https://doi.org/10.2352/ISSN.2470-1173.2020.9.IQSP-319

Prediction of performance of 2D DCT-based filter and adaptive selection of its parameters

Sergey Abramov, Vladimir Lukin, Oleksii Rubel, National Aerospace University, Kharkov, Ukraine
Karen Egiazarian, Tampere University, Tampere, Finland

Abstract

To enhance images, one often has to apply a filtering operation (denoising). However, there are several issues within the denoising. One of them is that sometimes denoising can be not efficient. Another issue regards a selection of an appropriate filter and setting of its parameters. As a particular case, we consider a 2D DCT-based filter with 8x8 pixel fully overlapping blocks where one of the parameters is a proportionality factor (PF) used in the threshold setup. We show that a performance of the considered filter in the sense of standard PSNR and visual quality metric PSNR-HVS-M can be predicted before applying image filtering procedure. This prediction is sufficiently faster than the denoising itself and accurate enough. We demonstrate that, having DCT statistics in a limited number of image blocks, such a prediction can be done for several values of PF. This allows deciding is it worth applying filtering to an image at hand. If the denoising is desired, it is also possible to select the PF optimal value for the considered image and noise intensity. Such a procedure, in some cases, can result in improvement of output PSNR or PSNR-HVS-M by up to 1 dB in comparison to the default parameters setup.

Keywords: DCT, denoising, filtering performance prediction, parameter optimization.

Introduction

A lot of images are acquired nowadays. They are of different quality and some images are corrupted by noise that is clearly visible degrading an image quality [1,2]. Such images have to be denoised in order to improve their visual appearance [2-4].

Image filtering (denoising) has been a hot topic for several decades [4-6]. The filters proposed so far belong to different families [5-7], where nonlocal and transform based techniques are among the best performing ones [7-9]. Meanwhile, efficiency of image denoising depends not only on the used filter but also on noise type and intensity, complexity of an image to be processed, filter parameters (e.g. scanning window or block size) set by a default or adjusted by a user [5-9]. Sometimes a filtering is useless according to a chosen criteria or according to a visual appearance [5, 7]. If one can predict such a situation, then the filtering can be skipped [5, 10-12]. This can save time and resources spent on image processing.

Analysis of a potential efficiency of removal of additive white Gaussian noise (AWGN) for nonlocal filtering has been first presented in [7] for the case when a true image is available. Several important conclusions have been drawn. First, it has been shown that a potential efficiency of denoising is low for highly textured images corrupted by noise that is not intensive. Second, it has been demonstrated that this potential efficiency is practically reached by the best existing nonlocal filters, such as, e.g., BM3D [13]. Meanwhile, there is a large space to improve the performance of existing filters for the case of processing of simple structure images corrupted by an intensive noise.

Later it has been shown [14, 15] that it is possible to estimate a potential efficiency of denoising without having the noise-free (reference) image. However, this requires considerable computations and, thus, its practical application is limited. Here it is worth recalling the main requirements to prediction of denoising efficiency. Prediction should be quite fast, i.e. sufficiently faster than the filtering itself. Prediction should be accurate enough, i.e. predicted values of an analyzed metric have to be close to the true values for each image.

An approach that more or less meets these requirements has been proposed in [10]. It presumes that noise statistics are a priori known or pre-estimated with a high accuracy [16, 17] and noise is AWGN. A simple statistic (two statistics have been studied) is determined in a limited number of 8x8 pixel blocks in DCT domain and used for the prediction. Prediction is carried out using approximating curve obtained in advance by curve fitting into scatter-plot [10, 18]. The scatter-plot is obtained by filtering images from a quite large set of test images corrupted by AWGN with the variance values varying in wide limits.

