

# Исследование мультиапертурных систем на основе дифракционной оптики и глубокого обучения

А.В. Никоноров<sup>1,2</sup>, Р.В. Скиданов<sup>1,2</sup>, В.В. Кутикова<sup>1,2</sup>, М.В. Петров<sup>1,2</sup>,  
А.П. Алексеев<sup>1,2</sup>, С.А. Бибилов<sup>1,2</sup>, Н.Л. Казанский<sup>1,2</sup>

<sup>1</sup>Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

<sup>2</sup>Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

**Аннотация.** Использование мультиапертурных камер - один из основных трендов в современных устройствах как бытовых, так и профессиональных. В настоящей работе предлагается создание мультиапертурных камер на основе длиннофокусных однолинзовых дифракционных объективов. Такие объективы в разы превосходят рефракционные по весу и стоимости, однако существенно уступают по качеству получаемого изображения, и поэтому нуждаются в применении программной реконструкции. В данной работе предложены различные схемы мультиапертурных дифракционных объективов, позволяющие увеличить как угол обзора, так и разрешение системы. Предложена сверточная нейронная сеть реконструкции изображений в мультиапертурных дифракционно-оптических системах.

## 1. Введение

Для изображающих оптических систем характерен компромисс между углом обзора и линейным разрешением системы. Широкоугольные системы обеспечивают низкое линейное разрешение, а длиннофокусные, высокоразрешающие системы обладают небольшим углом обзора. В последнее время появился ряд мультисенсорных и мультиобъективных систем, созданных чтобы избежать указанного компромисса. На рынке систем видеонаблюдения это мультифокальная матричная камера DallmeyerPanomera, а на рынке бытовых фотокамер – запатентованная четырехобъективная камера Nikon [3], а также камера Light L16 разработанная стартапом Light в 2016 году [4]. Камера Light L16 согласно опубликованным характеристикам обеспечивает 50 Мп разрешение в сочетании с огромной светочувствительностью за счет того, что использует при съемке 16 камер. Многообъективные конфигурации используются, как правило, для увеличения угла обзора при сохранении разрешения, для решения задач сверхразрешения по нескольким изображениям [2], [4], а также для более специфических задач, таких как построение разнофокальных систем и систем постфокусировкой [1], [2]. Несмотря на интересные перспективы использования таких систем масса объектива является существенным ограничением при построении многоапертурных конфигураций.

Преодолеть ограничение, связанное с массой объектива, позволяют системы на основе плоской дифракционной оптики. Успехи в развитии подобных систем продемонстрированы в работах

последних лет начиная с 2015 года [5]. Такая оптика позволяет существенно сократить вес и стоимость длиннофокусной оптической системы [6].

В настоящей работе предложены различные возможные конфигурации многоапертурных систем на основе однолинзовых дифракционных объективов. Отдельно рассмотрены оптические схемы для конфигурации, направленной на увеличение углового зрения системы, и для решения задачи сверхразрешения по нескольким изображениям. Задача сверхразрешения решается на основе глубокой сверточной нейронной сети.

