INTERNATIONAL CONFERENCE ON MACHINE VISION

Neural network regularization in the problem of few-view computed tomography

A.V. Yamaev [1,2], M.V. Chukalina [1,4,3], D.P. Nikolaev [1,3], L.G. Kochiev [5], A.I. Chulichkov [2]
[1] Smart Engines Service LLC, Nobelya, 7, Moscow, Russia;
[2] Moscow State University, Michurinsky Pr., 1, Moscow, Russia;
[3] Institute for Information Transmission Problems (Kharkevich Institute) RAS, Bolshoy Karetny Pereulok,12, stroenie 1, Moscow, Russia;
[4] FSRC "Crystallography and Photonics" RAS, Leninski prospekt, 59, Moscow, Russia;
[5] Simon Fraser University, 8888 University Dr, BC V5A 1S6, Burnaby, Canada

Abstract

The computed tomography allows to reconstruct the inner morphological structure of an object without physical destructing. The accuracy of digital image reconstruction directly depends on the measurement conditions of tomographic projections, in particular, on the number of recorded projections. In medicine, to reduce the dose of the patient load there try to reduce the number of measured projections. However, in a few-view computed tomography, when we have a small number of projections, using standard reconstruction algorithms leads to the reconstructed images degradation. The main feature of our approach for few-view tomography is that algebraic reconstruction is being finalized by a neural network with keeping measured projection data because the additive result is in zero space of the forward projection operator. The final reconstruction presents the sum of the additive calculated with the neural network and the algebraic reconstruction. First is an element of zero space of the forward projection operator. The second is an element of orthogonal addition to the zero space. Last is the result of applying the algebraic reconstruction method to a few-angle sinogram. The dependency model between elements of zero space of forward projection operator and algebraic reconstruction is built with neural networks. It demonstrated that realization of the suggested approach allows achieving better reconstruction accuracy and better computation time than state-of-the-art approaches on test data from the Low Dose CT Challenge dataset without increasing reprojection error.

Keywords: computed tomography, few-view tomography, artificial intelligence, neural network, U-Net, learned residual fourier reconstruction.

Citation: Yamaev AV, Chukalina MV, Nikolaev DP, Kochiev LG, Chulichkov AI. Neural network regularization in the problem of few-view computed tomography. Computer Optics 2022; 46(3): 422-428. DOI: 10.18287/2412-6179-CO-1035.

Acknowledgements: This work was partly supported by RFBR (grants) 18-29-26020 and 19-01-00790.

Introduction

X-ray computed tomography (CT) is the widely used non-destructive method for reconstructing the internal structure of an object. In medicine, CT is used for a non-invasive examination of the body structure [1]. This makes it possible to diagnose diseases such as cancer and COVID-19. However, a large absorbed dose of X-ray is harmful to human health. Therefore, it strives to minimize the dose of X-ray radiation received by a patient, but so that, based on the result of the reconstruction, the doctor can determines pathology and sets a diagnosis. The received dose is reduced, for example, in the following two ways: by decreasing the number of measured projections [2] and by decreasing the exposure time of measurement.

Reducing the exposure time of measurement leads to increase noise in the recorded data. As a consequence, this leads to noisy reconstruction. For noise suppression, algorithms of the following classes are used: based on statistical analysis, nonlinear filters, iterative optimization algorithms, and neural networks. With noisy data, neural networks can be as a noise reduction operation after reconstruction [3, 4] (post-processing), as a noise reduction operation on a set of projections (pre-processing) [5], and as a full-reconstruction operator [6, 7]. To solve these problems various neural network architectures are used, for example, convolutional neural networks [8], neural networks operating in wavelet space [9] (post-processing noise suppression), networks operating both in the reconstruction space and measured data space [7], generative neural networks (post-processing noise reduction) [10].

