Efficiency of texture image enhancement by DCT-based filtering

Aleksey Rubel, Vladimir Lukin, Mikhail Uss, Benoit Vozel, Oleksiy Pogrebnyak, Karen Egiazarian

HAL Id: hal-01240773

https://hal-univ-rennes1.archives-ouvertes.fr/hal-01240773

Submitted on 25 Jan 2016

Efficiency of texture image enhancement by DCT-based filtering

Aleksey Rubel [1], Vladimir Lukin[2], Mikhail Uss[3], Benoit Vozel[4], Oleksiy Pogrebnyak[5], Karen Egiazarian[6]

[1,2]National Aerospace University, Department of Receivers, Transmitters and Signal Processing, Kharkov, Ukraine
[3]National Aerospace University, Department of Radioelectronic Systems Design, Kharkov, Ukraine
[4]University of Rennes 1 – Enssat, Rennes, France
[5]Instituto Politecnico Nacional, Centro de Investigacion en Computacion, Ave. Juan de Dios Batiz S/N, Mexico, D.F.
[6]Tampere University of Technology, Department of Signal Processing, Tampere, Finland

edu.rubel@gmail.com, vladimlukin@yahoo.com, uss_ind@mail.ru
benoit.vozel@univ-rennes1.fr, olek@pollux.cic.ipn.mx, karen.egiazarian@tut.fi

Abstract. Textures or high-detailed structures as well as image object shapes contain information that is widely exploited in pattern recognition and image classification. Noise can deteriorate these features and has to be removed. In this paper, we consider the influence of textural properties on efficiency of image enhancement by noise suppression for the posterior treatment. Among possible variants of denoising, filters based on discrete cosine transform known to be effective in removing additive white Gaussian noise are considered. It is shown that noise removal in texture images using the considered techniques can distort fine texture details. To detect such situations and to avoid texture degradation due to filtering, filtering efficiency predictors, including neural network based predictor, applicable to a wide class of images are proposed. These predictors use simple statistical parameters to estimate performance of the considered filters. Image enhancement is analysed in terms of both standard criteria and metrics of image visual quality for various scenarios of texture roughness and noise characteristics. The discrete cosine transform based filters are compared to several counterparts. Problems of noise removal in texture images are demonstrated for all of them. A special case of spatially correlated noise is considered as well. Potential efficiency of filtering is analysed for both studied noise models. It is shown that studied filters are close to the potential limits.
Keywords: texture analysis, image denoising, DCT-based filtering, visual quality.

1. Introduction

Image texture features are widely exploited in numerous applications of pattern recognition [1], remote sensing [2], similarity search in large databases [3]. In such applications, acquired (original) images are often degraded by noise that, in fact, might be the main destructive factor that prevents solving all related practical problems. Indeed, image fine texture details can be essentially disguised by the noise, and denoising is often a desired stage in the image processing chain. However, alongside with a positive effect of noise removal, the filtering can distort texture images in a larger or smaller extent. Hence, denoising should be performed more carefully in the case of texture images or texture regions of real-life images. For this reason, the noisy data should be analysed to make a decision in the followed filtering stage [1-5].

One can argue that there are many efficient image denoising techniques proposed recently [4-11]. Indeed, in the case of additive white Gaussian noise (AWGN), several filters have demonstrated good performance on different test images [11, 12]. However, practically all of them run into difficulties in preserving texture. This regards partial differential equation based and total variation based denoising techniques [9, 10]. Similar problems also arise for sliding window filters and other modern approaches [4, 13-15]. Methods based on orthogonal transforms, in particular, discrete cosine transform (DCT) and wavelets [15-20] usually perform quite well [4]. The main

reasons to apply DCT are obvious. Firstly, DCT has a good compactness of signal energy or "sparseness". Secondly, DCT can be performed in blocks to be adapted to a local structure of processed images and to noise characteristics [4, 18]. By excluding small amplitude spectrum components of transformed image data, noise removal is attained on one hand. On the other hand, the DCT-based denoising methods are, anyway, not perfect in the sense of texture preservation. While for some textures and noise intensities considerable improvement in output mean square error (MSE) or peak signal to noise ratio (PSNR) is gained, there are also situations when practically no improvement is observed according to conventional criteria as output MSE or PSNR.

Another class of image denoising filters belongs to the group of so-called nonlocal filters [5, 6, 8, 11, 21-23]. These filters exploit self-similarity of image fragments (patches) and can be also equipped by other denoising mechanisms such as 3D DCT-based filtering for collected patches [21]. However, these filters also may result in distortion of image texture details, introduction of artefacts and they may run into difficulties for high-frequency data [12, 18, 19]. Recall that the worst efficiency of denoising has been observed for the most textural images Mandrill (also called Baboon) and Grass among the considered test images in [12, 24, 25]. Moreover, just for these images the potential limit of filtering efficiency for the case of additive white Gaussian noise (AWGN) is the highest (the worst). Thus, texture preservation is problematic even for the most advanced filters.

In this sense, several aspects are worth mentioning. Firstly, practically in all papers dealing with texture preservation, noise is supposed to be AWGN although it can be also spatially correlated [15, 25, 26]. For this reason, it is worth paying more attention to the case of spatially correlated noise. Secondly, there are different approaches to determine potential limits of noise filtering applicable for different groups of filters [12, 18, 26] and noise models, and analysis of these lower bounds could be interesting to understand texture images as well as the problems and limits in texture denoising efficiency. Thirdly, filter performance is mostly considered in terms of conventional criteria (metrics) as output mean square error (MSE) or peak signal to noise ratio (PSNR) [9, 11-23]. Meanwhile, visual quality metrics [7, 25] as well as other statistical parameters as moments [4] are worth using to characterize texture feature preservation by filters. Finally, in practice, it could be fine to predict how denoising can affect an image at hand. In particular, a natural question often arises: "Is denoising really needed for a given image or image fragment?" This question most frequently arises for texture containing images which is the main subject of our study.

First steps forward in attempts to answer this question have been made in the papers [18, 27-29]. In [18, 27] the authors have shown that there is a connection between simple statistics of DCT coefficients in 8x8 blocks and filtering efficiency. Due to this, these statistics can serve for predicting a parameter characterizing filtering efficiency as, e.g., the ratio of the output mean square error (MSE) and AWGN variance. Furthermore, a set of test images used to obtain such dependence did not contain enough texture images to get precise approximation and to reach the final goal of decision-making while denoising.

It has been demonstrated [25, 28] that other parameters describing denoising efficiency such as improvement of PSNR (*IPSNR*) and improvement of PSNR-HVS-M (IPSNR-HVS-M) (where PSNR-HVS-M is visual quality metric [30]) can be predicted as well. Moreover, very accurate

prediction can be achieved if a trained neural network (NN) is applied for approximating the dependence between statistical parameters of a noisy image and a parameter characterizing denoising efficiency [29]. Although textural images were present in the training set, thorough analysis of prediction accuracy just for texture images has not yet been performed.

Thus, there are several prime goals of this paper. First, we would like to present a wide set of simulation data characterizing denoising efficiency of two DCT-based filters [16, 21] for textures with different properties and various intensities of AWGN and spatially correlated noise. Filter performance is characterized not only by conventional criterion (*PSNR*) but also by a visual quality metric. The second goal is to analyse denoising efficiency and compare it to lower bounds that can be determined for a given noise model. The third goal is to consider how well efficiency of denoising can be predicted and how to exploit prediction data.

The structure of this paper is the following. Section 2 describes image/noise models as well as DCT-based filters used in our study. Section 3 discusses how lower bounds of filtering can be determined and what they are for different images. It also describes the considered metrics of denoising efficiency. Section 4 deals with approaches to prediction of filtering efficiency, in particular, of the method that uses a trained NN for this purpose. Some details concerning numerical simulations are presented in Section 5. Analysis of the obtained results is carried out in Section 6. Then, Conclusions follow.

2. Image/noise models and considered filters

In our study, we consider a simple yet conventional image/noise model. It is assumed that an observed noisy one-component (greyscale) image is

$$I_{ij}^{n} = I_{ij}^{tr} + n_{ij} \tag{1}$$

where ij are pixel indices, I_{ij}^{tr} and n_{ij} are true image value and noise, respectively, $i = 1,..., I_{\text{Im}}$ and $j = 1,..., J_{\text{Im}}$, I_{Im} and J_{Im} define image size.

Concerning the true image, it is assumed to be textural since texture preservation is of our main interest in this paper. These can be real-life or artificial textures. Noise is supposed to be zero mean and Gaussian having variance σ_{0}^{2}. Note that we do not restrict ourselves by considering only white noise. Instead, we assume that noise can be both white (independent identically distributed – i.i.d.) and spatially correlated. In the latter case, it can be characterised by 2D autocorrelation function or 2D power spectrum in Fourier or other orthogonal transform basis, e.g., DCT. For spatially correlated noise, we assume the following. First, far correlation does not exist and only neighbour pixels have essential correlation of noise values. Without loss of generality, we assume in simulations that noise has the same main (vertical and horizontal) cross-sections of 2D autocorrelation function. A more important assumption is that noise spatial correlation properties are supposed to be a priori known [7] or pre-estimated with an appropriate accuracy [15]. This allows taking them into account at the filtering stage.

In this paper, two DCT-based filters are considered. Before giving their brief description, let us explain why they have been chosen for our analysis. The BM3D filter [8, 21] is considered to be a state-of-the-art for suppressing AWGN. This filter is not especially suited for processing texture

images but it provides practically the best results for many test images including such textural test images as Baboon and Grass [10, 12] according to conventional quality metrics (output MSE or PSNR) and visual quality metrics, e.g., SSIM [31] – see data in [10]. Besides, BM3D has various modifications including those for spatially correlated noise [32, 33].

The sliding DCT-based filter [7, 16, 43], which is a particular case of BM3D, is much simpler and faster than BM3D. Meanwhile, this filter is able to preserve texture well enough [4] and performs close to the BM3D and other advanced filters [24], especially for texture images [25]. Besides, the standard DCT-based filter can be also easily modified to take into account available information on properties of spatially correlated noise, namely, normalised DCT spectrum in 8x8 pixel blocks $W_{norm}(k,l)$ where $k,l = 0,...,7$ are indices and $k = l = 0$ relate to a direct current (DC) component.

BM3D filter exploits two denoising mechanisms [21]. First, patches (blocks) of size 8x8 pixels most similar to each given (reference) block are found. Then, these patches are collected together and the obtained 3D array (usually of size $8x8x2^n$) is transformed to a spectral domain using separable 2D DCT transforms and vertical (1D) Haar transform.

Search for similar blocks and DCT-based denoising are two main operations although there are other auxiliary operations used to improve the filter performance (see [8, 21]). The modification of BM3D for spatially correlated noise removal [33] has two differences from the original one. Firstly, the Bray-Curtis distance [34] is applied for similarity search; it is defined as

$$d(Q,R) = \frac{\sum_{i=1}^{N} |Q_i - R|}{\sum_{i=1}^{N} |Q_i + R|}, \qquad (2)$$

where R denotes features of reference, Q are query features, N is a number of features. As features, DCT coefficients in compared blocks are used ($N = 64$). Secondly, frequency-dependent thresholds equal to $T(k,l) = \beta\sigma_0\sqrt{W_{norm}(k,l)}$, $k,l = 0,...,7$ (the recommended $\beta = 2.6$ [7, 16, 23] or slightly smaller [35]) are used in hard thresholding of DCT coefficients. These two modifications have allowed reaching about 2 dB improvement of output *PSNR* and *PSNR-HVS-M* compared to the direct application of the conventional BM3D filter [33].

The sliding DCT-based filter performs in DCT blocks of fixed size [43] where the size 8x8 pixels is a good choice [4, 18]. For each block position, direct DCT, hard thresholding of the obtained DCT coefficients and inverse DCT are performed. Next, for each image pixel, the obtained filtered values are averaged using overlapping blocks to which this pixel belongs. Full overlapping of the blocks (where neighbour blocks are shifted by only one pixel with respect to each other) produces the best performance in terms of most quantitative criteria (although faster processing is reached for partial overlapping of the blocks) [16]. The same standard hard threshold is applied for all spatial frequencies $T(k,l) = 2.6\sigma_0$, $k,l = 0,...,7$ in the case of AWGN but it should be frequency dependent ($T(k,l) = 2.6\sigma_0\sqrt{W_{norm}(k,l)}$, $k,l = 0,...,7$) for spatially correlated noise [7]. It is possible to vary a little the thresholds using a factor value slightly smaller than 2.6 [4] to improve the filter performance. However, the benefit of using β specially adapted to a given image and noise

intensity is not large compared to the case of using the fixed $\beta = 2.6$. Besides, adaptation mechanism is still unknown. Thus, all the data below (see Section 6) are obtained for $\beta = 2.6$ for both BM3D and sliding DCT-based filter.

3. Lower bounds of filtering efficiency

Researchers dealing with image denoising and filter design are well aware that performance of any filter depends upon properties of the true image. P. Chatterjee and P. Milanfar [12] provided lower bound of image denoising for nonlocal filtering under conditions that noise-free image is available, noise is additive white and its variance is known a priori. Analysis has been carried out in terms of output MSE (although it could be easily interpreted in terms of output PSNR).

The considered approach to determine the lower bounds has several limitations and shortcomings: it requires noise-free images be available (later this shortcoming has been overcome in [36]), an underlying noise has to be white and all the results in [12] are for Gaussian noise only.

Despite of this, the results obtained in [12] are very useful. In particular, it has been shown that a potential (limit) output mean square error (MSE_{out}^{pot}) [36] for a given noise variance σ_0^2 can vary in very wide limits. The ratio $MSE_{out}^{pot}/\sigma_0^2$ can be from 0.01 (and even less) for simple structure (cartoon) test images and high noise intensity to about 0.9 (and even higher) for highly textural test images corrupted by moderate and low intensity noise. This shows that there can be practical situations when filtering is in fact, useless, and being applied, it can lead to texture smoothing as shown in the paper [10].

