EXACT TRANSFORM-DOMAIN NOISE VARIANCE FOR COLLABORATIVE FILTERING OF STATIONARY CORRELATED NOISE

Ymir Mäkinen, Lucio Azzari, and Alessandro Foi

Tampere University, Finland

ABSTRACT

Collaborative filters perform denoising through transform-domain shrinkage of a group of similar blocks extracted from an image. Existing methods for collaborative filtering of stationary correlated noise have all used simple approximations of the transform noise power spectrum adopted from methods which do not employ block grouping. We note the inaccuracies of these approximations and introduce a method for the exact computation and effective approximations of the noise power spectrum. Unlike earlier methods, the calculated noise variances are exact even when noise in one block is correlated with noise in any of the other blocks. We discuss the adoption of the exact noise power spectrum within shrinkage, in similarity testing (block matching), and in aggregation. Extensive experiments support the proposed method over earlier crude approximations used by image denoising filters such as BM3D, demonstrating dramatic improvement in many challenging conditions.

***Index Terms*—** Image denoising, correlated noise, collaborative filtering, noise power spectrum, BM3D.

1. INTRODUCTION

Transform-based denoising algorithms perform noise removal in a chosen domain where the signal to recover is sparse, i.e. it can be represented with few coefficients significantly different from zero. Transform-based algorithms can be combined with the principles of nonlocal denoising (e.g., [1, 2]) to exploit the mutual similarity between blocks at different locations in the image. BM3D [3] is one of the leading methods in this hybrid class known as *collaborative filters*. Mutually similar blocks are jointly processed by applying first a 2-D transform $\mathcal{T}^{2D}$ to each block and then a 1-D transform $\mathcal{T}^{1D}$ across the obtained $\mathcal{T}^{2D}$-spectra. This results in a 3-D transform $\mathcal{T}^{3D}$ that decorrelates both local and nonlocal image regularity. The advantage of collaborative filtering lies in the enhanced sparsity in this $\mathcal{T}^{3D}$ domain where shrinkage is performed. However, the effectiveness of denoising hinges on a correctly set shrinkage threshold, which in turn requires knowledge of the noise variance in this transform domain.

To model the $\mathcal{T}^{3D}$ noise power spectrum, [3] and subsequent works, such as the BM3D filter for correlated noise [4], have adopted a simplified modelling borrowed from local filters like [5, 6], where the PSD in $\mathcal{T}^{3D}$ domain is calculated by merely replicating the $\mathcal{T}^{2D}$ PSDs. However, this model presumes $\mathcal{T}^{1D}$ is orthonormal and, most importantly, that noise in one block is always independent from that in any other block. The latter requirement often does not apply. As noted in [3] for i.i.d. noise, noise correlation between blocks may occur due to their overlap. Additionally, with stationary correlated noise, noise may be correlated across different blocks even if they do not overlap, potentially creating large inaccuracies in the simplified approximations.

We introduce a method for the exact computation of the noise variance in collaborative transform domain as well as effective approximations for faster computation. The new variance calculation is embedded into the BM3D algorithm, where it is used to improve shrinkage accuracy and block matching, and for weighting in the aggregation of the filtered blocks. This leads to dramatic improvement of denoising results for a wide variety of correlated noise types.

♡ This work was supported by the Academy of Finland (project no. 310779).

2. PROPOSED METHOD

2.1. Problem formulation

Let us consider a noisy observation $z : X \to \mathbb{R}$ of an unknown deterministic noise-free image y corrupted by additive stationary correlated noise η

$$z(x) = y(x) + \eta(x), \quad x \in X, \tag{1}$$

where $x \in X \subset \mathbb{Z}^2$ is the coordinate in the finite 2-D regular image domain X, and

$$\eta = \nu \circledast g, \qquad \nu(\cdot) \sim \mathcal{N}(0,1), \tag{2}$$

ν being zero-mean i.i.d. Gaussian noise with unit variance, and g being a convolution kernel representing the spatial correlation of the noise. Since $\mathrm{var}\{\nu\} = 1$, $\mathrm{var}\{\eta\} = \|g\|_2^2$. An equivalent way of representing correlated noise is by its PSD Ψ:

$$\Psi = \mathrm{E}\left\{|\mathscr{F}[\eta]|^2\right\} = \mathrm{var}\left\{\mathscr{F}[\eta]\right\} = |X|\,|\mathscr{F}[g]|^2, \tag{3}$$

$\mathscr{F}$ being the 2-D Fourier transform. When g is a scaled Dirac delta, (1)-(2) reduces to the additive white Gaussian noise (AWGN) model. The goal of denoising is to estimate y from z; we assume Ψ known.