The reconstruction problem in 2D case can be represented as follows. We denote by R the 2D distribution of the linear absorption coefficient of the investigated ob-

ject, A is the linear operator connecting R with a set of projections (sinogram). AR is the sum of linear absorption coefficient along trajectory ray describing the passing X-ray through the object. Since the sinogram is recorded with a certain error v, the measured value is presentedin the following form

$$P = AR + v. \quad (1)$$

The reconstruction goal is to estimate R from P. Typically, the 2D distribution of the linear absorption coefficient is specified in the form of an image represented by a limited number of pixels, the values of which correspond to linear absorption coefficients, and, as rule, it also exceeds the number of registered projections pixels. The values of the measured projections P and the reconstruction values R are assumed to be the Euclidean vectors, then due to the linearity of the transformation A and that dimensionality of P is less than R, the euclidean space of all images can be represented as the sum of zero-space of operator A. $N(A) = x : Ax = 0$ and orthogonal addition $N^{\perp}(A)$ to this space, moreover, N_A is no-zero dimensionality linear space, and R can be represented as the sum

$$R = R_N + R_{\perp}, \quad (2)$$

where $R_N \in N(A)$ and $R_{\perp} \in N^{\perp}(A)$ are orthogonal, moreover,

$$AR_N = 0. \quad (3)$$

Then in (1)

$$P = AR + v = AR_{\perp} + v, \quad (4)$$

where and P does not depend from R_N. This means that if additional information did not used, then by measuring P only the $R_{\perp}$ component of the estimated image R can be reconstructed, and the R_N component cannot be reconstructed, since R_N information is not contained in P. Therefore, for a qualitative reconstruction, the measurement model (1) must be supplemented with some information about the R_N component, if the $R_{\perp}$ estimate obtained from the measurement (1) is known. For this, the model of the dependency between the orthogonal components R_N and $R_{\perp}$ of the image R should be built. When using machine learning methods, we trained this dependency model by the neural network from a set of example pairs of P and R from the Low Dose CT Challenge dataset [11]. In our approach information on the $R_{\perp}$ component of the image R is taken from standard reconstruction methods and supplemented with data on the component R_N, obtained from the $R_{\perp}$ to R_N dependency model constructed by an artificial neural network.

1. Existence of R_N and its properties

If sinogram P is calculated from R in computation experiment with negligible error, then R_N component of the R image can be computed according equation

$$P = AR, \quad (5)$$

then R can be calculated using a pseudo solution of equation (5):

$$R_{\perp} = A^{-}P, \quad (6)$$

where A^{-} is the operator, pseudo-inverse to and $R_{\perp}$ can be obtained as a solution to the minimum problem [12, 13]

$$\inf_R \| P - AR \|^2, \quad (7)$$

where the square of the norm is calculated as the sum of the squares of all coordinates of the vector P . Then the component R_N is calculated as the difference

$$R_N = R - R_{\perp}. \quad (8)$$

To solve the problem (7), for example, the SIRT method is used [14]. In practice, noise-distorted sinograms are often event; it makes problems. There are widely used FBP reconstruction algorithm. In this algorithm, a sinogram is filtered with the ramp filtering operation [15] and the result is backprojected. Another approach based on iterative algorithms for algebraic reconstruction (for example SIRT [16]) is convenient for solving regularized versions of the problem (7), in which the functional

$$\inf_R \| P - AR \|^2 + w \| DR \|^2, \quad (9)$$

where D is linear operator, for example, differentiation, w is normalization coefficient. The solution to the problem (5) can be present in the form

$$R_w = (A^2 + wD^2) - AP. \quad (10)$$

It belongs to $N^{\perp}(A)$ for $D = I$, however, now R_w is not exactly equal to $R_{\perp}$ and differs from $R_{\perp}$, the more, the less $w > 0$. Correctly chosen regularization will allow to obtain the estimation closest to $R_{\perp}$. If there is an error $v0$ when registering the sinogram P in (1) and with a small number of angles, the reconstruction quality turns out to be low when using the FBP algorithm and when using iterative algorithms. The difference between the Phantom and the SIRT reconstruction of the few-angle sinogram of the Phantom is shown in fig. 1. R_N is