Another important conclusion in [12] is that existing advanced filters practically reach the lower bound of filtering efficiency for textural images if noise is intensive, i.e., if $\sigma_0^2 \geq 100$. Although noise of such intensity can be present in real-life images, we are more interested in less intensive noise which is more typical in practice and is still visible and annoying in observed images [7]. Besides, results in [12] are presented for only two really textural images (Mandrill and Grass) whilst a more thorough analysis is needed for textural images with a wider range of texture properties variation.

Note that the approach in [12] to determine denoising potential efficiency is not unique. Another approach, which is especially suited for texture images, is based on their description by the model of fractal Brownian motion (fBm) [26, 37]. The fBm model [38] has been found suitable to describe a wide range of natural textures, especially as they appear on remote sensing images. Recall that fBm model is described by only two parameters: Hurst exponent that relates to fBm-texture roughness (H varies from 0 for rough texture to 1 for smooth) and variance of fBm increments on unit distance σ_x^2 that describes the fBm amplitude. To provide an idea of what kind of textures can be represented by fBm-model, two examples of simulated textures are given in Fig. 1 for $H = 0.3$ and $H = 0.7$.
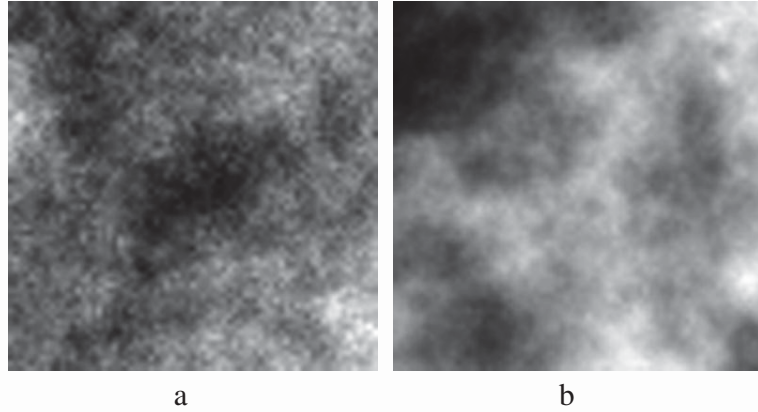
Fig. 1. Examples of isotropic fBm textures. $H = 0.3$ (a) and $H = 0.7$ (b)

To obtain the MSE filtering lower bound based on fBm-model, image filtering is viewed preferentially as a parameter estimation problem. The image is tiled into non-overlapping scanning windows. Pixels within each scanning window (SW) form a noisy observation of $\mathbf{Y}$. Each k-th element of $\mathbf{Y}$ has coordinates (i_k, j_k) with respect to SW centre. The central pixel has index k_0 and coordinates (i_{k0}, j_{k0}). The problem to solve is thus to estimate the intensity (true value) of the central pixel $x_0 = x(i_{k0}, j_{k0})$ together with the texture parameters (σ_x, H). The full parameter vector is, therefore, $\mathbf{\theta} = (\sigma_x, H, x_0)$. Having stated the previous definitions, the Cramér–Rao lower bound (CRLB) of estimate $\hat{\mathbf{\theta}}$ in terms of MSE can be derived based on the corresponding Fisher matrix $\mathbf{I}(\mathbf{\theta})$:

$$\mathbf{I}(\mathbf{\theta}) = \begin{pmatrix} I_{\sigma_x} & I_{H\sigma_x} & 0 \\ I_{H\sigma_x} & I_H & 0 \\ 0 & 0 & I_{x_0} \end{pmatrix}, \tag{2}$$

where $I_{x_0} = \sum_{k_1}\sum_{k_2} \mathbf{R_Y}^{-1}(k_1, k_2)$, $I_{\sigma_x} = \frac{1}{2} tr\left(\frac{\partial \mathbf{R_Y}}{\partial \sigma_x} \mathbf{R_Y}^{-1} \frac{\partial \mathbf{R_Y}}{\partial \sigma_x} \mathbf{R_Y}^{-1}\right)$, $I_{H\sigma_x} = \frac{1}{2} tr\left(\frac{\partial \mathbf{R_Y}}{\partial \sigma_x} \mathbf{R_Y}^{-1} \frac{\partial \mathbf{R_Y}}{\partial H} \mathbf{R_Y}^{-1}\right)$,

$I_{\sigma_x} = \frac{1}{2} tr\left(\frac{\partial \mathbf{R_Y}}{\partial \sigma_x} \mathbf{R_Y}^{-1} \frac{\partial \mathbf{R_Y}}{\partial \sigma_x} \mathbf{R_Y}^{-1}\right)$, $\mathbf{R_Y}$ is data sample $\mathbf{Y}$ correlation matrix. Then, the filtering lower bound on MSE (MSE_{fBm}) is finally obtained as the corresponding element of the inverse matrix $(\mathbf{I}(\mathbf{\theta}))^{-1}$: $MSE_{fBm} = 1/I_{x_0}$.

Note that $\mathbf{R_Y} = \mathbf{R_X} + \mathbf{R}_n$, where $\mathbf{R_X}$ is the correlation matrix of image noise-free texture and $\mathbf{R}_n$ is the noise correlation matrix. For the fBm-field correlation matrix, elements of the $\mathbf{R_X}$ matrix are given in [38] as

$$\mathbf{R_X}(k_1, k_2) = 0.5\sigma_x^2 \left[\left(i_{k_1}^2 + j_{k_1}^2\right)^H + \left(i_{k2}^2 + j_{k2}^2\right)^H - \left(\left(i_{k_1} - i_{k_2}\right)^2 + \left(j_{k_1} - j_{k_2}\right)^2\right)^H \right]. \tag{3}$$

Noise correlation matrix can take different forms. Among them, Gaussian and exponential noise are commonly used. Here we model correlated noise by applying box filter (filter support is $n_s \times n_s$ with $n_s = 3$) to the uncorrelated noise. In this case, the matrix $\mathbf{R}_n$ elements are found as follows:

$$\mathbf{R}_n(k_1, k_2) = \begin{cases} \sigma_n^2 \left(1 - \left|i_{k_1} - i_{k2}\right|/n_s - \left|j_{k_1} - j_{k2}\right|/n_s + \left|i_{k_1} - i_{k2}\right| \cdot \left|j_{k_1} - j_{k2}\right|/n_s^2\right), \left|i_{k_1} - i_{k2}\right| \le n_s, \left|j_{k_1} - j_{k2}\right| \le n_s \\ 0, \text{otherwise} \end{cases}.$$

(4)

The value of MSE_{fBm} for the whole image is found according to the following algorithm. For each p-th SW, fBm-model parameters are estimated (this can be done using either noise-free or noisy image, the noise parameters might be known or preliminary estimated). This allows to calculate correlation matrix $\mathbf{R}_Y$, Fisher matrix element I_{x_0} and $MSE_{fBm}(p)$ bound for this single SW. Finally, MSE_{fBm} is found by averaging $MSE_{fBm}(p)$ over all image scanning windows.

Note that the bound MSE_{fBm} has its own limitations that are important for our analysis. First, fBm-model is not adequate for all kinds of image textures. Edges, anisotropic textures, highly non-Gaussian textures do not match to this model. In this case, previous MSE_{fBm} bound overestimates potential filtering MSE. Second, MSE_{fBm} bound does not reflect the gain due to image filtering in overlapping scanning windows (that is the case for DCT and BM3D filters). This also leads to overestimated MSE_{fBm} values. Third, MSE_{fBm} bound is local and does not take into account non-local image redundancy (used by BM3D filter). Therefore, MSE_{fBm} bound is a pessimistic estimate of potential filtering MSE achieved by non-local filtering methods.

Although the approaches [12] and [26] considerably differ from each other, the main conclusions that follow from them basically coincide. Meanwhile, it could be interesting to compare the obtained lower bounds, at least, for the case of AWGN. An obvious advantage of the approach [26] is that it allows determining MSE_{out}^{pot} for the case of spatially correlated noise.

4. Neural network based prediction of filtering efficiency

Clearly, it is good to know potential efficiency of filtering in advance or to pre-estimate it. The Chatterjee and Milanfar approach [36] allows doing this quite accurately without having a noise-free image, but intensive computations are required. However, practically reachable MSE_{out} can differ from the potential error MSE_{out}^{pot} and this difference can be significant [36]. Then, it is desirable to have estimates (prediction) of denoising efficiency for existing filters. The main idea of such a prediction is the following [27-29]. Suppose that it is possible to quickly estimate one or several statistical parameters of a given image that are quite strictly connected with a parameter (or parameters) that characterize denoising efficiency. If such a dependence or approximation exists and it has been obtained in advance, then it is possible to calculate this statistical parameter(s) and to determine parameter(s) that characterize denoising efficiency. One possible action then is to

undertake a decision on whether or not it is worth applying filtering. Another decision could be to change filter settings compared to the conventional (recommended) ones.

Here we do not concentrate on decision undertaking. Instead, we show that there are statistical parameters and metrics that describe denoising efficiency for which essential correlation exists. Moreover, this correlation allows carrying out a quite accurate prediction. One possibility to characterize denoising efficiency is to use the ratio MSE_{out}/σ_0^2 [27], where

$$MSE_{out} = \sum_{i=1}^{I_{\mathrm{Im}}} \sum_{j=1}^{J_{\mathrm{Im}}} \left(I_{ij}^f - I_{ij}^{tr}\right)^2 / \left(I_{\mathrm{Im}} \cdot J_{\mathrm{Im}}\right), \tag{5}$$

I_{ij}^f denotes filtered image value in ij-th pixel. Note that the ratio MSE_{out}/σ_0^2 is strictly connected with another parameter called improvement of PSNR (*IPSNR*) expressed in dB and calculated as

$$IPSNR = 10\log_{10}\left(\sigma_0^2/MSE_{out}\right). \tag{6}$$

It has been stated many times [30, 31] that standard output MSE or PSNR do not adequately characterise visual quality of images, in particular, images after filtering [31]. Then, other metrics, the so-called visual quality indices or human vision system (HVS) metrics, are to be employed. One of them is $PSNR - HVS - M$ [30], defined as

$$PSNR - HVS - M = 10\log_{10}\left(255^2/MSE_{HVS-M}\right), \tag{7}$$

where MSE_{HVS-M} is a specific for human vision system MSE. Algorithm of MSE_{HVS-M} calculation takes into account two peculiarities – different sensitivity of HVS to distortions in low and high spatial frequencies as well as masking effects which are typical for texture (see [30] for more details). Then, it is possible to calculate the improvement of $PSNR - HVS - M$ which is the difference of *PSNR-HVS-M* values after and before filtering, expressed in dB:

$$\begin{aligned} IPSNR - HVS - M &= 10\log_{10}\left(255^2/MSE_{HVS-M}\right) - 10\log_{10}\left(255^2/MSE_{HVS-M}^{noisy}\right) \\ &= 10\log_{10}\left(MSE_{HVS-M}^{noisy}/MSE_{HVS-M}\right), \end{aligned} \tag{8}$$

where MSE_{HVS-M}^{noisy} denotes a specific for human vision system MSE between noisy and original image.

Thus, there are, at least, three parameters able to characterize denoising efficiency quantitatively. A question is what statistical parameters are able to jointly characterize image and noise properties that influence denoising efficiency. Originally, dependence between filtering efficiency and some statistics of DCT coefficients have been noticed in [18], where the AWGN case was studied. The statistical parameters considered later in [27] were probabilities $P_{2\sigma}$ and $P_{2.7\sigma}$. The probability $P_{2\sigma}$ is an estimate of probability that absolute DCT coefficient values in 8x8 blocks of a

given image do not exceed $2\sigma_0$. In turn, the probability $P_{2.7\sigma}$ is an estimate that absolute DCT coefficient values in 8x8 image blocks exceed the threshold of $2.7\sigma_0$. Direct current DCT coefficient is not used in this probability estimation. Noise standard deviation σ_0 is assumed to be known a priori or accurately pre-estimated.

Following an analysis in [27-29], it is better to use $P_{2\sigma}$ than $P_{2.7\sigma}$ (prediction is more accurate). Moreover, it is possible to employ other probabilities as, e.g., $P_{0.5\sigma}$, which is an estimate of probability that absolute DCT coefficient values in 8x8 blocks of a given image do not exceed $0.5\sigma_0$ [29]. Below we concentrate on considering mainly $P_{2\sigma}$ as the most studied statistic.

Here, the estimates of probabilities are considered because each image has a limited size and only one realization of the noise is available. It is not necessary to use in estimation all possible positions of the image blocks. Usually it is enough to have 500 non-overlapping or partly overlapping blocks to obtain an estimate, which is accurate enough for our purpose [29]. The reasons for this will become clear later. Note that $P_{2\sigma}$ can be, in fact, regarded as the mean of local estimates obtained in blocks. Alongside with the mean, it is also quite easy to calculate other statistics such as median, variance, and higher moments or statistical parameters [29].

Here it is important to mention that estimation of the aforementioned probabilities can be made quickly enough because DCT in 8x8 pixel blocks is a fast operation and is performed for not all blocks. Therefore, the probability estimation is much faster than the standard DCT-based filtering itself.

