

Получение цветных изображений системой на основе трех дифракционных линз

С.О. Степаненко^{1,2}, В.В. Евдокимова^{1,2}, М.В. Петров^{1,2}, Р.В. Скиданов^{1,2}, А.В. Никоноров^{1,2}

¹ Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Россия, г. Самара, Московское шоссе, д. 34;

² Институт систем обработки изображений РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН, 443001, Россия, г. Самара, ул. Молодогвардейская, д. 151

Аннотация

Возможность существенно снизить массу и стоимость систем технического зрения приводит к появлению большого числа работ, посвященных разработке новых изображающих систем на основе дифракционной оптики. В рамках настоящей работы предложена новая изображающая система из трех дифракционных линз, каждая из которых формирует один канал цветного RGB-изображения. Такой подход позволяет существенно сузить спектральный диапазон каждой линзы и таким образом существенно снизить искажения изображения, вызванные хроматической аберрацией, присущей дифракционной оптике. Показано, что такая схема позволяет провести нейросетевую реконструкцию, существенно повышающую качество получаемого изображения. Для минимизации артефактов реконструкции на изображениях реальных сцен в работе используется аугментация обучающей выборки, а также предложен критерий ложных контуров, позволяющий учитывать реальные данные в процессе обучения сети.

Ключевые слова: цветовая коррекция, цветовая стабилизация, глобальная оптимизация, нейросетевая реконструкция изображений.

Цитирование: Степаненко С.О. Получение цветных изображений системой на основе трех дифракционных линз / С.О. Степаненко, В.В. Евдокимова, М.В. Петров, Р.В. Скиданов, А.В. Никоноров // Компьютерная оптика. – 2023. – Т. 47, № 5. – С. 716-724. – DOI: 10.18287/2412-6179-CO-1258.

Citation: Stepanenko SO, Evdokimova VV, Petrov MV, Skidanov RV, Nikonorov AV. Color imaging using a system based on 3 diffractive lenses. Computer Optics 2023; 47(5): 716-724. DOI: 10.18287/2412-6179-CO-1258.

Введение

В настоящее время все больше исследований проводится в сфере создания систем технического зрения на основе плоских оптических элементов, получаемых на основе дифракционной нанофотоники или металлинз [1–14]. Использование плоской оптики позволяет значительно сократить вес и упростить конструкцию оптической системы, что является большим преимуществом перед классическими рефракционными аналогами. Особенно заметно такое преимущество перед рефракционной оптикой в многолинзовых системах [4, 8].

Многолинзовые системы на основе дифракционной оптики исследованы в целом ряде работ. Использование дифракционной оптики в составе многолинзовых систем, как правило, связано с требованием миниатюризации и снижения веса [11–14]. В работе [11] также показана возможность существенного сокращения габаритов системы за счет использования дифракционной оптики с переотражением между двумя поверхностями. В работах [15, 16] показаны преимущества дифракционно-оптических систем со спектральным разделением, позволяющие решать задачи ДЗЗ, такие как расчет вегетативных индексов [16] и индексов влажности [15]. В работе [4] впервые

показано, что использование бинокулярной дифракционно-оптической системы с нейросетевой реконструкцией изображений позволяет существенно повысить качество получаемых изображений.

Проблема снижения качества изображений при использовании дифракционных оптических элементов возникает вследствие присущих таким системам сильнейших хроматических аберраций. Хроматические аберрации для дифракционной оптики, как правило, невелики в узком спектральном диапазоне, близком к расчетной длине волны, и линейно возрастают по мере удаления от нее [14, 8]. Такие искажения могут быть в значительной степени устранены как оптическими [2, 3], так и алгоритмическими методами [6, 1], в том числе за счет сквозной нейросетевой реконструкции изображений [4, 5]. В работе [4] показано, что в комбинации с бинокулярной системой такой подход позволяет получать существенное улучшение качества.

В настоящей работе предложена новая оптическая схема, в которой для получения цветного RGB-изображения используются три дифракционных линзы с базовыми длинами волн в синем, красном и зеленом диапазонах. Таким образом, цветное изображение получается в трех узких спектральных диапазонах с уменьшенным влиянием хроматической абер-

рации. Далее три полученных канала алгоритмически совмещаются в единое изображение, для которого выполняется нейросетевая реконструкция.

1. Формирование изображений в трехлинзовой системе на основе дифракционной оптики

1.1. Оптическая схема и расчет системы

Мультиапертурная система состоит из трех дифракционных линз с фокусным расстоянием 100 мм, каждая из которых рассчитана на свой спектральный диапазон (рис. 1).

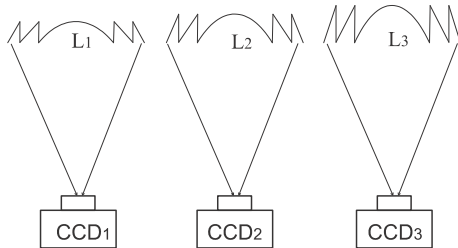


Рис. 1. Оптическая схема мультиапертурной системы

Рабочая длина волны каждой дифракционной линзы задается высотой микрорельефа.

Так, для формирования изображения на длине волны λ высота микрорельефа определяется по формуле:

$$h = \frac{\lambda}{n - 1}, \tag{1}$$

где n – показатель преломления материала линзы.

Для нормального режима работы были выбраны длины волн 450 нм, 550 нм и 650 нм, равномерно покрывающие видимый диапазон. При расчете были использованы значения показателя преломления резиста, полученные в результате измерения на эллипсомере M2000DI (рис. 2).

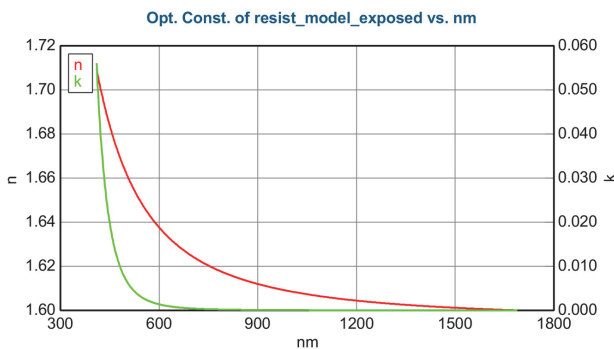


Рис. 2. Зависимость показателя и коэффициента поглощения преломления резиста ФП-40 от длины волны

Для длины волны 450 нм в расчете $n = 1,68$ расчетная высота микрорельефа составила 660 нм, для длины волны 550 нм в расчете $n = 1,64$ расчетная высота микрорельефа составила 860 нм, для длины волны 650 нм в расчете $n = 1,63$ расчетная высота микрорельефа составила 1030 нм.

