

COLLABORATIVE FILTERING BASED ON GROUP COORDINATES FOR SMOOTHING AND DIRECTIONAL SHARPENING

Lucio Azzari and Alessandro Foi

Department of Signal Processing, Tampere University of Technology
P.O. Box 553, FIN-33101 Tampere, Finland

ABSTRACT

Groups of mutually similar image blocks are the key element in nonlocal image processing. In this work, the spatial coordinates of grouped blocks are leveraged in two distinct parts of the transform-domain collaborative filtering within the BM3D algorithm. First, we introduce an adaptive 1-D transform for 3-D collaborative filtering based on sampling 2-D smooth functions at the positions of grouped blocks. This adaptive transform is applied for improved decorrelation of the 2-D spectra of the grouped blocks. Second, we propose a directional sharpening procedure whose strength varies adaptively according to the relative orientation of the transform basis functions with respect to the group coordinates. Experiments confirm the efficacy of the proposed adaptations, for denoising as well as for sharpening of noisy images.

Index Terms— BM3D, adaptive transforms, collaborative filtering, denoising, sharpening.

1. INTRODUCTION

BM3D [1] is one of the most effective nonlocal image denoising algorithm. Its state-of-the-art performance is based on the so-called grouping and collaborative filtering approach, which consists in 1) stacking mutually similar image blocks into 3-D *groups*, 2) a 3-D transformation of the groups into typically very sparse group spectra; 3) shrinkage of these spectra; 4) inversion of the 3-D transform and 5) aggregation of the resulting block estimates at their original local within the image.

A peculiarity of this procedure is that steps 2), 3), and 4) are carried out irrespective of the positions that the grouped blocks had in the image, with the 3-D transform realized as a separable composition of a 2-D block transform with a standard 1-D decorrelating transform along the stacking dimension of the group.

We here introduce a procedure for the adaptive design of a 3-D separable decorrelating transform based on the spatial coordinates of the similar blocks that enter the group, thus

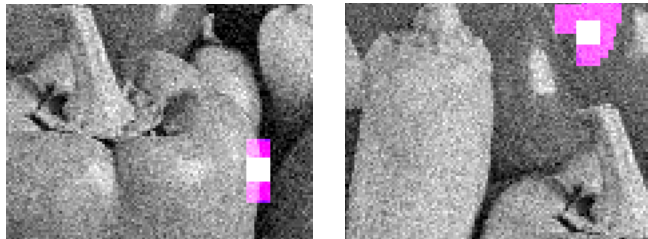


Fig. 1: Details of a noisy realization of the *Peppers* image. The purple area highlights patches similar to the white reference ones. Note how the patches that contain an edge are typically organized along the edge itself, in collinear relative positions (left); conversely, uniform regions yield groups of blocks scattered without a clear pattern (right).

leading to a spatially consistent collaborative filtering of the nonlocal features of the data.

The proposed procedure is designed to enhance BM3D performance whenever the extracted similar blocks spectra can be approximated as smooth functions of the blocks' spatial coordinates, while maintaining its usual performances in the generic case. We thus define, in Section 3, an adaptive orthonormal transform based on the spatial coordinates of the blocks entering the group.

Further, observing that the relative positions of the blocks in the group can be indicative of the block content (see Fig. 1), in Section 4 we also devise a sophisticated modification of collaborative sharpening [2] where both the position of the blocks and the dominant direction of each basis function modulate the strength of enhancement of the corresponding spectrum coefficient.

2. OBSERVATION MODEL AND BM3D

Let us consider a noisy observation z of a noise free signal y corrupted by additive i.i.d. Gaussian noise η ,

$$z(x) = y(x) + \eta(x), \quad (1)$$

where $x \in X \subset \mathbb{R}^2$ are the coordinates of the samples and $\eta(\cdot) \sim \mathcal{N}(0, \sigma^2)$. The goal is to estimate y from z .

Image blocks are denoted by a capital letter and a subscript that indicates the top-left coordinate of the block; e.g.,

This work was supported by the Academy of Finland (project no. 252547).