Analytical descriptions of the dependences between MSE_{out}/σ_0^2 (or *IPSNR* or *IPSNR-HVS-M*) and abovementioned probabilities have been obtained as follows. The main operation is to form a scatter-plot where vertical axis corresponds to MSE_{out}/σ_0^2 (or *IPSNR* or *IPSNR-HVS-M*) and horizontal (argument) axis relates to one of the probabilities. To get a better picture, Fig. 2 presents two examples of scatter-plots – *IPSNR* vs $P_{2\sigma}$ and *IPSNR-HVS-M* vs $P_{2\sigma}$ (AWGN case). Compactness of the scatter-plot points clearly indicates that there are strict connections between the considered parameters. The correlation is, obviously, larger for *IPSNR* vs $P_{2\sigma}$.
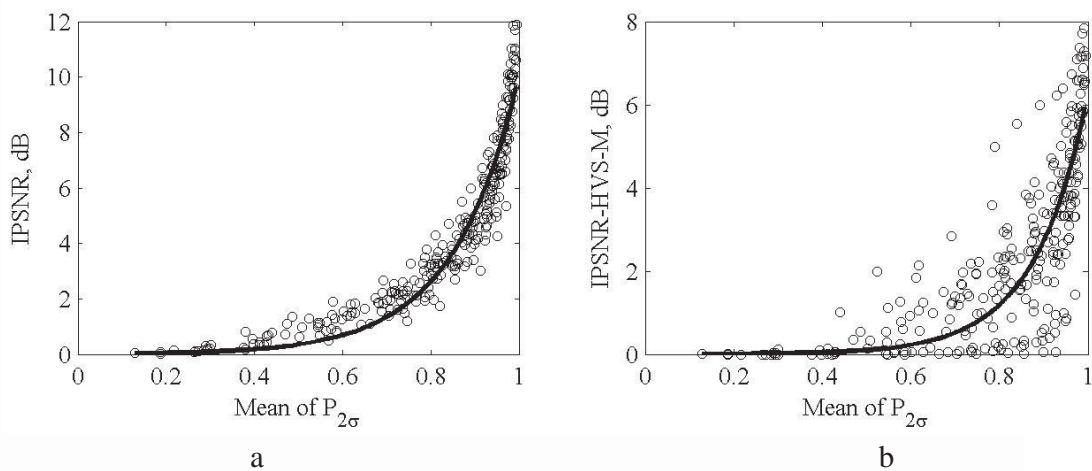

Fig. 2. Scatter-plots of *IPSNR* vs $P_{2\sigma}$ (a) and *IPSNR-HVS-M* vs $P_{2\sigma}$ (b).

Obtained approximations for DCT-based filter (shown above) are the following:

$$IPSNR_{pred} = 0.012 \cdot \exp(6.7 \cdot P_{2\sigma}), \tag{9}$$

$$IPSNR - HVS - M_{pred} = 0.002 \cdot \exp(8.3 \cdot P_{2\sigma}). \tag{10}$$

The scatter-plots in Fig. 2 have been obtained using a large number of test images corrupted by AWGN with variances from 4 to 625 and processed by the sliding DCT-based filter. More scatter-plots are presented in [27, 28] and they are more thoroughly studied there. Here, it is more important to stress the following. First, dependences can be both increasing or decreasing but they are monotonous. Second, there are areas of $P_{2\sigma}$ (smaller than 0.5) where it seems not expedient to perform filtering due to its low efficiency. Third, scatter-plots can be more or less clustered. In the latter case, prediction based on using only one statistic of DCT coefficients is less accurate.

The proposed filtering efficiency prediction starts with curve fitting into a scatter-plot. Then, estimating a selected statistical parameter (e.g., $P_{2\sigma}$) and having a fitted curve (e.g., its analytical expression) for a chosen metric (e.g., *IPSNR* (6)) one easily obtains the predicted estimate of denoising efficiency. Small errors in estimation of $P_{2\sigma}$ or $P_{2.7\sigma}$ due to a limited number of analysed image blocks lead to small errors in prediction. These errors can be ignored since points in a scatter-plot obtained for particular images stand more or less aside the fitted curve. In other words, prediction based on a parameter describing a noisy image in aggregate cannot be absolutely accurate.

The accuracy of the proposed prediction estimate depends on many factors, namely, what statistic is calculated in blocks and then processed, what metric is approximated, how a scatter-plot has been obtained (what test images and noise variance values have been used), what function has been used for fitting, etc. Some conclusions and recommendations in designing prediction methods and tools can be found in the papers [27-29] and they are the following. Properties of test images and noise variance values shall vary in wide limits in order to have arguments (values of $P_{2\sigma}$ or $P_{2.7\sigma}$) in wide limits covering values of these probabilities that can be encountered in practice. Not only mean probabilities characterize denoising efficiency. Parameters that describe filtering efficiency are connected with distribution of local estimates (in blocks) of the considered probabilities [28, 29]. Then, it is possible to employ not one but few parameters derived from the local estimates:

$$\hat{P}_{2\sigma} = \sum_{k=0}^{7} \sum_{l=0}^{7} \delta(n,k,l)/63, \tag{11}$$

where for AWGN $\delta(n,k,l) = \begin{cases} 1, \text{if } |D(n,k,l)| \leq 2\sigma_0 \\ 0, \text{ otherwise} \end{cases}$, n is the index of the considered image block,

$D(n,k,l)$ is kl-th DCT coefficient calculated for n-th block. If noise is spatially correlated, one has

$$\delta(n,k,l) = \begin{cases} 1, \text{if } |D(n,k,l)| \leq 2\sigma_0 \sqrt{W_{norm}(k,l)} \\ 0, \text{ otherwise} \end{cases}.$$

For the obtained set of local estimates $\hat{P}_{2\sigma}(n)$ of the probability $P_{2\sigma}$, it is possible to calculate different statistical characteristics and to use them together for a prediction. These statistics can be jointly used via pattern recognition. For this purpose, we propose to employ a simple feed-forward neural network (NN), the multilayer perceptron [39].

The results presented in the paper [29] show that a good set of input parameters includes mean, variance, skewness and kurtosis of the local estimates. Therefore, the proposed NN structure has four inputs, ten neurons in one hidden layer (chosen empirically), and one output neuron. The NN has been trained using Levenberg-Marquardt optimization algorithm for 128 test images distorted by noise with ten levels.

Goodness of fit R^2 [40] was used to characterize approximation ability of the NN. Recall that the values of R^2 larger than 0.9 are considered good. For *IPSNR* (6), prediction is already good even if one uses only $P_{2\sigma}$ (R^2 is about 0.95) and it increases by about 0.011 due to a use of the trained NN. For *IPSNR-HVS-M* (8), the scatter-plot based prediction using $P_{2\sigma}$ is considerably worse (R^2 is about 0.83). However, the use of the trained NN as an approximator allows providing sufficiently better R^2 (about 0.95 for the DCT-based filter and about 0.91 for BM3D [29]).

5. Numerical simulation details

We have carried out simulations for twelve test images, all of them are either textural or contain mainly textural regions. These images are presented in Fig. 3. Eight of these test images taken from the texture image database (USC-SIPI Image Database) have been earlier analysed in the paper [25] and they represent textures of different grain size and contrast (these images have indices from 1 to 8 assigned to all test images for convenience of their further analysis). For this reason, we have entered into analysis four other textural images (some of which are quite popular and have been already used in different papers) with indices from 9 to 12. The reason of adding them to the initial set was to consider images that are natural and have not anisotropic textures.

Three variance values have been considered for both i.i.d. (white) and spatially correlated noise cases. Noise variance equal to 25 (low level) corresponds to practically invisible white noise but noticeable spatially correlated noise [7], which is desirable to suppress. Noise variance equal to 100 can be referred to middle intensity noise. This noise is seen in visual inspection for both white and spatially correlated cases. Finally, the third value of the noise variance (225) relates to the case of intensive noise, which can be considered annoying for white and, especially, spatially correlated cases. Therefore, we study different practical situations that can be met in different applications.

To each test image, ten realizations of white and spatially correlated noise with aforementioned values of noise variance have been added. Thus, for each test image, there were six sets of its noisy versions. For each noisy image, local probabilities $\hat{P}_{2\sigma}(k), k = 1,...,K_{bl}$ have been obtained using all possible positions of blocks (if $I_{Im} = J_{Im} = 512$,

$K_{bl} = (512 - 7) \cdot (512 - 7) = 255025$). For each set, processing has been carried out by the DCT-based filter with fixed thresholds or with frequency-dependent thresholds for white and spatially correlated noise, respectively.
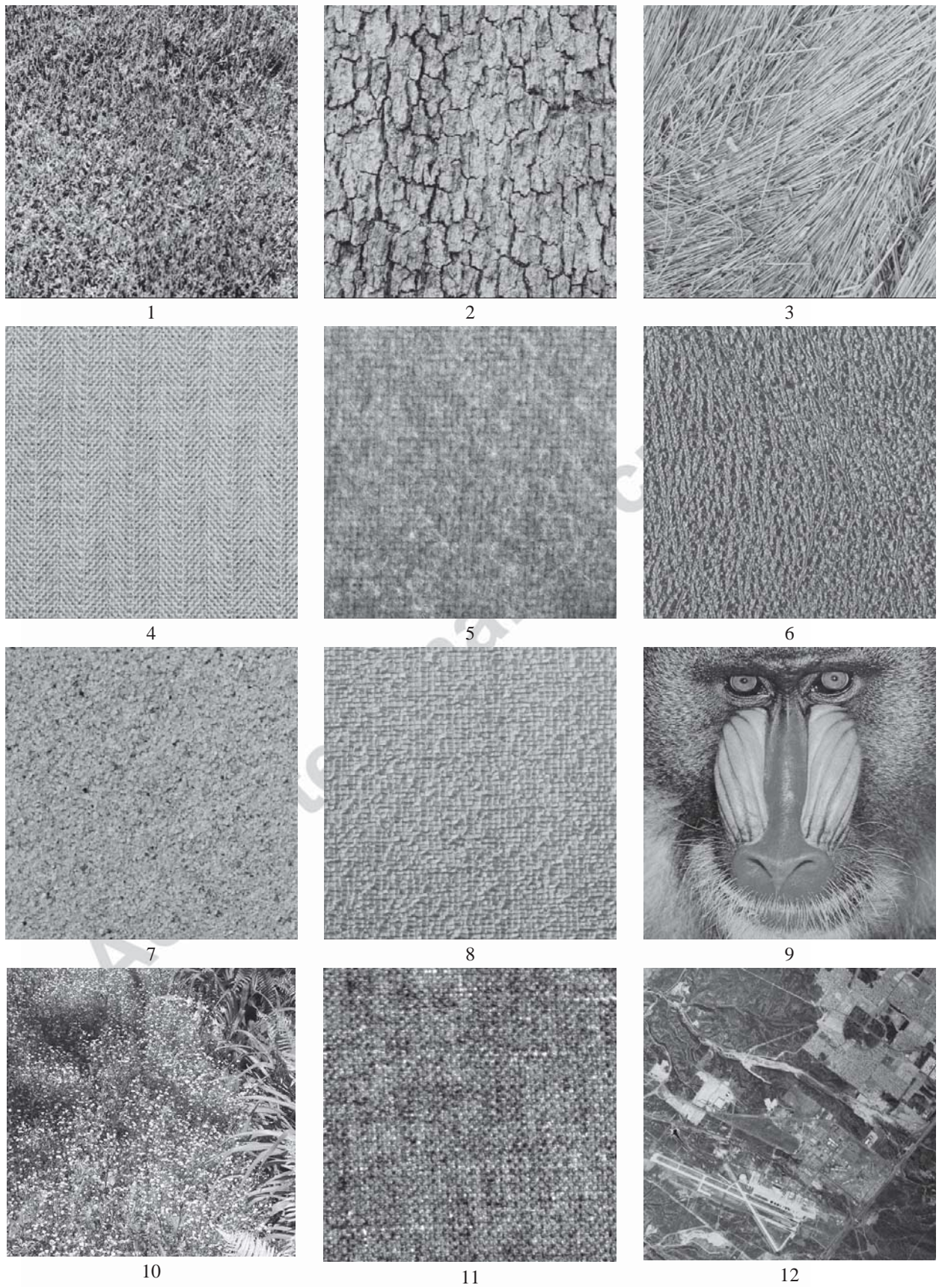
Fig. 3. Test texture images and their indices

Additionally, processing has been carried by the BM3D filter for AWGN or its modification (as described above) for spatially correlated noise.

For the denoised image, the following parameters have been determined:

- MSE_{out}, then $PSNR$ and $IPSNR$ (6);

- MSE_{HVS-M}, then $PSNR-HVS-M$ (7) and $IPSNR-HVS-M$ (8);

- predicted values $IPSNR_{pred}$ (9) and $IPSNR-HVS-M_{pred}$ (10) as it is described in Section 4;

- Milanfar&Chatterjee lower bound MSE_{out}^{pot} [12] (only for AWGN, see the first part of Section 3) that has been recalculated to

$$IPSNR^{pot} = 10\log_{10}\left(\sigma_0^2/MSE_{out}^{pot}\right);$$

- fBm-based lower bound MSE_{out}^{fBm} (for both AWGN and spatially correlated noise, see the second part of Section 3) that has been recalculated to

$$IPSNR^{fBm} = 10\log_{10}\left(\sigma_0^2/MSE_{out}^{fBm}\right).$$

Since MSE_{out}^{pot} and MSE_{out}^{fBm} are recalculated to $IPSNR^{pot}$ and $IPSNR^{fBm}$, respectively, we will further refer to an upper bound while considering $IPSNR^{pot}$ and $IPSNR^{fBm}$.

The obtained parameters have been later averaged for the considered noise realizations. From one to another realization, the values of the metrics $IPSNR$ (6), $IPSNR-HVS-M$ (8), $IPSNR_{pred}$ (9), and $IPSNR-HVS-M_{pred}$ (10) changed by no more than 0.2 dB. Thus, after averaging, we have accurate estimates of the considered metrics for providing reliable conclusions.

6. Analysis of the obtained results
6.1 AWGN case

Let us start from considering a case of AWGN. Simulation data for the smallest value of noise variance (25) is given in Fig. 4 where six values ($IPSNR$ and $IPSNR_{pred}$ for the DCT based filter, $IPSNR$ and $IPSNR_{pred}$ for the BM3D filter, $IPSNR^{pot}$ and $IPSNR^{fBm}$) are presented for each test image. Analysis of data allows concluding the following:

1) For all considered test images, analysed metric values are quite small, usually less than 3 dB.

2) There are no obvious differences between all six parameters for most test images. The exceptions are the test images ## 8, 9, and 12 where $IPSNR^{pot}$ is sufficiently larger than other values; for these images, either texture is quite regular (test image # 8) or there are quasi-homogeneous regions (## 9 and 12), this allows expecting benefits from finding similar patches used in nonlocal denoising;

3) Values of $IPSNR$ for the DCT-based and BM3D filters differ only a little. The differences are less than 0.5 dB for all test images. Therefore, the more complex BM3D filter does not produce obvious benefits compared to the standard DCT-based filter in the considered practical situation;

4) Moreover, *IPSNR* for both the DCT-based and BM3D filters is often smaller than 1 dB showing that, in fact, there is no reason to carry out denoising by any filter. These are examples when denoising is useless.

