# BM3D Image Denoising with Shape-Adaptive Principal Component Analysis

*Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian*

Department of Signal Processing, Tampere University of Technology
P.O. Box 553, 33101 Tampere, Finland. E-mail: firstname.lastname@tut.fi

***Abstract*— We propose an image denoising method that exploits nonlocal image modeling, principal component analysis (PCA), and local shape-adaptive anisotropic estimation. The nonlocal modeling is exploited by grouping similar image patches in 3-D groups. The denoising is performed by shrinkage of the spectrum of a 3-D transform applied on such groups. The effectiveness of the shrinkage depends on the ability of the transform to sparsely represent the true-image data, thus separating it from the noise. We propose to improve the sparsity in two aspects. First, we employ image patches (neighborhoods) which can have data-adaptive shape. Second, we propose PCA on these adaptive-shape neighborhoods as part of the employed 3-D transform. The PCA bases are obtained by eigenvalue decomposition of empirical second-moment matrices that are estimated from groups of similar adaptive-shape neighborhoods. We show that the proposed method is competitive and outperforms some of the current best denoising methods, especially in preserving image details and introducing very few artifacts.**

## I. INTRODUCTION

Image denoising is a vivid research subject in signal processing because of its fundamental role in many applications. Some of the most recent and successful advances are based on: Gaussian scale mixtures (GSM) modeling in overcomplete multiscale transform domain [9], [15], [8]; learned dictionaries of atoms to filter small square neighborhoods [6]; steering kernel regression [16] (also combined with learned dictionaries [2]); shape-adaptive DCT (SA-DCT) on neighborhoods whose shapes are adaptive to the image structures [7]; nonlocal filtering based on the assumption that there exist mutually similar blocks in natural images [1], [12], [3], [4]. Among these different strategies, the nonlocal filtering, which originates from the work of Buades et al. [1], is the one that demonstrates the biggest potential. In particular, the transform-based BM3D filter [3] can be considered state-of-the-art in image denoising [13].

The BM3D filter exploits a specific nonlocal image modeling [5] through a procedure termed *grouping and collaborative filtering*. Grouping finds mutually similar 2-D image blocks and stacks them together in 3-D arrays. Collaborative filtering produces individual estimates of all grouped blocks by filtering them jointly, through transform-domain shrinkage of the 3-D arrays (groups). In doing so, BM3D relies both on *nonlocal* and *local* characteristics of natural images, namely the abundance of mutually similar patches and that the image data is locally highly correlated. If these characteristics are verified, the group enjoys correlation in all three dimensions and a sparse representation of the true signal is obtained by applying a decorrelating 3-D transform on the group. The effectiveness of the subsequent shrinkage depends on the sparsity of the true signal; i.e. the true signal can be better separated from the noise when its energy is compactly represented in the 3-D transform domain. We have shown [3] that even if square image patches and a fixed 3-D transform are used, the denoising performance is very high and the obtained MSE results are still beyond the capabilities of most of the more recent and advanced algorithms. However, square image blocks containing fine image details, singularities, or sharp and curved edges are examples where a non-adaptive transform is not able to deliver a sparse representation. Thus, for these blocks, the BM3D filter may introduce certain artifacts and the denoising is not very effective. Unfortunately, these are often the very parts of the image where the visual attention is mainly focused.

In order to further increase the sparsity of the true signal in the 3-D spectra, in [4] we proposed a generalization of the BM3D filter, which uses *grouping* of mutually similar *adaptive-shape neighborhoods*. The employed 3-D transform there is a separable composition of the (2-D) SA-DCT and a 1-D orthonormal transform. This transform is applied on 3-D groups that are generalized cylinders with adaptive-shape cross sections (as opposed to square prisms in BM3D). The adaptive-shape neighborhoods enable local adaptivity to image features so that the true signal in such a neighborhood is mostly homogeneous. Thus, the spatial correlation improves the sparsity as compared with the BM3D. However, even though the neighborhoods have adaptive shapes, the SA-DCT basis are fixed for any given shape; i.e. the basis elements do not adapt to the signal within the grouped neighborhoods.

