UDC 551.568.85

ASSESMENT OF VISUAL QUALITY OF DENOISED IMAGES

A. S. Rubel, V. V. Lukin

National Aerospace University, Kharkiv

E-mail: rubel.andrew@gmail.com, lukin@ai.kharkov.com

The problem of image denoising is considered from the viewpoint of visual quality of filtered images. Special experiments with a large number of observers have been carried out to determine probability that a denoised image is preferable compared to the corresponding original noisy one. It has been found that there are many practical cases when observers prefer noisy images. This usually happens if an image is highly textural, noise has either quite low or too high intensity, and a used filter performs not efficiently. It has been also shown that modern metrics, even those that take into account peculiarities of human vision system, often perform not adequately. The cases when it is really worth to carry out image denoising are considered.

Keywords: *image denoising, visual quality, experimental assessment.*

ОЦІНКА ВІЗУАЛЬНОЇ ЯКОСТІ ЗОБРАЖЕНЬ ПІСЛЯ ФІЛЬТРАЦІЇ

А. С. Рубель, В. В. Лукін

Національний аерокосмічний університет, Харків

Фільтрацію зображень розглянуто як візуальну якість оброблених зображень. Виконано експерименти зі залученням великої кількості волонтерів, щоб визначити ймовірність того, що оброблене зображення краще за відповідне первинне. Виявлено, що є багато ситуацій, коли людина віддає перевагу первинному зображенню, що спотворене шумом. Зокрема, якщо зображення дуже текстурне, шум або дуже слабкий, або занадто інтенсивний, а фільтр працює неефективно. Також показано, що сучасні міри, навіть візуальної якості, часто працюють неадекватно. Розглянуто випадки застосовності фільтрації.

Ключові слова: *фільтрація зображень, візуальна якість, експериментальне оцінювання.*

It is known that noise is one of the main factors that degrades quality of images acquired by different systems [1, 2]. Because of this image filtering or denoising has become a standard operation in the image processing chain. Numerous filters have been designed so far (see [2–5] and references therein). Different types of noise have been considered [1, 6–8]. Various metrics including visual quality ones [9] have been applied to evaluate efficiency of image denoising [10, 11].

Meanwhile, efficiency of image filtering has to be evaluated from the viewpoint of purposes of image denoising. There are several purposes possible. One purpose could be to provide favorable pre-conditions for solving further tasks of image processing as, e.g., classification [12], segmentation, object detection [1], etc. Another purpose of filtering could be the removal of specific type of noise, i.e. removal of impulses. But the most common purpose of denoising is to improve visual quality of image, i.e. to enhance them [3–5, 10, 11].

In this sense the situation is such that many researchers (customers or observers) are unsatisfied by filtering results. There are quite many cases that observers consider visual quality of original images to be better or, at least, not worse than visual quality of the corresponding denoised ones although the metrics characterizing quality of original and filtered images evidence in favor of quality improvement due to denoising. This can be due to different reasons.

Firstly, a filter used really can be not efficient. In our studies [11, 13], it has been found that there are modern filters (including, nonlocal mean filter [14] that often produce decreasing of visual quality, in particular, for highly textural images). This is not surprising since it has been shown theoretically [2, 15] that there are potential limits of filtering efficiency for many practical situations (highly textural images corrupted by low and middle intensity noise) for non-local approach to denoising which is the state-of-the-art nowadays.

Secondly, alongside with positive effects of noise reduction, a filter can introduce artifacts or make annoying effects or produce images of unnatural appearance. In such cases a customer can be unsatisfied by filtering results although metrics of image quality might not reflect (indicate) the presence of these effects. Meanwhile, their presence can influence the opinions of observers who evaluate image quality.

Thirdly, metrics of image quality can be not adequate. Many researchers state that such standard metrics as output mean square error (MSE) or peak signal-to-noise ratio (PSNR) are often unable to characterize image visual quality well [9]. But specially designed visual quality metrics can be not adequate enough too [16].

Therefore, the goal of this paper is to analyze visual quality of the denoised images for two filters that belong to the most popular families: the standard DCT-based filter [17] which represents orthogonal transform based denoising and block-matching 3-dimensional (BM3D) filter [3] that is one of the best in the family of non-local de-noisers. We would like to understand the reasons why original images can be preferred by the observers who compare them to denoised counterparts.