5) Prediction performs well enough and the corresponding values of *IPSNR* and $IPSNR_{pred}$ differ by less than 0.7 dB. Hence, prediction is able to indicate when there is no reason to perform denoising.

6) The values of $IPSNR^{pot}$ are always larger than the corresponding $IPSNR^{fBm}$ where the latter upper limit values practically coincide with $IPSNR_{pred}$ for the sliding DCT-based filter.
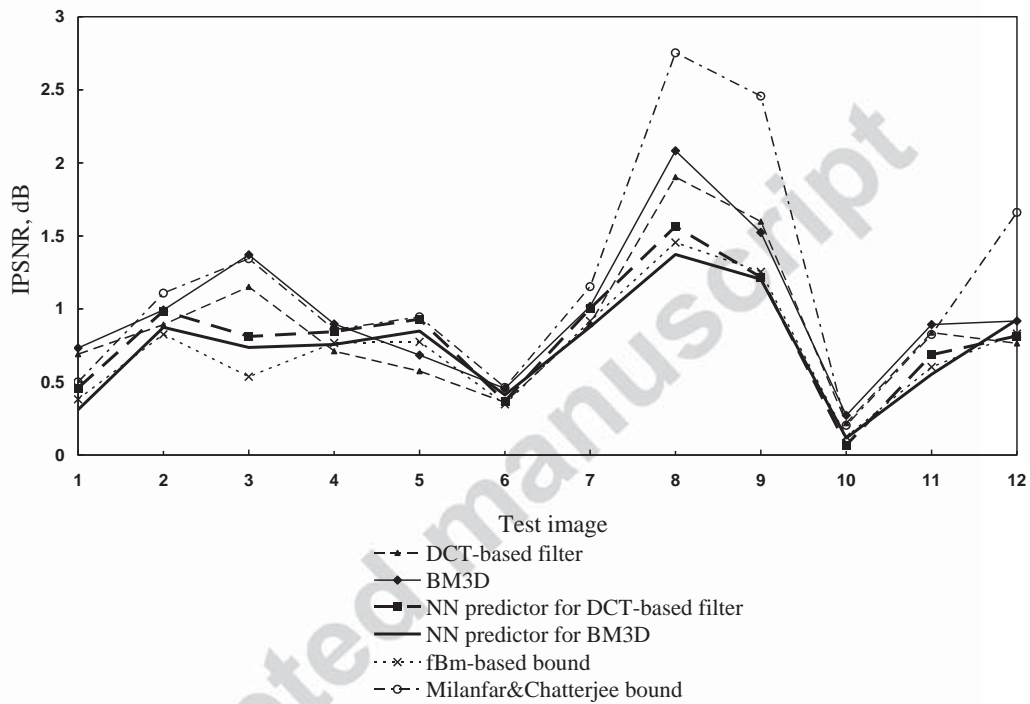

Fig. 4. *IPSNR* metric, predicted and upper bound values for AWGN with $\sigma_0^2 = 25$.
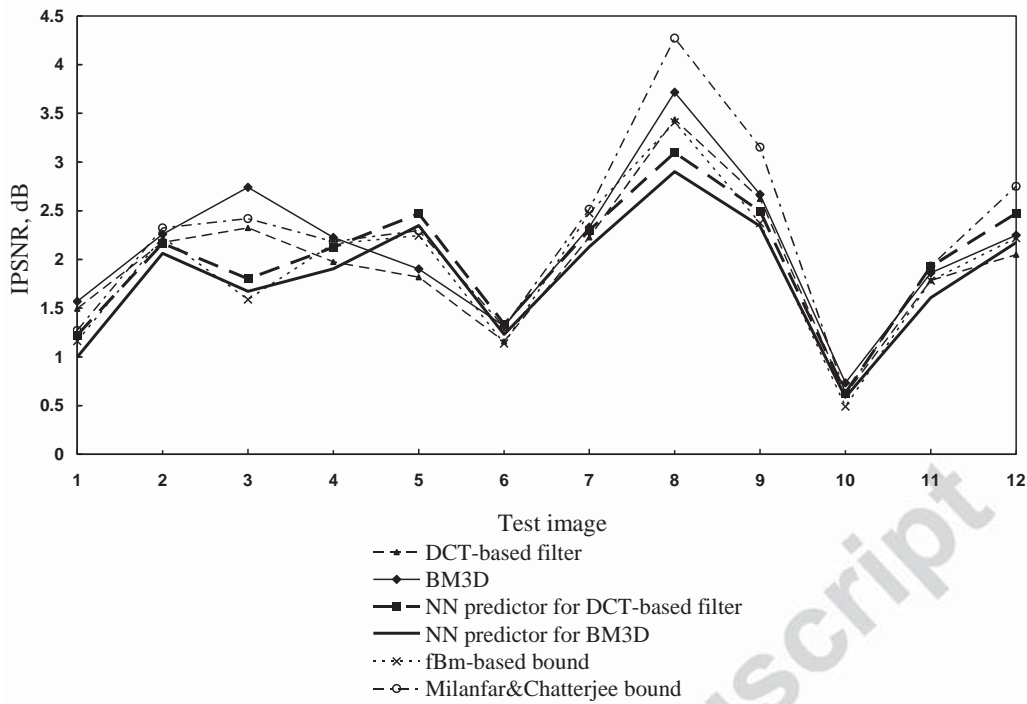
Fig. 5. *IPSNR* metric, predicted and upper bound values for AWGN with $\sigma_0^2 = 100$.
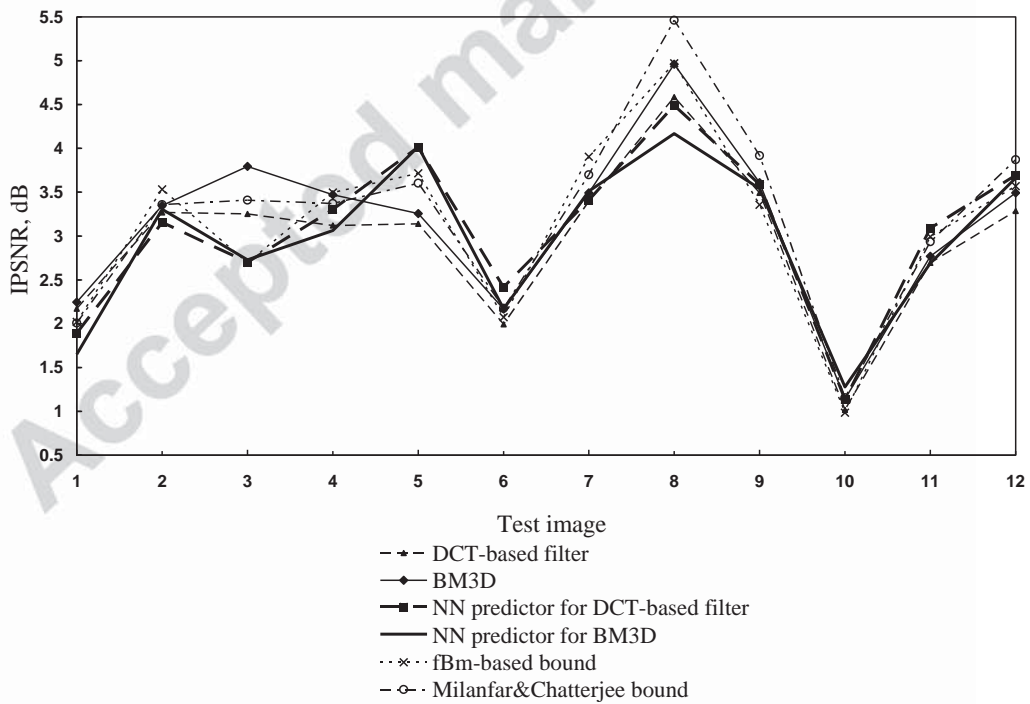

Fig. 6. *IPSNR* metric, predicted and upper bound values for AWGN with $\sigma_0^2 = 225$.

Fig. 5 represents the obtained data for middle level AWGN. Conclusions that can be drawn are practically similar to those given above. The difference is that practically all values

corresponding to each other in Figures 4 and 5 have increased. This confirms the known fact that it is easier to suppress a more intensive noise. It becomes, in general, more reasonable to apply denoising although there are still textures (test images ## 1, 6, 10) for which the positive effect due to filtering is quite small. There appear images for which data for the BM3D filter and MSE_{out}^{pot} practically coincide (e.g., test images ## 1-3).

The results obtained for the largest variance (225) are depicted in Fig. 6. Efficiency of denoising continues to increase. For most images, $IPSNR_{pred}$ for both the standard DCT-based and BM3D filters becomes not worse than 3 dB. Such improvement is often visible if one compares the original and processed image. There appear test images for which BM3D sufficiently outperforms the standard DCT-based filter. This is, in particular, the test image # 3, which has a regular texture. Due to this, BM3D finds similar blocks and exploits this similarity in denoising. Meanwhile, denoising efficiency for the test image # 10 is still low despite the high noise level.

One might expect that other filters perform better than these two considered above. To check this, we have analysed performance for a set of different filters. First of all, we have tested the Guided filter [22]. It performs almost perfectly if there are available noise-free patches that can be met in a filtered image. If such patches are absent (as this often happens in practice), the filter efficiency drops radically. Because of this, below we consider the following five filters. First, these are two wavelet based filters [19, 20] that we denote in Tables 1 and 2 as Translation invariant wavelet shrinkage (TI-WS) and BLS-GSM, respectively. Both filters exploit data on noise variance or noise standard deviation (STD) supposed to be a priori known. A third filter, a nonlocal means filter (NLM) [23] is a good representative of non-local denoising techniques. As a fourth filter, we analyse the BM3D filter [21] data which are given in Figures 4-6. This allows indirect comparison of other filters to data in Figures 4-6. Finally, the Geodesic path filter [14] (further denoted as Geodesic) was originally designed for processing colour images. But it also has a single-channel (2D) version (kindly provided to us by Prof. Bogdan Smolka) [14]. We also present data for noisy images subject to further processing.

The main conclusion is the following. No one of the considered counter-parts performs considerably better than the BM3D filter. There are few cases when Portilla's filter [20] (BLS-GSM) slightly outperforms BM3D but usually the BM3D filter is more efficient. The results for NLM and TI-WS filters are even worse, and the geodesic path filter competes only for large values of noise variance. Thus, the task of texture filtering is difficult for a wide set of the considered denoising techniques.

Consider now the metric $IPSNR-HVS-M$. Recall that we can determine $IPSNR-HVS-M$ for two filters and two predicted values for $IPSNR-HVS-M$. The results for $\sigma_0^2 = 25$ (STD=5) are presented in Fig. 7. The main observations are the following. All values of $IPSNR-HVS-M$ both derived and predicted are close to zero for all analysed twelve images. The results for the BM3D filter are slightly better than those for the sliding DCT-based denoising but the difference is not large (less than 0.2 dB). Prediction is good enough although prediction errors can reach 0.6 dB and are larger than for $IPSNR$. Small values of $IPSNR-HVS-M$ indicate that there is no reason to apply filtering to enhance images.

Table 1. Simulation results for test images and different filters, PSNR (dB)

