https://doi.org/10.2352/EI.2022.34.14.COIMG-152

Blind estimation of noise level based on pixels values prediction

Mykola Ponomarenko[1], Oleksandr Miroshnichenko[2], Vladimir Lukin[2] and Karen Egiazarian[1]
[1]Tampere University, Tampere, Finland
[2]National Aerospace University, Kharkiv, Ukraine

Abstract

Noise parameters estimation is required in various stages of digital image processing. Many efficient algorithms of noise estimation were proposed during last two decades. However, most of these algorithms are efficient only for a specific type of noise for which they are designed. For example, methods of variance estimation of additive white Gaussian noise (AWGN) will not work in the case of additive colored Gaussian noise (ACGN) or, in general, in the case of a noise with non AWGN distribution. In this paper, a totally blind method of noise level estimation is proposed. For a given image, a distorted image with a discarded portion of pixels (around 10%) is generated. Then an inpainting (or impulse noise removal) method is applied to recover those discarded pixels values. The difference between the true and recovered pixel values is used to robustly estimate image noise level. The algorithm is applied for different image scales to estimate a noise spectrum. In this paper, we propose a convolutional neural network called PIXPNet for effective prediction of values of missing pixels. A comparative analysis confirms that the proposed PIXPNet provides smallest error of recovered pixel values among all existing methods. A good efficiency of application of the proposed method in both AWGN and spatially correlated noise suppression is demonstrated.

Introduction

A task of noise level estimation is actual for many areas of digital image processing and analysis [1]. Quality of image denoising [1, 2] directly depends how precise a noise level is estimated. Noise level estimation is important in no-reference image visual quality assessment [3]. This task is also connected to estimation of unpredictability of image regions for human perception which is used in lossy image and video compression [4, 5]. Knowing a noise level one can estimate a level of lossy compression distortions which are visually inconspicuous [6].

Most efficient methods of noise level estimation are not blind. They are designed under condition that a noise specific distribution, type, and spectrum are known in advance.

Efficient algorithms of AWGN variance estimation [7-9] are designed under assumption that noise distribution is Gaussian, while noise variance and mean are constant for a whole image.

Algorithms for estimation of spectrum of ACGN [10,11] are also designed under assumption that distribution of the noise is Gaussian and its variance is a constant.

A map of standard deviations of non-stationary noise can be estimated with a good precision [12] if the noise distribution is Gaussian and noise spectrum is uniform.

In the case when an image is corrupted by a noise which distribution differs from distributions used for the design of the noise estimation algorithms, these algorithms will produce erroneous estimates. This problem is most contentious for noise level estimators which are based on deep convolutional neural networks (DCNNs) [11-13]. Such CNNs, pre-trained under specific noise (e.g., AWGN), may produce wrong results of prediction in real situations.

In practice, noise is often non-Gaussian and/or colored. For cases when a pre-processing is applied to an image, a variance of residual noise can be different for different image regions. In this case, most of the state-of-the-arts methods of estimation of variance of AWGN are not applicable. Because of this, designing robust methods not trained for a specific noise is a very important problem, which we address in this paper.

A method of totally blind noise level estimation was proposed in [5], where a Deep convolutional autoencoder (DCAE) with a small compression ratio introduces losses in a given image. Since this DCAE is trained on noise-free images, it preserves an image informative component and destroys a noise component. Because of this, a difference between input and output images of the DCAE are used for noise level estimation. The main drawback of the DCAE is possible overlearning during a training, where after a small number of iterations the DCAE packs two pixels values in one element of layer activations in the autoencoder's bottleneck, which limits an ability of DCAE to separate an image into noise and informative components.

In the paper, we propose a different fully blind approach of noise level estimation for prediction of an informative component of the image, and a use of a difference between input and predicted images for a noise level estimation.

The key element of the proposed method is a prediction of pixel values descarded by an impulse noise. Any efficient algorithm of image inpainting [14-21] or an algorithm of impulse noise removal [22-27] can be used for this purpose. However, in this paper we propose a novel DCNN for pixels values prediction. This DCNN, which we called PIXPNet, is optimized for this task and shows a state-of-the-art precision of pixel values prediction.

In this paper, we analyze efficiency of pixel values predictions by PIXPNet as well as an efficiency of suppression of different types of noise using estimates obtained by the proposed method.