2.2. Variance of the $\mathcal{T}^{3D}$ spectrum of a group of blocks

Collaborative filters operate on groups of similar blocks extracted from the image. Let $\left\{z_{x_1}, \ldots, z_{x_M}\right\}$ be such a group of M 2-D blocks extracted from z at coordinates[1] $x_1, \ldots, x_M$, respectively, where each block is composed of N elements. Let $\mathcal{T}^{2D}$ be a 2-D block transform, and denote by $s_i^{x_t} = \left\langle z_{x_t}, b_i^{2D}\right\rangle$ a generic $\mathcal{T}^{2D}$-spectrum coefficient of z_{x_t}, where b_i^{2D} is the i-th basis function of $\mathcal{T}^{2D}$. Further, we denote by $\left\{s_{i,j}^{x_1,\ldots,x_M},\ i{=}1,\ldots,N,\ j{=}1,\ldots,M\right\}$ the $\mathcal{T}^{3D}$ spectrum of the group $\left\{z_{x_1}, \ldots, z_{x_M}\right\}$, computed by applying a 1-dimensional transform $\mathcal{T}^{1D}$ of length M to $\left[s_i^{x_1}, \ldots, s_i^{x_M}\right]$, $i{=}1,\ldots,N$.

The core of this work is about the calculation and use of the variances $\mathrm{var}\left\{s_{i,j}^{x_1,\ldots,x_M}\right\}$ of the $\mathcal{T}^{3D}$ spectrum $s_{i,j}^{x_1,\ldots,x_M}$, which we denote by $v_{i,j}^{x_1,\ldots,x_M}$.

[1] As coordinate of a block we intend the coordinate of the pixel in the top-left corner.

2.2.1. Preliminaries

The noise variance in $\mathcal{T}^{2D}$ can be calculated from the PSD and the transform basis functions:

$$\operatorname{var}\left\{s_i^{x_t}\right\} = \operatorname{var}\left\{\left(v \circledast g \circledast \overleftrightarrow{b}_i^{2D}\right)(x_t)\right\} = \operatorname{var}\{v\}\left\|g \circledast \overleftrightarrow{b}_i^{2D}\right\|_2^2,$$

where the $\overleftrightarrow{\ }$ decoration denotes the reflection about the origin of $\mathbb{Z}^2$. We note that this variance does not depend on the coordinate x_t of the block; hence we can adopt the simple notation $v_i = \operatorname{var}\left\{s_i^{x_t}\right\}$, and since $\operatorname{var}\{v\} = 1$,

$$v_i = \left\|g \circledast \overleftrightarrow{b}_i^{2D}\right\|_2^2 = \left\|\,|X|^{-2}\,\Psi\left|\mathscr{F}\left[\overleftrightarrow{b}_i^{2D}\right]\right|^2\right\|_1, \tag{4}$$

where the last equality follows from Parseval's isometry and (3).

The $\mathcal{T}^{3D}$-spectrum coefficients are calculated through the direct tensor product of $\mathcal{T}^{2D}$ and $\mathcal{T}^{1D}$, as

$$\begin{aligned} s_{i,j}^{x_1,\dots,x_M} &= \left\langle \left[z_{x_1};\cdots;z_{x_M}\right], b_i^{2D}\otimes b_j^{1D}\right\rangle = \\ &= \left\langle \left[s_i^{x_1},\cdots,s_i^{x_M}\right], b_j^{1D}\right\rangle = \sum_{t=1}^{M} b_j^{1D}(t)\, s_i^{x_t}, \end{aligned} \tag{5}$$

where $b_j^{1D}(t)$ is the t-th element of the j-th basis function b_j^{1D} of $\mathcal{T}^{1D}$, and $[\cdot;\cdots;\cdot]$ denotes the stacking along the 3rd dimension.