Радиальные сечения микрорельефа для линз на эти длины волн были рассчитаны [17, 18] для диамет-

ра 10 мм (рис. 3). На рис. 3 приведены только центральные части микрорельефов до радиуса 1 мм, это сделано специально, чтобы исключить слишком высокочастотные краевые участки.

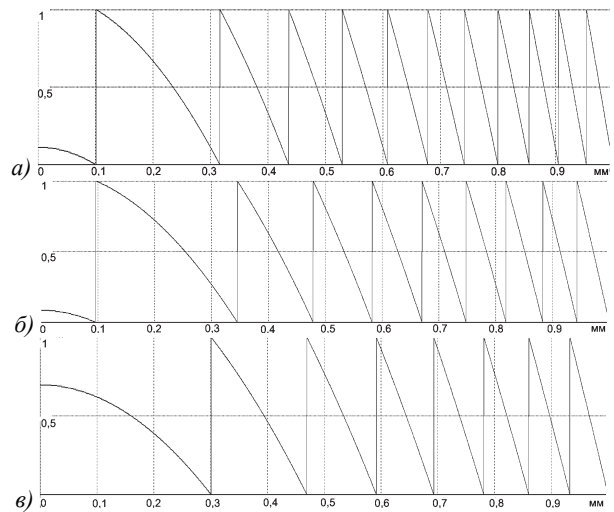


Рис. 3. Радиальные сечения микрорельефа до радиуса 1000 мкм для линз: на 450 нм (а), на 550 нм (б), на 650 нм (в)

По приведенным на рис. 3 расчетным микрорельефам методом прямой лазерной записи на фоторезисте ФП-40 [19, 20] были изготовлены дифракционные линзы. Профилограммы их центральных частей приведены на рис. 4.

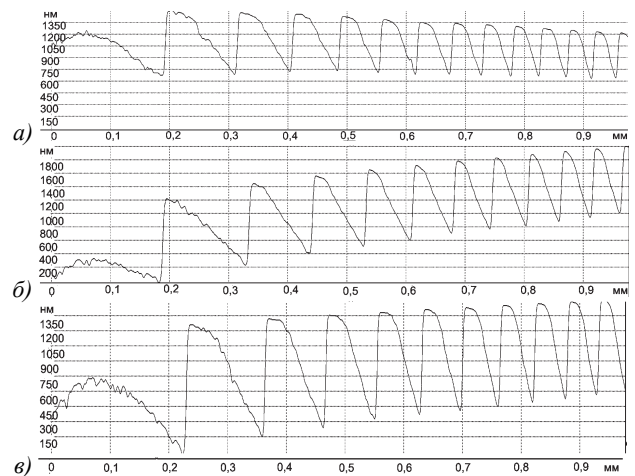


Рис. 4. Профилограммы центральных частей до радиуса 1000 нм для изготовленных линз, рассчитанных для длин волн: 450 нм (а), 550 нм (б), 650 нм (в)

1.2. Формирование набора изображений

Прототип трехлинзовой системы представляет собой три объектива и три USB 3.0 камеры Basler acA1920-40uc, закрепленные на единой площадке (рис. 5).

Юстировка каждого объектива после ручной фокусировки осуществляется с помощью трех подпружиненных регулировочных винтов таким образом, чтобы каждый из объективов смотрел приблизительно в одну и ту же точку.



Рис. 5. Прототип оптической системы на основе трех дифракционных линз

Набор данных состоит из 1857 RGB-изображений размера 1024×1024 и нескольких калибровочных изображений размера 1024×1024 , содержащих маркеры для сопоставления снятых кадров с цифровыми оригиналами (рис. 6а). Разработанное ПО с использованием SDK-камеры автоматизирует процесс вывода снимаемого изображения на экран монитора или проектора (в данной работе используется монитор с диагональю 27 дюймов и разрешением 3840×2160 пикселей) и его съемку с выставляемой задержкой захвата изображения. Съемка осуществлялась с расстояния 2,3 м при поле зрения по горизонтали около $7,5^\circ$.

Подбор коэффициента баланса белого для красного и синего каналов изображения осуществляется в два этапа. Сначала определяется минимум из максимальных значений выдержки каждой камеры, чтобы исключить возможность переэкспозиции («пересвета»), но и не допустить сужения динамического диапазона регистрируемого изображения. Для этого на экран монитора выводится RGB-изображение белого цвета с интенсивностью (255, 255, 255). Полученное значение выдержки фиксируется для каждой камеры, производится съемка белой заливки экрана, затем вычисляются коэффициенты.

Съемка набора данных осуществляется последовательно для каждого из объективов, в файл сохраняется определенный канал изображения.

Сопоставление калибровочного изображения и снятого кадра осуществляется для каждого из трех наборов полученных данных. Сначала осуществляется бинаризация изображения с автоматическим подбором значения порога, локализация угловых маркеров на бинарном изображении, затем вычисление параметров проективного преобразования. Полученные проективные преобразования применяются к каждому набору данных. После этого осуществляется

склейка полутоновых изображений в единое цветное RGB-изображение.

Съемка реального изображения производится по аналогичной схеме. Рядом с интересующим объектом размещается калибровочное изображение с маркерами (рис. 6а). Последовательно осуществляется съемка каждым из объективов и сопоставление снятого изображения с исходным. Далее полученные координаты маркеров используются для вычисления проективного преобразования красного и синего каналов изображения к зеленому. Полученный результат изображен на рис. 6б.