$$R_N = Phantom - R_{\perp}. \quad (11)$$

It can be seen that the component R_N is very significant. As the initial data, we used a sinogram obtained by a detector of 512 pixels; measurements are made at 64 angles. The size of the reconstructed image is 512×512. Thus, with a small number of angles and an error, a mathematical model is needed that allows one to estimate the R_N component from the measuring component $R_{\perp}$. In this work, the selection of such a model will be carried out by training a neural network.

2. Training data

To demonstrate the effectiveness of the proposed approach, two neural networks were trained: LPDR and our network. The neural network train- ing was carried out on

an open dataset Low Dose CT Challenge [17]. This dataset consists of X-ray views of the human body from medical tomographs and full and quaternary X-ray reconstructions. In total, the set contains data from 156 patients of various ages and health conditions. Full dose and quaternary dose reconstructions were constructed for one patient at the same time. For each patient, there were around 300 full doses and 300 low doses reconstructed images of the inner structures of their bodies. In addition, these reconstructions are linked with each other layer-by-layer, which makes it possible to train neural networks on the provided set of data pairs. Reconstructions from the Low Dose CT Challenge set are shown in Fig. 1. All reconstructions width in the dataset is 512 pixels.

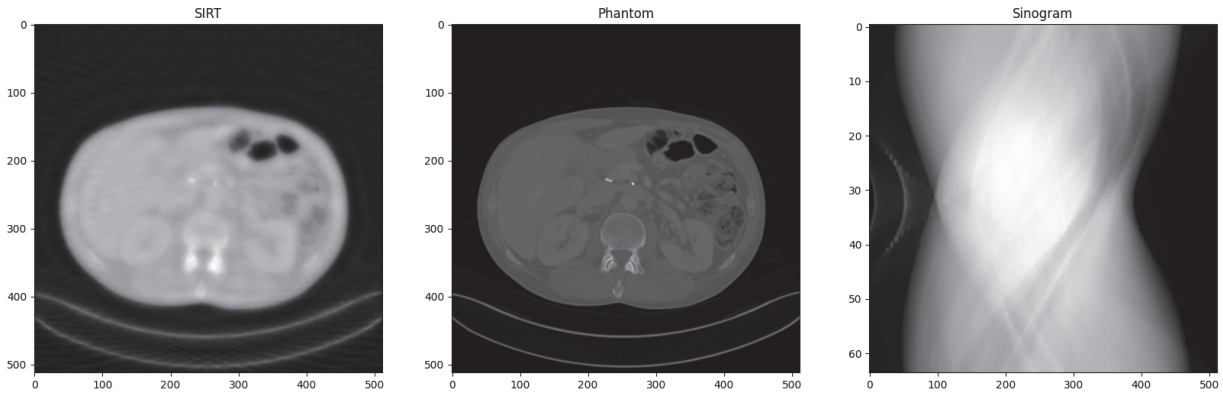

Fig. 1. Phantom - an example of a full-dose reconstruction presented in the dataset Low Dose CT Challenge [12]. SIRT - SIRT reconstruction from sinogram calculated with 64 projections from the Phantom. Sinogram is calculated with 64 projections from the Phantom

To investigate the problem of few-view tomography, sinograms with 64 projections were generated by reprojection from full-dose reconstructions. With the help of the constructed projections, the reconstruction was carried out by the SIRT algorithm. With 20 iterations of the SIRT algorithm, the reprojection error is less than 1e-8 in the $l2$ metric. For a hypothetical 16 bit X-ray detector, their value step is $1/2^{16} = 1.5e-5$. Our reprojection error is less than then a step of a hypothetical 16 bit detector, then on the measured sinogram by this detector, we could not detect reprojection error. Examples of Phantom, SIRT reconstruction from the calculated sinogram, the 64 angles calculated sinogram from the Phantom are shown in fig. 1. Calculated reconstructions and sinograms from full-dose images of the dataset Low Dose CT Challenge were used as data for training neural networks. It can be seen that the SIRT reconstruction in fig. 1 has easily distinguishable artifacts, for example, radial rays coming from the center of the reconstruction and general blurring of the reconstructed image.