2.2.2. Conventional methods for approximating $v_{i,j}^{x_1,\dots,x_M}$

The exact calculation of the variance of the $\mathcal{T}^{3D}$-spectrum coefficients $s_{i,j}^{x_1,\dots,x_M}$ may not be immediate from (5):

$$\begin{aligned} v_{i,j}^{x_1,\dots,x_M} &= \operatorname{var}\left\{s_{i,j}^{x_1,\dots,x_M}\right\} = \operatorname{var}\left\{\sum_{t=1}^{M} b_j^{1D}(t)\, s_i^{x_t}\right\} = \\ &= \sum_{t=1}^{M}\left(b_j^{1D}(t)\right)^2 \operatorname{var}\left\{s_i^{x_t}\right\} + \sum_{k\neq t} b_j^{1D}(k)\, b_j^{1D}(t)\operatorname{cov}\left\{s_i^{x_t}, s_i^{x_k}\right\} \quad (6) \\ &\approx v_i \sum_{t=1}^{M}\left(b_j^{1D}(t)\right)^2, \quad (7) \end{aligned}$$

where the term $\operatorname{cov}\{s_i^{x_t}, s_i^{x_k}\}$ represents the covariance between same $\mathcal{T}^{2D}$ spectrum coefficients for different blocks. The approximation (7) becomes an equality if the noise of the stacked blocks is independent; in other words, noise may be correlated within each block but not across distinct blocks. If $\mathcal{T}^{1D}$ is additionally orthonormal, then $\sum_{t=1}^{M}\left(b_j^{1D}(t)\right)^2 = 1$ and $v_{i,j}^{x_1,\dots,x_M} = v_i$, that is, the $\mathcal{T}^{3D}$ spectrum variances will be identical to those of the $\mathcal{T}^{2D}$ spectra.

This approximation is used in BM3D [3] and its derivative works [4, 7, 8, 9], which all use an orthonormal $\mathcal{T}^{1D}$. However, the requirement of independence between blocks is not always fulfilled: it obviously fails if the blocks are overlapping, but when noise is correlated it may fail even if they do not overlap as correlation may exceed the block boundaries. The failure of this requirement results in potentially significant errors in the calculation of the variances.

2.2.3. Exact calculation of $v_{i,j}^{x_1,\dots,x_M}$

Observe that $s_i^{x_t}$ can equivalently be written as[2]

$$s_i^{x_t} = \left\langle z, b_i^{2D}\circledast\delta_{x_t}\right\rangle, \tag{8}$$

[2]The first element of b_i^{2D} is used as its origin in convolution.

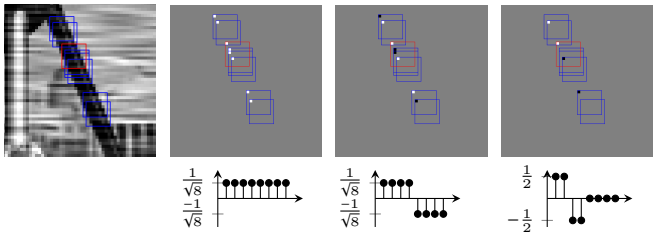


Fig. 1. A noisy image and a group of 8 blocks (reference block in red), and the corresponding $\tilde{b}_j^{1D}$ for $j = 1, 2, 3$, as well as the respective b_j^{1D} where $\mathcal{T}^{1D}$ is the Haar transform. Gray in the $\tilde{b}_j^{1D}$ images indicates the zero level.

where δ_{x_t} is a Dirac delta at the coordinate x_t. Therefore

$$s_{i,j}^{x_1,\dots,x_M} = \sum_{t=1}^{M} b_j^{1D}(t)\left\langle z, b_i^{2D}\circledast\delta_{x_t}\right\rangle = \tag{9}$$

$$= \sum_{t=1}^{M}\left\langle z, b_i^{2D}\circledast\left(b_j^{1D}(t)\,\delta_{x_t}\right)\right\rangle. \tag{10}$$