| Image index	STD	Noisy	TI-WS	BLS-GSM	NLM	BM3D	Geodesic
1	5	34.15	33.36	34.50	34.23	34.37	25.82
	10	28.16	27.89	29.25	28.75	29.05	24.64
	15	24.66	25.05	26.47	25.91	26.18	23.31
2	5	33.06	32.51	33.91	34.41	34.97	29.01
	10	27.55	27.92	29.36	29.49	30.20	27.17
	15	24.64	25.86	27.02	27.21	27.72	25.63
3	5	33.10	32.92	34.05	34.95	35.24	27.99
	10	27.54	28.04	29.48	30.08	30.57	26.44
	15	24.62	25.72	27.21	27.72	28.06	24.87
4	5	33.17	32.60	33.91	34.54	35.06	28.74
	10	27.79	28.12	29.37	29.67	30.29	27.06
	15	24.62	26.01	27.15	27.31	27.87	25.53
5	5	33.32	32.47	33.87	34.56	34.70	30.21
	10	27.68	27.65	29.05	29.81	29.70	27.99
	15	24.62	25.55	26.76	27.64	27.37	26.43
6	5	32.94	32.26	33.36	34.31	34.56	27.63
	10	27.54	27.12	28.60	29.09	29.35	26.15
	15	24.62	24.37	26.20	26.46	26.60	24.56
7	5	32.74	32.55	33.78	34.54	35.10	30.27
	10	27.16	27.93	29.22	29.88	30.30	28.17
	15	24.23	25.86	27.04	27.70	27.81	26.52
8	5	33.02	33.90	35.11	35.11	36.28	31.47
	10	27.54	29.40	30.97	30.97	31.82	28.99
	15	24.55	27.26	28.73	28.73	29.40	27.23
9	5	33.32	33.08	34.26	34.66	35.09	28.93
	10	27.64	28.45	29.77	29.99	30.39	27.27
	15	24.54	26.19	27.49	27.70	27.90	25.72
10	5	34.19	33.57	34.45	34.45	34.30	23.77
	10	28.21	27.68	28.93	28.93	28.60	23.13
	15	24.73	24.43	25.90	25.90	25.37	22.25
11	5	34.14	33.36	34.91	34.91	34.60	28.08
	10	28.14	28.30	29.94	29.94	29.56	26.56
	15	24.63	25.96	27.33	27.33	26.99	25.27
12	5	33.13	32.76	34.02	34.02	35.04	30.34
	10	27.71	28.26	29.74	29.74	30.27	28.20
	15	24.62	26.18	27.56	27.56	27.96	26.51
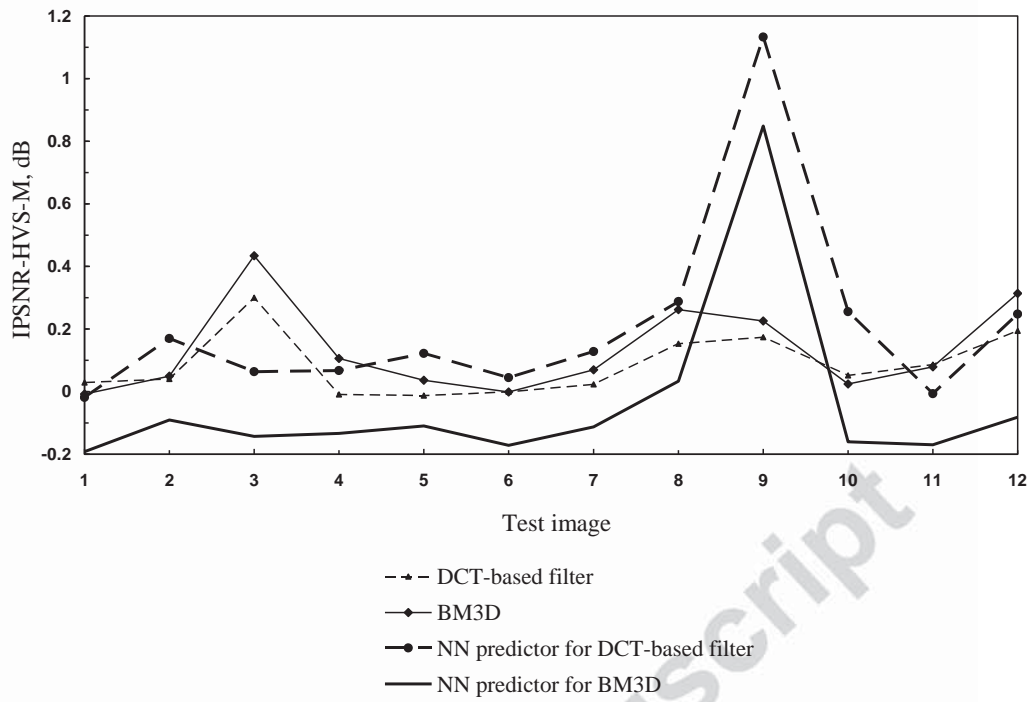
Fig. 7. $IPSNR - HVS - M$ metric, predicted values for AWGN with $\sigma_0^2 = 25$.
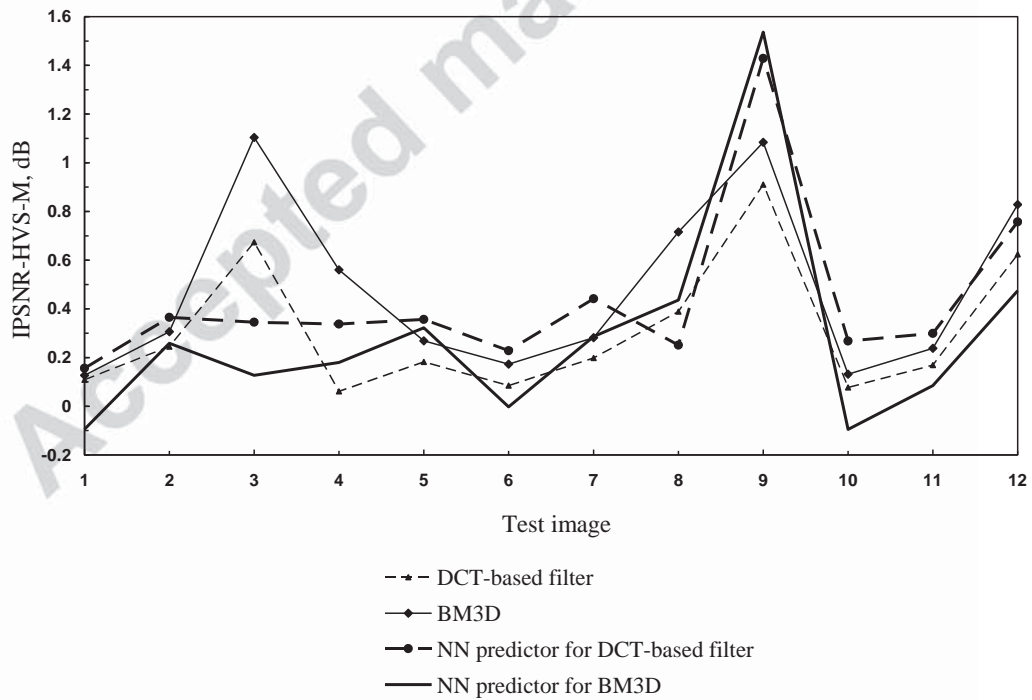

Fig. 8. $IPSNR - HVS - M$ metric, predicted values for AWGN with $\sigma_0^2 = 100$.
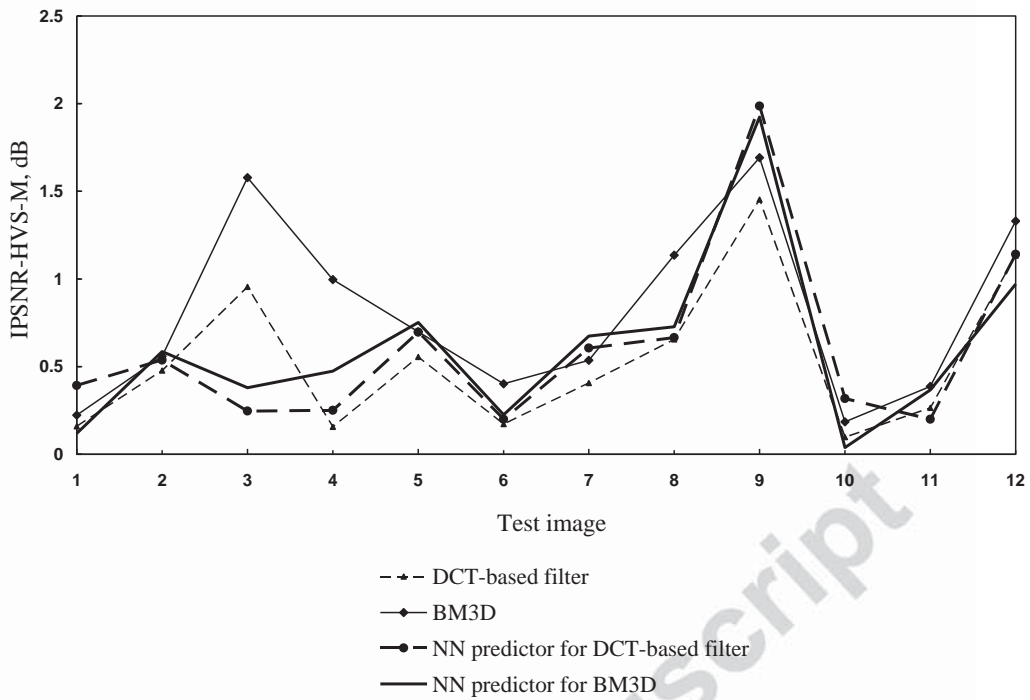
Fig. 9. $IPSNR - HVS - M$ metric, predicted values for AWGN with $\sigma_0^2 = 225$.

Fig. 8 presents data for $\sigma_0^2 = 100$ (STD=10). $IPSNR - HVS - M$ values have increased compared to the case of low intensity noise but they remain rather small. Only for the test images ## 3 and 9 it is worth applying denoising since $IPSNR - HVS - M$ is about 1 dB and such improvement can be noticed visually [35]. Prediction is accurate enough although up to 1 dB errors can take place (see data for the test image # 3 for BM3D filter). The reason for such an error could be that the parameter $P_{2\sigma}$ in no way takes into account regularity of texture exploited in denoising.

Finally, Fig. 9 shows data for the largest intensity noise (STD=15). $IPSNR - HVS - M$ values are larger than the corresponding values for smaller variances. For some test images (## 3 and 9), $IPSNR - HVS - M$ reaches 1.5 dB for the BM3D filter and this indicates that it is worth performing filtering. Meanwhile, there are images (## 1, 4, 6, 10, and 11) for which filtering seems useless.

Let us now analyse the results for other filters. They are presented in Table 2 (similarly to Table 1). For low intensity noise (STD=5), there is practically no improvement of visual quality characterized by the metric *PSNR-HVS-M*. Moreover, its values close to 45 dB observed for noisy images and output images for the best filters show that noise is invisible in noisy images whilst residual noise and distortions due to filtering are also not seen in filtered images.

The situation slightly changes for larger STD of the noise. The noise becomes visible for noisy images (this happens if *PSNR-HVS-M* is smaller than 40 dB). Residual noise and distortions after filtering are also visible. Even the best filters (BLS-GSM and BM3D) produce quite small improvement of visual quality. Other filters either do not considerably change visual quality (NLM and TI-WS filters) or make it worse (Geodesic).

Finally, for intensive noise ($\sigma_0^2 = 225$), BLS-GSM filter usually provides the best noise suppression and texture preservation in visual sense. The BM3D is slightly worse. Other filters can be ranked as NLM, sliding DCT, TI-WS, Geodesic in the order of performance reduction.

Table 2. Simulation results for test images and different filters, *PSNR-HVS-M* (dB)

Image index	STD	Noisy	TI-WS	BLS-GSM	NLM	BM3D	Geodesic
1	5	46.60	46.18	46.52	46.47	46.52	30.89
	10	36.77	35.74	36.89	36.41	36.73	28.90
	15	31.19	29.97	31.46	30.70	31.05	26.44
2	5	44.16	43.33	44.17	43.69	44.14	33.16
	10	34.73	33.63	35.10	34.45	34.76	29.96
	15	29.72	28.73	30.36	29.54	29.84	26.90
3	5	44.73	44.30	44.93	44.00	45.02	32.83
	10	34.96	34.55	35.72	34.95	35.79	29.95
	15	29.80	29.41	30.98	30.76	30.96	26.90
4	5	44.71	43.77	44.70	44.25	44.78	31.49
	10	34.85	33.55	35.18	34.31	35.20	28.78
	15	29.69	28.34	30.31	29.45	30.29	25.96
5	5	41.97	40.89	42.03	41.94	41.91	33.97
	10	33.00	31.82	33.40	33.34	32.93	30.15
	15	28.39	27.53	29.13	29.32	28.52	27.19
6	5	45.64	45.10	45.63	45.56	45.61	32.20
	10	35.66	34.67	35.86	35.65	35.71	29.95
	15	30.29	29.12	30.69	30.47	30.39	27.42
7	5	42.92	42.01	43.02	42.68	42.90	33.64
	10	33.66	32.62	34.10	33.63	33.57	29.89
	15	28.86	27.93	29.60	29.30	28.77	27.46
8	5	43.72	42.88	43.95	42.71	43.95	35.69
	10	34.18	33.13	34.84	33.13	34.71	30.53
	15	29.25	28.35	30.29	29.15	30.03	26.92
9	5	41.17	40.43	41.68	41.16	41.47	32.83
	10	32.88	32.65	34.09	33.32	33.78	29.90
	15	28.40	28.61	30.12	29.67	29.80	27.27
10	5	46.08	45.70	46.02	45.91	46.08	28.24
	10	36.56	35.81	36.72	36.23	36.50	27.21
	15	31.14	30.19	31.40	31.01	30.91	25.71
11	5	44.74	43.95	44.75	44.50	44.71	29.69
	10	34.99	33.68	35.27	34.76	34.94	27.97
	15	29.83	28.59	30.32	29.89	29.81	26.01
12	5	41.28	40.58	41.63	41.05	41.38	34.42
	10	32.78	32.38	33.72	33.11	33.26	30.76
	15	28.30	28.34	29.80	29.23	29.27	27.76

The results presented in Table 2 also show the following. Even if noise is intensive, denoising does not lead to any considerable improvement of image quality. The largest values of *IPSNR-HVS-M*

reach about 1.5 dB and this can be treated as noticeable but not great improvement of visual quality. Noise suppression is observed visually (especially for large-grain non-intensive textures). However, texture smearing takes place as well.

6.2 Spatially correlated noise case

Let us study now the results obtained for spatially correlated noise. Recall that MSE_{out}^{pot} and the corresponding $IPSNR$ are not available for this type of noise. Here, we have five parameters characterizing denoising efficiency for each test image and noise standard deviation: two $IPSNR$ values for the considered filters, two predicted values for these filters, and the upper bound for fBm-based approach [26]. Note that the NN predictor has been retrained for the case of spatially correlated noise. Its structure remained the same. Accuracy of prediction is at the same level as for AWGN, i.e. $IPSNR$ is predicted better than $IPSNR - HVS - M$ and prediction for the standard DCT-based filter is slightly better than for the BM3D filter.