## **2. Оптические схемы мультиапертурных систем технического зрения на основе дифракционной оптики**

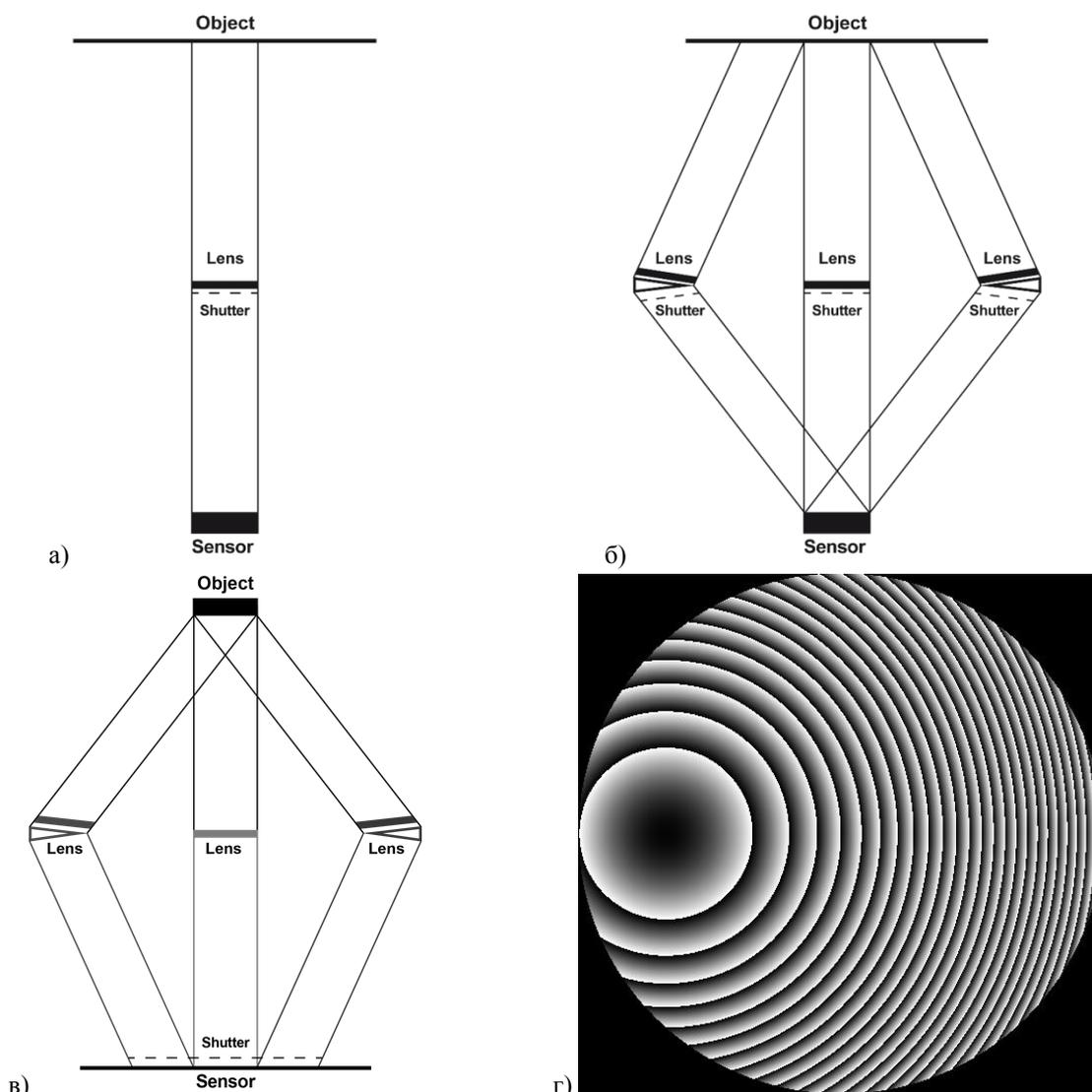
Как показано в ряде работ, дифракционная оптика позволяет многократно сократить вес длиннофокусных объективов [13]. Такие системы крайне актуальны для систем космического ДЗЗ, легких БПЛА, систем видеонаблюдения [6]. Для мультиапертурных систем вес оптической части имеет критическое значение. Так в работе [14] рассмотрено сокращение длины оптического тракта в четыре раза за счет использования девяти апертур в ИК диапазоне. Однако сложность изготовления и общий вес системы привел к тому, что авторами приведены только теоретические расчеты такой системы. В настоящей работе предлагаются две схемы мультиапертурных систем на основе дифракционной оптики направленные на расширение угла обзора системы и на повышение разрешения системы.

Принципиальные схемы объективов приведены на рисунке 1. В качестве базовой оптической схемы на рисунке 1а) приведен однолинзовый дифракционный объектив. Такой длиннофокусный дифракционный объектив является составным компонентом двух других оптических схем.