The sum in (10) can be finally seen as the convolution

$$s_{i,j}^{x_1,\dots,x_M} = \left(\overleftrightarrow{z}\circledast b_i^{2D}\circledast\tilde{b}_j^{1D}\right)(0), \tag{11}$$

where $\tilde{b}_j^{1D} = \sum_{t=1}^{M} b_j^{1D}(t)\,\delta_{x_t}$ is an array of the same size of z that is zero everywhere except at the coordinates x_t where it assumes the corresponding values $b_j^{1D}(t)$ (see Figure 1). Even though (11) is arithmetically identical to (5), it physically embeds the spatial locations $x_1,\dots,x_M$ that (5) had lost through the stacking. Hence,

$$\begin{aligned} v_{i,j}^{x_1,\dots,x_M} &= \operatorname{var}\left\{s_{i,j}^{x_1,\dots,x_M}\right\} = \operatorname{var}\left\{\left(\overleftrightarrow{z}\circledast b_i^{2D}\circledast\tilde{b}_j^{1D}\right)(0)\right\} = \\ &= \operatorname{var}\left\{\left(v\circledast\overleftrightarrow{g}\circledast b_i^{2D}\circledast\tilde{b}_j^{1D}\right)(0)\right\} = \\ &= \left\|\overleftrightarrow{g}\circledast b_i^{2D}\circledast\tilde{b}_j^{1D}\right\|_2^2 = \quad (12) \\ &= \left\|\,|X|^{-2}\,\Psi\left|\mathscr{F}\left[b_i^{2D}\right]\right|^2\left|\mathscr{F}\left[\tilde{b}_j^{1D}\right]\right|^2\right\|_1. \quad (13) \end{aligned}$$

Note how (12) incorporates a convolution against $\tilde{b}_j^{1D}$ which is instead completely missing from (4); this extra operation varies with the relative displacement of the blocks, and thus makes (12)-(13) differ between groups, as opposed to (4)-(7) which are the same for all groups.

Although discussed here in 2-D, the above procedure is applicable to nonlocal collaborative transforms used by filters for arbitrary d-dimensional data, for instance by BM3D for filtering images ($d=2$) and by BM4D for filtering volumetric data ($d=3$).

2.3. Application to BM3D

2.3.1. Block matching

In BM3D, forming groups prioritizes blocks which are the most similar to the reference block. Since matching is based on comparing noisy blocks, it is affected by correlation in the noise. It is thus useful to consider the variances between two blocks compared in the block matching phase.

The difference between two noisy blocks z_{x_R} and z_{x_j} can be

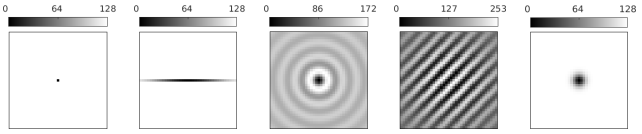


Fig. 2. Left to right: maps of $\sum_{i=1}^{N} v_{i,2}^{x_R,x_j}$ as a function of $x_R - x_j$ for white noise and the four noise types described in Table 1 with $\mathrm{var}\{\eta\} = 1$. Note how the center pixels are black as $v_{i,2}^{x_R,x_j} = 0$, $i = 1,\ldots,N$, when $x_R = x_j$.

written as

$$\left\|z_{x_R} - z_{x_j}\right\|_2^2 = 2\sum_{i=1}^{N}\left\langle\left[s_i^{x_R}, s_i^{x_j}\right], b_2^{1\mathrm{D}}\right\rangle^2 = 2\sum_{i=1}^{N}\left(s_{i,2}^{x_R,x_j}\right)^2, \quad (14)$$

where $b_2^{1\mathrm{D}} = \frac{1}{\sqrt{2}}[1,-1]$ and $s_i^{x_R}, s_i^{x_j}$ are spectra produced by an arbitrary orthonormal transform $\mathcal{T}^{2\mathrm{D}}$. We note that $\left(s_{i,2}^{x_R,x_j}\right)^2$ is a non-central chi-squared random variable with one degree of freedom and with mean and variance