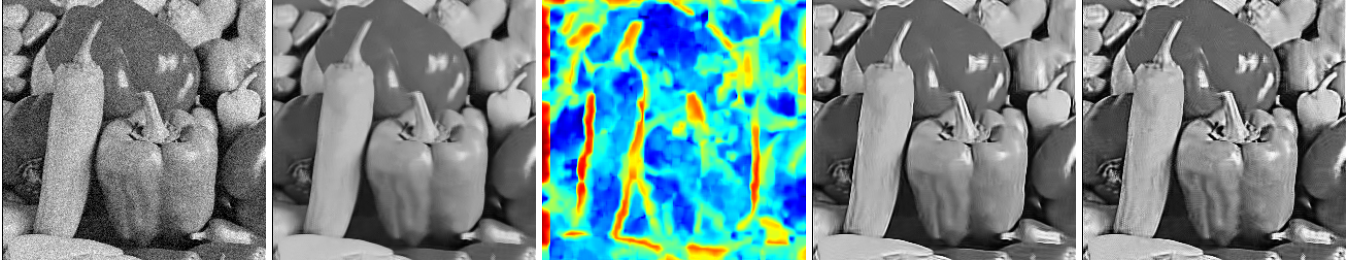


Fig. 2: Sharpening of *Peppers* corrupted by Gaussian noise with $\sigma = 20$. From left to right: *Peppers* noisy image; denoised image with no sharpening; ratio between the lengths of the principal axes of the group coordinates: higher values are shown in red, and lower values in dark blue; image denoised and sharpened by the proposed algorithm (adaptive $0.7 \leq \alpha \leq 1.8$); conventional alpha-rooting (constant $\alpha = 1.5$).

Z_x is a block extracted from z at position x .

In the basic BM3D formulation [1], the image z is raster scanned and for each position x_R and corresponding reference block Z_{x_R} we find the *ordered* set S_{x_R} of the coordinates of the N blocks most similar to Z_{x_R} . The similarity is measured by computing the ℓ_2 -distances of the blocks content. The blocks are then stacked together in a 3-D volume $\mathbf{Z}_{S_{x_R}}$, that is subsequently transformed through the 3D separable linear transform \mathcal{T}_{3D} , filtered, and synthesized with the 3D inverse transform \mathcal{T}_{3D}^{-1} :

$$\hat{\mathbf{Y}}_{S_{x_R}} = \mathcal{T}_{3D}^{-1} \left(\Upsilon \left(\mathcal{T}_{3D} \left(\mathbf{Z}_{S_{x_R}} \right) \right) \right), \quad (2)$$

where $\hat{\mathbf{Y}}_{S_{x_R}}$ is the filtered 3-D volume, and Υ is a shrinkage operator, such as hard-thresholding or a Wiener filter. For the success of the procedure it is important that the underlying unknown spectrum $\mathcal{T}_{3D} \left(\mathbf{Y}_{S_{x_R}} \right)$ is sparse, as this permits to effectively attenuate the noise without introducing severe bias [1]. The coordinates S_{x_R} do not play any role in the collaborative filtering (2) applied to $\mathbf{Z}_{S_{x_R}}$; S_{x_R} regain their significance only when the block estimates $\hat{Y}_x, x \in S_{x_R}$, are extracted from $\hat{\mathbf{Y}}_{S_{x_R}}$ and aggregated at their original position into the resulting image estimate \hat{y} .

3. ADAPTIVE GROUP TRANSFORM BASED ON GROUP COORDINATES

Although a number of works [3, 4] have discussed the performance limits of denoising, indicating that BM3D is essentially attaining a performance bound on complex natural images, it is otherwise established that nonlocal methods often yields suboptimal results when filtering simpler images, such those composed by large regular surfaces. Thus, in this work, we aim at improving the effectiveness of collaborative filtering on such large regular content by embedding a smooth local model within the 3-D transform \mathcal{T}_{3D} , while maintaining comparable performance on complex heterogeneous images.