The approach [10] has been further advanced in [5, 11, 12, 19, 20]. In particular, it has been shown that accuracy of prediction can be improved using two or more input parameters and more complicated approximators, such as, e.g., neural networks. It has been also demonstrated that a prediction can be carried out for other than AWGN types of noise under condition that its characteristics are known in advance [11, 20]. More important, a prediction can be performed not only for the DCT-based filters [8, 13, 21]. A prediction is possible not only for the widely used metrics such as MSE or improvement of peak signal-to-noise ratio (IPSNR) but also for the visual quality metrics [5] such as PSNR-HVS-M [22] and others, although prediction for visual quality metrics is usually less accurate.

Meanwhile, design of prediction methods and tools has been performed under assumption that filter parameters are fixed (set according to some recommendations by default). In particular, threshold value for the DCT-based filter [8] is set equal to 2.7σ where σ is an AWGN standard deviation assumed to be known in advance. At the same time, it is known that filter parameters can be varied for some particular goal, e.g., to provide a better visual quality of output images [23]. Such a variation can be both local and “global” where proportionality factor is set fixed but different than 2.7 for processing of all image blocks.

Below we consider a way to determine this proportionality factor in advance, at the stage of a prediction of image denoising efficiency. We show that no essential modifications into prediction procedure are needed under condition that a preliminary work for obtaining approximating curves is done in advance. Meanwhile, a sufficient positive effect is possible.

Image filtering and metrics used

Let’s consider the following general image/noise model:

$$I^{n}_{ij} = I^{t}_{ij} + n_{ij}(I^{t}_{ij}), i = 1,...,I_{Im}, j = 1,...,J_{Im}, \qquad (1)$$

where I^{t}_{ij} denotes the true image *ij*-th pixel value, $n_{ij}(I^{t}_{ij})$ is the zero mean AWGN in the *ij*-th pixel, I_{Im}, J_{Im} define the processed image size.

The 2D DCT-based filter has several modifications and parameters. It performs in square shape blocks [8] where the block size 8x8 pixels is preferable due to the several reasons. Filtering can be done in non-overlapping, partly overlapping, and fully overlapping blocks where the last setting is the most efficient in the sense of the provided efficiency of noise removal but it is the most time consuming (this drawback is usually not crucial since the DCT-based filter is among the fastest denoisers). Hence, below we consider the fully overlapping blocks. For each block position (each image pixel except 7x7 lower rightmost ones serves as the left upper corner of the corresponding block), 3D DCT is performed. Then, thresholding is done for all DCT-coefficients except the DCT-coefficient responsible for the block mean. Thresholding can be carried out in different manner where there are soft, hard, and combined thresholding algorithms. Below we consider hard thresholding which is more efficient than soft [21] but simpler than the combined thresholding. After the thresholding, inverse 2D DCT is performed and initial filtered values are, thus, obtained for all pixels that belong to a given block. Since a given pixel belongs to several block positions (excluding image pixels at four corners), initially filtered values "coming" from different blocks are then aggregated by averaging.

The threshold value is set equal to $\beta\sigma$ where the default setting for β is 2.7 [21, 23]. This is done according to the analysis performed for many test image corrupted by AWGN with different σ. Moreover, this analysis shows the following: 1) optimal value of the parameter β (β_{opt}) depends on a quality metric used to characterize the filtering efficiency; it is usually larger for IPSNR compared to the visual quality metrics (for example, IPHVSM (improvement of PSNR-HVS-M due to the filtering determined as the difference between PSNR-HVS-M for the filter output and PSNR-HVS-M for original (noisy) image)); 2) β_{opt} also depends on image complexity and noise intensity; β_{opt} increases for simpler structure images and larger noise intensity (smaller input PSNR and PSNR-HVS-M). It is worth recalling here that the metric PSNR-HVS-M is expressed in dB, varies in a wide range (close to the range of PSNR variation). PSNR-HVS-M takes into account two important features of human vision – less sensitivity to distortions in higher spatial frequencies and masking effect (noise is masked by textures). By default, β_{opt} can be set to 2.4 if PSNR-HVS-M is used as a quality metric [23].

