

## Нейросетевая реконструкция видеопотока в дифракционных оптических системах массового производства

В.В. Евдокимова<sup>1,2</sup>, М.В. Петров<sup>1,2</sup>, М.А. Ключева<sup>1,2</sup>, Е.Ю. Зыбин<sup>3</sup>, В.В. Косьянчук<sup>3</sup>,  
И.Б. Мищенко<sup>3</sup>, В.М. Новиков<sup>3</sup>, Н.И. Сельвесюк<sup>3</sup>, Е.И. Ершов<sup>4</sup>, Н.А. Ивлиев<sup>1,2</sup>,  
Р.В. Скиданов<sup>1,2</sup>, Н.Л. Казанский<sup>1,2</sup>, А.В. Никоноров<sup>1,2</sup>

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

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

<sup>3</sup> ФГУП Государственный научно-исследовательский институт авиационных систем,  
125319, Россия, г. Москва, ул. Викторенко, д. 7,

<sup>4</sup> Институт проблем передачи информации имени А.А. Харкевича РАН,  
127051, Россия, г. Москва, Большой Каретный переулок, д. 19, стр. 1

### Аннотация

Возможность существенно снизить массу и стоимость систем технического зрения привела к появлению большого числа работ, посвященных разработке новых оптических схем на основе дифракционной оптики и новых подходов к реконструкции получаемых изображений. Получаемые системы демонстрируют достаточное для прикладных систем технического зрения качество изображений. Однако при создании таких прикладных систем возможны источники дополнительных потерь качества получаемого видеопотока. В настоящей работе исследовано влияние на итоговое качество реконструируемого видеопотока таких факторов, как ограничения технологии массового производства дифракционной оптики, артефактов сжатия видеопотока с потерями, а также особенностей нейросетевого подхода к реконструкции. Предложена сквозная нейросетевая технология реконструкции изображений, позволяющая компенсировать дополнительные факторы потери качества и получить итоговый видеопоток с качеством, достаточным для решения прикладных задач технического зрения.

**Ключевые слова:** дифракционная оптика, дифракционные линзы, нейросетевая реконструкция, обработка изображений.

**Цитирование:** Евдокимова, В.В. Нейросетевая реконструкция видеопотока в дифракционных оптических системах массового производства / В.В. Евдокимова, М.В. Петров, М.А. Ключева, Е.Ю. Зыбин, В.В. Косьянчук, И.Б. Мищенко, В.М. Новиков, Н.И. Сельвесюк, Е.И. Ершов, Н.А. Ивлиев, Р.В. Скиданов, Н.Л. Казанский, А.В. Никоноров // Компьютерная оптика. – 2021. – Т. 45, № 1. – С. 130-141. – DOI: 10.18287/2412-6179-CO-834.

**Citation:** 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.

### Введение

В настоящее время все больше исследований проводится в области реконструкции изображений, полученных на основе плоских оптических элементов [1–5]. Использование плоской оптики позволяет значительно сократить вес и упростить конструкцию оптической системы, что является большим преимуществом перед классическими рефракционными аналогами. Наиболее популярным подходом к проектированию плоской оптики является использование т.н. металлинз [5], которые позволяют создать оптическую систему с субволновой толщиной. Другой пример плоской оптики – многоуровневые дифракционные линзы (МДЛ), которые стали использоваться для получения цветных изображений высокого разрешения только с 2015 года [3]. Подход, основанный на исполь-

зовании метаповерхностей, обеспечивает меньшую толщину линзы, однако МДЛ проще в изготовлении и позволяют получить изображения лучшего качества.

При создании плоских оптических элементов возникает проблема существенных искажений регистрируемых изображений. В работе [5] авторы предлагают устранять искажения при помощи дополнительных оптических элементов. В [1, 2, 3, 4, 6, 7] используется так называемый алгоритмический подход к реконструкции, при котором в оптических системах применяется комбинация дифракционной оптики и алгоритма реконструкции, предназначенного для компенсации возникающих на изображении искажений.

В работе [1] предлагается сквозная нейросетевая технология реконструкции изображений и впервые используется генеративно-состязательная сеть (GAN) для решения задачи цветовой коррекции.

В настоящей работе проводится анализ качества для различных конфигураций сетей, технологии изготовления МДЛ и параметров сжатия видеопотока. В результате получена эффективная архитектура полностью нейросетевой реконструкции с возможностью адаптации под особенности технологии изготовления МДЛ и алгоритмы компрессии видеопотока [8–10].

Полученные в работах [1, 4, 11, 24–26] результаты показывают, что после нейросетевой реконструкции качество изображений, полученных при помощи МДЛ, становится достаточным для решения прикладных задач технического зрения. Поэтому для данной сферы становятся актуальными исследования, связанные с решением прикладных систем и созданием систем массового производства на основе МДЛ. Подобные исследования применимости в прикладных задачах распознавания лиц, в частности, уже проводились для систем на основе безлинзовых камер [12]. В настоящей работе будут исследованы вопросы, предваряющие решение прикладных задач на массовых системах технического зрения на основе МДЛ, а именно: будет исследовано влияние на итоговое качество реконструируемого видеопотока таких факторов, как ограничения технологии массового производства дифракционной оптики, артефактов сжатия видеопотока с потерями, а также особенностей нейросетевого подхода к реконструкции.

Переход к массовому изготовлению МДЛ требует оптимизации и упрощения технологии изготовления. Так, в частности, представляет интерес возможность использования литографической технологии для изготовления МДЛ. Влиянию технологии изготовления дифракционной оптики на качество получаемых изображений посвящен целый ряд исследований [1]. Однако в этих исследованиях не принимался в расчет процесс реконструкции изображения. В настоящей работе исследуется ухудшение качества изображений, к которому приводит упрощение технологии изготовления и возможность компенсации этого ухудшения на этапе нейросетевой реконструкции.

Следующий вопрос, который требует исследования, – это влияние сжатия видеопотока с потерями, с применением распространенных видеокодеков, таких как H265/H264. Использование таких кодеков требуется при создании мобильных систем технического зрения с применением МДЛ, таких как камеры БПЛА, камеры видеонаблюдения, специализированные камеры мобильных телефонов. В предыдущих работах [6, 7] проводилось достаточно детальное исследование этапов получения и реконструкции изображений, однако без учета сжатия изображения с потерями. В настоящей работе исследована зависимость степени компрессии и качества нейросетевой реконструкции, предложена схема реконструкции, позволяющая минимизировать влияние компрессии на качество итогового видеопотока.

Указанные вопросы, связанные с качеством получаемых МДЛ изображений, требуют решения для перехода к массовому производству и использованию систем технического зрения на основе МДЛ. В настоящей работе проводится детальное исследование и предлагается новая сквозная нейросетевая технология, позволяющая в существенной мере компенсировать возникающее ухудшение качества изображений.

### 1. Технологические особенности изготовления гармонических линз

Фокусное расстояние МДЛ существенно зависит от длины волны фокусируемого света. Для обычной дифракционной линзы с высотой микрорельефа  $\lambda/(n-1)$ , где  $\lambda$  – расчетная длина волны, разница для фокусных расстояний в видимом диапазоне сопоставима с самим фокусным расстоянием на расчетной длине волны. В данном случае получить качественное изображение невозможно. Решением данной проблемы является использование гармонических линз, высота рельефа которых  $m\lambda/(n-1)$  [1], где  $m$  – показатель гармоничности. Так, для длины волны 510 нм, при высоте микрорельефа 6 мкм  $m=8$ . Для гармонической линзы изменение оптической силы гармонической линзы подчиняется следующей формуле [13]:

$$\frac{\Delta f}{f} = \frac{\Delta P}{P} = \frac{\Delta \lambda}{\lambda} \cong \frac{1}{m}, \quad (1)$$

где  $f$  – фокусное расстояние,

$P$  – оптическая сила;

$\Delta f$  – диапазон изменения фокусного расстояния при изменении длины волны;

$\Delta P$  – диапазон изменения оптической силы при изменении длины волны;

$\Delta \lambda$  – диапазон изменения длины волны.

Таким образом, для минимизации изменения фокусного расстояния на разных длинах волн необходимо увеличивать показатель гармоничности  $m$ , следовательно, увеличивать высоту микрорельефа.

Для формирования микрорельефа использовался позитивный фоторезист ФП-3535. На рис. 1 представлена зависимость показателя преломления данного резиста от длины волны, полученная на эллипсометре.