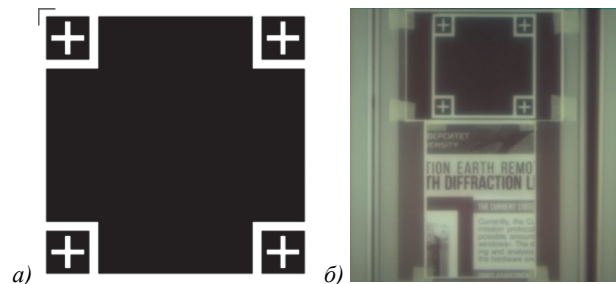


Рис. 6. Калибровочное изображение для сопоставления изображений (а) и снятое изображение реальной сцены после сопоставления кадров с трех линз (б)

2. Нейросетевая реконструкция изображений

2.1. Реконструкция изображений на основе глубокого обучения

Задача реконструкции изображений в дифракционных оптических системах схожа с задачей сверхразрешения по одному изображению (single image super resolution, SISR) [21–24]. Современные нейросетевые решения задачи сверхразрешения обеспечивают хорошее визуальное качество изображений высокого разрешения, а также высокие значения таких метрик качества, как пиковое отношение сигнала к шуму (PSNR) и индекс структурного сходства (SSIM) [21, 25]. Однако большинство из существующих подходов используют известную модель деградации, основанную на бикубической интерполяции изображения низкого разрешения или на размытии изображений различными ядрами [22, 23, 26–28]. Как правило, модели деградации для изображений реального мира неизвестны или сложны для моделирования [29], что приводит к появлению различных артефактов при обработке реальных изображений. В работах [24, 29, 30] предложены современные решения, позволяющие преодолеть проблему артефактов для задачи сверхразрешения. В работе [30] предлагается адаптация универсальных нейросетевых моделей для конкретного случая деградации изображения, основанная на процедуре метаобучения по нулевой обучающей выборке (MZSR). В [24] предложены архитектура и процедура обучения сети, позволяющие получить точное представление деградации для конкретных изображений. Для моделирования деградаций изображений реального мира в [29] предложена

процедура моделирования деградаций высоких порядков на основе простых моделей, таких как размытие изображения, изменение размера, добавление шума и др.

Артефакты, возникающие при реконструкции реальных изображений в дифракционных оптических системах, связаны с различиями в условиях съемки обучающего набора данных, описанных в пункте 1.2, и изображений реальных сцен. Экспозиция и динамический диапазон на изображениях реальных сцен значительно отличаются от результата съемки с экрана, которая проводилась в условиях средней освещенности в 100 люкс. Кроме того, причинами артефактов реконструкции могут быть усиление (gain) или светочувствительность (ISO) камеры и сжатие с потерями, исследуемые в работах [31, 32]

В данной работе мы предлагаем решение для борьбы с артефактами реконструкции на основе аугментации обучающего набора данных. Аугментация данных проводится путем моделирования двух видов деградаций изображений: шум, причиной которого является усиление или светочувствительность камеры (ISO шум), и изменение экспозиции изображения. Кроме того, мы предлагаем измерять уровень артефактов на реальном изображении в процессе обучения сети для поиска лучшей точки в пространстве параметров модели. Для измерения уровня артефактов мы вводим новую метрику, называемую уровнем ложных контуров (False edge level, FEL), которая позволяет использовать изображение реальной сцены в процессе обучения и не требует наличия соответствующего эталонного (ground truth) изображения.

В работах [4, 7] было показано, что деградация изображений, характерная для дифракционной оптики, включает локальные искажения, вызванные хроматическими аберрациями, и глобальные (контекстно-зависимые искажения, хроматический сдвиг), вызванные перераспределением энергии между дифракционными порядками линзы. Поскольку площадь данных искажений составляет более 200 пикселей [4], то в данной работе используется модификация архитектуры U-Net [33], рецептивное поле которой составляет более 200 пикселей. Данная модификация сети U-Net ранее успешно использовалась для реконструкции изображений в дифракционных оптических системах [31, 32, 34, 35].

2.2. Критерий ложных контуров

Артефакты реконструкции на изображении визуально имеют вид контуров (рис. 7а и рис. 7б). В настоящей работе предлагается ввести количественную характеристику для таких артефактов в виде критерия ложных контуров (False edge level, FEL), который определяется как процент пикселей, принадлежащих контуру, для фрагмента изображения, на котором визуально не должно быть контуров:

$$FEL = \frac{1}{NM} \sum_{\substack{0 \leq i \leq N-1 \\ 0 \leq j \leq M-1}} E_{ij} \cdot 100, \quad (2)$$

где E – это бинарная маска контуров размера $N \times M$, полученная в результате работы алгоритма Canny (параметры порога установлены равными 0 и 70). Значения маски контуров E должны быть нормализованы, то есть приведены к диапазону [0, 1].

На рис. 7 приведены результаты работы алгоритма Canny на фрагментах восстановленных изображений реальной сцены. Для каждого фрагмента приведены значения FEL.

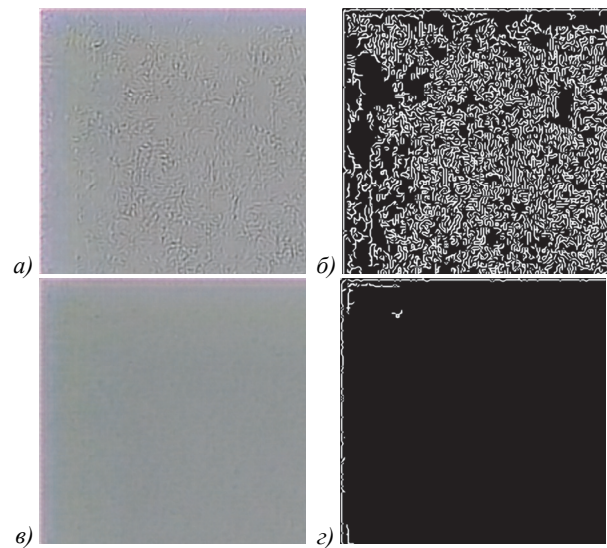


Рис. 7. Примеры работы алгоритма Canny: а) фрагмент восстановленного изображения реальной сцены (FEL – 30,69%), б) результат работы алгоритма Canny для а), в) фрагмент восстановленного изображения реальной сцены (FEL – 1,28%), г) результат работы алгоритма Canny для в)

2.3. Аугментация набора данных

В результате описанной в пункте 1.2 процедуры сформирован набор данных, состоящий из 1878 RGB-изображений размера 1024×1024 , который разделен на обучающую (1244 изображения), тестовую (613 изображений) и валидационную (21 изображение) подвыборки. Обучение нейронной сети на изображениях, снятых с экрана монитора, обеспечивает высокое среднее значение PSNR и хорошее визуальное качество реконструкции на тестовой выборке. Поскольку для обучающей выборки не характерны те типы деградации, которые возникают при съемке изображений реальных сцен, при обработке таких снимков появляются нежелательные артефакты реконструкции [31, 32].