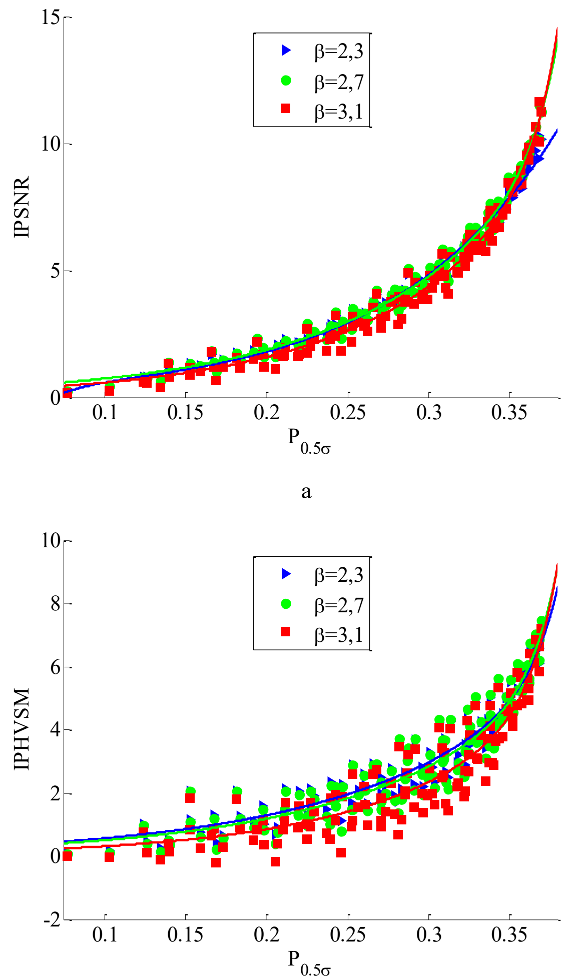
Analysis also shows that β_{opt} can vary in the wide limits. For example, according to the maximized IPSNR, it is possible that β_{opt} is in the limits 2.7±1. Similarly, according to maximized IPHVSM, β_{opt} varies in the limits 2.4±1.

Having this knowledge at our disposal, we have to analyze is it possible to predict the filter performance for the different values of β. Our idea is that if it is possible and if the difference in performance is sufficient, then it is possible to carry out prediction for several values of β and to choose the best value.

Prediction approach

The simplest way to carry out prediction is to use one input parameter [10]. Based on the results presented in [19], let us further use the probability $P_{0.5\sigma}$ – probability that absolute values of AC DCT coefficients do not exceed 0.5σ. Theoretically, $P_{0.5\sigma}$ varies in the limits from 0 to 0.38 where the upper limit is reached for fully homogeneous images (or images approaching to them as simple structure images corrupted by intensive AWGN).

An important preliminary stage needed to be passed once before using the efficiency prediction procedure is to get approximating dependences (curves). This can be done using scatter-plots and regression [18]. Examples are presented in Fig. 1.



b

Figure 1. Scatterplots and the fitted curves for IPSNR (a) and IPHVSM (b) on $P_{0.5\sigma}$ for three values of β

Each point of the scatterplot corresponds to one grayscale image corrupted by AWGN with one value of noise variance (standard deviation) and then filtered by 2D DCT-based filter with hard threshold set as $\beta\sigma$. Points for all three considered values of β are presented (they are marked by different colors and shapes). The curves are fitted by traditional tools available in Matlab or Excel separately for each value of β. The curves are marked by blue for β=2.3, green for β=2.7 and red for β=3.1. Color components of 15 first test images from the database TID2013 [24] have been used as the test images. Eight values of AWGN variance, namely, 4, 8, 16, 32, 64, 128, 256, and 512 have been employed to obtain noisy images.