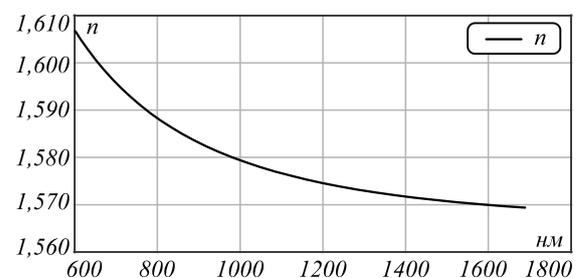


Рис. 1. Показатель преломления резиста ФП-3535

Максимальная толщина нанесения резиста составляет всего 2,5 мкм, таким образом, для получения

максимальной высоты рельефа необходимо трех-слойное нанесение резиста. Проведенные эксперименты по записи дифракционных линз на максимальную глубину показали, что при любом увеличении экспозиции не образуются структуры высотой более 7,5 мкм. Для определения причины такого явления на эллипсомере проведено измерение спектрального коэффициента пропускания резиста. Измерения проводились на пленке толщиной 7,5 мкм. Результаты измерений представлены на рис. 2.

Как видно из рис. 2, предельная высота формируемых на резисте ФП-3535 дифракционных структур объясняется очень просто, на длине волны 405 нм, на которой идет запись, наблюдается полное поглощение. Таким образом, записать более высокий рельеф на этом резисте не представляется возможным.

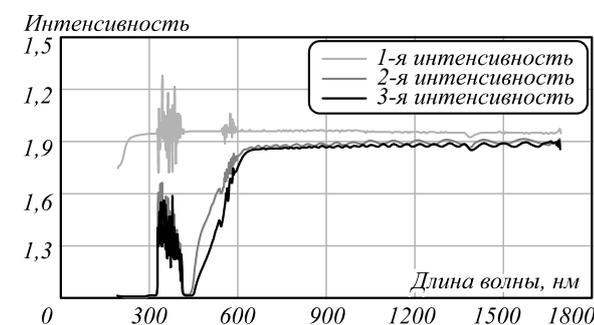


Рис. 2. Спектральный коэффициент пропускания резиста ФП-3535

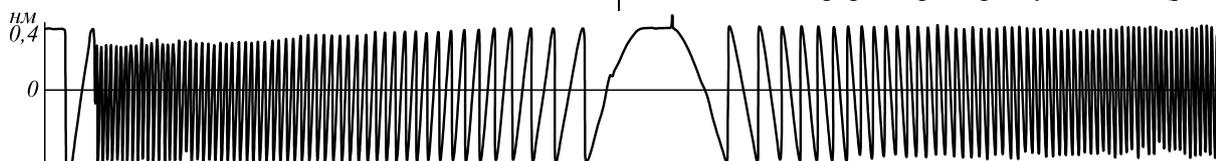


Рис. 3. Профилограмма дифракционной линзы с высотой микрорельефа 7,5 мкм

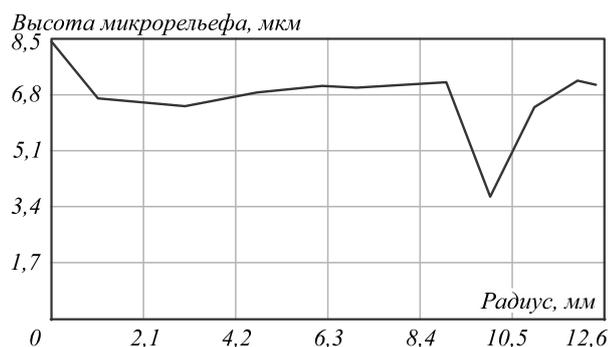


Рис. 4. Зависимость высоты микрорельефа от радиуса записи

На основе построенной зависимости и данных по коэффициенту поглощения резиста формировался новый подробный файл мощности, в котором точки переключения мощности располагались на расстоянии 1 мм. Новый файл мощности рассчитывался по формуле:

$$P(r) = \frac{h_0}{h} P_0(r), \quad (2)$$

По нашим сведениям, в настоящий момент этот резист – единственный доступный для формирования дифракционных структур с высоким рельефом на территории Российской Федерации, поэтому, несмотря на высокое поглощение, работы проводились с этим резистом.

Вывод на предельную глубину для ФП-3535 нежелателен по двум причинам: слишком большое время формирования дифракционных линз (20 – 30 часов при диаметре 25 мм) и сильная нелинейность записи на максимальной глубине (отклонение около 200 нм в центральной области и около 350 нм на краю элемента), приводящая к существенному искажению микрорельефа. На рис. 3 приведена профилограмма, демонстрирующая этот эффект.

На рис. 3 хорошо заметна нелинейность во впадинах микрорельефа. Для исключения этой нелинейности была выбрана высота рельефа 6 мкм, на которой нелинейность, вызванную высоким коэффициентом поглощения, можно скорректировать. Станция лазерной записи позволяет осуществлять ступенчатое изменение мощности записывающего пучка.

Для обеспечения высокой точности формирования микрорельефа изготовление каждой линзы осуществлялось в два этапа. На первом этапе формировался микрорельеф тестовой линзы с использованием общего файла мощности. На основании измеренной профилограммы микрорельефа строилась кривая зависимости высоты микрорельефа от радиуса записи (рис. 4).

где  $P_0$  – исходная мощность записывающего пучка;  
 $P$  – скорректированное значение мощности;  
 $h_0$  – высота микрорельефа в профилограмме;  
 $h$  – необходимая высота микрорельефа.

В дальнейшем эта процедура повторяется 2 – 3 раза с постепенным уточнением файла мощности. Попытка численно учесть влияние коэффициента поглощения не дала хорошего результата из-за сложности процесса взаимодействия между светом и резистом.

На рис. 5 представлен профиль дифракционной линзы с минимальным отклонением параметров от расчетных. Отклонение профиля на краю и в центре линзы не превышает 20 нм.

На рис. 5 наблюдается небольшой наклон профиля, который вызван клиновидностью оптической подложки, на которой была изготовлена линза.



Рис. 5. Общий вид профиля дифракционной линзы с фокусным расстоянием 70 мм и диаметром 10 мм

Были проведены измерения глубины микрорельефа от центра к краю дифракционной линзы, которые показали, что высота микрорельефа почти постоянна по области измерений. Заметно существенное отличие в центральной области, но это скорее недостаток метода измерения, поскольку размер центральной зоны превышает размер интервала измерения высоты. Таким образом, технологические особенности массового производства накладывают определенные ограничения на высоту воспроизводимого рельефа гармонической линзы, ограничивая тем самым и показатель гармоничности линзы (1). В дальнейших параграфах будет проведено исследование, как такое ограничение на высоту рельефа линзы влияет на качество получаемых изображений.

## 2. Нейросетевая реконструкция видеопотока

### 2.1. Нейросетевая технология реконструкции

Сквозная нейросетевая технология реконструкции изображений для дифракционно-оптических систем впервые предложена в работе [1]. Реконструкция изображений на основе предложенной технологии заключается в последовательной обработке изображений двумя нейронными сетями: генеративно-сопоставительной нейронной сетью (GAN) для устранения контекстно-зависимого цветового сдвига и сверточной нейронной сетью (CNN), компенсирующей хроматическое размытие изображений. В качестве GAN использовалась архитектура Pix2Pix [14], в качестве второй сети предлагались две архитектуры – модификация VDSR [15] и DRRN [4].

Как показали исследования, проведенные в работе [1], GAN позволяет значительно увеличить качество реконструкции изображений из тестовой выборки по сравнению с ранее используемыми методами [3, 4, 11, 17], однако недостатком GAN-обработки является появление артефактов на изображениях сцен реального мира. Проблема возникновения артефактов, характерная для сети Pix2Pix, в данной работе решается использованием сети U-Net. В пункте 3.2.1 приведено сравнительное исследование качества реконструкции изображений сцен реального мира на основе сетей U-Net и Pix2Pix. В данной работе предлагается модификация сквозной нейросетевой технологии реконструкции, предложенной в работе [1], для задачи реконструкции видеопотока.

На рис. 6 представлена схема реконструкции видеопотока. Вместо сети Pix2Pix, цветовая коррекция изображений выполняется сетью U-Net [18], которая выступает как сеть-генератор в Pix2Pix. Архитектуры VDSR и DRRN используются в том же виде, что и в работе [1]. Предложенная схема реконструкции включает этап предобработки, который необходим для формирования обучающей выборки. Данный этап заключается в имитации видеопотока, полученного на борту квадрокоптера.