The data for low intensity noise is presented in Fig. 10. The main conclusions are the following:

1) All five values of $IPSNR$ for a given test image are quite close;

2) This means that prediction works well enough and the BM3D filter does not have obvious advantages compared to the standard DCT-based filter;

3) Upper bound determined using the fBm-based approach practically coincides with denoising efficiency for the considered filters being closer to $IPSNR$ for the BM3D filter;

4) $IPSNR$ for most considered test images is small and, for $\sigma_0^2 = 25$, there is no reason to carry out filtering (note that prediction is able to indicate such situations before starting actual denoising of texture images);

5) Comparison of values that correspond to each other in Figures 4 (AWGN) and 10 (spatially correlated noise) shows that efficiency of denoising in the latter case is worse (it is more difficult to remove spatially correlated noise than AWGN under condition of the same variance).
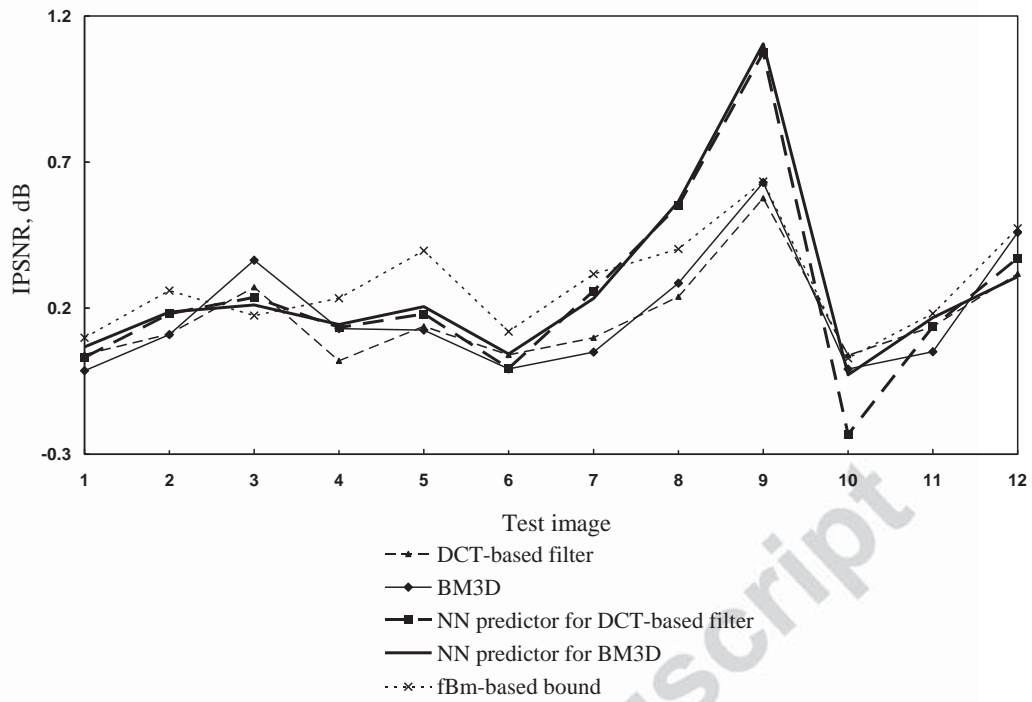
Fig. 10. *IPSNR* metric, predicted and fBm-based upper bound values for spatially correlated noise with $\sigma_0^2 = 25$.
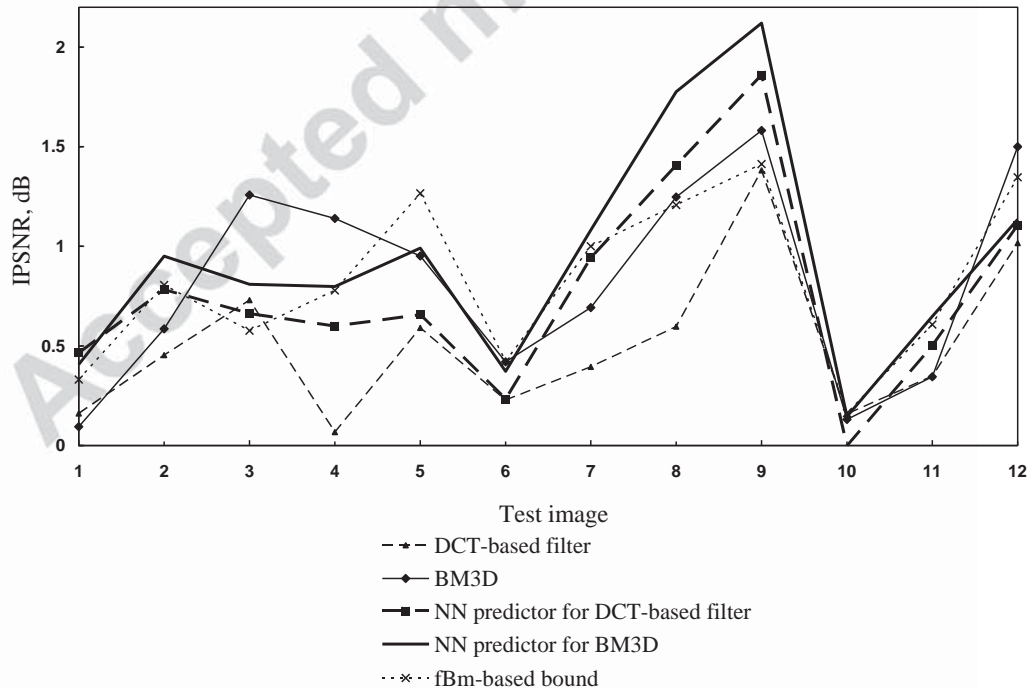

Fig. 11. *IPSNR* metric, predicted and fBm-based upper bound values for spatially correlated noise with $\sigma_0^2 = 100$.
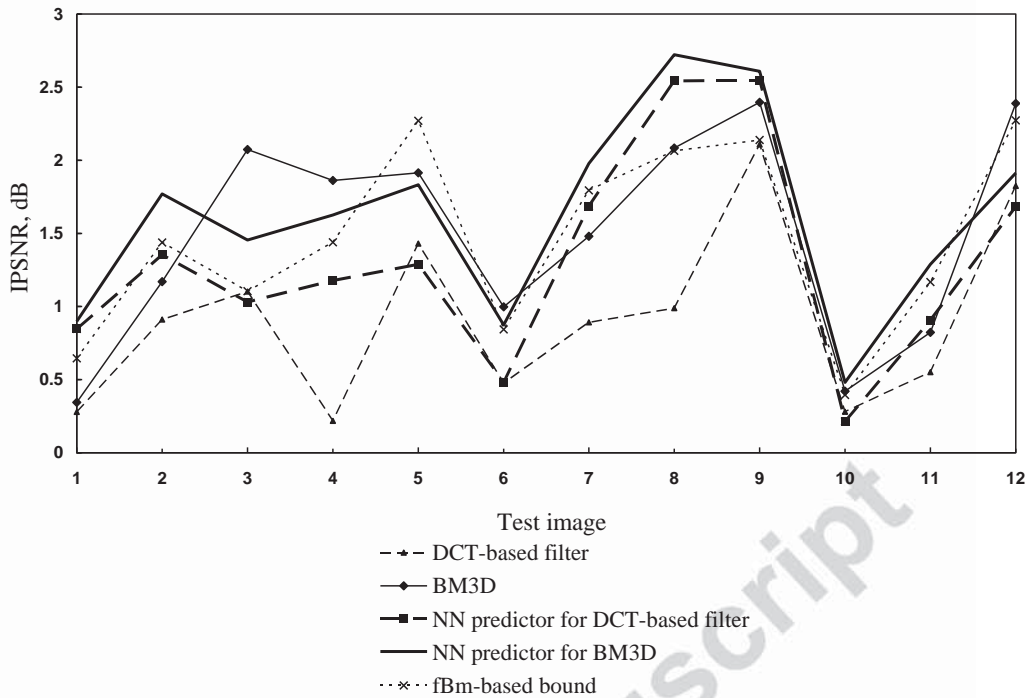
Fig. 12. *IPSNR* metric, predicted and fBm-based upper bound values for spatially correlated noise with $\sigma_0^2 = 225$.

The plots for middle and high intensity spatially correlated noise are given in Figures 11 and 12, respectively. The conclusions given above are mainly valid for them. The major difference is that *IPSNR* values increase if noise variance becomes larger (compare the corresponding data in Figures 10, 11, and 12). There appear test images for which BM3D outperforms the standard DCT filter sufficiently (for the test images ## 3 and 4 for $\sigma_0^2 = 225$) and for which it seems expedient to perform denoising (e.g., if *IPSNR* is larger than 2 dB).

It is worth comparing the DCT and BM3D filters to other known denoising techniques. Note that there are not many methods adapted to suppression of spatially correlated noise. On the other hand, the methods suited for removing AWGN usually lose their efficiency if they have to suppress spatially correlated noise. One method able to remove spatially correlated noise has been proposed recently [15]. This is a powerful blind denoising method able to cope with signal-dependent and spatially correlated noise. Additive case we analyse here can be considered as a particular case of signal-dependent one. Ability to perform filtering blindly is an advantage of the method [15]. However, there are several tuning parameters in it and, as we can conclude from our experiments, the default values of these parameters do not always provide the best efficiency. Because of this, here we present the best (potential) results provided by the method [15]. For this purpose, several sets of tuning parameters have been used for a given test image and noise standard deviation, the best result (the largest *IPSNR* or *IPSNR-HVS-M*) has been found and just these values are presented in Tables 3 and 4 (the method [15] is denoted as Col). In parallel, the values of *IPSNR* and *IPSNR-HVS-M* for the DCT-based and BM3D filters adapted to noise spatial correlation are given in Tables 3 and 4, respectively.

Consider the results for *IPSNR* presented in Table 3. The results for the filter [15] are very similar to those ones for the BM3D filter (and for the DCTF as well). Performance for all of them slightly improves for larger variance of the noise, but it is still difficult to denoise most of texture.

Table 3. *IPSNR* for BM3D, DCT-based and Colom's filter with optimal parameters, for three variance levels 25, 100 and 225

Image index		1	2	3	4	5	6	7	8	9	10	11	12
Col	25	0.06	0.19	0.53	0.36	0.23	0.16	0.19	0.35	0.64	0.15	0.08	0.41
DCTF		0.04	0.11	0.27	0.02	0.14	0.04	0.1	0.24	0.58	0.04	0.14	0.32
BM3D		-0.01	0.11	0.36	0.13	0.12	-0.01	0.05	0.29	0.63	-0.01	0.05	0.46
Col	100	0.29	0.55	1.24	0.89	0.86	0.36	0.75	1.03	1.52	0.33	0.40	1.18
DCTF		0.16	0.45	0.73	0.07	0.59	0.23	0.39	0.6	1.38	0.16	0.35	1.02
BM3D		0.09	0.59	1.26	1.14	0.95	0.42	0.69	1.25	1.58	0.13	0.35	1.50
Col	225	0.51	1.12	1.86	1.48	1.66	0.78	1.40	1.87	2.40	0.54	0.78	1.94
DCTF		0.28	0.91	1.1	0.22	1.43	0.48	0.89	0.99	2.11	0.28	0.55	1.82
BM3D		0.34	1.17	2.07	1.86	1.91	0.99	1.48	2.08	2.40	0.42	0.82	2.39

Here we would like to mention the following. For the filtered images, *IPSNR* about 1…2 dB does not necessarily mean that visual quality has improved [35] due to denoising compared to the original image. Besides, PSNR does not correlate well with visual perception of image quality. Meanwhile, HVS-based metric does it better but also not perfectly [35]. Our experience with the metric $PSNR-HVS-M$ [41] shows that improvement of visual quality becomes noticeable if $PSNR-HVS-M$ increases by, at least, 0.5…1 dB.

To make the final conclusions, consider simulation data for the metric $IPSNR-HVS-M$. Here, we have four values for each test image and noise standard deviation, namely, two $IPSNR-HVS-M$ values for the considered filters and two predicted values. Data for low intensity noise is represented in Fig. 13. It is seen that practically no improvement of visual quality is observed and there is no reason to carry out filtering. Prediction also indicates this.

The results for middle and high intensities of the noise are shown in Figures 14 and 15. Improvements of visual quality are observed for some test images (## 3, 4, 9, and 12) if BM3D filter is applied and noise has high intensity. Prediction performs well enough although prediction errors up to 1 dB are possible.
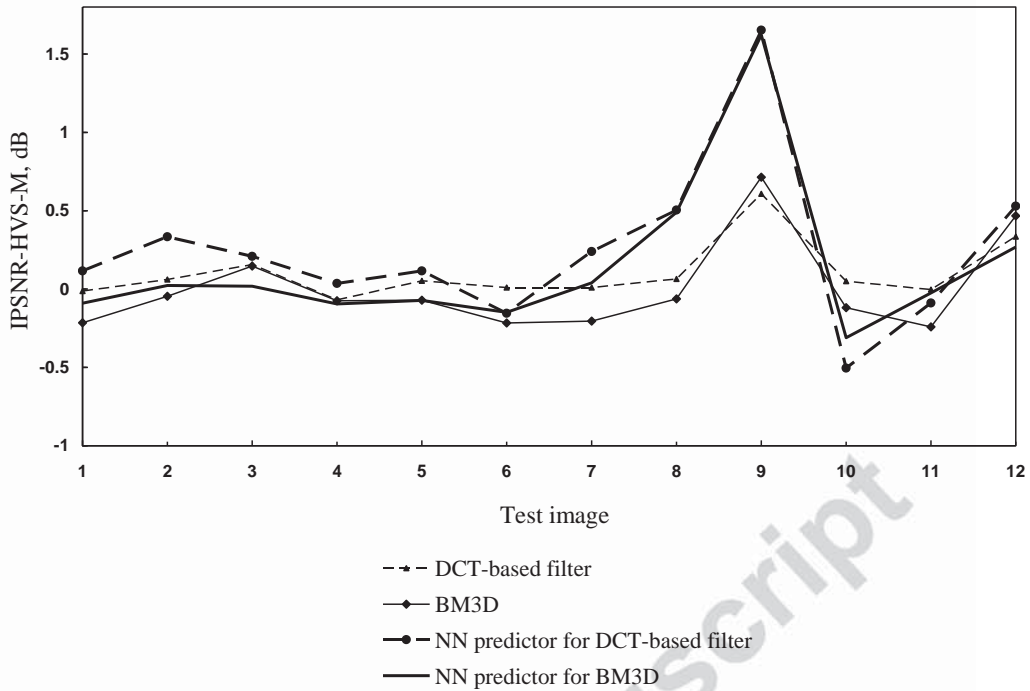
Fig. 13. $IPSNR - HVS - M$ metric and predicted values for spatially correlated noise with $\sigma_0^2 = 25$.

Additional data are presented in Table 4. Again, the results for the filter [15] and the BM3D filter are close. If noise variance increases, the performance of the filters improves.