A preliminary analysis of the plots in Fig. 1 shows the following:

1. There is an obvious tendency of filtering efficiency increase according to both metrics for $P_{0.5\sigma}$ increase; moreover, for $P_{0.5\sigma}$ smaller than 0.15…0.2, the improvement of quality due to filtering is negligible and filtering seems to be useless with any setting of β;
2. Scatter-plot points are placed in a compact manner around the fitted curves; the points in the second scatter-plots are placed less compactly; this means that IPSNR can be predicted better (with higher accuracy) than IPHVSM;
3. For the scatter-plots presented in Fig. 1 several functions can be used for fitting; analysis of opportunities is done in [5].

We have obtained the following approximations:

$$\mathrm{IPSNR}_{2.3}\left(\mathrm{P}_{0.5\sigma}\right)=0.2493\cdot e^{9.858\cdot P_{0.5\sigma}}-18.06\cdot e^{-52.58\cdot P_{0.5\sigma}} \quad (2)$$

$$\mathrm{IPSNR}_{2.7}\left(\mathrm{P}_{0.5\sigma}\right)=9.874\cdot 10^{-11}\cdot e^{64.4\cdot P_{0.5\sigma}}+0.2898\cdot e^{9.301\cdot P_{0.5\sigma}} \quad (3)$$

$$\mathrm{IPSNR}_{3.1}\left(\mathrm{P}_{0.5\sigma}\right)=6.103\cdot 10^{-11}\cdot e^{66.06\cdot P_{0.5\sigma}}+0.2102\cdot e^{10.08\cdot P_{0.5\sigma}} \quad (4)$$

$$\mathrm{IPHVSM}_{2.3}\left(\mathrm{P}_{0.5\sigma}\right)=3.008\cdot 10^{-10}\cdot e^{60.44\cdot P_{0.5\sigma}}+0.2484\cdot e^{8.238\cdot P_{0.5\sigma}} \quad (5)$$

$$\mathrm{IPHVSM}_{2.7}\left(\mathrm{P}_{0.5\sigma}\right)=3.71\cdot 10^{-10}\cdot e^{60.55\cdot P_{0.5\sigma}}+0.2127\cdot e^{8.59\cdot P_{0.5\sigma}} \quad (6)$$

$$\mathrm{IPHVSM}_{3.1}\left(\mathrm{P}_{0.5\sigma}\right)=2.015\cdot 10^{-10}\cdot e^{62.48\cdot P_{0.5\sigma}}+0.1146\cdot e^{10.01\cdot P_{0.5\sigma}} \quad (7)$$

Additional analysis of the fitted curves allows drawing the following conclusions:

1. All three curves in Fig. 1,a are close to each other for almost all values of $P_{0.5\sigma}$; only for $P_{0.5\sigma}$ about 0.27 filtering with β=2.3 and β=2.7 provides certain benefits compared to denoising using β=3.1; meanwhile, for $P_{0.5\sigma}>0.35$ the use of β=2.7 or even β=3.1 provides benefit up to 1 dB compared to processing with β=2.3;
2. According to the plots in Fig. 1,b, the use of β=3.1 is undesired for any images and noise intensities; the use of β<2.7, e.g., β=2.3 can be a reasonable choice.

Taking these conclusions into account, a procedure for setting a proper β can be as follows. Suppose we have approximating curves for several values of β. Then, for a given image, it is needed to determine $P_{0.5\sigma}$ and to calculate predicted IPSNR or IPHVSM for all values of β. After this, find the largest improvement and use β that corresponds to it. Note that input parameter is calculated only once, approximating curves are simple, and so it is easy to choose the best β according to a considered quality metric.

Some filtering examples

In fact, there are some practical situations when filtering is not needed at all. Fig. 2 presents such an example. The test image (Fig. 2,a) is corrupted by AWGN with the variance equal to 4. Such a noise is practically invisible (Fig. 2,b) and, thus, there is no need to suppress it. The filtered image (Fig. 2,c) looks practically the same as two other ones. Both IPSNR and IPHVSM are smaller than 1 dB and this additionally confirms that filtering is useless. Thus, we can recommend skipping filtering if $P_{0.5\sigma}$ does not exceed 0.15…0.17.