Общая схема многозатворного мультиапертурного объектива продемонстрирована на рисунке 1б). В этой схеме на сенсор попадают изображения трех участков наблюдаемого объекта, через три дифракционных линзы. С центральной линзы изображение участка объекта получается сенсором напрямую, тогда как с двух других линз получается отклоненное от оси изображение. Затворы после линз срабатывают последовательно, таким образом на сенсоре последовательно образуются изображения трех участков объекта, формируя составной кадр, из отдельных кадров, зарегистрированных в последовательные промежутки времени.

В предлагаемой схеме используется особенность дифракционного объектива, состоящая в том, что оптический путь в таком объективе на 99,9% свободен от оптических элементов. Таким образом, при компоновке сложной изображающей системы оптические пути могут пересекаться, что невозможно в обычном объективе, состоящем из большого числа рефракционных оптических элементов [14]. Также используется тот факт, что при съемке 25 кадров/сек формирование каждого кадра происходит в течении 40 мс. При этом характерные величины выдержки, достаточные для видеосъемки на открытом пространстве составляют 2-5 мс. При такой съемке, многообъективная конфигурация может обеспечивать формирование составного кадра из 8-20 отдельных кадров, составные кадры же будут формироваться с частотой 25 кадров/сек. Отклонение изображения в таких системах выполняется по ранее полученной методике [15], [6], пример рельефа дифракционного элемента с отклонением приведен на рисунке 1г).

На рисунке 1в) приведена принципиальная схема многоапертурного объектива направленного на повышение разрешения изображения за счет использования подхода многокадрового сверхразрешения [13]. Как и в работе [14], в такой схеме каждый отдельный объектив мультиапертурной схемы обеспечивает регистрацию частного изображения на части сенсора. При этом общее изображение с более высоким разрешением получается в результате вычислительной реконструкции. В настоящей работе рассматривается реконструкция изображений с применением сверточной нейронной сети.

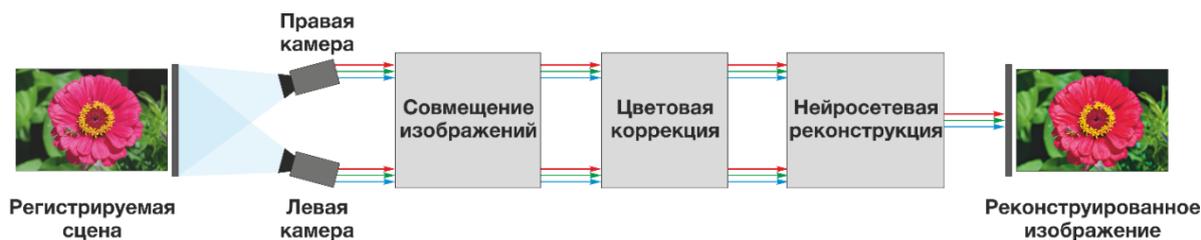


**Рисунок 1.** Схемы одно и многоапертурной дифракционно-оптических систем: а) базовый однолинзовый дифракционный объектив; б) мультиапертурный мультизатворный объектив; в) мультиапертурный объектив; г) дифракционный объектив, обеспечивающий формирования отклоненного от оси изображения.

**3. Реконструкция изображений в мультиапертурных дифракционно-оптических системах**

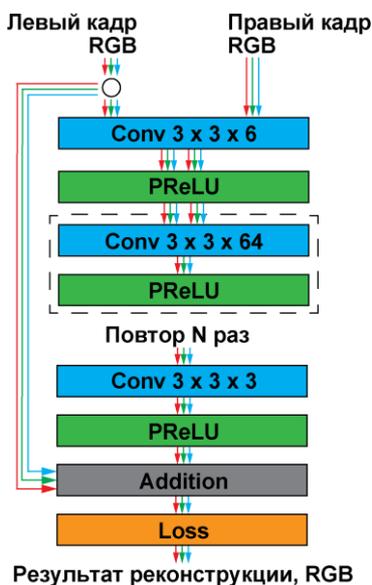
В качестве прототипа многоапертурной системы в рамках настоящего проекта использовалась оптическая схема, при которой съемка велась двумя независимыми камерами, оснащенными однолинзовыми дифракционными объективами. Зарегистрированные изображения имеют геометрические искажения проективного характера. Коррекция таких искажений осуществляется обратным проективным преобразованием, выполняемым на основе матрицы внутренней калибровки двухобъективной системы. Оба зарегистрированных изображения приводятся в систему координат левого изображения.