In this paper, to enable adaptivity of the applied shape-adaptive transform basis to the input data, we propose principal component analysis (PCA) as part of the 3-D transform used for collaborative filtering. For a 3-D group of adaptive-shape image patches, we obtain a PCA basis by eigenvalue decomposition of an empirical second-moment matrix computed from these patches. As principal components (PC), we select only the eigenvectors whose corresponding eigenvalues are greater than a threshold that is proportional to noise variance. Hence, the overall 3-D transform is a separable composition of the PCA (applied on each image patch) and a fixed orthogonal 1-D transform in the third dimension.

In the sequel we present the developed denoising method and show that it is competitive and outperforms some of the current best denoising methods, particularly in preserving image details and producing very few artifacts.

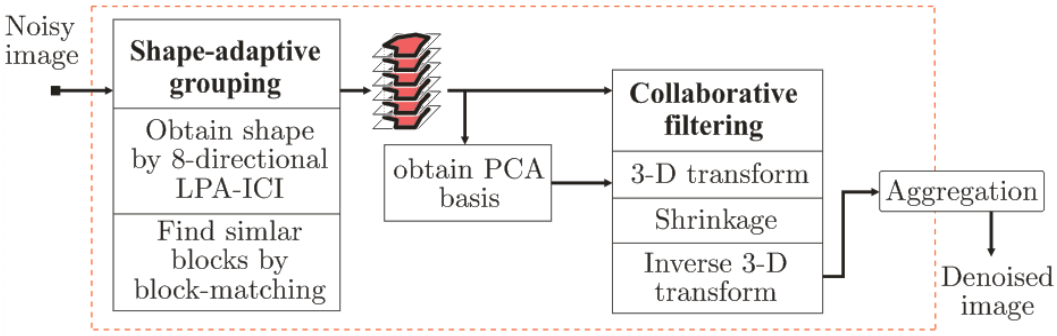



Fig. 1. Flowchart of the proposed BM3D-SAPCA image denoising method. Operations surrounded by dashed lines are repeated for each processed coordinate of the input image.

## II. BM3D-SAPCA ALGORITHM

### A. Algorithm outline

Following is an outline of the proposed algorithm, which we denominate BM3D-SAPCA, and which is illustrated in Figure 1. The input noisy image is assumed to be corrupted by an additive white Gaussian noise with variance $\sigma^2$ and zero mean.

- The input image is processed in raster scan where at each processed pixel the following operations are performed:

1) Obtain adaptive-shape neighborhood centered at the current pixel using the 8-directional LPA-ICI exactly as in [7], [4]. The neighborhood is enclosed within a fixed-size and non-adaptive square block, which we term *reference block*. We denote the number of pixels in the neighborhood by $N_{el}$.
2) Find blocks that are similar to the reference one using block-matching and extract an adaptive-shape neighborhood from each of these matched blocks using the shape obtained in Step 1. The number of matched blocks is denoted by $N_{gr}$.
3) Determine the transform to be applied on the adaptive-shape neighborhoods. We have two cases, depending on whether $\frac{N_{gr}}{N_{el}}$ is larger or smaller than a fixed threshold $\tau$.
   a) If $\frac{N_{gr}}{N_{el}} \geq \tau$, we consider that we have found a sufficient number of mutually similar neighborhoods to reliably estimate a second-moment matrix. The eigenvectors of this matrix constitute the shape-adaptive PCA basis. Subsequently, we retain only those eigenvectors whose corresponding eigenvalues are greater than a predefined threshold, thus obtaining a *trimmed* shape-adaptive PCA transform.
   b) If $\frac{N_{gr}}{N_{el}} < \tau$, we deem there are not enough similar neighborhoods to use as training data and we resort to the fixed (i.e. non data-adaptive) SA-DCT, exactly as in [4].
4) Form a 3-D array (called *group*) by stacking together the $\min(N_{gr}, N_2)$ adaptive-shape neighborhoods with highest similarity to the reference one, where $N_2$ is a fixed parameter that restricts the number of filtered neighborhoods.
5) Apply the transform obtained in Step 3 on each of the grouped adaptive-shape neighborhoods. Subsequently, apply a 1-D orthogonal transform (e.g., Haar wavelet decomposition) along the third dimension of the 3-D group.
6) Perform shrinkage (hard-thresholding or empirical Wiener filtering) on the 3-D spectrum.
7) Invert the 3-D transform from Step 5 to obtain estimates for all of the grouped adaptive-shape neighborhoods.
8) Return the obtained estimates to their original locations using weighted averaging in case of overlapping.