$$\mathrm{E}\left\{\left(s_{i,2}^{x_R,x_j}\right)^2\right\} = v_{i,2}^{x_R,x_j} + \mathrm{E}^2\left\{s_{i,2}^{x_R,x_j}\right\}, \quad (15)$$

$$\mathrm{var}\left\{\left(s_{i,2}^{x_R,x_j}\right)^2\right\} = 2\left(v_{i,2}^{x_R,x_j}\right)^2 + 4v_{i,2}^{x_R,x_j}\mathrm{E}^2\left\{s_{i,2}^{x_R,x_j}\right\}, \quad (16)$$

where $v_{i,2}^{x_R,x_j}$ can be calculated with (13) for the corresponding $\mathcal{T}^{2\mathrm{D}}$ and $\mathcal{T}^{1\mathrm{D}}$ transforms. Noting that $2\sum_{i=1}^{N}\mathrm{E}^2\left\{s_{i,2}^{x_R,x_j}\right\} = \left\|\mathrm{E}\left\{z_{x_R} - z_{x_j}\right\}\right\|_2^2$, we can express the expectation of (14) as

$$\mathrm{E}\left\{\left\|z_{x_R} - z_{x_j}\right\|_2^2\right\} = \left\|\mathrm{E}\left\{z_{x_R} - z_{x_j}\right\}\right\|_2^2 + 2\sum_{i=1}^{N} v_{i,2}^{x_R,x_j}, \quad (17)$$

which quantifies the positive bias in the block comparison and shows that this bias depends exclusively on the noise through $v_{i,2}^{x_R,x_j}$, $i,\ldots,N$, and can vary with the relative position between blocks, i.e. with $x_R - x_j$, as illustrated in Figure 2. By subtracting $2\sum_{i=1}^{N} v_{i,2}^{x_R,x_j}$ from $\left\|z_{x_R} - z_{x_j}\right\|_2^2$, we get an unbiased estimate of $\left\|\mathrm{E}\left\{z_{x_R} - z_{x_j}\right\}\right\|_2^2$. Our experiments indicate that the denoising quality is further improved by ranking potential matches according to

$$\left\|z_{x_R} - z_{x_j}\right\|_2^2 - 2\gamma\sum_{i=1}^{N} v_{i,2}^{x_R,x_j} \quad (18)$$

with $\gamma = 4$. Referring to Figure 2, block matching based on (18), for any $\gamma > 0$, prioritizes blocks in white positions over the darker ones if the differences $\left\|z_{x_R} - z_{x_j}\right\|_2^2$ between the noisy blocks are the same. We speculate that a $\gamma > 1$, beyond compensating the bias, facilitates the inclusion in the group of noisy blocks which differ from the reference block mainly due to a larger $\mathrm{var}\left\{\left\|z_{x_R} - z_{x_j}\right\|_2^2\right\}$.

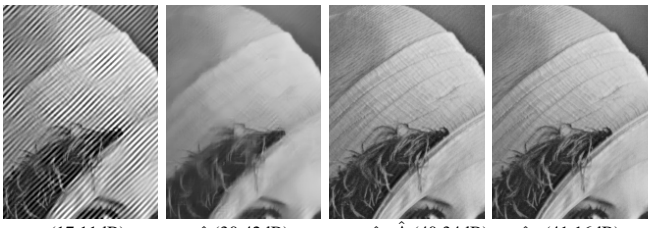
The term $\sum_{i=1}^{N} v_{i,2}^{x_R,x_j}$ can be precomputed since it depends only on $x_R - x_j$, which is limited by the fixed size of the search neighborhood where blocks are matched (e.g., 39×39 as seen in Figure 2).