Для решения этой задачи была использована модификация СНС разработанной в работах [6] и [12] для повышения разрешения цветных и гиперспектральных изображений. На входной слой сети подаются два зарегистрированных изображения, что составляет 6 цветовых каналов, рисунок 2. Цветовая коррекция предваряет этап нейросетевой реконструкции и выполняется также, как было описано в работе [6]. На выходном слое сети получаются три цветовых канала результирующего изображения высокого разрешения. Обучение сети ведется с использованием функции потерь, которая, как и в работе [6] учитывает межканальную связь.



**Рисунок 2.** Схема регистрации и реконструкции изображения в мультиапертурной системе.

Архитектура СНС основана на архитектуре сети VDSR, описанной в работе [8]. Замена функции активации ReLU на PReLU [9], а также сочетание архитектуры VDSR и функции потерь с учетом штрафа (grey-edgepenalty) за отклонение контура от ахроматичности позволили достичь хороших результатов в [7]. Трехкомпонентная СНС, принимающая на вход одно RGB изображение, обеспечивает лучшее качество реконструкции по сравнению с однокомпонентной СНС. Однако дальнейшее совершенствование архитектуры сети (рисунок 3) показало, что наиболее эффективной с точки зрения качества восстановления изображения является шестикомпонентная СНС.



**Рисунок 3.** Архитектурасверточной нейронной сети реконструкции изображений в мультиапертурных системах.

На последнем слое сети происходит сложение выхода предыдущего слоя, который состоит из трех компонент, и входного изображения, состоящего из шести компонент. В данной работе для сложения используется только одно (левое) входное RGB изображение.

Функция потерь использует межканальную связь, основанную на гипотезе об ароматичности контуров [10], и имеет следующий вид [7]:

$$L(\hat{p}, p^0, p_G^0, w) = \frac{1}{N} \sum_{j=1}^N d(\hat{p}^{(j)} - p^{0(j)}) + \lambda \frac{1}{N} \sum_{j=1}^N \|p_G^{0(j)} \nabla \hat{p}^{(j)} - \hat{p}^{(j)} \nabla p_G^{0(j)}\|_2 + \eta \|w\|_2$$

где  $N$  – количество изображений в одном минибатче обучения,  $j$  – индекс изображения в выборке,  $p^{0(j)}$  – эталонное RGB изображение,  $\hat{p}^{(j)}$  –восстановленное RGB изображение,  $p_G^{0(j)}$  - зеленый канал эталонного изображения. Первое слагаемое функции потерь обеспечивает минимизацию расстояния  $d()$  между эталонным изображением и изображением, полученным на выходе сети. Второе слагаемое реализует штраф за отклонение градиента от зеленого канала

$P_G^{0(j)}$  эталонного изображения. Последнее слагаемое накладывает ограничения на норму весовых коэффициентов сети (weight decay). Качество реконструкции оценивается на основе пикового соотношения сигнал-шум (PSNR), которое успешно использовалось в работе [7].

#### 4. Результаты экспериментов

Набор данных включает 148 пар изображений (изображения, снятые с двух разных позиций), над которыми предварительно проведена цветовая коррекция [7], и соответствующие эталонные изображения. Для увеличения объема выборки применялись зашумление и поворот имеющихся изображений. Перед обучением нейронной сети изображения разбивались с шагом 1 на фрагменты размера 41x41. Таким образом, объем обучающей выборки составил 185328 пар фрагментов изображений.

В качестве метода оптимизации в данной работе был использован алгоритм Adam[11]. Размер пакета был равен 64, а начальная скорость обучения составляла  $10^{-5}$  и каждые 10 эпох уменьшалась в два раза. В соответствии с исследованиями, проведенными в работе [7] коэффициент скорости обучения  $\lambda$ , параметр функции потерь  $\epsilon$  и коэффициент, ограничивающий норму весовых коэффициентов,  $\eta$  (weight decay) в начале обучения были установлены равными  $10^{-2}$ ,  $10^{-4}$ , и  $5 \cdot 10^{-2}$  соответственно, а в конце обучения  $\eta$  уменьшался до  $10^{-4}$ .