### B. Trimmed PCA

Since the main contribution of the proposed method is the application of the shape- and data-adaptive PCA transform on groups of adaptive-shape neighborhoods, we explain in detail what is done in Step 3a, while we refer the reader to our previous works [3], [7], [4] for details on the other steps of the algorithm. The input for Step 3a is a group of $N_{gr}$ adaptive-shape neighborhoods that are found to be mutually similar. We represent each of these 2-D neighborhoods as a 1-D column vector $\overrightarrow{v}_i$ of length $N_{el}$, $i = 1, \ldots, N_{gr}$. An $N_{el} \times N_{el}$ sample second-moment matrix is then computed by matrix multiplication,

$$C = \left[ \overrightarrow{v}_1 \;\; \overrightarrow{v}_2 \;\; \ldots \;\; \overrightarrow{v}_{N_{gr}} \right] \left[ \overrightarrow{v}_1 \;\; \overrightarrow{v}_2 \;\; \ldots \;\; \overrightarrow{v}_{N_{gr}} \right]^T, \quad (1)$$

and subsequently its eigenvalue decomposition yields

$$U^T C U = S = \operatorname{diag}\left(s_1, s_2, \ldots, s_{N_{el}}\right),$$

where $U$ is orthonormal matrix and $S$ is a diagonal matrix containing eigenvalues ordered by magnitude, $s_1 > s_2 > \cdots > s_{N_{el}}$. Finally, the PCs used for the decomposition of the adaptive–shape neighborhoods are the first $N_{trim}$ columns of $U$, where $N_{trim}$ is the number of eigenvalues greater than $\lambda\sigma^2$, $\lambda$ being a fixed threshold. An illustration of trimmed PCs is given in Figure 2.

### C. Iterative refinement

Similarly to [3], [7], [4], the above algorithm can be applied in more than one iteration. In this work, we propose a three-iteration approach. In the first iteration, the shrinkage is performed by hard-thresholding and the search for similar blocks is performed in the noisy image, since this is the only