2.3.2. Shrinkage

The core of BM3D is shrinkage performed on the $\mathcal{T}^{3\mathrm{D}}$ spectrum of the grouped noisy blocks. In particular, the standard implementation of BM3D includes two distinct denoising stages using first hard-thresholding, then empirical Wiener filtering as the shrinkage function. Both stages crucially depend on the noise variances $v_{i,j}^{x_1,\ldots,x_M}$,

Table 1. Base models for the four correlation kernels, where u_h and u_v denote the horizontal and vertical coordinates, and G_ς denotes a Gaussian with standard deviation ς centered at the origin. For each experiment, the kernel g is defined as $g = q_n/\|q_n\|_2^{-2}\sigma$, where σ and $n \in \{1,2,3,4\}$ determine the desired noise variance $\mathrm{var}\{\eta\} = \|g\|_2^2 = \sigma^2$ and correlation model.

q_1	$16 - \lvert u_h \rvert$	$-15 \le u_h \le -15,\ u_v = 0$
q_2	$\cos\left(\left(u_h^2 + u_v^2\right)^{1/2}\right) G_{10}(u_h, u_v)$	$-50 \le u_h, u_v \le 50$
q_3	$\cos(u_h + u_v) G_{10}(u_h, u_v)$	$-50 \le u_h, u_v \le 50$
q_4	$G_{1.2}(u_h, u_v)$	$-7 \le u_h, u_v \le 7$



z (17.11dB) $\hat{y}$ (30.42dB) $z_2 = \hat{y} + \hat{\Delta}$ (40.34dB) $\hat{y}_2$ (41.16dB)

Fig. 3. View of denoising of *Lena* corrupted by q_3 of Table 1 with $\sigma^2 = 0.02$. From left to right: the noisy image (z), result of collaborative hard-thresholding ($\hat{y}$), result of collaborative hard-thresholding combined with its residual (z_2), and result of refiltering z_2 ($\hat{y}_2$). PSNR values in parentheses are for the full images.

which affect both the threshold in hard-thresholding and the attenuation coefficients of the Wiener filter, and thus can be significantly improved by replacing the previous estimates v_i (7) with the exact variances (13).

Unlike $v_{i,2}^{x_R,x_j}$, the variances $v_{i,j}^{x_1,\ldots,x_M}$ cannot be precomputed before the block matching, because they depend on the relative spatial configuration of the entire group: for a typical size of search neighborhoods and number M of blocks in the group, the number of possible configurations is huge. Thus, $v_{i,j}^{x_1,\ldots,x_M}$ are computed online after each group is formed.

2.3.3. Aggregation

After shrinkage, the processed blocks are transformed by $(\mathcal{T}^{3\mathrm{D}})^{-1}$ and aggregated to form an estimate of y. In aggregation, denoised blocks from different groups are weighted based on the residual noise of the shrunk group spectrum; the weight of a group is calculated as the reciprocal of the sum of the products of variances and squared shrinkage attenuation factors [4]. These aggregation weights can be computed from the exact variances (13) used in shrinkage, in place of the conventional variance estimates (7).

2.3.4. Fast implementation

The increased complexity compared to the conventional methods can be seen from the additional global convolution in (12) or the multiplication and an additional Fourier transform in (13) compared to (4). Moreover, (13) has to be recomputed for each group, whereas (7) can be computed once at the beginning of the algorithm and reused for every group. Even using fast algorithms, computing either (12) or (13) for every group is not feasible for any reasonably sized image. To decrease the computational time, the following points may be leveraged:

1. Compute (13) with respect to a smaller Fourier transform with a downscaled PSD;

Table 2. Average PSNR (dB) for denoising of white noise and the four correlated noise types defined in Table 1, over *Barbara*, *Boat*, *Cameraman*, *Hill*, *House*, *Lena*, *Man*, and *Peppers*. In "proposed (refilter)", we apply also the refiltering process described in Section 2.4 after both hard-thresholding and Wiener stages.