3. Suggested approach realization with neural networks

The diagram of an algorithm that implements the proposed approach is shown in fig. 2. It contains three stages: interpolation in Fourier space (Conv Net 1), U-Net, and fix reprojection error in Fourier space (Conv Net 2 and FT Mul 3). All stages contain networks. The SIRT reconstruction is used as network input to calculate an additive to this SIRT reconstruction. According to the Projection Slice theorem [18], projections can be represented in a Fourier space. Each Fourier transformed projection is presented in the form of a straight line on the image in this space and the direction of their line corresponds to the projection angle. 2D inverse Fourier transformed image from this space is reconstructed image. This means, what would a sinogram pixel from the Additive result (fig. 2) has zero value, it needs to the Fourier image of the Additive result must have zero values in lines corresponding to projections of the sinogram by Project Slice theorem. In our approach, we use this corollary of the Projection Slice theorem to normalize U-Net output to the trained pixel-by-pixel normalization FT Mul 3. Normalization is performed by multiplying the input image by pixel by the FT Mul 3 image of the same size. In this case, the values of the pixels of the FT Mul 3 image lie in the range from 0 to 1. This is done to zero out those pixels of the Fourier transform of the additive that correspond to any projection line. Conv Net 1 and Conv Net 2 is a simple stacked convolutional neural network that allows to make local interpolation in Fourier space before U-Net and after U-Net, which improves results. Conv nets outputs are also normalized by pixel-by-pixel trained normalization FT Mul and FT Mul 2, the same as FT Mul 3. Normalization FT Mul and FT Mul 2 are needed to weaken convolutional neural network artifacts that arise due to nonzero values of the Fourier transform near the boundaries of its definition field and for suppressing changing of original data for Conv Net 1 output. The central block of network architecture – U-Net - was selected because this network often is used in image transformation tasks, has low overfitting ability, and can capture features of images in all scales. Last is important because on our SIRT reconstruction there can be artifacts scales to the full image, for example, linear artifacts. The exact architectures of U-Net, ConvNet 1, Con- vNet 2, as well as the entire network, are described using the PyTorch library at the link https://github.com/ayamaev-

se/ComputerOptics LRFR.git. The full neural network was trained using the following loss function

$$Loss\left(R,R_N\right)=\| R-R_\perp-R_N \|^2+0.1\| FP\left(R_N\right)\|^2+ \\ +SSIM(R,R_\perp+R_N), \quad (12)$$

where R_N is the result of the calculated proposed neural network from $R_\perp$, $R_\perp$ is the SIRT reconstruction, R is the phantom, FP is the forward projection operation. The norm from images is the per-pixel $l2$ norm. The neural network was trained using the AdamW optimization algorithm with cosine reduction learning rate [19]. The initial learning rate lr was $1e-3$. Validation loss function stopped increasing after 20 epoch of neural network training. Under our experiments, we observed that there were better metrics values for our approach after training with our loss function. But there were worse metrics values for LPDR with our loss function than the standard MSE loss function. In this case, we used usually the trained LPDR network.