Let us present a few filtering examples. They are given for spatially correlated noise with $\sigma_0^2 = 225$ in Fig. 16. Negative influence of the noise is seen well in Fig. 16, b. The DCT-based filter partly removes noise but also smears the textures (Fig. 16, c). Appearance of the processed images is better for the BM3D filter (Fig. 16, d) although this is clearly seen only for the textures 4 and 8. This agrees well with data in Fig. 15 where $IPSNR - HVS - M$ for BM3D are considerably larger than for the DCT-based filter only for the test images #4 and 8.
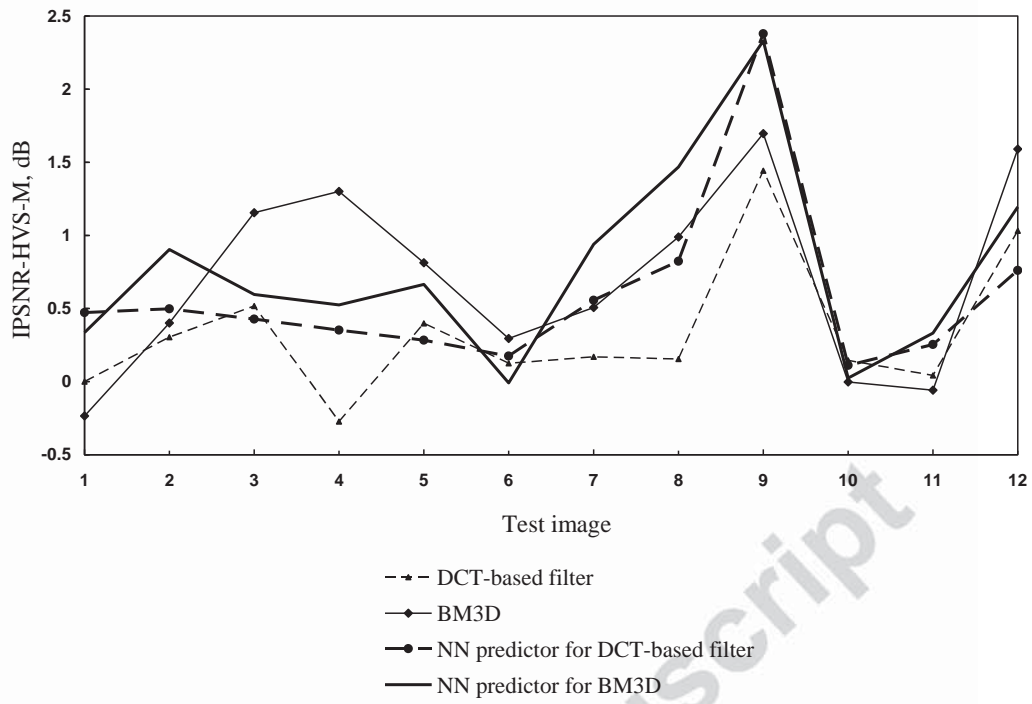
Fig. 14. *IPSNR − HVS − M* metric and predicted values for spatially correlated noise with $\sigma_0^2 = 100$
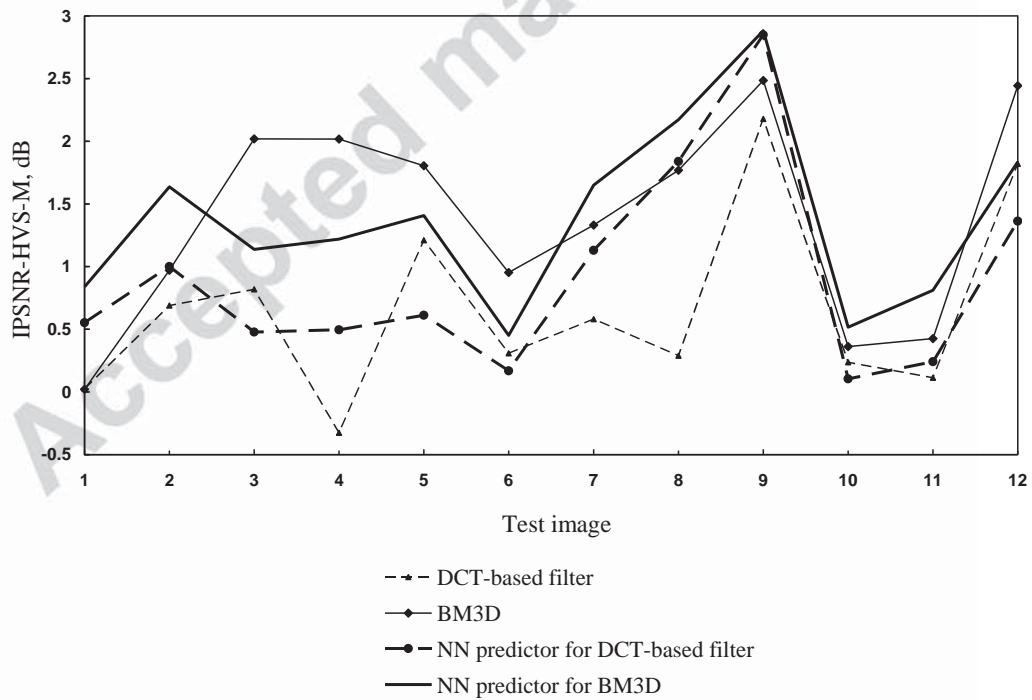

Fig. 15. *IPSNR − HVS − M* metric and predicted values for spatially correlated noise with $\sigma_0^2 = 225$

Table 4. IPSNR-HVS-M for BM3D, DCT-based and Colom's filter with optimal parameters

| Image index		1	2	3	4	5	6	7	8	9	10	11	12
Col	5	-0.1	0.05	0.47	0.33	0.14	0.00	0.10	0.15	0.78	-0.1	0.00	0.48
DCTF		-0.01	0.06	0.16	-0.07	0.05	0.01	0.01	0.07	0.61	0.05	0.00	0.34
BM3D		-0.2	-0.05	0.15	-0.07	-0.07	-0.22	-0.2	-0.06	0.71	-0.12	-0.24	0.47
Col	10	0.20	0.54	1.35	0.99	0.89	0.41	0.68	0.78	1.83	0.25	0.25	1.38
DCTF		0.00	0.3	0.52	-0.27	0.4	0.12	0.17	0.16	1.44	0.15	0.04	1.03
BM3D		-0.23	0.4	1.15	1.3	0.81	0.29	0.51	0.98	1.7	0.00	-0.06	1.59
Col	15	0.39	1.04	2.00	1.52	1.81	0.76	1.32	1.52	2.67	0.51	0.64	2.19
DCTF		0.03	0.69	0.82	-0.33	1.21	0.31	0.58	0.29	2.18	0.24	0.11	1.82
BM3D		0.02	0.97	2.02	2.02	1.8	0.95	1.33	1.77	2.49	0.36	0.43	2.44

| a | b | c | d |
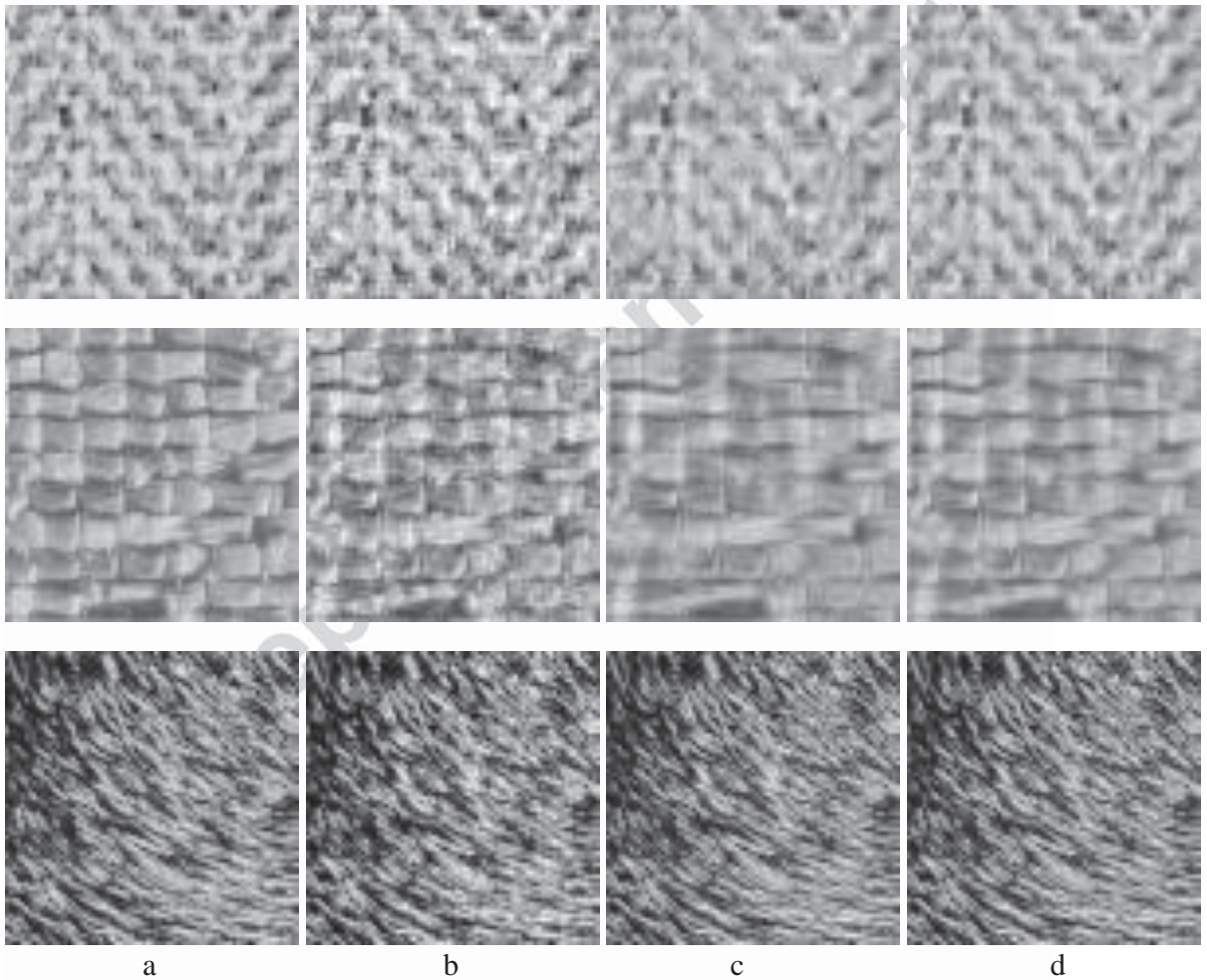
Fig. 16. Fragments of the test images ## 4, 8, and 9: a – original image, b – noisy image with $\sigma_0^2 = 225$, c – filtered image by DCT-based filter, d – filtered image by BM3D

Let us present one more example. A real image of neurons is presented in Fig. 17, a. This image is quite complex since it contains many small-sized objects. The noisy image for quite intensive AWGN is shown in Fig. 17,b (*PSNR*=28.55 dB, *PSNR-HVS-M*=34.25 dB). Noise is visible only in small quasi-homogeneous regions. The output of the DCT-based filter is represented in Fig.

17,c (*PSNR*=30.19 dB, *PSNR-HVS-M*=33.98 dB). This is an interesting case when *PSNR* has improved due to filtering whilst visual quality has not improved according to both *PSNR-HVS-M* metric and visual inspection. The output image for the BM3D filter is given in Fig. 17,d (PSNR=30.20 dB, PSNR-HVS-M=34.17 dB). Again, there is an improvement according to *PSNR*, but there is no noticeable improvement in visual quality. This example shows that improvement of *PSNR* due to denoising does not guarantee improvement of visual quality. Note also that predicted *IPSNR* is about 1.2 dB and *IPSNR-HVS-M* is about 0.5 dB for both filters. Thus, prediction also shows that it is not worth expecting high efficiency of denoising.
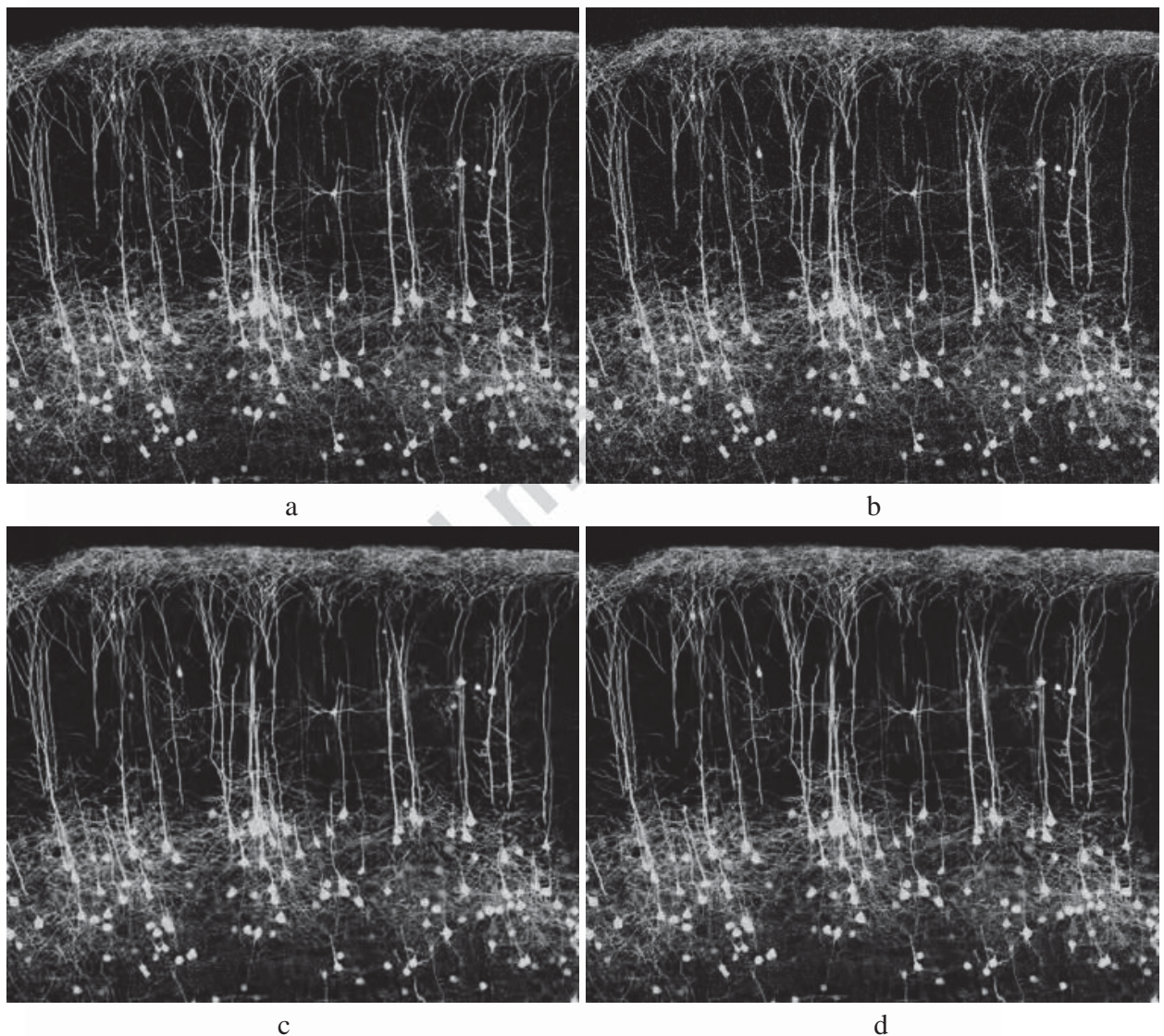

Fig. 17. Image of neurons: a – original image, b – noisy image, AWGN with $\sigma_0^2 = 100$, c – filtered image by the DCT-based filter, d – filtered image by the BM3D filter.