Для устранения артефактов реконструкции в данной работе предлагается аугментировать обучающую выборку путем моделирования двух типов деградации, характерных для изображений реальных сцен: моделирование шума, причиной которого является усиление или светочувствительность камеры (ISO

шум), и изменение экспозиции изображения. Для моделирования ISO-шума в данной работе предлагается использовать распределение Пуассона–Гаусса [36, 37], состоящее из Пуассоновской части, которая моделирует фотонный шум, и Гауссовской части, которая предназначена для моделирования оставшегося стационарного шума. Зашумление обучающих данных проводилось с вероятностью 0,5 (распределение Бернулли) и с использованием библиотеки *albumentations* [38], содержащей реализацию распределения Пуассона–Гаусса. Параметр интенсивности случайным образом выбирался из множества $\{0,1; 0,2; 0,3\}$ для каждого обучающего примера.

Изменение экспозиции изображений обучающей выборки проводилось с использованием алгоритма повышения качества изображения, снятого в условиях низкой освещенности, предложенного в работе [39]. Алгоритм основан на функции отклика камеры, связывающей освещенность сенсора со значениями пикселей изображения. Для повышения качества изображения алгоритм формирует оценку модели отклика камеры на основе гистограммных характеристик двух изображений с разной экспозицией и карту значений экспозиции. Алгоритм позволяет варьировать экспозицию изображения, не внося дополнительные искажения цвета. В рамках данной работы изменение экспозиции применялось с вероятностью 0,1 после аугментации ISO-шумом.

2.4. Архитектура сети

В рамках данной работы использовалась модификация сети U-Net [33], которая ранее успешно применялась для задачи реконструкции изображений в дифракционных оптических системах [21, 26]. Архитектура сети имеет следующий вид:

C64-C128-C256-C512-C512-C512-C512-C512
(Кодировщик),
CD512- CD512- CD512-C512-C256-C128-C64-C3
(Декодировщик), C3

где C_k – комбинация сверточного слоя, пакетной нормализации и функции активации ReLU, k – количество фильтров. Слой *dropout* добавлен перед слоем активации в CDk блоках. Размер ядра свертки равен 4×4 . Архитектура сети содержит связи между каждым i -м слоем в кодировщике и $(n-i)$ -м слоем декодировщика, n – общее число слоев. На последнем слое в качестве функции активации используется гиперболический тангенс (\tanh).

3. Экспериментальные исследования

3.1. Обучение сети с использованием критерия FEL

Для обучения сети, описанной в пункте 2.4, использовался алгоритм оптимизации ADAM с параметрами $\beta_1 = 0,5$ и $\beta_2 = 0,999$ и коэффициентом скорости обучения, равным 0,0002. В качестве функции потерь использовалась l_2 -норма.

Классический подход [31, 32] к выбору оптимальной точки в пространстве параметров сети в процессе обучения основан на среднем значении PSNR на валидационной выборке. Согласно данному подходу критерием выбора оптимальной точки является максимальное значение среднего PSNR на валидационной выборке (*max-PSNR* критерий). В данной работе предлагается альтернативный критерий, основанный на расчете метрики FEL для фрагмента изображения реальной сцены. Критерием выбора оптимальной точки в пространстве параметров является минимальное значение метрики FEL на фрагменте реального изображения (*min-FEL* критерий).

На вход нейронной сети подавались изображения размера 1024×1024 пикселей, снятые дифракционным объективом после процедуры склейки полутоновых изображений в единое цветное RGB-изображение.

В рамках проведенных экспериментов нейронная сеть обучалась в течение 200 эпох, поскольку при дальнейшем обучении наблюдается стагнация среднего значения PSNR на валидационной выборке. После каждой эпохи рассчитывались среднее значение PSNR на валидационной выборке и FEL на фрагменте (200×200 пикселей) реального изображения, представленного на рис. 8а. Используемый фрагмент выделен рамкой черного цвета. На выбранном фрагменте нет перепадов яркости и визуально не наблюдаются контуры. Поэтому все обнаруженные контуры на данном фрагменте относятся к артефактам реконструкции, и метрика FEL покажет объективную оценку уровня артефактов. После обучения сети были выбраны 2 точки пространства параметров на основе критериев *max-PSNR* и *min-FEL*. На рис. 8б и рис. 8в представлены фрагменты изображения реальной сцены, обработанные двумя отобранными моделями. В данном случае обучение проводилось на исходных неаугментированных данных. В случае выбора точки в пространстве параметров на основе критерия *max-PSNR* (рис. 8б) значение FEL составило 17,51%. При использовании критерия *min-FEL* значение FEL снизилось до 9,11%, что визуально соответствует меньшему количеству артефактов реконструкции на рис. 8в.

3.2. Обучение сети на аугментированных данных

В рамках проведенных исследований рассматриваются 3 случая обучения сети: без аугментации данных, аугментация обучающей выборки путем моделирования ISO-шума, а также аугментация моделированием ISO-шума с последующим изменением экспозиции. Зашумление данных проводилось с вероятностью 0,5 и значением параметра интенсивности, выбранным случайным образом из множества $\{0,1; 0,2; 0,3\}$. Изменение экспозиции проводилось с вероятностью 0,1 после аугментации ISO-шумом. В табл. 1 представлены результаты расчета среднего значения PSNR на тестовой выборке и FEL на фрагменте реального изображения для трех случаев обучения сети.

На рис. 9 указаны восстановленные фрагменты реального изображения, представленного на рис. 8а. Результаты в табл. 1 и на рис. 9 подтверждают эффективность критерия min-FEL для выбора точки в пространстве параметров модели, обеспечивающей снижение количества артефактов реконструкции на изображениях реальных сцен. Кроме того, лучший результат с точки зрения значения FEL на фрагменте реального изображения обеспечила аугментация на основе моделирования ISO-шума. Данный результат соответствует визуально меньшему числу артефактов реконструкции на рис. 9б.