Также в результате экспериментов оказалось, что лучшее качество и стабильную сходимость при обучении обеспечивает нейронная сеть, состоящая из 13 сверточных каскадов, то есть в схеме сети на рисунке 3 центральный блок из сверточного каскада  $3 \times 3 \times 64$  повторяется 13 раз.

В таблице 1 приведены средние значения PSNR для пяти тестовых изображений после обработки на основе трех- и шестикомпонентных СНС, а также после цветовой коррекции для первых 8 эпох обучения. Из таблицы видно, что шестикомпонентная СНС достигает 25,51 Дб, что превышает значение PSNR, полученное с помощью трехкомпонентной сети почти на 2 Дб.

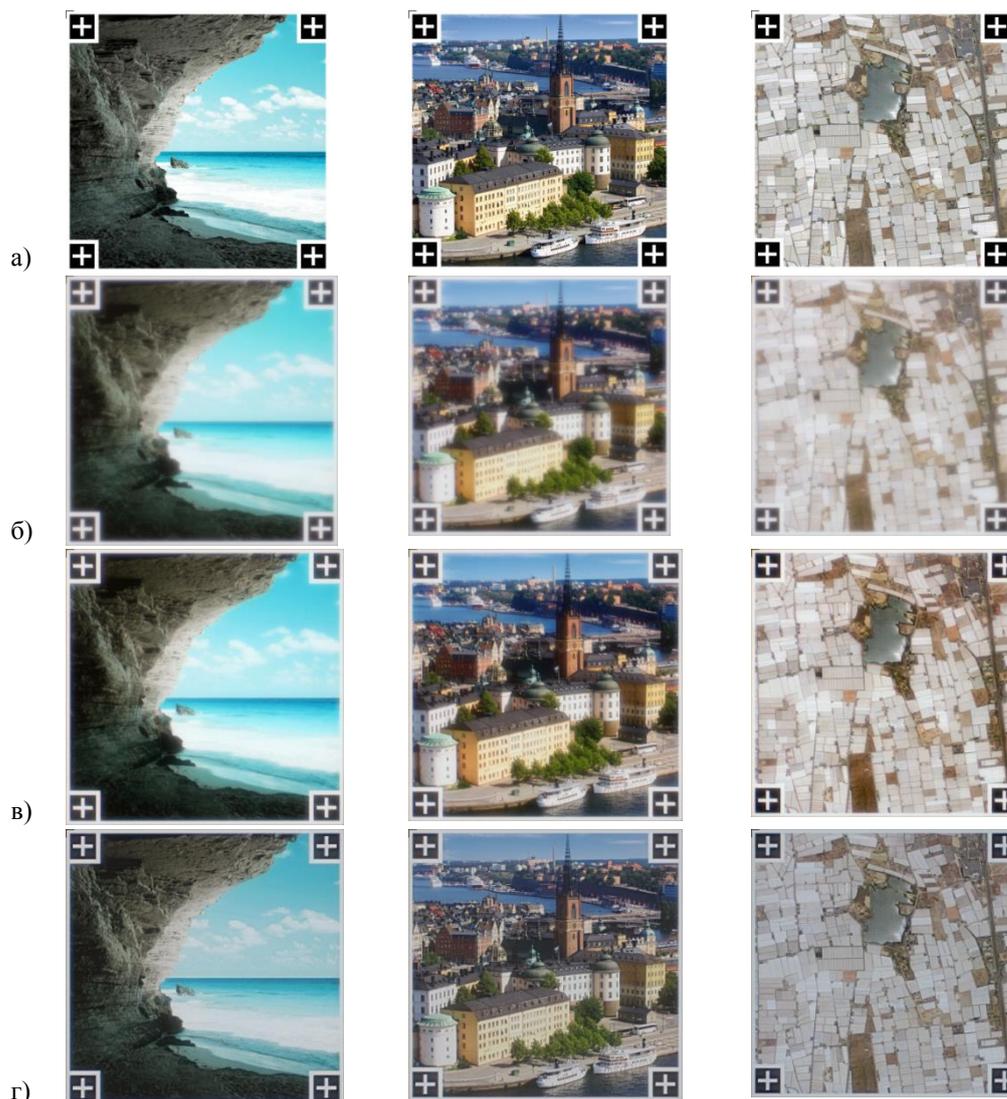
На рисунке 4 представлены примеры зарегистрированных и восстановленных с помощью СНС изображений, а также изображения после цветовой коррекции и соответствующие им эталонные изображения. Видно, что шестикомпонентная СНС обеспечивает лучшее визуальное качество реконструкции по сравнению с трехкомпонентной СНС.