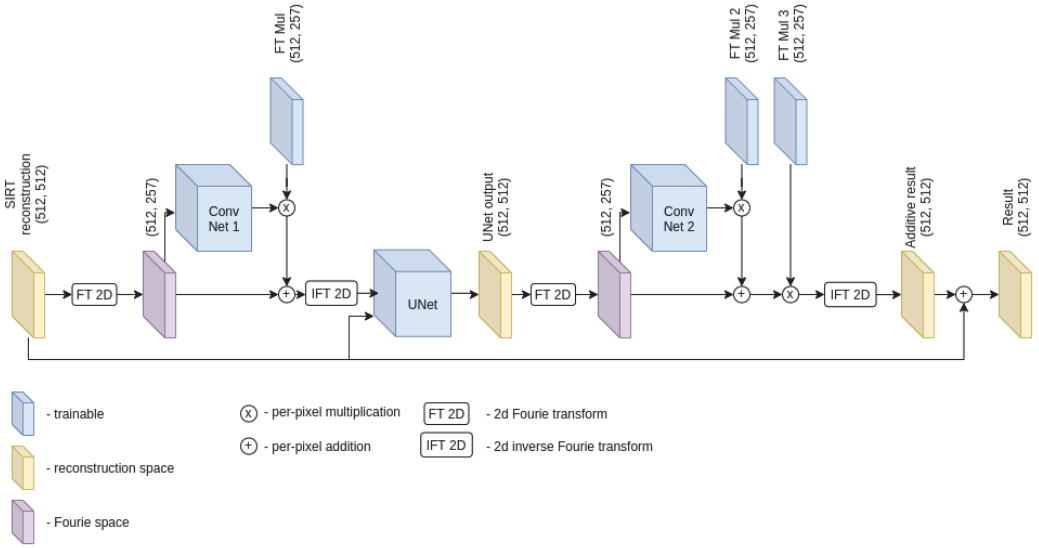

Fig. 2. Architecture of suggested approach

4. Results

On the Low Dose CT Challenge data, neural network from Learned Primal-Dual Reconstruction (LPDR) and a neural network from the pro- posed approach (LRFR - Learned Residual Fourier Reconstruction) were trained for the problem of few-view tomography. Examples of reconstruc- tions obtained using these networks are shown in fig. 3. It can be seen that the LPDR approach slightly retained the radial artifacts associated with low-angle tomography, which is not observed in the reconstruction of the image. Networks were tested on a part of patients which shown in Tab. 1. Also, the calculated metrics of working time and quality on the test sample are indicated in the tab. 2. As can be seen from the tab. 2, the proposed approach shows a higher reconstruction accuracy using the SSIM, PSNR met- rics and a shorter reconstruction time per image. Also, LPFR has a smaller count of additive and multiplication operations, but it has a bigger count of trainable parameters. To measure the reconstruction time, a computer with an AMD Ryzen 7 2700× processor and an Nvidia Titan Xp GPU was used. Fig. 4 shows the trained matrix FT Mul 2. As a result of training, some of the matrix elements FT Mul 2 lying on straight lines were zeroed. These matrix elements correspond to the projection lines from the Projection slice theorem. Therefore, following the theorem, LRFR lies in the zero space of the projection operator with some error. As you can see from the tab. 2, the reprojection error is 3e-6 in the l2 metric.

Tab. 1. Patient codes for neural network testing

Patient code
C002
C012
C027
C050
C067
C081
C099
C111
C121
C128

Tab. 2. Results

Approach	LPDR	LRFR (Ours)
PSNR	37.55	38.06
SSIM	0.879	0.891
Time per image (msec)	154	108
Reprojection MSE	2e – 5	3e – 6
Trainable params	2.5e + 5	7.3e + 6
Operation count	6e + 11	3e + 11

