sparse linear combination of basis elements (called ”atoms”) from a dictionary.

The gap between the local processing and the global need for a whole restored image is a major disadvantage of patch-processing in general, and the K-SVD denoising (and similar algorithms) in particular. More specifically, the K-SVD cleans the input image by breaking it into overlapping patches, restoring each patch using an adaptive sparsity model (local processing), and then reconstructing the full image by averaging the overlapping patches (the global need). As such, the denoising process is somewhat lacking – it estimates the clean patches independently while disregarding their inter-relations, overlooking the fact that these patches are sharing the very same pixels on the overlaps.

Motivated by this observation, series of recent papers came to close or at least narrow this local-global gap. Here we shall concentrate on several such techniques, designed specifically for the K-SVD. Our earlier work [10] suggests an iterative algorithm for improving the K-SVD [2]. The improvement is obtained by extracting the ”stolen” image content from the method-noise (the difference between the noisy image and its denoised version) image and then adding it back to the initial estimate. The approach in [10] suggests representing the stolen image content with the very same atoms used for the denoising of the patches. While this approach yields an improvement to the K-SVD, an evident shortcoming of it is the fact that it treats only part of the cause for deteriorated denoising performance. Another source of error is noise remaining in the denoised result, which [10] does not address at all.

From a different perspective, the work in [11] builds upon the EPLL framework [12], also treating the local-global gap. Sparsity-inspired EPLL encourages the patches of the global image (obtained by patch-averaging) to comply with the local sparsity prior. This is done by sequentially denoising the image with an assumption that the noise power is diminishing. The main challenge in this method is the need to evaluate the remaining (non-white) noise level in the filtered image, which is crucial for a successful deployment of the EPLL technique.

Naturally, a multi-scale framework can reduce the arti-

This research was supported by the European Research Council under EU’s 7th Framework Program, ERC Grant agreement no. 320649, and by the Intel Collaborative Research Institute for Computational Intelligence.

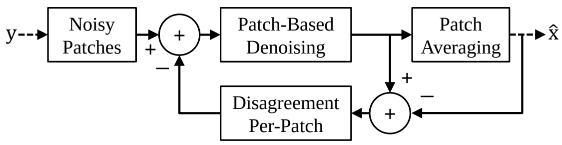


Fig. 1. A Block diagram of the proposed algorithm.

facts (especially for large smooth areas) that originates from the local patch-processing. The work in [13, 14] suggests a global MAP estimator for the denoised image in the wavelet domain. The core idea is to apply the K-SVD denoising per band, i.e., denoise the corresponding wavelet coefficients. Then, in a final stage, improved results are obtained by fusing the single-scale (original K-SVD) and multi-scale estimations.

Considering the similarity between image patches can also reduce the local-global gap. The LSSC [4] denoising combines the adaptive sparsity prior with the exploitation of non-local proximities. The latter assumes that each image patch may have similar patches within the image. Differently from the original K-SVD algorithm, which finds the representation of each patch independently, the LSSC finds the joint representation of a group of similar patches. As such, it considers not only the structure of the patches, but also the inter-relations between them. As such, this algorithm bares some resemblance to BM3D [3]. We should note, though, that despite the more effective local treatment, BM3D and LSSC both find themselves eventually averaging patches that exhibit disagreement, thus suffering from the same local-global problem, even if to a lesser extent.

In this paper we address the local-global gap in a different way from the above algorithms, and we focus our attention on the K-SVD image denoising [2]. Note that a similar approach can be deployed to other patch-based image denoising algorithms [1,3–5,11,12]. First, we define the "disagreement patch" – the difference between the local (intermediate) denoised result and the corresponding patch from the global (final) outcome. Due to the independent processing of the image patches, such disagreement naturally exists. Next, we suggest "sharing the disagreement" between the overlapping patches. As shown in Fig. 1, this is done by repeating the following procedure: (i) compute the disagreement per patch, (ii) subtract the disagreement from the noisy input patches, (iii) apply the K-SVD denoising on these patches, and (iv) reconstruct the global image. As a consequence, the proposed algorithm pushes the overlapping patches to share their local information (disagreements), thereby reducing the gap between the local processing and the global outcome.