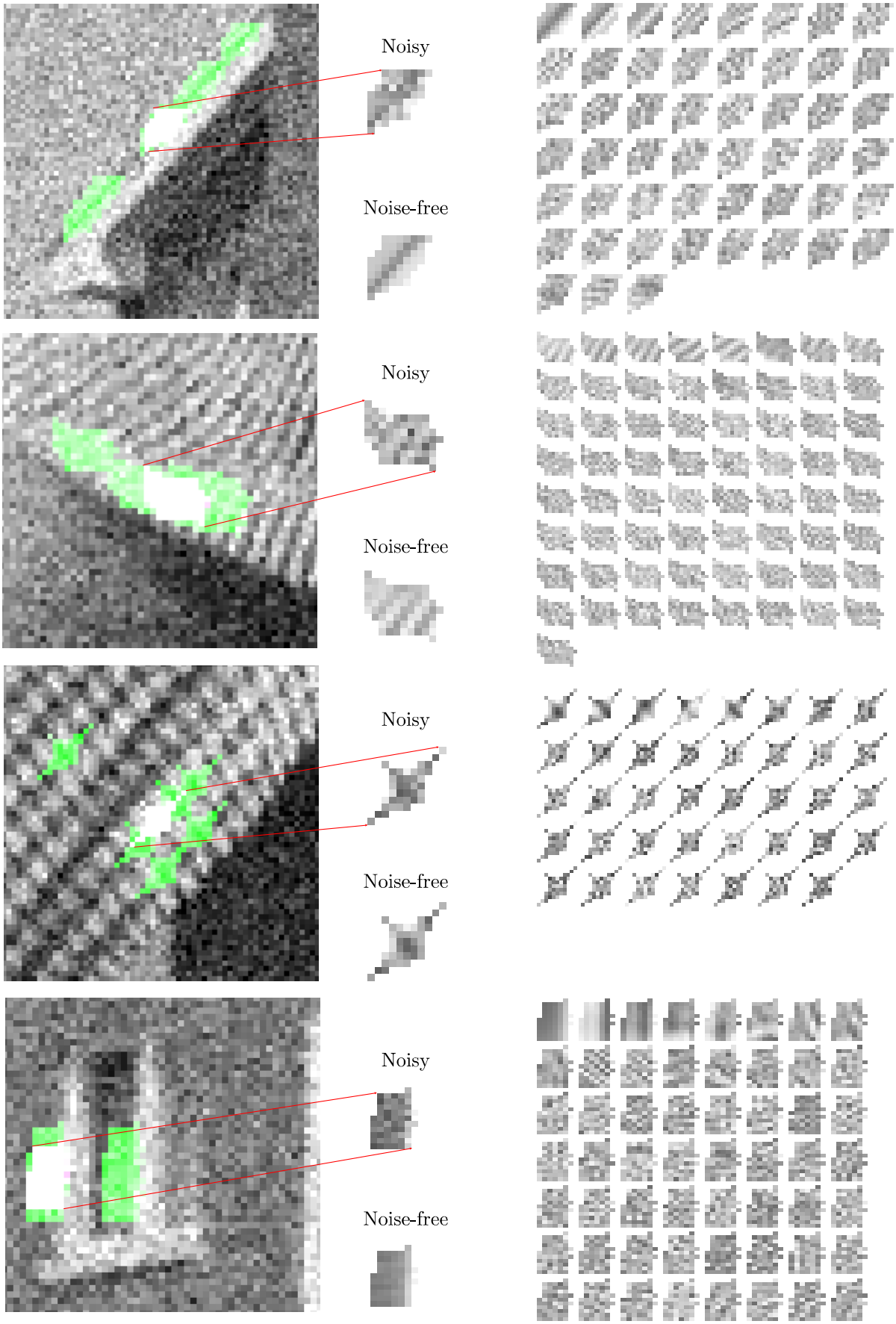



Fig. 2. Illustration of the PCs (shown on the right side) after trimming for four particular adaptive-shape neighborhoods. The green overlay is used to show the found similar neighborhoods used to form a 3-D group. The PCs are listed in decreasing magnitude of their corresponding eigenvalues. One can observe that the first few PCs have the strongest similarity with the noise-free signal in the neighborhood.

TABLE I

PSNR (dB) AND MSSIM RESULTS OF THE PROPOSED BM3D-SAPCA IMAGE DENOISING ALGORITHM. (THE MSSIM RESULTS ARE IN THE LOWER ROW OF EACH TABLE CELL.)

| $\sigma$/ PSNR | *C.man* $256^2$ | *House* $256^2$ | *Peppers* $256^2$ | *Montage* $256^2$ | *Lena* $512^2$ | *Barbara* $512^2$ | *Boats* $512^2$ | *Man* $512^2$ | *Couple* $512^2$ | *Hill* $512^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 5/ 34.16 | 38.53 | 39.99 | 38.30 | 41.40 | 38.82 | 38.34 | 37.47 | 38.03 | 37.63 | 37.30 |
| | 0.9622 | 0.9589 | 0.9563 | 0.9818 | 0.9454 | 0.9655 | 0.9434 | 0.9565 | 0.9529 | 0.9455 |
| 15/ 24.61 | 32.36 | 35.15 | 32.94 | 35.72 | 34.42 | 33.30 | 32.29 | 32.20 | 32.24 | 32.05 |
| | 0.9089 | 0.8987 | 0.9063 | 0.9555 | 0.8976 | 0.9253 | 0.8568 | 0.8740 | 0.8784 | 0.8455 |
| 25/ 20.18 | 29.81 | 32.95 | 30.43 | 32.96 | 32.22 | 30.99 | 30.03 | 29.81 | 29.82 | 29.96 |
| | 0.8635 | 0.8602 | 0.8685 | 0.9291 | 0.8643 | 0.8938 | 0.8037 | 0.8111 | 0.8214 | 0.7788 |
| 35/ 17.25 | 28.17 | 31.37 | 28.74 | 30.88 | 30.72 | 29.35 | 28.51 | 28.29 | 28.23 | 28.62 |
| | 0.8269 | 0.8374 | 0.8357 | 0.9030 | 0.8367 | 0.8605 | 0.7613 | 0.7633 | 0.7730 | 0.7299 |

available input. In the following two iterations, we utilize the estimate image from each previous iteration in the following manner:

- the search for similar blocks is done in the estimate image,
- the sample second-moment matrix (1) is computed from neighborhoods extracted from the estimate image,
- for the third iteration, the shrinkage is performed by empirical Wiener filtering.