Предлагается следующая модификация классической архитектуры U-Net, решающей задачу сегментации медицинских изображений [18]:

**Encoder:**C64-C128-C256-C512-C512-C512-C512-C512,  
**Decoder:**CD512- CD512- CD512-C512-C256-C128-C64,

где  $C_k$  – это комбинация сверточного слоя, пакетной (батч) нормализации и функции активации с количеством фильтров  $k$ ,  $CD_k$  означает, что перед активационным слоем добавляется dropout-слой. Размер окна свертки во всех случаях –  $4 \times 4$ . На каждом сверточном слое кодировщика (encoder) ширина и высота карт признаков уменьшаются в 2 раза, а слои декодировщика за счет транспонированной свертки напротив увеличивают размеры карт признаков, что позволяет получить на выходе сети восстановленное изображение такого же разрешения, как у входного изображения. В качестве функции активации в кодировщике используется Leaky ReLU, а в декодировщике – ReLU.

Отличием предложенной архитектуры от классической архитектуры U-Net [18] является наличие слоев батч-нормализации и dropout, которые используются для борьбы с эффектом переобучения. Кроме того, батч-нормализация позволяет скомпенсировать разницу в экспозиции между изображениями из обучающей выборки и изображениями сцен реального мира.

В рамках данной работы проведено экспериментальное исследование влияния функций потерь на качество цветовой коррекции на основе сети U-Net. В процессе обучения сети U-Net используются функции потерь, описанные в работе [19]. Это  $l_1$ ,  $l_2$ , SSIM [20], MS-SSIM [21], их комбинации, а также дополнительно исследуется функция Хьюбера [22]. Для оценки качества реконструкции предлагается использовать величину пикового отношения сигнала к шуму (PSNR), которая доказала свою эффективность в работах [1, 2, 3, 4, 11, 17].

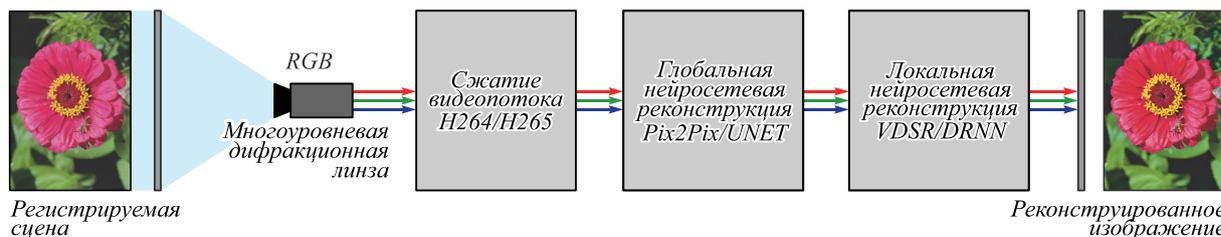


Рис. 6. Нейросетевая реконструкция видеопотока

## 2.2. Формирование данных для обучения

Формирование данных для обучения осуществлялось с использованием стенда в помещении. В состав стенда входят: откалиброванный монитор, разрешение которого составляет  $1920 \times 1080$  пикселей; камера Basler acA1920-40u, интерфейс USB 3.1 Gen 1, разрешение  $1920 \times 1200$  пикселей; ноутбук; объективы с резьбовым креплением C-mount, в которые установлены МДЛ. Штатив с камерой установлен таким образом, чтобы в область кадра камеры целиком умещался экран монитора. Внешние источники света были выключены, помещение затемнено. Выдержка и баланс белого установлены по выведенному на экран монитора изображению с заливкой белого цвета.

Фокусировка объектива осуществлена вручную с использованием миры «Siemens star», установленной в плоскости монитора на таком же расстоянии.

В набор изображений для съемки входит 1831 эталонное изображение и одно калибровочное изображение с маркерами. Съемка осуществлялась в один проход для каждой МДЛ.

Эталонное калибровочное изображение и его снимок на объектив с МДЛ используются для сопоставления всего набора эталонных изображений с полученными снимками. Маркеры предназначены для построения обратного проективного преобразования с целью коррекции геометрических искажений при регистрации камерой с МДЛ эталонных изображений.

В результате съемки сформированы два набора изображений, полученных на основе линз с показателями гармоничности 7 и 13 (1). Различные показатели гармоничности исследованы для того, чтобы определить влияние технологии производства на качество получения изображений и реконструкции. Каждый набор данных включает 1831 пару изображений (эталонное и снятое МДЛ) размера  $1024 \times 1024$  и случайным образом разбит на обучающую (1226 изображений) и тестовую (605 изображений) выборки.

## 2.3. Компенсация артефактов видеокодирования

Подход к формированию обучающей выборки, описанный в подпараграфе 2.2, использовался ранее при обработке изображений, снятых на основе МДЛ [1, 2, 3, 4, 11, 17]. Однако при реконструкции видеоизображений сетью U-Net, обученной на сформированном наборе данных, возникают артефакты, причиной которых является сжатие видеоизображений.

Для устранения артефактов сжатия схема реконструкции, представленная на рис. 7, включает этап предобработки обучающей выборки, который заключается в формировании видеопотока из изображений обучающей выборки по стандартам H.264/265 и последующей раскадровке полученного видео.

## 3. Экспериментальные исследования

В данном параграфе приведены экспериментальные исследования качества реконструкции на основе

схемы, предложенной на рис. 6: описана процедура обучения нейронных сетей, проведен сравнительный анализ качества реконструкции сетей U-Net и Pix2Pix, а также показаны результаты исследования артефактов при сжатии видеопотока без потерь и с потерями по стандартам H.264/265 [8, 9].

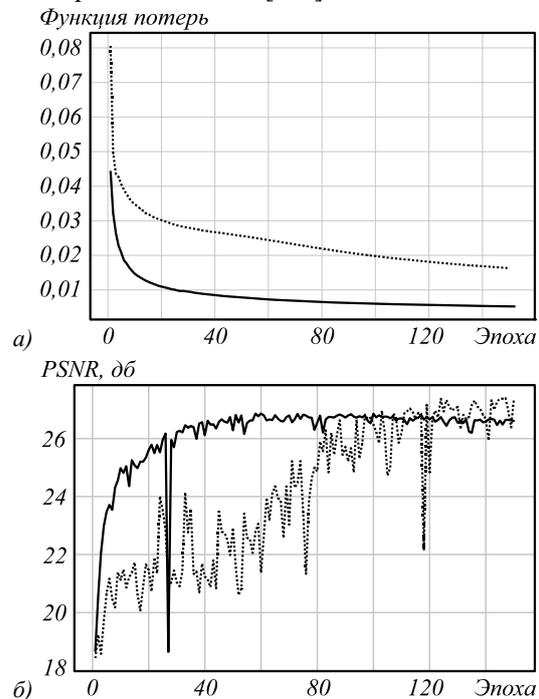


Рис. 7. Графики зависимости функции потерь (а) и PSNR (б) на валидационной выборке

## 3.1. Обучение нейросетевых моделей реконструкции изображений

Обучение U-Net проводилось на полноразмерных изображениях ( $1024 \times 1024$ ) обучающей выборки с помощью мини-пакетного стохастического градиентного спуска и метода оптимизации Adam [23]. Параметр скорости обучения (learning rate) составил  $0,0002$ ,  $\beta_1 = 0,5$ ,  $\beta_2 = 0,999$ ,  $\epsilon = 10^{-8}$ .

Для обучения VDSR изображения были разбиты на фрагменты размера  $51 \times 51$  с шагом 51 (stride). Кроме того, обучающая выборка была расширена путем поворота изображений на  $90^\circ$ ,  $180^\circ$ ,  $360^\circ$ , операции «отражения» (flip) по горизонтали и добавления Гауссовского шума с нулевым средним и дисперсией, равной 0,01. В результате для каждого фрагмента изображения из обучающей выборки было сформировано 15 дополнительных фрагментов.

Метод оптимизации Adam использовался также при обучении сети VDSR. Размер мини-пакета составил 64. Глубина сети, начальное значение скорости обучения,  $\beta_1$ ,  $\beta_2$  и  $\epsilon$  были установлены равными 18, 0,9,  $10^{-4}$ , 0,999 и  $10^{-8}$  соответственно.

Для обучения DRRN изображения были разбиты на фрагменты размера  $51 \times 51$ . Так же, как и в случае обучения VDSR, обучающая выборка была расширена путем поворота изображений на  $90^\circ$ ,  $180^\circ$ ,  $360^\circ$  и

операции «отражения» (flip) по горизонтали. В результате для каждого оригинального изображения из обучающей выборки было сформировано 7 дополнительных фрагментов.