Experiments with volunteers. Since humans are customers of original or denoised images it is worth carrying out experiments with quite many observers to understand when and why a denoised image is preferred. Such experiments have been performed [16]. Sixteen grayscale test images have been used where eight of them can be treated as natural scenes and eight others are textural. Four examples for each group are shown in Fig. 1 and 2.

Each observer was shown two images, original (noisy) and filtered ones, at high quality monitor in good illumination conditions and with favorable distance from an observer to the screen. The task was to choose a preferable image. Additive white Gaussian noise has been used as the noise model. Seven values of noise standard deviation have been considered, namely, 3, 5, 10, 15, 20, 25, and 30 where all noise-free images were represented as 8-bit data arrays.

Fig. 1. Natural scene test images with indices 1, 2, 4, and 8 in the created database.

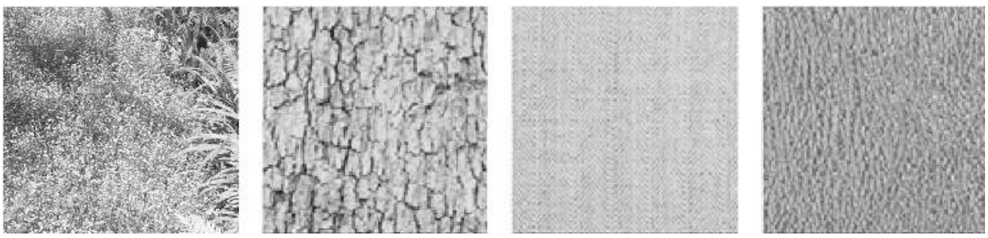

Fig. 2. Textural test images with indices 10, 12, 14, and 16 in the database.

Results of comparisons have been collected and averaged for all 145 observers that participated in experiments. As a result, probability of voting P_v in favor of denoised image has been calculated for each test image, for each value of noise standard deviation and for both filters. Probability tending to unity means that all observers prefer a filtered image. Probability of about 0.5 indicates that there is practically no difference in visual quality, Probability less than 0.5 evidences that filtering is useless.

Analysis of the obtained results. Let us show some results obtained for particular cases. Fig. 3*a* presents data for the test image # 4 that contains quite large homogeneous regions (walls) whilst Fig. 3*b* shows data for the textural image # 14. Dependences given in these Figures are typical for natural scene and texture images, respectively. The main findings are the following.

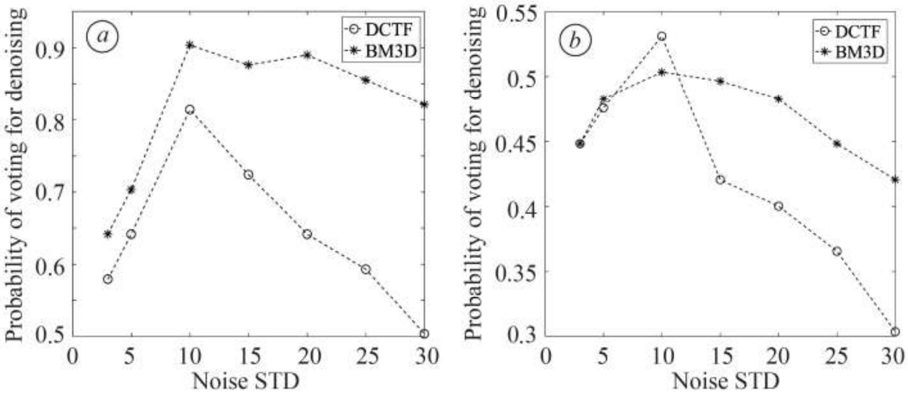

Fig. 3. Dependences of probability of voting on noise standard deviation
for two typical test images ## 4 and 14.

1) There are two parts of the curves, STD ≤ 10 where probabilities increase and reach maxima and STD ≥ 15 where dependences have the tendency to decrease.

2) This means that the main positive outcome from filtering can be expected if input PSNR is about 27 dB; however, even in this case positive effect can be not observed or it can be negligible; this happens for textural images.