The improvement contributed by the second and by the third iteration can be justified as follows. Because noise has already been attenuated in the estimate images, both the block-matching and the estimation of the second moments are more accurate when these operations are carried out on an estimate image. It results in sparser 3-D group spectra, as both the mutual similarity within each group and decorrelation due to PCA are enhanced. In addition, the empirical Wiener filtering is more effective than hard-thresholding when the estimate image from the second iteration is used for providing a reliable estimate of the power spectrum of the 3-D groups.

## III. RESULTS

We present results obtained with the proposed three-iteration BM3D-SAPCA. We use different values for $\tau$ and $\lambda$ in each iteration: $\tau = 1.3$, $\lambda = 49$ in the first iteration, $\tau = 1$, $\lambda = 13$ in the second iteration, and $\tau = 0.7$, $\lambda = 13$ in the third iteration. The complexity of the algorithm is linear with respect to the number of pixels in the input image and the computation time of our Matlab-only implementation (M-file) takes approximately 4 minutes on a 2.0-GHz Pentium machine for a 256×256 image.

The PSNR and the mean structural similarity index map (MSSIM) [17] results of the proposed method are provided in Table I. In Figure 3 we compare denoised images of the proposed method and the methods that it extends: BM3D [3], P.SA-DCT [7], and SA-BM3D [4]. Relatively high noise standard deviations ($\sigma = 35$ and $\sigma = 25$) were used in order to emphasize the differences in the results by each method. From the figure, we observe that the proposed method effectively reconstructs fine details and at the same time introduces less artifacts than the other methods. This observation is also supported by the improvement in terms of MSSIM and PSNR over the other methods.

## IV. DISCUSSION AND CONCLUSIONS

Regardless of the shape of the adaptive-shape neighborhood obtained in Step 2 of our algorithm, we use block-matching of *square blocks* to find similar neighborhoods, just as it is done in the original BM3D algorithm [3]. There arises the question why the matching is not done between adaptive-shape neighborhoods. In the presence of noise, the employed distance measure ($\ell^2$-norm of the difference between two neighborhoods) used for the matching can be reliably computed only if the number of elements in the neighborhood is greater than a number that depends on the signal-to-noise ratio (SNR) in these blocks (an illustration of such "breakdown point" can be seen in [3]). Thus, considering that a shape-adaptive neighborhood may consist only of a single pixel and that the SNR can be relatively low, we conclude that the matching cannot be done among the adaptive-shape neighborhoods with sufficient reliability.

It is worth discussing the condition for using the PCA instead of the fixed SA-DCT, i.e. $\frac{N_{gr}}{N_{el}} \geq \tau$. Let us remark that the effectiveness of the PCA crucially depends on the sample second-moment matrix $S$ (1). The ideal case would be to compute the second moments from the noise-free image (i.e., from an oracle), but in reality we can only estimate them either from the noisy data (as in the first iteration of the algorithm) or from otherwise distorted data (for the second and third iterations, the matrix is computed from the image estimate produced by the first and second iteration, respectively). A reasonable assumption is that greater $N_{gr}$ results in better estimation of the second-moment matrix. When $N_{gr}$ is small (say $N_{gr} < \tau N_{el}$), the second moments cannot be reliably estimated due to limited training data and we resort to the fixed SA-DCT transform. We recall that in the three iterations we use $\tau = 1.3, 1, 0.7$, respectively. Note that already in the case $N_{gr} < N_{el}$ (when $\tau < 1$), we have $N_{el} - N_{gr}$ null eigenvalues whose corresponding eigenvectors do not convey meaningful information. However, these eigenvectors will be trimmed since we basically keep only those eigenvectors corresponding to the larger eigenvalues. In fact, this trimming can be considered as part of the shrinkage as it is equivalent to unconditional truncation to zero of the transform coefficients that would have corresponded to the discarded eigenvectors. We note that the PC trimming approach, though a standard procedure in the practical application of the PCA (e.g., Section 2.8.5. of [10], Chapter 6 of [11]), is quite coarse and better ways could be devised to incorporate the eigenvalue

magnitudes in the shrinkage itself.