This paper is organized as follows: In Section 2 we provide brief background on sparse representation and dictionary learning, concentrating on the K-SVD denoising. In Section 3 we introduce our novel "sharing the disagreement" algorithm. Experiments are detailed in Section 4, demonstrating the achieved improvement to the K-SVD denoising. Conclusions and future research directions are given in Section 5.

2. BACKGROUND

In this section we bring a brief discussion on sparse representations and the K-SVD image denoising algorithm [2, 8, 9].

2.1. Sparse-Land Modeling and K-SVD Denoising

Sparsity-inspired algorithms assume that a signal (or an image patch in our case), denoted by $x \in \mathbb{R}^n$, can be represented as a sparse linear combination of dictionary atoms. As a result, the signal can be expressed by $x = \mathbf{D}\alpha$, where the dictionary $\mathbf{D} \in \mathbb{R}^{n \times m}$ is composed of $m \geq n$ atoms (leading to redundancy), and the vector $\alpha \in \mathbb{R}^m$ contains a few non-zero elements.

More specifically, given a noisy measurement y, estimating the underlying signal $\hat{x}$ can be done by estimating $\hat{\alpha}$ – the projection of y onto the set of low dimensional subspaces that $\mathbf{D}$ spans (up to an error bound ϵ). Then, the denoised signal is obtained by $\hat{x} = \mathbf{D}\hat{\alpha}$. Put formally, given $\mathbf{D}$ and ϵ, we seek for $\hat{\alpha}$, the solution of

$$\hat{\alpha} = \min_{\alpha} \|\alpha\|_0 \text{ s.t. } \|\mathbf{D}\alpha - y\|_2^2 \leq \epsilon^2, \tag{2}$$

where $\|\alpha\|_0$ counts the non-zero entries in α. Equation (2) is an NP-hard problem, thus, in practice $\hat{\alpha}$ is approximated by pursuit algorithms (e.g. OMP [15] as done in the K-SVD [2]).

An adaptation of the dictionary to the input may result in a sparser representation, compared to a fixed dictionary (for the same ϵ), thus leading to a better noise reduction. Given several noisy signals $\{y_i\}_{i=1}^N$, and their representations $\{\tilde{\alpha}_i\}_{i=1}^N$, the K-SVD suggests updating the dictionary by solving

$$\left[\hat{\mathbf{D}}, \{\hat{\alpha}_i\}_{i=1}^N\right] = \min_{\mathbf{D}, \{\alpha_i\}_{i=1}^N} \sum_{i=1}^{N} \|\mathbf{D}\alpha_i - y_i\|_2^2 \quad \text{s.t. } Supp\{\alpha_i\} = Supp\{\tilde{\alpha}_i\}, \tag{3}$$

where the resulting $\hat{\mathbf{D}}$ is the updated dictionary, and $Supp\{\tilde{\alpha}_i\}$ are the indices of the non-zero elements in $\tilde{\alpha}_i$. As such, the K-SVD adapts the dictionary to the measurements by alternating between sparse-coding (i.e. solving Equation (2) for each example) and dictionary-update (i.e. solving Equation (3)). This process is known as "Dictionary Learning".