	white noise				q_1			q_2			q_3			q_4		
$\mathrm{var}\{\eta\} = \|g\|_2^2$	0.001	0.01	0.02	0.05	0.001	0.01	0.02	0.001	0.01	0.02	0.001	0.01	0.02	0.001	0.01	0.02
proposed	35.82	30.42	28.83	**26.68**	36.05	28.23	25.70	35.39	29.12	27.42	37.41	31.78	30.32	33.36	25.15	22.55
proposed (refilter)	35.82	30.40	28.80	26.66	**36.45**	**28.68**	**26.11**	**36.42**	**30.66**	**28.99**	**42.40**	**41.59**	**41.17**	**33.49**	**25.30**	**22.69**
BM3D	**35.83**	**30.47**	**28.86**	25.79	35.93	26.64	23.24	34.36	25.36	22.46	32.42	21.76	18.67	33.21	24.26	21.44
NLM-C	34.06	28.64	27.02	24.79	30.44	21.25	18.46	30.62	21.71	19.01	29.96	20.47	17.56	19.25	19.19	19.12

2. $\tilde{b}_j^{1\mathrm{D}}$ is sparse, as well as the individual cascaded 1-dimensional FFTs required for the separable computation of $\mathscr{F}\big[\tilde{b}_j^{1\mathrm{D}}\big]$;
3. Since most of the signal is often compacted into few $s_{i,j}^{x_1,\ldots,x_M}$ with small j, we can compute $v_{i,j}^{x_1,\ldots,x_M}$ exactly only for $j \leq J \leq M$ and approximate the rest by exploiting the identity $\sum_{j=1}^{M} v_{i,j}^{x_1,\ldots,x_M} = M v_i$ that holds for an orthonormal $\mathcal{T}^{1\mathrm{D}}$.

These make it possible to achieve runtimes comparable to the original BM3D and BM3D for correlated noise [3, 4] without significant sacrifices in accuracy. We use a 32×32 PSD and $J = 4$ for the experiments in this paper, causing around 45% increase in runtime compared to [4]. This increase is almost entirely due to the online calculation of the variances used for shrinkage and aggregation.

2.4. Inherent limitations of transform-domain block filtering

Even with the exact variances (13), it may not be possible to adequately preserve image details while attenuating noise with certain PSDs. In particular, the used $\mathcal{T}^{2\mathrm{D}}$ may lack directional selectivity, e.g., the 2-D DCT cannot differentiate between diagonal and antidiagonal components. The small size of the blocks further limits the frequency resolution of $\mathcal{T}^{2\mathrm{D}}$. In practice, this results in oversmoothing the estimate.

An estimate of the details lost due to these factors may be obtained by comparing the global FFT spectrum of the residual $\Delta = z - \hat{y}$, where $\hat{y}$ is the denoised image, against the noise PSD:

$$\hat{\Delta} = \mathscr{F}^{-1}[\mathscr{F}[\Delta]\, H[\alpha_\Delta]],$$

where α_Δ is a three-sigma test

$$\alpha_\Delta = \begin{cases} 1 & \text{if } |\mathscr{F}[\Delta]| > 3\sqrt{\Psi} \\ 0 & \text{otherwise,} \end{cases}$$

and H is a morphological dilation filter to smooth the result of the test. A new noisy image z_2 with PSD $\Psi H[\alpha_\Delta]$ is defined as

$$z_2 = \hat{y} + \hat{\Delta}.$$

As demonstrated in Figure 3, by refiltering z_2 with the collaborative filter, we can restore additional details lost when filtering z. Iterative application of this process was not found to provide further benefits.

3. EXPERIMENTS

To assess the performance gain from the improvements described in Section 2.3, we compare denoising results of BM3D with the proposed improvements versus BM3D for correlated noise [4] as well as NLMeans-C [10]. We consider a variety of noise correlation types and strengths on a set of 8 natural test images. The results are reported in Table 2. We can see that the exact variances play a crucial role in enabling an effective denoising of not only correlated noise but also strong white noise. The results obtained with the refiltering described in Section 2.4 are also included in the table, with a few examples displayed in Figure 4: in most cases the improvement is significant, and it is especially dramatic for the diagonal pattern noise q_3, which exemplifies the limitations discussed in Section 2.4.