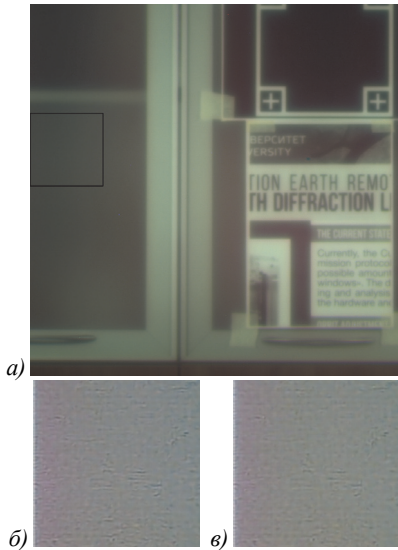


Рис. 8. Пример снятого и обработанного фрагмента изображения реальной сцены: а) полноразмерное изображение, снятое дифракционным объективом после процедуры склейки полутоновых изображений в единое цветное RGB-изображение, б) результат реконструкции фрагмента, представленного на рис. б) для случая выбора точки в пространстве параметров на основе критерия max-PSNR, в) результат реконструкции для случая критерия отбора min-FEL

На рис. 10 представлено сравнение результатов реконструкции изображения реальной сцены (рис. 8а) при использовании критерия min-FEL для случая аугментации ISO-шумом и при отсутствии аугментации. Видно, что в случае аугментации обучающей выборки ISO-шумом наблюдается меньшее количество артефактов и визуально они становятся менее заметными. Визуальную оценку качества можно провести после сравнения текстовых фрагментов на рис. 8а) и 10а). Видно, что текст, не читаемый на изображении, полученном при помощи дифракционной оптики, становится читаемым после проведенной нейросетевой реконструкции.

Среднее значение PSNR на тестовой выборке при использовании критерия min-FEL снижается на 0,35 дБ, однако этот негативный эффект компенсируется значительным улучшением визуального качества реконструкции изображения реальной сцены с точки зрения количества артефактов.

Табл. 1. PSNR и FEL для трех случаев обучения

| № | Аугментация | Критерий отбора точки в пространстве параметров модели | FEL (%) | PSNR на тестовой выборке (дБ) | Восстановленный фрагмент |
|---|--------------------------------|--|---------|-------------------------------|--------------------------|
| 1 | Отсутствует | Max-PSNR | 17,51 | 25,54 | Рис. 8б |
| | | Min-FEL | 9,11 | 25,52 | Рис. 8в |
| 2 | ISO шум | Max-PSNR | 1,20 | 25,67 | Рис. 9а |
| | | Min-FEL | 0,51 | 25,32 | Рис. 9б |
| 3 | ISO шум и изменение экспозиции | Max-PSNR | 4,70 | 25,23 | Рис. 9в |
| | | Min-FEL | 0,80 | 25,27 | Рис. 9г |

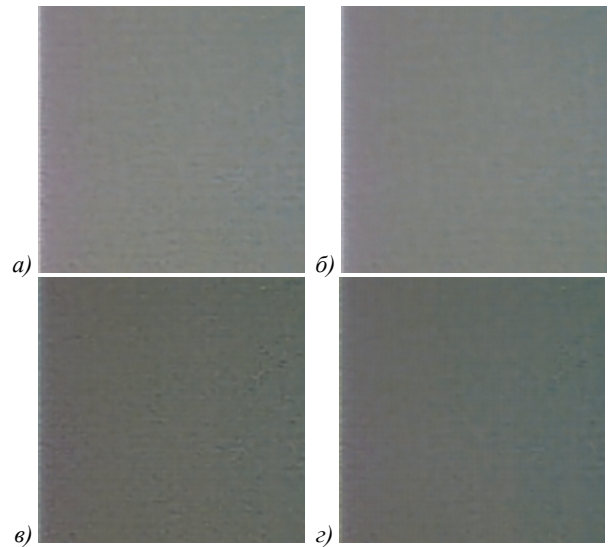


Рис. 9. Результат реконструкции фрагмента изображения реальной сцены, представленного на рис. 8а: а) аугментация обучающей выборки моделированием ISO-шума, использование критерия max-PSNR, б) аугментация обучающей выборки моделированием ISO-шума, использование критерия min-FEL, в) аугментация обучающей выборки моделированием ISO-шума и изменением экспозиции, использование критерия max-PSNR, г) аугментация обучающей выборки моделированием ISO-шума, использование критерия min-FEL

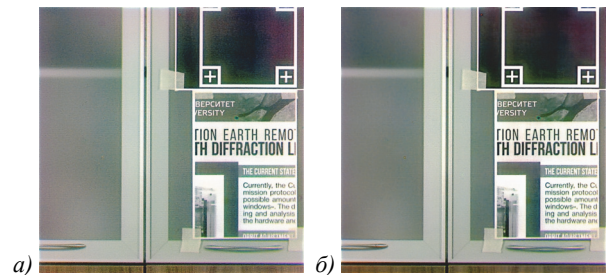


Рис. 10. Результаты реконструкции полноразмерного изображения реальной сцены, представленного на рис. 8а, при использовании критерия min-FEL: а) обучение сети без аугментации обучающей выборки, б) применение аугментации ISO-шумом

На рис. 11 представлено сравнение графиков обучения при отсутствии аугментации обучающих данных (табл. 1 строка 1) и в случае аугментации ISO шумом (табл. 1 строка 2). Видно, что использование аугментации обеспечивает более высокий уровень

PSNR на валидационной выборке (рис. 11а). С точки зрения FEL по графику на рис. 11б видно, что при обучении на исходных данных без аугментации появляется проблема переобучения, при которой значение FEL достигает некоторого локального минимума (около 80 эпохи), после чего начинает расти. При использовании аугментации ISO-шумом данная проблема исчезает.

В рамках данной работы нейросетевая реконструкция позволила увеличить среднее значение PSNR на тестовой выборке с 12,77 дБ до 25,67 дБ.

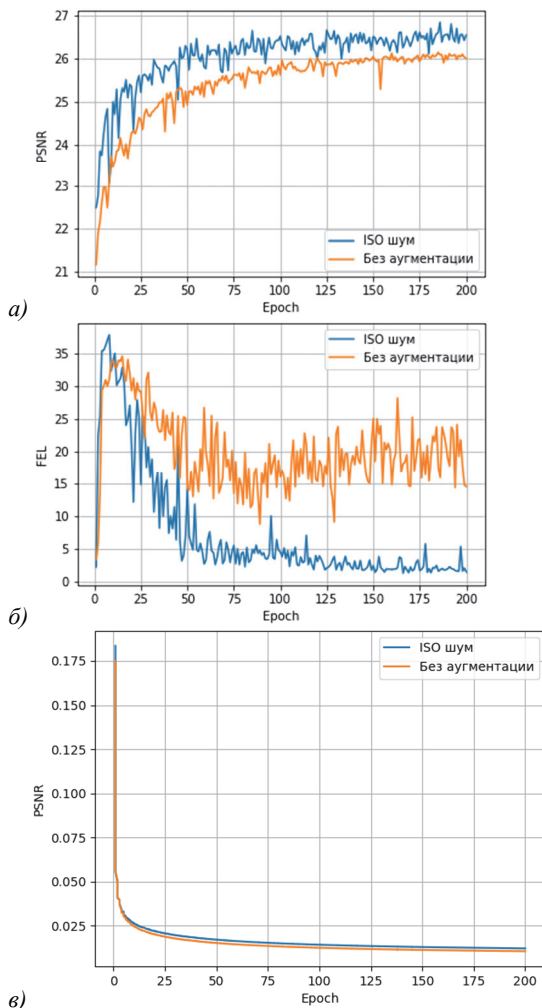


Рис. 11. Сравнение графиков обучения при отсутствии аугментации обучающей выборки и случая аугментации ISO-шумом: а) PSNR на валидационной выборке, б) FEL, рассчитанный по фрагменту изображения реальной сцены (рис. 8а), в) функция потерь на валидационной выборке