Proposed method of noise level estimation

For many practical situations, noise characteristics are unknown. In Fig. 1, two examples of such images are given. Image in the Fig.1,a is a result of image processing in phase imaging [28].

a b

Fig. 1. Examples of images with unknown noise characteristics

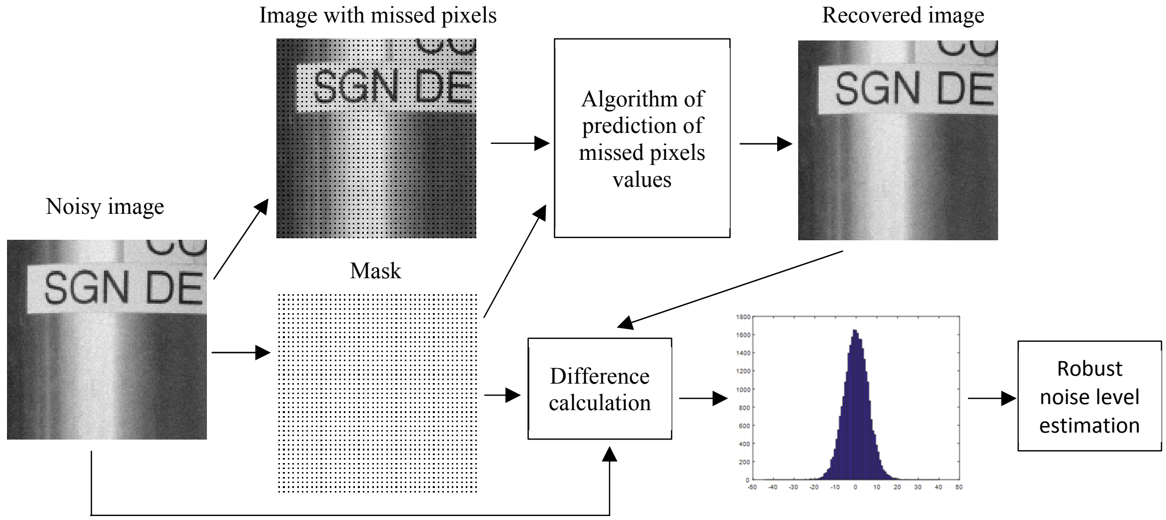

Fig. 2. Structural scheme of the proposed algorithm of noise level estimation

Noise on the image is spatially correlated (non-white) in the image center. The noise is almost white near image edges. The noise level is low in the image center and increases towards image edges. The noise distribution is near Gaussian in the image center and becomes non-Gaussian (heavy-tailed) near image edges. The noise on edges can be considered as an impulse noise.

Note, that all state-of-the-art methods for noise suppression [7-12] do not work for this image.

Another example is shown in the Fig.1,b. This image (and its enlarged fragment) is a result of image processing chain in Canon EOS 250D digital camera. It is automatically combined from several noisy images with a block matching. There is a visible residual noise on the image. The noise spectrum is non-white and unknown. Noise variance differs for different image regions. Noise is partially smoothed by lossy JPEG compression. A same situation is here: there is no good method which provides a reliable result.

Our goal is to propose a method which is able to estimate noise level map for such images with unknown and non-stationary noise characteristics.

A structural scheme of the proposed algorithm of noise level estimation is presented in Fig. 2.

For a given noisy image, a distorted image with removed set of pixels is generated as well as a mask indicating positions of these pixels. Then a method of restoration of the removed pixels values is applied. The method has two inputs: distorted image and the mask. The difference between recovered image and noisy image is calculated (only for the positions of removed pixels).

Our hypothesis is the following: if a good pixel predictor is used, then a noisy component of input noisy image will prevail in the difference, because the algorithm can predict pixel values but cannot predict noise values. Using this difference and a robust estimator (for example, median absolute deviation) one can estimate a noise level on the noisy image. The method will work for any noise distribution.

A key factor here is how efficient is the algorithm of prediction of missed pixel values. For an ideal prediction of pixels values, the difference between input and recovered images will contain only a noise component. For an inaccurate prediction there will be impurity in the difference conditioned by the presence of an information component. The larger impurity will result in larger errors in noise value estimation.

CNN for pixel values prediction

For missed pixel values prediction, we designed DCNN called PIXPNet (Fig. 3).