In practice, sparsity-inspired algorithms are limited in handling relatively small signals. As a consequence, the K-SVD image denoising breaks the noisy image into $\sqrt{n} \times \sqrt{n}$ overlapping patches, cleans these patches by iterating Equations (2) and (3), and then reconstructs the whole image by returning the denoised patches to their locations. As a final step, the K-SVD merges the denoised and noisy images by applying a weighted average. All this process approximates the solution of the following expression:

$$\left[\hat{\mathbf{x}}, \hat{\mathbf{D}}, \{\hat{\alpha}_i\}_{i=1}^N\right] = \min_{\mathbf{x},\mathbf{D},\{\alpha_i\}_{i=1}^N} \mu\|\mathbf{x}-\mathbf{y}\|_2^2 + \sum_{i=1}^N \|\alpha_i\|_0 \quad \text{s.t. } \forall i \ \|\mathbf{D}\alpha_i - \mathbf{R}_i\mathbf{x}\|_2^2 \le \epsilon^2, \tag{4}$$

where N denotes the number of image patches, and the matrix $\mathbf{R}_i \in \mathbb{R}^{n\times rc}$ extracts the i^{th} patch from the global image.

3. THE PROPOSED ALGORITHM

As mentioned in section 2, the K-SVD denoising [2] (and similar algorithms) breaks the noisy image into overlapping patches, and cleans each patch independently. Since the global image is more than a collection of independent patches, this treatment is somewhat lacking. Furthermore, the final image is obtained by a simple patch averaging, without considering the shared estimations on the overlaps.

Motivated by the game-theory concepts in general, and the "consensus and sharing" problem [16] in particular, the proposed algorithm aims to narrow the local-global gap by encouraging the overlapping patches to influence each other. More specifically, the "consensus" problem involves the minimization of a single global variable (the denoised image), where the objective and constraint terms split into N parts (the recovery of the overlapping patches). In addition, the closely related "sharing" problem involves the adjustment of local variables to minimize their own (local) cost function, as well as the shared (global) objective. Following these ideas, the proposed iterative method drives the overlapping patches towards an agreement by sharing the neighbors disagreements, thus called "sharing the disagreement".

In the context of the K-SVD denoising, we define the "disagreement-patch" as

$$\mathbf{q}_i^k = \mathbf{D}^k\alpha_i^k - \mathbf{R}_i\hat{\mathbf{x}}^k, \tag{5}$$

where $\mathbf{D}^k\alpha_i^k$ is the i^{th} locally denoised result, and $\mathbf{R}_i\hat{\mathbf{x}}^k$ is the corresponding part from the global estimate, both obtained at the k^{th} iteration. As a reminder, the sparse-coding step does not consider the relations between the patches, thus the energy of $\mathbf{q}_i$ is not negligible in general. Next, we encourage the overlapping patches to collaborate by modifying the input patches to the next denoising stage. This is done by subtracting $\mathbf{q}_i$ from $\mathbf{R}_i\mathbf{y}$, and then applying the K-SVD algorithm on these patches (including the image reconstruction). In practice, as detailed in Algorithm 1, we repeat this procedure several times.

Still following Algorithm 1, the new input patch can be expressed by

$$\begin{aligned} \mathbf{p}_i^k &= \mathbf{R}_i\mathbf{y} - \mathbf{q}_i^k \\ &= \mathbf{R}_i\mathbf{y} - (\mathbf{D}^k\alpha_i^k - \mathbf{R}_i\hat{\mathbf{x}}^k) \\ &= \mathbf{R}_i\hat{\mathbf{x}}^k + \mathrm{r}_i^k, \end{aligned} \tag{10}$$

Algorithm 1 : Sharing the disagreement.

Initialization:

1: Assign $k=0$, and $\mathbf{q}_i^0 = 0$.
2: $\mathbf{D}^0$ – an initial dictionary.