Заключение

В рамках настоящей работы предложена мультиапертурная новая изображающая система из трех дифракционных линз, каждая из которых формирует один канал цветного RGB-изображения. Такой подход позволяет на основе использования стандартных байеровских фильтров на цветной матрице существенно сузить спектральный диапазон каждой линзы

и таким образом существенно снизить искажения изображения, вызванные хроматической aberrацией, присущей дифракционной оптике. При этом наличие трех изображений вместо одного существенно расширяет возможности последующей нейросетевой коррекции.

Для разработанной изображающей системы нейросетевая реконструкция позволила повысить как визуальное качество реальных снимков, так и среднее значение PSNR на тестовой выборке (с 12,77 дБ до 25,67 дБ). Предложенный критерий на основе уровня ложных контуров FEL позволил найти точку в пространстве параметров модели, обеспечивающую минимальное количество артефактов реконструкции на реальном изображении. При использовании классического критерия отбора лучшей точки на основе PSNR значение FEL составило 17,51%. При использовании предложенного критерия значение FEL снизилось до 9,11%, что визуально соответствует меньшему количеству артефактов реконструкции на изображении реальной сцены.

Кроме того, в работе предложен алгоритм аугментации данных для компенсации различий в условиях съемки обучающего набора данных и изображений реальных сцен. Показано, что аугментация на основе изменения экспозиции изображений обеспечивает лучшее визуальное качество восстановленных изображений и позволяет значительно уменьшить количество артефактов реконструкции с точки зрения метрики FEL (с 9,11% до 0,51%) без значительного уменьшения PSNR на тестовой выборке.

В дальнейшем планируется исследование двух-этапной схемы реконструкции, объединяющей архитектуры U-NET и RCAN [21].

Мультиапертурные системы в совокупности с нейросетевой обработкой позволят в будущем создавать компактные высокоразрешающие системы технического зрения. Результаты, полученные в данной работе, также могут быть использованы в некоторых задачах неизображающей дифракционной оптики [40].

Благодарности

Работа выполнена при поддержке гранта РФФ 22-19-00364.

References

- [1] Chen MK, Liu X, Sun Y, Tsai DP. Artificial intelligence in meta-optics. *Chem Rev* 2022; 122(19): 15356-15413.
- [2] Genevet P, Capasso F, Aieta F, Khorasaninejad M, Devlin R. Recent advances in planar optics: from plasmonic to dielectric metasurfaces. *Optica* 2017; 4(1): 139-152.
- [3] Banerji S, et al. Imaging with flat optics: metalenses or diffractive lenses? *Optica* 2019; 6(6): 805-810.
- [4] Nikonorov A, Evdokimova V, Petrov M, Yakimov P, Bibikov S, Yuzifovich Y, Skidanov R, Kazanskiy N. Deep learning-based imaging using single-lens and multi-aperture diffractive optical systems. 2019 IEEE/CVF Int