Let us briefly compare our method with the K-SVD [6] and the local PCA denoising [14], both of which employ data-adaptive transform bases. The K-SVD performs training of a global (overcomplete, in general) dictionary of basis elements on small square patches; the training can be done on a set of noise-free images or directly on the noisy image. The efficacy of the subsequent denoising (performed on noisy image patches) depends on the ability of the dictionary elements to sparsely represent true-image data and thus separate it from the noise. The local PCA denoising [14] exploits PCA on square image blocks, where the covariance matrix used by the PCA is estimated from all blocks in a given neighborhood. The proposed approach exploits adaptive-shape neighborhoods, which allows for further adaptation to image details and it estimates the second-moment matrix for the PCA using *similar* adaptive-shape neighborhoods as training data — and not just any local neighborhoods. Furthermore, the sparsity of the true data is further increased by applying a transform along the third dimension of the grouped adaptive-shape neighborhoods.

The experimental results shown in Section III are very promising and demonstrate that by employing shape-adaptive PCA we can further improve the state-of-the-art denoising performance of the BM3D algorithm. Future work shall address novel shrinkage criteria, which are adaptive with respect to the utilized transforms, and the use of adaptive transforms for the third-dimension of the group.

## ACKNOWLEDGMENT

This work was supported by the Academy of Finland (project no. 213462, Finnish Programme for Centres of Excellence in Research 2006-2011, project no. 118312, Finland Distinguished Professor Programme 2007-2010, and project no. 129118, Postdoctoral Researcher's Project 2009-2011) and by Tampere Graduate School in Information Science and Engineering (TISE).

## REFERENCES

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

[2] P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," *IEEE Trans. Image Process.*, 2008, accepted for publication.

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

[4] ——, "A nonlocal and shape-adaptive transform-domain collaborative filtering," in *Proc. Local and Nonlocal Approx. in Image Process.*, Lausanne, Switzerland, September 2008.

[5] ——, "Spatially adaptive support as a leading model-selection tool for image filtering," in *Proc. WITMSE*, Tampere, Finland, August 2008.

[6] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. on Image Process.*, vol. 15, no. 12, pp. 3736–3745, December 2006.

[7] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise Shape-Adaptive DCT for high-quality denoising and deblocking of grayscale and color images," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1395–1411, May 2007.

[8] J. Guerrero-Colon and J. Portilla, "Two-level adaptive denoising using Gaussian scale mixtures in overcomplete oriented pyramids," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, Genova, Italy, September 2005.

[9] J. A. Guerrero-Colón, E. P. Simoncelli, and J. Portilla, "Image denoising using mixtures of Gaussian scale mixtures," in *Proc. IEEE Int. Conf. Image Process.*, October 2008.

[10] J. E. Jackson, *A User's Guide to Principal Components*. Wiley, 1991.

[11] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. Springer, 2002.

[12] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2866–2878, October 2006.