На рис. 7 представлены графики зависимости функций потерь (функция Хьюбера для  $m = 7$  и  $l_2$  для  $m = 13$ ) и среднего значения PSNR на валидационной выборке (15 изображений из тестовой выборки) в процессе обучения сети U-Net.

3.2. Экспериментальный анализ качества реконструкции

3.2.1. Сравнение качества реконструкции на основе сетей Pix2Pix и U-Net

Экспериментальные исследования качества реконструкции на основе сетей U-Net и Pix2Pix показыва-

ли, что сеть Pix2Pix вносит дополнительные артефакты на изображения сцен реального мира. На рис. 8 представлены результаты обработки изображений сцен реального мира сетью Pix2Pix и U-Net. Видно, что на фрагментах изображений, обработанных сетью Pix2Pix (4 столбец), появляются артефакты. Сеть U-Net не вносит дополнительных артефактов на изображения, снятые при искусственном освещении объективом с линзой с показателем гармоничности 13 (1 строка на рис. 8). Однако при обработке кадров, полученных на открытом воздухе, местами возникают цветовые искажения в виде серых пятен.

При визуальной оценке качества реконструкции сеть U-Net превосходит сеть Pix2Pix благодаря значительному сокращению количества артефактов на изображениях сцен реального мира.

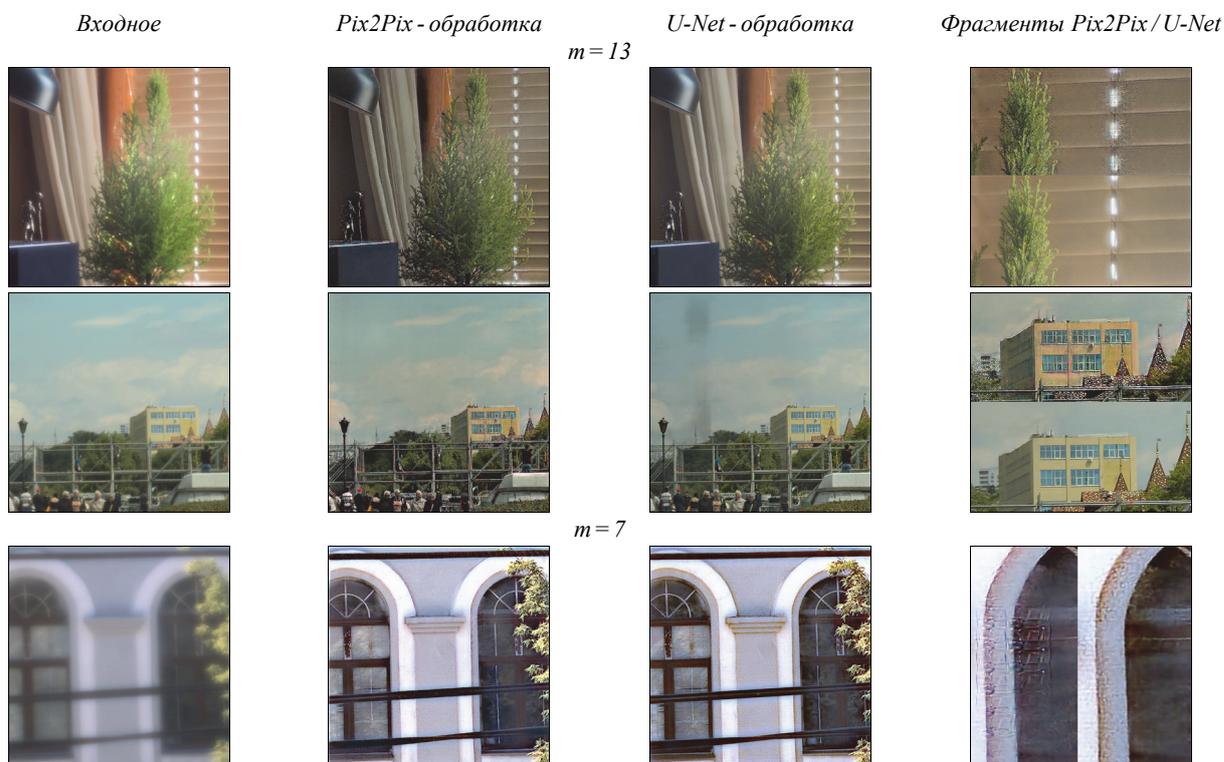


Рис. 8. Результаты цветовой коррекции изображений реальных сцен

3.2.2. Реконструкция при сжатии без потерь

В табл. 1 приведены результаты оценки качества реконструкции при обучении сети U-Net с использованием разных функций потерь.

Для оценки качества реконструкции использовалось среднее значение PSNR на тестовой выборке. Видно, что для линзы с показателем гармоничности 13 наилучшее качество реконструкции обеспечивает функция потерь  $l_2$ , а для линзы с показателем гармоничности 7 лучший результат достигнут благодаря использованию функции Хьюбера. Таким образом, функция потерь Хьюбера и  $l_2$  выбраны для обучения сети U-Net в задаче реконструкции видеопотока.

Табл. 1. Оценка качества реконструкции при использовании разных функций потерь

Функция потерь	PSNR для линзы с показателем гармоничности 13	PSNR для линзы с показателем гармоничности 7
$l_1$	24,67	26,54
$l_2$	27,02	26,44
SSIM [6]	25,6	24,94
MS-SSIM [6]	25,9	25,03
Комбинация MS-SSIM и $l_1$	26,28	26,57
Комбинация MS-SSIM и $l_2$	26,37	26,55
Ф-я Хьюбера	25,86	26,65

Оценка качества модификации нейросетевой технологии реконструкции, предложенной в данной работе,

приведена в табл. 2, где указаны средние значения PSNR, рассчитанные на тестовой выборке (605 изображений) для необработанных, цветокорректированных и восстановленных изображений. Кроме того, в табл. 2 указаны средние значения PSNR, рассчитанные для 25 лучших по критерию PSNR изображений (ТОП-25).

Табл. 2. Оценка качества нейросетевой технологии реконструкции

$m$	Исходное	Результат цветокоррекции (U-Net)	Результат реконструкции (DRRN/ VDSR)	ТОП-25 (VDSR)
13	19,68	27,02	27,06/27,26	31,94
7	18,13	26,65	26,66/26,82	31,28

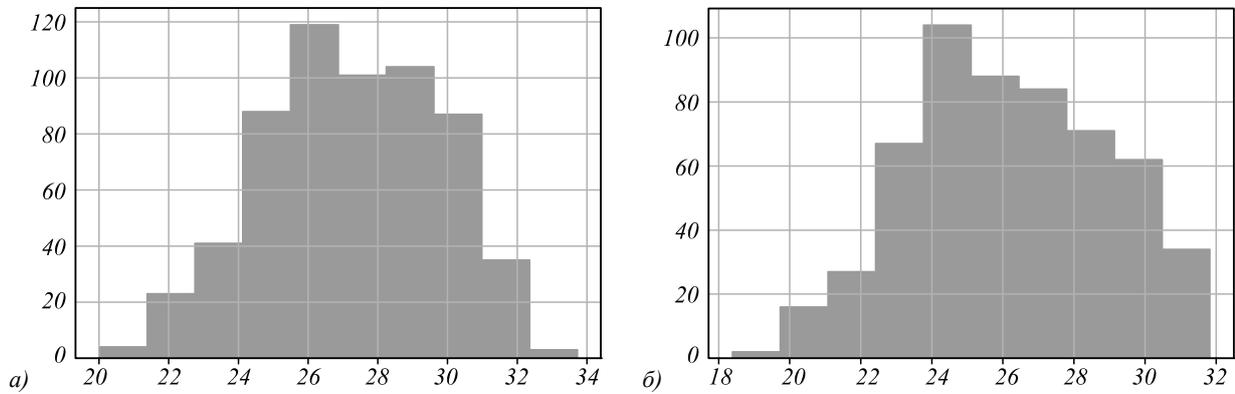


Рис. 9. Гистограммы распределения PSNR на тестовой выборке:  $m = 13$  (а),  $m = 7$  (б)

Снятое МДЛ                      U-Net-обработка                      DRRN/VDSR-обработка                      Эталонное



Рис. 10. Примеры изображений из тестовой выборки для  $m = 13$  и  $m = 7$

Из табл. 2 видно, что предложенная технология реконструкции не только обеспечивает высокий уровень PSNR, но и не зависит от качества входного набора данных, сокращая разницу между средними значениями PSNR для изображений, снятых двумя разными МДЛ, до 0,4 дБ. Также видно, что сеть VDSR превосходит по качеству реконструкции сеть DRRN, поэтому далее будет использоваться только комбинация сетей U-Net и VDSR. На рис. 9 представлены гистограммы распределения среднего значения PSNR на тестовой выборке, построенные после реконструкции сетями U-Net и VDSR.