For simplicity, we consider the usual separable decomposition of \mathcal{T}_{3D} into a spatial 2-D transform \mathcal{T}_{2D} (e.g., DCT, DWT [1], PCA [5]) followed by orthonormal 1-D transform

\mathcal{T}_{1D} along the stacking dimension, and focus our attention on the latter. In particular, a curious feature of \mathcal{T}_{1D} , is that its particular choice bears negligible influence on the denoising performance, as long as \mathcal{T}_{1D} is chosen from fixed non-adaptive transforms having a DC term [1]. However, here we consider an adaptive design for \mathcal{T}_{1D} .

Let $\Phi = \{\phi_j\}_{j=1}^{|S_{x_R}|}$ be a collection of bivariate functions $\phi_j : X \rightarrow \mathbb{R}$, such that the one-dimensional vectors

$$\{\phi_j(S_{x_R})\}_{j=1}^{|S_{x_R}|} \quad (3)$$

form a set of linearly independent generators for $\mathbb{R}^{|S_{x_R}|}$. We can then build the $|S_{x_R}| \times |S_{x_R}|$ matrix P whose columns are the aforementioned vectors. Since the functions ϕ_j are linearly independent, *i.e.* $\text{rank}(P) = |S_{x_R}|$, we can apply the QR decomposition (Gram-Schmidt orthonormalization) to P . In this way we obtain

$$QR = P, \quad (4)$$

where Q is an orthonormal matrix and R is an upper-triangular matrix. We can interpret Q as an orthonormal transform for inputs of length $|S_{x_R}|$. Most importantly, the columns of Q (*i.e.* the basis functions of the transform) inherit from Φ the regularity w.r.t. S_{x_R} ; such spatially adaptive orthonormalization of P can be seen as a particular case of shape-adapted transforms [6]. We use such Q as a direct replacement of \mathcal{T}_{1D} in BM3D.

By design, the proposed transform yields sparser spectra when the individual \mathcal{T}_{2D} -spectrum coefficients of the blocks in $\mathbf{Y}_{S_{x_R}}$ agree with the regularity of Φ over $x \in S_{x_R}$. The validity of this assumption can be different for different \mathcal{T}_{2D} -spectrum components and different S_{x_R} ; it is reasonable to expect that high-frequency components of the \mathcal{T}_{2D} spectrum meet this assumption to a lesser degree than the low-frequency ones. In practice, for each group, Φ is expressed with respect to the principal axes of S_{x_R} , and in the orthonormalization (4) the basis functions ϕ_j are sorted so to maximize their independence over S_{x_R} . The principal axes play an essential role also in the following adaptive sharpening strategy.