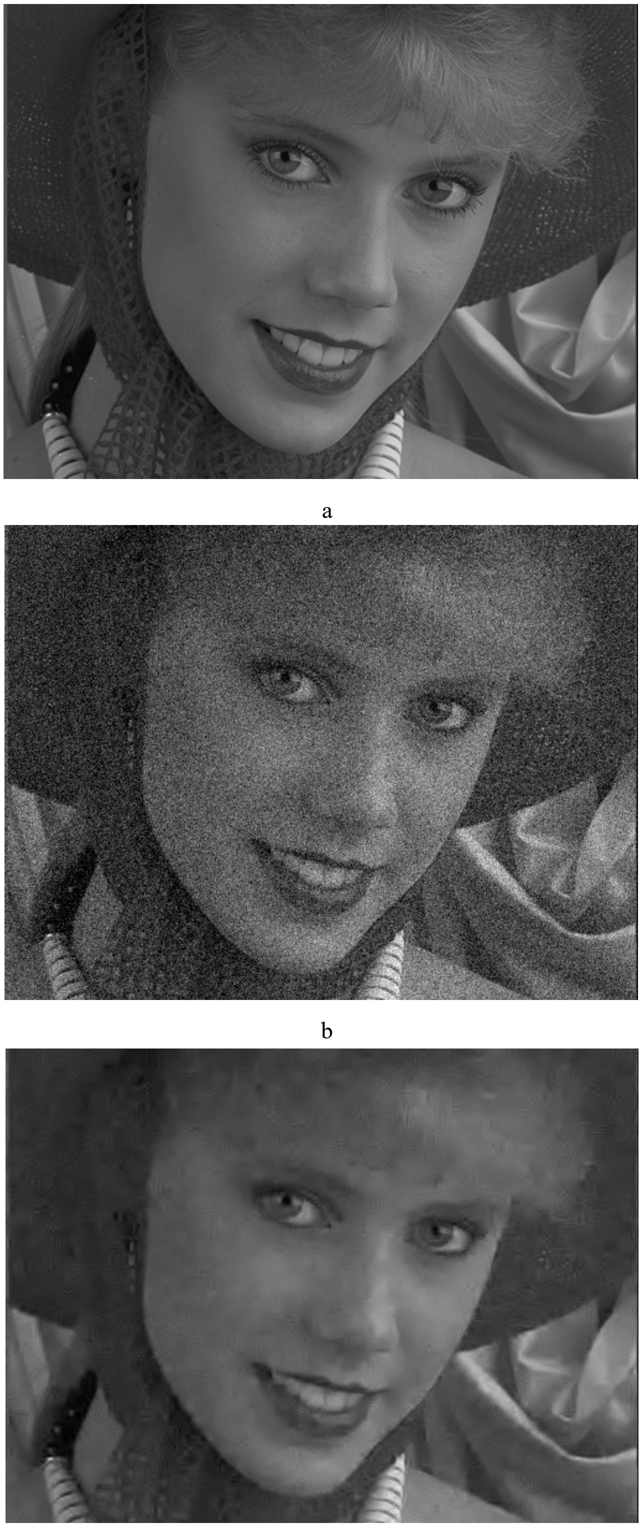
a

b

c

Figure 2. Noise-free (Test image TID#08) (a), noisy ($\sigma^2=4$; $P_{0.5\sigma}=0.13$) (b) and filtered ($\beta=2.3$; $\mathrm{IPSNR}=0.78$ dB; $\mathrm{IPHVSM}=0.46$ dB) (c) images

Consider now another example presented in Fig. 3.