На рис. 10 приведены примеры восстановленных изображений из тестовой выборки.

### 3.2.3. Реконструкция видеопотока при сжатии с потерями

В результате реконструкции видеопотока по схеме, приведенной на рис. 6, но без предобработки обучающей выборки возникают артефакты, причиной которых является использование алгоритмов видеокодирования. В данной работе предлагается решение проблемы артефактов сжатия, которое заключается в предварительной обработке обучающей выборки методом, описанным в подпараграфе 2.3.

Рис. 11 демонстрирует возникновение артефактов сжатия на примере изображений из тестовой выборки для линзы с показателем гармоничности 7. На рис. 11а представлен пример реконструкции изображения без предобработки алгоритмом сжатия. На рис. 11б показаны результаты реконструкции изображения после предобработки алгоритмом сжатия. Восстановление изображений проводилось по схеме, предложенной на рис. 6, но без предобработки обучающей выборки. Видно, что в случае имитации видеопотока (рис. 11б) на восстановленном изображении появляются артефакты. На основе проведенного эксперимента можно сделать вывод, что причиной артефактов является использование алгоритмов видеокодирования.



Рис. 11. Примеры изображений из тестовой выборки: без использования предобработки (а), с использованием предобработки (б)

Для устранения артефактов видеокодирования была проведена предобработка обучающей выборки согласно подпараграфу 2.3. На рис. 12 представлены графики зависимости среднего значения PSNR на тестовой выборке от значения коэффициента постоянного сжатия, CRF (Constant Rate Factor, CRF) [10].

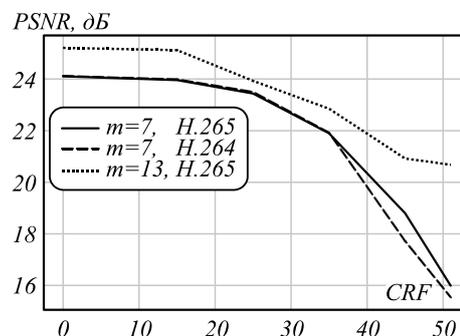


Рис. 12. График зависимости среднего значения PSNR на тестовой выборке от CRF

Чем ниже значение CRF, тем выше качество видеоизображения. В данном случае для задачи реконструкции видеопотока мы использовали CRF, равный 25, поскольку в таком случае на восстановленных видеокадрах, снятых с квадрокоптера, полностью отсутствовали артефакты видеокодирования. Графики для видеокодексов H.264 и H.265 практически совпадают, поэтому далее в работе приведены результаты только для H.265. В табл. 3 представлены результаты реконструкции изображений из тестовой выборки при использовании алгоритма сжатия с параметром CRF, равным 25.

Среднее значение PSNR на тестовой выборке составило 23,93 дБ для линзы с показателем гармоничности 13 и 23,44 дБ для линзы с показателем гармоничности 7.

Табл. 3. Оценка качества реконструкции при использовании сжатия с потерями

<i>m</i>	Исходное	Результат сжатия	Результат реконструкции
13	19,68	19,72	23,93
7	18,13	17,92	23,44

На рис. 13 изображены примеры реконструкции видеокадров, полученных на борту квадрокоптера на основе линзы с показателем гармоничности 7. Приведено сравнение фрагментов изображений, восстановленных по схеме, предложенной на рис. 7, с предобработкой обучающей выборки и без предобработки. Видно, что в случае предобработки обучающей выборки качество реконструкции видеоизображений значительно лучше, артефакты сжатия исчезают. В работе значения PSNR рассчитывались на независимых кадрах из тестовой выборки. При визуальном анализе восстановленной последовательности видеокадров не было обнаружено временных колебаний качества реконструкции. Таким образом, предложенная схема реконструкции обеспечивает согласованность последовательных кадров одной сцены.

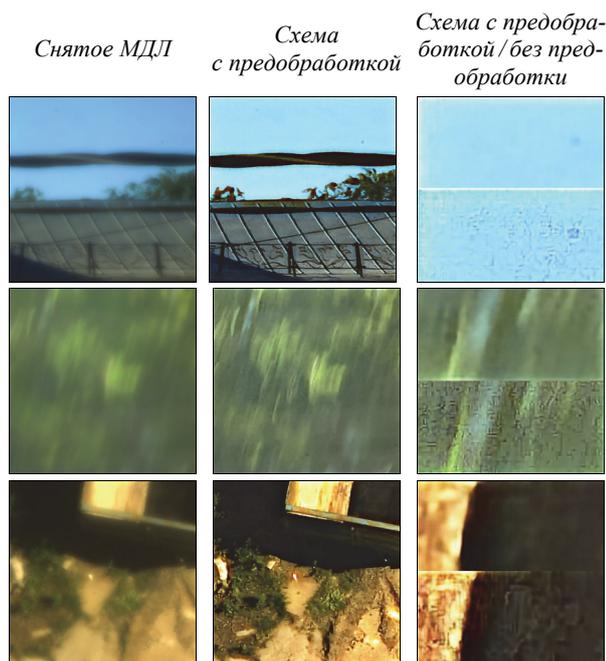


Рис. 13. Примеры обработки видеокадров, снятых на борту квадрокоптера

### Заключение

В настоящей работе исследованы артефакты и проблемы ухудшения качества изображений, получаемых в системах на основе МДЛ. Показано, что технологические особенности массового производства, приводящие к существенному уменьшению высоты рельефа МДЛ, а также артефакты видеокodирования заметно снижают качество получаемых изображений, до 18 и 19 дБ по показателю PSNR. Однако предложенная нейросетевая технология реконструкции позволяет в существенной степени нивелировать это различие, сведя его до менее чем 0,4 дБ. Таким образом, показатель гармоничности линзы не оказывает сильного влияния на качество изображений, восстановленных на основе предложенной технологии.

Получаемые в результате нейросетевой реконструкции видеопотока изображения демонстрируют среднее значение PSNR более 27 дБ при сжатии без потерь и почти 24 дБ при сжатии с потерями, что является приемлемым уровнем качества для решения задач технического зрения в системах видеонаблюдения и мобильных устройствах. Исследование эффективности МДЛ в конкретных задачах технического зрения, таких как трекинг объектов, распознавание лиц, инфраструктурных объектов, является задачей исследований ближайшего будущего.

### Благодарности

Теоретическая часть работы и разработка нейросетевых моделей выполнена при поддержке гранта РФФ 20-69-47110, экспериментальная часть выполнена при поддержке грантов РФФИ № 18-07-01390-А и № 19-29-01235-мк, а также в рамках государственного задания

ИСОИ РАН – филиала Федерального научно-исследовательского центра «Кристаллография и фотоника» РАН (соглашение № 007-ГЗ/Ч3363/26).