3) The difference in visual quality of images processed by DCTF and BM3D filter is not considerable for STD ≤ 10 but it becomes larger for larger STD where the use of BM3D is preferable;

4) Probability varies in rather wide limits – from 0.3 to 0.9 in examples presented in Fig. 3; meanwhile, in none experiment the probability reaches either unity or zero; this means that human’s opinions are quite different even if it seems that a situation can be treated in only one way.

One question that arises: why BM3D produces better visual quality of denoised images than DCTF? This happens mainly due to the better preservation of edges and fine details that attract human attention during comparisons and sufficiently influence the observer’s decision [16]. Meanwhile, it often happens for textures that similar patches used in BM3D filter are not found (or, more exactly, the found patches are not too similar) and, thus, texture preservation is not perfect. An example is given in Fig. 4 where it is unclear what among three images (Fig. 4*b*–*d*) has the best visual quality (is the most similar to the noise free image in Fig. 4*a*).

Fig. 5 presents probabilities of voting for denoising for noise STD equal to 15. It is seen that this probability of the BM3D filter is always larger than for DCTF, especially for test images ## 2 and 3 having a lot of edges that are preserved by BM3D sufficiently better than by DCTF. Meanwhile, there are such test images (the textural images with indices 11–16) for which it is not worth applying filtering (both DCTF or BM3D). In general, there are only a few test images for which it is expedient to apply DCTF to improve visual quality (images with indices 1–5 and 8).

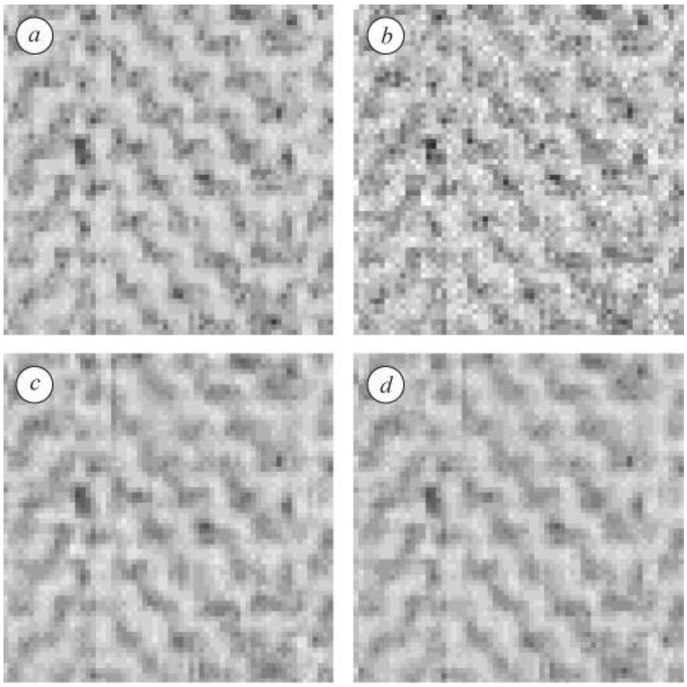

Fig. 4. Fragments of noise-free (*a*), noisy (*b*), filtered by DCTF (*c*), and denoised by BM3D (*d*) images.

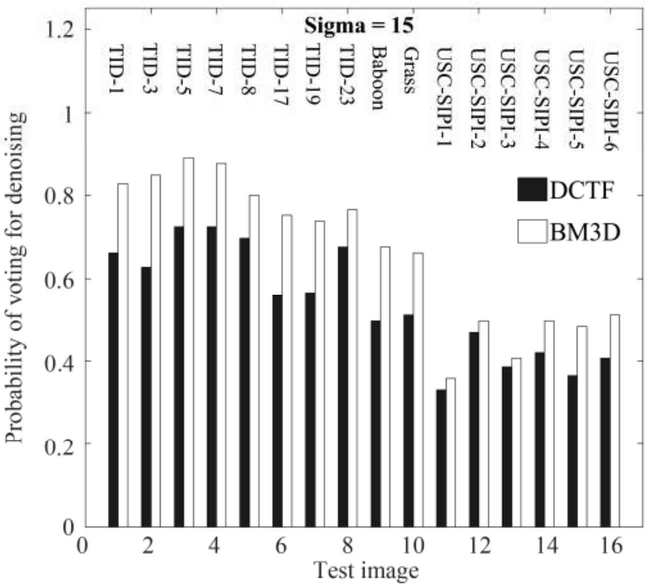

Fig. 5. Probabilities of voting for denoising for STD = 15 for two filters.