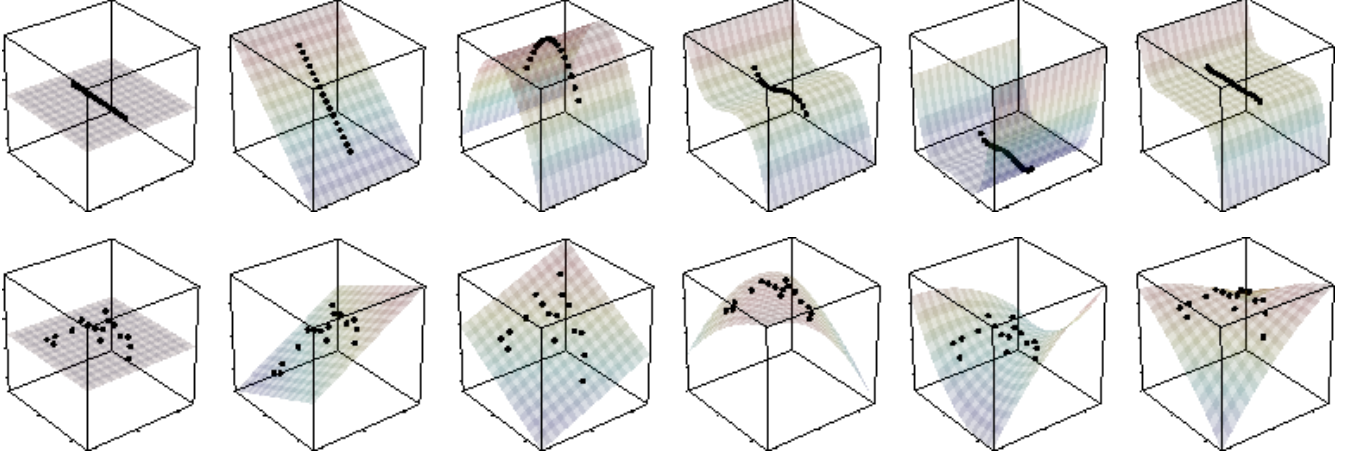


Fig. 3: Example of adaptive bases for the \mathcal{T}_{1D} transform generated by orthogonal polynomials. On the top row, we show the first 6 basis functions for the group in Fig. 1(left), whose blocks are aligned along an edge. On the bottom row, we show the first 6 basis functions for the group in Fig. 1(right), located within a uniform area. Observe how the bases change depending on the group coordinates.

4. DIRECTIONAL SHARPENING WITH ADAPTIVE TRANSFORM-DOMAIN ALPHA-ROOTING

The collaborative-filtering framework allows to conveniently sharpen the noisy image z without amplification of the noise through *alpha-rooting* of the thresholded \mathcal{T}_{3D} spectrum [2]. Here, instead of adopting an alpha-rooting of the image with constant α [2], we propose to amplify the \mathcal{T}_{3D} -spectrum coefficients $t(i)$ with adaptive $\alpha(i)$ that can be different for each i , i being the index within the \mathcal{T}_{3D} spectrum:

$$t_{\alpha(i)}(i) = \begin{cases} \text{sign}[t(i)] |t(0)| \left| \frac{t(i)}{t(0)} \right|^{\frac{1}{\alpha(i)}} & \text{if } t(0) \neq 0 \\ t(i) & \text{otherwise,} \end{cases} \quad (5)$$

where $t_{\alpha(i)}(i)$ is the amplified \mathcal{T}_{3D} coefficient and $t(0)$ is the \mathcal{T}_{3D} -spectrum DC.

Specifically, we adapt $\alpha(i)$ based on 1) the ratio between the major and minor axes of the group coordinates, as well as on 2) how the orientation of the major axis agrees with the orientation of the \mathcal{T}_{2D} basis functions that contribute to the individual coefficients $t(i)$ via \mathcal{T}_{1D} . If the major axis is sensibly longer than the minor one, the blocks are mostly distributed along the major axis: this occurs when the group includes pixels from a straight edge in the image. Conversely, when the axes' lengths are similar, the blocks are scattered, indicating that the group is likely from a wide regular region. This feature is shown in Fig. 2(center), where we show the ratio between major and minor axis for *Peppers*. Thus, when the ratio between the axes' lengths is large, we assume that the image features an intensity change in the direction orthogonal to the major axis, while when the ratio is close to 1, we assume that the group contains a regular region (see Fig. 1). Therefore, when sharpening the blocks content, it is reasonable 1) to use large α values only in the presence of an intensity change, i.e. when the ratio is large, and 2) to amplify only those coeffi-

cients whose corresponding basis functions are characterized by variations collinear with those of the image, i.e. orthogonal to the major axis. We thus compute individually the energy (ℓ_2 norm) of the derivative of each \mathcal{T}_{2D} -transform basis function in the direction of the major and minor axes, namely E_{\parallel} and E_{\perp} . If $E_{\perp} > E_{\parallel}$, we perform sharpening ($\alpha(i) > 1$), otherwise we soften the coefficient ($\alpha(i) < 1$). Hence, we modulate $\alpha(i)$ monotonically with $E_{\perp} - E_{\parallel}$ with rate proportional to the axes ratio.