- Conf on Computer Vision Workshop (ICCVW) 2019: 3969-3977. DOI: 10.1109/ICCVW.2019.00491.
- [5] Kazanskii NL, Khonina SN, Skidanov RV, Morozov AA, Kharitonov SI, Volotovskiy SG. Formation of images using multilevel diffractive lens. *Computer Optics* 2014; 38(3): 425-434. DOI: 10.18287/0134-2452-2014-38-3-425-434.
- [6] Nikonorov A, Skidanov R, Fursov V, Petrov M, Bibikov S, Yuzifovich Y. Fresnel lens imaging with post-capture image processing. 2015 IEEE Conf on Computer Vision and Pattern Recognition Workshops (CVPRW) 2015: 33-41. DOI: 10.1109/CVPRW.2015.7301373.
- [7] Nikonorov A, Petrov M, Bibikov S, Yakimov P, Kutikova V, Yuzifovich Y, Morozov A, Skidanov R, Kazanskiy N. Toward ultralightweight remote sensing with harmonic lenses and convolutional neural networks. *IEEE J Sel Top Appl Earth Obs Remote Sens* 2018; 11(9): 3338-3348. DOI: 10.1109/JSTARS.2018.2856538.
- [8] Nikonorov AV, Skidanov RV, Kutikova VV, Petrov MV, Alekseev AP, Bibikov SA, Kazanskiy NL. Towards multi-aperture imaging using diffractive lens. *Proc SPIE* 2019; 11146: 111460Y. DOI: 10.1117/12.2526923.
- [9] Nikonorov A, Petrov M, Bibikov S, Yuzifovich Y, Yakimov P, Kazanskiy N, Skidanov R, Fursov V. Comparative evaluation of deblurring techniques for fresnel lens computational imaging. 2016 23rd Int Conf on Pattern Recognition (ICPR) 2016: 775-780. DOI: 10.1109/ICPR.2016.7899729.
- [10] Peng Y, Fu Q, Amata H, Su Sh, Heide F, Heidrich W. Computational imaging using lightweight diffractive-refractive optics. *Opt Express* 2015; 23(24): 31393-31407.
- [11] Skidanov R, Strelkov Y, Volotovskiy S, Blank V, Ganchevskaya S, Podlipnov V, Ivliev N, Kazanskiy N. Compact imaging systems based on annular harmonic lenses. *Sensors* 2020; 20(14): 3914. DOI: 10.3390/s20143914.
- [12] Sales TRM, Morris GM. Diffractive-refractive behavior of kinoform lenses. *Appl Opt* 1997; 36: 253-257.
- [13] Mohammad N, Meem M, Shen B, Wang P, Menon R. Broadband imaging with one planar diffractive lens. *Sci Rep* 2018; 8: 2799.
- [14] Kim G, Domínguez-Caballero JA, Menon R. Design and analysis of multi-wavelength diffractive optics. *Opt Express* 2012; 20: 2814-2823.
- [15] Blank V, Skidanov R, Doskolovich L, Kazanskiy N. Spectral diffractive lenses for measuring a modified red edge simple ratio index and a water band index. *Sensors* 2021; 21(22): 7694. DOI: 10.3390/s21227694.
- [16] Blank VA, Skidanov RV, Doskolovich LL. Investigation of a spectral lens for the formation of a normalized difference vegetation index NDVI0.705. *J Opt Technol* 2022; 89(3): 137-141. DOI: 10.1364/JOT.89.000137.
- [17] Doskolovich LL, Golub MA, Kazanskiy NL, Khrarov AG, Pavelyev VS, Seraphimovich PG, Soifer VA, Volotovskiy SG. Software on diffractive optics and computer generated holograms. *Proc. SPIE* 1995; 2363: 278-284. DOI: 10.1117/12.199645.
- [18] Kazanskiy NL. Modeling diffractive optics elements and devices. *Proc SPIE* 2018; 10774: 107740O. DOI: 10.1117/12.2319264.
- [19] Kazanskiy NL. Research & Education Center of Diffractive Optics. *Proc SPIE* 2012; 8410: 84100R. DOI: 10.1117/12.923233.
- [20] Kazanskiy NL, Skidanov RV. Technological line for creation and research of diffractive optical elements. *Proc. SPIE* 2019; 11146: 111460W. DOI: 10.1117/12.2527274.
- [21] Zhang Y, Li Ku, Li K, Wang L, Zhong B, Fu Y. Image super-resolution using very deep residual channel attention networks. In Book: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, eds. *Computer Vision – ECCV 2018*. Cham: Springer Nature Switzerland AG; 2018: 286-301.
- [22] Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. 2016 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2016: 1646-1654.
- [23] Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y. Residual dense network for image super-resolution. 2018 IEEE/CVF Conf on Computer Vision and Pattern Recognition 2018: 2472-2481.
- [24] Wang L, Wang Y, Dong X, Xu Q, Yang J, An W, Guo Y. Unsupervised degradation representation learning for blind super-resolution. 2021 IEEE/CVF Conf on Computer Vision and Pattern Recognition (CVPR) 2021: 10576-10585. DOI: 10.1109/CVPR46437.2021.01044.
- [25] Wang Z, Chen J, Hoi SCH. Deep learning for image super-resolution: A survey. *IEEE Trans Pattern Anal Mach Intell* 2020; 43(10): 3365-3387. DOI: 10.1109/TPAMI.2020.2982166.
- [26] Dong C, Loy CC, He K, Tang X. Learning a deep convolutional network for image super-resolution. In Book: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. *Computer Vision – ECCV 2014*. Cham: Springer International Publishing Switzerland; 2014: 184-199.
- [27] Zhang Y, Li K, Li K, Wang L, Zhong B, Fu Y. Image super-resolution using very deep residual channel attention networks. In Book: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, eds. *Computer Vision – ECCV 2018*. Cham: Springer Nature Switzerland AG; 2018: 294-310.
- [28] Zhang K, Zuo W, Zhang L. Learning a single convolutional super-resolution network for multiple degradations. 2018 IEEE/CVF Conf on Computer Vision and Pattern Recognition 2018: 3262-3271.
- [29] Wang X, Xie L, Dong C, Shan Y. Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data. *IEEE/CVF Int Conf on Computer Vision Workshops (ICCVW)* 2021: 1904-1914.
- [30] Soh JW, Cho S, Cho NI. Meta-transfer learning for zero-shot super-resolution. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2020: 3513-3522. DOI: 10.1109/CVPR42600.2020.00357.
- [31] Evdokimova VV, Petrov MV, Klyueva MA, Zybin EY, Kosianchuk VV, Mishchenko IB, Novikov VM, Selvesiuk NI, Ershov EI, Ivliev NA, Skidanov RV, Kazanskiy NL, Nikonorov AV. Deep learning-based video stream reconstruction in mass production diffractive optical systems. *Computer Optics* 2021; 45(1): 130-141. DOI: 10.18287/2412-6179-CO-834.
- [32] Ivliev N, Evdokimova V, Podlipnov V, Petrov M, Ganchevskaya S, Tkachenko I, Abrameshin D, Yuzifovich Y, Nikonorov A, Skidanov R, Kazanskiy N, Soifer V. First Earth-imaging CubeSat with harmonic diffractive lens. *Remote Sens* 2022; 14(9): 2230. DOI: 10.3390/rs14092230.
- [33] Ronneberger O. U-Net: Convolutional networks for biomedical image segmentation. In Book: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical image computing and computer-assisted intervention – MICCAI*. New York, Dordrecht, London: Springer, 2015: 234-241. DOI: 10.1007/978-3-319-24574-4_28.
- [34] Dun X, Ikoma H, Wetzstein G, Wang Z, Cheng X, Peng Y. Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging. *Optica* 2020; 7: 913-922.

- [35] Peng Y, Sun Q, Dun X, Wetzstein G, Heidrich W, Heide F. Learned large field-of-view imaging with thin-plate optics. *ACM Trans Graph* 2019; 38(6): 219. DOI: 10.1145/3355089.3356526.
- [36] Foi A, Trimeche M, Katkovnik V, Egiazarian K. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Trans Image Process* 2008; 17(10): 1737-1754. DOI: 10.1109/TIP.2008.2001399.
- [37] Guo S, Yan Z, Zhang K, Zuo W, Zhang L. Toward convolutional blind denoising of real photographs. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2019: 1712-1722. DOI: 10.1109/CVPR.2019.00181.
- [38] Do more with less data. 2023. Source: <<https://albuquerque.ai/>>.
- [39] Ying Z, Li G, Ren Y, Wang R, Wang W. A new low-light image enhancement algorithm using camera response model. *IEEE Int Conf on Computer Vision Workshops (ICCVW)* 2017: 3015-3022. DOI: 10.1109/ICCVW.2017.356.
- [40] Karpeev SV, Alferov SV, Khonina SN, Kudryashov SI. Study of the broadband radiation intensity distribution formed by diffractive optical elements. *Computer Optics* 2014; 38(4): 689-694. DOI: 10.18287/0134-2452-2014-38-4-689-694.

Сведения об авторах

Степаненко Сергей Олегович, 1997 года рождения, в 2021 году окончил магистратуру Самарского национального исследовательского университета имени академика С.П. Королёва по направлению «Прикладная математика и информатика». Аспирант кафедры суперкомпьютеров и общей информатики Самарского университета. E-mail: serega.stepanenko.97@gmail.com.