Filter performance prediction. Recently, methods to predict some metrics of image quality, including visual quality ones [18] have been proposed [19, 20]. In particular, it has been shown that it is possible to predict improvement of PSNR (*IPSNR*) and improvement of the metric PSNR-HVS-M [18] (denoted as *IPHVS*, both are expressed in dB and are calculated as differences of the metric values for output and input images). Formally, positive values of *IPSNR* and *IPHVS* indicate that positive effect of filtering has been reached. But is this really so? And how accurate the prediction is?

Fig. 6 shows the real and predicted values of *IPHVS* for the noise STD = 15 and for two considered filters. The prediction method designed in [20] has been used. Prediction is performed using probability $P_{0.5\sigma}$ that absolute values of AC DCT coefficients determined in 500 randomly placed 8x8 pixel blocks do not exceed 0.5σ. It is supposed that noise standard deviation σ is either a priori known or pre-estimated with appropriate accuracy [21]. The probability $P_{0.5\sigma}$ is calculated for an image to be denoised and applied as input parameter for dependence of output parameter *IPHVS* on $P_{0.5\sigma}$ where this dependence has been obtained in advance in off-line mode. Note that prediction is very fast since: a) DCT in standard block size is used; b) amount of considered image blocks is limited; c) approximating dependences are simple and they have been obtained in advance.

It is seen that the real and predicted values corresponding to each other can differ by up to 1.5…2 dB. The predicted values for textural images (that have indices 9–16) are usually slightly larger than the real ones whilst the predicted values for images with large homogeneous regions (test images ## 2, 4, 8) are smaller than real ones.

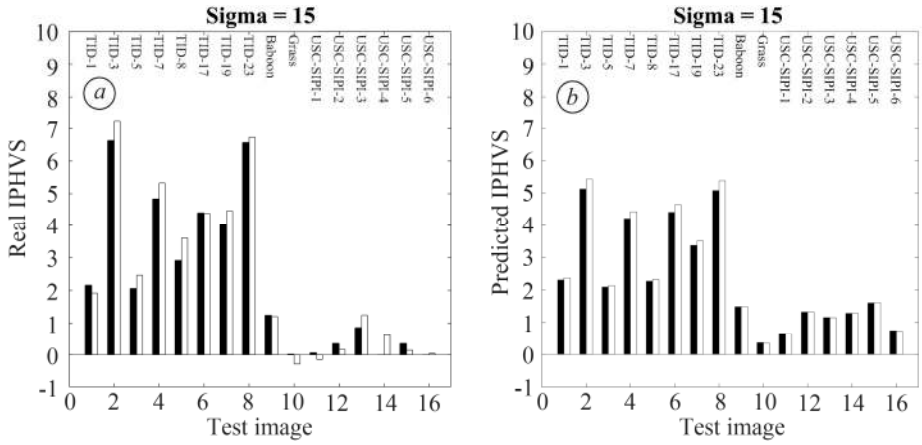

Fig. 6. Real (*a*) and predicted (*b*) values of IPHVS.

This shows that, on the one hand, further improvement of prediction accuracy is desired. But, on the other hand, the problem consists in the other phenomenon. As it follows from the joint analysis of data in Fig. 5 and 6, even the predicted values of IPHVS of about 1.5 dB do not guarantee (for STD = 15) that visual quality of images will improve due to denoising. Only if the predicted IPHVS exceeds 2…2.5 dB, one can expect with high probability that it is worth applying denoising. Based on these observations as well as on the analysis for other values of noise STD and the results obtained for the database TID2013, one can expect that $P_v > 0.5$ if