5. EXPERIMENTS

In our experiments we use bivariate orthogonal polynomials for Φ . In Fig. 3 one can appreciate how the bases given by Q adapt to the particular relative positions of the grouped blocks. On the top row, we show the first six basis elements of the \mathcal{T}_{1D} transform of the group of blocks from the edge in Fig. 1(left). The first six basis functions shown in the bottom row are for the group from the flat area in Fig. 1(right), and are visibly different from the former ones. In the figure, we also show the continuous surfaces obtained by applying R^{-1} to Φ .

In order to compare the performances of the proposed algorithm with respect to the standard BM3D, we denoise a standard set of eleven natural images (*Lena*, *Cameraman*, *Peppers*, etc.) corrupted by i.i.d. Gaussian noise with $\sigma = 15, 35$ (1). We further consider a piecewise smooth test image (Fig. 4, bottom) to validate the advantage of the proposed scheme in its ideal conditions.

To clearly discriminate the performances of the two implementations, and also because collaborative sharpening had not been designed within a two-stage algorithm [2], we restrict our analysis to the hard-thresholding only. Furthermore, in order to have a fair comparison, we use the same parameters for both implementations except for the parameter λ of

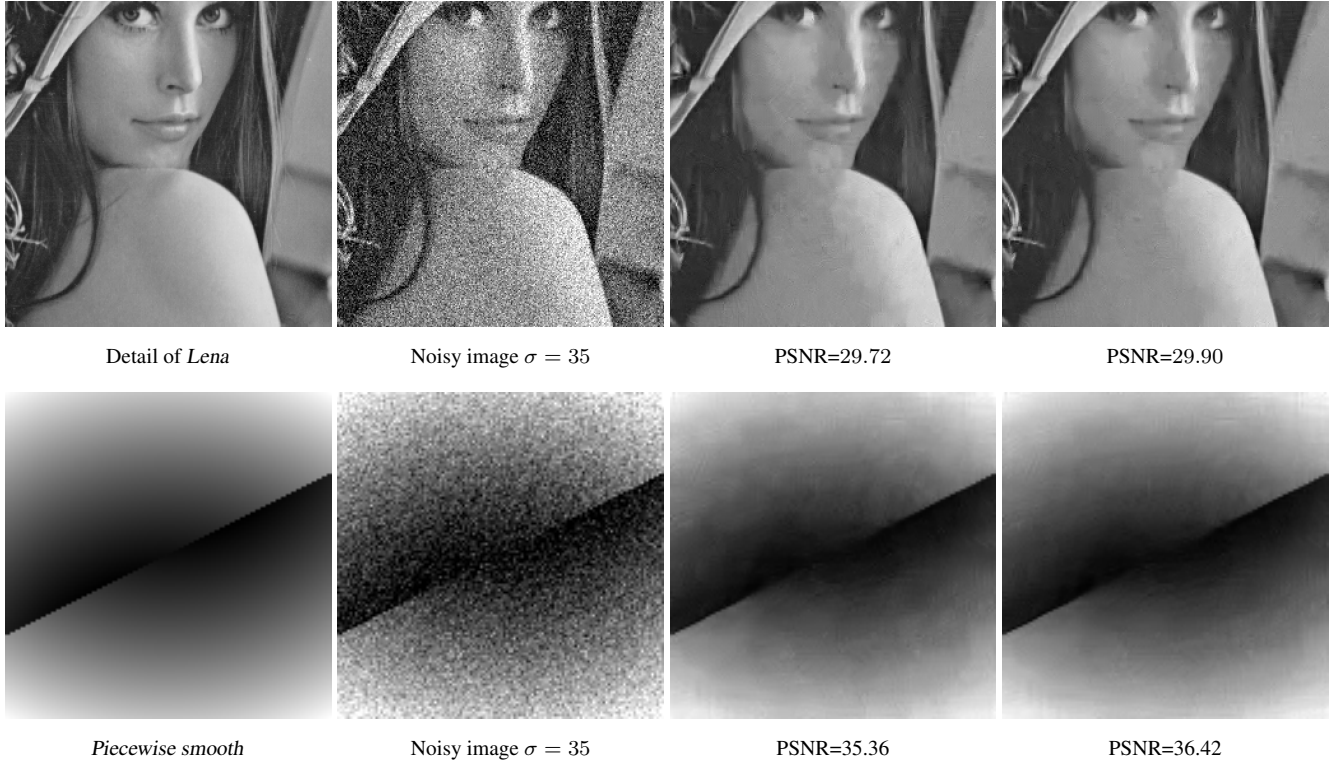


Fig. 4: Denoising of *Lena* (top) and piecewise-smooth test image (bottom), both corrupted by i.i.d. Gaussian noise with $\sigma = 35$. From left to right: original images, noisy observations, images denoised by the standard BM3D algorithm, images denoised by the proposed algorithm with collaborative filtering based on adaptive orthonormal polynomials (see Fig. 3). Notice the improvement, particularly in smooth regions, such as the shoulder of *Lena*.

the hard-thresholding operator Υ in (2), that has been optimized for the two algorithms independently, so to maximize the PSNR average over the dataset of natural images.

The PSNR differences between the proposed and the original implementation are, in terms of mean \pm standard-deviation over the natural images dataset, 0.076 ± 0.063 and 0.12 ± 0.065 for $\sigma = 15$ and $\sigma = 35$, respectively, i.e. there is a marginal but recurrent improvement. For the piecewise-smooth test image, the gain is instead $0.63dB$ and $1.28dB$ for the two noise levels, which is significant and confirms the advantage when the blocks spectra vary smoothly with respect to the group coordinates. Fig. 4 compares the results of the two implementations for *Lena* and for the test image, where one can appreciate the improvement on *Lena*'s shoulder as well as on the gradients of the bottom figure.