In contrary to the conventional CNN for noisy estimation, this predictor can be trained using any images including noise-free images. As a result, the proposed algorithm is not connected to any specific noise characteristics. In this sense it is fully blind and can be applied to any real-life noise level estimation.

The network combines U-Net and ResNet architectures and has two inputs: grayscale image with the missed pixels and the mask of these pixels' positions (known pixels are marked by 0, missed pixels are marked by 1). The output of PIXPNet is a recovered image.

The network architecture is very similar to DRUNet architecture [2] demonstrating state-of-the-art efficiency of AWGN suppression on images. Values of missed pixels in the input image are replaced by mean level of known neighbor pixels. PIXPNet was trained in Matlab using a custom training loop and MSE loss function. Input size for the training was 128x128 pixels. After the training PIXPNet can process images of arbitrary size.

The PIXPNet was pre-trained using 11000 images: noise-free images, photos with real-life distortions, clipart, cartoons, sketches, infographics. Due to this, a pre-trained PIXPNet can predict pixels of any images. We pretrained the PIXPNet for impulse noise probability 0..15% and for 200 000 iterations with minibatch size 16. Adam optimization was used with the initial learning rate 0.0001. After each 50 000 iterations the learning rate was decreased twice. To increase a number of difficult patches in the training, we have used a technique described in [29].

A pre-trained PIXPNet and Matlab's demo scripts are available in http://ponomarenko.info/pixpnet.

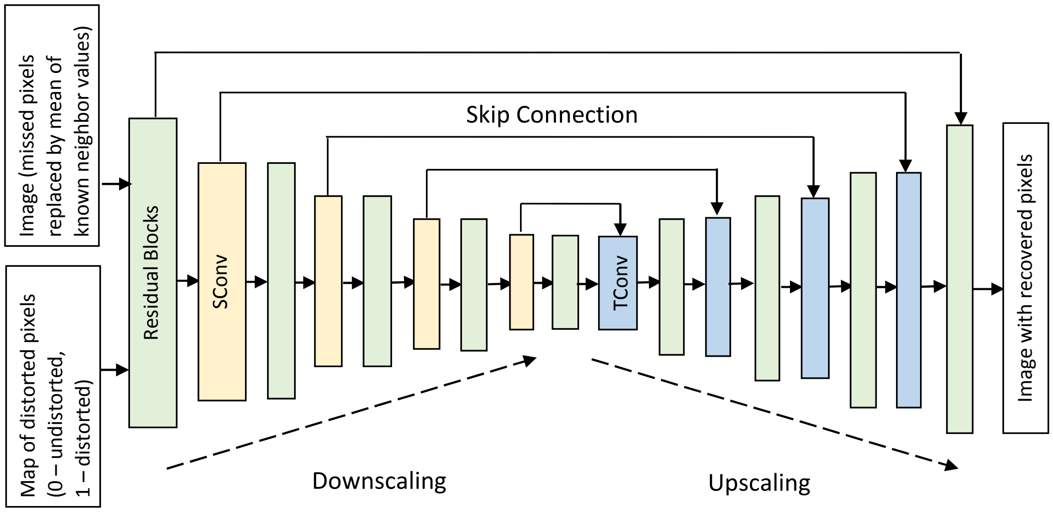

Fig. 3. Structural scheme of PIXPNet (SConv – Stride convolution, TConv – Transpose convolution)

Comparative analysis of pixel prediction efficiency

We compared a pre-trained PIXPNet with the several algorithms of impulse noise removal (for known positions of the corrupted pixels). The results obtained for noise-free Tampere17 [7] database images (not used for PIXPNet training) are given in Fig. 4.