a

b

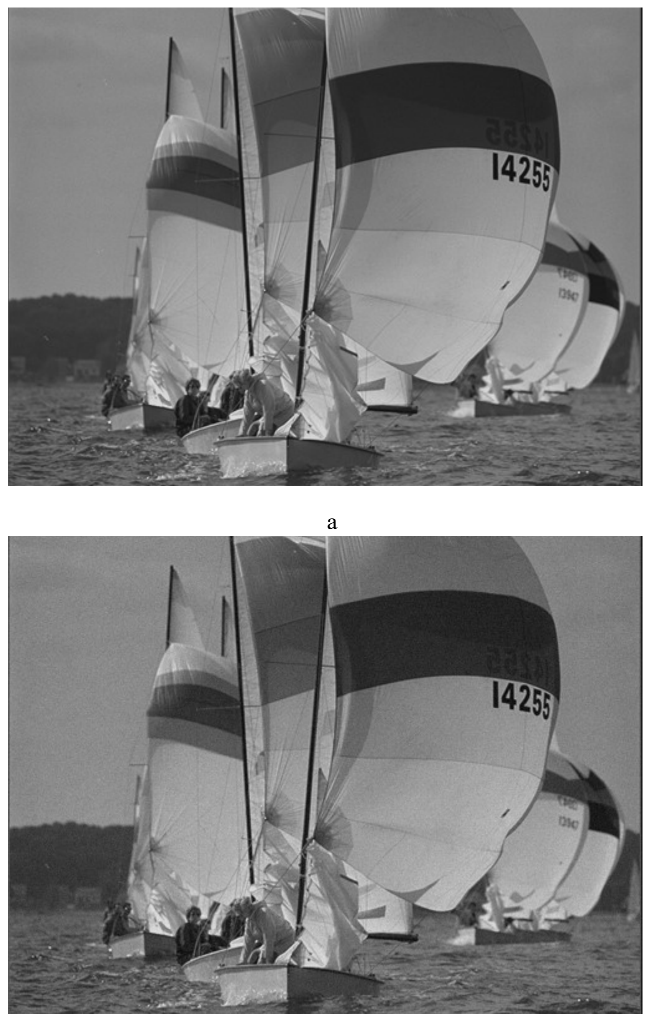
c

Figure 3. Noise-free (Test image TID#04) (a), noisy ($\sigma^2 = 512$; $P_{0.5\sigma} = 0.37$) (b) and filtered ($\beta = 2.7$; IPSNR=11.48 dB, IHVSM=6.86 dB) (c) images

Noise is intensive and it is clearly visible in the noisy image (Fig. 3,b) including the textured fragments. Noise is considerably suppressed by the filter (Fig. 3,b) but the details originally masked by the noise are not restored (note that other filters applied for this test image corrupted by such an intensive noise do not perform much better). Both metrics show that filtering is efficient (their values are quite large) but the filtering result might be unsatisfactory for some users.

Fig. 4 presents an example of "middle" situation. A noise-free test image (Fig. 4,a) is corrupted by middle intensity noise (Fig. 4,b) that is visible in homogeneous image regions and practically invisible in textural regions where the noise is masked.

Two filtered versions are given in Fig. 5. According to approximations in Fig. 1, it is better to apply β=2.7 than β=3.1. Both visual comparison of output images and quantitative data (metrics' values) confirm this. The image in Fig. 5(a) looks less smeared (edges and details are sharper) while noise is suppressed equally well.



a

b

Figure 4. Noise-free (Test image TID#09) (a), noisy ($\sigma^2 = 32$; $P_{0.5\sigma} = 0.31$) (b) images

a

b

Figure 5. Filtered images with β=2.7 (IPSNR=5.27 dB, IHVSM=4.34 dB) (a) and β=3.1 (IPSNR=5.00 dB, IHVSM=4.07 dB) (b)

Conclusions

This paper considers the problem of filtering efficiency prediction for 2D DCT-based filter and optimal parameter selection for this filter. It is shown that some metrics including visual quality ones can be predicted. Moreover, optimal parameter can be chosen without essential additional calculations. Prediction also shows situations when filtering can be skipped, saving time and resources.