### Литература

1. **Nikonorov, A.** Deep learning-based imaging using single-lens and multi-aperture diffractive optical systems / A. Nikonorov, V. Evdokimova, M. Petrov, P. Yakimov, S. Bibikov, Y. Yuzifovich, R. Skidanov, N. Kazanskiy // 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). – 2019. – P. 3969-3977. – DOI: 10.1109/ICCVW.2019.00491.
2. **Nikonorov, A.** Deep learning-based image reconstruction for multi-aperture diffractive lens / A. Nikonorov, V. Evdokimova, M. Petrov, S. Bibikov, A. Alekseev, R. Skidanov, N. Kazanskiy // Journal of Physics: Conference Series. – 2019. – Vol. 1368. – 052031. – DOI: 10.1088/1742-6596/1368/5/052031.
3. **Nikonorov, A.** Fresnel lens imaging with post-capture image processing / A. Nikonorov, R. Skidanov, V. Fursov, M. Petrov, S. Bibikov, Y. Yuzifovich // Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – 2015. – P. 33-41. – DOI: 10.1109/CVPRW.2015.7301373.
4. **Nikonorov, A.** Toward ultralightweight remote sensing with harmonic lenses and convolutional neural networks / A. Nikonorov, M. Petrov, S. Bibikov, P. Yakimov, V. Kutikova, Y. Yuzifovich, A. Morozov, R. Skidanov, N. Kazanskiy // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2018. – Vol. 11, Issue 9. – P. 3338-3348. – DOI: 10.1109/JSTARS.2018.2856538.
5. **Genevet, P.** Recent advances in planar optics: from plasmonic to dielectric metasurfaces / P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, R. Devlin // Optica. – 2017. – Vol. 4, Issue 1. – P. 139-152.
6. **Peng, Y.** Computational imaging using lightweight diffractive-refractive optics / Y. Peng, Q. Fu, H. Amata, Sh. Su, F. Heide, W. Heidrich // Optics Express. – 2015. – Vol. 23, Issue 24. – P. 31393-31407.
7. **Sun, T.** Revisiting cross-channel information transfer for chromatic aberration correction / T. Sun, Y. Peng, W. Heidrich // IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 3268-3276. – DOI: 10.1109/ICCV.2017.352.
8. **Rao, K.R.** Video coding standards. AVS China, H.264/MPEG-4 PART 10, HEVC, VP6, DIRAC and VC-1. / K.R. Rao, D.N. Kim, J.J. Hwang. – Dordrecht, Heidelberg, New York, London: Springer, 2013. – 499p.
9. **Naeem, R.** Quality of experience evaluation of H.265/MPEG-HEVC and VP9 comparison efficiency / R. Naeem, P. Zeeshan, A. Abbes // 2014 26<sup>th</sup> International Conference on Microelectronics (ICM). – 2015. – DOI: 10.1109/ICM.2014.7071846.
10. **Bienik, J.** Impact of constant rate factor on objective video quality assessment / J. Bienik, M. Uhrina, P. Kortis // Advances in Electrical and Electronic Engineering. – 2017. – Vol. 15, Issue 4. – P. 673-682. – DOI: 10.15598/aeee.v15i4.2387.
11. **Nikonorov, A.** Towards multi-aperture imaging using diffractive lens / A. Nikonorov, R. Skidanov, V. Kutikova, M. Petrov, A. Alekseev, S. Bibikov, N. Kazanskiy // Proceedings of SPIE. – 2019. – Vol. 11146. – 111460Y. – DOI: 10.1117/12.2526923.

12. **Tan, J.** Face detection and verification using lensless cameras / J. Tan, L. Niu, J.K. Adams, V.Boominathan, J.T. Robinson, R.G. Baraniuk, A. Veeraraghavan // IEEE Transactions on Computational Imaging. – 2018. – Vol. 5, Issue 2. – P. 180-194.
13. Sweeney, D.W. Harmonic diffractive lenses / D.W. Sweeney, G.E. Sommargren // Applied Optics. – 1995. – Vol. 34, Issue 14. – P. 2469-2475. – DOI: 10.1364/AO.34.002469.
14. **Isola, P.** Image-to-image translation with conditional adversarial networks / P. Isola, J.-Y. Zhu, T. Zhou, A. Efros. – arXiv Preprint v3. – 2018. – URL: <https://arxiv.org/abs/1611.07004> (request date 2.11.2020).
15. **Kim, J.** Accurate image super-resolution using very deep convolutional networks / J. Kim, J. Lee, K. Lee // IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – arXiv Preprint v2. – URL: <https://arxiv.org/abs/1511.04587> (request date 2.11.2020).
16. **Tai, Y.** Image super-resolution via deep recursive residual network / Y. Tai, J. Yang, X. Liu // IEEE Conference on Computer Vision and Pattern Recognition. – 2017. – P. 2790-2798. – DOI: 10.1109/CVPR.2017.298.
17. **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 // 2016 23<sup>rd</sup> International Conference on Pattern Recognition (ICPR). – 2016. – P. 775-780. – DOI: 10.1109/ICPR.2016.7899729.
18. **Ronneberger, O.** U-Net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox. – In: Medical image computing and computer-assisted intervention – MICCAI 2015 / ed. by N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi. – New York, Dordrecht, London: Springer, 2015. – P. 234-241. – DOI: 10.1007/978-3-319-24574-4\_28.
19. **Zhao, H.** Loss functions for image restoration with neural networks / H. Zhao, O. Gallo, I. Frosio, J. Kautz // IEEE Transactions on Computational Imaging. – 2016. – Vol. 3. – P. 47-57. – DOI: 10.1109/TCI.2016.2644865.
20. **Wang, Z.** Image quality assessment: from error visibility to structural similarity / Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli // IEEE Transactions on Image Processing. – 2004. – Vol. 13, Issue 4. – P. 600-612. – DOI: 10.1109/TIP.2003.819861.
21. **Wang, Z.** Multiscale structural similarity for image quality assessment / Z. Wang, E.P. Simoncelli, A.C. Bovik // Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers. – 2003. – Vol. 2. – P. 1398-1402. – DOI: 10.1109/ACSSC.2003.1292216.
22. **Huber, P.J.** Robust estimation of a location parameter / P.J. Huber // The Annals of Mathematical Statistics. – 1964. – Vol. 35, Issue 1. – P. 73-101.
23. **Kingma, D.P.** Adam: A method for stochastic optimization / D.P. Kingma, J. Ba // Proceedings of the 3<sup>rd</sup> International Conference on Learning Representations (ICLR 2015). – 2017. – arXiv Preprint v9. – URL: <https://arxiv.org/abs/1412.6980> (request date 2.11.2020).
24. **Dun, X.** Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging / X. Dun, H. Ikoma, G. Wetzstein, Z. Wang, X. Cheng, Y. Peng // Optica. – 2020. – Vol. 7, Issue 8. – P. 913-922. – DOI: 10.1364/OPTICA.394413.
25. **Li, R.** DeepUNet: A deep fully convolutional network for pixel-level sea-land segmentation / R. Li, W. Liu, L. Yang, S. Sun, W. Hu, F. Zhang, W. Li // Computer Science: Computer Vision and Pattern Recognition (cs.CV). – 2017. – arXiv Preprint. – URL: <https://arxiv.org/abs/1709.00201> (request date 2.11.2020).
26. **Peng, Y.** Learned large field-of-view imaging with thin-plate optics / Y. Peng, Q. Sun, X. Dun, G. Wetzstein, F. Heide // ACM Transactions on Graphics. – 2019. – Vol. 38, Issue 6. – P. 219. – DOI: 10.1145/3355089.3356526.

### Сведения об авторах

**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». Аспирант кафедры суперкомпьютеров и общей информатики Самарского университета. Область научных интересов: компьютерная обработка изображений, теория распознавания образов, интеллектуальный анализ данных, глубокое обучение. E-mail: [vkutikova94@gmail.com](mailto: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](mailto:max.vit.petrov@gmail.com).

**Ключева Марина Александровна**, родилась в 1995 году. В 2018 году окончила магистратуру Самарского национального исследовательского университета имени академика С.П. Королёва по направлению «Прикладная математика и информатика». Область научных интересов: распознавание образов и анализ изображений, глубокое обучение. E-mail: [golova.m@bk.ru](mailto:golova.m@bk.ru).

**Зыбин Евгений Юрьевич**, 1978 года рождения. В 2000 году окончил с отличием Военный авиационный технический университет по специальности «Инженер-электрик авиационного оборудования». В 2004 году защитил кандидатскую диссертацию в Высшей военной инженерной академии им. Н.Е. Жуковского. В 2016 году

защитил докторскую диссертацию в Научно-исследовательском институте стандартизации и унификации. Область научных интересов: матричные вычисления, интеллектуальный анализ данных, параллельные и распределённые вычисления, математическое моделирование, теория управления, диагностика технического состояния. Автор и соавтор более 140 научных трудов. E-mail: [evzybin@2100.gosniias.ru](mailto:evzybin@2100.gosniias.ru).

**Косьянчук Владислав Викторович**, 1970 года рождения. Окончил Военно-воздушную академию им. Н.Е. Жуковского и Ю.А. Гагарина, г. Москва. Профессор РАН, доктор технических наук (2003 год), профессор (2007 год). Работает заместителем генерального директора Государственного научно-исследовательского института авиационных систем по гражданской авиации. Специалист в области управления и навигации авиационных комплексов. Опубликовал более 250 научных трудов. E-mail: [kos.vl.v@gmail.com](mailto:kos.vl.v@gmail.com).

**Мищенко Ирина Борисовна**, родилась в 1995 году. В 2019 году окончила магистратуру Новосибирского государственного технического университета физико-технического факультета по направлению «Оптехника». Аспирант ФГУП «ГосНИИАС» по направлению подготовки «Управление в технических системах». Инженер подразделения «Перспективные комплексы бортового оборудования гражданской авиации» ФГУП «ГосНИИАС». В настоящее время основные сферы научных интересов: исследование и создание материалов и структур, активно применяемых в областях микроэлектроники и фотоники, в том числе фотонные интегральные схемы авиационного применения. Является автором и соавтором 8 научных трудов. E-mail: [mishchenko.ib@gmail.com](mailto:mishchenko.ib@gmail.com).