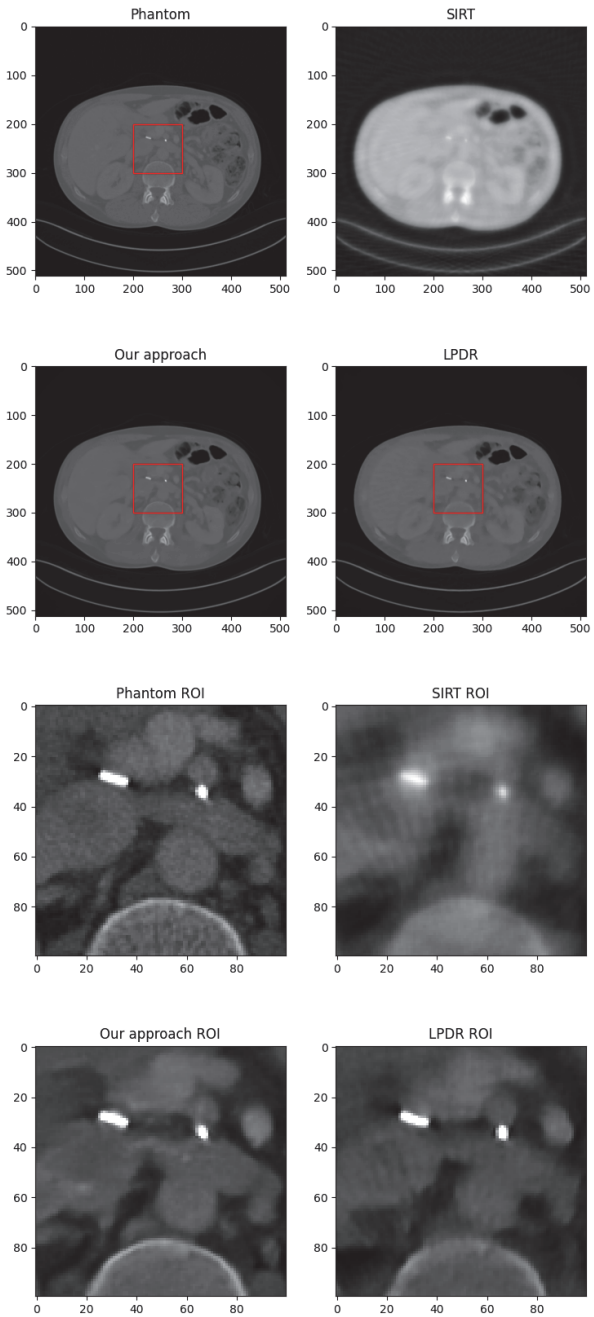

Fig. 3. Results. A slice of the torso is shown in this images. It can be seen that our approach is smoother and detailed than the LPDR result

Conclusion

In this work we presented the approach for constructing reconstructions from the small number of projections. The approach consists in the algebraic reconstruction with subsequent processing by the neural network, which in- cludes the computation in Fourier space. The reprojection error of our result is $3e-6$, which is less than $1.5e-5$, the step of a hypothetical 16-bit X-ray detector. Consequently, there will be no differences between the sinogram constructed by reprojection from the reconstruction and the sinogram measured by a detector. The implementation of the proposed approach shows the state-of-the-art results of reconstructions according to the PSNR and SSIM metric and has a shorter running time than the similar approach. Moreover, due to a specially selected architecture that minimizes the reprojection error, the proposed approach is fundamentally less inclined to create non-existent objects for reconstruction. Since the reprojection error of our approach is less than the step of the hypothetical 16-bit X-ray detector $step = 1/2^{16}$, the neural network's response lies in the zero space of this hypothetical detector. We suppose, that our network is useful for medical application because small reprojection error makes the network safer to add and remove elements in the reconstructed image displayed in the original projections. At the moment, about 80 percent of LRFR's time is spent on creating SIRT reconstruction. In the future, we plan to explore the possibility of creating fast reconstruction algorithms for few-view tomography with zero reprojection error, if the pro- posed method would be enhanced in terms of performance. For example, it can be single step algorithms or accelerated SIRT by fast linear summations [20].