Repeat

1: *Sparse-Coding Step:* Using the OMP [15], solve

$$\left[\{\hat{\alpha}_i^{k+1}\}_{i=1}^N\right] = \min_{\{\alpha_i\}_{i=1}^N} \sum_{i=1}^N \|\alpha_i\|_0 \ \text{ s.t. } \forall i \ \|\mathbf{D}^k\alpha_i - (\mathbf{R}_i\mathbf{y} - \mathbf{q}_i^k)\|_2^2 \le \epsilon^2. \tag{6}$$

2: *Dictionary-Update Step:* Solve

$$[\mathbf{D}^{k+1}, \{\alpha_i^{k+1}\}_{i=1}^N] = \min_{\mathbf{D},\{\alpha_i\}_{i=1}^N} \sum_{i=1}^N \|\mathbf{D}\alpha_i - (\mathbf{R}_i\mathbf{y} - \mathbf{q}_i^k)\|_2^2 \quad \text{s.t. } Supp\{\alpha_i\} = Supp\{\hat{\alpha}_i^{k+1}\}. \tag{7}$$

In practice, given $\mathbf{D}^k$, we obtain $\mathbf{D}^{k+1}$ using the K-SVD.

3: *Image Reconstruction Step:* Solve

$$\hat{\mathbf{x}}^{k+1} = \min_{\mathbf{x}} \sum_i \|\mathbf{D}^{k+1}\alpha_i^{k+1} - \mathbf{R}_i\mathbf{x}\|_2^2 + \mu\|\mathbf{x}-\mathbf{y}\|_2^2. \tag{8}$$

The outcome of this term leads to a simple averaging of the denoised patches on the overlaps, followed by a weighted average with the noisy image.

4: *Disagreement-Update Step:* Compute

$$\mathbf{q}_i^{k+1} = \mathbf{D}^{k+1}\alpha_i^{k+1} - \mathbf{R}_i\hat{\mathbf{x}}^{k+1}, \tag{9}$$

and set $k \leftarrow k+1$.

Until

Maximum quality is obtained, else return to "Sparse-Coding Step".

Output

$\hat{\mathbf{x}}^k$ – the last result.

---

where $\mathrm{r}_i^k = \mathbf{R}_i\mathbf{y} - \mathbf{D}^k\alpha_i^k$ represents the i^{th} local method-noise patch, obtained at the k^{th} iteration. As can be inferred, our algorithm aims to recover a patch from the global estimation $\hat{\mathbf{x}}^k$, corrupted by the method-noise patch r_i^k. Therefore, "sharing the disagreement" reduces the local-global gap in a different way compared to the EPLL [11, 12], which iteratively cleans the previous estimation, without considering the signal that resides in the method-noise image. In addition, this approach is different from our earlier work [10], which cleans the method-noise image and does not consider the noise that remains in the denoised image. A unique property of the new algorithm, when compared to earlier work, is the fact that it harnesses intermediate patch-denoising results, that are inner

σ	$\hat{\sigma}$	Barbara		Boat		House		Fingerprint		Peppers		Couple		Average
		Orig	New	Orig	New	Orig	New	Orig	New	Orig	New	Orig	New	New
10	1.12σ	34.55	34.55	33.63	**33.70**	36.06	**36.09**	32.37	**32.40**	34.81	**34.83**	33.57	**33.69**	0.05
20	1.06σ	30.86	**31.03**	30.40	**30.62**	33.16	**33.37**	28.48	**28.65**	32.31	**32.42**	30.05	**30.31**	0.19
25	1.06σ	29.61	**29.88**	29.31	**29.58**	32.30	**32.62**	27.29	**27.51**	31.50	**31.61**	28.96	**29.29**	0.25
50	1.04σ	25.38	**26.12**	25.92	**26.36**	27.93	**28.69**	23.31	**23.98**	28.16	**28.58**	25.28	**25.80**	0.59
75	1.02σ	22.89	**23.53**	23.94	**24.45**	25.22	**25.96**	20.00	**21.49**	25.80	**26.36**	23.65	**24.10**	0.73
100	1.02σ	21.83	**21.99**	22.86	**23.16**	23.63	**24.29**	18.28	**19.55**	24.26	**24.76**	22.63	**22.88**	0.52

Table 1. Comparison between the denoising results [PSNR] of the original K-SVD algorithm [2] and its "sharing the disagreement" outcome (Algorithm 1). The best results per each image and noise level are highlighted.

to the K-SVD algorithm's architecture.