Conclusions

It is confirmed that it is a difficult task to remove noise from texture images or textural fragments of real-life images. Essential benefit due to denoising is usually provided only if noise is intensive and white. If noise is spatially correlated, benefits due to denoising are worse. The problem is not in filtering itself. Both thoroughly considered DCT-based filters are quite close to potential limits (upper bound of *PSNR*) determined in two different ways – for nonlocal approach to denoising (applicable for AWGN case only) and for fBm-based approach that is able to deal with both AWGN and spatially correlated noise. Other considered filters produce similar or even worse results.

Filter performance in terms of, at least, two quality metrics can be predicted quite accurately, especially if the trained NN-based predictor is used. This prediction is fast (faster than the sliding DCT-based denoising) since it is carried out using simple statistical parameters determined in a limited number of image blocks. Based on this prediction, it seems possible to undertake the decision if it is worth filtering a given image or not. We expect that such a decision can be also undertaken for textural image fragments under the condition that they are detected within a locally adaptive approach to perform filtering.

Aknowledgement

References

[1] R. Haralick, D. Dori, "A Pattern Recognition Approach to Detection of Complex Edges", Pattern Recognition Letters, Vol. 16, No. 5, pp. 517-529, May, 1995.

[2] R. Schowengerdt, "Remote Sensing: Models and Methods for Image Processing", Academic Press, 2006, 560 p.

[3] F. Cheikh, B. Cramariuc, M. Gabbouj, "MUVIS: A System for Content-Based Indexing and Retrieval in Large Image Databases", Proceedings Workshop on Very Low Bit Rate Coding, VLBV 1998, pp. 41-44.

[4] Tsymbal O.V., Lukin V.V., Ponomarenko N.N., Zelensky A.A., Egiazarian K.O., Astola J.T., "Three-state Locally Adaptive Texture Preserving Filter for Radar and Optical Image Processing", EURASIP Journal on Applied Signal Processing, No. 8, 2005,pp. 1185-1204.

[5] C.-A. Deledalle, L. Denis, F. Tupin, "How to compare noisy patches? Patch similarity beyond Gaussian noise", International Journal of Computer Vision, Vol. 99, No. 1, 2012, pp. 86-102.

[6] M. Lebrun, M. Colom, A. Buades, and J. M. Morel, "Secrets of image denoising cuisine," Acta Numerica, Vol. 21, 2012, pp. 475–576.

[7] D. Fevralev, V. Lukin, N. Ponomarenko, S. Abramov, K. Egiazarian, J. Astola, "Efficiency analysis of color image filtering", EURASIP Journal on Advances in Signal Processing, Vol. 2011:412, 2011, 19 p..

[8] K. Dabov, "Image and Video Restoration with Nonlocal Transform-Domain Filtering: Thesis for the degree of Doctor of Technology", Tampere (Finland) 2010, 181 p..

[9] G. Gilboa, N. Sochen, Y.Y. Zeevi, "Variational Denoising of Partly-Textured Images by Spatially Varying Constraints", IEEE Transactions on Image Processing, Vol. 15, No. 8, 2006, pp. 2281-2289.

[10] W. Zuo, L. Zhang, C. Song, D. Zhang, "Texture Enhanced Image Denoising via Gradient Histogram Preservation", Proceedings of CVPR, Portland, OR, USA, 2013, pp. 1203-1210.

[11] A. Buades, A. Coll, J.M. Morel, "A non-local algorithm for image denoising", Computer Vision and Pattern Recognition (CVPR) IEEE Computer Society Conference, Vol. 2, 2005, pp. 60-65.

[12] P. Chatterjee, P. Milanfar, "Is Denoising Dead?", IEEE Trans. Image Processing, Vol. 19, No. 4, 2010, pp. 895-911.

[13] R. Lukac, "Adaptive vector median filtering", Pattern Recognition Letters, Vol. 24, No 12, 2003, pp. 1889-1899.

[14] M. Szczepanski, B. Smolka, K.N. Plataniotis, A.N. Venetsanopoulos, "On the geodesic paths approach to color image filtering", Signal Processing, Vol. 83, No 6, 2003, pp. 1309-1342.

[15] M. Lebrun, M. Colom, J.-M. Morel, The Noise Clinic: a Blind Image Denoising Algorithm, Image Processing On Line, 5 (2015), pp. 1–54. http://dx.doi.org/10.5201/ipol.2015.125

[16] V. Lukin, R. Oktem, N. Ponomarenko, K. Egiazarian, "Image filtering based on discrete cosine transform", Telecommunications and Radio Engineering, Vol. 66, No. 18, 2007, pp. 1685-1701.

[17] L. Sendur, I. Selesnick, "Bivariate Shrinkage With Local Variance Estimation", IEEE Signal Processing Letters, Vol. 9, No. 12, 2002, 4 p.

[18] O. Pogrebnyak, V. Lukin, "Wiener discrete cosine transform based image filtering", SPIE Journal of Electronic Imaging, Vol. 21(4), 2012, pp. 043020-1-043020-15 doi: 10.1117/1.JEI.21.4.043020

[19] R.R. Coifman and D. Donoho, Wavelets and Statistics, Springer Verlag, chapter "Translation-invariant de-noising", 1995, pp. 125–150.

[20] J. Portilla, V. Strela, M.J. Wainwright, E. Simoncelli, "Image Denoising Using Scale Mixtures of Gaussians in the Wavelet Domain", IEEE Transactions on Image Processing, Vol. 12, No 11, 2003, pp. 1338-1351.

[21] K. Dabov, A. Foi, V. Katkovnik, K Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering", IEEE Transactions on Image Processing, Vol. 16, No. 8, 2007, pp. 2080-2095.

[22] K. He, J. Sun, X. Tang, "Guided image filtering" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, No 6, 2013, pp. 1397-1409.

[23] A. Buades, B. Coll, J.-M. Morel. "A review of image denoising algorithms, with a new one", SIAM Journal on Multiscale Modeling & Simulation, 4.2, 2005, pp. 490-530.

[24] V. Lukin, S. Abramov, N. Ponomarenko, K. Egiazarian, J. Astola, "Image Filtering: Potential Efficiency and Current Problems", Proceedings of ICASSP, 2011, 4 p.

[25] A. Rubel, V. Lukin, O. Pogrebniak, "Efficiency of DCT-based denoising techniques applied to texture images", Proceedings of MCPR, Cancun, Mexico, 2014, pp. 261-270.

[26] M. Uss, A. Rubel, V. Lukin, B. Vozel, K. Chehdi, "Lower Bound on Image Filtering Mean Squared Error in the Presence of Spatially Correlated Noise", Proceedings of MRRS, Kiev, Ukraine, 2014, pp. 10-13.

[27] S. Abramov, S. Krivenko, A. Roenko, V. Lukin, I. Djurovic, M. Chobanu, "Prediction of Filtering Efficiency for DCT-based Image Denoising", 2-nd Mediterranean Conference on Embedded Computing (MECO) , 2013, 4 p.

[28] O. Rubel, V. Lukin, "An Improved Prediction of DCT-Based Filters Efficiency Using Regression Analysis," Information and Telecommunication Sciences, Kiev, Ukraine, Vol. 5, No. 1, 2014, pp. 30-41.

[29] A. Rubel, A. Naumenko, V. Lukin, "A Neural Network Based Predictor of Filtering", Proceedings of MRRS, Kiev, Ukraine, 2014, pp. 14-17.

[30] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, M. Carli, "New full-reference quality metrics based on HVS", Proceedings of the Second International Workshop on Video Processing and Quality Metrics, Scottsdale, USA, 2006, 4 p.

[31] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image processing, Vol. 13, No. 4, 2004, pp. 600–612.

[32] M. Matrecano, G. Poggi, L. Verdoliva, "Improved BM3D for correlated noise removal", Proceedings of International Conference on Computer Vision, Theory and Applications, Vol. 1, 2012, pp. 129-134.

[33] A.S. Rubel, V.V. Lukin, K.O. Egiazarian, "Metric performance in similar blocks search and their use in collaborative 3D filtering of grayscale images", Proceedings of SPIE, Vol. 9019, USA, 2014, 12 p.

[34] A. Kadir, L.E. Nugroho, A. Susanto, P.I. Santosa, "Experiments of distance measurements in a foliage plant retrieval system", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 5, No 2, 2012, 14 p.

[35] V. Lukin, N. Ponomarenko, K. Egiazarian, "HVS-Metric-Based Performance Analysis Of Image Denoising Algorithms", Proceedings of EUVIP, Paris, France, 2011, pp. 156-161.

[36] P. Chatterjee, P. Milanfar, "Practical Bounds on Image Denoising: From estimation to information", IEEE Transactions on Image Processing. Vol. 20, No. 5, 2011, pp. 1221-1233.

[37] M. Uss, B. Vozel, V. Lukin, K. Chehdi, "Potential MSE of color image local filtering in component-wise and vector cases", Proceedings of CADSM, Ukraine, 2011, pp. 91-101.

[38] B. Pesquet-Popescu and J. Lévy Véhel, "Stochastic fractal models for image processing", IEEE Signal Process. Mag. Vol. 19, No. 5, 2002, pp. 48-62.

[39] A. Badiru, J. Cheung, "Fuzzy Engineering Expert Systems with Neural Network Applications", Wiley-Interscience, 2002, 320 p.

[40] C. Cameron, A. Windmeijer, A. G. Frank, H. Gramajo, D. E. Cane, C. Khosla, "An R-squared measure of goodness of fit for some common nonlinear regression models", Journal of Econometricsm ,Vol. 77, No. 2, 1997. pp.329-342.

[41] V.V. Abramova, S.K. Abramov, V.V. Lukin, K.O. Egiazarian, J.T. Astola, "On required accuracy of mixed noise parameter estimation for image enhancement via denoising", EURASIP Journal on Image and Video Processing, 2014:3, 2014, 16 p., doi:10.1186/1687-5281-2014-3.

[42] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo, "Color Image Database TID2013: Peculiarities and Preliminary Results", Proceedings of EUVIP, Paris, France, 2013, pp. 106-111

[43] K. Egiazarian, J. Astola, M. Helsingius, and P. Kuosmanen, "Adaptive denoising and lossy compression of images in transform domain", Journal of Electronic Imaging, 07/1999, pp.233-245.

Biographies

Aleksey Rubel received the M.Sc. degree in digital image processing from the National Aerospace University,Department of Receivers, Transmitters and Signal Processing, Kharkiv, Ukraine, in February 2013, where he is currently pursuing the Candidate of Technical Science degree in digital signal processing for remote sensing. His interests include image enhancement and design of image processing algorithms. His current work focuses on image processing algorithms adaptation to different noise models and its effectiveness prediction.

Vladimir V. Lukin graduated from Kharkov Aviation Institute (now National Aerospace University) in 1983 and got Diploma with honor in radio engineering. Since then he has been with the Dept of Transmitters, Receivers and Signal Processing of National Aerospace University. He presented the thesis of Candidate of Technical Science and in 1988 and Doctor of Technical Science in 2002 in DSP for Remote Sensing. Since 1995 he has been in cooperation with Tampere University of Technology. Currently he is Dept Vice-chairman and Professor. His research interests include digital signal/image processing, remote sensing data processing, image filtering and compression.

Benoit Vozel obtained the state engineering degree and the MSc degree in control and computer science in 1991 and the PhD degree in 1994 from École Centrale de Nantes (France). As a PhD student, he was with the Signal Processing Group at Institut de Recherche en Communication et Cybernétique de Nantes, where he worked on the detection of abrupt changes in signals. Since 1995, he has been with the ENSSAT, where he is currently with the signal and multicomponent/multimodal image processing research team (TSI2M) within the Institute of Electronics and Telecommunications of Rennes (IETR). His research interests generally concern blind estimation of noise characteristics, image filtering and restoration, and adaptive image and remote sensing data processing.

Mikhail Uss received the Diploma (with honors) in radio engineering and the Candidate of Technical Science degree in DSP for remote sensing from the National Aerospace University, Kharkov, Ukraine, in 2002 and 2006, respectively, and the Ph.D. degree from the University of Rennes 1, Rennes, France, in 2011. Since 2002, he has been with the Department of Aircraft Radioelectronic Systems Design of National Aerospace University. He is currently Head of the Department of Aircraft Radioelectronic Systems Design. His research interests include the statistical theory of radiotechnical systems, digital signal/image processing, blind estimation of noise characteristics, and theory of fractal sets with applications to image processing and remote sensing.

Karen O. Egiazarian received the Ph.D. degree from Moscow M. V. Lomonosov State University, Russia, in 1986, and Doctor of Technology degree from Tampere University of Technology, Finland, in 1994. He is a leading scientist in signal, image, and video processing,with about 300 refereed journal and conference articles, three book chapters, and a book published by Marcel Dekker, Inc. His main interests are in the field of multirate signal processing, image and video denoising and compression, and digital logic. He is a Member of the DSP Technical Committee of the IEEE Circuits and Systems Society.

Oleksiy Pogrebnyak received the Ph.D. degree from Kharkov Aviation Institute (now National Aerospace University), Ukraine, in 1991. Currently he is with The Center for Computing Research of National Polytechnic Institute, Mexico. His research interests include digital signal/image filtering and compression, remote sensing.

*Photo of the author(s)