**Новиков Валерий Михайлович**, родился в 1954 году. В 1981 году окончил Московский авиационный институт по специальности «Электронные вычислительные машины». Начальник сектора «Разработка перспективной архитектуры комплекса бортового радиоэлектронного оборудования» ФГУП «ГосНИИАС». В настоящее время основными сферами научных интересов являются: создание архитектур перспективных комплексов БРЭО, исследование форм применения в комплексах БРЭО оптических компонентов на борту самолёта, в том числе фотонных интегральных схем авиационного применения. Является автором и соавтором 40 научных трудов и пяти патентов. E-mail: [vmnovikov@2100.gosniias.ru](mailto:vmnovikov@2100.gosniias.ru).

**Сельвесюк Николай Иванович**, 1973 года рождения. В 1994 году с отличием окончил Ставропольское высшее авиационное инженерное училище им. маршала авиации В.А. Судца. Доктор технических наук (2007 г.), доцент, профессор РАН. Работает заместителем генерального директора – руководителем научного комплекса государственного научного центра ФГУП «Государственный научно-исследовательский институт авиационных систем». Специалист в области линейной алгебры, многосвязных систем управления, анализа и обработки изображений. Опубликовал более 120 научных работ, 16 учебников и учебных пособий, 5 авторских свидетельств и патентов. E-mail: [nis@gosniias.ru](mailto:nis@gosniias.ru).

**Ершов Егор Иванович**, 1990 года рождения, кандидат физико-математических наук, сотрудник Лаборатории зрительных систем Института проблем передачи информации РАН. Окончил МФТИ в 2014 году. Область научных интересов: цветное компьютерное зрение, быстрое преобразование Хафа и Радона, визуальная одометрия. E-mail: [ershov@iitp.ru](mailto:ershov@iitp.ru).

**Ивлиев Николай Александрович**, родился в 1987 году. В 2010 году окончил Самарский государственный аэрокосмический университет, инженер по специальности «Проектирование и технология радиоэлектронных средств». Область научных интересов: физика поверхности твердого тела. Является автором и соавтором более 30 научных трудов. E-mail: [ivlievn@gmail.com](mailto:ivlievn@gmail.com).

**Скиданов Роман Васильевич**, 1973 года рождения. В 1995 году окончил физический факультет Самарского государственного университета. В 1998 году защитил кандидатскую диссертацию в Самарском государственном аэрокосмическом университете, а в 2007 году защитил докторскую диссертацию. В 2011 году получил награду «Лауреат Губернской премии в области науки и техники». В настоящее время основные сферы научных интересов: дифракционная оптика и нанофотоника. Р.В. Скиданов опубликовал более 95 научных трудов и 6 монографий. E-mail: [romans@ipsiras.ru](mailto:romans@ipsiras.ru).

**Казанский Николай Львович**, 1958 года рождения. В 1981 году с отличием окончил Куйбышевский авиационный институт (КуАИ, ныне – Самарский национальный исследовательский университет имени академика С.П. Королёва) по специальности «Прикладная математика». Доктор физико-математических наук (1996 год), профессор, работает руководителем Института систем обработки изображений РАН – филиала Федерального

научно-исследовательского центра «Кристаллография и фотоника» Российской академии наук (ИСОИ РАН), профессором кафедры технической кибернетики Самарского университета. Заведующий базовой (Самарского университета в ИСОИ РАН) кафедрой высокопроизводительных вычислений. Является членом международных научных обществ OSA, SPIE и IAPR. Н.Л. Казанский – специалист в области дифракционной оптики, математического моделирования, обработки изображений и нанофотоники. В списке научных работ Н.Л. Казанского 380 статей, 12 монографий, 53 авторских свидетельства и патента. Страница в интернете: <http://www.ipsi.smr.ru/staff/kazanskiy.htm> . E-mail: [kazanskiy@ssau.ru](mailto:kazanskiy@ssau.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](mailto:artniko@gmail.com) .

---

ГРНТИ: 28.23.15

Поступила в редакцию 12 ноября 2020 г. Окончательный вариант – 8 декабря 2020 г.

---

---

# Deep learning-based video stream reconstruction in mass-production diffractive optical systems

V. Evdokimova<sup>1,2</sup>, M. Petrov<sup>1,2</sup>, M. Klyueva<sup>1,2</sup>, E. Zybin<sup>3</sup>, V. Kosianchuk<sup>3</sup>, I. Mishchenko<sup>3</sup>, V. Novikov<sup>3</sup>, N. Selvesiuk<sup>3</sup>, E. Ershov<sup>4</sup>, N. Ivliev<sup>1,2</sup>, R. Skidanov<sup>1,2</sup>, N. Kazanskiy<sup>1,2</sup>, A. Nikonorov<sup>1,2</sup>

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

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

<sup>3</sup>Federal State Unitary Enterprise State Research Institute of Aviation Systems, 125319, Russia, Moscow, Viktorenko, 7,

<sup>4</sup>Institute for Information Transmission Problems, RAS, 127051, Moscow, Russia, Bolshoy Karetny per. 19, build 1

## Abstract

Many recent studies have focused on developing image reconstruction algorithms in optical systems based on flat optics. These studies demonstrate the feasibility of applying a combination of flat optics and the reconstruction algorithms in real vision systems. However, additional causes of quality loss have been encountered in the development of such systems. This study investigates the influence on the reconstructed image quality of such factors as limitations of mass production technology for diffractive optics, lossy video stream compression artifacts, and specificities of a neural network approach to image reconstruction. The paper offers an end-to-end deep learning-based image reconstruction framework to compensate for the additional factors of quality losing. It provides the image reconstruction quality sufficient for applied vision systems.

**Keywords:** diffractive optics, diffractive lenses, deep learning-based reconstruction, image processing.

**Citation:** 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.

**Acknowledgements:** The theoretical part and neural network models were developed with the support from the Russian Science Foundation under RSF grant 20-69-47110. The experimental part was executed with the support from the Russian Foundation for Basic Research under RFBR grant 18-07-01390-A and under the government project of the IPSI RAS – a branch of the Federal Scientific-Research Center "Crystallography and Photonics" of the RAS (agreement 007-Г3/Ч3363/26).

## References

- [1] [1] Nikonorov A, Evdokimova A, 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. *IEEE/CVF ICCV* 2019: 3969-3977. DOI: 10.1109/ICCV.2019.00491.
  - [2] Nikonorov A, Evdokimova V, Petrov M, Bibikov S, Alekseev A, Skidanov R, Kazanskiy N. Deep learning-based image reconstruction for multi-aperture diffractive lens. *J Phys Conf Ser* 2019; 1368: 052031. DOI: 10.1088/1742-6596/1368/5/052031.
  - [3] Nikonorov A, Skidanov R, Fursov V, Petrov M, Bibikov S, Yuzifovich Y. Fresnel lens imaging with post-capture image processing. *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* 2015: 33-41. DOI: 10.1109/CVPRW.2015.7301373.
  - [4] 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.
  - [5] 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.
  - [6] 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.
  - [7] Sun T, Peng Y, Heidrich W. Revisiting cross-channel information transfer for chromatic aberration. *IEEE International Conference on Computer Vision (ICCV)* 2017: 3268-3276. DOI: 10.1109/ICCV.2017.352.
  - [8] Rao KR, Kim DN, Hwang JJ. Video coding standards. AVS China, H.264/MPEG-4 PART 10, HEVC, VP6, DIRAC and VC-1. Dordrecht, Heidelberg, New York, London: Springer; 2013.
  - [9] Naem R, Zeeshan P, Abbes A. Quality of experience evaluation of H.265/MPEG-HEVC and VP9 comparison efficiency. *2014 26th International Conference on Microelectronics (ICM)* 2015. DOI: 10.1109/ICM.2014.7071846.
  - [10] Bienik J, Uhrina M, Kortis P. Impact of constant rate factor on objective video quality assessment. *Adv Elect Electron Eng* 2017; 15(4): 673-682. DOI: 10.15598/aece.v15i4.2387.
  - [11] Nikonorov A, Skidanov R, Kutikova V, Petrov M, Alekseev A, Bibikov S, Kazanskiy N. Towards multi-aperture imaging using diffractive lens. *Proc SPIE* 2019; 11146: 111460Y. DOI: 10.1117/12.2526923.
-