Finally, Fig. 2 compares the conventional sharpening algorithm [2] with constant $\alpha = 1.5$ against the proposed one with adaptive $0.7 \leq \alpha(i) \leq 1.8$. Although the two methods yield similar enhancement of the edges, the former one presents spurious artefacts particularly noticeable over smooth regions and in the vicinity of edges. These artefacts are instead mitigated by the proposed adaptive alpha-rooting, because the sharpening is strong only across edges, while softening occurs along edges (due to the different $\alpha(i)$ for differently oriented \mathcal{T}_{2D} -spectrum basis functions) and smooth

areas are merely denoised (due to the axes' length ratio being close to unity).

6. DISCUSSION AND CONCLUSIONS

We introduced an adaptive 1-D transform for 3-D collaborative filtering based on sampling 2-D smooth functions at the positions of grouped blocks. Experiments demonstrate a slight improvement over the standard BM3D on complex natural images and a substantial advantage on images characterized by large piecewise regular regions. These improvements are confirmed both quantitatively and qualitatively. This is consistent with results on the performance limits in image denoising, that predict only a small room for further improvement over BM3D on complex natural images. The proposed approach is further extended to a sharpening algorithm where the group coordinates determine an adaptive alpha-rooting whose strength can vary from group to group as well as within each group spectrum. Unlike the traditional alpha-rooting that amplifies the whole spectrum indiscriminately, the proposed method enables a directional sharpening which is only acting across edges, thus reducing the visibility of artefacts within smooth regions as well as along edges. The proposed scheme is not limited to denoising and sharpening, but it is relevant to any application of 3-D collaborative filtering.

7. REFERENCES

- [1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug 2007.
- [2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Joint image sharpening and denoising by 3D transform-domain collaborative filtering," in *Proc. 2007 Int. TICSP Workshop Spectral Meth. Multirate Signal Process., SMMSP*, 2007.
- [3] P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 895–911, April 2010.
- [4] A. Levin, B. Nadler, F. Durand, and W.T. Freeman, "Patch complexity, finite pixel correlations and optimal denoising," in *Proc. European Conference on Computer Vision (ECCV 2012)*, pp. 73–86. Springer, 2012.
- [5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," in *Proc. 2009 workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS'09)*, 2009.
- [6] W. Philips, "A fast algorithm for orthogonalizing polynomials on an arbitrarily shaped region (revised version)," *Multidimensional Systems and Signal Processing*, vol. 8, no. 4, pp. 409–421, 1997.