**Таблица 1.** Результаты реконструкции изображений на основе СНС.

Эпоха	PSNR после цветовой коррекции		PSNR после СНС	
	Входное изображение 1	Входное изображение 2	Трехкомпонентная	Шестикомпонентная
0	21,81	21,81	23,72	19,51
1				23,89
2				24,61
3				25,15
4				<b>25,51</b>
5				25,08
6				23,21
7				25,43

#### 5. Заключение

В работе проведено исследование мультиапертурных систем на основе однолинзовых дифракционных объективов. Такие системы призваны преодолеть необходимость компромисса между углом обзора и линейным разрешением оптической системы. Предложены две базовых схемы мультиапертурных систем. Первая схема с несколькими затворами позволяет расширить угол обзора системы. Вторая, мультиобъективная система, позволяет решить задачу сверхразрешения по нескольким кадрам.



**Рисунок 3.** Примеры изображений – а) эталонное; б) после цветовой коррекции; в) после трехкомпонентной СНС; г) после шестикомпонентной СНС.

Решение задачи сверхразрешения по нескольким кадрам было проведено для прототипа двухапертурной системы. Для реконструкции двухкадровой съемки была предложена архитектура шестикомпонентной сверточной нейронной сети для восстановления изображений. Уровень PSNR, достигнутый с помощью шестикомпонентной СНС, на 2 Дб превышает значение PSNR, которое обеспечивает трехкомпонентная СНС по одному кадру. Также шестикомпонентная СНС обеспечивает лучшее визуальное качество реконструкции по сравнению с трехкомпонентной СНС. В результате экспериментов оказалось, что лучшее качество и стабильную сходимость при обучении обеспечивает нейронная сеть, состоящая из 13 сверточных каскадов.

Предварительные исследования показывают принципиальную возможность повышения качества получаемых изображений за счет увеличения числа снятых кадров. Разработка таких систем, а также исследование возможности многокадрового сверхразрешения в движущихся системах, таких как БПЛА является перспективным направлением дальнейших исследований. Также вопросом дальнейшего изучения является производительность системы и более подробный анализ разрешения с использованием таких критериев как частотно-контрастная характеристика.