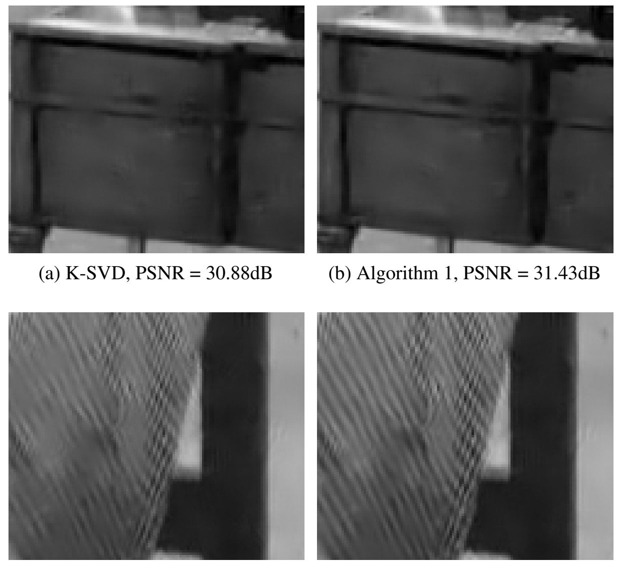
4. EXPERIMENTS

In this section, we present detailed results of the proposed method for several noise levels and test images (Barbara, Boat, Fingerprint, House, Peppers, and Couple). These images are corrupted by an additive zero-mean Gaussian noise with a standard-deviation σ. We evaluate the denoising performance using the Peak Signal to Noise Ratio (PSNR), defined as $20\log_{10}(\frac{255}{\sqrt{\text{MSE}}})$, where MSE is the Mean Squared Error between the original image and its denoised version.

The proposed "sharing the disagreement" algorithm does not involve any additional parameters over the original K-SVD denoising [2] parameters. However, we found that using $\hat{\sigma}$, a higher noise energy than σ, leads to better performance. This originates from the noise energy of $\mathbf{R}_i\hat{\mathbf{x}}^k + \mathrm{r}_i$ (see Equation (10)), which might be larger than σ. As such, we have tuned the parameter $\hat{\sigma}$ per each noise level σ. The denoising results of Table 1 are obtained by applying Algorithm 1 for 30 iterations, where each iteration includes 2 sparse-coding and dictionary-update steps. In addition, the initial dictionary is obtained by applying 20 iterations of the K-SVD algorithm (leading to what is referred to in the table as 'Orig' results).

Following Table 1, in terms of PSNR, "sharing the disagreement" algorithm improves the original K-SVD denoising [2] for all images and noise levels (especially for large σ). Visually, according to Fig. 2, the proposed method improves the recovery of edges, as shown in (a,b), and texture areas, as demonstrated in (c,d). In addition, the visual improvements are consistent with the PSNR increase.

We should note that the gain achieved by the proposed method is of similar extent to the other boosting techniques mentioned [4, 10–14]. Moreover, as a future study, it is interesting to combine our method with the above techniques.



(a) K-SVD, PSNR = 30.88dB (b) Algorithm 1, PSNR = 31.43dB

(c) K-SVD, PSNR = 28.56dB (d) Algorithm 1, PSNR = 29.59dB

Fig. 2. Visual and PSNR comparisons between the K-SVD denoising and "sharing the disagreement" outcomes for cropped regions from the images (a,b) Couple with $\sigma = 20$ and (c,d) Barbara with $\sigma = 25$.

5. CONCLUSIONS AND FUTURE DIRECTIONS

We presented a new algorithm termed "sharing the disagreement", which aims to improve the restoration performance of the K-SVD denoising [2]. The core idea behind our approach originates from game theory and consensus methods. When overlapping patches are treated separately, they naturally disagree on the final result. We take this disagreement and add it back to each of the patches in such a way that they are encouraged to get closer to an agreement in a consequent denoising stage. Our method does not require any additional parameters or internal modifications to the original K-SVD algorithm.