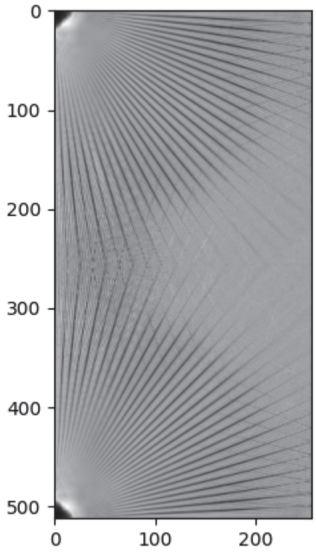

Fig. 4. FT Mul 3 values from network architecture. It can be seen that the image contains zero lines that it is agreed with the Projection Slice theorem

Discussion

The suggested approach is based on the Projection Slice theorem and theory of Measuring computing systems. Realization of the approach gives around zero error in reprojection by per-pixel normalization on the Fourier space. Also, it contains interpolating in Fourier and Reconstruction spaces.

Its first approach, which works in zero space of the forward projection operator and in Fourier space of Reconstruction space.

References

[1] Chukalina MV, Khafizov AV, Kokhan VV, Buzmakov AV, Senin RA, Uvarov VI, Grigoriev MV. Algorithm for post-processing of tomography images to calculate the di-

mension-geometric features of porous structures. Computer optics 2021; 45(1): 110-121. DOI: 10.18287/2412-6179-CO-781.
[2] Bulatov K, Chukalina M, Buzmakov A, Nikolaev D, Arlazarov VV. Monitored reconstruction: Computed tomography as an anytime algorithm. IEEE Access 2020; 8: 110759-110774. DOI: 10.1109/ACCESS.2020.3002019.
[3] Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, Kalra MK, Zhang Y, Sun L, Wang G. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. IEEE Trans Med Imaging 2018; 37(6): 1348-1357. DOI: 10.1109/TMI.2018.2827462.
[4] Zhu L, Han Y, Li L, Xi X, Zhu M, Yan B. Metal artifact reduction for X-ray computed tomography using U-net in image domain. IEEE Access 2019; 7: 98743-98754. DOI: 10.1109/ACCESS.2019.2930302.
[5] Yamaev A, Chukalina M, Nikolaev D, Sheshkus A, Chulichkov A. Lightweight denoising filtering neural network for FBP algorithm. Proc SPIE 2021; 11605: 116050L. DOI: 10.1117/12.2587185.
[6] Yang H-K, Liang K-C, Kang K-J, Xing Y-X. Slice-wise reconstruction for low-dose cone-beam CT using a deep residual convolutional neural network. Nucl Sci Tech 2019; 30(4): 59. DOI: 10.1007/s41365-019-0581-7.
[7] Adler J, Öktem O. Learned primal-dual reconstruction. IEEE Transactions on Medical Imaging 2018; 37(6): 1322-1332. DOI: 10.1109/TMI.2018.2799231.
[8] Nakai H, Nishio M, Yamashita R, Ono A, Nakao KK, Fujimoto K, Togashi K. Quantitative and qualitative evaluation of convolutional neural networks with a deeper u-net for sparse-view computed tomography reconstruction. Acad Radiol 2020; 27(4): 563-574.
[9] Kang E, Ye JC, et al. Wavelet domain residual network (WavResNet) for low-dose X-ray CT reconstruction. arXiv preprint, 2017. Source: ⟨https://arxiv.org/abs/1703.01383⟩.
[10] Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X, Curran WJ, Liu T, Yang X. Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. Med Phys 2019; 46(9): 3998-4009. DOI: 10.1002/mp.13656.
[11] McCollough CH, Bartley AC, Carter RE, Chen B, Drees TA, Edwards P, R Holmes III DR, Huang AE, Khan F, Leng S, McMillan KL, Michalak GJ, Nunez KM, Yu L, Fletcher JG. Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 low dose CT grand challenge. Med Phys 2017; 44(10): e339-e352. DOI: 10.1002/mp.12345.
[12] Pyts'ev YuP. Measurement-computation converter as a measurement facility. Autom Remote Control 2010; 71(2): 303-319. DOI: 10.1134/S0005117910020116.
[13] Buzmakov A, Zolotov D, Chukalina M, Ingacheva A, Sheshkus A, Asadchikov V. Iterative tomography reconstruction in a limited field of view. Proc SPIE 2020; 11433: 114331W. DOI: 10.1117/12.2557501.
[14] Usanov MS, Kulberg NS, Morozov SP. Usage of adaptive homomorphic filters for CT processing [In Russian]. Informatsionnye Tekhnologii i Vychslitel'nye Sistemy 2017; 2: 33-42.
[15] Wei Y, Wang G, Hsieh J. An intuitive discussion on the ideal ramp filter in computed tomography (I). Comput Math with Appl 2005; 49(5-6): 731-740. DOI: 10.1016/j.camwa.2004.10.034.
[16] Su BL, Zhang YH, Peng LH, Yao DY, Zhang BF. The use of simultaneous iterative reconstruction technique for electrical capacitance tomography. Chem Eng J 2000; 77(1-2): 37-41. DOI: 10.1016/S1385-8947(99)00134-5.
[17] McCollough C. TU-FG-207A-04: Overview of the low dose CT grand challenge. Med Phys 2016; 43(6Part35): 3759-3760. DOI: 10.1118/1.4957556.
[18] Devaney AJ. Generalized projection-slice theorem for fan beam diffraction tomography. Ultrasonic Imaging 1985; 7(3): 264-275. DOI: 10.1016/0161-7346(85)90006-9.
[19] Gotmare A, Keskar NS, Xiong C, Socher R. A closer look at deep learning heuristics: Learning rate restarts, warmup and distillation. arXiv preprint, 2018. Source: ⟨https://arxiv.org/abs/1810.13243⟩.
[20] Soshin KV, Nikolaev DP, Gladilin SA, Ershov EI. Acceleration of summation over segments using the fast hough transformation pyramid [In Russian]. Vestnik Yuzhno-Ural'skogo Universiteta. Seriya Matematicheskoe Modelirovanie i Programmirovanie 2020; 13(1): 129-140. DOI: 10.14529/mmp200110.