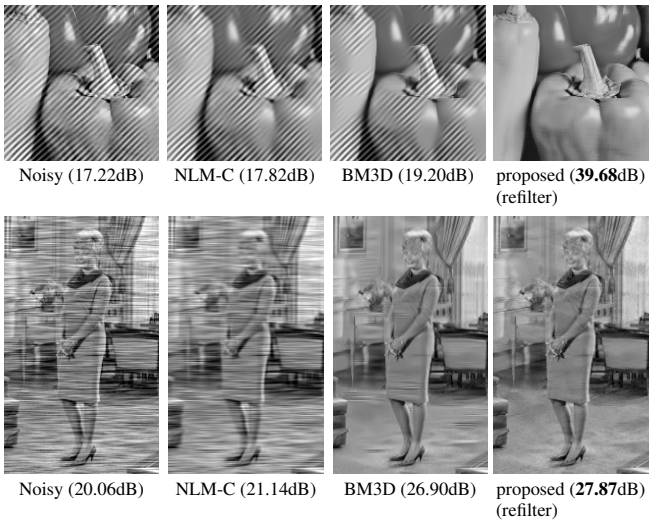


Fig. 4. A view of *Peppers* with diagonal pattern noise (q_3, $\sigma^2 = 0.02$) and a view of *Couple* with horizontal noise (q_1, $\sigma^2 = 0.01$) denoised by BM3D [4], NLMeans-C [10], and BM3D with the proposed improvements. PSNR values in parentheses are for the full images.

4. DISCUSSION AND CONCLUSIONS

We presented a method which allows for both the exact computation and effective approximations of the noise spectrum in nonlocal collaborative transforms by taking into account the relative positions of the matched blocks. The conducted experiments show that the presented method can yield significant improvements in BM3D denoising especially in the case of strongly correlated noise. Specifically, the exact variances allow us to both operate more accurate shrinkage and to avoid matching blocks that are strongly correlated in noise without being similar in the noise-free image. We note that some limitations inherent to the used transforms and small size of the block can be substantially ameliorated by refiltering the image estimate with an added thresholded residual.

Although neural networks are widely regarded as the state-of-the-art in white noise denoising, neural network based methods for correlated noise denoising have not yet caught up. Expensive retraining of the network is required especially in the case of structured noise with visible long range correlation. The proposed method can adapt to varying correlation without any prior training and can thus be utilized within filters and iterative recovery schemes that require online adjustment of the noise model such as [11, 12].

5. REFERENCES

[1] J. S. De Bonet, "Noise reduction through detection of signal redundancy," Rethinking Artificial Intelligence, MIT AI Lab, 1997.

[2] 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.

[3] 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, 2007.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," in *Proc. SPIE Electronic Imaging '08, vol. 6812, no. 681207, San Jose (CA), USA*, 2008, p. 681207.

[5] R. Neelamani, H. Choi, and R. Baraniuk, "ForWaRD: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 418–433, 2004.

[6] A. Foi, K. Dabov, V. Katkovnik, and K. Egiazarian, "Shape-adaptive DCT for denoising and image reconstruction," in *Proc. SPIE Electronic Imaging 2006, Image Processing: Algorithms and Systems V*, vol. 6064, 2006, pp. 203–214.

[7] M. Matrecano, G. Poggi, and L. Verdoliva, "Improved BM3D for correlated noise removal," in *2012 International Conference on Computer Vision Theory and Applications (VISAPP)*, 2012, pp. 129–134.

[8] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 3952–3966, 2012.

[9] A. Rubel, V. Lukin, and K. Egiazarian, "Block matching and 3D collaborative filtering adapted to additive spatially correlated noise," *Proc. International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM), Scottsdale, USA*, 2015.

[10] B. Goossens, H. Luong, A. Pizurica, and W. Philips, "An improved non-local denoising algorithm," in *Proc. 2008 International Workshop on Local and Non-Local Approximation in Image Processing (LNLA 2008)*, 2008, pp. 143–156.

[11] M. Maggioni, E. Sánchez-Monge, and A. Foi, "Joint removal of random and fixed-pattern noise through spatiotemporal video filtering," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4282–4296, 2014.

[12] N. Eslahi and A. Foi, "Anisotropic spatiotemporal regularization in compressive video recovery by adaptively modeling the residual errors as correlated noise," in *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, 2018, pp. 1–5.