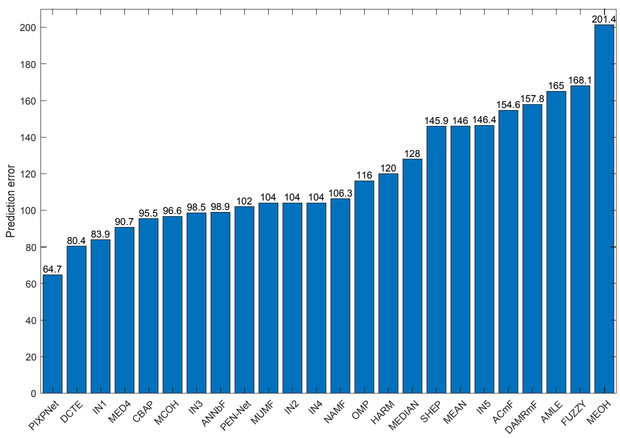

Fig. 4. Prediction errors for pixels recovering by different methods

We added to the comparison the following image inpainting methods. DCTE [14] reconstructs missed pixels based on minimization of entropy of coefficients of discrete cosine transform. IN1 (based on least-squares method), IN2 (based on linear system of equations), IN3 (IN2 with a better plate equation), IN4 (usage of a spring metaphor), IN5 (neighbors averaging) are algorithms implemented in “inpainting nans” library [15]. CBAP is a content based adaptive predictor used in lossless image compression [16]. MCOH is Matlab’s “inpaintCoherent” function. PEN-Net is a DCNN for image inpainting [17]. MUMF is an image inpainting algorithm based on the Mumford-Shah-Euler image model [18]. HARM is a harmonic inpainting [19]. SHEP is the Shepard inpainting [20]. AMLE is inpainting based on absolute minimization of Lipschitz extension [21]. MEOH is Matlab’s “inpaintExemplar” function.

Also, we added for the comparison the following methods of impulse noise removal. ANNbF is ANN-based Removal for Salt and Pepper Noise [22]. NAMF is a non-local adaptive mean filter [23]. OMP is a method based on orthogonal matching pursuit [24]. MEDIAN is median of eight neighbor pixels. MEAN is mean of eight neighbor pixels. ACmF is adaptive Cesáro mean filter [25]. DAMRmF is an adaptive modified Riesz mean filter [26]. FUZZY is an adaptive fuzzy filter for salt and pepper noise [27].

a	b	c
d	x	e
f	g	h

Fig.5. Illustration of the proposed MED4 impulse noise removal algorithm

In addition, we propose a new simple and fast method of pixel’s prediction combining four median edge detectors [30]. We called the method MED4. For prediction of a pixel x eight neighbor pixel values are used (see Fig. 5):

$$x = \text{median}\{a, a, e, e, b, b, g, g, d+b-a, b+e-c, d+g-f, e+g-h\} \quad (1)$$

Median edge detector for the left top corner of the Fig. 5 is calculated as $median\{d, b, d+b-a\}$ [30]. In (1), we combine median edge detectors for all four corners of 3x3 neighborhood of x. It is interesting that the proposed simple MED4 provides for this task superior results than most of the compared sophisticated and time-consuming methods.

It is clearly seen that PIXPNet significantly outperforms other methods producing 25% smaller prediction error than the nearest competitor DCTE which, in addition, is much slower.

Thus, PIXPNet is the best predictor to be used as the core of the proposed blind noise level estimation algorithm.

a b c

Fig. 7. Example of spatially correlated noise suppression: a) noisy image σ_b=0.8, σ=5, PSNR=34.2 dB, b) DRUNet+NLNet, PSNR=39.1 dB, c) DRUNet+PIXPNet, PSNR=39.2 dB

Usage of PIXPNet for estimation and suppression of AWGN noise

We estimated efficiency of the proposed method for AWGN suppression. For denoising, we have used DRUNet network [2] which requires noise standard deviation as one of the inputs.

Two cases were considered: an ideal denoising with a priori known noise standard deviation, and denoising with noise standard deviations using the proposed method based on pixel values predictions by PIXPNet.

Let us give details of the noise standard deviation estimation. PIXPNet is used to predict every ninth pixel of a given image (one pixel for each 3x3 patch is considered as a missing one and is predicted). Then the differences between pixel values and predicted values are calculated. Local variances in 7x7 sliding window are calculated for the recovered image. 95% of the differences with corresponding largest values of local variances are rejected from the analysis. Median absolute deviation is used to estimate standard deviation of noise on remained differences.

Results of denoising for Tampere17 set and different true noise standard deviations are given in Table 1.

Table 1. AWGN suppression, PSNR, dB

	σ=5	σ=10	σ=20
Noisy image	34.2	28.1	22.1
Denoising with true standard deviation	41.6	38.3	35.1
Denoising with estimation by PIXPNet	41.0	38.0	35.0

One can see that the proposed estimator provides PSNR of denoised images smaller than for an ideal case only by 0.3-0.6 dB.