Authors' information

Andrei Viktorovich Yamaev, (b. 1996) is PhD student in Moscow State University from 2018. Currently he works as the researcher at the "Smart Engine Service Ltd" company. Research interests are neural net-works, computed tomography, programming. E-mail: *rewin1996@gmail.com* .

Marina Valerievna Chukalina, (b.1965) graduated from Moscow Engineering Physics Institute in 1987, majoring in Mathematics. Currently she works as the senior researcher at the Institute of Crystallography – Branch of the FSRC "Crystallography and Photonics" RAS. Research interests are direct and inverse problems in computer tomography.
E-mail: *chukalinamarina@gmail.com* .

Dmitry Petrovich Nikolaev, (b. 1978), Ph. D. in Physics and Mathematics, a head of the laboratory at the IITP RAS. Graduated from Lomonosov M.V. in 2000. Research interests are machine vision, algorithms for fast image processing, pattern recognition. E-mail: *dimonstr@iitp.ru* .

Leon Guramievich Kochiev, (b.1998), studied at Moscow Institute of Physics and Technologies and Simon Fraser University. Works as a machine learning engineer at IntegraNT. Research interests: computer vision, 3D computer vision, medical imaging, natural language processing. Email: *kochiev.lg@phystech.edu* .

Alexey Ivanovich Chulichkov, (b. 1954) graduated from Physics faculty at Lomonosov Moscow State University in 1978. Received PhD degree in Physics and Mathematics from Lomonosov Moscow State University in 1983 and Doctor of Sciences degree in 1993. Professor of the Mathematical Modeling and Informatics department at the Physics faculty MSU from 1994 to 2019. Head of the Mathematical Modeling and Informatics department since 2020. Research interests: morphological methods of image analysis, fuzzy and indefinite fuzzy mathematics, analysis and interpretation of experimental data. Author of over 200 articles and 13 books. Email: *achulichkov@gmail.com* .

Received September 01, 2021. The final version – January 17, 2022.