To further improve the prediction performance, it is possible to use several input parameters. To our opinion, the proposed method can be also used for other filters in setting their parameters to decide is it worth to apply denoising.

References

[1] B.W. Keelan, Handbook of Image Quality, Marcel Dekker Inc., 2002.

[2] R.C. Gonzalez, R. Woods, Digital Image Processing, Pearson, 2018.

[3] D.M. Chandler, "Seven Challenges in Image Quality Assessment: Past, Present, and Future Research," ISRN Signal Processing, vol. 2013, pp. 1-53, 2013.

[4] M. Elad, Sparse and Redundant Representations. From Theory to Applications in Signal and Image Processing, Springer Science+Business Media, LLC, 2010.

[5] O. Rubel, S. Abramov, V. Lukin, K. Egiazarian, B. Vozel, A. Pogrebnyak, “Is Texture Denoising Efficiency Predictable,” International Journal on Pattern Recognition and Artificial Intelligence, vol. 32, 1860005, 32 p, 2018.

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

[7] P. Chatterjee, P. Milanfar, “Is Denoising Dead?,” IEEE Transactions on Image Processing, Vol. 19, No. 4, pp. 895–911, 2010.

[8] O. Pogrebnyak, V. Lukin, “Wiener discrete cosine transform based image filtering,” SPIE Journal of Electronic Imaging, Vol. 21, No. 4, pp. 1–15, 2012.

[9] X. Ma, P. Wu, Y. Wu, H. Shen, “A Review on Recent Developments in Fully Polarimetric SAR Image Despeckling,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 11, Issue 3, pp. 1-16, 2018.

[10] S. Abramov, S. Krivenko, A. Roenko, V. Lukin, I. Djurović, M. Chobanu, “Prediction of filtering efficiency for DCT-based image denoising,” Proc. of MECO 2013, Budva, Montenegro, pp. 97-100, 2013.

[11] O.S. Rubel, V.V. Lukin, F.S. de Medeiros, “Prediction of Despeckling Efficiency of DCT-based filters Applied to SAR Images,” Proceedings of DCOSS, Fortaleza, Brazil, pp. 159-168, 2015.

[12] A. Rubel, V. Lukin, M. Uss, B. Vozel, K. Egiazarian, O. Pogrebnyak, “Efficiency of texture image enhancement by DCT-based filtering,” Neurocomputing, Vol. 175, part B., pp. 948-965, 2016.

[13] 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, P. 2080-2095, 2007.

[14] X. Zhu, P. Milanfar, “Automatic parameter selection for denoising algorithms using a no-reference measure of image content,” IEEE Transactions on image processing, Vol. 19, No. 12, P. 3116-3132, 2010.

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

[16] M. Uss, B. Vozel, V. Lukin, K. Chehdi, “Maximum likelihood estimation of spatially correlated signal-dependent noise in hyperspectral images,” Optical Engineering, Vol. 51, No. 11, DOI: 10.1117/1.OE.51.11.111712, 2012.

[17] M. Colom, M. Lebrun, A. Buades, J.M. Morel, “A Non-Parametric Approach for the Estimation of Intensity-Frequency Dependent Noise,” IEEE International Conference on Image Processing (ICIP), DOI: 10.1109/ICIP.2014.7025865, 2014.

[18] 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 Econometrics, Vol. 77, No. 2, pp. 1790–1792, 1997.

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

[20] O. Rubel, V. Lukin, A. Rubel, K. Egiazarian, “NN-Based Prediction of Sentinel-1 SAR Image Filtering Efficiency,” Geosciences, Vol. 9, No 7, 22 p., 2019.

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

[22] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, V. Lukin, “On between-coefficient contrast masking of DCT basis functions,” CD-ROM Proceedings of VPQM, USA, 4 p., 2007.

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

[24] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battistid, C.-C. Jay Kuo, "Image database TID2013: Peculiarities, results and perspectives," Journal of Signal Processing: Image Communication, vol. 30, pp. 57–77, 2015.