Usage of PIXPNet for estimation and suppression of spatially correlated noise

We also estimated a case when the proposed estimator is used for suppression of spatially correlated nose. We used multiscale denoising scheme [11] showed in Fig. 2. DRUNet [2] is consistently applied to different image scales with a preliminary estimated noise value as one of the network's inputs. We compared the proposed estimator with NLNet estimator [11] which provides the state-of-the-art results for the task.

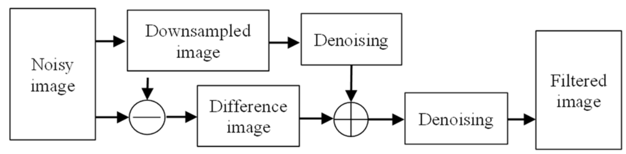

Fig. 6. Multiscale scheme of spatially correlated noise suppression (more scales can be included)

Table 2 shows the obtained results for Tampere17 set for different noise levels and different σ_b characterizing correlation between neighbor noise values [11]. PIXPNet shows good efficiency providing in average only 1 dB smaller PSNR than NLNet.

Table 2. Suppression of spatially correlated noise, PSNR, dB

	σ_b=0.5			σ_b=0.65			σ_b=0.8		
	σ=5	σ=10	σ=20	σ=5	σ=10	σ=20	σ=5	σ=10	σ=20
Noisy	34.2	28.1	22.1	34.2	28.1	22.1	34.2	28.1	22.1
DRUNet+ NLNet	40.3	36.7	32.7	39.4	35.7	31.8	38.5	34.8	31.0
DRUNet+ PIXPNet	39.2	36.3	33.1	39.1	34.5	31.0	37.6	34.0	29.5

Fig. 7 shows an example of noisy and processed images. PIXPNet estimates noise levels for three scales as 13.2 (smallest scale), 8.4, and 2.5 (full size).

Conclusions

This paper describes a novel algorithm of blind noise level estimation based on pixels values prediction. The pixels values predictor does not need to be pre-training for any specific noise (even noise-free images can be used for training). This allows to apply the proposed algorithm to estimate levels of various types of real-life noises.

It is shown that the proposed CNN PIXPNet provides a state-of-the-art efficiency of pixels values predictions. It can be used to recover pixels corrupted by impulse noise with a probability 1%-15%. It is also shown that the proposed algorithm can be used for both AWGN and ACGN levels estimation in image denoising providing the results close to an ideal or state-of-the-art denoisers.

References

[1] J. Astola, P. Kuosmanen, Fundamentals of Nonlinear Digital Filtering, CRC press, 288 p., 1997.

[2] K. Zhang, Y. Li, W. Zuo, L. Zhang, L.Van Gool, R. & Timofte, "Plug-and-play image restoration with deep denoiser prior", IEEE Transactions on Pattern Analysis and Machine Intelligence, 17 p, 2021.

[3] N. Ponomarenko, V. Lukin, O. Eremeev, K. Egiazarian, J. Astola, "Sharpness metric for no-reference image visual quality assessment", Image Processing: Algorithms and Systems X and Parallel Processing for Imaging Applications II. International Society for Optics and Photonics, vol. 8295, 11 p, 2012.

[4] O. Ieremeiev, N. Ponomarenko, V. Lukin, J. Astola, K. Egiazarian, Masking effect of non-predictable energy of image regions, Telecommications and Radio Engineering, 76 (8), pp. 685-708, 2017.

[5] S.G. Bahncmiri, M. Ponomarenko, K. Egiazarian, Deep Convolutional Autoencoder for Estimation of Nonstationary Noise in Images, In 2019 8th European Workshop on Visual Information Processing (EUVIP), pp. 238-243, 2019.

[6] V. Lukin, S. Abramov, N. Ponomarenko, M. Uss, M. Zriakhov, B. Vozel, K. Chehdi, J. Astola, Methods and automatic procedures for processing images based on blind evaluation of noise type and characteristics, Journal of applied remote sensing, 5(1):053502, 2011.

[7] M. Ponomarenko, N. Gapon, V. Voronin, K. Egiazarian, "Blind estimation of white Gaussian noise variance in highly textured images", Proceedings of International Conference Image Processing: Algorithms and Systems, 5p., 2018.

[8] S. Pyatykh, J. Hesser, and L. Zheng, "Image Noise Level Estimation by Principal Component Analysis", IEEE Transactions on Image Processing, pp. 687-699, 2013.