- 
- [12] Tan J, Niu L, Adams JK, Boominathan V, Robinson JT, Baraniuk RG, Veeraraghavan A. Face detection and verification using lensless cameras. *IEEE Trans Comput Imaging* 2018; 5(2): 180-194.
- [13] Sweeney DW, Sommargen GE. Harmonic diffractive lenses. *Appl Opt* 1995; 34(14): 2469-2475. DOI: 10.1364/AO.34.002469.
- [14] Isola P, Zhu J-Y, Zhou T, Efros A. Image-to-image translation with conditional adversarial networks. *arXiv Preprint v3* 2018. Source: (<https://arxiv.org/abs/1611.07004>).
- [15] Kim J, Lee J, Lee K. Accurate image super-resolution using very deep convolutional networks. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2016. *arXiv Preprint v2* 2016. Source: (<https://arxiv.org/abs/1511.04587>).
- [16] Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2017: 2790-2798. DOI: 10.1109/CVPR.2017.298.
- [17] Nikonov 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. *ICPR* 2016: 775-780. DOI: 10.1109/ICPR.2016.7899729.
- [18] Ronneberger O, Fischer P, Brox T. 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 2015*: 234-241. DOI: 10.1007/978-3-319-24574-4\_28.
- [19] Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE Trans Comput Imaging* 2016; 3: 47-57. DOI: 10.1109/TCI.2016.2644865.
- [20] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 2004; 13(4): 600-612. DOI: 10.1109/TIP.2003.819861.
- [21] Wang Z, Simoncelli EP, Bovik AC. Multiscale structural similarity for image quality assessment. *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers* 2003; 2: 1398-1402. DOI: 10.1109/ACSSC.2003.1292216.
- [22] Huber PJ. Robust estimation of a location parameter. *The Annals of Mathematical Statistics* 1964; 35(1): 73-101.
- [23] Kingma DP, Ba J. Adam: A method for stochastic optimization. *Proc 3<sup>rd</sup> Intl Conf on Learning Representations (ICLR)* 2015. *arXiv Preprint v9* 2017. Source: (<https://arxiv.org/abs/1412.6980>).
- [24] 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(8): 913-922. DOI: 10.1364/OPTICA.394413.
- [25] Li R, Liu W, Yang L, Sun S, W. Hu, F. Zhang, W. Li. DeepUNet: A deep fully convolutional network for pixel-level sea-land segmentation. *Computer Science: Computer Vision and Pattern Recognition (cs.CV)*. *arXiv Preprint* 2017. Source: (<https://arxiv.org/abs/1709.00201>).
- Peng Y, Sun Q, Dun X, Wetzstein G, Heide F. Learned large field-of-view imaging with thin-plate optics. *ACM Trans Graph* 2019; 38(6): 219. DOI: 10.1145/3355089.3356526.
- 

#### *Authors' information*

**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](mailto: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](mailto:max.vit.petrov@gmail.com).

**Marina Alexandrovna Klyueva** (b. 1995) graduated from Samara National Research University in 2018 with a Master's degree, majoring in Applied Mathematics and Informatics. Research interests: pattern recognition and image analysis, deep learning. E-mail: [golova.m@bk.ru](mailto:golova.m@bk.ru).

**Eugene Yurievich Zybin** (b. 1978) graduated with honors (2000) from the Military Aviation Technical University with a degree in Aviation Equipment Electrical Engineer, received Candidate in Engineering degree from Zhukovsky Air Force Engineering Academy (2004) and Doctor in Engineering degree from Scientific Research Institute for Standardization and Unification (2016). Research interests: matrix computing, data mining, parallel and distributed computing, mathematical modeling, control theory, health monitoring. He is author and co-author of over 140 scientific papers. E-mail: [evzybin@2100.gosniias.ru](mailto:evzybin@2100.gosniias.ru).

**Vladislav Viktorovich Kosianchuk** (b. 1970) graduated from Zhukovsky – Gagarin Air Force Academy in Moscow. He is Professor of the Russian Academy of Sciences, doctor of Technical Sciences (2003), Professor (2007). He works as the Deputy General Director of the State Research Institute of Aviation Systems for Civil Aviation. He is a specialist in the field of control and navigation of aviation complexes, and he has published more than 250 scientific papers. E-mail: [kos.vl.v@gmail.com](mailto:kos.vl.v@gmail.com).

---

---

**Irina Borisovna Mishchenko** (b. 1995). She graduated from Novosibirsk State Technical University with a Master's degree. She is a post-graduate student of the Federal State Unitary Enterprise "GosNIIAS" in the field of "Management in Technical Systems", and she is an engineer at the Advanced Civil Aviation Avionics Systems Division. Currently, the main areas of research interests: research and development of materials and structures that are actively used in the fields of microelectronics and photonics, including photonic integrated circuits for aviation applications. She is the author and co-author of 8 scientific papers. E-mail: [mishchenko.ib@gmail.com](mailto:mishchenko.ib@gmail.com).

**Valery Mikhailovich Novikov** (b. 1954). In 1981, he graduated from Moscow Aviation Institute with a specialization in Electronic Computing. He is the lead of the sector "Development of the Perspective Architecture of the Avionics Complex" of FSUE GosNIIAS. Currently, the main areas of research interests are: creating architectures for advanced avionics systems, research and application of optical components on board aircraft in avionics systems, including photonic integrated circuits for aviation applications. He is the author and co-author of 40 scientific papers and five patents. E-mail: [ymnovikov@2100.gosnias.ru](mailto:ymnovikov@2100.gosnias.ru).

**Nikolay Ivanovich Selvesiuk** (b. 1973). In 1994, he graduated from the Stavropol Higher Aviation Engineering School im. Marshal of Aviation V.A. Sudts. He is a Doctor of Technical Sciences (2007), associate Professor, Professor of the Russian Academy of Sciences. He is the Deputy General Director and head of the scientific complex of the State Scientific Center of the Federal State Unitary Enterprise "State Research Institute of Aviation Systems". He is a specialist in linear algebra, multi-connected control systems, and image analysis and processing. He has published more than 120 scientific papers, 16 textbooks and manuals, 5 author's certificates and patents. E-mail: [nis@gosnias.ru](mailto:nis@gosnias.ru).

**Egor Ivanovich Ershov** (b. 1990), Ph.D. in Physics and Mathematics. He is working in Vision Systems Lab at the Institute for Information Transmission Problems RAS since 2014. Graduated from Moscow Institute of Physics and Technology in 2014. Research interests: color computer vision, fast Hough and Radon transforms, visual odometry. E-mail: [ershov@iitp.ru](mailto:ershov@iitp.ru).

**Nikolay Alexandrovich Ivliev** (b. 1987), graduated from Samara State Aerospace University in 2010 (presently, Samara National Research University, short – Samara University), majoring in Design and Technology of Radioelectronic Equipment. Candidate of Engineering Sciences (2015). Currently he works as the researcher at the Image Processing Systems Institute of RAS – Branch of the FSRC "Crystallography and Photonics" RAS, assistant at Technical Cybernetics sub-department of Samara University. Research interests: surface physics, micro- and nanotechnology. E-mail: [ivlievn@gmail.com](mailto:ivlievn@gmail.com).

**Roman Vasilevich Skidanov** (b. 1973). Graduated with honors (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 SSU's Technical Cybernetics sub-department. He is co-author of 160 scientific papers, 7 monographs. His current research interests include diffractive optics, mathematical modeling, image processing, and nanophotonics. E-mail: [romans@ipsiras.ru](mailto:romans@ipsiras.ru).

**Nikolay Lvovich Kazanskiy** (b. 1958), graduated with honors (1981) from S.P. Korolyov Kuibyshev Aviation Institute (presently, S.P. Korolyov Samara National Research University), majoring in Applied Mathematics. He received his Candidate in Physics & Maths (1988) and Doctor in Physics & Maths (1996) degrees from Samara University. He is the director of Image Processing Systems Institute of the Russian Academy of Sciences (IPSI RAS), also holding a part-time position of a professor at Technical Cybernetics sub-department of Samara University, holding the chair at the sub-department of High-Performance Computing at IPSI RAS. He is an OSA-, SPIE and IAPR-member. He is co-author of 290 scientific papers, 12 monographs, 53 inventions and patents. His current research interests include diffractive optics, mathematical modeling, image processing, and nanophotonics. Website: <http://www.ipsi.smr.ru/staff/kazanskiy.htm>. E-mail: [kazanskiy@ssau.ru](mailto:kazanskiy@ssau.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](mailto:artniko@gmail.com).

---

---

---

---