## 6. Литература

- [1] Tanida, J. Thin observation module by bound optics (TOMBO): concept and experimental verification / J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, Y. Ichioka // *Applied Optics*. – 2001. – Vol. 40(11). – P. 1806-1813.
- [2] Tanida, J. Multi-aperture optics as a universal platform for computational imaging // *Optical Review*. – 2016. – Vol. 23(5). – P. 859-864.
- [3] Патент Японии № JP2013081156A, 09.04.2013 // 欣也加藤, 欣也加藤 Imaging device, 2013.
- [4] Sahin, F. Light L16 Computational Camera / F.E. Sahin, R. Laroia // *Imaging and Applied Optics (3D, AIO, COSI, IS, MATH, pcAOP)*. – OSA Technical Digest (online), 2017. – P. JTu5A.20.
- [5] Nikonorov, A. Fresnel Lens Imaging with Post-Capture Image Processing / A. Nikonorov, R. Skidanov, V. Fursov, M. Petrov, S. Bibikov, Y. Yuzifovich // *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2015. – P. 33-41.
- [6] Nikonorov, A. Toward Ultralightweight Remote Sensing With Harmonic Lenses and Convolutional Neural Networks / A.V. Nikonorov, M.V. Petrov, S.A. Bibikov, P.Y. Yakimov, V.V. Kutikova, Y.Y. Yuzifovich, A.A. Morozov, R.V. Skidanov, N.L. Kazanskiy // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2018. – Vol. 11(9). – P. 3338-3348.
- [7] Никоноров, А.В. Реконструкция изображений в дифракционно-оптических системах на основе сверточных нейронных сетей и обратной свертки / А.В. Никоноров, М.В. Петров, С.А. Бибииков, В.В. Кутикова, А.А. Морозов, Н.Л. Казанский // *Компьютерная оптика*. – 2017. – Т. 41, № 6. – С. 875-887. DOI: 10.18287/2412-6179-2017-41-6-875-887.
- [8] Kim, J. Accurate image super-resolution using very deep convolutional networks / J. Kim, J.K. Lee, K.M. Lee // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. – P. 1646-1654. DOI: 10.1109/CVPR.2016.182.
- [9] He, K. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification / K. He, X. Zhang, S. Ren, J. Sun // *IEEE International Conference on Computer Vision (ICCV)*, 2015. – P. 1026-1034. DOI: 10.1109/ICCV.2015.123.
- [10] Van de Weijer, J. Edge-Based Color Constancy / J. van de Weijer, T. Gevers, A. Gijsenij // *IEEE Transactions on Image Processing*. – 2007. – Vol. 16(9). – P. 2207-2214.
- [11] Kingma, D.P. Adam: A method for stochastic optimization / D.P. Kingma, J.L. Ba // *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.
- [12] Nikonorov, A. Deep Learning-Based Enhancement of Hyperspectral Images Using Simulated Ground Truth / A. Nikonorov, M. Petrov, S. Bibikov, V. Kutikova, P. Yakimov, A. Morozov, R. Skidanov, N. Kazanskiy // *10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS)*, 2018. – P. 1-9.
- [13] Tao, X. Detail-revealing Deep Video Super-resolution / X. Tao, H. Gao, R. Liao, J. Wang, J. Jia // *IEEE International Conference on Computer Vision (ICCV)*, 2017. – P. 4482-4490.
- [14] Carles, G. Compact multi-aperture imaging with high angular resolution / G. Carles, G. Muyo, N. Bustin, A. Wood, A.R. Harvey // *Journal of the Optical Society of America A*. – 2015. – Vol. 32(3). – P. 411-419.
- [15] Nikonorov, A. Comparative Evaluation of Deblurring Techniques for Fresnel Lens Computational Imaging / A. Nikonorov, M. Petrov, S. Bibikov, Y. Yuzifovich, P. Yakimov, N. Kazanskiy, R. Skidanov, V. Fursov // *International Conference on Pattern Recognition (ICPR)*, 2016. – P. 775-780.

## Благодарности

Работа выполнена при поддержке грантов РФФИ (проекты 16-47-630721 p\_a, № 17-29-03112-офи-м, № 18-07-01390-А, № 16-29-09528-офи\_м).

# Multi-aperture systems on the base of diffractive lens and deep neural networks

A.V. Nikonorov<sup>1,2</sup>, R.V. Skidanov<sup>1,2</sup>, V.V. Kutikova<sup>1,2</sup>, M.V. Petrov<sup>1,2</sup>, A.P. Alekseyev<sup>1,2</sup>, S.A. Bibikov<sup>1,2</sup>, N.L. Kazanskiy<sup>1,2</sup>

<sup>1</sup>Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

<sup>2</sup>Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

**Abstract.** The use of multi-aperture cameras is one of the main trends in modern devices, both casual and professional. This paper presents the creation of multi-aperture cameras based on long-focus simple diffraction lenses. Such lenses are several times better than the common refraction lenses in terms of weight and cost, but they are significantly inferior in quality of the resulting image, and therefore they require special image enhancement. We introduce various schemes of multi-aperture diffraction lenses, allowing to increase both the viewing angle and the resolution of the imaging system. We propose a convolutional neural network for reconstruction of images in multi-aperture diffraction optical systems.