[9] X. Liu, M. Tanaka and M. Okutomi, "Noise Level Estimation Using Weak Textured Patches of a Single Noisy Image", IEEE International Conference on Image Processing, 2012, pp. 665-668.

[10] N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola, A method for blind estimation of spatially correlated noise characteristics, Proceedings of SPIE Conference Image Processing: Algorithms and Systems VII, 2010.

[11] M. Ponomarenko, O. Miroshnichenko, V. Lukin, K. Egiazarian, "Blind Estimation and Suppression of Additive Spatially Correlated Gaussian Noise in Images", European Workshop on Visual Information Processing, 6 p., 2021.

[12] S.G. Bahnemiri, M. Ponomarenko, K. Egiazarian, "Learning-based Noise Component Map Estimation for Image Denoising", submitted to IEEE Signal Processing Letters, 2022.

[13] M. Uss, B. Vozel, V. Lukin, K. Chehdi, NoiseNet: Signal-Dependent Noise Variance Estimation with Convolutional Neural Network, In International Conference on Advanced Concepts for Intelligent Vision Systems, pp. 414-425, 2018.

[14] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, J. Astola, Reconstruction of missing pixels in images using entropy of DCT coefficients, In Information, Telecommunication and Sofware Tools of Digital Signal Processing, pp. 5-20, 2008.

[15] J. D'Errico, Inpaint_nans, MATLAB Central File Exchange, https://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans), 2012.

[16] F. Golchin, K. Paliwal, A context-based adaptive predictor for use in lossless image coding, Proceedings of IEEE Region 10 Annual Conference Speech and Image Technologies for Computing and Telecommunications, pp. 711-714, 1997.

[17] Y. Zeng, J. Fu, H. Chao, B. Guo, Learning pyramid-context encoder network for high-quality image inpainting, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1486-1494), 2019.

[18] S. Esedoglu, J. Shen, Digital inpainting based on the Mumford-Shah-Euler image model, European Journal of Applied Mathematics, 13(04), 353-370, 2002.

[19] Shen, J., & Chan, T. F. Mathematical models for local nontexture inpaintings. SIAM Journal on Applied Mathematics, 62(3), 1019-1043, 2002.

[20] J. Ren, L. Xu, Q. Yan, W. Sun, Shepard convolutional neural networks, Advances in Neural Information Processing Systems, pp. 901-909, 2015.

[21] V. Caselles, J. Morel, C. Sbert, An axiomatic approach to image interpolation, IEEE Transactions on Image Processing, 7(3), pp. 376-386, 1998.

[22] B. Turan, ANN Based Removal for Salt and Pepper Noise, Global Conference on Engineering Research, GLOBCER'21, 2021.

[23] H. Zhang, Y. Zhu, H. Zheng, NAMF: A Non-local Adaptive Mean Filter for Salt-and-Pepper Noise Removal, arXiv preprint arXiv:1910.07787, 2019.

[24] J. Tropp, A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Transactions on Information Theory, 53(12), pp. 4655–4666, 2007.

[25] S. Enginoğlu, U. Erkan, S. Memiş, Adaptive cesáro mean filter for salt-and-pepper noise removal, El-Cezeri Journal of Science and Engineering, 7(1), pp. 304-314, 2020.

[26] S. Memiş, U. Erkan, Different Adaptive Modified Riesz Mean Filter For High-Density Salt-and-Pepper Noise Removal in Grayscale Images, Avrupa Bilim ve Teknoloji Dergisi, (23), 359-367, 2021.

[27] V. Singh, Vikas, et al. "Adaptive type-2 fuzzy approach for filtering salt and pepper noise in grayscale images." IEEE Transactions on Fuzzy Systems 26.5 (2018): 3170-3176.

[28] P. Kocsis, I. Shevkunov, V. Katkovnik, H. Rekola, K. Egiazarian, "Single-shot pixel super-resolution phase imaging by wavefront separation approach", Optics Express, 29(26):43662-43678, 2021.

[29] M. Ponomarenko, S.G. Bahnemiri, K. Egiazarian, "Deep Convolutional Network for Spatially Correlated Rayleigh Noise Suppression on TerraSAR-X Images", in Proceeding of IEEE Ukrainian Microwave Week (UkrMW), pp. 458-463, 2020.

[30] Weinberger, M.J., Seroussi, G. & Sapiro, G. The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS. IEEE Transactions on Image processing, 9(8):1309-1324, 2000.