$$IPHVS > 9.33 – 0.337 PSNR_{inp}.$$

To be more "sure" that it is worth using denoising (i.e. to have P_v about 0.6 or larger), one has to check the condition that the predicted $IPHVS > 10.5 – PSNR_{inp}/2$ to have some reserve for possible inaccuracies of estimated parameters. Recall that $PSNR_{inp}$ can be estimated having an estimate of σ as $PSNR_{inp} = 10 \log_{10}(255^2/\sigma^2)$. Besides, it is not worth applying filtering if $PSNR_{inp} > 35$ dB, i.e. if noise standard deviation is smaller than 4…5 (see data in Fig. 7). This is because the condition $PSNR_{inp} \approx$ ≈ 35 dB approximately corresponds to the limit of noise visibility in noisy images. In fact instead of condition $PSNR_{inp} > 35$ dB it is better to use $PHVS_{inp} > 40$ dB [22].

The proposed procedure to undertake decision concerning applying denoising or skipping it is quite approximate. As it follows from the analysis of data in Fig. 5 and 6, it might be so that the same values of *IPHVS* can correspond to both images for which it is worth applying filtering (the test image # 10, BM3D filter) or it is worth to skip denoising (the test image # 11). Most probably, the reason for such difference deals with different properties of textures. Then this aspect should be studied more in detail.

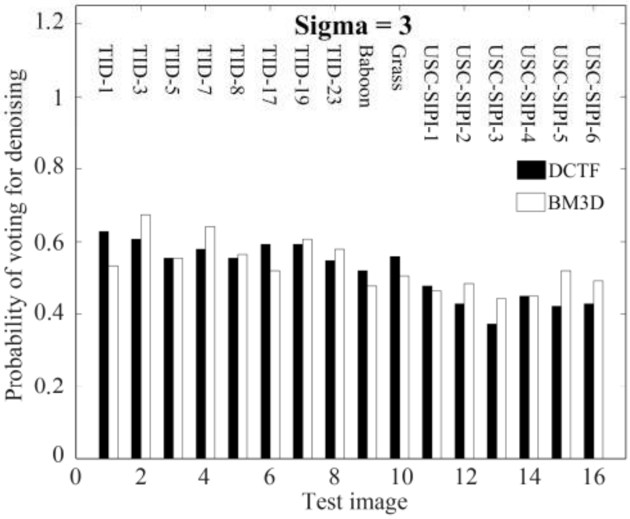

Fig. 7. Probabilities of voting for denoising for STD = 3 for two filters.

Besides the predicted parameters *IPSNR* and *IPHVS* are integral measures determined for entire image. It is assumed that the same parameter(s) as, e.g., threshold is (are) used for all positions of image blocks. Meanwhile the use of locally adaptive settings might improve the filter performance.

CONCLUSIONS

It is demonstrated that there are quite many practical situations when it is worth applying image denoising while there are also many cases where filtering can be skipped without losing image quality but with saving time and resources. Filtering efficiency can be predicted in advance for two filters (in fact, prediction can be done for many filters that belong to different families [13]). However, prediction of metrics considered to be able to characterize image visual quality quite adequate does not guarantee that, based on their analysis, it is possible to undertake a decision to filter a given image or to skip denoising.

1. *Schowengerdt R. A.* Remote Sensing: Models and Methods for Image Processing. – Academic Press. – 2006. – 560 p.
2. *Chatterjee P. and Milanfar P.* Is Denoising Dead? // IEEE Trans. Image Processing. – 2010. – **Vol. 19**, no. 4. – P. 895–911.
3. *Image* denoising by sparse 3D transform-domain collaborative filtering / K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian // IEEE Transactions on Image Processing. – 2007. – **Vol. 16**, no. 8. – P. 2080–2095.
4. *Secrets* of image denoising cuisine / M. Lebrun, M. Colom, A. Buades, J. M. Morel // Acta Numerica. – 2012. – **Vol. 21**, no. 1. – P. 475–576.
5. *Elad M. and Aharon M.* Image denoising via sparse and redundant representations over learned dictionaries // IEEE Transactions on Image Processing. – 2006. – **Vol. 15**, no. 12. – P. 3736–3745.
6. *Solbo S. and Eltoft T.* Homomorphic wavelet-based statistical despeckling of SAR images // IEEE Transactions on Geoscience and Remote Sensing. – 2004. – **Vol. 42**, no. 4. – P. 711–721.
7. *Pižurica A.* Image Denoising Algorithms: From Wavelet Shrinkage to Nonlocal Collaborative Filtering // Wiley Encyclopedia of Electrical and Electronics Engineering. – 2017. – 17 p.
8. *Enriquez A. and Ponomaryov V.* Image denoising using block matching and discrete cosine transform with edge restoring // 2016 Int. Conf. on Electronics, Communications and Computers (CONIELECOMP). – 2016. – P. 140–147.