Евдокимова Виктория Витальевна, 1994 года рождения, в 2017 году окончила магистратуру Самарского национального исследовательского университета имени академика С.П. Королёва по направлению «Прикладная математика и информатика». Ассистент кафедры суперкомпьютеров и общей информатики Самарского университета. Научный сотрудник лаборатории интеллектуального анализа видеоданных Института систем обработки изображений РАН – филиала Федерального государственного учреждения «Федеральный научно-исследовательский центр «Кристаллография и фотоника» Российской академии наук. Область научных интересов: компьютерная обработка изображений, теория распознавания образов, интеллектуальный анализ данных, глубокое обучение. E-mail: vkutikova94@gmail.com.

Петров Максим Витальевич, родился в 1989 году. В 2013 году окончил Самарский государственный аэрокосмический университет. Ассистент кафедры суперкомпьютеров и общей информатики Самарского университета, научный сотрудник лаборатории интеллектуального анализа видеоданных Института систем обработки изображений РАН – филиала Федерального государственного учреждения «Федеральный научно-исследовательский центр «Кристаллография и фотоника» Российской академии наук. Является автором и соавтором 20 научных трудов. Область научных интересов: распознавание образов и анализ изображений, параллельные и распределённые вычисления, вычисления с использованием графических процессоров. E-mail: max.vit.petrov@gmail.com.

Скиданов Роман Васильевич, 1973 года рождения. В 1990 году с отличием окончил Самарский государственный университет (СамГУ) по специальности «Физика». Доктор физико-математических наук (2007 год), работает заведующим лабораторией микро- и нанотехнологий Института систем обработки изображений РАН (ИСОИ РАН), профессором кафедры технической кибернетики Самарского университета. Р.В. Скиданов – специалист в области дифракционной оптики, математического моделирования, обработки изображений и нанофотоники. В списке научных работ Р.В. Скиданова 250 статей, 8 монографий. E-mail: romans@ipsiras.ru.

Никоноров Артем Владимирович, родился в 1979 году. В 2005 году окончил Самарский государственный аэрокосмический университет. Доктор технических наук (2016), профессор кафедры суперкомпьютеров и общей информатики Самарского университета. Опубликовал более 80 работ, две монографии. Область научных интересов: распознавание образов и анализ изображений, идентификация систем, параллельные и распределённые вычисления, вычисления с использованием графических процессоров. E-mail: artniko@gmail.com.

ГРНТИ: 28.23.15

Поступила в редакцию 6 декабря 2022 г. Окончательный вариант – 14 апреля 2023 г.

Color imaging using a system based on 3 diffractive lenses

S. Stepanenko¹, V. Evdokimova^{1,2}, M. Petrov¹, R. Skidanov^{1,2}, A. Nikonorov^{1,2}

¹ Samara National Research University, 443086, Samara, Russia, Moskovskoye Shosse 34;

² IPSI RAS – Branch of the FSRC “Crystallography and Photonics” RAS,
443001, Samara, Russia, Molodogvardeyskaya 151

Abstract

The possibility of essentially reducing the weight and production cost of computer vision systems has led to the publication of a large number of research works dealing with the development of new imaging systems based on diffractive optics. This study proposes a new imaging system composed of three diffractive lenses, with each forming a separate channel of the color RGB image. This approach allows us to significantly narrow the spectral range of each lens, thus significantly reducing the image distortion caused by chromatic aberration inherent in diffractive optics. It shows that this scheme allows us to perform the neural network-aided image reconstruction, providing a significantly improved resulting image quality. The study proposes a false edge level criterion (FEL) for evaluating the neural network-aided reconstruction.

Keywords: color correction, color stabilization, global optimization.

Citation: Stepanenko SO, Evdokimova VV, Petrov MV, Skidanov RV, Nikonorov AV. Color imaging using a system based on 3 diffractive lenses. *Computer Optics* 2023; 47(5): 716-724. DOI: 10.18287/2412-6179-CO-1258.

Acknowledgements: This work was financially supported by RSF grant #22-19-00364.

Authors' information

Sergei Olegovich Stepanenko (b. 1997) graduated from Samara National Research University in 2021 with a Master's degree, majoring in Applied Mathematics and Informatics. Now he is a postgraduate student at Supercomputing and Computer Science sub-department of Samara University. E-mail: serega.stepanenko.97@gmail.com.

Viktoriia Vitalievna Evdokimova (b. 1994) graduated from Samara National Research University in 2017 with a Master's degree, majoring in Applied Mathematics and Informatics. Now she is a postgraduate student at Supercomputing and Computer Science sub-department of Samara University. Current research interests are in computer image processing, pattern recognition, data science, and deep learning. E-mail: vkutikova94@gmail.com.

Maksim Vitalyevich Petrov (b. 1989). He graduated from SSAU in 2013. Currently he works as an assistant at Supercomputing and Computer Science sub-department of Samara University, a researcher at the Image Processing Systems Institute of the Russian Academy of Sciences – Branch of the Federal Scientific Research Centre “Crystallography and Photonics” RAS. He is the co-author of 20 scientific papers. Field of scientific interest: pattern recognition and image analysis, parallel and distributed programming, GPGPU programming. E-mail: max.vit.petrov@gmail.com.

Roman Vasilevich Skidanov (b. 1973). Graduated with honours (1990) from Samara State University (SSU), majoring in Physics. He received his Doctor in Physics & Maths (2007) degrees from Samara State University. He is the head of Micro- and Nanotechnologies laboratory of the Image Processing Systems Institute of RAS – Branch of the FSRC “Crystallography and Photonics” of the Russian Academy of Sciences, holding a part-time position of professor at Samara National Research University Technical Cybernetics sub-department. He is co-author of 250 scientific papers, 8 monographs. His current research interests include diffractive optics, mathematical modeling, image processing, and nanophotonics. E-mail: romans@ipsiras.ru.

Artem Vladimirovich Nikonorov (b. 1979). He graduated from SSAU in 2002. He received his PhD degree in 2016. Now he works as a professor at Supercomputing and Computer Science sub-department of Samara University. He has more than 70 publications, two monographs. Field of scientific interest: pattern recognition and image analysis, system identification, parallel and distributed programming, GPGPU programming. E-mail: artniko@gmail.com.

Received December 6, 2022. The final version – April 14, 2023.