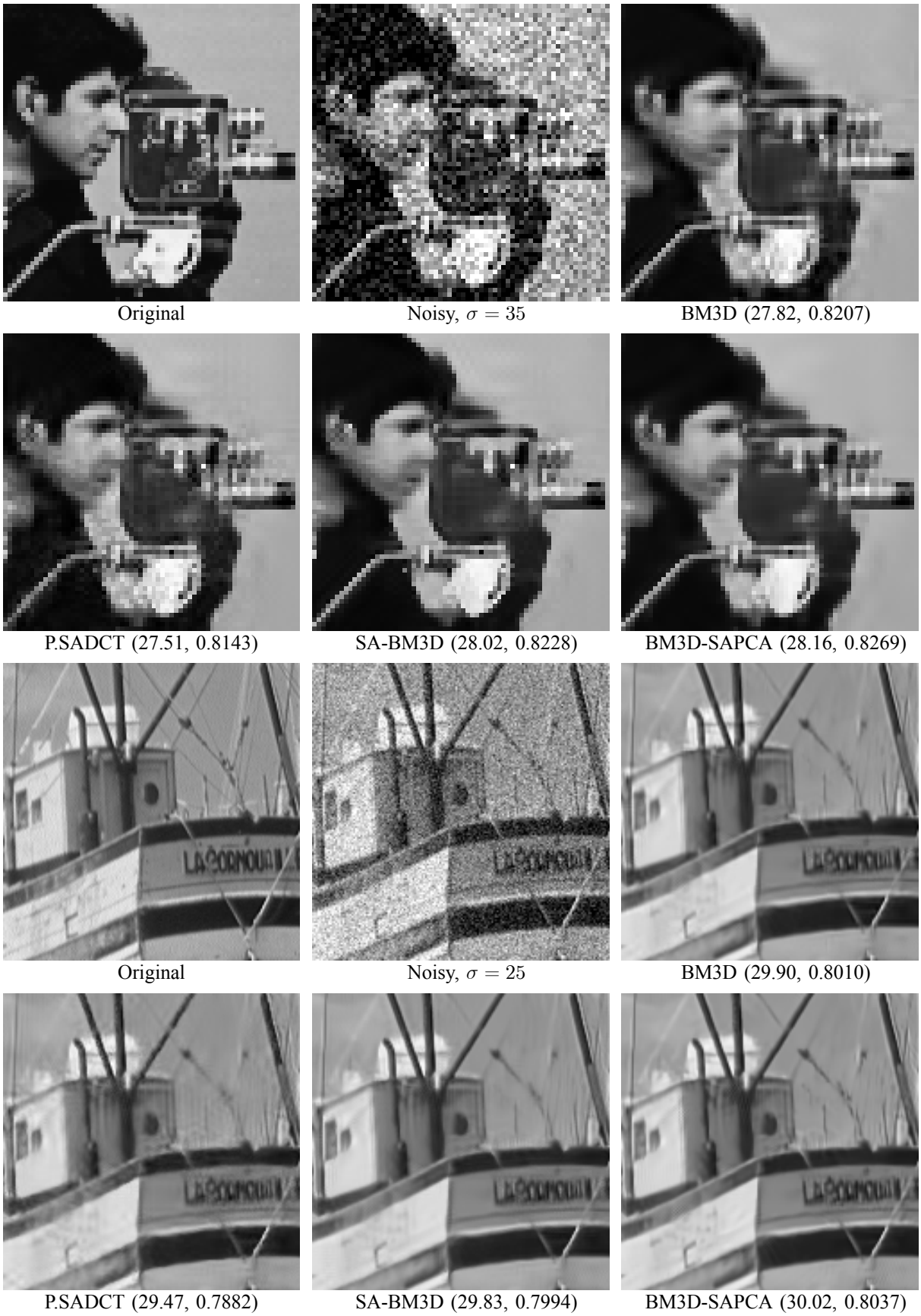
[13] S. Lansel, "DenoiseLab," http://www.stanford.edu/~slansel/DenoiseLab.

[14] D. Muresan and T. Parks, "Adaptive principal components and image denoising," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, September 2003.

[15] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, "Image denoising using a scale mixture of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, November 2003.

[16] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, February 2007.

[17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, April 2004.



Original | Noisy, $\sigma = 35$ | BM3D (27.82, 0.8207)

P.SADCT (27.51, 0.8143) | SA-BM3D (28.02, 0.8228) | BM3D-SAPCA (28.16, 0.8269)

Original | Noisy, $\sigma = 25$ | BM3D (29.90, 0.8010)

P.SADCT (29.47, 0.7882) | SA-BM3D (29.83, 0.7994) | BM3D-SAPCA (30.02, 0.8037)

Fig. 3. Zoomed fragments of *Cameraman* and *Boats* images filtered with the methods that we compare with. The numbers listed in brackets are PSNR [dB] and MSSIM results, respectively.