In a follow-up paper [17] we generalize the patch-disagreement formulation discussed here, and show that it leads to an overall algorithm that can treat the given denoising method as a "black-box". The work in [17] also studies the convergence properties of the resulting algorithm, and deploys it to a series of leading denoising methods, showing for each a gain in performance. As a future work, we hope to propose a similar scheme for improving other restoration tasks (e.g. [18–20]).

6. REFERENCES

[1] A. Buades, B. Coll, and J.-M. Morel, “A review of image denoising algorithms, with a new one,” *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[2] M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” *IEEE Tran. on Image Proc.*, vol. 15, no. 12, pp. 3736–3745, 2006.

[3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Trans. on Image Proc.*, vol. 16, no. 8, pp. 2080–2095, 2007.

[4] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, “Non-local sparse models for image restoration,” in *Int. Conf. on Computer Vision*. IEEE, 2009, pp. 2272–2279.

[5] G. Yu, G. Sapiro, and S. Mallat, “Solving inverse problems with piecewise linear estimators: from gaussian mixture models to structured sparsity,” *IEEE Trans. on Image Proc.*, vol. 21, no. 5, pp. 2481–2499, 2012.

[6] H. C. Burger, C. J. Schuler, and S. Harmeling, “Image denoising: Can plain neural networks compete with BM3D?,” in *IEEE Conf. on Computer Vision and Pattern Recog.*, 2012, pp. 2392–2399.

[7] I. Ram, M. Elad, and I. Cohen, “Image processing using smooth ordering of its patches,” *IEEE Trans. on Image Proc.*, vol. 22, no. 7, pp. 2764–2774, 2013.

[8] A. M. Bruckstein, D. L. Donoho, and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” *SIAM review*, vol. 51, no. 1, pp. 34–81, 2009.

[9] M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*, Springer, 2010.

[10] Y. Romano and M. Elad, “Improving K-SVD denoising by post-processing its method-noise,” in *IEEE Int. Conf. on Image Proc.*, Sept 2013, pp. 435–439.

[11] J. Sulam and M. Elad, “Expected patch log likelihood with a sparse prior,” in *submitted to Energy-Minimization-Methods workshop, Hong-Kong*, January 13-16 2015.

[12] D. Zoran and Y. Weiss, “From learning models of natural image patches to whole image restoration,” in *IEEE Int. Conf. on Computer Vision*, 2011, pp. 479–486.

[13] J. Sulam, O. Boaz, and M. Elad, “Image denoising through multi-scale learnt dictionaries,” in *IEEE Int. Conf. on Image Proc.* IEEE, 2014.

[14] B. Ophir, M. Lustig, and M. Elad, “Multi-scale dictionary learning using wavelets,” *Selected Topics in Signal Proc.*, vol. 5, no. 5, pp. 1014–1024, 2011.

[15] Y. C. Pati, R. Rezaiifar, and PS Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” in *Conf. Record of The 27 Asilomar Conf. on Signals, Systems and Computers*. IEEE, 1993, pp. 40–44.

[16] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[17] Y. Romano and M. Elad, “SOS boosting of image denoising algorithms,” *submitted to SIAM Journal on Imaging Sciences*, 2015.

[18] R. Zeyde, M. Elad, and M. Protter, “On single image scale-up using sparse-representations,” in *Curves and Surfaces*, pp. 711–730. Springer, 2012.

[19] R. Giryes and M. Elad, “Sparsity based Poisson denoising with dictionary learning,” *to appear in IEEE Trans. on Image Proc.*, 2014.

[20] Y. Romano, M. Protter, and M. Elad, “Single image interpolation via adaptive nonlocal sparsity-based modeling,” *IEEE Trans. on Image Proc.*, vol. 23, no. 7, pp. 3085–3098, 2014.