9. *Lin W. and Jay Juo C.-C.* Perceptual visual quality metrics: A survey // J. of Visual Communication and Image Representation. – 2011. – **Vol. 22**, no. 4. – P. 297–312.
10. *Rubel A., Lukin V. and Pogrebniak O.,* Efficiency of DCT-based denoising techniques applied to texture images // Proceedings of MCPR, Cancun, Mexico. – June 2014. – P. 111–120.
11. *Efficiency* of texture image enhancement by DCT-based filtering / A. Rubel, V. Lukin, M. Uss, B. Vozel, K. Egiazarian, O. Pogrebnyak // Neurocomputing. – Jan. 2016. – **Vol. 175**. – P. 948–965.
12. *Analysis* of classification accuracy for pre-filtered multichannel remote sensing data / V. Lukin, S. Abramov, S. Krivenko, A. Kurekin, O. Pogrebnyak // J. of Expert Systems with Applications. – 2013. – **Vol. 40**, no. 16. – P. 6400–6411.
13. *Is texture* denoising efficiency predictable? / O. Rubel, S. Abramov, V. Lukin, K. Egiazarian, B. Vozel, A. Pogrebnyak // Int. J. on Pattern Recognition and Artificial Intelligence. – 2018. – **Vol. 32**, no. 1. – 32 p.
14. *Buades A., Coll B., and Morel J.-M.* A review of image denoising algorithms, with a new one // J. on Multiscale Modeling and Simulation. – 2005. – **Vol. 4**, no. 2. – P. 490–530.
15. *Chatterjee P. and Milanfar P.* Practical Bounds on Image Denoising: From Estimation to Information // IEEE Transactions on Image Processing. – 2011. – **Vol. 20**, no. 5. – P. 1221–1233.
16. *Rubel A. and Lukin V.* Regression-based Analysis of Visual Quality for Denoised Images // Fourth Int. Scientific-Practical Conf. Problems of Infocommunications Science and Technology (Kharkiv, Ukraine). – 2017. – P. 219–222.
17. *Image* filtering based on discrete cosine transform / V. Lukin, R. Oktem, N. Ponomarenko, K. Egiazarian // Telecommunications and Radio Engineering. – 2007. – **Vol. 66**, no. 18. – P. 1685–1701.
18. *On between-coefficient* contrast masking of DCT basis functions / N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, V. Lukin // Proc. of the Third Int. Workshop on Video Processing and Quality Metrics, USA. – Jan. 2007. – 4 p.
19. *Prediction* of filtering efficiency for DCT-based image denoising / S. Abramov, S. Krivenko, A. Roenko, V. Lukin, I. Djurovic, M. Chobanu // 2nd Mediterranean Conf. on Embedded Computing (MECO) (Budva, Montenegro). – 2013. – P. 97–100.
20. *Rubel O. and Lukin V.* An Improved Prediction of DCT-Based Filters Using Regression Analysis // Information and Telecommunications Sciences. – 2014. – **Vol. 5**, no. 1. – P. 30–41.
21. *Image* Informative Maps for Estimating Noise Standard Deviation and Texture Parameters / M. Uss, B. Vozel, V. Lukin, S. Abramov, I. Baryshev, K. Chehdi // EURASIP J. on Advances in Signal Processing. – 2011. – **Vol. 2011**, no. 1. – 12 p.
22. *Analysis* of HVS-Metrics' Properties Using Color Image Database TID2013 / V. Lukin, N. Ponomarenko, K. Egiazarian, J. Astola // Proceedings of ACIVS, Italy. – October 2015. – P. 613–624.

Received 12.10.2018